# A molecular dynamics study of micro-heterogeneities in aqueous alcohol solutions

Larisa Zoranic

# THESE DE DOCTORAT DE
# L'UNIVERSITE PIERRE ET MARIE CURIE
# L'UNIVERSITE DE ZAGREB

Spécialité

**PHYSIQUE THEORIQUE des LIQUIDES**

Présentée par Larisa ZORANIĆ

Pour obtenir le grade de

**DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE**

# ETUDE PAR DYNAMIQUE MOLECULAIRE DE LA MICRO-HETEROGENEITE DANS LES MELANGES EAU-ALCOOLS

soutenue le *18 Decembre 2008* devant le jury composé de :

| | | |
|---|---|---|
| Mr. | Aurélien PERERA | Directeur de thèse |
| Mr. | Srećko KILIĆ | Directeur de thèse |
| Mr. | Vladimir PAAR | Presidente du Jury |
| Mr. | Redha MAZIGHI | Rapporteur |
| Mr. | Franjo SOKOLIĆ | Rapporteur |

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO-MATEMATIČKI FAKULTET
POSLIJEDIPLOMSKI STUDIJ PRIRODNIH ZNANOSTI
FIZIKA

Larisa Zoranić

# Ispitivanje mikroheterogenosti u vodenim otopinama alkohola metodom molekularne dinamike
# A Molecular Dynamics study of micro-heterogeneities in aqueous alcohol solutions

Doktorski rad
predložen Fizičkom odsjeku
Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu
radi stjecanja akademskog stupnja
doktora prirodnih znanosti (fizika)

Doctoral Thesis
submitted to the Department of Physics
Faculty of Science, University of Zagreb
for the academic degree of
Doctor of Natural Sciences (Physics)

Zagreb, 2008.

*To my mother and father*

*Mojim roditeljima*

*À ma mère et mon père*

# RESUME

Thèse doctorat
Université Pierre et Marie Curie
Laboratoire de Physique Théorique de la Matière Condensée

**Étude par Dynamique Moléculaire de la micro-hétérogénéité dans les mélanges eau-alcools**

Larisa Zoranić
Faculté des Sciences, Université de Split, Nikole Tesle 12, Split, Croatie
Laboratoire de Physique Théorique de la Matière Condensée, Université Pierre et Marie Curie, Tour 24, Boîte 121, 4, Place Jussieu, 75252 Paris Cedex 05, France

Ce travail concerne l'analyse structurale des liquides associés comme l'eau et les alcools et leur mélanges. Nous sondons par la Dynamique Moléculaire la micro-structure de ces liquides. A partir du calcul des fonctions de corrélations et des facteurs de structures associés, ainsi que d'un parametre d'ordre effectifs, nous proposons une vision consistente de la microhétérogénéité au sein des liquides macroscopiquement homogénes. Cette analyse nous permet de distinguer entre la micro-structure dans les liquides purs associés, de celle de micro-ségrégation dans leurs mélanges binaires, alors que le mécanisme commun est bien la liaison hydrogéne. Celle-ci structure différamment les sites partiels concernés (hydrogéne et oxygéne), tandis que les sites inertes méthyls sont purement désordonnés. Ainsi, la micro-hétérogénéité apparait comme une propriéte universelle des mélanges de liquides associés. Ce type d'ordre local n'appartient pas tout-é-fait la classe du désordre, pas plus qu'é celle de l'ordre global. Il apparait donc comme une nouvelle forme d'ordre et défie nos méthodes pour le mettre en évidence, tant du point de vue expérimental que théorique.

(180 pages, 51 figures, 3 tables, 106 références, l'original dans l'anglais)

**Thèse déposée dans** National and University Library, Hrvatske bratske zajednice 4, 10000 Zagreb; Les bibliothèques de l'Université Pierre et Marie Curie, 4, Place Jussieu-75252 Paris, France

**Mots clé:** (Dynamique Moléculaire, micro-hétérogénéité, liquides associés )

Directeur de thèse: Prof. dr. sc. Srećko Kilić
Dr. sc. Aurélien Perera
Rapporteur: Prof. dr. sc. Vladimir Paar
Prof. dr. sc. Stjepan Marčelja
Dr. sc. Sanja Tomić
Prof. dr. sc. Franjo Sokolić
Dr. sc. Redha Mazighi

Thèse acceptée : le 9 décembre 2008.

## BASIC DOCUMENTATION CARD

Doctoral Thesis
University of Zagreb
Faculty of Science
Department of Physics

# A Molecular Dynamics study of micro-heterogeneities in aqueous alcohol solutions

Larisa Zoranić
Faculty of Science, University of Split, Nikole Tesle 12, Split, Hrvatska
La Laboratoire de Physique Théorique de la Matière Condensée, Université Pierre et Marie Curie, Tour 24, Boîte 121, 4, Place Jussieu, 75252 Paris Cedex 05, France

In this thesis we focus on the structural refinements present in the associated liquids. Using the molecular dynamics simulations we probe the microstructure of these liquids, namely the neat systems and alcohol-water mixtures. From the correlation functions in direct and reciprocal space and as well as the cluster distribution functions we create a consistent picture of the structural specificity that exist in the associated liquids. We made a distinction between the structural organization in pure liquids and in binary systems, first defined through the local self-association of molecules, and second corresponding to a local segregation of the components of the binary system. In each case the highly anisotropic site-site hydrogen bonding has main role. In the pure liquids the association is induced through specific site-interactions that lead to an inhomogeneous distribution ofsites, while the molecular distribution preserves the liquid-like behavior. In the binary system the micro-heterogeneity is due to the local immiscibility of the species. Micro-heterogeneity appears to be universal feature of aqueous solutions, predominantly driven by hydrogen bonding. Micro-heterogeneous order does not quite fit in the class of disorder, but neither in that of order. It appears like a new form or order, and challenges our techniques to put this in evidence, both experimentally or through simulations.

(180 pages, 51 figures, 3 tables, 106 references, original in English)

**Thesis deposited in** National and University Library, Hrvatske bratske zajednice 4, 10000 Zagreb; Les bibliothèques de l'Université Pierre et Marie Curie, 4, Place Jussieu-75252 Paris, France

**Keywords: (Molecular dynamics, microheterogeneity, associated liquids)**

Supervisor: Prof. dr. sc. Srećko Kilić
Dr. sc. Aurélien Perera
Reviewers: Prof. dr. sc. Vladimir Paar
Prof. dr. sc. Stjepan Marčelja
Dr. sc. Sanja Tomić
Prof. dr. sc. Franjo Sokolić
Dr. sc. Redha Mazighi

Thesis accepted: 9th December 2008.

# TEMELJNA DOKUMENTACIJSKA KARTICA

---

Doktorska disertacija
Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Fizički odsjek

## Ispitivanje mikroheterogenosti u vodenim otopinama alkohola metodom molekularne dinamike

Larisa Zoranić
Prirodoslovno-matematički fakultet, Sveučilište u Splitu, NikoleTesle12, Split, Hrvatska
Laboratorij za teorijsku fiziku čvrste tvari, Sveučilište Pierre i Marie Curie, Tour 24, Boîte 121, 4, Place Jussieu, 75252 Paris Cedex 05, France

Osnovna tema doktorata jest ispitivanje strukturne organizacije u vodenim otopinama. Naime, na nano-skali ovi sustavi pokazuju lokalnu uređenost koja ne narušava globalnu homogenost. Korištenjem metode molekularne dinamike ispitivali smo mikrostrukturu u tekućinama i otopinama alkohola i vode. Analizom korelacijskih funkcija u recipročnom i direktnom prostoru te distribucijske funkcije klastera odredili smo posebitosti njihove strukturne organizacije. U tekućinama, djelovanjem vodikove veze molekule se povezuju u klastere ili veće asocijacije kao što je npr. struktura mreže. Ovakva strukturna organizacija vodi do nehomogene distribucije atoma koji sudjeluju u gradnji vodikove veze, dok molekularna distribucijska funkcija, kao i funkcije hidrofobnih atoma, zadržava svojstva homogene tekućine. U otopinama, uzrok lokalne heterogenost je također vodikova veza, naime mikroheterogenost se odnosi na lokalno ne-mješanje komponenata otopina. Mikroheterogenost je opće svojstvo vodenih otopina. Ovakva uređenost sustava na lokalnom nivou ne spada u klasu reda ni nereda. Tako da bolja definicija ovog novog svojstva predstavlja izazov eksperimentalnim i teorijskim metodama i tehnikama.

(180 stranica, 51 slika, 3 tablica, 106 literaturnih navoda, jezik izvornika engleski )

**Rad je pohranjen** u Nacionalnoj i sveučilišnoj knjižnici, Hrvatske bratske zajednice 4, 10000 Zagreb; Les bibliothèques de l'Université Pierre et Marie Curie, 4, Place Jussieu-75252 Paris, France

**Ključne riječi (molekularna dinamika, mikroheterogenost, vodene otopine )**

Mentor: Prof. dr. sc. Srećko Kilić, redoviti profesor, PMF, Split
Dr. sc. Aurélien Perera
Ocjenjivači: Prof. dr. dc. Vladimir Paar, akademik, PMF, Zagreb
Prof. dr. sc. Stjepan Marčelja, redoviti profesor, PMF, Split
Dr. sc. Sanja Tomić, znanstvena savjetnica, IRB, Zagreb
Prof. dr. sc. Franjo Sokolić, izvaredni profesor, PMF, Split
Dr. sc. Redha Mazighi

Rad prihvaćen: sjednica Vijeća Fizičkog odsjeka i Vijeća Geofizičkog odsjeka održana 9. prosinca 2008.

---

# Table of Contents

# Acknowledgements

When I first came to Paris, the idea was to implement 9-month project. Gradually, what was 9-month long fellowship, grew into joint PhD thesis. I would like to deeply thank all those people who, during past several years, shared their time and their love with me.

To my supervisor, Aurélien Perera, I would like to express my sincere gratitude for the opportunity to work with him and for the knowledge and experience he shared with me. His energy and enthusiasm were motivating and pushed things strongly forward. His friendship and help in big or small issues were encouraging.

My thanks goes also to my supervisor from Croatia Prof. Srećko Kilić for his substantial support throughout this work.

I also wish to acknowledge the contributions, advice and suggestions of Prof. Franjo Sokolić. His constant help and guidance has been of great value to me.

I would like to express my thanks to Redha Mazighi for constructive discussions which clarified my thinking on this and other matters.

I am grateful to the Prof. Stjepan Marčelja for the care with which he read this thesis and for his correction and suggestions which made this work better.

I would like to express my thanks to all people in the Laboratoire de Physique Théorique de la Matieére Condensée. Special thanks goes to the secretaries Sylviane, Sylvie and Martine for their kindness and support.

My thanks go to the director of LPTMC Bertrand Guillot, for making me feel welcome in this lab.

To write thanks to all the persons who touched my life in these past three years is a

really difficult task. It is impossible to speak of each and every one of you, but you all know how much you are important for me. And to my family, my gratitude goes beyond words.

This research was supported by the following grants: 9-months long fellowship for foreign students supported by French Government; one-year fellowship of The National Foundation for Science, Higher Education and Technological Development of the Republic of Croatia; and financial aid from Croatian Ministry of Science, Education and Sport for the PhD students that study abroad.

# Remerciements

Intitialement j'tais arrive a Paris avec un contrat de recherche de 9 mois. Mais comme cette recherche s'est rvle passionnante, j'ai finalement dcid de transformer ce contrat en une these en cotutelle. Je voudrais remercier ici, et de tout mon coeur, toutes les personnes qui, ces dernieres annes, ont partag leurs temps et leur amiti avec moi.
A mon directeur de these, Aurlien Perera, j'aimerais exprimer toute ma gratitude de m'avoir permis de travailler avec lui et d'avoir partag avec moi son savoir et son exprience. Son nergie et son enthousiasme ont toujours t tres motivants pour faire avancer mes travaux. Son amiti et son soutien m'ont constamment encourage, aussi bien devant les petites que les grandes difficults.
Je voudrais galement remercier mon directeur de these en Croatie, Monsieur Srećko Kilić, pour son soutien considerable tout au long de ma carriere.
Un grand merci a Monsieur Franjo Sokolić, dont les contributions, conseils et suggestions ont toujours t un grand soutien pour moi.
Je voudrais galement remercier Monsieur Redha Mazghi pour nos discussions constructives, qui m'ont aide a mettre de l'ordre dans mes rflexions, dans ma recherche et de aussi de maniere gnrale.
Je suis reconnaissante envers Monsieur Stjepan Marčelja, pour l'attention avec laquelle il lu ma these, pour ses corrections et ses suggestions qui l'ont considerablement amliore.
J'adresse mes remerciement a tous les membres du Laboratoire de Physique Thorique de la Matiere Condense. Je tiens a remercier tout spcialement les secrtaires Sylviane,

Sylvie et Martine de leur gentillesse et leur soutien.
Merci au directeur du LPTMC, Monsieur Bertrand Guillot, qui m'a chaleureusement accueillie dans son laboratoire.
Parler de toutes ces personnes qui ont touch ma vie ces dernieres annes n'est pas une tache facile. Elles se reconnaitront, et je leur dis : " Vous etes tous tres importants pour moi ! "
Pour remercier ma famille, tous les mots ne suffiraient pas.
Mes recherches ont t finances par les organismes suivants: -la bourse de 9 mois pour les tudiants trangers a t offerte par l'Ambassade de France en Croatie, -la bourse d'un an a t dcerne par la fondation Nationale pour la Science, l'Education Suprieure et le Dveloppement Technique de la Rpublique de Croatie -le soutien financier propos par le Ministere de la Science, de l'Education et du Sport croate destin aux tudiants croates qui prpare leur these a l' tranger.

# Résumé


**ETUDE PAR DYNAMIQUE MOLECULAIRE DE LA MICRO-HETEROGENEITE DANS LES MELANGES EAU-ALCOOLS**

**1. Introduction**

L'eau et les alcools sont des liquides associs caractrisés par l'existence de la liaison hydrogne, qui favorise l'apparition de structures agrégées par l'intermédiaire de cette interaction hautement directionnelle. L'étude par simulation en Dynamique Moléculaire de ces liquides purs et de leurs mélanges permet de confirmer l'existence de telles structures et de déterminer le comportement structurel et thermodynamique de ces liquides. L'homogénéité globale de ces systmes implique que le paramètre d'ordre qui caractérise l'apparition de l'ordre local est la fonction de corrélation de paire. Dans un premier temps, nous étudions les liquides purs au travers de divers modèles de champs de force. L'analyse des fonctions de distributions atomiques, ainsi que les facteurs de structure associés, montre une différence notable entre la micro-structure de l'eau et celle des alcools simples. Ces derniers forment des phases agrégats de topologie très riche, alors que la signature de l'ordre dans l'eau se révèle étonnamment pauvre en comparaison. Nous associons ce dernier comportement à une structuration associative plus étendue, sans doute de type réseau. Ensuite, nous analysons la structure des mélanges, notamment par les intégrales de Kirwood-Buff (KBI) et leur dépendance en concentration d'alcool. Les mélanges font apparatre une

forme différente de micro-structuration, de type micro-ségrégation. Cette structure se traduit à la fois par l'apparition d'un pré-pic dans le facteur de structure, ainsi que par des fortes fluctuations des KBI, qui se font à l'échelle du nanomètre et de la nanoseconde, c'est-àdire àun ordre de magnitude supérieure de celui des échelles moléculaires. Cette étude montre la richesse surprenante de la micro-hétérogénéité au sein même de l'homogénéité et du désordre apparent de ces états liquides particuliers. Ce travail concerne l'analyse structurale des liquides associés comme l'eau et les alcools et leur mélanges. Nous sondons par la Dynamique Moléculaire la micro-structure de ces liquides. A partir du calcul des fonctions de corrélations et des facteurs de structures associés, ainsi que d'un paramètre d'ordre effectifs, nous proposons une vision consistante de la micro-hétérogénéité au sein des liquides macroscopiquement homogènes. Cette analyse nous permet de distinguer entre la micro-structure dans les liquides purs associés, de celle de micro-ségrégation dans leurs mélanges binaires, alors que le mécanisme commun est bien la liaison hydrogène. Celle-ci structure différemment les sites partiels concernés (hydrogène et oxygène), tandis que les sites inertes méthyles sont purement désordonnés. Ainsi, la micro-hétérogénéité apparat comme une propriété universelle des mélanges de liquides associés. Ce type d'ordre local n'appartient pas tout-à-fait à la classe du désordre, pas plus qu'à celle de l'ordre global. Il apparat donc comme une nouvelle forme d'ordre et défie nos méthodes pour le mettre en évidence, tant du point de vue expérimental que théorique.

## 2. La physique statistique

A la température ambiante les mélanges d'alcool et d'eau sont dans l'état liquide. D'habitude, ces mélanges sont modélises comme les liquides classiques. L'approche théorique est alors développée en utilisant la physique statistique classique et les liens entre la mécanique statistique et la thermodynamique. La physique statistique fournit un cadre pour associer les propriétés microscopiques des molécules à celles macroscopique du liquide. Ainsi, elle explique la thermodynamique comme le résultat naturel de l'approche statistique (classique et quantique) au niveau microscopique.

L'état dynamique d'un liquide monoatomique un composant à chaque instant est défini par les 3N coordonné es $\mathbf{r}^N = \{\mathbf{r}_1, ..., \mathbf{r}_N\}$ et 3N moments des particules $\mathbf{p}^N = \{\mathbf{p}_1, ..., \mathbf{p}_N\}$. Ces valeurs définissent un point de phase dans un espase de phase de 6N dimensions. Un point de phase correspond à celui de l'état d'un système. Etant donné les coordonné es et les moments des particules à quelque instant, leurs valeurs plus tard (ou plus tt) dans le temps, peuvent être calculé par l'équations de mouvements de Newton , lesquelles, en absence d'un champ externe, sont de la forme:

$$m\mathbf{r}_i = -\nabla_i V_N\left(\mathbf{r}^N\right) \tag{1}$$

Dans la mécanique statistique toutes les propriétés observables d'un système sont calculées comme les moyennes sur les trajectoires de phase (la méthode de Boltzmann) ou comme les moyennes sur un ensemble des systèmes (la méthode de Gibbs). Dans la formulation de Gibbs la distribution de les points de phase dans un ensemble sont décrits par la densité de la probabilité $f^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N; t\right)$:

$$\int\int f^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N; t\right) d\mathbf{r}^N d\mathbf{p}^N = 1, for\ all\ t. \tag{2}$$

La densité de la probabilité d'équilibre est définie par les paramètres macroscopiques qui caractérisent l'ensemble. Par exemple, l'ensemble canonique décrit les systèmes qui ont le volume V, le nombre de particules N et température T constants, pour lequel la densité de la probabilité est donnée par:

$$f_o^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N\right) = \frac{1}{h^{3N}N!}\frac{e^{-\beta H}}{Q_N} \tag{3}$$

oú h est la constante de Planck, $\beta = 1/(k_B T)$ oú $k_B = 1.3810^{-23} J/K$ est la constante de Boltzmann et la constante de la normalisation $Q_N$ est la fonction de la partition canonique, définie comme:

$$Q_N = \frac{1}{h^{3N}N!}\int\int e^{-\beta H} d\mathbf{r}^N, d\mathbf{p}^N \tag{4}$$

Les quantités les plus importantes dans notre analyse sont la densité de paire et la fonction de la distribution de paire, aussi bien que le facteur de structure qui

correspond à la transformation de Fourier de la fonction de distribution de paire. Elles peuvent être calculés en utilisant les données de simulation et aussi par les expériences de diffraction. Ces fonctions contiennent les renseignements sur des corrélations entre les particules dans le système, et servent ainsi à décrire leur micro-structure.

$$\rho_N^{(n)}\left(\mathbf{r}^n\right) = \frac{N!}{(N-n)!}\frac{1}{h^{3N}N!Q_N}\int\int e^{-\beta H}d\mathbf{r}^{(N-n)}d\mathbf{p}^{(N-n)} = \frac{N!}{(N-n)!}\frac{1}{Z_N}\int e^{-\beta V_N}d\mathbf{r}^{(N-n)} \tag{5}$$

L'équation 5 définit la probabilité de trouver N particules du système avec les coordonnées dans l'élément de volume, sans tenir compte de la position d'autres particules et sans tenir compte de tous les moments. Les densités de particule et les fonctions de distribution, définies ci-dessous, fournissent une description complète de la structure d'un liquide. La connaissance de la distribution de particules, en particulier de la densité de paire $\rho_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right)$ est souvent suffisant pour calculer l'équation d'état et d'autres propriétés thermodynamiques. La fonction de distribution de n-particules est définie par la densité de particule correspondante par:

$$g_N^{(n)}\left(\mathbf{r}^n\right) = \frac{\rho_N^{(n)}\left(\mathbf{r}_1, ..., \mathbf{r}_n\right)}{\prod_{i=1}^{n}\rho_N^{(1)}\left(r_i\right)} \tag{6}$$

qui pour un système homogenous est réduit à

$$\rho^n g_N^{(n)}\left(\mathbf{r}^n\right) = \rho_N^{(n)}\left(\mathbf{r}_1, ..., \mathbf{r}_n\right) \tag{7}$$

Nous discuterons plus en détail la fonction de distribution de paire définie comme:

$$g_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right) = \frac{\rho_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right)}{\rho_N^{(1)}\left(\mathbf{r}_1\right)\rho_N^{(1)}\left(\mathbf{r}_2\right)} \tag{8}$$

Si le système est homogène et isotrope, la fonction de distribution de paire $g_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right)$ est la fonction seulement de la séparation $r = |\mathbf{r}_2 - \mathbf{r}_1|$. Sa moyenne angulaire est appelée fonction de distribution radiale (RDF). Les oscillations de cette fonction indiquent essentiellement l'empilement des particules au voisinage de la particule à

l'origine. Mais, elles indiquent aussi la nature des corrélations entre voisins et leur interprétation demande souvent un effort de compréhension de la structure sous-jacente, surtout dans le cas ou des corrélations angulaires sont également présente.

La définition de $g(r)$ implique que le nombre moyen de particules qui est dans la gamme entre $r$ et $r + dr$ de la particule de référence est $4\pi\rho g\,(r)$. Les pics dans $g\,(r)$ représentent les couches de voisins autour de la particule de référence. L'intégration de $4\pi r^2\rho g\,(r)$ jusqu'àla position du premier minimum fournit donc une estimation du nombre de coordination, CN. Le concept d'une couche et de nombre de coordination sont plus appropriés pour les solides que pour les liquides, mais ils fournissent une mesure de la structure d'un liquide. Pour l'argon nous pouvons estimer que le nombre de coordination est $CN \approx 12.2$, qui montre qu'en moyenne chaque atome de l'argon est entouré par 12 voisins.

Nous avons exploré aussi la fonction de distribution radiale dans l'espace de Fourier, que l'on appelle le facteur de structure. Le facteur de structure est une mesure de la réponse de la densité du système, au départ en l'équilibre, à un petite perturbation externe de longueur d'onde $2\pi/k$ [1].

Cette mesure permet aux moyens de déterminer la distribution radiale par l'utilisation de la transformation de Fourier. I y a donc un lien direct entre la transformation de Fourier et la mesure par diffusion (SANS ou SAXS). Donc, le S(k) est donné par la transformation inverse et le $g(r)$ est donné par la transformation inverse:

$$\rho g\,(\mathbf{r}) = (2\pi)^{-3} \int (S\,(\mathbf{k}) - 1)\, e^{-i\mathbf{kr}} d\mathbf{k} \tag{9}$$

Quand le potentiel intermoléculaire est décrit comme étant la somme des interactions entre sites atomiques, la façon naturelle de décrire la structure du liquide est par les fonctions de distribution site-site. Si la position du site $\alpha$ sur la molécule $i$ est dénoté par $r_{i\alpha}$ et celui du site $\beta$ sur la molécule $j$ par $r_{j\beta}$, alors la fonction de distribution site-site radiale, sRDF $g_{\alpha\beta}\,(r)$ est défini comme:

$$g_{\alpha\beta}\,(r_{\alpha\beta}) = \frac{\langle \rho_{i\alpha}\,(r_{i\alpha})\,\rho_{j\beta}\,(r_{j\beta})\rangle}{\rho_\alpha \rho_\beta} \tag{10}$$

oú la densité $\rho_\alpha = N_\alpha/V$ est la densité d'un site $\alpha$.

Les fonctions de distribution site-site ont une interprétation physique analogue à celle de la RDF. Elles sont aussi directement rattachés aux facteurs de structure, mesuré par la diffusion de rayons X et les neutrons. D'autre part, l'intégration sur tous les angles implique une perte irrémédiable de renseignements. Ainsi, la fonction de distribution totale ne peut être reconstruite exactement à partir de l'ensemble fini de toutes les fonctions de distribution site-site.

## 3. La Théorie de Kirkwood-Buff

En 1951 Kirkwood et Buff [2] ont proposé une théorie des fluctuations de concentration dans les solutions. Ils ont montré que les propriétés thermodynamiques comme les volumes molaire partiels, la compressibilité isotherme et les dérivées du potentiel chimique par rapport à la concentration, peuvent être exprimé comme les intégrales des fonctions de distribution radiales de toutes les paires moléculaires présentent dans les solutions. Quelques décades plus tard, Ben-Naim [3] a suggéré que les mesures expérimentales des quantités thermodynamiques peuvent, à leur tour, donner les renseignements sur les distributions moléculaires. L'intégrale de Kirkwood-Buff (KBI) est définie comme:

$$G_{ij} = 4\pi \int_0^\infty \left( g_{ij}^{(2)}(r) - 1 \right) r^2 dr \tag{11}$$

oú $i$ et $j$ représent le soluté et/ou le solvant et $g_{ij}(r)$ la fonction de distribution radiale correspondante.

Les équations valables pour un système binaire sont données dans le paragraphe suivant, mais la relation semblable peut être écrite pour chaque système à plusieurs composants. Il est convenable d'exprimer les KBIs($G_{ij}$) dans les unités de $\left[\frac{cm^3}{mol}\right]$:

$$G_{ij} = -G_{12}\delta_{ij} + (1-\delta_{ij}) \left( RT\chi_T - \frac{V_1 V_2}{V_m \widetilde{D}} \right) + \frac{\delta_{ij}}{\chi_i} \left( \frac{V_i}{\widetilde{D}} - V_m \right) \tag{12}$$

oú R représente la constante de gaz $R = 8.3145 JK^{-1}mol^{-1}$, $\chi_T$ est la compressibilité thermique, $V_m$ est le volume molaire, $V_j$ est le volume molaire partiel et $\mu_j$ est le potentiel chimique pour la fraction molaire $\chi_j$ du mélange et le $\widetilde{D} = \chi_j \dfrac{\partial \beta \mu_j}{\partial \chi_j}$.

## 4. Dynamique moléculaire

En théorie des liquides, on considère souvent que la simulation fournit des données presque exactes et quasi-expérimentales, à la condition que les modèles bien définis. Ceci est particulièrement vrai des modèles les liquides simples purs [1]. Les deux méthodes numériques les plus largement utilisées sont: la méthode de la Dynamique Moléculaire (MD) et la méthode de Monte-Carlo (MC) [4] [5].

Pour cette thèse, les simulations ont été exécutées en utilisant le programme de DLPOLY de la Dynamique Moléculaire qui a été développée dans le laboratoire de Daresbury à Cardiff [6]. Plusieurs différents modèles ont été utilisés pour la simulation des liquides associés, comme l'eau, méthanol et tert-butanol, aussi bien que de mélange alcools-eau. Les détails sur les champs de force ainsi que les acronymes des modèles sont énumérés dans la Table I. Tous les modèles sont des modèles site-site et les approximations de corps rigide ont été utilisées pour modéliser les liaisons intramoléculaire. Le corps rigide est une collection d'atomes dont la géométrie locale est invariante dans le temps (pas de polarisation). La dynamique de corps rigide est décrite par le mouvement translationnel du centre de masse et la rotation du corps rigide. Les interactions inter-moléculaires sont exprimées comme une somme de potentiels de Lennard-Jones et potentiels de Coulomb. Donc le potentiel entre deux molécules $i$ et $j$ est défini comme la somme d'interactions entre les sites $\alpha$ et $\beta$, sur les molécules $i$ et $j$, respectivement, avec les charges partielles $q_{\alpha i}$ et $q_{\beta j}$, et le diamètre Lennard-Jones $\sigma_{\alpha i \beta j}$, l'énergie $\varepsilon_{\alpha i \beta j}$ et les distances relatives $r_{\alpha i \beta j}$.

$$U_{ij}(r) = \sum 4\varepsilon_{\alpha i \beta j} \left[ \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{12} - \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{6} \right] + \sum \frac{1}{4\pi\varepsilon_o} \frac{q_{\alpha i} q_{\beta i}}{r_{\alpha i \beta j}} \qquad (13)$$

Les paramètres intermoléculaires Lennard-Jones ont été calculés par l'utilisation des règles Lorentz-Berthelot pour mélane $\varepsilon_{\alpha\beta} = \sqrt{\varepsilon_{\alpha\alpha}\varepsilon_{\beta\beta}}$ et $\sigma_{\alpha\beta}=(\sigma_{\alpha\alpha} + \sigma_{\beta\beta})/2$.

Les conditions des limites périodiques ont été utilisées pour imiter un système pseudo-infini. Nous avons exécuté la simulation dans les ensembles canonique (NTV constant) et isobare-isotherme (NPT constant). Les conditions standard ont été imposées: la température T=300 K et la pression de 1 atm, maintenus par le thermostat et barostat de Berendsen avec les temps de relaxation de 0.1 ps et de 0.5 ps, respectivement. Le pas de temps d'intégration a été fixé à 2 fs. Pour l'équilibration les périodes de au moins 100 ps ont été nécessaires. Nous avons aussi vérifié la convergence de l'énergie intérieure, le volume, la pression, et la stabilisation des fonctions de distribution. Le calcul des valeurs thermodynamiques, aussi bien que d'autre fonctions, a été exécuté pour un temps d'au moins 64 ps, pris après l'équilibration. Le pas de temps de 0.5 ps a été utilisé pour la collecte des fonctions de distribution.

- **Table I**. Champ de Force

| | Méthanol OPLS [7] | | | Méthanol WS [8] | | |
|---|---|---|---|---|---|---|
| | $H$ | $O$ | $CH_3$ | $H$ | $O$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.71131 | 0.86612 | 0.088 | 0.6506 | 0.8672 |
| $\sigma$(Å) | 0.0 | 3.071 | 3.775 | 1.58 | 2.664 | 3.748 |
| $q$(e) | 0.435 | -0.7 | 0.265 | 0.52 | -0.82 | 0.3 |
| | *tert*butanol OPLS [7] | | | | | |
| | $H$ | $O$ | $C$ | $CH_3$ | $CH_3$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.71172 | 0.20936 | 0.67073 | 0.67073 | 0.67073 |
| $\sigma$Å | 0.0 | 3.07 | 3.80 | 3.91 | 3.91 | 3.91 |
| $q$(e) | 0.435 | -0.7 | 0.265 | 0.0 | 0.0 | 0.0 |
| | Eau SPC/E [9] | | | Eau TIP4P [10] | | |
| | $HW$ | $OW$ | | $HW$ | $OW$ | M |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.650 | | 0.0 | 0.648 | 0.0 |
| $\sigma$(Å) | 0.0 | 3.116 | | 0.0 | 3.15365 | 0.0 |
| $q$(e) | 0.4238 | -0.8476 | | 0.5200 | 0.0 | -1.0400 |

| | Eau TIP5P [11] | | | |
|---|---|---|---|---|
| | $HW$ | $OW$ | $M1$ | $M2$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.6694 | 0.0 | 0.0 |
| $\sigma$Å | 0.0 | 3.12 | 0.0 | 0.0 |
| $q$(e) | 0.241 | 0.0 | -0.241 | -0.241 |
| | Acétone OPLS [12] | | | |
| | $C$ | $O$ | $CH_3$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.440 | 0.879 | 0.67 | 0.67 |
| $\sigma$Å | 3.75 | 2.96 | 3.91 | 3.91 |
| $q$(e) | 0.3 | -0.424 | 0.62 | 0.62 |

## 5. Fluctuations de densité dans l'ensemble N-constant

Afin de clarifier plusieurs problèmes importants qui concernent le calcul des propriétés structurelles, nous avons examiné la contribution des corrélations de paires au-delàdes premiers voisins pour les systèmes purs en utilisant les calculs de KBIs et les calculs d'énergies configurationelles.

La forme asymptotique de RDF aussi bien que la fonction de densité de paire pour un système fini, est d'habitude décrite avec un correction de 1/N. La correction générale a été d'abord trouvé par Ornstein et Zernike [13]:

$$\lim_{(r\to\infty)} \rho\left(1,2\right) = \rho^2\left(1 - \frac{\rho k_B T \chi_T}{N}\right) \tag{14}$$

Il a été prouvé de façon plus générale par Lebowitz et Percus [14] dans le cadre de la thermodynamique statistique. Nous avons évalué comment ce fait affecte les résultats de la simulation, c'est--dire comment les résultats dépendant de la taille du système. D'autres questions importantes concerne le calcul de l'intégrale de Kirkwood-Buff. KBIs simulés sont obtenus par l'hypothèse que:

$$G_{\alpha\beta} = 4\pi\int_0^\infty \left(g_{\mu VT}\left(r\right) - 1\right) r^2 dr = 4\pi\int_0^R \left(g_{NVT}\left(r\right) - 1\right) r^2 dr \tag{15}$$

Donc, on suppose que dans l'ensemble N-constant nous avons les fluctuations de densité apparente qui donnent le même résultats que celles dans le système infini. Lebowitz et Percus [14] discutent aussi sur la nature locale de la densité, et ils montrent qu'une densité qui varie lentement peut être interprété en représentant chaque élément liquide comme un système ouvert qui échange des particules avec les éléments liquides adjacents (le nombre de particules total étant maintenu). Cela donne le sens à un potentiel chimique local aussi bien que àdes fluctuations de la densité locale dans des systèmes à N-constant. Donc, on peut utiliser les simulation dans l'ensemble à N-constant pour le calcul des intégrales de Kirkwood-Buff qui est rigoureusement défini seulement pour les ensembles ouverts de type Grand Canonique.

Le calcul de RDF et de KBI pour les liquides moléculaires est fait en utilisant les fonctions site-site. En suivant comment les valeurs asymptotiques sont atteintes, on peut présenter les quantités courantes (running). L'intégrale courante de Kirkwood-Buff pour les sites $\alpha$ et $\beta$ $G_{\alpha\beta}(R)$ est defini par:

$$G_{\alpha\beta}(R) = 4\pi \int_0^R \left(g^{(2)}_{\alpha\beta}(r) - 1\right) r^2 dr \tag{16}$$

D'après le comportement asymptotique d' intégrales courante KBI on peut constater qu'aux distances intermoléculaires courtes cette quantité oscille, reflétant l'arrangement microscopique de molécules, tandis que pour des distances les plus grandes elle ne varie plus, atteignant ainsi sa valeur asymptotique.

$$G_{\alpha\beta} = \lim_{(r\to\infty)} G_{\alpha\beta}(R) = G \textit{ pour les sites } \alpha \textit{ et } \beta. \tag{17}$$

De la meme facon, la compressibilité courante peut etre obtenue par l'équation:

$$\chi_T(R) = \frac{1 + \rho G(R)}{\rho k_B T} \tag{18}$$

avec $\chi_T(R \to \infty) = \chi_T$ .

L'énergie interne d'excès d'un liquide moléculaire qui est d'ecrit par les interactions additives par paire des sites $v_{\alpha\beta}(r)$ et les RDFs site-site, correspondant, est donné par:

$$U_{\alpha\beta}(R) = 2\pi\rho^2 \int_0^R v_{\alpha\beta}(r)\, g_{\alpha\beta}(r)\, r^2 dr \tag{19}$$

oú les énergies de configuration totales sont égales:

$$\frac{U^{ex}}{N} = \sum_{\alpha,\beta} U_{\alpha\beta}(R \to \infty) \tag{20}$$

Le terme $U_{\alpha\beta}(R)$ est une nouvelle quantité et il calcule la contribution courante de la distribution de énergie. Il est fondé sur l'analogie avec l'intégrale de Kirkwood-Buff

courante (l'équation 16). L'équation 20 permet de calculer séparément les contributions des interactions Lennard-Jones interactions, et celles des charges partielles. On peut alors étudier comment chacun des type d'interaction contribue à l'établissement global de l'énergie du système macroscopique.

L'énergie configurationelle peut être obtenu aussi, en utilisant des méthodes standard, comme la troncature du potentiel et les sommes d'Ewald. Nous avons évalué la consistance de ces deux calculs, quand la taille de systèmes a été changée. Cette étude inclut une simulation de dynamique moléculaire de plusieurs modèle d'eau, aussi bien que les liquides organiques comme l'acétone et le méthanol, àla température de 300 K et la pression de 1 atm et l'utilisation de plusieurs tailles de système, depuis N = 256 jusqu'à N = 10976 pour l'eau et N = 256-2048 pour d'autre liquides organiques. Ici, nous présentons comme un exemple, les résultats de la simulation de l'eau, mais les résultats semblable ont étés aussi obtenus pour d'autres systèmes étudiés, comme le méthanol et l'acétone. La Figure 1 montre les KBI courants (l'équation 16) calculés pour les modèles différents d'eau, aussi bien que la valeur expérimentale $G = -16.9\, cm^3/mol$. Cette valeur a été déduite de la compressibilité expérimentale à la température ambiante qui est $\chi_T = 0.4566\, GPa^{-1}$[15]. La valeur pour la compressibilité réduite est alors $\chi_T^* = \frac{\chi_T}{\chi_T^0} = 0.0623$. La correspondance RDFs, qui sont égaux àsRDF pour les sites O-O, sont montré dans le haut insert.

L'énergie de Coulomb totale résultante est montrée dans la figure 2, tous les deux pour les modèles SPC/E et TIPnP. Nous remarquons d'abord que le RDF (l'équation 19) s'accorde très bien avec le calcul direct, tant pour le Lennard-Jones que pour les énergies de Coulomb.

Notre but principal dans cette étude est la structure de liquides moléculaires comme l'eau, le méthanol et l'acétone, notamment au-delàdu premier pic des RDF. Nous avons montré que ces corrélations, qui reflètent les fluctuations de densité, donnent des KBIs (ou autrement dit la compressibilité) qui sont constant. Ce fait indique que les simulations à N-constant n'affectent pas les corrélations dans un large intervalle de distance, qui va jusqu'àla demi-bote. En plus, pour tous les modèles étudiés

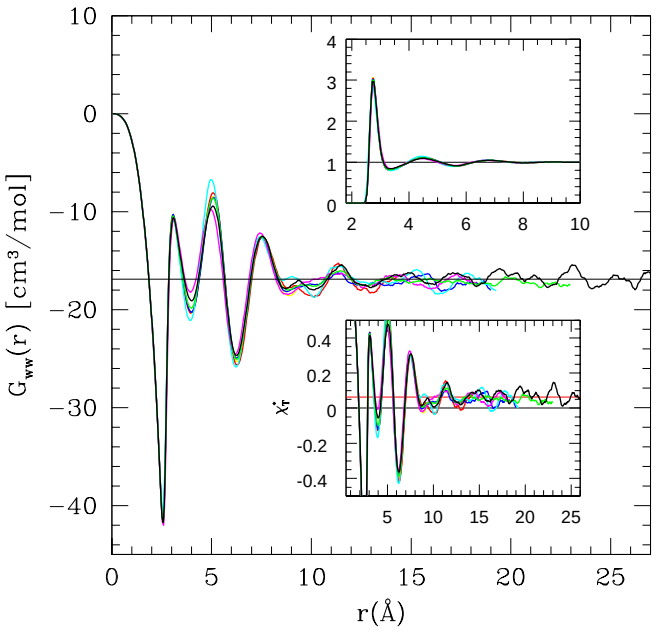


Fig.1 Zoranic et al.

Figure 1: L'intégrale de Kirkwood-Buff courante pour les modèles d'eau et pour de différentes grandeurs de système (SPC/E: N=256 (jaune), 864 (rouge), 2048 (bleu), 4000 (vert), 10,976 (noir); TIP4P: N=2048 (magenta); TIP5P: N=2048 (cyan)). L'insert supérieur montre le RDF avec memes conventions en couleur. Plus bas l'insert montre running la compressibilité (réduite) (l'équation 18) pour tous modèles avec memes conventions en couleur. La ligne horizontale rouge est la compressibilité expérimentale.

ici, la valeur numérique obtenue pour le KBI des simulations présente un accord tout à fait remarquable avec les valeurs expérimentales. A lui seul, ce résultat indique que, non seulement les modèles reproduisent les fluctuations de densité correctement, mais aussi dans un grand intervalle de distances intermoléculaires.

Nous avons montré que la dépendance de la grandeur n'affecte pas les fonctions de corrélations. Finalement, nous analysé la distribution d'énergie en comparant la moyenne avec celle obtenue par le calcul direct des RDF site-site. Cette analyse révè le que la convergence des interactions LJ est environ 15 Å, alors que l'énergie de

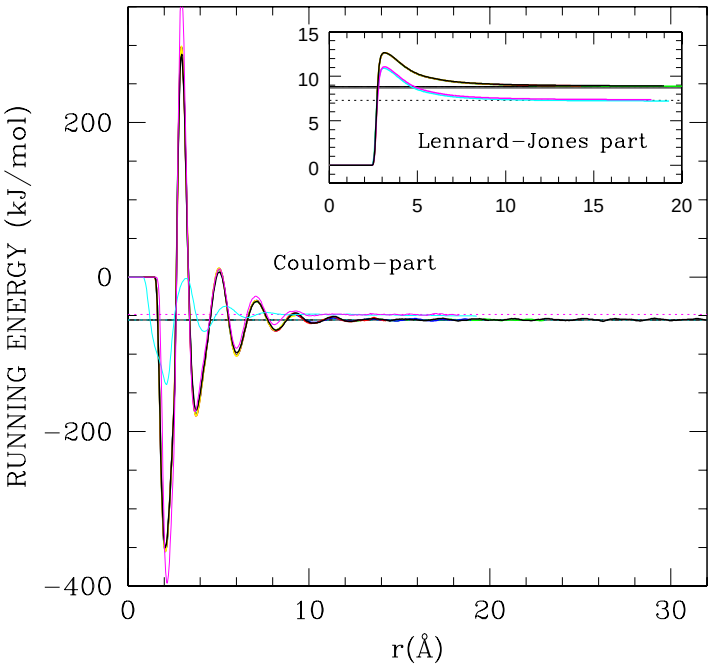


Figure 2: Les contributions “running” Coulomb total et Lennard-Jones (montré dans l’insert) les énergies pour les modèles d’eau. Les conventions en couleur sont le meme comme dans la figure 1. Les lignes horizontales (tous les deux dans le principal et l’insert) dénote les valeurs des énergies calculé directement dans les simulations.

Coulomb montre des oscillations régulières avec une période sans doute rattachée à l’alternance de charges de site positifs et négatifs, alors que la décroissance semble contrôlée par la force d’écrantage. Ce calcul fournit quelques renseignements indirects sur la structure du liquide pur. Nous avons prouvé que la simulation permet d’obtenir des résultats concernant la micro-structure et les propriétés qui sont rattachées aux fluctuations. Donc, elle peut être utilisée pour aborder les fluctuations spatiales et l’organisation de la micro-hétérogènes dans les liquides pur et dans leur mélanges.

## 6. Alcools

Les alcools sont des molécules organiques dans lequel un groupe hydroxyle (-OH) est attaché àun atome de carbone (-C) d'un alkyle ou groupe alkyle substitué. La formule générale pour un simple l'alcool est $C_nH_{2n+1}OH$. Ils sont classifiés dans les trois sous-ensembles importants: primaire, secondaire et tertiaire, basé sur le nombre de carbones avec qui -C est attaché. Le plus petit alcool primaire simple est le méthanol. L'alcool secondaire le plus simple est de l'alcool isopropyl (propan-2-ol), et un alcool tertiaire simple est de l'alcool de tert-butyle ou tert-butanol (2-methylpropan-2-ol). Pour ces alcools les propriétés chimiques et physiques sont fortement sous l'influence de la nature opposée de leurs composants: le groupe hydroxyle polaire et la queue hydrophobe de carbone. La dominance d'une affinité sur l'autre définit aussi la solubilité d'alcools dans l'eau ou d'autres solvants. Dans le cas d'alcools simples, la tendance du groupe - OH à former les liaisons hydrogènes permet de vaincre la résistance à l'eau de la partie hydrophobe, les rendant ainsi solubles à toutes les concentrations. Dans notre recherche, nous nous concentrons sur le méthanol et tert-butanol. Le méthanol comme le plus simple de tous les alcools et le tert-butanol car, en plus du fait d'avoir la géométrie caractéristique de le la partie hydrophobe, il est aussi le plus grand alcool qui est complètement soluble avec l'eau.

## 7. Micro-structure de liquides purs

On pense généralement aux liquides comme étant macroscopiquement homogènes, surtout quand ils sont considérés loin des transitions de phase. Pourtant, dans les liquides associés, la liaison hydrogène qui est extrêmement directionnelle, a tendance à modifier localement leur structure, en menant à la formation d'immiscibilité locale par petits groupements dans la phase homogène à équilibre. A savoir, la topologie de la liaison hydrogène dépend de l'electronegativité atomique et la forme de la molécule. La molécule de méthanol, par exemple, a une connectivité linéaire:

un groupe de hydroxyle OH est en moyenne en liaison avec deux hydrogènes des molécules adjacentes. Les molécules d'eau forment des liaisons tetrahedrales qui permettent quatre voisins. Pourtant, le fait qu'un système a une topologie anisotrope des voisins au contact n'est pas suffisant pour le classer cela comme micro-structuré ou micro-hétérogène. La définition de micro-structure implique l'existence d'associations spécifiques de molécules sur une plus grande échelle que le contact de paire. C'est toujours une définition vague, puisque que nous n'avons pas définit les propriétés qui devraient être sensibles à un tel agrégation, les "signatures" (thermodynamique ou structurel) qui émergeraient d'une telle organisation locale. Nous avons donc étudié par dynamique moléculaire les liquides purs, le méthanol, le tert-butanol ainsi que l'eau, pour extraire les renseignements sur les propriétés de agrégation.

Nous nous sommes concentrés sur les fonctions de corrélation de densité: fonction de distribution radiale et facteur de structure. Les propriétés du facteur de structure sont bien comprises, comme par exemple dans le cas des cristaux. Il a une forme caractérisée par les interférences constructives et destructives des interférences sur les atomes périodiquement arrangés. De même, le facteur de structure dans les liquides indique une périodicité, dans le sens en fait de moyenne, de l'arrangement des molécules dans le liquide. Donc, il devrait fournir des renseignements sur les structures locales dans les liquides associés. Nous avons considéré les fonctions de distribution radiales site-site, afin de produire partiellement les informations angulaires et aussi la fonction de distribution radial du centre de masse. Les RDF site-site sont définis comme:

$$g_{\alpha\beta}\left(r_{\alpha\beta}\right) = \frac{\left\langle \rho_{\alpha}\left(r_{\alpha}\right)\rho_{\beta}\left(r_{\beta}\right)\right\rangle}{\rho_{\alpha}\rho_{\beta}} \tag{21}$$

En plus de la fonction de site de site $g_{\alpha\beta}\left(r_{\alpha\beta}\right)$, nous avons aussi calculé les facteurs de structure correspondants:

$$S_{\alpha\beta}\left(k\right) = 1 + \rho\int dre^{-ikr}g_{\alpha\beta}\left(r\right) \tag{22}$$

Aussi, notre analyse consiste à compter de différentes tailles et des formes de

groupes liés par la liaison hydrogène. Pour les liquides simples, la distribution de groupe est un résultat de la fluctuation du nombre de particules et nous utiliserons cette caractéristique comme une référence. La différence entre le distributions de groupe calculée et la référence montreront agrégation spécifique qui est la conséquence de l'heterogenéité locale. Les probabilités de trouver un groupe de grandeur sont définies comme:

$$p(n) = \frac{\sum_{k} s(k,n)}{\sum_{n,k} s(k,n)} \tag{23}$$

oú $s(k,n)$ représente le nombre de clusters de la taille $n$ dans la configuration $k$. Donc, $p(n)$ représente la distribution de taille de agrégat. Pour cette analyse nous avons choisi la définition de Stillinger [16] oú deux particules sont liés s'ils sont séparés par moins qu'une certaine valeur d'approche. En pratique, cette distance est choisie pour correspondre au contact de paire moyen comme décrit par la fonction de distribution radiale.

La figure 3 montre la fonction de distribution radiale au site pour les sites de méthanol. Le panneau supérieur montre des corrélations entre les sites de méthyle aussi bien que la fonction de distribution radiale entre centre de masse. Les autres panneaux montrent les RDFs des sites qui sont impliqués dans la liaison hydrogène. La première caracté ristique remarquable sont les oscillations prononcées dans $g_{MM}(r)$ aussi bien que dans $g_{cm}(r)$, que pour la distance plus grande que ≈6 Å sont presque superposé. Le $g_{cm}(r)$ a aussi un premier pic étroit qui est la signature de liaison hydrogène. Donc, la RDF du centre de masse comprend la corrélations à courte portée en raison de la forte interaction entre les molécules adjacentes et les oscillations dues à l'empilement moléculaire typiques pour les liquides simples. Les fonctions de distribution impliquant les sites qui sont lies avec la liaison hydrogène ont aussi des caractéristiques de courte portée caractéristiques: le fort premier le pic qui correspond à une liaison hydrogène. La valeur maximale dans $g_{OH}(r)$ est à la distance 2.5 Å qui est la valeur entrée dans l'usage général pour la liason hydrogène.

Ce qui est surprenant c'est que toutes les fonctions de distribution des HB-sites

montrent un manque de corrélations pour les distances moyennes et grandes. Les RDFs de ces sites pour les distances plus grandes que ≈8 Å sont apparemment égaux à un, pendant que le $g_{cm}(r)$, pour la même distance, a toujours la structure oscillatoire typique d'un liquide. Le manque d'oscillations pour les distances de médium-á-longue portée est d'une manière caractéristique trouvée dans la fonction de distribution monomère-monomère dans les liquides de polymères [17].

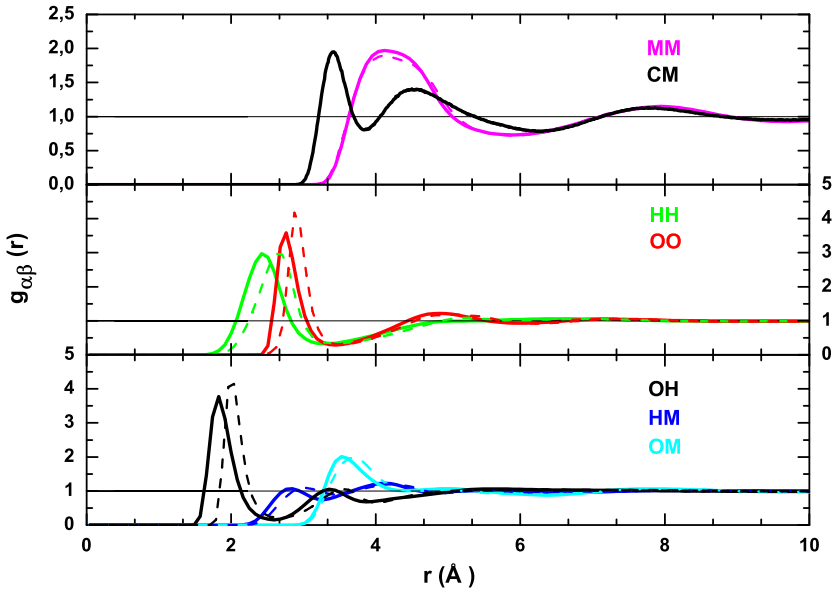


Figure 3: Les RDFs site de site pour OPLS et les modèles de WS de méthanol. Panel du haut: pour le sites MM (magenta) et le centre de masse (noir). Panel du milieu: pour OO (rouge) et les sites HH(verts). Panel du bas: pour OH le panel noir, OM (cyan) et les sites HM (bleus). Pour tous les données, les courbes complètes sont pour le modèle d'OPLS et anéanties pour WS modèle.

Donc, ces résultats peuvent être interprétés comme: les molécules MetOH ont tendance à former des chaînes localement, O et les sites H étant fortement en corrélation le long les chaînes, pendant que les sites M sont distribués aléatoirement autour des

chaînes. La distribution de sites M correspond àune distribution uniforme et n'est pas apparemment influencée par la formation de chaîne. Les corrélations de site de méthyle aussi bien que le centre de masse ont des caractéristiques semblables au liquide simples, la structure oscillatoire prononcée indiquant l'empilement typique du liquide dense, pendant que l'oxygène et les sites hydrogènes ont les corrélations anisotropes typiques d'une structure de chaîne.

Dans le cas du tert-butanol, comme montré dans la figure 4, tous les sites montrent le pic pointu typique de la liaison hydrogène. Cela indique que la liaison hydrogène influence aussi les corrélations à courte portée de sites qui ne sont pas impliqués dans ces interaction. Les RDF des centres de masses sont presque identiques à que indirectement évalue les positions centrales de l'atome de carbone dans la molécule. A part le premier pic, toutes les corrélations dans le panneau de gauche suit les oscillations d'empilement des liquides simples, et ils sont superposés dans les distances médium-á-grandes. On peut imaginer que ces oscillations estiment grossièrement le diamètre de particule LJ-sphériques du centre de masse qui ont donc des empilements semblables. Donc, nous attribuons cette valeur à la grandeur moléculaire moyenne.

Par contraste avec le méthanol, toutes les fonctions de distribution de tert-butanol affichent le comportement oscillatoire, avec la différence principale dans la période d'oscillation: les sites qui sont liés avec la liaison hydrogène ont une période plus grande que le périodicité de corrélation des centres de masses. Alors qu'on s'attend à ce que la structure d'empilement soit de la grandeur moléculaire, une plus grand période indique que les corrélations dues àla liaison hydrogène sont modulées par la supra-structure résultante. Le fait que cette dernière structure oscillatoire est absente pour le méthanol est indicatif des différences dans l'organisation locale entre deux liquides: les molécules de TBA forment des micelles sphérique, donc la modulation du RDF est plus semblable à celle d'un liquide et plus structurée que pour cas de méthanol, dont nous avons vu la structure de type polymère. Nous nous attendons donc à observer dans les facteurs de la structure les différences dans la structure affichées par les RDFs.

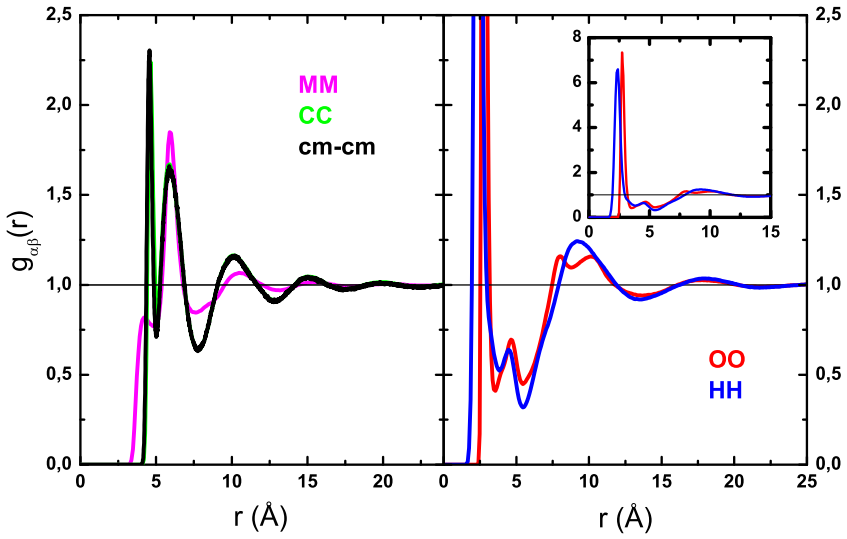


Figure 4: Les RDFs site de site pour OPLS tert-butanol; panel gauche: MM (magenta), CC (vert) et noir pour le centre de masse; panel juste: OO (rouge), HH (bleu) (l'insert montre le détail de le pics).

La figure 5 montre les facteurs de structure pour le méthanol site de site. Dans les panneaux supérieurs nous montrons les facteurs de structure pour les sites de méthyle et le centre de masse, pendant que le panneau plus bas représente les sites liés par la liaison hydrogène. Les facteurs de structure dans le panneau supérieur ressemblent à ceux d'un liquide de Lennard-Jones-type. Le premier pic dans $S_{MM}(k)$ est à $k_{eff} = 1.75$ Å$^{-1}$, dont nous pouvons extraire $\sigma_{eff} = 3.42$ Å, cela correspond grossièrement à la grandeur du site de méthyle. Le pic principal dans les facteurs de structure d'habitude indique la structure empilement. Dans notre cas cela correspond aux premiers pics du centre de masse aussi bien que le premier pic dans le facteur de structure de méthyle.

Par contraste avec cela, les facteurs de structure impliquant les sites qui sont liés avec la liaison hydrogène (le panneau plus bas) ont une forme très particulière: ils

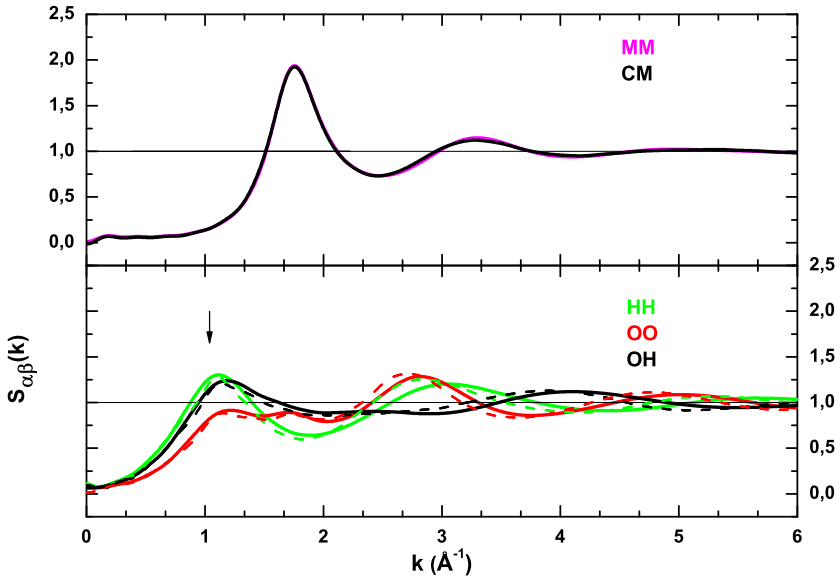


Figure 5: Les facteurs de structure de site du RDFs (OPLS et les modèles de WS) montré dans la figue 3, avec la meme couleur et convention de ligne; panel supérieur: le MM site et le centre de masse; fond panel: OO, OH et les HH. Le pré-pic est indiqué par la flèche.

ont un pic à $k_p$ plus petit alors $k_{eff}$. La plus petite valeur de k correspond à la plus grande grande de la structure, par exemple $S_{OO}(k)$ a $k_p = 1.25$ Å$^{-1}$ qui correspond à un périodicité de $\sigma_p \approx 5.1$ Å. Nous appellerons pré-pic ce pic pour souligner que sa valeur en k est plus petite que celle du pic principal (ce pic est aussi appelé le pic intérieur dans la littérature [18]). Conformément à l'interprétation spatiale précédente, nous associons ce pré-pic à la structure semblable à la chaîne de sites liée par liaison hydrogènes dans le méthanol liquide. La figure 6 montre le facteur de structure du tert-butanol. De nouveau, les facteurs de structure semblables àcelui d'un liquide typique sont observés pour tous les sites hydrophobes. Le pic principal de $S_{cm}(k)$ aussi bien que $S_{CC}(k)$ et $S_{MM}(k)$, est au vecteur de signe $k_{eff} = 1.34$

$Å^{-1}$ qui correspond à un estimé de la taille moléculaire $\sigma_{eff}$ ≈5.6 Å. La facteur de la structure de site hydrogène affiche un pré-pic conforme à la longueur 7.85 Å. Donc, ces deux longueurs correspondent aux deux périodes différentes de la structure oscillatoire dans la figure 4. Le pré-pic est plus haut que le pic principal, qui accentue la forte influence de l'association sur les caractéristiques structurel dans le TBA, par opposition au cas de méthanol liquide, oú le facteur de structure pour le pic principal a la valeur la plus élevée.

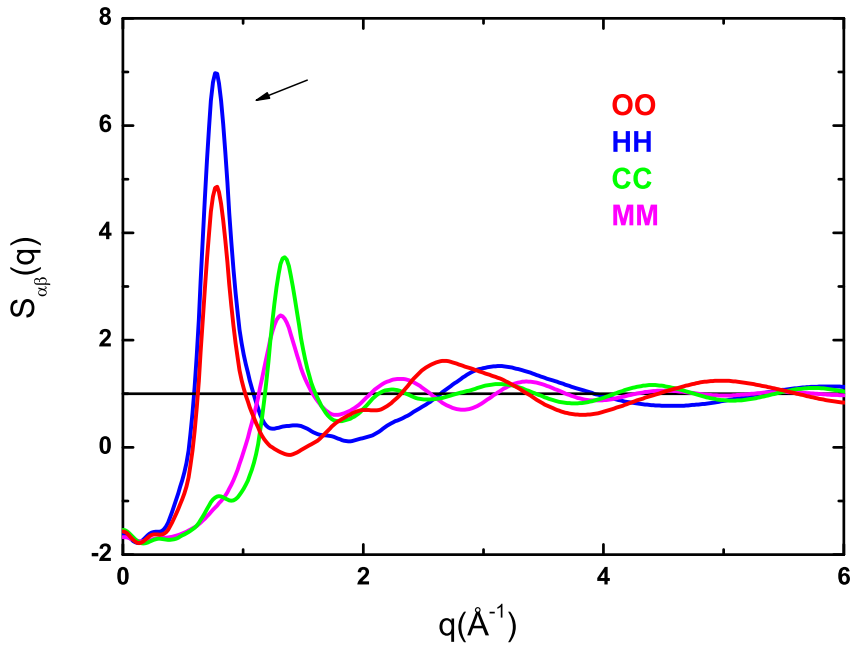


Figure 6: Les facteurs de structure de site-site pour un RDFs pour OPLS tert-butanol montré dans figure 4, avec la méme couleur convention. Le pré-pic est indiqué par la flèche.

Dans la figure 7 nous présentons les résultats pour de probabilité de distribution d'agrégats comme une fonction de taille de ces agrégats, pour les modèle de WS et OPLS du méthanol. Le panneau inséré montre $p^{MM}(n)$ pour les agrégations n¡250

pour le groupe méthyle et pour deux différentes valeurs de contact $l_c$. Les valeurs s'étendent les distances autour des premiers minimums de $g_{cm}(r)$ (voir la figure 3). La courbe suit la décroissance de type exponentiel pour de petites grandeurs de agrégat $l_c$. Les agrégats de grande taille apparaissent pour de plus grandes valeurs de $l_c$. Cela est en accord avec ce qu'on trouve dans le cas de liquides simples. Le panneau àgauche dans la figure 7 montre l'agrégation $p^{OO}(n)$ pour les sites oxygènes. La forme globale est semblable à celle trouvée pour l'agrégation du groupement méthyle, pourtant, nous observons un pic évident autour de n=5. Cette caractéristique est robuste au choix de $l_c$ pour valeurs autour des premiers minimums.

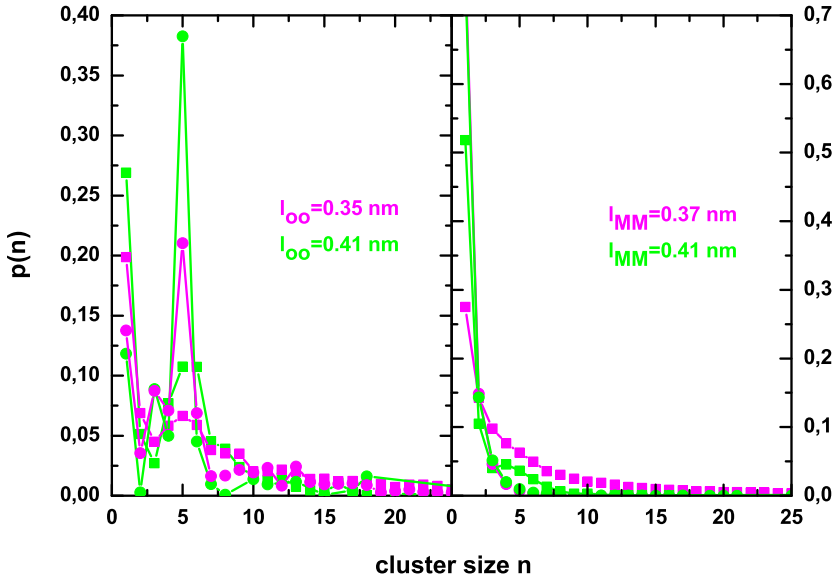


Figure 7: La probabilite de clusters de site pour OPLS et WS méthanol: panels gauches: clusters de site oxygène. panel du milieu: cluster de site méthyle. Pour chaque clustering les résultats pour deux distances sont montrées. Les données pour le modèle d'OPLS sont dans les carrés et WS dans les cercles.

Les probabilités élevées pour les tailles de cluster environ 5 indiquent que ces clusters apparaissent plus fréquemment que prévu, indiquant ainsi les structures préférées. De meme, l'analyse pour le cas de tert-butanol montre un pic à n égale à 4. Pour l'oxygène, la probabilité d'aggrégat à n=1 est plus élevée que à n=4, indiquant la plus grande probabilité de monomère. L'algorithme detecte des motifs aggrégés en forme de boucle, mais aussi des chaines ouvertes. L'analyse est robuste au choix de la valeur d'approche minimale des sites. Nous négligeons délibérement les clusters de taille infinie en nous concentrant sur ceux de taille petite. Donc les résultats plus significatifs sont des motifs structurels trouvés dans le clusters de tailles favorables. Pour le cas de méthanol, dans la gamme n=3-7, nous trouvons que 81 % des clusters forment les chaînes ouvertes et que seulement 19 % apparaissent sous forme des boucles qui connectent les atomes d'oxygène. Pour le tert-butanol, les meme calculs ont montré, que dans la taille de $n = 3 - 7$, environ 65 % des OO clusters sont des boucles, et 35 % forment une structure de chaîne. En cas de deux alcohols nets simulés, méthanol et tert-butanol, nous avons montré que leurs phases liquides sont partiellement ou complètement micro-structuré dans les conditions ambiantes. Nous insistons sur le fait que ces liquides désordonnés sont en fait hautement microstructurés sans être pour autant en instabilité. C'est attesté par la petite valeur du facteur de structure à $k = 0$. Nous avons montré que les fonctions de corrélation de densité de site-site traduisent une haute structuration locale. Les fonctions de site hydrophobes ressemblent a ceux d'un liquide simple, tandis que les sites qui participent à la liaison hydrogène traduisent la structuration spécifique. Les facteurs de structure confirment l'existence de structures locales en accord avec l'analyse en clusters. Le méthanol apparait comme plus faiblement associé tandis que le TBA est fortement associé en micelles.

En examinant le cas d'eau liquide, nous n'avons pas trouvé d'évidence de clustering par l'analyse utilisée dans ce travail, ce qui est aussi en accord avec les conclusions d'autres auteurs [19]. Par rapport à ce point, il est intéressant de noter que le comportement aux petits k du facteur de structure d'eau montre une augmentation

remarquable à la température ambiante [20]. Ceci semble suggérer qu'un mécanisme de type fluctuation de densité se substitue à celui de clustering pour ce liquide si particulier.

## 8. Micro-hétérogénéité dans les mélanges eau-alcool

L'existence d'inhomogeneities à l'echelle du nanomètre dans les solutions aqueuses est maintenant acceptée comme une évidence depuis les dernières 10 années, surtout grace aux simulations numériques, mais ils sont aussi supportés par toute une série de résultats expérimentaux. En effet, l'entropie très basse de mélange dans les solutions de méthanol et l'eau (plus basse que pour un mélange ideal) a été expliquée en utilisant le concept d'immiscibility locale entre l'eau et de méthanol [21]. Les études spectroscopiques confirment aussi le clustering local d'espèces dans les solutions aqueuses [22]. Donc, c'est la compréhension des propriétés microscopique qui nous aide à dévoiler les caractéristiques macroscopiques. L'idée principale est que la ségrégation locale entre l'eau et l'alcool aide à préserver l'homogénéité globale ainsi que le désordre macroscopique.

La comparaison entre valeuers expérimentales et simulées des propriétes thermodynamiques telles que le volume, la densité et enthalpies, montre que les simulations sont en bonne mesures pour capturer ces propriétes particulières. Pourtant, l'analyse des quantités d'excès, qui sondent directement les contribution non-idéales, indiquent que, sur ce niveau, les modèles de simulation ont manqué de suivre les résultats expérimentaux. En effet, l'excès d'enthalpie a un maximum pour la concentration molaire environ 0.3, que le modèles de simulation ont du mal à reproduire correctement [23], [24]. L'appréciation directe des inhomogeneités locales se fait par les fonctions de corrélation: la fonction de distribution radiale et le facteur de structures $S(k)$, définis comme dans les chapitres précédents. La fonction de site mesure maintenant la corrélation entre les sites de chaque espèce, mais aussi les corrélations de inter-espèces. Alors, le RKBI est égal :

$$G_{\alpha i\beta j}(R) = 4\pi \int_0^R \left(g^{(2)}_{\alpha i\beta j}(r) - 1\right) r^2 dr \tag{24}$$

oú $\alpha i$ et $\beta j$ sites correspond au site des espèces $i$ et $j$, àsavoir $\alpha i$ et $\beta j$ peuvent etre égaux àn'importe lequel d'H, O, C, M, OW et sites HW. De méme les facteurs de structure, pour les sites $\alpha i$ et $\beta j$, sont définis comme:

$$S_{\alpha i\beta j}(k) = 1 + \rho_{\alpha i\beta j} \int dr e^{-ikr} g_{\alpha i\beta j}(r) \tag{25}$$

oú $\rho_{\alpha je\beta j}$ est la densité de sites $\sqrt{\rho_{\alpha j}\rho_{\beta j}}$ et $\rho_{\alpha j} = N_{\alpha j}/V$.

La figure 8 montre la dépendance de concentration pour le rapport molaire 0.2, 0.5 et 0.8: le panel supérieur montre la distribution radiale de site, pendant que le panel plus bas montre les "running" KBIs. Nous avons choisi le site de méthyle et le site d'oxygène d'eau puisque ces sites suivent les distributions de centre des masse, donc c'est un bon choix pour contrôler les corrélations d'espèces.

Les fonctions de corrélations de méthyle montrent des oscillations semblables au liquide et il n'y a presque aucune variation de ces fonctions en ce qui concerne la concentration. Ce qui est intéressant est le comportement de l'eau. La hauteur du premier pic du RDF augmente comme la concentration d'eau diminue, indiquant la majoration de la structure première-adjacente. Donc, les molécules d'eau ont une forte tendance à s'associer et c'est valide meme pour les petites fractions d'eau. Nous avons aussi évalué le comportement des sites à liaison hydrogènes du méthanol. Pour les concentrations de méthanol basses les corrélations entre les sites d'oxygène sont très faibles et le self-clustering de sites d'oxygène augmente avec l'augmentation des concentrations de méthanol. Cela montre que dans les solutions riches d'eau, la compétition pour HB entre l'eau et le méthanol détruit les chaînes de OO et c'est seulement quand la concentration de méthanol augmente que les sites d'oxygène restituent la self-association anisotrope. Meme si ces résultats semblent au premier coup d'oeil attendu, il est important de remarquer que les sites à liaison hydrogène dans le méthanol montrent une tendance différente que dans le cas de l'eau. Les molécules d'eau ont la plus forte préférence pour les self-associations , tandis

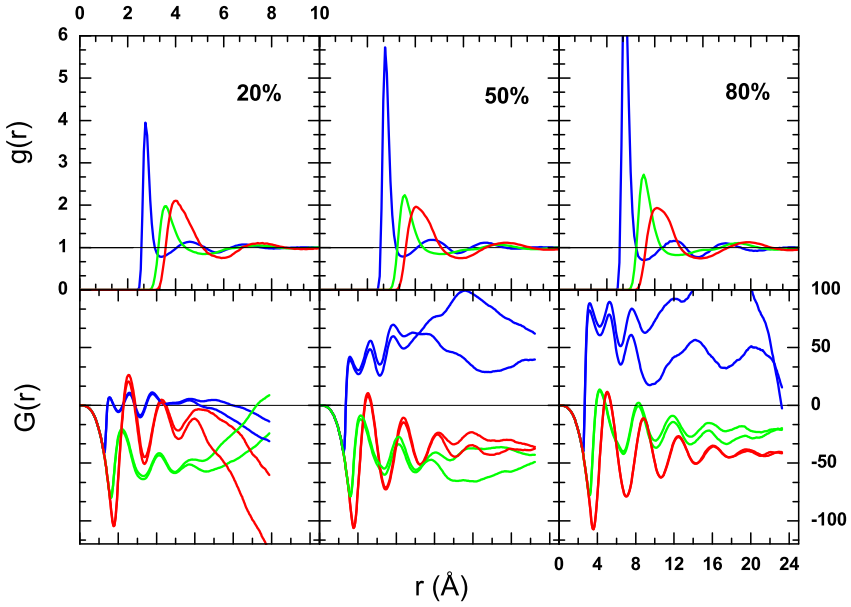


Figure 8: La dépendance de concentration du fonctions de distribution radial sont affichées par les panels supérieurs et les RKBIs correspondants dans les panels plus bas. Les corrélations de méthanol sont montrées dans les fonctions rouges, d'eau dans bleu et corrélations de trans-espèces dans vert. Les trois panels verticaux correspondez aux fractions molar d'alcool 0.2, 0.5 et 0.8 respectivement. Les RKBIs sont montrés pour deux différentes courses.

que les sites àliaison hydrogènes dans le méthanol agissent en faveur du clustering de l'eau, en soutenant plutôt qu'en détruisant les cluster d'eau. Les informations des RDFs sont plus faciles de comprendre aux courtes distances courte, puisque les corrélations àlongues portées sont cachées dans les queues apparemment sans structuration évidente. Donc, la corrélation pour les distances à moyenne portée sera plus évidente dans RKBI, oú les variations petites seront amplifiées dans le processus d'intégration.

Les corrélations àl ongue portée pour tous les sites de méthanol (le méthyle aussi bien que l'oxygène hydroxyl) montrent un comportement semblable, ils ont une valeur

asymptotique bien définie et de petites déviations de cette conduite sont observées pour la concentration de méthanol basse. Les corrélations d'eau, sont au contraire amplifiées. EN effet, KBIs de l'eau montrent la grande variation de la partie asymptotique et il n'est pas clair si ceux-ci tendent vers une limite asymptotique bien définie. De plus, les résultats des différents calculs par simulation reproduisent les données qui sont significativement différentes. Ces résultats peuvent etre expliqués par le fait que l'eau est hautement associée et les différentes simulations échantillonnent sur de différentes réalisations de ces associations.

Pourquoi alors les fonctions de méthanol ne montrent pas de tels signes de l'association? Les instantanés des simulations sont montrés dans la figure 10. On voit clairement que le méthanol est homogènement distribué, pendant que les clusters d'eau forment une structure semblable àla structure d'éponge. Nous pouvons suggérer que méthanol est homogènement distribué d'une telle façon que la majorité de molécules de méthanol sont aussi liées àcelles des domaines d'eau, et que juste quelques molécules de méthanol sont piégées dans les domaines d'eau. Donc, la micro-ségrégation d'espèces est implicite derrière les résultats montrés ici. C'est quand meme énigmatique que l'eau et le méthanol, en tant que donneurs et les accepteurs de liaisons hydrogènes, ne forment pas un mélange plus homogène. Les différences entre les organisations d'eau et de méthanol sont meme plus apparenten regard des facteurs de structure. Nous nous intéressons particulièrement aux comportement aux petits k-vecteurs, qui fournissent les informations sur l'organisation à grande échelle. Nous montrons des résultats puor concentration 0.2,0.5 et 0.8 (la figure 9)

Nous remarquons que tous les facteurs de structure d'eau montrent des pics aux petit k-vecteur qui indiquent l'existence d'associations à grande d'échelle. Les résultats des différentes runs montrent aussi la variation des pics aux petits-k, qui peuvent alors etre attribués à la réalisation spécifique d'un clustering de l'eau en réseau de type éponge. Le fait que le pré-pic dans le cas du méthanol, ou dans le cas de l'eau, possède un petit épaulement dans $S_{ww}(k)$, qui varie avec la concentration, montre qu'il y a une contribution venant de l'interface eau-alcool. D'habitude, la haute

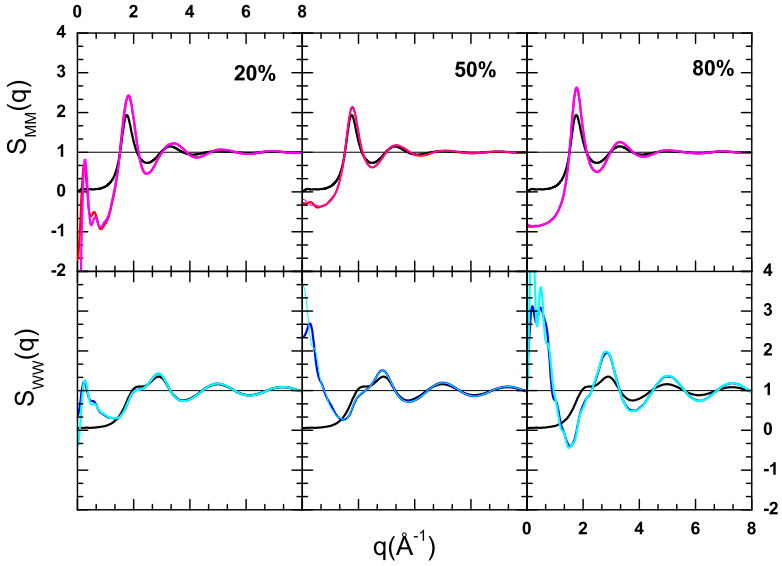


Figure 9: Les dépendances de concentration de les facteurs de structure. Les fonctions de méthyle sont montrées dans les panels supérieurs et les facteurs de structure d'eau dans panels plus bas. Dans chaque graphique sont montrées les résultats de deux calculs successifs avec temps de 64 ps: les fonctions de méthanol sont en rouge et magenta, tandis que les fonctions pour l'eau sont montrées en bleu et cyan. Comme le cas précédent, les facteurs de structure de chaque système pur sont présentés en noi : dans le panel supérieur le facteur de structure de méthanol et dans le panel plus bas le facteur de structure d'eau pure. Les trois panels verticaux correspondent aux fractions molaires d'alcool 0.2, 0.5 et 0.8 respectivement.

valeur à $S(k \to 0)$ indique que les fluctuations de concentration sont élevées et que le système est près de la séparation de phase. On peut conclure, que dans notre simulation, la taille du système est trop petite pour décrire les fluctuations qui mèneront le système à la séparation de phase, et donc tous les résultats sont évocateurs des structures résultantes. Quand meme, la demixion possible du système est indiquée par l'augmentation de la fluctuation pour toutes les composantes, qui n'est pas le cas ici. Donc, nous trouvons l'autre explication plus appropriée. A savoir que ces mélanges sont stables mais très fortements micro-structurés. Cette immiscibilité locale est due au fait que les liaisons l'hydrogènes entre les espèces identiques et croisées

ont tendance àse former pour des raisons énergétique et non pas entropiques, de sorte que le réseau forme une structure localement hétérogène. Notre analyse montre que cela s'accompagne aussi d'augmentation des fluctuations de concentrations, simplement à cause la répartition spatiale inégale des espèces moléculaires, qui est donc la caractéristique inhérente de ces liquides.

Pourtant, les facteurs de structure expérimentaux ne montrent aucun spécificité dans la région des petits-k-région, pour les mélanges aqueux [21]. Comment nous pouvons relier dans ce cas la simulation et résultats expérimentaux? Chaque état petit d'un liquide ordinaire semble hétérogène en raison des fluctuations de nombre de particule. Dans l'approche statistique des liquides simples, en faisant en moyenne, les fluctuations de densité ou de la concentration disparaissent. Pour les mélanges aqueux, la MH est l'inhomogénéité locale oú dans chacun l'état les composantes exposent leur immiscibilité locale, et c'est la caractéristique permanente de chaque réalisation de ce type de système. On peut donc supposer, que la MH correspond àun assemblage des espèces, mais avec un temps de vie bien défini. Donc, pour l'échantillonnage étendu, la signature due à la MH disparaîtra puisque la variété de formes et de grandeurs des clusters d'espèces aura une contribution statistique moyenne. Par contre, dans notre simulation, en raison de la petite grandeur de système et de aussi de la petite échelle de temps, nous ne sommes capable de mesurer qu'une partie de la réalisation de la MH, qui produit produit une variation du facteur de structure au petit k-vecteur. Donc, nous avons un échantillonnage sur la dynamique interne de la micro-heterogeneité. C'est la première fois que la MH est ainsi rattachée à une quantité mesurable, bien que seulement par l'intermédiaire des résultats de simulation. L'idée que nous proposons est que la micro-heterogeneité est la caractéristique inhérente des solutions aqueuses. L'étude de la MH nécessite des grandes tailles de système et de l'échantillonnage les temps dans les simulations des solutions aqueuses. Les grand tailles de système sont nécessaires afin de bien echantilloner toutes les conformations statistiques de la micro-hétérogénéité. Nous croyons que cette étude est un premier pas dans la direction que nous croyons est

important pour l'enquete des solutions aqueuses. Nous allons meme jusqu'à postuler une analogie entre les mélanges aqueux et les micro-émulsions (ME). En effet, les solutions aqueuses sont un peu semblables àdes systèmes micellaires, dans lequel le comportement structurelle dominant est gouverné par la micelle en tant que méta-molécules. La différence principale est que les molécules solute sont plus petites et ne produit pas un tel morphologies comme micelle. Les micro-émulsions ont été bien étudiées entre les années 70 et le début des années quatre-vingt, tant du point de vue expérimental que du point de vue théorique et beaucoup de livres ont été écrits sur ce sujet. Les vraies approches microscopiques à ME, basé sur la mécanique statistique, sont assez rares [25] [26] [27]. Le fait fondamental qui a permit une formulation théorique de ces systèmes est que la différence d'échelle entre les motifs formé (environ 1 $\mu$ m) et la taille des clusters (aux limites de peu Å à peu de dizaines de Å). Autrement dit, les forces moléculaires qui produisent la variété de motifs sont les memes que dans les systèmes que nous étudions, à l'exception du fait que souvent les micro-émulsions sont des systèmes au moins ternaires, alors nous avons des systèmes binaires. Pourtant, les expériences récentes avec les rayons X et les neutrons ont également été faites sur les systèmes binaires d'eau et d'alcool [28], avec pour but de detecter des analogies. Quand on s'approche de l'échelle moléculaire, il devient plus dur de définir bien les concepts géométriques de la courbure, qui sont important pour les micelles et les vésicules. Dès que ce point important est admis, l'importance de la présente analogie devient plus clair. Ici, nous pouvons parler d'un état fondamental pour l'auto-assemblage, quand il commence juste à se produire à l'échelle moléculaire. Cela offre de nouveaux défis auxquels faire face, tant expérimentalement que théoriquement.

## 8. Conclusion

Dans cette thèse, nous étudions l'organisation structurelle dans les liquides associés, les systèmes purs et mélangés, en utilisant les théories microscopiques comme la

physique statistique de liquides et les simulations de dynamique moléculaire classiques. Les systèmes étudiés sont de l'eau, alcohols, comme méthanol et tert-butanol et les mélanges d'alcool et d'eau. Nous distinguons entre l'organisation structurelle dans les liquides pur et celle dans les systèmes binaires, d'abord défini par l'association entre molécules de meme espèce, et ensuite celle du mélange qui suit celle conforme à la ségrégation locale des composantes du système binaire. Dans chaque cas, la liaison hydrogène hautement anisotropique joue le rôle principal, et pourtant sur de différents niveaux. Dans les liquides pur l'association se fait par des interactions spécifiques de entre sites et l'inhomogénéité résulte du clustering de sites. C'est cette caractéristique que nous avons appelé la micro-structuration [29] [30]. Dans les systèmes binaires, l'inhomogénéité est en raison d'une ségrégation locale des espèces et cela représente ce nous avons appelé la micro-hétérogénéité [31].

Les simulations numérique de mieux comprendre le comportement des système simulés à la nano-échelle dans l'espace et la pico-échelle en temps. C'est une façon unique d'étudier le comportement des liquides au niveau d'un microétat. Donc, la première partie de notre investigation devait évaluer la possibilité pour les simulations pour reproduire correctement la microstructure. La microstructure des systèmes est directement liée aux fonctions de distribution de paires de particule, car pour un système homogène, la fonction d'un corps -la densité- a une valeur constante. La connexion entre les données simulées et les propriétés de système réel a été faite par le calcul de l'intégrale de Kirkwood-Buff (KBI).

Le KBI peut être calculé tant de la simulation s'ensuit que le thermodynamique mesures. Nous avons évalué le comportement des systèmes purs et la reproductibilité du KBI et les fluctuations de concentration. Cette étude montre que les simulation des systèmes purs reproduit correctement la valeur de KBI (ou encore la compressibilité isothermal du système). Donc, la microstructure des systèmes simulés peuvent être raccordés à la microstructure dans les systèmes réels, au moins sur le niveau des quantités que nous exploré. Les différences entre le facteur de structure entre avec liaison hydrogènes et les sites hydrophobes montrent la différence dans l'organisation

locale de ces sites. A savoir, le pré-pic dans les facteurs de structure de sites H-bonded indique le clustering sur les tailles plus grandes que la distance des premiers voisins, tandis que les corrélations entre les sites de méthyles montrent seulement le pic principal, qui correspond à la structure d'empilement et indique une organisation homogène de ces sites non-liés.

Ces conclusions sont corroborées par le comportement des RDFs. L'utilisation de l'analogie de polymère en cas du méthanol et de l'analogie avec le système de micelle pour le cas du tert-butanol. De plus, le calcul de distribution de clusters permet de définir la taille la plus probable clusters. Dans le cas sites hydrophobes, aussi bien que le centre de la masse, les distributions de cluster montrent seulement des caractéristiques des distribution de particules homègenement réparties.

La simulation des melange d'eau et de méthanol montre des résultats encore plus intéressants. Le comportement aux petits-k indiquele clustering de l'eau et donc l'immiscibilité local des espèces. Cela est aussi montré aussi par la non-convergence du RKBIs pour les corrélations d'eau. Nous avons liés ces conclusions avec l'existence de la micro-hétérogénéité. Donc, le pré-pic dans les corrélations de sites indiquent les préférences structurelles des sites liés avec la liason hydrogène et les pics aux plus petits-k avec les corrélations d'espèces (tous les sites qui appartiennent à une meme espèce) indiquant le clustering des espèces.

Pourquoi insiter sur la MH? Nous avons analysé les données qui ont été accumulées pendant un temps d'échantillonnage petit (sur les 100 ps), pendant lequel nous avons essayé seulement quelques réalisations topologiques de l'agrégation. Par cette voie nous pourrions analyser l'évolution du MH par les pas de 100 ps. Cette dynamique de MH est vu par les variations du facteur de structure d'eau aux petit-k et aux travers d'é chantillonnages différents. Le petit temps d'é chantillonnage permet alors l'approche instantané du clustering des espèces, dans un sens que nous regardons le système sur une période courte comparée aux temps de relaxation véritables de la MH (que nous pensons être à l'echelle d'au moins 1 ns). Au contraire, le facteur de structure expérimental d'eau ne montre aucun grand variations dans la région des

petits-k. Donc nous attendons à ce que, avec l'augmentation des temps de simulation et surtout les tailles de systèmes simulés, les résultats de simulation s'approcheront de ceux expérimentaux. Quand même, en utilisant les arguments présentés ici, nous avons montré qu'il vaut la peine de faire ces calculs observer le système sur une echelle plus petite, qui révèle toute la richesse de la micro-structuration dans ces mélange.

L'idée de l'inhomogénéité locale est en dehors de la nécessité des grandes tailles de système et des grands temps de simulation. Par exemple, la pico-seconde est le temps de la relaxation du mouvement moléculaire, mais la réorganisation de MH, basé sur nos résultats de simulation, semble être dans la gamme de la nanoseconde. Puisque le système est dans l'état désordonné, nous sommes devant le problème de décrire une forme d'ordre local àl'intérieur du désordre global. C'est ce qui rend les solutions aqueuses fascinantes, malgré le fait que l'aspect physico-chimique-physique a été complètement étudié au siècle passé. Nous avons suggéré l'analogie entre la micro-émulsion et les mélanges de liquides associés: les deux systèmes sont àl'échelle globale homogène alors qu'ils sont localement hétérogènes, sauf que dans la micro-émulsion nous avons les motifs nettement définies comme les micelles, tandis que dans les mélanges de liquides associés, les grandeurs, les formes et les temps des agrégations ne sont pas bien définis.

L'analogie dérivée dans cette étude est intéressante d'un autre point de vue fondamental: il permet analyse très détaillée et explicite de la structure àune echelle microscopique, ce qui est presque impossible pour les micro-émulsions, oú la différence dans l'échelle entre le solvant, surfactant et la micro-structure est bien plus grand.

Les résultats obtenes dans cette thèse peuvent aider àla construction d'une approche théorique plus microscopique pour les liquides associés et leurs mélanges. C'est notre conviction qu'une telle approche pourrait avoir des résultats inattendus dans d'autres domaines de la physique de la matière condensé et de la physico-chimie.

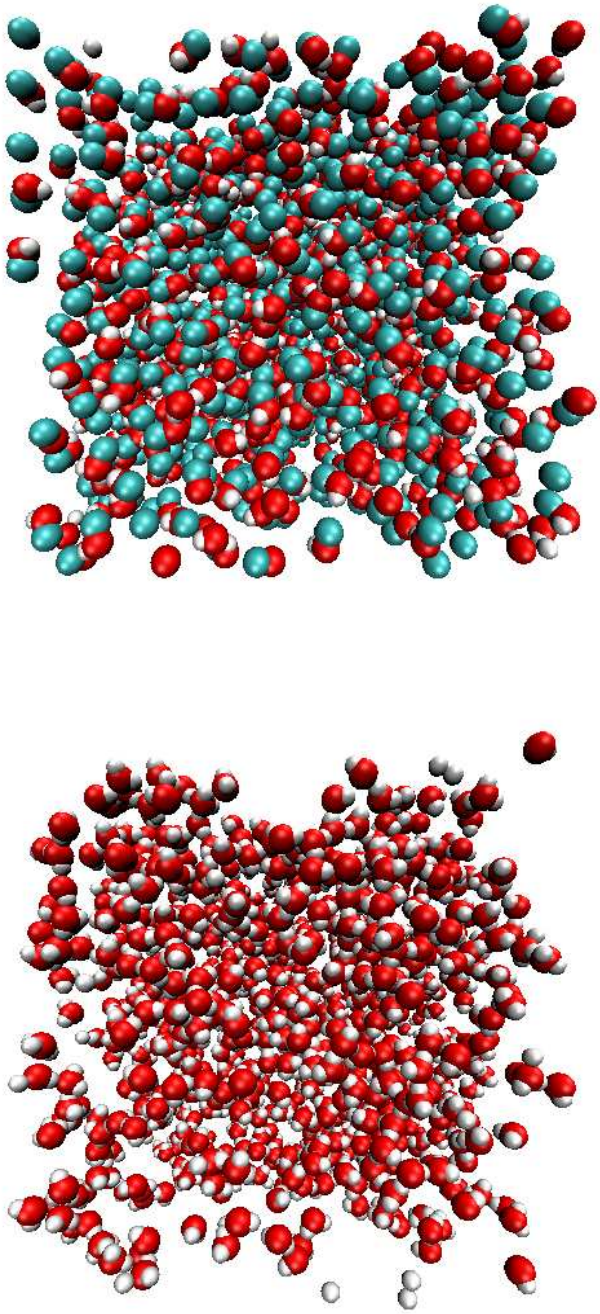

Figure 10: Les instantanés des simulations aux fractions molaires 0.5: panels haut méthanol; panels fond l'eau [32].

# Chapter 1

# Introduction

Water is the essential medium for the life processes and also one of the most investigated liquids. From the scientific and the philosophical point of view it holds a key to the beginning and the existence of life. Regardless of many theories that have been proposed to explain its anomalous features, water and aqueous solutions still hold many secrets and present hard problem for present day science.

The hydrogen bond is certainly the key to understanding the behavior of hydrogen-bonded liquids also called associated liquids. The strong directional hydrogen bonding induces the anisotropic connectivity between molecules and therefore the associated liquids tend to be locally more organized than ordinary liquids. Liquids are disordered, as opposed to solids, and a fundamental question for our understanding of liquid-systems is how the local heterogeneity can exist at the same time that the global homogeneity is preserved. Also, it is not clear how the local properties would get translated in the measurable average quantities or what are the macroscopic properties that are sensitive to this small scale behavior of the system. These are the questions that we will address in this thesis.

Our systems of interest are water-alcohol mixtures. The alcohols are smallest

amphipathic molecules. The amphipathicity comes from the opposite nature of their constituents, in this case the hydrophilic and hydrophobic part. Due to their asymmetrical preference for water in the aqueous mixture the amphipathic molecules form specific types of structural organization. The common explanation, for this structural organization, is hydrophobic effect that buries the hydrophobic part inside the core leaving the hydrophilic part in touch with the water molecules. The types of aggregates are more apparent for the cases of large amphipathic molecules, for example for surfactant (surface-active-agents) in the emulsions. Also the hydrophobic effect is commonly invoked in processes that involve large macromolecules such as folding of the proteins. However, keeping in mind these "large-scale behavior", we start from the simpler molecular system in order to study the features which on the microscopic scale have same cause as the above mentioned phenomena. Therefore, the main aim of this thesis is to investigate the structural organization in water-alcohol mixtures using the microscopic theories such as the statistical theory of liquids and classical computer simulations. The common opinion was that these liquids are disordered and isotropic, where alcohol and water, both hydrogen donors and acceptors, mix well. However, recent findings show that these constituents actually exhibit nano-scale immiscibility [21]. This local inhomogeneity of mixtures we will call micro-heterogeneity.

The local heterogeneities are well understood in the case of mixing-demixing phase transitions. What happens in case of demixing is that the system starts to develop strong concentration fluctuations, which lead to the final separation of the components. During this metastable state, the persisting large local heterogeneities appear and the system is no longer homogeneous. Therefore, the enhancement of concentration fluctuations are a signal of the instability of the system and this metastable

state of the system looks micro-heterogeneous.

The aqueous solutions show micro-heterogeneities in equilibrium state far away from the phase transitions. Namely, in the associated system, the hydrogen bonding stabilizes the local fluctuations and induces local ordering. Therefore the systems close to demixing and associated liquids both have features of micro-heterogeneity but for different reasons: the first one because of strong concentration fluctuations and the second due to a strong local interaction. This is the first important point of this thesis. In line with this we will distinguish the concentration fluctuation (CF) and micro-heterogeneity (MH): the first one that controls the phase behavior and the second one describes the local inhomogeneities in equilibrium systems.

One can also ask does a MH system have different order-properties than disordered liquids. There are few liquids that are not disordered systems such as nematic liquids which have orientational order and smetic liquids with both orientational and positional order. These liquids have global order that is achieved after liquid-liquid phase transition. MH systems have only one liquid phase, that has no global order. In other words, nano-scale order of MH system does not propagate on a larger scale, and these systems preserve the fundamental disordered nature of liquids.

Therefore, MH systems are more reminiscent of systems such as isotropic micro-emulsion (ME). MEs are also macroscopically in the homogeneous phase, but have a more "organized" local heterogeneity, as e.g. micelles. The quantitative difference between these two systems, as mentioned before is: micro-emulsions of binary systems occur for amphiphilic molecules with rather large oily tails [28], while the solutes that we consider here are smaller molecules such as simple alcohols and amides. Therefore, the comparison between these molecules and ME may seem far fetched, and one does

not expect any domains or micellar-like structure to be formed in such systems. However, in this thesis, we wish to show that the frontier between ME and MH systems is, in reality very diffuse, if it exists at all. First of all, from a thermodynamical point of view, ME are considered to be in the isotropic disordered phase, and the morphological changes that occur in them, such as micellar or bicontinuous structure, are still considered to be disordered phases [33]. Only in very specific conditions do ME form true ordered phases, such as lyotropic liquid-crystalline or lamellar phases, or when liquid-gas type phase separation occurs for micellar fluid itself [25]. So, in the isotropic state, ME are truly random mixtures, much like our systems that we consider here. Therefore, the frontier between ME and MH systems herein, appears to be somewhat blurred, and the differences must be essentially microscopic in nature, and not thermodynamical.

Clearly, the difference between CF and MH should help to characterize the difference between the ME and the binary mixtures we consider here. Concentration fluctuation is a well defined concept of statistical mechanics: it is related to the statistical averages defining species density fluctuations $\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle$ where $N_\alpha$ is the number of particles of species $\alpha$, and the average is performed in the Grand Canonical ensemble which allows for fluctuations in the number of particles. This quantity is related to the small wave vector limit $k \to 0$ of the structure factor $S_{\alpha\beta}(k)$, hence the long range behavior of the pair distribution function $g_{\alpha\beta}(r)$. We already mentioned that concentration fluctuations govern the stability of particular phase and have several well defined thermodynamical signatures, such as the divergence of a response function near spinodal line [34]. On the other hand the MH is a nouvelle concept, not yet well defined.

Can MH be viewed as a particular case of CF? Any system when viewed instantaneously is locally inhomogeneous for purely statistical reasons. The statistical average of such inhomogeneity is precisely due to concentration fluctuations. In the MH system the inhomogeneities will persist after averaging over all microstates. What is the structure that does not vanish in averaging? Particular case in ME are micelles and clearly this structure will leave a signature in statistically averaged values. For our mixture of smaller molecules, the micro-heterogeneities do not have well defined shapes. Also, the idea of persistence is ill-defined. The hydrogen bond interactions are responsible for MH: they bind molecules very directionally and over a short period of time (0.1 picosecond). Therefore, over a short period of time the local order will be destroyed and reformed again, and this frequency will also influence the time averaging.

The connection between this fact and the existence of MH is not clearly defined at present. The manifestation of micro-heterogeneities over thermodynamical quantities is also not very well characterized. Clearly the MH should influence also the thermodynamical properties of the system. The most recent example is the study of methanol-water solutions where the unusually small entropy of mixing is explained using concepts of local immiscibility of water and methanol [21]. Another example is also the large Kirkwood-Buff values reported for some specific aqueous mixtures. The Kirkwood-Buff value is directly connected to a concentration fluctuation, and its large value is indication of the enhancement of CF therefore the instability of phase. We would like to show that these systems are also influenced by the micro-heterogeneous structuring which enables such systems to be in a stable phase.

At present, it is not possible to disentangle the contributions from CF and MF to

any of the measurable quantities. However, the computational simulations provide unique tools to tackle the microscopic properties of systems. It is a unique way of having an open window to a small time- and space-scale behavior. In other words, computer simulations allow us to observe systems on the level of one microstate. This is especially important in our case since the problem of interest is the manifestation of the nano-scale order which at the present is not clear how or if it will influence the macroscopic properties. Therefore, we will use primarily the results of computational simulations to provide clues to clarify the role of micro-heterogeneities in aqueous solutions. However, computer simulations present new problems specific to the application of this method to micro-heterogeneous systems. Computational models, despite the imposed constraints, reproduce reliable results, as we will see when comparing the macroscopic observable such as enthalpies, volumes or densities. It is the properties such as excess values that are more sensitive to microscopic details where the deviation from real values are mostly observed. Another quantity, also very sensitive to simulation constraints, is the Kirkwood-Buff integral which corresponds to the integral of radial distribution function. The obvious reason is the limited number of particles as well as system sizes that are presently possible to simulate. Also, it is known that aqueous solutions are very "difficult" to simulate and the large system sizes as well as extensive time averaging is obligatory [8][35]. Spurious demixing of water-solute systems such as is a case for acetone-water mixture [36] were also reported. One may argue that the stability of aqueous solutions is governed both by the MH and CF. We need large system to accommodate the long-ranged correlation of micro-heterogeneity. Also, the MH system has a local structure on the time scale that is not yet well understood, it is clear that if we do not use a sufficiently long

simulation time, the true homogeneity of our system will not be reached. We will explore the variety of specific simulation results and try to address several problems mentioned above.

The organization of the doctorate is as follows. We start by the theoretical background of the statistical physics of liquids, following the theoretical framework of Kirkwood-Buff theory. The brief introduction of the molecular dynamics simulation and simulation models as well as the features of the DLPOLY package is given in the fourth chapter. The following chapters present the results of this thesis. In the fifth chapter the reliability of computer simulation in studying the microstructure of the liquid system is discussed. Brief introduction on the alcohols is in the sixth chapter. The results that consider the behavior of a neat system, as well as the definition of microstructure and the correlation functions that probe the local inhomogeneity are presented in the seventh chapter. The micro-heterogeneity of the alcohol-water mixture is described in the eight chapter. Finally, overall conclusion is given in the last chapter.

# Chapter 2

# Statistical physics of simple liquids

At room temperature the mixture of alcohol and water is in the liquid state. The liquid state of matter is intuitively perceived as being intermediate in nature between a gas and a solid. Its complexity can not be described using simple ideal models, as the ideal gas model is used for gasses or the model of harmonic oscillators for solid states. In this chapter, we give an outline of the statistical physics of liquids which provides a theoretical framework for the description of a liquid state.

## 2.1 The liquid state

A starting point for discussion of the properties of any given system is the relationship between pressure P, number density $\rho$ and temperature T in different phases, summarized in the equation od state $f(P, \rho, T) = 0$.

In the one-component phase diagram in $\rho - T$ plane typical of a simple liquid presented at figure 2.1, the region of existence of liquid is bounded above by the critical point and below by the triple point. Above the critical point there is only a single fluid phase. The coexistence curve separates liquid, gas and the solid states. The curve

inside the coexistence curve is called spinodal. The region bounded in-between the spinodal and coexistence curve is the region of meta-stable physical states. The states inside spinodal are unstable states, which are physically not allowed. The liquid-gas and liquid-solid transitions have discontinuous paths and pass over the region inside the coexistence curve. Similar diagrams are also valid for the mixing-demixing phase transitions.

Therefore, even if the computational simulation is the tool to study liquids, one should be very careful to put the simulated model in the conditions which correspond to a liquid state. Otherwise the simulated system can be in a fluid or glassy state or even in a physically unreal state such as states inside a coexistence curve. The first step, therefore, is to verify that all the physical and chemical parameters of the simulated system agree with those for a liquid state, otherwise the applicability of the statistical-physical-theory is put in question.

Here we will focus on the classical liquids, meaning we use classical equations and force fields to describe the behavior of a liquid state. Using the classical approximation, the contribution to thermodynamic properties which arises from thermal motion can be separated from those due to interactions between particles. The separation of kinetic and potential terms suggests a simple means of characterizing the liquid state. Let the $V_N$ be the total potential energy of a system and let $K_N$ be the total kinetic energy. Then in the liquid state we find that $K_N/|V_N| \approx 1$, where as $K_N/|V_N| \gg 1$ corresponds to the dilute gas and $K_N/|V_N| \ll 1$ to the low-temperature solid. In the liquid state the energy due to thermal motion is of the same order as the energy of the interaction between particles. Therefore, the structure of the liquid results from both of these contributions. In ideal gas atoms are non-interacting

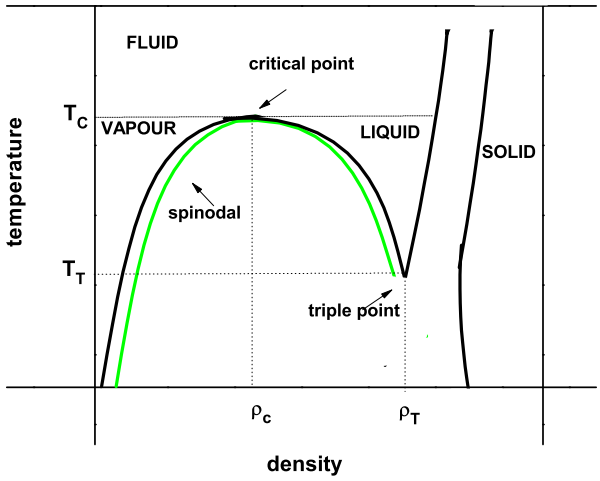


Figure 2.1: Schematic phase diagram of a typical monoatomic system like argon.

and thermodynamic quantities have only contributions arising from thermal motion. In the low-temperature solid the energy of thermal motion is small, the structure is rigidly defined mainly by the potential energy. Liquids and dense fluids are also distinguished from dilute gases by the greater importance of collision processes and short-range correlations, and from solids by the lack of long-range order. Nevertheless we will question the presence of long-range order in the associated liquids, and try to expand this oversimplified view of the liquid state.

In a simple monoatomic fluid such as argon $Ar$ or simple molecular system like nitrogen $N_2$ the correlations are short-ranged and the structure is mainly governed by the packing requirements. Associated liquids like water or alcohols have another level of complexity due to the attraction (or equally to the repulsion) of specific parts of the molecule. The structure of such systems is not defined only by the packing of molecule

but it combines several contributions such as attraction, repulsion and steric effects. These contributions result in a build-up of long-range correlations like a hydrogen-network in water, or towards the local ordering that we call micro-heterogeneity.

We will use a well established theoretical background of the theory of simple liquids. We will see how the phenomena such as micro-heterogeneity fit into this framework.

## 2.2 Statistical physics: introduction

Theoretical background is developed using classical statistical physics and the links between statistical mechanics and thermodynamics. We follow the outline given in the book by Hansen and McDonald, "Theory of simple liquids" [1].

Statistical physics provides a framework for relating the microscopic properties of individual atoms and molecules to the macroscopic or bulk properties of matter, therefore it explains thermodynamics as a natural result of statistics and mechanics (classical and quantum) at the microscopic level.

The dynamical state of a one-component monoatomic fluid at any instant is defined by the 3N coordinates $\mathbf{r}^N = \{\mathbf{r}_1, ..., \mathbf{r}_N\}$ and 3N momenta $\mathbf{p}^N = \{\mathbf{p}_1, ..., \mathbf{p}_N\}$ of the particles. The values of these 6N variables define a phase point in a 6N-dimensional phase-space. One phase-point corresponds to one state of a system. The Hamiltonian of the system in a absence of an external field can be written as:

$$H\left(\mathbf{r}^N, \mathbf{p}^N\right) = K_N\left(\mathbf{p}^N\right) + V_N\left(\mathbf{r}^N\right) \tag{2.1}$$

where kinetic energy $K_N\left(\mathbf{p}^N\right) = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m}$ and $V_N\left(\mathbf{r}^N\right)$ is the inter-particle potential energy. The motion of phase point along its phase trajectory is determined by

Hamilton's equations:

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} \quad \dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} \tag{2.2}$$

The trajectory of a system could be equally defined using Newton's equations. Given the coordinates and momenta of the particles at some instant, their values later (or earlier) in time can be in principle calculated as solution to Newton's equations of motions i.e. to a set of 3N coupled second-order, differential equations, which in absence of an external field, have the form:

$$m\mathbf{r}_i = -\nabla_i V_N\left(\mathbf{r}^N\right) \tag{2.3}$$

In statistical mechanics all observable properties of a system are calculated as averages over phase trajectories (the method of Boltzmann) or as averages over an ensemble of system (the method of Gibbs). In the Gibbs formulation the distribution of phase points of the ensemble is described by a phase-space probability density $f^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N; t\right)$:

$$\int\int f^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N; t\right) d\mathbf{r}^N d\mathbf{p}^N = 1, for\ all\ t. \tag{2.4}$$

A statistical ensemble is an arbitrary large collection of imaginary systems, each of which is a replica of the physical system of interest and characterized by the same macroscopic properties. The systems of the ensemble differ from each other in the microscopic realization of the coordinates and momenta of the particles. Given a complete knowledge of the probability density it would be possible to calculate the average value of any function of the coordinates and momenta. Usually we are not interested in a full phase-space probability, but only in the behavior of the subset of particles of size n, and redundant information can be eliminated by integration of $f^{[N]}$ over the coordinates and momenta of the other $(N-n)$ particles. We define a

reduced phase-space distribution function:

$$f^{[n]}\left(\mathbf{r}^n, \mathbf{p}^n; t\right) = \frac{N!}{(N-n)!} \int\int f^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N; t\right) d\mathbf{r}^{(N-n)} d\mathbf{p}^{(N-n)} \tag{2.5}$$

For simplicity we use index $N$ as particle number in further text. Let the equilibrium averages of the function $B\left(\mathbf{r}^N, \mathbf{p}^N\right)$ be $\langle B \rangle$. Then time average $\langle B \rangle_t$ for a given dynamical history is:

$$\langle B \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau B\left(\mathbf{r}^N(t), \mathbf{p}^N(t)\right) dt \tag{2.6}$$

and the equilibrium ensemble average $\langle B \rangle_e$ is defined as integral over equilibrium probability density $f_o^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N\right)$:

$$\langle B \rangle_e = \int\int B\left(\mathbf{r}^N, \mathbf{p}^N\right) f_o^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N\right) d\mathbf{r}^N d\mathbf{p}^N \tag{2.7}$$

The time average and the ensemble average are identical if the system is ergodic. The hypothesis of ergodicity is crucial to our comparison of the simulation and theoretical data. The ergodicity means that after a suitable lapse of time the phase trajectory of the system will have passed an equal number of times through every phase-space element defined by the probability density function. In the experiments, the measurements are usually done as time averages. Similarly, the results of molecular dynamics simulation are also time averages, but in the Monte Carlo modeling the calculation of mean values corresponds to the ensemble average.

The equilibrium probability density is defined by the macroscopic parameters that characterize the ensemble. The canonical ensemble describes systems that have constant volume V, number of particles N and temperature T and corresponding probability density is given by:

$$f_o^{[N]}\left(\mathbf{r}^N, \mathbf{p}^N\right) = \frac{1}{h^{3N} N!} \frac{e^{-\beta H}}{Q_N} \tag{2.8}$$

where h is Planck constant, $\beta = 1/(k_B T)$ where $k_B = 1.3810^{-23} J/K$ is the Boltzmann constant and the normalization constant $Q_N$ is the canonical partition function, defined as:

$$Q_N = \frac{1}{h^{3N} N!} \int \int e^{-\beta H} d\mathbf{r}^N, d\mathbf{p}^N \tag{2.9}$$

The inclusion of factor $\frac{1}{h^{3N}}$ ensures that we have dimensionless function and also for the consistency with the corresponding quantities of quantum statistical mechanics, while division by $N!$ ensures that microscopic states are correctly counted.

The thermodynamic potential appropriate to the $(N, V, T)$ constant condition is Helmholtz free energy F. The term potential is introduced because, the equilibrium at constant $(N, V, T)$ is reached when F is a minimum with respect to the variation of any internal parameter. The link between statistical mechanics and thermodynamics is established via relation between thermodynamical potential and the partition function:

$$F = -k_B T ln Q_N \tag{2.10}$$

Separation into the kinetic and potential terms in the manner of integration over all momenta in relation 2.9 allows the partition function to be rewritten as:

$$Q_N = \frac{1}{N!} \frac{Z_N}{\Lambda^{3N}} \tag{2.11}$$

where $\Lambda$ is the de Broglie wavelength $\Lambda = \sqrt{\frac{2\pi\beta\hbar^2}{m}}$ (m is the mass of an atom and $\hbar$ is the reduced Planck constant divided by $2\pi$) and configurational integral is:

$$Z_N = \int e^{-\beta V_N} d\mathbf{r}^N \tag{2.12}$$

For ideal gas approximation the potential $V_N$ is equal to 0 and configuration integral $Z_N = V^N$. Hence the partition function of the uniform ideal gas is:

$$Q_N^{id} = \frac{1}{N!} \frac{V^N}{\Lambda^{3N}} \tag{2.13}$$

Therefore, the partition function for the canonical ensemble can be written in the form:

$$Q_N = Q_N^{id} \frac{Z_N}{V^N} \tag{2.14}$$

Then, upon taking the logarithm of both sides, the Helmholtz free energy separates naturally into "ideal" and "excess" part:

$$F = F^{id} + F^{ex} = k_B T \left( ln\Lambda^3 \rho - 1 \right) - k_B T ln \frac{Z_N}{V^N} \tag{2.15}$$

where in calculation of the ideal part we used Stirling's approximation for $lnN!$.

The ideal part is the free energy of the system where the interactions between particles are small and therefore can be neglected and the excess part contains the contribution to the free energy that arises from interaction between particles. A similar division into ideal and excess parts can be made of any thermodynamical function obtained by differentiation of F with respect to either V or T. For example, the internal energy is equal to the sum of kinetic (ideal gas energy) and potential energy:

$$U = U^{id} + U^{ex} = \frac{3}{2} N k_B T + \frac{1}{Z_N} \int V_N e^{-\beta V_N} d\mathbf{r}^N \tag{2.16}$$

In the isothermal-isobaric ensemble, the fixed parameters are pressure P, temperature T, and number of particle N, while the thermodynamical potential that characterizes this system is the Gibbs free energy G:

$$G = F + PV = -k_B T ln \Delta_N \tag{2.17}$$

The link with thermodynamics as written in the relation 2.17, is made through the isothermal-isobaric partition function $\Delta_N$:

$$\Delta_N = \frac{\beta P}{h^{3N} N!} \int_0^\infty dV \int \int e^{-\beta(H+PV)} d\mathbf{r}^N d\mathbf{p}^N = \beta P \int_0^\infty e^{-\beta PV} Q_N dV \tag{2.18}$$

Similar conversion can be made between other different ensembles and it is defined by Laplace transform for partition functions $Q_A$ and Legendre transform for thermodynamic potential $\Psi_A$:

$$Q_{new} = \int Q_A e^{-\beta a A} dA \tag{2.19}$$

$$\Psi_{new} = \Psi_A + aA$$

where $a$ is an intensive variable such as pressure and temperature and $A$ is an extensive variable such as total energy and volume.

Our outline so far considers only closed systems, ie. uniform systems containing a fixed number of particles. The thermodynamic state of an 'open" system corresponds to a grand canonical ensemble and is defined by specifying the values of chemical potential $\mu$, volume V and temperature T. The thermodynamic potential of a grand canonical ensemble is grand potential $\Omega$:

$$\Omega = F - N\mu = -PV \tag{2.20}$$

The phase space of the grand canonical ensemble is the union of the phase space corresponding to all values of the variable N with constant T and $\mu$. The open system means that the equilibrium is achieved by exchanging both heat and matter with the surroundings. The equilibrium ensemble probability density is now a function of the number of particles N:

$$f_o\left(\mathbf{r}^N, \mathbf{p}^N; N\right) = \frac{e^{-\beta(H-N\mu)}}{\Xi} \tag{2.21}$$

where grand canonical partition function is:

$$\Xi = \sum_{N=0}^{\infty} \frac{e^{N\beta\mu}}{h^{3N} N!} \int \int e^{-\beta H} d\mathbf{r}^N d\mathbf{p}^N \tag{2.22}$$

Therefore the link with thermodynamics is established through the relation:

$$\Omega = -k_B T ln \Xi \tag{2.23}$$

and the ensemble average of microscopic variable $B\left(r^N, p^N\right)$ is:

$$\langle B\rangle_e = \sum_{N=0}^{\infty} \frac{1}{h^{3N}N!} \int\int B\left(\mathbf{r}^N, \mathbf{p}^N\right) f_o\left(\mathbf{r}^N, \mathbf{p}^N; N\right) d\mathbf{r}^N, d\mathbf{p}^N \tag{2.24}$$

For example, the average number of particles in the system is

$$\langle N\rangle = \sum_{N=0}^{\infty} Np\left(N\right) \tag{2.25}$$

where $p\left(N\right)$ is the probability that at equilibrium a system of the ensemble contains N particles irrespective of their coordinates and momenta:

$$p\left(N\right) = \frac{1}{h^{3N}N!} \int\int f_o\left(\mathbf{r}^N, \mathbf{p}^N; N\right) d\mathbf{r}^N, d\mathbf{p}^N \tag{2.26}$$

The more detailed discussion of grand canonical ensemble is in chapter 3.

## 2.3 Fluctuations

The definition of the ensemble means that we have some fixed parameters, while other thermodynamic parameters are allowed to fluctuate. A measure of the fluctuation about its average value $\langle B\rangle$ is provided by a mean-square deviation $\sigma_B^2 = \langle B^2\rangle - \langle B\rangle^2$. The fluctuations are related to intensive thermodynamic quantities that measure changes of one thermodynamic parameter with respect to another under a certain equilibrium condition. For example the internal energy for a canonical ensemble is obtained as:

$$U = \left\langle H\left(\mathbf{r}^N, \mathbf{p}^N\right)\right\rangle = \frac{1}{h^{3N}N!Q_N} \int\int He^{-\beta H} d\mathbf{r}^N, d\mathbf{p}^N \tag{2.27}$$

The heat capacity at a constant volume (the amount of energy that is given to a system to rise a temperature for one unit under the constant volume) is defined:

$$c_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{1}{k_B T^2}\left(\frac{\partial U}{\partial \beta}\right)_V \tag{2.28}$$

The derivation of equation 2.27 with respect to $\beta$ yields:

$$\left(\frac{\partial U}{\partial \beta}\right) = -\frac{1}{h^{3N}N!}\left[\frac{\int\int H^2 e^{-\beta H} d\mathbf{r}^N d\mathbf{p}^N}{Q_N} + \left(\frac{\int\int H e^{-\beta H} d\mathbf{r}^N d\mathbf{p}^N}{Q_N}\right)^2\right]$$

$$= \left\langle H^2 \right\rangle - \left\langle H \right\rangle^2 \tag{2.29}$$

Thus, the heat capacity at a constant volume is equal to the fluctuation of internal energy:

$$c_V = \frac{1}{k_B T}\left(\left\langle H^2 \right\rangle - \left\langle H \right\rangle^2\right) \tag{2.30}$$

Similarly, the isothermal compressibility $\chi_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$ correspond to the fluctuation in number of particle:

$$\frac{\left\langle N^2 \right\rangle - \left\langle N \right\rangle^2}{\left\langle N \right\rangle} = \rho k_B T \chi_T \tag{2.31}$$

The equations 2.30 and 2.31 and other fluctuation formulae can be also derived by purely thermodynamical arguments using the theory of fluctuations [1].

## 2.4 Particle densities and distribution functions

The most important quantities in our analysis are pair density and pair distribution function (as well as structure factor which corresponds to the pair distribution function in reciprocal space). Firstly, they can be calculated using the simulation data and also measured using the thermodynamic measurements or spectroscopic and diffraction experiments. Secondly, these functions are directly related to the microstructure of the systems and they contain the information of the correlations between particles in the system. We emphasize this point further in the following chapters.

The separation of the equilibrium phase-space probability density into the kinetic and potential terms leads to a division of thermodynamic properties into ideal and excess parts. A similar factorization can be made for reduced phase space distribution 2.5 for a canonical ensemble, which yields:

$$f_o^{(n)}\left(\mathbf{r}^n, \mathbf{p}^n\right) = \rho_N^{(n)}\left(\mathbf{r}^n\right) f_{Max.}^{(n)}\left(\mathbf{p}^n\right) \tag{2.32}$$

where

$$f_{Max.}^{(n)}\left(\mathbf{p}^n\right) = \frac{1}{\left(2\pi m k_B T\right)^{3n/2}} e^{-\beta \sum_{i=1}^{n} \frac{|\mathbf{p}_i|^2}{2m}} \tag{2.33}$$

is the product of n independent Maxwell distributions and the $\rho_N^{(n)}\left(\mathbf{r}^n\right)$ the equilibrium n-particle density:

$$\rho_N^{(n)}\left(\mathbf{r}^n\right) = \frac{N!}{(N-n)!} \frac{1}{h^{3N} N! Q_N} \int\int e^{-\beta H} d\mathbf{r}^{(N-n)} d\mathbf{p}^{(N-n)} = \frac{N!}{(N-n)!} \frac{1}{Z_N} \int e^{-\beta V_N} d\mathbf{r}^{(N-n)} \tag{2.34}$$

The equation 2.34 defines the probability of finding n particles of the system with coordinates in the volume element $d\mathbf{r}^n$, irrespective of the position of the remaining particles and irrespective of all momenta. The particle densities and the related, equilibrium particle–particle distribution functions, defined below, provide a complete description of the structure of a fluid. The knowledge of the low-order particle distribution functions in particular of the pair density $\rho_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right)$ is often sufficient to calculate the equation of state and other thermodynamic properties. The definition of the n-particle density follows:

$$\int \rho_N^{(n)}\left(\mathbf{r}^n\right) d\mathbf{r}^n = \frac{N!}{(N-n)!} \tag{2.35}$$

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

$$\rho_N^{(1)}\left(\mathbf{r}\right) = \frac{N}{V} = \rho \tag{2.36}$$

In the special case of a uniform, ideal gas the pair density is equal:

$$\rho_N^{(2)} = \rho^2 \left(1 - \frac{1}{N}\right) \tag{2.37}$$

The n-particles distribution function is defined in terms of the corresponding particle density by:

$$g_N^{(n)}(\mathbf{r}^n) = \frac{\rho_N^{(n)}(\mathbf{r}_1, ..., \mathbf{r}_n)}{\prod_{i=1}^n \rho_N^{(1)}(r_i)} \tag{2.38}$$

which for a homogenous system reduces to

$$\rho^n g_N^{(n)}(\mathbf{r}^n) = \rho_N^{(n)}(\mathbf{r}_1, ..., \mathbf{r}_n) \tag{2.39}$$

The particle distribution functions measure the extent to which the structure of a fluid deviates from complete randomness.

The particle densities are also expressible in terms of a $\delta$-function of position in a form that is very convenient for later purposes. From the definition of $\delta$-function it follows that:

$$\langle \delta(\mathbf{r} - \mathbf{r}_1) \rangle = \frac{1}{Z_N} \int \delta(\mathbf{r} - \mathbf{r}_1) e^{-\beta V_N(\mathbf{r}_1, ..., \mathbf{r}_N)} d\mathbf{r}^n = \frac{1}{Z_N} \int ... \int e^{-\beta V_N(\mathbf{r}, ..., \mathbf{r}_N)} d\mathbf{r}_2 ... d\mathbf{r}_n \tag{2.40}$$

The ensemble average in 2.40 is a function of the coordinates $\mathbf{r}$ but it is independent of the particle labels (here taken to be 1). The sum over all particles, therefore is equal to N times the contribution from any one particle. Comparison with the definition 2.34 then shows:

$$\rho_N^{(1)}(\mathbf{r}) = \left\langle \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle \tag{2.41}$$

which represents the ensemble average of a microscopic particle density $\rho(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$ where $\rho(\mathbf{r})$ is the density of one microstate of the system. Equally, the

microscopic density can be regarded as instantaneous density where $\rho(\mathbf{r})$ is the density of system in one instant of time, that corresponds to one configuration calculated in the simulation.

### 2.4.1 Pair distribution function

We will discuss in more details the pair distribution function defined as:

$$g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho_N^{(1)}(\mathbf{r}_1)\,\rho_N^{(1)}(\mathbf{r}_2)} \tag{2.42}$$

where similarly to equation 2.41 the pair density is expressible as the average of a product of two $\delta$-function:

$$\rho_N^{(2)}(\mathbf{r}, \mathbf{r'}) = \left\langle \sum_{\substack{i=1 \\ i\neq j}}^{N} \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i)\,\delta(\mathbf{r'} - \mathbf{r}_j) \right\rangle \tag{2.43}$$

If the system is homogenous and also isotropic, the pair distribution function $g_N^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ is function only of separation $r = |\mathbf{r}_2 - \mathbf{r}_1|$ and it is usually called radial distribution function, RDF. The $\delta$-function formalism also yields some useful expressions for the radial distribution function of the homogeneous and isotropic system:

$$\left\langle \frac{1}{N} \sum_{\substack{i=1 \\ i\neq j}}^{N} \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j + \mathbf{r}_i) \right\rangle = \left\langle \frac{1}{N} \int \sum_{\substack{i=1 \\ i\neq j}}^{N} \sum_{j=1}^{N} \delta(\mathbf{r'} - \mathbf{r} - \mathbf{r}_j)\,\delta(\mathbf{r'} - \mathbf{r}_i)\,d\mathbf{r'} \right\rangle \tag{2.44}$$

$$= \frac{1}{N} \int \rho_N^{(2)}(\mathbf{r'} + \mathbf{r}, \mathbf{r'})\,d\mathbf{r'} = \frac{\rho^2}{N} \int g_N^{(2)}(\mathbf{r}, \mathbf{r'})\,d\mathbf{r'} = \rho g(r)$$

The radial distribution function is a key quantity in the physics of liquids. The $g(r)$ is measurable by radiation-scattering experiments and also in indirect form by thermodynamic measurement as explained in the chapter 3. The $g(r)$ for liquid argon is pictured in figure 2.2.

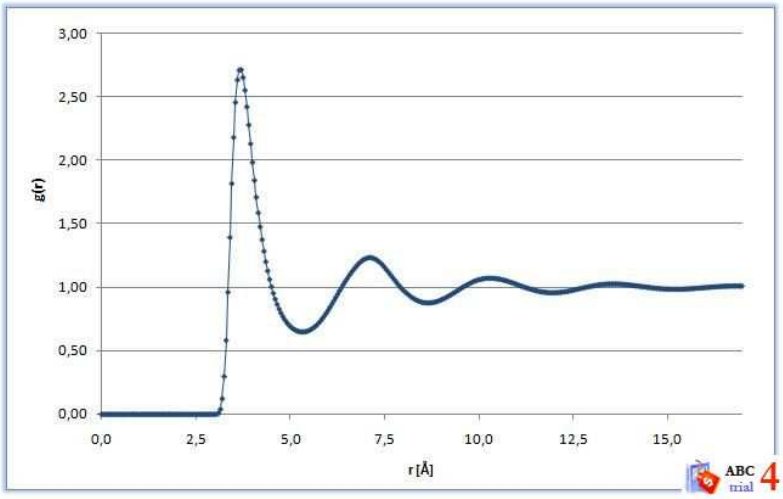


Figure 2.2: Radial distribution function of argon for system of N=864 particle at reduced temperature 0.833 (100 K) and reduced density 0.785 which corresponds to liquid state argon. Results from diploma thesis of M. Mijaković[37].

The $g(r)$ shows oscillation with the period that is approximately equal to a separation of atoms. This pattern is characteristic of all monoatomic liquids: the oscillation that indicates the packing of atoms. For large $r$, $g(r)$ tends to unity and vanishes as $r \to 0$ as a consequence of a strongly repulsive force that act at small particle separations. The definition of $g(r)$ implies that on the average the number of particles lying within the range $r$ to $r + dr$ from the reference particle is $4\pi\rho g\,(r)$ and peaks in $g\,(r)$ represent "shells" of neighbors around the reference particles. Integration of $4\pi r^2 \rho g\,(r)$ up to the position of the first minimum therefore provides an estimate of the "coordination number", CN. The concept of a shell of neighbors and coordination numbers is more appropriate for solids than for liquids, but it nevertheless provides a useful measure of the structure of a liquid. From the data for argon we can estimate that the coordination number is $CN \approx 12.2$, which shows that on the average each

atom of argon is surrounded by 12 neighbors.

Additionally, if the particles interact through the pairwise-additive forces, thermodynamic properties can be expressed in terms of integrals over $g(r)$. Consider a uniform fluid for which the total potential energy is given by the sum of pair interactions:

$$V_N\left(\mathbf{r}^N\right) = \sum_{i=1}^{N}\sum_{j>i}^{N} v\left(r_{ij}\right) \tag{2.45}$$

According to 2.16 and 2.45 the excess internal energy is equal:

$$U^{ex} = \frac{N\left(N-1\right)}{2}\int\int v\left(r_{12}\right)\left(\frac{1}{Z_N}\int ... \int e^{-\beta V_N} d\mathbf{r}_3 ... d\mathbf{r}_n\right) d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.46}$$

the double sum over i,j in 2.46 gives rise to $\frac{N(N-1)}{2}$ terms, each of which leads to the same result after integration. The definition of the pair density 2.34 and of the pair distribution function 2.42 allows 2.46 to be rewritten as:

$$U^{ex} = \frac{N^2}{2V^2}\int\int v\left(r_{12}\right) g_N^{(2)}\left(\mathbf{r}_1, \mathbf{r}_2\right) d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.47}$$

We can take the position of particle $\mathbf{r}_1$ as the origin of coordinates, set $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ and integrate over the coordinate $\mathbf{r}_1$ (which yields a factor V) to give:

$$U^{ex} = \frac{N^2}{2V^2}\int\int v\left(r_{12}\right) g\left(r_{12}\right) d\mathbf{r}_1 d\mathbf{r}_2 = \frac{N^2}{2V}\int v\left(r\right) g\left(r\right) d\mathbf{r} \tag{2.48}$$

or

$$\frac{U^{ex}}{N} = 2\pi\rho\int_0^\infty v\left(r\right) g\left(r\right) r^2 dr \tag{2.49}$$

Similarly, it is also possible to express the equation of pressure (called also the equation of state) as integral over $g(r)$:

$$\frac{P\beta}{\rho} = 1 - \frac{2\pi\rho\beta}{3}\int_0^\infty v'\left(r\right) g\left(r\right) r^3 dr \tag{2.50}$$

In the grand canonical ensemble the n-particle density is now a function of the probability $p(N)$ (relation 2.26):

$$\rho^{(n)}(\mathbf{r}^n) = \sum_{N \geqslant n}^{\infty} p(N)\, \rho_N^{(n)}(\mathbf{r}^n) \tag{2.51}$$

and the corresponding distribution function is:

$$g^{(n)}(\mathbf{r}^n) = \frac{\rho^{(n)}(\mathbf{r}_1, ..., \mathbf{r}_n)}{\prod_{i=1}^{n} \rho^{(1)}(\mathbf{r}_i)} \tag{2.52}$$

In a similar manner, as in equation 2.35 the integration over the coordinates yields for a pair density:

$$\int\int \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1 d\mathbf{r}_2 = \left\langle N^2 \right\rangle - \langle N \rangle \tag{2.53}$$

The energy equation and the pressure equation are also valid in the grand canonical ensemble. In the grand canonical ensemble we also can define compressibility equations, which express $\chi_T$ as an integral over $g(r)$. The normalization of equation 2.53 and corresponding equation for single-particle density yields:

$$\int\int \left[\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - \rho^{(1)}(\mathbf{r}_1)\, \rho^{(1)}(\mathbf{r}_2)\right] d\mathbf{r}_1 d\mathbf{r}_2 = \left\langle N^2 \right\rangle - \langle N \rangle - \langle N \rangle^2 \tag{2.54}$$

For the homogeneous system follows:

$$\langle N \rangle + \rho^2 \int\int \left[g^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - 1\right] d\mathbf{r}_1 d\mathbf{r}_2 = \left\langle N^2 \right\rangle - \langle N \rangle^2 \tag{2.55}$$

and by using the same procedure as in 2.44 and the definition for fluctuation in number of particles 2.31, it follows:

$$1 + \rho \int \left[g(r) - 1\right] dr = \frac{\left\langle N^2 \right\rangle - \langle N \rangle^2}{\langle N \rangle} = \rho k_B T \chi_T \tag{2.56}$$

We will broaden our discussion on the radial distribution function in the following chapters.

### 2.4.2 Structure factor

We explored also the radial distribution function in inverse space which is called the structure factor. Namely the structure factor is a measure of the density response of a system, initially in equilibrium, to a weak external perturbation of wavelength $2\pi/k$ [1]. For example in neutron scattering, $S(k)$ is proportional to the total scattered intensity in direction of a beam determined by the momentum transfer $h\mathbf{k}$ between beam and sample. This measurement enables means to determine radial distribution functions by use of the Fourier transformation. Namely the structure factor is more generally defined in terms of the Fourier transform of the pair correlation function:

$$S(\mathbf{k}) = \left\langle \frac{1}{N} \rho_k \rho_{-k} \right\rangle \tag{2.57}$$

where $\rho_k$ is a Fourier component of the microscopic density:

$$\rho_k = \int \rho(\mathbf{r}) e^{-i\mathbf{kr}} dr = \sum_{i=1}^{N} e^{-i\mathbf{kr_i}} \tag{2.58}$$

Using the $\delta$-function representation, the structure factor in the homogenous case implies:

$$\begin{aligned} S(\mathbf{k}) &= \left\langle \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} e^{-i\mathbf{kr_i}} e^{i\mathbf{kr_j}} \right\rangle = 1 + \left\langle \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i}^{N} e^{-i\mathbf{k}(\mathbf{r}_i - \mathbf{r}_j)} \right\rangle \\ &= 1 + \left\langle \frac{1}{N} \sum_{i=1}^{N} \sum_{j \neq i}^{N} e^{-i\mathbf{k}(\mathbf{r}-\mathbf{r}')} \delta(\mathbf{r} - \mathbf{r_i}) \delta(\mathbf{r}' - \mathbf{r}_j) d\mathbf{r} d\mathbf{r}' \right\rangle \\ &= 1 + \int \int e^{-ik(\mathbf{r}-\mathbf{r}')} \rho_N^{(2)}(\mathbf{r} - \mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ &= 1 + \rho \int g(r) e^{-i\mathbf{kr}} d\mathbf{r} \end{aligned} \tag{2.59}$$

Therefore the $g(r)$ is given by the inverse transform:

$$\rho g(\mathbf{r}) = (2\pi)^{-3} \int (S(\mathbf{k}) - 1) e^{-i\mathbf{kr}} d\mathbf{k} \tag{2.60}$$

## 2.5 Molecular distribution functions

The above description applied to an atomic system. The description of the structure of a homogeneous molecular fluid in terms of particle densities and distribution functions can be developed along similar lines. Let $R_i$ be the translational coordinates of molecule $i$ and let $\Omega_i$ be the orientation of $i$ in the laboratory-fixed frame of reference. If the molecule is linear, $\Omega_i = (\theta_i, \phi_i)$ where $\theta_i, \phi_i$ are usual polar angles; if it is non-linear, $\Omega_i = (\theta_i, \phi_i, \chi_i)$, where $\theta_i, \phi_i, \chi_i$ are the Euler angles. Then the generalized molecular pair density is defined as:

$$\rho^{(2)}(R, R', \Omega, \Omega') = \left\langle \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} \delta(R - R_i)\, \delta(R' - R_j)\, \delta(\Omega - \Omega_i)\, \delta(\Omega' - \Omega_j) \right\rangle \quad (2.61)$$

and the molecular pair distribution function as:

$$g(R_{12}, \Omega_1, \Omega_2) = \left(\frac{\Omega}{\rho}\right)^2 \rho^{(2)}(R_{12}, \Omega_1, \Omega_2) \quad (2.62)$$

where $\Omega \equiv \int d\Omega_i$. From the definition of $\Omega$ we get that $\Omega = \int\int d(\cos\theta_i)\, d\phi_i = 4\pi$ for linear case and $\Omega = \int\int d(\cos\theta_i)\, d\phi_i d\chi_i = 8\pi$ for non-linear case. The coordinates $R_i$ are often taken to be those of the molecular centre of mass or some other point of high symmetry in the molecule, but the choice of molecular centre is entirely arbitrary. To simplify the notation it is convenient to use symbol $i \equiv (R_i, \Omega_i)$ to denote both the coordinates of the molecular centre and the orientation of molecule $i$. Integration of the pair distribution function over variables $\Omega_i, \Omega_2$ yields a function:

$$g_c(R) = \frac{1}{\Omega^2} \int\int g(R_{12}, \Omega_1, \Omega_2)\, d\Omega_1 d\Omega_2 = \langle g(1,2) \rangle_{\Omega_1, \Omega_2} \quad (2.63)$$

When an interaction-site model is used to represent the intermolecular potential the natural way to describe the structure of the fluid is in terms of site-site distribution

functions. If the coordinates of site $\alpha$ on molecule $i$ are denoted by $r_{i\alpha}$ and those of site $\beta$ on molecule $j$ $(j \neq i)$ by $r_{j\beta}$, then the site-site radial distribution function, sRDF $g_{\alpha\beta}(r)$ is defined as:

$$g_{\alpha\beta}(r_{\alpha\beta}) = \frac{\langle \rho_{i\alpha}(r_{i\alpha})\,\rho_{j\beta}(r_{j\beta})\rangle}{\rho_\alpha \rho_\beta} \tag{2.64}$$

where the density $\rho_\alpha = N_\alpha/V$ is the number density of a site $\alpha$. Similarly, the site-site radial distribution function can be defined in terms of $\delta$-function:

$$\rho g_{\alpha\beta}(r) = \left\langle \frac{1}{N}\sum_{\substack{i=1\\ i\neq j}}^{N}\sum_{j=1}^{N}\delta(r + r_{2\beta} - r_{1\alpha}) \right\rangle = \langle\ (N-1)\delta(r + r_{2\beta} - r_{1\alpha})\rangle \tag{2.65}$$

The definition 2.65 can be used to relate the site-site distribution function to the molecular pair distribution function. Let $l_{i\alpha}$ be vector displacement of site $\alpha$ in molecule i from the molecular centre $R_i$ $(l_{i\alpha} = r_{i\alpha} - R_i)$ Then $g_{\alpha\beta}(r)$is given by the integral of $g(1,2)$ over all coordinates, subject to the constraint that the vector separation of sites $\alpha, \beta$ is equal to $r$:

$$\begin{aligned} g_{\alpha\beta}(r) &= \frac{1}{\Omega^2}\int\int\int\int dR_1 dR_2 d\Omega_1 d\Omega_2 g(1,2) * \delta\left[R_1 + l_{1\alpha}(\Omega_1)\right]\delta\left[R_2 + l_{2\beta}(\Omega_2) - r\right] \\ &= \frac{1}{\Omega^2}\int\int\int\int dR_1 dR_2 d\Omega_1 d\Omega_2 g(1,2) * \delta\left[R_{12} + l_{2\beta}(\Omega_2) - l_{1\alpha}(\Omega_1) - r\right] \end{aligned} \tag{2.66}$$

The site-site distribution functions have a simple physical interpretation. They are also directly related to the structure factors measured in x-ray and neutron-scattering experiments. On the other hand, the integration in 2.66 involves an irretrievable loss of information, and $g(1,2)$ cannot be reconstructed exactly from any finite set of site-site distribution functions. Many quantities that are expressible as integrals over $g(1,2)$ can also be written in terms of site-site distribution functions. If the intermolecular potential is of the interaction-site form and the site-site potential is

spherically symmetric, the excess internal energy is given:

$$\frac{U^{ex}}{N} = 2\pi\rho \sum_{\alpha} \sum_{\beta} \int_{0}^{\infty} v_{\alpha\beta}(r)\, g_{\alpha\beta}(r)\, r^2 dr \tag{2.67}$$

The equation of state can be determined by integration of compressibility equation 2.56. Because the choice of molecular centre is arbitrary, and need not be the same for each molecule, we can write:

$$\rho k_B T \chi_T = 1 + \rho \int \left[ g_{\alpha\beta}(r) - 1 \right] dr \tag{2.68}$$

where $\alpha, \beta$ refer to any pair of sites.

# Chapter 3

# Kirkwood-Buff theory

In this chapter, we give a presentation of the Kirkwood-Buff (KB) theory. It is a general statistical theory of solutions that connects the molecular distribution function in integrated form with macroscopic properties that could be measured using standard thermodynamic techniques. It is developed using the statistical physics and the theory of composition fluctuation in the grand canonical ensemble. In the last section we give several physical insights and critically address the experimental measurement of the Kirkwood-Buff integral.

## 3.1 Introduction

In 1951 Kirkwood and Buff [2] proposed a theory of concentration fluctuation in solutions. They have shown that the thermodynamic properties such as partial molar volumes, isothermal compressibility, and concentration derivatives of the chemical potential, could be expressed in terms of integrals of radial distribution functions of the several types of molecular pairs present in solutions. After a few decades Ben-Naim [3] rediscovered the power of this theory suggesting that the experimental measurements

of the latter properties provide information of the molecular distribution functions and therefore give an insight into the microstructure of a system. This especially comes in handy in exploring the results of the computational simulations, which can directly probe the microstructural properties.

The KB theory was introduced using the theory of the grand canonical ensemble. It is a very general theory valid for any kind of particles, not necessarily spherical and applicable to all types of intermolecular interaction. The grand canonical ensemble is an ensemble of an open system that has constant volume, temperature and chemical potential. We consider a multi-component open system that contains $\mathbf{N} = \{N_1, ..., N_\nu\}$ ($N = \sum_{i=1}^{\nu} N_i$) molecules of $\nu$-species in volume V at temperature T with chemical potential $\boldsymbol{\mu} = \{\mu_1, ...\mu_\nu\}$.

The chemical potentials are defined as the thermodynamic variables conjugate to the numbers of particles of each species and expressible as derivatives of any of the thermodynamic potentials. The chemical potential of species $\alpha$ is equal to (with abbreviation U internal energy, F Helmholtz free energy, G Gibbs energy and H enthalpy):

$$\mu_\alpha = \left(\frac{\delta U}{\delta N_\alpha}\right)_{V,S,\mathbf{N}'_\alpha} = \left(\frac{\delta F}{\delta N_\alpha}\right)_{V,T,\mathbf{N}'_\alpha} = \left(\frac{\delta G}{\delta N_\alpha}\right)_{P,T,\mathbf{N}'_\alpha} = \left(\frac{\delta H}{\delta N_\alpha}\right)_{T,S,\mathbf{N}'_\alpha} \tag{3.1}$$
$$\mathbf{N}'_\alpha = \{N_1, ..., N_\nu\}/N_\alpha$$

where $N_\alpha$ is number of particles of species $\alpha$.

The partial molar quantities are usually defined for a multi-component system being the derivatives with respect to a mole-number of each species. For example, the volume of a mixture, in general, is not equal to a sum of volumes occupied by the separate components (at the same pressure and temperature) before mixing. However, we can define the increase in volume on adding the infinitesimal amount of

species $i$ to a large amount of mixture of known composition. Then, the partial molar volume of species $i$ is equal to: $V_i = \frac{\partial V}{\partial n_i}$, where $n_i$ is mole-number. Total volume can be expressed in terms of partial molar volumes as: $V = \sum_{i=1}^{\nu} V_i n_i$ . Similar if we define partial quantities $V_{N_i}$ as derivatives with respect to the particle-number, then $V = \sum_{i=1}^{\nu} V_{N_i} N_i$. The partial molar properties, which are intensive properties and depend only on the pressure, temperature and composition of mixture, define how much of an extensive properties, in this case volume, is to be ascribed to each component. The function such as compressibility, heat capacity, coefficient of thermal expansion are also called response functions, namely the compressibility is the volume change as a response to a pressure change and heat capacity is the amount of energy required to raise system temperature by one unit.

Using separation into "ideal" and "excess" part, the Gibbs energy and chemical potential can be written as follows:

$$G = G^0 + k_B T \sum_{i=1}^{\nu} \rho_i ln \rho_i + G^{ex} \tag{3.2}$$

$$\mu_\alpha = \mu_\alpha^0 + k_B T ln \rho_\alpha + \mu_\alpha^{ex}$$

where the superscripts *0* and *ex* indicate the standard state and the excess from the ideal gas, respectively. Density $\rho_\alpha$ is number density of species $\alpha$.

For mixtures, excess quantities such as excess molar Gibbs energy $G_m^E$ or excess molar volume $V_m^E$ are often introduced by removing the linear mole fraction dependence in terms of the pure substance quantities, for example the excess molar enthalpy is equal to:

$$H_m^E(\boldsymbol{\chi}) = H_m(\boldsymbol{\chi}) - \sum_{i=1}^{\nu} \chi_i H_m(\chi_i = 1) \tag{3.3}$$

The $\boldsymbol{\chi} = \{\chi_1, ..., \chi_\nu\}$ and $H_m\left(\chi_i = 1\right)$ are molar fraction in mixture and molar enthalpy of pure system for species $i$. In the following chapters we will use the convention: $G^E$ as excess from ideally mixed solutions and the $G^{ex}$ as excess from the ideal-gas approximation.

## 3.2 Short presentation of Kirkwood-Buff theory

In this section, we will introduce the KB theory. The equations of the grand canonical ensemble are employed first to connect the pair correlation functions to fluctuation in number of particles, and second to relate composition fluctuation to derivatives of the chemical potential. Through these relations thermodynamic properties are linked to molecular distribution functions. Using the same annotation as in chapter 2 the number density of species $\alpha$ is:

$$\rho_\alpha^{(1)}\left(\mathbf{r}\right) = \left\langle \rho_\alpha^{(1)}\left(\mathbf{r}\right)\right\rangle = \frac{\sum\limits_{N_1,...N_\nu=0}^{\infty} e^{\frac{\boldsymbol{\mu}\mathbf{N}}{k_B T}} \int d\mathbf{r}_1 ... d\mathbf{r}_N \sum\limits_{i_\alpha=1} N_\alpha \delta\left(\mathbf{r} - \mathbf{r}_{i_\alpha}\right) Q_{\mathbf{N}}\left(V,T\right)}{\Xi\left(\boldsymbol{\mu},V,T\right)}, \quad (3.4)$$

where the grand canonical partition function is equal to

$$\Xi\left(\boldsymbol{\mu},V,T\right) = \sum_{N_1,...N_\nu=0}^{\infty} e^{\frac{\boldsymbol{\mu}\mathbf{N}}{k_B T}} Q_{\mathbf{N}}\left(V,T\right). \quad (3.5)$$

Similarly, the pair density of species $\alpha$ and $\beta$ is (same annotation was used for a site-site correlation in chapter 2):

$$\rho_{\alpha\beta}^{(2)}\left(\mathbf{r},\mathbf{r}'\right) = \left\langle \rho_\alpha^{(1)}\left(\mathbf{r}\right) \cdot \rho_\beta^{(1)}\left(\mathbf{r}'\right)\right\rangle = \left\langle \sum_{i_\alpha=1}^{N_\alpha} \sum_{j_\beta=1}^{N_\beta} \delta\left(\mathbf{r} - \mathbf{r}_{i_\alpha}\right) \cdot \delta\left(\mathbf{r}' - \mathbf{r}_{j_\beta}\right)\right\rangle \quad (3.6)$$

These densities posses by nature of their definition the following integrals:

$$\int d\mathbf{r} \rho^{(1)}\left(\mathbf{r}\right) = \left\langle \mathbf{N}_\alpha \right\rangle \quad (3.7)$$

$$\int\int d\mathbf{r}'d\mathbf{r}\rho^{(2)}_{\alpha\beta}\left(\mathbf{r},\mathbf{r}'\right)=\left\langle N_\alpha N_\beta\right\rangle-\delta_{\alpha\beta}\left\langle N_\alpha\right\rangle$$

Combining the above relations we obtain expression:

$$\int d\mathbf{r}'d\mathbf{r}\left[\rho^{(2)}_{\alpha\beta}\left(\mathbf{r},\mathbf{r}'\right)-\rho^{(1)}_{\alpha}\left(\mathbf{r}\right)\rho^{(1)}_{\beta}\left(\mathbf{r}'\right)\right]=\left\langle N_\alpha N_\beta\right\rangle-\left\langle N_\alpha\right\rangle\left\langle N_\beta\right\rangle-\delta_{\alpha\beta}\left\langle N_\alpha\right\rangle \tag{3.8}$$

Using the definition for pair distribution function, that correspond to the correlation between species $\alpha$ and $\beta$, $g^{(2)}_{\alpha\beta}\left(\mathbf{r},\mathbf{r}'\right)=\frac{\rho^{(2)}_{\alpha\beta}(\mathbf{r},\mathbf{r}')}{\rho^{(1)}_{\alpha}(\mathbf{r})\rho^{(1)}_{\beta}(\mathbf{r}')}$ and relation 3.8 the integrated form of pair distribution function can be expressed as function of particle number fluctuation:

$$\int\left(g^{(2)}_{\alpha\beta}\left(r\right)-1\right)d\mathbf{r}=V\left[\frac{\left\langle N_\alpha N_\beta\right\rangle-\left\langle N_\alpha\right\rangle\left\langle N_\beta\right\rangle}{\left\langle N_\alpha\right\rangle\left\langle N_\beta\right\rangle}-\frac{\delta_{\alpha\beta}}{\left\langle N_\alpha\right\rangle}\right] \tag{3.9}$$

This relation is valid for homogenous and isotropic systems, where the pair distribution function is a function of only the separation $r=|\mathbf{r}-\mathbf{r}'|$. Namely, the particle number fluctuations are related to the radial part of the total pair distribution function, the radial distribution function $g^{(2)}_{\alpha\beta}\left(r\right)$. We introduce the Kirkwood-Buff integral (KBI) of species $\alpha$ and $\beta$ as:

$$G_{\alpha\beta}=4\pi\int_0^\infty\left(g^{(2)}_{\alpha\beta}\left(r\right)-1\right)r^2dr \tag{3.10}$$

Therefore, the KBIs can be also defined as a function of the fluctuation in the particle number:

$$G_{\alpha\beta}=V\left[\frac{\left\langle N_\alpha N_\beta\right\rangle-\left\langle N_\alpha\right\rangle\left\langle N_\beta\right\rangle}{\left\langle N_\alpha\right\rangle\left\langle N_\beta\right\rangle}-\frac{\delta_{\alpha\beta}}{\left\langle N_\alpha\right\rangle}\right] \tag{3.11}$$

Next one needs to establish the relation between the composition fluctuation and the chemical potential derivatives. We start from the ensemble average for the number of species $\alpha$ defined through the grand canonical probability density:

$$\left\langle N_\alpha\right\rangle=\frac{\sum_{N_1,\ldots N_\nu=0}^{\infty}e^{\frac{\boldsymbol{\mu}\mathbf{N}}{k_BT}}Q_{\mathbf{N}}\left(V,T\right)}{\Xi\left(\boldsymbol{\mu},V,T\right)}=k_BT\left(\frac{\partial ln\Xi\left(\boldsymbol{\mu},V,T\right)}{\partial\mu_\alpha}\right)_{T,V,\boldsymbol{\mu}'_\alpha} \tag{3.12}$$

$$\boldsymbol{\mu}_{\alpha}^{'} = \{\mu_1, ..., \mu_\nu\} / \mu_\alpha$$

Simple derivation of the relation 3.12 together with the definition of ensemble averages and the symmetry arguments gives the relations:

$$k_B T \left( \frac{\partial \langle N_\alpha \rangle}{\partial \mu_\beta} \right)_{T,V,\boldsymbol{\mu}_{\boldsymbol{\beta}}^{'}} = k_B T \left( \frac{\partial \langle N_\beta \rangle}{\partial \mu_\alpha} \right)_{T,V,\boldsymbol{\mu}_{\boldsymbol{\alpha}}^{'}} = \langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle \tag{3.13}$$

To simplify the calculations, the matrices $\mathbf{B}$ and $\mathbf{A}$ are introduced:

$$B_{\alpha\beta} = \frac{k_B T}{V} \left( \frac{\partial \langle N_\alpha \rangle}{\partial \mu_\beta} \right)_{T,V,\boldsymbol{\mu}_{\boldsymbol{\beta}}^{'}} = k_B T \left( \frac{\partial \left\langle \rho_\alpha^{(1)} \right\rangle}{\partial \mu_\beta} \right)_{T,V,\boldsymbol{\mu}_{\boldsymbol{\beta}}^{'}} \tag{3.14}$$

$$A_{\alpha\beta} = \frac{V}{k_B T} \left( \frac{\partial \mu_\beta}{\partial \langle N_\alpha \rangle} \right)_{T,V,\mathbf{N}_\beta^{'}} = \frac{1}{k_B T} \left( \frac{\partial \mu_\beta}{\partial \left\langle \rho_\alpha^{(1)} \right\rangle} \right)_{T,\boldsymbol{\rho}_{\boldsymbol{\beta}}^{'}}$$

We notice that from the definition of these matrices it follows $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$ where $\mathbf{I}$ is unit-matrix, therefore

$$A_{\alpha\beta} = \frac{|\mathbf{B}|^{\alpha\beta}}{|\mathbf{B}|} \tag{3.15}$$

where $|\mathbf{B}|^{\alpha\beta}$ is cofactor in the determinant $|\mathbf{B}|$.

It is through the definition of the matrix $\mathbf{B}$ and relation 3.13, that we can express the elements of matrices $\mathbf{B}$ in terms of particle number fluctuation, and therefore introducing the definition 3.11 the $B_{\alpha\beta}$ are equal:

$$B_{\alpha\beta} = G_{\alpha\beta} \frac{\langle N_\alpha \rangle \langle N_\beta \rangle}{V^2} - \delta_{\alpha\beta} \frac{\langle N_\alpha \rangle}{V} \tag{3.16}$$

Usually liquid solutions are considered as homogeneous and isotropic and their average density is equal to a constant number, therefore $\rho_\alpha^{(1)}(\mathbf{r}) = \frac{\langle N_\alpha \rangle}{V} = c_\alpha$ where $c_\alpha$ is bulk molecular concentration of species $\alpha$. Then the elements of the matrix $\mathbf{B}$ can be expressed in terms of concentrations:

$$B_{\alpha\beta} = c_\alpha \delta_{_{\alpha\beta}} + c_\alpha c_\beta G_{\alpha\beta} \tag{3.17}$$

In order to obtain expressions for the variables measured at the constant pressure we employ the relation that connects the derivatives for (T,V)-constant and (T,P)-constant ensemble:

$$\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,V,\mathbf{N}'_\alpha}=\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,P,\mathbf{N}'_\alpha}+\left(\frac{\partial\mu_\beta}{\partial P}\right)_{T,P,\mathbf{N}}\left(\frac{\partial P}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,V,\mathbf{N}'_\alpha} \tag{3.18}$$

Using the identity $\left(\frac{\partial P}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,V,\mathbf{N}'_\alpha}\left(\frac{\partial\left\langle N_\alpha\right\rangle}{\partial V}\right)_{T,P,\mathbf{N}}\left(\frac{\partial V}{\partial P}\right)_{T,V,\mathbf{N}}=-1$ and expressions for partial molar volume per particle and thermal compressibility $V_{N_\alpha}=\left(\frac{\partial V}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,V,\mathbf{N}'_\alpha}=\left(\frac{\partial\mu_\alpha}{\partial P}\right)_{T,\mathbf{N}}$ and $\chi_T=-\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,\mathbf{N}}$ we have:

$$\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,P,\mathbf{N}'_\alpha}=\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,V,\mathbf{N}'_\alpha}-\frac{V_{N_\alpha}V_{N_\beta}}{V\chi_T} \tag{3.19}$$

The right part of the equation 3.19 is linked to the $A_{\alpha\beta}$ and through the definition 3.15 it is related to the elements of the matrix $\mathbf{B}$. Therefore, the link with thermodynamics is established combining the relation 3.14 and 3.19:

$$\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,P,\mathbf{N}'_\alpha}=\frac{|\mathbf{B}|^{\alpha\beta}}{|\mathbf{B}|}-\frac{V_{N_\alpha}V_{N_\beta}}{V\chi_T} \tag{3.20}$$

Using straightforward mathematical transformation of expression 3.20 and a well-known Gibbs-Duhem equality $\sum_{i=1}^{\nu}N_i\left(\frac{\partial\mu_i}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,P,\mathbf{N}'_\alpha}=0$, one can obtain relations:

$$\begin{gathered}k_BT\chi_T=\frac{|\mathbf{B}|}{\sum_{\alpha,\beta}^{\nu}|\mathbf{B}|^{\alpha\beta}}\\ V_{N_\alpha}=\frac{\sum_{\beta}^{\nu}|\mathbf{B}|^{\alpha\beta}}{\sum_{\beta,\gamma}^{\nu}|\mathbf{B}|^{\beta\gamma}}\\ \frac{V}{k_BT}\left(\frac{\partial\mu_\beta}{\partial\left\langle N_\alpha\right\rangle}\right)_{T,P,\mathbf{N}'_\alpha}=|\mathbf{B}|^{\alpha\beta}-\frac{\sum_{\gamma}^{\nu}c_\gamma|\mathbf{B}|^{\gamma\beta}\sum_{\gamma}^{\nu}c_\gamma|\mathbf{B}|^{\gamma\alpha}}{\sum_{\gamma,\sigma}^{\nu}c_\gamma c_\sigma|\mathbf{B}|^{\gamma\sigma}}\end{gathered} \tag{3.21}$$

The mathematical inversion of the equation 3.21 together with the definition 3.17 gives the inversion of the Kirkwood-Buff theory that are the relations that express the molecular distribution function in terms of the thermodynamical properties measured at constant pressure and temperature. We also notice that relation 3.21 connects the number fluctuation that are hidden in the **B**-defined terms and the thermodynamic variables. The equations valid for a binary system are given in the following paragraph, but the similar relation can be derived for any multi-component system. Mathematical inversion of relations 3.21 gives the KBI of species *1* and *2*:

$$\begin{aligned} G_{12} &= k_B T \chi_T - \frac{V_{N_1} V_{N_2}}{VD} \\ G_{11} &= G_{12} + \left( \frac{V_{N_2}}{D} - V \right) \frac{1}{N_1} \\ G_{22} &= G_{12} + \left( \frac{V_{N_1}}{D} - V \right) \frac{1}{N_2} \end{aligned} \tag{3.22}$$

Where D depends on the derivatives of the chemical potential:

$$D = \frac{N_1}{N_2} \left( \frac{\partial \beta \mu_1}{\partial N_1} \right)_{P,T,N_2} = \frac{N_2}{N_1} \left( \frac{\partial \beta \mu_2}{\partial N_2} \right)_{P,T,N_1} = - \left( \frac{\partial \beta \mu_1}{\partial N_2} \right)_{P,T,N_1} = - \left( \frac{\partial \beta \mu_2}{\partial N_1} \right)_{P,T,N_2}$$

It is convenient to express KBIs in terms of mol-fraction derivatives and in $G_{ij}$ $\left[ \frac{cm^3}{mol} \right]$ units:

$$G_{ij} = -G_{12} \delta_{ij} + (1 - \delta_{ij}) \left( RT \chi_T - \frac{V_1 V_2}{V_m \widetilde{D}} \right) + \frac{\delta_{ij}}{\chi_i} \left( \frac{V_i}{\widetilde{D}} - V_m \right) \tag{3.23}$$

where R stands for the gas constant $R = 8.3145 JK^{-1} mol^{-1}$ and $\widetilde{D} = \chi_i \dfrac{\partial \beta \mu_i}{\partial \chi_i} = DN$.

## 3.3 The experimental measurement of Kirkwood-Buff integral

The KBIs are extremely sensitive to the experimental procedure or to the numerical treatment of the data [38][39][40]. The most critical parameter is the chemical potential, which is connected with the energy of the insertion of a solute particle in the mixture. This insertion energy will strongly depend, not only on the direct interaction between the particles, but also on the local organization of the mixture [39]. For liquid solutions, the isothermal compressibility is usually small when compared to the other terms of the expression 3.23. The partial molar volumes can be obtained from the total molar volume by the slope intercept method or by differentiation of the fitted volume. The excess chemical potential is usually calculated using the measurements for partial pressure. The partial pressures $p_i$ are commonly determined by measuring the total pressure $P$, with the liquid mixture of known composition, and converting the $(P, \chi_i)$ data set into $(p_i, \chi_i)$ by the method due to Barker [41] whereby an analytic function for the excess Gibbs energy of the mixture is curve-fitted to obtain the partial pressures. The overall accuracy of KBI strongly depends on the applicability of the fitting procedure [39].

Apart from thermodynamics, scattering techniques also provide a way to obtain concentration fluctuations through the structure factor at zero wave vector (see chapter 2). Following Bhatia and Thornton [42] one can relate density fluctuations to concentration fluctuations through the concentration-concentration structure factor. They showed that the scattering function is generally expressible in terms of three structure factors $S_{\chi\chi}(q)$, $S_{NN}(q)$ and $S_{N\chi}(q)$ constructed from the Fourier transforms

of the local number density and concentration. These structure factors have the property that at temperatures above the Debye temperature and in the long-wavelength limit ($q \to 0$) $S_{\chi\chi}(q)$ and $S_{NN}(q)$ represent, respectively, the mean square thermal fluctuations in the particle number and concentration, and $S_{N\chi}(q)$ the correlation between these two fluctuations:

$$\begin{aligned} S_{NN}(0) &= \frac{\left\langle (\triangle N)^2 \right\rangle}{N} = \frac{N}{V} k_B T \chi_T + \delta^2 S_{\chi\chi}(0) \\ S_{\chi\chi}(0) &= N \left\langle (\triangle \chi)^2 \right\rangle = \frac{N k_B T}{\left(\frac{\partial^2 G}{\partial \chi^2}\right)_{T,P,N}} \\ S_{N\chi}(0) &= \langle \triangle N \triangle \chi \rangle = -\delta S_{\chi\chi}(0) \end{aligned} \tag{3.24}$$

where $\left\langle (\triangle \chi)^2 \right\rangle$ is the mean square fluctuation, $\triangle \chi$ is defined as $\triangle \chi = \frac{[(1-\chi)\triangle N_1 - \chi \triangle N_2]}{N}$ for binary system with the molar fractions $\chi = \frac{N_1}{N} = \chi_1$ and $1 - \chi = \frac{N_2}{N} = \chi_2$ and $\triangle N = \triangle N_1 + \triangle N_2$. Factor $\delta = \frac{V_{N_1} - V_{N_2}}{\chi V_{N_1} + (1-\chi) V_{N_2}} = \frac{N}{V}(V_{N_1} - V_{N_2})$ is a dilatation factor and $V_{N_i}$ are partial molar volumes per particle.

The mole fraction structure factor $S_{\chi\chi}(q)$ is simply related to a linear combination of the site-site structure factors $S_{ij}$:

$$S_{\chi\chi}(q) = \chi_1 \chi_2 \left[1 + \chi_1 \chi_2 \left(S_{11}(q) + S_{22}(q) - 2S_{12}(q)\right)\right] \tag{3.25}$$

Similar, relations for density-density and cross structure factors are:

$$S_{NN}(q) = \chi_1^2 S_{11}(q) + \chi_2^2 S_{22}(q) + 2\chi_1 \chi_2 S_{12}(q) \tag{3.26}$$

$$S_{N\chi}(q) = \chi_1 \chi_2 \left[\chi_1 \left(S_{11}(q) - S_{12}(q)\right) - \chi_2 \left(S_{22}(q) - 2S(q)\right)\right]$$

Above relations provide the calculational background for experimental evaluation of the KBIs. It is interesting to note that scattering and thermodynamic experiments probe quantities that are inverse of each other. This is already an internal source of

discrepancy between the two results, since small errors in one representation will be magnified in the other. Scattering experiments probe directly the concentration fluctuations, while thermodynamic measurements probe indirectly variations in chemical potentials [39].

Matteolli and Lepori [38] were among the first to provide extensive measurements of KBIs for various aqueous mixtures. Their data show a rich display of behavior that indicates the complexity of aqueous solutions. Nishikawa [43] provided the first X-ray scattering measurements of KBI. The KBIs were also exploited for reproducing the more accurate force field (see chapter 4). However, there is a considerable variation in the published results [38][44][40]. Possible sources of inaccuracies, as discussed in Perera et al.[39], could have experimental origins and also could come from numerical treatment of the data.

# Chapter 4

# Molecular dynamics

In recent years, existence of micro-heterogeneous organization has become more and more widely recognized in the description of the aqueous solution. To study it we need a reliable tool that will enable us to peer into the microstructure of our solutions. This is where the computer simulation comes in handy. These simulations can provide many sets of configurations which allow a direct insight into the micro-structure of the simulated system.

In this chapter, we give a brief introduction to computation methods such as molecular dynamics, and an outline of the recent models of the molecular system of our concern. At the end, we will describe the DL-POLY package that has been used in this thesis.

## 4.1 Computer simulations

Computer simulations are one of the most important tools for studying liquid systems due to the fact that we lack a simple model that could describe the liquid state. From the standpoint of the liquid-state theory, computer simulation provide almost

exact, quasi-experimental data on well-defined models, particularly on those that are prototypical models of simple liquids [1]. Two numerical methods widely used are: Molecular Dynamics method (MD) and Monte Carlo method (MC) [4][5].

The MD simulation is based on the calculation of the time dependent set of configurations using classical equations of motion. The configuration of system of N-particles implies the 3N coordinates $\mathbf{r}^N = \{\mathbf{r}_1, ..., \mathbf{r}_N\}$ and 3N momenta $\mathbf{p}^N = \{\mathbf{p}_1, ..., \mathbf{p}_N\}$ of the particles at some instant of time. From the initial configuration, the trajectory of the particles, or next configuration is calculated using the 3N coupled, second order differential equations 2.3.

The initial configuration contains the N particle-coordinates within a simulation cell of volume V, usually a cubic box and assigned velocities usually from a Maxwell distribution appropriate to the temperature of interest and selected in such a way that the total linear momentum is equal to zero. The boundary conditions are regulated using periodically repeated cells which surround the main cell. In mirror-cells the position and velocity of each particle is periodically repeated and this convention is called periodic boundary condition (PBC).

The interaction between particles is usually a pairwise and additive. The computation of short-range interaction is calculated up to certain cutoff distance and an analytic expression (see equation 4.2) is used to estimate the contributions beyond this cutoff distance. This truncated potential is usually calculated in a sphere radii half-cell size. The interactions are computed using the technique called the "nearest-neighbor" convention [1]. The method such as Ewald summations is applied to calculate the long-range interaction such as the Coulomb interaction. The time-evolution of the system is calculated for each chosen time step, using the Verlet or

related algorithms. The properties of the simulated system, than, are computed as the time averages over a large number of configurations that are collected after the system has reached the equilibrium state.

The Monte Carlo simulation starts also with an initial set of coordinates within periodic boundary conditions, that interacts through some known potential. A sequence of configurations is then generated by successive random displacements of the particles. The new configurations are accepted in such a way that configurational space is asymptotically sampled according to the probability density corresponding to a particular statistical mechanical ensemble [1]. The MC method is used to calculate only static properties, and it essentially represents the Boltzmann averaging over the microstates of the systems.

Computer simulation turns out to be an extremely powerful tool, since it allows one to perform cheap statistical experiments for a small number of particles, in nearly experimental conditions-at least this is the credo that prevails among simulators, and which has been confirmed for the case of simple liquids. Nevertheless, one should be aware of all approximations that are inherent to the simulations. Strictly, simulations are used to validate statistical theories and computational models, and are confirmed for the description of real systems only by the validation of proposed models through comparison with real-system properties. For example, in the simulation, in general we have constant-number ensembles and the pseudo-infinite systems introduced through periodic boundary condition. In experiments we are dealing with an open system with a number of particles of the order of Avogadro's number. Usually, force-field models are defined through pair interactions, for example, the hydrogen-bonding is modeled by the site-site Coulomb interaction and the sterical contributions by Lennard-Jones

potential. This is an oversimplified version of all interaction present in a real system, since it neglects the high-order interaction or polarization effects. Nevertheless, these models have been proven to reproduce correctly the thermodynamical properties for large number of real systems. In addition, the simulation of aqueous solutions imposes problems of its own. Empirically, it has been found that associated systems are very sluggish to equilibrate [23][35]. The spurious demixing of acetone-water solutions were reported by Perera and Sokolić [36]. Therefore the simulations of aqueous mixtures represent another level of complexity which has not yet been properly addressed.

## 4.2 Computational models

The reliability of the computational simulations depends strongly on the models, which means basically the force fields between components. These force fields are expressed in terms of properties like bond lengths, angles, torsion angles, Lennard-Jones potential, as well as partial atomic charges. These parameters are derived from spectroscopic and diffraction experiments, quantum chemical calculation or empirical parameterization used to reproduce liquid properties such as density and molar enthalpy. The final testing for a force field is to reproduce other quantities which were not used in the parameterization process, for example, the thermodynamic properties of the mixed systems, diffusion constants or dynamic properties such as relaxation times. In addition, the good transferability of the force field is also required. This is achieved by minimizing the number of sites or pseudo-atoms that are needed for any particular molecule, and by using the same parameters for given sites in all types of molecules [45].

One of the widely used potentials is the Optimized Intermolecular Potential for Liquid Solutions (OPLS) proposed by Jorgensen [7]. The set of intermolecular potential functions is derived primarily by directly fitting experimental thermodynamic and structural data on pure organic liquids, water, water solutions and others [7]. In OPLS alcohol, for example, there is one site for each atom, except for the $CH_n$ groups which are taken as single sites. The Lennard-Jones parameters for $CH_n$ group are taken directly from OPLS hydrocarbons, in such a way, that prolonging the alcohol chains reflects as adding the contribution of $CH_n$ groups in the force field. The charge distribution is the same for all alcohols and is assigned to O (negative charge) and H and $\alpha$-C, $CH_n$ (positive charges). This is a simplification that ignores the influence of the hydrophobic tail on the electronic structure of the hydroxyl group and also, using the fixed charges neglects many-body polarization effects [45]. Internal rotational degrees of freedom are defined by the Intermolecular Rotational Potential Function using the Fourier expansion. The OPLS models proved to be very good in modeling the variety of organic liquids [12][7], however, it has been found that OPLS alcohols are not transferable to the long-chain alcohols and to elevated temperatures [46]. The group of Siepmann [45] developed The Transferable potential for phase equilibria (TraPPE) force field. The parameterization follows the same line as OPLS models fitting the Lennard-Jones parameters for $\alpha$-C and $CH_n$ groups. It is aimed to reproduce the thermodynamic properties over a wide range of physical conditions, starting with the single-component vapor-liquid coexistence curve.

Reproducing the macroscopic quantities has been proven to be a good quality assessment for the validation of different force fields. However, recent studies emphasized the importance of probing the reliability of microscopic features, namely,

the local microscopic structure. Weerasinghe and Smith [47], proposed to use the Kirkwood-Buff theory to access structural properties. They developed a series of force fields which are designed to reproduce the experimental Kirkwood-Buff integrals for condensed and liquid phase solution mixtures. These KB derived force fields (KBFF) have been shown to reproduce reasonably not only the Kirkwood-Buff integrals, but also other thermodynamic and physical properties of aqueous solution mixtures [47]. A similar approach was used by Lee and van der Vegt in modeling the aqueous tert-butanol solutions. They focused on reproducing the solution thermodynamics and aggregation behavior of tert-butanol-water mixtures [35].

There are more than 100 models for water and none of them so far has been able to reproduce all the thermodynamic properties or the phase-behavior of water. A recent review by Guillot listed 46 distinct models and indirectly indicated their lack of success in quantitatively reproducing the properties of real water [48]. The most used models are Single-Point-Charge models such as SPC [49] and the more recent SPC/E [9] and the Transferable Intermolecular Potential models TIP3P, TIP4P [10][50] and most recent TIP5P [11]. A detailed review of water models is available at http://www.lsbu.ac.uk/water/models.html.

## 4.3 Simulations details

For this thesis, the simulations were performed using the DLPOLY program that was developed in the Daresbury laboratory, Cardiff UK [6]. Several different models were used for simulation of associated liquids, namely, water and alcohols methanol and tert-butanol, as well as alcohols-water mixtures; and the details and abbreviations

are listed in Table I. All models are site-site models and as default the rigid body approximations were used for modeling the intra-molecular bonding. A rigid body unit is a collection of point atoms whose local geometry is time invariant. The dynamics of rigid body is described in terms of the translational motion of the center of mass and rotation about the center of mass. The calculation of force are solved using the quaternion formalism.

The inter-molecular interactions are expressed as a sum of Lennard-Jones potentials and Coulomb potentials. That is the force field between two molecules $i$ and $j$ is defined as a sum of interaction between the sites $\alpha$ and $\beta$, on molecules $i$ and $j$, respectively, with partial charges $q_{\alpha i}$ and $q_{\beta j}$, corresponding Lennard-Jones diameter $\sigma_{\alpha i \beta j}$, energy $\varepsilon_{\alpha i \beta j}$ and relative distances $r_{\alpha i \beta j}$.

$$U_{ij}(r) = \sum 4\varepsilon_{\alpha i \beta j} \left[ \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{12} - \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{6} \right] + \sum \frac{1}{4\pi\varepsilon_o} \frac{q_{\alpha i} q_{\beta i}}{r_{\alpha i \beta j}} \tag{4.1}$$

The Lennard-Jones intra-molecular parameters were calculated using the Lorentz-Berthelot mixing rules $\varepsilon_{\alpha\beta} = \sqrt{\varepsilon_{\alpha\alpha}\varepsilon_{\beta\beta}}$ and $\sigma_{\alpha\beta} = (\sigma_{\alpha\alpha}\sigma_{\beta\beta})/2$.

The cubic periodic boundary conditions (PBC) were used to mimic the pseudo-infinite system. We performed the simulation in canonical (NTV constant) and isobaric-isothermal (NPT constant) ensembles. The standard conditions, if not specified otherwise, were temperature T=300 K and a pressure of 1 atm, that were maintained through the Berendsen thermostat and barostat with relaxation times of 0.1 ps and 0.5 ps, respectively. The integration time step was fixed at 2 fs. For the equilibration periods at least 100 ps were required. We also verified the convergence of the internal energy, volume and pressure, and the stabilization of the distribution functions. The calculation of thermodynamic values, as well as other functions, was

performed for at least 64 ps long runs, taken after the equilibration. The time step of 0.5 ps was used for the collection of the distribution functions. Other simulation details are given with corresponding results in the following chapters.

The DLPOLY package computes the configurational part of the internal energy, by direct averaging of the pair interactions for the Lennard-Jones interaction and by the Ewald summation techniques for the Coulomb interaction. The Lennard-Jones interaction contributions were accumulated within the range of certain cutoff value, usually taken around the half-cell size, and corrected for the part beyond the cutoff [6]. This correction is analytically calculated using expression:

$$U_{LJij}(r \to \infty) = \int dr \sum \varepsilon_{\alpha i \beta j} \left[ \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{12} - \left( \frac{\sigma_{\alpha i \beta j}}{r_{\alpha i \beta j}} \right)^{6} \right] \tag{4.2}$$

- **Table I**. Force field parameters

| | Methanol OPLS [7] | | | Methanol WS [8] | | |
|---|---|---|---|---|---|---|
| | $H$ | $O$ | $CH_3$ | $H$ | $O$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.71131 | 0.86612 | 0.088 | 0.6506 | 0.8672 |
| $\sigma$(Å) | 0.0 | 3.071 | 3.775 | 1.58 | 2.664 | 3.748 |
| $q$(e) | 0.435 | -0.7 | 0.265 | 0.52 | -0.82 | 0.3 |
| | *tert*butanol OPLS [7] | | | | | |
| | $H$ | $O$ | $C$ | $CH_3$ | $CH_3$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.71172 | 0.20936 | 0.67073 | 0.67073 | 0.67073 |
| $\sigma$(Å) | 0.0 | 3.07 | 3.80 | 3.91 | 3.91 | 3.91 |
| $q$(e) | 0.435 | -0.7 | 0.265 | 0.0 | 0.0 | 0.0 |
| | Water SPC/E [9] | | | Water TIP4P [10] | | |
| | $HW$ | $OW$ | | $HW$ | $OW$ | M |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.650 | | 0.0 | 0.648 | 0.0 |
| $\sigma$(Å) | 0.0 | 3.116 | | 0.0 | 3.15365 | 0.0 |
| $q$(e) | 0.4238 | -0.8476 | | 0.5200 | 0.0 | -1.0400 |

| | Water TIP5P [11] | | | |
|---|---|---|---|---|
| | $HW$ | $OW$ | $M1$ | $M2$ |
| $\varepsilon(kJmol^{-1})$ | 0.0 | 0.6694 | 0.0 | 0.0 |
| $\sigma$(Å) | 0.0 | 3.12 | 0.0 | 0.0 |
| $q$(e) | 0.241 | 0.0 | -0.241 | -0.241 |
| | Acetone OPLS [12] | | | |
| | $C$ | $O$ | $CH_3$ | $CH_3$ |
| $\varepsilon(kJmol^{-1})$ | 0.440 | 0.879 | 0.67 | 0.67 |
| $\sigma$(Å) | 3.75 | 2.96 | 3.91 | 3.91 |
| $q$(e) | 0.3 | -0.424 | 0.62 | 0.62 |

The Ewald summation technique enables the calculation of energy for the long-range potentials such as electrostatic potential. The first step is to add the opposite charge distributions usually gaussian distributions, for each charged site. This way, the effect of screening transforms the long-range potential into short-range potential, for which the contribution is calculated by direct summations of the screened pair interaction using the truncation of potential at the half-cell distance. Now one needs to subtract the contribution of the added charges. This potential is obtained from Poisson's equation and is solved as a Fourier series in reciprocal space. The complete Ewald sum requires an additional correction, known as the self energy correction, which arises from a gaussian acting on its own site, and is constant. Ewald's method therefore replaces a potentially infinite sum in real space by two finite sums: one in real space and one in reciprocal space; and the self energy correction [6].

# Chapter 5

# Density fluctuations in N-constant ensemble

The structural information, in the integrated form, is partially contained in the Kirkwood-Buff integrals. These quantities can be derived both from the thermodynamic measurements and scattering experiments, as well as from computer simulation calculations of radial distribution functions.

In this chapter, we will discuss the reliability of the radial distribution functions calculated from computer simulations. We will address several important questions: a) the reliability of the asymptotic limit of the RDF that is connected to number-particle fluctuations in the pseudo-infinite system; b) the influence of the system size and the time of sampling on the simulation results; c) the reproducibility of density and energy distributions.

## 5.1 The radial distribution function

The radial distribution function is the central quantity of our investigation. The RDF provides a means to calculate the Kirkwood-Buff integral and also the structure factor,

which are both experimentally measurable quantities. It provides direct information about the structural organization in the simulated systems. However, the various features of the RDF are not often easy to interpret, even for simple Lennard-Jones liquids [51]. Many studies require the high accuracy of the computed radial distribution function. Nevertheless, the simulation technique imposes constraints which are the cause of errors in simulated data.

The errors in simulated data according to Kolafa et al. [52] could be summarized as follows: a) systematic errors due to a finite number of particles in simulation cell, usually called finite-size errors; b) closely related with size and density of simulated system is also the maximum range for which the calculation of RDF is valid, i.e. the so-called the tail errors; c) errors due to a histogram calculation or grid errors, which become important in calculation of discrete functions of RDF, such as the calculation of structure factor using the Fourier transformation, and; d) statistical errors on collected data which are mainly due to a correlation of accumulated runs in average calculations.

The finite-size and also tail-errors are mainly due to the fact that the contribution for the RDF from the maximum radius to infinity is usually approximated as one. The grid-errors treated with the appropriate mathematical calculation can be minimized [52]. In statistical averaging we required that each contribution is independent (or uncorrelated), and this depends on the frequency of measurements. This frequency compromise between simulation times and overall efficiency expressed as overall statistical error. Kolafa et al.[52] discuss in detail the contribution of each error in calculating the bridge function of hard spheres by direct inversion of simulation data. They uncover that even for a hard-sphere model the contributions of size

and tail effect play an important part in calculation the inverse functions of RDF. To overcome this effect the correction functions are suggested in order to eliminate the "error" contribution from the simulation data.

It is clear that the calculation which is sensitive to the long range limit of RDF requires extensive simulation or specific treatment of the data. However, in our research this imposes constraints that lead to a physically incompatible calculations. Usually the simulations are performed in the canonic or isothermal-isobaric ensemble, which do not allow for the global fluctuation of the number of particles. The constant-N constraint imposes a fixed behavior of the asymptotic part of the RDF and it is not clear how we can use the N-constant RDF to compute the KBI which is strictly defined only for the grand canonical ensemble.

In addition, the radial distribution function corresponds to angular average of the pair distribution function; therefore it lacks information about angular correlations. The molecular liquids that we study have anisotropic interaction such as hydrogen-bonding, and angular correlations are an important part of the structural information. To account for the angular correlations we compute also site-site radial distribution functions (sRDFs) that allows us to tackle the anisotropy indirectly. All formulae are given in chapter 2.

## 5.2 Large N-limit of the pair distribution function

Let us examine in more detail the asymptotic behavior of the pair distribution function. The pair distribution function $g_N^{(2)}(r_1, r_2)$ is defined by relation 2.42 which is

more conveniently to write as:

$$g\left(1,2\right)=\frac{\rho\left(1,2\right)}{\rho^{2}} \tag{5.1}$$

where $\rho\left(1,2\right)=\rho_{N}^{(2)}\left(r_{1},r_{2}\right)$ and $\rho^{2}=\rho_{N}^{(1)}\left(r_{1}\right)\rho_{N}^{(1)}\left(r_{2}\right)$ for homogenous system.

This definition implies that when particles are uncorrelated $\rho\left(1,2\right)=\rho^{2}$, which corresponds to:

$$\lim_{(r\rightarrow\infty)}g\left(1,2\right)=1 \tag{5.2}$$

This relationship is trivially valid for an infinite system, but must be revisited for a finite system with N-particles. These are the conditions for micro-canonical (constant NVE), canonical (NVT constant) or isobaric-isothermal (NPT constant) ensembles in which most simulations are performed.

The asymptotic form of RDF as well as pair density function for a finite-N system, is usually described in terms of the 1/N correction. The general correction term to the pair density function for a systems with constant density, was first found by Ornstein and Zernike [13]:

$$\lim_{(r\rightarrow\infty)}\rho\left(1,2\right)=\rho^{2}\left(1-\frac{\rho k_{B}T\chi_{T}}{N}\right) \tag{5.3}$$

This was proven in more general way by Lebowitz and Percus [14] both through thermodynamic and statistical mechanic route. They investigated the corrections to the representation of the joint distribution $q+l$ particles $\rho^{(2)}(q+l)$ to the product $\rho(q)\rho(l)$ for large separation between sets of q and l particles. They found that the 1/N correction term to a simple product, depends on the derivation of the densities $\rho(1)$ and $\rho(2)$ with respect to the total density $\rho=N/V$:

$$\lim_{(r\rightarrow\infty)}\rho\left(1,2\right)=\rho(1)\rho(2)-\frac{\rho k_{B}T\chi_{T}}{N}\left(\rho\frac{\delta\rho(1)}{\delta\rho}\right)\left(\rho\frac{\delta\rho(2)}{\delta\rho}\right) \tag{5.4}$$

When particles 1 and 2 are both in the interior of a uniform fluid, equation 5.4 leads to the Ornstein-Zernike relation:

$$\lim_{(r\to\infty)} g\,(1,2) = 1 - \frac{\rho k_B T \chi_T}{N} \tag{5.5}$$

For the non-interacting particles this asymptotic limes is equal to:

$$\lim_{(r\to\infty)} g\,(1,2) = 1 - \frac{1}{N} \tag{5.6}$$

which is the exact result for the ideal-gas limit in the canonical ensemble [1]. Similar calculation for a grand canonical ensemble evaluates the asymptotic value to be exactly one.

$$\lim_{(r\to\infty)} g_{\mu VT}\,(1,2) = 1 \tag{5.7}$$

The basic assumption which is made implicitly in the derivation of equations 5.4 and 5.3 is the absence of the long-range correlation in a system, that is the fluctuation behavior does not extend over all volume of the system.

Therefore, the simulation RDFs should contain the N-dependent term. However, this correction depends strongly on the reduced compressibility (see equation 5.8) of the system. For our systems of interest the reduced compressibility is of order $10^{-2}$, therefore we expect that this correction will not influence our results since it is in the range of the statistical error of the simulation data. We will verify this assumption through the inspection of system size dependence on the RDFs in the following sections.

The 1/N term in RDFs was discussed also by Ben-Naim [53], where he states that for a reasonable N the 1/N term is a negligible quantity except when the integration over RDF extends to infinity. Therefore, we find that this observation at present has not been adequately addressed, moreover, some algorithms include correction factors

to force the 1/N-dependent limits, neglecting the correct formula which depends, also, on the compressibility of the system.

## 5.3 The density fluctuation in N-constant ensemble simulations

The asymptotic behavior of RDF directly concerns the calculation of Kirkwood-Buff integrals, as well as indirectly, properties which depend on the fluctuation of particle number such as compressibility. The expression 3.11 can be rewritten through a fluctuation-dissipation type relation:

$$\frac{\chi_T}{\chi_T^0} = \frac{\partial \rho}{\partial \beta P} = \frac{\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle}{\langle N_\alpha \rangle} = 1 + \rho G_{\alpha\beta} \tag{5.8}$$

where $\chi_T^0 = \beta/\rho$ is the ideal gas compressibility, $\rho$ number density and $G_{\alpha\beta}$ is the Kirkwood-Buff integral. This formula is strictly applicable only to the grand canonical ensemble and to emphasize we write the physically incorrect formula for N-constant ensemble, where the 5.8 has trivial and not physical solution, since the global fluctuation is equal to zero (the total number of particles $N$ does not fluctuate):

$$\left( \frac{\chi_T}{\chi_0} = \frac{\partial \rho}{\partial \beta P} = \frac{\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle}{\langle N_\alpha \rangle} = 1 + \rho G_{\alpha\beta} = 0 \right)_{NVT} \tag{5.9}$$

Therefore, a natural question is: when we perform simulations with a fixed number of particles, can we reproduce correct fluctuation of particle number or correct KBI. It turns out that, even with this constraint the results of a simulation reproduce correctly the values of the KBI and corresponding compressibility of an open system. This is in line with a well known calculation of the chemical potential where the chemical potential is evaluated from the simulation in the N-constant ensemble [54].

The procedure consists of evaluating the insertion free energy of an additional particle, which implicitly supports the existence of the local density fluctuations. These local fluctuations give rise to the apparent global density fluctuation in the very absence of any such macroscopic feature. Such behavior is observed equally in mixtures, where, in addition to density fluctuations, concentration fluctuations are equally important. We are particularly interested in the structural properties and the quantities which strongly depend on the fluctuation of concentration and fluctuation of the number of particles in the associated systems.

What is the origin of the density fluctuation in the N-constant ensemble? To clarify this issue one can look more closely to the properties of a sub-volume of the box with volume V containing the N particles at the temperature T, keeping in mind that the configuration of the particles in this cell represents one microstate of NVT constant ensemble. The intensive properties such as temperature and pressure should not be changed since the sub-volumes are in equilibrium with the rest of the box, yet they can fluctuate depending on the size of the sub-volumes. Size of the sub-volumes is also arbitrary and can be chosen as constant. Therefore, every sub-volume keeps the T, P or V constant property. However, any chosen sub-volume can exchange the particles with the rest of the box, that implicates that the N-constant property is lost. Therefore, in N-constant ensemble the sub-volumes have the features of grand canonical ensemble, which is the origin of the local fluctuation of particle number in N-constant ensemble.

The radial distribution function is directly obtained by computing the histogram $H(r, \Delta r)$ which counts the pairs of particles (or specific sites) at distances between $r$

and $r + \Delta r$. This histogram is normalized and the resulting expression is:

$$g(r) = \frac{H(r, \Delta r)}{N^2 \Delta V(r, \Delta r)} \tag{5.10}$$

where $\Delta V(r, \Delta r) \approx \frac{4\pi^2 \Delta r}{V}$ is the volume of spherical shell centered at each particle used in counting, reduced by the total volume $V$ of the simulation cell. In practice, this calculation is performed and averaged over several configurations. It is very important to note that 5.10 is valid for any statistical ensemble. It is the nature of fluctuations of particle ensembles that will affect the form of histogram $H(r, \Delta r)$ and determine the corresponding asymptotic behavior of $g(r)$. The computer simulations with an N constant number of particles are not, strictly speaking, simulations of finite or closed systems, since the periodic boundary conditions are used to mimic the infinite system. This pseudo-infinite system has relevant structural information up to the half size of the cell that contains the N-particles. The information for large distances is artificially repeated using the PBCs. Therefore, the RDF-histogram is usually computed using the spherical shells until the half-box length radii. For each shell the remaining volume acts as a reservoir of particles, and this is valid also for shells at the half-cell distances. The total, usually cubic, volume is bigger than the spherical volume that is covered within computation of the histogram, resulting in fact that the total number of particles counted in each run could be different (some particles can be left in the corners of box-cell), therefore we do have a fluctuation even of the total number of particle. This lifts the mystery of the origin of the global density fluctuation in the N-constant ensemble. First, the PBCs allow simulation of a pseudo-infinite system and second, with the appropriate counting formula the fluctuations of particle number are computed. This is in line with more rigorous discussion on the local nature of the density functions by Lebowitz and Percus [14],

where they show that a slow varying density can be interpreted by representing each fluid element as an open system exchanging particles with neighboring fluid elements (total particle number being maintained). This gives meaning to a local chemical potential as well as the local density fluctuation in N-constant systems.

## 5.4 Calculation of the Kirkwood-Buff integral

The KBIs correspond strictly to integrals of RDF in the $\mu VT$ ensemble. The fact that the local fluctuations in the N-constant simulations give rise to the apparent global density fluctuation, allows one to use the simulated N-constant RDFs to calculate the KBIs. The simulated KBIs are obtained by assuming that:

$$G_{\alpha\beta} = 4\pi \int_0^{\infty} \left(g_{\mu VT}\left(r\right) - 1\right) r^2 dr = 4\pi \int_0^{R} \left(g_{NVT}\left(r\right) - 1\right) r^2 dr \tag{5.11}$$

This assumption includes that $g_{\mu VT}\left(r\right)$ at small distances is approximately equal to $g_{NVT}\left(r\right)$ (or equally for NPT-ensemble simulations, $g_{NPT}\left(r\right)$ ) and that for R, the cutoff distance, $g_{NVT}\left(r\right)$ is essentially unity.

In practice, the latter condition is difficult to achieve unless one uses very large systems. The similarity between the RDF in canonical and grand canonical ensembles has been illustrated by Weerasinghe and Pettit[55], however, only within the short range distances and using the simulation of the Lennard-Jones system. Lyubartsev and Marčelja [56] estimate the radial distribution function of the infinite system from its N-particle finite system counterpart as equal to:

$$g_{open}\left(r\right) = g_N\left(r\right) + \frac{\rho k_B T \chi_T}{N} \tag{5.12}$$

One needs to decide, for which distances the approximation in equation 5.11 is valid, or if it is necessary to use a correction term as in the equation 5.12.

Chitra and Smith[57] used a cutoff distance equal to a range over which the intermolecular forces dominate the distribution of particles. The above approximation is equivalent in spirit to the assumption of a correlation volume used to obtain preferential solvation parameters through inversion of the KBI theory using experimental data obtained at constant P and T, and the experimental estimates of the correlation volume suggest that the sphere of influence of one molecule over another extends for several molecular diameters [53] [57] [58]. Weerasinghe and Smith suggested that a reasonable approximation to determine KBI is typically one molecular diameter [47]. For the methanol-water mixture they reported that all RDFs were essentially unity beyond 1.0 nm and they calculated KBIs by averaging between 0.95 nm and 1.2 nm, arguing that good agreement with experimental values was achieved [23]. However, we anticipate that large cutoff distances are necessary for aqueous solutions, since the pair-correlations are also affected by the micro-heterogeneous structural organization which is on a larger scale than the first-neighbor distances.

Also, in practice, the $r^2$ weighting in the KBI calculation indicates that long range effects may dominate, and that numerical errors from the simulation may be amplified. Due to this reason, one has to have small error in the accumulated data, and one way to obtain this is to have long runs which would minimize the statistical error. Weerasinghe and Smith investigated the convergence properties of the KBI as a function of simulation time [47]. They showed that the reasonable values of the KBI were obtained after 1 ns, although for the minority species small variations are seen even after 2 ns. However, they did not discus what is behind the necessity of

such extensive simulation in order to reach convergence.

Many simulations have been performed using different system sizes and simulation times. For example, several studies of methanol-water solutions have system sizes such as: N=2000 for a calculation of KBIs [23]; N=600 to study cluster formation [59]; N=500 for a calculation of the excess properties using the MC simulation [60]; and the simulation times of the order of several ns. This naturally imposes questions why are the simulation scales so important for aqueous solutions? We will debate the idea that aqueous solutions have local persistent inhomogeneities as inherited feature. Therefore, the structure as well as the stability of such systems is due to the fine balance between concentration fluctuation and micro-heterogeneity and any size effect could strongly influence the equilibrium properties. For example, the system, due to the small size, could be trapped in an unstable state and kept from demixing, or a small size system could not be able to accommodate a range of correlations between locally formed clusters. Larger time scales are needed, apart from statistical reasons, due to the coupling of the two different temporal dynamics, the evolution time of pair correlations, and the larger evolution time of clusters dynamics. Therefore, the computer simulations of water solutions should be revisited in line with the recent concepts of micro-heterogeneities.

## 5.5 Density and energy distribution

In order to clarify several important issues that we mentioned above, we first examined the results of simulations of neat liquids. We evaluated the contribution of the pair correlations beyond the first neighbors for neat systems using the calculations of KBIs

and the calculation of configurational energies. We anticipate that these results will provide information of the characteristic behavior that is precursor to what is found in mixtures.

### 5.5.1 Theoretical considerations

The computation of RDF and KBI for molecular liquids can be done in terms of the site-site functions as described in chapter 2. To follow how the asymptotic values are attained, one can introduce the running quantities. The running Kirkwood-Buff integral for $\alpha$ and $\beta$ sites $G_{\alpha\beta}(R)$ is defined as:

$$G_{\alpha\beta}(R) = 4\pi \int_0^R \left( g_{\alpha\beta}^{(2)}(r) - 1 \right) r^2 dr \tag{5.13}$$

Following the behavior of running integrals one can see: the range of the intermolecular distances over which this quantity fluctuates, reflecting the microscopic arrangement of molecules; and for which distances it does not vary anymore, matching the expected asymptotic value. An important remark concerns the site-site KBIs. Since the center of a molecule is arbitrary, all the site-site KBIs are identical and equal to molecular KBI that is calculated from the molecular RDF (see chapter 2).

$$G_{\alpha\beta} = \lim_{(r\to\infty)} G_{\alpha\beta}(R) = G \; for \; any \; sites \; \alpha \; and \; \beta. \tag{5.14}$$

Similar, a running compressibility can be obtained from equation 5.13:

$$\chi_T(R) = \frac{1 + \rho G(R)}{\rho k_B T} \tag{5.15}$$

and one has $\chi_T(R \to \infty) = \chi_T$ .

The excess internal energy of a molecular liquid, modeled using the additive pair-wise site-site interaction $v_{\alpha\beta}(r)$ , can be written as the sum of the site-site interactions in terms of the corresponding sRDFs (see equation 2.49):

$$U_{\alpha\beta}(R) = 2\pi\rho^2 \int_0^R v_{\alpha\beta}(r)\, g_{\alpha\beta}(r)\, r^2 dr \tag{5.16}$$

where the total configuration energies are equal:

$$\frac{U^{ex}}{N} = \sum_{\alpha,\beta} U_{\alpha\beta}(R \to \infty) \tag{5.17}$$

The term $U_{\alpha\beta}(R)$ is a new quantity and it computes the running contribution from the energy distribution. It is based on the analogy with the running Kirkwood-Buff integral 5.13. The equation 5.17 allows one to compute separately the running integrals for the site-site Lennard-Jones interaction, as well as those from partial charges. One can then study the way each type of interaction contributes to the global establishment of the energy of the macroscopic system. Apart from the RDF route (equation 3.11), the configurational energy can be obtained also, using standard methods, such as truncation potential and Ewald summations (chapter 4). We tested the consistency of these two calculations, when the systems size was changed.

### 5.5.2 Simulation details

We conducted a molecular dynamics study of several water models, as well as organic liquids such as acetone and methanol, at a temperature of 300 K and a pressure of 1 atm. The simulation models are listed in Table I (chapter 4). We used the DLPOLY2 program [6]. The Berendsen algorithm was used for constant temperature and pressure simulations. The integration time step was fixed at dt = 2 fs. Several

system sizes were investigated, ranging from N = 256 up to N = 10 976 for water, and N = 256-2048 for other organic liquids. Each system was equilibrated for 100 ps and the statistics for the RDF were collected for another 100 ps. The DLPOLY package contains the computation of the configurational part of the internal energy, by a direct averaging of the pair interactions over the whole cell for the Lennard-Jones interaction, and by the Ewald summation techniques for the Coulomb part. The energy and pressure contributions from the long range part of the Lennard-Jones interactions were corrected, as usual by analytically expressed contributions beyond the cutoff which is taken to be 15 Å.

### 5.5.3 Neat water

We studied the SPC/E water model [9] together with the OPLS TIP4P [10] and more recent TIP5P [11] models. Figure 5.1 shows the running KBIs (equation 5.13) calculated for different water models, together with the experimental value $G = -16.9\, cm^3/mol$. This value has been deduced from the experimental compressibility at room temperature which is $\chi_T = 0.4566\, GPa^{-1}$[15]. The value for the reduced compressibility is then $\chi_T^* = \frac{\chi_T}{\chi_T^0} = 0.0623$ . The corresponding RDFs, which are equal to sRDF for O-O sites, are shown in the top inset.

Although, the integrated structure beyond 8 Å appears as a noisy signal, it nevertheless seems reproduced to some extent by most models. This may correspond to the extent of correlation of the water hydrogen bond network. The lower inset shows the running compressibility together with the experimental value (red line). It is seen that all models reproduce this value quite well. The fact that the correlations reflect the fluctuations until almost the half-width of the simulation box, tends to confirm

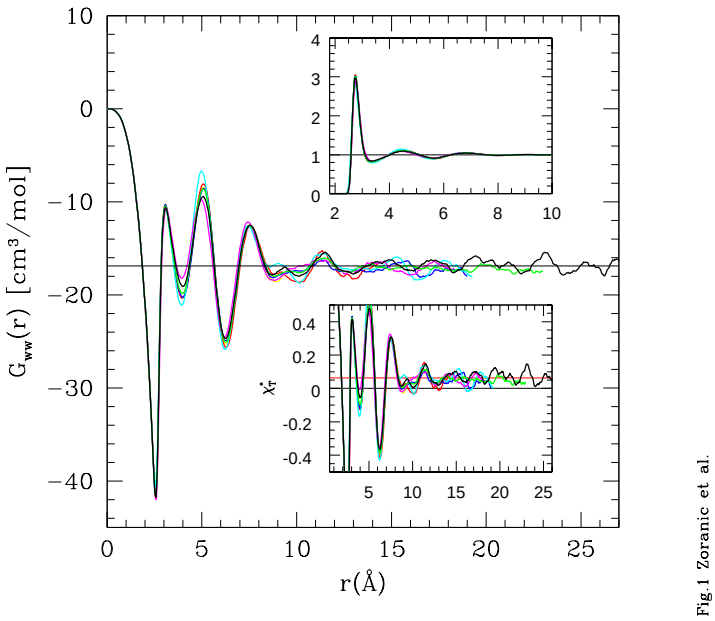


Figure 5.1: The running Kirkwood-Buff integral for water models and for different system sizes (SPC/E: N=256 (yellow), 864 (red), 2048 (blue), 4000 (green), 10,976 (black)); TIP4P: N=2048 (magenta); TIP5P: N=2048 (cyan)). The upper inset shows the RDF with same color conventions. The lower inset shows the running (reduced) compressibility (equation 5.15)for all models with same color conventions. The red horizontal line is the experimental compressibility

that the asymptotic behavior of the RDF is not affected by the constant-N ensemble artifact 5.5. From this observation, we tend to believe that the constant-N ensemble simulations are able to reproduce the same RDF as those calculated in the grand canonical ensemble. This means that the local fluctuations that one obtains from N-constant ensemble simulations, are similar to that of a fully fluctuating ensemble until almost the half-box width.

Figure 5.2 shows the comparison of the running site-site partial Coulomb energies,

and this for several models and system sizes. By partial Coulomb energy, it is meant that the sRDFs in $g_{\alpha\beta}$ in equation 5.16 is replaced by $g_{\alpha\beta} - 1$. This manipulation avoids the divergence of the individual site-site Coulomb energies at large r-values. This is justified through the fact that the total Coulomb energy is not affected by this manipulation, because of the global neutrality of site charges on each molecule. In figure 5.2 it is seen that the running site-site energies have a structure similar to the underlying RDF, albeit in an integrated form. The differences between the 3 water models simply reflect the differences in the partial charges, as discussed below.

The resulting total Coulomb energy is shown in figure 5.3, both for the SPC/E models as well as the TIPnP models. We first observe that the RDF route (equation 5.16) matches very well with that of the direct route, both for the Lennard-Jones and for the Coulomb energies. Although there appears to be almost no difference between the SPC/E and TIP4P models, even in the short range structure, the TIP5P model shows very different amplitude in oscillations. The exact shape of the short-range correlation function depends on the geometry and charge distribution. In TIP4P, negative charge is larger than in SPC/E and is placed more towards the center of the mass, which than leads to slightly higher electrostatic energy at the short range, and increase of the repulsive part of the LJ energy, since due to the electrostatic attraction sites tend to come closer. On the other hand, distribution the negative charges on the two meta sites, as in the case of TIP5P model, resulted in noticeable change of the short-range correlation as well as short-range behavior of the running energies.

The regularity of the oscillations in all models seems to indicate the existence of a sum rule, possibly due to some kind of alternate distribution of site charges. The

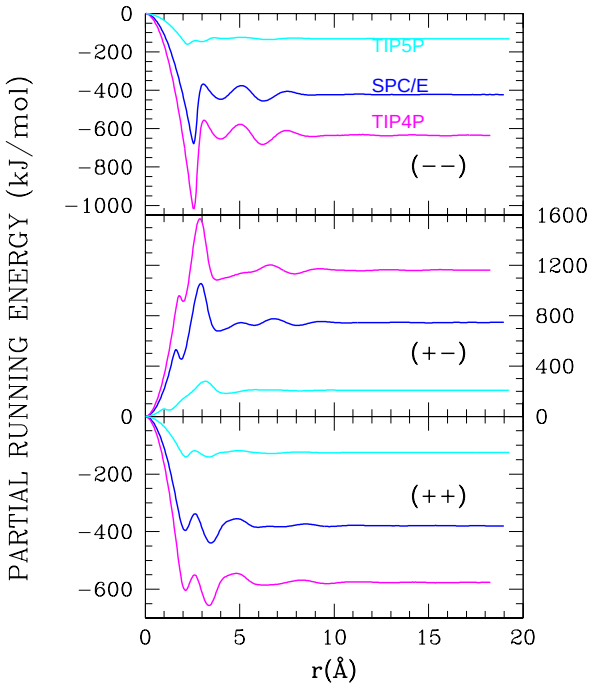


Figure 5.2: The partial running energies for the water models. The color conventions are the same as in Fig. 5.1 except that only the N=2048 are shown. Top panel: (- -) energies, middle panel: (+-) energies and lower panel: (++) energies.

period of oscillations seems close to the water diameter size $\sigma_W = 3.1$ Å. The magnitude of oscillations indicates the differences in the geometry and charges between the 3 charged sites models (SPC/E and TIP4P) and the 4 charged sites model TIP5P (figure 5.3). We note that, despite differences in the total Coulomb energies and in the partial Coulomb energies between the SPC/E and TIP4P models (fig. 5.2), their short range oscillations are very similar. The inset shows the total running Lennard-Jones energy. The larger magnitude of the Coulomb energy $-48.5\ kJ/mol$), when

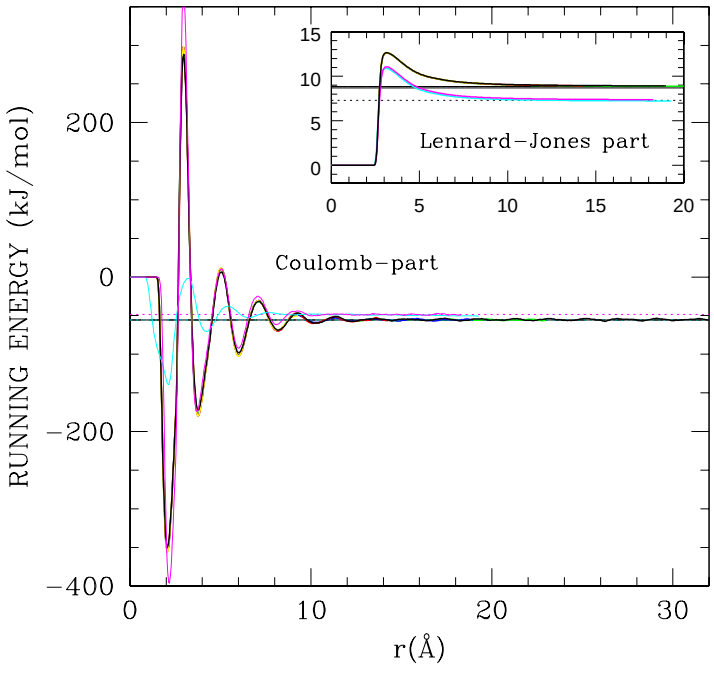


Figure 5.3: The total running Coulomb and Lennard-Jones (shown in the inset) energies for water models. The color conventions are the same as in figure 5.1. The horizontal lines (both in the main and inset) denotes the values of the energies calculated by the direct route in the simulations.

compared to that of the LJ term (about 9 $kJ/mol$), is due to the strength of the hydrogen bonding. We can roughly estimate that both types of energies decay to the bulk value in the range of four to five molecular diameters. This brings the optimal system size up to N = 1000-2000 in order to achieve convergence between the direct and statistical routes for the evaluation of the configurational energy.

Finally, there is a well known $5kJ/mol$ difference between the total SPC/E configurational energy and that of the TIPnP models, which are closer to the experimental

value. Figure 5.3 shows the partial sources for this difference: the Lennard-Jones energies differ by $1.8\,kJ/mol$ between the SPC/E and both TIPnP models. Although it is not easily seen on the plot, there is about a $3.2\,kJ/mol$ difference in the Coulomb parts as well. This means that the usual Berendsen argument that this energy difference in the SPC/E model is due to gas to liquid transfer polarization energy, accounts also for the Lennard-Jones energy for about nearly one third of the difference [49].

### 5.5.4 Neat methanol

Figure 5.4 shows the various sRDFs for the OPLS-methanol model, for the N = 256 and N = 2048 system sizes. Perhaps the most noticeable feature is that the correlations between the methyl sites, and to some extent those involving this site, tend to show more pronounced oscillations. This would mean that these sites reflect the packing structure of the liquid. In contrast, the absence of such a structure in the O and H site correlations, specifically at long range, would reflect the hydrogen bond network correlations. This observation recoups that made on water, and therefore tends to support our interpretation.

Figure 5.5 shows the corresponding running site-site KBIs. We observe that the size dependence is almost unnoticeable, except near the end points of the smallest system r = 8.5 Å. The correlations between the larger neutral methyl groups, display clear long ranged oscillations that extend beyond the range of the largest system size studied here N = 2048. In any case, it is clear that this latter size seems clearly more appropriate if the correlations have to be studied. It is seen, that all the site-site running KBIs converge to the same value at large separations, in accordance with the expected theoretical result in 5.9. This value is also in good agreement with

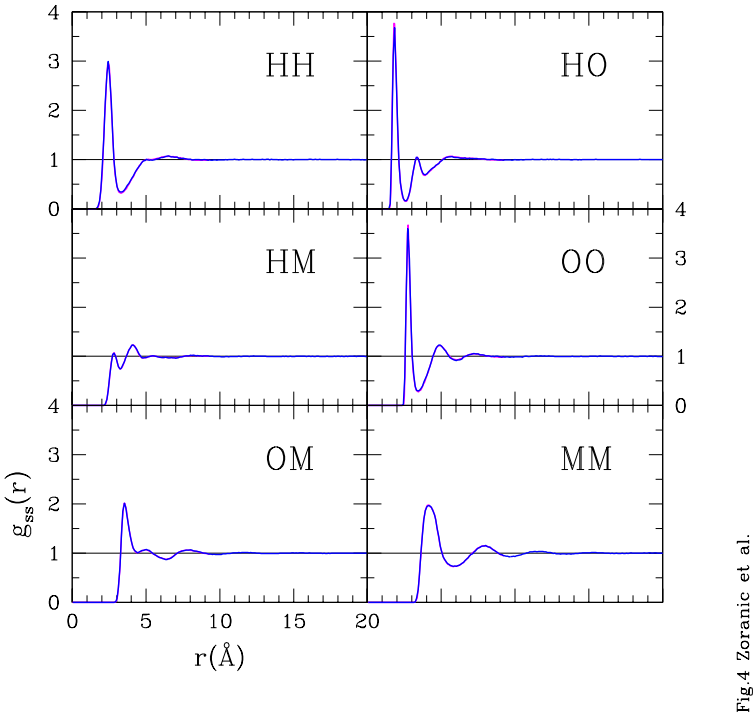


Figure 5.4: The site-site radial distribution functions of the OPLS methanol model for two different system sizes (N=256 (magenta) and 2048(blue)). Each panel shows the correlations for the site pairs indicated in the upper corner).

the experimental value (about $-40\ cm^3/mol$), which again is a direct test of the compressibility of the pure fluid as well as equation 5.9. Figure 5.6 shows the running total Coulomb and Lennard-Jones (inset) energies. Once again the matching of the direct (horizontal lines) and RDF route for the energies is near perfect.

We observe that the Coulomb part has regular oscillations with a period of 2.5 Å, which is smaller than the diameter of the methanol molecule (about 4 Å). The convergence of the latter energy is achieved in a shorter range (about 1 nm) than for

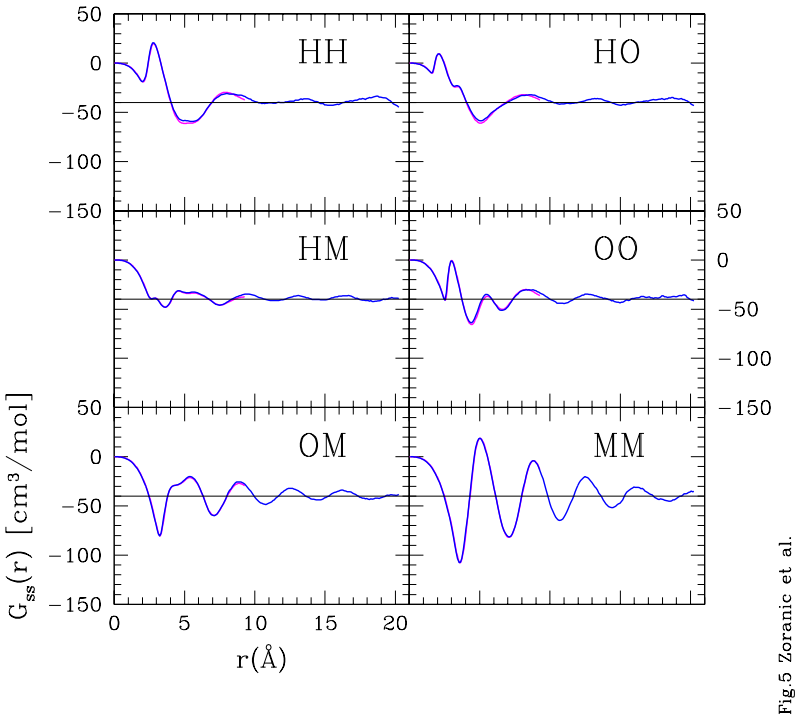


Figure 5.5: The integrated site-site radial distribution functions of the OPLS methanol model. The horizontal lines correspond to the experimental KBI. The color convention is same as in figure 5.4.

the LJ part (about 15 Å). This latter part is similar to the range for water, while it is shorter for the Coulomb part. This might be in line with the fact that the dipolar and quadrupolar interactions are weaker for methanol than for water. The Coulomb energy (about $-30\,kJ/mol$) is not as different from the LJ part ($-6\,kJ/mol$) as it was for water, which means that the hydrogen bonding is weaker in this liquid, as expected.

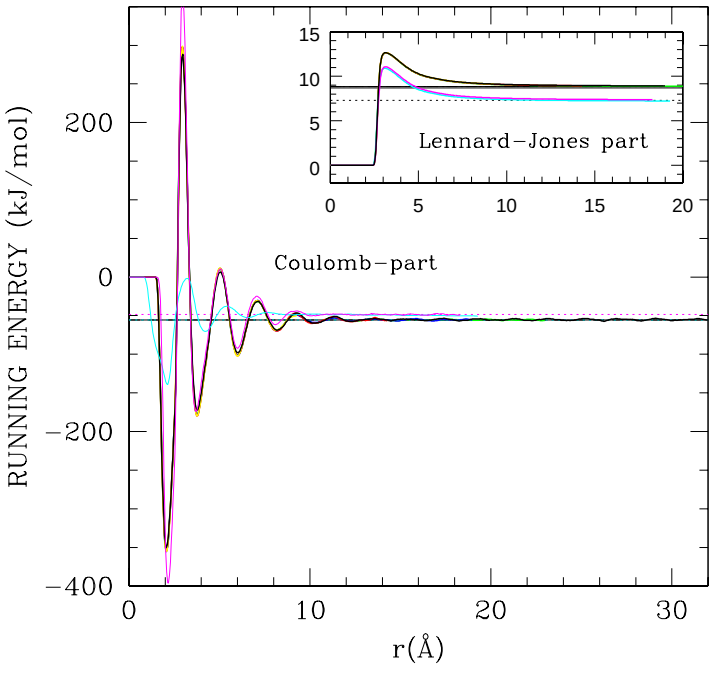


Figure 5.6: The total running Coulomb and Lennard-Jones (shown in the inset) energies for the OPLS methanol model. The color conventions are the same as in Fig. 5.4

### 5.5.5 Neat acetone

We conducted similar studies for the OPLS acetone model [12]. Despite several deficiencies, this model appears as good as the others in what concerns the energy and the density [36]. Figure 5.7 shows the RDF in the main panel, as well as a zoom of the long range behavior in the upper inset and the running KBI in the lower inset, for two system sizes N=864 and N=2048. The oscillations in the running Coulomb energy extend further than 25 Å, while the LJ energy converges at about 15 Å, as in

the case of water or methanol. The period of oscillations is around 5 Å, in accordance with the molecular diameter of acetone.

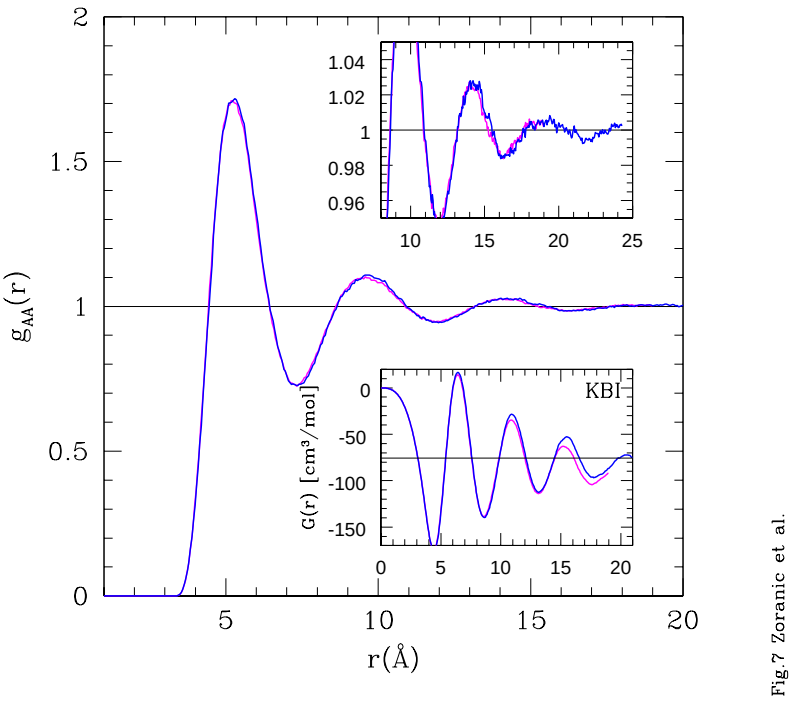


Figure 5.7: The integrated radial distribution of the OPLS acetone models for different system sizes (N=256 (magenta) and 2048(blue)). The upper inset shows a zoom of the large distance behavior, and the lower inset shows the running KBI (the horizontal line is the experimental value).

Figure 5.8 shows the running LJ and Coulomb energies that can now be shown in the same panel, since they are about the same order of magnitude. The total Coulomb term is $-6.5\,kJ/mol$ while the LJ part is about $-21\,kJ/mol$. The oscillations in the running Coulomb energy extend further than 25 Å, while the LJ energy converges at about 15 Å, as in the case of water or methanol. The period of oscillations is around

5 Å, in accordance with the molecular diameter of acetone.

The oscillations in the running Coulomb energy extend further than 25 Å, while the LJ energy converges at about 15 Å, as in the case of water or methanol. The period of oscillations is around 5 Å, in accordance with the molecular diameter of acetone.

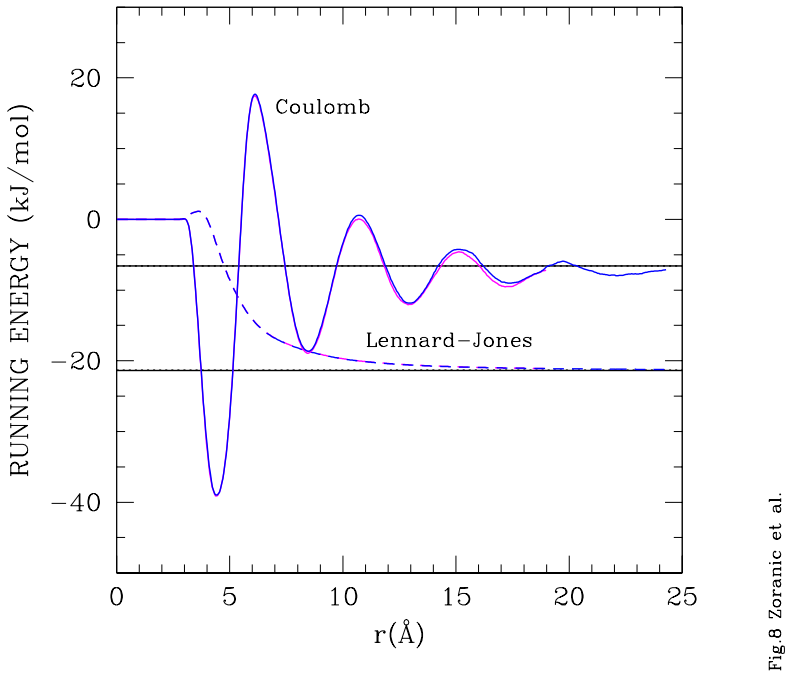


Figure 5.8: The total running Coulomb and Lennard-Jones energies for the OPLS acetone model. The color conventions are the same as in figure 5.7.

### 5.5.6 Discussion

In view of the fact that the period of oscillations of the running Coulomb energy for water and acetone seems to match their respective diameter, one may wonder why those of methanol are even smaller than the diameter of water. One explanation may be that the alternate charge distribution, corresponding to the O and H sites, may match the size of the molecule for the former ones, while it may match the O-H distance in the case of methanol, with the methyl group standing out of the way of the charge alignment. This is further enlarged in the following chapters where the structures of water and methanol are studied in more detail. However, this study of the running energy has already provided indirect information over the microstructure of molecular fluids. Other observation concerns the range required for the Coulomb force to reach uniformity. Namely, for water, and to a greater extent for methanol, the Coulomb force is highly anisotropic, and this anisotropy is seen through the apparent lack of the long range contribution. A similar observation is valid for the sRDF of the sites that are involved in the hydrogen bonding (for example O-O correlations for water and methanol). Namely, the sites that have distributions that corresponds to the packing structure have apparently long-range correlation seen through the long-lasting oscillations of sRDF, in contrast to the hydrogen-bonded sites that show uniformity within the third or forth neighboring shell. The reason behind this contradiction is as follows: the process of angular averaging, for the anisotropic site distribution leads to the apparent loss of the long-range radial correlation, because the angular dependance destroys the spherical-symmetric shell packing and correlations are diminished to a very small variation of the RDF. Therefore, the radial part of the total pair correlation function masks the correct range of the correlation in

the case of the anisotropic site distribution. This is also valid for the calculation of Coulomb energy, and therefore based on this calculation, the estimation of sufficient system size is not totally correct.

To end, we observe that as we go from fully hydrogen bonded water to acetone, the respective trends in the Coulomb and Lennard-Jones energies are reversed: the structure of the networked liquids is governed mainly by the electrostatic interactions, in the sense that these represent the hydrogen bonding according to our modeling.

## 5.6 Conclusion

In this chapter, we analyzed the structure of molecular liquids such as water, methanol and acetone, beyond the first peaks of the RDF. We have shown that these correlations, which reflect the density fluctuations, give rise to running KBIs (or alternatively compressibility) that is constant in a quite large range. This fact indicates that the N-constant artifacts, such as 5.5 do not affect the correlations in a wide range, which is until half-box width range. In addition, for all models studied here, the numerical value obtained for the KBI from the simulations agrees quite well with the experimental value. This alone indicates that, not only do the models reproduce the correct density fluctuations, but they also do so over quite a long range. We have shown that the size dependence does not affect the correlation function and no N-dependent limit is observed in the computed RDFs. Finally, we analyzed the energy distribution by comparing the average with that obtained by direct calculation from the site-site RDF. This analysis reveals that the convergence of LJ interaction is around 15 Å, while the Coulomb energy shows regularly damped oscillations with a period probably

related to the alternation of positive and negative site charges, while the decay seems controlled by the strength of the screening. This calculation provides some indirect information about the structure of the pure liquid. We have proven that the simulation produces reliable results concerning the micro-structure and the properties which are related to fluctuations. Therefore they can be used to tackle the spatial fluctuation and micro-heterogeneous organizations in pure and mixed associated fluids.

# Chapter 6

# Alcohols

The alcohols are organic compound in which a hydroxyl group (-OH) is bonded to a carbon atom ($\alpha$-C) of an alkyl or substituted alkyl group. The general formula for a simple alcohol is $C_nH_{2n+1}OH$. They are classified in the three major subsets: primary, secondary and tertiary, based upon the number of carbons the $\alpha$-C carbon is bonded to. The smallest alcohol methanol is a simple primary alcohol. The simplest secondary alcohol is isopropyl alcohol (propan-2-ol), and a simple tertiary alcohol is tert-butyl alcohol or tert-butanol (2-methylpropan-2-ol). For the alcohols the chemical and physical properties are strongly influenced by the opposite nature of their constituents: the polar hydroxyl group and the hydrophobic carbon tail. The dominance of one affinity over the other also defines the solubility of alcohols in water or other solvents. In case of the simple alcohols the tendency of -OH group to form hydrogen bonds wins over the resistance of the hydrophobic part, making them miscible with water on all concentration.

In our research we namely focus on the methanol and tert-butanol. The methanol as the simplest of all alcohols and tert-butanol having the characteristic geometry of the hydrophobic part and being the largest alcohols that is fully miscible with water.

In the solid state, methanol has an infinite chain intermolecular association with two H bonds per monomer where the intermolecular O-O separation is about 2.66 Å with monomers arranged in parallel, hydrogen-bonded chains [61]. In gaseous state, there is evidence for the special stability of hydrogen bonded cyclic methanol tetramers [62]. However, the structure of the liquid methanol it is still puzzling.

In the following paragraphs, we will mention several experiments that were performed in order to elucidate this structural organization in a liquid alcohols. Early investigations were targeted on the properties of hydrogen-bonding, that is the near neighbor contact but recent interest have been oriented to a specificity of long-range structural organization. Namely, the stability of structure over a wide range of temperature, small viscosity coefficient, the observation of vapor-phase non-ideality and also measurements of the heat capacity suggest the formation of clusters in alcohol solutions. The clustering is also very well supported by the results of the computational simulations [63][64][59]. However, there is on going controversy whether this association is in linear or cyclic form and what are the proportions of each form.

In the 1967 edition of his book "The Nature of Chemical Bonding" L. Pauling based on the presence of cyclic tetrameres in gaseous state, anticipates the dominance of more energetically stable cyclic hexamer units in the liquid methanol [65]. Following this idea, Sarkar and Joarder [66] employed a combined analysis of x-ray and neutron diffraction data to investigate possible clustering of methanol monomers in the liquid state at room temperature. With assumption that the methanol liquid is an aggregate of small clusters of various sizes composed of molecules, the total structural function was written in terms of the cluster structure function. Using the three plausible clustering of methanol monomers: straight and closed tetramer chains and closed

hexamer chain, all of them as planar structure, they fitted the experimental data. This analysis strongly favored closed hexamer chain structure with the respect to other clusters models [66].

Going one step further, Yamaguchi et al. [67] reported a detailed study of the structure of pure liquid methanol at $-80\,^oC$ and $+25\,^oC$ using the pulsed neutron diffraction with isotope substitution on the hydroxyl hydrogens. This technique enables the extraction of the composite partial structure factors $XX, XH$ and $HH$, where $X$ represents a weighted sum of correlation from carbon, oxygen and methyl hydrogen atoms on the methanol molecule. The empirical potential structure refinement (EPSR) [68] computer simulation of the liquid at both temperatures was used to analyze the data. Model distributions of molecules consistent with these data was used to estimate the individual site-site radial distribution functions, the coefficients of spherical harmonics expansions of the orientational pair correlation functions and the length of possible chains of methanol molecules formed in the liquid. The analysis showed that for both temperatures very few molecules have 3 or more hydrogen bonds, and that the distributions are peaked strongly at 1.3 hydrogen bond per molecule, which is against the closed ring structure or even network of molecules. Also, they showed that the average number of molecules that forms chains is 2.7 on average [67].

Another diffraction experiment was reported by Weitkamp et al. [69]. They did seven independent diffraction experiments: neutron diffraction on the isotopic species $CD_3OD$, $CD_3OH$, $CD_3OM$, $CZ_3OD$, $CZ_3OZ$, $CM_3OD$ and high energy X-ray diffraction measurement on $CH_3OH$ (M is mixture of H and D and Z the H/D 'zero mixture' with zero coherent scattering length), which led to seven partial structural functions $H_OH_O, H_OH_C, H_CH_C, H_O(C/O), H_C(C/O), C(O/C)$ and $O(O/C)$. From

this function they deduced molecular geometry. The position and the angle of $O - H_O...O$ is closed to tetrahedral angle, that suggested parametrization of the methanol structure in continuous random network, which, when applied successfully described the intermolecular short distance order. Also they point out that the features of the intermediate distances agreed better with cluster-like than network-like structure [69].

Adya et al. analyzed near-neighbor correlation using the neutron diffraction measurement on liquid methanol ($CD_3OD$, $CD_3OH$, $CD_3O(D/H)$, $CD_3OH$) under ambient conditions. They obtained the total (intra+intermolecular) and intermolecular radial distribution function for three samples. The H/D substitution on hydroxyl-hydrogen was used to extract the partial distribution functions $XH_O$ and $XX$ ($X = C, O$ and H-methyl hydrogen) from the difference techniques of neutron diffraction at both the total and intermolecular level. This analysis showed that the distance $O - H_O$ is about 0.98 Å, the distance of H-bond ($O...H_O$) is between 1.75-1.95 Å, and the coordination number $O...H_O$ is 0.82 Å or 0.51 Å using two different models for extracting the intra-molecular structure [70].

Apart from the scattering experiments, the association in the liquid methanol were studied by other experimental technique. For example, using a combination of density functional calculations of molecular clusters with a quantum cluster equilibrium model Ludwig provided the evidence that liquid methanol is dominated by cyclic and/or lasso structure. He showed that only cluster populations of these structures fit the measured thermodynamic and spectroscopic properties such as heat of vaporization, heat capacity, NMR (Nuclear Magnetic Resonance) chemical shift and quadrupole coupling constants [71].

One of the first studies of tert-butanol was x-ray diffraction experiment on the

liquid tert-butanol at 26 $^oC$ of Narten and Sandler. The data show that distance of hydrogen-bond $O...O$ is around 2.74 Å and that each hydroxyl group has two nearest neighbors at this distance [72]. Zimmermann et al. explored hydrogen bonded rings, chains and lassos in the case of t-butyl alcohol clusters. They investigate infrared OH stretching spectra of hydrogen bonded tert-butanol clusters by ragout-jet FTIR spectroscopy. It is found that the cyclic tetramer of tert-butanol is particularly stable, also they found that lasso structures are also energetically competitive with simple ring structures [73]. Using the second-order difference neutron scattering with hydrogen/deuterium isotopic substitution Bowron et al. measured the intermolecular structural correlation in pure liquid tert-butanol. For analysis they used the EPRS method. The results supports strongly hydrogen bonded intermolecular correlations with each tertiary butanol molecule bonded to two others via hydroxyl group interactions [68].

# Chapter 7

# Micro-structure of neat liquids

Liquids are generally thought of as macroscopically homogenous when they are considered far from phase transitions. However, in associated liquids, the highly directional hydrogen-bonding interaction tends to enhance locally their structure.

In this chapter, we analyze the formation of micro-structure in neat associated liquids, namely, alcohols methanol and tert-butanol. The introduction of the micro-structure is given in the first section. The main part of this chapter is the analysis of the results of MD simulations of neat systems. We conclude with the description of micro-structure in the case of alcohols, and discuss the absence of such clustering in the case of water.

## 7.1 Introduction

The unambiguous determination of clustering in any liquid is a problem in itself, which touches the fundamentally disordered nature of the system. If the system is homogeneous and disordered the associations of molecules have no distinguishing features. The clustering becomes important, for example, when a system is close to a

phase transition, and it is the enhancement of the fluctuation of the number of particle (for a neat system) or concentration fluctuations (for a mixture) that give rise to a local heterogeneity. In associated systems, the cluster formation in the homogeneous phase at equilibrium is at the same time obvious and puzzling. Obvious, because highly directional interaction will tend to strongly associate molecules, and puzzling because the resulting local ordering shows on the macroscopic features, as we will explain later.

The main distinguishing feature of the associated liquids is that they have a strong local ordering as opposed to simple liquids. Namely, simple liquids have a highly symmetric structure, which is governed, mainly by the excluded volume effects. For example, the coordination number for a Lennard-Jones liquid is 12, which is the maximal number of possible near neighbors and it indicates a dense symmetric packing of spherical particles. In associated liquids first-neighbor topology is due to a highly directional hydrogen bonding. Namely, the hydrogen bond is often described as an electrostatic interaction, for example, in water, between the negative partial charge on the oxygen atom on one water molecule and the positive partial charge on the hydrogen atom of neighboring molecule. However, it also has some features of covalent bonding: it is directional, strong, produces interatomic distances shorter than the sum of van der Waals radii, and usually involves a limited number of interaction partners, which can be interpreted as a kind of valence [33]. Therefore, the topology of hydrogen bonding depends on the atomic electronegativity and shape of the molecule. The methanol molecule, for example, has linear connectivity: one hydroxyl group -OH forms in average two hydrogen bonds with neighboring molecules. The water molecules form a tetrahedral shape of hydrogen bonding that allows four

near neighbors. However, just the fact that a system has an anisotropic topology of near-neighbor contact is not sufficient to classify it as micro-structured or micro-heterogeneous.

The definition of micro-structure implies existence of specific self-associations of molecules on a larger scale than a pair-contact. It is still a vague definition, since we did not define the properties which should be sensitive to such clustering or the "signs" (thermodynamical or structural) that would emerge from such local organization. This was the first step in our investigation, to determine the properties of associated liquids that directly probe the clustering. We performed molecular dynamics study of methanol, tert-butanol and also water, in order to extract the information on the local clustering. We focused on the density-density correlation functions: radial distribution function and structure factor. The properties of the structure factor are well understood, for example in the case of crystals. It has a shape of a distinct pattern that is formed by the constructive and destructive interferences on the periodically arranged atoms. Similarly, the structure factor in liquids indicates "periodic", in the averaged sense, arrangement of the molecules in the liquid. Therefore, it should provide information on the local patterns in the associated liquids. Also, counting different sizes and shapes of clusters induced by hydrogen bonding, comes naturally into this analysis. For simple liquids, cluster distribution is a result of the fluctuation of the number of particles, and we will use its features as a reference. The difference between computed cluster distributions and the reference will point to the specific clustering that is consequence of the local heterogeneities.

## 7.2 Simulation details

We have used two models for methanol $CH_3OH$ (MetOH) namely, the OPLS model [7] and the more recent Weerasinghe-Smith (WS) model [8] which is built to reproduce the Kirkwood-Buff integrals of aqueous methanol mixtures. It turns out that both models have very similar local structure as we will show in the later section. We have equally used the OPLS model for tert-butanol $(CH_3)_3COH$ (TBA) [7]. We simulated also the Lennard-Jones (LJ) system using the LJ parameters for a carbon atom. Water has been modeled using the SPC/E parameters [9]. All models are site-site models, with three sites for MetOH, namely, oxygen site O, hydrogen site H, and one single site for the methyl group $M = CH_3$, and six sites for TBA, with the additional carbon site C. We have only considered rigid models herein. Table I (chapter 4) gives the corresponding force field parameters used in this work.

All our computer simulations were performed in the constant NPT ensemble, with the DLPOLY2 package [6]. All reported results concern system size of N=2048 molecules, with the exception of global snapshots of the simulation boxes that are more readable with N=256 molecules. The conditions were fixed at T=300 K and a pressure of 1 atm, which were maintained through the Berendsen thermostat and barostat with relaxation times of 0.1 and 0.5 ps, respectively. A time step of 2 fs was taken. From the analysis of the convergence of the internal energies and volumes, as well as from the stability of the correlation functions we found that the required equilibration time was in general about 100 ps. After this equilibration step, the latter functions were computed over 1000 configurations separated by 0.5 ps, which ensured smooth curves, in particular, in view of the numerical Fourier transforms that

are performed on them. In comparison, the cluster counting needs to be averaged over a lesser number of configurations in order to ensure similar smoothness in the data, about 200 configurations sampled every 20 ps, the larger time step selected to ensure that clusters have been broken and reformed elsewhere from their previous configurations.

## 7.3 Results

The values of the configurational energies and molar volumes, as obtained from the simulations, have been checked for consistency with the experimental enthalpy of liquefaction and densities. This is reported in Table II. It can be seen that the agreement is generally qualitatively good. Note that the enthalpy for the WS MetOH model is not corrected here for the additional polarization factor mentioned in reference [8].

- **Table II**. Enthalpies, molar volumes and densities.

| | Methanol | | | *tert*–butanol | |
|---|---|---|---|---|---|
| | *Expt* | *OPLS* | *WS* | *Expt* | *OPLS* |
| $\Delta H(kJmol^{-1})$ | -37.3 | -35.43 | -41.51 | -46.74 | -45.74 |
| $V_m(cm^3mol^{-1})$ | 40.74 | 44.26 | 42.27 | 94.6 | 93.25 |
| $\rho(gcm^{-3})$ | 0.7869 | 0.7239 | 0.75194 | 0.78086 | 0.79487 |

### 7.3.1 Information from the pair distribution functions

The one body density of homogeneous and disordered liquids is just a constant. Therefore, the information about the local structure can only be retrieved from two-body density function and higher order density functions. The pair-distribution function $g(r)$ measures the correlations between any two particles, and any manifestation of the heterogeneity will be hidden among the usual features of this quantity. It is not straightforward to extract information about clustering amongst other features that are generally specific to a regular molecular distribution. Moreover, in our calculations we have only the radial part of the total pair distribution function and angular correlations are lost in the process of averaging. We considered the site-site radial distribution functions, in order to retrieve partly the information of angular preferences and also the radial distribution function of the center of mass. The site-site RDFs are defined as in 2.66:

$$g_{\alpha\beta}(r_{\alpha\beta}) = \frac{\langle \rho_\alpha(r_\alpha)\, \rho_\beta(r_\beta) \rangle}{\rho_\alpha \rho_\beta} \tag{7.1}$$

In addition to site-site function $g_{\alpha\beta}(r_{\alpha\beta})$, we have equally computed the corresponding structure factors:

$$S_{\alpha\beta}(k) = 1 + \rho \int dr e^{-ikr} g_{\alpha\beta}(r) \tag{7.2}$$

There are alternative definitions, where for example the density $\rho$ can be replaced by the density of sites $\sqrt{\rho_\alpha \rho_\beta}$, but these distinctions do not affect the major features of this function. Direct comparison of the resulting measured and calculated structure factors often leads only to qualitative similarities. Namely, the calculations show that this quantity is very sensitive to interaction parameters meaning that size and, also

position of the peaks can vary depending on the model parameters. Nevertheless, the major features such as local organization should appear on the both measured and calculated quantities.

The structure factor nicely complements the radial distribution function by providing information about medium and long range part of density correlation. Namely, the RDF in direct space and the corresponding structure factor in reciprocal space, form a pair related by the Fourier transform. This mathematical operation relates the long range tail in the radial distribution function into the small k range of the structure factor. For example, asymptotic value of the RDF corresponds to $S(k)$ at value $k \to 0$.

In figure 7.1 a typical radial distribution function as well as the structure factor of an ordinary Lennard-Jones liquid are shown. It has a simple-liquid like feature: the long-lasting regular oscillation indicating the densely packed homogenous system. This oscillatory structure corresponds to the main peak in structure factor.

The position of the first peak and the periodicity of peaks in RDF is $\sigma \approx 3.5\mathring{A}$. The $k_{eff} = 2\pi/\sigma_{eff} \approx 2.3\mathring{A}^{-1}$, corresponds to a $\sigma_{eff} = 2\pi/k_{eff} \approx 2.7\mathring{A}$. We note that neither the period of oscillations nor the inverse position of the main peak is not exactly equal to the LJ radius but rather corresponds to the average value of first contact. This value is estimated to be $1.4\sigma_{LJ}$ from the period of oscillation in RDF[74], which is then equal to a smaller value in the k-space as shown from the calculation of the peaks position. What do we expect for associated molecular liquids?

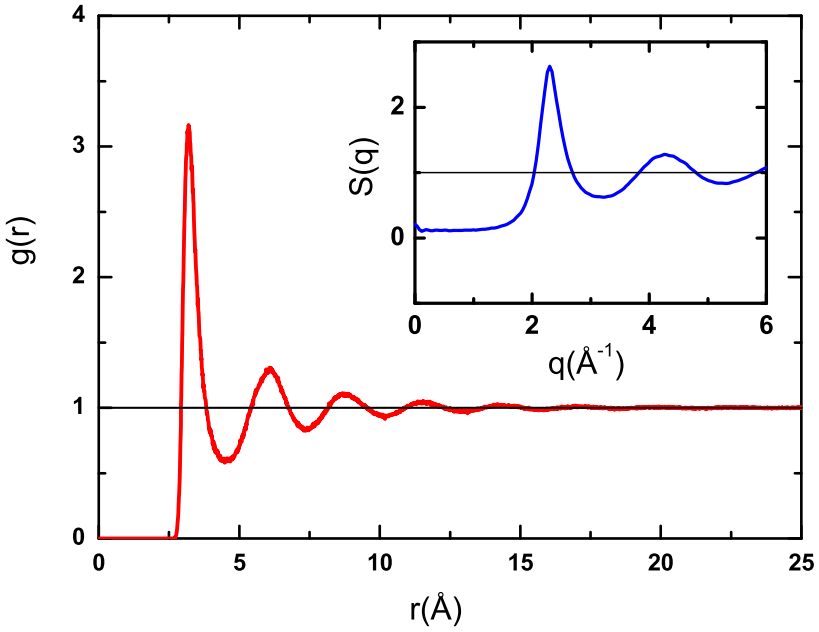


Figure 7.1: RDF and structure factor for Lennard-Jones liquid.

**Results for methanol**

Figure 7.2 shows the site-site radial distribution function for methanol-sites. The upper panel shows correlations between the methyl-methyl sites as well as the center of mass radial distribution function. The other panels show the RDFs of sites that are involved in hydrogen bonding. The first distinguishing features are pronounced oscillations in $g_{MM}(r)$ as well as in $g_{cm}(r)$, which for the distance larger than $\approx 6$ Å are almost superposed. These oscillations are typical for simple liquids as presented in figure 7.1. The $g_{cm}(r)$ has also a narrow first peak which is a signature of hydrogen bonding. Therefore, RDF of center of mass comprises the short-range behavior due

to the strong interaction between neighboring molecules and medium-to-large range oscillations typical for simple liquids.

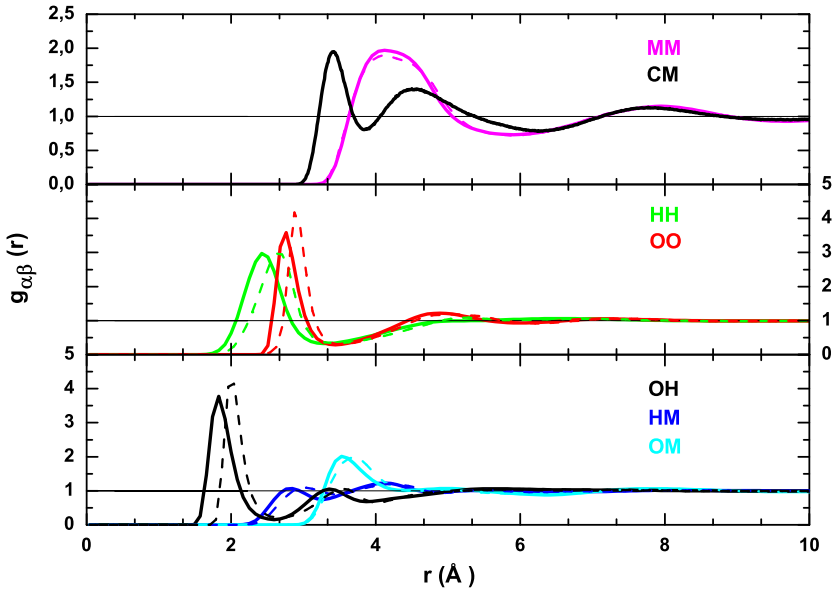


Figure 7.2: Site-site RDFs for OPLS and WS models of methanol. Top panel: for MM (magenta) sites and center of mass (black). Middle panel: for OO (red) and HH (green) sites. Bottom panel: for OH black panel, OM (cyan) and HM (blue) sites. For all data, full curves are for the OPLS model and dashed for WS model.

The distribution functions involving the hydrogen bonded sites have also characteristic short range features: the strong narrow peak that corresponds to a hydrogen-bonding (HB). The first peak in $g_{OH}(r)$ is at the distance 2.5 Å which is the generally accepted value for hydrogen bond. What is surprising is that all distribution functions of HB-sites show a lack of correlations for medium and large distances. The RDFs of these sites for distances larger than $\approx 8$ Å are apparently equal to one, while the

$g_{cm}(r)$, for the same distance, still shows the liquid-like oscillatory structure. Therefore the sites which have stronger correlation due to the hydrogen bonding appear less correlated than randomly distributed sites. The lack of oscillatory structure in the medium-to-long distances is typically found in monomer-monomer distribution functions in polymeric fluids [17], and can be understood in the following way. It is the chained alignment of sites, or monomers that influence the density correlations preferentially along the chain and the near neighbor correlations are weaker. Due to the averaging of angular correlations in RDFs this is translated in the first narrow peaks indicating the strong bond between neighboring sites and the lack of medium-to-large correlations due to the highly anisotropic chainlike structure. The packing effects in chains are felt only for gel-type very dense polymer phases [17].

In our case, this observation can be translated to the following interpretation: the MetOH molecules tend to form chains locally, with O and H sites strongly correlated along the chains, while Met sites are distributed around the chains. The distribution of Met sites corresponds to a uniform and random distributions and is apparently not "influenced" by the chain formation. The methyl site correlations as well as the center of mass correlations have simple liquid-like features, the pronounced oscillatory structure indicating the dense-liquid packing, while oxygen and hydrogen sites have highly anisotropic correlations typical for a chain structure. It is important to note that the information about chain formation is hidden in the shapeless feature of the RDF at large distances, and should be better seen in the Fourier space at short wave vectors.

Figure 7.3 reports the structure factors corresponding to site-site RDF in figure 7.2. In the upper panels we show the structure factors for methyl sites and center of

mass, while the lower panel pictures the hydrogen-bonded sites. The structure factors in the top panel look very much like dense Lennard-Jones-type structure factors (see fig. 7.1). The first peak in $S_{MM}(k)$ is at $k_{eff} = 1.75\,\mathring{A}^{-1}$, from which we can extract $\sigma_{eff} = 3.42$ Å, that roughly corresponds to the size of the methyl site. The main peak in structure factors usually indicates the packing structure. In our case this corresponds to the first peaks of center of mass as well as the first peak in methyl structure factor.

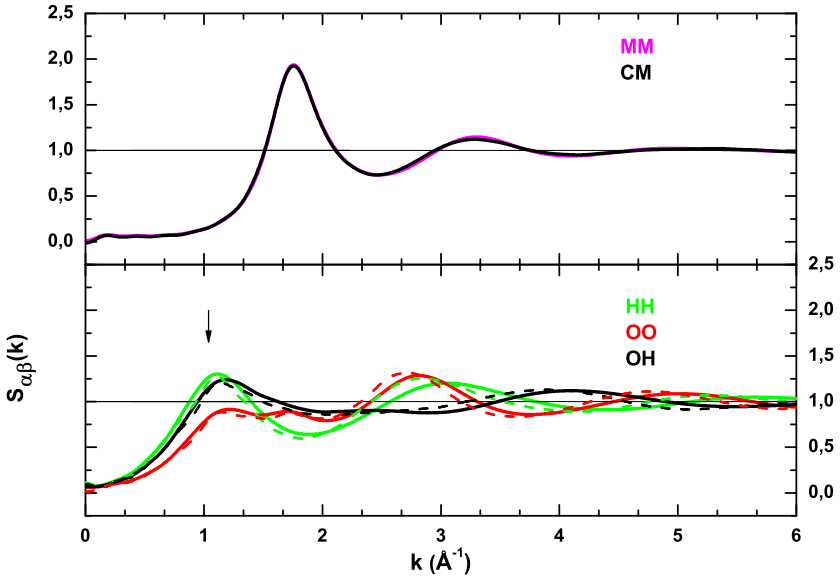


Figure 7.3: Site-site structure factors for some RDFs for OPLS and WS models shown in fig. 7.2, with the same color and line convention. Top panel: MM and center of mass data. Bottom panel: OO, OH and HH data. Pre-peak is indicated by the arrow.

In contrast to this, the structure factors involving the hydrogen bonding sites (lower panel) have a very peculiar shape: they exhibit a peak at the wave vector

$k_p$ smaller then $k_{eff}$. The smaller k value corresponds to larger size structure, for example $S_{OO}(k)$ has $k_p = 1.25\ \mathring{A}^{-1}$ that corresponds to a periodicity of $\sigma_p \approx 5.1$ Å. We will call this peak the pre-peak to emphasize that its k-value is smaller than that of the main peak (this feature is also called the inner peak in the literature [18]). In line with previous direct space interpretation, we associate this pre-peak with the chainlike structure of hydrogen-bonded sites in liquid methanol.

It is important not to mix the features of hydrogen bond between first neighbors and the chainlike structure that is induced by hydrogen bonding. The signature of hydrogen bond is the first narrow peak in the radial distribution function, and should be translated to a large k-vector distance. The small k-vector features of structure factors are due to the large $r$ behavior of radial distribution function. Therefore the pre-peak indicates the structure that is hidden in the apparent shapeless tail of RDF. It is indicative that the HB peak in $g_{cm}(r)$ left almost no sign in the corresponding $S_{cm}(k)$ . Namely, the presence of the HB peak in RDF does not imply necessarily the appearance of the pre-peak, which emphasizes that these two features are distinct. The previous literature is not clear about this problem, and these two features are often confused for one another. Namely, the narrow first peak in RDF and the pre-peak in structure factor have the same origin: strong hydrogen bond between the concerned molecules. However, the peak in RDF expresses direct near neighboring pairing, while the pre-peak expresses the existence of local organization on a larger length scale, induced by hydrogen bonding. This directly tackles the puzzle of the existence of stable local heterogeneity inside macroscopic homogenous fluids. It confronts also the inherited viewpoint of the traditional representation of interaction, where macroscopic properties are usually built from pair interactions.

**Results for tert-butanol**

Figure 7.4 shows the site-site distribution functions for tert-butanol. The correlations between the hydrogen bonded sites are in the right panel, and the other site-site distribution functions in the left panel. In contrast to the case of methanol, all sites show a sharp peak corresponding to hydrogen bonding, except for the methyl group where only a small bump is observed. This indicates that the hydrogen bonding influences also the short-range correlations of sites that are not involved in these interactions. The RDF between the center of masses is almost identical to $g_{CC}(r)$, which indirectly assesses the central positions of the carbon atom in the molecule. Apart from the first peak, all correlations in the left panel follow the simple-liquid packing oscillations, and they are superposed within medium-to-large distances. One can imagine that these oscillations estimate roughly the diameter of LJ-spherical particle centered at the center of the mass that have similar packing requirements. Therefore, we will attribute this value to average molecular size.

The hydrogen bonded sites show a very large narrow first peak, as compared to the methanol RDFs, indicating that the correlations between first neighbors are stronger in TBA than in the case of methanol. The maximum in $g_{OO}(r)$, and to some extent $g_{OH}(r)$, shows two distinct peaks. The modeling of TBA as a rigid molecule imposes two preferential positions for OH group which correspond to a split-peak feature in $g_{OO}(r)$ .

All distribution functions display oscillatory behavior, with a main difference in the period of oscillation: the hydrogen bonded sites have a period larger than the periodicity of correlation function between the centers of masses. While the structural packing is expected to be about the molecular size, a larger period indicates

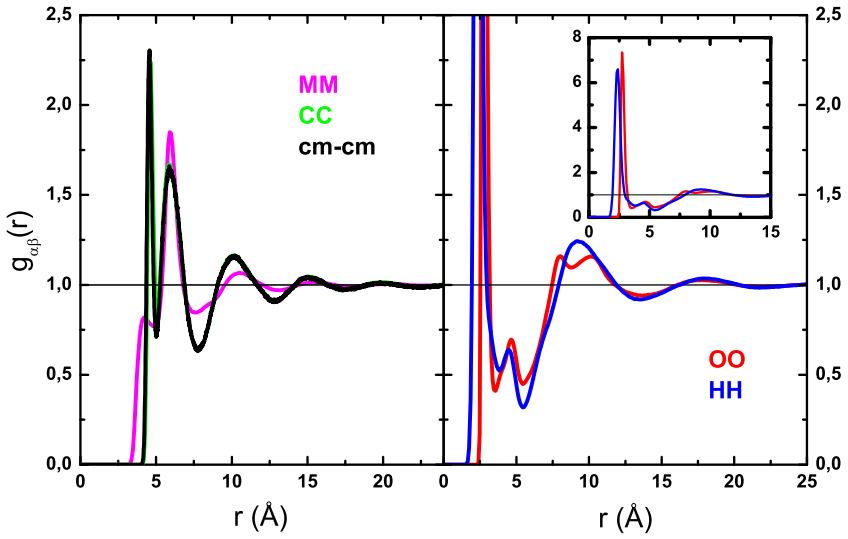


Figure 7.4: Site-site RDFs for OPLS tert-butanol. Left panel: MM (magenta), CC (green) and black for the center of mass RDF. Right panel: OO (red), HH (blue) (the inset shows detail of the peaks).

that hydrogen bonding is modulated by the supra-structure formed by the molecules though this mechanism. The fact that this latter oscillatory structure is absent for methanol is indicative of the differences in local organizations between two liquids: the TBA molecules form a spherical micelle-like structure, so the modulation of the RDF is more liquid-like and more enhanced than in the case of methanol, which we have seen to display polymer-like structures.

Hence, we expect to observe in the structure factors the differences in structuring displayed in the RDFs. Figure 7.5 shows the structure factor of tert-butanol. Again, the typical liquid-like structure factors are observed for all non-bonded sites. The main peak of $S_{cm}(k)$ as well as $S_{CC}(k)$ and $S_{MM}(k)$, is at the wave vector $k_{eff} =$

$1.34\,\mathring{A}^{-1}$ which corresponds to a estimated molecular size $\sigma_{eff} \approx 5.6$ Å. The structure factor of hydrogen bonded sites displays a pre-peak on the $k_p \approx 0.8\mathring{A}^{-1}$, corresponding to the length scale $\sigma_{eff} \approx 7.85$ Å. Therefore, these two lengths correspond to the two different periods of the oscillatory structure in figure 7.4. The size of the pre-peak is higher than that of the main peak, which highlights the strong influence of the association on the structural features in the TBA, as opposed to the case of liquid methanol, where the structure factor for the main peak has the highest value.

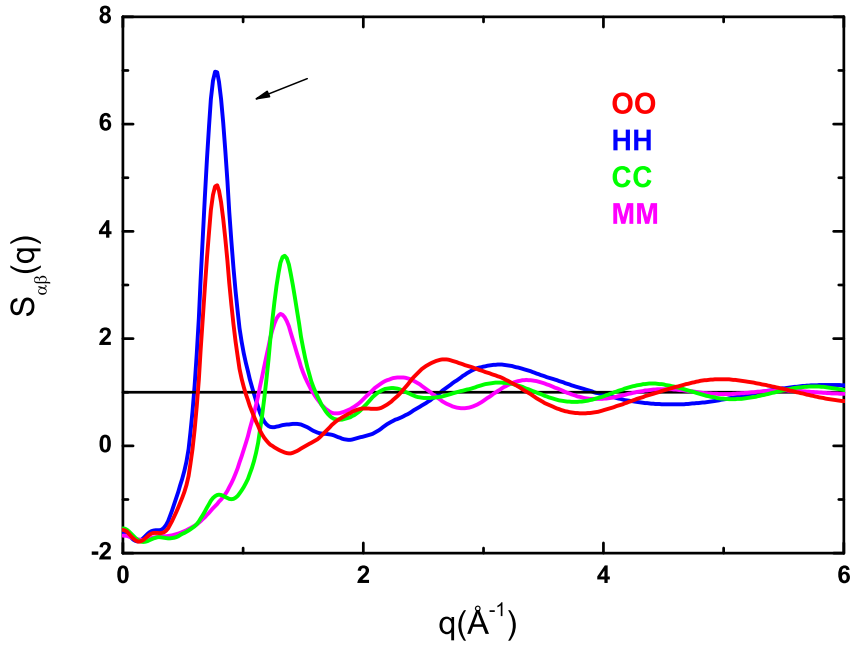


Figure 7.5: Site-site structure factors for some RDFs for OPLS tert-butanol shown in figure7.4, with the same color convention. Pre-peak is indicated by the arrow.

In order to assess the origin of the pre-peak, we have simulated the same model TBA under the same temperature and pressure condition, but without the partial

charges. We verified that this simulation still corresponds to a dense liquid. The radial distribution function and structure factor are shown in figure 7.6. It is striking that the pre-peak has vanished and all site-site distribution functions show similar behavior: all peaks in the corresponding structure factors match the wave vector of the main peak. We note that there is a weak remainder of the pre-peak that attests to the existence of some ordering due to excluded volume effects, the smaller OH groups tending to cluster together.

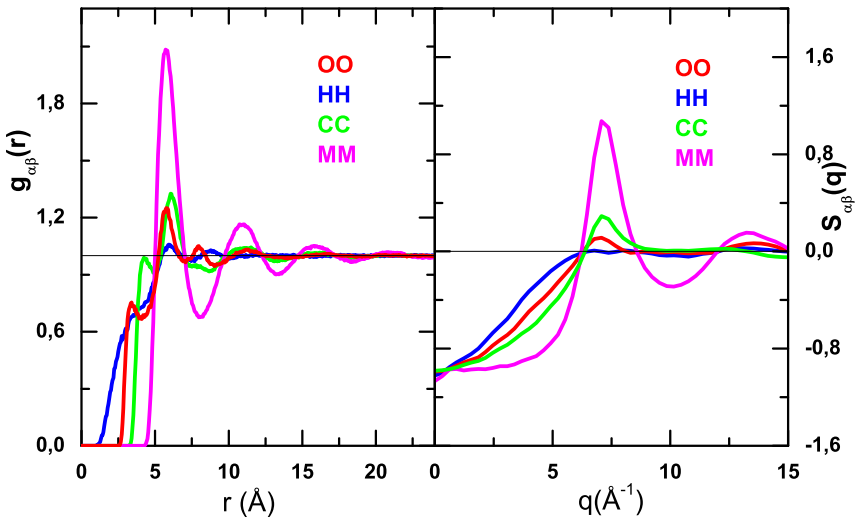


Figure 7.6: Distribution functions for OPLS tert-butanol modeled without charges. Left panel: sRDFs for MM, CC, HH, OO. Right panel: Site-site structure factors for same sites. The color convention is same as in figure 7.4.

An important point that we want to stress out here is the difference in small-k behavior of all the structure factors, between the charged and uncharged cases. Figure 7.6 shows clearly that, in addition to the appearance of a pre-peak, we observe an

increase of $S(k)$ near and at $k = 0$, which is absent in the inset. The latter increase corresponds to an increase in density fluctuations and compressibility. Hence, we can conclude that micro-structured fluids tend to exhibit larger concentration fluctuations as well.

### 7.3.2 Information from the cluster distribution functions

Figure 7.7 pictures a snapshot of a simulation cell of methanol. This snapshot is only one realization of the system and presents one microstate of system. We can associate one microstate with instantaneous density $\rho(\mathbf{r})$ that is calculated using the instantaneous positions of the particle (see equation 2.41):

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \tag{7.3}$$

where $\mathbf{r} = (\mathbf{R}, \Omega)$ are generalized coordinates: $R$ is the position vector and $\Omega$ the set of Euler angles describing the orientation from some arbitrary origin; $\mathbf{r}_i = (\mathbf{R}_i, \Omega_i)$ are coordinates of molecule i of N particles in a system.

The relevant quantities are time averages, which for a uniform and homogeneous system correspond to a constant number density:

$$\langle \rho(\mathbf{r}) \rangle_t = \frac{1}{T} \int dt \rho(\mathbf{r}, t) \longrightarrow \langle \rho(\mathbf{r}) \rangle_t = \rho \tag{7.4}$$

A snapshot of any liquid displays some local heterogeneities. Therefore, instantaneous density for each microstate could be written as a composition of average density $\rho$ and $\delta\rho(\mathbf{r})$ the quantity which measures the local deviation from uniformity: $\rho(\mathbf{r}) = \rho + \delta\rho(\mathbf{r})$. For a homogeneous system the time average of local deviation is equal to $\langle \delta\rho(\mathbf{r}) \rangle_t = 0$ , because $\rho(\mathbf{r}) = n(\mathbf{r})/V$, the local number of particles per volume,

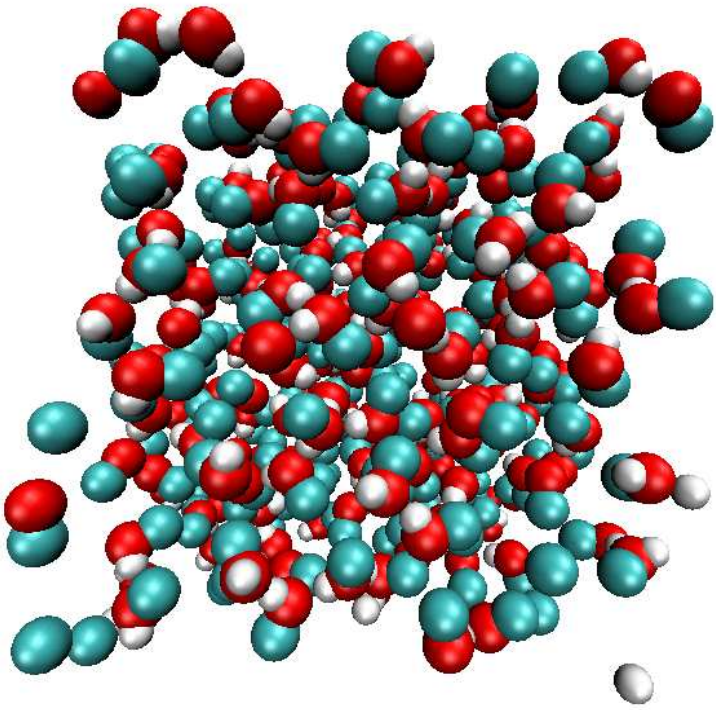

Figure 7.7: Snapshot of the N=256 OPLS methanol system [32].

fluctuates around uniform value throughout the system. The function $n(\mathbf{r})$ measures distribution of particles from some arbitrarily chosen centre. However, instead of an arbitrary partition of the system, the total volume could be divided according to certain constraints. In figure 7.8 we show 2D examples of different partitions of a system. The presented case is a trivial and oversimplified representation of a 2D microstate; however it gives the visual representation of differences in system-partitioning. The random partition yields the random distribution of subsystems. The partitioning due to the association of the particles separates volume into dense

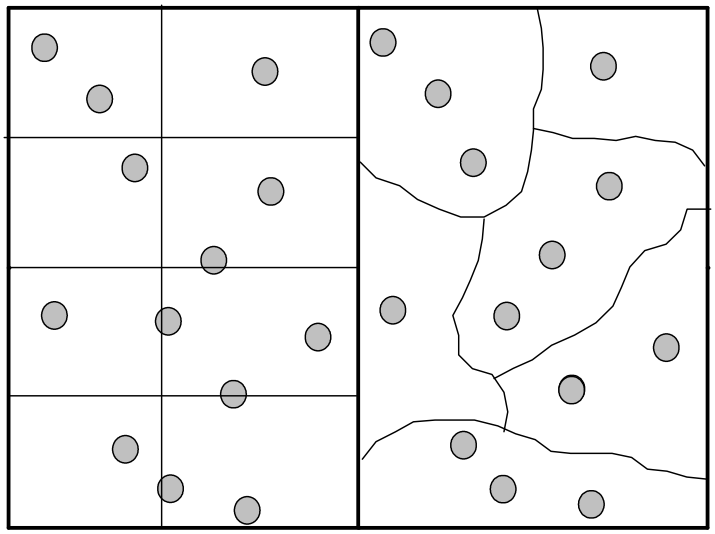

Figure 7.8: Schematic representation of the two types of the partition of the 2D system. In the left panel uniform partition of the system where average number of particles in a chosen sub-volumes is 2. In the right panel partition of system according to the cluster distribution, where is clear that system prefers to form three particle clusters.

and less dense regions, making the local heterogeneities more visible.

Then, the local number of particles can be written as a combination of two functions: the particle-number function of the particles that are bonded in clusters and the particle-number function of the non-connected particles:

$$n\left(\mathbf{r}\right) = n_c\left(\mathbf{r}\right) + n_{nc}\left(\mathbf{r}\right) \tag{7.5}$$

where $n_c\left(\mathbf{r}\right)$ is local number particles per volume of the associated sites, and $n_{nc}\left(\mathbf{r}\right)$ is local number particles of the remaining sites. The time average of instantaneous

density separated as $n(\mathbf{r})/V = n_c(\mathbf{r})/V + n_{nc}(\mathbf{r})/V = \rho_c(\mathbf{r}) + \rho_{nc}(\mathbf{r})$ and $\delta\rho(\mathbf{r}) = \delta\rho_c(\mathbf{r}) + \delta\rho_{nc}(\mathbf{r})$, will result in non vanishing deviation $\langle\delta\rho_c(\mathbf{r})\rangle_t \neq 0$ for a system that has strong local heterogeneities. Then counting for each configuration how many particles appear connected in a cluster of a given size $n$ is a good way to account for local heterogeneities which will not average to zero when computed over many realizations. The probability of finding a cluster of size $n$ is defined as:

$$p(n) = \frac{\sum_k s(k,n)}{\sum_{n,k} s(k,n)} \tag{7.6}$$

where $s(k,n)$ represents the number of clusters of the size $n$ in the configuration $k$. Therefore, $p(n)$ represents cluster size probability function. This definition includes the arbitrariness through the definition as to when two particles are mutually bonded within a cluster. For this analysis we chose the Stillinger[16] definition where two particles are bonded if they are separated for less than certain cutoff value $l_c$. In practice, this distance is chosen to correspond to the average pair contact probability as described by the corresponding radial distribution function. Many such distances can be defined starting from the first-peak distance to a first minimum of radial distribution function. The distance that corresponds to the first minimum of the radial distribution function is the most appropriate considering the temporal stability of clusters [75]. This refers to clustering in the LJ system, however the strongly bonded HB clusters will have a good temporal stability even for distances that are smaller than the first minimum. Therefore for associated liquids, we will consider cluster distribution for a wider ranges of distances between the first peak and the first minimum of corresponding RDFs.

What is the relation between $s(k,n)$ and the instantaneous density? Calculation of $s(k,n)$ comprises counting of the particle and the instantaneous density is a function of positions of the particles. Therefore, the total number of particle is equal:

$$N = \int dns(k,n) \tag{7.7}$$

$$N = \int dV \rho_k(\mathbf{r})$$

Mapping the instantaneous cluster distribution $s(k,n)$ into the instantaneous density $\rho_k(\mathbf{r})$ is not possible because in the counting process we lose the information of the particle positions, however we gain the distribution function that includes information of the local organization. It is important to note that $s(k,n)$ should not be confused with the pair-functions: it is strictly one-body information relative to the instantaneous local heterogeneities of the system. Namely, the counting of clusters involves criteria that are defined by the relative distance of two particles and that correspond to a two-body function. However, the center of counting is constantly changing since we impose the distance criteria on each particle within a cluster. In this process the information of the two-body correlation is lost. The cluster counting gives only the partitioning of number of particles according to specific criteria, which is by nature of definition a one-body function.

Associated to a liquid-gas phase transition, in the Landau formalism, the order parameter is density. In the disordered liquids the density is just a constant, therefore one needs local parameter of order that would give the information about possible small-scale structuring, such as clustering or even network-forming structure. This is concept that goes beyond the traditional description of the order parameter which are global quantities that varies through the phase transitions. We define a cluster distribution function as local order parameter, since it probes if the system has

some preferential clustering. The cluster distribution, as explained before has similar features as density, in a sense that they are both one-body distribution function. Another advantage of a cluster distribution function is that it could be calculated using the results of computer simulations.

Since the liquids studied here are molecular, it is important to account for the anisotropy of the distribution functions. We will compute the cluster distribution for molecular centers of masses and also cluster distribution for specific sites, and label them with the site name formula. The site-site clusters probability function $p^{XX}(n)$ are counted in such a way that the distance criteria are imposed on the distance between the sites, while for $p^{cm}(n)$ the bonded pair is defined using corresponding distances between the centers of mass.

The site-site clustering could be related to like-site and unlike-site distributions. The computation algorithm used in these two cases differs, due to the nature of the counting. For like sites we used the standard Stoddard algorithm [76]. The unlike sites posed a new computational problem, therefore we created a new algorithm that calculates the cluster distribution of any unlike sites.

Several parameters were additionally checked before the starting analysis. In figure 7.9 we present the cluster size probability function as a function of cluster size calculated over a different numbers of configuration.

Even for the small size system (N=256), averaging over 100 configurations gave reliable results. Our choice for further calculations were system sizes N=2048 and the number of configuration used in the averaging process was more than 100. The time sampling of the configuration was 20 ps. This relatively large time ensured that the clusters had been broken and reformed elsewhere from their previous configuration.

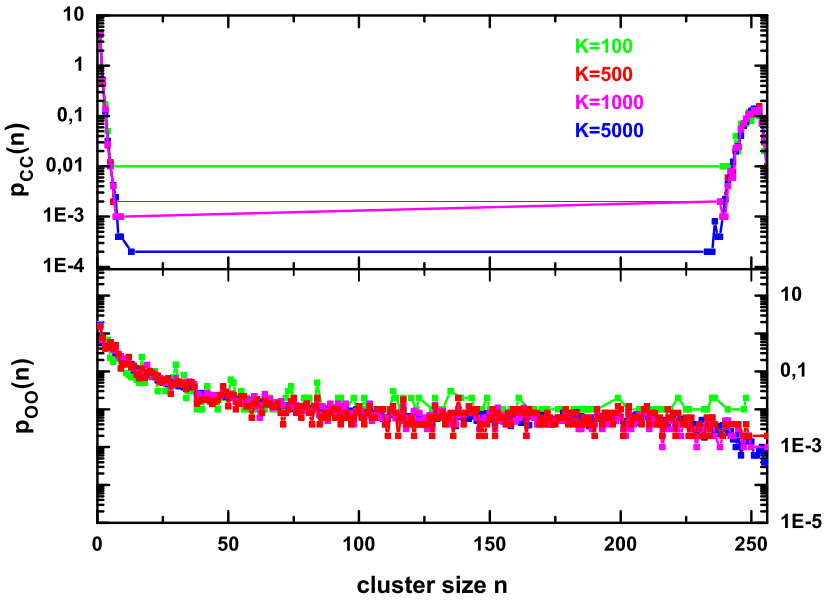


Figure 7.9: Cluster size probability averaged over $k = 100, 500, 1000, 5000$ configuration, using same cut-off distance, of OPLS methanol system with size N=256. In the upper panel we show the carbon-carbon, and in the lower panel oxygen-oxygen site-site cluster distribution.

The results for a Lennard-Jones liquid are pictured in the figure 7.10. It is intuitively apparent that the single molecular cluster n=1 will be the most important and the $p\,(n)$ will decrease very fast as n increases. The theory of percolation has exact results for cluster size probability for points distributed on a 3D lattice and numerical results are known for Lennard-Jones-type fluids [77]. The percolation is used to describe the systems which can be spanned using the connectivity criteria, and the density for which this feature appears is called the percolation density. The simplest case of percolation is when the cluster distribution comprises the clusters that are

approximately equal to a total number of particles. At the percolation density the cluster probability obeys the power law $p(n) = An^{-z}$, where the exponent $z$ has universal value [78]. The cluster probability decays more slowly in the dense liquid phase than in the gas or fluid phase, but even then, the decay in the percolation phase is exponential. Heyes and Melrose got for the Stilinger cluster in a Lennard-Jones system $z = 2.1 \pm 0.1$ [79]. The cluster distribution depends, also on the value of the cutoff $l_c$. Figure 7.10 displays expected results: for small cutoff-distances only the small size cluster span the system; as we increase $l_c$ the bigger-size clusters appear. This agrees with the shape of the corresponding radial distribution function in figure 7.1. For a dense packed LJ liquid the majority of particles is connected if the connectivity is in the range of the first minimum. This shows, also, the ambiguity of the definition of the percolation, where all systems could be classified as percolated if the connectivity criteria is large enough.

What do we expect for the associated liquids? There has been considerable investigation for water, and the results showed that the cluster size probability function decays exponentially like in ordinary liquids. It is a puzzling result considering that water is highly associated by hydrogen bonding. We show that the cluster-size probability for pure alcohols have more specific features.

**Local one-body distribution function of methanol**

In figure 7.11 we present the results for cluster size probability as a function of cluster size for pure methanol OPLS and WS model. The right panel shows $p^{MM}(n)$ for methyl clustering for the range of cluster sizes $n < 250$ and for two different values $l_c$. The values span the distances around the first minima of $g_{cm}(r)$ (refer to figure

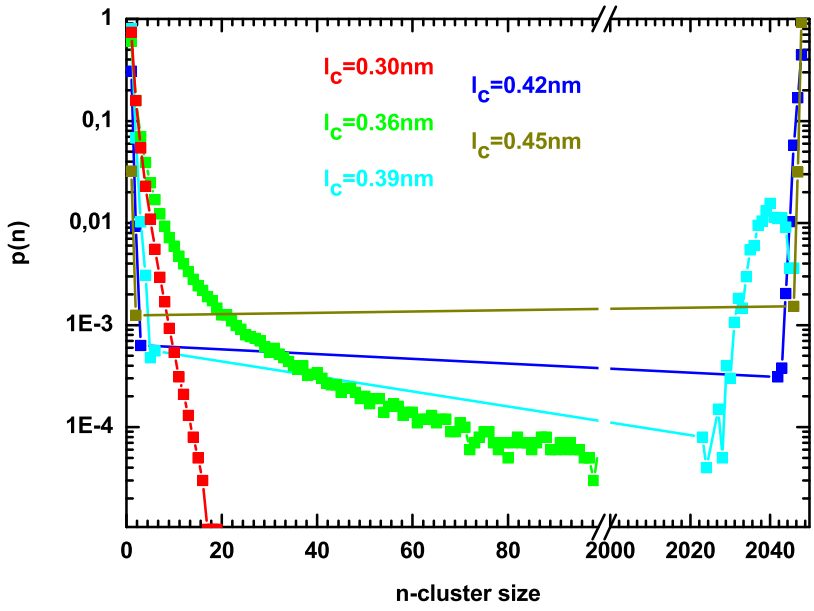


Figure 7.10: Cluster size probability function calculated on Lennard-Jones system N=2048 and averaged over $k = 100$ configurations using different cutoff values.

7.2). The curve follows the exponential-type decay for small cluster sizes. Clusters with larger n appear for larger values of cutoff $l_c$. This agrees with curves for the case of simple liquids. The left panel in figure 7.11 shows $p^{OO}(n)$ for the oxygen-atom clustering. The global shape is similar to that found for methyl clustering, however, we observe a clear bump around n=5. This feature is robust to the choice of $l_c$ for values around the first minima.

The high probability for cluster sizes around 5 indicates that these clusters appear more often than we would expect. This unusually high probability for specific cluster sizes highlights local structural preferences in our system. We note that the sizes

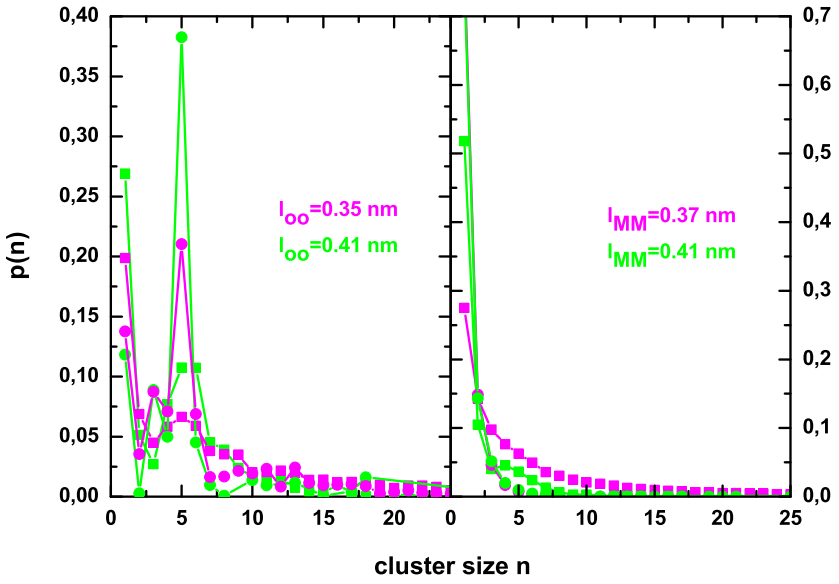


Figure 7.11: Site cluster probabilities for OPLS and WS methanol: Left panels: oxygen site clusters. Right panel: methyl site cluster. For each clustering the results for two cutoff distances is showed. The data for OPLS model is in the squares and WS model in the circles.

around 5 are preferential for all sites which are involved in hydrogen bonding as shown in figure 7.12. The hydrophobic site as well as the cluster distribution for center of mass (figure 7.11) behaves similar to LJ-type liquids.

We also note that two different models for methanol produce the similar clustering information as presented in figure 7.11. In view of their differences in thermodynamical properties from table II, we can conclude that local heterogeneities observed in local structural preferences are inherent features of liquid menthol.

Methanol has recently attracted interest in computer simulation studies of its mixture with water. Some of these studies indicate that water-methanol mixtures form

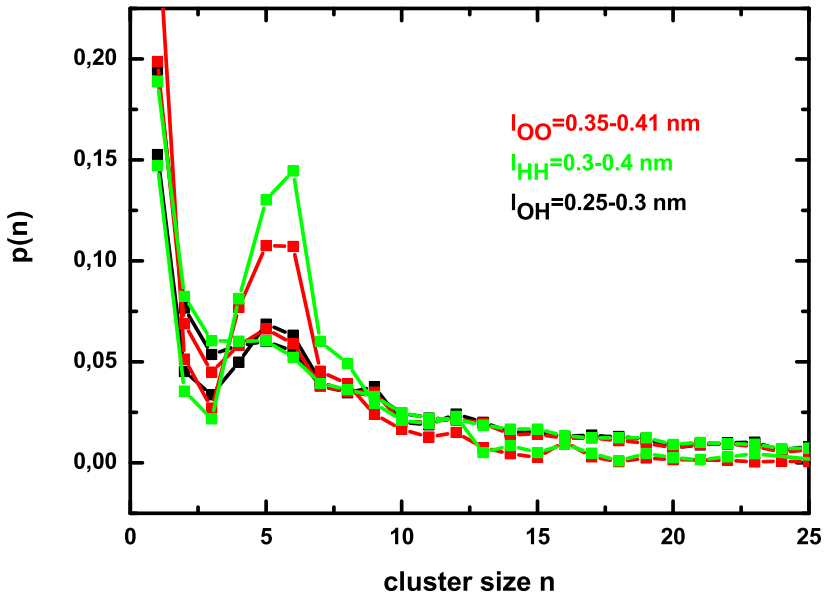


Figure 7.12: Site cluster probabilities for OPLS methanol for hydrogen bonded site OO (red), OH (black), HH (green) clusters. The distribution functions are shown for two cutoff distances.

bipercolating phases with local immiscibility [64]. Spectroscopical experiments show predominance of chains and rings with six and/or eight molecules in liquid methanol [22]. These findings are contested by the prediction based on the simulations. The Monte Carlo studies by Bako et al.[80] found no predominance of hexameric rings. They predicted that cycle structures comprise only about one third of total clusters in methanol. The average cycle size was predicted to be 4, as well as the average length of hydrogen-bonded chains. 15% of the methanol molecules existed as single molecules, eq. singeltons. The MD study by Allison et al.[63] predicted similar results. The most common size for a cycle structure was found to be 3 or 4, where 30%

of all clusters contain one or more cyclic motifs.

In line with this ongoing controversy we look in more detail at the structural motifs in oxygen clustering. The calculations of cluster size probability isolated the specific cluster sizes, but it lacked the information about cluster shapes. The structural motifs were investigated by our algorithm that screens previously defined clusters verifying if the cluster-bonded sites span close loops or compound close loops (comprising the smaller cyclic loops), or rather the chain-like open shapes. This analysis is rather robust dividing the cluster shapes only into two categories: open or closed shapes; however it is a good indicator for the preferential structural motifs. We specifically focus on the favorable size clusters, deliberately neglecting the total cluster distribution. Namely, the total cluster distribution comprised also "randomly" formed clusters. Therefore the more significant results are structural motives found in the clusters of favorable sizes. In the range n=3-7, we find that 81% of cluster form the open chains, and that only 19% of clusters appear in the form of close loops of connected oxygen atoms.

We could discuss whether these results contest or confirm recent findings. However, the cluster probability functions, as well as the favorable structural motifs are very sensitive to the means of description and applied calculations. It is clear that by using different simulation models, the results are similar, but not equal as we can see by comparing the cluster size probability for two different model of methanol. Therefore, we will rather emphasize, what all these results have in common: the local ordering on the scale larger than the first neighbors contact that corresponds to local heterogeneities in the macroscopically homogenous and disordered system.

Important question considers is , also, the dynamics of clustering. Pugnaloni

and Vericat proposed the time dependent cluster analysis of so-called chemical and physical clusters [75]. The implementation of this or similar methods will be part of our future investigation.

So far, we have not discussed the nature of clustering. The local ordering of the OO sites is due to the hydrogen bonding and all sites that form HB show similar patterns (fig. 7.12). Nevertheless, we have checked if the cluster size probability $p^{OO}(n)$ corresponds to that of hydrogen-bonded clusters. We use tighter geometrical requirements, suggested by Pagliai et al. [81], to eliminate unphysical clusters, including a OH-O angle smaller than $30^o$. This value is generally considered to be within the acceptable value for hydrogen bond. In 20 randomly chosen configurations, less than 2% of the bonds showed an angular deviation above $30^o$. This result testifies that a purely electrostatic modeling of the H bond is satisfactory. We considered also that the instantaneous topology of clusters is equally important as the average one; in particular, it is relevant to decipher the relations between molecular interaction and the average cluster shape, as well as its unambiguous detection and description from available statistical quantities. Also, it is known that the energy of various cluster shapes is not the same [71], therefore, the clustering should influence the thermodynamical properties of a system. Figure 7.13 shows some typical clusters found in our simulations.

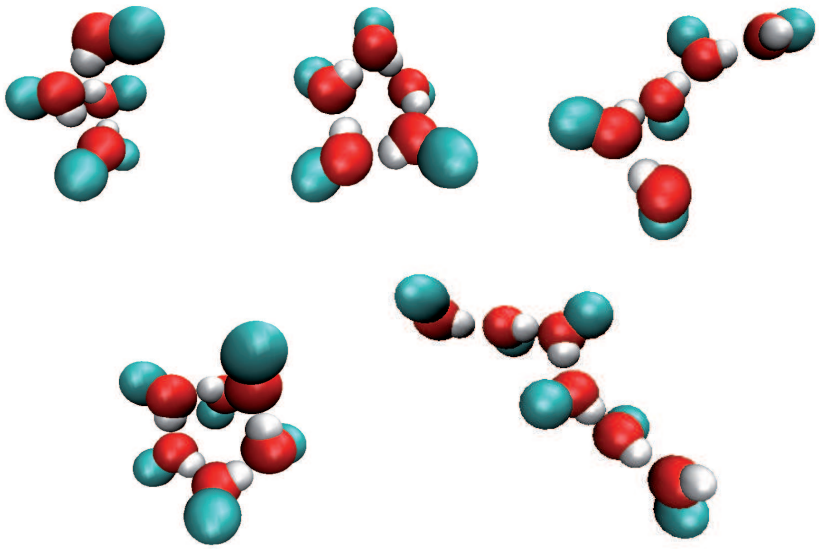

Figure 7.13: Some characteristic clusters for OPLS methanol [32].

**Local one-body distribution function of tert-butanol**

Figure 7.14 shows cluster probability functions for the methyl sites, carbon sites and oxygen sites of tert-butanol. While all curves show usual clustering tendencies of ordinary liquids, we observe, in the left panel, that $p^{OO}(n)$ and to some extent $p^{CC}(n)$ (in inset) show radically different behavior at small n values. The distant peak has emerged at n equal to 4. The differences in comparison with oxygen clustering in the case of methanol are the following: the single site cluster probability at $n = 1$ is no longer the highest and the peak at the $n = 4$ has the highest probability. The bump in the $p^{CC}(n)$ shows weaker clustering than the oxygen clustering and is induced through geometrical constraints of TBA molecule. These features are robust against the choice of $l_c$ for values around the first minima in corresponding RDFs. The right panel shows no specific clustering for methyl-methyl clusters, hinting that these sites are randomly distributed, while oxygen and indirectly carbon clusters show specific associations.

We conducted an analysis of the OH-O alignment similar to that for methanol, with similar statistic, confirming again that the OO clusters correspond to H-bonded clusters. The shape classifications showed, contrary to that for methanol, that in the range of $n = 3 - 7$ clusters, about 65% of OO clusters come in close loops, and 35% form a chain structure. This result agrees well with the idea of globular clustering in the case of tert-butanol. A snapshot of a simulation cell of TBA for N=2048 particle is shown in figure 7.15. One can clearly see that the oxygen and hydrogen are segregated throughout the sample from the methyl group. This segregation is more apparent than that for methanol, due to the cyclic clustering that enhances segregation of sites.

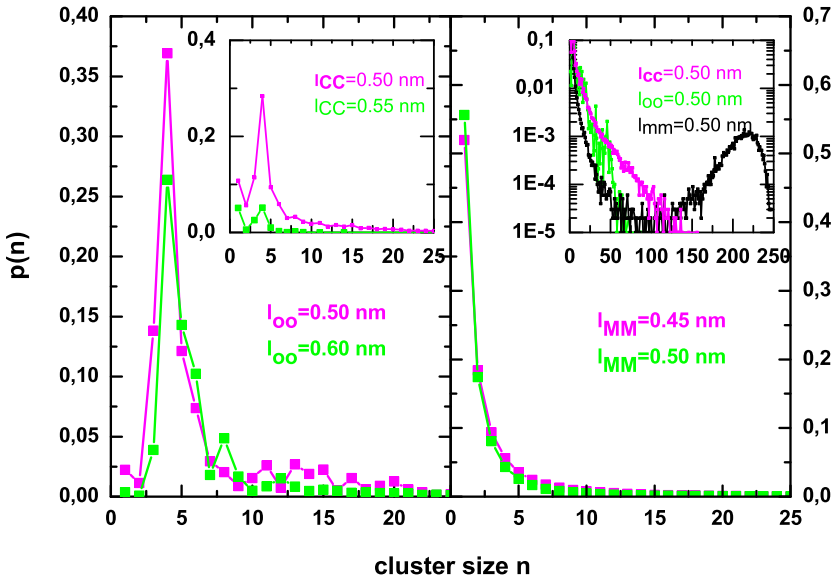


Figure 7.14: Site cluster probabilities for OPLS tert-butanol: Left panels: oxygen and carbon (inset) for two $l_c$ values. Right panel: methyl site clusters for two $l_c$ values. Inset shows all three probabilities for the whole system size at the same cutoff for N=256.

The underlying cluster structure is displayed in figure 7.16. We note that some clusters may come in interconnected two pieces that look like a chain of semi-open loops. These shapes strongly resemble the micelle formation, which is consistent with other findings.

We notice that the structural motifs for methanol and TBA differ strongly. The hydrogen bonding of hydroxyl groups has in both cases a liner character: On average the OH group is connected to two first neighbors. Therefore these differences are determined by the shape of the hydrophobic part of the molecule. Methyl group in methanol is small enough to accommodate more open shapes of HB connected sites.

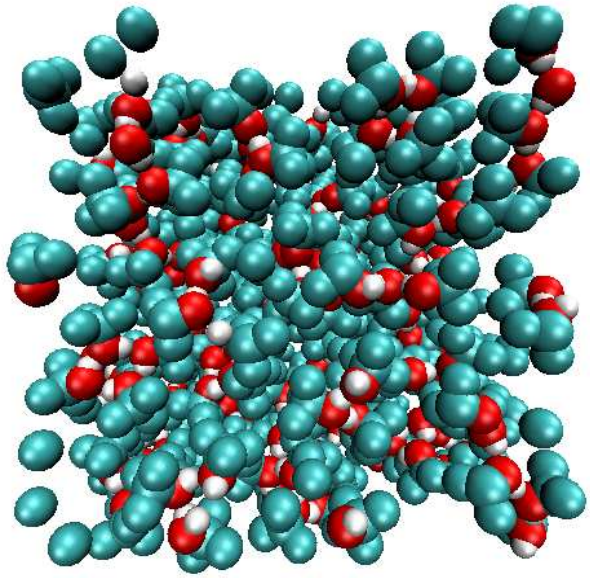

Figure 7.15: Snapshot of N=2048 OPLS tert-butanol system [32].

The globular shape of the methyl groups in TBA does not allow much flexibility, and in the resulting clustering the hydroxyl groups are buried into the core of clusters.

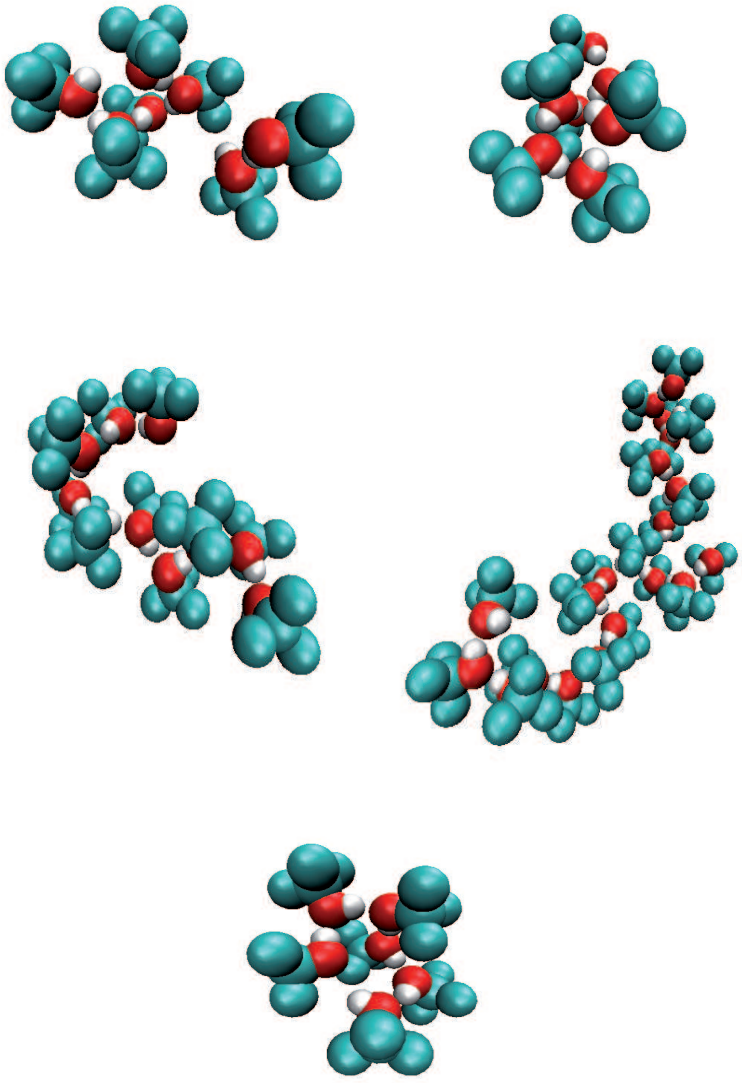

Figure 7.16: Some characteristic clusters for OPLS tert-butanol [32].

### One-body cluster distribution of the centers of mass

The analysis of the clustering as viewed from the center of mass of the molecules for both alcohols is equally interesting. Figure 7.17 shows that clustering is observed for the values of $l_c$ that lie around the first minima region of radial distribution function of the center of mass (figures 7.2 and 7.4). The $g_{cm}(r)$ shows distinct first peaks which correspond to the hydrogen bonding, and medium-to-large distances display liquid like oscillatory structure. This asymmetry is also present in the cluster probability function. The cluster bumps appear for a cutoff distance equal to HB distances in the $g_{cm}(r)$. Smaller or bigger $l_c$ produces a distribution of simple liquids.

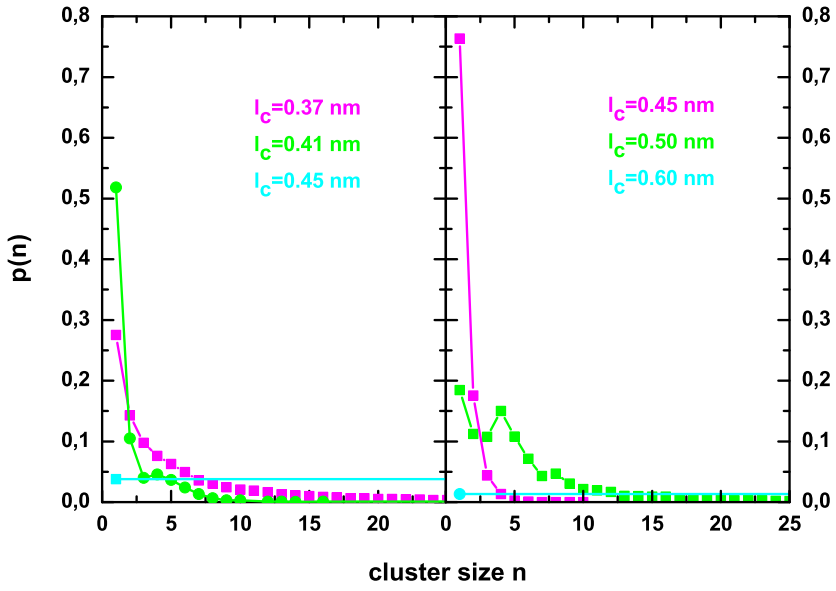


Figure 7.17: Center of mass cluster probabilities for different values of $l_c$. Left panel: methanol. Right panel: tert-butanol.

Despite the fact that the cluster peak is smaller than observed for the specific sites, it is remarkable that there is a noticeable clustering at the level of the center of mass. We note that a single molecule (n=1) has the highest probability for both alcohols, which indicates that the center mass clustering is weaker than for a specific sites, particularly for TBA. This observation shows that clustering is highly anisotropic, as should be expected for these systems.

### 7.3.3 Liquid water

It is equally interesting to accompany previous findings with those for water, since water is the most important associated liquid [82]. Figure 7.18 shows the structure factors for water-sites (see corresponding sRDF in fig. 5.1).

We notice that the water radial distribution functions display a similar behavior as hydrogen bonded sites for methanol; therefore we expect a strong anisotropic local organization. However, the structure factors show only a weak shoulder in $S_{OO}(k)$ at $k \approx 2.1\ \mathring{A}^{-1}$, which corresponds to the first peak in $S_{OH}(k)$. The wave vector $k \approx 2.1\ \mathring{A}^{-1}$ corresponds to $\sigma \approx 3.0$ Å which is close to the diameter of SPC/E water $\sigma_{SPC/E} \approx 3.6$ Å. Clearly this does not indicate any supra-molecular clustering or chain formation, because the $\sigma$ value is approximately equal to the size of the molecule. In line with this, the cluster distributions function in figure 7.19 shows only ordinary liquid features.

In literature similar analysis of water clustering point out to a strong pairing in tetrahedral configuration and possible clustering on the larger scale is not addressed [83]. Many percolation studies of water only stress the fact that water looks percolated, but not specifically clustered [19].

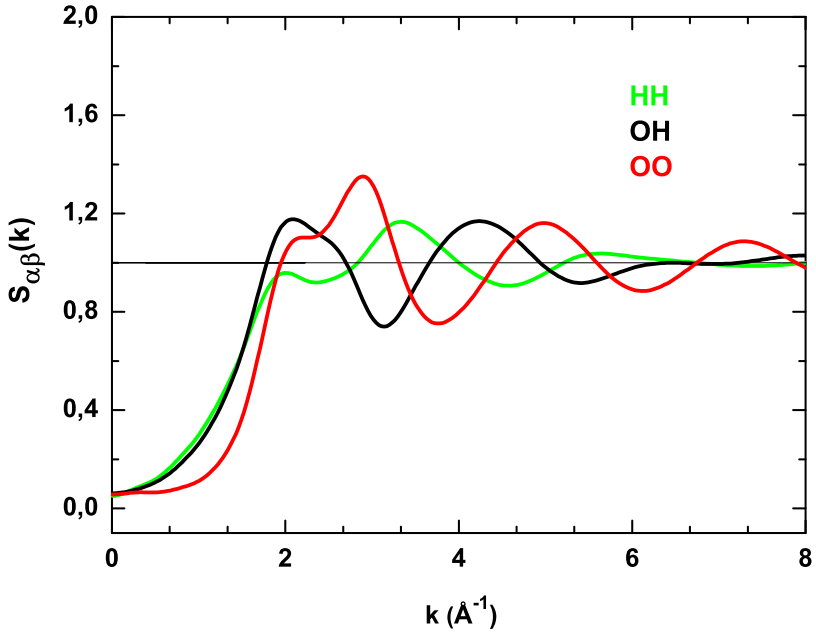


Figure 7.18: Structure factors for SPC/E water: OO (red), OH (black) and HH (green).

In view of these results water looks even more puzzling. The structure of water is mainly governed by the strong directional interaction, and excluded volume effects are negligible. However, the results show that hydrogen bonding in water does not induce formation of any specific association. The explanation may be that micro-clustering in water consists of topological conformations that do not get trivially sampled in the two quantities that we have explored herein. This may be due to the tetrahedrality of water hydrogen bonds that opens many bifurcations, and thus get sampled to nearly zero due to a global sphericity of the resulting distribution, indicating the network-like rather than cluster-like associations.

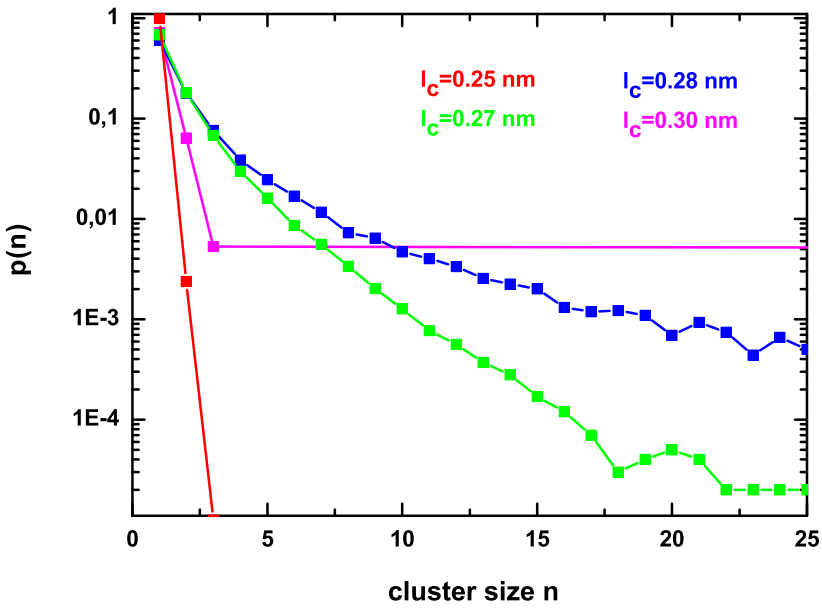


Figure 7.19: Oxygen cluster probability for SPC/E water for various $l_c$ values.

## 7.4 Discussion and conclusion

In the case of two simulated neat alcohols, namely, methanol and tert-butanol, we have shown that their liquid phases are partially or fully micro-structured under ambient conditions. We stressed the fact that despite being micro-structured, these liquids remain disordered, and far away from any phase transition instabilities. This is attested by the small value of the structure factor at the $k = 0$. We have shown that the site-site density correlation functions incorporate the information of inherent local structural patterns. The hydrophobic site functions followed the simple-liquid like features, while the hydrogen bonded sites correlation functions are modified according

to the specific self-associations of molecules. The analysis of the sRDF stressed out the differences in sites structuring and using the corresponding analogies, the features of sRDFs were connected to the specific clustering. The structure factor analysis further confirmed this differences in behavior. The structure factors showed two different structural periodicities, the first corresponding to the packing requirement, and the second to self-association of the sites on the scales larger than the first-neighbor contact. This is equally valid for the local one-body functions, cluster size probability, from which the preferable sizes of clusters were extracted. Both the site pair functions in reciprocal and direct space and clusters distributions exhibited features that point to the same micro-clustering characteristic. Namely, methanol is found to be weakly clustered, forming various chainlike patterns, while tert-butanol is almost entirely associated and forms a micelle-like primary pattern.

Examining the case of liquid water, we found no apparent clustering through the analysis used in this work, which is consistent with other authors' findings [19]. In that respect, it is interesting to note that the small-k behavior of the structure factor of water shows a notable increase at room temperature [20]. It may then indirectly point to some type of micro-clustering, which is not clearly probed through the cluster counting and the structure factor.

The important issue is also the distinction of the particle number fluctuation and local heterogeneity. Density fluctuation as well as concentration fluctuation are directly connected to the value of the structure factor at $k = 0$. Pre-peak that indicated the micro-structuring was on the small finite k-value. However, we noted also that micro-structuring is accompanied by the small enhancement of the density fluctuations. Therefore the micro-structuring may help to constrain the increase of

the fluctuations that will drive the liquid out of the stability range. However, the relation between these two features is still unclear.

In order to position these puzzling facts, it may be interesting to compare them to other cases where local order competes with global disorder. Let us start with liquid crystalline molecules that would form a nematic phase. Such molecules are strongly anisotropic and tend to align parallel to each other even in the dense disordered isotropic phase. This is related to pre-transitional phenomena. Their phase instability is driven by the $k = 0$ behavior of the orientational structure factor [84]. It is only near the spinodal that micro-domains are formed. Otherwise one goes from global disorder to global order through the ordering transition. Since, in our system we only have a local order, these liquid crystal systems do not provide a good comparison point.

Next, supercritical fluids also exhibit considerable clustering, which is even used to micro-segregate solute particles at industrial level [85]. However, these effects are entirely due to the strong concentration fluctuations near $k = 0$ and high compressibility because of the proximity to the critical point. These liquids do not exhibit any pre-peak, only a large peak near $k = 0$ which is due to the high compressibility of this state. So this case is also ruled out as a comparison point. We note that both these cases point to an underlying phase transition, which is totally absent from our case herein. Model fluids may be closer to our case: the dipolar hard sphere fluid DHSF, for example. The DHSF tends to form long chains[86] although this feature is not seen in the RDF. In particular, there is no strong main peak, and the structure looks more like that of a disordered fluid [87]. This may be due to the $1/r^3$ dependence of the directional interaction. Despite this, we would like to point out

that this fluid does not have a clear gas-liquid phase transition as it should. One of the interpretations[88] is that the chain formation blocks the condensation of a true liquid phase. The main feature that we retain from this model is that directional interactions can suppress the phase transition, and, like in our case, directional H bonding can create a stable phase with a considerable sustained microstructure.

# Chapter 8

# Micro-heterogeneity in water-alcohol mixture

The existence of nano-scale inhomogeneities of the aqueous solutions have become apparent in the last decade, thanks mainly to computer simulations. The mere idea that aqueous solutions are locally immiscible is very intriguing, since it questions the fundamental features of liquid solutions, which are homogeneity and disorder.

In this chapter, we will describe the micro-heterogeneous distribution of species in water-alcohol mixture. The introductory part contains the experimental findings that support the idea of local inhomogeneity. In subsequent sections, computer simulations are used to explore further this nano-scale heterogeneities. The analogy between micro-emulsion and aqueous alcohols is discussed in the closing sections.

## 8.1 Introduction

The research of the alcohol-water mixtures, at first glance, seems to be well advanced. Indeed, considerable knowledge has been gathered about the nature of the aqueous solutions (for example thermodynamical properties such as molar volumes, partial

molar volumes, densities, enthalpies, Gibbs free energies, chemical potentials, as well as dynamical quantities). Namely, the alcohol-water mixtures have been investigated using a variety of experimental techniques, and also several computation simulations have been aimed to explore these solutions (see chapter 6). In spite of all this research, we still lack the physical insight that would explain the behavior of such systems. We can add one more reason why studying the alcohol-water mixture is so important: they are the simplest molecules that show the hydrophobic effect, which is the basis of many biologically important processes such as folding of proteins or construction of membranes.

One of the puzzles is that when a simple alcohol is mixed with water, the entropy of the system increases less than expected for an ideal solution of randomly mixed molecules [89]. This effect has been explained by the fact that the hydrophobic head-groups create a ice-like or clathrate-like structures in the surrounding water [90]. However, the recent findings suggest that the alcohol molecules in a water mixture cluster together [21].

Using neutron diffraction with H/D isotope substitution Dixit at al. [21] probe the molecular-scale structure of a 7:3 molar methanol-water solution. The data indicates that most of the water molecules exist as small hydrogen-bonded strings and clusters in a "fluid" of close-packed methyl groups, with water clusters bridging neighboring methanol hydroxyl groups through hydrogen bonding. This behavior suggests that the anomalous thermodynamics of water-alcohol systems arises from incomplete mixing at the molecular level and from retention of remnants of the three-dimensional hydrogen-bonded network structure of bulk water [21].

This finding has been further confirmed by the report of the Guo et al. [22].

They investigated the structural properties of pure liquid methanol and methanol-water solutions using X-ray absorption (XA) and selectively excited X-ray emission (XE). XA and XE spectra reflect the local electronic structure of the various conformations, for example, the oxygen line shape is sensitive to the hydrogen bonding configurations. They found that a combination of pure liquid XE spectra reproduces the equimolar solution XE spectrum almost in detail. Due to the fact that these spectra are very sensitive to the changes in the local electronic structure, this observation indicates incomplete mixing at the microscopic level. They point out that any appreciable amount of "free-swimming" water molecules without hydrogen bonds would have given a completely different result. The XA spectrum of solution shows similarities to the spectra of the two pure liquids, but with distinct structures prior to the main absorption edge which is absent in the spectra of both pure liquids. These structures directly reflect the local electronic interaction between water and methanol molecules. They concluded, that specific bonded structures, involving both water and methanol, are responsible for this spectral behavior, and this feature is explained using the model where water molecules bridge methanol chains to form rings [22].

The deviation from the ideality of water-alcohol mixture is best expressed in terms of excess functions (see chapter 3). It is known that the concentration and temperature dependencies of the excess functions for alcohol solutions are quite complicated [91], part of it is related to a weakening of hydrogen bonds with the increase of temperature. The water-alcohol mixtures, also, have unusually large negative excess volumes and positive excess heat capacities. The current explanation is that the large negative excess volumes are due to a solute-water bonding, and the positive heat capacities

are due to the increase of the orientational freedom of both components as the temperature is raised [92]. A large negative excess volume indicates that the water and alcohol are packed tighter to each other than they are in neat systems. If we suppose that water and methanol are randomly mixed, than it is not clear how they can have such large excess volume, since the hydrophobic part of alcohol does not "like" water and in the process of mixing this would necessarily cause an increase of the excess volume. Therefore, the explanation that the large negative excess volume is just due to solute-water bonding is, clearly not sufficient.

The water alcohol mixture has three regimes of mixing: at the high alcohol concentrations the structure is governed by the H-bonded alcohol aggregates; at high water concentrations the structure is dominated by the tetrahedral-like structure of water; and at the compositions in between the micro-inhomogeneous regions of both, the characteristic alcohol and water structures coexist and smooth continuous transitions between two regimes are observed. However, there is no consensus about what is the underlying microstructure. For example, there are two main leading views regarding the water distribution in a mixture with high alcohol concentration. The first defends the hypothesis of the water-rich micro-inhomogeneities in terms of small "water pockets", where water is localized in the hydrophilic regions formed by alcohols. The second speaks in favor of a more uniform distribution of water over the system, with the hydration of the hydrophilic -OH chains and of the enhancement of the alcohol self-associations, which is therefore a more loose structure, where the -OH groups are either bonded to water or forming clusters characteristic of alcohols. Recent simulation studies by Tomščić et al.[93] strongly support the second scenario. Namely,

their results show, for pure alcohols, the non-existence of the large hydrophilic regions;and in the case of alcohol solutions, no huge enlargement of the hydrophilic regions that would point out to a formation of water pockets was observed. This study corroborates well with our results which we will show in the later sections.

Wakisaka et al. [94] probe the cluster structure in solution studying the clusters isolated from liquid droplets by using the mass spectrometry. They found that molecular clustering changes with varying the alcohol-water mixing ratio and proposed the three regions of mixing: at the small alcohols proportion the inherent water-cluster structure is preserved and alcohol molecules act as substitutions of water molecules in the hydrogen-bonding network of water; the middle region is characterized by layer structure and destruction of water network; and with the increase of the alcohol contents, the alcohol self-aggregation structure becomes dominant [94].

However, it was Koga [95], who pointed out first, that there are three concentration regions within a single-phase domain in aqueous solutions, in each of which the mixing scheme is qualitatively different from those in the other regions. Moreover, in his work, the crossover from one scheme to the next is associated with anomalies in the thermodynamic quantities that are proportional to the third or the fourth derivatives of the Gibbs free energy [95]. The similar idea, has been reported for the small mole fraction of tert-butanol [96]. Namely, the maximum of light scattering as well as maximum for the specific heat and the minimum for the compressibility has been found at the concentration 0.03-0.05 mole fraction of alcohol and these anomalies has been attributed to the fluctuation of the structure [96]. Nishikawa et al. [97] analyzed the tert-butanol-water structure at low concentration based on the concentration fluctuation obtained from the x-ray scattering. At the concentration about 0.04 the

formation of clathrate- hydrate $TBA(H_2O)_n$ is proposed, and as the concentration increase the formation of larger clusters of $TBA_n(H_2O)_n$. The analysis of the KBIs indicated that both species have tendency of the self-association. However, more convincing scenario, according to Nishikawa et al. [97], is the exitance of the free TBA regions and the $TBA_n(H_2O)_n$ domains. Bowron and Moreno [98] investigated the high alcohol concentration of TBA-water mixture using the neutron diffraction with hydrogen/deuterium isotope substitution. The results highlighted the creation of water pockets within the solution structure. Also, these results show that only a small amount of water is required to drive this solution in the direction of hydrophobic behavior, more commonly associated with the water-rich solution composition [98].

Therefore, it is the understanding of the microscopic properties, that we lack in order to uncover the behavior of the macroscopic features. The leading idea is that of the local segregation of water and alcohol which preserves the macroscopic homogeneity and disorder. This is close to the features of micro-emulsion, which have highly heterogeneous distribution of molecules that form aggregates such as micelles, but remain macroscopically in the homogenous phase. Micro-emulsions consist of a water component, an oil component and a surfactant, in which the latter forms the interfacial area between the two otherwise immiscible components. Clearly, they are more complex systems than a binary mixtures, but nevertheless, this analogy will prove to be valid for the description of the association of self-species in water-alcohol mixtures.

## 8.2 Simulation details

We have used the OPLS model for methanol $CH_3OH$ (MetOH) and equally the OPLS model for tert-butanol $(CH_3)_3COH$ (TBA) [7], together with the SPC/E model for water [9]. As in previous study all models are site-site models: model for MetOH comprises the oxygen site O, hydrogen site H, and one single site for the methyl group $M = CH_3$, TBA as six site model with the additional carbon site C, and SPC/E water with OW oxygen and two HW hydrogen sites. All details are presented in Table I (chapter 4). The computer simulations were performed in constant NPT ensemble with the DLPOLY2 package [6]. The number of particle was 2048, and the simulations were done for a nine different mole fractions from 0.1, 0.2 to 0.9. The conditions were fixed at T=300 K and a pressure of 1 atm, which were maintained through the Berendsen thermostat and barostat with relaxation times of 0.1 and 0.5 ps, respectively. A time step of 2 fs was taken. The sampling time for each simulation run was 64 ps, and number of runs were from four to twenty two (for equimolar concentration) for different concentrations of methanol in water-methanol mixtures. The alcohol mole fraction 0.2 in tert-butanol-water mixture was simulated for thirty two runs of 64 ps, and other concentrations for one or more runs of 64 ps.

We took additional care in the process of the equilibration of the systems. Namely, each concentration was equilibrated independently of the others. Starting configuration was constructed from the equilibrated configuration of the nearest larger alcohol concentration in which molecules of alcohol were swaped for the molecules of water to reach desired concentration. We started this process from the pure alcohol configuration towards the configurations of the smaller alcohol concentrations. Each system

was equilibrated for et least 64 ps.

Since, we are probing the local inhomogeneities, it is extremely important, to have well equilibrated system, in a sense that we do not have any memory of the initial configuration. However, this is not straight forward process, since the time and the space scale of the relaxation of the strongly bonded micro-heterogenous systems are not known in advance. Also, standard tests such as the convergence of the energy or volume is not sufficient to confirm the equilibration of the MH-systems, as we will point out in the case of water-tert-butanol mixture.

## 8.3 Theoretical considerations

The computation of RDF, structure factors as well as RKBI is done in terms of the site-site function as already defined in previous chapters. The site-site function measure now correlation between sites of each species, but also the cross-species correlations. Then, the RKBI is equal to:

$$G_{\alpha i\beta j}(R) = 4\pi \int_0^R \left( g^{(2)}_{\alpha i\beta j}(r) - 1 \right) r^2 dr \tag{8.1}$$

where $\alpha i$ and $\beta j$ sites corresponds to site of species $i$ and $j$, namely $\alpha i$ and $\beta j$ can be equal to any of H, O, C, M, OW and HW sites. Similarly, the structure factors, for sites $\alpha i$ and $\beta j$, are defined as:

$$S_{\alpha i\beta j}(k) = 1 + \rho_{\alpha i\beta j} \int dr e^{-ikr} g_{\alpha i\beta j}(r) \tag{8.2}$$

where $\rho_{\alpha i\beta j}$ is density of sites $\sqrt{\rho_{\alpha i}\rho_{\beta j}}$ and $\rho_{\alpha i} = N_{\alpha i}/V$.

## 8.4 Micro-heterogeneity in methanol

In the table III, we show the number of runs and total simulation time for each alcohol concentration, not including the equilibration-runs. The results presented in the next sections are taken from the last run for each configuration, accept in case of the $\chi_a = 0.5$ where the overall results are showed for the $1.08\,ns$ run.

- **Table III**. Number of runs and the simulation time for the $\chi_a$ alcohol mole fraction

| $\chi_a$ | Number of runs | Total simulation time $[ps]$ |
|---|---|---|
| 0.1 | 5 | 320 |
| 0.2 | 7 | 448 |
| 0.3 | 4 | 256 |
| 0.4 | 6 | 384 |
| 0.5 | 22 | 1408 |
| 0.6 | 5 | 320 |
| 0.7 | 4 | 256 |
| 0.8 | 7 | 448 |
| 0.9 | 5 | 320 |

### 8.4.1 The thermodynamical properties

In this section, we discuss the thermodynamical properties of the simulated system. We compare the OPLS model with the two other simulation studies of the water-methanol mixture: the simulation data from Ferrario et al., FHMK model [24] and

the Weerasinghe and Smith (WS) model [8]. We equally make relation with the experimental results, which are also shown in each figure [8][99][100].

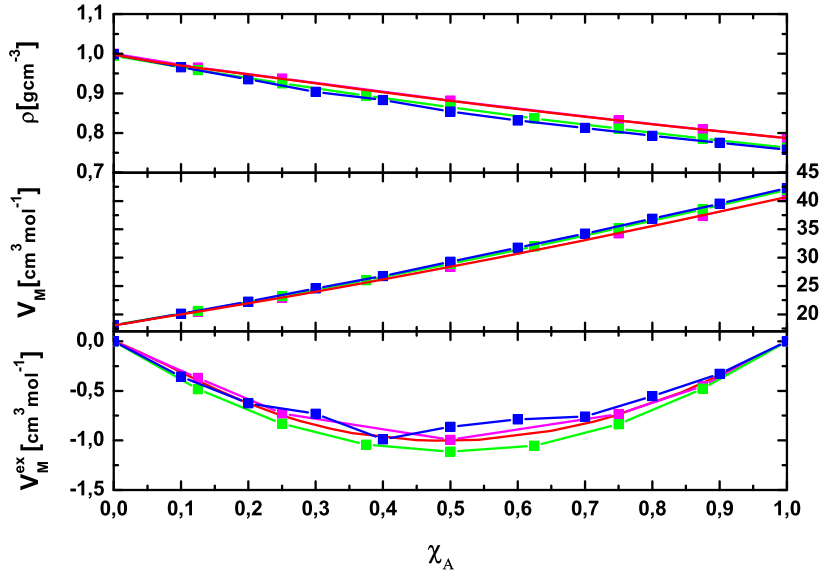


Figure 8.1: The concentration dependence of alcohol-water mixture for the OPLS model shown in blue, the FHMK simulation data in magenta, the WS simulation results in green and experimental values are shown as red line [100][99]: upper panel density; middle panel molar volume; and lower panel excess molar volume.

In figure 8.1, density as a function of concentration is displayed in the upper panel, molar volume in the middle panel, and variation of the excess molar volume with mole fraction is shown in the lower panel. The correct values of the excess molar volumes are closely related to the correct description of the microstructure since the excess volume directly probes the non-ideal contributions of the component

mixing. We notice that there is a general good agreement of all simulation data with experimental values. The FHMK model shows the best agreement, because the parameter of this models were defined in such a way to reproduce the experimental density and volume. We notice that WS and OPLS model overestimate the volume of the neat alcohol system, which is then translated to the small increase of the volume in the high alcohol concentration mixture. The excess molar volumes are also well reproduced by all models. Again the FHMK-curve is almost identical to the experimental function, while other data show small deviations. We notice that OPLS model show small jaggedness for the central concentration region.

The internal energy calculated in the simulation is compared usually with the heats of the solution that are measured experimentally. These latter quantities are in fact enthalpies resulting from the dilution of a given quantity of solvent in water. As far as we are concerned, we are more interested in the enthalpy difference between the solution and pure water. This enthalpy difference is related to the corresponding internal energy difference by $\Delta H = \Delta U + P\Delta V$, where $\Delta V$ is the volume change resulting from the dilution of solvent. The experiments are conducted at ambient pressure conditions, and the MD simulations are performed as close as possible to the liquid side of the liquid-vapor coexistence curve, that is, at zero pressure condition, and both the corresponding thermodynamical states are well below the critical point. Therefore, in the NPT-constant ensemble the $P\Delta V$ term is negligible when compared to the energy term, so the internal energy can be directly related to the measured enthalpy [101]. Since, the kinetic part of the internal energy is constant (in T-constant ensemble is $RkT$ term) the change of the internal energy of the simulated system is, than related to the configurational energy. Also, the correct SPC/E energies for neat

water are lower than the experimental ones by $\delta E = 5\, kJ/mol$, which corresponds to the water polarization contribution [49]. Following an argument by Berendsen [49] we have corrected the displayed water component energies by this term by the following expression $E_{new} = E_{old} + \delta E(1 - \chi_a)$. The corrected results are shown in figure 8.2. We presented the concentration dependence of the configurational energy (the upper panel) and as well the excess configurational energy (the lower panel).

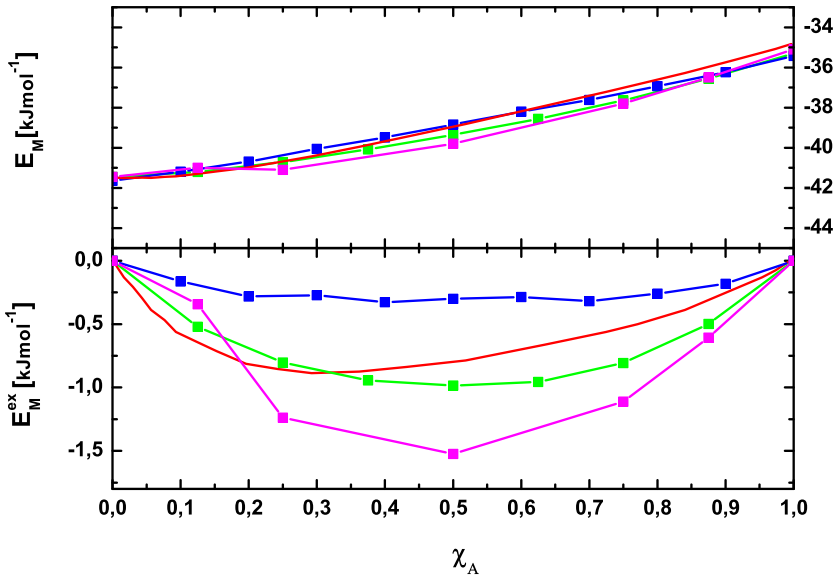


Figure 8.2: The concentration dependence of alcohol-water mixture for three different simulation data and the experimental results [23]: upper panel configurational energy; and lower panel excess configurational energy. Color convention is the same as in figure 8.1.

The agreement of the simulation and the experimental results is less good than in the case of the volumes and the densities. The largest deviations is observed for the FHMK model. Namely, the FHMK model was designed to reproduce experimental

density, which then resulted in a less good reproducibility of the configurational energy. This emphasize the fact that it is difficult to produce a model that is equally good in reproducing all the thermodynamical properties of the systems. The important point is also the large variation of the simulated excess energies between the results of the different models and as well as between the simulation data and the experimental measurements. We notice that none of the models are able to reproduce the correct value or the correct position of the minimum of the excess energy. All simulation data shows the symmetrical distribution with the minimum at the equimolar concentration. The experimental data have a shifted minimum at the alcohol mole fraction around $\chi_a = 0.3$. The literature offers several explanation why the excess enthalpy function is highly asymmetric, however there is still no consensus. This is one more puzzle of the aqueous solutions. Even more important question is what the alcohol or water models "miss" in order to reproduce correct excess energy.

Let us take a closer look at the energetical contributions of the OPLS simulations. In figure 8.3, we show in the upper panel the Lennard-Jones (LJ) energy, the Coulomb energy and their sum which is the total configurational energy, together with the experimental data; and, in the lower panel corresponding excess energies. The excess Coulomb energy displays the smooth symmetrical curve. The excess LJ energy is close to zero, and it has positive contribution for the central mole fraction region. Therefore, we are tempted to explain the jaggedness of the excess volume using the arguments that consider Lennard-Jones energy. In the system which has non-bonded and bonded sites resulting LJ energy is a sum of the two opposite contributions that in terms of the excess energy correspond to: negative excess energy for the non-bonded sites and positive excess energy from the bonded sites. In our case, this leads to the

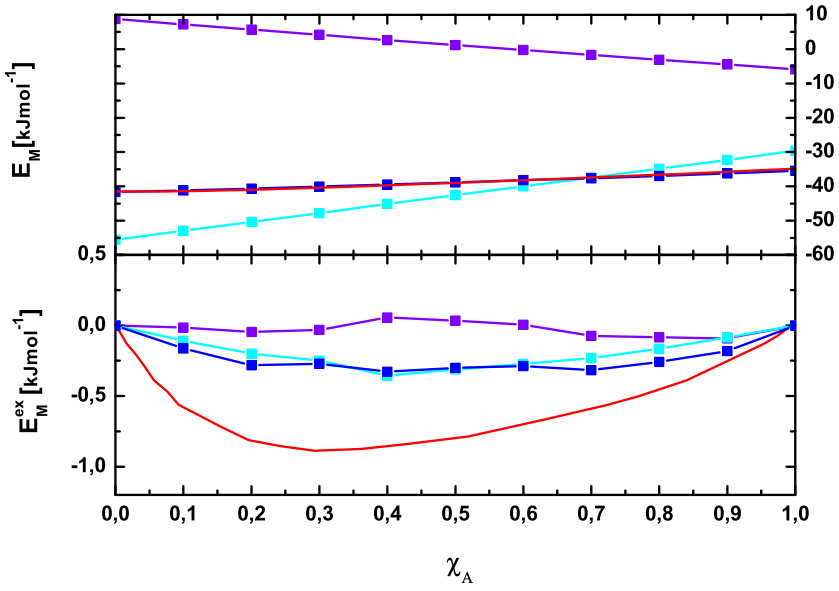


Figure 8.3: The concentration dependence of alcohol-water mixture for the OPLS model: the Lennard Jones potential energy is shown in violet, the Coulomb energy in cyan and the sum of the LJ and Coulomb contribution, the total configurational energy is shown in blue. The experimental values are shown as red line. The upper panel displays the energies and the lower panel corresponding excess values.

excess energy which is close to zero. The positive excess LJ energy indicates that the molecules on the average are closer than the LJ radii, which leads also to a decrease of the excess volume. It is well known that, due to the strong electrostatic attraction, the average distance of the first neighbors in hydrogen bonded systems is less that the van der Waals radii [92]. Important question is why the excess energy becomes slightly positive in the central molar fraction region. If the system is homogeneous, than the excess energy should scaled according to the concentration, and the curve should be smooth. Therefore, we give a purely hypothetical explanation, starting

from the assumption that the system exhibits segregation of the species. The particles that are in the interfaces are even more confined, then particles which are in the species clusters, and they contribute to the positive part of LJ energy. Since, for the concentration around 0.5, the interfaces between specie clusters is the biggest, the excess LJ energy becomes slightly more positive.

### 8.4.2 The pair correlation function in direct space

Figures 8.4, 8.5 and 8.6 show concentration dependence of the selected site-site distribution functions: the upper panel shows site-site radial distribution functions, while the lower panel shows the corresponding running KBIs. The previous simulation results confirmed that these sites follow the center of the mass distributions, therefore there are good choice to monitor the component correlations.

We notice that the methanol-methanol RDFs look similar for all concentrations: all functions show liquid-like oscillations, and positions as well as size of the peaks do not vary much. Just looking at these functions, one may assume that the organization of molecules of methanol is not affected by the presence of water. This could be an indication of methanol being segregated in pockets. On the contrary, water clearly displays strong concentration dependence. The height of the first peak increases as we go from water rich towards water poor solutions, indicating the enhancement of the first-neighbor structure. This is, also, true for methanol correlations, but to much smaller extent.

Namely, RDF measures the pair-correlations normalized with respect to the uniform density of the species (equation 2.39). Therefore, the enhancement of the first peak indicates that the molecules tend to cluster, rather than being disperse over

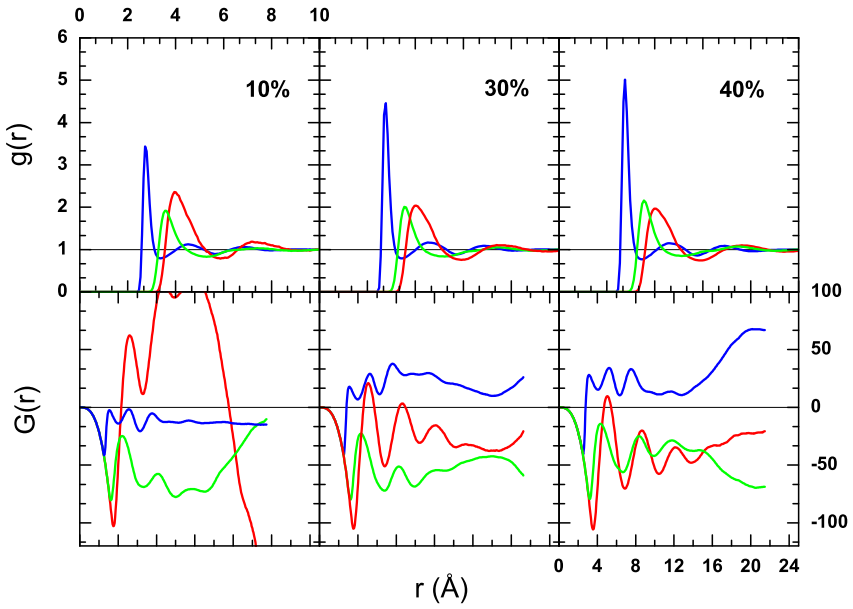


Figure 8.4: The concentration dependence of the radial distribution functions are displayed in the upper panels, and the corresponding RKBIs in the lower panels. The methanol correlations are shown in red, water functions in blue, and cross-species correlations in green. The three vertical panels correspond to the alcohol mole fractions 0.1, 0.3 and 0.4 respectively.

total volume. The increase of the RDF peaks as the concentration of water decrease shows that molecules of water will self-associate even for small concentrations. As expected only the height of the peaks changes and not the positions which are equal to a standard HB 2.5 Å.

Cross-species RDFs show expected behavior: the correlations slowly increase as the water content decreases.

In order to access the HB correlations in methanol, in figure 8.7, we display

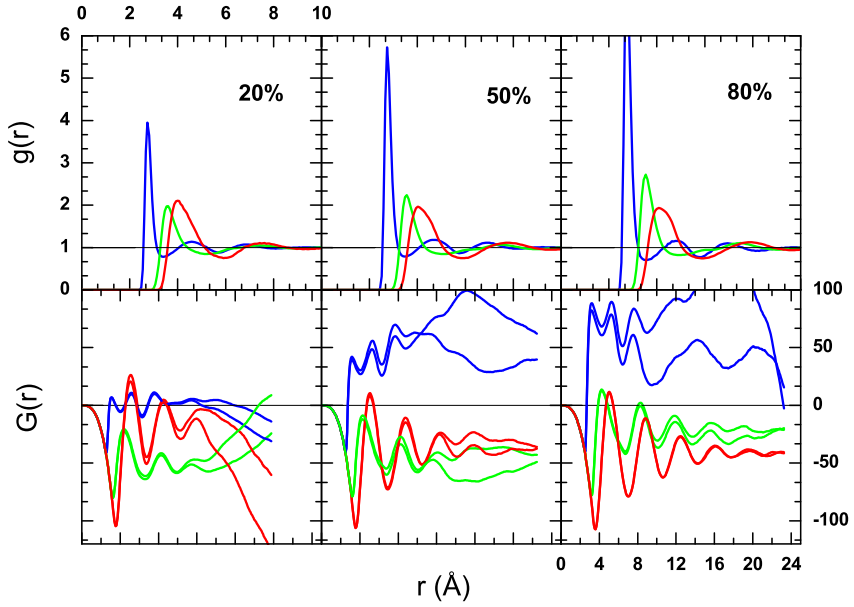


Figure 8.5: The concentration dependence of the RDFs and RKBIs, with the same color convention as in Figure 8.4. The three vertical panels correspond to the alcohol mole fractions 0.2, 0.5 and 0.8 respectively. The RKBIs are shown for two different runs.

the methanol oxygen correlations, as well as water and cross-species correlation (O-water)-(methanol-O). The hydroxyl-oxygen correlations follow the same trend as water correlations. For low methanol concentrations the correlations between oxygen sites are very weak, and the self-clustering of oxygen sites increases as concentrations of methanol increase. This is an expected result, since, in water rich solutions, the competition for HB between water and methanol destroys the chain-like OO clustering, and only as the concentration of methanol increases the oxygen sites restore anisotropic self-association.

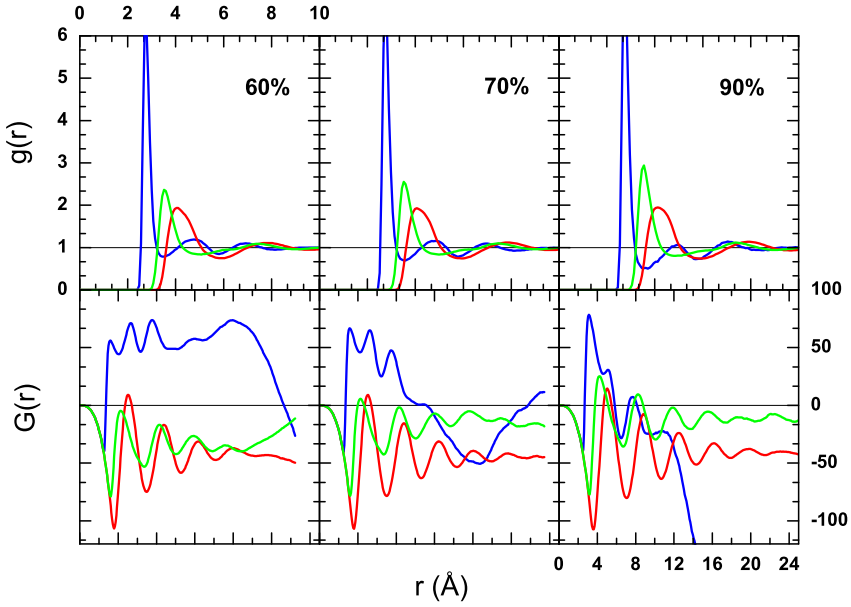


Figure 8.6: The concentration dependence of the RDFs and RKBI, with the same color convention as in Figure 8.4. The three vertical panels correspond to the alcohol concentrations 0.6, 0.7 and 0.9 respectively.

The RDFs information is easier to understand within short range distances, since the long-range correlations are hidden in apparently shapeless tails. Therefore, the medium-to-long range correlation will be more noticeable in RKBI, where the small variations will be amplified in the process of integration.

What do we expect starting from on the previous findings? The simulations of pure systems show that the convergence of RKBI is achieved within 3-4 molecular diameters. Therefore, we expect that all sites RKBIs will tend smoothly towards a well defined KBI value. Deviations may appear, only for the component in the low concentration region, that is purely for the statistical reasons.

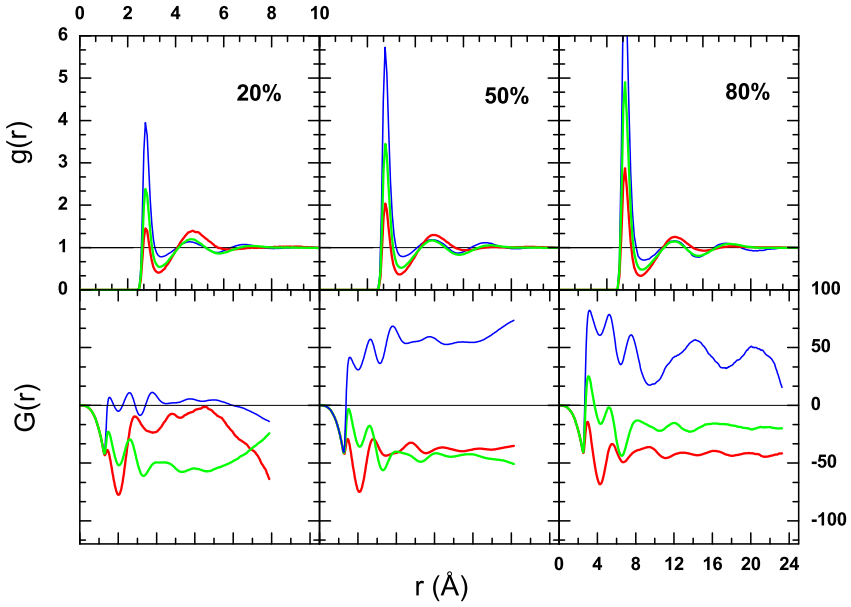


Figure 8.7: The concentration dependence of the RDFs and RKBIs. The oxygen-oxygen methanol correlations are shown in thick red line, the water oxygen correlation in thin blue line, and the cross correlation (O-water, O-methanol) are shown in green line. The three vertical panels correspond to the alcohol mole fractions 0.2, 0.5 and 0.8 respectively.

As expected, when the solute component is majority, its RKBI displays the regular long-lasting oscillations around well defined value of KBI, and small deviations from this behavior are observed for the low methanol concentration. This is equally valid for the methyl and the oxygen sites, therefore it is a feature of all methanol-site correlations. However, the water RKBIs show unexpected results. For almost all concentrations, medium-to-long range tail of RKBI displays oscillatory-like behavior comprising half ($\chi_a = 0.6 - 0.5$ of water), one ($\chi_a = 0.4$) or even two periods

($\chi_a = 0.2$). These results show that water and methanol apparently have two different kinetics of self-associations in water-methanol mixture. Namely, while the simulation time was large enough for methanol functions to reach uniformity, this was not sufficient for the water function to converge.

In addition, in figure 8.5 for solute mole fractions 0.2, 0.5 and 0.8 we present two successive runs. Different statistical runs give the more or less same results for the methanol functions, which is compatible with the idea of reproducibility and equilibration of the simulated system. The variations are present only for small solute concentrations. On the contrary, different runs for water functions displayed totally different tail variations for all concentrations.

Usually one expects the problems with the sampling for the components with a smaller number of particles, which is even more valid if these particles tend to cluster. However, it seems that for water the reproducibility of runs is highly perturbed for almost every concentrations. Each of these functions has been calculated by sampling every 0.04 ps for runs typically over 64 ps, which represents 1 600 configurations. This is quite large when we compare the sampling time needed for a simple liquids (1000 samplings is generally enough), but it is 5 to 10 times smaller (on a 1-2ns run-times basis) than what seems suggested by recent calculations by other authors [47][35] for mixtures of associated liquids.

Therefore, we simulated the equimolar concentration with the total simulation times around 2 ns. The correlation functions were accumulated for 1 ns run, which was followed by successive runs of the 64 ps until the total time of simulations was around 2 ns. These results are pictured in figures: water-water functions in 8.8; and in figure 8.9: in the lower panel solute-solute functions, and in the upper panel

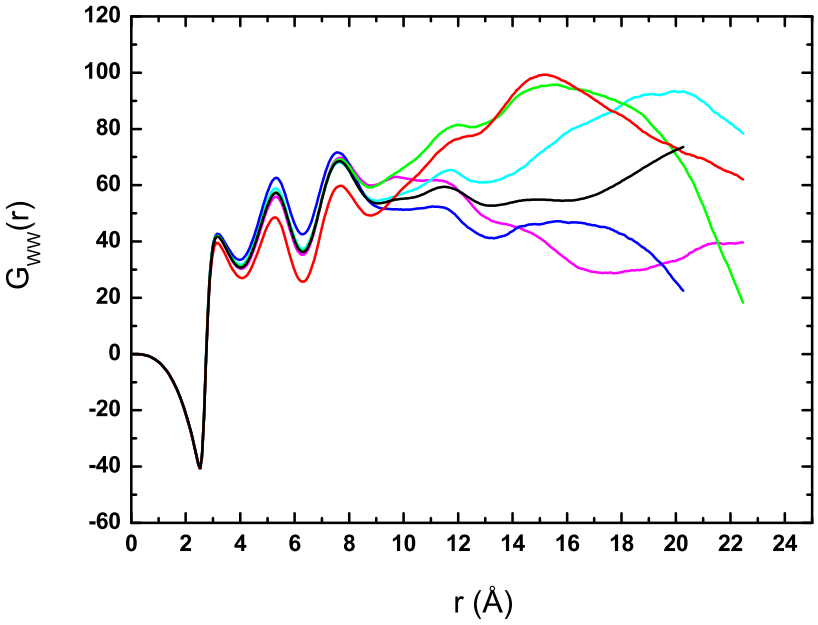


Figure 8.8: The water-water RKBIs for the equimolar concentration using different simulation runs. The black line shows the results from the 1.02 ns sampling time and the other curves are average values from successive simulations runs of the 64 ps. The order of the runs is following: black-red-green-blue-magenta-cyan line.

cross-component functions.

Again, the solute-solute correlation do not differ much considering the results of different runs. On the contrary, the water-water functions oscillate above and bellow the data of the longest run of 1 ns. Therefore, if we applied longer statistics the erratic behavior of RKBI will eventually vanish, and this large sampling will reproduce the correct asymptotic limit. This agrees also with the experimental findings that KBI has the well defined experimental value. Thus, we agree that it is necessary to have extended simulations for water mixtures in order to reach the correct KBI. However,

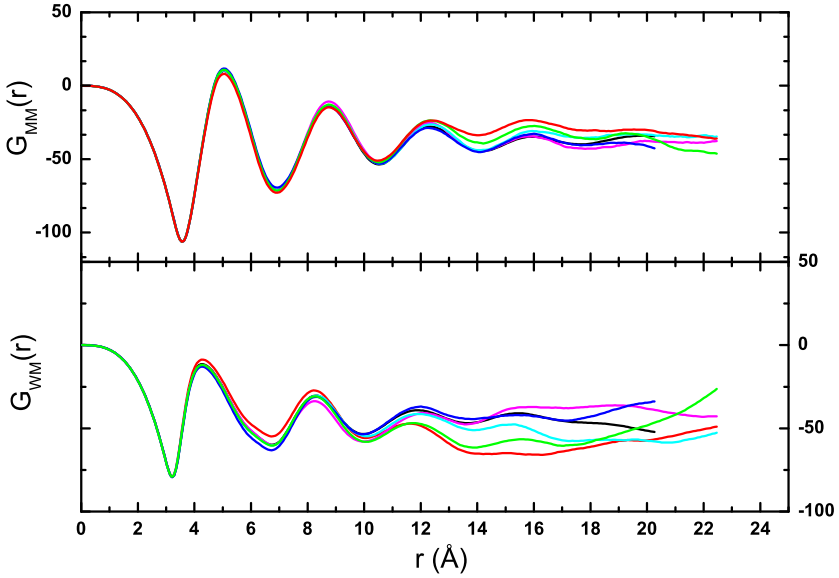


Figure 8.9: The upper panel displays the methyl-methyl RKBIs for the equimolar concentration calculated using different simulation runs, and the lower panel show same results for the cross-species functions. The color convention is the same as in figure 8.8.

we find more interesting to elucidate what is the reason behind this erratic behavior of RKBI.

The irreproducibility of different simulations runs in the case of water, can be explained by the fact that water is highly associated, and the different simulation runs actually sample over different realizations of these associations. On the contrary, the methanol RDF and RKBI show no signal of clustering, apart at the low methanol concentrations. The snapshots from the simulations are shown in figure 8.10. One sees clearly that methanol is homogeneously distributed, while water forms sponge-like structure. This, at first, is hard to imagine, since if one species is clustered, just

for the sterical reasons, the other species should also be clustered. We can suggest that methanol is homogenously distributed in such a way that it is bonded to the other molecule of methanol and to the water molecule, therefore, the majority of methanol molecules are also clustered, just some molecules of methanol are trapped in the water domains, which than results in the apparent homogenous distribution of methanol.

Therefore, micro-segregation of species is omnipresent behind the results shown here. It is, nevertheless, puzzling why water and methanol, both being hydrogen donors and acceptors, would not form a homogeneous mixture. The mixing is result of the competition of two types of hydrogen bonding, that between self-species and that between cross-species. Nevertheless, we do not expect that the small methyl hydrophobic sites will order the mixture as much as that found in a true micro-emulsion. Apparently, there is no reason why these species should form an uniform mixture. Nevertheless, the simulation results, as well as experimental results, indicate that the equilibrium structure of methanol-water is, actually micro-heterogeneous.

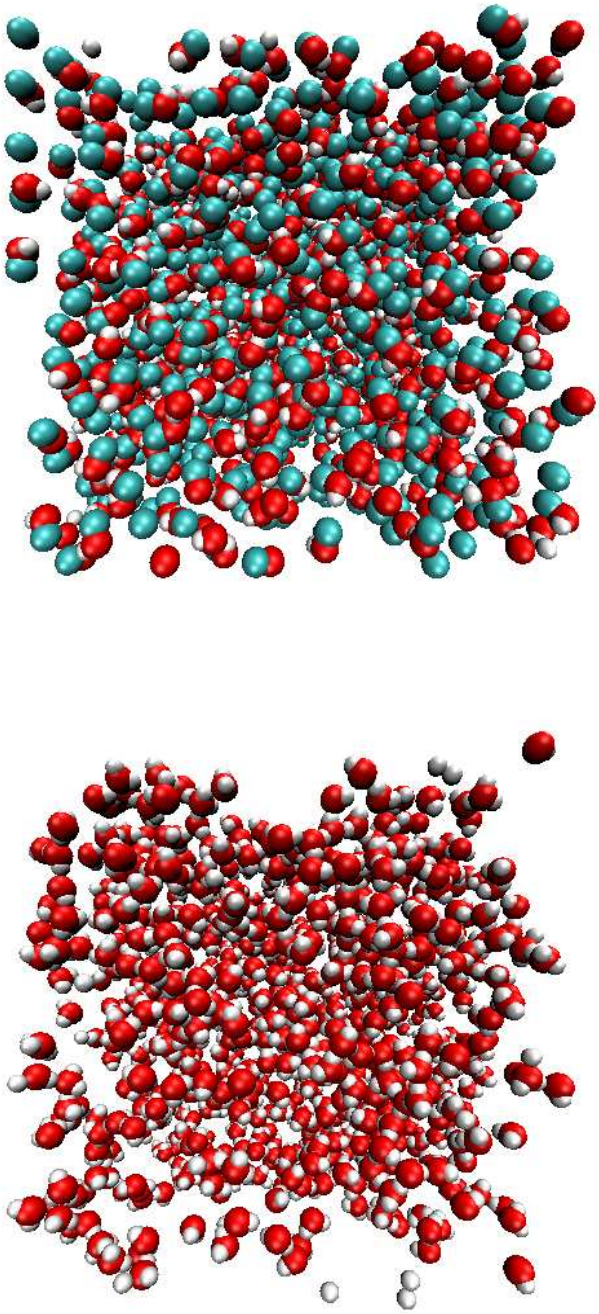

Figure 8.10: The snapshots of water-methanol mixture for mole fraction 0.5 and N=2048 particles. In the upper figure methanol molecules are displayed and in the lower water molecules [32].

### 8.4.3 The pair structure in the reciprocal space

The site-site structure factors are discussed in the following sections. We are particularly interested in behavior of small k-vectors, which provide the information of the large scale organization of the mixture.

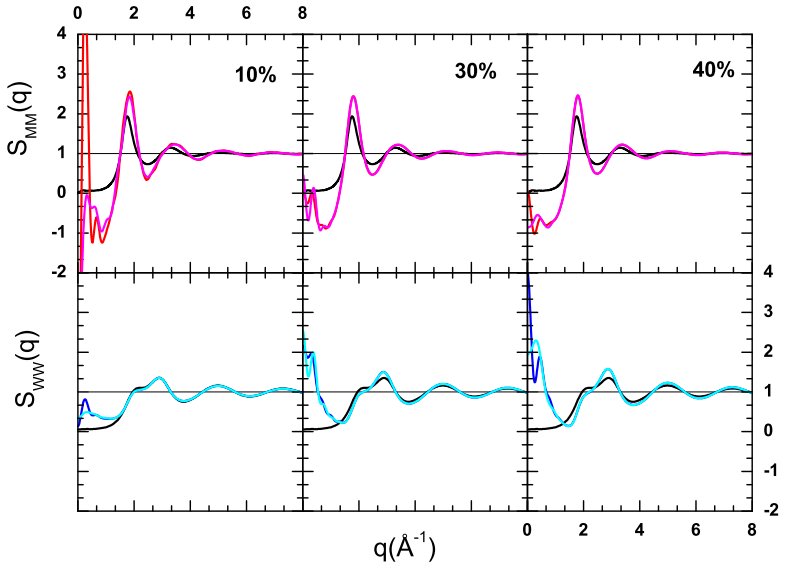


Figure 8.11: The concentration dependencies of the structure factors. The methyl-methyl functions are shown in the upper panels and the water-water structure factors in the lower panels. In each graph the results from two successive 64 ps long runs are shown: methanol functions are in red and magenta, and the water function are shown in blue and cyan. As referent case, the structure factors of each neat system are presented in black: in the upper panel the methanol structure factor and in the lower panel the structure factor of neat water. The three vertical panels correspond to the alcohol mole fractions 0.1, 0.3 and 0.4 respectively.

Figures 8.11, 8.12 and 8.13 show concentration dependence of the structure factors: the upper panel show methanol functions, and the lower panel shows water functions.

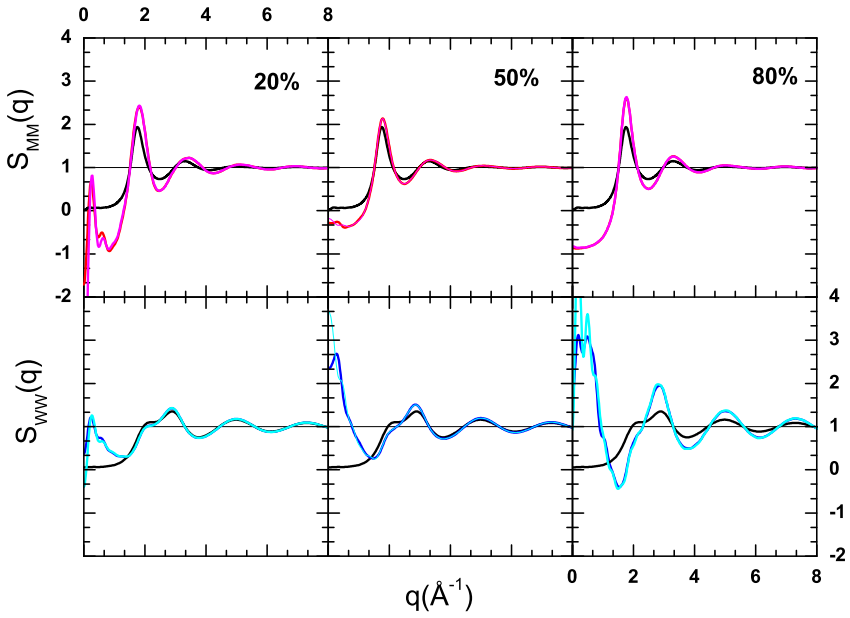


Figure 8.12: The concentration dependence of the structure factors. The functions and the color convention are same as in figure 8.11. The three vertical panels correspond to the alcohol mole fractions 0.2, 0.5 and 0.8 respectively. For the mole fraction 0.5 the results of only one run are displayed since, the detail run-evolution is shown in figures 8.17.

As before, we equally show two successive runs, together with the structure factor of the neat liquid (in black line), which helps to measure the departure from homogeneity due to the presence of the opposing component.

The differences between methanol and water organizations are even more apparent looking at the corresponding structure factors. Water structure factors, at all concentrations, show large peaks at the small k-vector. Similar peaks for methanol are only present in the low solute concentration. Again, the results from successive run are almost superposed for the methyl sites, and show large variation for all water

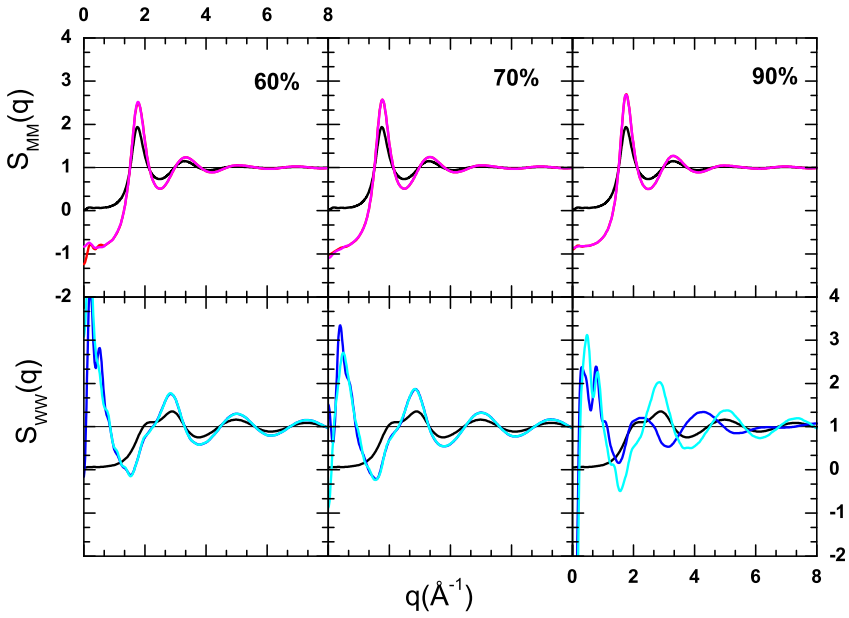


Figure 8.13: The concentration dependence of the structure factors. The functions and the color convention are same as in figure 8.11. The three vertical panels correspond to the alcohol concentrations 0.6, 0.7 and 0.9, respectively.

functions. Apart from the persisting peak at the small k-vector, the shape of the peaks in water structure factors also changes. Namely, small shoulder present in a neat system structure factor almost disappears for small water content.

The concentration dependence of the methyl-methyl structure factor are shown in figure 8.14, and in figure 8.15 oxygen-oxygen of hydroxyl group. We equally show how the water structure factors vary with the concentration in figure 8.16. The main peak of methyl site does not show any concentration dependence. However, this is not true for the structure factor of the hydrogen-bonding site. We also notice that the variation of the small k-vector peaks is similar for methyl and oxygen functions.

Therefore, the small k-vector behavior corresponds to the species features and the changes in pre-peak shows that methanol does not preserve same structure as in neat system. The water has a strong concentration dependence, which is seen trough the small-k variation and, also the change of the main peaks.

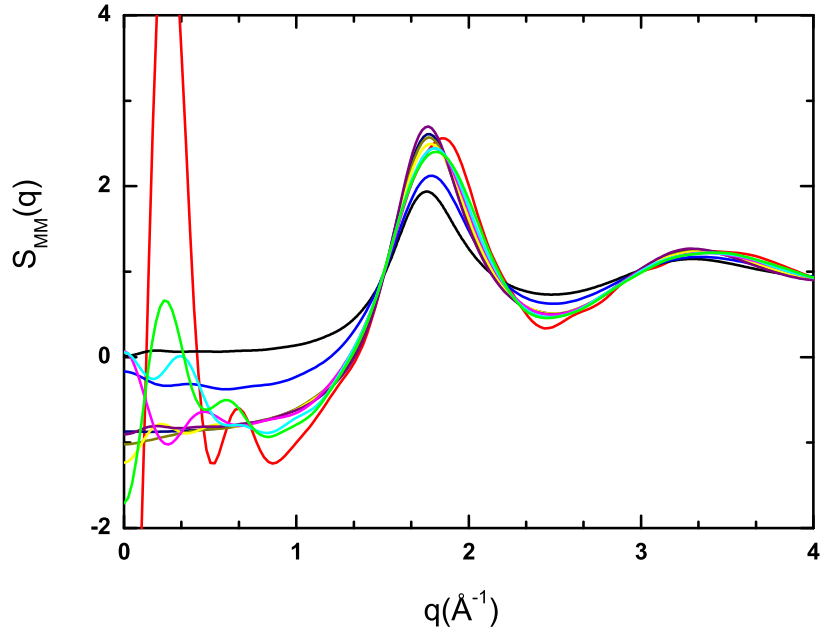


Figure 8.14: The concentration dependence of the methyl-methyl structure factor. The methanol functions are show for the mole fractions 0.1, 0.2 to 1.0., and the corresponding colors are red-green-cyan-magenta-blue-yellow-dark yellow-navy-purple-black. For example, result for the mole fraction 0.5 is shown in blue and methyl-methyl structure factor of the neat system is in black.

All extracted information from structure factors can be summarized as follows. The small k-vector peaks indicate the appearance of the large scale associations which clearly support the idea of self-clustering and micro-heterogeneity. Such peaks are a

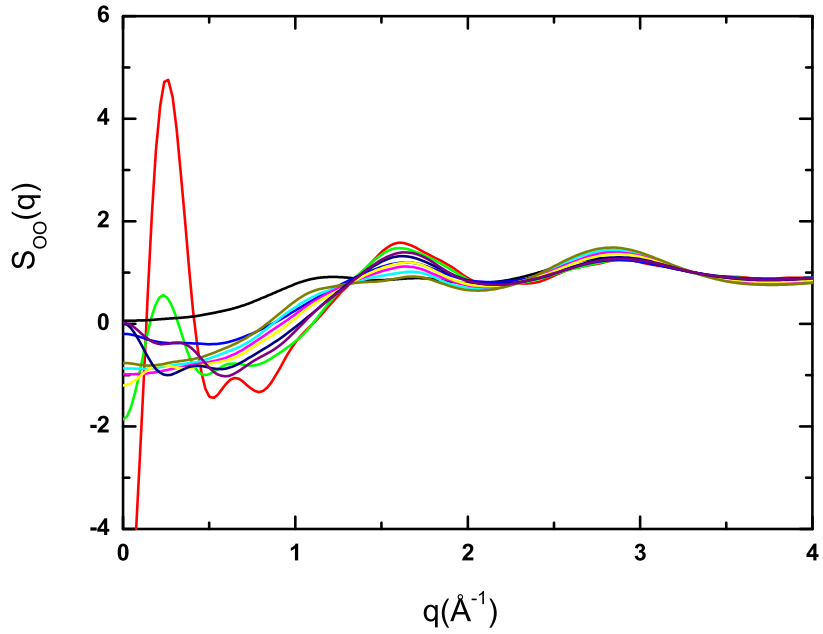


Figure 8.15: The concentration dependence of the oxygen-oxygen methanol structure factor. The color convention are same as in figure 8.14. Black line is the OO structure factor of a neat methanol

characteristic feature of all methanol-sites and water functions, therefore, they are an indication of species association. The fact that the pre-peak, or in the case of water the small shoulder in $S_{ww}(k)$, vary with concentration, directly asses the change in hydrogen bonding, showing that the microstructure results from the self-species and cross-species hydrogen bonding at the cluster interface.

Let us turn now to the data from the successive runs. Figure 8.17 displays the water structure factors for 2 ns runs for equimolar concentration.We notice that $k = 0.25\ \mathring{A}^{-1}$ corresponds to a $\sigma = 25.12$ Å, which is approximately equal to the

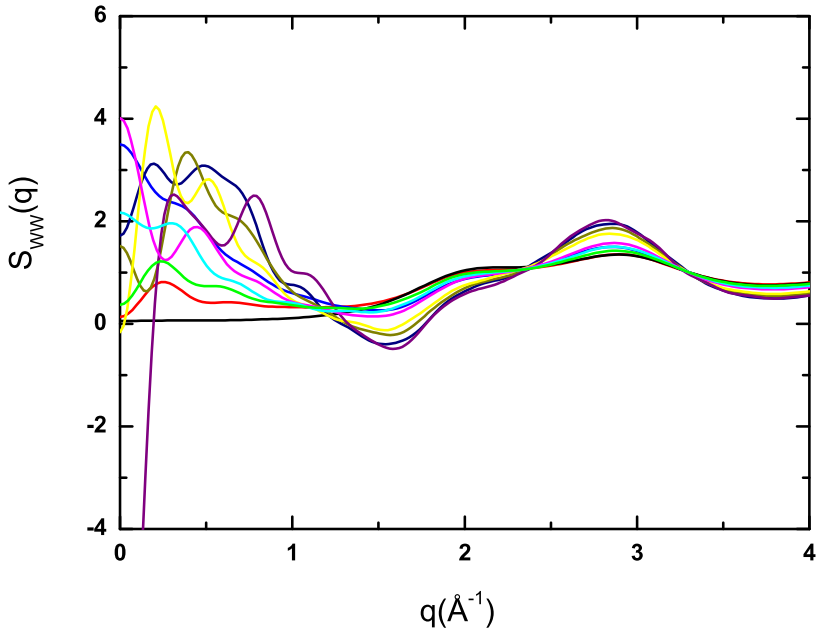


Figure 8.16: The concentration dependence of the water-water structure factor. The color convention are same as in figure 8.14, only difference is that the black line corresponds to the neat water correlation.

half-system sizes (the largest distances that we can measure). This, at first, may be interpreted as if the Fourier transformation was not done properly. However, if this was the case indeed, than other structure factors would have the same fault. Therefore, we conclude, that the small-k behavior of $S_{ww}(k)$ is necessarily due to the water self-association.

Again, we can attribute the peaks from different runs to specific realization of sponge-like water clustering. The remaining question is: will this signal disappear if we use additional statistics? We notice that even for the longest run (shown in

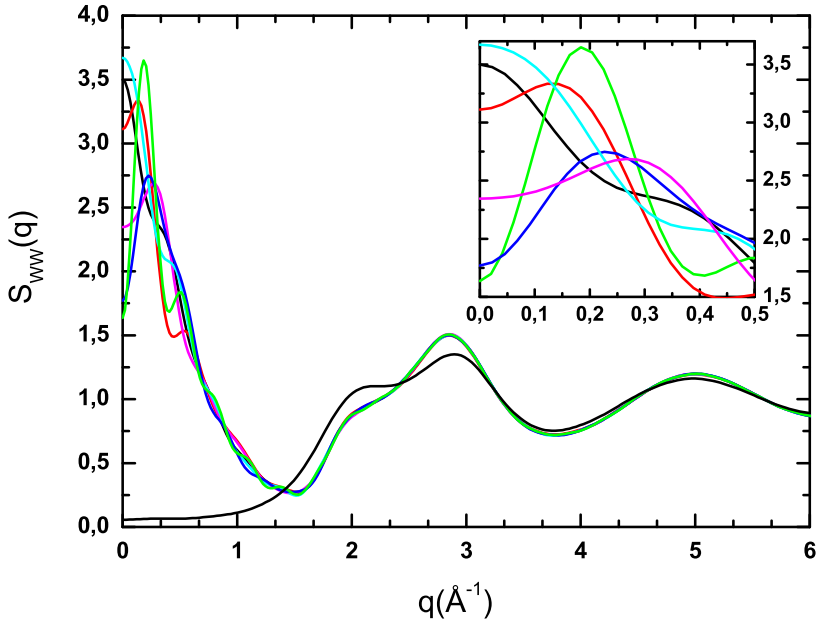


Figure 8.17: The water-water structure factors for the equimolar concentration using different simulation runs. The color convention is same as in figure 8.8: the black line show the results from the 1.02 ns sampling time and the other curves are results from successive simulations runs of 64 ps. The order of the runs is the following: black-red-green-blue-magenta-cyan line. In inset the small-k vector deviation are zoomed in.

black) the $S(k)$ has unusually high value at $k = 0$. Therefore, these results are not conclusive. However, we expect that extensive simulation will give results that is closer to the experimental findings in which we do not observe such large values of structure factor at the small k-vectors.

Usually, the high value at $S(k \to 0)$ indicates that the concentration fluctuations are high and that system is close to the phase separation. One may conclude that, in our simulation, the size of the system is too small to accommodate the fluctuations

that will lead system to phase separation and all results are reminiscent of this trapped structure. Nevertheless, the demixion of the system is indicated by the increase of the fluctuation for all components which is not the case here. Therefore, we find the other explanation more appropriate. The equilibrium structure mixtures of associated liquids is micro-heterogeneous, meaning that on the micro-scale these liquids do not mix well, while they nevertheless form globally homogeneous mixture. This local immiscibility is due to the hydrogen-bonding between self-species and cross-species that tend to form the energetically and entropically most stable hydrogen-bonded network which corresponds then to a locally heterogeneous structure. Our analysis shows that this also leads to the enhancement of the concentration fluctuation, simply because of the unequal spatial partition of the molecular species, that is inherent feature of these liquids.

## 8.5 Micro-heterogeneity in tert-butanol

The results for tert-butanol are shown only for the alcohol molar ratio 0.2. The reason is continuous growth of the large range correlation with simulation time. The simulation is organized in such a way that each run has a sampling time of 64 ps and successive runs are numbered starting from 1 to 32. This comprises a total simulation time of 2.05 ns ($64 * 32 = 2048ps = 2.048ns$). Figure 8.18 shows the time evolution of the radial distribution functions: in the upper panel the water-water RDFs and in the lower panel the carbon-carbon functions. We equally show the time evolution of the corresponding RKBIs: figure 8.20 shows the water-water functions and figure 8.21 in the upper panel carbon-carbon correlation and in the lower panel carbon-water

functions as cross-component correlations. These functions are chosen to represent the molecular correlations based on the previous studies of the neat systems. In each figure the arrow indicates the direction of the time evolution.

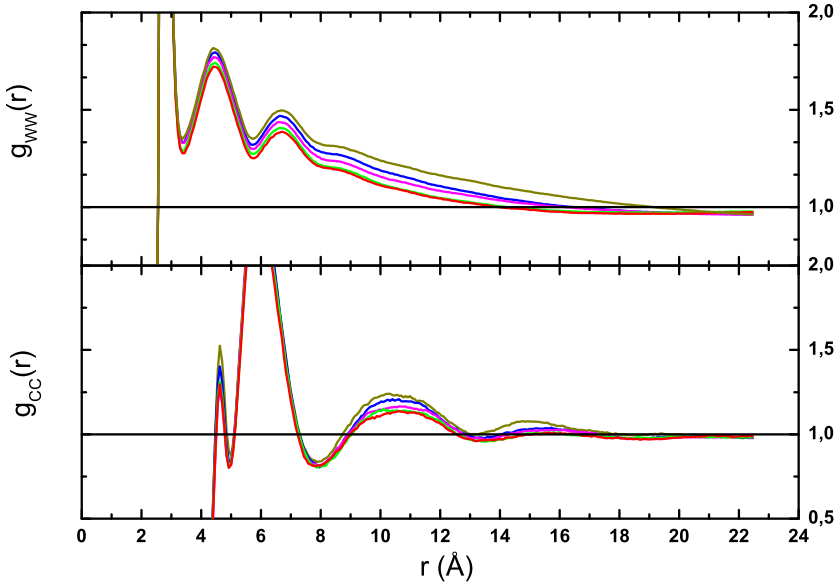


Figure 8.18: The water-water radial distribution functions for the mole fraction 0.2 calculated using different simulation runs. The water-water correlations are displayed in the upper panel, and in the lower panel the tert-butanol carbon-carbon RDFs. The run numbered 16 is in red, 20 is in green, 24 is in magenta, 28 is in blue, and run number 32 is in dark yellow. The number of run indicates the order of successive runs with a sampling time of 64 ps.

We notice that the growth of the RDFs starts already at the first neighboring shell. One other characteristic is that tails of the RDFs go under one. This is equally valid for the water as well as tert-butanol. This behavior translated in terms of RKBIs shows a large constant growth of the RKBIs with the simulation time. The increase is

most dramatic for the water-water RKBIs which rises to the value of $2000 cm^3 mol^{-1}$, but nevertheless present also in the carbon-carbon RKBIs. This is quite large when compared to the experimental value which is in the range $700 - 1000 cm^3 mol^{-1}$ for water-water KBI in water-tert-butanol mixture [97][38]. The cross-species correlations show the same trend with a different direction of time evolution, namely, the functions decrease as the time of the simulation increases.

The growth of the RDFs indicates the building up of the association of the molecules. Therefore, one may assume that system is still evolving. Moreover, the increase of the self-species correlation together with the decrease of the cross-species correlation may be an indication that the system evolves toward a separation of the species. We notice that this behavior is opposite to the simulation results for the methanol-water mixture, where RKBIs oscillate above and bellow the asymptotic curve.

However, the visual inspection of the system, presented in figure 8.19 (for run numbered 6), does not show any clear separation of the species. Moreover, the thermodynamical properties of the system are quite stable over extensive simulation. Namely, the internal energy and the volume fluctuates around the average value and the fluctuation are of the order of standard deviation. Since, these quantities are standard assessment to the stability of the system, than it is not so clear what is happening. Therefore, we hesitate between two opposite scenarios, either the system is slowly going towards demixion of the species, or this results show the kinetics of the micro-heterogeneity.

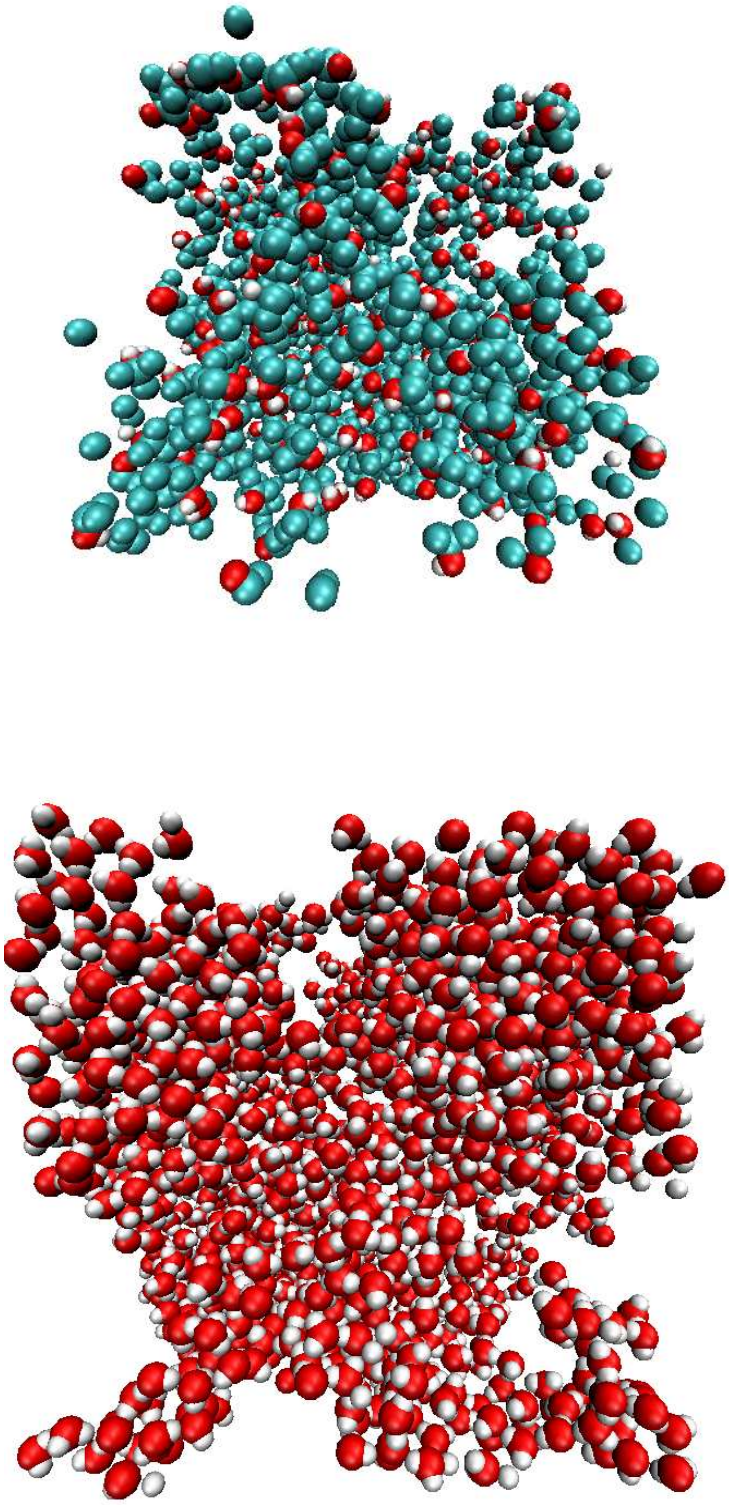

Figure 8.19: The snapshots of tert-butanol-water mixture for $\chi_a = 0.2$ concentration and N=2048 particles. In the upper figure, we show tert-butanol molecules and in the lower water molecules [32].

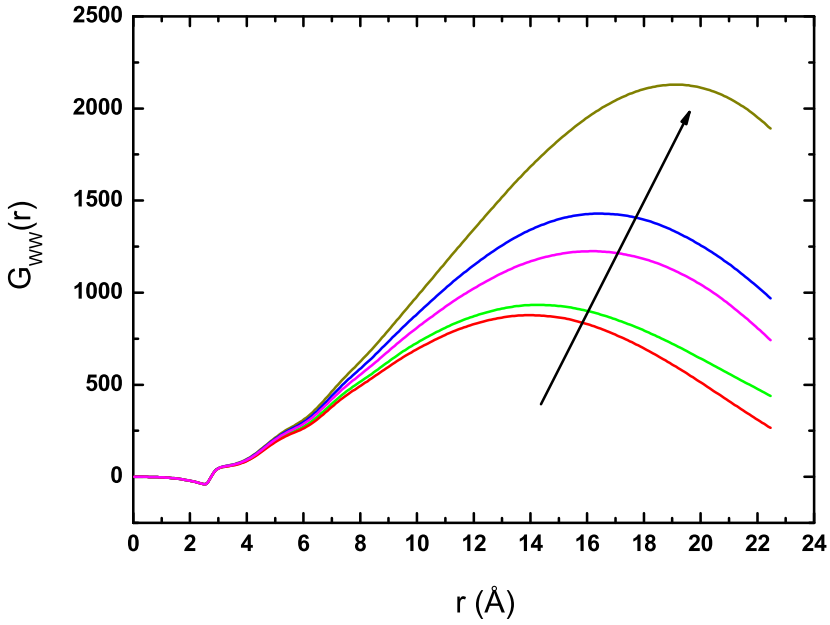


Figure 8.20: The water-water RKBIs for the mole fraction 0.2, calculated using different simulation runs. The run numbered 16 is in red, 20 is in green, 24 is in magenta, 28 is in blue, and run number 32 is in dark yellow. The number of the run indicates the order of successive runs with a sampling time of 64 ps. The arrow shows the time evolution.

The structure factors, presented in figure 8.22, show large values at the small k-vector. It is interesting to follow the evolution of the small to medium k-vector range functions. In figure 8.22, the lower graph zooms in the small k-vector behavior, and the upper graph, medium k-range. The small k-vector behavior in each plot shows a constant increase of the $S(k)$ at $k \rightarrow 0$ which is characteristic for the process of phase transitions indicating an enhancement of the concentration fluctuation. The pre-peak and main peak region shows that water has reached the structural organization of

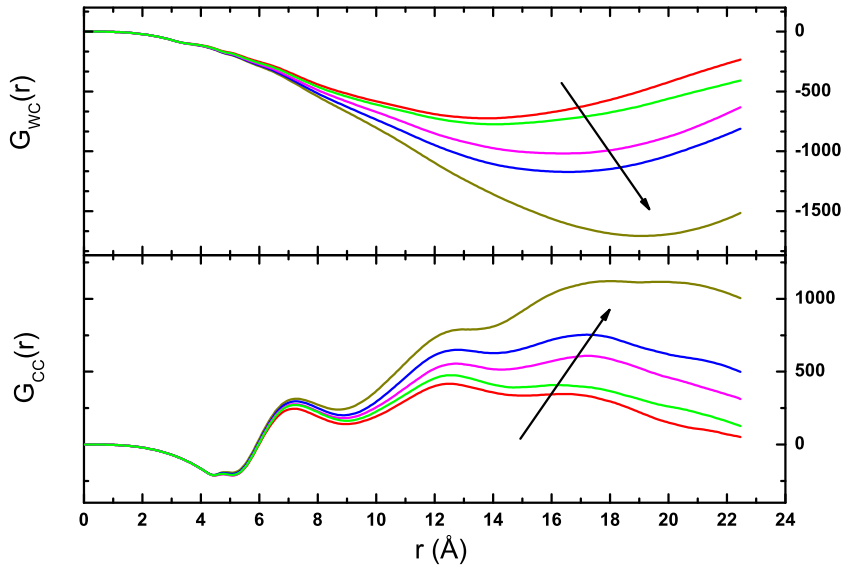


Figure 8.21: The upper panel cross-species (carbon-water function) RKBIs and in the lower panel tert-butanol RKBIs (carbon-carbon function). The functions and the color convention are same as in figure 8.21.

a neat system, namely, the shoulder and the main peak of $S_{ww}(k)$ agree well with the structure factor of a neat water. However, this is not true for the tert-butanol structure factors. The oxygen-oxygen functions have a pre-peaks, that are broader and shifted towards larger k-values than the pre-peak of a neat system. These results indicate that we have a micro-heterogeneous organization of components, but in addition, the concentration fluctuations are also enhanced and these two contributions cannot be easily separated.

The self-association in the tert-butanol-water mixtures is even more enhanced than in the case of the methanol-water solutions. The reason for this is a larger

hydrophobic tail of tert-butanol, which contributes to the segregation of species, and strongly influences the local structure of the mixture. Due to this, these systems are even more difficult to simulate. Firstly, good force fields are needed in order to correctly estimate the electrostatic and hydrophobic contribution. Secondly, the finite-size effects are extremely important, due to the large species associations which can be artificially suppressed in small-size systems. Thirdly, the dynamics of clusters now becomes the dominant factor in the time averaging and the simulation time order of even 10 ns becomes necessary, especially for a low concentrations.

Kusalik at al.[102] compared two different models: the rigid three-site model and fully flexible all-atom 15-site model with the O-H stretching described with an anharmonic potential well for pure tert-butanol. Generally, it is found that the 15-site model is superior. Similar results are valid, also, for the mixtures [103]. Therefore, the flexibility of model is important factor, since the observed structure is a tightly packed structure. However, in order to have simulation of such extensive model, they used system-sizes of no more that 600 particles, and this might be not sufficient to reproduce correct long-range correlations.

The TBA models are also very sensitive to the magnitude of the site charges, and hence the strength of the H-bonding [102]. The previous studies of OPLS tert-butanol [7] and GROMOS force field [104] with SPC water [49] indicated excessive aggregation of the alcohol and water molecules with an almost tenfold overestimation of the alcohol-alcohol and water-water correlations. These observations indicate that both models are too much hydrophobic. The Lee and van der Vegt [35] developed a new force field for TBA using re-parameterizations of the distribution and magnitude of the partial atomic charges of the GROMOS model. They show that the new model

reproduces the experimental KBI and their compositional dependence. However, they noted that the converged value of the KBI were obtained only after 7–10 ns sampling time. They showed, also, that the finite size effects can be the cause of a large errors in calculation of KBI.

Our choice of the OPLS models was due to the following reasons: OPLS models are available for all simple alcohols, and their force fields are built by the process of adding the $CH_n$ group contributions and having the same partial charges of the sites in hydroxyl group that model the hydrogen bonding. This facilitates the analysis of the electrostatic and hydrophobic contributions in the cases of different alcohol-water mixtures. Several reasons could be behind the inability of OPLS simulation to reproduce correct KBI for mixtures: the finite size effects as well as the small simulation times, but also the transferability of OPLS parameters, might be not good for the case of the large alcohols. We will continue to study the tert-butanol-water mixture using other models in order to study the forming of the micro-heterogeneity.

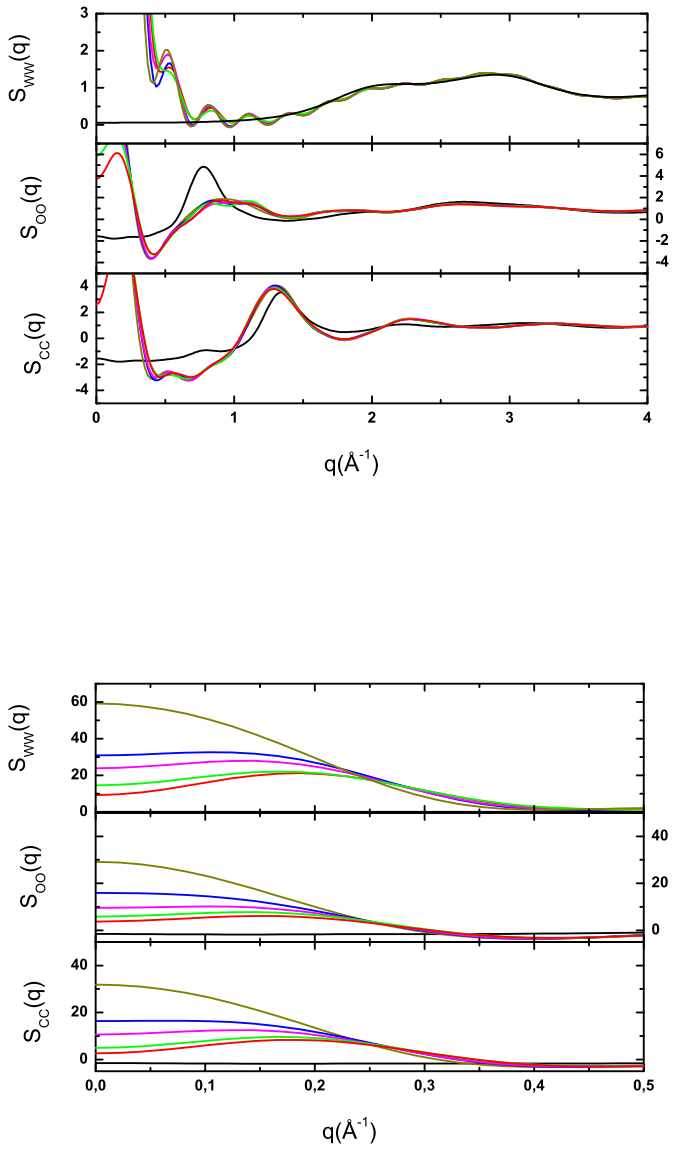


Figure 8.22: The structure factors for the mole fraction 0.2 calculated using different simulation runs. In the first plot we show region around main peak and in the second plot behavior at the small k-vectors. In each plot the upper panel shows water structure factors, in the middle panel structure factors of oxygen-oxygen of the hydroxyl group of tert-butanol and in the lower panel carbon-carbon structure factors. The color convention is same as in figure 8.21, apart from the structure factors of neat systems, which are shown in black.

## 8.6 Discussion and conclusion

In this chapter, we presented the results from the molecular dynamics simulation of aqueous alcohol solutions. The methanol-water mixture was used as a template to introduce the concept of the micro-heterogeneity: the local inhomogeneities of species due to the local immiscibility of the components of the mixture.

Discussion of the thermodynamical properties of these systems signaled the underlying structural specificity. Namely, the thermodynamical properties such as volume, density and enthalpies that are standard access to the reliability of the simulation results showed very good agreement with the experimental results, indicating that the used models are able to reproduce correctly the real system properties. However, the analysis of the excess quantities, which directly probe the non-ideal contribution, indicate that, on this level, simulation models failed to follow the experimental results.

The main assessment to the local inhomogeneities were correlation functions: radial distribution function and $S(k)$. The correlation functions of water as opposed to methanol function showed unexpected variation of the long-range correlation seen as the variation of tail in RKBIs and equally the variation of the small k-vector region in $S\,(k)$. Why is such behavior connected to the micro-heterogeneity?

To explain it let us come back to the differentiation of concentration fluctuation (CF) and micro-heterogeneity (MH). The CF is connected to the asymptotic values of RKBI and $S\,(k)$ (zero limit). The water structure factors however, exhibit the variation at the vectors which are close, but not equal to zero, which indicates the presence of the large clusters of the species. Therefore, we attributed the peaks at the small but non-zero k-vector to micro-heterogeneity. Similarly, large variation of tails

of the RKBI also highlights the aggregation of species. One more reason why we can attribute these variation to MH and not to CF are the correct thermodynamical values for mixtures such as internal energy and densities, which are stable over extensive simulation (50% mixture simulation results).

However, the experimental structure factors do not show any specificity for the small k-region in the spectra of aqueous solutions [21]. How we can link simulation and experimental results? Again, we will take the CF-analogy. Every micro state of an ordinary liquid looks heterogeneous due to the fluctuations of number of particle. In statistical averaging, for simple liquids, density or concentration fluctuations vanish. MH is local inhomogeneity where in each micro state the components exhibit local immiscibility, it is a permanent feature of every realization of a system. One can argue, that MH corresponds to a self-assembly of the species, but with not well defined shape or life-time. Therefore, for the extensive sampling the MH will also vanish since the variety of shapes and sizes of the species-clusters will not give a clear statistically averaged signal. In our simulation due to the small system size, and also small time scale, we are able to measure the part of the realization of MH, which then produces a variation of the structure factor at the small k-vector. Therefore, we have a sampling over dynamics of the micro-heterogeneity. This is first time that the MH has been related to a measurable quantity, albeit only from the simulation results.

The idea that we propose is that micro-heterogeneity is the inherent feature of the aqueous solutions. The MH is also behind the necessity of the large system sizes and the sampling times in the simulations of the aqueous solutions. The large system sizes are required to accommodate the heterogeneous structuring of the components of the mixtures, and the extensive simulation times to sample over the time-scales of

the dynamics of the MH. We think that this study is one step forward in the direction which we believe to be an important one for the investigation of the aqueous solutions. We will go even further and argue about the analogy between the aqueous solutions and the micro-emulsions (ME). Namely, the aqueous solutions are somewhat similar to the micellar systems, where the dominant structural behavior is governed by the micelles acting as meta-molecules. The principal difference is that the solute molecules are smaller and does not produce such apparent morphologies as micelles.

Micro-emulsions have been well studied between the late seventies and the early eighties, both from experimental and theoretical point of view, and many books have been written on this subject. True microscopic approaches to ME, based on statistical mechanics, are rather scarce [25][26][27]. The fundamental fact that has enabled a theoretical formulation of these systems is that scale difference between the patterns formed (about 1 $\mu m$) and the size of the constituents (ranging from few Å to few tens of Å). Otherwise, the molecular forces that produce the variety of patterns is the same as in the present system, with the exception that often micro-emulsions are at least ternary systems, while here we have binary systems. However, recent x-ray and neutron scattering work has equally been focused on binary alcohol-water systems as well [28]. When molecular scales are approached, it becomes harder to define clear-cut geometrical concepts about curvature and interfacial sheet. Once this is realized, the importance of the present analogy becomes clearer. Here, we can speak about a ground basis for self assembly, when it just starts to occur at molecular scale. This offers new challenges to face, both experimentally and theoretically.

What is crucially missing is also a theoretical framework in which the local heterogeneities are taken explicitly. The analogy with micro-emulsion might give a clue

where to start. The structure factor can be determined experimentally from measurements of the scattering intensity $I(q)$ [1]. For the ordinary liquids the long wavelength limit of $I(q)$ is described by Ornstein-Zernike (OZ) behavior:

$$I(q) \sim \frac{1}{1 + q^2 \xi_{oz}} \tag{8.3}$$

where $\xi_{oz}$ is correlation length. Teubner and Strey (TS) [105] proposed a new phenomenological formula for the micro-emulsion:

$$I(q) \sim \frac{1}{a_2 + c_1 q^2 + c_2 q^4} \tag{8.4}$$

Fourier transformation of the equation 8.4 is equal to the correlation function:

$$\gamma(r) = \frac{d}{2\pi r} e^{-\frac{r}{\xi}} sin\left(\frac{2\pi r}{d}\right) \tag{8.5}$$

In contrast to OZ behavior with it's single length scale, one now finds two length scale $\xi$ and $d$. Parameter $d$ is characteristic for the domain sizes, while $\xi$ is the correlation length. Teubner and Strey fitted the experimental data and calculated the characteristic scales and the $c_1$ coefficient as they vary with composition and systems [105]. They concluded that the scattering formula 8.4, predicts well two characteristic properties experimentally observed in small angle scattering: a single broad peak and $g^{-4}$ decay at large $q$. This theory permits to describe the single scattering curves of a variety of ME practically within experimental error with only three fit parameters. They reported also that there are slight deviations in some cases which have to be discussed in a more refined version of the theory. D'Arrigo at al. [28] used the Teubner and Strey phenomenological formula to obtain a measure for the amphiphilicity strength of several studied systems. They analyzed small angle neutron scattering (SANS) on binary aqueous solution of some short-chain amphiphiles at room temperature. They

found that even short-chain amphiphiles have a miscibility gaps and form some sort of micellar aggregates or short-lived micelle-like structures. Their conclusion is that the SANS spectra of the investigated short-chain amphiphiles can be satisfactorily represented by means of a Teubner and Strey equation (eventually with the addition of a OZ term). According to these studies, the theoretical models for the associated system, may comprise two different correlation lengths one corresponding to a standard correlation length that describes fluctuations and other related to the cluster formation. This is one of the possible pathways for the continuation of our study.

# Chapter 9

# Conclusion

In this thesis, we study the structural organization in pure and mixed associated liquids using the microscopic theories as statistical physics of liquids and the classical molecular dynamics simulations. The studied systems are water, alcohols, namely methanol and tert-butanol and water-alcohol mixtures.

We proposed the idea that micro-heterogeneity is inherent feature of the associated liquids meaning that these liquids are locally inhomogeneous while they preserve the fundamental disordered and homogenous nature of liquids.

We distinguish between the structural organization in pure liquids and in binary system, first defined through the local self-association of molecules, and second corresponding to a local segregation of the components of the binary system. In each case the highly anisotropic site-site hydrogen bonding interaction has main role, however on different levels. In neat liquids the association is induced through specific site-interactions and inhomogeneity is due to the clustering of the hydrogen-bonded sites. This feature we named micro-structuring [29] [30]. In the binary system the inhomogeneity is due to a local segregation of species and this features we called micro-heterogeneity [31].

The computer simulations offer an insight in the nano-space scale and pico-time scale behavior of simulated system, which is a unique way to study the behavior of the liquids on the level of one microstate. Therefore, the first part of our investigation was to test the reliability of computer simulation to reproduce correctly the microstructure of the simulated systems. The microstructure of system is directly connected to the particle distribution functions, namely the pair-distribution functions, since for homogenous system the one-body function, the density, has a constant value. The connection between the behavior of the radial distribution function and the real system was made through the calculation of the Kirkwood-Buff integral (KBI). The KBI can be calculated both from the simulation results and the thermodynamical measurements. Precisely, the KBI is directly connected to the concentration fluctuation. We tested the behavior of neat system and the reproducibility of the KBI and concentration fluctuations. This study revels that simulation of a neat system reproduce correctly the KBI value or thermal compressibility of the system. Therefore, microstructure of the simulated systems can be connected to the microstructure in the real systems, at least on the level of the quantities that we explored [106].

Second step was to define the simulation functions that can probe the local heterogeneities. In this case the structure factor which is Fourier transform of RDF and measurable quantity in scattering experiments proves to be extremely valuable. Differences between site-site structure factor of hydrogen-bonded sites and hydrophobic sites show the difference in the local organization of these sites. Namely, the pre-peak in the structure factors of H-bonded sites indicates the local clustering on the sizes larger than the first-neighbor distance, while the correlations between the methyl sites show only the main peak, which corresponds to the packing structure and indicates a

homogeneous organization of these non-bonded sites. This findings are corroborated by the behavior of the RDFs using the polymer analogy in case of methanol and analogy with micelle system in case of tert-butanol. Moreover, the calculation of cluster distribution highlights the size of the preferential clusters of hydrogen bonded sites. In the case of hydrophobic-sites as well as center of the mass, cluster distributions show only features characteristic for homogenous distribution of particles.

The simulation of methanol-water mixture show interesting results. The small k-vector behavior indicates clustering of water, and therefore the local immiscibility of species. This is also shown in the non-convergence of the RKBIs for the water correlations. We connected these findings with the existence of micro-heterogeneity. Therefore, the pre-peak in the site-site correlations indicates the structural preferences of hydrogen bonded sites and the peaks at the smaller k-vectors in the species correlations (all sites that belong to one species) indicate the clustering of the species. Why this this related to MH? We analyzed the data which were accumulated during the small sampling time (over the 100 ps), during which we sampled only some specific topological realizations of the self-aggregating configurations. This way we could analyze the evolution of the MH by steps of 100 ps. This dynamics of MH is seen by the variations of the structure factor of water at the small k-vector across various samplings. Namely, small sampling times allow access to "instantaneous" clustering of the species, in a sense that we look at the system over short period of time comparing to the relaxation times of MH (which we approximate to be around 1 ns). On the contrary , the experimental structure factor of water does not show any large variations in the small-k range. So we expect that with increasing simulation

times, and especially simulation sizes, the simulation results will approach the experimental ones. Nevertheless, using the arguments presented here, we showed that it is worthwhile to observe the system on the smaller scales, which reveals all the richness of the micro-structuring in these mixtures.

The idea of the persistent local inhomogeneity stressed out the necessity of the large system sizes and the large simulation times. For example, pico-second is the time scale of the relaxation of the molecular motion, but the reorganization of MH, based on our simulation results seems to lie in the nanosecond range. Since the system is disordered, we are facing the problem of describing a form of local order within global disorder. This is what makes aqueous mixtures fascinating, despite the obvious physical-chemical aspect, that has been mostly and thoroughly studied in the past century. We suggested the analogy between the micro-emulsion and the associated liquids: both system are globally homogeneous while they are locally heterogeneous, except that in micro-emulsion we have a well-defined shapes such as micelle, while in associated liquids the sizes, shapes and time scales of the aggregates are not well defined. The analogy drawn in this study is interesting from another fundamental point of view: it allows a very detailed and meaningful analysis of the short range details in the structure, while this is nearly impossible for micro-emulsions, where the difference in scale between the solvent, surfactant and micro-structure is not amenable to molecular level simulation while keeping a detailed statistics of the large structure at the same time. For these reasons, we find it interesting to coin the term nano-emulsions to aqueous mixtures.

The study puts another important concept forward, the distinction between concentration fluctuations and micro-heterogeneity. Micro-heterogeneity is presented as

a universal feature of aqueous mixtures. It results from several competing interactions, mostly based on the H-bond interaction. First, the direct interaction between water molecules which tend to self-aggregate, but they can also bond to a solute, and the solutes can also bond between themselves. Then, indirect interactions, such as the repulsion of water to the part of the solute that cannot H-bind, and similarly for the composite solutes themselves. Finally, the usual excluded volume effects and the Van der Waals dispersive interactions, the two of which are usual of any ordinary liquid.

It is the first two interactions that carry and contribute mostly to the micro-heterogeneous features. All 4 interactions contribute to concentration fluctuation. Therefore it is difficult to disentangle these two contributions. We have shown that the analysis of the correlation functions, both in the direct and reciprocal spaces, was an invaluable tool to appreciate both contributions.

The insights provided by the results displayed in this thesis can help building a more microscopic theoretical approach to these simple yet rich systems which are associated liquids and their mixtures. It is our belief that such an approach could have unexpected and unifying fall backs in other areas of condensed matter physics and chemistry.

# Bibliography


[1] Hansen J.-P. and McDonald I. R. *Theory of simple liquids*. Academic Press, Elsevier, 2006.

[2] Kirkwood J. G. and Buff F. P. *J Chem Phys*, 10:774, 1951.

[3] Ben-Naim A. *J Chem Phys*, 67:4884, 1977.

[4] Haile J. M. *Molecular Dynamics Simulation*. Wiley -interscience publication, 1992.

[5] Allen M. P. and Tildesley D. J. *Computer simulation of liquids*. Oxford University Press, 1987.

[6] Smith W., Forester T.R., and Todorov I.T. *The DLPOLY2.0 User Manual*. Daresbury Laboratory, United Kingdom, 2004.

[7] Jorgensen W. L. *J Phys Chem*, 90:1276, 1986.

[8] Weerasinghe S. and Smith P. E. *J Phys Chem B*, 109:15080, 2005.

[9] Berendsen H. J. C., Grigera J. R., and Straatsma T. P. *J Phys Chem*, 91:6269–6271, 1987.

[10] Jorgensen W. L. and Madura J. D. *Mol Phys*, 56:1381–1392, 1985.

[11] Mahoney M. W. and Jorgensen W. L. *J Chem Phys*, 112:8910–8922, 2000.

[12] Jorgensen W. L.and Briggs J. M. and Contreras M. L. *J Phys Chem*, 94:1683, 1990.

[13] Ornstein L. S. and Zernike F. *Proc Amsterdam Acad Sci*, 17:793, 1914.

[14] Lebowitz J. L. and Percus J. K. *Physical Review*, 122(6):1675, 1961.

[15] Kell G. S. *J Chem ENG Data*, 20:97, 1975.

[16] Stillinger F. H. *J Chem Phys*, 38:1486, 1963.

[17] Curro J. G. and Schweizer. *J Chem Phys*, 87:1842, 1987.

[18] Warren B. E. *Phys Rev*, 44:969, 1933.

[19] Stanley H. E., Teixeira J., Geiger A., and Blumberg R. L. *Physica A*, 106:260, 1981.

[20] Bosio L., Teixeira J., and Stanley H. E. *Phys Rev Lett*, 46:597, 1981.

[21] Dixit S., Crain J., Poon W. C. K., Finney J. L., and Soper A. K. *Nature*, 416:829, 2002.

[22] Guo J.-H., Y. Luo Y., A. Auggustsson A., Kashtanov S., Rubensson J.-E., Shuh D. K., Agren H.J., and Nordgren J. *Phys Rev Lett*, 91:157401, 2003.

[23] Weerasinghe S. and Smith P. E. *J Phys Chem B*, 109:15080, 2005.

[24] Ferrario M., Haughney M., McDonal I. R., and Klein M. L. *J Chem Phys*, 93:5156, 1990.

[25] Gomper G. and M. Schick M. *Phase Transitions and Critical Phenomena*. Academic Press, 1994.

[26] Stillinger F. H. *J Chem Phys*, 78:4654, 1983.

[27] Puvvada S. and Blankschtein D. *J Chem Phys*, 92:3710, 1990.

[28] D'Arrigo G., Giordano R., and Teixeira J. *Eur Phys J E*, 10:135, 2003.

[29] Zoranić L., Sokolić F., and Perera A. *J Chem Phys*, 127:024502, 2007.

[30] Perera A., Sokolić F., and Zoranić L. *Phys Rev E*, 75:060502, 2007.

[31] Zoranić L., Redha M., Sokolić F., and Perera A. *J Phys Chem C*, 111:15586, 2007.

[32] All snaphshots are made using VMD molecular visualization program. http://www.ks.uiuc.edu/research/vmd/.

[33] Tanford C. *The hydrophobic effect: formation of micelles and biological membranes*. Wiley, New York, 1980.

[34] Chaikin P. M. and Lubensky T. C. *Principles of Condensed Matter Physics*. Cambridge University Press, Cambridge, 1995.

[35] Lee M. E. and van der Vegt N. F. A. *J Chem Phys*, 122:114509, 2005.

[36] Perera A. and Sokolić F. *J Chem Phys*, 121:11272, 2004.

[37] Mijaković M. *Diplomski rad: Teorija dvokomponentnih Lennard-Jones sustava*. Prirodoslovno matematički fakultet, Sveučilište u Splitu, 2007.

[38] Matteoli E. and Lepori L. *J Chem Phys*, 80:2856, 1984.

[39] Perera A., Sokolić F., Almasy L., Westh P., and Koga Y. *J Chem Phys*, 123:024503, 2006.

[40] Perera A., Sokolić F., Almasy L., and Koga Y. *J Chem Phys*, 124:124515, 2006.

[41] Barker J. A. *Aust J Chem*, 6:207, 1953.

[42] Bahtia A. B. and Thornton D. E. *Phys Rev B*, 2:3004, 1970.

[43] Nishikawa K. *Chem Phys Lett*, 132:50, 1986.

[44] Lepori L. *J Phys Chem B*, 101:9800, 1997.

[45] Chen B. and Potoff J. J.and Siepmann J. I. *J Phys Chem B*, 105:3093, 2001.

[46] van Leeuwen M. E. *Mol Phys*, 87:87, 1996.

[47] Weerasinghe S. and Smith P. E. *J Phys Chem B*, 107:3891, 2003.

[48] Guillot B. *J Mol Liq*, 101:219–260, 2002.

[49] Berendsen H. J. C., Postma J. P. M., van Gunsteren W. F., and Hermans J. *Intermolecular Forces*. Reidel, Dordrecht, 1981.

[50] Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., and Klein M. L. *J Chem Phys*, 79:926–935, 1983.

[51] Whitford P. C. and Philliesa G. D. J. *J Chem Phys*, 122:044508, 2005.

[52] Kolafa J., Labik S., and Malijevsky A. *Mol Phys*, 100(16):2629–2640, 2002.

[53] Ben-Naim A. *Statistical Theromodynamics for Chemists and Biochemists*. Plenum Press, New York, 1992.

[54] Widom B. *J Chem Phys*, 39:2808, 1963.

[55] Weerasinghe S. and Pettitt B. M. *Mol Phys*, 82:897, 1994.

[56] Lyubartsev A. P. and Marčelja S. *Phys Rev E*, 65:041202, 2002.

[57] Chitra R. and Smith P. E. *J Chem Phys*, 114(1):426, 2001.

[58] Chitra R. and Smith P. E. *J Phys Chem B*, 105:11513, 2001.

[59] Soper A. K., Dougan L., Crain J., and Finney J. L. *J Phys Chem B*, 10:3472, 2006.

[60] Gonzalez-Salgado D. and Nezbeda I. *Fluid Phase Equlibria*, 240:161, 2006.

[61] Tauer K. J. and Lipscomb W. N. *Acta Crystallogr*, 5, 1952.

[62] Jr W. Weltner and Pitzer K. S. *J Am Chem Soc*, 73:2606, 1951.

[63] Allison S. K., Fox J. P., Hargreaves R., and Bates S. P. *Phys Rev B*, 71:024201, 2005.

[64] Dougan L., Bates S. P., Hargreaves R., Fox J. P., Crain J., Finney J. L., Rat V., and Soper A. K. *J Cem Phys*, 121:6456, 2004.

[65] Pauling L. *The Nature of the Chemical Bond 3rd ed.* Oxford, University of Oxford, 1967.

[66] Sarkar S. and Joarder R. N. *J Chem Phys*, 99(3):2032, 1993.

[67] Yamaguchi T., Hidaka K., and Soper A. K. *Molecular Phys*, 96(8):1159–1168, 1999.

[68] Bowron D. T., Finney J. L., and Soper A. K. *Mol Phys*, 93(4):531, 1998.

[69] Weitkamp T., Neuefeind J., Fischer H. E., and Zeidler M D. *Molecular Phys*, 98(3):125–134, 2000.

[70] Adya A. K., Bianchi L., and Wormald C. J. *J Chem Phys*, 112(9):4231–4241, 2000.

[71] Ludwig R. *Chem Phys Chem*, 6:1369–1375, 2005.

[72] Narten A. H. and Sandler S. I. *J Chem Phys*, 71(5):2069–2073, 1979.

[73] Zimmermann D., H aber TH., Schaal H., and Suhm M. A. *Mol Phys*, 99(5):413, 2001.

[74] Ben-Naim A. *Molecular theory of solutions*. Oxford University Press Inc., New York, 2006.

[75] Pugnaloni L. A. and Vericat F. *J Chem Phys*, 116:3, 2002.

[76] Stoddard S. D. *J Comput Phys*, 27:291, 1978.

[77] Sator N. *Phys Rep*, 376:1, 2003.

[78] Stauffer D. *Introduction to Percolation Theory*. Taylor and Francis, London, 1985.

[79] Heyes D. M. and Melrose J. R. *Mol Phys*, 66:1057, 1989.

[80] Bako I., Jedlovsky P., and Palinkas G. *J Mol Phys*, 87:243, 2000.

[81] Pagliai M., Cardini G., Righini R., and Schettino V. *J Chem Phys*, 119:6655, 2003.

[82] Stanley H. E., Buldyrev S. V., Mishima O., Sadr-Lahijany M. R., Scala A., and Starr F. W. *J. Phys. Condens. Matter A*, A12:403, 2000.

[83] Kusalik P. G. and Svishchev I. M. *Science*, 265:1219, 1994.

[84] Perera A., P. G. Kusalik P. G., and Patey G. N. *J Chem Phys*, 87:1295, 1987.

[85] Eckert C. A., Knutson B. L., and Debenedetti P. G. *Nature, London*, 383:313, 1996.

[86] Levesque D. and Weis J. J. *Phys Rev E*, 49:5131, 1994.

[87] Patey G. N., Levesque D., and Weis J. J. *Mol Phys*, 45:733, 1982.

[88] Levin Y. *Phys Rev Lett*, 83:1159, 1999.

[89] Murrell J. N. and Jenkis A. D. *Properties of liquids and solutions 2nd edn 102-106*. Wiley, Chichester, 1994.

[90] Frank H. S. and Evans M. W. *J Chem Phys*, 13:507–532, 1945.

[91] Marsh K. N. and Richards A. E. *Aust J Chem*, 33:2121, 1980.

[92] Rowlison J. S. and Swinton F. L. *Liquids and Liquids Mixtures*. Butterworts Scientific, 1982.

[93] Tomšićc M., Jamnik A., Fritz-Popovski G., Glatter O., and Vlcek L. *J Phys Chem B*, 11:1738, 2007.

[94] Wakisaka A., Komatsu S., and Usui Y. *J Mol Liq*, 90:175, 2001.

[95] Koga Y. *J Phys Chem*, 100:5172, 1996.

[96] Vuks M. F. and Shurupova L. V. *Optics Communications*, 5(4):277, 1972.

[97] Nishikawa K., Kodera Y., and Iijima T. *J Phys Chem*, 91:3694, 1987.

[98] Bowron D. T. and Moreno S. D. *J Chem Phys*, 117(8):3753, 2002.

[99] Koga Y. *private communication*, 2005.

[100] Landalt-Borstein. *Neue Serie Group IV Vol1 part B*. Springer Verlag, 1977.

[101] Sokolić F., A. Idrissi, and Perera A. *J Chem Phys*, 116:1636, 2002.

[102] Kusalik P. G., Lyubartsev A. P., Bergman D. L., and Laaksonen A. *J Phys Chem B*, 104:9526, 2000.

[103] Kusalik P. G., Lyubartsev A. P., Bergman D. L., and Laaksonen A. *J Phys Chem B*, 104:9533, 2000.

[104] vanGunsteren W.F., Billeter S.R., Eising A.A., Hünenberger P.H., Krüger, Mark A.E., Scott R.P., and Tironi I.G. *Biomolecular Simulation:The GROMOS96 Manual-user guide*. vdf Hochshulverslag,ETH Zürich Switzerland, 1996.

[105] Teubner M. and Strey R. *J Chem Phys*, 87(5):3195, 1987.

[106] Zoranić L., Sokolić F., and Perera A. *J Mol Liquids*, 136:199, 2007.

## ETUDE PAR DYNAMIQUE MOLECULAIRE DE LA MICRO-HETEROGENEITE DANS LES MELANGES EAU-ALCOOLS

**Ce travail concerne l'analyse structurale des liquides associés comme l'eau et les alcools et leur mélanges. Nous sondons par la Dynamique Moléculaire la micro-structure de ces liquides. A partir du calcul des fonctions de corrélations et des facteurs de structures associés, ainsi que d'un parametre d'ordre effectifs, nous proposons une vision consistente de la microhétérogénéité au sein des liquides macroscopiquement homogénes. Cette analyse nous permet de distinguer entre la micro-structure dans les liquides purs associés, de celle de micro-ségrégation dans leurs mélanges binaires, alors que le mécanisme commun est bien la liaison hydrogéne. Celle-ci structure différamment les sites partiels concernés (hydrogéne et oxygéne), tandis que les sites inertes méthyls sont purement désordonnés. Ainsi, la micro-hétérogénéité apparait comme une propriéte universelle des mélanges de liquides associés. Ce type d'ordre local n'appartient pas tout-é-fait la classe du désordre, pas plus qu'é celle de l'ordre global. Il apparait donc comme une nouvelle forme d'ordre et défie nos méthodes pour le mettre en évidence, tant du point de vue expérimental que théorique.**

**DYNAMIQUE MOLECULARE, SOLUTIONS AQUOUSES, ALCOOL, MICRO-HETEROGENEITE, FUNCTIONS DE CORRELATIONS, LA LIASON HYDROGENE, L'ORDRE LOCAL**

## MOLECULAR DYNAMICS STUDY OF MICRO-HETEROGENEITIES IN AQUEOUS SOLUTIONS

**In this thesis we focus on the structural refinements present in the associated liquids. Using the molecular dynamics simulations we probe the microstructure of these liquids, namely the neat systems and alcohol-water mixtures. From the correlation functions in direct and reciprocal space and as well as the cluster distribution functions we create a consistent picture of the structural specificity that exist in the associated liquids. We made a distinction between the structural organization in pure liquids and in binary systems, first defined through the local self-association of molecules, and second corresponding to a local segregation of the components of the binary system. In each case the highly anisotropic site-site hydrogen bonding has main role. In the pure liquids the association is induced through specific site-interactions that lead to an inhomogeneous distribution ofsites, while the molecular distribution preserves the liquid-like behavior. In the binary system the micro-heterogeneity is due to the local immiscibility of the species. Micro-heterogeneity appears to be universal feature of aqueous solutions, predominantly driven by hydrogen bonding. Micro-heterogeneous order does not quite fit in the class of disorder, but neither in that of order. It appears like a new form or order, and challenges our techniques to put this in evidence, both experimentally or through simulations.**

**MOLECULAR DYNAMICS, AQUEOUS SOLUTIONS, ALCOHOL, MICRO-HETEROGENEITY, CORRELATION FUNCTION, HYDROGEN BOND, LOCAL ORDER**