From solutions to molecular emulsions
Aurélien Perera

To cite this version:
Conference paper

Aurélien Perera*

From solutions to molecular emulsions

DOI 10.1515/pac-2015-1201

Abstract: Concentration fluctuations play an important role in the statistical description of the stability of liquids, particularly in the neighborhood of phase transitions. Classical thermodynamics is blind to fluctuations, and statistical thermodynamics is required to fully understand quantities such as the isothermal compressibility or heat capacity, by linking them to fluctuations of appropriate statistical microscopic quantities and showing that they are response functions. This is illustrated by the seminal Kirkwood–Buff theory of solutions. However, the existence of micro-heterogeneous structures, particularly in aqueous mixtures, which leads to large Kirkwood–Buff integrals, suggest that micro-heterogeneity is a form of concentration fluctuation. This interpretation becomes difficult to accept when extrapolated to larger micro-heterogeneous structures such as micellar aggregates in micro-emulsions. By analyzing how different methods, experimental, computer experiments and theoretical approaches deal with the underlying duality behind these two physical manifestations, we put in evidence the need to reconsider the description of liquids by incorporating the description of emergent “objects”, such as the micro-heterogeneous structures from a molecular point of view. On this path, the concept of “molecular emulsion” allows to describe in a unified way all type of disordered liquids, from solutions to the organized liquids of soft matter.

Keywords: aggregation; ICSC-34; liquid phases; modeling; molecular dynamics; solution chemistry; water.

Introduction

All liquids and mixtures are naturally endowed with local fluctuations in the number of particles or species, in other words density or concentration fluctuations [1, 2]. Specific type of interactions, such as the hydrogen bond, produce an additional local heterogeneity by enhancing particle association. A natural question to ask is whether these two forms of heterogeneity should be distinguished from one another as corresponding to different physical phenomena [3]. The research in our group has focused on this question and we have pointed its pertinence through the study of several cases [4]. In this introductory part, we illustrate in Fig. 1 the type of problems posed by this question with two examples. The first is about the type of heterogeneity found in aqueous mixtures of surfactant molecules, when the latter is varied from small alcohol, such as t-butanol in water [5] – Fig. 1(a), all the way to the micelle forming monomer molecule, such as n-alkyl polyglycol ether, for example, as sketched in Fig. 1(b). The first case belongs to the category of solutions [6], while the second to the category of emulsions [6–8]. Is there a physically meaningful separation between the two? In the first case, one finds either small nanometer sized inform clusters of t-butanol in water, and in the second case, the heterogeneity takes the form of a micron-sized micelle, which is usually considered as a particle [9–11]. The distinction between concentration fluctuation, cluster or micelle, seems to depend on the different cases, and it would desirable to have precise criteria to classify these different forms of heterogeneity. The second
example concerns the model liquid of dipolar hard spheres. This model has attracted the attention, in the past decade, of the liquid state physicists [12, 13], and it is still perplexing [14, 15]. Indeed, the mean field interaction corresponding to this model is
\[ w(1, 2) = -k_B T \ln(\exp(-\frac{v(1, 2)}{k_B T})) \],
where \( v(1, 2) \) is the dipole-dipole interaction between dipolar hard spheres 1 and 2, \( \langle \ldots \rangle \) designates a statistical ensemble average, \( k_B \) is the Boltzmann constant and \( T \) the temperature, and it looks similar to a Lennard–Jones interaction with a short range repulsion and a long range attraction [16]. Hence, one would expect that this model would show a gas–liquid phase transition. Instead, more than a decade of computer simulations of this system, show the formation labile chains of aligned dipoles, which seems to suppress this transition [12, 15]. Figure 2 shows a system of neutral spheres on the left panel, and the same system under the same box orientation, but when dipolar interactions are added in: one clearly sees the chain patterns. One can ask if these chains are local density fluctuations or specific clusters, and the fact that they suppress a phase transition illustrates both the ambiguous status of these “objects” and their importance as a physical manifestation.

Although a decade of computer simulation results have confirmed the universal property of the micro-heterogeneity of systems such as aqueous mixtures [5, 17–21], we still ignore what is the importance of this micro-heterogeneity as far as their impact on macroscopic physical properties of these systems. In the recent years, the study of room temperature ionic liquid has revived the importance of micro-segregation, since these liquids are characterized by the existence of segregated polar and apolar domains [22–26]. One of the

Fig. 2: Snapshots of 2048 dipolar soft-spheres in a box of \( L = 120 \, \text{Å} \): (left) without the dipole-dipole interactions; (right) with dipole-dipole interaction turned on. Each molecule is represented as a red/blue dimer representing the orientation of the dipole moment.
A recurrent finding in these systems is the existence of a homogeneity pre-peak in the scattering intensity [23]. This finding leads one to ask whether or not such inhomogeneity pre-peak should also appear in micro-heterogeneous aqueous mixtures. Our own computer simulations studies indicate the existence of such a pre-peak in some of the atom-atom scattering functions [5, 27, 28]. However, the detection of this pre-peak is subject to calculation issues related to the correct statistical sampling of the micro-heterogeneity through computer simulations [5, 28]. We have been concerned in our research about the questionable confusion between concentration fluctuations and micro-segregation. Interestingly, statistical theory of liquids, such as integral equation theories, often predict large concentration fluctuations in place of observed micro-heterogeneity [29–31]. In view of these facts, the question asked in the beginning seems at the heart of our investigation tools themselves, and it seems necessary to clarify all these issues.

The purpose of this talk is to show what each approach – thermodynamics, computer simulations and theory – can tell us about the status of fluctuations versus local heterogeneity, both from the established corpus of knowledge and our recent investigations in this domain, and what is problematic in the information we gather from each of the methods of investigations. In particular, I put forward the necessity of considering micro-heterogeneity as distinct from concentration fluctuations by separating the manifestation of the first as the pre-peak feature observed in the structure factors, from the manifestation of the second as the $k = 0$ (zero wave vector) part of the same structure factors.

### Theoretical considerations

This section mostly recapitulates various problems met in the theoretical tools when studying micro-heterogeneous systems, which we have previously mentioned in our works. The existence of such problems in confirmed methodological approach illustrates the limits of these methods, hence the pertinence of our investigations. In order to better illustrate our point of view, we start with the problems met by classical thermodynamics versus statistical thermodynamics.

#### Classical thermodynamics

Classical thermodynamics is a very robust domain of Physics built around unshakable principles and laws [32]. However, the very existence of fluctuations in quantities such as energy, volume or concentration, strongly indicate that thermodynamics is an emergent science, with a lower microscopic level of description [32]. To illustrate this, consider derivative quantities such as the isothermal compressibility $\chi_T$, or the heat capacity $C_v$. It is only through statistical thermodynamics that the anomalous behavior of these quantities close to a second order phase transition can be fully understood. This is because these quantities are in fact related to a more microscopic description in terms of the fluctuations of number of particle for the first [2]

$$
\chi_T = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T = \frac{1}{\rho k_B T} \left< \left< N^2 \right> - \left< N \right>^2 \right>
$$

(1)

(where $\rho = N/V$ is the number density and $N$ is the number of particle per volume $V$, $T$ the temperature and $k_B$ is the Boltzmann constant and $\left< .. \right>$ designates an appropriate ensemble average) or in terms of the fluctuation of the microscopic inter-particle energy $u$ for the second (where $U$ is the internal energy) [2]

$$
C_v = \left( \frac{\partial U}{\partial T} \right)_v = \frac{1}{k_B T^2} \left[ <u^2> - <u>^2 \right]
$$

(2)

Following these example, one can introduce a new microscopical variable, namely the microscopic density of a system of $N$ particles, defined as [2]
\[ \rho(1) = \sum_{i=1}^{N} \delta(1-i) \]  

(3)

where we use the shortcut notation with \( i \) designating both the particle index and the set of variables defining its position and orientation inside the Dirac delta function. In an isotropic and homogeneous system, one can define all the moments related to this new variable, such as the average density

\[ \rho = \langle \rho(1) \rangle = \frac{\langle N \rangle}{V} \]  

(4)

as well its fluctuation between two particles 1 and 2 \[2\]

\[ \langle \rho(1) \rho(2) \rangle - \langle \rho(1) \rangle \langle \rho(2) \rangle = \rho^2(g(1, 2) - 1) \]  

(5)

where \( g(1, 2) \) defines the correlation function. This function is the most useful statistical tool to get the knowledge of the distribution of particles in a given system. It can be measured by computer simulation experiments and by the statistical theory of liquids, albeit through various approximations that we discuss below. In practice, \( g(1, 2) \) depends both in the relative distance \( \vec{r} = \vec{r}_2 - \vec{r}_1 \) between the two particles 1 and 2, and their orientations. In an alternative description, one uses atom-atom (or site-site) correlation functions \( g_{ij}(r) \) which depend on the atomic sites \( i \) and \( j \) of molecules 1 and 2, as well as their distance \( r = |\vec{r}| \).

### Computer simulations

The credo of the science of computer simulation is that, if interactions between particles are properly modeled, then one can obtain an exact description of any given realistic system \[33\]. Of course, there are obvious limitations such as the choice between quantum or classical level of description. To make things clear, we will limit ourselves to the classical case. For example, it is possible to produce a relatively good interaction model of argon or methane that would reproduce many of its macroscopic properties through simulations \[33\]. But one cannot say the same for water \[34\]. This is because water has many anomalous properties all over its phase diagram \[35, 36\], and it seems impossible to describe all of them with a single classical model \[34, 36\]. This is the first level of problem with simulations.

The second level in more intrinsic to the methodology, and concerns the asymptote of the correlation function \( g_{ij}(r) \) in a finite size system subject to periodic boundary conditions, as discovered by Lebowitz and Percus \[27, 37\]:

\[ \lim_{r \to \infty} g_{ij}(r) = 1 - \frac{e_0}{N} \]  

(6)

with

\[ e_0 = \frac{1}{\rho x_i \chi_j} \left( \frac{\partial \rho_i}{\partial \mu_j} \right)_{N=N_i} \]  

(7)

where \( \chi_i = N_i/N \) is the mole fraction of species to which atom \( i \) belongs (\( N_i \) is the corresponding number of molecules), \( \rho_i = N_i/V \) and \( \mu_j \) is the chemical potential of species to which atom \( j \) belongs (\( \beta = 1/k_B T \) is the Boltzmann factor). Equation (6) indicates that the \( 1/N \) correction can never be healed by considering very large systems, particularly when computing the Kirkwood–Buff integrals (KBI) \[38\]

\[ G_{ij} = \int d\vec{r} [ g_{ij}(r) - 1 ] \]  

(8)
This can be particularly well illustrated by considering the running KBI (RKBI) defined by

$$G_i(r) = 4\pi \int_0^r ds \; s^2 \left[ g_i(s) - 1 \right]$$

From eq. (6) it is clear that the incorrect evaluation of this integral will lead to an incorrect asymptote instead of $G_i$. We have indicated several recipes to correct for this problem [4, 27, 39]. This correction is intrinsic to any computer simulation and cannot be avoided if correct KBI must be evaluated. The Kirkwood–Buff theory [38, 40–42] shows that the term $\epsilon_{ij}$ is related to the concentration fluctuations, and consequently that the evaluation of the KBI from computer simulation is sensitive to these fluctuations. It turns out that systems, from macroscopic thermodynamics point of view, that are micro-segregated are a priori indistinguishable from those having strong concentration fluctuations. So micro-segregated systems are particularly sensitive to the evaluation of the KBI through computer simulations, as was empirically discovered by several authors [43–46].

The third level of problem is related to the simulation of system being micro-segregated. In such systems, the local heterogeneity establishes an additional length scale related to the inhomogeneity domain size. It is therefore necessary to take this in account when choosing the system size, such that the distribution of the inhomogeneity is properly sampled. This problem cannot be avoided by simply increasing the sampling time with small system sizes. Indeed, since the inhomogeneity scale distribution is not properly sampled, long runs lead to oscillations in the KBI values that never settle in the course of time [5, 28]. In order to avoid prohibitive system sizes, we have introduced a method which accounts for the domain oscillations in $g_{ij}(r)$ within simulations of reasonably small systems [5, 28]. This method is based on the formal analogy between the structural heterogeneity in micro-emulsions and that of the systems with micro-heterogeneity. This analogy is implemented through the form of the structure factors, which are the Fourier transform of the $g_{ij}(r)$

$$S_i(k) = \delta_{0k} + \rho \sqrt{\pi} \frac{1}{X} \int dr \; \exp(ikr) \left[ g_{ij}(r) - 1 \right]$$

In the case of micro-emulsions, the small-$k$ behavior of the $S_i(k)$ has a specific form called the Teubner–Strey structure factor [3, 47, 48]. This way, the initially very large values predicted for the KBI evaluated through computer simulations can be brought down to values closer to experiments [5, 28]. This method predicts that the domain pre-peak should appear in the structure factor. This prediction is confirmed directly by computer simulations for few systems that we mention below. It is important to note that the appearance of a pre-peak in the partial structure factors does not imply that the experimental total scattered intensity should equally show such a pre-peak like in micro-emulsions. This is due to spurious compensation effect, such as those observed in room temperature ionic liquids for example [22, 23, 25]. The origin of this compensation is an open topic for further investigation. In particular, it is not clear if this cancellation is a physical or a mathematical property related to each type of systems.

**Statistical theory of liquids**

The formalism of the theory of liquids allows one to compute the $g_{ij}(r)$ from the sole knowledge of the atom-atom interactions in the molecular liquids, as well as state parameters such as the partial densities $\rho_i$ and the temperature $T$. However, the $g_{ij}(r)$ are computed at the price of often uncontrollable approximations. They give reasonable prediction for the so-called simple liquids, but are quite useless to predict phase transitions because they do not account for high order correlation. In what concerns more complex liquids such as aqueous mixtures, these methods often fail to produce numerical solutions for state parameters (temperature, density) for which the corresponding systems are stable [49, 50]. There is a physical reason for this behavior, which is particularly illuminating about the properties of the systems of interest. Indeed, these approximate methods are unable to distinguish between concentration fluctuations and micro-segregation [31]. This is related to fundamental problem that we evoked in the Introduction. It is interesting to observe that approximate theoretical methods stumble on the same problem. This confusion in the theory arises
A. Perera: From solutions to molecular emulsions

because the various approximations miss the contribution of particularly complex high order correlations [51]. For this reason, these theories are unable to account properly the structure of complex liquids such as hydrogen bonded liquids [29, 52]. It turns out, however, that it certain particular cases which we discuss below, these theories are able to account for the complex segregation behavior of some systems, often model systems, which capture the mechanisms of the real systems. We describe below in Section “Insights into complex disorder” particular cases when this happens and we attempt to provide some insights for application to realistic systems.

The site-site integral equation theory formalism [52–56] used in this work allows to compute the $g_{ij}(r)$ from the following two equations. The first is the site-site Ornstein–Zernike (SSOZ) equation which can be cast into a matrix form:

$$\text{(W + } \rho H)(W^{-1} - \rho C) = I \quad \text{(11)}$$

where every matrix element $X_{\alpha \beta ij}$ is defined in terms of the atomic sites $(i_a, j_b)$ belonging to molecules of species $a$ and $b$, and corresponds to a wave vector $k$-dependent function. This way, the three matrix elements correspond, respectively to the intra-molecular correlations $W_{\alpha \beta ij}(\sigma_\alpha \sigma_\beta \delta_{ij})$, the inter-molecular correlations $H_{\alpha \beta ij}(\sigma_\alpha \sigma_\beta \exp(ik \cdot \vec{r}))$, and the direct correlations $C_{\alpha \beta ij}(\sigma_\alpha \sigma_\beta \exp(ik \cdot \vec{r}))$.

The second equation corresponds to a closure and is written as

$$\beta \rho \nu_{\alpha \beta ij}(r)(\rho C_{\alpha \beta ij}(r) - 1) \quad \text{(12)}$$

and involves the site-site interactions $\nu_{\alpha \beta ij}(r)$ as well as the temperature through the Boltzmann factor $\beta$. These two equations must be solved by an iterative process as detailed in [52].

Simple and complex disorder

The digressions in the previous subsections indicate that some systems are easier to deal with theoretical approaches than other, whether these are simulations or statistical theory. This way, simple atomic liquids such as nitrogen, but also molecular liquids such as CCl4 and even weakly polar liquids such as acetone fall in this category. Generally speaking, liquids involving strong directional interaction, such as hydrogen bonded liquids pose problems. This is because such liquids tend to self-segregate by forming self-bonded molecular domains. From this simple remark, we will introduce two categories of liquids, by referring to the type of disorder they convey, simple and complex disorder. What is the purpose of introducing this distinction? For example, such distinction could be introduced at the level of the interactions, such as isotropic interactions and directional interactions. However, the resulting macroscopic liquids come in variety of combinations of disorder which may not be easily categorized. Therefore it seems more suitable to introduce the categorisation at a higher level of organization. This has a heuristic meaning because it decouples the microscopic interactions from the resulting macroscopic manifestations. We will come back to this point in the last section after examining different cases.

A conjecture about thermodynamical signature of disorder

When one looks at snapshots (Fig. 3) taken from two different systems such as benzene+pentane (top figures) and aqueous t-butanol (bottom figures), and at same different concentrations of the second component, is the striking heterogeneity found in the second system. In particular, when one looks at the concentration dependence of the snapshots in both cases, the first system look more or less homogeneous, while the second
have a striking variation in the nature of the aggregation. In other words, it is tempting to ask if macroscopic thermodynamical quantities, measured as the function of the concentration of the solute, would reflect these marked differences. There is a considerable corpus of such experimental data. We focus here in quantities $F(x)$ such as the excess enthalpy, the excess volume – as a function of the solute mole fraction $x$. These data are abundantly reported for various systems. What we have generically observed is that, for the systems in the first category, these excess quantities are always smoothly varying functions of $x$, with no particular accidents in the second derivative. In other words, these quantities would be either convex or concave functions of $x$ and will never change for given thermodynamical conditions. In other words, it would seem that a good thermodynamical argument for simple disorder would be:

$$\frac{\partial^2 F(x)}{\partial x^2} \text{ keeps same sign in the interval } x \in [0, 1]$$

(13)

This is illustrated in Fig. 4, for $F$ being either the excess volume (left panel) or excess enthalpy (right panel), in cyclohexane(solvent)-hexane isomers(solute) [57], $x$ being the solute mole fraction. While the excess enthalpies seem to obey the above recipe in eq. (13), the excess volume show some exceptions for the case of the branched ones, 2-methylpentane and 3-methylpentane, but not for 2,3-dimethylbutane, which is also branched. We come back to this type of exception below.

Conversely, in the case of aqueous mixtures, when looking at various excess quantities (Fig. 5) [58–62], we note sharp changes occurring at small $x$ (mostly $x < 0.2$) and a weaker ones (mostly for $x > 0.5$). These changes could correspond to marked micro-structural changes. Intuitively, since these systems are characterized by H-bonded domains, the changes could correspond to a particular change in the connectivity of the domains, one occurring at small solute concentration and the other being the symmetric at larger mole fractions. In the case of water, the changes are stronger on the denser side of water concentration. These sharp variations in the second derivative are connected to at least one change of convexity, possibly three. Therefore, a good thermodynamical characterisation of complex disorder would be:

$$\frac{\partial^2 F(x)}{\partial x^2} \text{ changes sign in the interval } x \in [0, 1]$$

(14)
Fig. 4: Thermodynamical properties of a simply disordered system: cyclohexane – hexane isomer mixtures [57]. (left) Excess enthalpies and (right) excess volumes. Numbers correspond to hexane isomers: (1) n-hexane – (2) 2-methylpentane – (3) 3-methylpentane – (4) 2,3-dimethylbutane – (5) 3,3-dimethylbutane.

Fig. 5: Thermodynamical properties of a complex disordered system: aqueous mixtures: Excess enthalpies for various alcohols [57] (methanol, ethanol, 1-propanol and 2-propanol); Excess partial molar volumes [58]; Excess sound speed [59]; (c) Excess heat capacity [60, 61].

One way to understand the existence of such two changes is consider that the complex mixture percolates around the equimolar concentration or possible the equal volume occupancy (this point is not clear from experimental data). By percolation, we refer to a separation of the system in two sub-domains, but without demixing [63, 64], so this is a loose concept of percolation. The two points where convexity changes could then correspond to typical concentrations where one of the component starts to cluster appreciably. This percolation concept necessitates that at least one component is connecting to itself in priority (associating...
molecule). Indeed, the equimolar mixture of a random mixing two component system does not show any percolation behavior.

Now, we can try to interpret exceptions for the first type of disorder, such as that we found for the excess volumes of the mixtures of branched alkanes with cyclohexane. This type of exception can be explained in terms of the complexity brought by packing difficulties due to branching. We note that such changes are very small when compared with those that are truly complex such as aqueous mixtures.

We insist on the fact that these two criteria are conjectured on the basis of experimental data. Even though they are inspired from the look of snapshots from computer simulations, they do not depend on the accuracy problems about force fields and other such issues often brought forward in the community of simulators.

The second point we want to make concerns the domain of validity of these conjectures, since we saw some exceptions above. Clearly, when phase transition regions are approached, fluctuations and in particular concentration fluctuations are likely to grow and affect these criteria, in particular the first one. Indeed, large fluctuations could appear as heterogeneity and change the convexity of some properties that would otherwise obey eq. (13). In other words, we assert that violations of eq. (13) for simple disorder systems are likely to happen close to regions of phase transitions. In fact, this could be a criteria to detect the approach of a phase transition for these system.

At this point, it is interesting to ask if it is possible to justify such conjectures on the basis of theoretical arguments, from the molecular details for example.

A structural signature of complex disorder

Since complex disorder is characterized by the appearance of cluster structures, one seeks a structural evidence for it. The same way the distribution of molecules in a system modulates the molecular and site-site distribution functions, one expects the existence of domains to induce an additional supra-molecular modulation of the same functions. In the Fourier space, the molecular/atomic modulation produces a main peak in the structure factors, centered at a wave vector $k_m = \frac{2\pi}{\sigma}$ where $\sigma$ is the molecule/atom size. Similarly, one expects the domain modulation to produce a pre-peak centered at wave vector $k_p = \frac{2\pi}{d}$, where $d$ is the mean domain size. Consequently, permanent local heterogeneity should produce a distinct pre-peak in some partial structure factors. This conjecture is important since it provides a structural distinction between concentration fluctuations and microscopic heterogeneity. Indeed, the statistical theory of liquids shows that macroscopic density or concentration fluctuations provide a raise of the structure factor at $k = 0$ \cite{42, 65}. However, the existence of permanent heterogeneity should influence the behavior of the structure factor in the range of wave vector $k \in [0, k_p]$ and on expects that the raise at $k = 0$ could be influenced by the heterogeneity. It is in this sense that macroscopic density or concentration fluctuations and local heterogeneity can be related to each other. The existence of such micro-heterogeneity pre-peak in partial structure poses two problems. First, it is often difficult to detect in computer simulations because the system sizes required to sample properly the distribution of the micro-heterogeneity necessitate simulation box sizes of $L \gg 10$ nm. Most simulations in the literature are with box sizes $L \approx 0.5$ nm, which is sufficient to obtain accurate values of many thermodynamic quantities. In the next section we will show cases for which the pre-peak is detected in small box sizes. The second problem is more serious: most scattering experiment, X-ray or neutron, do not show conclusive evidence of a pre-peak in the total scattered intensity for most aqueous mixtures of small solute molecules. However, the total structure factor is a sum over partial structure factors \cite{25, 66}, involving the atomic form factors $f_i(k)$

$$I(k) = \sum_{i,j} \sqrt{x_i x_j f_i(k) f_j(k)} [S_{ij}(k) - \delta_{ij}]$$  \hspace{1cm} (15)

It turns out that, in some cases, the cross structure factors $S_{ij}(k)$ with $i \neq j$ have a negative pre-peak that cancels the positive pre-peak \cite{22, 23} in the like-structure factors $S_{ii}(k)$. The exact origin of this cancellation is currently not well documented, as mentioned above in Section “A conjecture about thermodynamical
signature of disorder”. For example, the structural pre-peak experimentally is observed in the case of alcohols [67, 68] and room temperature ionic liquids [22, 23], and is attributed to the domain segregation between charged atomic groups and neutral groups (such as the methyl groups). In such cases, the correspondence between this experimental pre-peak and those observed in computer simulations is well established [25]. Since such peaks arise because of the local segregation of different types of atoms, it is natural to expect such pre-peaks in the case of the micro-segregation observed in aqueous mixtures. This expectation is enforced by the experimental evidence for such pre-peaks in micro-emulsions, with micelles or lamellar domains. Moreover, eq. (15) shows that different types of scattering experiments (SANS, SAXS, light scattering) differ by the associated form factors $f_i(k)$ for atom of type $i$. These factors are always positive [66]. Therefore, the pre-peak features intrinsic to the various atom-atom structure factors $S_{ij}(k)$ being independent of the type of experiments, the cancellation coming from a series of specific atom pairs will always occur, but only with a difference in magnitude intrinsic to the associated form factors. In particular, specific techniques, such as isotopic substitution will change the appropriate $f_i(k)$, but will not affect the cancelations in every pair of terms involving the same isotope $i$. In view of these remarks, the reason why such pre-peaks have not been reported in aqueous mixtures, which are obviously micro-heterogeneous, is very intriguing.

**Insights into complex disorder**

In this section, we report cases where the complexity of disorder, as defined in the previous section, can be either explicitly observed directly from computer simulations, or from the theoretical approaches. These particular cases are more interesting than the generic aqueous mixtures since they are free from the constraints imposed from the methodology used to probe them. The results shown below are closely related to our previous works [27, 69, 70], where we have explored each of the cases separately in greater details. By gathering some of the details below, we hope to illustrate the purpose and the coherence of our project.

**Simple models of complex disorder**

What are the elementary ingredients to generate complex disorder, and which can be studied safely by simulations and theory without any of the pitfalls indicated previously? Following the generic patterns of systems that produce micro-heterogeneity, one can model molecules that have the desired features in the following way. We consider a 3-sites molecule made of joint spherical atoms, with a head atom of type A and large tail atom of type B. These dimer molecules are immersed in a mono-atomic solvent of type C. We consider the case where A–C atoms interactions are of the standard Lennard–Jones type, while all other interactions are purely repulsive. Figure 6 provide an illustration of the model and the interactions used. This way, we enforce the A atoms to associate with the solvent C. This should create segregated regions where the A and C atoms are grouped, separated by regions filled with the B tail atoms, which cannot be phase separated since they are attached to the A atoms, as illustrated in the snapshot in Fig. 7. We investigate such system both by computer simulations and the RISM integral equation techniques explained
in Section “Statistical theory of liquids”. We consider this binary mixture for the following parameters: the diameters are set to $\sigma_A = \sigma_C = 3 \text{ Å}$ and $\sigma_B = 6 \text{ Å}$, the molar volume of the dimer liquid and that of the solvent are taken as equal $V_D = V_C = 170 \text{ cm}^3/\text{mol}$, the temperature $T = 225 \text{ K}$, the inter-site distance in the dimer molecule is $L = 1.75 \sigma_A$. A snapshot of the equimolar mixture shows clearly micro-segregated parts [69], where the A and C atoms regroup, while the B atoms are segregated. Site-site correlation functions $g_{ij}(r)$ are shown in Fig. 7, with corresponding structure factors. The existence of a domain pre-peak is quite visible. The agreement between the simulations and the theory is remarkable. In addition, the effective site-site interactions $W_{ij}(r)$ obtained by inverting $g_{ij}(r) = \exp(-W_{ij}(r))$ from the expression eq. (12), shown in Fig. 6 (right panel) indicates that this interaction has a particular form, with a short range attraction followed by a long range repulsion – the SARL (short range attraction long range repulsion) type interaction [71–73] – which has been used in recent works to explain the aggregation of colloids. Our microscopic model allows to emerge such effective interaction directly from a more microscopic representation. Such type of interaction is characterized by the existence of two length scales, one of which can be associated to the formation of clusters [73]. This is entirely consistent with the description of H-bonded clusters in aqueous mixtures. The two-length scale description is particularly relevant to explain critical crossovers in hydrogen bonded liquids [74].

In view of the near impossibility to obtain similar agreement with aqueous mixtures of alcohol for example [29], one is lead to examine the differences in the interactions. In both cases there are attractive interactions between the bonding parts which dominate the weaker ones between the non-bonding parts. One difference resides in the magnitude of the Coulomb interactions in aqueous mixtures compared with the much smaller LJ interactions in the present simple model. However, as seen in the next subsection, Coulomb interactions are rather well described by integral equation theories. Therefore one is led to consider that it is the directionality of the H-bond interaction that is not well described by the theory. Indeed, symmetry broken interactions enhance the high order correlations and reduce cancellation effects that occur in their contributions. This argument holds equally for the computer simulations because the clusters are less isotropic and a
better sampling is required to account for all possible cluster shapes. This directionality of the H-bond seems to be at the heart of the complexity and the problems encountered in various methodologies.

**Ionic liquids: disorder through charge ordering**

Mixtures of Lennard–Jonesiums exhibit simple disorder and some amount of concentration fluctuation is inherently present. One would naively expect that a true homogeneous mixture would have various atoms dispersed in perfect alternated order. Is this a mirage? In fact, ionic liquids achieve exactly this type of order-in-disorder, which is called charge ordering: the snapshot in Fig. 8 shows the difference between an homogeneous Lennard–Jonesium binary mixture and it can be seen that the ionic liquid shows a near perfect alternate disposition of oppositely charged ions, making the system ideally homogeneous. Therefore, ideal homogeneity could mask very particular interactions and consequently an underlying and hidden form of order.

In this section we consider a simple model of ionic liquids in order to illustrate the inherent complexity of this type of systems. It consists of cation made of a jointed 3-sites molecule with a positive charge on one end site, and a spherical anion. All sites have the same size \( \sigma = 3 \text{ Å} \) and interact through both a repulsive LJ-type interaction and the Coulomb interaction:

\[
v_{ij}(r) = 4\epsilon\left(\frac{\sigma}{r}\right)^{12} + \frac{e^2 z_i z_j}{4\pi\varepsilon_0 r}
\]

where we consider \( \epsilon = 84 \text{ K} \). This system is studied at a temperature of \( T = 3000 \text{ K} \) and a molar volume of \( V_m = 45 \text{ cm}^3/\text{mol} \), both by computer simulations and integral equation techniques. Figure 8 shows the snapshot for a temperature of \( T = 3000 \text{ K} \) and a molar volume of \( V_m = 45 \text{ cm}^3/\text{mol} \), at which this system is liquid [75]. It shows the typical segregation of the neutral tails from the ionic sites which tend to cluster due to the Coulomb interaction. One can notice that the charge ordering in these latter clusters. Figure 9 shows the correlations functions between the charged sites as well as the corresponding structure factor. Simulation results are shown in dashed lines and theory in full lines. The charge ordering is visible in the \( g_{xy}(r) \) through the out-of-phase correlations between the like and opposite charge. This equally manifests itself in the positive and negative main peaks of the structure factor. In addition, one notices the pre-peak feature at \( k = 0.6 \text{ Å}^{-1} \), which corresponds to the segregated domains. The domain pre-peak can be further enhanced by using the Bhatia–Thornton [76] structure factor representation developed for alloys, which relates the charged atom-atom structure factors to the concentration-concentration structure factors. We have done this in [75]. However, this approach is currently not adapted to other types of mixtures involving molecular systems, although one can devise some ways to adapt it as we have shown the case of the system studied in Section “Simple models of complex disorder” above [69].

Fig. 8: (left) Snapshot of a neutral binary mixture. (middle) Snapshot of a binary mixture of charged spheres illustrating charge ordering. (right) Snapshot of a ionic liquid model showing segregation of the oily tails (black) and charged atoms (yellow anion, orange cation).
What is perhaps the most striking is the near perfect agreement between the theory and the simulations. In addition, the theory is able to predict the pre-peak in very good agreement with the simulations. In view of the poor agreement between the two approaches usually found for complex systems [29, 52], one can wonder why this is not the case here. The reasons are twofold. First, the theory is generally very good for charged hard spheres [77]. Although the present system is more complex, the charge-charge correlations seem to obey a similar trend. Next, when charges are split across many sites, the complexity of the interactions increases, with substantial frustration between the different molecular positions. This requires an increased accuracy in the correlations, which the theory is no more able to provide. In other words, the theory is able to provide a good agreement provided the interactions are kept simple, while nevertheless capturing many features of complex systems. For more realistic cases, such as associating molecules, where the theory generally fails, the present case provides clues to improve the theory: one needs to describe better only those charge-charge interactions which dominate the behavior of the system. This could be done, for example, by an effective representation of the charges in the system. Work along such lines is in progress.

Influence of solvent fluctuations

Since water as a solvent is problematic, it is interesting to consider a more neutral solvent such as benzene or hexane, for example. In such a solvent, we expect the anisotropic shape to play an important role. In a recent study [27], we have shown that binary mixtures such benzene-hexane behaved very similarly to ideal mixtures, at least in what concerns concentration fluctuations, showing in particular ideal KBI. When considering binary mixture of ethanol and benzene, we found that ethanol formed very compact clusters [27], specially at low concentration, unlike in water, where the ethanol clusters were more fuzzy [78]. The snapshot in Fig. 10 illustrates this point very clearly: it shows 20% ethanol molecules in water and in benzene. Only the oxygen and hydrogen atoms of the ethanol molecules are shown. It is clear that the ethanol molecules in benzene are distinctly clustered in loops and chains, while in water they are more dispersed. As a result the oxygen–oxygen correlations and the corresponding structure factors display a very different behavior in each systems, as illustrated in Fig. 11. In order to clarify the situation of the water-ethanol system, we have
conducted simulation of a larger system with twice the size (hence eight times more molecules: 16,000 molecules instead of 2000). The snapshot in Fig. 10 illustrates clearly the clustering of ethanol, even if it is not as apparent as that of the ethanol-benzene system. The structure factor of the larger system shows a pre-peak that could not be unambiguously detected in the smaller system – although 2000 particles is considered to be sufficient for such types of systems. The zoom on the tail correlation functions in Fig. 12 shows the long range domain oscillations, which cannot be unambiguously handled in the small system. We have provided in [5] a method based on the analogy with emulsions, principally on the Teubner–Strey form [47] of the structure factor, which allows to extend the long range correlations in order to account for the missing domain oscillations. It is seen from the example provided here that this method, while heuristically useful, will not replace a proper simulation on a larger system. This remark poses the problem of the proper size required in simulations of complex systems, even the most simple ones.

Finally, before closing this subsection, we can ask why the small system was enough to detect a pre-peak in the case of benzene-ethanol, and not in the case of water-ethanol. In view of the results above, this question amounts to ask why ethanol was more fuzzily clustered in water than in benzene. One answer would be to invoke once again the emulsion-molecular emulsion analogy and compare the clustering of ethanol in water to direct micelles and that of ethanol in benzene to inverse micelles. It seems more stable for ethanol to
group its binding sites at the core and leave the methyl tails out, than to bind to water outside and shy the oily tails inside. So, the unambiguous existence of a site-site pre-peak is intimately related to the stability issue of underlying aggregates.

**Influence of the solute shape**

Heterogeneity in most aqueous mixtures is particularly characterized by clusters of water with no particular shapes. This is the case in aqueous ethanol, t-butanol, acetone and amide mixtures that we have previously studied [5, 28, 39, 63]. In a recent work, we have examined aqueous DMSO mixtures and we have found that water forms a very different type of cluster, mostly linear in shape [70], as shown in Fig. 13. This intriguing effect can only be due to this particular solute molecule. When comparing its shape with the other solute molecules, we note that the sulfur molecule is relatively large and that all the bonding and non-bonding sites are mostly positioned on one side of the sulfur atom, leaving an open non-bonding spherical space. This strong difference in the symmetry of the molecule influence the way water molecules bind. In addition, in [70] we have observed that the linear shape of the water cluster produce a structure factor pre-peak which is obtained directly from the simulations of both usual size with $N = 2048$ and larger size $N = 16,000$ molecules (see Fig. 14). This finding shows that the structural complexity depends on the solute molecule and its signature depends on the form of the water aggregated domains. These findings seem to be independent of the nature of the force field models used in [70].
Discussion and conclusion

In this presentation, we have proposed that the classification of the nature of liquids needs to be reconsidered on the basis of their inherent structural micro-heterogeneity. Usually, liquids are classified according to their physical and chemical properties. This classification leads to an artificial distinction between ordinary liquids and those in emulsions and soft-matter, which appears as arbitrary when seen from the point of view of the interactions between the constituents and their microscopic heterogeneity. Without entirely abandoning the usual classification, it seems heuristically interesting to consider a classification based on the local heterogeneity seen as an intermediate state between the molecular constituents and the macroscopic manifestation. This way, one distinguishes between simple form of disorder which does not have a permanent local heterogeneity, such as in liquid argon or carbon-tetrachloride for example, and those systems which are inherently micro-heterogeneous, and which we proposed to call molecular emulsions [5, 28, 79, 80]. This distinction is based on the difference between concentration fluctuations and the local heterogeneity. A measure of the first can be obtained through quantities available by thermodynamical investigations, while the second requires scattering experiments. Therefore, simple disorder can be measured by macroscopic thermodynamics, while complex disorder requires more specifically a microscopic tool. The conjecture formulated in Section “A conjecture about thermodynamical signature of disorder” could be a strong alternative to this latter requirement.

This distinction between simple and complex disorder has important issues related to the fundamental aspects of Statistical Physics. Density or concentration fluctuations are important since they control the mechanical stability of the system, hence its phase transitions [81]. On the other hand, micro-heterogeneity tends to prevent phase separation: antagonist interactions in molecules are grouped in such way that the global frustration is resolved by local segregation of species instead of full phase separation. This is the principal reason for the complexity of this type of disorder. The phenomena of self-assembly is intimately related to this type of disorder: it is just another way to overcome the local frustration between molecular groups. However, before tackling complex microscopic issues related to auto-assembly, the approach outlined in this presentation can provide a more systematic way to understand progressively the mechanisms, starting from how small molecules organize themselves under energetic and spatial constraints.

Aside from these heuristic aspects, we have seen that the complexity of the disorder requires some precautions in order to be handled properly by the usual methodology used for liquids, such as simulations. It turns out that the problems met with the use of these tools allow a finer investigation of the various forms of complex disorder, such as the shape of the water clusters for example. The importance of such findings may not be apparent for now, but could be potentially useful in subsequent microscopic investigations of...
biological matter [82, 83]. Indeed, biological media are essentially non-heterogeneous systems, with various forms of heterogeneity; covalently bound, hydrogen bound, with ionic interactions, strongly directional interactions. The hidden ordering of such form of matter could have mechanisms which could be unveiled by the type of microscopic approach that we propose here.

References

[33] For an insightful review about water science, visit http://www1.lsbu.ac.uk/water/.