Statistical and thermodynamical analysis of the micro-structure in molecular emulsions

Kežić-Lovrinčević, Bernarda

THESE DE DOCTORAT DE
L'UNIVERSITE PIERRE ET MARIE CURIE
EN COTUTELLE AVEC
L'UNIVERSITE DE SPLIT (CROATIE)

Spécialité
Physique Théorique et Physique Chimie

Présentée par
Bernarda KEŽIĆ-LOVRINČEVIĆ

Pour obtenir le grade de
DOCTEUR des UNIVERSITÉS PIERRE
ET MARIE CURIE ET DE SPLIT

Sujet de la thèse:

Analyse statistique et thermodynamique de la micro-structure dans les émulsions moléculaires

(Statistical and thermodynamical analysis of the micro-structure in molecular emulsions)

Soutenue le 23 Octobre 2013

devant le jury composé de :

Mme.	Hong	XU	(Rapporteur)
M.	Pascal-Henry	FRIES	(Rapporteur)
M.	Thomas	ZEMB	(Examinateur)
M.	Xavier	ASSFELD	(Examinateur)
M.	Marc	SIMON	(Examinateur)
M.	Gerhard	KAHL	(Examinateur)
M.	Franjo	SOKOLIĆ	(co-directeur)
M.	Aurélien	PERERA	(co-directeur)

Thèse préparée au
Laboratoire de physique théorique de la matière condensée (UMR CNRS 7600).
Université Paris VI Pierre et Marie Curie
75 005 Paris CEDEX

et

Department de physique
Université de Split
21000 Split, Croatie

Résumé

Les mélanges aqueux diffèrent des mélanges de liquides simples, du fait que les molécules d'eau tendent à s'associer préférentiellement entre elles au travers de la liaison hydrogène, de ce fait ségréguant les molécules de soluté sans pour autant entraîner une démixtion totale des mélanges. Même des mélanges aussi simples que l'eau et le méthanol exhibent ce type d'ordre. La micro-hétérogénéité (MH) résultante apparaît comme étant distincte des fluctuations de concentration (CF), qui jouent un rôle important dans la stabilité des mélanges. En établissant la correspondance entre les micro-émulsions, et ce mélanges que nous appelons ”émulsions moléculaire“ à partir de la théorie Ornstein-Zernike aux petits vecteurs d’onde, nous montrons comment certains problèmes intrinsèques aux simulations, et qui gênaient par exemple le calcul des intégrales de Kirkwood-Buff, peuvent être résolus en recalculant les corrélations à longue portée. De même, l’introduction des fonctions bridges des corps purs extraits des simulations, permet de mieux comprendre le comportement des équations intégrales pour les mélanges, en particulier en ce qui concerne la distinction entre CF et MH.

Mots-clefs: mélanges aqueux, micro-hétérogénéité, fluctuations de concentration, simulations numériques, émulsion moléculaire, équations intégrales.

STATISTIČKA I TERMODINAMIČKA ANALIZA MIKRO-STRUKTURE U MOLEKULARNIM EMULZIJAMA

Sažetak

Vodene mješavine se razlikuju od mješavina jednostavnih tekućina, jer se molekule vode između sebe povezuju vodikovom vezom te uzrokuju segregaciju otopljene tvari na molekularnoj razini. Ovakva vrsta mikro-strukture postoji čak i u vodenoj mješavini najjednostavnijeg alkohola, a to je metanol. Rezultirajuća mikro-heterogenost (MH) se razlikuje od koncentracijskih fluktuacija (KF), koje imaju važnu ulogu u stabilnosti mješavina. Pomoću analogije između mikro-emulzija i ovih mješavina koje smo nazvali “molekularnim emulzijama” u Ornstein-Zernike teoriji za male valne vektore, pokazat ćemo kako se određeni problemi računalnih simulacija, primjerice problemi vezani za račun Kirkwood-Buff integrala, mogu rješiti tako da se uzmu u obzir dugodosežne korelacije. Iz računalnih simulacija jednokomponentnih tekućina izračunat ćemo korelacije višeg reda koje će nam omogućiti bolje razumijevanje ponašanja integralnih jednadžbi za mješavine tekućina, a prvenstveno razlikovanje između MH i KF.

Ključne riječi: vodene mješavine, mikro-heterogenost, koncentracijske fluktuacije, računalne simulacije, molekularna emulzija, integralne jednadžbe.

Contents

Acknowledgements/Remerciements/Zahvale

There are many people to whom I owe my gratitude for being enormously caring and supportive to me for the past 3 years I spent both in France and in Croatia. First of all, I would like to thank my supervisor, Aurélien Perera for guiding me through this thesis. His energy and love for science were a true inspiration to me. He has been a kind friend to me from the first day I joined the lab in Paris. Working with him, I have learned a great deal about Liquid Theory, but I have also grown more mature as a scientist.

To my Croatian supervisor, prof. Franjo Sokolić, I would like to express my gratitude for the support and scientific advice. I will always appreciate his faith in me. I am grateful to Larisa Zoranić and Marijana Mijaković for constructive discussions and encouragement. I am happy to have them by my side. I also wish to acknowledge the support of prof. Davor Juretić from the Doctoral School of Biophysics in Split.

I would like to thank all the kind people from *Laboratoire de Physique Théorique de la Matière Condensée* (LPTMC) at *Université Pierre et Marie Curie*, especially the director, Pascal Viot for welcoming me in the lab. Special thanks goes to the secretaries of LPTMC for their help in administrative tasks.

I would like to thank the French Government for an allocation of a three years research grant.

Finally, and most importantly, I thank my parents Marija and Goran, my sisters Zrinka and Dubravka, and my husband Dejan for their infinite support and love.

List of Figures

List of Tables

CHAPTER 1

Introduction

What makes liquid state different from gas and solid state of matter? In a gas state, molecules move randomly and independently of each other, with rare collision between them. In a solid state, molecules are sitting on the knots of a regular lattice and are allowed only to vibrate around their fixed positions. Liquid state is characterized by the greater importance of collision processes and short–range positional correlations [1]. For example, in liquid acetone, since molecules are polar, they tend to lie parallel or antiparallel to each other, but this order is lost beyond a certain distance [15]. The concept of order is described in Statistical Mechanics through the use of an order parameter. Generally, this is the one–body density function $\rho(1)$ specifying the positions and orientations of all molecules in the liquid. If the liquid is homogeneous and isotropic, then this is just the bulk number density $\rho = N/V$, where N is the number of particles closed in a volume V. Correlations between two particles are described by the two–body density function $\rho^2(1,2)$ and the range of correlations is associated to a single length scale called the correlation length ξ. It is possible to relate this quantity to density fluctuations in a single–component liquid and concentration fluctuations in a two–component liquid. When the correlation length diverges, density (or concentration) fluctuations grow and a phase transition occurs. When the correlation length grows, the system is metastable. There is a known effect of critical opalescence that happens in some binary mixtures, such as the mixture of methanol and cyclohexane [2], when the correlation length reaches the visible light wave length at some temperature. The light is scattered by fluctuations and as a consequence the mixture appears fuzzy. In scattering experiments, it is possible to determine the static structure factor $S(k)$ which is determined by the two–body correlation function in the absence of the external field. This is an example of the fluctuation–dissipation theorem [1].

In a mixture of species *a* and *b*, the concentration fluctuations can be written as [3]

$$\int (g_{ab}(r)-1)d\vec{r} = V\frac{\langle N_a N_b\rangle - \langle N_a\rangle\langle N_b\rangle}{\langle N_a\rangle\langle N_b\rangle},$$

where $g_{ab}(r)$ is the radial distribution function. This function represents the probability of finding particle 2 at a distance *r* from particle 1 fixed at the origin, if the system is uniform and isotropic. Otherwise, it depends also on the mutual orientation of the particles. Integrals of the pair distribution functions are called the Kirkwood–Buff integrals (KBI) [3] and they are directly related to some thermodynamic properties, such as partial molar volumes. They provide also a measure for concentration fluctuations, but I will show in this thesis that this link is not as straightforward as one may think.

In 1984 Matteoli and Lepori [4] published an article where they compared the experimental values of KBI for a variety of aqueous mixtures, and the results are surprising. Figure 1.1 shows KBI for 9 different aqueous mixtures.

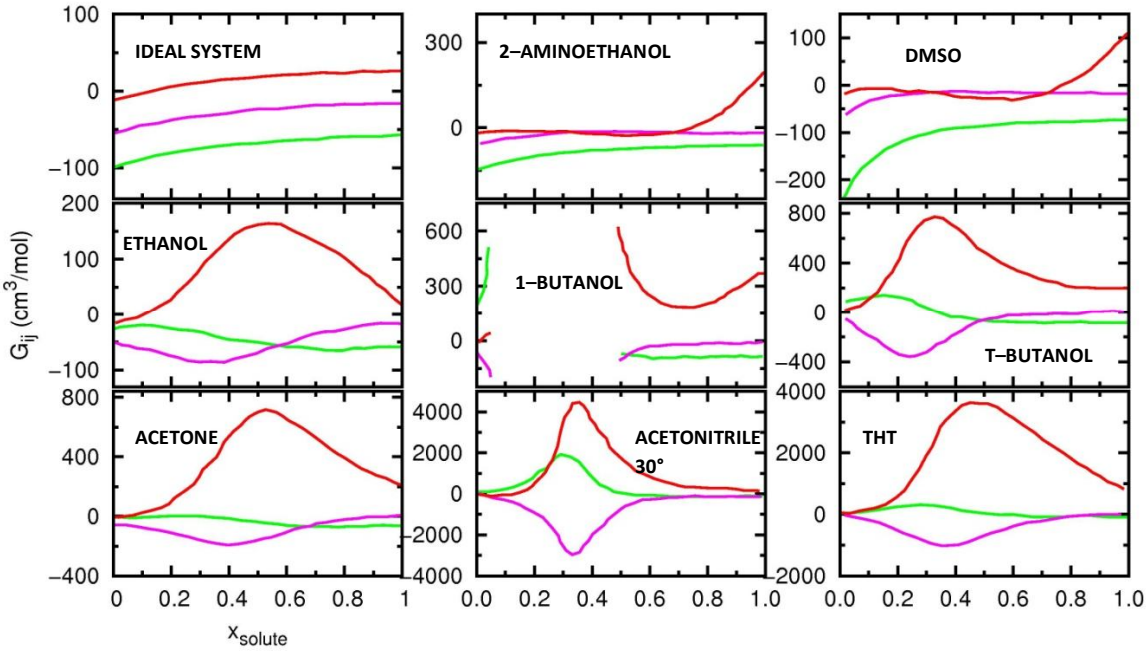

Figure 1.1: The KBI from Ref. [4]. The water–water pair is in red, the solute–solute pair is in green and the cross pair is in magenta.

Aqueous mixture of 1–butanol demixes, as one observes in the middle panel. Some KBI look almost as if the mixture were ideal (non–interacting particles), such as water–DMSO, while most

others look as if the mixture were metastable and would demix with a small decrease of temperature. This concerns in particular mixtures such as water–*tert*–butanol (TBA), water–acetone, or water–acetonitrile. We know that these mixtures are stable at room temperature, so the origin of high KBI values cannot be in high concentration fluctuations. In particular, high KBI values are seen for the water–water pair. The decay of the pair correlation function is given by the Ornstein–Zernike form

$$g_{ab}(r) \to 1 + \exp\left(-\frac{r}{\xi}\right)/r .$$

If there are no significant concentration fluctuations, this means that the correlation length ξ is not very large. This suggests that the high KBI value does not come from the long–range part of $g_{ab}(r)$, but rather the short–range part. Non–trivial local structure produces high $g_{ab}(r)$. In other words, this is an indication of a strong segregation of water molecules, which is driven by the H–bonding interaction. Associated liquids, such as water or alcohols, are therefore different from simple liquids in a sense that the same phenomenon (high KBI) is not necessarily explained in the same way, by invoking concentration fluctuations.

There is a related and equally intriguing experimental feature, known from various calorimetric experiments. When mixing 2 liquids, where the solvent is an associated liquid (for example, water), then various quantities such as excess enthalpy, speed of sound, excess heat capacity, viscosity, show two weak singularities, one at low solute mole fraction (typically below $x = 0.30$, when $x = N_S / N$, where N is the total number of particles and N_S is the number of solute molecules) and the other at high solute mole fraction (typically above $x = 0.60$). These singularities are weak since they represent a rapid variation of the derivative near these 2 points. For example, excess enthalpy of aqueous ethanol solution [5] has two slope variations: one at $x \approx 0.20$ and another at $x \approx 0.70$ mole fraction of ethanol, while for cyclopenthane–tetrachloroethylene [6] mixture this function has almost an ideal shape (Figure 1.2).

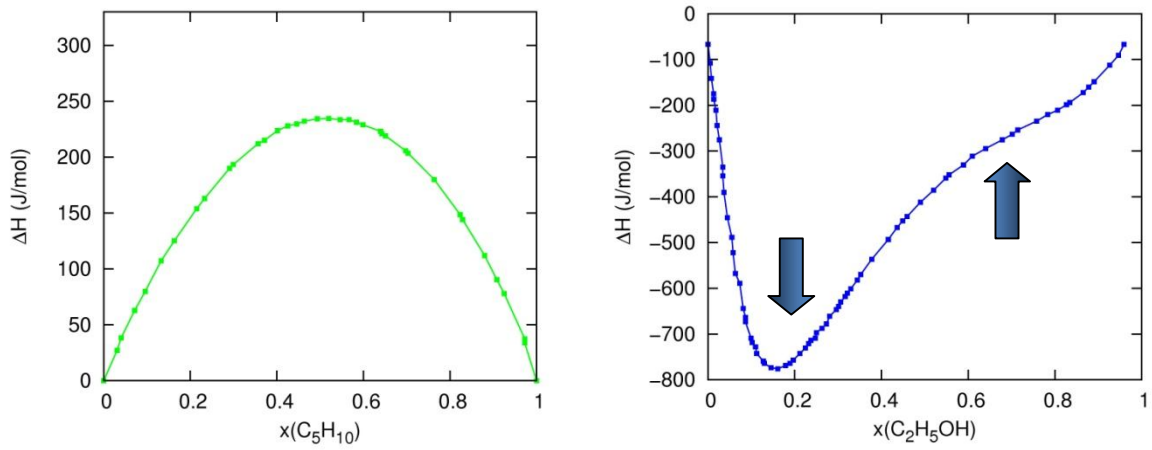

Figure 1.2: Experimental excess enthalpies for cyclopenthane–tetrachloroethylene mixture [6] (left panel) and aqueous ethanol [5] (right panel).

These changes in slope are not signs of any known types of phase transition, since no discontinuities are observed in the first and the second derivatives of the Gibbs free energy [7]. Instead, they are attributed to the changes of the microscopic order in the system and the way molecules reorganize themselves at a certain concentration [8]. Studies on aqueous mixtures of ethanol [9, 10] and TBA [11] suggest that the anomalous behaviour of thermodynamic quantities is a signature of the structural complexity these systems exhibit, due to the H–bonding ability of water and the amphyphilic nature of the solute molecules.

How can we study the microscopic origin of this complexity? Computer simulations are a valuable tool to study the molecular structure in details, but they are not free from restrictions. Let us consider for example the dynamics of micro–emulsions. Micro–emulsion is a mixture of water, oil and a surfactant, where oil and water form domains that are in continuous movement and collision with each other, while the surfactant molecules help stabilizing these objects by sitting in between water and oil domains [12]. An average collision time between two domains is 1 μs, which is not out of reach of modern computers, but is still far from desktop calculations. Each domain is regarded as a spherical droplet which can contain tens of thousands of molecules and if we take for example 50 droplets, the total number of molecules in the simulation box becomes of the order of 10^6, which is beyond the reach of the current computational power.

It is seen that even mixtures of water and higher alcohols, such as TBA, require larger simulation box with more than 2048 particles, which is generally accepted to be enough in the study of simple aqueous mixtures. Indeed, this mixture is known for its micro–heterogeneous structure,

where water and alcohols form domains or micelles and therefore can be regarded as a baby micro–emulsion. In fact, in this thesis, I present a new concept of molecular–emulsion, having in mind the difference in size between the objects that form in micro–emulsion and in molecular–emulsion. The analogy I find between the two systems is quite simple. Let us consider a mixture of water and a molecule with a large oily tail. This type of emulsion can form two different structural profiles. First one is the oil–in–water emulsion, where oil droplets are dispersed in water, and the second is the water–in–oil emulsion with water being in the dispersed phase. This mixture will eventually de–mix unless there are surfactant molecules that allow for its stability. Now, imagine that the oil molecule becomes smaller and smaller, having at the end a short hydrophobic tail and keeping its hydrophilic head. What kind of structure will take place in such an emulsion? Clearly, there is no large size difference between the water and the solute molecules and the effect of concentration fluctuation becomes of significance, but there will still be domain formation. It is not obvious how one can distinguish between these two phenomena. One possible answer is hidden in the long–range part of the correlation functions and the associated small–k behaviour of the structure factors. Although this solution appears to be rather simple, it has been shown by many researches in the field of computer simulations that these systems are not easy to study, since they require large sizes and simulation times to accommodate both size and dynamics of the objects forming in the system. When simulating systems characterized by the domain or micelle formation, it is possible in principle to modify the interaction parameters in order to accommodate the micro–objects within a reasonably large box. However, it is not clear whether this method distorts the "true" micro–segregation.

This brings us to the heart of the problem computer simulations have in the study of ordered systems with fluctuations. If we consider that one realisation (or one micro-state) of the system is an instantaneous snapshot, then fluctuations allows us to appreciate how different one micro-state is from another. There are 2 major problems. The first is the fact that one snapshot does not represent the whole macroscopic system, but only a part of it. So, this representation is possible only if the part of the system is not different from the rest of the whole system. This is more valid for simple liquids with truly random disorder than complex liquids that have extended local order. In such systems, the periodic boundaries can affect the way the system is connected inside the box, and hence alter its statistical properties. This may not be so important in thermodynamic quantities, but it can be a subtle effect. The second problem is closely related to the first, since it appears when the system shows a complex spatial organisation. This is the case of mixtures such as aqueous mixtures. Clear patterns

appear which are not defined by a simple geometrical method. The box will then distort these patterns to accommodate them inside the periodic boundary conditions. Both problems pose the question of the "coherence" of the system with respect to box size and time scale. If the system is de–correlated before reaching the boundaries, then the usual statistical analysis is adequate. A third problem, but which is not directly related to simulations, is how this "coherence" can be measured in real experiments.

An alternative technique to study these systems is to use Integral Equations [1, 13] (IE), which are free from those constraints. The main problem of such equations is that they contain severe approximations on the many body contributions (the so called bridge terms) and therefore, they do not describe fluctuations properly. So, a major issue is to see what the difference between fluctuations and local heterogeneity is from the point of view of the Liquid State Theory. It is known that near the critical point, where fluctuations become important, IE either lead to spurious solutions either do not converge numerically in the thermodynamic limit of interest [1, 13].They also fail in producing the numerical solution mostly in the high density–low temperature regime. For instance, the behaviour of the hypernetted–chain theory (HNC) in the approach to the phase boundary is characterized by an increasing instability of the numerical algorithm, until a limiting density (or temperature) is reached beyond which no solution can be found [13]. This line of no solution is not the true phase boundary of the system. It is obvious that when fluctuations exceed a certain spatial range, these theories loose solution. Now, if spatial heterogeneity is akin to some sort of fluctuation, how would they be described by IE? If IE do not make any difference between fluctuations and heterogeneity, they will cease to have solution by mistakenly taking heterogeneity for a big fluctuation. So, in a way, the behaviour of these theories would inform us about the status of fluctuations versus micro–heterogeneity, albeit in an indirect way.

The central equation of IE is the exact Ornstein Zernike (OZ) equation. In the molecular form, it is the Molecular Ornstein Zernike equation (MOZ), and in terms of site–site correlations it is the site–site Ornstein Zernike equation (SSOZ). The SSOZ equation is approximate equation where the molecular direct correlation functions are written as a sum of the site–site direct correlation functions. This is exactly true for the pair interaction, and since the asymptotical form of the direct correlation function is exactly that of the pair interaction, this approximation is justified in this limit. However, it is not true for medium and short range distances, so the SSOZ approximation suffers from many inconsistencies. For example, the results obtained

from the equation of state and those obtained for quantities that depend on angular correlations in the fluid, such as dielectric constant, are incorrect [1, 14]

Both MOZ and SSOZ equations have to be solved in conjecture with another equation, called the closure equation. This equation contains the uncontrolled approximation since it does not take into account the correct form of the many–body correlations. Even now, after 50 years of using IE, it is still not clear why some approximations are well suited to describe some systems and badly for other systems. Concerning the accuracy of the solution, the MOZ equation gives better results than the SSOZ equation [15]. However, when applied to liquid water at room temperature, both equations fail severely and none of them is able to describe the correct tetrahedral structure of water [16-22]. Figure 1.3 shows the comparison between the two IE theories and the Molecular Dynamics simulation results for SPC/E water. As one can see, both theories are missing some features of the site–site radial distribution functions. In the MOZ formalism, the full molecular distribution function has one distance and 5 angle variables and its representation extends over an infinite set of functions [1, 23]. This is difficult to represent, even if one considers a finite set of such functions. This is why the site–site representation is so appealing – one can see how pairs of sites are correlated and there is always a finite set of these functions. Since both MOZ and SSOZ suffer from approximations contained in the closure equation, it is unclear whether any of them would in fact give the correct description of fluctuations and micro–heterogeneity. In addition, both theories are seen to fail to describe correlations beyond the range of a few molecular diameters for some systems, so it is easier to try first with the easier to implement SSOZ theory.

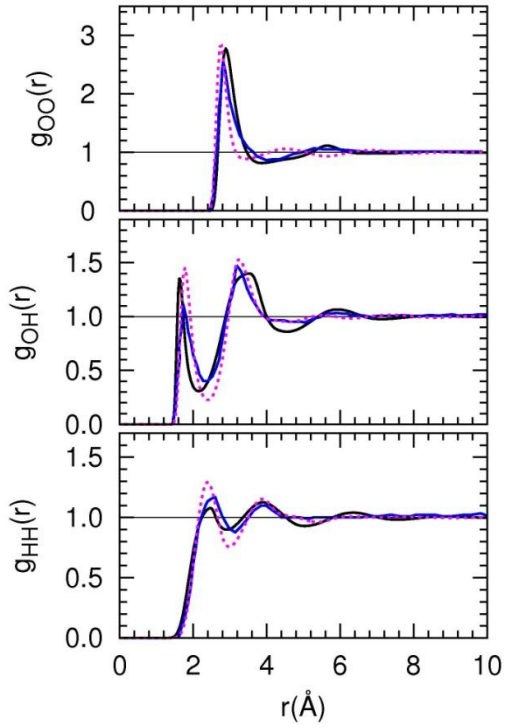

Figure 1.3: Comparison between of the pair distribution functions from theory and simulations for SPC/E water at ambient conditions. Magenta dotted line is for the MD simulation results, blue full line is for the MOZ/HNC results [22] and black full line is for the SSOZ/HNC results.

The thesis is organized as follows. Chapter 2 presents a summary of IE theories, namely the SSOZ and the MOZ theory. Chapter 3 presents the results of the SSOZ equation for various pure liquids, such as acetone, carbon–tetrachloride, formamide, lower alcohols and water. Several closure relations (HNC, KH, PY) are used and both structure and thermodynamics are tested against the simulation results. A new procedure of extracting the bridge functions from the simulation data is also discussed in this Chapter. In Chapter 4, results of the SSOZ equation for selected aqueous mixtures are discussed, as well as the role of fluctuations and micro–heterogeneity. The mixtures are aqueous methanol and a mixture of water and modified water model. Molecular simulation results for aqueous ethanol are presented and the structure and thermodynamics of this mixture is analysed in Chapter 5. The concept of molecular–emulsion is discussed in Chapter 6. In Chapter 7, aqueous mixtures of TBA and acetone are analysed through the concept of molecular–emulsion. Computer simulation results for structure and thermodynamics of these mixtures are presented and discussed. Mixture of water and three different modified water models was studied through computer simulations. The results are discussed in Chapter 8. OZ equation was applied to a binary mixture of core–soft fluids and the

results are compared with computer simulations in Chapter 9. Finally, the conclusion of this thesis is given in Chapter 10. Exact derivations of the important equations in this thesis are shown in Appendices A, B, C and D.

CHAPTER 2

Integral Equation Theory

2.1 Introduction

I give an overview of the basic definitions of the correlation functions and their relations to the thermodynamic properties in the Liquid Theory, following the standard textbook by Hansen and McDonald "Theory of simple liquids" [1].

Consider a one–component monoatomic liquid that consists of N particles enclosed in a volume V. Position and momenta of all the particles in the system are $\vec{r}^{\,N} = \{\vec{r}_1, \ldots, \vec{r}_N\}$ and $\vec{p}^{\,N} = \{\vec{p}_1, \ldots, \vec{p}_N\}$, respectively. The equilibrium particle density $\rho_N^{(n)}$ is defined as

$$\rho_N^{(n)}(\vec{r}^{\,n}) = \frac{N!}{(N-n)!}\frac{1}{h^{3N}N!Q_N}\iint \exp(-\beta H)\, d\vec{r}^{\,(N-n)} d\vec{p}^{\,N} = \frac{N!}{(N-n)!}\frac{1}{Z_N}\int \exp\left(-\beta U_N\right) d\vec{r}^{\,(N-n)}, \quad (2.1)$$

where $\beta = 1/(k_B T)$, k_B is the Boltzmann constant and T is the absolute temperature. Here H and U_N are the Hamiltonian and the total potential energy of the system.

The canonical partition function Q_N is defined as

$$Q_N = \frac{1}{h^{3N}N!}\iint \exp(-\beta H) d\vec{r}^{\,N} d\vec{p}^{\,N} \quad (2.2)$$

and the configurational integral $Z_N = \int \exp(-\beta U_N) d\vec{r}^{\,N}$.

The quantity $\rho_N^{(n)}(\vec{r}^{\,n}) d\vec{r}^{\,n}$ is the probability of finding *n* particles of the system with coordinates in the volume element $d\vec{r}^{\,n}$ irrespective of the positions of the remaining particles and

irrespective of all momenta. The particle densities and the equilibrium particle distribution functions $g_N^{(n)}$ completely describe the structure of a fluid. Low–order particle distribution functions, the pair distribution function $g_N^{(2)}(\vec{r}_1, \vec{r}_2)$ in particular, are sufficient to calculate the equation of state and other thermodynamic properties of the system in the case of pairwise additive intermolecular energy potentials. The definition of the *n*–particle density means that

$$\int \rho_N^{(n)}(\vec{r}^{\,n}) d\vec{r}^{\,n} = \frac{N!}{(N-n)!} \ . \tag{2.3}$$

The single particle density of a uniform fluid is therefore equal to the number density ρ :

$$\rho_N^{(1)}(\vec{r}) = \frac{N}{V} = \rho \ . \tag{2.4}$$

The particle densities can also be expressed in terms of δ–functions:

$$\rho_N^{(1)}(\vec{r}) = \left\langle \sum_{i=1}^{N} \delta(\vec{r} - \vec{r}_i) \right\rangle \tag{2.5}$$

and

$$\rho_N^{(2)}(\vec{r}, \vec{r}^{\,'}) = \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N}{}' \delta(\vec{r} - \vec{r}_i) \delta(\vec{r}^{\,'} - \vec{r}_j) \right\rangle, \tag{2.6}$$

where the prime on the summation indicates that terms for which $i = j$ must be omitted.

The *n*–particle distribution function $g_N^{(n)}$ is defined in terms of the corresponding particle densities:

$$g_N^{(n)}(\vec{r}^{\,n}) = \frac{\rho_N^{(n)}(\vec{r}_1, \ldots, \vec{r}_n)}{\prod_{i=1}^{n} \rho_N^{(1)}(\vec{r}_i)} \ . \tag{2.7}$$

For a homogeneous system, Eq. (1.14) reduces to

$$\rho^n g_N^{(n)}(\vec{r}^{\,n}) = \rho_N^{(n)}(\vec{r}^{\,n}) \, . \tag{2.8}$$

The particle distribution functions measure the extent to which the structure of a fluid deviates from complete randomness. If a system is also isotropic, the pair distribution function $g_N^{(2)}(\vec{r}_1, \vec{r}_2)$ is a function only of the separation $r_{12} = \left| \vec{r}_2 - \vec{r}_1 \right|$. Then it is usually called the radial

distribution function and written as $g(r)$. This function is measurable by radiation–scattering experiments and it contains important information about the structure of a liquid.

Equations (2.1) and (2.2) can be extended to the grand–canonical ensemble (ensemble with constant chemical potential μ, volume V and temperature T). It can be shown that in the dilute gas limit $\rho \to 0$, one has

$$g(r) \to \exp(-\beta U(r)). \tag{2.9}$$

Consider a uniform fluid for which the total potential energy is given by a sum of the pair terms:

$$U_N(\vec{r}^N) = \sum_{i=1}^{N} \sum_{j>i}^{N} U(r_{ij}), \tag{2.10}$$

where $U(r_{ij})$ is the pairwise interaction potential between particles *i* and *j* and $r_{ij} \equiv \left| \vec{r}_i - \vec{r}_j \right|$.

The excess internal energy U^{ex} per particle of this system is given as

$$\frac{U^{ex}}{N} = 2\pi\rho \int_0^\infty U(r) g(r) r^2 dr, \tag{2.11}$$

Similarly, it is possible to express the equation of pressure (called also the equation of state):

$$\frac{\beta P}{\rho} = 1 - \frac{2\pi\beta\rho}{3} \int_0^\infty U'(r) g(r) r^3 dr, \tag{2.12}$$

where $U'(r) \equiv dU(r)/dr$.

In a grand canonical ensemble isothermal compressibility κ_T is expressed as an integral over $g(r)$ via the so called compressibility equation:

$$1 + \rho \int [g(r) - 1] d\vec{r} = \rho k_B T \kappa_T. \tag{2.13}$$

Eq. (2.13) is written for a homogenous system and it is applicable regardless the pairwise additivity of the interparticle forces. This equation also relates the fluctuations in the system with the isothermal compressibility.

In radiation–scattering experiments the measured quantity is the structure factor $S(k)$ which has the following form in the homogeneous case:

$$S(k) = 1 + \rho \int g(r) \exp(-i\vec{k}\vec{r}) d\vec{r} \,. \tag{2.14}$$

Therefore, $g(r)$ is given by the inverse transform of $S(k)$:

$$\rho\, g(r) = (2\pi)^{-3} \int (S(k) - 1) \exp(-i\vec{k}\vec{r}) d\vec{k} \,. \tag{2.15}$$

In a uniform and isotropic fluid, the so called Ornstein–Zernike (OZ) relation is defined as

$$h(r) = c(r) + \rho \int c\left(\left|\vec{r} - \vec{r}'\right|\right) h(r') d\vec{r}' \,, \tag{2.16}$$

where ρ is the number density of a fluid, $h(r) = g(r) - 1$ is the pair correlation function and $c(r)$ is the pair direct correlation function between two particles separated by the distance r. Here the integration is taken over the position $\vec{r}'$ of a third particle. On taking the Fourier transform of Eq. (2.16), one obtains

$$\tilde{h}(k) = \frac{\tilde{c}(k)}{1 - \rho \tilde{c}(k)} \,, \tag{2.17}$$

where $\tilde{a}(k)$ is defined as $\tilde{a}(k) = \int a(\vec{r}) \exp(-i\vec{k}\vec{r}) d\vec{r}$, $a = h, c$.

In an isotropic fluid, from Eq. (2.13) and Eq. (2.14) at the limit $k = 0$, one has

$$S(0) = \rho k_B T \kappa_T \,. \tag{2.18}$$

From Eq. (2.17) and Eq. (2.18), it is easily shown that

$$\rho \tilde{c}(0) = 1 - \frac{1}{\rho k_B T \kappa_T} \,. \tag{2.19}$$

Eq. (2.19) suggests that $\tilde{c}(0)$ is finite so the function $c(r)$ decays as $1/r^3$ or more rapidly.

If one iteratively substitutes $h(r)$ inside the kernel of Eq. (2.16), one obtains

$$h(r) = c(r) + \rho \int c(|\vec{r} - \vec{r}'|) c(\vec{r}') d\vec{r}' + \rho^2 \int c(|\vec{r} - \vec{r}'|) c(|\vec{r}' - \vec{r}''|) c(\vec{r}'') d\vec{r}' d\vec{r}'' + \ldots \,. \tag{2.20}$$

This result has an obvious physical explanation: the pair correlation function $h(r)$ is a sum of different contributions, first one being the direct correlation function $c(r)$, and the indirect correlations propagated via increasingly large number of intermediate particles.

The OZ equation alone is not sufficient since both the pair correlation function $h(r)$ and the direct correlation function $c(r)$ are unknown. Therefore, another equation, the so called closure equation, is needed to complement the OZ equation. For simple fluids, the exact form of the closure equation is given by

$$g(r) = \exp(-\beta U(r) + h(r) - c(r) + b(r)), \tag{2.21}$$

where $b(r)$ is the bridge function, which can be formally expressed as infinite sum over integrals involving higher order correlation functions. Since the direct correlation function itself is a second functional derivative of excess free energy, the bridge function is very difficult to determine and approximations are needed.

Speaking in terms of diagrams, the direct correlation function is an infinite sum of diagrams and the bridge function is an infinite sum of graphs which are a subset of these diagrams. If the bridge function is completely neglected and we set $b(r) = 0$, then the closure becomes of the form

$$g(r) = \exp(-\beta U(r) + h(r) - c(r)). \tag{2.22}$$

This is the so called hypernetted–chain closure (HNC). The HNC closure has given good results for systems with long–range interactions, such as electrolyte solutions [24]. If one expands the power term of $\exp\left(h\left(r\right) - c\left(r\right)\right)$ in (2.2) and keeps just the first two terms of the expansion, one obtains the Percus–Yevick (PY) closure

$$g(r) = \exp(-\beta U(r))(1 + h(r) - c(r)). \tag{2.23}$$

Comparison of Eq. (2.22) and Eq. (2.23) reveals that the PY closure can actually be obtained by linearization of the HNC result with respect to $(h(r) - c(r))$. In diagrammatic analysis the PY approximation sums a smaller class of diagrams than the HNC approximation. The PY closure has proven to be more successful than the HNC when the potential is strongly repulsive and short ranged, such as that of a hard sphere fluid and the Lennard–Jones fluid.

The mean spherical approximation (MSA) is defined in terms of the pair distribution function $g(r)$ and the direct correlation function $c(r)$ as

$$g(r) = 0, \qquad r < d \tag{2.24}$$

$$c(r) = -\beta U_1(r), \; r > d \tag{2.25}$$

where d is the hard–sphere diameter and the interparticle potential is given by

$$\begin{aligned} U(r) &= \infty, \qquad r < d \\ &= U_1(r), \; r > d. \end{aligned} \tag{2.26}$$

Eqs. (2.24) and (2.25) complement the OZ equation; the first equation is exact and the second is an approximation. The MSA gives better results for a square–well fluid than the HNC or the PY closure [25].

The Kovalenko–Hirata (KH) closure [26, 27] combines the HNC and the MSA closure in the following way:

$$\begin{aligned} g(r) &= \exp(d(r)), \quad d(r) < 0 \\ &= 1 + d(r), \qquad d(r) \geq 0, \end{aligned} \tag{2.27}$$

where $d(r) = -\beta U(r) + h(r) - c(r)$.

The KH closure has been successfully employed in the study of water in contact with a metal surface [26] and in the study of the phase behaviour of SPC water [27]. Although it is an empirical closure and therefore diagrammatically incorrect, the KH closure is often used for systems in which the HNC closure fails to produce a numerical solution [28].

There are other closures, such as the Martynov–Sarkisov closure (MaS) [29], the Roger–Young closure (RY) [30], the hybridized mean spherical approximation (HMSA) [31] the soft mean spherical approximation (SMSA) [32] and the Ballone–Pastore–Galli–Gazzillo (BPGG) [33] closure. All these closures provide useful information about specific systems, depending on the nature of the interaction between particles.

The equivalent of the OZ equation for a multicomponent fluid is given by

$$h_{ab}(r) = c_{ab}(r) + \sum_c \rho_c \int c_{ab}\left(|\vec{r} - \vec{r}'|\right) h_{ab}(r') d\vec{r}', \tag{2.28}$$

where $h_{ab} = g_{ab} - 1$ and c_{ab} are the pair and the direct correlation functions between two particles of species *a* and *b*, respectively, and the summation is taken over all species.

The closure relation (2.21) for multicomponent fluid is given by

$$g_{ab}(r) = \exp(-\beta U_{ab}(r) + h_{ab}(r) - c_{ab}(r) + b_{ab}(r)). \tag{2.29}$$

The partial structure factors $S_{ab}(k)$ are defined as

$$S_{ab}(k) = \delta_{ab} + (\rho_a \rho_b)^{1/2} \int d\vec{r} (g_{ab}(r) - 1) \exp(-i\vec{k}\vec{r}), \tag{2.30}$$

where ρ_a is the number density of species *a*.

The equation of state for mixtures is given by

$$\frac{\beta P}{\rho} = 1 - \frac{2\pi}{3} \sum_{a,b} \rho_a \rho_b \int_0^\infty U_{ab}'(r) g_{ab}(r) r^3 dr, \tag{2.31}$$

with the summation taken over all the component species.

It is very instructive to study the integrals of the species pairs of the radial distribution functions.

The running Kirkwood–Buff integrals (running KBI), $G_{ij}(r)$ are defined as [3]

$$G_{ij}(r) = 4\pi \int_0^r \left(g_{ij}(r') - 1 \right) r'^2 dr', \tag{2.32}$$

where *i* and *j* are the species index. The KBI are defined as the large–*r* limit of the running Kirkwood–Buff integrals as

$$G_{ij} = \lim_{r\to\infty} G_{ij}(r) = 4\pi \lim_{r\to\infty} \int_0^r \left(g_{ij}(r') - 1 \right) r'^2 dr'. \tag{2.33}$$

2.2 The Molecular Ornstein–Zernike Equation

Any pair correlation function $a(1,2)$ can be expanded as [1, 23]

$$a(1,2) = a(\vec{R}_{12}, \Omega_1, \Omega_2) = \sum_{m,n}^{\infty} \sum_{l=|m-n|}^{m+n} \sum_{\mu=-m}^{m} \sum_{\nu=-n}^{n} a_{\mu\nu}^{mnl}(R_{12}) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{R}_{12}), \tag{2.34}$$

where $\vec{R}_{12} = \vec{R}_1 - \vec{R}_2$ is the distance between the center of molecules 1 and 2. Set of angles $\Omega_i \equiv (\theta_i, \phi_i, \chi_i)$, i = 1,2 define the orientation of molecule 1 and 2, respectively.

Functions $\Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{R}_{12})$ are called rotational invariants and they are defined by

$$\Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{R}_{12}) = \sum_{\mu'\nu'\lambda} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda \end{pmatrix} R^m_{\mu'\mu}(\Omega_1) R^n_{\nu'\nu}(\Omega_2) R^l_{\lambda 0}(\hat{r}_{12}), \tag{2.35}$$

where $R^m_{\mu'\mu}(\Omega_1)$ is Wigner's generalized spherical harmonic [34].

On taking the Fourier transform of Eq. (2.34) we get:

$$\tilde{a}(1,2) = \tilde{a}(\hat{k}, \Omega_1, \Omega_2) = \sum_{\substack{m,n,l \\ \mu,\nu}} \tilde{a}^{mnl}_{\mu\nu}(k) \Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{k}), \tag{2.36}$$

where the terms $\tilde{a}^{mnl}_{\mu\nu}(k)$ are defined as Fourier–Hankel transform of $a^{mnl}_{\mu\nu}(r)$ as

$$\tilde{a}^{mnl}_{\mu\nu}(k) = 4\pi i^l \int_0^\infty j_l(kr) a^{mnl}_{\mu\nu}(r) r^2 dr, \tag{2.37}$$

and $j_l(kr)$ are spherical Bessel functions.

The MOZ equation in *k*–space can therefore be written as [35]

$$\sum_{m,n,l} \sum_{\mu,\nu} \left[\tilde{h}^{mnl}_{\mu\nu}(k) - \tilde{c}^{mnl}_{\mu\nu}(k) \right] \Phi^{mnl}_{\mu\nu}(\Omega_1, \Omega_2, \hat{k}) = \frac{\rho}{\hat{\Omega}} \int d\Omega_3 \sum_{\substack{m_1,n_1,l_1 \\ m_2,n_2,l_2}} \sum_{\substack{\mu_1,\nu_1 \\ \mu_2,\nu_2}} \tilde{h}^{m_1 n_1 l_1}_{\mu_1\nu_1}(k) \tilde{c}^{m_2 n_2 l_2}_{\mu_2\nu_2}(k)$$

$$\times \Phi^{m_1 n_1 l_1}_{\mu_1\nu_1}(\Omega_1, \Omega_3, \hat{k}) \Phi^{m_2 n_2 l_2}_{\mu_2\nu_2}(\Omega_2, \Omega_3, \hat{k}). \tag{2.38}$$

where $\hat{\Omega} = \int d\Omega$ is the measure of the angular space defining the molecular orientation Ω. We have $\hat{\Omega} = 1$, 4π, and $8\pi^2$ for molecules with spherical, linear, and arbitrary symmetry, respectively. If we define $\eta(1,2) = h(1,2) - c(1,2)$, then the above expression can be written in the form [36]

$$\sum_{m,n,l}\sum_{\mu,\nu}\tilde{\eta}_{\mu\nu}^{mnl}(k)\Phi_{\mu\nu}^{mnl}(\Omega_1,\Omega_2,\hat{k})=\frac{\rho}{\hat{\Omega}}\int d\Omega_3\sum_{\substack{m_1,n_1,l_1\\m_2,n_2,l_2}}\sum_{\mu_1,\nu_1\ \mu_2,\nu_2}\tilde{c}_{\mu_1\nu_1}^{m_1n_1l_1}(k)\left[\tilde{\eta}_{\mu_2\nu_2}^{m_2n_2l_2}(k)+\tilde{c}_{\mu_2\nu_2}^{m_2n_2l_2}(k)\right]$$
$$\times\Phi_{\mu_1\nu_1}^{m_1n_1l_1}(\Omega_1,\Omega_3,\hat{k})\Phi_{\mu_2\nu_2}^{m_2n_2l_2}(\Omega_2,\Omega_3,\hat{k})\,. \tag{2.39}$$

Fries and Patey [36] used the so called χ–transformations [37, 38] to reduce the matrix calculation in Eq. (2.39) and to rewrite the MOZ equation in a matrix form. The χ–transformations are defined as

$$\tilde{a}_{\mu\nu}^{mn\chi}(k)=\sum_{l=|m-n|}^{m+n}\begin{pmatrix}m & n & l\\ \chi & -\chi & 0\end{pmatrix}\tilde{a}_{\mu\nu}^{mnl}(k)\,, \tag{2.40}$$

where $a=\eta,c$, and they greatly reduce the dimensionality of the matrices which must be inverted numerically. The matrices $\tilde{N}_\chi(k)$ and $\tilde{C}_\chi(k)$ can be defined in such a way that the elements (i,j) of these matrices are the quantites $\tilde{n}_{\mu\nu}^{mn\chi}(k)$ and $\tilde{c}_{\mu\nu}^{mn\chi}(k)$, the indices of which satisfy the equations:

$$\begin{aligned} i&=m(m+1)+\mu+1\\ j&=n(n+1)+\nu+1\quad . \end{aligned} \tag{2.41}$$

With the definition of the $P=\left[P_{ij}\right]$ matrix as

$$P_{ij}=\begin{cases}(-1)^{\mu} & for\quad i=m(m+1)-\mu+1\\ & \qquad j=m(m+1)+\mu+1,\\ 0 & otherwise\end{cases} \tag{2.42}$$

The MOZ equation (2.39) can be rewritten in this form of [36]

$$\tilde{N}_\chi(k)=(-)^{\chi}\rho\tilde{C}_\chi(k)P\tilde{C}_\chi(k)\left[I+(-)^{\chi+1}\rho P\tilde{C}_\chi(k)\right]^{-1}\,, \tag{2.43}$$

where I is the identity matrix. In many cases Eq. (2.42) will be reduced by molecular symmetry and the number of independent coefficients can be drastically reduced. This explains why the MOZ equation has been applied only to simple molecular geometries [15, 22, 39].

The MOZ equation accompanied by the closure relation provides a solution for molecular fluids in terms of molecular distribution functions $g(1,2)$ and the thermodynamic quantities can be

obtained by weighted integration of $g(1,2)$. The excess internal energy per particle U^{ex}/N is computed by a configurational average of the pair potential U as [1]

$$\frac{U^{ex}}{N} = \frac{\rho}{2\Omega^2}\iiint d\vec{R}_{12}d\Omega_1 d\Omega_2 U(1,2)g(1,2). \tag{2.44}$$

The pressure equation is given by [1]

$$\frac{\beta P}{\rho} = 1 - \frac{2\pi\beta\rho}{3}\int_0^\infty dR_{12}d\Omega_1 d\Omega_2 U'(1,2)g(1,2)R_{12}^3. \tag{2.45}$$

The isothermal compressibility κ_T is computed by [1]

$$\rho k_B T\kappa_T = 1 + \rho\int_0^\infty d\vec{R}_{12}d\Omega_1 d\Omega_2 \left(g(1,2)-1\right). \tag{2.46}$$

The Kirkwood factor is defined as [15]

$$\begin{aligned} g_K &= 1 - \frac{1}{\sqrt{3}}\frac{\rho}{\Omega^2}\iiint d\vec{R}_{12}d\Omega_1 d\Omega_2 \Phi_{00}^{110}(\vec{R}_{12},\Omega_1,\Omega_2)g(\vec{R}_{12},\Omega_1,\Omega_2) \\ &= 1 - \frac{4\pi}{\sqrt{3}}\rho\int_0^\infty g_{00}^{110}(R_{12})R_{12}^2 dR_{12}, \end{aligned} \tag{2.47}$$

and used to calculate the dielectric constant ε from the Kirkwood equation

$$\frac{(\varepsilon-1)(2\varepsilon+1)}{9\varepsilon} = y_D g_K, \tag{2.48}$$

where $y_D = 4\pi\rho\mu^2/(9k_BT)$ and μ is the dipole moment.

The MOZ equation for a binary mixture of species *a* and *b* is given by [40]

$$h_{ab}(1,2) = c_{ab}(1,2) + \sum_c \frac{\rho_c}{\hat{\Omega}_c}\int d\vec{R}_3 d\Omega_3 c_{ac}(1,3)h_{cb}(3,2), \tag{2.49}$$

where the summation goes over all species and the integration is taken over the position $\vec{R}_3$ and the orientation Ω_3 of particle 3. For linear molecules, one has $\hat{\Omega}_c = \iint d(\cos\theta_i)d\varphi_i = 4\pi$ and

for nonlinear molecules $\hat{\Omega}_c = \iiint d(\cos\theta_i)d\varphi_i d\chi_i = 8\pi^2$. $\rho_c = N_c / V$ is the number density of species *c*.

The excess internal energy per particle U^{ex}/N for mixture is given by [40]

$$\frac{U^{ex}}{N} = \frac{1}{2}\rho\sum_{a,b}\frac{x_a x_b}{\hat{\Omega}_a\hat{\Omega}_b}\iiint d\vec{R}_{12}d\Omega_1 d\Omega_2 U_{ab}\left(\vec{R}_{12},\Omega_1,\Omega_2\right)g_{ab}\left(\vec{R}_{12},\Omega_1,\Omega_2\right), \tag{2.50}$$

where $x_i = \rho_i/\rho$, $i = a, b$ and $\rho = \sum_c \rho_c$.

The compressibility factor $\beta P/\rho$ can be obtained from the virial expression [40]

$$\frac{\beta P}{\rho} = 1 - \frac{\rho\beta}{6}\sum_{a,b}\frac{x_a x_b}{\hat{\Omega}_a\hat{\Omega}_b}\int d\vec{R}_{12}d\Omega_1 d\Omega_2 \frac{\partial U_{ab}(\vec{R}_{12},\Omega_1,\Omega_2)}{\partial R_{12}}g_{ab}(\vec{R}_{12},\Omega_1,\Omega_2). \tag{2.51}$$

The isothermal compressibility κ_T is related to the pair distribution function by the equation [40]

$$\sum_{a,b}\rho_a\rho_b(F^{-1})_{ab} = \frac{\beta}{\kappa_T}, \tag{2.52}$$

where F^{-1} is the inverse of the matrix F . The elements of F are defined by [40]

$$\begin{aligned} F_{ab} &= \rho_a\delta_{ab} + \frac{\rho_a\rho_b}{\hat{\Omega}_a\hat{\Omega}_b}\int d\vec{R}_{12}d\Omega_1 d\Omega_2 h_{ab}(\vec{R}_{12},\Omega_1,\Omega_2) \\ &= \rho_a\delta_{ab} + \rho_a\rho_b\int h_{ab}^{000}(R_{12})d\vec{R}_{12} \\ &= \rho_a\delta_{ab} + \rho_a\rho_b\tilde{h}_{ab}^{000}(k=0) \end{aligned}, \tag{2.53}$$

where δ_{ab} is the Dirac delta function.

2.3 The Site–Site Ornstein Zernike Equation

In the Interaction Site Model (ISM) [41], the interaction energy between molecule 1 and 2, $U(1,2)$ may be approximated by a sum of interactions between specific sites in molecules 1 and 2 which are regarded as perfectly rigid. If we denote the potential between site α on molecule 1

and site β on molecule 2 as $U_{\alpha\beta}(r_{\alpha\beta})$, where $r_{\alpha\beta} = \left|\vec{r}_\alpha^{(1)} - \vec{r}_\beta^{(2)}\right|$ is the distance between these sites, then $U(1,2)$ is given by [42]

$$U(1,2) = \sum_{\alpha,\beta=1}^{m} U_{\alpha\beta}(r_{\alpha\beta}), \tag{2.54}$$

where we assume that each molecule has *m* interaction sites. Chandler and Andersen [41] considered potentials of the form:

$$U_{\alpha\beta}(r) = U_{\alpha\beta}^{(0)}(r) + U_{\alpha\beta}^{(1)}(r), \tag{2.55}$$

where $U_{\alpha\beta}^{(0)}(r)$ is the hard sphere potential with diameter $\sigma_{\alpha\beta}$ and $U_{\alpha\beta}^{(1)}(r)$ is the perturbation. In the Reference Interaction Site Model (RISM) sites interact through the potential of the form [41]:

$$\begin{aligned} U_{\alpha\beta}(r) &= \infty, \quad r < \sigma_{\alpha\beta} \\ &= 0, \quad r > \sigma_{\alpha\beta}, \end{aligned} \tag{2.56}$$

where $\sigma_{\alpha\beta}$ is the interaction distance between the pair of sites.

As given by Eq. (2.34), the pair correlation functions $h(1,2)$ and direct correlation function $c(1,2)$ are functions of both positions and orientations of molecules 1 and 2. Since molecules interact via site–site interactions, the natural way to describe the structure of a fluid is by using the site–site correlation functions $h_{\alpha\beta}(r)$ which are defined as [1]

$$\begin{aligned} h_{\alpha\beta}(\vec{r}) &= \frac{1}{\Omega^2}\iint\iint d\vec{R}_1 d\vec{R}_2 d\Omega_1 d\Omega_2 h(1,2)\delta\left[\vec{R}_1 + \vec{l}_{1\alpha}(\Omega_1)\right]\delta\left[\vec{R}_2 + \vec{l}_{2\beta}(\Omega_2) - \vec{r}\right] \\ &= \frac{1}{\Omega^2}\iint\int d\vec{R}_{12} d\Omega_1 d\Omega_2 h(1,2)\delta\left[\vec{R}_{12} + \vec{l}_{2\beta}(\Omega_2) - \vec{l}_{1\alpha}(\Omega_1) - \vec{r}\right] \end{aligned}, \tag{2.57}$$

where $\vec{l}_{i\alpha}$ is the vector displacement of site α on molecule *i* from the molecular centre $\vec{R}_i$, $\vec{l}_{i\alpha} = \vec{r}_{i\alpha} - \vec{R}_i$, for *i* = 1, 2. Chandler and Andersen [41] were able to derive the site–site Ornstein–Zernike (SSOZ) equation by making an assumption that the direct correlation function between molecules 1 and 2 can be written as a sum of the pair terms

$$c(R_{12}) = \sum_{\alpha,\beta} c_{\alpha\beta}\left(\left|\vec{r}_\alpha^{(1)} - \vec{r}_\beta^{(2)}\right|\right). \tag{2.58}$$

The site–site Ornstein Zernike (SSOZ) Equation has the following matrix form in *k*–space:

$$\widetilde{H} = \widetilde{W}\,\widetilde{C}\,\widetilde{W} + \rho\,\widetilde{W}\,\widetilde{C}\,\widetilde{H}\,, \tag{2.59}$$

where ρ is the number density of the fluid, the matrices $\widetilde{H} = \left[\widetilde{h}_{\alpha\beta}(k)\right]$ and $\widetilde{C} = \left[\widetilde{c}_{\alpha\beta}(k)\right]$ contain three dimensional Fourier transforms of the site–site pair and direct correlation functions respectively and the elements of $\widetilde{W} = \left[\widetilde{w}_{\alpha\beta}(k)\right]$are given by

$$\widetilde{w}_{\alpha\beta}(k) = \delta_{\alpha\beta} + (1-\delta_{\alpha\beta})\frac{\sin(kl_{\alpha\beta})}{kl_{\alpha\beta}}, \tag{2.60}$$

where $l_{\alpha\beta}$ is the fixed distance between sites α and β within the same molecule. Eq. (2.60) is a three dimensional Fourier transform of $w_{\alpha\beta}(r)$ which is defined as

$$w_{\alpha\beta}(\vec{r}) = \delta_{\alpha\beta}\delta(\vec{r}) + (1-\delta_{\alpha\beta})\frac{\delta(r-l_{\alpha\beta})}{4\pi\, l_{\alpha\beta}^2}. \tag{2.61}$$

Function $w_{\alpha\beta}(\vec{r})$ is called the site–site intramolecular correlation function and it is proportional to the probability density of finding site β at distance $\vec{r}$ from site α [23]. The SSOZ equation can be written for mixtures as

$$\widetilde{H} = \widetilde{W}\widetilde{C}\widetilde{W} + \widetilde{W}\widetilde{C}R\widetilde{H}\,, \tag{2.62}$$

where $R = \left[\rho_c\right]$ is the matrix of the numerical densities. The derivation of the SSOZ equation from the MOZ equation for one–component and *n*–component systems is given in Appendix A.

As in the case of the OZ equation, the SSOZ equation needs the closure relation in order to be solved analytically. Chandler and Andersen [41] suggested an approximate closure of the form:

$$h_{\alpha\beta}(r) = -1, \qquad r < \sigma_{\alpha\beta} \tag{2.63}$$

$$c_{\alpha\beta}(r) = -\beta U_{\alpha\beta}(r). \quad r > \sigma_{\alpha\beta} \tag{2.64}$$

Eq. (2.63) is exact, while Eq. (2.64) is an approximation based on the analogy with the MSA for central force potentials [43]. In practice, any closure can be chosen to complement the SSOZ equation and obtain results for the site–site distribution functions $g_{\alpha\beta}(r)$. Various thermodynamic properties can be calculated as integrals of the site–site functions $g_{\alpha\beta}(r)$,

$h_{\alpha\beta}(r)$ and $c_{\alpha\beta}(r)$. The excess internal energy per particle for single–component liquids is given by [1]

$$\frac{U^{ex}}{N} = 2\rho \sum_{\alpha,\beta} \int_0^\infty U_{\alpha\beta}(r) g_{\alpha\beta}(r) d^3 r \,, \tag{2.65}$$

where N is the total number of particles and the summation is taken over all pairs of sites (α, β). For mixtures of n–components, U^{ex}/N is given by [44]

$$\frac{U^{ex}}{N} = 2 \sum_{\alpha,\beta,a,b} \rho_\alpha^{(a)} \rho_\beta^{(b)} \int_0^\infty U_{\alpha\beta}^{(ab)}(r) g_{\alpha\beta}^{(ab)}(r) d^3 r \,, \tag{2.66}$$

where $g_{\alpha\beta}^{(ab)}(r)$ and $U_{\alpha\beta}^{(ab)}(r)$ denote the pair radial distribution function and the pair potential between sites α and β on molecules of species a and b, respectively. $\rho_\alpha^{(a)}$ and $\rho_\beta^{(b)}$ are the density of sites of type α and β on molecules of species a and b, respectively and the summation goes over all pairs of species and all pairs of sites.

The excess Helmholtz free energy A^{ex} can be calculated by using the formula of Hiroike [45]

$$\frac{A^{ex}}{V} = \int_0^1 d\zeta \left\langle U^{exc} \right\rangle_\zeta = \frac{1}{2} \sum_{\alpha,\beta,a,b} \rho_\alpha^{(a)} \rho_\beta^{(b)} \int d^3 r \int_0^1 d\zeta \frac{\partial U_{\alpha\beta}^{(ab)}(r;\zeta)}{\partial \zeta} g_{\alpha\beta}^{(ab)}(r), \tag{2.67}$$

where ζ is the coupling parameter of the interaction $\left\langle U^{exc} \right\rangle_\zeta$ between all the pairs of molecules. Following diagrammatic arguments, Morita and Hiroike [46] wrote an expression for A^{ex} which involves the integration of the pair distribution function. Singer and Chandler [47] derived a formula which is consistent with the HNC closure

$$\begin{aligned} \frac{\beta A^{ex}}{V} = & -\frac{1}{2} \sum_{\alpha,\beta,a,b} \rho_\alpha^{(a)} \rho_\beta^{(b)} \int d^3 r \left\{ c_{\alpha\beta}(r) - \frac{1}{2} h_{\alpha\beta}(r) \right\} \\ & + \frac{1}{2} \int \frac{d^3 k}{(2\pi)^3} \left\{ Tr \tilde{W}(k) \tilde{C}(k) + \ln \det \left[1 - \tilde{W}(k) \tilde{C}(k) \right] \right\}. \end{aligned} \tag{2.68}$$

The excess chemical potential μ^{ex} is defined by [48]

$$\mu_a^{ex} = \int_0^1 d\zeta' \left\langle \varepsilon_{uv} \right\rangle_{\zeta'} = \sum_{\alpha\beta b} \rho_\beta^{(b)} \int d^3 r \int_0^1 d\zeta' \frac{\partial U_{\alpha\beta}^{(ab)}}{\partial \zeta'} g_{\alpha\beta}^{(ab)}(r;\zeta') \,, \tag{2.69}$$

where ζ' is the coupling parameter of the interaction $\langle \varepsilon_{uv} \rangle_{\zeta'}$ between a single molecule of type a with all the molecules in the system. Morita and Hiroike [46] derived μ^{ex} for simple fluids and Lue generalized it to the case of interaction site fluids [49]. When the HNC closure is used, the expression from Ref. [49] reduces to the expression obtained by Singer and Chandler [47]

$$\beta\mu^{ex} = \rho \sum_{\alpha,\beta} \int d^3 r \left(\frac{1}{2} h_{\alpha\beta}^2(r) - \frac{1}{2} h_{\alpha\beta}(r) c_{\alpha\beta}(r) - c_{\alpha\beta}(r) \right). \tag{2.70}$$

The compressibility κ_T can be computed from any site–site distribution function $g_{\alpha\beta}$ as [1]

$$\rho k_B T \kappa_T = 1 + 4\pi\rho \int_0^\infty \left(g_{\alpha\beta}(r) - 1 \right) r^2 dr \tag{2.71}$$

The dielectric constant of SSOZ is given by [1]

$$\varepsilon = 1 + 3 y_D. \tag{2.72}$$

One sees from Eq. (2.72) that the dielectric constant in SSOZ theory does not depend of the site–site correlation functions.

2.4 Previous applications of the Integral Equation Theories

The OZ equation complemented by different integral equations has been applied for various simple systems, namely the hard sphere fluid [50, 51], the charged hard sphere fluid [52-54], the hard–core Yukawa fluid [55-57], the mixtures of hard spheres [58-60] and the mixtures of charged hard spheres [40]. The OZ equation in conjunction with the HNC closure has been used to study dipolar hard–spheres [36], [61], hard ellipsoids [62], spherocylinders [63], homonuclear dumbells [64, 65], as well as realistic fluids such as hydrogen chloride [66], sulphor dioxide [67], acetone [15, 39], chloroforme [15], acetonitrile [39], formamide [68], N–methylformamide [68] and dimethyl–formamide [68]. Systems consisting of hard dumbbells [69-73], charged hard dumbells [14, 74] and the mixtures [75] of the two were studied within the SSOZ formalism. Previous application of the SSOZ equation include as well the study of polar and quadrupolar fluids [76], water [16], methanol [77], sulphor dioxide [78], liquid oxygen [79], nitrogen [79, 80], bromine [79], benzene [80], polymer fluids [81, 82], aqueous solutions of salt [83, 84], aqueous solutions of alkanes [85], aqueous solutions of monoatomic solutes [86], aqueous

solutions of methanol [87] and aqueous solutions of *t*–butanol [88]. Generally, it has been shown that the integral equation theories are successful in describing the structure and thermodynamics of simple fluids and their mixtures, but for associated liquids they do not provide a correct description.

2.5 Conclusion

Integral equation theories are a valuable tool for studying structural and thermodynamic properties of model and realistic liquids, since they do not suffer from constraints such as statistical errors in the thermodynamic limit and restrictions concerning size of the system and simulation times. The OZ equation is the basic equation for all these theories and is accompanied by another integral equation called the closure equation. This pair of equations is solved for two unknown functions, the pair correlation function $h(r)$ and the direct correlation function $c(r)$, whose relation is exact in the OZ equation, but not in any of the existing closures. Diagrammatically proper closure contains an infinite sum over integrals involving higher order direct correlation functions, or the so called bridge function. The definition of the direct correlation function includes computing the second functional derivative of the excess free energy and one cannot build these many–body correlations intuitively, so the approximation of the bridge functions seems to be an appropriate way to solve this problem. Unfortunately, all known approximations fail to some extent in describing the dense liquid region and one can hardly expect that improving the closure would lead to a general accurate description of realistic fluids. In the case of liquid water at ambient conditions for example, all of these theoretical approaches never produce the tetrahedral structure of the liquid and the origin of this inefficiency is the lack of the many–body correlations. In view of this inherent drawback of integral equations, it is absolutely required to perform experiments or simulations and fit the theoretical input.

CHAPTER 3

Application of the Site–Site Ornstein Zernike Equation to pure liquids

3.1 Introduction

Direct computation of the bridge functions in the closure relation turned out to be an impossible task, for they are defined through functional derivatives of the excess free–energy. Hence, approximation was desperately needed in the closure to make the theory operational. Unfortunately, none of these approximations gave results in par of simulations, especially for the dense liquid region [13]. Two papers have previously explored the possibility of improving RISM by computing the missing bridge diagrams. First, Lue and Blackstein [19] have explicitly computed the first few diagrams of the bridge function for water at ambient conditions, but this has not affected much the accuracy of results. It has been found previously that the addition of the first few bridge diagrams does not improve the results significantly [89]. This is because the diagrammatic expansion has an alternate sign and one might need more diagrams in the expansion. In the second work, Roux and co–workers [90] have fitted the bridge function using computer simulation data for solvation of polar and non–polar molecules in water. Functional form of the bridge was implemented and it needs careful adjustment depending on the type of solute in water. Kolafa and co–workers [91] inverted the OZ equation and computed the bridge functions from the simulation data for a system of hard spheres. However, this has never been done so far for realistic fluids. In this Chapter, I present a procedure for extracting the bridge functions from Molecular Dynamics simulations. This procedure consists of inverting the SSOZ equation and taking the radial distribution functions from simulations as a first guess in the iteration process. I also compute the bulk thermodynamic properties of each system and compare the results with and without the bridge functions.

3.2 Models, simulation and computational details

Molecular Dynamics simulations have been conducted in this work for several realistic liquids, namely: acetone (CH_3OCCH_3), carbon–tetrachloride (CCl_4), formamide (CH_3NO), methanol (CH_3OH), ethanol ($CH_3\ CH_2OH$), 1–propanol ($CH_3\ CH_2\ CH_2OH$) and water (H_2O). In all cases, NPT ensemble was treated at ambient conditions of pressure p = 1 atm and temperature T = 300 K using the DLPOLY–2 code [92]. Number of particles in the simulation box, N = 2048 and the size of the box, $L \approx 30$ Å. Berendsen thermostat and barostat were used, with relaxation times of 0.1 ps and 0.5 ps, respectively. Equations of motion were integrated with Δt = 2 fs.

The interaction potential between any site a on molecule 1 and site b on molecule 2 at distance r is given by

$$U_{ab}(r) = U_{ab}^{LJ}(r) + U_{ab}^{Coul}(r) = 4\varepsilon_{ab}\left[\left(\frac{\sigma_{ab}}{r}\right)^{12} - \left(\frac{\sigma_{ab}}{r}\right)^{6}\right] + \frac{q_a q_b}{4\pi\varepsilon_0 r}. \qquad (4.1)$$

The first term in the sum accounts for the Lennard–Jones (LJ) potential and the second corresponds to the electrostatic interaction due to the partial charges on molecules. LJ interaction was truncated at one quarter of the box size and Ewald summation was used for the Coulomb interaction. The LJ parameters for different types of sites were calculated according to the usual Lorentz–Berthelot rules, $\varepsilon_{ab} = \sqrt{\varepsilon_{aa}\varepsilon_{bb}}$ and $\sigma_{ab} = (\sigma_{aa} + \sigma_{bb})/2$. OPLS (Optimized Potential for Liquid Simulations) force field was used to model interactions in acetone [93] and carbon–tetrachloride [94]. Cordeiro [95] force field was used for formamide. Alcohols were modelled with OPLS [96] and water with SPC/E [97] (Simple Point Charge Extended) force field. Force field parameters are given in Tables 3.1, 3.2 and 3.3.

Table 3.1: SPC/E force field for water

SPC/E [97]		
	O	*H*
$\varepsilon(kJ/mol)$	0.65	0.0
$\sigma(\AA)$	3.165	0.0
$q(e)$	– 0.8476	0.4238

Table 3.2: OPLS force field for alcohols.

	Methanol OPLS [96]			Ethanol OPLS [96]			
	H	*O*	CH_3	*H*	*O*	CH_2	CH_3
$\varepsilon(kJ/mol)$	0.0	0.71131	0.86612	0.0	0.71355	0.33823	0.86887
$\sigma(\AA)$	0.0	3.071	3.775	0.0	3.071	3.85	3.775
$q(e)$	0.435	–0.7	0.265	0.435	–0.7	0.265	0.0

	1–Propanol OPLS [96]				
	H	*O*	$CH_2^{(1)}$	$CH_2^{(2)}$	CH_3
$\varepsilon(kJ/mol)$	0.0	0.71131	0.33823	0.33823	0.86887
$\sigma(\AA)$	0.0	3.071	3.85	3.85	3.775
$q(e)$	0.435	–0.7	0.265	0.0	0.0

Table 3.3: Force fields for acetone, formamide and carbon–tetrachloride.

	Acetone OPLS [93]		
	C	*O*	CH_3
$\varepsilon(kJ/mol)$	0.440	0.879	0.67
$\sigma(\AA)$	3.75	2.96	3.91
$q(e)$	0.3	–0.424	0.062

	Formamide [95]				
	H	*O*	*C*	*N*	*H*
$\varepsilon(kJ/mol)$	0.1591	0.8792	0.4396	0.7117	0.0
$\sigma(\AA)$	2.75	0.296	0.375	0.325	0.0
$q(e)$	0.12	–0.46	0.34	–0.83	0.415

	Carbon–tetrachloride [94]	
	C	*Cl*
$\varepsilon(kJ/mol)$	0.1016	0.2034
$\sigma(\AA)$	4.6	3.5
$q(e)$	0.0	0.0

As described in Section 2.2 of Chapter 2, the SSOZ equation can be written in terms of matrices as

$$\widetilde{H} = \widetilde{W}\,\widetilde{C}\,\widetilde{W} + \rho\,\widetilde{W}\,\widetilde{C}\,\widetilde{H}\,, \tag{3.2}$$

where the matrix elements are defined by $\widetilde{H} = \left[\widetilde{h}_{ab}(k)\right]$, $\widetilde{C} = \left[\widetilde{c}_{ab}(k)\right]$, $\widetilde{W} = \left[\widetilde{w}_{ab}(k)\right]$ and $\rho = N/V$ is the number density of the system. This equation can be rewritten in order to express the $\widetilde{H}$ matrix as:

$$\widetilde{H} = (1 - \rho\widetilde{W}\,\widetilde{C})^{-1}\,\widetilde{W}\,\widetilde{C}\,\widetilde{W} \tag{3.3}$$

and solved by iteration once an initial guess for the direct correlation function matrix $\widetilde{C}$ is provided. It is solved on a grid of 2048 points with a step of $\Delta r = 0.02\sigma_W$, where $\sigma_W = 3.165$ Å is the diameter of the water molecule. The initial guess for the $c_{ab}^{(0)}(r)$ in the SSOZ equation are the Mayer functions of the form:

$$c_{ab}^{(0)}(r) = \exp(-\beta U_{ab}(r)) - 1\,, \tag{3.4}$$

where $U_{ab}(r) = U_{ab}^{LJ}(r)$ and $\beta = 1/k_B T$. The long range Coulomb part $U_{ab}^{Coul}(r)$ is removed from the interactions and fitted by an error function following the method of Ng [98]. This is done because the long–range part of the direct correlation function has the following decay:

$$c_{ab}(r) \to -\frac{\beta\, q_a q_b}{r}, \text{ as } r \to \infty. \tag{3.5}$$

The Fourier transform $\widetilde{c}_{ab}(k)$ in the small–*k* limit assumes the form of:

$$\widetilde{c}_{ab}(k) \to -\frac{4\pi\beta\, q_a q_b}{k^2}, \text{ as } k \to 0\,. \tag{3.6}$$

The divergence of the $\widetilde{c}_{ab}(k)$ is therefore avoided by fitting the $c_{ab}(r)$ with the function:

$$f_{ab}(r) = -\frac{\beta\, q_a q_b}{r} erf(\alpha r) \tag{3.7}$$

whose Fourier transform $\widetilde{f}_{ab}(k)$ is given by:

$$\tilde{f}_{ab}(k) = -\frac{4\pi\beta q_a q_b}{k^2}\exp\left(-\frac{k^2}{4\alpha^2}\right), \tag{3.8}$$

where α is an arbitrary parameter. In this work, α is set to 1 Å^{-1}.

The short–range part of the direct correlation function is transformed into *k*–space and the fitted part is added to obtain the functions $\tilde{c}_{ab}^{(0)}(k)$.

Further, functions $\tilde{h}_{ab}^{(0)}(k)$ are calculated from (3.3) and functions $h_{ab}^{(0)}(r)$ are calculated by inverse Fourier transform. After setting $\eta_{ab}^{(0)}(r) = h_{ab}^{(0)}(r) - c_{ab}^{(0)}(r)$, one of the three closure relations is used. The HNC closure is defined as:

$$g_{ab}(r) = \exp(-\beta U_{ab}(r) + \eta_{ab}(r)), \tag{3.9}$$

the PY closure as:

$$g_{ab}(r) = \exp(-\beta U_{ab}(r))(1 + \eta_{ab}(r)). \tag{3.10}$$

and the KH closure as:

$$\begin{aligned} g_{ab}(r) &= \exp(-\beta U_{ab}(r) + \eta_{ab}(r)), \quad -\beta U_{ab}(r) + \eta_{ab}(r) < 0 \\ &= \exp(-\beta U_{ab}(r))(1 + \eta_{ab}(r)), \quad -\beta U_{ab}(r) + \eta_{ab}(r) \geq 0 \end{aligned} \tag{3.11}$$

Solution of one of the three closures gives the function $g_{ab}^{(0)}(r)$ from which the next solution $c_{ab}^{(1)}(r)$ is built as $c_{ab}^{(1)}(r) = g_{ab}^{(0)}(r) - \eta_{ab}^{(0)}(r) - 1$.

The first and the second solution, $c_{ab}^{(0)}(r)$ and $c_{ab}^{(1)}(r)$, are then mixed to get the new input for the SSOZ equation:

$$c_{ab}^{new} = \lambda c_{ab}^{(0)}(r) + (1 - \lambda) c_{ab}^{(1)}(r), \tag{3.12}$$

where $\lambda = 0.75$. This is the so called Picard iteration [99]. Both closures are used and the data are compared with the simulation results.

Next problem concerns the inversion of the SSOZ equation in order to extract the exact bridge functions, $b_{ab}(r)$ from simulation results. The SSOZ equation can be rewritten in a computationally more convenient way by expressing the $\tilde{C}$ matrix explicitly as

$$\tilde{C} = \tilde{W}^{-1} \tilde{H} \tilde{S}^{-1} \quad , \tag{3.13}$$

where $\tilde{S}$ is the matrix of the total structure factor which contains both intra–molecular and inter–molecular contribution, $\tilde{S} = \tilde{W} + \rho \tilde{H}$ and $\tilde{H}$ is the total correlation function matrix obtained from Molecular Dynamics calculations.

One can determine the $\tilde{c}_{ab}(k=0)$ values analytically, since $\tilde{w}_{ab}(k=0) = 1$ and $\tilde{h}_{ab}(k=0) = h$. The first relation comes from the definition of $\tilde{w}_{ab}(k)$ and the second is a property of site–site functions whose integral cannot depend on the choice of any particular pair of sites [1]. Direct use of Eq. (3.13) would cause numerical problems in the small–*k* region, because the $\tilde{W}$ matrix is singular at *k* = 0. However, this can be easily treated by interpolating the $\tilde{c}_{ab}(k)$ in the small–*k* limit.

Before applying this procedure, one should remove the divergent $1/k^2$ Coulomb part from the $\tilde{c}_{ab}(k)$. In other words, the extrapolation from exact *k* = 0 values holds only for the short range part of the site–site direct correlation functions. Once an initial guess of the functions $c_{ab}(r)$ is found one can compute the corresponding site–site bridge functions, $b_{ab}(r)$ using the following equation:

$$b_{ab}(r) = \ln(g_{ab}(r)) - h_{ab}(r) - c_{ab}(r) + \beta U_{ab}(r) \ . \tag{3.14}$$

The bridge functions $b_{ab}(r)$ can only be computed for distances above the hard core region, because the interaction term $\beta U_{ab}(r)$ creates divergence inside the core and also $\ln(g_{ab}(r))$ diverges for $r \le \sigma_{ab}$. In practice, the values of $b_{ab}(r)$ are kept until distances for which the effective interaction $\beta U_{ab}(r) - b_{ab}(r)$ is smaller than 10. This allows for sufficiently repulsive steepness to consider $g_{ab}(r) \approx 0$ beneath this distance. With the current estimate of $b_{ab}(r)$ it is necessary to converge the solution, since the functions $\tilde{c}_{ab}(k)$ have been extrapolated in the small–*k* limit. This is done by adding the bridge function to the first guess $c_{ab}^{(0)}(r)$ and calculating the next solution $c_{ab}^{(1)}(r)$ through the closure relation. After a few iterations, one obtains the solution that is consistent with $b_{ab}(r)$ and remains very close to $g_{ab}(r)$ from simulations. Details of this procedure are illustrated in Fig. 3.1 in the case of pure methanol.

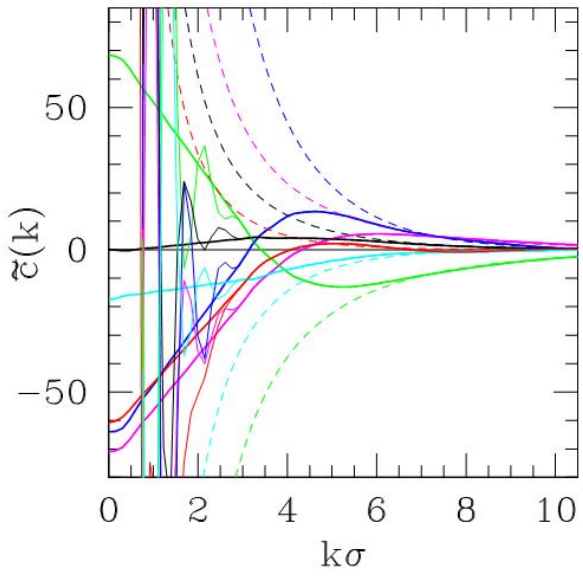

Figure 3.1: Details of $\tilde{c}_{ab}(k)$ in the small–*k* range. Thick lines show the final converged values of . Thin lines show as initially computed by direct inversion of the SSOZ equation. Dashed lines show the Coulomb part of . Magenta is for H–H, green for O–H, blue for O–O, black for H–M, cyan for O–M and red for M–M.

We examine in detail the inversion of the SSOZ equation for the case of an asymmetric dumbbell molecule. The $\tilde{W}$ matrix is defined as

$$\tilde{W} = \begin{bmatrix} 1 & \tilde{w}_{12}(k) \\ \tilde{w}_{12}(k) & 1 \end{bmatrix}, \tag{3.15}$$

where $\tilde{w}_{12}(k) = \sin(kl_{12})/kl_{12}$ and l_{12} is the distance between sites 1 and 2 of the same molecule. Similarly, the $\tilde{H}$ matrix is defined as

$$\tilde{H} = \begin{bmatrix} \tilde{h}_{11}(k) & \tilde{h}_{12}(k) \\ \tilde{h}_{12}(k) & \tilde{h}_{22}(k) \end{bmatrix}. \tag{3.16}$$

The exact result for the $\tilde{C}$ matrix from Eq. (3.13) is

$$\tilde{C} = \frac{1}{D}\begin{bmatrix} \tilde{h}_{11}\tilde{s}_{22} + \tilde{w}_{12}\tilde{h}_{22} - \tilde{h}_{12}(\tilde{w}_{12} + \tilde{s}_{12}) & \tilde{w}_{12}\tilde{h}_{11}\tilde{s}_{11} - \tilde{h}_{12}(1 + \tilde{w}_{12}\tilde{s}_{12}) + \tilde{w}_{12}\tilde{h}_{22} \\ \tilde{w}_{12}\tilde{h}_{11}\tilde{s}_{11} - \tilde{h}_{12}(1 + \tilde{w}_{12}\tilde{s}_{12}) + \tilde{w}_{12}\tilde{h}_{22} & \tilde{h}_{22}\tilde{s}_{11} + \tilde{w}_{12}\tilde{h}_{11} - \tilde{h}_{12}(\tilde{w}_{12} + \tilde{s}_{12}) \end{bmatrix}, \tag{3.17}$$

where $D=(\tilde{w}_{12}^2-1)(\tilde{s}_{11}\tilde{s}_{22}-\tilde{s}_{12}^2)$. It is clear that divergence occurs at k = 0, because of $\tilde{w}_{12}(k=0)=1$, but since it is known that the $\tilde{C}$ matrix is well behaved everywhere, there must be some cancellation of the divergent part by the terms inside the matrix in Eq. (3.17).

Indeed, all the three projections should be equal at k = 0, $\tilde{h}_{11}(k=0)=\tilde{h}_{12}(k=0)=\tilde{h}_{22}(k=0)=h$. This comes from the fact that in the k = 0 limit the choice of the molecular centre does not depend on any particular site [1]. In simulations, there are often small differences in these values so the inversion procedure becomes impossible to handle. Writing the three projections exactly at k = 0, gives from Eq. (3.17) the $\tilde{c}_{ab}(k=0)$ values as

$$\tilde{c}_{11}(k=0)=\tilde{c}_{12}(k=0)=\tilde{c}_{22}(k=0)=h/\left[(1+\tilde{w}_{12})(1+\tilde{w}_{12}+2\rho h)\right]. \tag{3.18}$$

Obviously, these values are not divergent at k = 0. Higher order matrices would appear if one considers more complex molecules, but the overall procedure remains the same.

3.3 Site–site distribution functions from Molecular Dynamics

Site–site distribution functions $g_{ab}(r)$, where indices a and b denote site a and b on molecules 1 and 2, respectively, where calculated using Molecular Dynamics simulations. It is a known fact that the periodic boundary conditions in simulations are implemented to mimic the properties of an infinitely large system. This simple mathematical solution works well for many simple systems, though one can question whether the fact that the observed system is confined in a box with finite size distorts the correlations among molecules and gives a different microscopic picture.

This is indeed true, as it was proved decades ago by Lebowitz and Percus [100] who found that in a one component system of N particles in a finite sized box, the tail of $g_{ab}(r)$ in the limit $r\to\infty$ does not go to 1, but is shifted by a certain correction factor

$$g_{ab}(r)=1-\frac{1}{N}\left(\frac{\partial\rho}{\partial\beta P}\right)_T=1-\frac{1}{N\rho}\left(\frac{\partial\rho}{\partial\beta\mu}\right)_T, \tag{3.19}$$

where ρ is the number density, P is the pressure, T is the absolute temperature, μ is the chemical potential and $\beta = 1/k_B T$. If the system is indeed infinite, $N \to \infty$ and $g_{ab}(r) = 1$.

For a finite system, the correction term is not negligible and has an impact on the Fourier transform of $g_{ab}(r)$ in the small *k*–limit.

A way to handle this problem is to shift the tail of $g_{ab}(r)$ by adding an asymptote, while keeping the short–range behaviour of the pair correlation function intact, as it was proposed in the work of Perera and co–workers [101]. Therefore, $g_{ab}(r)$ becomes:

$$g_{ab}^{new}(r) = g_{ab}(r)\left[1 + \frac{\alpha}{2}\left(1 + \tanh\left(\frac{r - r_c}{\kappa}\right)\right)\right], \tag{3.20}$$

where α is the shift of the function, r_c and κ are the position in $g_{ab}(r)$ where the shift starts and the smoothness of the shift, respectively. The α–value is found empirically by looking at the original data and setting them to unity. The other two parameters were also found empirically: $r_c \approx 2r_\tau$, where r_τ is the mean period of the first oscillations in $g_{ab}(r)$ and in all cases, κ is set to 1. While the behaviour of the tail of $g_{ab}(r)$ doesn't seem to be important for the structural features of a liquid, where the first two peaks correspond to the first and the second neighbours shell, it influences extremely the value of the KBI. As mentioned before, the tail of $g_{ab}(r)$ determines the $k \to 0$ limit of the structure factor $S_{ab}(k)$, which is defined as

$$S_{ab}(k) = 1 + \rho(\tilde{g}_{ab}(r) - 1). \tag{3.22}$$

Fig. 3.2 illustrates the procedure for $g_{OO}(r)$ and $g_{MM}(r)$ in OPLS methanol. As seen in the upper panel, there is not much of a difference between the corrected and the uncorrected pair correlation function. However, the difference between two sets of running KBI values is striking. Moreover, KBI without the correction don't converge to a single value. Instead, they have a descending, unphysical behaviour after about 10 Å. Adding the asymptotic shift solves the problem, which is proved by the convergence of the KBI close to the experimental value. The structure factor $S_{ab}(k)$ is shown in the lower panel, for which the difference arises in the $k \to 0$ limit.

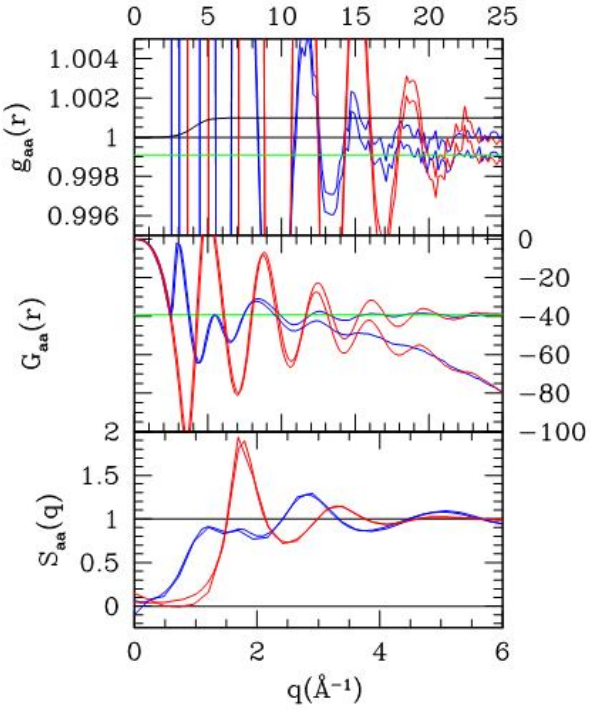

Figure 3.2: Detail of the long–range correction of $g_{ab}(r)$ of OPLS methanol. Top panel shows the zoom on the tail of the two site–site functions between oxygen atoms (O) $g_{OO}(r)$ in blue and the methyl site (M = CH_3) $g_{MM}(r)$ in red. The green line is the wrong asymptote as produced by the simulations. Thick lines for corrected data and thinner lines for uncorrected data. The thick black line shows the switching function used in Eq. (3.3). Middle panel shows $G_{OO}(r)$ and $G_{MM}(r)$ with same line and color convention as in top panel. Bottom panel shows the structure factors $S_{OO}(k)$ and $S_{MM}(k)$ with the same conventions as in top panel.

3.4 Application of RISM and the SSOZ inversion procedure to realistic liquids

3.4.1 Acetone

The site–site distribution functions $g_{ab}(r)$ and the site–site structure factors $S_{ab}(k)$ are computed for liquid acetone and presented in Figs 3.3 and 3.4, respectively. As it was previously observed by Fries and co–workers [15], the agreement between the simulation data and the data produced by both HNC and PY closure is quite good. Some discrepancies are seen in $g_{OO}(r)$, $g_{OC}(r)$ and $g_{OM}(r)$ which indicates that here the higher order correlations should not be neglected. I show

also the bridge–corrected results which are practically indistinguishable from the simulation results. The agreement is even better in *k*–space (Fig. 3.4) where the differences in the short range *r*–region correspond to the ones in the long range part. Instead of showing the bridge functions directly, I plot the effective interaction, $U_{ab}^{eff}(r)$ to see how much the bridge term affects the original interaction. $U_{ab}^{eff}(r)$ is defined as

$$U_{ab}^{eff}(r) = U_{ab}^{SR}(r) - b_{ab}(r) \ , \tag{3.23}$$

where $U_{ab}^{SR}(r)$ is the short–range part of the interaction potential. In all cases, $U_{ab}^{SR}(r)$ is simply the LJ potential. Fig. 3.5 [102] shows some of the LJ interactions and effective interactions between sites. It is seen that adding the bridge function in the interaction makes it more attractive in the short range. For the CC interaction, there is a nontrivial repulsive part in the long range. Having in mind all these subtle, but significant changes of the potential, I conclude that the local order in acetone is affected by the many–body effects.

Molar excess internal energy U^{exc}, molar excess Helmholtz free energy A^{exc}, molar excess chemical potential μ^{exc} and the isothermal compressibility κ_{T} were calculated and listed in Table 3.4. Results obtained by solving the HNC equation with the bridge functions and without them are compared with available simulation and experimental data. Experimental molar excess internal energies for all substances were calculated by using the formula $\Delta H_{vap} = -U^{exc} + RT$ and the enthalpies of vaporization ΔH_{vap} were found in Ref. [103]. Unfortunately, I have not found experimental and simulation data for molar excess Helmholtz energy and molar excess chemical potential, but it is worth reporting the calculated results. As one can see from Table 3.4, molar excess internal energies U^{exc} are close to the simulation result and the result is even slightly better from the HNC equation, than with the addition of the bridge.

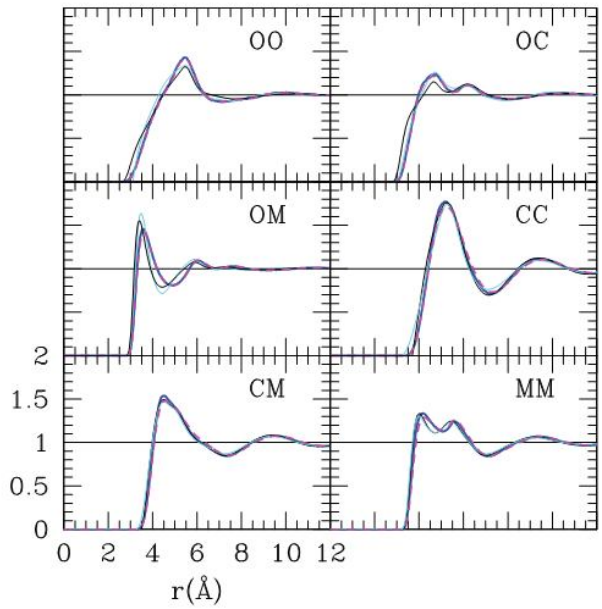

Figure 3.3: Site–site distribution functions $g_{ab}(r)$ for OPLS acetone at ambient conditions. Thick blue line is for current results (RISM+BRIDGE obtained from simulation), dashed magenta line for simulation results, cyan thin line for PY, and black thin line for HNC.

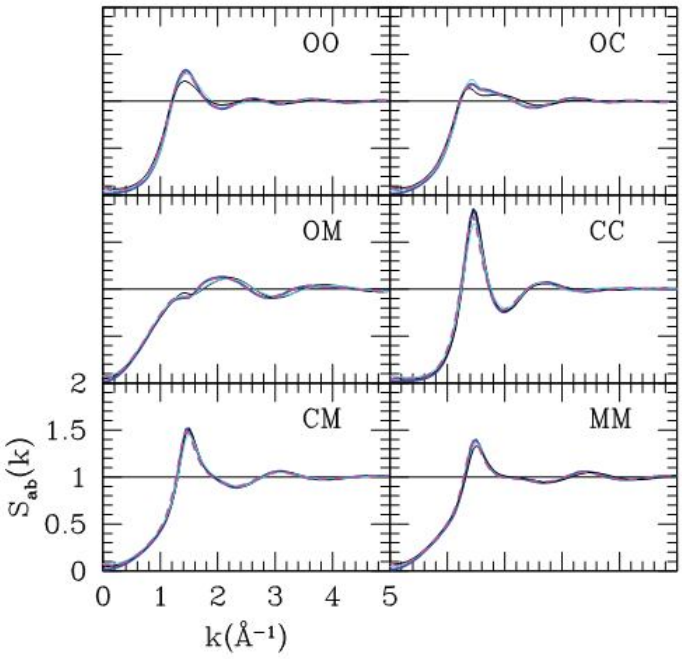

Figure 3.4: Structure factors $S_{ab}(k)$ for the OPLS acetone at ambient conditions. Line and colour conventions are as in Fig. 3.3.

If one looks closely, one will see small differences between the HNC+Bridge results and the simulation results. These small differences get blown by the r^2 – term in the integral of the molar excess internal energy (Eq. 2.63) and we get a difference in the final result. We cannot get exactly the same results, because of the iterative procedure of the SSOZ equation which treats the exact solution from simulations only as a first guess. This is another confirmation that the SSOZ equation is an approximate equation. The HNC result shows good agreement with the simulations, despite of the fact that the structure does not exactly match the simulations. This effect is found in results of the MSA equation for electrolyte solutions were it is seen that the excess internal energies match the simulations, while the structure is poorly reproduced [1]. A significant difference is seen between the values of the isothermal compressibility κ_T from HNC and HNC+Bridge equation. The fact that this value is more than 5 times bigger as calculated from the HNC equation suggests that the addition of the bridge function diminishes the density fluctuations. The values of A^{exc} from the HNC and the HNC+Bridge calculation differ by 12 kJ/mol, while the value of μ^{exc} from the HNC+Bridge calculation is about 7 times the one from the HNC calculation.

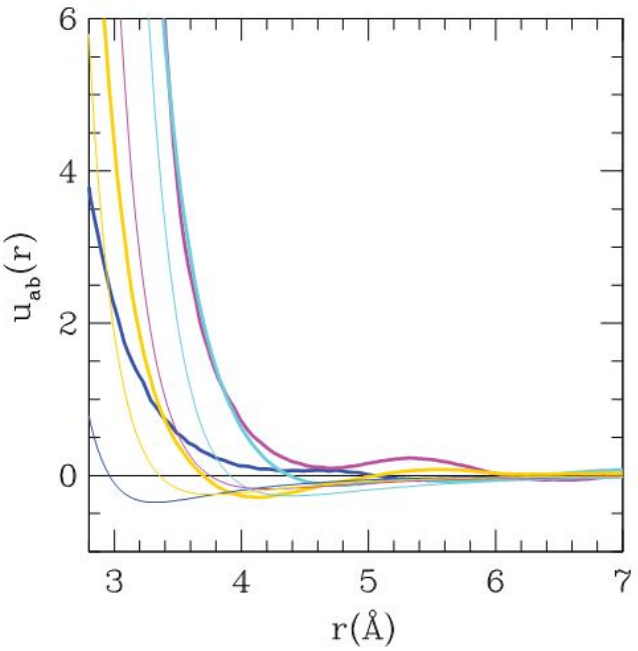

Figure 3.5: LJ interaction (thin curves) and bridge corrected effective interactions (thick curves) for OPLS acetone. Blue lines for O–O, magenta for C–C, cyan for M–M, and gold for O–C.

Table 3.4: Predictions of some thermodynamic properties of the SSOZ equation for liquid acetone at ambient conditions and comparison with experimental values and simulations.

	SSOZ (HNC+Bridge)	SSOZ (HNC)	Expt.	MD
$-U^{exc}$ (kJ/mol)	**29.16**	**26.6**	28.5 [103]	24.32
$-A^{exc}$ (kJ/mol)	**41.28**	**29.47**	–	–
μ^{exc} (kJ/mol)	**133.74**	**17.60**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**42.16**	**225.50**	124 [103]	–

3.4.2 Carbon tetrachloride

CCl_4 has been previously studied integral equation theories [104, 105]. HNC and PY closures were used to solve the SSOZ equation and the results were compared with the simulation data. Bridges from simulations were extracted and added inside the HNC closure as well. Figs. 3.6 and 3.7 show $g_{ab}(r)$ and $S_{ab}(k)$, respectively. As in the case of acetone, both HNC and PY theory reproduce the correlations very close to the ones from simulations, while the bridge–corrected results fit exactly the simulations.

Fig. 3.8 shows the site–site LJ and the effective interaction for CCl_4. The bridge term has pushed the attractive interaction part further in *r*–space and has grown a weak repulsion, especially noticeable in the Cl–Cl potential. These differences could hardly be reproduced by changing the original correlation function in any way, which has been done in many closures such as the Verlet bridge [106], the Rogers–Young closure (RY) [30], the Hybridized Mean Spherical Approximation (HMSA) [31] closure, the Martynov–Sarkisov closure (MS) [29] and the Kovalenko–Hirata (KH) [27] closure. The weak brakes of the slope in the repulsive part of the effective interaction are caused by small incertitude in the simulation pair correlation function in the region of strong repulsion. One has to take the $\ln(g_{ab})$ term in order to calculate the bridge function which then affects the effective potential.

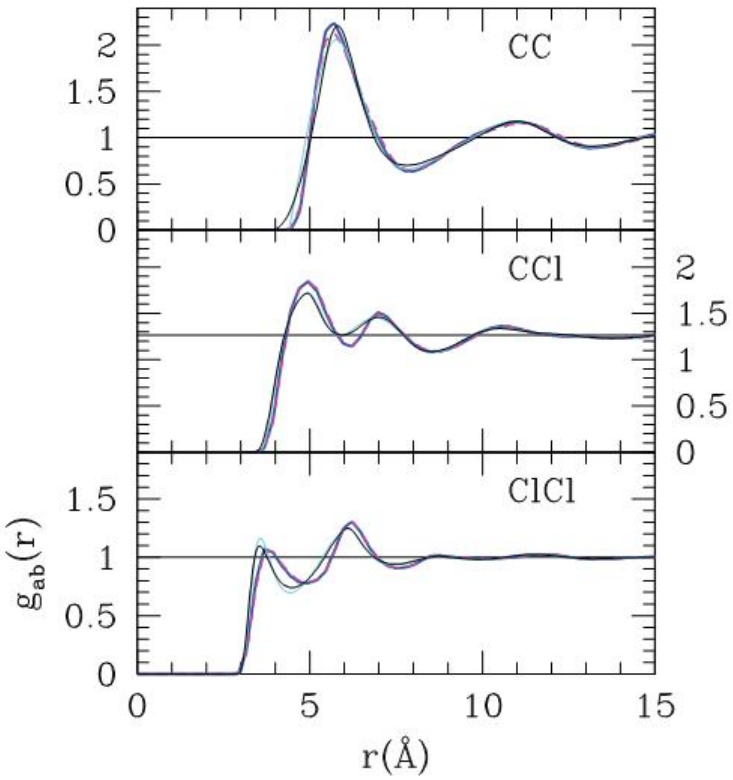

Figure 3.6: Site–site distribution functions $g_{ab}(r)$ for OPLS carbon–tetrachloride at ambient conditions. Line convention: thick blue (RISM+Bridge), dashed magenta (simulation), cyan thin (PY) and black thin (HNC).

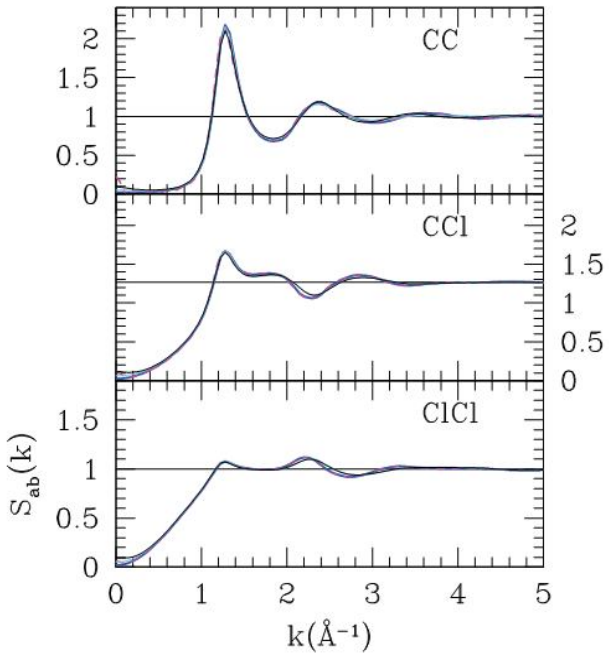

Figure 3.7: Structure factors $S_{ab}(k)$ for OPLS carbon–tetrachloride at ambient conditions. Line and colour conventions are as in Fig. 3.6.

Thermodynamic properties are calculated for liquid carbon–tetrachloride using the bridge corrected HNC closure and the HNC closure. Results are compared with the available simulation data and experimental values in Table 3.5. As in the case of liquids acetone, the value of U^{exc} is lower from the HNC+Bridge calculation than the HNC calculation. The isothermal compressibility κ_T from the HNC equation is by a factor of 4 bigger than the value from the HNC+Bridge calculation. This suggests that the density fluctuations are lowered by the addition of the missing bridge diagrams. The calculated values of A^{exc} and μ^{exc} are different. The HNC result for A^{exc} is about 14 kJ/mol more positive than the HNC+Bridge result and the HNC result for μ^{exc} is about 5 times lower than the HNC+Bridge result.

Table 3.5: Predictions of some thermodynamic properties of the SSOZ equation for liquid CCl_4 at ambient conditions and comparison with experimental values and simulations.

	SSOZ (HNC+Bridge)	SSOZ (HNC)	Expt.	MD
$-U^{exc}$ (kJ/mol)	**29.34**	**28.45**	29.94 [103]	27.32
$-A^{exc}$ (kJ/mol)	**46.70**	**33.18**	–	–
μ^{exc} (kJ/mol)	**124.89**	**26.10**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**72.15**	**289.06**	105 [103]	–

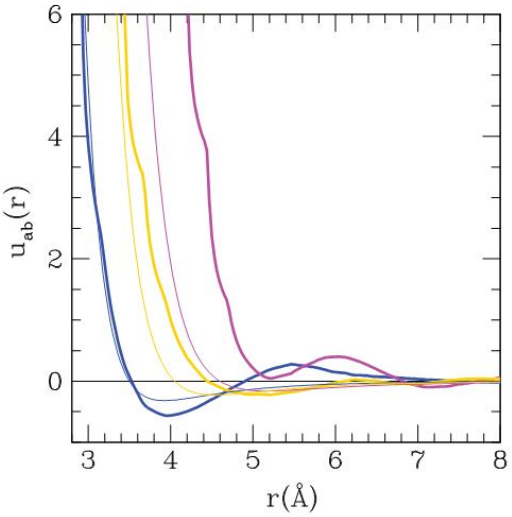

Figure 3.8: LJ interaction (thin curves) and bridge corrected effective interactions (thick curves) for OPLS CCl_4. Blue lines for C–C, magenta for Cl–Cl, and gold for C–Cl.

3.4.3 Formamide

Formamide is a hydrogen bonding liquid similar to water in such a way that it forms micro–emulsions with appropriate surfactants, although to a lesser extent than water [107]. Figs. 3.9 and 3.10 show site–site distribution functions $g_{OH}(r), g_{OO}(r), g_{NN}(r), g_{CC}(r)$ in r– and k–space, respectively. As seen in these figures, bridge corrected HNC theory agrees quite well with the simulations, while HNC theory alone, could not give the proper structure of formamide. The agreement is better in k–space, as seen in Fig. 3.10. The PY closure clearly does not capture the structure of formamide properly, as it is observed in Fig. 3.9. Thermodynamic properties are calculated and listed in Table 3.6 for liquid formamide at ambient conditions. The two calculated values for U^{exc} are quite close, but 10 kJ/mol more positive than the simulation value. The values of A^{exc} are different by 6 kJ/mol and the HNC value of μ^{res} is two times lower than with the addition of the bridge. Density fluctuations are lower after the addition of the bridge which is seen by comparing the values of κ_T.

Table 3.6: Predictions of some thermodynamic properties of the SSOZ equation for liquid formamide at ambient conditions and comparison with experimental values and simulations.

	SSOZ (HNC+Bridge)	SSOZ (HNC)	Expt.	MD
$-U^{exc}$ (kJ/mol)	**45.77**	**45.97**	57.65 [103]	55.75
$-A^{exc}$ (kJ/mol)	**31.48**	**37.73**	–	–
μ^{res} (kJ/mol)	**81.13**	**41.35**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**35.05**	**52.81**	41.1 [108]	–

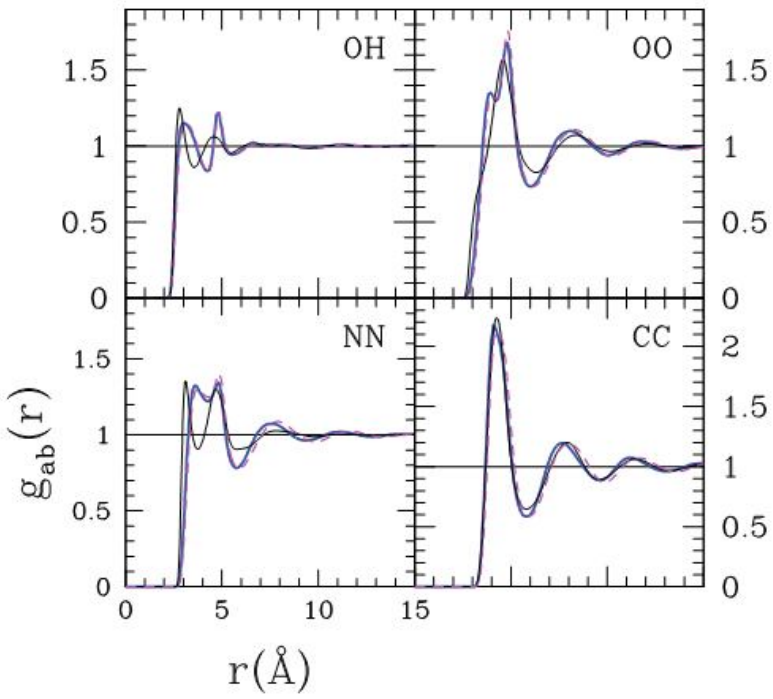

Figure 3.9: Site–site distribution functions $g_{ab}(r)$ in r–space for Cordeiro formamide. Line and colour conventions are as in Fig. 3.6.

.

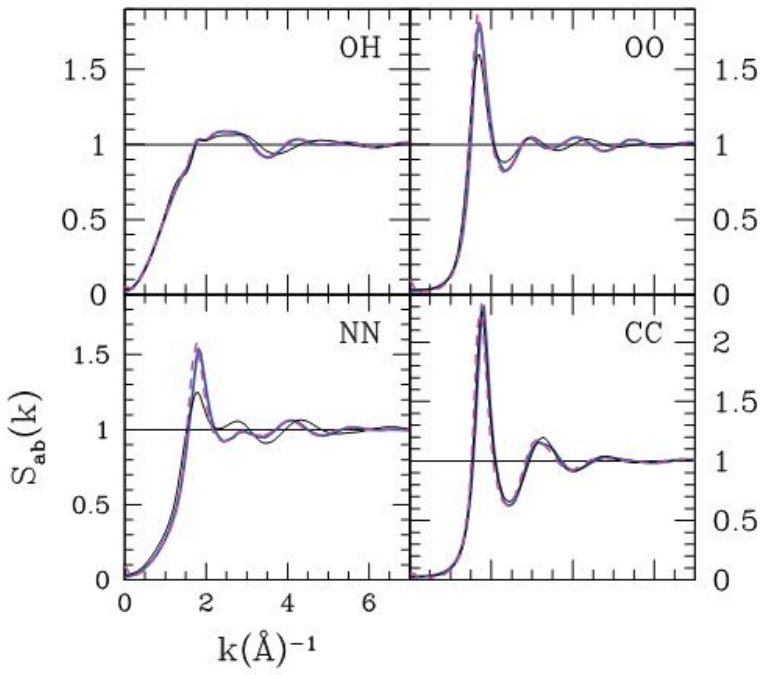

Figure 3.10: Site–site distribution functions $S_{ab}(k)$ in k–space for Cordeiro formamide. Line and colour conventions are as in Fig. 3.6.

3.4.4 Methanol, ethanol and 1–propanol

Alcohols are hydrogen bonding liquids that form very specific clusters, such as chains and loops [109]. Since this type of clustering is highly non–trivial, one expects that the many–body correlations will play an important role in the short–range structure of these liquids. Therefore, the bridge function is supposed to give special features in the pair correlation function. RISM theory has been previously used in the study of alcohols [77, 105, 110]. When handling the pair interactions involving the H–atom only the repulsive $1/r^{12}$ part was kept, with parameters $\sigma_{HH} = 1\text{Å},\ \varepsilon/k_B = 50K$ and Lorentz–Berthelot mixing rules. The HH LJ–interaction was set to zero. This trick was used by Pettit and Rossky [16] and by Lue and Blackstein [19], in order to numerically avoid the O–H Coulomb attraction core–collapse. Fig. 3.11 shows three specific site–site distribution functions $g_{OH}(r)$, $g_{OO}(r)$ and $g_{MM}(r)$ for methanol, ethanol and 1–propanol. It is seen that the results of the KH closure and the HNC closure are almost indistinguishable for site–site distribution functions of all three liquids. Bridge corrected data match the simulations perfectly, as seen in the example of methanol (upper left panel) and ethanol (upper right panel). Fig. 3.12 shows the same set of functions as Fig. 3.11, but in *k*–space. It is seen that all the bridge corrected $S_{ab}(k)$ for methanol (upper left panel of Fig. 3.12) match the simulation results. However, this is not true for ethanol O–O correlations (upper right panel of Fig. 3.12) and 1–propanol correlations in the small–*k* region (lower panel of Fig.3.12). Thermodynamic properties are calculated and listed in Table 3.7 for liquid methanol, ethanol and 1–propanol at ambient conditions. For methanol, the value of U^{exc} from the HNC+Bridge calculation matches the simulation value, but for ethanol and 1–propanol is about 10% more positive. The values of A^{exc} from the HNC+Bridge calculation are lower than the ones from the HNC equation and the values of μ^{exc} are bigger than the HNC results for all alcohols. Isothermal compressibility, κ_T from the bridge calculations is closer to the experimental value than the HNC result for all three alcohols.

Table 3.7: Predictions of some thermodynamic properties of the SSOZ equation for liquid methanol, ethanol and 1–propanol at ambient conditions and comparison with experimental values and simulations.

	SSOZ (HNC+Bridge)	SSOZ (HNC)	Expt.	MD
Methanol				
$-U^{exc}$ (kJ/mol)	**32.14**	**36.86**	35.78 [103]	32.00
$-A^{exc}$ (kJ/mol)	**39.15**	**36.44**	–	–
μ^{exc} (kJ/mol)	**45.20**	**–9.77**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**43.04**	**178.82**	121.4 [103]	–
Ethanol				
$-U^{exc}$ (kJ/mol)	**30.76**	**32.16**	36.06 [103]	34.44
$-A^{exc}$ (kJ/mol)	**70.33**	**36.70**	–	–
μ^{exc} (kJ/mol)	**6.43**	**–1.26**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**88.34**	**347.67**	111.9 [103]	–
1–Propanol				
$-U^{exc}$ (kJ/mol)	**32.41**	**32.83**	44.95 [103]	36.10
$-A^{exc}$ (kJ/mol)	**47.60**	**38.31**	–	–
μ^{exc} (kJ/mol)	**46.82**	**7.13**	–	–
$10^{11}\kappa_T$ (Pa^{-1})	**109.98**	**383.47**	84.3 [103]	–

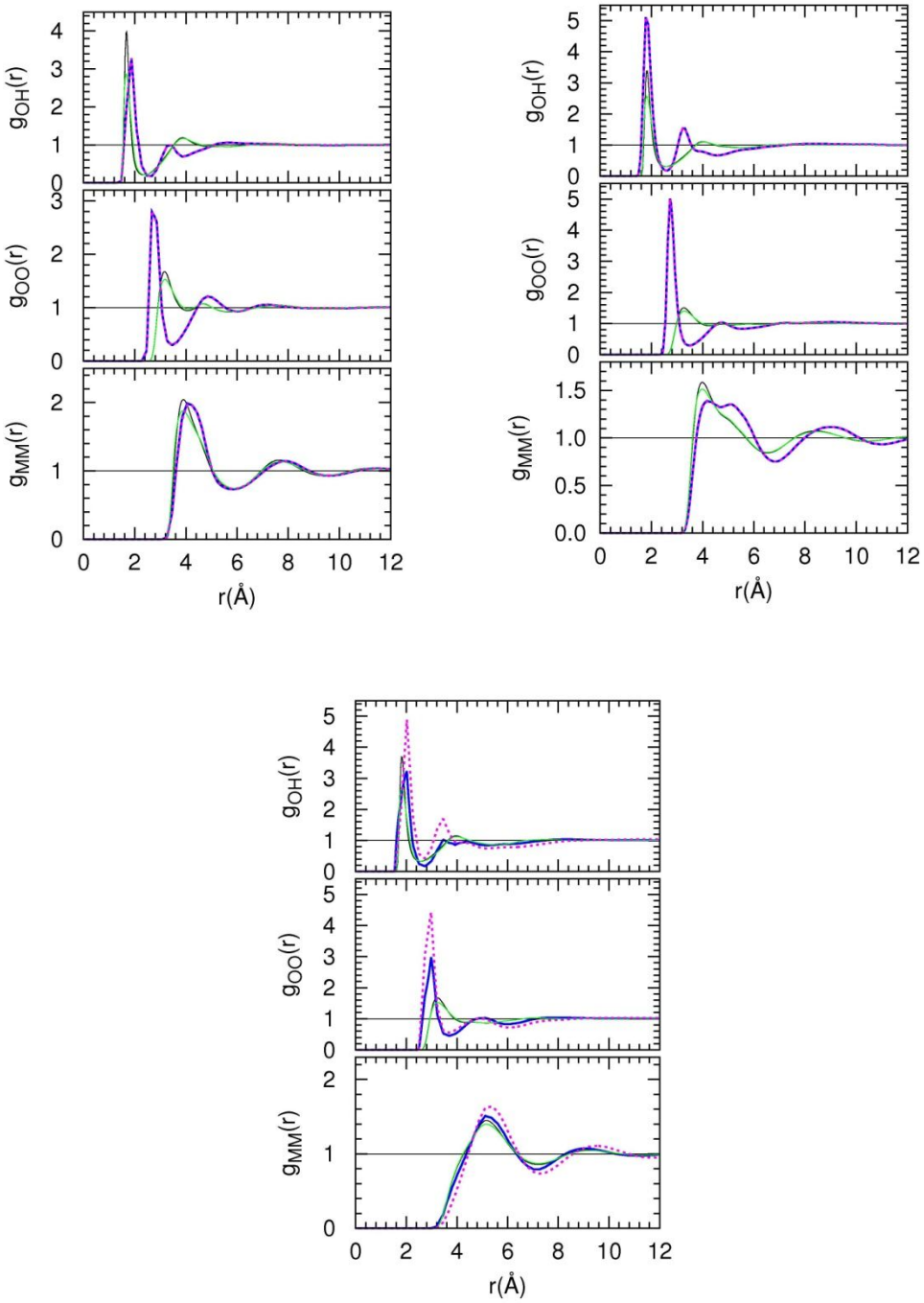

Figure 3.11: Oxygen–hydrogen, oxygen–oxygen and methyl–methyl correlation functions in *r*–space for OPLS methanol (upper left panel), ethanol (upper right panel) and 1–propanol (lower panel). Thick blue line is for current results (RISM+BRIDGE obtained from simulation), dashed magenta line for simulation results, green thin line for KH, and black thin line for HNC closure.

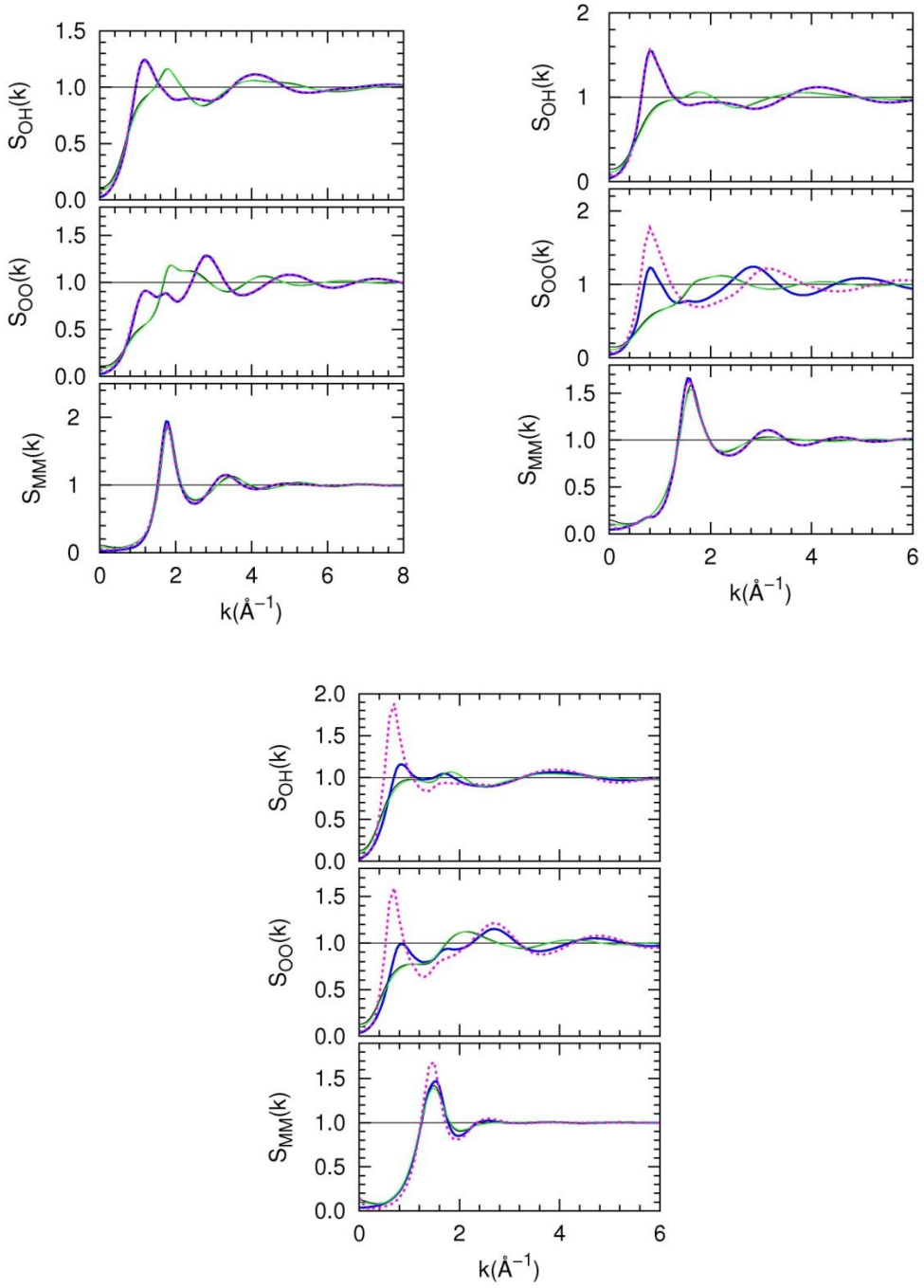

Figure 3.12: Oxygen–hydrogen, oxygen–oxygen and methyl–methyl correlation functions in *k*–space for OPLS methanol, ethanol and 1–propanol. Panel, line and colour conventions are as in Fig. 3.11.

As seen from these figures, the quality of the bridge calculation technique is perfect for methanol, quite good for ethanol, but significantly low for 1–propanol. Obviously, it depends on the number of methyl groups in the molecule, indicating that this procedure is not appropriate for higher alcohols, hence larger bio–molecules. There is absolutely no logical reason why this

procedure works well for small molecules and poorly for bigger ones, unless one takes into account the approximations within the SSOZ equation. Fig. 3.13 shows O–O, O–H and M–M effective interactions for three alcohols. It is seen that the hydrogen bonding O–O effective interaction develops a stepwise repulsion in all three alcohols, indicating that the inclusion of the bridge function produces a core–softening effect, very similar to the one found in the case of water [111].

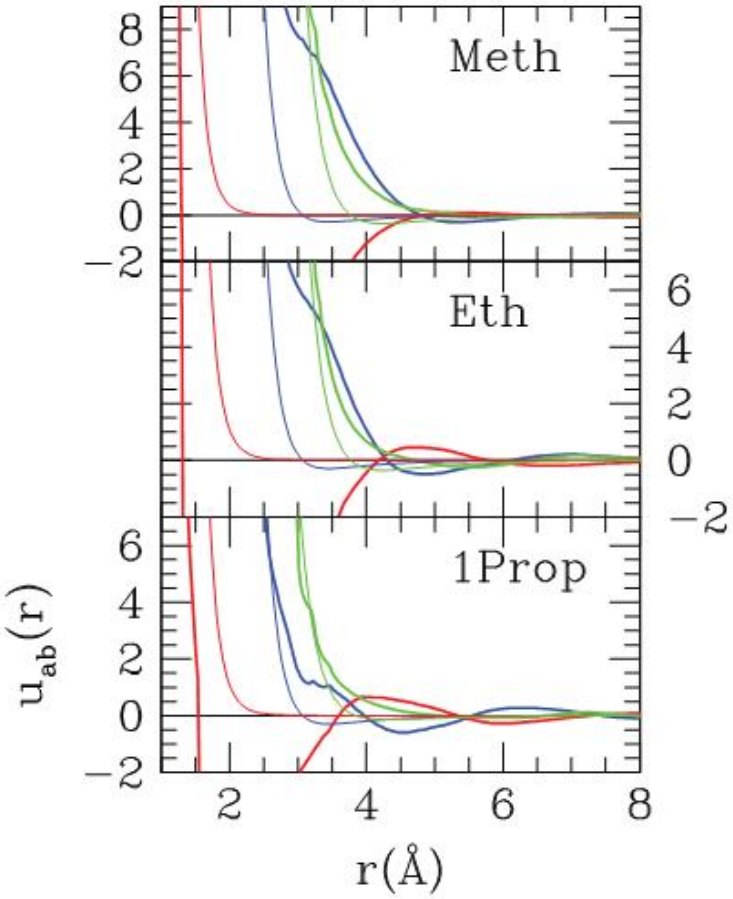

Figure 3.13: LJ interaction (thin curves) and bridge corrected effective interactions (thick curves) for OPLS alcohols. Blue lines for O–O, red for O–H and green for M–M.

.

3.4.6 Water

Water has always been an important test ground for both force field models and theories. Generally, it is known that none of the integral equation theories are able to produce accurate pair correlation functions for this liquid [14, 17-21]. As in the case of alcohols, only the $1/r^{12}$ part in the OH interaction was kept, with parameters $\sigma_{HH} = 1\text{Å}$ and $\varepsilon / k_B = 5K$. The HH interaction was set to zero. Table 3.8 presents the thermodynamic properties of water as calculated by RISM and compared to literature data. The value of U^{exc} from the HNC equation is 2 kJ/mol more negative than the simulation value and the value from the HNC+Bridge calculation is about 5 kJ/mol more negative. Since the SSOZ equation contains approximations, the result after the addition of the bridge functions does not match the simulations. There are also different simulation values of U^{exc}, because the long–range interactions between molecules are treated differently [19]. The value of excess free energy A^{exc} from the HNC equation is closer to the simulation value than the value from the HNC+Bridge calculation. The theory predicts wrong values of the excess chemical potential μ^{exc}, when compared to the experimental value, although the HNC+Bridge calculation produce a result with a positive sign. The isothermal compressibility κ_T from the HNC+Bridge calculation produces a value only 10% higher than the experimental value and it is closer to it then the HNC value.

Figs. 3.14 and 3.15 show all the site–site distribution functions $g_{ab}(r)$ and $S_{ab}(k)$ in direct and *k*–space, respectively. It is seen that the bridge modified results agree almost perfectly with the simulations. Both HNC and PY theory do not capture properly the structure of water. Small discrepancies are seen in Fig. 3.15 in the small–*k* part of $S_{OO}(k)$. This illustrates the incapability of the SSOZ equation to find the correct long–range behaviour of the O–O correlation in water. Fig. 3.16 shows the effective interaction for all three pairs of sites. As for alcohols, the O–O effective potential grows a repulsive part and has a rather unpredictable oscillatory structure. The O–H and the H–H effective potentials have also a non–trivial behaviour.

Table 3.8: Predictions of some thermodynamic properties of the SSOZ equation for liquid methanol, ethanol and 1–propanol at ambient conditions and comparison with experimental values and simulations.

	SSOZ (HNC+Bridge)	SSOZ (HNC)	Expt.	MD
$-U^{exc}$ (kJ/mol)	**37.73**	**44.26**	41.42 [112]	42.91
		41.42		42.59 [113]
		(46.44) [49]		37.7 [114]
				45.19 [114]
$-A^{exc}$ (kJ/mol)	**7.09**	**18.69**	24.02 [112]	23.01 (26.77)
		20.63 (25.86) [49]		[115]
μ^{exc} (kJ/mol)	**0.74**	**–13.10**	24.46 [116]	–
		–10.45 (–15.77) [49]		
$10^{11}\kappa_T$ (Pa^{-1})	**48.09**	**54.72**	44.65 [117]	26.65 [118]
		54.87 [49]		

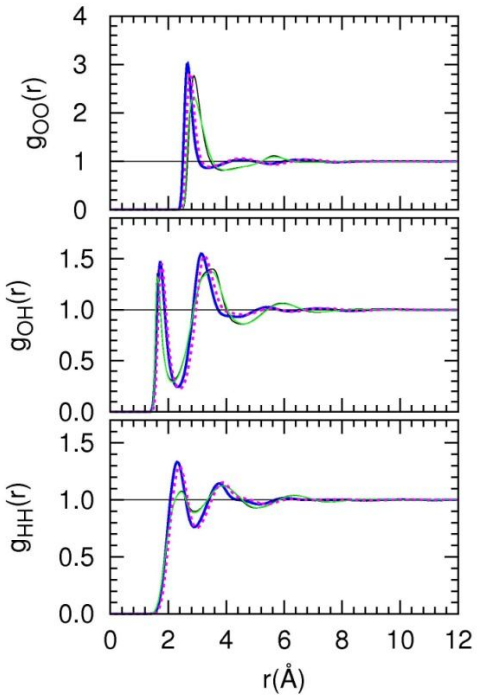

Figure 3.14: Site–site distribution functions $g_{ab}(r)$ for SPC/E water at ambient conditions. Thick blue (current results), dashed magenta (simulations), thin black (HNC), thin green (KH), thin black dashed (PY).

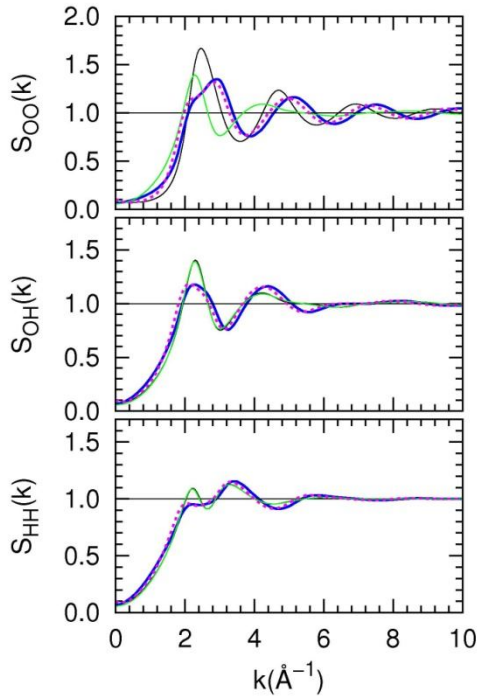

Figure 3.15: Site–site distribution functions $S_{ab}(k)$ in k–space for SPC/E water at ambient conditions. Line and colour conventions are as in Fig. 3.14.

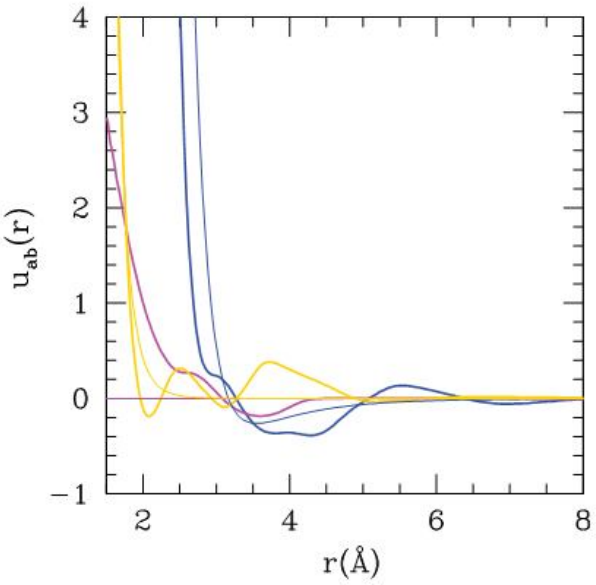

Figure 3.16: LJ interaction (thin curves) and bridge corrected effective interactions (thick curves) for SPC/E water. Blue lines for O–O, gold for O–H and magenta for H–H.

3.5 Conclusion

In this Chapter, I presented a procedure for extracting the exact bridge functions, which have been missing so far in the closure relation, from the pair correlation functions from Molecular Dynamics simulations. Since they contain higher order correlations, these functions are extremely important in the description of the structure of liquids. Both HNC and PY theory were compared with simulation data and it is found that the agreement is better for liquids with a simple, Lennard–Jones–like structure (acetone, carbon–tetrachloride), than for H–bonded liquids (alcohols, water).

This methodology allowed calculating new, bridge–corrected pair correlation functions for several liquids. In the case of small H–bonding liquids, an excellent agreement was found between the calculated data and the simulations. However, this is not the case when studying larger molecules, such as 1–propanol, where the bridge function alone cannot cure the insufficiency made by the approximations in the SSOZ equation itself. This throws a serious doubt in the ability of integral equations to serve as a tool in describing large molecular liquids or even bio–molecular systems [28]. The only way to solve this problem is to correct the SSOZ equation rather than to improve the closure relation, as it was attempted here. One could avoid using the SSOZ equation and use the diagrammatically proper CSL equations [119, 120] instead, but these equations are more difficult to handle, since they split each of the bridge functions in 4

parts, in a physically non–intuitive way. Therefore, this procedure could not be easily implemented in the CSL equations.

CHAPTER 4

Application of the Site–Site Ornstein Zernike Equation to aqueous mixtures

4.1 Introduction

The structure of aqueous mixtures is known to possess a micro–heterogeneous nature, besides the density and concentration fluctuations that exist in any mixture. It is particularly challenging to find observables that will allow differentiating between these phenomena in order to describe the correct structural behaviour of these mixtures. Computer simulation studies have brought important insights about aqueous mixtures and have successfully complemented experimental measurements. However, there are systems, such as aqueous mixtures of higher alcohols for example, for which inherent problems of computer simulations, become a serious obstacle in detailed analysis [121]. Previous IE studies on realistic mixtures include the study of methanol–water mixture [87], TBA–water mixture [110] and methanol–carbon–tetrachloride mixture [105]. Tanaka, Walsh and Gubbins [87] used the SSOZ equation complemented by the HNC equation for aqueous methanol, where they renormalized the Coulomb interaction in the HNC equation, but did not explore the problem of micro–heterogeneity. Yoshida *et al.* [110] used the SSOZ equation with the KH closure to reproduce the short–range structure of the correlations in aqueous TBA mixtures over the whole concentration range. The SSOZ/HNC study of methanol–carbon–tetrachloride mixture [105] showed good agreement with the Molecular Dynamics results, with a better accuracy in description of the apolar liquid than the alcohol. On the other hand, the site–site structure factors of this system are seen to overestimate the simulation results. In this Chapter, the SSOZ equation is used with two closures, namely the BHNC and the KH closure, to study aqueous mixtures of methanol and "weak water". A new closure is introduced, named the BHNC closure, which uses the bridges from pure liquids in the HNC closure

belonging to the site–site correlations that correspond to those liquids. I compare the results of both closures against the simulations.

4.2 Details of the computational procedure

In Chapter 3, a procedure that extracts the exact bridge functions from the simulation results for a variety of pure liquids is proposed and solved the SSOZ equation with the closure containing all the missing many–body correlations. However, this methodology cannot be applied to aqueous mixtures since it requires very accurate pair radial distribution functions $g_{ij}(r)$ from simulations. The problem lies in the long–range behaviour of $g_{ij}(r)$, since it has been seen that, even in aqueous methanol, the asymptote of the correlations does not stabilize [122]. Fig. 4.1 from Ref. [122] shows all pairs of running Kirkwood–Buff integrals $G_{ij}(r)$ for aqueous methanol at x = 0.20 mole fraction of methanol. Simulation results of two systems, N = 2048 and N = 16 384 particles, are presented and also the results after the LP correction of the erroneous asymptote of the correlations. The differences between the results from simulation of larger and smaller system are striking, especially for methanol, which is the minority species. One observes apparent oscillations of $G_{ij}(r)$ in the long–range region of the larger system for all three pairs and the periodicity seems to be matched. These oscillations confirm the existence of methanol clusters and the matching periodicity means that methanol clusters leads to voids in water domains. Similar effects in the running $G_{ij}(r)$ appear in mixtures of x = 0.50 and x = 0.80 mole fraction of methanol. Clearly, aqueous methanol is a system that requires intensive computational effort in order to be described properly, so the exact bridge functions for this mixture remain unreachable.

An alternative way is introduced to improve the use of integral equations in the study of aqueous mixtures. The SSOZ equation is complemented by a new closure, called the BHNC closure, of the following form:

$$g_{ab}(r) = \exp\left[-\beta U_{ab}(r) + h_{ab}(r) - c_{ab}(r) + B_{ab}(r)\right], \tag{4.1}$$

where $B_{ab}(r)$ is the effective bridge function which is taken from that of the neat liquids and used only in conjunction to the site–site correlations related to the species component in the mixture.

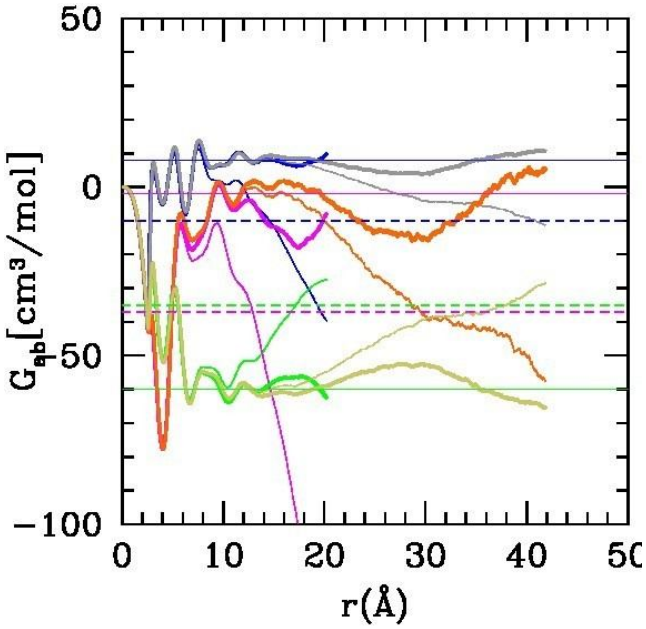

Figure 4.1: (from Ref. [122]) Running Kirkwood–Buff integrals $G_{ij}(r)$ for aqueous methanol at x = 0.20 mole fraction of methanol. Thin lines are uncorrected results and thick lines are corrected results. Water–Water pair: blue (N=2048) and grey (N=16 384). Cross pair: green (N=2048) and jade (N=16 384). Methanol–methanol pair: magenta (N=2048) and orange (N=16 384). Horizontal dashed lines are expt. results [172] with the same colour codes as N=2048 systems and full lines are actual predictions from the N=16 384 system.

If the bridge function in Eq. (4.1) is neglected, one recovers the usual HNC equation. Bridge functions $B_{ab}(r)$ between sites belonging to the same species are calculated as:

$$B_{ab}(r) = a_W (1-x) B_{ab}^{(W)}(r) + a_S B_{ab}^{(S)}, \tag{4.2}$$

where $B_{ab}^{(W)}$ is the site–site bridge function for pure water and $B_{ab}^{(S)}$ is the site–site bridge function for pure solute. In Eq. (4.1), x is the mole fraction of the solute and parameters a_W, a_S serve as scaling factors that allow optimising the numerical solutions. For aqueous methanol mixtures in the water–rich region, a_W is set to 1 and a_S is set to 0. For the equimolar mixture, a_W is chosen to be 0.50 and a_S is 0, while for the solute–rich region, $a_W = 0$ and $a_S = 0.90$.

For the water–"weak water" mixtures, the total bridge function of water is put for all concentrations. The cross bridge functions are absent from calculations for each system in this study.

The SSOZ equation is coupled by the BHNC closure and solved on a grid of 2048 points, both in *r* and in *k* space. The bridge function for pure liquids are used from our previous work on pure liquids [102], where the bridge functions is extracted from the simulation data. The procedure of solving the SSOZ equation is the following. First, a solution is obtained at a high temperature where all the integral equations produce numerical solutions. Then, the solution is used as a first guess in a standard iterative procedure [99]. The temperature is slowly decreased until room temperature is reached. When the solution for a given concentration x is obtained, it is used in the iteration process for another value of x. The whole concentration range is studied, from $x = 0$ to $x = 1$, with steps of 0.1. In addition, the SSOZ equation complemented by the KH closure is solved for these systems. This closure is often used for micro–heterogeneous systems [26, 28], since it offers numerical solution over the whole concentration range, contrary to the HNC closure, which tends to enhance the long–range correlations, loosing solution as a consequence. The HNC closure could not produce numerical solutions, except at very high temperatures, due to the increasing correlations, indicated by a rapid growth of site–site structure factors $S_{ab}(k)$ near $k = 0$. Previous HNC studies [123, 124] showed that this closure looses solution passed a certain point were structure factors reach some high value, but do not diverge. When the bridge functions from pure liquids were inserted in the HNC closure, numerical solutions over the whole concentration range were obtained.

4.3 Structural results of the SSOZ Equation for aqueous methanol mixtures

Fig. 4.2 shows the radial distribution functions $g_{ij}(r)$ between oxygen sites of water–water, water–methanol and methanol–methanol molecules for mixtures at $x = 0.20$, 0.50 and 0.80 mole fraction of methanol. Results from the SSOZ/HNC equation and the SSOZ/KH equation are compared against simulations. One can conclude from these figures that the addition of bridge functions from pure liquids inside the HNC closure produces a structure very similar to those from simulations, as far as the like–like correlations are concerned. This suggests that the correlations between the like–like species in the mixture are not so different from those of pure

liquids. In other words, this is an indirect proof for micro–segregation of species into domains. The KH closure, on the other hand, gives a solution in poor agreement with the simulations.

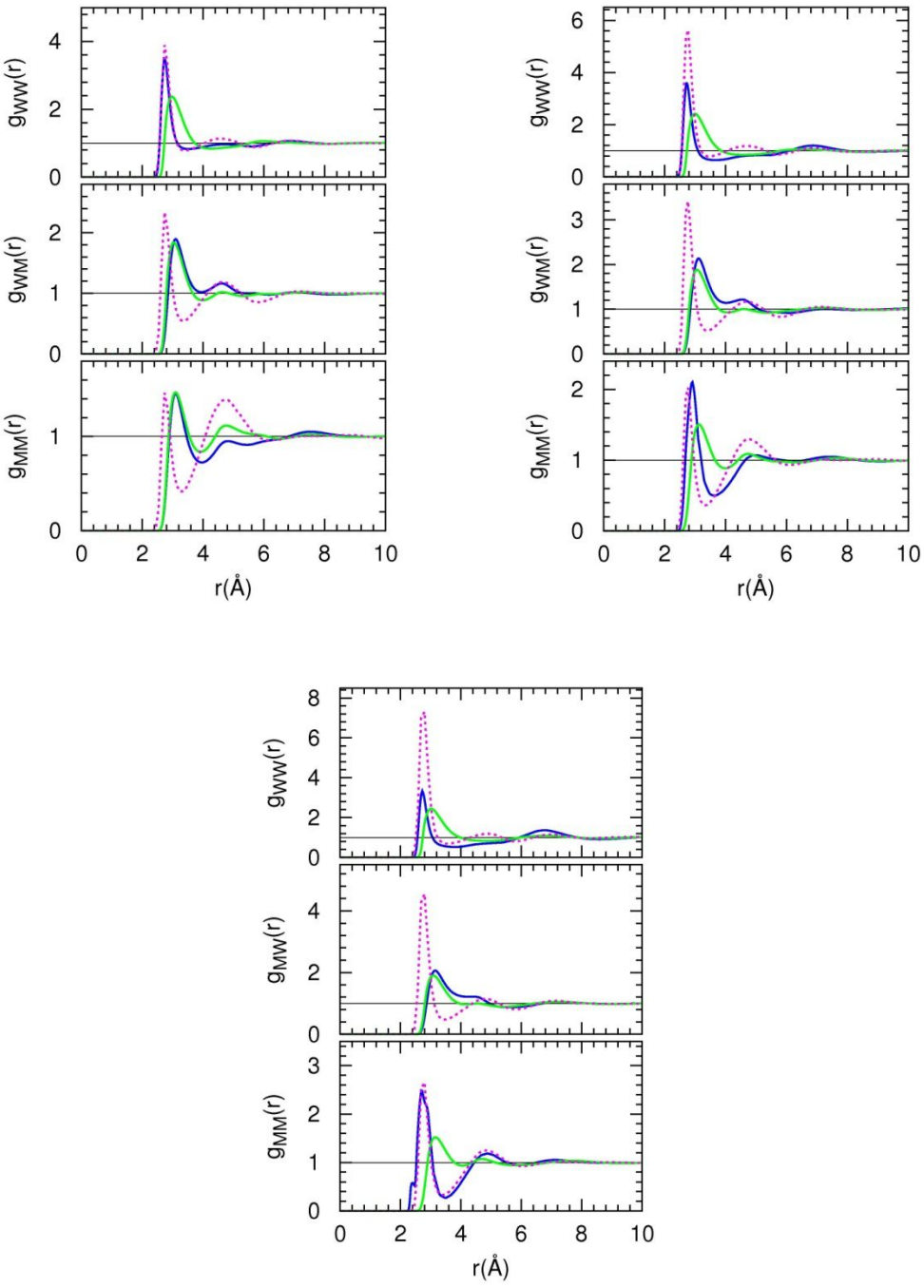

Figure 4.2: Oxygen–oxygen correlation functions for all three pairs of correlations for $x = 0.20$ mixture (upper left panel), $x = 0.50$ mixture (upper right panel) and $x = 0.80$ (lower middle panel). On each panel, upper inset is for water–water correlations, middle one is for cross correlations and lower one is for methanol–methanol correlations. Blue lines are BHNC results, green lines are KH results and dashed magenta lines are simulation results.

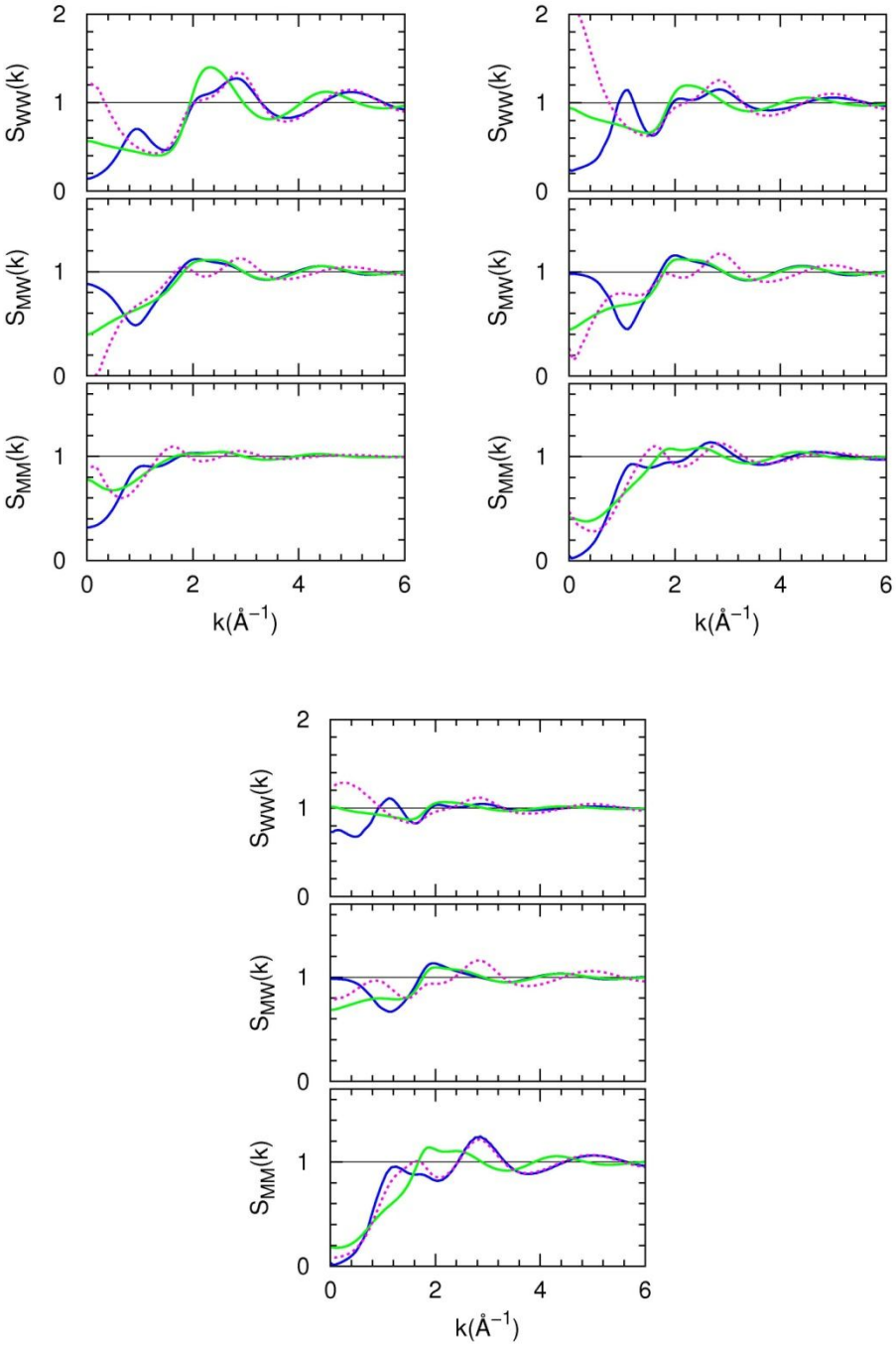

Figure 4.3: Oxygen–oxygen structure factors corresponding to the pair correlation functions in Fig. 4.2. Panel, line and colour convention is the same as in Fig. 4.2.

Fig. 4.5 shows the site–site structure factors $S_{ij}(k)$ that correspond to $g_{ij}(r)$ from Fig. 4.4, and as one can see, the long–range correlations are not properly described by the BHNC closure, though some pairs show better agreement with the simulations that the others. A closer look at the top panel of Fig. 4.5 for the x = 0.20 mixture, reveals that $S_{WW}(k)$ beyond k = 2 Å^{-1} actually

assumes the correct behaviour. One observes the same thing for $S_{MM}(k)$ in the middle panel of Fig. 4.6 that shows the results for x = 0.50 mixture and for $S_{MM}(k)$ in the bottom panel of Fig. 4.7 that concerns the x = 0.80 mixture. This is a direct consequence of the excellent agreement these correlations have with simulations in the real space. However, the BHNC results show a weak pre–peak at k = 1 Å^{-1} of all $S_{WW}(k)$, which is absent from both simulation and KH results, meaning that it promotes the existence of water aggregates size of about 6 Å.

4.4 Structural results of the SSOZ Equation for aqueous "weak water" mixtures

The SSOZ equation is solved in conjunction with the BHNC and the KH closure for a selected mixture of water and "weak water" model. A detailed simulation study of this aqueous mixture is conducted and results are discussed in Chapter 7. "Weak water" model has the same geometry as water, only the partial charges on sites are scaled by a factor λ = 4/5. It is found that this model solute is miscible with water in all concentrations and micro–heterogeneity is observed, which is similar to realistic aqueous mixtures. Bridges from pure SPC/E water are added to water–water correlations using Eq. (4.1) with parameter a_W set to one. Fig. 4.4 shows three different pairs of oxygen–oxygen correlation functions for mixtures with x = 0.20, 0.50 and 0.80 mole fraction of "weak water" compared with the simulation results. As seen in Fig. 4.4, the BHNC results show good agreement with the simulation results, especially concerning water–water correlations, where the position of the first peak is properly reproduced. However, the height of the first peak is underestimated, which is particularly observed in the x = 0.80 mixture. Clearly, the correlations between water molecules in the mixture where water is the minority species are stronger than those of the pure liquid. The KH closure is seen to produce results almost the same as BHNC for the cross and "weak water" correlations, while those for water are far from the correct structure. Fig. 4.5 shows the same functions as in Fig. 4.4, but seen in k–space. Similar to the case of aqueous methanol, BHNC closure does not predict the correct long–range correlations, as seen in the small–k behaviour of the site–site structure factors. The KH closure is seen to be slightly closer to simulation results in the small–k region and there is no weak pre–peak that one observes in $S_{WW}(k)$ from BHNC in all three mixtures. Knowing that the same effect is observed in the methanol–water mixture, one can conclude that adding the effective bridge from pure water to the water–water correlations promotes the self aggregation of water.

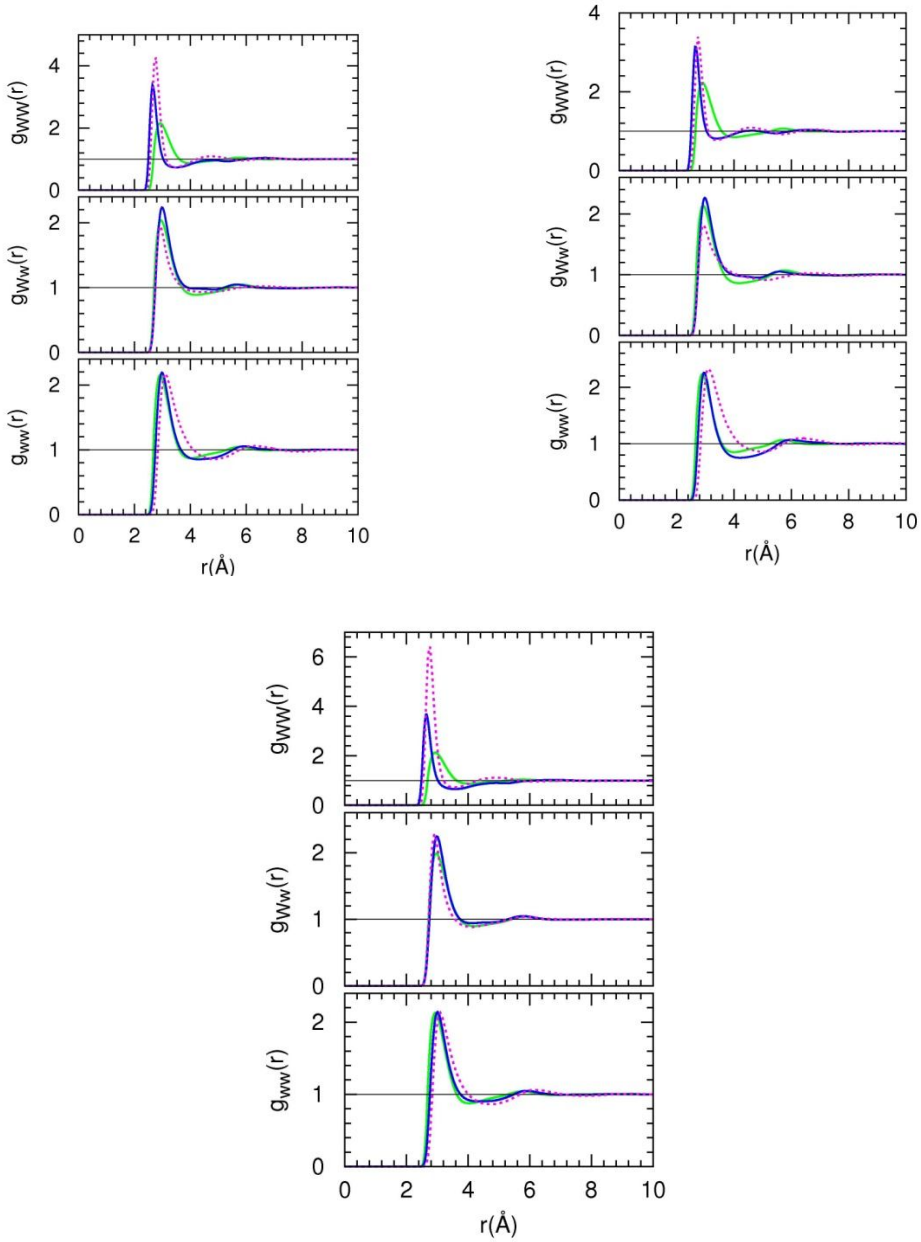

Figure 4.4: Oxygen–oxygen distribution functions for all three pairs of correlations for x = 0.20 mixture (upper left panel), x = 0.50 mixture (upper right panel) and x = 0.80 (lower middle panel). Panel, line and colour convention is the same as in Fig. 4.2.

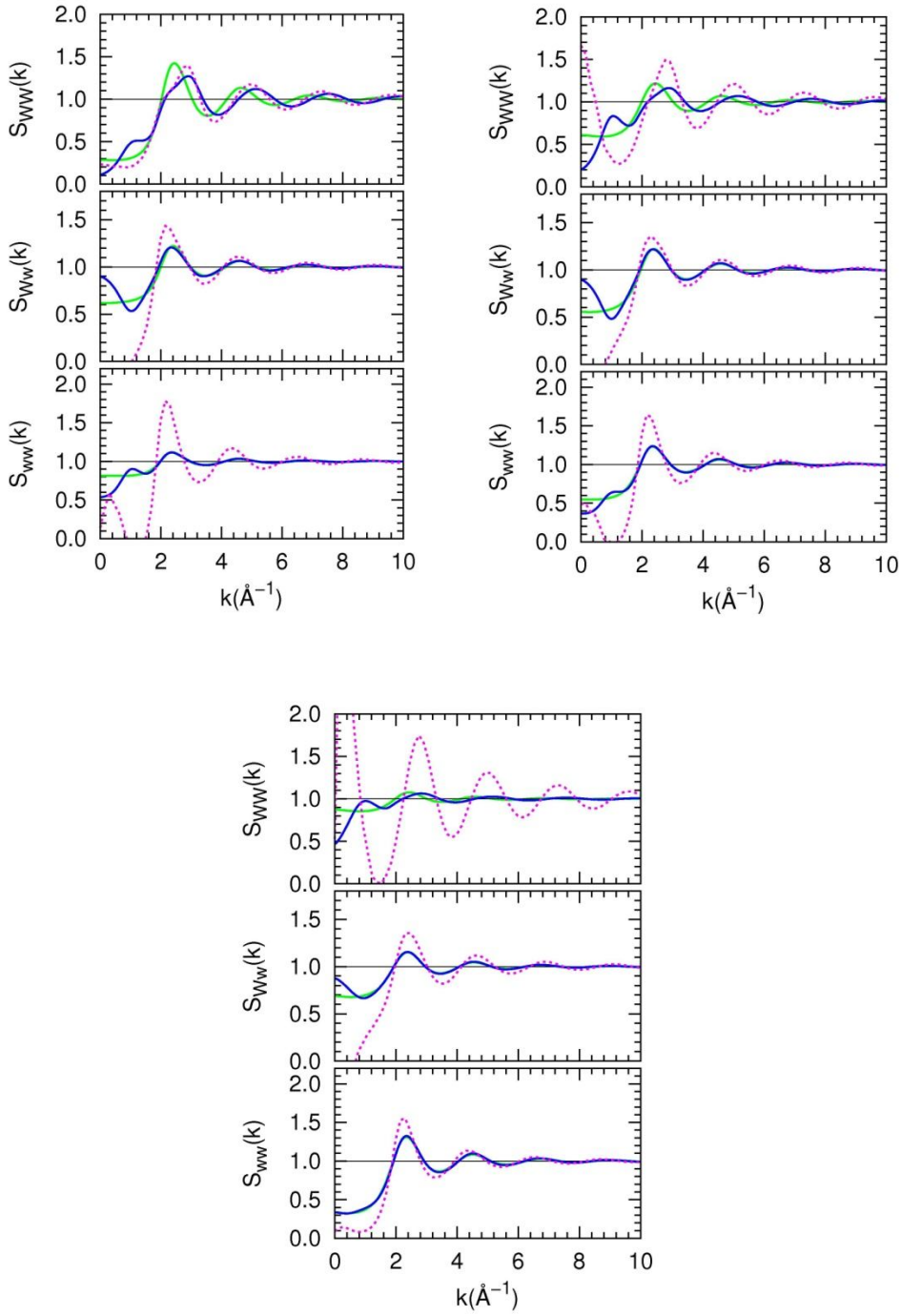

Figure 4.5: Oxygen–oxygen structure factors corresponding to the pair correlation functions in Fig. 4.2. Panel, line and colour convention is the same as in Fig. 4.2.

4.5 Conclusion

In the simulations of aqueous methanol mixtures [122], the tail of the radial distribution function does not stabilize to an asymptote, which indicates the existence of concentration fluctuations at temporal and spatial scales larger than the time of the simulation and the size of simulation box. This behaviour was already observed in simulations of higher alcohols such as TBA [121, 125] and now one faces the same difficulties in aqueous methanol. This indicates that there is a universal structural mechanism of aqueous mixtures and its origin certainly lies in the hydrogen–bonding interactions between molecules. Results of the study on aqueous mixtures of methanol and "weak water" have shown that the Integral Equation Theory does not describe the correct structural behaviour of these mixtures. The KH closure is useful in the sense that it provides a numerical solution over the whole range of concentrations, but since it is a diagrammatically incorrect closure, the results are in poor agreement with the simulations. The results of the new closure, called the BHNC closure, are accurate in the short–range part of the correlations, but fails in describing the long–range structure, which is seen in the small–k part of the site–site structure factors, $S_{ab}(k)$. The HNC closure does not have numerical solutions for aqueous mixtures, because the correlations in this closure grow too rapidly so the value of $S_{ab}(k)$ at $k = 0$ becomes too large. The BHNC closure suppresses the growth of correlations by including the bridge functions from neat liquids for those site–site correlations that belong to these species components in the mixture. Unfortunately, the BHNC closure does not produce the correct $S_{ab}(k = 0)$ behaviour. While the simulations show a raise of $S_{ab}(k = 0)$ indicating concentration fluctuations, the BHNC closure misses this feature, which suggests the importance of the diagrammatical links that are not included in the HNC closure. It is important to note that the HNC and the KH closure produce almost the same results for neat water and neat methanol, as shown in Chapter 3 of this thesis, while they have completely different results for aqueous methanol. This comes as a consequence of the incorrect description of the cross correlations that each of these closures has.

CHAPTER 5

Structure and thermodynamics of aqueous ethanol mixture

5.1 Introduction

Ethanol–water mixture has been studied by various experimental techniques [126-130], as well as computer simulations [131-134]. What makes it different from a random mixture of two liquids is the fact that it has a particular organisation on a molecular scale, with clusters of alcohol and water that are similar to micelles [135]. Nontrivial molecular structure was also found in the aqueous mixture of methanol, which is even smaller molecule than ethanol [136, 137]. Based on their neutron scattering data, Dixit and co–workers [138] conclude that there are coexisting methanol and water clusters and that the local immiscibility of the two species results in anomalous thermodynamics of this mixture. This anomalous behaviour of aqueous alcohol mixtures concerns in the first place the entropy of the system which increases less than expected for an ideal solution of randomly mixed molecules [139]. It is generally believed that these properties are a direct result of a specific interaction that occurs between amphiphilic alcohol molecules and water molecules where alcohol molecules group together their hydrophobic (methyl) parts [140]. D'Arrigo and Paparelli [127] gave a quantitative explanation for the volume and adiabatic compressibility of the ethanol–water mixture through the concept of clathrate–like structures filled with ethanol. Egashira and Nishi [126] proposed a model of ethanol clustering in water based on their Raman spectroscopy measurements. They introduce a picture where ethanol molecules form a polymer cluster with ethyl groups stacked alternatively along the hydrogen–bonding chain of alcohols. Noskov and co–workers [132] performed Molecular Dynamics study of the ethanol–water mixture by using a polarizable force field which revealed the complexity of water clusters in the mixture over the whole composition range. For $x < 0.30$ mole fraction of ethanol, water forms large clusters which break with the increasing

ethanol content. At the concentration slightly less than x = 0.50 mole fraction of ethanol, there are water clusters of various sizes and in the ethanol–rich regions, water exists mostly as dimers and monomers.

5.2 Models and simulation details

OPLS flexibile force field [96] was used for ethanol and SPC/E force field [97] was used for water. The parameters of these force fields are shown in Tables 2.1 and 2.2 in Chapter 2 of this thesis. All the simulations were conducted using the DLPOLY–2 simulation package [92] for a constant NPT ensemble with a total of N = 2048 particles at ambient conditions of temperature (T = 300 K) and pressure (p = 1 atm). The condition of pressure and temperature were maintained through Berendsen thermostat and barostat with relaxation times of 0.1 ps and 0.5 ps, respectively. Ewald summation was used to account for the electrostatic interactions in periodic conditions. The integration time step was set to 2 fs. The mixtures were studied over the whole concentration range from x = 0.1 to 1, with a step of 0.1. Typical equilibration was in the 0.5–1 ns range, which was followed by a series of production runs of 0.5 ns. The total accumulated statistics were in the range of 1–2 ns, which was found sufficient to obtain smooth distribution functions in the long–range part [101, 107].

5.3 Molecular Dynamics results

5.3.2 Structure of neat liquids

Alcohols have a very specific microstructure which was confirmed by the direct cluster study [109, 141]. It was found that methanol forms chain–like clusters, while TBA forms micellar–type structures. The site–site structure factors $S_{ab}(k)$ between O–O, CH_2–CH_2 and CH_3–CH_3 sites of OPLS ethanol and between O–O and H–H of SPC/E water were computed. On the left panel of Fig. 5.1, the site–site structure factors between methyl and methylene sites show a typical Lennard–Jones type structure with large oscillatory structure behind the main peak. In contrast, the O–O structure factor shows a peak at $k \approx 0.75$ Å^{-1}, which corresponds to O–O domains size of $r = 2\pi / k \approx 8$ Å. This suggests the presence of small clusters in pure ethanol. However, for water there is no such distinct pre–peak. There may be an H–bonded network of water molecules [142, 143], but the structure of water is still not well understood.

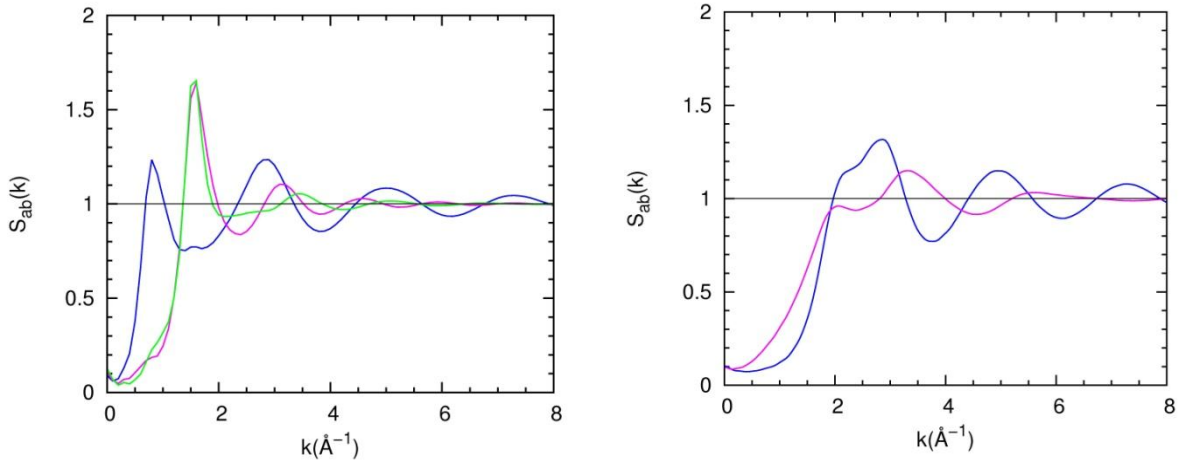

Figure 5.1: Site–site structure factors, $S_{ab}(k)$ of pure ethanol (left panel) and water (right panel). Blue is for O–O, magenta is for CH_2–CH_2 and green is for CH_3–CH_3 for ethanol. For water, blue is for O–O, magenta is for H–H.

5.3.3 Structure and thermodynamics of aqueous ethanol mixture

Microscopic structure of liquids is completely described by the site–site correlation functions, which are direct observables of the correlations in liquids. In Figs. 5.2–5.4, selected site–site correlation functions are reported for several mole fractions of ethanol. Fig. 5.1 shows the $g(r)$ between two oxygen sites on water molecules for various concentrations from $x = 0$ to $x = 0.90$. It is seen that the first peak increases in magnitude as water becomes more dilute which is a sign of increasing water H–bond correlations. The correlations beyond the first peak are seen to first increase until $x \approx 0.60$–0.70 and then sharply decrease in the second neighbour shell. This can be interpreted by the increasing H–bond network which starts to break into smaller clusters at a certain threshold. Fig. 5.3 shows O–O, CH_2–CH_2 and CH_3–CH_3 correlations between the two ethanol molecules. Ethanol O–O correlations behave in the inverse fashion to those of water with respect to the ethanol content and they decrease with less ethanol in the mixture, indicating the fact that H–bond induced clustering in ethanol is not of the same percolating nature as that of the water network [142, 144].

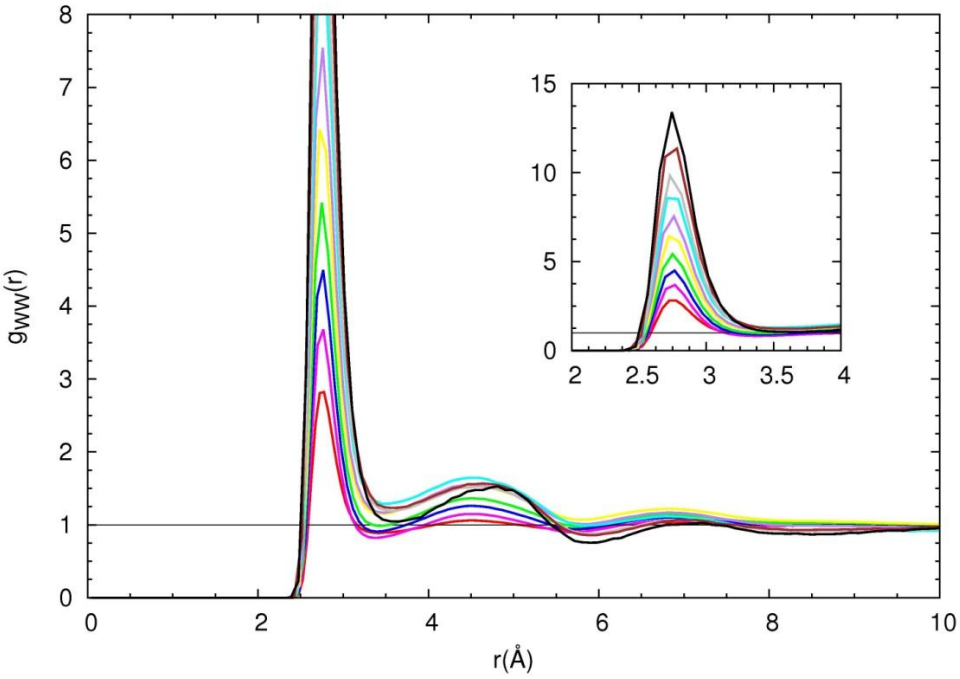

Figure 5.2: Water O–O site distribution functions for ethanol mole fractions from x = 0 (red curve) to x = 0.90 (black curve). Inset shows the zoom on the first peak of the $g_{WW}(r)$. Colour convention: x = 0.80 (brown), x = 0.70 (gray), x = 0.60 (cyan), x = 0.50 (purple), x = 0.40 (yellow), x = 0.30 (green), x = 0.20 (blue), x = 0.10 (magenta).

Middle and bottom panel of Fig. 5.3 show a nontrivial behaviour of the CH_2–CH_2 and CH_3–CH_3 correlations. The first ones exhibit a clear narrowing in the first neighbour shell as x increases, while the second ones broaden and decrease with the same trend. In both cases, the position of the first peak does not change with x. The rate at which all these correlations vary is faster at smaller x values and tends to saturate after about x = 0.60, which is a signature of important clustering changes at lower content to ethanol.

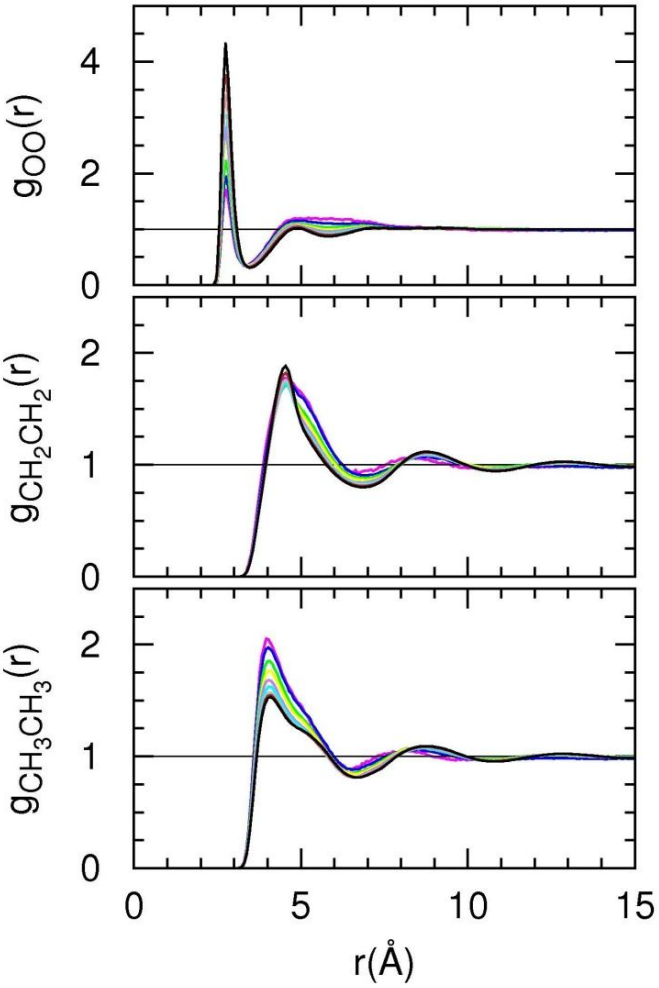

Figure 5.3: Selected ethanol site–site distribution functions for various concentrations. Top panel (O–O), middle panel (CH_2–CH_2) and bottom panel (CH_3–CH_3). Colour convention: x = 0.90 (black), x = 0.80 (brown), x = 0.70 (gray), x = 0.60 (cyan), x = 0.50 (purple), x = 0.40 (yellow), x = 0.30 (green), x = 0.20 (blue) and x = 0.10 (magenta).

Ethanol–water cross correlations are presented in Fig. 5.4. The x dependence of the cross oxygen–oxygen correlations is similar to that of water molecules: the correlations increase as one adds more ethanol to the mixture. Cross methyl–oxygen and methylene–oxygen correlations reveal an interesting behaviour of ethanol molecules as one goes from the ethanol–poor to the ethanol–rich side of the mixture. From x = 0.10 to x = 0.40, the CH_3–O_W correlations decrease in magnitude, after which they increase again. The way these correlations change with concentration indicates that ethanol molecules are shielded from the water surrounding in the low ethanol region, just as if they were forming micelles. These micelles then break when the mixture becomes rich with ethanol and the methyl groups are re–exposed to water.

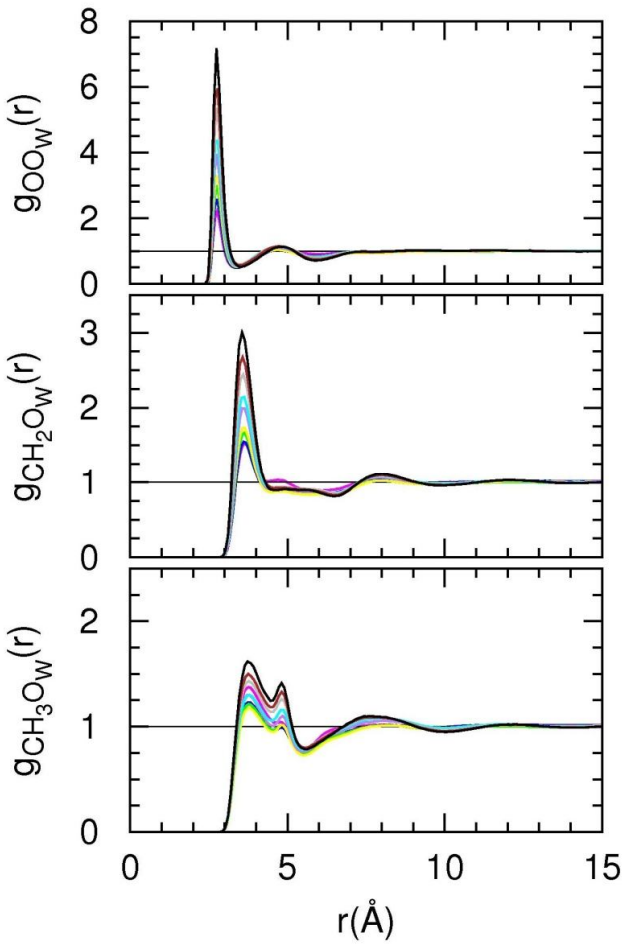

Figure 5.4: Selected cross site–site distribution functions for various concentrations. Top panel ($O–O_W$), middle panel ($CH_2–O_W$) and bottom panel ($CH_3–O_w$). Colour convention is the same as in Fig 5.3.

In the analysis of micro–heterogeneous mixtures, it can be quite instructive to look at various snapshots and to observe whether there is a dominant topology in the system. Three snapshots are taken for systems at $x = 0.10$, $x = 0.50$ and $x = 0.80$ mole fraction of ethanol and given in Fig. 5.5. In the snapshot for $x = 0.10$ mixture, there are small ethanol clusters of about 2 to 4 molecules which pack their methyl groups together. Driving mechanism for the formation of these clusters is the creation of H–bond between alcohol OH groups and water molecules, and the hydrophobic CH_3 groups are buried inside the clusters. For the equimolar mixture, one observes many segregated domains of water and ethanol. For the $x = 0.80$ mixture, water molecules are seen to form H–bonded clusters, with water monomers present as well. These clusters are different from ethanol clusters because they are more filament–like and every water molecule inside the cluster bonds to a surrounding ethanol molecule with the free OH bond.

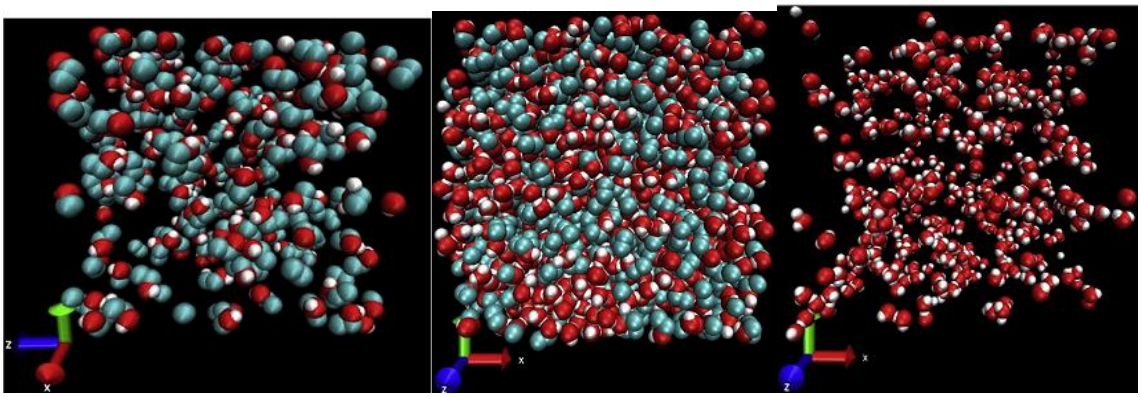

Figure 5.5: Snapshots for x = 0.1 (left), x = 0.5 (middle) and x = 0.8 (right) mole fraction of ethanol. For snapshots at x = 0.1 and x = 0.8, only the minority species is shown. Colour convention: oxygen atom (red), hydrogen atom (white), methyl and methylene group (cyan).

Densities, volumes and excess volumes for ethanol–water mixture are shown in Fig. 5.6. As seen from Fig. 5.6, the results agree quite well with the experimental results for density and molar volume, but slightly worse for excess molar volume than the simulation results from Ref. [145] were the authors used a constant NVT ensemble.

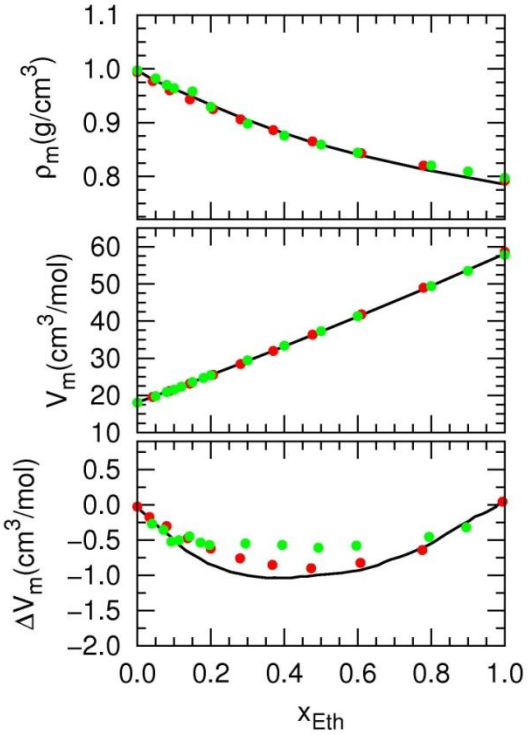

Figure 5.6: Densities (top panel), molar volumes (middle panel) and excess molar volumes (bottom panel) for ethanol–water mixture. Colour and line convention: experimental results [146] (lines), simulation results from Ref. [145] (red dots), calculations (green dots).

Fig. 5.7 shows enthalpies and excess enthalpies for ethanol–water mixture. Experimental excess enthalpies are negative, as shown in the bottom panel of Fig. 5.7, which indicates that the mixing of the two species is energetically favourable. SPC/E enthalpies for pure water are lower than the experimental ones by the factor of $\delta E \approx 5$ kJ/mol, which corresponds to the water polarization contribution. Following the argument by Berendsen [97], the pure water enthalpy is corrected by adding this term and the mixture enthalpies by adding the term $5(1-x)$ kJ/mol, where x is the ethanol mole fraction. As a consequence of a low pure ethanol enthalpy, the results show positive excess energies. However, they do capture the nontrivial composition dependence of experimental excess enthalpy, with the changes in slope at x = 0.15, x = 0.40 and x = 0.60–0.70. It suggests that these models are able to capture energy changes in the mixture at proper concentrations and just the absolute value of excess enthalpy is affected by the inaccurate result for pure ethanol enthalpy.

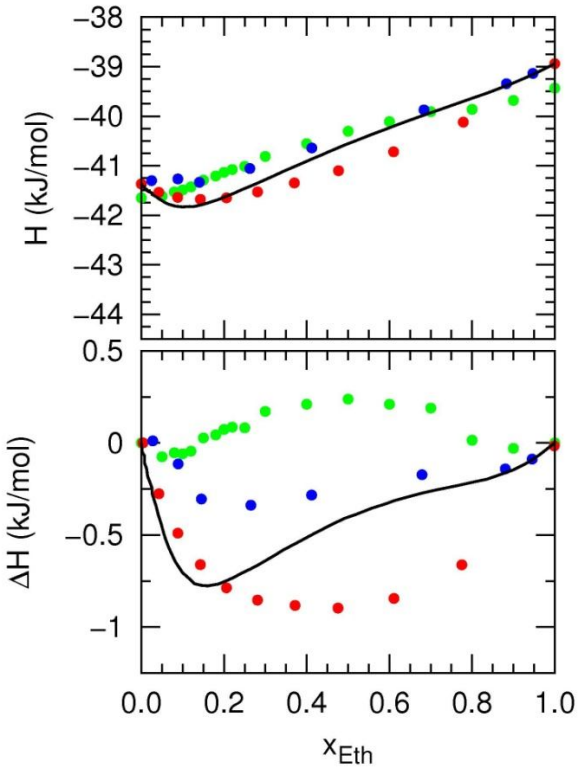

Figure 5.7: Enthalpies (top panel) and excess enthalpies (bottom panel) for ethanol–water mixture. Colour and line conventions: fitted experimental results [5] (lines), simulation results from Ref. [145] (blue dots), simulation results from Ref. [131] (red dots), calculations (green dots). The results of Wensink et al. [131] have been recalculated using their Eq. 9 in conjunction with the results given in their Table II.

Left panel of Fig. 5.8 shows that the experimental excess enthalpy shape from Ref. [5] is very similar to the shape of the negative excess ultrasonic and hypersonic speeds [10]. Since the minimum of these curves is at about the same position at $x \approx 0.18$, one can conclude that the microscopic mechanism behind the changes of these macroscopic values is in fact the same. Right panel of Fig. 5.8 shows the experimental heat capacity from experimental studies [5, 147, 148] and the maximum is again about the value $x \approx 0.18$. The second change of the heat capacity slope is at $x \approx 0.70$, which is the same point were one observes the change of the excess enthalpy slope. The analysis of the oxygen–oxygen correlations of water (Fig. 5.2) over the entire concentration range shows that there is a decrease of the correlations in the second neighbour shell at $x \approx 0.60$–0.70. This re–structuring of the mixture is due to the breaking of the water network into smaller water clusters.

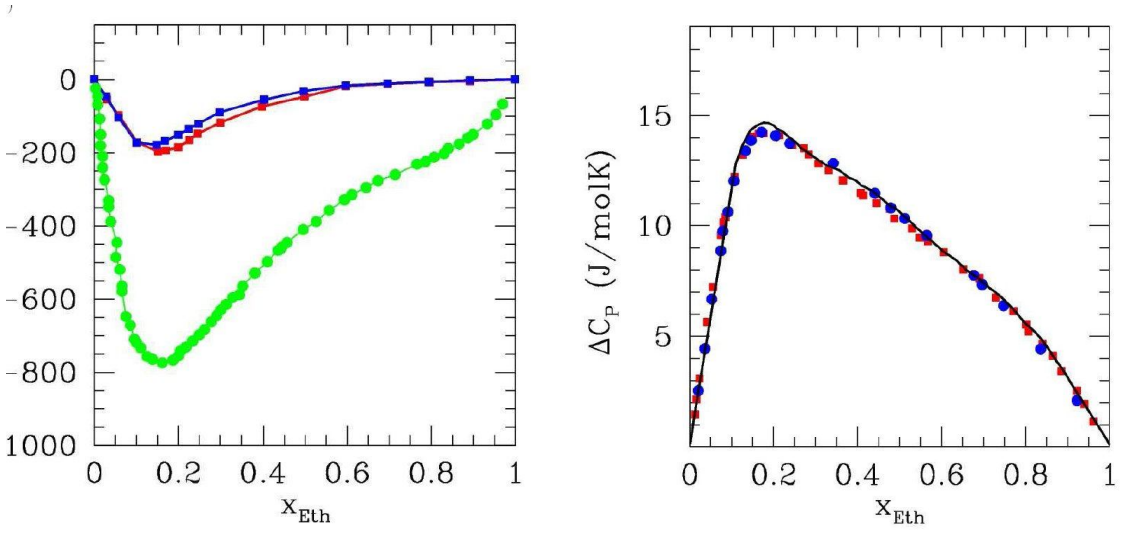

Figure 5.8: Experimental excess enthalpy [5] in units J/mol (green circles) and negative excess sound speed [10] (blue squares for ultrasonic speed and red squares for hypersonic speed in units m/s) in the left panel. Excess heat capacity in the right panel (line from Ref. [5], red squares from Ref. [147] and blue circles from Ref. [148]).

In Chapter 3 of this thesis, the importance of the LP correction, which consists of shifting the tail part of $g(r)$ to unity, is discussed. This behaviour of the long–range correlations was noticed also in aqueous ethanol, but the LP correction did not lead to the correct KBI values, due to the fact that the duration of the simulation runs was inadequate. The results suggest that the simulations should be longer than 2 ns so that the configurational space is properly sampled and the structure one obtains after 2 ns is only a transient one. Figs. 5.9 and 5.10 show examples of the LP correction for mixtures at $x = 0.10$ and $x = 0.90$ mole fraction of ethanol. One observes in

these figures that the tail part of $g_{ij}(r)$ (top panels) does not go to the correct asymptote, which leads to wrong KBI values (bottom panels). This is particularly seen for the minority component (upper left panel of Fig. 5.9 and upper right panel of Fig. 5.10), were the tail fluctuates around some asymptote and one cannot tell where it exactly stabilizes. The corresponding KBI values also do not stabilize. On the other hand, $g_{ij}(r)$ that belong to other pairs (the cross and the majority pair) seem to show a clear asymptote and the LP shift allows the running KBI to reach a horizontal asymptote.

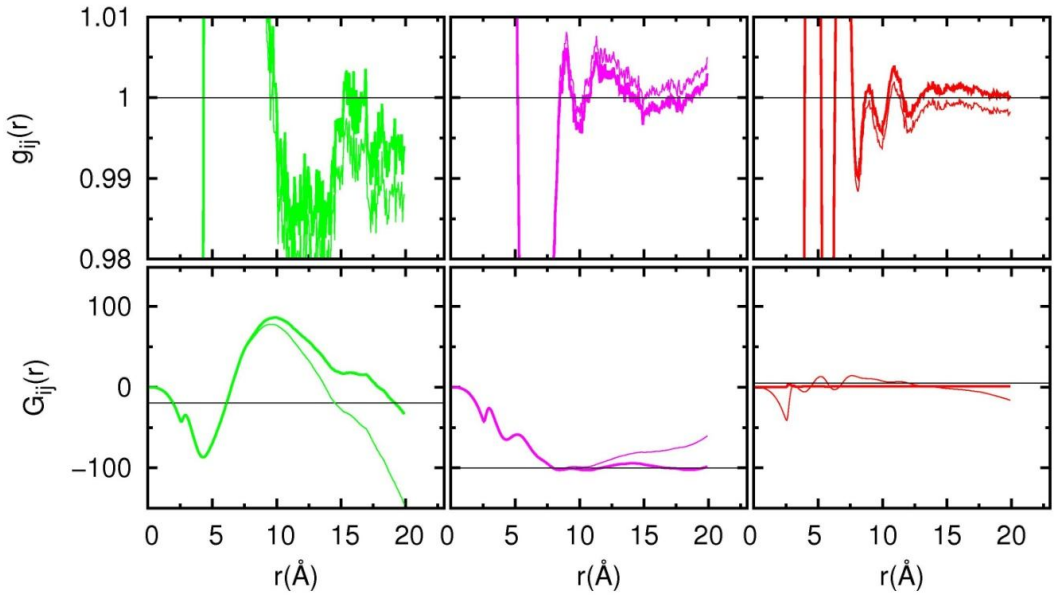

Figure 5.9: Example of the LP correction for aqueous ethanol mixture x = 0.10. Colour convention: ethanol–ethanol (green), ethanol–water (magenta), water–water (red). Panel convention: $g_{ij}(r)$ (top), $G_{ij}(r)$ (bottom). Thin lines are the uncorrected data and thick lines are the corrected data. Thin black lines in the lower panels are the estimated values of the corrected KBI.

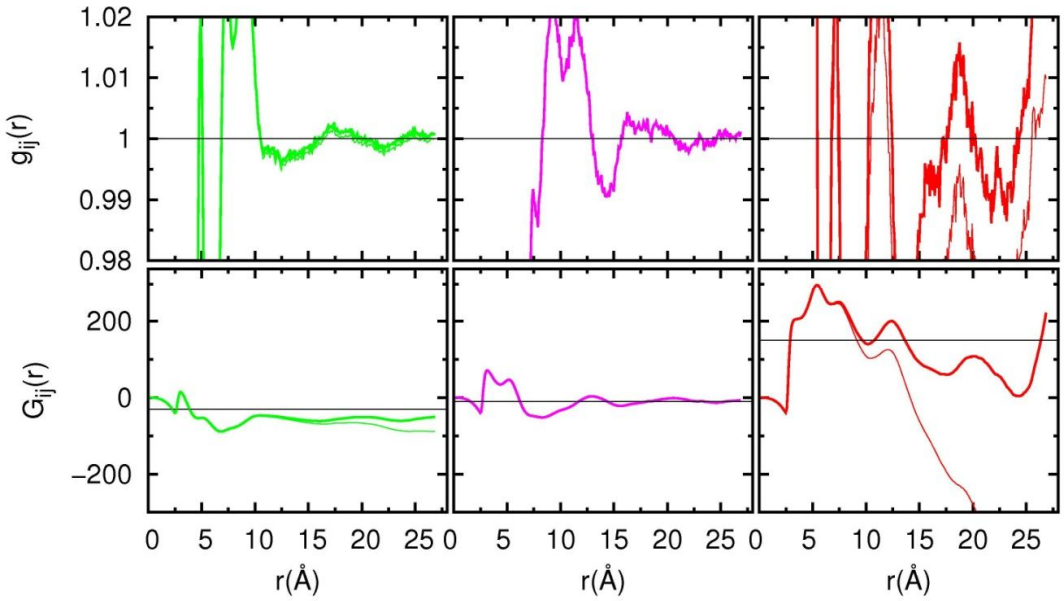

Figure 5.10: Example of the LP correction for aqueous ethanol mixture $x = 0.90$. Colour, line and panel convention is the same as in Fig. 5.8. Thin black lines in the lower panels are the estimated values of the corrected KBI.

5.4 Conclusion

The mixture of ethanol–water was studied over the whole concentration range by using Molecular Dynamics simulations. This mixture is different from the mixture of simple liquids, since one observes changes in the slopes of thermodynamic quantities which indicate structural re–organization. This cannot be a local effect, since local rearrangements also happen in mixtures of simple liquids and they do not lead to changes of global variables. There are two such changes for aqueous ethanol: one at low mole fraction of ethanol, at $x \approx 0.18$ and another at $x \approx 0.70$. First one corresponds to the differences in solute clustering and the other to the breaking of water network. Analysis of the tail of $g_{ij}(r)$ showed that the current simulation statistics are not adapted to the study of this mixture, since the tails do not stabilize in the thermodynamic limit. As a consequence, the KBI does not have the corresponding asymptotic value. The application of the LP correction did not solve this problem which suggests that a more extensive simulation study is required. The short–range structure of the mixture is however successfully described within the current study.

CHAPTER 6

Concept of molecular–emulsion

6.1 Introduction

Micro–emulsion is a well defined term in soft matter physics. It corresponds to a thermodynamically stable ternary solution made of water, oil and surfactant [149]. If mixed together, oil and water would phase separate by creating a horizontal interface between them. Addition of a surfactant, which is an amphiphilic molecule, leads to the creation of oil–water domains or micelles. These objects are on a nanometer scale and their structure depends on temperature and the composition of a mixture, as shown in Fig. 6.1 from Ref. [150]. Oil droplets can be dispersed in water forming micelles with the surfactant molecules (Fig. 6.1a). In Fig. 6.1b, we see a two phase coexistence of oil and water domains with surfactant sitting at the borders. Inverse micelles can form when water is dispersed in oil, as shown in Fig. 6.1c. There are several approaches describing the microscopic structure of micro–emulsions [151-154]. In this chapter, I study the Teubner and Strey (TS) model based on the Ginzburg–Landau theory [155]. Using neutron scattering technique, Teubner and Strey [156] fitted the scattering intensity of a micro–emulsion to the following form:

$$I(k) = \frac{I(0)}{a_2 + k^2 c_1 + k^4 c_2}, \tag{6.1}$$

where k is the modulus of the wave vector and a_2, c_1 and c_2 are phenomenological parameters.

Since the objects in these solutions are far bigger than the molecules they consist of, they can be described by continuous fields. Therefore, Teubner and Strey used the Ginzburg–Landau field theory where free energy, F of a system in an ordered state can be expressed in the terms of an order parameter $\psi(\vec{r})$ as [155]

$$F[\psi(\vec{r})] = \int dr^3 \left[a_0 + a_1\psi + a_2\psi^2 + a_3\psi^3 + \ldots + c_1(\nabla\psi)^2 + c_2\left(\nabla^2\psi\right)^2 + \ldots \right]. \qquad (6.2)$$

The point where the term under integration stops depends on the considered problem and the accuracy one wishes to achieve. The order parameter $\psi(\vec{r})$ describes the spatial distribution of matter by neglecting molecular details.

As seen from Eq. (6.2) only even terms of $\nabla^i\psi$, where *i* = 1, 2, ... are kept. This ensures the free energy invariance under the transformation $\psi \to -\psi$.

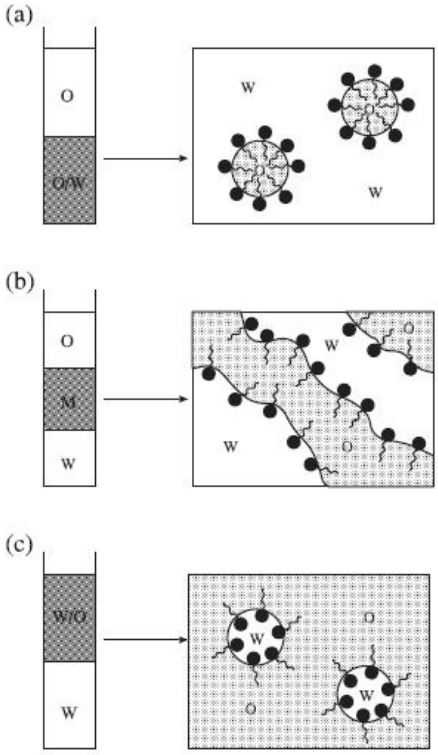

Figure 6.1: (from Ref. [150]) The structure of micro–emulsion. a) Oil–in–water (O/W) droplet phase coexisting with excess oil (O) phase, b) Middle phase (M) coexisting with excess oil (O) and excess water phase (W) with surfactant molecules at the interfaces, c) Water–in–oil droplet phase coexisting with an excess water (W) phase.

Teubner and Strey chose the order parameter $\psi(\vec{r})$ to be the difference in the local volume fractions of oil and water. They neglect higher orders than the second in $\psi(\vec{r})$ because of the small fluctuations in a homogeneous isotropic phase and keep only $a_2 > 0$ in Eq. (6.2). The term c_1 is also kept and taking all the other *c*–terms to be zero, would lead to the Ornstein–Zernike expression for the scattering intensity near the critical point [157]

$$I(k) = \frac{1}{a_2 + c_1 k^2}, \tag{6.3}$$

where the correlation length ξ is defined as $\xi = \sqrt{c_1 / a_2}$. Teubner and Strey obtained the correlation function $g(r)$ by taking the inverse Fourier transform of Eq. (1.36) as

$$g(r) = \frac{d}{2\pi r} e^{-r/\xi} \sin\left(\frac{2\pi r}{d}\right), \tag{6.4}$$

where they introduced two characteristic length scales, d and ξ. d is the size of the domains and it is defined as

$$d = 2\pi \left[\frac{1}{2}\left(\frac{a_2}{c_2}\right)^{1/2} - \frac{1}{4}\frac{c_1}{c_2} \right]^{-1/2}, \tag{6.5}$$

while ξ is the correlation length of the form

$$\xi = \left[\frac{1}{2}\left(\frac{a_2}{c_2}\right)^{1/2} + \frac{1}{4}\frac{c_1}{c_2} \right]^{-1/2}. \tag{6.6}$$

The TS approach established a direct connection between an experimental quantity $I(k)$ and the microscopic parameters d and ξ for a broad spectrum of micro–emulsions only by varying three fitting values. It is possible to derive the TS form of the structure factor from the MOZ equation for molecular fluids that consist of linear molecules. The exact derivation is given in Appendix B.

6.2 Methodology of the TS extension

Strongly micro–heterogeneous systems, such as aqueous TBA, need large scale simulations with several thousands of particles so that the structure can be correctly sampled over the configurational space [158-160]. Simulation of the TBA–water mixture [121] in this work has been conducted in a box with N = 2048 particles, which was enough to conclude that there are oscillations in the pair correlation function over the half–width of the simulation box, indicating a nontrivial long–range behaviour. The structure and thermodynamics of this mixture will be discussed in details in Chater 6 of this thesis. I chose to test the TS fit on a mixture of TBA and

water by prolongating the incomplete ending in the long–range part of the correlations between (*i,j*) pair of sites which belong to two distinct molecules

$$t_{ij}(r \geq L) = A_{ij} \frac{\exp\left(-(r-L)/\xi\right)}{r} \sin\left(\frac{r - R_{ij}}{\bar{d}}\right), \tag{6.7}$$

where L is the half–width of the simulation box and the place at which the fit starts. The parameters (A_{ij}, R_{ij}) are adjusted to match the value from the simulation at the point $g_{ij}(r = L)$, as well as to keep the derivative continuous at the merging point. The reduced domain size parameter $\bar{d} = d / 2\pi$ corresponds to the size of the oscillations, while the correlation length ξ is the parameter that damps these oscillations. Parameters (A_{ij}, R_{ij}) are different for each site–site correlation function, but $(\bar{d}, \xi)$ are not, as demonstrated from the exact derivation in Appendix B.

Figure 6.2 shows the TS extension of the function $g_{oo}^{(W)}(r)$ at x = 0.20 mole fraction of OPLS [96] TBA in SPC/E [97] water, the corresponding running KBI (RKBI) $G_{oo}^{(W)}(r)$ and $S_{oo}^{(W)}(k)$. The same is shown in Figure 6.3 but for O–O correlations of TBA. The top panel of each figure shows a zoom on the tail of $g_{oo}^{(W)}(r)$ from simulations and the first peak is given in the inset. The simulation RKBI, as seen in the middle panel of both figures, goes towards a value far greater that the experimental one and the TS correction brings the RKBI closer to the expected value. Structure factors $S_{ab}(k)$ are given in the bottom panels of the two figures and they are plotted against the structure factor of neat liquids whose correlation is analyzed. What is particularly interesting is the k = 0 behaviour of the structure factor, because it reflects the long–range structure of the liquid. In neat water and TBA, $S_{ab}(k)$ vanishes in the $k \to 0$ limit, contrary to what is observed in the corrected $S_{ab}(k)$ for TBA–water mixture. The structure factor of water $S_{oo}^{(W)}(k)$ (bottom panel of Fig.6.2) in this mixture shows a peak at small–k, indicating the formation of water–water domains. The oxygen–oxygen TBA structure factor $S_{oo}^{(A)}(k)$, grows in the small–k limit and reaches a non–zero value (bottom panel of Fig.6.3), which means that there are fluctuations of concentration in the mixture.

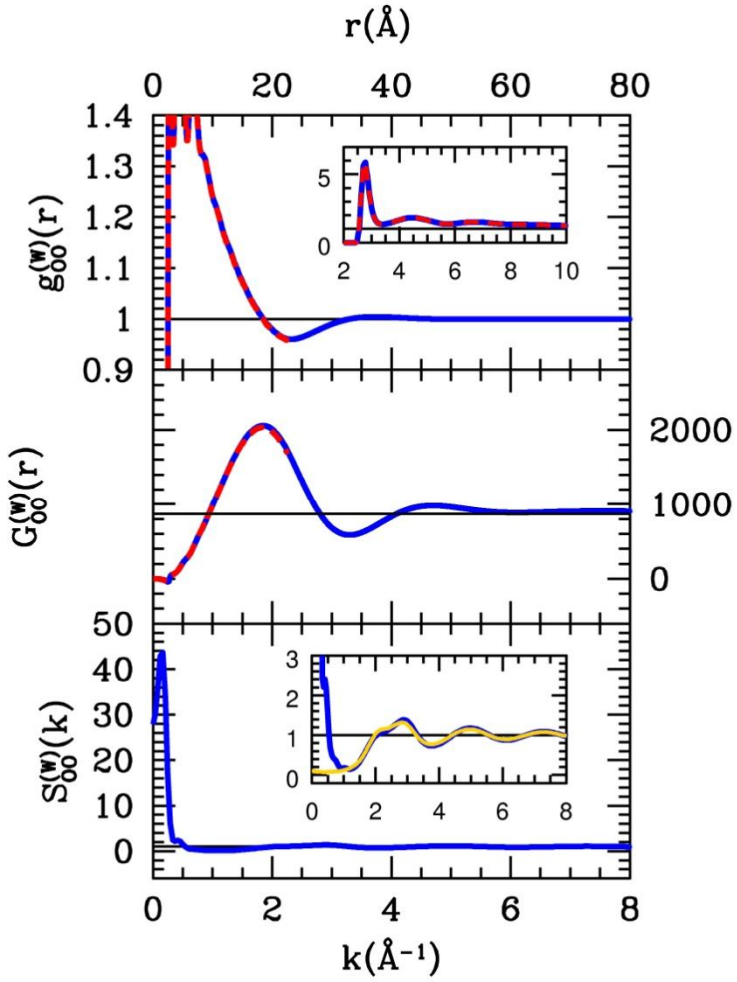

Figure 6.2: Details of TS extension procedure for the OPLS TBA oxygen–oxygen correlations for the TBA mole fraction x = 0.20. Top panel, tail part of $g_{oo}^{(W)}(r)$ (the short range part shown in the inset): red dashed line for the uncorrected data, blue line after TS correction. Middle panel, RKBI $G_{oo}^{(W)}(r)$ with same line conventions as in top panel. Black line is the experimental data from Ref. [4]. Lower panel, structure factor $S_{oo}^{(W)}(k)$ with same line conventions as in top panel. The gold curve is the inset is the structure factor of neat water.

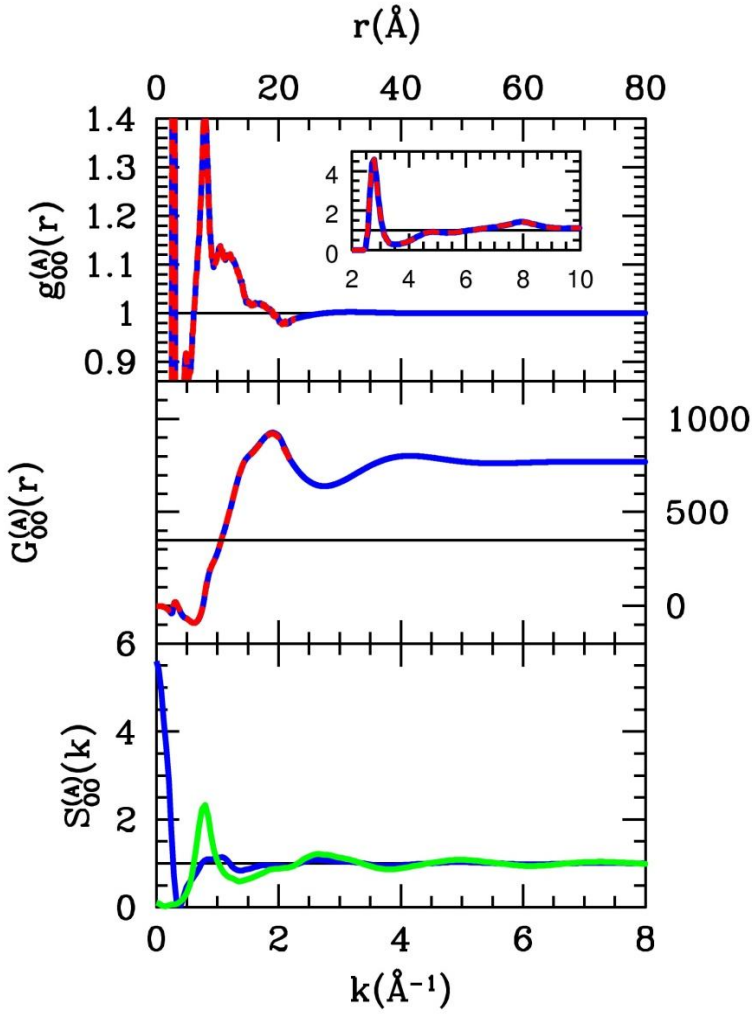

Figure 6.3: The same as Fig. 6.2 but for SPC/E water oxygen–oxygen correlations $g_{oo}^{(A)}(r)$ for x = 0.20. Black line in the middle panel is the experimental data from Ref. [4].The green curve in the lower panel is the structure factor of neat OPLS TBA.

.

TS extension is also presented at x = 0.80 mole fraction of TBA in Figs. 6.4 and 6.5, so one can see what happens with the correlations when water is the minority species. It is seen from the middle panel of Fig. 6.4 that the uncorrected RKBI have a slanted asymptote. Therefore, it is essential to perform an LP correction that consists of shifting the tail of $g_{ab}(r)$ before applying the TS extension. Another important point refers to the corrected structure factors $S_{ab}(k)$. The corrected water–water structure factor, $S_{oo}^{(W)}(k)$ as seen in the bottom panel of Fig. 6.4, shows a noisy pre–peak at a small–k value due to the noise in the tail of $g_{ab}(r)$ from simulations. As seen from this panel, $S_{oo}^{(W)}(k)$ from mixture differs significantly from the neat water $S_{oo}^{(W)}(k)$.

In the bottom panel of Fig. 6.5, the corrected TBA structure factor $S_{oo}^{(A)}(k)$ is compared with the neat TBA $S_{oo}(k)$ from simulations to visualize the difference in the structure between the neat liquid and the majority component in the mixture.

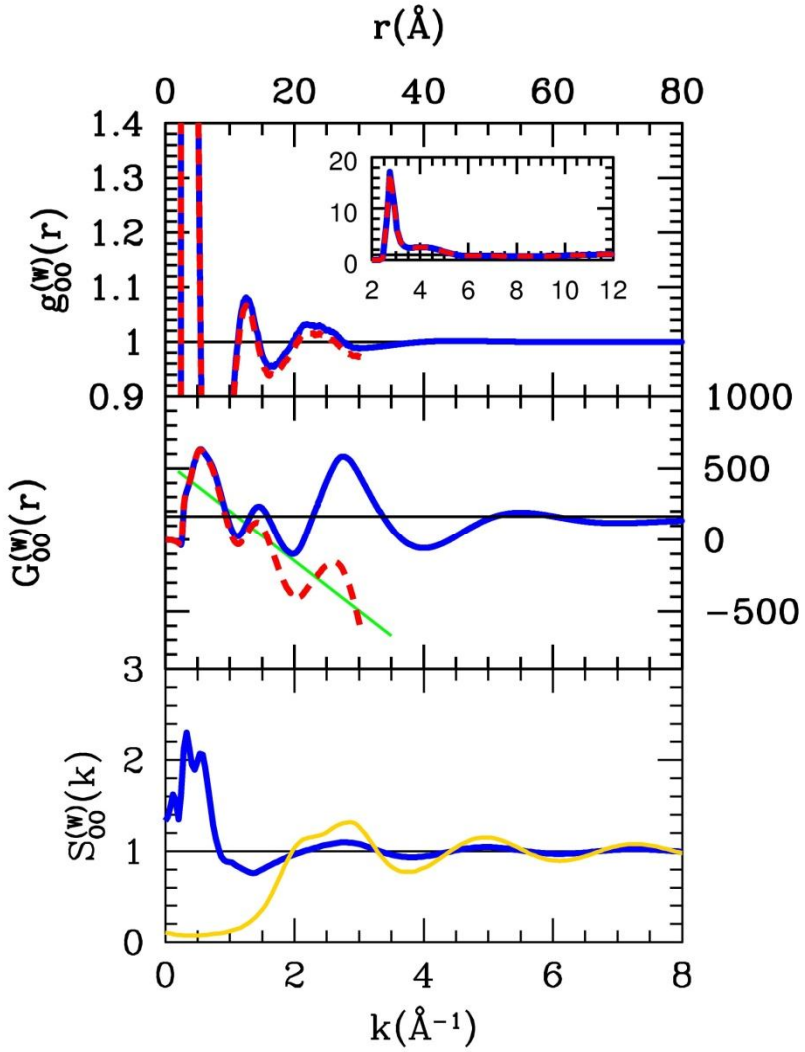

Figure 6.4: The same as Fig. 6.2, but for x = 0.80. The green slanted line in the middle panel shows the wrong slope of the uncorrected data before the LP correction. Black line in the middle panel is the experimental data from Ref. [4]. The gold curve in the lower panel is the structure factor of neat SPC/E water.

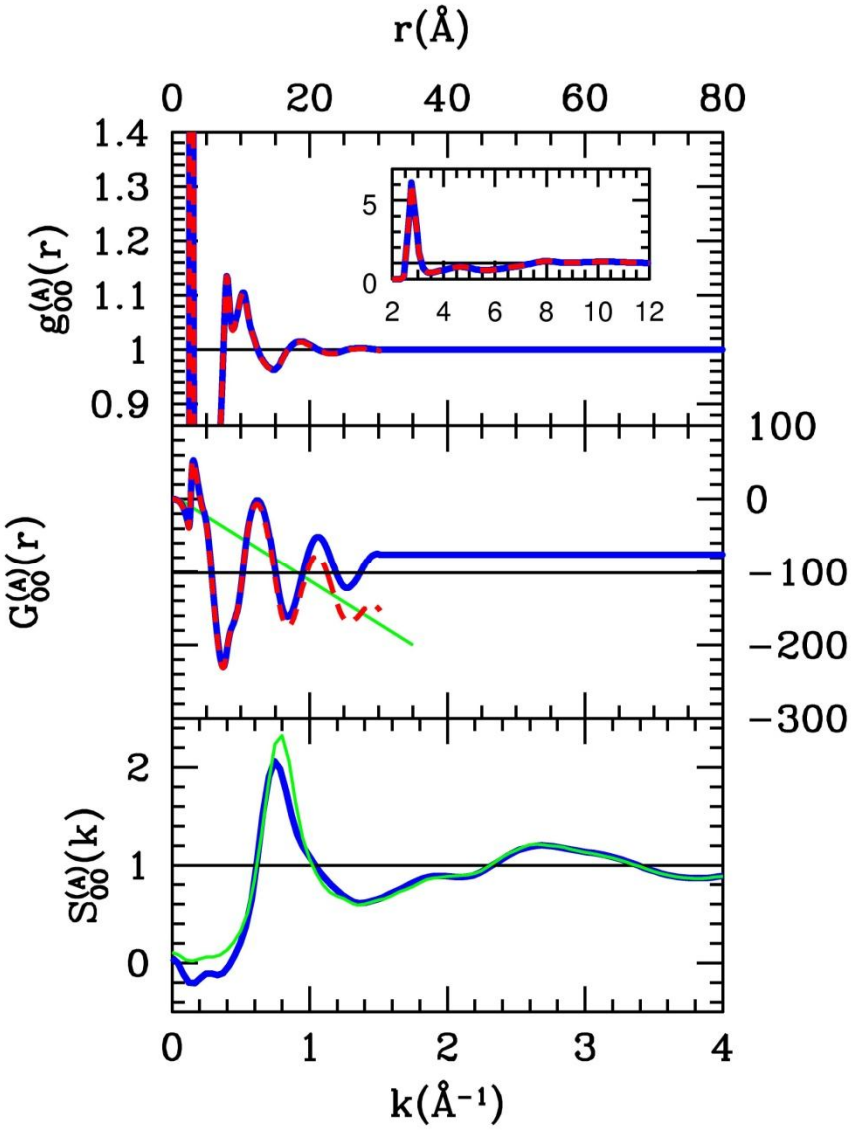

Figure 6.5: The same as Fig. 6.3, but for x = 0.80. The green slanted line has the same meaning as in Fig. 6.4. Black line in the middle panel is the experimental data from Ref. [4]. The green curve in the lower panel is the structure factor of neat OPLS TBA.

As seen from each of these four figures, the tail of the correlation function from simulations stops somewhere below 1, with a clear tendency to extend beyond 23 Å. I found that these is no large margin of choice for parameters $(\bar{d},\xi)$. As a matter of fact, the size of oscillations controlled by the $\bar{d}$ –parameter, is already hinted by the tail of the function itself. However, this is not the case for the ξ –parameter, since one can chose when the oscillations will stop. We empirically found $\bar{d} \approx 30\,\text{Å}$ and $\xi \approx 12\text{Å}$ for nearly all concentrations in the range x = 0.20–0.80. These mixtures exhibit strong micro–heterogeneity which has the same topology, as it is implied the fitting parameters.

6.3 Conclusion

In this Chapter, I introduce the concept of molecular–emulsion, which is based on the analogy with real emulsions were domains form at a micrometer scale. Real emulsions consist of oil, water and surfactant, while in molecular–emulsions domains are being formed from water and small amphiphilic molecules, such as TBA, for example. The concept of molecular–emulsion can be applied to systems with strong micro–heterogeneous structure and it allows recovering the experimental value of the KBI. In the simulation of such systems, it is important to have large enough statistics so that one can clearly see the asymptotical decay of the radial distribution functions and the oscillatory structure they assume in the long–range part. Only then one can apply the TS extension procedure.

If the correlations in these systems are not captured correctly in the thermodynamic limit, this affects the KBI value by making it far too large compared to the experimental value. The TS extension involves two parameters, the domain size d which is a measure for micro–heterogeneity and the correlation length ξ which corresponds to concentration fluctuations. If these two parameters approach the same value, it means that micro–heterogeneity starts to compete with the concentration fluctuations and this competition affects the small–k behaviour of the structure factor $S(k)$. The formation of a new object that appears from concentration fluctuations is observed in the pre–peak of $S(k)$ and this is exactly the result of the TS extension, since the small–k behaviour of the correlations in k–space corresponds to the long–range oscillations in the real space. The TS extension is tested against simulation results also for a mixture of water and modified water and for water–acetone mixture. The results will be discussed in Chapters 7 (for acetone–water mixture) and 8 (water –"weak water" mixture).

CHAPTER 7

Structure and thermodynamics of molecular–emulsions

7.1 Introduction

TBA is considered to be the most hydrophobic of all lower alcohols and it is the largest alcohol that is fully miscible with water in all proportions [161]. This mixture has received attention of both experimentalists [161-165] and simulators [160, 166, 167]. Nishikawa [165] used small angle X–ray scattering technique to study TBA–water mixture and confirmed the existence of clathrate–like structures of TBA surrounded by water molecules for dilute solutions of TBA. Bowron and Diaz–Moreno [168] analyzed the structure of 0.86 mole fraction of TBA in water by neutron diffraction with hydrogen/deuterium isotope substitution. They found that water molecules form pockets of about two or three molecules on average in this particular mixture. The same technique was used by Bowron et al. [161] for the study of aqueous TBA mixtures at x = 0.06, 0.11 and 0.16 mole fraction of TBA, which demonstrated that alcohol molecules associate between themselves through their non–polar methyl regions, rather than via polar interactions of their hydroxyl groups. Lee and van der Vegt [160] performed molecular dynamics simulations which revealed strong solute self–aggregation in the low TBA region and water–water aggregation over the whole concentration range. Simulation studies on aqueous TBA revealed that this system needs large scale simulations, mostly beyond N = 1000 particles, as well as run lengths of the order of several nanoseconds, in order to sample correctly both the molecular scale and the domain scale dynamics. It was also noticed that most of the existing force fields, the OPLS force field in particular, is not adapted to describe the micro–heterogeneity in aqueous TBA [121]. This is seen in the excessively large KBI values, which indicate that the distribution function has a pathological behaviour in the thermodynamic limit. This behaviour persists even in the simulations of N = 2000–3000 particles, and some authors

proposed the modifications of the solute force field in order to match the experimental KBI values [160]. This was also suggested for aqueous acetone. Many combinations of acetone and water models, namely, OPLS [93], FHMK (Ferrario, Haughney, McDonald and Klein) [169] for acetone and SPC/E [97] and TIP4P (Transferable Intermolecular Potential 4–Point Charge) [113] produced KBI 6–7 times higher than the experimental values [170]. Perera and Sokolić [170] found that strong–clustering in a system with N = 864 molecules becomes a clear phase separation for N = 2048 molecules, for OPLS acetone mole fraction of 0.30 in SPC/E water. Weerasinghe and Smith [159] developed a new force field model for acetone (WS) that is miscible with SPC water and produces the KBI near the experimental vales. However, when plotted against the mole fraction of acetone x, these KBI have a wrong shape. They increase monotonously until x = 1, while they should have a maximum around x = 0.60. The importance of this model lies in the fact that it was able to reduce micro–heterogeneity in this system, as it was also observed in snapshots [170], hence providing KBI on an acceptable scale. In this Chapter, another acetone force field is used, the TraPPE–UA (Transferable Potential for Phase Equilibria–United Atoms) model, that was built to reproduce accurately the liquid–gas coexistence of acetone mixtures [171]. This model mixes with SPC/E water at all mole fractions, but the KBI reach high values.

Instead of performing large–scale simulations, the concept of molecular–emulsion is used to study these mixtures. In micro–emulsions, oil and water domains modulate the long–range part of the correlation functions, which leads to a peak in the scattering intensity at small wave vector. Micro–heterogeneity in aqueous mixtures can be considered as domain formation, although at a much smaller scale than in micro–emulsions. Based on the Teubner and Strey (TS) approach to micro–emulsions, which is described in Chapter 6, the long–range part of the radial distribution function is extended in order to account for these domain oscillations that are currently unavailable from computer simulations. As a result, the KBI values are brought to reasonable size, which shows the importance of the long–range structure in the calculation of the KBI.

7.2 Models and simulation details

OPLS [93] and TraPPE–UA [171] force field models were used for TBA and SPC/E [97] water was used to model water. TraPPE–UA force field was used for acetone. Tables 7.1 and 7.2 show TBA force fields and acetone force fields, respectively. The differences one observes between

force field parameters for the same solute lead to large differences in the description of micro–heterogeneity. All the simulations were conducted using the DLPOLY–2 simulation package [92] for a constant NPT ensemble with a total of N = 2048 particles at ambient conditions of temperature (T = 300 K) and pressure (p = 1 atm). The condition of pressure and temperature were maintained through Berendsen thermostat and barostat with relaxation times of 0.1 ps and 0.5 ps, respectively. Ewald summation was used to account for the partial charges in periodic conditions. The integration time step was set to 2 fs. The TBA–water mixtures were studied over the whole concentration range from x = 0.05 to 1, with a step of 0.1. Each concentration was initially randomly generated in a box whose volume corresponds to the experimental one. It was then equilibrated for 100 ps in a constant NVT ensemble with all partial charges set to zero to ensure a homogeneous mixture. Next, the charges are turned on and the system is first equilibrated for 200 ps in an NVT ensemble and then in an NPT ensemble for time varying from 1–1.5 ns, depending of the concentration. Longer equilibration times were needed around x = 0.1 to 0.3 mole fraction of TBA, were the maximum of concentration fluctuations occur.

Table 7.1: TBA force field models.

σ (Å)	CH_3	C	O	H
OPLS	3.91	3.8	3.07	0
TraPPE–UA	3.75	5.8	3.02	0
Ref. [160]	3.75	6.64	2.95	0
ε (kJ/mol)				
OPLS	0.6699	0.2094	0.7117	0
TraPPE–UA	0.814	0.0415	0.7730	0
Ref. [160]	0.8671	0.0070	0.8496	0
q (e)				
OPLS	0	0.265	–0.7	0.435
TraPPE–UA	0	0.265	–0.7	0.435
Ref. [160]	0	0.337	–0.76	0.423

Table 7.2: Acetone force field models.

σ (Å)	O	C	CH_3
OPLS	2.96	3.75	3.91
TraPPE–UA	3.05	3.82	3.75
WS	3.10	3.36	3.75
ε (kJ/mol)			
OPLS	0.879	0.440	0.67
TraPPE–UA	0.656	0.224	0.814
WS	0.56	0.33	0.867
q (e)			
OPLS	–0.424	0.3	0.062
TraPPE–UA	–0.564	0.662	–0.049
WS	–0.565	0.565	0

7.3 Molecular Dynamics results

7.3.1 Structure of neat liquids

The site–site structure factors $S_{ab}(k)$ between O–O, C–C and CH_3–CH_3 sites of pure TBA and between O–O, C–C and CH_3–CH_3 sites of pure acetone are computed and shown in the left and right panel of Fig. 7.1, respectively. The site–site structure factors between methyl and methylene sites show a typical Lennard–Jones type structure, which is also observed in neat ethanol. The peak of the O–O structure factor in TBA indicates clustering and the existence of O–O clusters that appear as closed loops or micelles in TBA was reported previously in Ref. [141]. Each of the structure factors shown for neat acetone (right panel of Fig. 7.1) has a Lennard–Jones type behaviour. There are no pre–peaks for the selected pairs and this suggests that these are no clusters in neat acetone. The height of the peak at $k \approx 1.5$ Å^{-1} is the biggest for the C–C pair, meaning that these correlations are the strongest ones.

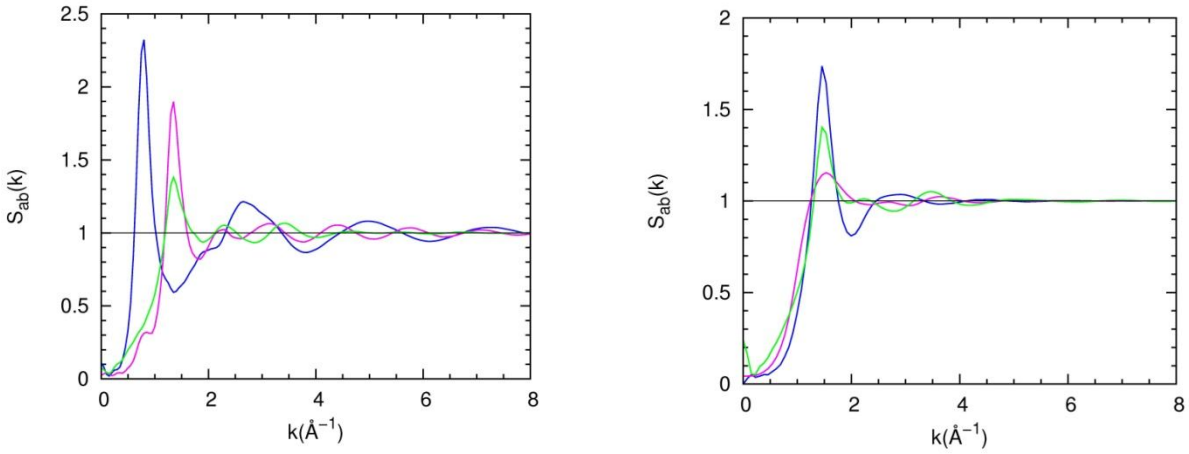

Figure 7.1: Site–site structure factors, $S_{ab}(k)$ of pure TBA (left panel) and acetone (right panel). Blue is for O–O, magenta is for CH_2–CH_2 and green is for CH_3–CH_3 of TBA. Blue is for C–C, green is for CH_3–CH_3 and magenta is for O–O of acetone.

7.3.2 Structure and thermodynamics of aqueous TBA mixture

The analysis of the site–site radial distribution functions for the TBA–water mixture is presented. Results on Figs. 7.2–7.4 are calculated for the OPLS force field of TBA. Fig. 7.2 shows the water oxygen–oxygen distribution functions for different concentrations of TBA. As seen on this plot, the first peak of the $g_{WW}(r)$ increases as water becomes more rarefied, like in aqueous ethanol. However, the maximum of the water–water correlations is not at $x = 0.90$, but rather at $x = 0.80$, which is an indication of strong H–bond clustering at that particular concentration. Correlations in the second coordination shell also increase and become sharper as one adds more TBA, but only up to $x = 0.50$ where they start to decrease. Another important feature of $g_{WW}(r)$ is the fact that it reaches the highest value after $r \approx 6$ Å for the $x = 0.20$ mixture, indicating that water molecules are very strongly correlated even beyond the second coordination shell.

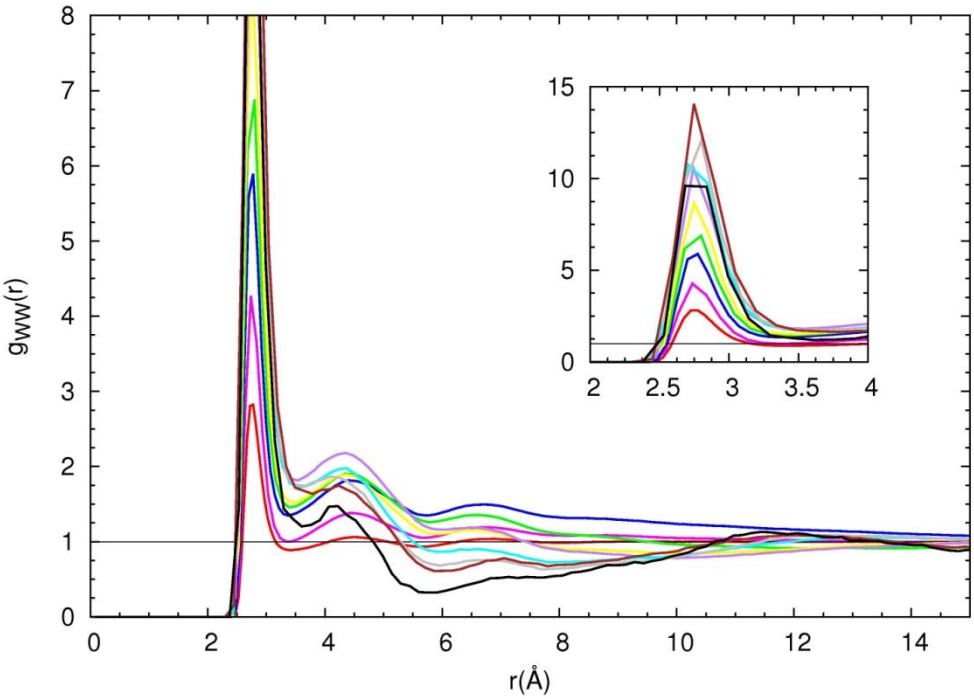

Figure 7.2: Water O–O site distribution functions for OPLS TBA mole fractions from x = 0 (red curve) to x = 0.90 (black curve). Inset shows the zoom on the first peak of the $g_{WW}(r)$. Colour convention: x = 0.90 (black), x = 0.80 (brown), x = 0.70 (gray), x = 0.60 (cyan), x = 0.50 (purple), x = 0.40 (yellow), x = 0.30 (green), x = 0.20 (blue) and x = 0.10 (magenta).

Top panel of Fig. 7.3 shows the O–O correlations between the two TBA molecules. At x = 0.90, the first peak has the highest value, but there is depletion in the second and the third coordination shell, as TBA becomes more correlated with water than with its own kind beyond the first neighbours. C–C and CH_3–CH_3 correlations are shown in the middle and the bottom panel of the same figure. As seen on both panels, there are two first neighbour distances where the height of $g(r)$ depends on the concentration. Interestingly, these correlations are the strongest for x = 0.20, after which they start to become smaller and smaller. This is an indication of cluster formation of TBA in water where clusters are formed by turning the methyl and carbon sites towards each other. Further addition of TBA molecules in the mixture causes these correlations to decrease and favours O–O correlations instead.

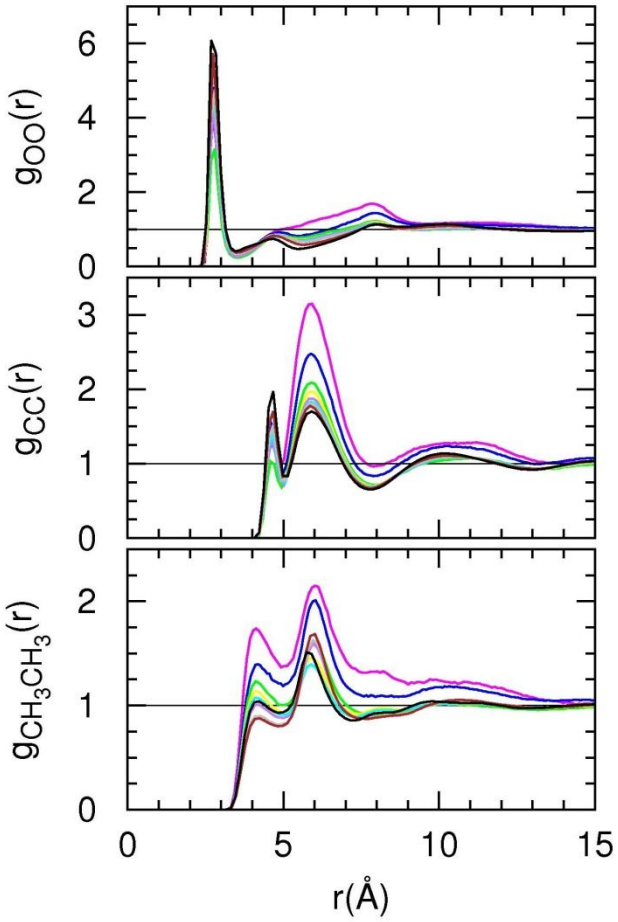

Figure 7.3: Selected OPLS TBA site–site distribution functions for various concentrations. Top panel (O–O), middle panel (C–C) and bottom panel (CH_3–CH_3). Top panel (O–O), middle panel (CH_2–CH_2) and bottom panel (CH_3–CH_3). Colour convention: x = 0.90 (black), x = 0.80 (brown), x = 0.70 (gray), x = 0.60 (cyan), x = 0.50 (purple), x = 0.40 (yellow), x = 0.30 (green), x = 0.20 (blue) and x = 0.10 (magenta).

The O–O_w, C–O_w and CH_3–O_w correlations are reported in Fig. 7.4. These correlation smoothly grow from x = 0.10 to x = 0.90, as cross correlations become more enhanced when going to the TBA–rich side of the mixture. The middle and bottom panel of this figure may be regarded as complement panels of the Fig. 7.3. Indeed, at x = 0.20 the C–O_w and CH_3–O_w correlations are very low which supports the picture of carbon and methyl sites that are shielded from the surrounding water. In Fig. 7.5, the short range features of $g(r)$ from the two mixture models studied here is compared to that from Ref. [160] where force field was altered to reproduce the proper KBI within the system size. Various site–site correlations are examined and it is clearly seen that all model combinations reproduce similar features. Model from Ref. [160] reproduces

results that are closer to OPLS than to TraPPE–UA results. Differences are observed in the short–range part of $g_{CC}(r)$ (top panel) and $g_{OH}^{AW}(r)$ (bottom panel) and in the long–range part of water–water correlations $g_{OO}^{W}(r)$ (bottom panel). For OPLS model and model from Ref. [160], the water O–O correlations drop below 1 after 8 Å, while for TraPPE–UA model it extends above 1, which makes the corresponding KBI value quite large. In conclusion, the TraPPE–UA model overestimates the domain correlations.

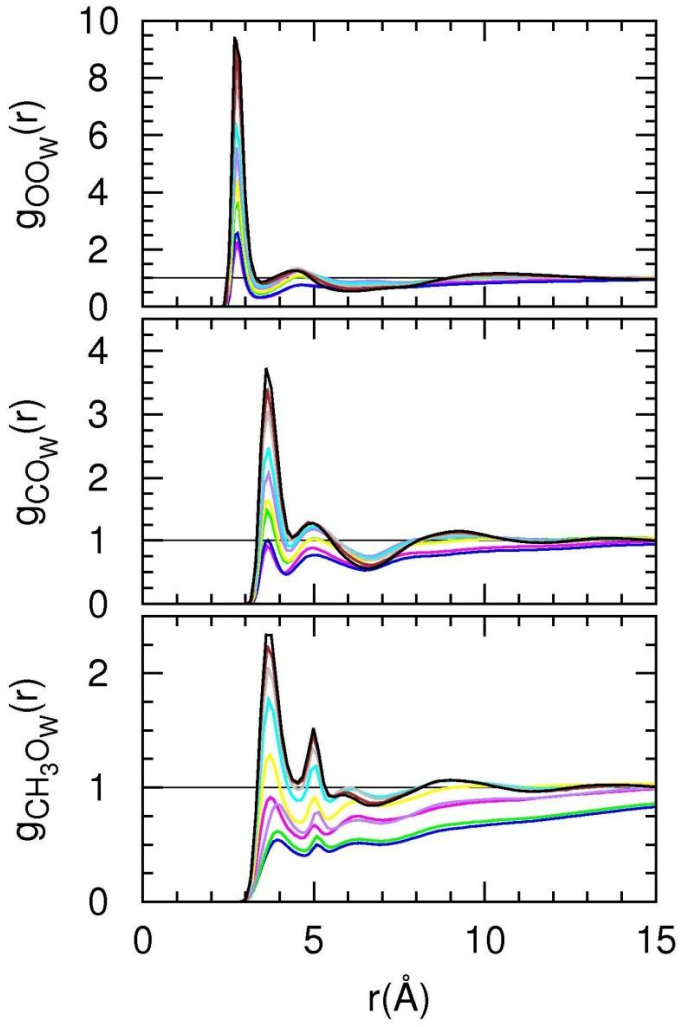

Figure 7.4: Selected cross site–site distribution functions for various concentrations. Top panel ($O–O_w$), middle panel ($C–O_w$) and bottom panel ($CH_3–O_w$). Colour convention is the same as in Fig 7.3.

Snapshots of mixtures at x = 0.20, 0.50 and 0.80 mole fraction of TBA are shown in Fig. 7.6. At x = 0.20 mole fraction of TBA, alcohol molecules form tight pockets inside the water network. A

bi–continuous structure of TBA and water is formed at x = 0.50 and at x = 0.80, water molecules appear as dimers and monomers.

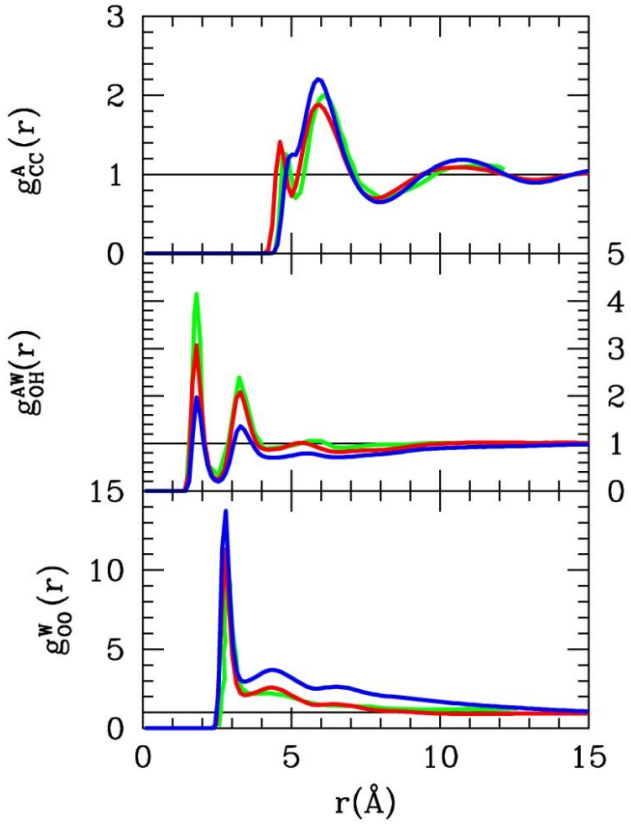

Figure 7.5: Selected site–site distribution functions for mixture x = 0.50. Top panel: TBA C–C correlations, middle panel: cross correlations O–H_w and bottom panel: water O_w–O_w correlations. Red is for OPLS model, blue is for TraPPE–UA and green is for the mixture model from Ref. [160].

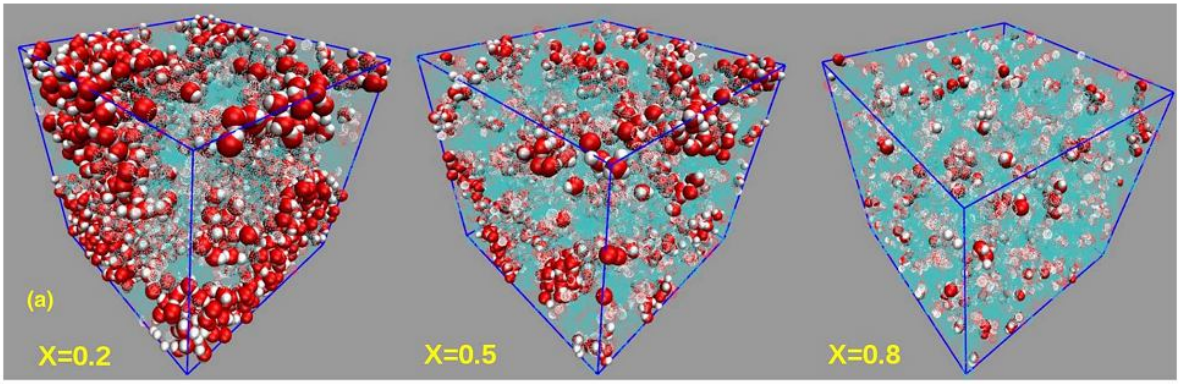

Figure 7.6: Snapshots for x = 0.20 (left), x = 0.50 (middle) and x = 0.80 (right) mole fraction of TBA. Colour convention: oxygen atom (red), hydrogen atom (white) and methyl group (cyan). TBA is shown as semi–transparent molecule in all snapshots.

Volumes, enthalpies and excess quantities for the TBA–water mixture are shown in Fig. 7.7. Data from two different TBA force fields (TraPPE–UA and OPLS) with SPC/E water are compared with experimental [11, 163] and simulation [160] results. Due to the disagreement of $\delta E \approx 10$ kJ/mol between TraPPE–UA energy and the experimental data for pure TBA, I added the correction of $10x$ kJ/mol to the calculated data for the mixture, similar to the correction for pure SPC/E water. It is seen that all the models capture quite well the thermodynamics of the real aqueous TBA mixture. However, there are noticeable differences in the excess quantities, especially the excess enthalpies, where none of the models captures the experimental S–shaped curve. This suggests that these models do not capture various energy changes at the proper concentrations in the TBA–water mixture.

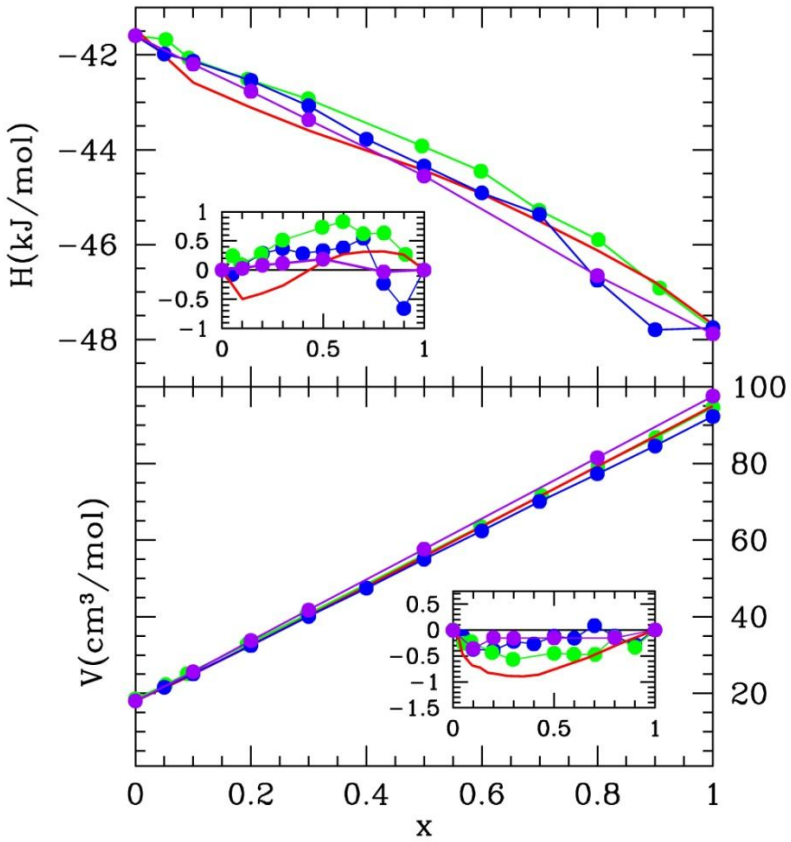

Figure 7.7: Enthalpies (top panel), excess enthalpies (inset on the top panel) volumes (bottom panel) and excess volumes (inset on the bottom panel) for the TBA–water mixture. Line and colour convention: experimental data from Ref. [163](top panel) (red lines) and Ref. [11] (bottom panel) (red lines), simulation results from Ref. [160] (green dots), OPLS TBA model (blue dots), TraPPE–UA TBA model (purple dots). Lines connecting dots are guidelines.

Fig. 7.8 shows the KBI as obtained from the TS correction procedure, by using both OPLS and TraPPE–UA TBA models, and data from SAXS [165], SANS [172] and calorimetric experiments [4, 172]. In view of the large dispersion of the various experimental results, the KBI data agree quite well with the experimental trends.

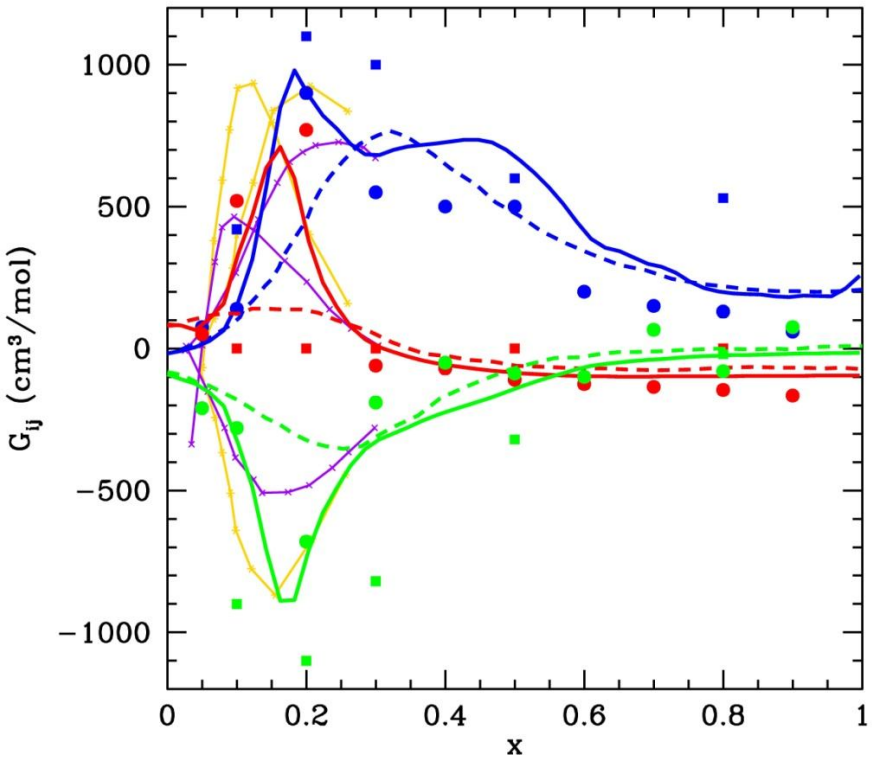

Figure 7.8: Corrected KBI for the two TBA models. Symbol convention: OPLS TBA model (filled dots), TraPPE–UA TBA model (filled squares). Colour convention: G_{WW} (blue), G_{AW} (green), G_{AA} (red). Line convention: experimental results from Ref. [172] (continuous lines), experimental data from Ref. [4] (dashed lines), SAXS data from Ref. [165] (magenta lines), SANS data from Ref. [172] (yellow lines).

7.3.3 The structure and thermodynamics of aqueous acetone mixture

Analysis of the radial distribution functions for aqueous acetone is presented over the whole concentration range. Fig. 7.9 shows the radial distribution function between the two oxygen sites on water molecules. It is seen that the first peak of $g_{WW}(r)$ gradually increases as water becomes mole dilute and reaches its maximum height for the mole fraction of acetone x = 0.80. Water molecules become more correlated as water network rarefies, as it is expected. Water correlations also increase in the second and the third neighbour shell, but decay when reaching the concentration x = 0.90. Fig. 7.10 shows O–O, C–C and CH_3– CH_3 correlations between the two acetone molecules. One observes that all the acetone–acetone correlations grow as acetone becomes the minority species in the mixture. This suggest the clustering of acetone molecules

occurs in the same way as of water molecules, although the rate of increase of the correlations is slower and the maximum of the correlations is observed for the mole fraction of acetone x = 0.10.

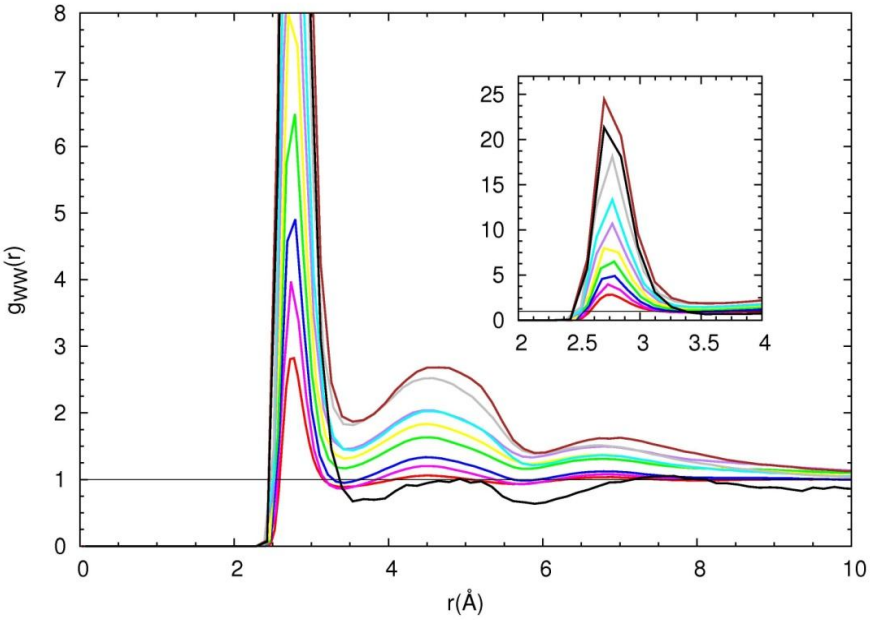

Figure 7.9: Water O–O site distribution functions for acetone mole fractions from x = 0 (red curve) to x = 0.90 (black curve). Inset shows the zoom on the first peak of the $g_{WW}(r)$. Colour convention is the same as in Fig. 7.2.

Oxygen–oxygen acetone correlations (top panel of Fig. 7.10) increase with decreasing x, while the position of the split first peak remains the same. The behaviour of the carbon–carbon and methyl–methyl correlations with respect to mole fraction x (middle and bottom panel of Fig. 7.10) is similar to that of the O–O correlations. Because of the increasing acetone content, the number of H–bonds that acetone forms with water increases, resulting with decreasing correlations of acetone molecules with its own kind. Fig. 7.11 shows the O–O_W, C–O_W and CH_3–O_W cross correlations. Oxygen–oxygen cross correlations (top panel) grow as acetone mole fraction x increases, which is consistent with the fact that O–O acetone correlations decrease with the same trend. The C–O_W and CH_3–O_W correlations increase with increasing x, reaching their maximum height at x = 0.90.

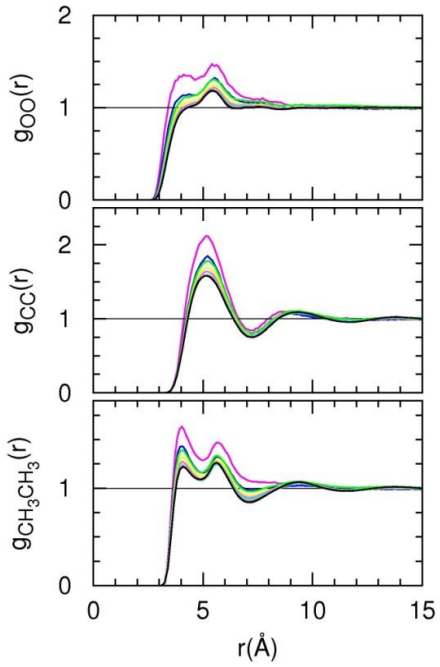

Figure 7.10: Selected acetone site–site distribution functions for various concentrations. Top panel (O–O), middle panel (C–C) and bottom panel (CH_3–CH_3). Colour convention: x = 0.90 (black), x = 0.80 (brown), x = 0.70 (gray), x = 0.60 (cyan), x = 0.50 (purple), x = 0.40 (yellow), x = 0.30 (green), x = 0.20 (blue) and x = 0.10 (magenta).

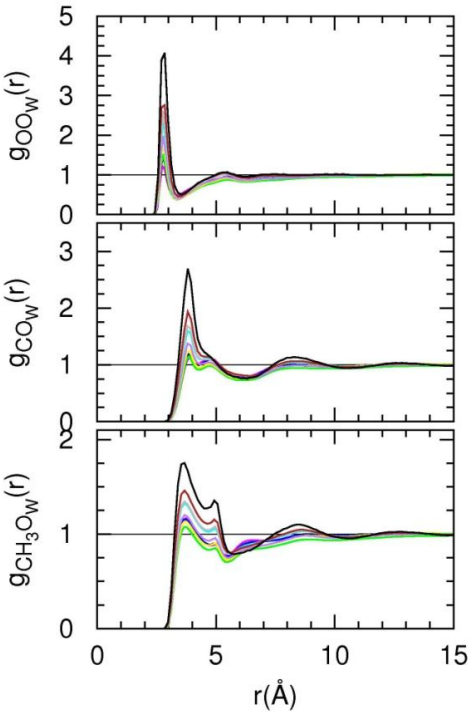

Figure 7.11: Selected cross site–site distribution functions for various concentrations. Top panel (O–O_W), middle panel (C–O_W) and bottom panel (CH_3–O_w). Colour convention is the same as in Fig 7.10.

Fig. 7.12 shows shapshots of mixtures at x = 0.20, 0.50 and 0.80, where micro–heterogeneity is seen. For mixture with x = 0.20 in the left panel, one observes that there are holes inside water network where water segregates acetone molecules. This type of Swiss–cheese structure is kept for all mixtures until x > 0.50, when water becomes minority and makes closely packed linear and globular clusters.

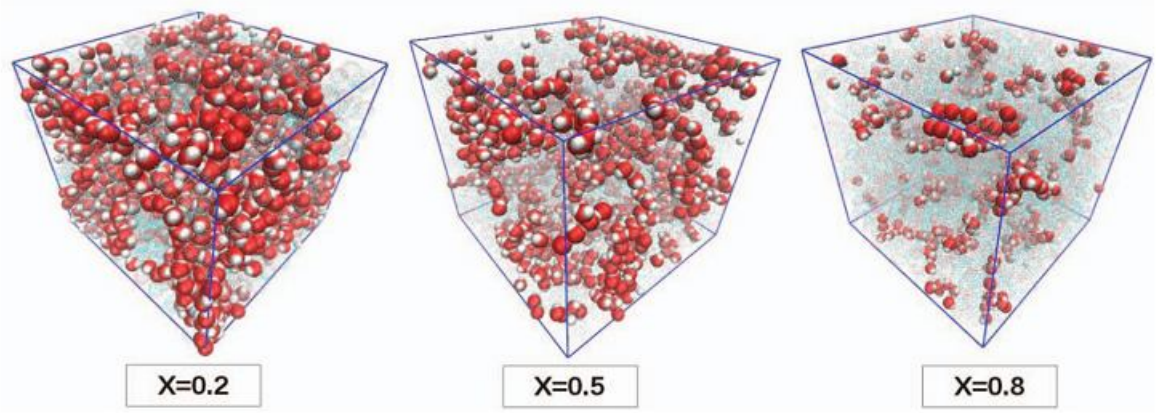

Figure 7.12: Snapshots of acetone–water mixtures for acetone mole fractions x = 0.20 (left), x = 0.50 (middle) and x = 0.80 (right). Oxygen is shown in red, hydrogen in white and carbon and methyl groups in cyan. Acetone molecules are shown as semi–transparent in all shapshots.

Enthalpies, volumes, and their excess quantities were studied for acetone–water mixture over the whole concentration range. These data are shown in Fig. 7.13 where the results of TraPPE–UA acetone and SPC/E water model mixture are compared with the experimental results [173, 174], OPLS–SPC/E and WS–SPC/E model mixtures from Ref. [170]. As seen in the inset of the top panel in Fig. 7.13, TraPPE–UA model agrees in an excellent way with the experimental enthalpy, while this is not true for other models. It also captures well the shape of experimental excess volume, although slightly worse than the WS model. TraPPE–UA model was initially designed to reproduce the entire phase behaviour of acetone [171] so it is not surprising that it gives good results for the thermodynamic parameters of acetone aqueous mixtures.

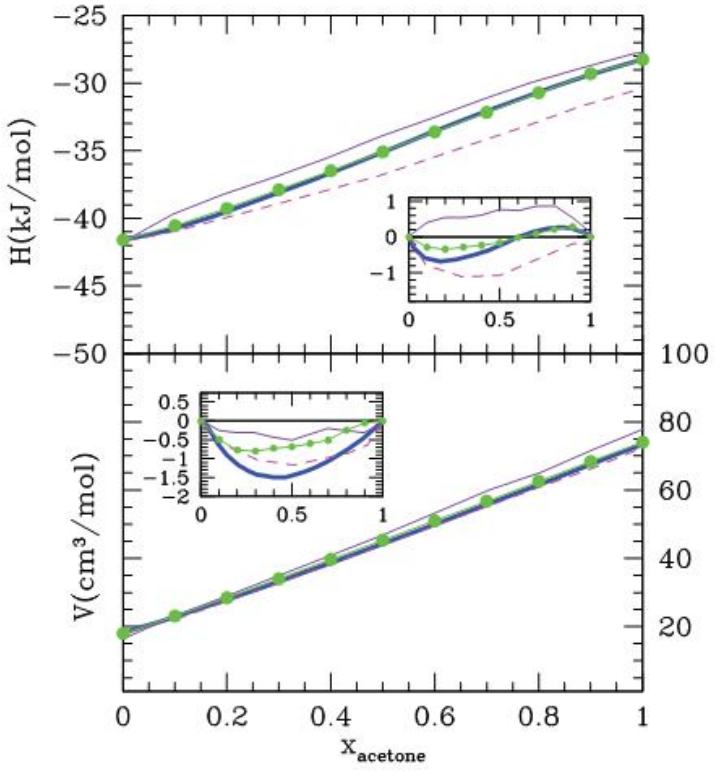

Figure 7.13: Top panel: enthalpy. Bottom panel: volume. Excess quantities are shown in insets. Symbols: TraPPE–UA–SPC/E (dots), experimental results from Ref. [173] in top panel and from Ref. [174] in bottom panel (blue line), OPLS–SPC/E [170] (magenta line), WS–SPC/E [170] (magenta dashed line).

7.4 TS extension for aqueous acetone mixture

The TS extension procedure was described in details in Chapters 6 for aqueous TBA mixture. I have demonstrated that in the case of aqueous TBA mixtures, the tail of the site–site correlation functions can be modulated by an exponentially decaying sine function. The function has the following form:

$$g_{ij}(r \to \infty) = \frac{A_{ij}}{r} \exp\left(-\frac{r}{\xi}\right) \sin\left(\frac{r}{\bar{d}}\right), \tag{7.1}$$

where d is the mean domain size, with $\bar{d} = d / 2\pi$, ξ is the correlation length, and A_{ij} is a factor that depends on each of the site–site functions. Each of the radial distribution functions $g_{ij}(r)$ is

extrapolated from $r = L/2$, where L is the box size. Factors A_{ij} are chosen to match the value of $g_{ij}(r)$ at end point and keep it a smooth function,

Figs. 7.14 and 7.15 show two examples of the TS extension procedure applied to the pair correlation functions in TraPPE–UA acetone and SPC/E water mixture. Fig. 7.14 shows the water–water correlations at x = 0.20 mole fraction of acetone and Fig. 7.15 shows acetone–acetone correlations at x = 0.30. Both correlations are between oxygen sites. Top panels of both figures show that the TS extension of $g_{ij}(r)$ needs to be a smooth continuation of the simulation result which brings the KBI shown in the middle panels, to a value much closer to the experimental KBI. Structure factors are shown in the bottom panels and it is seen that the TS procedure leads to a pre–peak at small–k. This is typical for micro–emulsions and the formation of domains. It is found empirically that the parameter d varies from 10 Å to 20 Å for different concentrations, with no particular trend and $\xi \approx 10$Å. The behaviour of d may be explained by the inappropriate box size which could not accommodate the micro–heterogeneous structure. Larger box size may alter both parameters, but I conjecture that ξ will remain close to the present value and d might change in a smoother way with x. Before applying the TS extension procedure, in some cases the asymptote of the radial distribution function had to be shifted to the correct value. As discussed in Chapter 3, the correlation functions from simulations do not always tend to unity, but they are slightly shifted due to the finite size box. Lebowitz and Percus [100] demonstrated that the shift depends both on the system size N and the partial compressibility. Therefore, the tail of the radial distribution functions has to be shifted to unity through the procedure described in Chapter 3. Without this correction, the corresponding KBI does not reach the proper horizontal asymptote but curves downwards. In practice, the LP correction is used for the correlations of the majority component. This is the $x < 0.20$ region for water and $x > 0.70$ for acetone.

Fig. 7.16 shows the KBI for TraPPE–UA–SPC/E model mixture with the TS procedure, two sets of experimental [4, 175] KBI and the simulation results [170] for the TIP4P/FMKH model mixture at x = 0.30. It is found that the TS extension brings the KBI close to their experimental values. The TIP4P/FMKH model mixture for one single concentration produces KBI that agrees well with the other data for the same concentration. As seen from Fig. 7.16, the water–water KBI are not in such good agreement with experiments as the two other sets of KBI, but the experimental data themselves differ for this particular set of KBI.

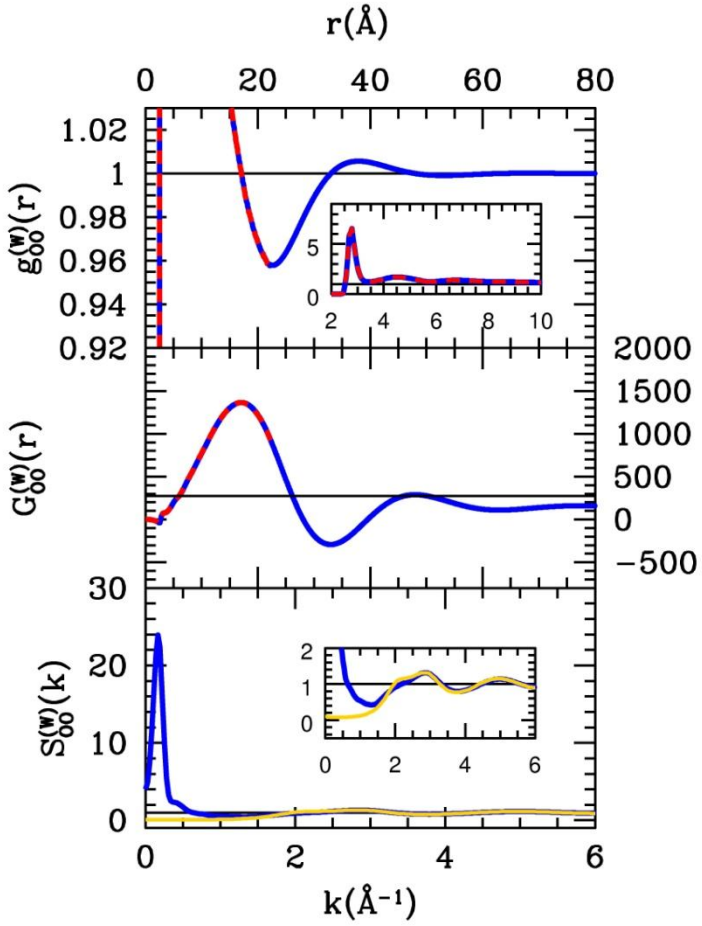

Figure 7.14: Illustration of the TS extension procedure for water–water correlation function $g_{oo}^{(W)}(r)$ at acetone mole fraction x = 0.20. Top panel: tail of the $g_{oo}^{(W)}(r)$. Inset shows full correlation function. Colour convention: simulation data (dashed red line), TS extension (blue line) with parameters d = 20 Å and $\xi = 10$Å. Middle: the RKBI corresponding to top panel. Horizontal line is the experimental KBI from Ref. [4]. Bottom: site–site structure factor. The yellow line is the structure factor of neat SPC/E water.

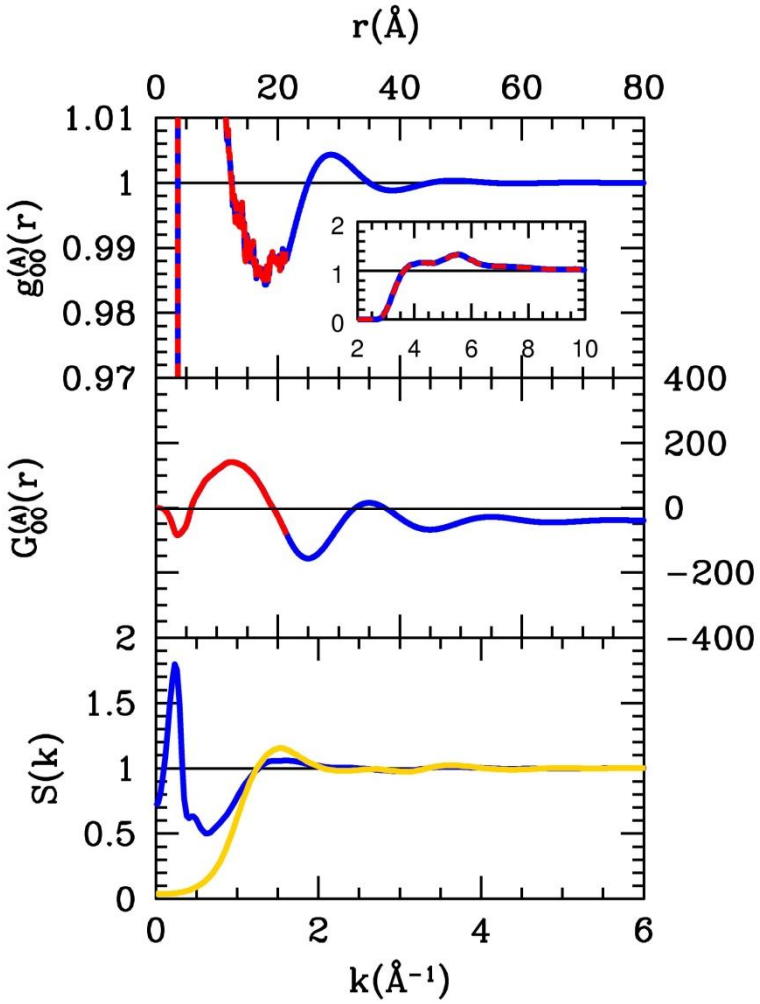

Figure 7.15: Illustration of the TS extension procedure for acetone–acetone correlation function $g_{oo}^{(A)}(r)$ at acetone mole fraction x = 0.30. The TS parameters are $d = 40$ Å and $\xi = 10$Å. Colour and line conventions are the same as in Fig. 7.14. Horizontal line is the experimental KBI from Ref. [4]. Bottom panel shows the site–site structure factor of neat TraPPE–UA acetone.

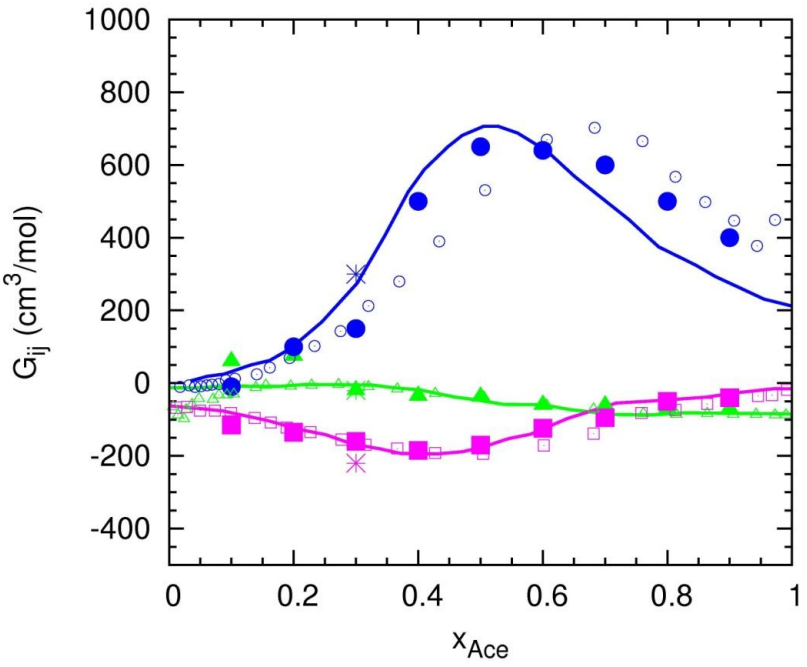

Figure 7.16: KBI of aqueous acetone mixtures. Symbol convention: present results with TS extension (filled symbols), experimental results from Ref. [4] (lines), experimental results from Ref. [175] (open symbols), TIP4P/FMKH model mixture from Ref. [170] at x = 0.30 (stars). Colour convention: blue circles for G_{WW}, magenta squares for G_{AW}, green triangles for G_{AA}.

7.5 Conclusion

Structural and thermodynamic properties of aqueous TBA and aqueous acetone were studied by using computer simulations. In particular, the concept of molecular–emulsion was used to describe the micro–heterogeneous structure of these mixtures. This concept allows one to extend the long range part of the correlations, by taking into account its oscillatory behaviour. Due to micro–heterogeneity, it is required to simulate these systems for a size that allows one to see how domain correlations start to develop from the short range molecular correlations. If the system size is too small, these correlations could be distorted by the periodic boundary conditions and the results in the long range might not correspond to the real correlations. Large scale simulations were not performed, but with the system size of N = 2048 particles, the results are good enough to show the short range structure of the mixture and the long range part was fitted through the TS extension. I conjecture that the results for a larger system would not differ significantly from the current ones, knowing that the values associated to the long range correlations are quite small. The analogy between these aqueous mixtures and micro–emulsions

is a legitimate one. Micro–emulsions are ternary mixtures of water, oil–type solute and a surfactant and a simplified micro–emulsion consists just of water and a surfactant with a large oily tail [149]. In the limit of a very small surfactant molecule, one can consider acetone or TBA molecule. Both molecules have CH_3 groups that repel water. Acetone has a CO group and TBA has an OH group that attracts it. Due to the amphiphilic structure of acetone and TBA molecules, these aqueous mixtures are micro–heterogeneous. This type of structure differs from the one in micro–emulsions because the oil–water domains are quite small so they become indistinguishable from concentration fluctuations that exist in any mixture. Information about micro–segregation is contained in the short range part of the correlation functions, while the long range part accounts for concentration fluctuations. If the integration of the pair correlation function includes only the short range part, the KBI reach excessively high values, indicating strong self–aggregation. Therefore, it is essential to sample enough domain statistics to obtain information about concentration fluctuations. The methodology used in this Chapter allows extending the correlations in the long range part which brings the corresponding KBI close to their experimental values. During the simulation of these aqueous mixtures and more generally, of all aqueous mixtures that exhibit micro–heterogeneity, clusters of water molecules coarsen very slowly, leading to domain formation which requires large simulation times.

Some authors have modified the TBA force fields in order to scale the micro–heterogeneity and bring the KBI to the proper values [160]. These modifications did not affect the thermodynamic properties, such as enthalpy for example, because the major contribution to these quantities comes from the short range correlations. In order to decide between the two different approaches, clear experimental criterion is required. For example, if the experiments show a pre–peak in the structure factor $S(k)$, this will confirm the presence of domains inside the mixture and validate the present approach. If this is not the case, then reducing the micro–segregation by altering the solute force field is the proper way of treating these systems. There are two sets of experimental data available; first set is from SANS experiments [161, 162, 168] and the second is from SAXS experiments [165, 176]. In the first data set, it is very hard to tell from the scale of the plots weather there is a pre–peak in the small–k range of $S(k)$ or not. In Ref. [176] the SAXS results are displayed in log–scale and it seems like there could be a pre–peak at $k \approx 0.0015$ Å^{-1} for $x = 0.085$ mole fraction of TBA. It is unfortunate that the available experimental data are not solid enough to support our conjecture. Therefore, experimental revisiting of these systems is required and a search for any evidence of domain formation akin to the one found in micro–emulsions.

CHAPTER 8

Fluctuations and micro–heterogeneity in aqueous mixtures of modified water models

8.1 Introduction

Water is considered to be a complex disordered liquid with a local tetrahedral order [177]. The nature of that order is still not well understood and there are two main approaches to this subject. The first one is inspired by the earlier work of Franks [178] where water is considered to be made of hydrogen–bonded domains mixed with disordered molecules. It might even be that water itself is a mixture of an ordered and disordered form of the same liquid and the corresponding liquid–liquid critical point is hidden somewhere in the phase diagram [144]. Alternative way of interpreting the particularity of the water structure is by emphasizing its smallness in such manner that size effects play a leading role. Angell and co–workers [179] have examined the importance of geometry and tetrahedral interactions in substances such as water and silica. The effects of geometry were further studied by Linden–Bell and co–workers [180-182] as they varied the geometry of charges on a water molecule, while keeping the charge magnitudes the same.

Similar to this latter approach, I studied a mixture of real water and a water model which has been modified in such a way that the molecule's geometry is the same as water, but the charges are multiplied by a factor λ [183, 184]. Three different models have been examined, with $\lambda = 1/3$, 2/3 and 4/5, which offered a clear insight in the importance of electrostatic interactions, namely the hydrogen bond, in the structural and thermodynamic properties of aqueous mixtures. For the first value of $\lambda = 1/3$, it was found that the model does not mix with water, except at solute mole fractions of $x = 0.1$ and $x = 0.9$. The second model, $\lambda = 2/3$, mixes well with water at all concentrations and shows a strong micro–segregation of the like species. In the mixture of

water and $\lambda = 4/5$ model, moderate micro–heterogeneity has been observed over the entire concentration range.

This study has demonstrated that one does not have to necessarily have hydrophobic groups to produce a hydrophobic effect, as observed in many alcohol–water mixtures [126, 138, 161], but it can also occur as a consequence of solely interactions (electrostatic ones being more important than the Lennard–Jones) and without invoking any size difference between the two types of molecules.

8.2 Models and simulation details

The SPC/E model was used to model the water interactions. The simplest solute miscible with water is a molecule of the same geometry as SPC/E water, but with partial charges scaled down by a factor λ. The oxygen site in SPC/E water has the parameters $(\varepsilon_0, \sigma_0) =$ (0.65 kJ/mol, 3.165 Å), where ε_0 is the energy parameter in the OO LJ interaction and σ_0 is the diameter of oxygen. Partial charge on site H is $q_H = 0.4238e$ and on site O is $q_O = -0.8476e$, where e is the charge of the electron. Each H site is distant $d = 1$ Å from the central O site, with the angle $\angle HOH = 109.47°$. In what follows, I will denote by a subscript "W" all quantities related to ordinary water and with "w", those related to "weak water". "Weak water" models have the same geometry as water, but the charges on sites are $q_Y = \lambda q_H$ for site Y which corresponds to H site in real water and $q_X = \lambda q_O$ for site X which corresponds to O site. Three different values of λ have been studied, λ = 1/3, 2/3 and 4/5. Table 7.1 gives the model parameters for all liquids in the study.

Since all these "waters" have smaller charges than real water, the LJ parameters had to be larger than in real water in order to keep them in the liquid state at ambient conditions. Water being in the liquid state is a result of the H–bond interaction which is an electrostatic effect. If one removes the charges on SPC/E water by setting λ = 0 under ambient conditions, it would become a gas like nitrogen. Therefore, the LJ values for "weak waters" were set in order to keep their molar volumes somewhere between 28 cm^3/mol and 18 cm^3/mol. The first value is the molar volume of many liquids close to their triple point and the second value is the molar volume of neat water.

For "water" with $\lambda = 1/3$, ε_0 was set to 2.269 kJ/mol, which is 3.5 times that of the SPC/E water, for "water" with $\lambda = 2/3$, ε_0 was set to 1.292 kJ/mol, double the value for the SPC/E water and for "water" with $\lambda = 4/5$, ε_0 was set to 0.972 kJ/mol, which is 1.5 times that of the SPC/E water. In the third case, the molar volume was closer to the one for real water and in the first two cases the value was closer to 28 cm^3/mol. It was found that it is quite hard to keep these model solutes in the dense liquid region at $T = 300$ K and $p = 1$ atm while keeping their molar volumes close to that of water.

This is due to the fact that anisotropic (tetrahedral) interactions play a key role in the structural organization of water and if one changes them, it is not simply by increasing the isotropic LJ interaction that one can always keep the system in the liquid state.

Table 8.1: Model parameters

λ	σ (Å)	ε_o (kJ/mol)	q (H)	θ (°)
0	3.165	2.601	0	109.47
1/3	3.165	2.269	0.1413	109.47
2/3	3.165	1.292	0.2825	109.47
4/5	3.165	0.972	0.3390	109.47
1 (SPC/E)	3.165	0.65	0.4238	109.47

All the simulations were conducted in the NPT ensemble under ambient conditions with N = 2048 particles. The DLPOLY–2 simulation package [92] was used. The integration time step was set to 2 fs and Ewald summation part technique was used to handle the electrostatic part of the interactions. Mixtures of "weak water" and SPC/E water were studied over the whole concentration range of "weak water", from x = 0.1 to 1, with step size Δx = 0.1, where x is the mole fraction of the solute. In a few cases, intermediate points were studied. Each mixture was first allowed to reach equilibrium for 0.5 ns and the production runs were ranging from 1 to 3 ns. Mixtures with x = 0.2, 0.5 and 0.8 were studied in details, since they represent equimolar mixture and mixtures rich with one of the two components. The values of excess energy and volume for these systems did not fluctuate much over the number of runs, which is in contrast with the previous simulations of realistic models [10]. This can be attributed to the fact that smaller molecules of "weak water" can easily explore the phase space and find the proper arrangement

much faster than bigger molecules. Indeed, it is necessary to wait exceedingly long times until the system that contains many–site molecules gains smooth thermodynamic values.

Model with λ = 1/3 was found to de–mix already during the equilibration runs, except for x = 0.1. The λ = 2/3 and 4/5 models were studied over the whole concentration range. The monitored quantities were the molar volumes, the configurational energies and the LJ and the Coulomb contributions to them, as well as the site–site radial distributional functions. Different micro–states of the system were also monitored by looking at snapshots. This way it was possible to immediately tell that the system with λ = 1/3 was phase separating. The other two models exhibit micro–segregation, so it was insightful to look at snapshots.

8.3 Molecular Dynamics results

8.3.1 Structure of neat liquids

The upper panel of Fig. 8.1 shows the oxygen–oxygen radial distribution functions, $g_{OO}(r)$ for 3 different "weak water" models, SPC/E water and the λ = 0 case. All the $g_{OO}(r)$ functions for "weak water" show a typical Lennard–Jones–like oscillatory behaviour. The first peak of the $g_{OO}(r)$ in SPC/E water is narrower than the others, because of the specific H–bond interaction which results with 4, rather than 12 first neighbours around a molecule. The second peak is stretched and out of phase compared to the others. This comparison proves the uniqueness of the water structure, as reflected through the $g_{ab}(r)$.The 4/5 model shows an interesting behaviour, as it is closer to 2/3 model than to SPC/E water, which is surprising given the fact that it differs from water by the factor of just 1/5.

The inset in the top panel of Fig. 7.1 shows the corresponding coordination numbers, defined as:

$$C(r) = 4\pi\rho\int_0^r g(r')r'^2\,dr', \tag{8.1}$$

where ρ is the number density of the liquid. Coordination number is the number of neighbours within the first shell around a molecule, represented by the first peak of the $g(r)$.Thus, the integration is taken up to the first minimum of the $g(r)$. This gives 4 for water, but for λ = 4/5 and λ = 2/3, it becomes 8–9. For λ = 1/3 and λ = 0, one has 10–11 and 12, respectively. Specific structure of water can be spot out also by looking at the Fourier transform of the $g(r)$, which is

the structure factor, $S(k)$ (bottom panel of Fig. 8.1). It is seen that only real water has the split peak in the small–k region, while all the others show a behaviour very close to that of the LJ liquid.

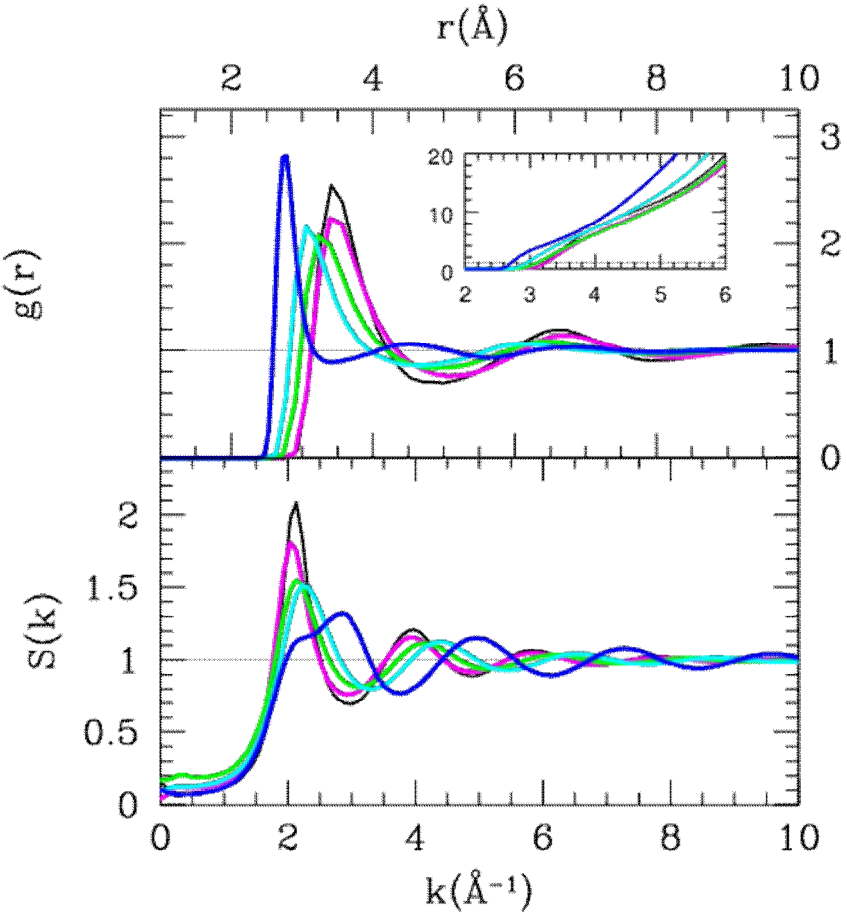

Figure 8.1: Oxygen–oxygen radial distribution function, $g_{OO}(r)$ (top panel) and oxygen–oxygen structure factor, $S_{OO}(k)$ (bottom panel) for 4 models and SPC/E water. Color convention: SPC/E water (blue), λ = 4/5 (cyan), λ = 2/3 (green), λ =1/3 (magenta), λ =0 (black). The inset in the top panel shows the coordination numbers.

8.3.2 Structure of mixtures

Structural properties of water–"weak water" mixtures were studied through the analysis of the site–site radial distribution functions and the Kirkwood–Buff integrals. Thermodynamic analysis was also conducted through the study of volumes, energies, their excess quantities and the diffusion coefficients.

Fig. 8.2 shows $g_{WW}(r)$, $g_{Ww}(r)$ and $g_{ww}(r)$ between two oxygen sites for different mole fractions of λ = 1/3 model in water, x = 0.20, 0.50 and 0.80. Snapshots of these mixtures can be seen on Fig. 8.3. Although, this model shows de–mixing for concentrations above x = 0.10, it is meaningful to examine the differences between $g_{ij}(r)$ of a micro–segregated and a phase separated system. It is seen that the tails of $g_{WW}(r)$ and $g_{ww}(r)$ develop a long range behaviour, which is a signature of concentration fluctuations at large distances. Perera and Sokolić [170] demonstrated that the competing attractive and repulsive interaction will result with cluster formation if the competition is high enough, but leads to phase separation if attraction prevails. In the present case, it is the attraction of water molecules trying to maintain their H–bond connectivity that completely segregates the "weak water" molecules. In consequence, $g_{Ww}(r)$ shows depletion, which is typical in simulation of a de–mixed system. One can also see on snapshots how the system phase separates.

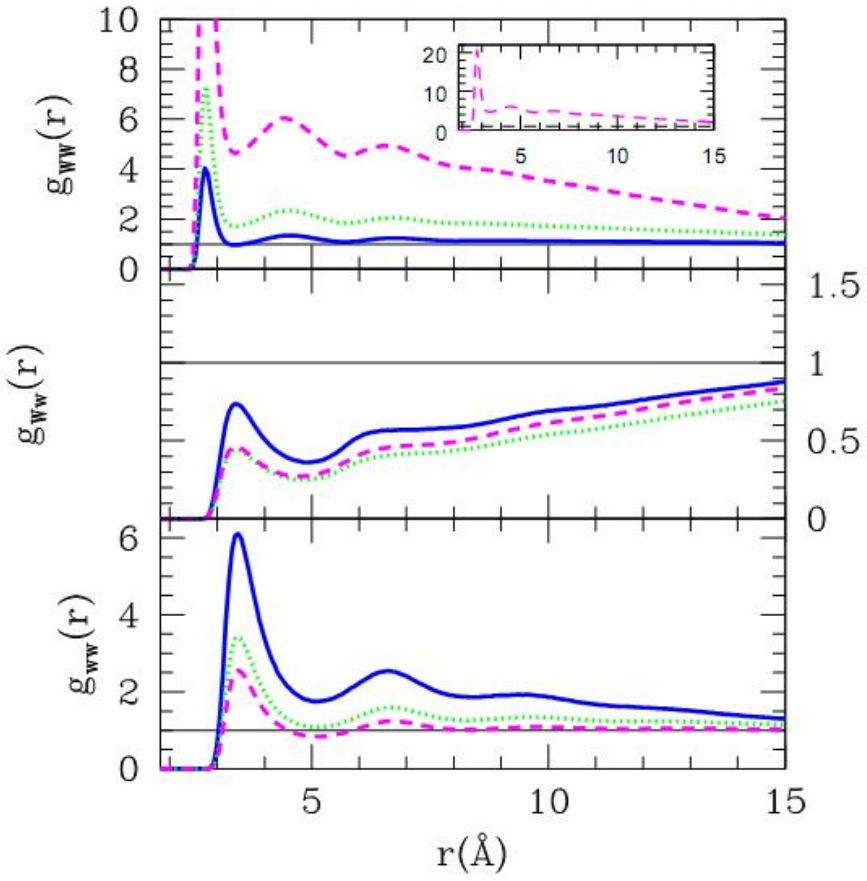

Figure 8.2: Oxygen–oxygen radial distribution functions $g_{ij}(r)$ for λ = 1/3 model at three mole fractions of "weak water". x = 0.20 (blue lines), x = 0.50 (green dotted lines), x = 0.80 (magenta dashed lines).

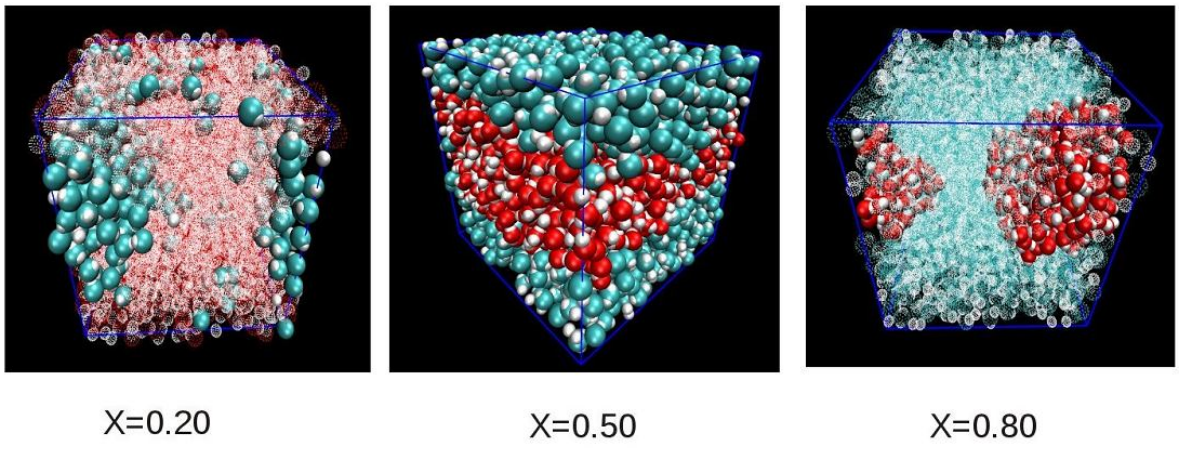

Figure 8.3: Snapshots of the $\lambda = 1/3$ model. Left panel: $x = 0.20$, middle panel: $x = 0.50$, right panel: $x = 0.80$. The oxygen atoms are in red, hydrogen in white and "weak water" oxygen in cyan. On the left and right panels the majority species is shown as semi–transparent.

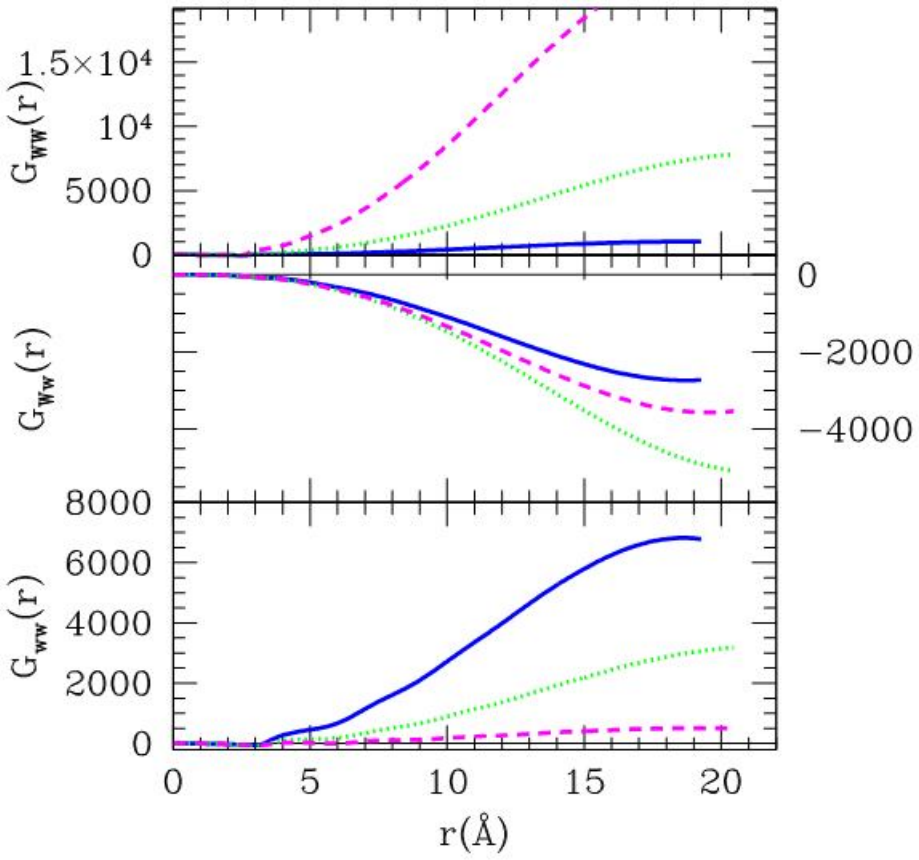

Figure 8.4: Running Kirkwood–Buff integrals, $G_{ij}(r)$ corresponding to the $g_{ij}(r)$ in Fig. 8.3.

Fig. 8.4 shows the corresponding running Kirkwood–Buff integrals, $G_{ij}(r)$. It is seen that the asymptotes are not reached which is another indication of the system phase separation. The same quantities were examined for λ = 4/5 model in water, which is the model that seems to mix the best with water. Fig. 8.5 shows three different $g_{ij}(r)$ for mole fractions x = 0.20, 0.50, 0.80 between two oxygen sites and Fig. 8.6 shows the corresponding $G_{ij}(r)$. As seen in Fig. 8.6, all the functions reach an asymptote. However, this asymptote is not reached in the simulations, since it has been shown by Lebowitz and Percus [100] that the radial distribution function in the finite size system does not go to 1, but rather to $1-1/\left[\left(\rho_a\sqrt{N_aN_b}\right)\left(\partial\rho_a/\partial\beta\mu_b\right)\right]$. Just by looking at the simulation $g_{ij}(r)$, one can conclude that they need a correction factor. For this purpose, the following expression was used:

$$g_{ij}^{corrected}(r)=g_{ij}^{MD}(r)\left[1+\frac{\delta_{ij}}{2}\left(1+\tanh\left(\frac{r-\kappa_{ij}}{\alpha_{ij}}\right)\right)\right], \tag{8.2}$$

where $g_{ij}^{MD}(r)$ is the uncorrected site–site radial distribution function from Molecular Dynamics simulation, the coefficients δ_{ij}, κ_{ij} and α_{ij} depend on the species pairs (i, j). δ_{ij} is defined as $\delta_{ij}=1/g_{ij}^{MD}(L/2)-1$, where L is the simulation box size. In practice, $\alpha_{ij}=1$ Å and $\kappa_{ij}=2.5d_{ij}$, where d_{ij} is the distance under the first peak of $g_{ij}^{MD}(r)$. This assures that the correction applies only past the first two layers around the molecule. The parameter δ_{ij} was set to 10^{-3} which is small enough to bring the asymptote to the proper value, as seen previously in these works [10, 101].

Without this correction, one would get the wrong KBI that show a maximum and then they decay, instead of showing a clear asymptote. This method has been applied to all the mixtures and the values are presented in Fig. 8.7. Other than statistically, one can also compute the KBI using the following three equations [3]:

$$G_{Ww}=\int d\vec{r}\left(g_{Ww}(r)-1\right)=k_BT\kappa_T-\frac{\overline{V}_W\overline{V}_w}{VD}$$

$$G_{WW}=\int d\vec{r}\left(g_{WW}(r)-1\right)=G_{Ww}+\frac{1}{1-x}\left(\frac{\overline{V}_w}{D}-V\right) \tag{8.3}$$

$$G_{ww} = \int d\vec{r}\left(g_{ww}(r) - 1\right) = G_{Ww} + \frac{1}{x}\left(\frac{\overline{V}_W}{D} - V\right),$$

where $\overline{V}_i$ is the partial molar volume of species *i*, defined as $\overline{V}_i = V - x_j \partial V / \partial x_j$, with $x_j = 1 - x_j$, V is the volume of the mixture, D is defined as $D(x) = x(\partial \beta \mu_w / \partial x)_{PT}$, where μ_w is the chemical potential of "weak water" and $\beta = 1/k_B T$.

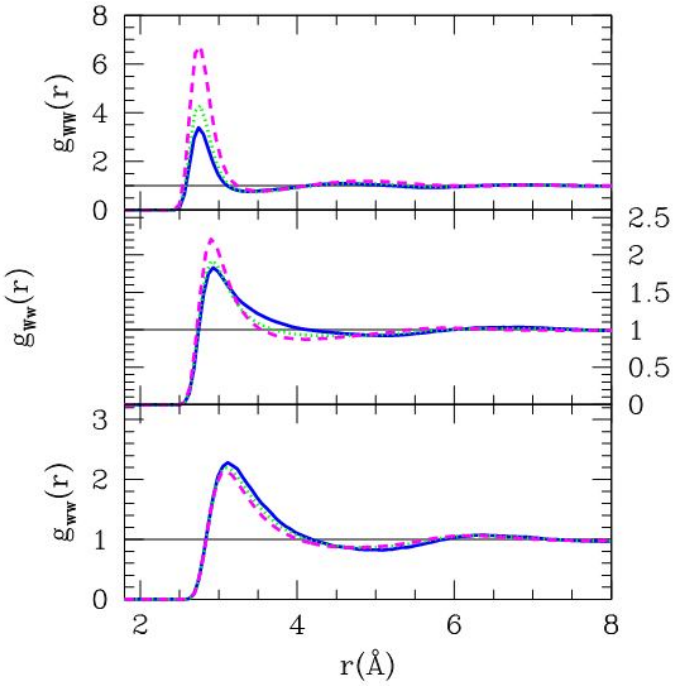

Figure 8.5: Oxygen–oxygen radial distribution functions $g_{ij}(r)$ for λ = 4/5 model at three mole fractions of "weak water". x = 0.20 (blue lines), x = 0.50 (green dotted lines), x = 0.80 (magenta dashed lines).

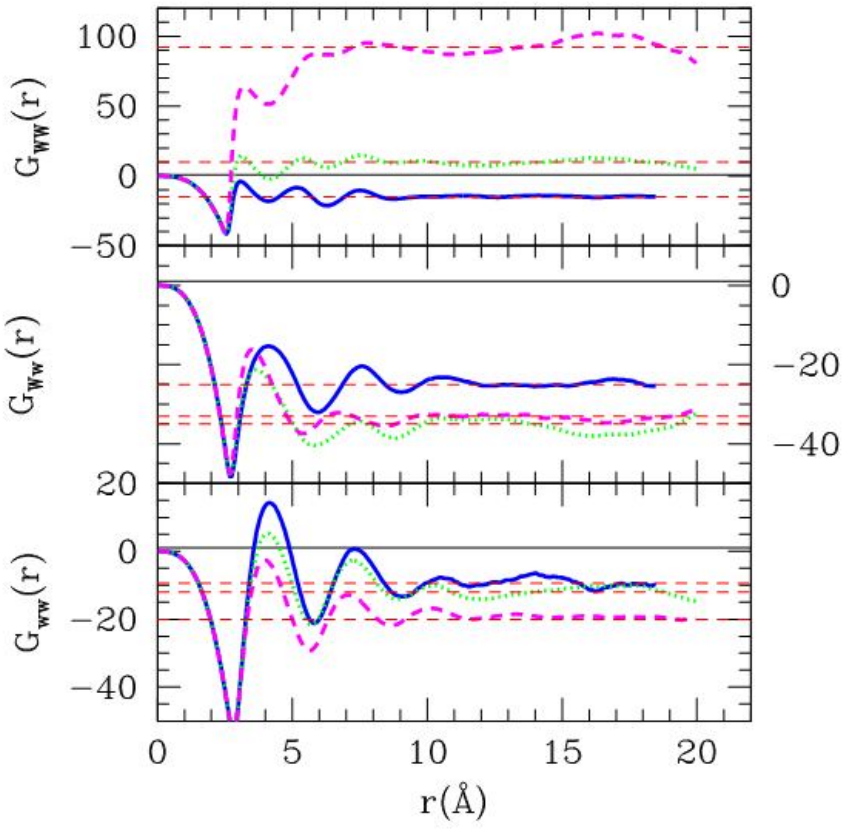

Figure 8.6: Running Kirkwood–Buff integrals $G_{ij}(r)$ corresponding to $g_{ij}(r)$ in Fig. 8.5. Dashed red lines correspond to the corrected asymptote (see text).

The first equation contains the isothermal compressibility term κ_T, but it is usually neglected since the liquids are incompressible. Combining the other two terms, one can write these expressions for $D(x)$:

$$D(x) = \frac{\overline{V}_W(x)\left[V(x) - x\overline{V}_w(x)\right]}{V(x)\left[x(G_{ww}(x) - k_B T\kappa_T) + V(x)\right]} \tag{8.4}$$

$$D(x) = \frac{\overline{V}_w(x)\left[V(x) - (1-x)\overline{V}_W(x)\right]}{V(x)\left[(1-x)(G_{WW}(x) - k_B T\kappa_T) + V(x)\right]}. \tag{8.5}$$

These two equations give the same result for $D(x)$, given the properly converged $G_{ij}(r)$ values. The molar volumes were fitted and the partial molar volumes were derived from it. The resulting KBI are shown in Fig. 8.7. Two sets of KBI were compared and it is seen that they match in the case of λ = 4/5.

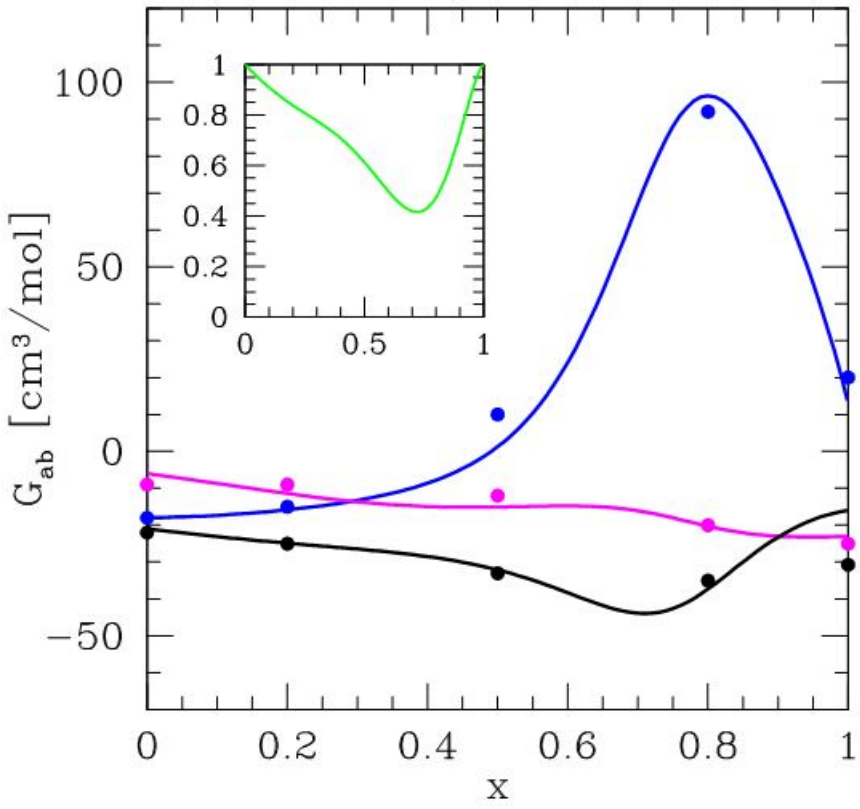

Figure 8.7: KBI of the λ = 4/5 mixture. Colour and line convention: G_{WW} in blue, G_{Ww} in black, G_{ww} in magenta. Dots are from direct integration of $g_{ij}(r)$ and lines from the thermodynamic route. The inset shows $D(x)$ as obtained from the procedure described in the text.

As seen from the snapshots in Fig. 8.8, this model shows moderate micro–heterogeneity when mixed with water. It seems that the two species behave symmetrically.

As seen in Fig. 8.9, oxygen–oxygen radial distribution functions, $g_{ij}(r)$ in the λ = 2/3 model are not so different from the 4/5 model, except for low water content where the function $g_{WW}(r)$ does not seem to go to unity, which indicates de–mixing behaviour. This is further confirmed in the Fig. 8.11 where KBI are shown. Fig. 8.10 shows the corresponding running KBI for three different mole fractions of the solute. One can see from Fig. 8.11 that in the rich "weak water" region, $G_{WW}(r)$ are very high, which is typical for realistic mixtures, where water–water correlations are exaggerated.

The KBI for the 2/3 model were computed by using two different routes, as in the 4/5 case. There is a pronounced discrepancy in the $x > 0.50$ region between the two sets of data for the KBI, as shown in Fig. 8.11. However, for mixtures with $x < 0.50$, the data match nicely, which

confirms the good quality of the integrated KBI and also the accuracy of the thermodynamic procedure for computing the KBI. The following empirical expression was used for $D(x)$:

$$D(x) = \left(1 - 3.85x(1-x)\right)\left(\frac{0.5 + 0.561}{\sqrt{(0.25+x)(1.05-x)}}\right). \tag{8.7}$$

Similar expression was used in the works of Perera et al. [172, 175], where the authors showed that the experimental functions for $D(x)$ could well be approximated by a simple deviation from the theory of regular solution, of the generic form $D(x) = 1 + \alpha\, x(1-x)$. This expression ensures $D(x) = 1$ for $x = 0$ and $x = 1$, as it should be. The α–parameter is determined from the experimental fit.

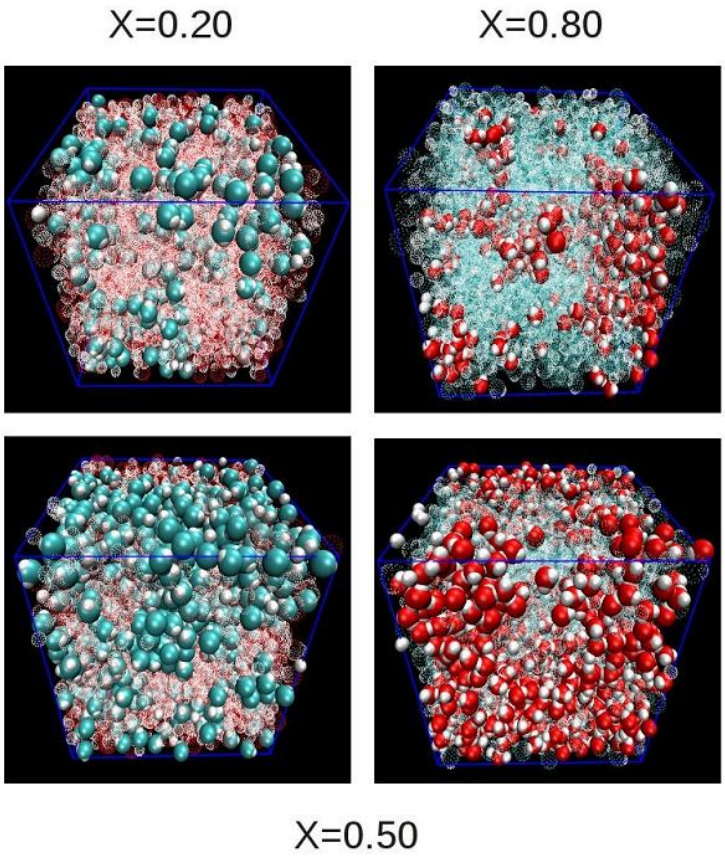

Figure 8.8: Snapshots of the 4/5 model. The colour convention is the same as in Fig.8.3. Top left panel: x = 0.20, top right: x = 0.80, two bottom panels: x = 0.50, the water and "weak water" molecules are shown in separate shots taken under the same angle.

The inset of Fig. 8.11 shows the fitted $D(x)$ and the one obtained from the integrated KBI. As one can clearly see on the inset, the data from the KBI calculation deteriorate seriously for $x >$ 0.50, which is due to the fact that the integrals did not reach the proper asymptotic value. The

third set of the KBI data in Fig. 8.11 (shown in filled shapes) is obtained from the procedure that is developed in analogy with micro–emulsions, where it is observed that the tails of $g_{ij}(r)$ are modulated by oil–water domains. This is the so called TS behaviour which is explained in details in Chapter 6 of this thesis. It predicts the following expression for the decay of the water–water distribution function:

$$g_{WW}(r) \to \frac{A}{r}\exp\left(-\frac{r}{\xi}\right)\sin\left(\frac{r}{\bar{d}}\right), \tag{8.8}$$

where ξ is the correlation length and $\bar{d} = d/2\pi$, d being the domain mean size. The exact derivation of Eq. (8.8) from the MOZ equation is given in Appendix B.

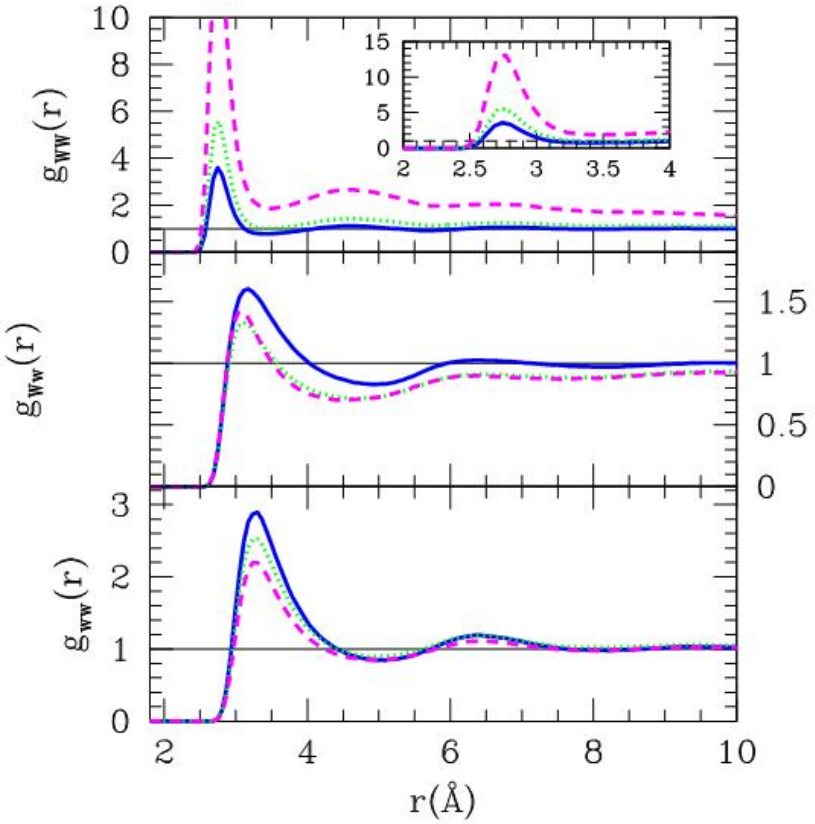

Figure 8.9: Oxygen–oxygen radial distribution functions $g_{ij}(r)$ for λ = 2/3 model at three mole fractions of "weak water". x = 0.20 (blue lines), x = 0.50 (green dotted lines), x = 0.80 (magenta dashed lines).

Following Eq. (8.8), all $g_{ij}(r)$ from simulations were extrapolated with this expression:

$$g_{ij}^{TS}(r) = g_{ij}^{MD}(r)\left[1 + \frac{A_{ij}}{r}\exp\left(-\frac{r}{\xi}\right)\sin\left(\frac{r - R_d}{\bar{d}}\right)H(r - R_c)\right], \tag{8.9}$$

where A_{ij}, R_d, ξ and $\bar{d}$ are fitting parameters that allow a smooth extrapolation of the current $g_{ij}(r)$, from $r = R_C$, with R_C as the half box size. All the parameters are the same for each $g_{ij}(r)$, except for the parameter A_{ij} which depends on the pair of species.

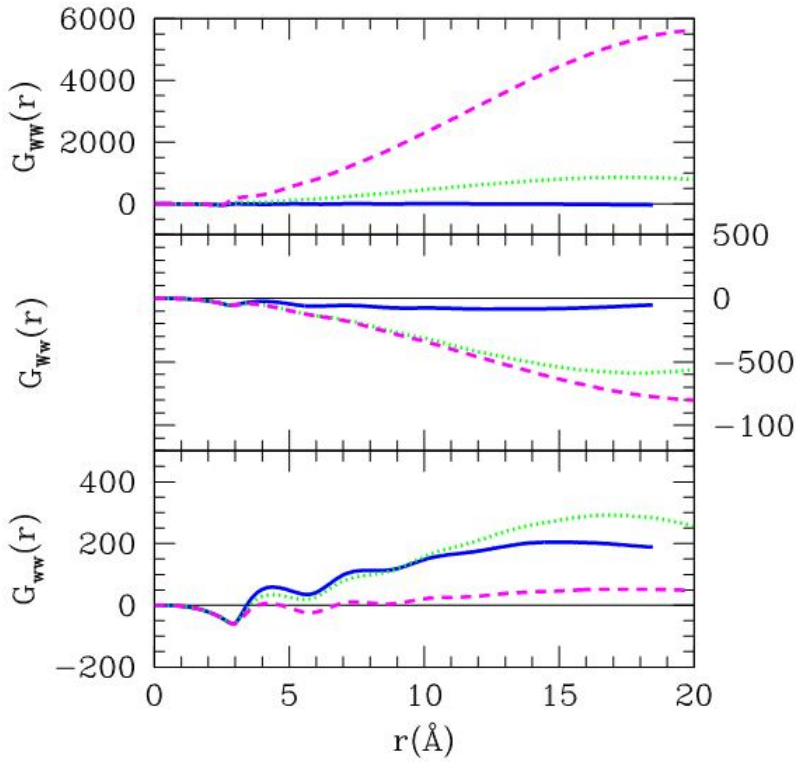

Figure 8.10: Running Kirkwood–Buff integrals, $G_{ij}(r)$ corresponding to the $g_{ij}(r)$ in Fig. 8.9.

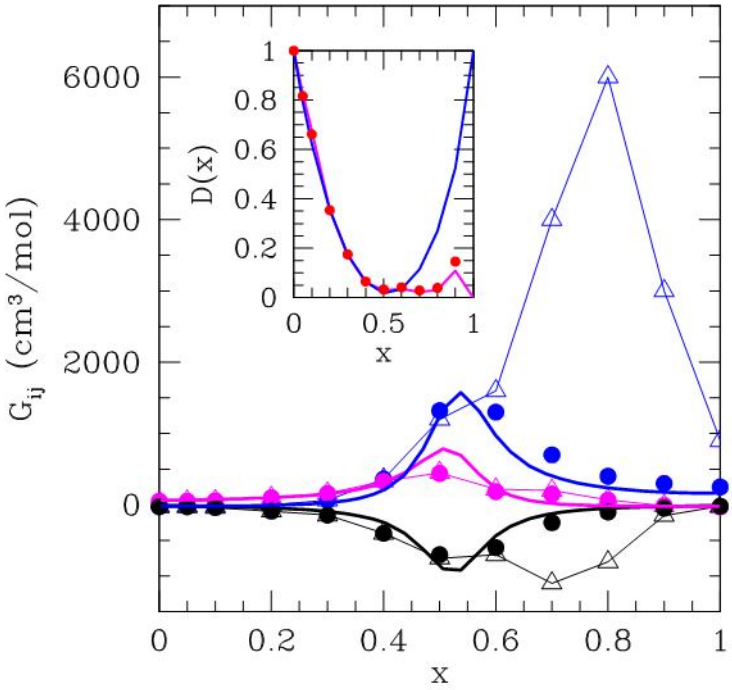

Figure 8.11: KBI of the λ = 2/3 mixture model. Colour and line convention: G_{WW} in blue, G_{Ww} in black, G_{ww} in magenta. Open triangles are from direct integration of $g_{ij}(r)$, thick lines from the thermodynamic route and filled dots are from the TS procedure (see text). The inset shows $D(x)$ as obtained from the integration of the KBI (dots) and from the fit (line).

Fig. 8.12 shows the TS extrapolation of water–water distribution function between two oxygen sites $g_{WW}(r)$, as well as the corresponding KBI and the structure factor $S_{WW}(k)$. One sees in the top panel of Fig. 8.12 that $g_{WW}(r)$ does not go to 1. If one just sets the function to 1, this will result with an exceedingly high KBI value, while with the TS extension, it is not the case, as seen in the middle panel of Fig. 8.12. Moreover, the TS extension leads the KBI value to a reasonably lower value and this happens after the transient regime which is about half of the box size. It should be noted that the TS extension has to be performed also for $g_{Ww}(r)$ and $g_{ww}(r)$ with the same parameters $\bar{d}$ and ξ in order to get consistent KBI values. In this particular case, it was set $\bar{d}$ = 16 Å and ξ = 5 Å. The value of A_{ij} is adapted for a particular *i*–*j* correlation and R_d is adjusted to match the sine function near $r = R_d$. The entire procedure has to be performed for mixtures with $x > 0.50$ and the LP correction is not applied in these cases, since the asymptote of $g_{ij}(r)$ is not well defined. This is due to the small box size which cannot accomodate systems

micro–heterogeneous behaviour. In the bottom panel of Fig. 8.12 , two forms of structure factor $S_{WW}(k)$ are shown.

$S_{WW}(k)$ which corresponds to the TS extension does not have the Ornstein–Zernike (OZ) form of the structure factor. Instead, there is a pre–peak in the small–k region, which is a proof of domain formation and the position of the pre–peak corresponds to the average size of the domains.

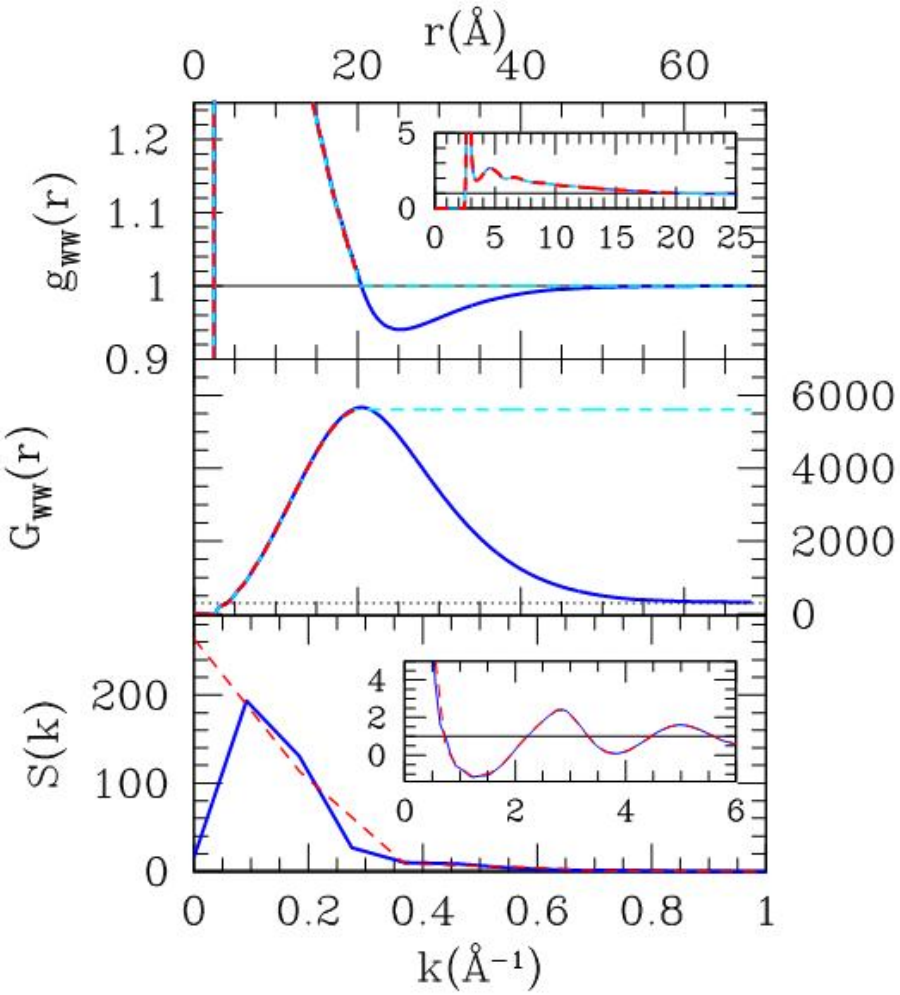

Figure 8.12: Detail of the TS procedure (see text) applied to water–water radial distribution function between two oxygen sites $g_{WW}(r)$ for the λ = 2/3 mixture at x = 0.80. Top panel: zoom on the tail extension with damped domain oscillations. The red dashes are the original $g_{WW}(r)$ from simulations, the cyan dashes are the trivial extension setting $g_{WW}(r)$ to 1 and the blue line is the TS extension (with $\overline{d}$ = 16 Å and ξ = 5 Å). Middle panel: the RKBI corresponding to $g_{WW}(r)$ shown in the top panel. The horizontal dashed line is the KBI expected from the thermodynamical route. Bottom panel: structure factors $S_{WW}(k)$ corresponding to $g_{WW}(r)$ shown in the upper panel, with the same color conventions. The inset shows the details near the main peak at $k_m = 2\pi/\sigma_W \approx 3$ Å^{-1}.

Looking at snapshots of this model, shown in Fig. 8.13, one does not notice clear domains, just compact water clusters. This suggests that this structure, whose signature is given by the form of $S_{ij}(k)$, could be seen if we simulate larger systems. The present results clearly show that the number of a few thousands of particles, generally considered to give meaningful results in the study of aqueous mixtures, is very limited as far as the long range correlations are concerned. The RKBI shown in Fig. 8.10 seem to have a transient form for at least 3 times the given box size, which gives the factor 3^3 and an overall number of 50 000 particles to unambiguously describe the long range structure of these mixtures, particularly aqueous mixtures [185, 186].

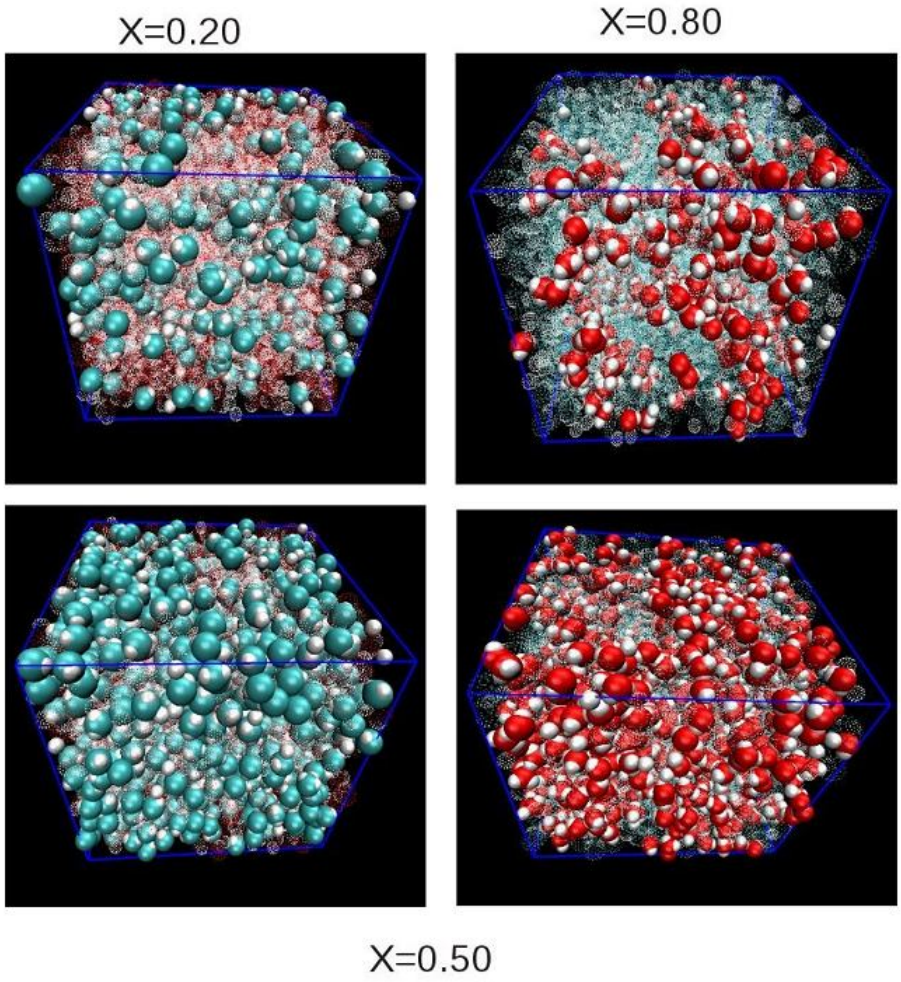

Figure 8.13: Snapshots of the 2/3 model. The colour convention is the same as in Fig. 8.3. Top left panel: x = 0.20, top right: x = 0.80, two bottom panels: x = 0.50, the water and "weak water" molecules are shown in separate shots taken under the same angle.

In order to confirm the validity of the TS extension, an MD simulation has been conducted for the 2/3 mixture model at x = 0.80 mole fraction of "weak water" where the system had twice the

number of particles of the previous simulations, N = 16 384. As observed in the inset of Fig. 8.14, there is a large number of small water domains in the larger box and these domains do not differ greatly from the ones in the smaller box. It was noticed that water molecules first form string–like aggregates which during simulation coarsen and become domains. This process takes a lot of time and after 1 ns, the domains still appear to be coarsening which affects the long range part of the correlation functions. This is the reason why small discrepancies are observed in the long range part of $g_{WW}(r)$. However, it is quite clear that the function indeed has the long range oscillatory structure, as suggested by the TS extension and that the parameters $\overline{d}$ = 16 Å and ξ = 5 Å are chosen properly.

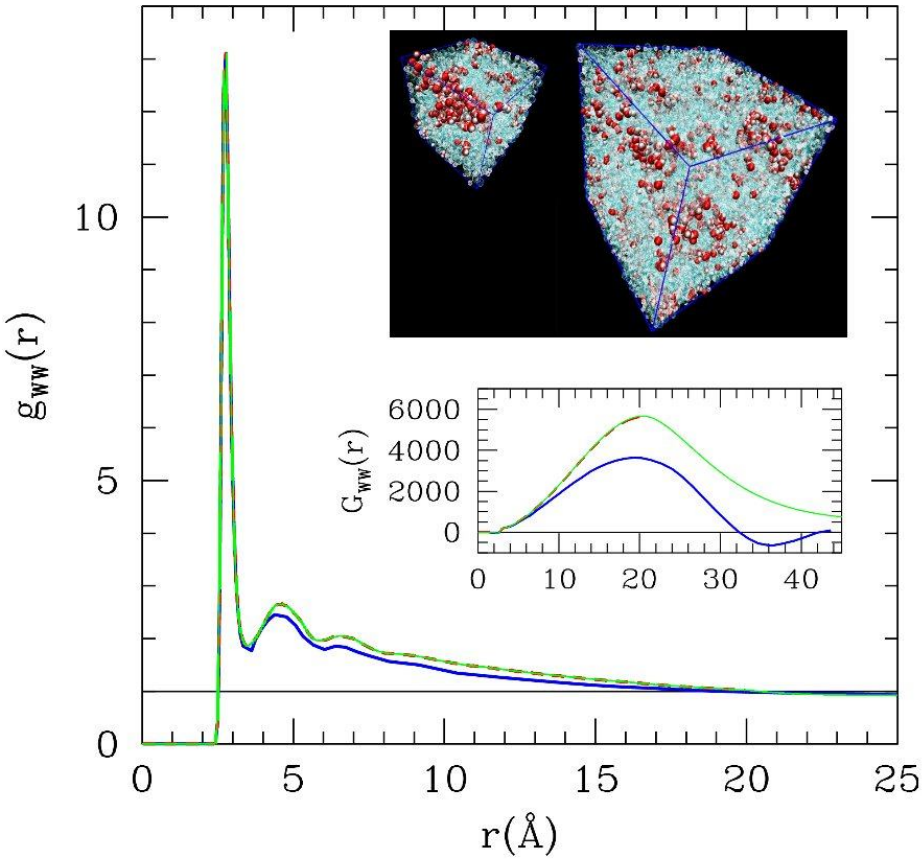

Figure 8.14: Comparison of the structure between the N = 2048 and N = 16 384 molecule systems, for λ = 2/3 at x = 0.8 mole fraction of "weak water". The main panel: $g_{WW}(r)$ between two oxygen sites. Colour convention: red dashed line for the small system, blue line for the large system and green line for the TS extension. The lower inset shows the corresponding KBI and the upper inset shows the snapshots of both systems (colour convention is the same as in Fig.8.3).

8.3.3 Thermodynamic analysis of mixtures

Volumes, energies, their excess quantities and diffusion coefficients were studied for all three types of mixtures. Excess quantity (volume or energy) is defined as $A^{ex}(x) = A(x) - A_W(1-x) - A_w x$, where A_W and A_w are values of this quantity for the pure substance, water and "weak water", respectively. First, the volumes and excess volumes are presented for mixtures with λ = 1/3, shown in Fig. 8.15. One sees at once how the volumes coming from two opposite sides of the concentration range do not converge to the same value at equimolar concentration. This could be due to the fact that this mixture is biphasic, which is also demonstrated by snapshots at Fig.8.3. Because of that, this mixture was not studied for concentrations other than for x = 0.10. The other two cases show negative excess volumes, as any mixture of hard spheres [8]. In hard sphere mixtures, negative excess volumes are driven by excluded volume effects, but in this example, both species have the same volume. It is the electrostatic property of the solute, namely the value of λ parameter, which has the leading role in the spatial organization of the mixture, leading to negative excess volumes. As seen in the bottom panel of Fig. 8.15, the 2/3 model has excess volumes that are more negative than the ones of the 4/5 model. The pure 2/3 model has a larger molar volume than the pure 4/5 model. Hence, molecules of the 2/3 model occupy more space and if mixed with water, they will have more free volume than the 4/5 molecules.

Total configurational energy, its van der Waals and Coulomb contribution, as well as their excess values for the 2/3 and the 4/5 model are shown in Figs. 8.16 and 8.17, respectively. These figures reveal a similar energetic portrait between the two models, showing positive total excess energies and the domination of the Coulomb part both in energy and in excess energy.

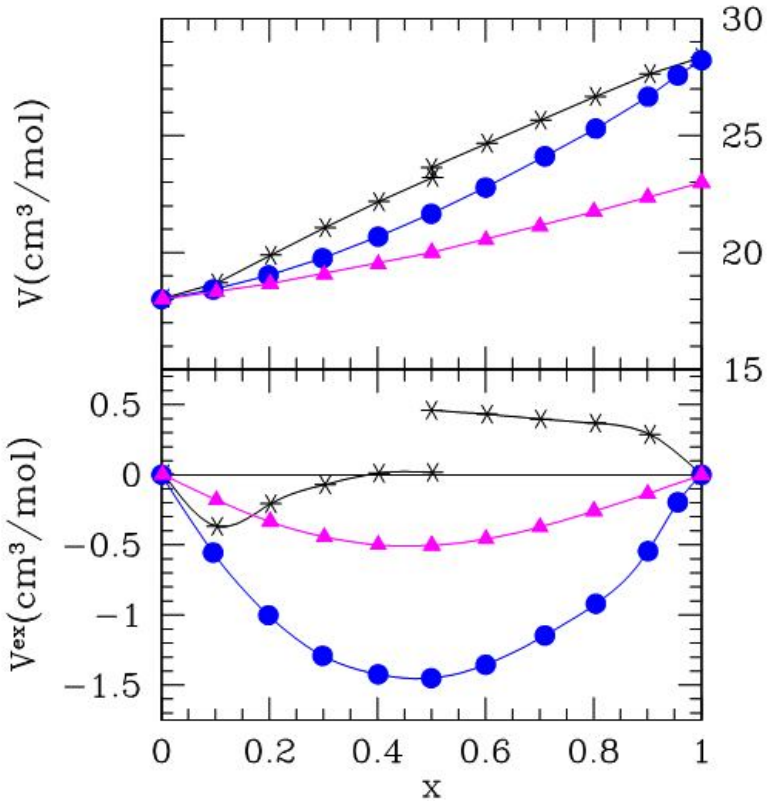

Figure 8.15: Volumetric properties of the 3 mixtures. Top panel: volumes, lower panel: excess volumes. Symbols: 1/3 model (stars), 2/3 model (dots), 4/5 model (triangles).

However, a considerable difference occurs in the position of the maximum of excess Coulomb energy, which is placed in the low concentration region for the 2/3 model (Fig. 8.16, lower panel) and in the high concentration region for the 4/5 model (Fig. 8.17, lower panel). The position of the extrema and generally, inflection points in the excess properties such as energy and heat capacity is related to the morphological changes in the mixtures [10]. As seen in the bottom panel of Fig. 8.16, there are two inflection points in the excess Coulomb energy of the 2/3 model. The first inflection point is around $x = 0.40$ and it would correspond to the clustering of the solute molecules. The second one is around $x = 0.90$, where the water clusters start to form.

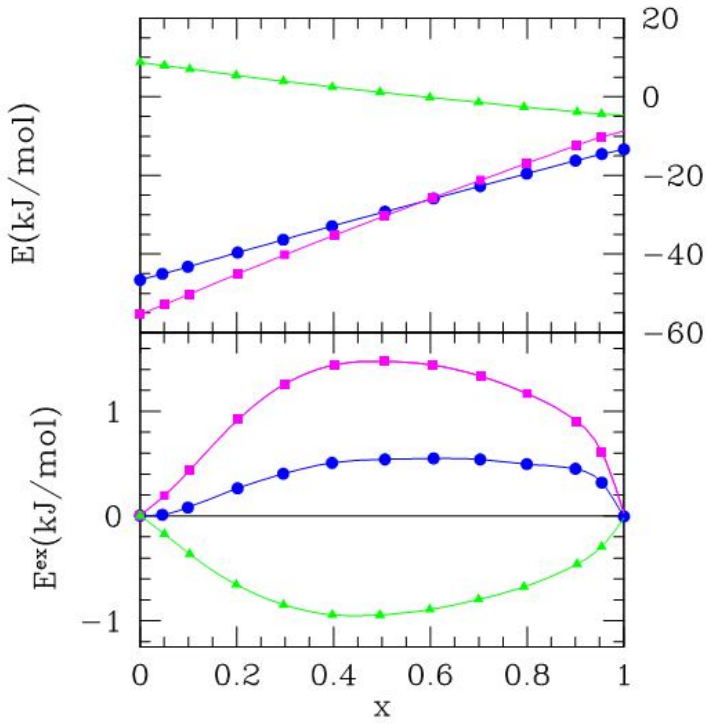

Figure 8.16: Energetic properties of the 2/3 model. Top panel: energy, lower panel: excess energy. Symbols: van der Waals part (triangles), Coulomb part (squares) and total configurational energy (dots).

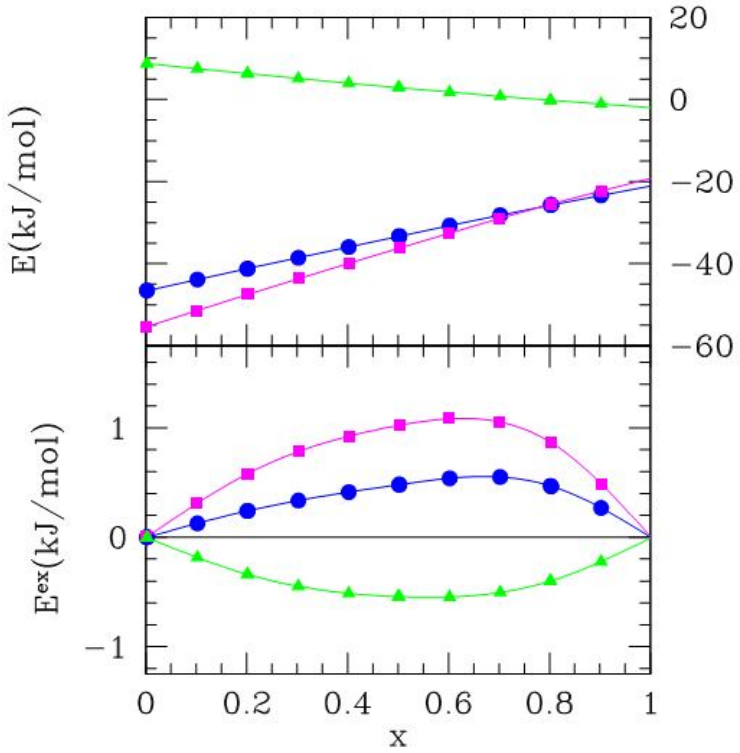

Figure 8.17: Energetic properties of the 4/5 model. Top panel: energy, lower panel: excess energy. Symbol convention is the same as in Fig. 8.16.

In the 4/5 model, as shown in the bottom panel of Fig.8.17, the maximum of the excess Coulomb energy is around $x = 0.70$, indicating that at this particular concentration water molecules start to aggregate themselves. Having in mind that there is no size difference between the solute and water, as well as between the two kinds of solute, one has to attribute these changes solely to the electrostatic properties of the solute.

The self diffusion coefficients for water and the three models are shown in Fig. 8.18. The self diffusion coefficient D_i is defined by the Green–Kubo relation (velocity autocorrelation function) [187]

$$D_i = \frac{1}{3}\int_0^\infty \langle v_i(t_0) v_i(t+t_0) \rangle dt \tag{8.10}$$

and the Einstein relation (mean square displacement) [187]

$$D_i = \frac{1}{6}\lim_{t\to\infty} \frac{d}{dt} \left\langle [r_i(t_0) - r_i(t_0 + t)]^2 \right\rangle, \tag{8.11}$$

where $v_i(t)$ is the velocity of the centre of mass of molecule i at time t and $r_i(t)$ is the position of the centre of mass of molecule i at time t.

Diffusion of the oxygen site, which is the central site of the water molecule, corresponds to the transport of the water molecule itself. On the other hand, diffusion of the hydrogen site corresponds to the degree of rotation of the molecule. If one considers a water molecule that is rotating around a fixed axis which passes through the central site, then the displacement of the hydrogen site would give a measure of the molecule's rotation. As seen in the top panel of Fig. 8.17, the diffusion of hydrogen and oxygen of real water stays about the same through the concentration range of the 1/3 model. It means that water molecules are not sensitive in terms of motion to the presence of the solute, as they are segregated from the solute species.

When it comes to the mixtures of the other two models, the trend of the two curves is very similar, indicating higher diffusion of hydrogen and oxygen in the low water region. The sharp increase in the graphs is more pronounced for the 2/3 model with an inflection point around $x = 0.70$, while for the 4/5 model, the values grow in a linear fashion. In the pure solute region, the values of hydrogen and oxygen diffusion constants become closer to each other, corresponding to the translational–rotational coupling as water becomes more monomeric.

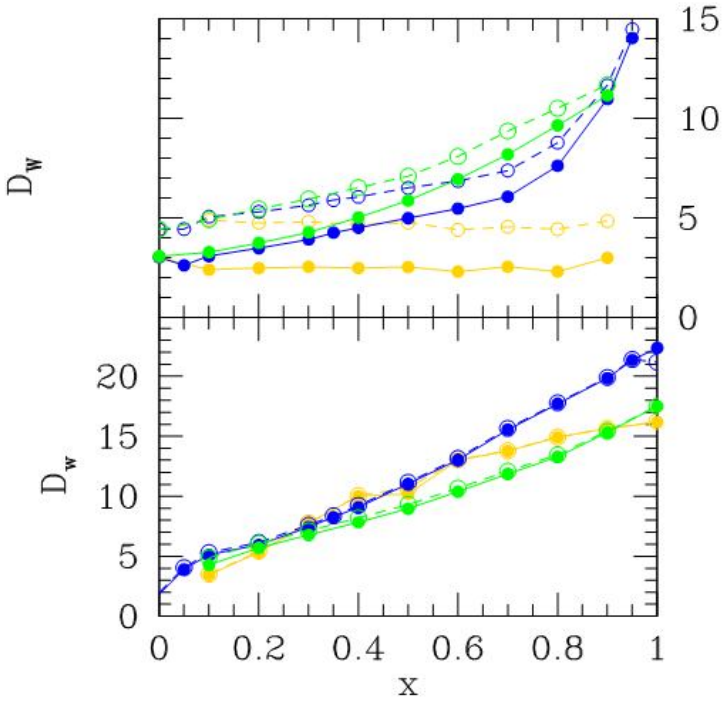

Figure 8.18: Oxygen and hydrogen site diffusion coefficients (in $cm^2/s \cdot 10^{-5}$) for water (top panel) and "weak water" (bottom panel). Symbols: hydrogen (open dots and dashed lines), oxygen (filled dots and full lines). Lines serve as guidelines. Colour convention: 1/3 model (gold), 2/3 model (blue), 4/5 model (green).

Looking at the bottom panel of Fig. 8.18, one concludes that the diffusion coefficients for the 4/5 model are closer to the ones of water, shown in the top panel. This is not surprising given the fact that the 4/5 model is the most similar to water of all three models. The differences in water diffusion between the two models can be explained in terms of micro–heterogeneity. In a strongly micro–segregated system, such as the mixture of the 2/3 model with water, the clustering of water molecules in the low solute region reflects in their enhanced translational and rotational motion and in the mixture of the 4/5 model with water, the system is more homogenous and this effect is not observed.

8.4 Conclusion

Different types of aqueous mixtures of modified water models were studied using Molecular Dynamics simulations. It was discovered that these mixtures exhibit many features of real aqueous mixtures. The water models differ between themselves in the amount of the partial charge on a molecule, ranging from 0 to 4/5 of the partial charge on water sites. Based on

statistical and thermodynamic results, a correspondence was found between each water model in study and a realistic solute, regarding their behaviour in water. In particular, a model with no charges behaves like a Lennard–Jones particle. The 1/3 model does not mix with water and it would represent solutes such as benzene. Strong micro–heterogeneity is observed in the 2/3 model mixture, corresponding to aqueous mixtures of alcohols [10, 101] or acetone [170]. Solutes such as formamide [107] or dimethyl sulphoroxide [188] (DMSO) form a moderately micro–segregated mixture with water, similar to what was found in aqueous mixtures of the 4/5 model. The modelling used in this study allowed us to eliminate effects rising from the size difference between the solvent and the solute. In other words, there is only one length scale and that is the diameter of water, $\sigma_W \approx 3$Å. It was demonstrated as well that hydrophobic effect can appear even if there are no hydrophobic groups. This comes as a consequence of water self–aggregation in a mixture due to electrostatic interactions. KBI have been extensively explored for the 3 mixtures in study, both by direct integration of $g_{ij}(r)$ and by thermodynamic route. For the model that exhibits moderate micro–heterogeneity, simulation gave KBI far higher than the ones obtained thermodynamically. This is not a simulation artefact, but rather a genuine feature these systems have, because of their specific morphology.

The domain formation in the system causes a transient behaviour in the running KBI function, so in order to recover the correct value of the KBI, one has to extend $g_{ij}(r)$ past the domain size through the procedure described in this Chapter 5. This approach is inspired by the analogy with micro–emulsions where there are domains of oil and water on a micro–meter scale, which is 1000 greater than the structures in our systems. Despite that, there is a difference between the natures of micro–heterogeneity of these two systems. Micro–emulsions are dominated by their coarse–grained heterogeneous structure, while concentration fluctuations play the leading role in aqueous mixtures. However, these two systems may not be that different, given the fact that there is a point in the phase diagram of micro–emulsions, called the Lifshitz point, where domains arise from concentration fluctuations [189]. This effect is exactly what is observed in the studied systems.

CHAPTER 9

Application of the Ornstein Zernike Equation to core–soft mixtures

9.1 Introduction

The core–softened interaction potential was first introduced by Hemmer and Stell [190, 191] who discovered that the addition of the repulsive ramp to the hard core potential (or core–softening) can lead to the appearance of a second liquid–liquid critical point if there is a liquid–vapour critical point. Subsequent studies of phase transitions of fluids with core–softened interactions include the study of fluids in dimension 1 [192, 193], dimension 2 [194-198] and dimension 3 [199-201]. Malescio and Pellicane [195] observed rich pattern formation in a 2D fluid were particles interact through a radially symmetric pair potential which consists of an impenetrable hard core plus a repulsive square shoulder. Camp [196-198] discovered a variety of structures in a 2D system were particles interact through the Lennard–Jones potential plus a $(1/r^3)$–term. Core–softened model fluids were studied also by Integral Equation Theory [201, 202]. Choudhury and Ghosh [202] extracted the bridge functions from the simulation data for a core–softened fluid. Phase behaviour and thermodynamic anomalies of such fluids were studied by Lomba et al. [201] using both simulations and the integral equation approach. The main reason why the core–softened fluid attracts so much attention is the possible relation between the core–softening of the potential interaction and the anomalous properties of realistic liquids, such as water for example. Jagla [203, 204] has shown that a core–softened potential can be used to reproduce water–like properties. The structure of liquid water was successfully reproduced by a soft–core $(1/r^{12})$ repulsive interaction together with a Gaussian repulsive core by both Monte Carlo simulations and integral equation theories [205]. It was found that the HNC closure was able to reproduce very accurately the experimental features of water at room temperature.

In this Chapter, the OZ equation complemented by the HNC equation is solved for two different binary mixtures where the solvent particles are water-like and they interact by a core-softened potential with the same parameters as in Ref. [205]. Despite the fact that all the interactions in these mixtures are repulsive, they produce Kirkwood–Buff integrals similar to real mixtures that show moderate micro–heterogeneity, such as DMSO–water [4]. Both Monte Carlo simulations and integral equations confirm that there is local micro-segregation, but there are no significant concentration fluctuations.

9.2 Mixtures of core–softened particles

The solvent–solvent interaction is modelled as

$$U_{WW}(r) = 4\varepsilon\left(\frac{\sigma_W}{r}\right)^{12} + 30\varepsilon\exp\left[-3.7\left((r-0.85)/\sigma_W\right)^2\right], \tag{9.1}$$

where the gaussian part serves to soften the soft–core part $(1/r^{12})$. The parameters σ_W and ε are both set to 1. This pair potential introduces two interaction length scales: the diameter of the solvent σ_W, and $\lambda_{CS} > \sigma_W$ which would correspond to the half-width of the gaussian part (blue dashed curve in the upper and bottom panels of Fig. 9.1). The solute–solute interactions are defined as

$$U_{SS}(r) = 4\varepsilon\left(\frac{\sigma_S}{r}\right)^{12}, \tag{9.2}$$

where the energy parameter ε is taken the be the same as that of the water model and $\sigma_S = 1.2\sigma_W$ is the solute diameter. For the first mixture, the solute–solvent interaction is defined according to the Lorentz-Berthelot rules:

$$U_{SW}(r) = 4\varepsilon\left(\frac{\sigma_{SW}}{r}\right)^{12}, \tag{9.3}$$

where $\sigma_{SW} = (\sigma_W + \sigma_S)/2$ and the energy parameter is the same as in Eqs. (9.1) and (9.2). In the following text, we refer to the mixture with the solute–solvent interaction defined by Eq. (9.3) as the Core–Soft/Soft–Core mixture. For the second mixture, the solute–solvent interaction is defined by

$$U_{SW}(r) = 4\varepsilon\left(\frac{\sigma_{SW}}{r}\right)^{12} + 10.5\varepsilon \exp\left[-3.7\left((r-0.85)/\sigma_{SW}\right)^2\right], \quad (9.4)$$

with the same energy parameter as in Eqs. (9.1)–(9.3). We refer to this mixture as the Core–Soft/Core–Soft mixture. Fig. 9.1 shows the potentials $U_{ij}(r)$ for the Core–Soft/Soft–Core mixture on the upper panel and for the Core–Soft/Core–Soft mixture in the lower panel.

Both binary mixtures are defined by the total reduced density $\rho = (N/V)\sigma_W^3$, where $N = N_W + N_S$ is the total number of particles (N_W and N_S is the numbers of water and solute molecules, respectively) and V is the volume of the system. Mixtures are studied for solute mole fraction $x = N_S / N$ varying from x = 0.1 to x = 0.9 by steps of 0.1. The total packing fraction η is fixed at $\eta = 0.42$, which is typical for a dense liquid. It is defined as $\eta = (N/V)(N_W v_W + N_S v_S)$, where $v_W = (\pi/6)\sigma_W^3$ is the volume of one water molecule and $v_S = (\pi/6)\sigma_S^3$ is the volume of one solute molecule. The reduced temperature is defined by $T^* = k_B T/\varepsilon$, where T is the absolute temperature and k_B is the Boltzmann constant. The reduced density is defined by $\rho^* = (6/\pi)\eta\sigma_W^3/(1 + x(\kappa^3 - 1))$, where $\kappa = \sigma_S/\sigma_W$.

The mixtures were studied by constant NVT Monte Carlo (MC) simulations with N = 4000 particles, which is sufficient to capture the long–range correlations in the system. The radial distribution functions were obtained by sampling every 20 global moves on a total of 50 000 steps. More statistics was needed for the solute–rich and the solvent–rich regions. The OZ equation was solved in conjunction with the HNC closure on a grid of N = 1024 points with an r-step of $\Delta r = 0.02$.

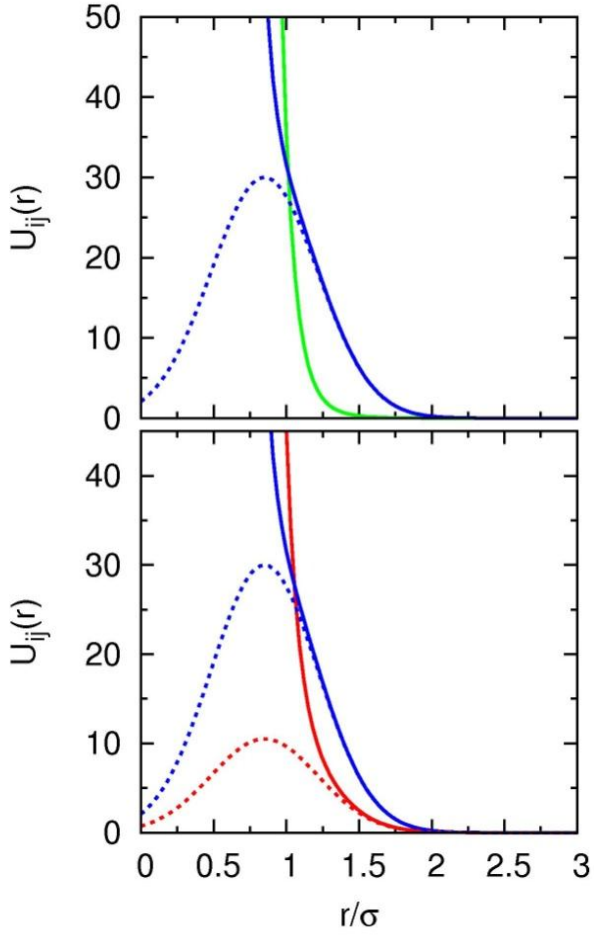

Figure 9.1: Pair potentials $U_{AB}(r)$ of the core–softened mixtures. Top panel: cross potential (green line), solvent–solvent potential (blue full line), Gaussian part of the solvent–solvent potential (blue dashed line) for the Core–Soft/Soft–Core mixture. Bottom panel: cross potential (red line), Gaussian part of the cross potential (red dashed line), solvent–solvent potential (blue full line) and Gaussian part of the solvent–solvent potential (blue dashed line) for the Core–Soft/Core–Soft mixture.

9.2.1 Radial distribution functions from the OZ/HNC theory and MC simulations

Fig. 9.2 presents the radial distribution functions $g_{ij}(r)$ from the OZ/HNC equation compared with Monte Carlo simulations at x = 0.10 mole fractions of the solute. Left side of this figure corresponds to the Core–Soft/Soft–Core mixture and right side corresponds to the Core–Soft/Core–Soft mixture. The same functions are presented in k–space in Fig. 9.3. Results for mixtures with x = 0.50 and 0.90 mole fractions of the solute are given in Figs. 9.4, 9.5, 9.6 and 9.7 both in real and in k–space.

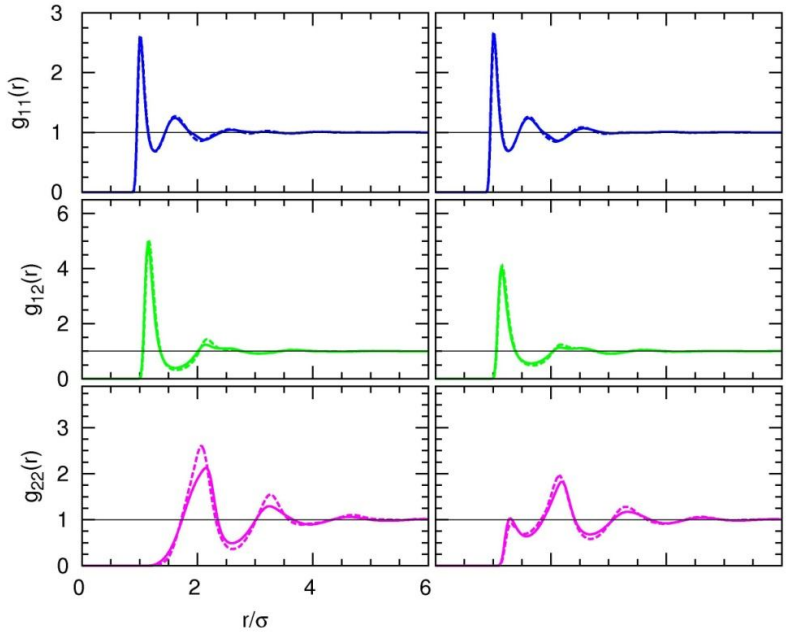

Figure 9.2: The radial distribution functions $g_{ij}(r)$ for mixture with x = 0.10 mole fraction of the solute. Blue is for solvent–solvent correlation $g_{11}(r)$, green is for cross correlation $g_{12}(r)$ and magenta is for solute–solute correlation $g_{22}(r)$. Left column: Core–Soft/Soft–Core mixture. Right column: Core–Soft/Core–Soft mixture. Full line is for OZ/HNC data and dotted line is for MC simulations.

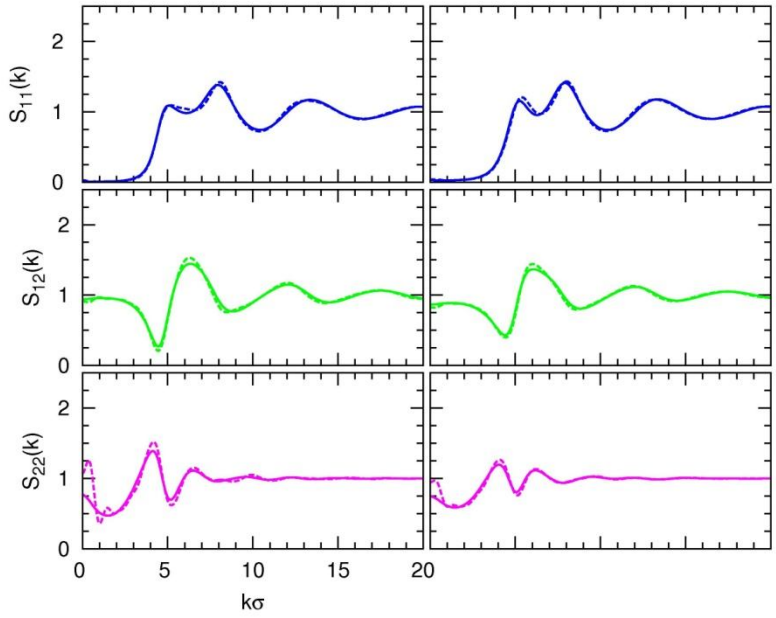

Figure 9.3: The same functions as in Fig. 9.2 given in *k*–space. Colour, line and panel convention is the same as in Fig.9.1.

.

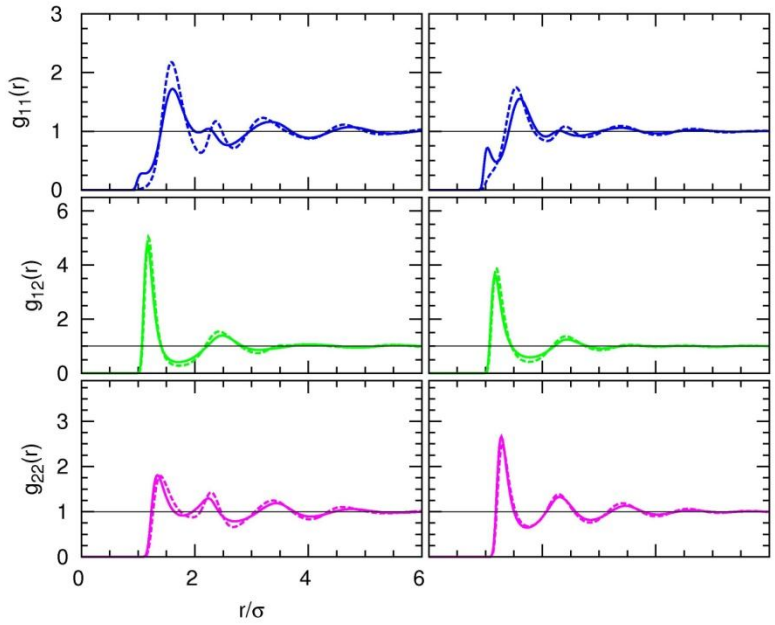

Figure 9.4: The radial distribution functions $g_{11}(r)$, $g_{12}(r)$ and $g_{12}(r)$ for the x = 0.50 mixture. Colour, line and panel convention is the same as in Fig.9.1.

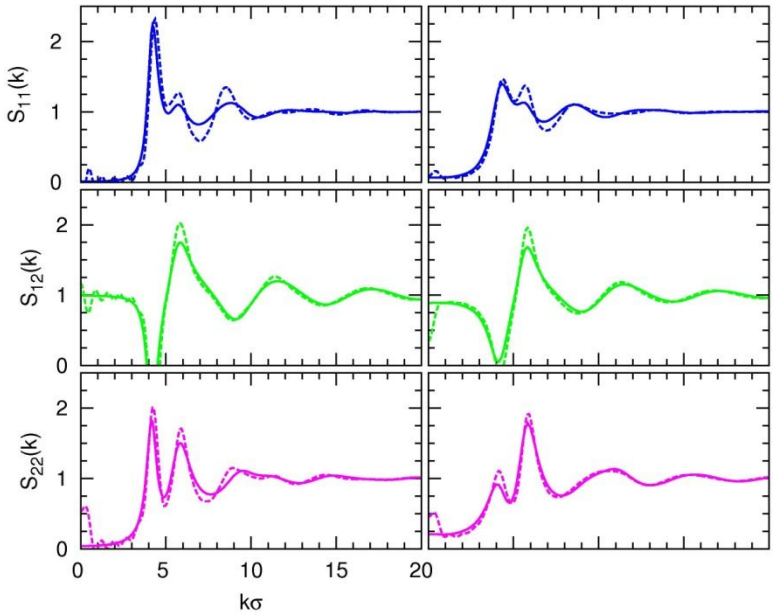

Figure 9.5: The same functions as in Fig. 9.3 given in *k*–space. Colour, line and panel convention is the same as in Fig.9.1.

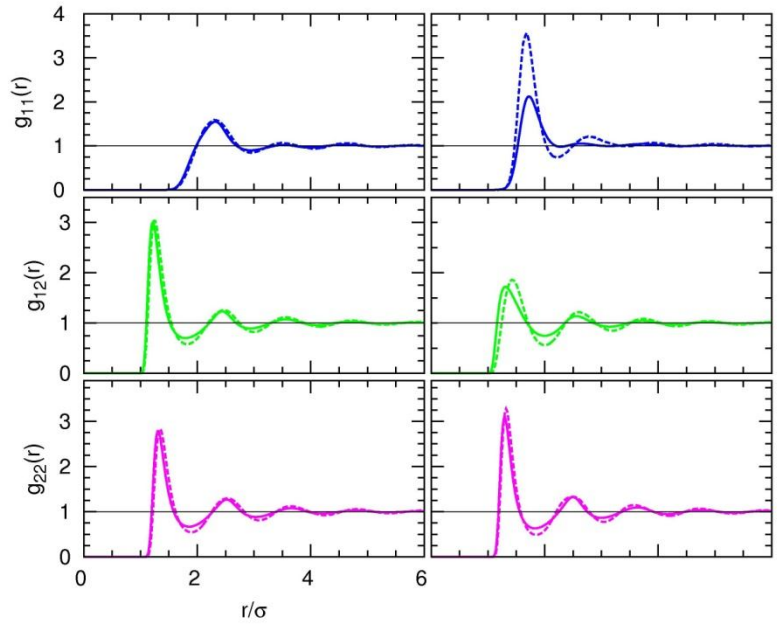

Figure 9.6: The radial distribution functions $g_{11}(r)$, $g_{12}(r)$ and $g_{12}(r)$ for the x = 0.90 mixture. Colour, line and panel convention is the same as in Fig.9.1.

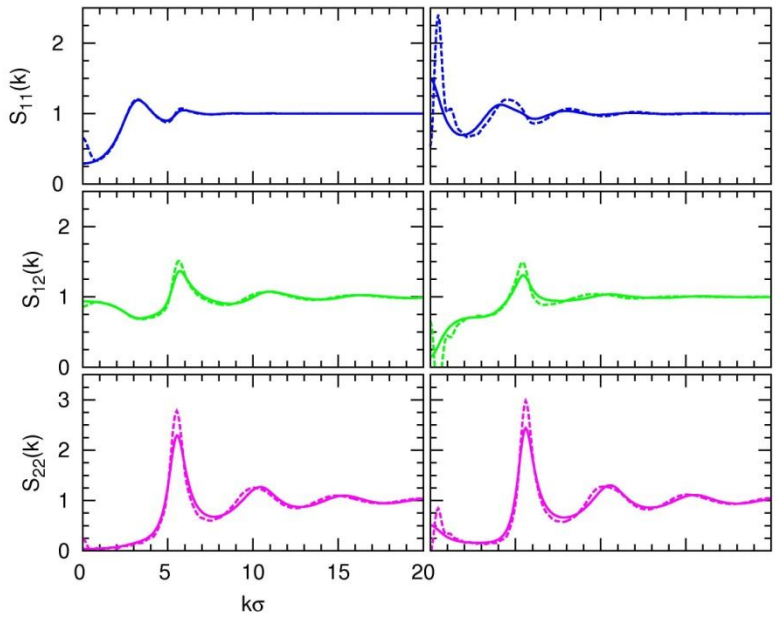

Figure 9.7: The same functions as in Fig.9.5 given in *k*–space. Colour, line and panel convention is the same as in Fig. 9.1.

As seen in Figs. 9.1–9.7, the OZ/HNC results show qualitatively good agreement with the simulations, especially in the real space. For the x = 0.10 mixture, as seen in Fig. 9.3, the match between the two sets of data is almost perfect. Some discrepancies are seen in the small–*k*

behaviour of the structure factors $S_{12}(k)$, as shown in Fig. 9.4. This indicates that there are differences in the long–range part of the correlations, which is not surprising given the fact that the theory misses a certain set of cluster diagrams [1]. Correlations of the equimolar mixture are presented in Figs. 9.3 and 9.4. It is seen that the oscillatory structure of the solvent–solvent correlation functions $g_{11}(r)$ from theory does not match the simulations. In *k*–space, the theory underestimates the second peak of $S_{11}(k)$ which is positioned at the contact distance of the two particles, meaning that the correlations at $r = 2\pi / k \approx 1\sigma_W$ are underestimated.

I present correlations in real and *k*–space for the solute–rich mixture, x = 0.90, in Figs 9.6 and 9.7, respectively. It is seen that the theory captures well the cross and the solute–solute correlations in both mixtures, but underestimates the height of the first peak of solvent–solvent correlations $g_{11}(r)$ in the Core–Soft/Core–Soft mixture. However, both simulation and theory show the same position of this peak, at $r/\sigma_W \approx 1.7$. Another striking difference is in the pre–peak of $S_{11}(k)$ from simulations (upper right panel of Fig. 9.7) at the position were the theory shows a raise. This suggests that there are water–water clusters which are not seen by the theory.

9.2.2 Concentration fluctuations and micro–heterogeneity in core–soft mixtures

Deeper insight into the structure of these mixtures is assured if one considers the concentration–concentration structure factor $S_{CC}(k)$ defined as [206]

$$S_{CC}(k) = x_1 x_2 \left(x_1 S_{11}(k) + x_2 S_{22}(k) - 2\sqrt{x_1 x_2} S_{12}(k) \right) \tag{9.2}$$

This quantity is a measure of the concentration fluctuations inside the mixture and its divergence at $S_{CC}(k=0)$ indicates demixing. It was found previously by other authors [124, 207] that the HNC equation overestimates the concentration fluctuations and looses solution as one approaches to the spinodal line, which bounds the mechanically unstable region in the phase diagram. Fig. 9.8 shows $S_{CC}(k)$ for mixtures at x = 0.10, 0.50 and 0.90 mole fraction of the solute. It is seen that both OZ/HNC results and simulations agree well on the small–*k* part of $S_{CC}(k)$, showing no particular raise at $S_{CC}(k=0)$ and therefore no concentration fluctuations. However, the theory does not describe correctly the double–peak structure of $S_{CC}(k)$, as seen on

the top panels of Fig.9.8. The theory also underestimates the height of the peaks of $S_{CC}(k)$ at $k\sigma_W = 5.5$, as observed in the middle and bottom panels of Fig.9.8.

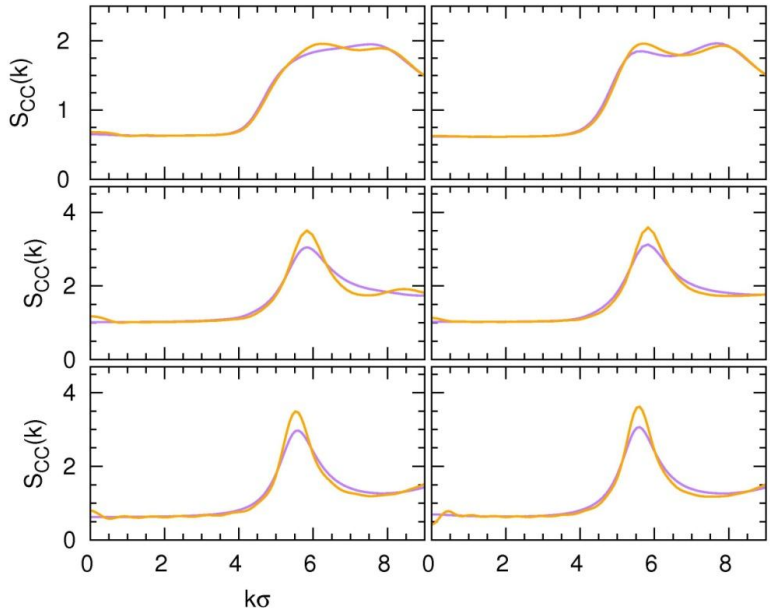

Figure 9.8: The concentration–concentration structure factors $S_{CC}(k)$ for mixtures at at x = 0.10 (top panels), 0.50 (middle panels) and 0.90 (bottom panels). On the left: Core–Soft/Soft–Core mixture. On the right: Core–Soft/Core–Soft mixture. Purple line is for OZ/HNC data and orange line is for MC simulations.

Another observable that probes the concentration fluctuations is the KBI. Figs 9.9, 9.10 and 9.11 show running KBI for x = 0.10, 0.50 and 0.90 mole fraction of the solute, respectively. As seen in the bottom panels of Fig. 9.9, there is a large discrepancy between theory and simulations in the long–range part of the solute–solute running KBI. This implies that the many–body correlations, which are missing from the theory, are essential if we are considering the behaviour of the solute molecules in the solvent–rich region. Contrary to the x = 0.10 mixture, the running KBI are fairly well reproduced by the theory for the equimolar mixture, as seen in Fig. 9.10., which indicates that these concentration fluctuations are more trivial and do not require the inclusion of higher order correlations. In other words, the equimolar mixture exhibits no particular structure. Fig. 9.11 presents the running KBI for the x = 90 mixture and it is seen, similar to the case of x = 0.10, that the theoretical running KBI for the minority species show a flat asymptote. Running solvent–solvent KBI for the Core–Soft/Core–Soft mixture do not

stabilize in the simulations, which is an effect that has been observed previously in aqueous mixtures [10, 125].

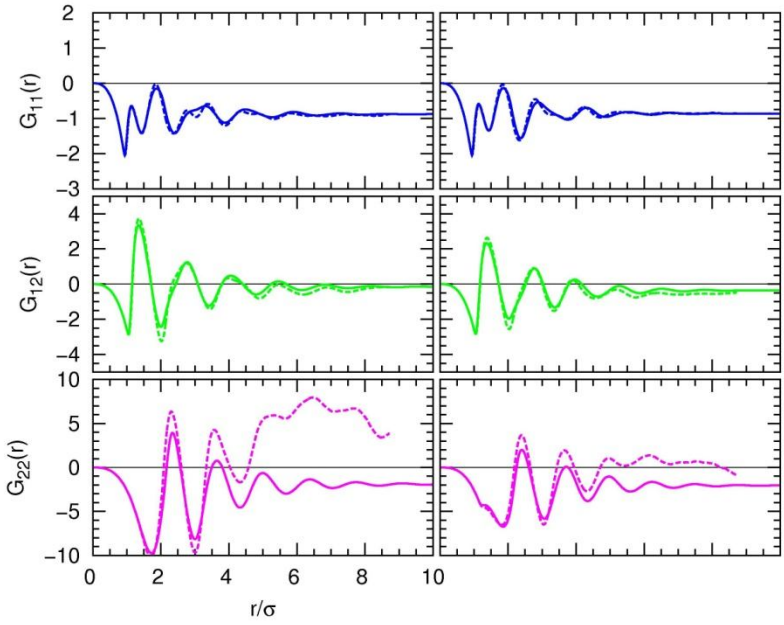

Figure 9.9: The running KBI for solute–solute pair, cross pair and solvent–solvent pair correlations for $x = 0.10$ mixture. Colour, line and panel convention is the same as in Fig. 9.3.

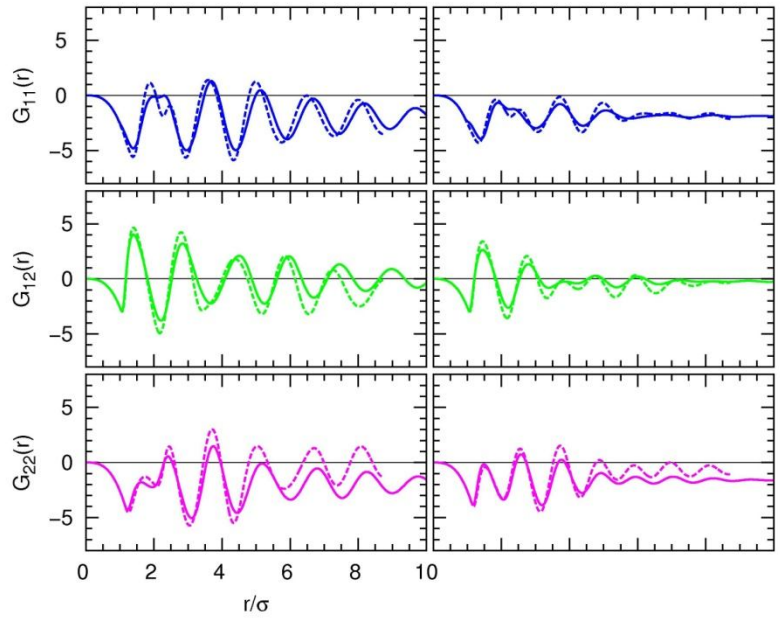

Figure 9.10: The running KBI for solute–solute pair, cross pair and solvent–solvent pair correlations for $x = 0.50$ mixture. Colour, line and panel convention is the same as in Fig. 9.3.

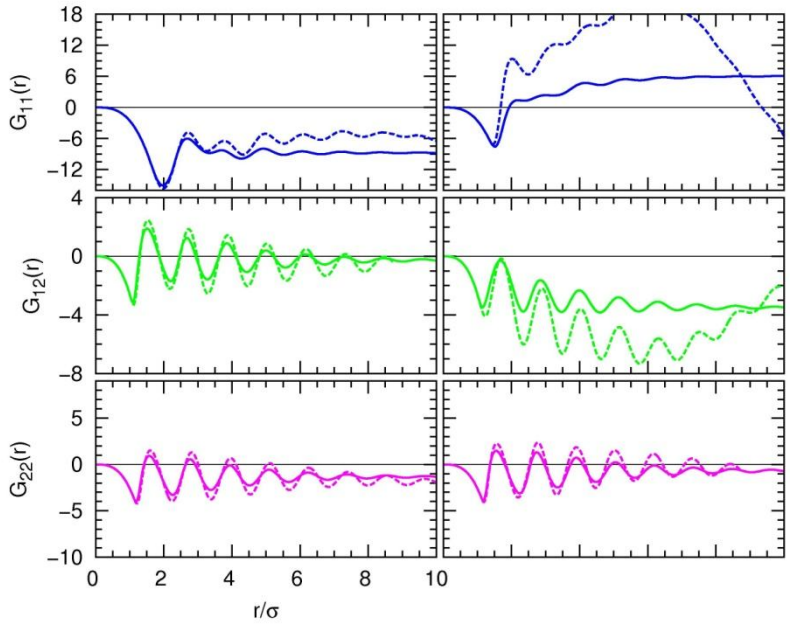

Figure 9.11: The running KBI for solute–solute pair, cross pair and solvent–solvent pair correlations for $x = 0.90$ mixture. Colour, line and panel convention is the same as in Fig. 9.3.

Fig. 9.11 shows the Kirkwood–Buff integrals G_{ij} plotted versus the mole fraction of the solute, x. As seen on both panels of Fig. 9.11, the theoretical KBI match those from simulations over the range $x = 0.20$–0.80. Some discrepancies are seen in the solvent–rich and solute–rich regions for the minority component, where the HNC theory underestimates the KBI values. This indicates that the HNC equation diminishes the segregation of particles.

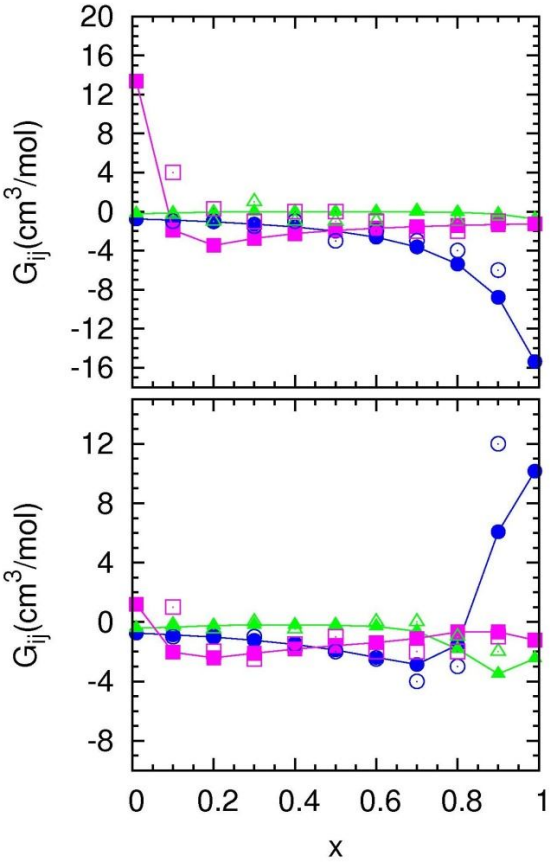

Figure 9.12: The Kirkwood–Buff integrals G_{ij} for the Core–Soft/Soft–Core mixture (upper panel) and the Core–Soft/Core–Soft mixture (lower panel). Open symbols: MC simulations. Filled symbols: OZ/HNC theory. Blue circles are G_{11}, green triangles are G_{12} and magenta squares are G_{22}.

Snapshots of all three concentrations for both mixtures are presented in Fig. 9.13. Left column corresponds to the Core–Soft/Soft–Core mixture and right column to the Core–Soft/Core–Soft mixture. For x = 0.10 mixtures in the top panels of Fig. 9.12, it is very difficult to see any particular structure and the particles seem randomly positioned. Middle panels show x = 0.50 mixtures where the water-like particles appear to form clusters. For x = 90 mixtures in the bottom panels, there are certainly solvent clusters, especially for the Core–Soft/Core–Soft mixture where one can see a few chain–like clusters.

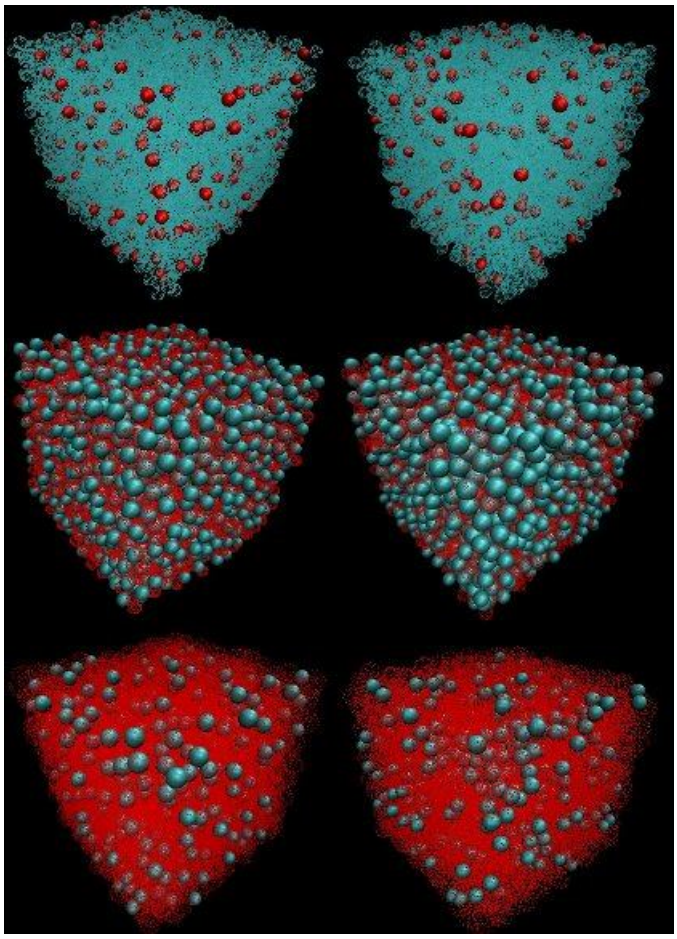

Figure 9.13: Snapshots of three concentrations: $x = 0.10$ (top panels), $x = 0.50$ (middle panels) and $x = 90$ (bottom panels). Left side: Core–Soft/Soft–Core mixtures. Right side: Core–Soft/Core–Soft mixtures. In $x = 0.10$ and $x = 0.90$ mixtures, the majority species is shown as semi–transparent. Blue particle is the solvent and red particle is the solute. In $x = 0.50$ mixtures, solute species is shown as semi–transparent.

9.3 Conclusion

An integral equation and Monte Carlo simulation study were performed for two types of binary systems containing spheres which interact by a core–soft potential. It was found that the HNC closure successfully describes the structure of mixtures in the intermediate region of concentrations, from $x = 0.20$ to $x = 0.80$ mole fraction of the solute. However, for concentrations of $x = 0.10$ and $x = 0.90$, the HNC theory does not show a good agreement with the simulations, which is seen particularly in the case of Core–Soft/Core–Soft mixture for the solvent–solvent correlations at $x = 0.90$. The fact that HNC underestimates these correlations, as seen from the short–range structure of the pair correlation function and the small–k behaviour of

structure factor $S_{11}(k)$, shows that the absence of the many–body contribution effects in this model mixture does not lead to the growth of these correlations, contrary to the case of realistic mixtures that exhibit micro–segregation [121]. In other words, the enhancement of the correlations or clustering that we observe from the short–range structure of the simulation $g_{11}(r)$ and confirm by snapshots is of a different nature compared to the clusters of micro–heterogeneous mixtures, whose structure is often described through the behaviour of the KBI. These integrals are seen to exhibit large values in micro–heterogeneous mixtures in the range of intermediate solute concentrations so one cannot clearly distinguish between concentration fluctuations and micro–heterogeneity only by looking at the KBI values [125, 170]. Since the KBI for our studied systems show a rather flat behaviour, these structural effects are not present here. High value of G_{11} in the solvent-rich region of the Core–Soft/Core–Soft mixture comes from the fact that the water-like particles are strongly micro-segregated, but there are minimal concentration fluctuations, as confirmed also by the small-*k* behaviour of the concentration-concentration structure factor $S_{CC}(k)$.

CHAPTER 10

Conclusion

In this thesis work, I have tried to show how the micro–heterogeneous nature of complex liquids and their mixtures is difficult to handle through usual statistical approach of liquids. The key feature I have tried to assert is the difference between concentration fluctuations and micro–heterogeneity, which is a form of local order. The principal difficulty both simulation and theories meet is the precisely the fact that they do not distinguish these manifestations as being distinct.

In the Introduction, I have mentioned how intriguing local order is, as opposed to global order, for which we have robust theories from Physics, based on the concept of order parameter. Phase transitions have been described by Thermodynamics even before the concept of atom was accepted among physicists. The statistical mechanics invented by Maxwell and Boltzmann have provided a new insight into these transitions, greatly helped by the microscopic picture provided by molecules. The modern key concept to phase transitions is the idea of fluctuations and associated response functions. It is noteworthy that the concept of fluctuation is also contained in classical thermodynamics, as for example in the isothermal compressibility, a quantity known prior to the work of Boltzmann. It is however the microscopic picture that provides a new insight into the true meaning of response functions.

The concept of local transitions is not clear in modern statistical physics: we do not know which response function to associate to such changes. If such concept exists, it remains to be invented and appropriate observable are required in order to track such transitions in experiments. With that respect, the new idea of topological phase transitions that occurs in quantum liquids and other related systems in condensed matter are quite appealing, since they are precisely not described by the standard Landau phase transition formalism, where there is a order parameter

and a symmetry to break [208]. Topological order is richer since it allows to have phase transitions that do not break any symmetry and allows to explain phenomenon occurring at very low temperatures such as fractionalization, for example [209, 210]. In particular, it explains several forms of quantum collective phenomena. However, its physics is also very remote from what I examined in this thesis.

In the present work, I have tried to track down local order through various systems, some realistic and some models. The system that is the closest to the form of local order we are talking about is the micro–emulsion. Micro–emulsions are considered to be homogeneous [211] although the segregated oil–water domains are in the range of 0.1 μm. This is because there is no global order to these systems, and global order is synonymous of heterogeneous. However, these systems can exhibit rich form of orders, such as a variety of lamellar and sponge phases. They also exhibit micelle phase, which is not a true phase transition. In contrast to these systems, most of the simple aqueous mixtures stay in the homogeneous phase. The only "transition" they can have is when a pre–peak emerges in the scattering function when domains become well defined. This is analogous to what happens in micro–emulsions, with the difference that the pre–peak corresponds to ordered domains such as lamellas. The pre–peak witnesses the fact that a new length scale has emerged into the system, for example the distance between lamella. This is an important insight that is very helpful to get further understanding into describing local order.

Generally, liquids are characterized by the length scale set by the size of the molecules. An additional length scale comes from the fluctuations, as can be deduced from the OZ equation. There is only one such length, regardless of the number of components, since this length is deduced from the small–k expansion of the OZ (under matrix form for a multi–component and angular dependent system) at order k^2. If we go to order k^4we can have an additional length, which corresponds to domain formation. So, the OZ equation contains all the information to explain local order from the point of view of local domains. This is what we have used by invoking the concept of molecular emulsion.

Pre–peaks in structure factors are not always associated to positional order such as in lamella. For example, cluster forming systems have a pre–peak. Such types of systems are found in reality with colloids coated with polymers [212] which have a short–range attractive interaction and a long range repulsive interaction. This type of system can be modelled by a 2–Yukawa representation with one attractive and the other repulsive [213]. The pre–peak of these systems is built in, as it can be shown by the direct Fourier transform of the interaction itself [189]. Other systems include those with two repulsive cores [214, 215]. Both of these systems are pseudo

one–component systems, with two length scale interactions.

Both simulations and integral equations can deal with these systems without any of the problems mentioned for micro–heterogeneous systems. It probably helps that the cause for domain formation is embedded within the interaction itself with the two length scales. So, it looks like the problems linked with micro–heterogeneity come from the fact that the domain length scale is left free to emerge from the system itself. This is a matter of pseudo–potentials, just like the Yukawa is a screened version of the direct Coulomb interaction, and it is difficult to obtain a Yukawa screened interaction from the correlations measured in a simulation.

By imposing the TS extension, we help the system producing the domain modulation in the correlation functions. This works well for molecules that produce well defined domains such as acetone or TBA, but not for smaller molecules such as methanol or ethanol, when domains are fuzzy (compare snapshots in Chapters 5 and 7). The high scattering in $S(k=0)$ is due to micro–heterogeneity and not concentration fluctuations, but it seems difficult to tell them apart without an accurate small–k dependence of $S(k)$. Indeed, if $S(k)$ has the OZ Lorentzian form $A/(k^2+\xi^{-2})$ then the high scattering is entirely due to concentration fluctuations. But if there is a non–negligible k^4 dependence, then despite the absence of a pre–peak, there is domain dependence. The simulation data is not accurate enough to tell unambiguously what the exact status of $S(k)$ is. The KH based integral equation theory always predicts an OZ form, while if we insert the pure water bridge, we predict a non–OZ behaviour. The problem is that this insertion also diminishes dramatically the k = 0 contribution. It is clear that important physics of this simple system is captured by neither the simulations nor the integral equations.

How can we progress beyond this point? We cannot invert the simulation data without having a clear small–k behaviour, like we did for pure liquids with good success. This imposes simulating systems of more than 10 000 particles. Here again, it is not clear that the tail of the correlations will stabilize in order to produce the proper KBI. It seems that the inter–domain correlations are quite large, even in a system as simple as water–methanol.

To conclude, I would like to put my work in perspective with the more general approach in Physics. Classical physics has made important progress when dealing with well–defined "objects" like planets, balls or atoms. It has also dealt successfully with continuum type media such as fluids, when described macroscopically (for example the Navier–Stokes or Boltzmann equations). This is because one can write equations that capture the generic behaviour of such systems. When dealing with systems where one has to encompass both levels of description, we have only the statistical approach, and we are missing relations involving the moments of the

statistical distribution that would be in par with the generic descriptions we have obtained in the cases mentioned above. In consequence, we are forced to deal with the details of fluctuations and emergent forms –such as micro–heterogeneity – without having generic equations to solve. Field theory provides such an approach through the Landau–deGennes–Ginzburg Lagrangian. But this type of approach is better suited to describe phase transitions, and not local transitions as those that encountered in aqueous mixtures. It is not obvious that such equations exist for the type of system we considered.

On the other hand, the fact that biological systems are so well–organized, using only fluctuations and molecules of various size, indicate the possible existence of generic equations –yet to be discovered– that would allow one to predict the paths of organization of such complex systems, just like we can predict the motion of planets despite the hidden generic chaos of the many body system. It is indeed like Albert Einstein said at the beginning of his Science paper in 1940 [216]: "Science is the attempt to make the chaotic diversity of our sense–experience correspond to a logically uniform system of thought".

Publications issued from this work

- M. Mijaković, B. Kežić, L. Zoranić, F. Sokolić, A. Asenbaum, C. Pruner, E. Wilhelm, and A. Perera, "Ethanol–water mixtures: ultrasonics, Brillouin scattering and molecular dynamics", *J. Mol. Liq.* **2011,** 164, (1–2), 66–73.

- B. Kežić and A. Perera, "Towards a more accurate reference interaction site model integral equation theory for molecular liquids", *J. Chem. Phys.* **2011**, 135, (23), 234104–11.

- A. Perera, R. Mazighi, and B. Kežić, "Fluctuations and micro–heterogeneity in aqueous mixtures", *J. Chem. Phys.* **2012**, 136, (17), 174516–10.

- B. Kežić and A. Perera, "Aqueous tert–butanol mixtures: A model for molecular–emulsions", *J. Chem. Phys.* **2012**, 137, (1), 014501–12.

- A. Perera, B. Kežić, F. Sokolić, and L. Zoranić, "Micro–heterogeneity in complex liquids" in "Molecular Dynamics–Studies of Synthetic and Biological Molecules", Ed. Wang, L., InTech, Rijeka, 2012.

- B. Kežić and A. Perera, "Revisiting aqueous–acetone mixtures through the concept of molecular emulsions", *J. Chem. Phys.*, **2012**, 137, (13), 134502–6.

- A. Asenbaum, C. Pruner, E. Wilhelm, M. Mijaković, L. Zoranić, F. Sokolić, B. Kežić, and A. Perera, "Structural changes in ethanol–water mixtures: Ultrasonics, Brillouin scattering and molecular dynamics studies", *Vib. Spectrosc.* **2012**, 60, (0), 102–106.

- B. Kežić, R. Mazighi, and A. Perera, "A model for molecular emulsions: Water and „weak water" mixtures", *Phys. A,* **2013,** 392, (4), 567–582.

- A. Perera, and B. Kežić, "Fluctuations and micro-heterogeneity in mixtures of complex liquids", *Faraday Discuss.*, **2013**, DOI: 10.1039/C3FD00072A.

Résumé

Ce travail de thèse porte sur l'analyse et la compréhension, au travers des outils de la théorie des liquides, de la nature de l'ordre local des mélanges liquides, et en particulier aqueux. En effet, les mélanges aqueux diffèrent des mélanges de liquides simples, du fait que les molécules d'eau tendent à s'associer préférentiellement entre elles au travers de la liaison hydrogène, de ce fait ségréguant les molécules de soluté sans pour autant entraîner une démixtion totale des mélanges. Même des mélanges aussi simple que l'eau et le méthanol exhibent ce type d'ordre [138]. La micro–hétérogénéité résultante apparaît comme étant distincte des fluctuations de concentration, qui jouent un rôle important dans la stabilité des mélanges. La micro–hétérogénéité est au cœur des processus microscopiques dans les milieux de type biologique, dans lesquels l'eau joue un rôle important ; comme par exemple l'auto–assemblage de membranes à partir de molécules de lipides [217]. Les simulations numériques seraient l'outil par excellence pour aborder le détail des structures moléculaires micro–hétérogènes. Elles ne sont cependant pas exemptes de diverses contraintes, telles que la taille de la boîte ainsi que le temps d'échantillonnage [121]. Ces problèmes deviennent des obstacles sérieux à l'étude des mélanges aqueux, car les corrélations à moyenne portée –précisément celles limitées par les tailles habituelles de boites– ne sont pas bien décrites. Nous illustrons cette difficulté dans notre étude au travers d'exemples spécifiques. Nous introduisons le concept d' »émulsion moléculaire » afin de décrire les mélanges aqueux. Les équations intégrales représentent une alternative à l'étude de ces systèmes, mais elles souffrent également d'autres types de contraintes. Le problème principal provient de leur inadéquation à décrire les corrélations à N–corps, et par conséquent les fluctuations complexes [1, 13].La description des corrélations est au centre de ces théories. Il y a deux types d'approches concurrentes. La première consiste à décrire toutes les corrélations angulaires explicitement en développant ces fonctions sur une base d'invariants rotationnels [35]. Elle se base sur l'équation exacte de Ornstein–Zernike (MOZ). La seconde consiste à calculer les corrélations entre sites atomiques des molécules, et se base sur une version approchée de MOZ appelée SSOZ. Nous décrivons ces deux approches ainsi que leurs mérites respectifs au chapitre 2.Dans le chapitre 3,

l'approche basée sur l'équation SSOZ est utilisée pour l'étude des corrélations dans plusieurs liquides moléculaires purs (acétone, tétrachlorure de carbone, amides et alcools). L'utilisation des relations de clôture HNC et PY montre qu'il est nécessaire d'introduire les contributions à N–corps en inversant la relation SSOZ en partant des corrélations obtenues par les simulations numériques. Les propriétés thermodynamiques sont comparées avec celles des simulations et des expériences.Au chapitre 4, l'équation SSOZ est utilisée pour l'étude des mélanges aqueux. On trouve des différences plus marquées entre les différentes relations de clôture telles que HNC et KH que dans le cas des corps purs pour lesquels ces deux clôtures donnent des résultats quasi–identiques. La clôture HNC n'a souvent aucune solution, du fait qu'elle ne distingue pas la micro–hétérogénéité des fluctuations de concentrations et conduit la plupart du temps à la démixtion des mélange. L'introduction des corrélations à N–corps des corps purs permet d'obtenir des solutions, et en particulier à bien décrire les corrélations à courte portée. Cependant la structure à moyenne portée est très mal décrite . La relation de clôture KH donne des solutions dans la plupart des cas, mais elle décrit très mal toutes les corrélations. Cette étude montre le rôle prépondérant joué par la différence entre les fluctuations de concentration et la micro–hétérogénéité. Dans le chapitre 5 nous étudions tout particulièrement le mélange eau–éthanol par dynamique moléculaire, pour plusieurs concentrations allant de $x = 0$(eau pure) à $x = 1$(éthanol pur), par pas de $x = 0{,}1$, où x est la fraction molaire d'éthanol. Nous abordons en particulier les soucis causés par les modèles de champs de force et les techniques de simulation, notamment en regard de la micro–hétérogénéité. Nous détaillons comment ces problèmes se reflètent dans le calcul des intégrales de Kirkwood–Buff (KBI).

Le concept d'émulsion moléculaire est introduit au chapitre 6, afin de décrire l'hétérogénéité en domaines, par analogie avec les micro–émulsions. Ces dernières sont des mélanges aqueux au moins ternaires, composé d'eau, d'huile et de surfactant. Ces systèmes font apparaître de nouvelles entités, telles que les micelles et les lamelles [208]. Les corrélations dans ces systèmes ont une description mathématique qui fait intervenir 2 échelles de longueurs [157], la longueur de corrélation et la taille de domaine. En établissant une analogie formelle avec les micro–émulsions, nous dérivons la forme de décroissance asymptotique des fonctions de corrélations à partir de l' équation MOZ et retrouvons ainsi la forme Teubner–Strey (TS). En utilisant cette forme asymptotique, nous montrons que l'on peut étendre les fonctions de corrélation calculées pour une taille finie en simulation, et de cette manière nous retrouvons un bon accord avec les résultats expérimentaux des KBIs. Au chapitre 7, nous illustrons le concept d'émulsion moléculaire et de l'extension TS, au travers de l' étude par dynamique moléculaire des mélanges

eau+acétone et eau+tbutanol. En effet, les simulations de ces mélanges prédisent une grosse micro–hétérogénéité et donc des KBI bien plus grand que ceux expérimentaux, de facteurs 3 ou plus. L'extension module les corrélations et permet de ramener ces valeurs très proches de celles expérimentales.Au chapitre 8 nous étudions des modèles d'eau « faible », obtenues en diminuant les charges partielles du modèle SPC/E par des facteurs 1/3 , 2/3 et 4/5. En les mélangeant avec l'eau SPC/E, nous montrons que les diverses structures micro–hétérogènes ne sont pas dues aux tailles plus grandes des solutés réalistes, mais bien à des interactions faibles de liaison hydrogènes. Dans le cas 2/3 nous montrons la pertinence de l'extension TS appliquée à ce cas de mélange modèle. Au chapitre 9, nous étudions des mélanges modèles avec un modèle d'eau purement répulsive à 2 cœurs, mais qui reproduit les particularités des corrélations de l'eau réelle. Et éliminant ainsi les effets de charges, et ne retenant que les effets d'organisation spatiale proche de l'eau réelle, nous montrons les différences structurales dues aux fluctuations de concentration et à la micro–hétérogénéité. Nous comparons notamment les corrélations entre les simulations et la clôture HNC, qui est remarquablement bonne, sauf quand à la prédiction de la différence entre fluctuations de concentration et micro–hétérogénéité, c'est à dire dans le domaine des petits k des facteurs de structure. Dans notre conclusion au chapitre 10, nous récapitulons les points essentiels qui émergent de notre étude, à savoir que les simulations et la théorie ont du mal à décrire la différence entre fluctuations de concentration et micro–hétérogénéité, et que la modulation en domaines introduite par la micro–hétérogénéité conduit à la présence d'un pré–pic dans le facteur de structure eau–eau. Ce pré–pic existe dans les micro–émulsions [208]. On retrouve également ce type de manifestation dans les systèmes de colloïdes enduits de polymères adsorbés à leurs surface, et qui contribuent à donner à l'interaction entre colloïde une attraction à très courte portée mais également une répulsion à longue portée. La modulation de domaine se présente comme une répulsion à longue portée des molécules de même espèces, provoquant ainsi la micro–ségrégation. L'introduction de la modulation au travers de l'extension TS nous a permis de décrire correctement la micro–ségrégation des systèmes aqueux réels, en dépit des limitations imposées par les simulations numériques.

Sažetak

Tema ovog doktorata je ispitivanje strukture čistih tekućina i njihovih mješavina s vodom. U vodenim otopinama određenih tvari, kao što su primjerice alkoholi, zapažena je mikro–segregacija molekula. Drugim riječima, molekule alkohola i vode zauzimaju različite dijelove prostora unutar otopine. Ta organiziranost karakterizira čak i vodenu otopinu metanola [138], koji je najjednostavniji od svih alkohola. Glavno pitanje je kako razlikovati koncentracijske fluktuacije, koje postoje u svim mješavinama, od mikro–heterogenosti, unutar okvira Teorije tekućina. Analiza mikro–heterogenosti u vodenim otopinama je neophodna za potpuno razumijevanje uloge vode u biološki važnim procesima, kao što je primjerice formiranje membrane agregacijom lipidnih molekula u vodenom mediju [217]. Računalne simulacije su općenito prihvaćen alat za analizu mikro–strukture sustava. Međutim, simulacije imaju određena ograničenja, kao što su ograničenja veličine simulacijske kutije i ograničeno simulacijsko vrijeme [121]. Zbog tih problema, simulacijski rezultati su često netočni u opisu dugodosežne strukture mikro–heterogenih sustava, primjerice vodene otopine *t*–butanola. Za takve sustave u ovome radu uveden je pojam molekularne emulzije. Integralne jednadžbe predstavljaju alternativu računalnim simulacijama, no ni one nisu savršen alat. Njihov glavni problem je to što sadrže brojne aproksimacije korelacija višeg reda pa ne mogu ispravno opisati fluktuacije [1, 13].

Jednadžba koja povezuje par korelacijsku funkciju $h(1,2)$ i totalnu korelacijsku funkciju $c(1,2)$ naziva se Ornstein–Zernike (OZ) jednadžba [1]

$$h(1,2) = c(1,2) + \rho \int c(1,3) h(3,2) d3 \,, \tag{1}$$

gdje se integracija provodi po položaju i orjentaciji čestice označene brojem 3.

Sama OZ jednadžba nije dovoljna za opis korelacija pa se koristiti i tzv. jednadžba „zatvaranja“ (eng. closure) koja je definirana kao

$$g(1,2) = \exp(-\beta U(1,2) + h(1,2) - c(1,2) + b(1,2)), \quad (2)$$

gdje je $b(1,2)$ tzv. funkcija „most“ (eng. bridge) koja uključuje korelacije višeg reda. Budući da je tu funkciju gotovo nemoguće direktno izračunati, koriste se brojne aproximacije izraza, kao što su Percus–Yewick jednadžba (PY) i Hypernetted–chain jednadžba (HNC), kao najpoznatije. Postoje dva različita pristupa unutar teorije integralnih jednadžbi. Prvi pristup je preko Molekularne Ornstein–Zernike jednadžbe (MOZ) koja računa molekularnu distribucijsku funkciju. Drugi pristup je preko Ornstein–Zernike jednadžbe za parove atoma (site–site Ornstein Zernike, SSOZ) koja uključuje korelacijske funkcije za parove atoma. Teorijski uvod sadržan je u drugom poglavlju.

U trećem poglavlju, SSOZ jednadžba korištena je u analizi nekoliko jednokomponentnih sustava tekućina, a to su: aceton, tetrakloruglјik, formamid, metanol, etanol, 1–propanol i voda. Korištene su PY i HNC jednadžba, kao i jednadžba koja sadrži korelacije višeg reda iz simulacija. Struktura jednokomponentnih tekućina dobro je opisana rezultatima SSOZ jednadžbe sa svim jednadžbama zatvaranja, no najbolji rezultat je postignut s jednadžbom koja sadrži korelacije višeg reda.

U četvrtom poglavlju, SSOZ jednadžba korištena je u analizi dvokomponentnih sustava: voda i metanol, te voda i modificirana voda. Pokazalo se da SSOZ jednadžba ne predviđa ispravno strukturu vodenih mješavina. To znači da aproksimacije unutar integralnih jednadžbi ozbiljno štete točnosti rezultata strukture mikro–heterogenih sustava.

U petom poglavlju, korištena je metoda Molekularne Dinamike za proučavanje vodenih otopina etanola. Analizirana je struktura i termodinamika za svaku koncentraciju etanola od $x = 0$ do $x = 1$, s prirastom od 0.10. Primijećeno je 2048 molekula u simulacijskoj kutiji te 2–3 ns simulacijskog vremena nisu dovoljni za ispravan opis dugodosežnih korelacija, budući da radijalna distribucijska funkcija ne konvergira u jednu jedinstvenu vrijednost. Zbog toga je teško izračunati vrijednosti integrala radijalne distribucijske funkcije (KBI).

U šestom poglavlju, uveden je pojam molekularne emulzije za sustav kojeg karakterizira stvaranje domena molekula. Taj pojam razvijen je u analogiji s mikro–emulzijama koje su definirane kao mješavine ulja, vode i emulgatora. U mikro–emulzijama se stvaraju objekti poput domena i micela. Radijalna distribucijska funkcija u mikro–emulzijama ima poseban matematički zapis u termodinamičkom limitu koji uključuje dvije relevantne duljine [156]. Taj zapis je iskorišten za nastavak radijalne distribucijske funkcije preko veličine simulacijske kutije.

U sedmom poglavlju, proučavane su struktura i termodinamička svojstva vodena otopine *t*–butanola i vodene otopine acetona metodom Molekularne Dinamike. U analizi tih otopina također je korišten koncept molekulane emulzije. Pomoću nastavka radijalne distribucijske funkcije u područje dugog dosega, izračunate su KBI vrijednosti koje su u blizini eksperimentalnih vrijednosti. Analizirane su termodinamičke veličine, kao što su volumeni, entalpije i njihove „excess" vrijednosti za obje vodene otopine te su uspoređene s vrijednostima iz literature.

U osmom poglavlju, proučavane su vodene otopine tri različita modela modificirane vode metodom Molekularne Dinamike. Zbog mikro–heterogenosti tih sustava, primijenjen je koncept molekularne emulzije, a strukturalni rezultati se dobro slažu s rezultatima većih simulacija. KBI vrijednosti izračunate preko koncepta molekularne emulzije slažu se s vrijednostima izračunatim termodinamičkim putem.

U devetom poglavlju proučena je mještavina tekućina sastavljenih od sfernih čestica koje međudjeluju preko „soft–core" potencijala. Korištena je Monte Carlo metoda te OZ jednadžba u paru s HNC jednadžbom. Rezultati obje metode se slažu za sve koncentracije, no odstupanja u opisu strukture se javljaju kod onih koncentracija gdje postoji najizraženija mikro–heterogenost.

U desetom poglavlju predstavljen je zaključak. Simulacije i teorija imaju poteškoća u opisu mikro–heterogenosti. U strukturnom faktoru $S(k)$ mikro–heterogenih sustava zapažen je maksimum između valnog vektora $k = 0$ i položaja glavnog maksimuma. Mikro–emulzije su primjer takvih sustava [211], te neki koloidni sustavi [212]. Nastavak radijalne distribucijske funkcije $g(r)$, koji je predstavljen u ovome radu, pomaže da se u korelacijskim funkcijama vodenih otopina pojave oscilacije koje upućuju na stvaranje domena. Taj postupak se može uspješno primijeniti na vodene otopine kod kojih postoje dobro definirane domene, kao što su vodene otopine *t*–butanola ili acetona, no ne i na vodene otopine metanola i etanola gdje su domene slabije definiranog oblika. Mikro–heterogenost je vrsta lokalnog reda te nije dobro definirana unutar Statističke Fizike, za razliku od globalnog reda za kojeg postoje teorije utemeljene na parametru reda.

Appendix A

The derivation of the SSOZ equation

A.1 Introduction

The SSOZ equation can be derived formally from the MOZ equation. We follow the derivation of Gray and Gubbins from „Theory of molecular fluids“ [23]. The site–site pair correlation function $h_{\alpha\beta}(r_{\alpha\beta}) = g_{\alpha\beta}(r_{\alpha\beta}) - 1$ is defined in *r*–space by Eq. (1.11), and is given in *k*–space as

$$\widetilde{h}_{\alpha\beta}(k) = \frac{1}{\Omega^2}\iint d\Omega_1 d\Omega_2 \widetilde{h}(1,2)\exp\left[i\vec{k}(\vec{r}_{2\beta}(\Omega_2) - \vec{r}_{1\alpha}(\Omega_1))\right]. \tag{A.1}$$

We define $c_{\alpha\beta}(r_{\alpha\beta})$ in *r*–space analogously to Eq. (1.11) and in *k*–space as

$$\widetilde{c}_{\alpha\beta}(k) = \frac{1}{\Omega^2}\iint d\Omega_1 d\Omega_2 \widetilde{c}(1,2)\exp\left[i\vec{k}(\vec{r}_{2\beta}(\Omega_2) - \vec{r}_{1\alpha}(\Omega_1))\right] \ . \tag{A.2}$$

We will use a shorthand notation for easier manipulation later by writing one of these equations as

$$\widetilde{h}_{\alpha\beta}(k) = \overline{P}_{1\alpha} P_{2\beta} \widetilde{h}(1,2), \tag{A.3}$$

where the bar indicates a complex conjugate and the generic projector notation is defined as

$$P_{1\alpha} = \frac{1}{\Omega}\int d\Omega_1 \exp\left[i\vec{k}(\vec{r}_{1\alpha}(\Omega_1))\right]. \tag{A.4}$$

Next, we assume that the direct molecular correlation function $c(1,2)$ can be written as a sum of partial site–site direct correlation functions, just like one does for the pair interaction in Eq. (1.7). We write an approximate equation for $c(1,2)$ as

$$c(1,2) = \sum_{\alpha\beta} \bar{c}_{\alpha\beta}(r_{\alpha\beta}), \tag{A.5}$$

where we expect that $c_{\alpha\beta} \neq \bar{c}_{\alpha\beta}$ in general. The Fourier transform of Eq. (A.5) is given by

$$\tilde{c}(1,2) = \sum_{\alpha\beta} \int d\vec{r}\, \bar{c}_{\alpha\beta}(r_{\alpha\beta}) \exp(i\vec{r}\vec{k}). \tag{A.6}$$

The integral on the right–hand side is for fixed orientations, so we have fixed $\vec{r}_{1\alpha}$ and $\vec{r}_{2\beta}$. Transforming the integration variable to $\vec{r}_{\alpha\beta}$ using $\vec{r}_{\alpha\beta} = \vec{r} + \vec{r}_{2\beta} - \vec{r}_{1\alpha}$ and substituting $d\vec{r} = d\vec{r}_{ab}$, we get

$$\begin{aligned} \tilde{c}(1,2) &= \sum_{\alpha\beta} \left(\int d\vec{r}_{\alpha\beta} \bar{c}_{\alpha\beta}(r_{\alpha\beta}) \exp(-i\vec{r}_{\alpha\beta}\vec{k}) \exp\left[i\vec{k}(\vec{r}_{1\alpha} - \vec{r}_{2\beta})\right] \right) \\ &= \sum_{\alpha\beta} \tilde{\bar{c}}_{\alpha\beta}(k) \exp\left[i\vec{k}(\vec{r}_{1\alpha} - \vec{r}_{2\beta})\right]. \end{aligned} \tag{A.5}$$

We now work out the relation between the two types of site–site direct correlation functions, $c_{\alpha\beta}$ and $\tilde{\bar{c}}_{\alpha\beta}$. Substituting Eq. (A.5) in Eq. (A.2) we get

$$\tilde{c}_{\alpha\beta}(k) = \sum_{\alpha'\beta'} \tilde{\bar{c}}_{\alpha'\beta'}(k) \left(\frac{1}{\Omega} \int d\Omega_1 \exp\left[i\vec{k}(\vec{r}_{1\alpha'} - \vec{r}_{1\alpha})\right] \right) \left(\frac{1}{\Omega} \int d\Omega_2 \exp\left[i\vec{k}(\vec{r}_{2\beta} - \vec{r}_{2\beta'})\right] \right). \tag{A.6}$$

We define the real and symmetric matrix $\tilde{W} = \left[\tilde{w}_{\alpha'\alpha}(k)\right]$ by

$$\tilde{w}_{\alpha'\alpha}(k) = \frac{1}{\Omega} \int d\Omega_1 \exp\left[i\vec{k}(\vec{r}_{1\alpha'} - \vec{r}_{1\alpha})\right], \tag{A.7}$$

which for rigid molecules becomes

$$\tilde{w}_{\alpha'\alpha}(k) = \frac{\sin kl_{\alpha'\alpha}}{kl_{\alpha'\alpha}}, \tag{A.8}$$

where $l_{\alpha'\alpha} = \left|\vec{r}_{1\alpha'} - \vec{r}_{1\alpha}\right|$ is the intramolecular $\alpha\alpha'$ separation. Thus, from Eqs. (A.6) and (A.8), we have

$$\tilde{c}_{\alpha\beta}(k) = \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}(k) \tilde{\bar{c}}_{\alpha'\beta'}(k) \tilde{w}_{\beta'\beta}(k), \tag{A.9}$$

or, in matrix notation:

$$\widetilde{C} = \widetilde{W}\widetilde{\widetilde{C}}\widetilde{W} \,. \tag{A.10}$$

This equation therefore relates the true site–site direct correlation functions to the approximate version assumed from the additivity assumption of the direct correlation functions.

A.2 Derivation: one–component system

We start from the MOZ equation written in *k*–space:

$$\widetilde{h}(1,2) = \widetilde{c}(1,2) + \frac{\rho}{\Omega}\int d\Omega_3 \widetilde{c}(1,3)\widetilde{h}(3,2) \,. \tag{A.11}$$

One can project this equation on a pair of sites following Eq. (A.1) and using the projector notation:

$$\widetilde{h}_{\alpha\beta}(k) = \widetilde{c}_{\alpha\beta}(k) + \frac{\rho}{\Omega}\int d\Omega_3 \overline{P}_{1\alpha}\widetilde{c}(1,3)P_{2\beta}\widetilde{h}(3,2) \,, \tag{A.12}$$

where we can replace $\widetilde{c}(1,3)$ by the approximate expression using Eq.(A.5):

$$\widetilde{c}(1,3) = \sum_{\alpha\beta}\widetilde{\widetilde{c}}_{\alpha\beta}(k)\exp\left[i\vec{k}(\vec{r}_{1\alpha}(\Omega_1) - \vec{r}_{3\beta}(\Omega_3))\right]. \tag{A.13}$$

Inserting Eq. (A.13) into Eq. (A.12), we get

$$\widetilde{h}_{\alpha\beta}(k) = \widetilde{c}_{\alpha\beta}(k) + \frac{\rho}{\Omega}\int d\Omega_3 P_{1\alpha}\sum_{\alpha'\beta'}\widetilde{\widetilde{c}}_{\alpha'\beta'}(k)\exp\left[i\vec{k}(\vec{r}_{1\alpha'} - \vec{r}_{3\beta'})\right]P_{2\beta}\widetilde{h}(3,2) \,. \tag{A.14}$$

We can integrate the two exponential terms over Ω_1 to get

$$\widetilde{w}_{\alpha'\alpha}(k) = \frac{1}{\Omega}\int d\Omega_1 \exp\left[i\vec{k}\left(\vec{r}_{1\alpha'}(\Omega_1) - \vec{r}_{1\alpha}(\Omega_1)\right)\right] \tag{A.15}$$

and inserting it into Eq. (A.14), we have

$$\widetilde{h}_{\alpha\beta}(k) = \widetilde{c}_{\alpha\beta}(k) + \rho\sum_{\alpha'\beta'}\widetilde{w}_{\alpha'\alpha}(k)\widetilde{\widetilde{c}}_{\alpha'\beta'}(k)\frac{1}{\Omega}\int d\Omega_3 \exp\left[-i\vec{k}(\vec{r}_{3\beta'})\right]P_{2\beta}\widetilde{h}(3,2) \,. \tag{A.16}$$

Writing the last integral as

$$\frac{1}{\Omega^2}\iint d\Omega_2 d\Omega_3 \widetilde{h}(3,2)\exp\left[i\vec{k}(\vec{r}_{2\beta}(\Omega_2) - \vec{r}_{3\beta'}(\Omega_3))\right] = \widetilde{h}_{\beta'\beta}(k) \,, \tag{A.17}$$

we rewrite Eq. (A.16) as

$$\tilde{h}_{\alpha\beta}(k) = \tilde{c}_{\alpha\beta}(k) + \rho \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}(k)\tilde{\bar{c}}_{\alpha'\beta'}(k)\tilde{h}_{\beta'\beta}(k) . \tag{A.18}$$

Finally, from Eqs.(A.9) and (A.18), we get

$$\tilde{h}_{\alpha\beta}(k) = \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}(k)\tilde{\bar{c}}_{\alpha'\beta'}(k)\tilde{w}_{\beta'\beta}(k) + \rho \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}(k)\tilde{\bar{c}}_{\alpha'\beta'}(k)\tilde{h}_{\beta'\beta}(k) \tag{A.20}$$

or in matrix notation:

$$\tilde{H} = \tilde{W}\tilde{\bar{C}}\tilde{W} + \rho\tilde{W}\tilde{\bar{C}}\tilde{H} . \tag{A.21}$$

This is the well known SSOZ equation.

A.3 Derivation: *n*–component system

We consider a mixture of *n* different species. The projector technique allows projecting each of the correlation functions over sites (α,β) belonging to species (a,b) respectively, so we get

$$\tilde{h}_{ab}^{\alpha\beta}(k) = P_{1\alpha}^{a} P_{2\beta}^{b} \tilde{h}(1,2) \tag{A.23}$$

where the projector notation is given by

$$P_{1\alpha}^{a} = \frac{1}{\Omega}\int d\Omega_1 \exp\left[i\vec{k}(\vec{r}_{1\alpha}^{\,a}(\Omega_1))\right],$$

$$P_{2\beta}^{b} = \frac{1}{\Omega}\int d\Omega_2 \exp\left[i\vec{k}(\vec{r}_{2\beta}^{\,b}(\Omega_2))\right], \tag{A.24}$$

where $\vec{r}_{1\alpha}^{\,a}$ and $\vec{r}_{2\beta}^{\,b}$ denote the position of site α and β on molecule of species *a* and *b*, respectively. We apply this projector to the MOZ equation and we get

$$\tilde{h}_{ab}^{\alpha\beta}(1,2) = \tilde{c}_{ab}^{\alpha\beta}(1,2) + \sum_{c} \frac{\rho_c}{\Omega}\int d\Omega_3 P_{1\alpha}^{a}\tilde{c}_{ac}(1,3)P_{2\beta}^{b}\tilde{h}_{cb}(3,2) . \tag{A.25}$$

As for the one–component system, we write the direct correlation function in terms of pairs of sites:

$$c_{ab}(1,2) = \sum_{\alpha\beta} \bar{c}_{ab}^{\alpha\beta}(r_{\alpha\beta}) . \tag{A.26}$$

The Fourier transform of Eq. (A.26) is given by

$$\tilde{c}_{ab}(1,2) = \sum_{\alpha\beta} \tilde{\tilde{c}}_{ab}^{\alpha\beta}(k) \exp\left[i\vec{k}\left(\vec{r}_{1\alpha}^{\,a}(\Omega_1) - \vec{r}_{2\beta}^{\,b}(\Omega_2)\right)\right]. \tag{A.27}$$

The site–site direct correlation function $\tilde{c}_{ab}^{\alpha\beta}(k)$ can be written applying Eq.(A.2) as

$$\tilde{c}_{ab}^{\alpha\beta}(k) = \frac{1}{\Omega^2}\iint d\Omega_1 d\Omega_2 \tilde{c}_{ab}(1,2) \exp\left[i\vec{k}(\vec{r}_{2\beta}^{\,b}(\Omega_2) - \vec{r}_{1\alpha}^{\,a}(\Omega_1))\right] \tag{A.28}$$

Using Eqs. (A.27) and (A.28), we get

$$\tilde{c}_{ab}^{\alpha\beta}(1,2) = \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}^{a}(k) \tilde{\tilde{c}}_{ab}^{\alpha'\beta'} \tilde{w}_{\beta'\beta}^{b}(k)\,, \tag{A.29}$$

where the function $\tilde{w}_{\alpha\alpha'}^{a}$ is defined as

$$\tilde{w}_{\alpha\alpha'}^{a}(k) = \frac{1}{\Omega}\int d\Omega_1 \exp\left[i\vec{k}\left(\vec{r}_{1\alpha'}^{\,a}(\Omega_1) - \vec{r}_{1\alpha}^{\,a}(\Omega_1)\right)\right] = \frac{\sin(kr_{\alpha\alpha'}^{a})}{kr_{\alpha\alpha'}^{a}}. \tag{A.30}$$

Since the integration in Eq. (A.29) includes only the sites belonging to the same species, the matrix $\tilde{W} = \left[\tilde{w}_{\alpha\alpha'}^{ab}(k)\right]$ is defined as

$$\tilde{w}_{\alpha\alpha'}^{ab}(k) = \tilde{w}_{\alpha\alpha'}^{a}(k)\delta_{ab}\,, \tag{A.31}$$

The angular integral over Ω_2 and Ω_3 from Eq. (A.25) can be written as

$$\frac{1}{\Omega^2}\iint d\Omega_2 d\Omega_3 \tilde{h}_{cb}(3,2) \exp\left[i\vec{k}(\vec{r}_{2\beta}^{\,b}(\Omega_2) - \vec{r}_{3\beta'}(\Omega_3))\right] = \tilde{h}_{\beta'\beta}^{cb}(k) \tag{A.32}$$

and we rewrite Eq. (A.25) as

$$\tilde{h}_{ab}^{\alpha\beta}(1,2) = \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}^{a}(k) \tilde{\tilde{c}}_{ab}^{\alpha'\beta'} \tilde{w}_{\beta'\beta}^{b}(k) + \sum_{c} \rho_c \sum_{\alpha'\beta'} \tilde{w}_{\alpha\alpha'}^{a}(k) \tilde{\tilde{c}}_{ac}^{\alpha'\beta'}(k) \tilde{h}_{\beta'\beta}^{cb}(k)\,. \tag{A.33}$$

If we define the species sub–matrices as $\tilde{H}_{ab} = \left[\tilde{h}_{ab}^{\alpha\beta}(k)\right]$, $\tilde{\tilde{C}}_{ab} = \left[\tilde{\tilde{c}}_{ab}^{\alpha\beta}(k)\right]$ and $\tilde{W}_a = \left[\tilde{w}_{\alpha\alpha'}^{a}(k)\right]$, then Eq. (A.33) becomes

$$\tilde{H}_{ab} = \tilde{W}_a \tilde{\tilde{C}}_{ab} \tilde{W}_b + \sum_{c} \rho_c \tilde{W}_a \tilde{\tilde{C}}_{ac} \tilde{H}_{cb}\,, \tag{A.34}$$

Inserting the full matrix notation, $\tilde{H} = \left[\tilde{H}_{ab}\right]$, $\tilde{\tilde{C}} = \left[\tilde{\tilde{C}}_{ab}\right]$, $\tilde{W} = \left[\tilde{W}_a\right]$ and $R = \left[\rho_c\right]$ reduces Eq. (A.34) to the following expression:

$$\tilde{H} = \tilde{W}\tilde{\tilde{C}}\tilde{W} + \tilde{W}\tilde{\tilde{C}}R\tilde{H}\,, \tag{A.35}$$

which is the SSOZ equation for the multicomponent mixture. To make Eq. (A.35) even simpler, we can define the generalized site–site structure factor matrix S as

$$S = \tilde{W} + R\tilde{H} \tag{A.36}$$

and the matrix M as

$$M = \tilde{W}\tilde{C}\,. \tag{A.37}$$

Inserting Eqs. (A.36) and (A.37) into Eq. (A.35), we get

$$\tilde{H} = \tilde{W}\tilde{\tilde{C}}(\tilde{W} + R\tilde{H}) = MS\,. \tag{A.38}$$

Appendix B

The derivation of TS functional form of the correlations from the MOZ equation

For a one component fluid, the MOZ matrix equation in *k*–space is given by the expression (1.31). Now we define $\widetilde{N}_\chi(k) = \widetilde{H}_\chi(k) - \widetilde{C}_\chi(k)$, insert into Eq. (1.31) and after a few rearrangements, we have

$$\widetilde{H}_\chi(k) = \widetilde{C}_\chi(k)\left[I - (-)^\chi \rho \widetilde{C}^*_\chi(k)\right]^{-1}, \tag{B.1}$$

where ρ is the number density, the matrices $\widetilde{H}_\chi(k) = \left[\widetilde{h}^{mn\chi}_{\mu\nu}(k)\right]$ and $\widetilde{C}_\chi(k) = \left[\widetilde{c}^{mn\chi}_{\mu\nu}(k)\right]$ contain χ–transformations of the total and the direct pair correlation functions, respectively. Here we use the notation $PC_\chi(k) = C^*_\chi(k)$ because the matrix P which is defined by Eq. (1.30) changes the sign of the elements of in $C_\chi(k)$. We want to examine the small–*k* behaviour of the function $h_{ij,mn\chi}(k)$. Any function $h_{ij,mn\chi}(k)$ will necessarily have the determinant of the matrix $\widetilde{D}(k) = \left[I - (-)^\chi C^*_\chi(k)\right]$ in the denominator. This determinant can be written in short as $1 - \widetilde{\gamma}_\chi(k)$, where the function $\widetilde{\gamma}_\chi(k)$ contains all of the $\widetilde{c}_{ij,mn\chi}(k)$ functions and their various products. Therefore, $\widetilde{h}_{ij,mn\chi}(k)$ will be of the form:

$$\widetilde{h}_{ij,mn\chi}(k) = \frac{\widetilde{t}_{ij,mn\chi}(k)}{1 - \widetilde{\gamma}_\chi(k)}, \tag{B.2}$$

where functions $\widetilde{t}_{ij,mn\chi}(k)$ are elements of a matrix which is a product of $\widetilde{C}_\chi(k)$ and the adjugate matrix of $\widetilde{D}(k)$. The direct correlation function in the *k*–space $\widetilde{C}(1,2)$ can be written as the sum of

products of functions $\widetilde{c}_{ij,mn\chi}(k)$ and the rotational invariants, as given in Eq. (1.22). Since the function $\widetilde{C}(1,2)$ is short–ranged, any function $\widetilde{c}_{ij,mn\chi}(k)$ in the sum is also short–ranged. Therefore, we can expand the function $\widetilde{\gamma}_{\chi}(k)$, which contains only short–ranged functions and their products as Taylor series around $k = 0$:

$$\tilde{\gamma}_{\chi}(k \to 0) = \tilde{\gamma}_{\chi;0} + k^2\tilde{\gamma}_{\chi;2} + k^4\tilde{\gamma}_{\chi;4} + k^6\tilde{\gamma}_{\chi;6} + \dots, \tag{B.3}$$

where coefficients are determined through the expansion of the Fourier–Hankel transform. Since it is quite difficult to express each of the $\widetilde{\gamma}_{\chi;n}(k)$ terms as an integral over $\widetilde{\gamma}_{\chi}(k)$, which is an infinite sum of all Fourier–Hankel transforms $\widetilde{c}_{ij,mn\chi}(k)$ defined in Eq. (1.24) and their products, we will keep these terms are they are. We note that there are only even powers of k in Eq. (B.3) because of the symmetry of all correlation functions where $a(-r) = a(r)$. If we retain in Eq. (B.3) up to the k^2–term and insert into Eq. (B.2), we will get the Ornstein–Zernike expression:

$$\widetilde{h}_{ij,mn\chi}(k \to 0) \approx \frac{\widetilde{t}_{ij,mn\chi}(k \to 0)}{1 - \widetilde{\gamma}_{\chi;0}(k) - k^2\widetilde{\gamma}_{\chi;2}(k)} = \frac{A_{ij;mn\chi}}{\xi^{-2} + k^2}, \tag{B.4}$$

where ξ is the correlation length expressed as $\xi = \sqrt{\dfrac{-\widetilde{\gamma}_{\chi;2}}{1 - \widetilde{\gamma}_{\chi;0}}}$.

The inverse Fourier transform of the above function is

$$\lim_{r\to\infty} h_{ij,mn\chi}(r) \approx A_{ij;mn\chi} \exp\frac{(-r/\xi)}{r}. \tag{B.5}$$

The correlation length in Eq. (B.5) is the same for all projections and it is related to the density fluctuations in a pure liquid or to the concentration fluctuations in a mixture [1].

What happens if we retain more terms than the k^2–term in Eq. (B.3)? Since the k^4 –expansion has been successfully used both in micro–emulsions and in aqueous mixtures of relatively short chain molecules [218, 219], we have a good reason to try this approximation first. The TS approximation of the MOZ equation follows from Eqs. (B.2) and (B.3):

$$\widetilde{h}^{(TS)}_{ij,mn\chi}(k \to 0) = \frac{\widetilde{t}_{ij,mn\chi}(0)}{1 - \widetilde{\gamma}_{\chi;0}(k) - k^2\widetilde{\gamma}_{\chi;2}(k) - k^4\widetilde{\gamma}_{\chi;4}(k)} = \frac{\widetilde{t}_{ij,mn\chi}(0)}{a_2 + k^2 c_1 + k^4 c_2}, \tag{B.6}$$

where we have adopted the original TS notation [156] in the denominator of the second equality. The correlation function $h_{ij,mn\chi}^{(TS)}(r)$ has the following form [156]

$$h_{ij,mn\chi}^{(TS)}(r) \approx \frac{\bar{d}}{r}\exp\left(-\frac{r}{\xi}\right)\sin\left(\frac{r}{\bar{d}}\right), \tag{B.7}$$

where ξ is the correlation length, d is the domain size and $\bar{d} = d/2\pi$ .Taking the Fourier transform of the function above leads to the following:

$$\tilde{h}_{ij,mn\chi}^{(TS)}(k) = \int \exp(-i\vec{k}\vec{r}) h_{ij,mn\chi}^{(TS)}(r) d\vec{r} = 4\pi \int_0^\infty \frac{\sin(kr)}{kr} h_{ij,mn\chi}^{(TS)}(r) r^2 dr. \tag{B.8}$$

Next, we insert Eq. (B.7) into Eq. (B.8) and calculate $\tilde{h}_{ij,mn\chi}^{(TS)}(k)$ as

$$\tilde{h}_{ij,mn\chi}^{(TS)}(k) = 4\pi \frac{2\bar{d}(\bar{d}\xi)^3}{(\bar{d}^2+\xi^2)^2 + 2(\bar{d}\xi)^2(\bar{d}^2-\xi^2)k^2 + (\bar{d}\xi)^4 k^4} \tag{B.9}$$

Comparing Eqs. (B.6) and (B.9), we can define the TS coefficients in the following manner (REF):

$$\begin{aligned} a_2 &= 1 - \tilde{\gamma}_{\chi;0}(k) = (\bar{d}^2+\xi^2)^2, \\ c_1 &= -\tilde{\gamma}_{\chi;2}(k) = 2(\bar{d}\xi)^2(\bar{d}^2-\xi^2), \\ c_2 &= -\tilde{\gamma}_{\chi;4}(k) = (\bar{d}\xi)^4. \end{aligned} \tag{B.10}$$

We can now write the TS correlation function with two length scales ($\bar{d}, \xi$) as the inverse Fourier transform of the form:

$$h_{ij,mn\chi}^{(TS)}(r) = \int \exp(i\vec{k}\vec{r}) \tilde{h}_{ij,mn\chi}^{(TS)}(k) d\vec{k} = \frac{\tilde{t}_{ij,mn\chi}(0)}{\pi^2(\bar{d}\xi)^3} \frac{\exp(-r/\xi)}{r} \sin(r/\bar{d}). \tag{B.11}$$

Appendix C

The derivation of the LP correction

C.1 The one component system

We give a simple derivation of the Lebowitz–Percus (LP) correction [100] of the long–range behaviour of the radial distribution function $g(r)$ in the NVT ensemble. While this derivation stands exactly for this ensemble, it is not obvious how it can be derived in the NPT ensemble.

Isothermal compressibility $\chi_T = (1/\rho)(\partial\rho/\partial P)_T$ is related to the integral of the $g(r)$ by the expression [1]:

$$S(0) = 1 + \rho \int d\vec{r}\,(g(r) - 1) = 1 + \rho\,\tilde{h}(0) = \chi_T^*, \tag{C.1}$$

where $\chi_T^* = \chi_T / \chi_T^0$ is the reduced compressibility and $\chi_T^0 = 1/(\rho k_B T)$ is the compressibility of the ideal gas. One can write the exact $g(r)$ as the sum of the $g_{MD}(r)$ evaluated in the NVT ensemble simulations and a function $\delta(r)$ which is a step function that is zero inside the core and raises to the LP correction value ε_{LP} around the first peaks:

$$g(r) = g_{MD}(r) + \delta(r) \text{ or} \tag{C.2}$$

$$\lim_{r\to\infty} g_{MD}(r) = 1 - \varepsilon_{LP}. \tag{C.3}$$

Combining Eqs. (C.1) and (C.2), we have:

$$\tilde{h}_{MD}(0) = \frac{\chi_T^*}{\rho} - \frac{1}{\rho} - D, \tag{C.4}$$

where $D = \int d\vec{r}\,\delta(r) \approx V\varepsilon_{LP}$. Here, V is the volume of the entire system and the approximation is obtained by neglecting the small contribution due to core integration. Using $\rho = N/V$, we rewrite the MD contribution as

$$\tilde{h}_{MD}(0) = \frac{\chi_T^*}{\rho} - \frac{1}{\rho} - \frac{N\varepsilon_{LP}}{\rho}. \tag{C.5}$$

The definition of the *n*–body correlation function $\rho^{(n)}(1,2,\ldots,n) \equiv \rho^{(n)}(n)$ in a N–constant ensemble is

$$\rho^{(n)}(n) = \frac{N!}{(N-n)!}\frac{1}{Z_N}\int d(n+1)\ldots dN \exp\left[-\beta U(N)\right], \tag{C.6}$$

where $Z_N = \int d1\ldots dN \exp\left[-\beta U(N)\right]$ is the canonical ensemble configurational integral and $U(N) = U(1,2,\ldots,N)$ is the total interaction energy in the system. The *n*–body distribution function $g^{(n)}(n)$ is defined from the *n*–body correlation function as:

$$g^{(n)}(n) = \frac{\rho^{(n)}(n)}{\prod_{i=1}^{n}\rho^{(1)}(i)}\ . \tag{C.7}$$

The integral of $\rho^{(n)}(n)$ over the positions of all *n* particles is given by:

$$\int d1\ldots dn\rho^{(n)}(n) = \frac{N!}{(N-n)!}. \tag{C.8}$$

For the pair correlation function $\rho^{(2)}(1,2)$ one obtains:

$$\int d1d2\rho^{(2)}(1,2) = N(N-1). \tag{C.9}$$

If the system is homogeneous, then $\rho^{(1)}(1) = \rho$ and the pair correlation function depends only on the radial distance between the two particles, $g^{(2)}(1,2) = g(r)$, where $r = \left|\vec{r}_1 - \vec{r}_2\right|$ is the interparticle distance. From Eq. (C.7) the pair correlation function is given by $\rho^{(2)}(1,2) = \rho^2 g(r)$ and the normalisation condition (C.8) is given by

$$\rho^2\int d1d2g(r) = \rho^2 V\int d\vec{r}g(r) = N(N-1), \tag{C.10}$$

which is simply

$$\int d\vec{r} g(r) = V(1 - \frac{1}{N}). \tag{C.11}$$

Using $g(r) = h(r) + 1$, we obtain

$$\widetilde{h}(0) = \int d\vec{r} h(r) = -\frac{1}{\rho}. \tag{C.12}$$

From the condition $\widetilde{h}(0) = \widetilde{h}_{MD}(0)$, Eq. (C.12) and Eq. (C.5), we get the value ε_{LP} of the LP correction in the NVT ensemble as

$$\varepsilon_{LP} = \frac{1}{N}\chi_T^* = \frac{1}{N}(\frac{\partial\rho}{\partial\beta P})_T, \tag{C.13}$$

which can also be written as

$$\varepsilon_{LP} = \frac{1}{N}(\frac{\partial\rho}{\rho\,\partial\beta\mu})_T, \tag{C.14}$$

where we used the exact thermodynamical relation involving the chemical potential μ :

$$1 - \rho\widetilde{c}(0) = \frac{\partial\beta P}{\partial\rho} = \rho\frac{\partial\beta\mu}{\partial\rho}. \tag{C.15}$$

Inserting Eq. (C.14) into Eq. (C.3), we obtain the asymptotic form of the LP correction for $g_{MD}(r)$:

$$\lim_{r\to\infty} g_{MD}(r) = 1 - \frac{1}{N}(\frac{\partial\rho}{\partial\beta P})_T = 1 - \frac{1}{N}(\frac{\partial\rho}{\rho\,\partial\beta\mu})_T. \tag{C.16}$$

C.2 The *n*–component system

We follow a procedure similar to that in the one component system and derive the LP correction for a *n*–component system. In the μVT ensemble one has [3]

$$B_{ij} = \sqrt{\rho_i\rho_j}(\delta_{ij} + \sqrt{\rho_i\rho_j}\,\widetilde{h}_{ij}(0)) = \sqrt{\rho_i\rho_j}S_{ij}(0) = \frac{\sqrt{\langle N_i\rangle\langle N_j\rangle}}{V}S_{ij}(0). \tag{C.17}$$

Matrix $B = \lfloor B_{ij} \rfloor$ is the inverse matrix of $A = \lfloor A_{ij} \rfloor$:

$$B_{ij} = \frac{cof\, A_{ij}}{\det A} = \langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle . \tag{C.18}$$

Combining Eqs. (C.17) and (C.18), we get

$$S_{ij}(0) = V \frac{\langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle}{\sqrt{\langle N_i \rangle \langle N_j \rangle}}, \tag{C.19}$$

where we have used $\rho_i = \langle N_i \rangle / V$. We note that there is a factor V in Eq. (C.19) which should not be there. Indeed, since A is dimensless, B is also and it cannot be defined by the equation that has a factor $1/V$ precisely. So we redefine B_{ij} as

$$B_{ij} = V \sqrt{\rho_i \rho_j} (\delta_{ij} + \sqrt{\rho_i \rho_j}\, \tilde{h}_{ij}(0)) = V \sqrt{\rho_i \rho_j} S_{ij}(0) \tag{C.20}$$

and we now get

$$S_{ij}(0) = \frac{\langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle}{\sqrt{\langle N_i \rangle \langle N_j \rangle}} . \tag{C.21}$$

From Ref. [3] we also have

$$A_{ij} = (\frac{\partial \beta \mu_i}{\partial N_j})_{TVN_k} = \frac{1}{V} (\frac{\partial \beta \mu_i}{\partial \rho_j})_{TVN_k} \tag{C.22}$$

and the inverse matrix is given by

$$B_{ij} = V (\frac{\partial \rho_i}{\partial \beta \mu_j})_{TVN_k} . \tag{C.23}$$

From Eqs. (C.20) and (C.23), we get the relation equivalent to Eq. (C.1):

$$S_{ij}(0) = \frac{B_{ij}}{\sqrt{\langle N_i \rangle \langle N_j \rangle}} = \frac{1}{\sqrt{\rho_i \rho_j}} (\frac{\partial \rho_i}{\partial \beta \mu_j})_{TVN_k} . \tag{C.24}$$

From Eqs. (C.20) and (C.24), we have

$$\tilde{h}_{ij}(0) = \frac{1}{\rho_i \rho_j}\left(\frac{\partial \rho_i}{\partial \beta \mu_j}\right)_{TVN_k} - \frac{\delta_{ij}}{\sqrt{\rho_i \rho_j}}. \tag{C.25}$$

If we now express the exact $h_{ij}(r)$ in terms of the MD results as

$$h_{ij}(r) = h_{ij;MD}(r) + \delta_{ij}(r) \tag{C.26}$$

this leads to

$$\tilde{h}_{ij}(0) = \tilde{h}_{ij;MD}(0) + V\varepsilon_{ij}, \tag{C.27}$$

where ε_{ij} is the LP correction. Eq. (C.27) was obtained with the same assumption as in the one component case where the core part contribution to the integral over the $\delta_{ij}(r)$ function is negligible. Next, we demonstrate the exact limit of $\tilde{h}_{ij;MD}(0)$ for the NVT ensemble. For a mixture of *n*–species, the configurational integral Z_N is defined as

$$Z_N = Z_N(N_1, \ldots, N_n) = \int dN_1 dN_2 \ldots dN_n \exp\left[-\beta U(N_1, \ldots, N_n)\right], \tag{C.28}$$

where $dN_k = d1_k d2_k \ldots dN_k$ is the short hand notation for the differential elements running over all the coordinates i_k of molecules of species k. The pair correlation function $\rho_{ij}^{(2)}(1,2)$ is defined as

$$\rho_{ij}^{(2)}(1,2) = N_i(N_j - \delta_{ij})\frac{1}{Z_N}\int dN_1 \ldots dN_i^{(-1)} \ldots dN_j^{(-1)} \ldots dN_n \exp\left[-\beta U(N_1, \ldots N_n)\right], \tag{C.29}$$

where the integration is taken over all variables except those that are selected in the correlation function.

The radial distribution function $g_{ij}^{(2)}(1,2)$ is defined as

$$g_{ij}^{(2)}(1,2) = \frac{\rho_{ij}^{(2)}(1,2)}{\rho_i^{(1)}(1)\rho_j^{(1)}(2)}. \tag{C.30}$$

For a homogeneous mixture, one has $\rho_k = N_k / V$ which is the bulk density of species k and the $g_{ij}^{(2)}(1,2)$ is a function of the distance r between the two particles, $g_{ij}^{(2)}(1,2) = g_{ij}(r)$. Integration of $\rho_{ij}^{(2)}(1,2)$ leads to

$$\int d1d2\rho_{ij}^{(2)}(1,2) = \rho_i \rho_j V \int d\vec{r} g_{ij}(r) = N_i (N_j - \delta_{ij}) . \tag{C.31}$$

Using $g_{ij}(r) = h_{ij}(r) + 1$, the condition $\tilde{h}_{ij;MD}(0) = \tilde{h}_{ij}(0)$ and Eq. (C.31), we get

$$\sqrt{\rho_i \rho_j}\, \tilde{h}_{ij;MD}(0) = \sqrt{\rho_i \rho_j} \int d\vec{r} h_{ij}(r) = -\delta_{ij} . \tag{C.32}$$

Combining Eqs. (C.25), (C.27) and (C.32) and using the mole fractions x_i defined as $x_i = \langle N_i \rangle / N$, we obtain the result for the LP correction as

$$\varepsilon_{LP} = \frac{1}{N\rho} \frac{1}{x_i x_j} \left(\frac{\partial \rho_i}{\partial \beta \mu_j}\right)_{TVN_k} . \tag{C.33}$$

Finally, from Eqs. (C.33) and (C.26), we have

$$\lim_{r\to\infty} g_{MD,ij}(r) = 1 - \frac{1}{N x_i x_j} \frac{1}{\rho} \left(\frac{\partial \rho_i}{\partial \beta \mu_j}\right)_{TVN_k} . \tag{C.34}$$

Appendix D

A tentative derivation of the CSL equations from the MOZ equation

The CSL (Chandler-Silbey-Ladanyi) [119] equations are thought to be exact equations. As such, one should be able to derive them from the MOZ equation. This is an attempt of such derivation.

As we know, the pair correlation function in *k*–space $\widetilde{h}(1,2)$ in the laboratory frame can be written as [1, 23]

$$\widetilde{h}(1,2) = \sum_{\substack{m,n,l \\ \mu,\nu}} \widetilde{h}_{\mu\nu}^{mnl}(k)\Phi_{\mu\nu}^{mnl}(\Omega_1,\Omega_2,\hat{k}), \tag{D.1}$$

which can be rewritten in the intermolecular frame (where k is parallel to z) as

$$\widetilde{h}(1,2) = \sum_{\substack{m,n,\chi \\ \mu,\nu}} \widetilde{h}_{\mu\nu}^{mn\chi}(k)R_{\chi\mu}^{m}(\Omega_1)R_{\underline{\chi}\nu}^{n}(\Omega_2), \tag{D.2}$$

where $R_{\chi\mu}^{m}(\Omega_1)$ is a Wigner generalized spherical harmonic [34] and $\underline{\chi} \equiv -\chi$.

The Raleigh expansion is given by [23]

$$\exp(i\vec{k}\vec{L}) = \sum_{l} i^l j_l(kL) \sum_{\lambda} R_{\lambda 0}^{l}(\hat{k})\overline{R}_{\lambda 0}^{l}(\hat{L}). \tag{D.3}$$

Since $R_{\lambda 0}^{l}(\hat{k}) = \delta_{\lambda 0}$ in the intermolecular frame, Eq. (D.3) becomes

$$\exp(i\vec{k}\vec{L}) = \sum_{l} i^l j_l(kL)\overline{R}_{00}^{l}(\hat{L}) \quad . \tag{D.4}$$

The site–site function $h_{\alpha\beta}(r)$ between site α on molecule 1 and site β on molecule 2 can be written as [1]

$$h_{\alpha\beta}(\vec{r}) = \frac{1}{\Omega^2}\iiint d\vec{R}_{12} d\Omega_1 d\Omega_2 h(1,2)\delta\left[\vec{R}_{12} + \vec{L}_{2\beta}(\Omega_2) - \vec{L}_{1\alpha}(\Omega_1) - \vec{r}\right], \tag{D.5}$$

where $\vec{L}_{i\alpha}$ is the vector displacement of site α on molecule i from the molecular centre $\vec{R}_i$, $\vec{L}_{i\alpha} = \vec{r}_{i\alpha} - \vec{R}_i$, for i = 1, 2.

The Fourier transform of Eq. (D.5) is given by

$$\tilde{h}_{\alpha\beta}(\vec{k}) = \frac{1}{\Omega^2}\iint d\Omega_1 d\Omega_2 \tilde{h}(1,2)\exp\left[-i\vec{k}\vec{L}_{2\beta}(\Omega_2) + i\vec{k}\vec{L}_{1\alpha}(\Omega_1)\right]. \tag{D.6}$$

When we insert Eqs. (D.1) and (D.4) into Eq. (D.6), we get

$$\begin{aligned}\tilde{h}_{\alpha\beta}(\vec{k}) = \frac{1}{\Omega^2}\iint d\Omega_1 d\Omega_2 \sum_{\substack{m,n,\chi \\ \mu,\nu}} \tilde{h}^{mn\chi}_{\mu\nu}(k) R^m_{\chi\mu}(\Omega_1) R^n_{\underline{\chi}\nu}(\Omega_2) \sum_{l_1} i^{l_1} j_{l1}(kL_{1\alpha}) \overline{R}^{l_1}_{00}(\hat{L}_{1\alpha}) \\ \times \sum_{l_2} i^{l_2} j_{l_2}(kL_{2\beta}) \overline{R}^{l_2}_{00}(\hat{L}_{2\beta})\end{aligned} . \tag{D.7}$$

Eq. (D.7) involves two angular integrals of type:

$$I_\alpha = \int d\Omega R^m_{\chi\mu}(\Omega) \overline{R}^l_{00}(\hat{L}_\alpha) \tag{D.8}$$

and in order to evaluate this, we have to relate $\hat{L}_\alpha$ to Ω. Typically, all sites correspond to a fixed rotation $\omega_\alpha = (\theta_\alpha, \varphi_\alpha, \beta_\alpha)$ of the molecular axis. If this molecular axis is rotated by angle Ω, we have the general composition rule:

$$\overline{R}^l_{\mu\nu}(\hat{L}_\alpha = \omega_\alpha \Omega) = \sum_{\mu'} R^l_{\mu'\mu}(\omega_\alpha^{-1}) \overline{R}^l_{\mu'\nu}(\Omega) . \tag{D.9}$$

Eq. (D.9) in our particular case assumes the form of

$$\overline{R}^l_{00}(\hat{L}_\alpha = \omega_\alpha \Omega) = \sum_{\mu'} R^l_{\mu'0}(\omega_\alpha^{-1}) \overline{R}^l_{0\mu'}(\Omega) \quad , \tag{D.10}$$

so that the angular integral given by the Eq. (D.8) now becomes:

$$I_\alpha = \sum_{\mu'} R^l_{\mu'0}(\omega_\alpha^{-1}) \int d\Omega R^m_{\chi\mu}(\Omega) \overline{R}^l_{0\mu'}(\Omega) = \frac{8\pi^2}{(2m+1)} R^m_{0\mu}(\omega_\alpha^{-1}), \tag{D.11}$$

where we used the formula [23]

$$\int d\Omega D^{l_2}_{m_2 n_2}(\Omega) D^{l_1}_{m_1 n_1}(\Omega) = \frac{8\pi^2}{(2l_1+1)} \delta_{l_1 l_2} \delta_{m_1 m_2} \delta_{n_1 n_2}. \tag{D.12}$$

We insert Eq. (D.11) into Eq. (D.7) and obtain the following expression:

$$\widetilde{h}_{\alpha\beta}(\hat{k}) = \sum_{\substack{m,n \\ \mu,\nu}} \widetilde{h}^{mn0}_{\mu\nu}(k) i^{m+n} j_m(kL_{1\alpha}) j_n(kL_{2\beta}) R^m_{0\mu}(\omega_\alpha^{-1}) R^n_{0\nu}(\omega_\beta^{-1}), \tag{D.13}$$

where we incorporated the term $1/\Omega^2$ in the terms $\widetilde{h}^{mn0}_{\mu\nu}(k)$. Eq. (D.13) can be written in terms of matrices as

$$\widetilde{H} = J_{1\alpha} \widetilde{H}_0 J_{2\beta}, \tag{D.14}$$

where the matrix $\widetilde{H} = \left[\widetilde{h}_{\alpha\beta}(\hat{k})\right]$, the matrix $J_{1\alpha}$ has for elements $i^m j_m(kL_{1\alpha}) R^m_{0\mu}(\omega_\alpha^{-1})$ and the elements of matrix $\widetilde{H}_0$ are given by the χ–transformation [37] for $\chi = 0$:

$$\widetilde{h}^{mn0}_{\mu\nu}(k) = \sum_l \begin{pmatrix} m & n & l \\ 0 & 0 & 0 \end{pmatrix} h^{mnl}_{\mu\nu}. \tag{D.15}$$

The matrix $\widetilde{H}_0$ is the same matrix from the MOZ matrix equation (B.1):

$$\widetilde{H}_\chi(k) = \widetilde{C}_\chi(k) + \rho \widetilde{H}_\chi(k) \widetilde{C}^*_\chi(k). \tag{D.16}$$

We will start from the CSL equations and try to derive the corresponding site–site equations by using Eq. (D.14). The CSL decomposition assumes that one can write the site–site functions $h_{\alpha\beta}(r)$ as a sum of four terms:

$$h_{\alpha\beta}(r) = \sum_{i=0}^{3} h^{(i)}_{\alpha\beta}(r), \tag{D.17}$$

where each term has a diagrammatic significance [119]. The idea is to write the molecular function given by Eq. (D.2) as a sum of four terms that correspond to those from the CSL equations. We split the molecular function in the following way:

$$h(1,2) = h(\vec{R}_{12}, \Omega_1, \Omega_2) = h_0(\vec{R}_{12}) + h_1(\vec{R}_{12}, \Omega_1) + h_2(\vec{R}_{12}, \Omega_2) + h_d(\vec{R}_{12}, \Omega_1, \Omega_2), \quad \text{(D.18)}$$

with the definitions

$$\begin{aligned} h_0(\vec{R}_{12}) &= h_0^{000}(R_{12}) \\ h_1(\vec{R}_{12}, \Omega_1) &= \sum_{m,\mu,\chi} h_{\mu 0}^{m0\chi}(r) R_{\mu\chi}^{m}(\Omega_1) \\ h_2(\vec{R}_{12}, \Omega_2) &= \sum_{n,\nu,\chi} h_{0\nu}^{0n\chi}(r) R_{\nu\bar{\chi}}^{n}(\Omega_2) \\ h_d(\vec{R}_{12}, \Omega_1, \Omega_2) &= \textit{remaining contributions from } h(1,2). \end{aligned} \quad \text{(D.19)}$$

With this convention, any matrix A in the MOZ equation (D.16) becomes a non–square bloc matrix

$$A = \begin{pmatrix} a_0 & A_1 \\ A_2 & A_d \end{pmatrix}. \quad \text{(D.20)}$$

We can rewrite the MOZ equation as a set of four block matrix equations

$$\begin{aligned} \tilde{h}_0 &= \tilde{c}_0 + \tilde{h}_0 \tilde{c}_0^* + \tilde{H}_1 \tilde{C}_2^* \\ \tilde{H}_1 &= \tilde{C}_1 + \tilde{H}_1 \tilde{c}_0^* + \tilde{H}_d \tilde{C}_1^* \\ \tilde{H}_2 &= \tilde{C}_2 + \tilde{h}_0 \tilde{C}_2^* + \tilde{H}_2 \tilde{C}_d^* \\ \tilde{H}_d &= \tilde{C}_1 + \tilde{H}_1 \tilde{C}_2^* + \tilde{H}_d \tilde{C}_d^* \quad . \end{aligned} \quad \text{(D.21)}$$

In order to go from Eqs. (D.21) to the site–site equations, we have to multiply each of the terms with the vectors $J_{1\alpha}$ and $J_{2\beta}$. For the first term in the first expression of (D.21), we get

$$\tilde{h}_{\alpha\beta}^{0}(k) = J_{1\alpha} \tilde{h}_0 J_{2\beta}, \quad \text{(D.22)}$$

However, for the term $\tilde{h}_0 \tilde{c}_0^*$, we do not get the splitting of the terms that appears in the CSL equations, since we have

$$J_{1\alpha} \tilde{h}_0 \tilde{c}_0^* J_{2\beta} \neq \sum_{\gamma} (J_{1\alpha} \tilde{h}_0 J_{2\gamma})(J_{1\gamma} \tilde{c}_0^* J_{2\beta}). \quad \text{(D.23)}$$

The result of Eq.(D.23) suggests that approximations are needed to go from the MOZ equation to the set of CSL equations.

Bibliography

[1] J.-P. Hansen, and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, Elsevier, Amsterdam, 3rd ed.,2006).

[2] P. Debye, B. Chu, and H. Kaufmann, J. Chem. Phys. **36**, 3378 (1962).

[3] J. G. Kirkwood, and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).

[4] E. Matteoli, and L. Lepori, J. Chem. Phys. **80**, 2856 (1984).

[5] J. B. Ott *et al.*, J. Chem. Thermodyn. **19**, 337 (1987).

[6] J. Polak *et al.*, J. Chem. Eng. Data **15**, 323 (1970).

[7] Y. Koga, *Solution Thermodynamics and its Application to Aqueous Solutions* (Elsevier, Amsterdam, 2007).

[8] J. S. Rowlinson, and F. L. Swinton, *Liquids and Liquid Mixtures* (Butterworth Scientific, London, 1982).

[9] T. Sato, A. Chiba, and R. Nozaki, J. Chem. Phys. **110**, 2508 (1999).

[10] M. Mijaković *et al.*, J. Mol. Liq. **164**, 66 (2011).

[11] K. Nakanishi, N. Kato, and M. Maruyama, J. Phys. Chem. **71**, 814 (1967).

[12] M. A. Lopez-Quintela *et al.*, Curr. Opin. Colloid Interface Sci. **9**, 264 (2004).

[13] C. Caccamo, Phys. Rep. **274**, 1 (1996).

[14] G. P. Morriss, and J. W. Perram, Mol. Phys. **43**, 669 (1981).

[15] J. Richardi *et al.*, Mol. Phys. **93**, 925 (1998).

[16] B. M. Pettitt, and P. J. Rossky, J. Chem. Phys. **77**, 1451 (1982).

[17] R. A. Thuraisingham, and H. L. Friedman, J. Chem. Phys. **78**, 5772 (1983).

[18] T. Ichiye, and A. D. J. Haymet, J. Chem. Phys. **89**, 4315 (1988).

[19] L. Lue, and D. Blankschtein, J. Chem. Phys. **102**, 5427 (1995).

[20] M. Lombardero *et al.*, J. Chem. Phys. **110**, 1148 (1999).

[21] G. Reddy *et al.*, J. Chem. Phys. **119**, 13012 (2003).

[22] J. Richardi, C. Millot, and P. H. Fries, J. Chem. Phys. **110**, 1138 (1999).

[23] C. G. Gray, and K. E. Gubbins, *Theory of Molecular Fluids* (Oxford University Press, New York, 1984).

[24] J. C. Rasaiah, D. N. Card, and J. P. Valleau, J. Chem. Phys. **56**, 248 (1972).

[25] W. R. Smith, D. Henderson, and Y. Tago, The Journal of Chemical Physics **67**, 5308 (1977).

[26] A. Kovalenko, and F. Hirata, J. Chem. Phys. **110**, 10095 (1999).

[27] A. Kovalenko, and F. Hirata, Chem. Phys. Lett. **349**, 496 (2001).

[28] F. Hirata, *Molecular Theory of Solvation* (Kluwer Academic, Dordrecht, 2003).

[29] G. A. Martynov, and G. N. Sarkisov, Mol. Phys. **49**, 1495 (1983).

[30] F. J. Rogers, and D. A. Young, Phys. Rev. A **30**, 999 (1984).

[31] G. Zerah, and J.-P. Hansen, J. Chem. Phys. **84**, 2336 (1986).

[32] W. G. Madden, and S. A. Rice, J. Chem. Phys. **72**, 4208 (1980).

[33] P. Ballone *et al.*, Mol. Phys. **59**, 275 (1986).

[34] A. Messiah, *Quantum Mechanics, Vol. II.* (Whiley, New York, 1962).

[35] L. Blum, and A. J. Torruella, J. Chem. Phys. **56**, 303 (1972).

[36] P. H. Fries, and G. N. Patey, J. Chem. Phys. **82**, 429 (1985).

[37] L. Blum, J. Chem. Phys. **57**, 1862 (1972).

[38] L. Blum, J. Chem. Phys. **58**, 3295 (1973).

[39] P. H. Fries, J. Richardi, and H. Krienke, Mol. Phys. **90**, 841 (1997).

[40] D. Levesque, J. J. Weis, and G. N. Patey, J. Chem. Phys. **72**, 1887 (1980).

[41] D. Chandler, and H. C. Andersen, J. Chem. Phys. **57**, 1930 (1972).

[42] P. T. Cummings, and G. Stell, Mol. Phys. **46**, 383 (1982).

[43] J. L. Lebowitz, and J. K. Percus, Phys.Rev. **144**, 251 (1966).

[44] L. L. Lee, *Molecular Thermodynamics of Nonideal Fluids* (Butterworths, Boston, 1987).

[45] K. Hiroike, J. Phys. Soc. Jpn. **15**, 771 (1960).

[46] T. Morita, and K. Hiroike, Progress of Theoretical Physics **23**, 1003 (1960).

[47] S. J. Singer, and D. Chandler, Mol. Phys. **55**, 621 (1985).

[48] T. L. Hill, *Statistical Mechanics: Principles and Selected Applications*. (McGraw-Hill, New York, 1956).

[49] L. Lue, *Integral Equation Theories for Complex Fluids*, PhD Thesis (MIT, Cambridge, 1994).

[50] L. J. Lowden, and D. Chandler, J. Chem. Phys. **59**, 6587 (1973).

[51] L. J. Lowden, and D. Chandler, J. Chem. Phys. **62**, 4246 (1975).

[52] E. Waisman, and J. L. Lebowitz, J. Chem. Phys. **56**, 3086 (1972).

[53] E. Waisman, and J. L. Lebowitz, J. Chem. Phys. **56**, 3093 (1972).

[54] G. N. Patey, Mol. Phys. **34**, 427 (1977).

[55] P. T. Cummings, and E. R. Smith, Chem.Phys. **42**, 241 (1979).

[56] P. T. Cummings, and G. Stell, J. Chem. Phys. **78**, 1917 (1983).

[57] P. T. Cummings, and P. A. Monson, J. Chem. Phys. **82**, 4303 (1985).

[58] B. A. Malijevsky, and M. Barosova, Mol. Phys. **91**, 65 (1997).

[59] L. L. Lee, and A. Malijevsky, J. Chem. Phys. **114**, 7109 (2001).

[60] P. Attard, and G. N. Patey, J. Chem. Phys. **92**, 4970 (1990).

[61] E. Lomba, C. Martín, and M. Lombardero, Mol. Phys. **77**, 1005 (1992).

[62] A. Perera, P. G. Kusalik, and G. N. Patey, J. Chem. Phys. **87**, 1295 (1987).

[63] A. Perera, and G. N. Patey, J. Chem. Phys. **89**, 5861 (1988).

[64] F. Lado, Mol. Phys. **47**, 283 (1982).

[65] E. Lomba, M. Lombardero, and J. L. F. Abascal, J. Chem. Phys. **91**, 2581 (1989).

[66] C. Martin *et al.*, J. Chem. Phys. **100**, 1599 (1994).

[67] F. Lado, E. Lomba, and M. Lombardero, J. Chem. Phys. **103**, 481 (1995).

[68] J. Richardi, H. Krienke, and P. H. Fries, Chem. Phys. Lett. **273**, 115 (1997).

[69] G. P. Morriss, J. W. Perram, and E. R. Smith, Mol. Phys. **38**, 465 (1979).

[70] G. P. Morriss, and J. W. Perram, Mol. Phys. **41**, 1463 (1980).

[71] P. T. Cummings, G. P. Morriss, and C. C. Wright, Mol. Phys. **43**, 1299 (1981).

[72] G. P. Morriss, and P. T. Cummings, Mol. Phys. **49**, 1103 (1983).

[73] P. T. Cummings, G. P. Morriss, and G. Stell, Mol. Phys. **51**, 289 (1984).

[74] G. P. Morriss, and D. J. Isbister, Mol. Phys. **52**, 57 (1984).

[75] G. P. Morriss, and D. J. Isbister, Mol. Phys. **59**, 911 (1986).

[76] F. Hirata, B. M. Pettitt, and P. J. Rossky, J. Chem. Phys. **77**, 509 (1982).

[77] B. M. Pettitt, and P. J. Rossky, J. Chem. Phys. **78**, 7296 (1983).

[78] B. Borštnik, and D. Janežič, Chem.Phys. **116**, 187 (1987).

[79] C. S. Hsu, D. Chandler, and L. J. Lowden, Chem.Phys. **14**, 213 (1976).

[80] E. Johnson, and R. P. Hazoume, J. Chem. Phys. **70**, 1599 (1979).

[81] J. G. Curro, and K. S. Schweizer, J. Chem. Phys. **87**, 1842 (1987).

[82] K. S. Schweizer, and J. G. Curro, Phys. Rev. Lett. **58**, 246 (1987).

[83] G. Hummer, and D. M. Soumpasis, Mol. Phys. **75**, 633 (1992).

[84] G. Hummer, D. M. Soumpasis, and M. Neumann, Mol. Phys. **77**, 769 (1992).

[85] T. Ichiye, and D. Chandler, J. Phys. Chem. **92**, 5257 (1988).

[86] H.-A. Yu, B. Roux, and M. Karplus, J. Chem. Phys. **92**, 5020 (1990).

[87] H. Tanaka, J. Walsh, and K. E. Gubbins, Mol. Phys. **76**, 1221 (1992).

[88] K. Yoshida *et al.*, J. Phys. Chem. B **106**, 5042 (2002).

[89] J. S. Perkyns, K. M. Dyer, and B. M. Pettitt, J. Chem. Phys. **116**, 9404 (2002).

[90] Q. Du, D. Beglov, and B. Roux, J. Phys. Chem. B **104**, 796 (2000).

[91] J. Kolafa, S. Labik, and A. Malijevsky, Mol. Phys. **100**, 2629 (2002).

[92] DL_POLY simulation package, CCLRC Daresbury Laboratory, see www.ccp5.ac.uk/DL_POLY/.

[93] W. L. Jorgensen, J. M. Briggs, and M. L. Contreras, J. Phys. Chem. **94**, 1683 (1990).

[94] E. M. Duffy, D. L. Severance, and W. L. Jorgensen, J. Am. Chem. Soc. **114**, 7535 (1992).

[95] J. M. M. Cordeiro, Int. J. Quantum Chem. **65**, 709 (1997).

[96] W. L. Jorgensen, J. Phys. Chem. **90**, 1276 (1986).

[97] J. C. Berendsen *et al.*, *Intermolecular Forces* (Reidel, Dortrecht, 1981).

[98] K.-C. Ng, J. Chem. Phys. **61**, 2680 (1974).

[99] W. H. Press *et al.*, *Numerical Recipes in FORTRAN: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 2nd ed., 1992).

[100] J. L. Lebowitz, and J. K. Percus, Phys.Rev. **122**, 1675 (1961).

[101] A. Perera *et al.*, J. Mol. Liq. **159**, 52 (2011).

[102] B. Kežić, and A. Perera, J. Chem. Phys. **135**, 234104 (2011).

[103] D. R. Line, *CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, 75th ed.,1994.).

[104] L. J. Lowden, and D. Chandler, J. Chem. Phys. **61**, 5228 (1974).

[105] G. Munao *et al.*, J. Chem. Phys. **132**, 084506 (2010).

[106] L. Verlet, Mol. Phys. **41**, 183 (1980).

[107] L. Zoranić *et al.*, J. Phys. Chem. C **111**, 15586 (2007).

[108] J. H. Riddick, W. B. Bunger, and T. K. Sakano, *Organic Solvents* (Wiley, New York, 4th ed., 1986).

[109] A. Perera, F. Sokolić, and L. Zoranić, Phys. Rev. E **75**, 060502 (2007).

[110] K. Yoshida *et al.*, J. Phys. Chem. B **106**, 5042 (2002).

[111] S. V. Buldyrev *et al.*, Proc. Natl. Acad. Sci. U.S.A. **104**, 20177 (2007).

[112] V. G. Dashevsky, and G. N. Sarkisov, Mol. Phys. **27**, 1271 (1974).

[113] W. L. Jorgensen *et al.*, J. Chem. Phys. **79**, 926 (1983).

[114] M. Prevost *et al.*, Mol. Phys. **71**, 587 (1990).

[115] M. Mezei, S. Swaminathan, and D. L. Beveridge, J. Am. Chem. Soc. **100**, 3255 (1978).

[116] A. Ben-Naim, and Y. Marcus, J. Chem. Phys. **81**, 2016 (1984).

[117] G. S. Kell, J. Chem. Eng. Data **20**, 97 (1975).

[118] K. A. Motakabbir, and M. Berkowitz, J. Phys. Chem. **94**, 8359 (1990).

[119] D. Chandler, R. Silbey, and B. M. Ladanyi, Mol. Phys. **46**, 1335 (1982).

[120] P. J. Rossky, and R. A. Chiles, Mol. Phys. **51**, 661 (1984).

[121] B. Kežić, and A. Perera, J. Chem. Phys. **137**, 014501 (2012).

[122] A. Perera, and B. Kežić, Faraday Discuss., DOI: 10.1039/C3FD00072A, (2013).

[123] M. I. Guerrero, G. Saville, and J. S. Rowlinson, Mol. Phys. **29**, 1941 (1975).

[124] L. Belloni, J. Chem. Phys. **98**, 8080 (1993).

[125] A. Perera *et al.*, in *Molecular Dynamics-Studies of Synthetic and Biological Molecules*

ed. L. Wang (InTech, Rijeka, 2012).

[126] K. Egashira, and N. Nishi, J. Phys. Chem. B **102**, 4054 (1998).

[127] G. D'Arrigo, and A. Paparelli, J. Chem. Phys. **88**, 405 (1988).

[128] G. D'Arrigo, and A. Paparelli, J. Chem. Phys. **88**, 7687 (1988).

[129] O. Conde, J. Teixeira, and P. Papon, J. Chem. Phys. **76**, 3747 (1982).

[130] S. Mashimo, T. Umehara, and H. Redlin, J. Chem. Phys. **95**, 6257 (1991).

[131] E. J. W. Wensink *et al.*, J. Chem. Phys. **119**, 7308 (2003).

[132] S. Y. Noskov, G. Lamoureux, and B. Roux, J. Phys. Chem. B **109**, 6705 (2005).

[133] L. Zhang *et al.*, J. Chem. Phys. **125**, 104502 (2006).

[134] S. Alavi *et al.*, J. Chem. Phys. **133**, 074505 (2010).

[135] A. Cipiciani, G. Onori, and G. Savelli, Chem. Phys. Lett. **143**, 505 (1988).

[136] L. Dougan *et al.*, J. Chem. Phys. **121**, 6456 (2004).

[137] A. K. Soper, and J. L. Finney, Phys. Rev. Lett. **72**, 948 (1994).

[138] S. Dixit *et al.*, Nat. **416**, 829 (2002).

[139] J. N. Murrell, and A. D. Jenkins, *Properties of Liquids and Solutions, 2nd edn.* (John Whiley & Sons, Chichester, 1994).

[140] F. Franks, and D. J. G. Ives, Q. Rev. Chem. Soc. **20**, 1 (1966).

[141] L. Zoranić, F. Sokolić, and A. Perera, J. Chem. Phys. **127**, 024502 (2007).

[142] H. E. Stanley *et al.*, J. Phys. Condens. Matter **12**, A403 (2000).

[143] J. D. Eaves *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 13019 (2005).

[144] K. Stokely *et al.*, Proc. Natl. Acad. Sci. U.S.A. **107**, 1301 (2010).

[145] C. Zhang, and X. Yang, Fluid Phase Equilib. **231**, 1 (2005).

[146] E. W. Washbrun, *International Critical Tables of Numerical Data, Physics, Chemistry and Technology* (Knovel, New York, 2003).

[147] J. P. E. Grolier, and E. Wilhelm, Fluid Phase Equilib. **6**, 283 (1981).

[148] G. Benson, P. D'Arcy, and O. Kiyohara, J. Solution Chem. **9**, 931 (1980).

[149] R. Nagarajan, and E. Ruckenstein, Langmuir **16**, 6400 (2000).

[150] S. Komura, J. Phys. Condens. Matter **19**, 463101 (2007).

[151] B. Widom, J. Chem. Phys. **84**, 6943 (1986).

[152] M. Schick, and W.-H. Shih, Phys. Rev. Lett. **59**, 1205 (1987).

[153] C. Kan *et al.*, J. Phys. C: Solid State Phys. **20**, L361 (1987).

[154] T. P. Stockfisch, and J. C. Wheeler, J. Phys. Chem. **92**, 3292 (1988).

[155] P. M. Chaikin, and T. C. Lubensky, *Principles of Consensed Matter Physics* (Cambridge University Press, Cambridge, 1995).

[156] M. Teubner, and R. Strey, J. Chem. Phys. **87**, 3195 (1987).

[157] M. E. Fisher, J. Math. Phys. **5**, 944 (1964).

[158] E. A. Ploetz, N. Bentenitis, and P. E. Smith, Fluid Phase Equilib. **290**, 43 (2010).

[159] S. Weerasinghe, and P. E. Smith, J. Chem. Phys. **118**, 10663 (2003).

[160] M. E. Lee, and N. F. A. van der Vegt, J. Chem. Phys. **122**, 114509 (2005).

[161] D. T. Bowron, J. L. Finney, and A. K. Soper, J. Phys. Chem. B **102**, 3551 (1998).

[162] D. T. Bowron, A. K. Soper, and J. L. Finney, J. Chem. Phys. **114**, 6203 (2001).

[163] Y. Koga, W. W. Y. Siu, and T. Y. H. Wong, J. Phys. Chem. **94**, 7700 (1990).

[164] J. Finney, L., D. Bowron, T., and A. Soper, K. , J. Phys. Condens. Matter **12**, A123 (2000).

[165] K. Nishikawa, Y. Kodera, and T. Iijima, J. Phys. Chem. **91**, 3694 (1987).

[166] K. Nakanishi *et al.*, J. Chem. Phys. **80**, 1656 (1984).

[167] P. G. Kusalik *et al.*, J. Phys. Chem. B **104**, 9533 (2000).

[168] D. T. Bowron, and S. Diaz Moreno, J. Phys. Condens. Matter **15**, S121 (2003).

[169] M. Ferrario *et al.*, J. Chem. Phys. **93**, 5156 (1990).

[170] A. Perera, and F. Sokolić, J. Chem. Phys. **121**, 11272 (2004).

[171] G. Kamath, G. Georgiev, and J. J. Potoff, J. Phys. Chem. B **109**, 19463 (2005).

[172] A. Perera *et al.*, J. Chem. Phys. **124**, 124515 (2006).

[173] B. Löwen, and S. Schulz, Thermochim. Acta **262**, 69 (1995).

[174] Landolt-Börnstein, *New Series, Group IV, Vol.1* (Springer-Verlag, New York, 1977).

[175] A. Perera *et al.*, J. Chem. Phys. **123**, 024503 (2005).

[176] H. D. Bale, R. E. Shepler, and D. K. Sorgen, Phys. Chem. Liq. **1**, 181 (1968).

[177] P. Ball, Nature **452**, 291 (2008).

[178] F. Franks, *Water: A Comprehensive Treatise, Vol.1* (Plenum Press, New York, 1975).

[179] C. A. Angell *et al.*, Phys. Chem. Chem. Phys. **2**, 1559 (2000).

[180] D. L. Bergman, and R. M. Lynden-Bell, Mol. Phys. **99**, 1011 (2001).

[181] R. M. Lynden-Bell, and P. G. Debenedetti, J. Phys. Chem. B **109**, 6527 (2004).

[182] R. M. Lynden-Bell, J. Phys. Condens. Matter **22**, 284107 (2010).

[183] A. Perera, R. Mazighi, and B. Kežić, J. Chem. Phys. **136**, 174516 (2012).

[184] B. Kežić, R. Mazighi, and A. Perera, Physica A **392**, 567 (2013).

[185] R. Gupta, and G. N. Patey, J. Chem. Phys. **137**, 034509 (2012).

[186] R. Gupta, and G. N. Patey, J. Phys. Chem. B **115**, 15323 (2011).

[187] M. P. Allen, and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, New York, 1987).

[188] I. A. Borin, and M. S. Skaf, J. Chem. Phys. **110**, 6412 (1999).

[189] A. Ciach, and W. T. Gozdz, Annu. Rep. R. Soc. Chem. Sect. C. Phys. Chem. **97**, 269 (2001).

[190] P. C. Hemmer, and G. Stell, Phys. Rev. Lett. **24**, 1284 (1970).

[191] G. Stell, and P. C. Hemmer, J. Chem. Phys. **56**, 4274 (1972).

[192] M. R. Sadr-Lahijany *et al.*, Phys. Rev. E **60**, 6714 (1999).

[193] J. M. Kincaid, G. Stell, and C. K. Hall, J. Chem. Phys. **65**, 2161 (1976).

[194] A. Scala *et al.*, Phys. Rev. E **63**, 041202 (2001).

[195] G. Malescio, and G. Pellicane, Nature Mater. **2**, 1476 (2003).

[196] P. J. Camp, Phys. Rev. E **68**, 061506 (2003).

[197] P. J. Camp, Phys. Rev. E **71**, 031507 (2005).

[198] P. J. Camp, J. Mol. Liq. **127**, 10 (2006).

[199] J. M. Kincaid, G. Stell, and E. Goldmark, J. Chem. Phys. **65**, 2172 (1976).

[200] J. M. Kincaid, and G. Stell, J. Chem. Phys. **67**, 420 (1977).

[201] E. Lomba *et al.*, J. Chem. Phys. **126**, 244510 (2007).

[202] N. Choudhury, and S. K. Ghosh, Phys. Rev. E **66**, 021206 (2002).

[203] E. A. Jagla, Phys. Rev. E **58**, 1478 (1998).

[204] E. A. Jagla, J. Chem. Phys. **111**, 8980 (1999).

[205] A. Perera, Mol. Phys. **107**, 487 (2009).

[206] A. B. Bhatia, and D. E. Thornton, Phys. Rev. B **2**, 3004 (1970).

[207] G. Malescio, J. Chem. Phys. **95**, 1198 (1991).

[208] X.-G. Wen, *Quantum Field Theory of Many-body Systems* (Oxford University Press, New York, 2007).

[209] X.-G. Wen, Phys. Rev. B **65**, 165113 (2002).

[210] M. A. Levin, and X.-G. Wen, Phys. Rev. B **71**, 045110 (2005).

[211] K. W. Mittal, *Micellization, Solubilization, and Micro-emulsions* (Plenum, New York, 1977).

[212] J. Groenewold, and W. K. Kegel, J. Phys. Chem. B **105**, 11702 (2001).

[213] M. Broccio *et al.*, J. Chem. Phys. **124**, 084501 (2006).

[214] A. J. Archer *et al.*, J. Phys. Condens. Matter **20**, 415106 (2008).

[215] A. J. Archer *et al.*, J. Chem. Phys. **126**, 014104 (2007).

[216] A. Einstein, Sci. **91**, 487 (1940).

[217] S. J. Marrink *et al.*, J. Am. Chem. Soc. **123**, 8638 (2001).

[218] G. D'Arrigo, R. Giordano, and J. Teixeira, Eur. Phys. J. E **10**, 135 (2003).

[219] G. D'Arrigo, R. Giordano, and J. Teixeira, Eur. Phys. J. E **29**, 37 (2009).