# Methanol-ethanol "ideal" mixtures as a test ground for the computation of Kirkwood-Buff integrals

Bernarda Lovrinčević, Adrien Bella, Isham Le Tenoux-Rachidi, Martina Požar, Franjo Sokolić, Aurélien Perera

# Methanol-ethanol "ideal" mixtures as a test ground for the computation of Kirkwood-Buff integrals

Bernarda Lovrinčević[2], Adrien Bella[1], Isham Le Tenoux-Rachidi[1],
Martina Požar[2], Franjo Sokolić[2] and Aurélien Perera[1]

[1]Sorbonne Université, Laboratoire de Physique Théorique de la Matière Condensée (UMR CNRS 7600), 4 Place Jussieu, F75252, Paris cedex 05, France.

[2]University of Split, Faculty of Sciences, Ruđera Boškovića 33, 21000, Split, Croatia

## Abstract

Mixtures of 1-alkanols are a textbook example of the concept of ideal mixtures. Yet, such mixtures have a very strong local order due to the hydrogen bonding interactions, with a strong tendency for chain formation. Despite this apparent non-ideality, the Kirkwood-Buff integrals of such system exhibit near ideal behaviour. This dual property can be used to test the calculations of the Kirkwood-Buff integrals in a controlled mixing situation, and clarify many points, in particular the statistical problems that can be encountered. By studying the methanol-ethanol mixtures, we uncover an interesting physical asymmetry between low methanol and low ethanol concentrations, which can produce statistical artifacts in the calculation of Kirkwood-Buff integrals, illustrating and exemplifying some of the difficulties encountered in such calculations. Finally, liquid state integral equations results for these mixtures are reported. They help demonstrate that thermodynamic ideality hides complex correlations and microscopic non-ideality.

1

# 1 Introduction

The relationship between microscopic interactions and statistical correlation functions is essential to understand the physical and structural properties of liquids from a theoretical point of view [1, 2]. The Kirkwoof-Buff integrals (KBI) [3] are a perfect tool to analyse this relationship, since they correspond to the integral of the pair correlation functions, and they can be related to various thermodynamic properties, such as the isothermal compressibility, the molar and partial molar volumes and the chemical potentials [3, 4, 5]. In particular, they allow to measure the deviation from ideality of mixtures in terms of the excess chemical potentials [3, 5]. However, the ideality of the KBI for this type of mixtures poses several problems. A binary mixture can be considered ideal when both components interact quite similarly. Mixtures of alcohols, particularly mono-ols, are considered as ideal [6, 7] since the interaction between the hydroxyl groups dominates entirely all the other atomic interactions, and therefore those between the methyl groups can be considered as a small perturbation. Indeed, in a previous study [8] we have found that the KBI of the methanol-ethanol mixtures are nearly ideal, although there was considerable local structural ordering, due to the hydrogen bonding interactions.

We revisit here these mixtures, with the particular aim of sorting out how the differences in alkyl tails affects the pseudo-ideality of these mixtures, as well as the statistics. We report a surprising difference in statistics between low methanol and low ethanol concentrations, to which we provide a physical explanation. For the same level of statistical sampling, the low methanol content mixture seems to show methanol segregation, contradicting the expected ideality. This behaviour is absent from the mixtures with low ethanol content. Instantaneous snapshots seem to confirm this difference in homogeneity. However, when longer statistics are performed from the low methanol content side, the expected ideality is restored. This finding indicates that ideality is a global concept and could be violated at short times. Since this is not observed for low ethanol content, it introduces a kinetic parameter in the statistics. In addition, this difference is supported by a marked difference between the oxygen-oxygen and oxygen-hydrogen pair correlations, which are the markers of the role played by the hydroxyl groups in the hydrogen bonding clustering and chaining phenomena. We will discuss this situation in the Results section below.

Another way to test the ideality of such mixtures is to solve approximate integral equation theories (IET), such as the reference interaction site method (RISM) theory[9]. Since they are approximations which neglect high order correlations, they are very sensitive to local order and global organisation.

Hence, they are indirect indicators of the complexity of a system. Indeed, the resulting equations cannot be solved for disordered systems which exhibit strong local order, such as aqueous mixtures, for example. In order to solve these IET, it is necessary to make approximations so drastic, that the resulting correlations are nearly featureless, when compared to those from computer simulations[10, 11]. In addition, these theories do not have the size dependence problems which computer simulations have. Anticipating the results of Section 4, we can state that methanol-ethanol mixture exhibit very strong correlations and local order, which these theories cannot capture. This poses the question of what thermodynamic ideality really means.

The present paper differs from the previous one[8] in that it presents a detailed account of the numerical determination of the KBI, a subject that has been of interest in recent investigations by many authors, as well as the above mentioned IET study.

The remainder of this paper is structured the following way. In the next section, we summarise the methods used in this work, and in particular the technical problems concerning the evaluation of the KBI. In the Results section, we examine in detail the calculation of the KBI by computer simulation and IET methods. Our conclusions are gathered in the last section.

# 2 Methodology

## 2.1 Computer simulations

The simulation protocol used herein is the same used in our previous study of the methanol-ethanol mixtures in Ref.[8]. We focus on the TraPPe models of both mono-ols [12], since the previous study has indicated that the OPLS model [13] results are not so much different, as far as ideality properties of the mixtures are concerned. We use the GROMACS 4.5.5 package [14]. N=2048 total number of molecules is considered, with a simulation time step of 2fs. Ambient conditions of T=300K and 1bar (0.98 atm) are imposed through the Nose-Hoover [15, 16] and Parrinello-Rahman [17, 18] algorithms, with relaxation times of 0.1ps and 1ps, respectively. The simulation box, with an initial volume estimated from the interpolation of experimental molar volumes of respective neat alcohols, is filled with both alcohols with the use of the PACKMOL program [19]. The system is then energy-minimised and 200ps NVT run is performed for an initial equilibration, followed by a 1ns NPT step, in order to ensure ambient conditions. Typically 2ns runs are performed for statistics, with 1000 independent configurations saved for calculation of the structural properties. In the cases of low methanol content

10ns runs were required, as discussed in the next section. These statistics are good enough to provide smooth looking correlations $g_{i_a j_b}(r)$ between atom $i_a$ belonging to molecular species $a$ and atom $j_b$ of molecular species $b$. The atom-atom structure factors defined as :

$$S_{i_a j_b}(k) = \delta_{ab} + \rho\sqrt{x_a x_b} \int d\vec{r}\,[g_{i_a j_b}(r) - 1]\exp\left(i\vec{r}.\vec{k}\right) \qquad (1)$$

are calculated by numerical fast Fourier transform. This calculation implies that the $g_{i_a j_b}(r)$ have been properly treated to have the correct asymptote of 1. Neglecting to make this correction will produce serious artifacts at very small $k$ values.

The Xray and neutron scattered intensities have been equally calculated using the expression derived previously [20] from the Debye expression [21, 22]:

$$I(k) = \rho \sum_{ab} \sqrt{x_a x_b} \sum_{i_a j_b} f_{i_a} f_{j_b} S_{i_a j_b}^{(T)}(k) \qquad (2)$$

where the summations runs over all the species $a$ and $b$, and atoms $i_a$ and $j_b$ within each species, $\rho = N/V$ is the total density, $f_m$ is the form factor of atom $m$ and depends on the nature of the scattered radiation (taken from crystallographic tables [23]) and $S_{i_a j_b}^{(T)}(k)$ is the total structure factor, which accounts for the intra-molecular correlations and selects atoms which belong to the same molecular species, and is defined similarly to Eq.(1)

$$S_{i_a j_b}^{(T)}(k) = W_{i_a j_b}(k)\delta_{ab} + \rho\sqrt{x_a x_b} \int d\vec{r}\,[g_{i_a j_b}(r) - 1]\exp\left(i\vec{r}.\vec{k}\right) \qquad (3)$$

where $W_{mn}(k) = j_0(kd_{mn})$ is the intra-molecular correlation function, which is related to the spherical Bessel function $j_0(x) = \sin(x)/x$ in case of rigid molecule (an approximation we shall use here), and where $d_{mn} = |\vec{r}_m - \vec{r}_n|$ is the distance between atomic sites $m$ and $n$ belonging to the same molecule.

It is noteworthy that the scattered intensities reported in our earlier work [8] were based on the Pings-Waser expression [24], which is misleading since it is missing the intra-molecular part $W_{ab}(k)$, which is crucial to match the experimental data. This deficiency is often "cured" by plotting $kI(k)$, which allows to minimise the small-$k$ problems produced by this incomplete expression. Hence, the present paper offers a more appropriate account of scattering properties than the previous one, showing the importance of including intra-molecular correlations.

In this work, we express the Xray scattering intensities in cm$^{-1}$ . In this case, the formula for the scattering in Eq(2) needs to be multiplied by

4

$r_0^2$, where $r_e \approx 2.8210^{-13}$cm is the electron radius. The appropriate factor to convert $I(k)$ expressed as in Eq.(2) to cm$^{-1}$ is the multiplicative factor $\kappa = r_e^2 N/L^3/100$, where N is the number of molecules in the simulations (here $N = 2048$) and $L$ is the box size for a given mixture. The neutron scattering intensities are plotted in units the squared form factors, as given by the cristallographic tables [23].

## 2.2 Integral equations

We briefly recall the RISM IET methodology[9]. It consists in solving the site-site Ornstein-Zernike (SSOZ) equation , together with the hypernetted chain (HNC) equation, in order to find the atom-atom correlation functions $g_{ab}(r)$, as well as the corresponding direct correlation functions $c_{ab}(r)$. The SSOZ equation can be written in the Fourier space in a matrix form. For this purpose, we introduce the matrix element notation $F_{ij}(k) = \rho\sqrt{x_i x_j}\tilde{f}_{ij}(k)$, where $f_{ij}(r)$ is either $h_{ij}(r) = g_{ij}(r) - 1$ or $c_{ij}$, and the tilde designates the Fourier transform $\tilde{f}_{ij}(k) = \int d\vec{r} f_{ij}(r)\exp\left[i\vec{k}.\vec{r}\right]$, hence defining new matrix elements $H_{ij}(k)$ and $C_{ij}(k)$. Further introducing the generalised structure factor $S_{ij}^{(T)}(k) = W_{ij}(k) + H_{ij}(k)$ defined in Eq.(3) and the matrix $\mathbf{M} = \mathbf{W^{-1}} - \mathbf{C}$, the SSOZ equation can be written in a compact form:

$$\mathbf{S^{(T)}M = I} \tag{4}$$

where $\mathbf{I}$ is the identity matrix.

The closure relation can be written as

$$g_{ij}(r) = \exp\left[-\beta v_{ij}(r) + h_{ij}(r) - c_{ij}(r) + b_{ij}(r)\right] \tag{5}$$

where $b_{ij}(r)$ is the site-site bridge function which contain higher rank direct correlation functions[25]. Neglecting this function leads to the HNC closure

$$g_{ij}(r) = \exp\left[-\beta v_{ij}(r) + h_{ij}(r) - c_{ij}(r)\right] \tag{6}$$

In a previous work [26], we have extracted the bridge function for several neat liquids, and in particular methanol and ethanol. In the present work, we test the application of equations (6) and (5). This second equation is solved by weighting the neat alcohol bridge functions with the concentration as follows:

$$b_{i_M j_M}(r) = a_{MM}(1-x)b_{i_M j_M}^{(\text{neat})}(r) \tag{7}$$

$$b_{i_E j_E}(r) = a_{EE}(1-x)b_{i_E j_E}^{(\text{neat})}(r) \tag{8}$$

for methanol and ethanol, respectively, and the cross bridge term is approximated as a mix in terms of the neat liquid bridge functions as:

$$b_{i_M j_E}(r) = a_{ME}(1-x)b_{i_M j_M}^{\text{(neat)}}(r) + a_{EM} x b_{i_E j_E}^{(neat)}(r) \qquad (9)$$

where the index $M$ and $E$ stand for atoms of the methanol and ethanol molecules, respectively, $x$ is the ethanol mole fraction, and the 2 neat liquid bridge functions in Eq.(9) are chosen such that the same pairs of atoms match. In the case of the methanol methyl and ethanol methylene united atoms, we have used the neat methanol methyl-methyl bridge function instead. The 4 coefficients $a_{MM}$, $a_{EE}$, $a_{ME}$ and $a_{EM}$ should ideally be 1. However, in practice, in many cases, it is not possible to bring some or all of these coefficients to 1 without loss of convergence. The reason for this problem is found in the anomalous raise of the $S_{ij}(k=0)$, indicating that the mixture is subject to large concentration fluctuations from the HNC theory point of view. This anomaly is typical of the HNC closure, and the KH closure form is supposed to cure it. This closure is similar to the HNC closure, except that, in the exponential Eq.(6), whenever the argument is positive, it is replaced by the MSA expression for $g_{ij}(r)$. This way, the tendency of HNC to overestimate positive correlations, is tamed down by a lower level closure. This procedure is known to provide numerical solutions whenever HNC fails to do so. However, the resulting correlations are weak and lack the typical structure of strongly correlated liquids. There is no systematic documentation of these problems, since it is not understood why these IET generally fail or are insufficient for complex liquids [27].

## 2.3   On the calculation of the KBI

In recent years, there has been a increased interest in the computation of Kirkwood-Buff integrals (KBI) from computer simulations. These quantities are defined as:

$$G_{ab} = \int d\vec{r}\,[g_{i_a j_b}(r) - 1] \qquad (10)$$

where $g_{i_a j_b}(r)$ is the pair distribution function between atoms $i_a$ and $j_b$, belonging to molecular species $a$ and $b$, respectively, and separated by the distance $r$. The KBI $G_{ab}$ do not depend on specific atomic sites since the integral must be invariant to the choice of any particular centre in the molecules. The atom-atom distribution functions $g_{i_a j_b}(r)$ can be computed from the standard histogram method [28], and the functions are usually sampled in the interval $r \in [0, L/2]$, where $L$ is the system cubic box size. Therefore, in view

of Eq.(1), it does not seem particularly difficult to numerically compute the desired integrals, which come down to evaluating:

$$G_{ab} = 4\pi \int_0^{L/2} dr\, r^2 \left[ g_{i_a j_b}(r) - 1 \right] \tag{11}$$

Yet, this simple task is made difficult by the fact, discovered decades ago by Lebowitz and Percus [29], that $g_{ij}(r)$ in a finite size system would not converge to the expected limit, which is 1, and that the distortion would induce an asymptotic shift, which in the canonical ensemble would be:

$$g_{ij}(r \to +\infty) \to 1 - \frac{\epsilon_{ij}}{N} \tag{12}$$

where $\epsilon_{ij}$ is related to the macroscopic concentration fluctuation $< N_i N_j > - < N_i >< N_j >$ through the relation, expressed in the Grand canonical ensemble as [3, 30]

$$\epsilon_{ij} = V \frac{< N_i N_j > - < N_i >< N_j >}{\sqrt{< N_i >< N_j >}} = \frac{1}{\rho x_i x_j} \left( \frac{\partial \rho_i}{\partial \beta \mu_j} \right)_{TV\mu_k} \tag{13}$$

where $V$ is the volume of the system, $\rho =< N > /V$ is the total number density ($N = \sum_i N_i$ where $N_i$ is the number of particles of species $i$), $x_i =< N_i > / < N >$ is the mole fraction of species $i$, $\rho_i = x_i \rho$ and $\beta = 1/k_B T$ is the Boltzmann factor ($T$ is the temperature and $k_B$ the Boltzmann constant). The relation for the isobaric ensemble is more involved[28], and although the asymptote shift remains similar to Eq.(12), the $\epsilon_{ij}$ term is more complicated.

If the spurious shift in the asymptote is not properly corrected, then the so-called running KBI (RKBI) defined as

$$G_{ab}(r) = 4\pi \int_0^r ds\, s^2 \left[ g_{i_a j_b}(s) - 1 \right] \tag{14}$$

instead of tending to the limit defined by Eqs.(10,11) would have a r-dependent cubic curvature $\sim \epsilon_{ij} r^3$ set by the spurious limit of Eq.(12).

It is important to underline that it is the problem mentioned above which is responsible for the difficulties met with the obtention of the KBI through simulations, and not the finite size problems which are inherent to these methods. These problems can be suppressed by considering large enough systems. This problematic was previously analysed by Sarracuze and co-workers[31, 32], who clearly state that the fluctuation problems play the most important role. Recent works on KBI [33, 34, 35, 36] go back to putting emphasis on finite size problems, and considering them in par with the fluctuation problems. We believe it is more important to correct for the asymptote problem of Eq.12.

In many papers [37, 38, 39], we have proposed to correct for the shifted asymptotes by simply shifting the atom-atom pair correlation functions $g_{ab}(r)$. There are some rules that should be obeyed, however. First, it seems reasonable to assume that the short range part of the $g_{ab}(r)$ are ensemble independent and should not be affected by fluctuations. This could cover the first and the second neighbour distances $0 < r < 2\sigma_{ab}$. Then, there should be a smooth continuation to the distance from which the asymptote may be defined. In earlier publications, we propose to use a multiplicative hyperbolic tangent function to provide the correct asymptote, with the shifted function given by

$$g_{ab}^{(\text{shifted})}(r) = g_{ab}(r) \times \frac{1}{2\alpha_{ab}}[1 + \tanh(\frac{r - \lambda_{ab}}{\kappa_{ab}})] \qquad (15)$$

with $\alpha_{ab} = \lim_{r \to L} g_{ab}(r)$ is the unshifted asymptote, $\lambda_{ab} \approx \sigma_{ab}$ and $\kappa_{ab}$ chosen such that the rate of change of the hyperbolic tangent covers 1 range of $\sigma_{ab}$. These 3 parameters are very empirical, but testing the method shows that these parameters should be chosen in a very narrow range of values if the final result should have a physical meaning. The shift is monitored by trial and error choice of the shift parameter $\alpha_{ab}$ by plotting the running KBI (RKBI) defined as

$$G_{ab}(r) = 4\pi \int_0^r ds\, s^2 \left[ g_{ab}^{(\text{shifted})}(s) - 1 \right] \qquad (16)$$

a quantity which should reach an asymptote defined by Eq.(14) which is the KBI for the pair function $g_{ab}(r)$.

In practice, it turns out that the shifting operation through the simpler recipe

$$g_{ab}^{(\text{shifted})}(r) = \frac{1}{\alpha_{ab}} g_{ab}(r) \qquad (17)$$

is equally efficient, since the shift is so close to 1 that small error due to shifting the $g_{ab}(r)$ in the first neighbour region does not lead to noticeable changes, and is always much smaller than the error bars of estimating the proper KBI. In practice, and for all the cases we have tested so far, one has $\alpha_{ab} = 1 + \epsilon_{ab}$ with $|\epsilon_{ab}| \ll 0.01$. In the next section we will show a detailed analysis of the application of this method to the methanol ethanol mixtures, as well the various problems encountered.

In the cases when one of the species is in low content, the proper determination of the long range correlation may be difficult to obtain because long statistics may be required to ensure the sampling of the macroscopic homogeneity. There seems to be no precise rule for this. In such cases, the asymptote will be affected by the noisy sampling of the real homogeneity,

and it may be possible to determine it properly. In such cases, one can use the following trick, which we have used in many of our previous papers. It consists in noting that, when one species is in small concentration, the other is in majority, and consequently its KBI is well defined through the asymptote shifting method. On such case, one can profitably use the analytical expression of the KBI, which can be approximated as follows

$$G_{11} = G_{12} + \frac{1}{1-x}\left[\frac{V_2}{D} - V\right] \tag{18}$$

$$G_{12} = \frac{V_1 V_2}{V D} \tag{19}$$

$$G_{22} = G_{12} + \frac{1}{x}\left[\frac{V_2}{D} - V\right] \tag{20}$$

where $x = x_2$ is the mole fraction of species 2, $V_1$ and $V_2$ are the molar volumes of the neat species, $V = (1-x)V_1 + xV_2$ is the volume of the mixture, and $D = \rho_a \partial \beta \mu_a / \partial \rho_a$ $(a = 1, 2)$ is the term which describes the deviation from ideality of the mixture. When $D = 1$, the chemical potentials $\mu_a = k_B T \ln x_a \rho$ $(a = 1, 2)$ are ideal and the mixture is ideal. In the expressions above, we have neglected several terms, such as the compressibility of the mixture, which is usually very small since dense liquids are generally incompressible, and we replaced the partial molar volumes by the neat liquid volumes. In practice, we found that the simpler equations above are quite accurate against experimental data [40, 30].

These expressions allow to extract the term D from the well behaved KBI (the one corresponding to majority species) and replace it in the analytical expression for the problematic KBI. We found empirically that this method gives excellent results in place of the problematic asymptote extraction.

# 3 Results

Many structural properties for this particular mixture have been previously reported in Ref.[8]. Therefore, in this section, we will focus on the KBI problem and the integral equation approach to these mixtures, in relation to their ideality.

## 3.1 KBI from simulations

In this section we present the information related to the KBI of the methanol-ethanol mixtures, by using the asymptote shifting method Eq.(17).

First of all, it is important to verify that the asymptotes of the various atom-atom correlations are affected by fluctuations the same way for each species pairs. This is an internal check of the consistency of the calculations. We illustrate this consistency in Fig.1(a-b) for the 60% ethanol mixture.

Fig.1a shows all of the atom-atom structure factors (Eq.(1)), with different colors for each species pairs, blue for all the methanol pairs, magenta for the ethanol pairs, and green for the cross correlations. The atomic details do not matter for this demonstration, which is why these are not specified. We see that the various atom-atom structure factor $S_{ab}(k)$ are different until nearly $k \approx 0$, where they converge to 3 distinct values related to the 3 KBI. It is noteworthy that, to the numerical estimate, the convergence point seems to be $k = 0$, and there is no superposition starting at some non-zero $k$ value. This is the case when micro-segregation occurs [30, 41], which is not the case here: the methanol-ethanol mixtures are therefore microscopically homogeneous.

Fig.1b shows the same information, but from the RKBI point of view, as defined in Eq.(16). All the RKBI corresponding to all the appropriately shifted atom-atom pair correlation functions, are shown with the same color convention as Fig.2a. It is seen that there are 3 distinct asymptotes, corresponding to the 3 KBI, namely $G_{mm}$ (blue), $G_{me}$ (green) and $G_{ee}$(magenta). Although the individual atom-atom pair correlations differ at short and medium distance, they all converge to the same fluctuation influenced KBI value. Both the shifted and unshifted values are shown. These correspond to 10ns statistics.

Fig.1c illustrates the same RKBI, but for 2ns statistics. Since this mixture corresponds to rather low methanol content (40%), we see how the spurious clustering kinetics influences the statistics and the asymptotes, which cannot be made horizontal, even by shifting the shifting procedure. There is a residual curvature, which witnesses the inhomogeneity in the outer reservoir provided by large r-values, as explained in Section 3.1. The statistic problem is so serious that we even observe an nonphysical crossover of the methanol and cross correlation RKBI curves.

Fig.2 shows the RKBI $G_{O_m O_m}$ for the methanol oxygen atom $O_m$ from the correlation functions $g_{O_m O_m}(r)$, for various ethanol mole fractions $x$ ranging from $x = 0$(pure methanol) to $x = 0.95$. These functions are shown to illustrate the asymptote shifting method (the $G_{O_m O_m}$ corresponding to the non-shifted $g_{O_m O_m}(r)$ are shown in gray lines, and are always tilting down to various degrees, depending on the magnitude of the shift). It is seen that the shifted RKBI reach a reasonably horizontal asymptotes for the small x values corresponding to the rich methanol side, where we expect less concentration fluctuations of this species. However, above the equimolar mixture $x > 0.5$,

10

it is seen that the asymptotes deteriorate and start to show spurious noisy behaviour, which correspond to large heterogeneity in the distribution of the methanol molecules and small methanol clusters. 10ns runs are not sufficient to clarify the sampling problems. In such cases, the determination of the asymptotes is impossible. However, we can use the trick explained at the end of Sub-section 3.1. The ideal KBI are equally shown in Fig.3 as green horizontal lines for all concentrations ranging from $x = 0$ to $x = 1$. It is seen that for x<0.6 the asymptotes nearly superpose to these ideal KBI, indicating that the mixture is indeed ideal. This cannot be asserted for x>0.7, but we will see below that these mixtures are equally ideal.

Fig.3 shows the RKBI $G_{O_m O_e}$ for the cross methanol-ethanol oxygen atom $O_m$ and $O_e$ obtained from the correlations $g_{O_m O_e}(r)$. In this case, we see that all the RKBI converge to their ideal values, although some noise is still present in some asymptotes, particularly at low methanol content when $x > 0.7$. These results again assert that the mixtures are nearly ideal.

Fig.4 shows RKBI $G_{O_e O_e}$ for the ethanol oxygen atom $O_e$ , obtained from the correlations $g_{O_e O_e}(r)$, for various ethanol mole fractions $x$ ranging from $x = 0.05$ to $x = 1$(pure ethanol). This time around, it is the asymptotes for small ethanol content which are noisy and not well defined, since this is now the minority species, and dispenses problems to sample the macroscopic homogeneity within a small system and microscopic sampling times. However, we note that the problem appear less serious than for the case of low methanol content.

This asymmetry between the low ethanol and low methanol content, witness the near ideal character of the mixture, since each molecules differ by one methyl atom. Nevertheless, Fig.5a shows, that when trick for correcting the deficient asymptotes is used, then the KBI are near perfectly ideal, as they superimpose to the ideal curves obtained from Eqs.(18-20) by setting $D = 1$. If one wants to get the proper KBI from a computer simulation, for these particular cases where it seems difficult to obtain them without the trick, then one need to simulate very large systems, and for even longer times, in order to sample the proper homogeneity of these mixtures, which tend to have long lived heterogeneity. This is a real operational problem, which can be circumvented in the present case, but which are very much problematic for aqueous mixtures [42, 43, 44, 45], for example. The present study help understand the nature of the practical problem which appear under other context.

Fig.5b illllustrates the LP shifts obtained from the simulations as $\epsilon_{ij} x_i$ versus ethanol concentration, where $\epsilon_{ij}$ is given by Eqs.(12,13). In practice, we numerically shift by the quantity $\theta_{ij} = \epsilon_{ij}/N$ which is very small because of $N$ (here $N = 2048$). We can always express the chemical potential of

11

species $i$ in a binary mixture as $\beta\mu_i = \ln(\rho_i) + \rho_i B_{ii} + \rho_j B_{ij}$ which represent the ideal term and the contribution from excess term [1, 6, 7], and where $B_{ii}$ and $B_{ij}$ contains all the high order density terms for a binary mixture. Then, using Eq.(13), we see that, for like species ($i = j$) we have $\epsilon_{ii} \approx (1 - \rho_i B_{ii})/x_i$ to first order in density. Fig.5b shows that, for methanol (in blue) and in ethanol (in magenta), $\epsilon_{ii} x_i \approx 1 - x_i B_{ii}$ despite the noisy data. If the mixture were ideal, we would expect this relation to hold exactly, and where $B_{ii}$ would be the second virial coefficient. As for the cross term, the cross derivatives in Eq.(13) can be evaluated through the Gibbs-Duhem equation and, for ideal mixtures, one indeed expects that $x_i \epsilon_{ij} = -1$. Therefore, Fig.5b confirms directly the ideality of the mixture, despite the noisy data coming from the empirical procedure of shifting the asymptote visually to 1. It serves to illustrate both the procedure, its shortcomings and its validity in this particular case. We wish to stress that Eq.(13) does not hold for the NPT ensemble, and yet we could get very convincing results.

## 3.2    Integral equation results

Fig.6, shows the 3 oxygen-oxygen correlation functions for the 20% ethanol mixture, as obtained from computer simulations (in blue), HNC (in green), HNC+bridge (in red). It is seen that the HNC approximation alone cannot reproduce the strong chain correlations which lead to the sharp first peak indicated by computer simulations, and the second peak is entirely missed. However, adding the approximated bridge (as explained in Section 2.2) allows to catch up to the first peak to some extent while fitting the second peak perfectly. This means that the neat liquid bridge alone are not enough to reproduce the strong correlations which lead to the first peak in the mixtures. As for the structure factor, the HNC+bridge approximation is relatively good, except for a notable rapid increase very close to k=0, which is not observed for the HNC approximation. This means that adding the new bridge, in addition to bringing a better agreement, also increases the concentration fluctuations. We find this coupling between the good description of local order with the spurious increase of concentration fluctuations, quite noteworthy. It certainly reveals how local heterogeneity and disorder are coupled.

Fig.7 shows the same information as Fig.6, but for the equimolar mixture. In addition, we show the results of the KH+bridge closure (dashed purple lines), which is seen to be overall weaker than the HNC+bridge closure, but is seems better for the small k part of the structure factors. However, it suffers also of the spurious $k = 0$ raise, hence showing that the trick underlying this approximation works only in particular cases. Since HNC+bridge is clearly

superior, this analysis shows that the KH closure should be used, since it has no theoretical justifications. The values obtained for the coefficients in Eqs.(7,9,8) are $a_{MM} = 1$, $a_{ME} = a_{EM} = 0.90$ and $a_{EE} = 0.85$ for the HNC+bridge solution. It was not possible to raise them further to 1, and trying to do so led to dramatic increase of the k=0 part of the structure factors, suggesting anomalous concentration fluctuation. However, this could be achieved for the KH closure, with the modest k=0 increase seen in Fig.7.

Fig.8 shows IET results for the 80% ethanol content. Very similar observation to Fig.6 and Fig.7 can be made. It is interesting that the ethanol correlations for all 3 concentrations are much better reproduced by the bridge correction than the methanol oxygen correlations. This is much against intuition, since one would expect the less complex molecule to be better described than the more complex one. The bare HNC results are very similar between the 3 cases. These 2 results combined, suggest that the methanol bridge is more complex than that provided by our approximation in Eq.(7) and to some extend Eq.(9).

The difficulties of solving IET for these mixtures without the need of the bridge function, clearly suggest the crucial role of the irreducible higher order rank correlations. For simply disordered liquids - typically Lennard-Jones mixtures, neglecting the bridge function does not alter dramatically the structural properties. However, for complex disorder liquids - typically hydrogen bonded liquids and mixtures, it is often impossible to solve IET. In the present case, it is still possible to solve HNC, but with very weak structural features. This is somewhat in contradiction with the fact such mixtures should be thermodynamically ideal, which usually suggests that the interactions are not strong. So, these mixtures are both ideal from thermodynamics, and highly non-ideal as far as structural features are concerned. The thermodynamic ideality arises from a possible cancellation of very strong but opposing non-ideal contributions. The present IET analysis brings an interesting complementary view on this problem, which we hinted to in a previous study [8].

## 3.3   Xray and Neutron scattering

Fig.9a-b shows the Xray and neutron intensities, as computed from the structure factors derived from the atom-atom correlation functions obtained from computer simulations. First of all, we observe that the scattering pre-peak, witnessing chain clustering of the hydroxyl groups, is clearly observed for the Xray intensities, and seen to continuously evolve from a shoulder at $k \approx 1$ Å$^{-1}$ for pure methanol, to a well defined peak at $k \approx 0.7$ Å$^{-1}$ for pure ethanol. It is seen in both cases that the shape of the main peak,

in the range $k \approx 1.5 - 1.8$ Å$^{-1}$, has an interesting cross-over when going from pure methanol to pure ethanol. Upon increasing ethanol mole fraction, this main peak first decreases, then shifts to smaller k-values, and then increases monotonously. This cross-over corresponds to a change in the pre-peak, which goes from a shoulder to a pronounced pre-peak. So, it should correspond to a structural change when going from methanol type to the ethanol type clustering. This should be detected in spectroscopy studies, but the currently available information [46] does not provide any conclusive evidence. Interestingly, the Xray diffraction experiments shown in Fig.9a [47, 48], clearly show the crossover since the main peaks agree with the calculated data. We observe that the 2 experimental data for ethanol differ considerably, indicating that unambiguous experimental Xray diffraction is not achieved. The discrepancy with the calculated data can be justified by modeling inaccuracies. Nevertheless, the shapes of the pre-peak behaviour for methanol and ethanol are quite well reproduced, which is not obvious since these correspond to emergent (cluster) structures, and should be more sensitive to model dependence. The TraPPE model [12] seems to capture these features quite well.

# 4   Conclusion

In this work, we have revisited the methanol-ethanol mixtures, in order to demonstrate how strong directional hydrogen bonding interactions affect the evaluation of the correlation functions, through the statistical problem posed by the existence of long lived hydrogen bonded structure. These structures create an artificial local inhomogeneity which will not melt on usual simulation time scales, and requiring longer runs and statistics. This problem is illustrated by the low concentration asymmetry observed for small statistics, more biased towards the small methanol content side. The existence of such structures, mostly chain-like oxygen clusters, homogeneously mixing both types of oxygen atoms, has been demonstrated by a cluster study in our previous work [8]. These structures create a strongly non-ideal system, from the point of view of correlations, but their integrals, namely the KBI, behave as if it was an ideal mixture. The apparently contradictory behaviour indicates that the thermodynamic signature is incomplete, and can be biased by canceling contributions. We have also illustrated the methodological problems in evaluating the KBI from computer simulations, together with the problems related to the misconception of the importance of the LP shift. We have shown how the evaluation of the KBI are affected both by the LP shift in Eq.(12) as well by the statistical problems inherent to this mixture.

This study, through the problems posed by the statistics and approximate IET approaches, indicates that substantial non-ideality is hidden behind the apparent thermodynamic ideality.

# Acknowledgments

# References

[1] Hill, T. L. *Statistical Mechanics;* Dover, New York, 1956.

[2] Chaikin, P., Lubensky, T., Eds. *Principles of Condensed Matter Physics;* Cambridge University Press, 2000.

[3] Kirkwood, J.; Buff, F. *The Journal of Chemical Physics* **1951**, *19*, 774.

[4] Ben-Naim, A. *The Journal of Chemical Physics* **1977**, *67*, 4884.

[5] Matteoli, E.; Lepori, L. *The Journal of Chemical Physics* **1984**, *80*, 2856.

[6] Koga, Y. *Solution Thermodynamics and its Application to Aqueous Solutions;* Elsevier Science, 1st ed., 2007.

[7] Rowlinson, J.; Swinton, F. *Liquids and Liquid Mixtures;* Butterworth Scientific, London, 1982.

[8] Požar, M.; Lovrinčević, B.; Zoranić, L.; Mijaković, M.; Sokolić, F.; Perera, A. *The Journal of Chemical Physics* **2016**, *145*, 064509.

[9] Hansen, J.-P.; McDonald, I. *Theory of Simple Liquids;* Academic Press, Elsevier, Amsterdam, 3rd ed., 2006.

[10] Perera, A.; Kežić, B. *Faraday Discussions* **2013**, *167*, 145.

[11] Munao, G.; Costa, D.; Saija, F.; Caccamo, C. *The Journal of Chemical Physics* **2010**, *132*(8), 084506.

[12] Chen, B.; Potoff, J.; Siepmann, J. *The Journal of Physical Chemistry B* **2001**, *105*, 3093.

[13] Jorgensen, W. *The Journal of Physical Chemistry* **1986**, *90*, 1276.

[14] Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.; Smith, J.; Kasson, P.; van der Spoel, D.; Hess, B.; Lindahl, E. *Methods of Biochemical Analysis* **2013**, *29*, 845.

[15] Nose, S. *Molecular Physics* **1984**, *52*, 255.

[16] Hoover, W. *Physical Review A* **1985**, *31*, 1695.

[17] Parrinello, M.; Rahman, A. *Physical Review Letters* **1980**, *45*, 1196.

[18] Parrinello, M.; Rahman, A. *Journal of Applied Physics* **1981**, *52*, 7182.

[19] Martínez, J.; Martínez, L. *Journal of Computational Chemistry* **2003**, *24*, 819.

[20] Perera, A. *Physical Chemistry Chemical Physics* **2017**, *19*, 1062.

[21] Debye, P. *Annalen Der Physik* **1915**, *351*, 809.

[22] Debye, P. In *The collected papers of Peter J.W. Debye;* Interscience Publishers, 1954.

[23] Prince, E., Ed. *International Tables for Crystallography, vol. C;* International Union of Crystallography, 2006.

[24] Pings, C.; Waser, J. *The Journal of Chemical Physics* **1968**, *48*, 3016.

[25] Perera, A. *Molecular Physics* **2009**, *107*, 2251.

[26] Chandler, D.; Andersen, H. *The Journal of Chemical Physics* **1972**, *57*, 1930.

[27] Perera, A. *Pure and Applied Chemistry* **2016**, *88*, 189.

[28] Allen, M.; Tildesley, D. *Computer Simulation of Liquids;* Clarendon Press Oxford, 1987.

[29] Lebowitz, J.; Percus, J. *Physical Review* **1961**, *122*, 1675.

[30] Požar, M.; Seguier, J.-B.; Guerche, J.; Mazighi, R.; Zoranić, L.; Mijaković, M.; Kežić-Lovrinčević, B.; Sokolić, F.; Perera, A. *Physical Chemistry Chemical Physics* **2015**, *17*, 9885.

[31] Salacuse, J. J.; Denton, A. R.; Egelstaff, P. A. *Physical Review E* **1996**, *53*, 2382.

[32] Salacuse, J. J.; Denton, A. R.; Egelstaff, P. A.; Tau, M.; Reatto, L. *Physical Review E* **1996**, *53*, 2390.

[33] Krüger, P.; Schnell, S.; Bedeaux, D.; Kjelstrup, S.; Vlugt, T.; Simon, J.-M. *The Journal of Physical Chemistry Letters* **2013**, *4*, 235.

[34] Dawass, N.; Krüger, P.; Schnell, S.; Bedeaux, D.; Kjelstrup, S.; Simon, J.; Vlugt, T. *Molecular Simulation* **2018**, *44*, 599.

[35] Ganguly, P.; van der Vegt, N. *Journal of Chemical Theory and Computation* **2013**, *9*, 1347.

[36] Milzetti, J.; Nayar, D.; van der Vegt, N. *The Journal of Physical Chemistry B* **2018**, *122*, 5515.

[37] Perera, A.; Zoranić, L.; Sokolić, F.; Mazighi, R. *Journal of Molecular Liquids* **2011**, *159*, 52.

[38] Kežić, B.; Perera, A. *The Journal of Chemical Physics* **2012**, *137*, 134502.

[39] Perera, A.; Kežić, B.; Sokolić, F.; Zoranić, L.; InTech, 2012; chapter Micro-Heterogeneity in Complex Liquids, pages 193,220.

[40] Perera, A.; Sokolić, F.; Almásy, L.; Koga, Y. *The Journal of Chemical Physics* **2006**, *124*, 124515.

[41] Požar, M.; Lovrinčević, B.; Zoranić, L.; Primorac, T.; Sokolić, F.; Perera, A. *Physical Chemistry Chemical Physics* **2016**, *18*, 23971.

[42] Allison, S.; Fox, J.; Hargreaves, R.; Bates, S. *Physical Review B* **2005**, *71*, 024201.

[43] Mijaković, M.; Kežić, B.; Zoranić, L.; Sokolić, F.; Asenbaum, A.; Pruner, C.; Wilhelm, E.; Perera, A. *Journal of Molecular Liquids* **2011**, *164*, 66.

[44] Gupta, R.; Patey, G. N. *The Journal of Chemical Physics* **2012**, *137*, 034509.

[45] Banerjee, S.; Furtado, J.; Bagchi, B. *The Journal of Chemical Physics* **2014**, *140*, 194502.

[46] Wrzeszcz, W.; Tomza, P.and Kwaniewicz, M.; Mazurek, S.; Szostak, R.; Czarnecki, M. *RSC Advances* **2016**, *6*, 37195.

[47] Vahvaselkä, K. S.; Serimaa, R.; Torkkeli, M. *Journal of Applied Crystallography* **1995**, *28*, 189.

[48] Tomšič, M.; Jamnik, A.; Fritz-Popovski, G.; Glatter, O.; Vlček, L. *The Journal of Physical Chemistry B* **2007**, *111*, 1738.
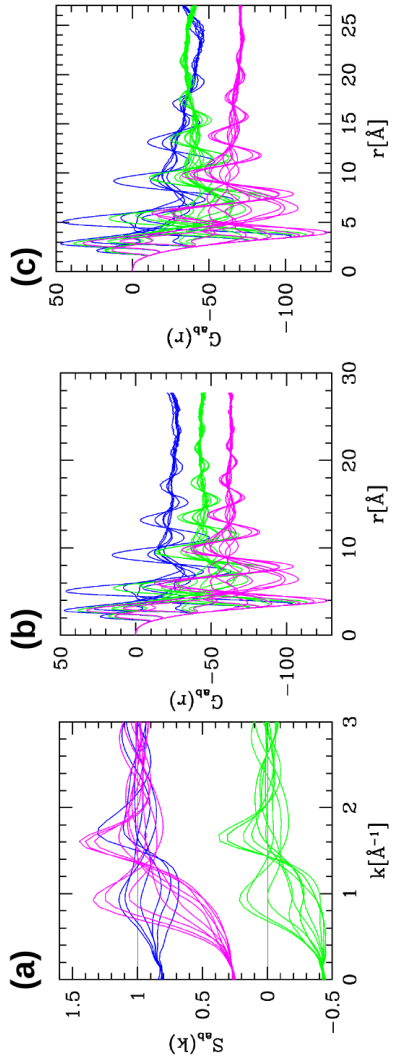
# Figure captions

Fig.1    Illustration of the self-consistent convergence of all the atom-atom RKBI of a mixture to the 3 common asymptotes which represent the 3 KBI values (here for ethanol mole fraction $x = 0.6$ of the methanol-ethanol mixture.). Blue for methanol, green for ethanol and magenta for methanol-ethanol cross correlations. (a) Convergence of the atom-atom structure factors at k=0 into the 3 independant KBI related value. (b) Proper RKBI asymptotes for 10ns statistics. (c) Incorrect RKBI asymptotes for 2ns statistics. See text in Section 3.1

Fig.2    RKBI for the methanol-methanol oxygen atom correlations, for all the ethanol concentrations (curves shifted by +10 for clarity). Blue lines for data from simulations (from bottom to top $x = 0.05$, 0.1 to 0.9 by steps of 0.1, and 0.95); green lines for ideal KBI (from bottom to top $x = 0$, 0.5, 0.1 to 0.9 by steps of 0.1, 0.95 and 1), and gray lines for RKBI from uncorrected asymptotes (see text).

Fig.3    Same as Fig.2, but for methanol-ethanol oxygen atom correlations. Blue and Grey lines for $x = 0.05$, 0.1 to 0.9 by steps of 0.1, and 0.95. Green lines as in Fig.3.

Fig.4    Same as Fig.2, but for ethanol-ethanol oxygen atom correlations. Blue and Grey lines for $x = 0.05$,, 0.1 to 0.9 by steps of 0.1, 0.95 and 1. Green lines as in Fig.3.

Fig.5    (a) KBI as obtained from computer simulations as full symbols, and lines for ideal values. Open symbols represent the anomalous values (from direct extraction from Fig.2-Fig.4 - see text). (b) The scaled LP shift $\epsilon_{ij}x_i$ (see Eqs.(12,13). Symbol for values from simulations, lines for ideal behaviour. In both panels, blue is for methanol, magenta for ethanol and green for cross contributions.

Fig.6    Integral equation results for the 20% ethanol mixture, for the 3 oxygen-oxygen correlation functions $g_{O_M O_M}(r)$ (a), $g_{O_M O_E}(r)$ (b) and $g_{O_E O_E}(r)$ (c). Blue curves are computer simulation results for reference, the green curve is HNC Eq.(6), the red curve is the HNC+bridge described by the Eqs.(5,7,8,9) The insets show the corresponding 3 structure factor $S_{O_X O_Y}(k)$ with same color conventions.

Fig.7    Same as Fig.6, but for the equimolar mixture. The dashed purple curve is for the KH approximation with the same bridges as the HNC+bridge approximation.

Fig.8    Same as Fig.6, but for the 80% ethanol mixture.

Fig.9    Xray (a) and Neutron (b) scattered intensities from computer simulation structure factors. Neat methanol data ($x = 0$) is in dark brown and neat ethanol data ($x = 1$) in black, $x = 0.05$ in red, 0.1 in orange, 0.2 in gold, 0.3 in lime, 0.4 in green, 0.5 in jade, 0.6 in azur, 0.7 in blue, 0.8 in cobalt blue, 0.9 in magenta and 0.95 in purple. Symbols are experimental Xray diffraction data from Ref.[47] for methanol (orange triangles) and ethanol (blue triangles), from Ref.[48] for ethanol (red dots).

Fig.1 - Illustration of the self-consistent convergence of all the atom-atom RKBI of a mixture to the 3 common asymptotes which represent the 3 KBI values (here for ethanol mole fraction $x = 0.6$ of the methanol-ethanol mixture. ). Blue for methanol, green for ethanol and magenta for methanol-ethanol cross correlations. (a) Convergence of the atom-atom structure factors at k=0 into the 3 independant KBI related value. (b) Proper RKBI asymptotes for 10ns statistics. (c) Incorrect RKBI asymptotes for 2ns statistics. See text in Section 3.1

Fig.2 - RKBI for the methanol-methanol oxygen atom correlations, for all
the ethanol concentrations (curves shifted by +10 for clarity). Blue lines for
data from simulations (from bottom to top $x = 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.9$
and 0.95); green lines for ideal KBI (from bottom to top $x = 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$
and Grey lines for RKBI from uncorrected asymptotes (see text).

Fig.3 - Same as Fig.2, but for methanol-ethanol oxygen atom correlations. Blue and Grey lines for $x = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.9$ and 0.95. Green lines as in Fig.2.
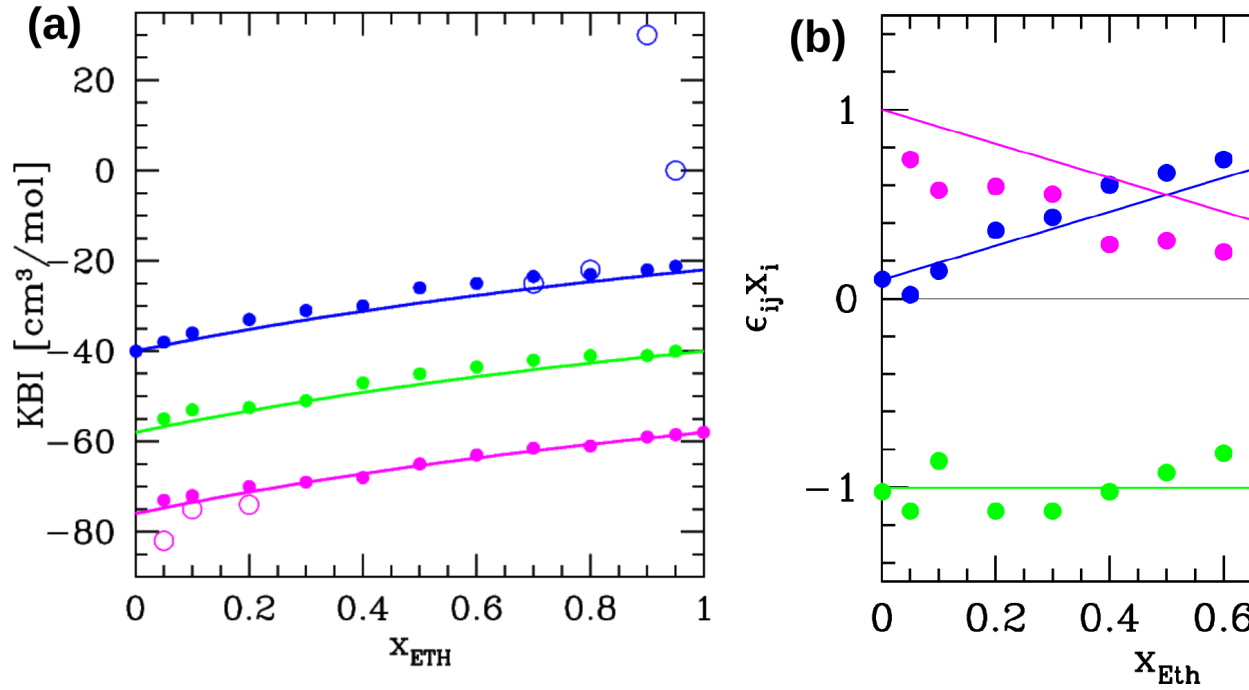
23

Fig.4 - Same as Fig.2, but for ethanol-ethanol oxygen atom correlations. Blue and Grey lines for $x = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.9, 0.95$ and 1. Green lines as in Fig.2.

24

Fig.5 - (a) KBI as obtained from computer simulations as full symbols, and lines for ideal values. Open symbols represent the anomalous values (from direct extraction from Fig.2-Fig.4 - see text). (b) The scaled LP shift $\epsilon_{ij}x_i$ (see Eqs.(12,13). Symbol for values from simulations, lines for ideal behaviour. In both panels, blue is for methanol, magenta for ethanol and green for cross contributions.
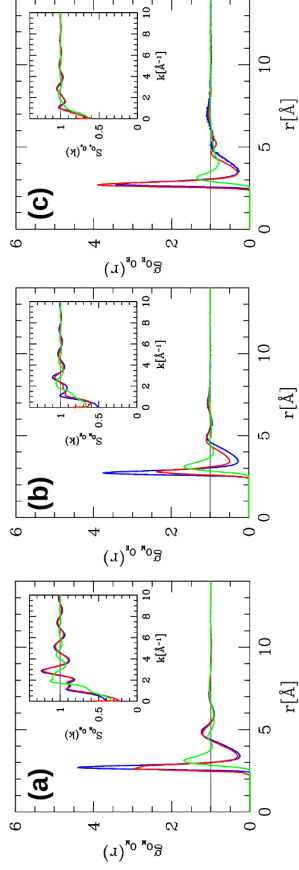
Fig.6 - Integral equation results for the 20% ethanol mixture, for the 3 oxygen-oxygen correlation functions $g_{O_M O_M}(r)$ (a), $g_{O_M O_E}(r)$ (b) and $g_{O_E O_E}(r)$ (c). Blue curves are computer simulation results for reference, the green curve is HNC Eq.(6), the red curve is the HNC+bridge described by the Eqs.(5,7,8,9) The insets show the corresponding 3 structure factor $S_{O_X O_Y}(k)$ with same color conventions.
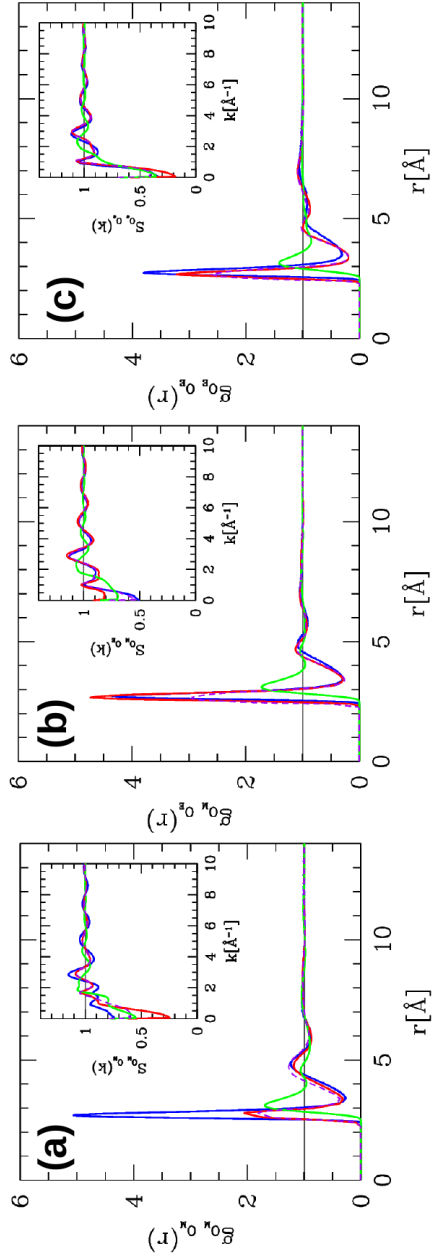
Fig.7 - Same as Fig.6, but for the equimolar mixture. The dashed purple curve is for the KH approximation with the same bridges as the HNC+bridge approximation.
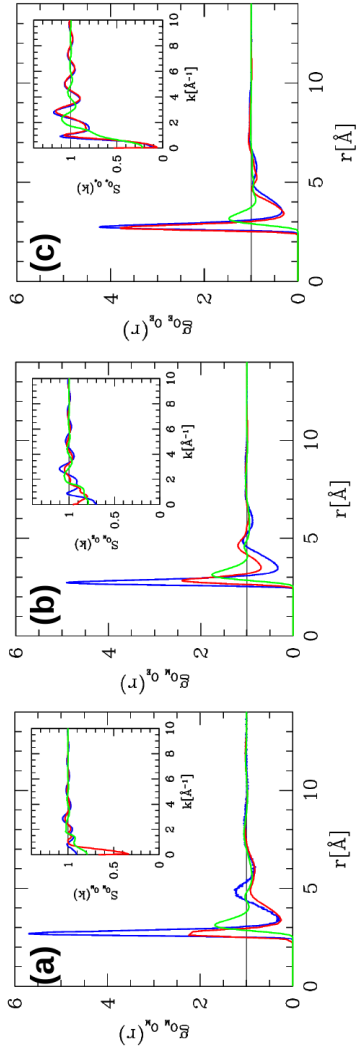
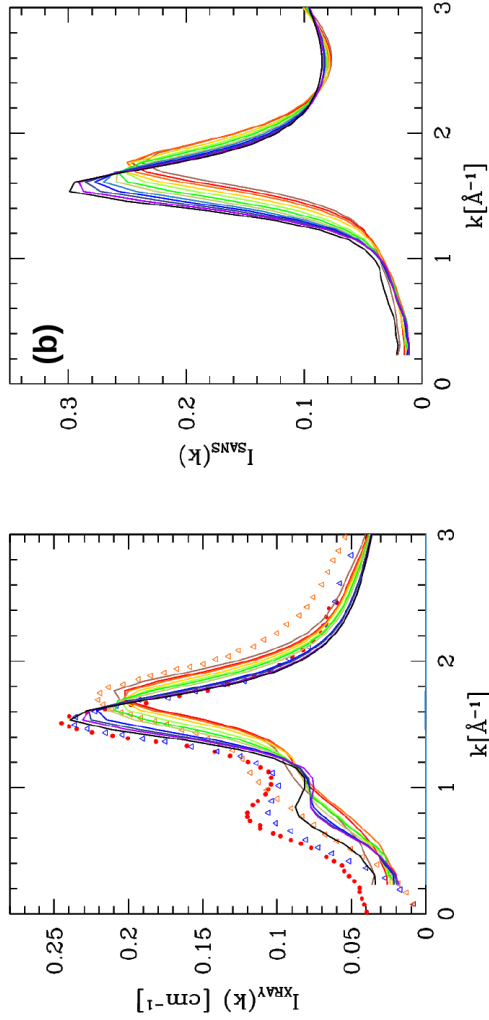27

Fig.8 - Same as Fig.6, but for the 80% ethanol mixture.

Fig.9 - Xray (a) and Neutron (b) scattered intensities from computer simulation structure factors. Neat methanol data ($x = 0$) is in dark brown and neat ethanol data ($x = 1$) in black, $x = 0.05$ in red, 0.1 in orange, 0.2 in gold, 0.3 in lime, 0.4 in green, 0.5 in jade, 0.6 in azur, 0.7 in blue, 0.8 in cobalt blue, 0.9 in magenta and 0.95 in purple. Symbols are experimental Xray diffraction data from Ref.[47] for methanol (orange triangles) and ethanol (blue triangles), from Ref.[48] for ethanol (red dots).

29