



# Une théorie de la fonctionnelle de la densité moléculaire pour la solvation dans l'eau

Guillaume Jeanmairet

## ► To cite this version:

Guillaume Jeanmairet. Une théorie de la fonctionnelle de la densité moléculaire pour la solvation dans l'eau. Chimie analytique. Université Pierre et Marie Curie - Paris VI, 2014. Français. <NNT : 2014PA066122>. <tel-01067993>

**HAL Id: tel-01067993**

**<https://tel.archives-ouvertes.fr/tel-01067993>**

Submitted on 24 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## THÈSE

Pour obtenir le grade de

**Docteur de l'Université Pierre et Marie Curie**

Spécialité : **Chimie Physique et Chimie Analytique**

Présentée par

**Guillaume Jeanmairet**

Thèse dirigée par **Daniel Borgis**

préparée au **Pôle de Physicochimie Théorique UMR 8640 PASTEUR**  
et dans l' **ED 388**

# Une théorie de la fonctionnelle de la densité moléculaire pour la solvation dans l'eau

Thèse soutenue publiquement le **16 Juillet 2014**,  
devant le jury composé de :

**Jean-François Dufrêche**

Professeur, Rapporteur

**Gerhard Kahl**

Professeur, Rapporteur

**Jean-Pierre Hansen**

Professeur émérite, Examineur

**Michel Mareschal**

Professeur, Examineur

**Julien Toulouse**

Maître de conférence, Examineur

**Daniel Borgis**

Directeur de Recherche, Directeur de thèse





# Remerciements

En premier lieu, je souhaite remercier les personnes ayant accepté d'évaluer ces travaux de thèse. Merci à Jean-François Dufrêche et Gerhard Kahl qui m'ont fait l'honneur de rapporter le présent manuscrit. Merci également à Jean-Pierre Hansen, Michel Mareschal et Julien Toulouse d'avoir pris part au jury de cette thèse.

Je remercie également Ludovic Jullien, directeur de l'UMR 8640 PASTEUR, qui m'a permis d'effectuer ce travail dans un environnement de grande qualité.

Ce travail a été réalisé sous la direction de Daniel Borgis à qui j'adresse mes sincères remerciements pour sa pédagogie, sa rigueur scientifique, son enthousiasme et la confiance qu'il m'a accordé. Toutes ces qualités ont permis de faire de ces trois années une aventure exaltante scientifiquement et humainement.

Merci également à Maximilien Levesque qui a participé à mon encadrement et qui a été à mon écoute tout au long de cette thèse. Sa sympathie, sa gentillesse et nos discussions scientifiques resteront un souvenir précieux.

Je souhaite également remercier les différents membres du pôle que j'ai eu le plaisir de côtoyer pendant ces trois ans. Parfois, les discussions autour d'un café sont plus fructueuses que la lecture d'un article.

Ces trois années ont également été pour moi l'occasion d'enseigner en tant qu'agrégé préparateur. Je remercie donc mes collègues de l'ENS et de la préparation à l'agrégation, en particulier Jérôme Quérard, mon partenaire de galère avec la « vieille polarographie ».

J'aimerais aussi remercier mes copains de promo à l'ENS pour nos débats, scientifiques ou non, où le respect n'avait d'égal que la mauvaise foi.

Enfin, je souhaite terminer par une pensée pour ceux qui me soutiennent depuis longtemps, ma famille mais aussi Fanny pour sa patience, ses encouragements, ses relectures, ses talents de coiffeuse et tout le reste.

# Table des matières

<b>Résumé</b>	<b>1</b>
<b>I. État de l'art</b>	<b>3</b>
<b>1. Modélisation des propriétés de solvation</b>	<b>4</b>
1.1. Solvation . . . . .	4
1.1.1. Aspect énergétique de la solvation . . . . .	5
1.1.2. La structure de solvation . . . . .	6
1.2. Modèles explicites . . . . .	8
1.2.1. Méthodes de solvant explicite . . . . .	8
1.2.2. Bases de thermodynamique statistique . . . . .	9
1.3. Méthodes de solvant implicite . . . . .	11
1.3.1. Milieu diélectrique continu . . . . .	12
1.3.2. Équations intégrales . . . . .	14
1.3.3. RISM (Reference Interaction Site Model) . . . . .	16
1.3.4. 3D-RISM . . . . .	17
<b>II. Théorie</b>	<b>20</b>
<b>2. La théorie de la fonctionnelle de la densité classique</b>	<b>21</b>
2.1. La théorie de la fonctionnelle de la densité électronique . . . . .	21
2.2. La théorie fonctionnelle de la densité classique (cDFT) . . . . .	24
2.2.1. Exemple introductif, le cas du fluide idéal . . . . .	24
2.2.2. Principe variationnel pour le grand potentiel . . . . .	27
2.3. Fonctions de corrélation directe et écriture de la fonctionnelle d'excès . . . . .	29
2.3.1. Approximation du fluide homogène de référence (HRF) . . . . .	31
<b>3. Une théorie de la fonctionnelle de la densité moléculaire pour l'eau</b>	<b>34</b>
3.1. Modèle d'eau et de soluté . . . . .	34
3.1.1. Modèle d'eau rigide . . . . .	34
3.1.2. Modèles de solutés rigides . . . . .	35
3.2. Un traitement strictement dipolaire de la polarisation . . . . .	35
3.2.1. Écriture de la fonctionnelle . . . . .	35

3.2.2. Structure de solvation . . . . .	40
3.2.3. Énergies de solvation . . . . .	47
3.3. Au delà de l'ordre dipolaire . . . . .	51
3.3.1. Écriture de la polarisation multipolaire . . . . .	51
3.3.2. Structures de solvation . . . . .	57
<b>4. Correction à trois corps</b>	<b>60</b>
4.1. Retour sur les solutés chargés . . . . .	60
4.2. Terme de correction à 3 corps . . . . .	62
4.3. Application à la solvation . . . . .	66
4.3.1. Structure . . . . .	67
<b>5. Hydrophobicité et couplage multi-échelle</b>	<b>74</b>
5.1. La solvation des solutés hydrophobes aux échelles microscopique et mésoscopique	74
5.2. La théorie MDFT/HRF avec HSB . . . . .	75
5.3. Description de l'hydrophobicité à différentes échelles dans MDFT . . . . .	77
5.3.1. Théorie . . . . .	77
5.3.2. Résultats . . . . .	83
<b>6. Implémentation numérique de la MDFT</b>	<b>86</b>
6.1. Discrétisation de la densité sur une double grille spatiale et angulaire . . . . .	86
6.2. Fonctionnement du code mdft . . . . .	87
6.2.1. Description moléculaire du soluté et du solvant . . . . .	87
6.2.2. Calcul du potentiel extérieur . . . . .	88
6.2.3. Initialisation de la densité . . . . .	90
6.2.4. Fonctions de corrélation directe . . . . .	90
6.2.5. Minimiseur et minimisation . . . . .	91
6.3. Calculs des différentes intégrales rencontrées . . . . .	92
6.3.1. Intégrales sur le volume . . . . .	92
6.3.2. Intégrales sur la grille angulaire . . . . .	92
6.3.3. Intégrales doubles sur le volume . . . . .	94
6.3.4. Calcul des densités de charges et de polarisation microscopique . . . . .	96
6.4. Transformées de Fourier rapides . . . . .	96
6.5. Influence du maillage sur la convergence . . . . .	97
6.6. Conclusion . . . . .	99
<b>III. Applications</b>	<b>100</b>
<b>7. Applications</b>	<b>101</b>
7.1. Calculs de potentiels de force moyenne (PMF) . . . . .	102
7.1.1. Fonctionnelle . . . . .	102
7.1.2. Résultats . . . . .	103

7.2. Étude de la solvation d'une argile . . . . .	107
7.2.1. Solvation par le fluide de Stockmayer . . . . .	108
7.2.2. Solvation par l'eau . . . . .	116
7.3. Étude d'une molécule d'intérêt biologique : le lysozyme . . . . .	122
<b>8. Conclusions et perspectives</b>	<b>124</b>
8.1. Conclusions . . . . .	124
8.2. Perspectives . . . . .	126
<b>IV. Appendices</b>	<b>129</b>
<b>A. Notions de fonctionnelle et de dérivation fonctionnelle</b>	<b>130</b>
A.1. Fonctionnelle . . . . .	130
A.2. Dérivée fonctionnelle . . . . .	130
<b>B. Démonstration : Le grand potentiel est une fonctionnelle des densités de particule et de polarisation</b>	<b>132</b>
B.1. La densité qui minimise le grand potentiel est la densité d'équilibre . . . . .	132
B.2. L'énergie libre intrinsèque est une fonctionnelle unique des densités de particule et de polarisation . . . . .	132
B.3. Démonstration : $\mathcal{F}$ est une fonctionnelle unique de $n(\mathbf{r})$ et $\mathbf{P}(\mathbf{r})$ . . . . .	133
<b>C. Fonctionnelle de la densité du fluide de sphères dures</b>	<b>135</b>
<b>D. Dérivées des fonctionnelles</b>	<b>137</b>
<b>E. Couplage densité-polarisation dans la fonctionnelle d'excès multipolaire</b>	<b>139</b>
<b>Bibliographie</b>	<b>144</b>

## Une théorie de la fonctionnelle de la densité moléculaire pour la solvation dans l'eau

La théorie de la fonctionnelle de la densité classique est utilisée pour étudier la solvation de solutés quelconques dans le solvant eau. Une forme approchée de la fonctionnelle d'excès pour l'eau est proposée. Cette fonctionnelle nécessite l'utilisation de fonctions de corrélation du solvant pur. Celles-ci peuvent être calculées par simulations numériques, dynamique moléculaire ou Monte Carlo ou obtenues expérimentalement. La minimisation de cette fonctionnelle donne accès à l'énergie libre de solvation ainsi qu'à la densité d'équilibre du solvant. Différentes corrections de cette fonctionnelle approchée sont proposées. Une correction permet de renforcer l'ordre tétraédrique du solvant eau autour des solutés chargés, une autre permet de reproduire le comportement hydrophobe à longue distance de solutés apolaires. Pour réaliser la minimisation numérique de la fonctionnelle, la théorie a été implémentée sur une double grille tridimensionnelle pour les coordonnées angulaires et spatiales, dans un code de minimisation fonctionnelle écrit en Fortran moderne, mdft. Ce programme a été utilisé pour étudier la solvation en milieu aqueux de petits solutés atomiques neutres et chargés et de petites molécules polaires et apolaires ainsi que de solutés plus complexes, une argile hydrophobe et une petite protéine. Dans chacun des cas la théorie de la fonctionnelle de la densité classique permet d'obtenir des résultats similaires à ceux théoriquement exacts obtenus par dynamique moléculaire, avec des temps de calculs inférieurs d'au moins trois ordres de grandeurs.

---

## A molecular density functional theory to study solvation in water

A classical density functional theory is applied to study solvation of solutes in water. An approximate form of the excess functional is proposed for water. This functional requires the knowledge of pure solvent direct correlation functions. Those functions can be computed by using molecular simulations such as molecular dynamic or Monte Carlo. It is also possible to use functions that have been determined experimentally. The functional minimization gives access to the solvation free energy and to the equilibrium solvent density. Some correction to the functional are also proposed to get the proper tetrahedral order of solvent molecules around a charged solute and to reproduce the correct long range hydrophobic behavior of big apolar solutes. To proceed the numerical minimization of the functional, the theory has been discretized on two tridimensional grids, one for the space coordinates, the other for the angular coordinates, in a functional minimization code written in modern Fortran, mdft. This program is used to study the solvation in water of small solutes of several kind, atomic and molecular, charged or neutral. More complex solutes, a neutral clay and a small protein have also been studied by functional minimization. In each case the classical density functional theory is able to reproduce the exact results predicted by MD. The computational cost is at least three order of magnitude less than in explicit methods.



# Notations

Les notations suivantes sont utilisées dans ce manuscrit.

- $(\theta, \phi)$  désigne les deux angles d'Euler des coordonnées sphériques,  $\theta$  est la colatitude qui varie entre 0 et  $\pi$ , et  $\phi$  la latitude qui varie entre 0 et  $2\pi$ .
- $\psi$  est le troisième angle d'Euler désignant la rotation propre autour du vecteur radial.
- $\boldsymbol{\Omega}$  est une notation compacte des trois angles d'Euler  $(\theta, \phi, \psi)$ .
- $(x, y, z)$  désigne les coordonnées cartésiennes.
- $\mathbf{r}$  est le vecteur position, c'est donc une notation compacte des coordonnées cartésiennes.
- $\star$  désigne le produit de convolution, par exemple  $[f \star g](x) = \int_{-\infty}^{\infty} f(x-u)g(u)du$ .
- $\hat{f}(\mathbf{k}) = \iiint_{R^3} f(\mathbf{r}) \exp(-2i\pi \mathbf{k} \cdot \mathbf{r}) d\mathbf{r}$  désigne la transformée de Fourier de la fonction  $f$ .
- $\mathbf{k}$  est le vecteur réciproque.
- $\beta = (k_B T)^{-1}$  est l'inverse de l'énergie thermique.
- $\tilde{\mathbf{u}}$  désigne le vecteur unitaire  $\mathbf{u} / \|\mathbf{u}\|$ .
- $\bar{f}$  désigne la fonction gros grains obtenue à partir de la fonction  $f$ .
- $\mathcal{F}[n(\mathbf{r})]$  désigne une fonctionnelle de la fonction  $n$ , de variable  $\mathbf{r}$ . La notation en crochets est réservée aux fonctionnelles, qui sont généralement écrites en caractères calligraphiés.
- $\frac{\delta \mathcal{F}[n(\mathbf{r})]}{\delta n(\mathbf{r})}$  désigne la dérivée fonctionnelle par rapport à la fonction  $n$ .
- $\delta(x)$  désigne la fonction de Dirac.
- $\|\cdot\|$  désigne la norme euclidienne.
- $|\cdot|$  est la fonction valeur absolue.

Première partie .

État de l'art

# 1. Modélisation des propriétés de solvation

## 1.1. Solvation

La solvation est définie par l'Union Internationale de Chimie Pure et Appliquée (IUPAC)[1] comme :

*Toute interaction stabilisante d'un soluté et d'un solvant [...]. De telles interactions mettent en jeu généralement des forces électrostatiques et de Van der Waals, ainsi que des effets plus spécifiques chimiquement, tels que la formation de liaisons hydrogènes.*

Une solution est un mélange homogène dans lequel une entité chimique est présente en grande quantité (le solvant), et où une ou plusieurs autres entités chimiques (les solutés) sont présentes en plus faible quantité. Bien que les solutions solides existent, elles ne seront pas du tout abordées dans ce manuscrit où on se limitera à l'étude des solutions liquides.

Dans le cadre du traitement complet des interactions entre noyaux et électrons réalisé en chimie quantique, les forces agissant entre solutés et molécules de solvant ont toutes pour origine l'interaction électrostatique. On peut s'interroger dès lors sur la signification des termes *forces de Van-der-Waals* et *liaisons hydrogènes* qui, dans la définition, semblent venir en plus des interactions électrostatiques. Un traitement complet et précis des interactions électrostatiques est complexe. Historiquement, les chimistes ont été amenés à les représenter par des modèles plus simples d'interactions comme les forces de Van-der-Waals et les liaisons hydrogènes.

Il y a formation d'une phase homogène si les interactions qui se développent entre solutés et solvant sont globalement plus favorables que la somme des interactions solutés-solutés et solvant-solvant perdues lors de la mise en solution. Ces interactions entre solutés et solvant jouent un rôle clé dans la mise en solution mais aussi dans la mise en œuvre des réactions chimiques qui se produisent en solution. Comme nous allons le voir dans la suite de ce chapitre, l'estimation de ces interactions a depuis longtemps intéressé les physico-chimistes. Le but de cette thèse est le développement d'une méthode théorique et son implémentation numérique, pour étudier les questions liées à la solvation en milieu aqueux. Il nous faut donc définir plus précisément les grandeurs qui vont être utilisées pour décrire la solvation. On s'intéressera d'abord à l'aspect énergétique mentionné auparavant puis à la structure de solvation, qui décrit l'arrangement spatial des molécules de solvant autour des solutés.

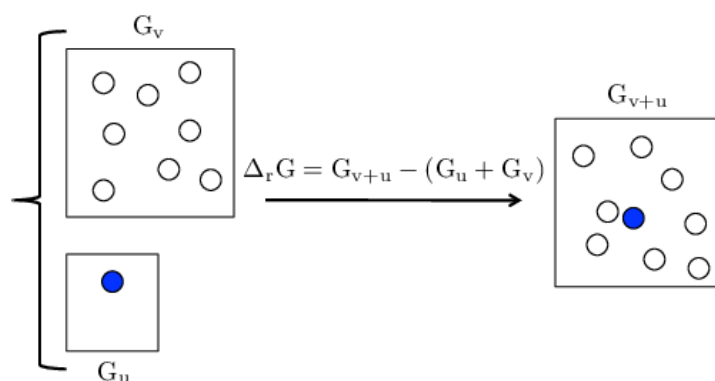
### 1.1.1. Aspect énergétique de la solvation

Nous avons vu que la solvation est la conséquence d'interactions (ou de forces) intermoléculaires se développant entre toutes les molécules de la solution, solvant et solutés.

Au premier abord, on pourrait penser qu'une bonne estimation de l'effet de la solvation est de considérer la différence entre la somme de toutes ces interactions dans le cas de chacun des constituants pris dans des phases pures et la somme de ces mêmes interactions dans la solution. Le signe de cette différence renseigne directement sur la stabilisation ou la déstabilisation causée par la mise en solution. Cependant, cette énergie microscopique dépend de la position de chacune des molécules. Ainsi, à chaque configuration différente (description détaillée des positions de toutes les molécules) est associée une valeur différente de l'énergie. De plus, il est évident qu'en procédant de cette façon, c'est-à-dire en ne considérant que les termes énergétiques, on manque un terme important qui est le terme entropique. Les observables appropriées sont des grandeurs macroscopiques qui sont des moyennes des configurations microscopiques et qui ne dépendent plus des positions instantanées.

La bonne façon de quantifier la solvation est de considérer plutôt les différentes énergies libres mises en jeu lors de ce processus. Pour le physico-chimiste la grandeur macroscopique la plus couramment considérée est l'énergie libre de Gibbs, aussi appelée enthalpie libre. C'est la mesure pertinente du travail d'un système à nombre de particules, température et pression fixés.

On considérera des variations (d'énergie, d'entropie, etc) entre l'état final de la réaction (ici les espèces mises en solution) et son état initial (les espèces prises dans leurs phases pures respectives), ces deux états étant pris à l'équilibre thermodynamique, voir fig. 1.1.1. On utilise généralement des grandeurs intensives, les grandeurs molaires. Toutes ces grandeurs peuvent être mesurées expérimentalement afin de conclure sur la stabilisation relative obtenue par la mise en solution.



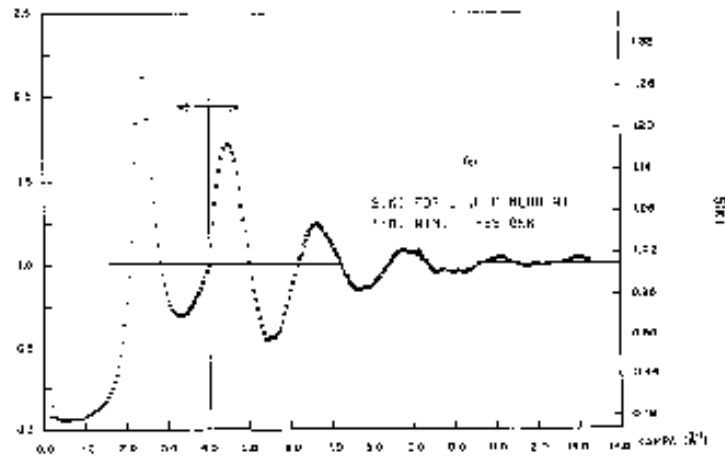
**FIGURE 1.1.1.:** Vue schématique de l'enthalpie libre de solvation, définie comme la différence entre l'enthalpie libre de la solution ( $G_{v+u}$ ) et la somme des enthalpies libres des molécules de solvants ( $G_v$ ) et solutés ( $G_u$ ) purs.

L'énergie libre de solvation (dimensionnée comme une énergie divisée par une quantité de matière) est parfaite pour mesurer à quel point la mise en solution est favorable, mais ne donne en

revanche aucune information sur l'arrangement spatial des molécules. L'organisation des molécules joue pourtant un rôle important dans l'explication des propriétés physiques et chimiques des réactions en solution. Il est donc judicieux de s'intéresser à des grandeurs informant sur l'organisation des molécules dans la solution pour connaître l'influence du soluté sur la structure spatiale du solvant.

### 1.1.2. La structure de solvation

En ce qui concerne l'arrangement tridimensionnel des molécules dans la solution, le saut microscopique-macroscopique effectué dans la section précédente est encore plus pertinent. En effet, la connaissance exacte de la position des molécules à chaque instant est bien sûr la donnée contenant le plus d'information sur la structure. Cependant, les positions des molécules sont sans cesse changeantes ce qui rend l'information de leurs positions instantanées très peu lisible et surtout impossible à mesurer expérimentalement. Il est donc indispensable d'introduire des grandeurs permettant de mesurer l'ordre moyen relatif de la solution. Expérimentalement ceci peut être fait en déterminant le facteur de structure  $S$  qui est lié à l'intensité du signal lumineux diffracté par l'échantillon. Comme pour un solide celui-ci est mesuré par des expériences de diffraction de neutrons ou de rayons X[2]. Contrairement à un cristal, le liquide est isotrope. De ce fait ce facteur de structure ne dépend que de la norme du vecteur de diffraction  $\|\mathbf{k}\|$  et non de son orientation. Si l'on représente le facteur de structure en fonction de la norme du vecteur de diffraction, un exemple est donné en fig. 1.1.2 pour le néon, on remarque que celui ci comporte des pics larges témoins de l'existence d'un ordre à courte distance.



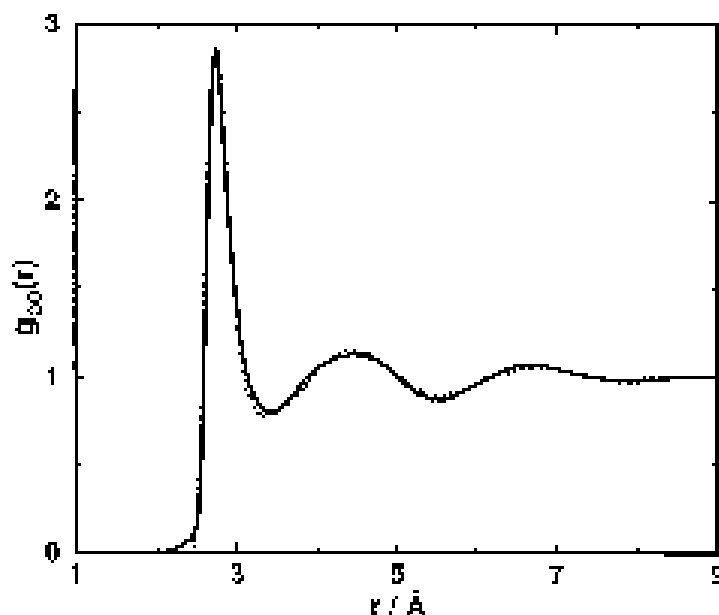
**FIGURE 1.1.2.:** Facteur de structure du néon liquide à 35.05K à une densité  $\rho = 0.03469 \text{ \AA}^{-3}$  obtenue par diffraction de neutrons [3] en fonction de la norme du vecteur de diffraction.

Dans le cadre de la physique des liquides, on utilise plus souvent la fonction de distribution radiale, notée  $g$ , qui est liée au facteur de structure par la relation[4] :

$$S(k) = 1 + \rho \iiint_{\mathbb{R}^3} e^{-i\mathbf{k}\cdot\mathbf{r}} g(r) d\mathbf{r}, \quad (1.1.1)$$

où  $k$  est le nombre d'onde du vecteur diffraction et  $\rho$  la densité du liquide homogène. Cette fonction de distribution radiale mesure le nombre moyen de molécules dans une coquille de rayon  $r$  et d'épaisseur  $dr$ , centrée sur une molécule donnée, normalisée par la quantité de molécule dans cette même coquille. Cette définition justifie donc que la fonction de distribution radiale tende vers 1 pour les grandes valeurs de  $r$ .

Une façon intuitive de voir cette fonction est de constater qu'elle représente la probabilité à l'équilibre thermodynamique, de trouver une molécule à une distance  $r$  d'une molécule donnée.



**FIGURE 1.1.3.:** Fonction de distribution radiale entre les oxygènes des molécules d'eau. La courbe noire est obtenue expérimentalement par diffraction de rayons X. La courbe grise est obtenue par dynamique moléculaire avec le modèle d'eau TIP5P[5].

Un exemple de fonction de distribution radiale entre les oxygènes des molécules d'eau dans le solvant pur est donné en fig. 1.1.3. On remarque que cette fonction présente un premier pic relativement fin autour de 3 Å. Ce pic correspond à une probabilité plus importante que dans un liquide parfaitement homogène de rencontrer deux atomes d'oxygènes séparés de cette distance. Ce premier pic correspond à une première sphère de solvation bien définie. L'eau voisine de la molécule d'eau que l'on regarde est très ordonnée. La déplétion observée juste après ce premier pic s'explique par des considérations stériques, il est en effet moins probable de trouver des atomes d'oxygène près de la région de haute densité précédente. On observe ensuite des oscillations avec des extrema de moins en moins importants jusqu'à atteindre la valeur asymptotique 1. On peut en conclure que la fonction de distribution radiale est une grandeur pertinente lorsqu'on s'intéresse à la structure de solvation. La courbe grise qui apparaît sur la fig. 1.1.3 est obtenue par dynamique moléculaire en utilisant le modèle d'eau non-polarisable TIP5P, un des modèles les plus récents, parmi les centaines de modèles qui ont été proposés pour l'eau. Cette courbe reproduit parfaitement les résultats expérimentaux.

On constate que l'accès aux structures et aux grandeurs énergétiques de solvation par des

méthodes numériques est donc possible. Comme mis en évidence sur la fig. 1.1.3 il faut pour cela développer et utiliser des méthodes ou théories qui permettent, pour un modèle de molécule donné, d'obtenir ces propriétés physiques à l'équilibre thermodynamique.

L'étude des propriétés de solvation, c'est-à-dire la bonne représentation des interactions entre molécules de soluté et de solvant est un problème physico-chimique important. Le but étant de reproduire ou prédire des résultats expérimentaux et notamment les propriétés énergétiques et structurales de la solvation.

Dans les deux prochaines parties de ce chapitre sont présentées différentes méthodes permettant de calculer les propriétés de solvation. Les simulations moléculaires, dites explicites, décrivent le solvant par ces configurations microscopiques. On peut également utiliser des méthodes de solvant implicite où le solvant est traité à un niveau théorique où certains détails moléculaires sont omis.

## 1.2. Modèles explicites

Pour calculer les propriétés de solvation on peut envisager d'utiliser une représentation microscopique réaliste du système étudié, c'est-à-dire représenter explicitement les molécules de solvant et de soluté. En choisissant un modèle des interactions entre molécules, on peut placer un ensemble de molécules à des positions diverses, et calculer l'énergie de cette configuration microscopique. La théorie physique microscopique pertinente pour déduire les propriétés macroscopiques à l'équilibre à partir de la connaissance des configurations microscopiques est la physique statistique. On se propose d'en rappeler quelques points clés dans le prochain paragraphe

### 1.2.1. Méthodes de solvant explicite

Dans les techniques de simulations de solvant explicite, on représente la solution par une collection de molécules. Celles-ci interagissent au travers de champ de force, ou de manière équivalente, d'une énergie potentielle  $\mathcal{U}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  dont dérive ces forces, modélisant les interactions entre molécules.

On peut décomposer cette énergie potentielle en la somme de deux composantes : une partie intramoléculaire, décrivant les interactions entre atomes d'une même molécule, et une partie intermoléculaire décrivant les interactions entre atomes de molécules différentes. Il existe de nombreux types de champs de forces pouvant décrire ces deux composantes. Dans le cadre de cette thèse, on ne considérera que des modèles de solvants rigides, c'est-à-dire que les longueurs et angles de liaison des molécules sont fixes. Une telle approximation revient à considérer que l'énergie intramoléculaire n'est jamais modifiée, ce qui permet de ne pas avoir à la décrire. Pour ce qui concerne les interactions intermoléculaires, on se limite à deux types d'énergie potentielle.

L'interaction entre deux sites chargés séparés d'une distance  $r$  est décrite par l'interaction cou-

lombienne, d'énergie potentielle  $V_C$ ,

$$V_C(r) = \frac{z_1 z_2 e^2}{4\pi\epsilon_0 r}, \quad (1.2.1)$$

où les charges des atomes sont respectivement  $z_1 e$  et  $z_2 e$ , avec  $e \approx 1.602 \cdot 10^{-19} \text{ C}$  la charge élémentaire, et  $z_1, z_2$  les charges relatives,  $\epsilon_0 \simeq 8.854 \cdot 10^{-12} \text{ F.m}^{-1}$  la permittivité du vide.

L'autre potentiel utilisé est le potentiel de Lennard-Jones, qui modélise l'ensemble des interactions de Van der Waals, à savoir les forces de Keesom, Debye et London, ainsi que la répulsion de Pauli à courte distance. Ce potentiel d'interaction entre deux atomes séparés d'une distance  $r$  a pour expression,

$$V_{LJ}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (1.2.2)$$

où  $\epsilon$  est la profondeur du puit de potentiel, et  $\sigma$  est relié à la position du minimum de ce puit de potentiel. Une représentation de ce potentiel d'interaction est donnée en fig. 1.2.1.

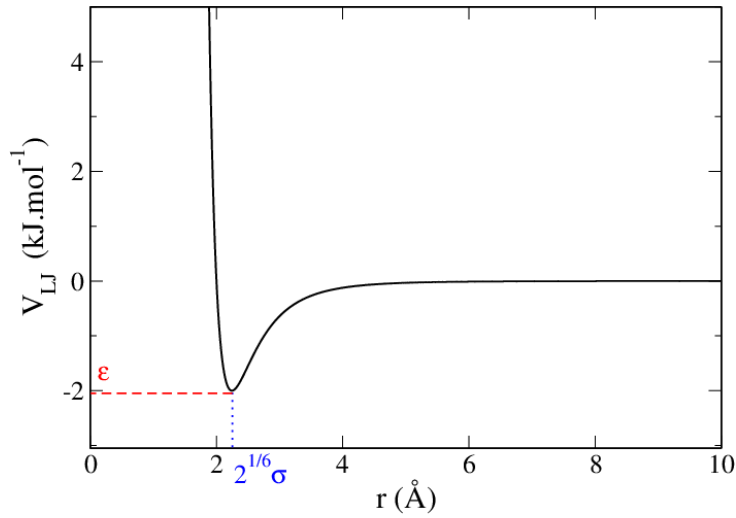


FIGURE 1.2.1.: Potentiel de Lennard-Jones avec  $\epsilon = 2.0 \text{ kJ.mol}^{-1}$ , et  $\sigma = 2 \text{ Å}$

### 1.2.2. Bases de thermodynamique statistique

Supposons que l'on ait une collection de  $N$  particules dans un volume  $V$  à une température  $T$  donnée, interagissant au travers d'une d'énergie potentielle  $\mathcal{U}$ . Si l'on décrit la position de chaque particule  $i$  par leur vecteur position  $\mathbf{r}_i$ , l'énergie totale associée à cette configuration est la somme d'une partie cinétique et d'une partie potentielle due aux forces entre molécules. Elle peut être décrite par la fonction de Hamilton, appelée par la suite hamiltonien,

$$\mathcal{H}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + \mathcal{U}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N). \quad (1.2.3)$$



Où  $\mathbf{p}_i$  est la quantité de mouvement de la molécule  $i$  de masse  $m_i$  et  $\mathcal{U}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  est l'énergie potentielle totale de la configuration étudiée.

Cet hamiltonien dépend de  $2N$  variables à valeurs dans  $\mathbb{R}^3$ . Il existe un espace à  $6N$  dimensions, appelé espace des phases, tel que chaque point (de coordonnées  $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N$ ) de cet espace représente une configuration de la collection de molécules.

Si le nombre de degrés de liberté, (le nombre de particules) est grand, il est alors impossible de connaître exactement la configuration microscopique. On utilise donc une description statistique des configurations microscopiques. La distribution de probabilité des configurations de l'espace des phases obéit à une statistique de Maxwell-Boltzmann[6],

$$P(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) = \frac{\frac{1}{N!h^{3N}} e^{-\beta \mathcal{H}(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)}}{\frac{1}{N!h^{3N}} \iiint_{\mathbb{R}^3} \dots \iiint_{\mathbb{R}^3} e^{-\beta \mathcal{H}(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)} d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N}, \quad (1.2.4)$$

où  $\beta = 1/k_B T$ , avec  $T$  la température absolue,  $k_B = 1.38066 \cdot 10^{-23} \text{ J.K}^{-1}$  la constante de Boltzmann et  $h = 6.62606957 \cdot 10^{-34} \text{ J.s}$  la constante de Planck. Le terme au dénominateur est d'une importance particulière, on le nomme fonction de partition et on le note usuellement  $Q$ . On peut remarquer que cette distribution de probabilité est normalisée. La démonstration de ce résultat n'est pas présentée ici mais on pourra se reporter à [7] pour l'y trouver.

L'énergie libre de Helmholtz,  $F$ , (aussi appelée énergie libre) est reliée à la fonction de partition  $Q$ ,

$$F = -k_B T \ln Q. \quad (1.2.5)$$

On remarque que  $F$  est définie à une constante près. Si l'on connaît toutes les configurations accessibles au système, c'est-à-dire si on est capable de proposer une méthode permettant de visiter entièrement l'espace des phases, on peut, en utilisant l'éq. 1.2.5, calculer l'énergie libre du système.

L'énergie libre est une grandeur particulièrement difficile à calculer grâce à des simulations moléculaires car elle ne se réduit pas à un calcul de valeur moyenne. Aussi on calcule toujours des différences d'énergie libre entre deux systèmes. Ceci revient, pour une transformation comme celle décrite sur la fig. 1.1.1, à calculer l'énergie libre de solvation.

En ce qui concerne les propriétés structurales, telles que les fonctions de distribution radiale, celles-ci sont facilement calculables en utilisant des méthodes d'histogrammes.

La physique statistique est un outil puissant pour calculer les grandeurs d'intérêt si on connaît l'espace des phases. Pour échantillonner cet espace on utilise principalement deux grandes familles de simulations moléculaires[6].

La première est la dynamique moléculaire qui consiste, à partir d'une configuration initiale de molécules à visiter l'espace des phases en résolvant les équations de la mécanique classique. Pour simplifier, à partir d'une configuration à un instant  $t$ , on calcule l'ensemble des forces agissant sur les molécules grâce aux potentiels d'interaction choisis. À partir du calcul de ces forces on

utilise une version discrète des lois de Newton pour déterminer la position des molécules au pas suivant, c'est-à-dire à l'instant  $t + \Delta t$ , avec  $\Delta t$  le pas de temps de la simulation. On refait alors un nouveau calcul de forces et on itère le procédé.

L'espace des phases peut également être échantillonné par la méthode de Monte-Carlo, qui est une méthode stochastique. Dans cette méthode, l'espace des phases est visité par des déplacements aléatoires qui sont acceptés ou refusés en fonction de la probabilité d'existence des états atteints par ces déplacements.

Une représentation microscopique réaliste du solvant nécessite d'utiliser un grand nombre de molécules, jusqu'à plusieurs centaines de milliers, pour étudier la mise en solution d'une molécule biologique. Une grande partie de la puissance de calcul utilisée est principalement consacrée à l'évaluation des interactions solvant-solvant, plutôt qu'à celles mises en jeu entre les solutés et le solvant. De plus pour que l'espace des phases soit correctement échantillonné il faut générer un grand nombre de configurations ce qui rallonge d'autant plus le temps de calcul. Pour ces raisons l'utilisation de simulations numériques pour étudier les propriétés de solvation est numériquement coûteux. Le calcul de l'énergie libre de solvation est encore plus problématique puisqu'il nécessite l'utilisation de méthodes d'intégrations thermodynamiques telles que l'Umbrella Sampling[8], qui impose la réalisation de plusieurs simulations moléculaires sur des systèmes intermédiaires entre l'état de départ (solvant bulk) et l'état d'arrivée (soluté en solution).

Il peut être intéressant de développer des méthodes numériquement plus efficaces, notamment pour étudier des systèmes de taille importante. Les méthodes de solvants implicites sont une alternative intéressante puisque le solvant y est décrit avec moins de détails, ce qui diminue le coût numérique de l'évaluation des interactions entre molécules de solvant.

### 1.3. Méthodes de solvant implicite

Conceptuellement, on peut décomposer l'enthalpie libre de solvation en une somme de trois termes,

$$\Delta G_{\text{solv}} = \Delta G_{\text{elec}} + \Delta G_{\text{VdW}} + \Delta G_{\text{cav}}, \quad (1.3.1)$$

où  $\Delta G_{\text{elec}}$  est la contribution due à l'électrostatique, qui provient du travail nécessaire pour créer la distribution de charge du soluté dans la solution et pour polariser cette distribution de charge sous l'influence du solvant,  $\Delta G_{\text{VdW}}$  est la contribution due aux forces de Van-der-Waals (dispersion et répulsion),  $\Delta G_{\text{cav}}$  correspond au coût nécessaire à la création de la cavité contenant le soluté. Ce dernier terme contient le coût entropique de réorganisation des molécules de solvant et la résistance à la pression du solvant pour former la cavité. Ces deux derniers termes sont généralement réunis en une contribution non polaire,

$$\Delta G_{\text{solv,np}} = \Delta G_{\text{cav}} + \Delta G_{\text{VdW}}. \quad (1.3.2)$$

Il existe plusieurs théories permettant d'estimer cette contribution à l'énergie libre (*Scaled Particle Theory*[9, 10, 11, 12] et *Solvent-Exposed Surface Area*[13]) en la décomposant en plusieurs termes dont le terme principal est essentiellement un terme de tension de surface,

$$\Delta G_{\text{solv,np}} \approx \gamma_v A, \quad (1.3.3)$$

où  $A$  est la surface accessible au solvant et  $\gamma_v$  la tension de surface liquide-gaz du solvant. Cette expression est quasi-exacte pour un soluté de sphères dures de grandes tailles. En pratique, pour des solutés plus complexes, les paramètres sont ajustés pour reproduire les énergies de solvation expérimentales. La partie électrostatique peut être évaluée à partir de nombreux modèles de solvants implicites, qui diffèrent essentiellement par l'échelle de description du solvant.

### 1.3.1. Milieu diélectrique continu

#### 1.3.1.1. Poisson-Boltzmann

Cette méthode suppose que le soluté est fixe et qu'il est défini comme un milieu de faible permittivité diélectrique,  $\epsilon^u$  (variant typiquement de 1 à 8). Il forme une cavité dans un continuum diélectrique polarisable caractérisé par une permittivité diélectrique élevée  $\epsilon^v$  (80 pour l'eau). Les charges du soluté sont traitées explicitement. Le problème se ramène à un problème d'électrostatique courant qui peut être formulé grâce à une équation de Poisson,

$$\nabla \cdot \mathbf{D}(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\epsilon_0}, \quad (1.3.4)$$

où  $\sigma$  est la densité de charge,  $\epsilon_0$  la permittivité du vide et  $\mathbf{D}$  le déplacement diélectrique défini comme,

$$\mathbf{D}(\mathbf{r}) = -\epsilon(\mathbf{r})\nabla\Phi(\mathbf{r}), \quad (1.3.5)$$

où  $\epsilon(\mathbf{r})$  est la permittivité relative et  $\Phi$  le potentiel électrostatique. Si on connaît  $\Phi$ , on peut calculer l'énergie libre de solvation du soluté,

$$\Delta G_{\text{elec}} = \frac{1}{2} \iiint_{\mathbb{R}^3} \sigma(\mathbf{r}) (\Phi(\mathbf{r}) - \Phi_0(\mathbf{r})) d\mathbf{r}, \quad (1.3.6)$$

avec  $\Phi_0(\mathbf{r})$  le potentiel électrostatique créé par le soluté dans le vide.

Si le système à modéliser est un soluté entouré d'un solvant contenant des ions en solution ( $N_I$  types d'ions), la densité de charge totale peut s'écrire,

$$\sigma(\mathbf{r}) = \sigma^{\text{IA}}(\mathbf{r}) + \sigma^u(\mathbf{r}), \quad (1.3.7)$$

où  $\sigma^u(\mathbf{r})$  désigne la densité de charge du soluté et  $\sigma^{\text{IA}}(\mathbf{r})$  celle de l'atmosphère ionique. Il est possible de faire l'hypothèse d'une distribution de Boltzmann pour les ions en solution à l'équilibre

(d'où le nom de la méthode) sous l'effet du champ électrique,

$$\sigma^{\text{IA}}(\mathbf{r}) = \sum_{i=1}^{N_I} z_i e n_i^0 \exp\left(\frac{z_i e \Phi(\mathbf{r})}{k_B T}\right), \quad (1.3.8)$$

où  $n_i^0$  est la densité ionique de l'ion  $i$  en l'absence d'interaction et  $z_i$  la valence de l'ion  $i$ . L'éq. 1.3.4 se réécrit alors,

$$\nabla \cdot [\epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r})] = \sum_{i=1}^{N_I} \frac{z_i e}{\epsilon_0} n_i^0 \exp\left(\frac{z_i e \Phi(\mathbf{r})}{k_B T}\right) + \frac{\sigma^u(\mathbf{r})}{\epsilon_0}, \quad (1.3.9)$$

qui est l'équation de Poisson-Boltzmann.

Cette équation est exacte et constitue la référence pour les modèles de solvants implicites. Malheureusement, elle n'est pas résoluble analytiquement avec des systèmes d'intérêt pour lesquels la densité de charge a une forme non-triviale. Il existe néanmoins de nombreux développements permettant de résoudre cette équation de manière numérique[14], basés par exemple sur la méthode des éléments finis[15] ou des différences finies[16, 17, 18, 19, 20]. Il existe de plus des codes mis à la disposition de la communauté par exemple le code DelPhi[21, 22] ou le code APBS[15]

### 1.3.1.2. Méthode de Born généralisée

La méthode de Born généralisée est une approximation de la méthode de Poisson-Boltzmann qui vise à donner des résultats comparables avec un coût numérique inférieur. On souhaite évaluer l'énergie libre électrostatique d'un ensemble de  $N_{\text{tot}}$  particules, de rayon de Born  $\alpha_i$  et de charge  $q_i$  dans un milieu de constante diélectrique  $\epsilon$ . Celle-ci est définie comme la somme de l'énergie de Coulomb dans un milieu diélectrique et d'une correction de Born,

$$\Delta G_{\text{elec}} = -\frac{1}{8\pi} \left(1 - \frac{1}{\epsilon}\right) \sum_{i=1}^{N_{\text{tot}}} \frac{q_i^2}{\alpha_i} \quad (1.3.10)$$

Cette expression est vraie si les rayons de Born des particules ne se recouvrent pas. Malheureusement, pour un soluté réel on ne connaît pas les rayons de Born des différentes particules (ou sites chargés).

Une généralisation empirique pour des particules dont les rayons de Born sont autorisés à se recouvrir a été proposée [23],

$$\Delta G_{\text{elec}} = \frac{1}{8\pi\epsilon} \sum_{i=1}^{N_{\text{tot}}} \sum_{j>i}^{N_{\text{tot}}} \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp\left(\frac{-r_{ij}^2}{4\alpha_i \alpha_j}\right)}} \left(1 - \frac{1}{\epsilon}\right) \quad (1.3.11)$$

La détermination des rayons de Born  $\alpha_i$  des atomes  $i$  se fait en calculant l'énergie libre électrostatique de l'atome  $i$ , avec les charges de toutes les autres particules égales à 0,  $\Delta G_{\text{elec},0}^i$ . Notons

cependant que le diélectrique est toujours égal à celui du soluté pour les particules de charges nulles. Par exemple, ceci peut être fait en utilisant l'équation de Poisson-Boltzmann. Le rayon de Born de cette particule est  $\alpha_i$  telle que l'énergie libre électrostatique calculée avec la formule de Born simple d'une particule sphérique calculée dans un continuum,  $\Delta G_{\text{Born}}^i$

$$\Delta G_{\text{Born}}^i = -\frac{q^2}{8\pi\epsilon_0\alpha_i} \left(1 - \frac{1}{\epsilon}\right) \quad (1.3.12)$$

soit la même que celle obtenue,  $\Delta G_{\text{elec},0}^i$ . On trouve donc,

$$\alpha_i = -\frac{1}{8\pi\epsilon_0} \left(1 - \frac{1}{\epsilon}\right) \frac{1}{\Delta G_{\text{elec},0}^i}. \quad (1.3.13)$$

Le coût numérique de cette méthode réside dans la détermination de ces rayons de Born  $\alpha_i$ , qui requiert pour chaque particule le calcul d'une énergie libre électrostatique pour laquelle on a éteint les charges de toutes les autres particules. De plus, si on désire utiliser cette méthode dans une simulation dynamique il faut recommencer le processus à chaque pas. D'autres déterminations du rayon de Born, plus efficaces numériquement, ont donc été proposées. La principale est l'approximation CFA (Coulomb Field Approximation) qui consiste à supposer que le déplacement diélectrique est égal au champ électrique coulombien dans le vide, ceci revient à négliger la réponse due au solvant puisque l'on néglige l'effet du continuum diélectrique ce qui tend à sous-estimer les rayons de Born et les énergies libres de solvation. Les rayons de Born peuvent désormais être déterminés numériquement ce qui rend la méthode extrêmement rapide et donc utilisable dans des simulations de dynamique moléculaire. Cependant cette théorie décrit le solvant par un continuum, et ne tient donc pas compte de la nature moléculaire de celui-ci. D'autres théories conservent ce niveau de description, on en décrit rapidement trois dans les paragraphes suivants.

### 1.3.2. Équations intégrales

Les théories des équations intégrales sont basées sur l'équation de Ornstein-Zernike, qui pour un fluide uniforme et isotrope s'écrit,

$$h(r) = c(r) + \rho \iiint_{\mathbb{R}^3} c(\|\mathbf{r} - \mathbf{r}'\|) h(r') d\mathbf{r}' \quad (1.3.14)$$

avec  $h$  la fonction de corrélation totale,  $c$  la fonction de corrélation directe et  $\rho$  le densité du fluide. La fonction  $h$  décrit la variation de la densité locale de particule à une distance  $r$  d'une particule située à l'origine. Elle est reliée à la fonction de distribution radiale par,

$$h(r) = g(r) - 1. \quad (1.3.15)$$

L'éq. 1.3.14 a une interprétation physique forte, la fonction de corrélation totale  $h$  entre deux

particules situées à une distance  $r$  est due à une interaction directe entre ces deux particules  $c$ , ainsi qu'à la somme de toutes les contributions indirectes propagées par l'ensemble de toutes les autres particules (le terme intégrale).

La fonction  $h$  caractérise la structure du liquide à l'équilibre thermodynamique et permet d'en calculer les propriétés d'équilibre. Pour résoudre l'éq. 1.3.14 et calculer les fonctions de corrélation il est néanmoins nécessaire d'utiliser une autre équation qui relie  $h$  et  $c$  appelée relation de fermeture.

On connaît un certain nombre de relations de fermeture relativement adaptées à l'étude de fluide atomique. On peut citer, par exemple, la relation de fermeture HNC, pour hypernetted-chain,

$$g(r) = \exp [-\beta v(r) + h(r) - c(r)]. \quad (1.3.16)$$

Cette relation relie les fonctions  $g$ ,  $h$  et  $c$  à une perturbation extérieure modélisée par un potentiel de paire  $v$ . L'éq. 1.3.16 est une approximation, elle néglige en particulier des corrélations à courte distance. Avec cette relation de fermeture il est désormais possible de résoudre l'éq. 1.3.14.

Lorsque le fluide est composé de molécules, les interactions de paires entre elles ne dépendent plus uniquement de la position relative des centres de masse des deux molécules, mais aussi de leurs orientations respectives. Pour décrire un tel fluide il est donc nécessaire de réécrire une version moléculaire de l'équation de Ornstein-Zernike (MOZ).

$$h(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) = c(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) + \frac{1}{8\pi^2} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \iiint_{\mathbb{R}^3} c(r_{13}, \mathbf{\Omega}_1, \mathbf{\Omega}_3) \rho h(r_{32}, \mathbf{\Omega}_3, \mathbf{\Omega}_2) d\mathbf{r}_3 d\mathbf{\Omega}_3 \quad (1.3.17)$$

On peut alors utiliser une voie similaire à l'étude des fluides simples pour essayer de résoudre MOZ. Il faut à nouveau utiliser une relation de fermeture, par exemple l'éq. 1.3.16 précédente. Le problème est désormais bien plus complexe puisqu'on étudie un fluide moléculaire. En effet, les fonctions  $g, h, v$  et  $c$  intervenant dans HNC sont également des fonctions moléculaires, elles dépendent donc de l'orientation relative et de la distance séparant les deux molécules considérées. La résolution de MOZ avec ces fonctions est numériquement très complexe. Des développements ont néanmoins été réalisés pour contourner ce problème. L'idée générale est de réaliser une projection sur une base d'invariants rotationnels des fonctions mises en jeu [24, 25]  $g, h, c, v$ . Par exemple,

$$c(\mathbf{r}_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) = \sum_{m,n,l,\mu,\nu} c_{\mu\nu}^{mnl}(r) \Phi_{\mu\nu}^{mnl}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \tilde{\mathbf{r}}), \quad (1.3.18)$$

avec les invariants rotationnels,  $\Phi_{\mu\nu}^{mnl}$  :

$$\begin{aligned} \Phi_{\mu\nu}^{mnl}(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \tilde{\mathbf{r}}) &= \sqrt{(2m+1)(2n+1)} \sum_{\mu', \nu', \lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} \\ &\times R_{\mu'\mu}^m(\boldsymbol{\Omega}_1) R_{\nu'\nu}^n(\boldsymbol{\Omega}_2) R_{\lambda'0}^l(\tilde{\mathbf{r}}). \end{aligned} \quad (1.3.19)$$

Les coefficients  $\begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix}$  sont les symboles  $3-j$  de Wigner[26], et les  $R_{\mu'\mu}^m(\boldsymbol{\Omega})$  sont les harmoniques sphériques généralisées de Wigner,  $\tilde{\mathbf{r}} = (\mathbf{r}_1 - \mathbf{r}_2)/\|\mathbf{r}_1 - \mathbf{r}_2\|$  désigne le vecteur unitaire séparant les deux particules considérées. Des projections similaires sont réalisées pour les autres fonctions  $h, v, g$ .

Les invariants rotationnels sont indépendants du référentiel (d'où leur intérêt) et forment une base orthogonale. Ils dépendent de l'orientation relative des deux molécules, donc de 5 angles d'Euler, caractérisés par les cinq indices  $m, n, l, \mu, \nu$ . En principe le développement sur la base des invariants rotationnels est infini mais pour des molécules linéaires (e.g., le fluide Stockmayer) trois angles d'Euler sont suffisants :  $m + n + l$  est pair et  $\mu = \nu = 0$ . On peut donc simplifier la notation dans ce cas précis de molécules linéaires. De plus, il est supposé et vérifié que l'on peut limiter ce développement à un nombre maximal de projections, caractérisées par un doublet  $(m, n)$  pour décrire quantitativement la dépendance angulaire des corrélations. En utilisant ces projections on peut alors résoudre MOZ avec la relation de fermeture HNC. Des développements ont été réalisés dans ce formalisme qui permettent une description précise des corrélations des liquides moléculaires[27].

### 1.3.3. RISM (Reference Interaction Site Model)

La théorie RISM postule que dans une molécule, constituée de sites atomiques, le potentiel d'interaction entre les molécules peut être décrit par une somme de potentiels de paire, qui ne dépendent que de la distance entre sites atomiques. Ceci est une approximation extrêmement courante en simulation moléculaire. On peut réécrire l'éq. 1.3.17 en faisant intervenir des fonctions de corrélation de paire entre sites des molécules  $h_{\alpha\beta}(r)$ , au lieu des fonctions moléculaires. On les calcule en fixant la distance entre sites d'interaction et en faisant une moyenne angulaire. Moyennant une série d'approximations pour pouvoir faire cette moyenne angulaire dans l'éq. 1.3.17, on obtient l'équation matricielle fondamentale de RISM,

$$\mathbf{h}(r) = \boldsymbol{\omega} \star \mathbf{c} \star \boldsymbol{\omega}(r) + \rho \boldsymbol{\omega} \star \mathbf{c} \star \mathbf{h}(r), \quad (1.3.20)$$

$\star$  désigne un produit matriciel et un produit de convolution et  $\mathbf{h}, \mathbf{c}, \boldsymbol{\omega}$  sont des matrices dont les éléments sont des fonctions décrivant des paires de sites moléculaires, qu'on labélise  $\alpha\gamma$ . Ainsi,  $h_{\alpha\gamma}$  est la fonction de corrélation de paire entre les sites  $\alpha$  et  $\gamma$ .  $\omega_{\alpha\gamma}$  est définie par

$$\omega_{\alpha\gamma} = \rho \delta_{\alpha\gamma}(r) + \rho s_{\alpha\gamma}(r), \quad (1.3.21)$$

où  $s_{\alpha\gamma}$  est la fonction de distribution de paire site-site intramoléculaire. Elle décrit la probabilité de trouver un site  $\gamma$  à une position  $\mathbf{r}$  sachant qu'un site  $\alpha$  se trouve à l'origine. Cette équation a été introduite pour des fluides composés de sphères dures mais son utilisation a été généralisée à des fluides composés de sphères Lennard-Jones et portant des charges. Le modèle de molécule se choisit au travers de la relation de fermeture utilisée conjointement à MOZ pour la résolution du problème, c'est-à-dire la détermination des fonctions de corrélation de paire. Le formalisme tel que décrit ici s'applique à un fluide pur.

Si on veut étudier un système soluté-solvant, on peut généraliser l'éq. 1.3.20 pour un mélange de fluide,

$$\rho \mathbf{h} \rho(r) = \boldsymbol{\omega} \star \mathbf{c} \star \boldsymbol{\omega}(r) + \boldsymbol{\omega} \star \mathbf{c} \star \rho \mathbf{h} \rho(r), \quad (1.3.22)$$

où les matrices  $\mathbf{h}, \mathbf{c}, \boldsymbol{\omega}$  sont désormais formées d'éléments  $h_{\alpha_M \gamma_{M'}}, c_{\alpha_M \gamma_{M'}}, \omega_{\alpha_M \gamma_{M'}}$ . L'indice  $\alpha_M$  désigne le  $\alpha^{ième}$  site de la molécule M et  $\gamma_{M'}$  le  $\gamma^{ième}$  site de la molécule M'. La matrice  $\rho$  est la matrice diagonale des densités en nombre des espèces,  $\rho_{\alpha_M \gamma_{M'}} = \delta_{\alpha\gamma} \delta_{MM'} \rho_M$ . On peut réorganiser cette équation matricielle en séparant les termes liés au soluté, désigné par l'indice  $u$ , et ceux liés au solvant, désignés par l'indice  $v$ . Dans la limite où la dilution du soluté est infinie, c'est à dire quand  $\rho_u \rightarrow 0$ , on obtient le jeu d'équations suivants,

$$\mathbf{h}_{vv} = \mathbf{w}_v \star \mathbf{c}_{vv} \star \mathbf{w}_v + \mathbf{w}_v \star \mathbf{c}_{vv} \star \rho_v \mathbf{h}_{vv} \quad (1.3.23)$$

$$\mathbf{h}_{uv} = \mathbf{w}_u \star \mathbf{c}_{uv} \star \mathbf{w}_v + \mathbf{w}_u \star \mathbf{c}_{uv} \star \rho_v \mathbf{h}_{uv} \quad (1.3.24)$$

$$\mathbf{h}_{uu} = \mathbf{w}_u \star \mathbf{c}_{uu} \star \mathbf{w}_u + \mathbf{w}_u \star \mathbf{c}_{uu} \star \rho_v \mathbf{h}_{uv}, \quad (1.3.25)$$

où  $\mathbf{w}_v = \boldsymbol{\omega}_v / \rho_v$  et  $\mathbf{w}_u = \boldsymbol{\omega}_u / \rho_u$ . En utilisant cette théorie avec une relation de fermeture adaptée il devient possible de calculer les fonctions de corrélation du système et d'en déduire des grandeurs physiques d'intérêt, dont l'énergie libre de solvation, via une intégration thermodynamique nécessitant la connaissance des fonctions de corrélation de paires à différents états thermodynamiques entre le solvant pur et le soluté en solution.

Une des limitations de ce modèle est qu'il ne permet pas une description de solutés de géométries complexes car les fonctions de corrélation site-site sont moyennées sur les orientations des molécules.

Pour ces raisons, une généralisation tridimensionnelle de RISM (3D-RISM) autour d'un soluté non sphérique a été proposée.

### 1.3.4. 3D-RISM

Cette théorie utilise des fonctions de corrélation tridimensionnelle pour les sites du solvant proches du soluté. On écrit d'abord l'équation moléculaire de Ornstein-Zernike dépendante des



angles.

$$\begin{aligned}
h^{uv}(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) = & c^{uv}(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) \\
& + \frac{1}{8\pi^2} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \iiint_{\mathbb{R}^3} c^{uv}(r_{13}, \mathbf{\Omega}_1, \mathbf{\Omega}_3) \rho^v h^{vv}(r_{32}, \mathbf{\Omega}_3, \mathbf{\Omega}_2) d\mathbf{r}_3 d\mathbf{\Omega}_3
\end{aligned} \tag{1.3.26}$$

Contrairement à RISM où cette équation était complètement moyennée sur l'ensemble des orientations du soluté et du solvant, la méthode 3D-RISM conserve la description angulaire du soluté. Les fonctions de corrélation de paire des sites de solvants autour des sites  $\gamma$  du soluté ont donc une dépendance angulaire en plus de leur dépendance radiale,

$$h_{\gamma}^{uv}(\mathbf{r}_{1\gamma}) \equiv h_{\gamma}^{uv}(r_{1\gamma}, \mathbf{\Omega}) = \frac{1}{\Omega} \iiint_{\pi^2} h^{uv}(\|\mathbf{r}_{1\gamma} - \mathbf{r}_{2\gamma}\|, \mathbf{\Omega}_1, \mathbf{\Omega}_2) d\mathbf{\Omega}_1, \tag{1.3.27}$$

où  $\mathbf{r}_{1\gamma}$  est le vecteur intermoléculaire entre une molécule de soluté dont le centre de masse est situé en  $\mathbf{r}_1$  et un site de solvant situé en  $\mathbf{r}_{\gamma}$ , et  $\mathbf{r}_{2\gamma}$  le vecteur intramoléculaire entre le centre de masse du solvant situé en 2 et le site  $\gamma$  de la molécule de solvant ayant l'orientation  $\mathbf{\Omega}_2$ . Pour pouvoir effectuer la moyenne angulaire de l'équation de OZ avec dépendance angulaire, l'hypothèse de base de 3D-RISM est que la fonction de corrélation directe 6D  $c^{uv}(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2)$  peut se décomposer en la somme des contributions partielles des sites de la molécule de solvant 2,

$$c^{uv}(r_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) = \sum_{\alpha} c_{\alpha}^{uv}(\mathbf{r}_{1\alpha}). \tag{1.3.28}$$

On transforme l'équation moléculaire de OZ dans l'espace de Fourier et on la moyenne sur  $\mathbf{\Omega}_2$ , on obtient alors l'équation fondamentale de 3D-RISM pour le soluté écrite dans l'espace réciproque,

$$\hat{h}_{\gamma}^{uv}(\mathbf{k}) = \hat{c}_{\alpha}^{uv}(\mathbf{k}) \hat{\omega}_{\alpha\gamma}^{vv}(\mathbf{k}) + \hat{c}_{\alpha}^{uv}(\mathbf{k}) \rho^v \hat{h}_{\alpha\gamma}^{vv}(\mathbf{k}) \tag{1.3.29}$$

Cette équation peut être calculée numériquement pour obtenir les fonctions de corrélation totale  $\hat{h}_{\alpha\gamma}^{vv}(\mathbf{k})$  et directes  $\hat{c}_{\alpha}^{uv}(\mathbf{k})$  par discrétisation sur une grille 3D uniforme et utilisation de transformées de Fourier rapides (FFT) pour calculer les produits de convolution. L'éq. 1.3.23 décrivant les corrélations entre molécules de solvant est inchangée. Il faut bien entendu toujours se munir de relations de fermetures pour pouvoir résoudre cette équation.

La théorie 3D-RISM a d'abord été développée par Roux et collaborateurs avec la relation de fermeture HNC[28] pour étudier la solvation de solutés polaires, eau et N-méthylacétamide, dans le solvant eau. Cette théorie permet de reproduire les propriétés structurales de solvation.

Malgré ce succès, la relation de fermeture HNC converge mal, notamment pour l'étude d'un liquide associé près d'une interface cristalline. Pour contourner ce problème, une version alternative de 3D-RISM a été proposée par Kovalenko et Hirata[29] basée sur une version linéarisée de HNC, la fermeture KH. Cette fermeture permet l'étude d'une interface métal-eau, le métal

étant décrit par DFT électronique.

La théorie 3D-RISM est une manière rapide d'accéder aux propriétés de solvation tout en conservant un niveau de description moléculaire, c'est pourquoi elle est de plus en plus utilisée. Parmi les utilisations notables de cette théorie on peut signaler : le calcul de potentiels de force moyenne[30, 31] et l'étude de molécules biologiques[32, 33, 34]. Le couplage entre DFT électronique et 3D-RISM mentionné précédemment a aussi été utilisé pour étudier l'effet du solvant sur des équilibres de tautomérisation en solution[35]. Récemment des corrections de volume molaire partiel aux calculs 3D-RISM ont été proposé pour la prédiction d'énergies libres de solvation[36, 37, 38].

Deuxième partie .

Théorie

## 2. La théorie de la fonctionnelle de la densité classique

Dans ce chapitre est fait un bref rappel de la théorie de la fonctionnelle de la densité électronique pour souligner les points communs entre cette théorie très connue des chimistes théoriciens avec son analogue classique. On présente ensuite les bases théoriques et certaines des équations fondamentales ainsi que diverses approximations de la théorie de la fonctionnelle de la densité classique. À la fin du chapitre est proposé un tableau récapitulatif des définitions et équations clés pour la lecture de la suite de ce manuscrit.

### 2.1. La théorie de la fonctionnelle de la densité électronique

En chimie quantique, un des objectifs est de connaître la structure électronique et l'énergie d'un édifice moléculaire, molécule, ion, etc. Pour cela, on peut chercher à résoudre directement l'équation de Schrödinger indépendante du temps. Résoudre cette équation revient à trouver la fonction d'onde électronique  $\Psi$  et l'énergie du système  $E$ , solutions de cette équation aux valeurs propres, en utilisant par exemple la méthode de Hartree-Fock[39]. On peut remarquer que la fonction d'onde électronique est généralement un vecteur de haute dimension puisque c'est une fonction qui dépend des coordonnées de chacun des électrons, soit un vecteur à  $3N$  dimensions (si on oublie le spin), où  $N$  est le nombre d'électrons. Du fait de ce nombre important de degrés de liberté, les méthodes de détermination de la fonction d'onde sont numériquement coûteuses.

Trouver la fonction d'onde, qui apparaît explicitement comme une variable, semble être le moyen le plus naturel pour résoudre cette équation. Celle-ci n'est pas directement une observable physique mais un objet mathématique qui contient toutes les informations du système. La densité de probabilité électronique, qui est égale à la norme au carré de la fonction d'onde est, en revanche, une observable physique. Du fait de l'indiscernabilité des électrons, on peut laisser de côté cette probabilité qui dépend de  $3N$  variables. On introduit alors une densité de probabilité de présence des électrons à une position  $\mathbf{r}$ . Celle-ci peut être définie comme

$$n(\mathbf{r}) = \sum_{i=1}^N \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}), \quad (2.1.1)$$

où  $\psi_i^*$  désigne l'adjoint de la fonction d'onde  $\psi_i$ . Bien que cette fonction que l'on nomme densité électronique, soit beaucoup plus simple que la fonction d'onde puisqu'elle ne dépend plus que de 3 variables spatiales, elle contient encore la plupart de l'information physique.

La théorie de la fonctionnelle de la densité électronique (eDFT, une définition du terme fonctionnelle est donnée en Appendice A), introduite par Kohn et Hohenberg[40], et développée par Kohn et Sham[41] au milieu des années 1960, est basée sur deux principaux théorèmes.

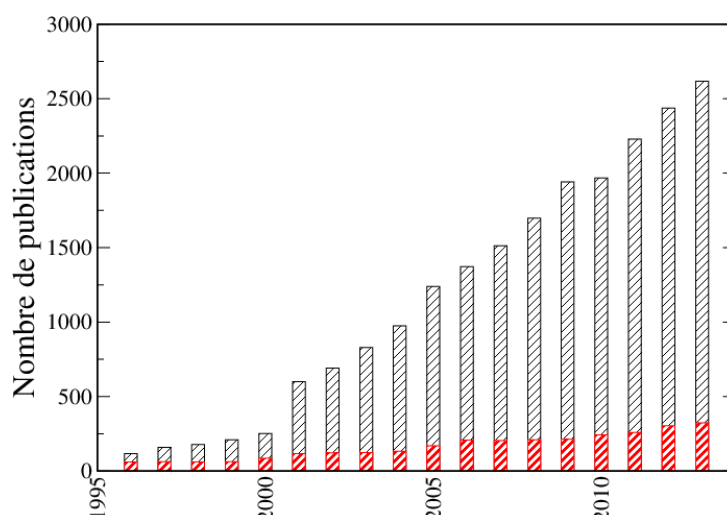
Le premier stipule que *l'énergie d'un système polyélectronique est une fonctionnelle unique de la densité électronique*, introduite dans l'éq. 2.1.1. Ce théorème montre qu'il existe une bijection entre la fonction d'onde décrivant l'état fondamental et la densité électronique de ce même état. Toutes les propriétés de l'état fondamental du système peuvent être connues via la connaissance de la densité électronique au lieu de celle de la fonction d'onde. Ceci est d'une grande importance puisque l'on ramène la résolution d'un problème à  $3N$  variables à celle d'un problème à 3 variables, ce qui explique l'efficacité numérique de la eDFT et sa large utilisation en chimie quantique et en physique du solide. Soulignons tout de même qu'à cause d'approximations techniques qu'il n'est pas pertinent de développer ici, les algorithmes de calculs de eDFT ont en réalité une complexité qui augmente comme  $\mathcal{O}(N)$ [42].

Actuellement, on ne connaît pas d'expression exacte de la fonctionnelle. Le second théorème de Kohn et Hohenberg donne cependant une information importante sur celle-ci, *la densité d'électron qui minimise la fonctionnelle est la densité électronique de l'état fondamental*. Ce théorème stipule que si on connaît une expression exacte de la fonctionnelle, on pourrait, en minimisant celle-ci, trouver la densité électronique de l'état fondamental. De nombreux développements ont été réalisés durant les dernières décennies pour proposer une expression approchée de la fonctionnelle[43], dont la forme exacte reste inconnue malgré la preuve de son existence. La diffusion de programmes de calcul de eDFT a permis de ne pas limiter l'utilisation de cette méthode aux seuls chimistes théoriciens mais de démocratiser son utilisation auprès des expérimentateurs. Le développement d'approximations efficaces de la fonctionnelle permet le calcul de propriétés structurales et énergétiques d'un grand nombre de système à un coût numérique raisonnable. Ceci explique les succès de cette méthode et la place qu'elle occupe aujourd'hui dans la production scientifique et la littérature[44].

Un an après l'article fondateur de Kohn et Hohenberg, Mermin étend la théorie de la fonctionnelle de la densité au cas du gaz inhomogène d'électron à température non-nulle[45]. Ceci revient en fait à réécrire la DFT dans le cadre de la physique statistique. Bien que toujours appliquée au cas d'un hamiltonien quantique dans cet article, ceci ouvre la porte à l'utilisation de la DFT avec un hamiltonien classique.

On peut alors imaginer son utilisation pour l'étude de la solvation dans un fluide décrit par un champ de force classique, ce qui est l'objectif de ces travaux de thèse. Il peut apparaître surprenant que, malgré leurs découvertes quasi-simultanées, la théorie fonctionnelle de la densité classique (cDFT), au-delà de modèles atomiques simples, n'ait pas connu le même succès que son analogue quantique. On peut expliquer ce constat par le fait qu'il existait déjà des méthodes permettant l'étude de problèmes à  $N$  corps classiques avant la formulation de la DFT, comme les simulations Monte-Carlo ou de dynamique moléculaire. Ces méthodes peuvent tout à fait être mises en oeuvre pour étudier des systèmes complexes (de taille importante). Elles ont un coût numérique élevé mais acceptable et sont théoriquement exactes (elles permettent en prin-

cipe, après un échantillonnage complet de l'espace des phases, d'obtenir le résultat prévu par la physique statistique). C'est pourquoi ces méthodes sont largement utilisées dans la communauté. Ceci peut expliquer l'intérêt relativement limité de la communauté au développement de la cDFT en chimie puisque les physicochimistes possèdent déjà des méthodes ayant fait leurs preuves pour étudier des systèmes classiques. En conséquence, pour l'instant, la cDFT n'a pas encore atteint le même niveau de précision dans la prédiction de propriétés physicochimiques d'intérêt pour les biologistes et les chimistes que la eDFT pour l'étude des systèmes quantiques. De plus, contrairement à la eDFT, il n'existe pas de programmes de cDFT largement répandus et accessibles à tous. Néanmoins, certains systèmes de grande taille restent inaccessibles aux simulations numériques, et il se peut que la cDFT, moins coûteuse numériquement, soit un bon moyen d'étude de ces systèmes. Ceci peut expliquer le regain d'intérêt dans la littérature ces dix dernières années pour le développement de la cDFT. Sur la fig. 2.1.1 on donne le nombre de publications par an ayant pour sujet « Classical Density Functional Theory » qui reste néanmoins deux ordre de grandeurs plus faible que celui ayant pour sujet « Density Functional Theory ».



**FIGURE 2.1.1.:** Nombre de publications par an ayant pour sujet, « Classical Density Functional Theory » en rouge et « Density Functional Theory » en noir, selon le site Web of Science. Le nombre de publications ayant pour sujet « Density Functional Theory » a été divisé par 10.

Dans les parties suivantes, nous exposons les principes de la théorie fonctionnelle de la densité classique dans le cas du fluide idéal. Nous présenterons ensuite la construction de la fonctionnelle dans le cas de l'approximation du fluide homogène de référence qui sera utilisée dans toute la suite de ce manuscrit.

## 2.2. La théorie fonctionnelle de la densité classique (cDFT)

### 2.2.1. Exemple introductif, le cas du fluide idéal

Les deux prochaines sous-parties présentent brièvement la DFT classique. Dans le cas du fluide idéal et des ensembles canonique et grand-canonique, on va montrer que l'on peut écrire des fonctionnelles de la densité de solvant qui permettent, par minimisation fonctionnelle, de retrouver les résultats exacts prévus par la thermodynamique statistique.

#### 2.2.1.1. L'ensemble canonique (N,V,T)

On considère ici le cas de l'ensemble canonique, c'est-à-dire un ensemble où le volume  $V$ , la température  $T$  et le nombre de particule  $N$  sont fixés. L'énergie du système est décrite par la donnée de l'hamiltonien décrivant un système à  $N$  particules de masse  $m$ , à la température  $T$ ,

$$\mathcal{H}_N = \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \mathcal{K}(\mathbf{p}^N) + \mathcal{U}(\mathbf{r}^N) + v(\mathbf{r}^N), \quad (2.2.1)$$

avec  $\mathcal{K}$  l'énergie cinétique,  $\mathcal{U}$  l'énergie potentielle d'interaction et  $v$  un potentiel extérieur, qui dépendent des positions et des impulsions  $(\mathbf{r}^N, \mathbf{p}^N) = (\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$  des  $N$  particules. Le potentiel extérieur est une énergie qui a pour origine une perturbation qui ne provient pas des particules du système, par exemple si un champ électrique non nul est imposé, celui ci interagit avec les particules si celles-ci sont chargées.

La densité de probabilité pour ce système à  $N$  particules est <sup>1</sup>

$$f_0 = Q^{-1} \exp [-\beta (\mathcal{H}_N)]. \quad (2.2.3)$$

Dans le cas où le système étudié est le fluide idéal isolé, c'est-à-dire que les particules n'interagissent pas entre elles ( $\mathcal{U} = 0$ ) et que le potentiel extérieur est nul ( $v = 0$ ), on peut calculer analytiquement la fonction de partition, on trouve,

$$Q_{\text{id}} = \left( \frac{V}{\Lambda^3} \right)^N \frac{1}{N!}. \quad (2.2.4)$$

Ce résultat a une interprétation immédiate en analyse combinatoire. En effet, la longueur d'onde de de Broglie  $\Lambda = \left( \frac{2\pi m k_B T}{h^2} \right)^{3/2}$ , avec  $h = 6.62606957 \cdot 10^{-34} \text{ m}^2 \cdot \text{kg} \cdot \text{s}^{-1}$  la constante de Planck, représente l'étalement spatial de la particule.  $\Lambda^3$  est donc le plus petit volume élémentaire que l'on puisse considérer dans l'espace des phases (en raison du principe d'incertitude d'Heisenberg).

---

1. On rappelle ici l'expression de la fonction de partition[7] :

$$Q = \frac{1}{h^{3N} N!} \int \cdots \int_{\mathbb{R}^{3N}} \int \cdots \int_{\mathbb{R}^{3N}} e^{-\beta \mathcal{H}_N} d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N. \quad (2.2.2)$$

On peut alors « découper » le volume  $V$  en  $V/\Lambda^3$  éléments et, comme les particules n'interagissent pas, il y a pour chaque particule  $V/\Lambda^3$  possibilités pour placer la particule. Le facteur  $1/N!$  provient de l'indiscernabilité des particules.

Si on réinjecte l'expression de la fonction de partition dans l'éq. 1.2.5 et que l'on utilise la formule de Stirling<sup>2</sup>, on obtient pour l'énergie libre de Helmholtz,

$$F_{\text{id}} = k_B T [N \ln (\rho \Lambda^3) - N], \quad (2.2.5)$$

où on a introduit  $\rho = \frac{N}{V}$ , la densité du fluide homogène.

On pose une fonctionnelle de la densité inhomogène,

$$\mathcal{F}_{\text{id}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) [\ln (\rho(\mathbf{r}) \Lambda^3) - 1] d\mathbf{r}. \quad (2.2.6)$$

En dérivant analytiquement cette fonctionnelle (voir Appendice A pour la définition d'une dérivée fonctionnelle) et en cherchant les racines de la dérivée obtenue, on trouve que la fonctionnelle est minimale pour une densité d'équilibre homogène  $\rho_{\text{eq}} = N/V$ .

On a montré ici que pour le fluide idéal isolé, étudié dans l'ensemble canonique, il est possible d'introduire une fonctionnelle de la densité de ce fluide idéal. Cette fonctionnelle est égale à l'énergie libre de Helmholtz du système homogène lorsqu'elle est minimale. De plus, ce minimum est atteint pour une densité du fluide qui est celle du fluide homogène. En minimisant cette fonctionnelle on retrouve donc les propriétés énergétiques et structurales du système à l'équilibre thermodynamique.

### 2.2.1.2. L'ensemble grand-canonique ( $\mu, V, T$ )

On se place maintenant dans l'ensemble grand-canonique, c'est-à-dire un ensemble où le volume  $V$ , la température  $T$  et le potentiel chimique  $\mu$  sont fixés.

Dans ce cas, la densité de probabilité d'un système à  $N$  particules à l'équilibre thermodynamique s'exprime,

$$f_{\text{eq}} = \Xi^{-1} \exp [-\beta (\mathcal{H}_N - \mu N)]. \quad (2.2.7)$$

La fonction de grand-partition,  $\Xi$ , est donnée par,

$$\Xi = \text{Tr}_{\text{cl}} \exp [-\beta (\mathcal{H}_N - \mu N)], \quad (2.2.8)$$

où  $\text{Tr}_{\text{cl}}$  désigne la trace classique, c'est-à-dire

$$\text{Tr}_{\text{cl}} = \sum_{N=0}^{\infty} \frac{1}{h^{3N} N!} \int \cdots \int_{\mathbb{R}^{3N}} \int \cdots \int_{\mathbb{R}^{3N}} d\mathbf{r} \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N, \quad (2.2.9)$$

---

2.  $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$



Dans le cas du fluide idéal en l'absence de potentiel extérieur, on a,

$$\Xi_{\text{id}} = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N! \Lambda^{3N}} \int \cdots \int_{\mathbb{R}^{3N}} d\mathbf{r}^N = e^{zV}, \quad (2.2.10)$$

en posant  $z = \exp(\beta\mu)/\Lambda^3$  l'activité.

L'expression du grand potentiel dans cet ensemble est

$$\Omega = -k_B T \ln \Xi, \quad (2.2.11)$$

ce qui dans le cas du gaz idéal donne,

$$\Omega_{\text{id}} = -k_B T z V. \quad (2.2.12)$$

Notons que le grand potentiel est également lié à l'énergie libre de Helmholtz par la transformée de Legendre,

$$\Omega = F - \mu N \quad (2.2.13)$$

On peut introduire, comme en sec. 2.2.1.1 et dans l'éq. 2.2.6, une fonctionnelle qui dépend de la densité non-homogène,

$$\Omega[\rho(\mathbf{r})] = \mathcal{F}_{\text{id}}[\rho(\mathbf{r})] - \mu \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) d\mathbf{r}, \quad (2.2.14)$$

En minimisant cette fonctionnelle, c'est-à-dire en cherchant les racines de la dérivée fonctionnelle de l'éq. 2.2.14, on trouve  $\rho_{\text{eq}} = \exp(\beta\mu)/\Lambda^3$ . En réinjectant cette expression de la densité à l'équilibre, on retrouve de façon cohérente l'expression du grand potentiel pour le fluide idéal donnée en éq. 2.2.12.

On retrouve donc ici un résultat similaire à celui obtenu dans l'ensemble canonique pour le fluide idéal.

On peut donc généraliser les principes de la théorie de la fonctionnelle de la densité classique dans l'ensemble grand-canonique au cas du fluide quelconque. La suite de notre étude sera menée exclusivement dans l'ensemble grand-canonique.

- Dans l'ensemble grand-canonique, pour un fluide donné, éventuellement soumis à un potentiel extérieur, on peut écrire une fonctionnelle unique de la densité de solvant.
- Cette fonctionnelle est minimale pour la densité de solvant correspondant à l'équilibre thermodynamique, elle est alors égale au grand-potentiel du système.

Ce résultat est intéressant puisqu'il permet, si on connaît une expression de la fonctionnelle, de déterminer par principe variationnel la densité à l'équilibre thermodynamique et le grand-potentiel, c'est-à-dire les propriétés énergétiques et structurales d'équilibres du système. Réaliser une minimisation fonctionnelle pourrait s'avérer beaucoup moins coûteux numériquement que de

réaliser l'échantillonnage de l'espace de phases. Dans la partie suivante on présente la démonstration générale de ces deux points, qui a été proposée par Evans[46].

### 2.2.2. Principe variationnel pour le grand potentiel

On a introduit la probabilité d'équilibre en éq. 2.2.7. Supposons qu'il existe une autre distribution de probabilité  $f$  avec la seule condition que celle-ci soit normée, c'est-à-dire  $\text{Tr}_{\text{cl}} f = 1$ . On définit alors une fonctionnelle de cette probabilité comme,

$$\Omega[f] = \text{Tr}_{\text{cl}} \left[ f \left( (\mathcal{H}_N - \mu N) + k_B T \ln f \right) \right], \quad (2.2.15)$$

avec  $\mathcal{H}_N$  le hamiltonien défini en éq. 2.2.1. Il est évident que pour  $f = f_{\text{eq}}$ , on a :

$$\Omega[f_{\text{eq}}] = \Omega = -k_B T \ln \Xi. \quad (2.2.16)$$

Supposons que  $f \neq f_{\text{eq}}$ , comme ces deux distributions de probabilité sont normalisées, on a :

$$\text{Tr}_{\text{cl}} f = \text{Tr}_{\text{cl}} f_{\text{eq}}. \quad (2.2.17)$$

On peut multiplier chaque membre de cette égalité par  $\Omega$  qui est un scalaire,

$$\text{Tr}_{\text{cl}} (\Omega f) = \text{Tr}_{\text{cl}} (\Omega f_{\text{eq}}). \quad (2.2.18)$$

Soit, en utilisant l'éq. 2.2.7 et l'éq. 2.2.16,

$$\text{Tr}_{\text{cl}} [f (v + \mathcal{K} + \mathcal{U} - \mu N + k_B T \ln f_{\text{eq}})] = \text{Tr}_{\text{cl}} [f_{\text{eq}} (v + K + U - \mu N + k_B T \ln f_{\text{eq}})], \quad (2.2.19)$$

et donc avec l'éq. 2.2.15,

$$\Omega[f] = \Omega[f_{\text{eq}}] + k_B T \text{Tr}_{\text{cl}} \left( f \ln \frac{f}{f_{\text{eq}}} \right) \quad (2.2.20)$$

On peut ajouter au membre de droite  $\text{Tr}_{\text{cl}} f_0$  et retirer  $\text{Tr}_{\text{cl}} f$  en conservant l'égalité grâce à l'éq. 2.2.17. L'éq. 2.2.20 devient

$$\Omega[f] = \Omega[f_{\text{eq}}] + k_B T \text{Tr}_{\text{cl}} \left[ f_{\text{eq}} \left( \frac{f}{f_{\text{eq}}} \ln \left( \frac{f}{f_{\text{eq}}} \right) - \frac{f}{f_{\text{eq}}} + 1 \right) \right]. \quad (2.2.21)$$

Il suffit de remarquer que la fonction entre parenthèses,  $x \mapsto x \ln x - x + 1$ , définie sur  $\mathbb{R}_+$  est positive et ne s'annule qu'en  $x = 1$  pour montrer que la fonctionnelle  $\Omega[f]$  est toujours strictement supérieure à  $\Omega$ , sauf en  $f = f_{\text{eq}}$  où elle vaut  $\Omega$ .

On vient de démontrer un théorème clé de la DFT classique : la fonctionnelle  $\Omega[f]$  introduite est égale au grand potentiel  $\Omega$  à son minimum qui est atteint pour la distribution de probabilité

d'équilibre  $f_{\text{eq}}$ .

On peut montrer que cette distribution de probabilité d'équilibre  $f_{\text{eq}}$  est une fonctionnelle unique de la densité d'équilibre  $\rho_{\text{eq}}(\mathbf{r})$ . Pour cela, il suffit de remarquer que la densité de probabilité d'équilibre est, de manière évidente, une fonctionnelle du champ extérieur  $v$ . La densité d'équilibre, en tant que fonctionnelle de la densité de probabilité d'équilibre est donc une fonctionnelle du champ extérieur  $v$ . On peut démontrer (ceci est un cas simplifié du calcul présenté en sec. B.2) que la densité de particule à l'équilibre est associée de manière unique à un potentiel extérieur. En conséquence, la distribution de probabilité d'équilibre est une fonctionnelle unique de la densité d'équilibre. De ce fait, le potentiel extérieur est une fonctionnelle unique de la densité d'équilibre.

On introduit alors la fonctionnelle intrinsèque

$$\mathcal{F}[\rho(\mathbf{r})] = \text{Tr}_{\text{cl}} \left[ f [(\mathcal{K} + \mathcal{U}) + k_B T \ln f] \right] \quad (2.2.22)$$

qui est elle aussi une unique fonctionnelle de la densité, même si on n'en connaît pas de forme qui dépende explicitement de la densité. On peut ainsi définir une nouvelle fonctionnelle :

$$\Omega_v [\rho(\mathbf{r})] = \mathcal{F}[\rho(\mathbf{r})] + \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} - \mu \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) d\mathbf{r} \quad (2.2.23)$$

$$= \mathcal{F}[\rho(\mathbf{r})] - \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} \quad (2.2.24)$$

où l'on a introduit  $\psi(\mathbf{r}) = \mu - v(\mathbf{r})$  le potentiel chimique intrinsèque.

Il est évident que pour une densité  $\rho(\mathbf{r})$  correspondant à une distribution de probabilité  $f$ , on a l'égalité suivante :

$$\Omega[f] = \Omega_v [\rho(\mathbf{r})] . \quad (2.2.25)$$

En se servant des propriétés démontrées pour  $\Omega[f]$ , on prouve directement que  $\Omega_v [\rho(\mathbf{r})]$  est égal au grand potentiel  $\Omega$  à son minimum, atteint pour la densité d'équilibre à une particule  $\rho_{\text{eq}}(\mathbf{r})$ . Cependant, nous ne connaissons pas d'expression de la fonctionnelle intrinsèque  $\mathcal{F}$ . Remarquons néanmoins que pour un fluide sans interactions (i.e.,  $\mathcal{U} = 0$ ) cette fonctionnelle est égale à la fonctionnelle  $\mathcal{F}_{\text{id}}$  introduite en sec. 2.2.1.1. Dès lors, on peut réécrire cette fonctionnelle intrinsèque comme la somme d'une partie idéale, qui décrit l'entropie d'un fluide sans interaction à la même densité, et d'un terme d'excès, défini comme la différence entre la partie intrinsèque et la partie idéale.

$$\mathcal{F}[\rho(\mathbf{r})] = \mathcal{F}_{\text{id}}[\rho(\mathbf{r})] + \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] . \quad (2.2.26)$$

En dérivant l'éq. 2.2.24 par rapport à  $\rho(\mathbf{r})$  on trouve immédiatement l'égalité suivante qui lie dérivée de la fonctionnelle intrinsèque et potentiel chimique intrinsèque

$$\frac{\delta \mathcal{F}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = \psi(\mathbf{r}) . \quad (2.2.27)$$

On a également, en dérivant l'éq. 2.2.24 par rapport au potentiel chimique intrinsèque :

$$\frac{\delta \Omega_v [\rho(\mathbf{r})]}{\delta \psi(\mathbf{r})} = -\rho(\mathbf{r}). \quad (2.2.28)$$

Ceci permet de voir la fonctionnelle  $\Omega_v$  comme obtenue par une transformée de Legendre généralisée, à partir de la fonctionnelle intrinsèque. Les variables conjuguées  $\psi(\mathbf{r})$  et  $\rho(\mathbf{r})$  étant les analogues de  $\mu$  et  $N$  dans le cas homogène.

## 2.3. Fonctions de corrélation directe et écriture de la fonctionnelle d'excès

La partie d'excès de la fonctionnelle est génératrice d'une hiérarchie de fonctions de corrélation directe  $c^{(i)}$ . La fonction à une particule est définie comme la première dérivée fonctionnelle du terme d'excès par rapport à la densité,

$$c^{(1)}(\rho; \mathbf{r}) = -\frac{\delta \beta \mathcal{F}_{\text{exc}} [\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})}, \quad (2.3.1)$$

remarquons la dépendance à la densité de cette fonction.

En dérivant l'éq. 2.2.26 et en y injectant l'éq. 2.3.1, on peut obtenir une expression de la densité d'équilibre,

$$\rho_{\text{eq}}(\mathbf{r}) = \frac{e^{\beta \mu}}{\Lambda^3} \exp \left[ -\beta v(\mathbf{r}) + c^{(1)}(\mathbf{r}) \right] \quad (2.3.2)$$

En comparant ce résultat avec celui de la densité d'équilibre dans le cas du gaz idéal isolé obtenu en sec. 2.2.1.2, on constate l'apparition d'un terme dû au potentiel extérieur. De plus, l'effet des interactions entre particules est décrite par cette fonction de corrélation à une particule.

La fonction de corrélation directe de paire est définie comme la dérivée de la fonction à une particule,

$$c^{(2)}(\rho; \mathbf{r}, \mathbf{r}') = \frac{\delta c^{(1)}(\mathbf{r})}{\delta \rho(\mathbf{r}')} = -\frac{\delta^2 \beta \mathcal{F}_{\text{exc}} [\rho(\mathbf{r})]}{\delta \rho(\mathbf{r}') \delta \rho(\mathbf{r})}. \quad (2.3.3)$$

C'est cette fonction de distribution de paire que l'on retrouve dans une des équations clés de la théorie des liquides, l'équation de Ornstein-Zernike[4].

Les fonctions de corrélation d'ordre plus élevé sont, de la même manière, générées par des dérivées fonctionnelles successives de la partie d'excès.

Pour tenter de trouver une expression pour la partie d'excès, on peut intégrer l'éq. 2.3.1 par rapport à la densité. Si  $\rho_0(\mathbf{r})$  est la densité dans un état de référence du système (par exemple le système homogène ou le fluide de sphères dures), et  $c^{(1)}(\rho_0; \mathbf{r}) = c_0^{(1)}(\mathbf{r})$  la fonction de corrélation à une particule dans ce même état de référence, on peut choisir un chemin d'intégration linéaire

entre l'état de référence et l'état final du système de densité  $\rho(\mathbf{r})$  de sorte que

$$\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \lambda \Delta \rho(\mathbf{r}), \quad (2.3.4)$$

avec  $\Delta \rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0(\mathbf{r})$  et  $\lambda$  un paramètre d'intégration variant entre 0 et 1. On obtient par intégration l'expression suivante pour la partie d'excès,

$$\begin{aligned} \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] - \mathcal{F}_{\text{exc}}[\rho_0(\mathbf{r})] &= k_B T \int_0^1 \left[ \iiint_{\mathbb{R}^3} \frac{\partial \rho(\mathbf{r}; \lambda)}{\partial \lambda} c^{(1)}(\rho_\lambda; \mathbf{r}) d\mathbf{r} \right] d\lambda \\ &= k_B T \int_0^1 \left[ \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}) c^{(1)}(\rho_\lambda; \mathbf{r}) d\mathbf{r} \right] d\lambda. \end{aligned} \quad (2.3.5)$$

On peut évaluer la fonction de corrélation  $c^{(1)}(\rho_\lambda; \mathbf{r})$  par une intégration similaire à celle qui vient d'être utilisée, en partant de l'éq. 2.3.2,

$$c^{(1)}(\rho_\lambda; \mathbf{r}) - c^{(1)}(0; \mathbf{r}) = \int_0^\lambda \left[ \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}) c^{(2)}(\rho_{\lambda'}; \mathbf{r}, \mathbf{r}') d\mathbf{r}' \right] d\lambda'. \quad (2.3.6)$$

Finalement, en réinjectant l'éq. 2.3.6 dans l'éq. 2.3.5 on obtient,<sup>3</sup>

$$\begin{aligned} \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] - \mathcal{F}_{\text{exc}}[\rho_0(\mathbf{r})] &= -k_B T \int \Delta \rho(\mathbf{r}) c_0^{(1)}(\mathbf{r}) d\mathbf{r} \\ &- k_B T \int_0^1 (1 - \lambda) \left[ \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}') \Delta \rho(\mathbf{r}) c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}' \right] d\lambda. \end{aligned} \quad (2.3.7)$$

On signale que si le chemin d'intégration utilisé a été choisi pour sa simplicité, le résultat obtenu est indépendant de ce chemin puisque  $\mathcal{F}_{\text{exc}}$  est une fonctionnelle unique de la densité. On peut alors exprimer le grand potentiel à l'aide des éq. 2.3.7, éq. 2.2.24 et éq. 2.2.6,

$$\begin{aligned} \Omega_v[\rho(\mathbf{r})] &= k_B T \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) [\ln(\rho(\mathbf{r}) \Lambda^3) - 1] d\mathbf{r} - k_B T \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}) c_0^{(1)}(\mathbf{r}) d\mathbf{r} \\ &+ k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}') \Delta \rho(\mathbf{r}) C^{(2)}(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ &- \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) \psi(\mathbf{r}) d\mathbf{r} + \mathcal{F}_{\text{exc}}[\rho_0(\mathbf{r})], \end{aligned} \quad (2.3.8)$$

avec,

$$C^{(2)}(\mathbf{r}, \mathbf{r}') = \int_0^1 (1 - \lambda) c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}') d\lambda. \quad (2.3.9)$$

Si on prend comme référence un fluide homogène qui possède le même potentiel chimique que le

---

3. En utilisant le fait que  $\int_0^1 \left[ \int_0^\lambda f(\lambda') d\lambda' \right] d\lambda = \int_0^1 (1 - \lambda) f(\lambda) d\lambda$  [4]

fluide réel de densité  $\rho_b$ , son grand potentiel peut se trouver en posant  $\rho(\mathbf{r}) = \rho_b$  dans l'éq. 2.3.8,

$$\Omega_b = k_B T \iiint_{\mathbb{R}^3} \rho_b(\mathbf{r}) [\ln(\rho_b \Lambda^3) - 1] d\mathbf{r} + \mathcal{F}_{\text{exc}}[\rho_b(\mathbf{r})] - \mu \iiint_{\mathbb{R}^3} \rho_b(\mathbf{r}) d\mathbf{r}. \quad (2.3.10)$$

En remarquant que dans le cas du fluide homogène la dérivée de l'éq. 2.2.26 donne

$$\frac{\delta \mathcal{F}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = \frac{\delta \mathcal{F}_{\text{id}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} + \frac{\delta \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = k_B T \ln(\rho_b \Lambda^3) - k_B T c_b(\mathbf{r}) = \mu, \quad (2.3.11)$$

on peut reformuler de manière plus compacte l'éq. 2.3.8 en

$$\begin{aligned} \Omega_v[\rho(\mathbf{r})] &= k_B T \iiint_{\mathbb{R}^3} \left[ \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \Delta \rho(\mathbf{r}) \right] d\mathbf{r} + k_B T \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \\ &+ k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}') \Delta \rho(\mathbf{r}) C^{(2)}(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}' + \Omega_b \end{aligned} \quad (2.3.12)$$

$$= \Omega_b + \mathcal{F}_{\text{id}}[\rho(\mathbf{r})] + \mathcal{F}_{\text{ext}}[\rho(\mathbf{r})] + \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] \quad (2.3.13)$$

avec

$$\mathcal{F}_{\text{id}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \left[ \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \Delta \rho(\mathbf{r}) \right] d\mathbf{r} \quad (2.3.14)$$

$$\mathcal{F}_{\text{ext}}[\rho(\mathbf{r})] = \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \quad (2.3.15)$$

$$\mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta \rho(\mathbf{r}') \Delta \rho(\mathbf{r}) C^{(2)}(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (2.3.16)$$

où  $\Delta \rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_b$ .

### 2.3.1. Approximation du fluide homogène de référence (HRF)

La partie d'excès décrite par l'éq. 2.3.16 est en fait difficilement évaluable telle quelle car elle fait intervenir la fonction  $C^{(2)}(\mathbf{r}, \mathbf{r}')$  qui, pour être calculée, nécessite de connaître la fonction de corrélation à deux corps à toutes les densités entre la densité du fluide de référence et la densité du système (pour toutes les valeurs de  $\lambda$ ). Pour simplifier considérablement le calcul de cette fonction on peut considérer qu'elle est égale à celle du fluide homogène de référence pour toutes les densités, soit

$$c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}') = c^{(2)}(\rho_0; \mathbf{r}, \mathbf{r}') = c^{(2)}(\mathbf{r}, \mathbf{r}') \quad (2.3.17)$$

En utilisant l'éq. 2.3.9 on obtient,

$$C^{(2)}(\mathbf{r}, \mathbf{r}') = -\frac{1}{2} c^{(2)}(\mathbf{r}, \mathbf{r}'). \quad (2.3.18)$$

Cette approximation est raisonnable si la densité reste proche de celle du fluide homogène, ce qui peut ne pas être le cas.

On peut montrer que cette approximation est en fait équivalente à l'approximation HNC (Hypernetted chain) employée dans l'étude des fluides homogènes[4].

Il y a une autre façon de voir cette formulation comme équivalente à HNC.

En repartant de l'éq. 2.3.7, on peut écrire une expression exacte de la fonction  $c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}')$  comme une série des fonctions de corrélation directe du fluide de référence d'ordres supérieurs.

$$\begin{aligned}
 c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}') &= c^{(2)}(\rho_0; \mathbf{r}, \mathbf{r}') + \lambda \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}_3) \frac{\delta c^{(2)}(\rho_0; \mathbf{r}, \mathbf{r}')}{\delta\rho(\mathbf{r}_3)} d\mathbf{r}_3 \\
 &\quad + \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \frac{\delta^n c^{(2)}(\rho_0; \mathbf{r}, \mathbf{r}')}{\delta\rho(\mathbf{r}_3) \cdots \delta\rho(\mathbf{r}_{n+2})} \Delta\rho(\mathbf{r}_3) \cdots \Delta\rho(\mathbf{r}_{n+2}) d\mathbf{r}_3 \cdots d\mathbf{r}_{n+2} \\
 c^{(2)}(\rho_\lambda; \mathbf{r}, \mathbf{r}') &= c^{(2)}(\rho_0; \mathbf{r}, \mathbf{r}') + \lambda \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}_3) c^{(3)}(\rho_0; \mathbf{r}, \mathbf{r}', \mathbf{r}_3) d\mathbf{r}_3 \\
 &\quad + \sum_{n=2}^{\infty} \frac{\lambda^n}{n!} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \Delta\rho(\mathbf{r}_3) \cdots \Delta\rho(\mathbf{r}_{n+2}) c^{(n)}(\rho_0; \mathbf{r}, \mathbf{r}', \mathbf{r}_3, \cdots, \mathbf{r}_{n+2}) d\mathbf{r}_3 \cdots d\mathbf{r}_{n+2}
 \end{aligned}$$

On voit donc que la fonction de corrélation de paire exacte du fluide étudié dépend de toutes les fonctions de corrélation directe d'ordre  $n$  du fluide de référence. L'approximation HNC consiste à se limiter aux corrélations d'ordre 2 dans l'équation de Ornstein-Zernicke et donc à tronquer cette équation à l'ordre 2. On retrouve alors l'approximation de l'éq. 2.3.18.

En clair, l'approximation réalisée revient à négliger toutes les corrélations d'ordre strictement supérieur à deux, c'est donc une approximation importante.

Les équations clés de ce chapitre sur la théorie fonctionnelle de la densité classique dans l'approximation du fluide de référence sont récapitulées dans l'encart suivant

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_b$$

$$\mathcal{F}[\rho(\mathbf{r})] = \Omega_v[\rho(\mathbf{r})] - \Omega_b = \mathcal{F}_{\text{id}}[\rho(\mathbf{r})] + \mathcal{F}_{\text{ext}}[\rho(\mathbf{r})] + \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})]$$

$$\mathcal{F}_{\text{id}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \left[ \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \Delta\rho(\mathbf{r}) \right] d\mathbf{r}$$

$$\mathcal{F}_{\text{ext}}[\rho(\mathbf{r})] = \iiint_{\mathbb{R}^3} \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r}$$

$$\mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}') \Delta\rho(\mathbf{r}) C^{(2)}(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'$$



### 3. Une théorie de la fonctionnelle de la densité moléculaire pour l'eau

L'écriture de la fonctionnelle est adaptée à l'étude de la solvation. En effet, dans cette écriture est incluse la réponse du solvant à une perturbation décrite par un potentiel extérieur  $v_{\text{ext}}$ . Ce potentiel, dimensionné comme une densité d'énergie, peut avoir plusieurs origines physiques : un champ électrique, une contrainte mécanique, etc. Dans le cas de cette thèse, le potentiel extérieur est une perturbation créée par un soluté.

#### 3.1. Modèle d'eau et de soluté

##### 3.1.1. Modèle d'eau rigide

Le modèle d'eau utilisé est le modèle SPC/E (Single Point Charge/Extended). Ce modèle d'eau est rigide et se compose d'un site Lennard-Jones sur l'oxygène et de trois charges partielles placées sur l'oxygène et les deux hydrogènes. Les paramètres Lennard Jones et les charges partielles sont indiquées dans le tab. 3.1

site	$\epsilon$ (kJ.mol <sup>-1</sup> )	$\sigma$ (Å)	$q$ (e)
O	0.65	3.165	-0.8476
H	0.0	0.0	0.4238

TABLE 3.1.: Paramètres du modèle d'eau SPC/E

La géométrie du modèle est donnée en fig. 3.1.1.



FIGURE 3.1.1.: Géométrie du modèle d'eau SPC/E, l'oxygène est en rouge, les hydrogènes en blanc.

Avec cette géométrie et ces paramètres, le moment dipolaire d'une molécule d'eau est de  $\mu = 2.35$  D, orienté conventionnellement des charges négatives (l'oxygène) vers les charges positives (le centre

de masse des hydrogènes). Ce modèle a été choisi car c'est un modèle rigide, il ne possède donc pas de degré de liberté interne qui compliquerait beaucoup l'écriture de la fonctionnelle. De plus, il est très utilisé en simulation moléculaire car il reproduit correctement les fonctions de distribution radiale, la densité et le coefficient de diffusion de l'eau liquide obtenus expérimentalement[47] tout en étant moins coûteux numériquement que des modèles plus évolués. Du fait de son abondante utilisation on pourra comparer les résultats de la cDFT à ceux obtenus par MD ou MC dans la littérature. Je n'ai réalisé aucun calcul de dynamique moléculaire ou de Monte Carlo durant ma thèse.

### 3.1.2. Modèles de solutés rigides

Les solutés sont eux aussi décrits par des modèles moléculaires rigides, composés de différents sites Lennard-Jones et de charges partielles ou entières. Le potentiel extérieur  $v_{\text{ext}}$  agissant sur une molécule de solvant à la position  $\mathbf{r}$  (on repère en fait la position du seul site Lennard-Jones de notre modèle d'eau, l'oxygène) avec une orientation  $\mathbf{\Omega}$  est donc décrit par :

- un potentiel Lennard-Jones entre le solvant et tous les sites du solutés ;
- un potentiel électrostatique provenant du champ électrique créé par le soluté.

L'expression de ces potentiels extérieurs et leurs calculs sont décrits dans le chapitre 6.

## 3.2. Un traitement strictement dipolaire de la polarisation

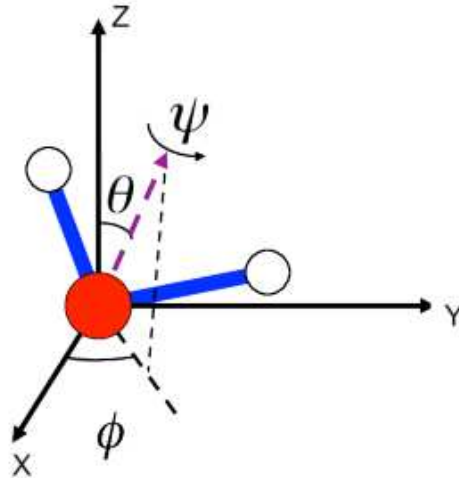
### 3.2.1. Écriture de la fonctionnelle

Pour un solvant moléculaire, contrairement aux cas présentés dans les chapitres précédents, le potentiel extérieur  $v_{\text{ext}}$  ne dépend plus uniquement des coordonnées spatiales  $\mathbf{r}$ , mais également des orientations des molécules, définies dans le repère fixe du laboratoire par les trois angles d'Euler  $\mathbf{\Omega} = (\theta, \phi, \psi)$ . Ces angles sont définis en fig. 3.2.1.

On minimise donc une fonctionnelle de la densité de solvant  $\rho(\mathbf{r}, \mathbf{\Omega})$  qui dépend également des orientations. Cependant, toutes les expressions démontrées dans le chapitre 2, sont facilement généralisables dans le cas de cette densité.

Avant mon arrivée au laboratoire, le groupe avait déjà proposé une fonctionnelle développée pour l'étude de la solvation en milieu aqueux dans l'approximation du fluide homogène de référence[48, 49, 50, 51, 52, 53]. Celle-ci consiste à partir d'une décomposition de la fonctionnelle en trois termes

$$\mathcal{F}[\rho(\mathbf{r}, \mathbf{\Omega})] = \mathcal{F}_{\text{id}}[\rho(\mathbf{r}, \mathbf{\Omega})] + \mathcal{F}_{\text{ext}}[\rho(\mathbf{r}, \mathbf{\Omega})] + \mathcal{F}_{\text{exc}}[\rho(\mathbf{r}, \mathbf{\Omega})]. \quad (3.2.1)$$



**FIGURE 3.2.1.:** Représentation des angles d'Euler décrivant l'orientation d'une molécule d'eau dans le repère fixe du laboratoire.

Les deux premiers ont pour expression,

$$\mathcal{F}_{\text{id}}[\rho(\mathbf{r}, \boldsymbol{\Omega})] = k_B T \iiint_{\mathbb{R}^3} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \left[ \rho(\mathbf{r}, \boldsymbol{\Omega}) \ln \left( \frac{\rho(\mathbf{r}, \boldsymbol{\Omega})}{\rho_b} \right) - \rho(\mathbf{r}, \boldsymbol{\Omega}) + \rho_b \right] d\mathbf{r} d\boldsymbol{\Omega} \quad (3.2.2)$$

et

$$\mathcal{F}_{\text{ext}}[\rho(\mathbf{r}, \boldsymbol{\Omega})] = \iiint_{\mathbb{R}^3} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \rho(\mathbf{r}, \boldsymbol{\Omega}) v_{\text{ext}}(\mathbf{r}, \boldsymbol{\Omega}) d\mathbf{r} d\boldsymbol{\Omega} \quad (3.2.3)$$

Il s'agit simplement des réécritures des éq. 2.3.14 et éq. 2.3.16 avec une densité qui dépend aussi des angles. Notons que dans cette formule,  $\rho_b$  n'est pas la densité en nombre du fluide homogène (que l'on notera à partir de maintenant  $n_b$ ), mais la densité en nombre par unité d'angle solide,  $\rho_b = n_b/(8\pi^2)$ .

Dans le cadre du développement par rapport au fluide homogène de référence, la partie d'excès peut s'écrire à partir de la fonction de corrélation directe du solvant homogène.

$$\mathcal{F}_{\text{exc}}[\rho(\mathbf{r}, \boldsymbol{\Omega})] = -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \iiint_{8\pi^2} \iiint_{8\pi^2} \Delta\rho(\mathbf{r}_1, \boldsymbol{\Omega}_1) c^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) \Delta\rho(\mathbf{r}_2, \boldsymbol{\Omega}_2) d\mathbf{r}_1 d\boldsymbol{\Omega}_1 d\mathbf{r}_2 d\boldsymbol{\Omega}_2, \quad (3.2.4)$$

où pour une écriture plus compacte on a défini :  $\iiint_{8\pi^2} \equiv \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi}$ . L'expression donnée en éq. 3.2.4 n'est pas utilisée directement pour deux raisons. D'une part on ne connaît pas facilement la fonction de corrélation  $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$ . D'autre part cette fonctionnelle s'avère numériquement coûteuse à calculer car elle requiert une intégration double sur le volume et la sphère unité. Pour ces deux raisons, il n'est pas immédiatement envisagé de calculer le terme d'excès en utilisant cette approche.

On peut réaliser un développement en invariants rotationnels de la fonction de corrélation,

$$c^{(2)}(\mathbf{r}_{12}, \mathbf{\Omega}_1, \mathbf{\Omega}_2) = \sum_{m,n,l,\mu,\nu} c_{\mu\nu}^{mnl}(r) \Phi_{\mu\nu}^{mnl}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}), \quad (3.2.5)$$

où  $\Phi_{\mu\nu}^{mnl}$  sont les invariants rotationnels déjà introduits au chapitre 1. Le développement en invariants rotationnels a été formulé pour séparer les parties radiale et angulaire [25, 24].

Dans l'hypothèse d'un solvant strictement dipolaire, c'est-à-dire où l'on considère que la polarisation microscopique d'une molécule d'eau peut être représentée par un dipôle orienté selon l'axe  $C_2$  de la molécule, on peut, avec une bonne approximation [24, 54, 25], se limiter au développement à l'ordre 2 avec  $\mu = \nu = 0$ . On ne considère alors que les invariants suivants :

$$\begin{aligned} \Phi^{000}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}) &= 1 \\ \Phi^{101}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}) &= -\mathbf{\Omega}_1 \cdot \tilde{\mathbf{u}}_{12} \\ \Phi^{011}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}) &= \mathbf{\Omega}_2 \cdot \tilde{\mathbf{u}}_{12} \\ \Phi^{110}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}) &= \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2 \\ \Phi^{112}(\mathbf{\Omega}_1, \mathbf{\Omega}_2, \mathbf{r}_{12}) &= 3(\mathbf{\Omega}_1 \cdot \tilde{\mathbf{u}}_{12})(\mathbf{\Omega}_2 \cdot \tilde{\mathbf{u}}_{12}) - \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2, \end{aligned}$$

avec  $\tilde{\mathbf{u}}_{12} = (\mathbf{r}_2 - \mathbf{r}_1) / \|\mathbf{r}_2 - \mathbf{r}_1\|$  le vecteur unitaire reliant les positions des deux molécules. Cela donne, si on réinjecte ce développement dans la partie d'excès :

$$\begin{aligned} \mathcal{F}_{\text{exc}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] &= \\ &- \frac{1}{2} k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} [c_{000}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_1) \Delta n(\mathbf{r}_2)] d\mathbf{r}_1 d\mathbf{r}_2 \\ &- \frac{1}{2\mu} k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} [c_{101}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_1) \mathbf{P}(\mathbf{r}_2) \cdot \tilde{\mathbf{u}}_{12}] d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ \frac{1}{2\mu} k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} [c_{011}(\|\mathbf{r}_{12}\|) \mathbf{P}(\mathbf{r}_1) \Delta n(\mathbf{r}_2) \cdot \tilde{\mathbf{u}}_{12}] d\mathbf{r}_1 d\mathbf{r}_2 \\ &- \frac{1}{2\mu^2} k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} [c_{110}(\|\mathbf{r}_{12}\|) \mathbf{P}(\mathbf{r}_1) \cdot \mathbf{P}(\mathbf{r}_2)] d\mathbf{r}_1 d\mathbf{r}_2 \\ &- \frac{1}{2\mu^2} k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{112}(\|\mathbf{r}_{12}\|) [3(\mathbf{P}(\mathbf{r}_1) \cdot \tilde{\mathbf{u}}_{12})(\mathbf{P}(\mathbf{r}_2) \cdot \tilde{\mathbf{u}}_{12}) - \mathbf{P}(\mathbf{r}_1) \cdot \mathbf{P}(\mathbf{r}_2)] d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ \mathcal{F}_{\text{cor}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] \end{aligned} \quad (3.2.6)$$

où  $n(\mathbf{r})$  est la densité de solvant à une particule qui ne dépend plus que des coordonnées d'espace,

$$n(\mathbf{r}) = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \rho(\mathbf{r}, \mathbf{\Omega}) d\mathbf{\Omega}, \quad (3.2.7)$$

et  $\mathbf{P}(\mathbf{r})$  le champ de polarisation dipolaire défini comme,

$$\mathbf{P}(\mathbf{r}) = \mu \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \tilde{\mathbf{n}} \rho(\mathbf{r}, \mathbf{\Omega}) d\mathbf{\Omega}, \quad (3.2.8)$$

avec  $\boldsymbol{\mu}$  le moment dipolaire d'une molécule d'eau de norme  $\mu$  et  $\tilde{\mathbf{n}} = \boldsymbol{\mu}/\mu$  le vecteur orientation unitaire de ce moment dipolaire.

Le premier terme du membre de droite de l'éq. 3.2.6 est un terme qui corrèle les densités entre elles, il est donc lié au facteur de structure. Les deux termes suivants sont des termes de couplage entre polarisation et densité. Les deux termes suivants couplent les densités de polarisation et peuvent donc être reliés aux permittivités diélectriques non-locales du fluide pur.

Enfin, le dernier terme  $\mathcal{F}_{\text{cor}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]$  est le terme correctif de cette approximation, qui est supposé ne dépendre que de  $n(\mathbf{r})$  et  $\mathbf{P}(\mathbf{r})$ . Il est défini comme la différence entre la partie d'excès exacte et les termes du développement à l'ordre 2. On l'appellera fonctionnelle de bridge par identification au terme de bridge des équations intégrales.

Dans la théorie des liquides pour un fluide homogène et isotrope, la fonction de corrélation directe,  $c^{(2)}$ , et la fonction de corrélation de paires  $h$ ,

$$g(r_{12}) = h(r_{12}) + 1, \quad (3.2.9)$$

sont reliées par l'équation de Ornstein-Zernike.

$$h(r_{12}) = c^{(2)}(r_{12}) + n_b \iiint_{\mathbb{R}^3} c^{(2)}(r_{13}) h(r_{32}) d\mathbf{r}_3. \quad (3.2.10)$$

Si le potentiel d'interaction entre particules du fluide est un potentiel de paire additif,  $u(r_{12})$ , une autre relation exacte peut être écrite par expansion diagrammatique[55, 4],

$$h(r) = \exp[-\beta u(r) + h(r) - c(r) + B(r)] - 1, \quad (3.2.11)$$

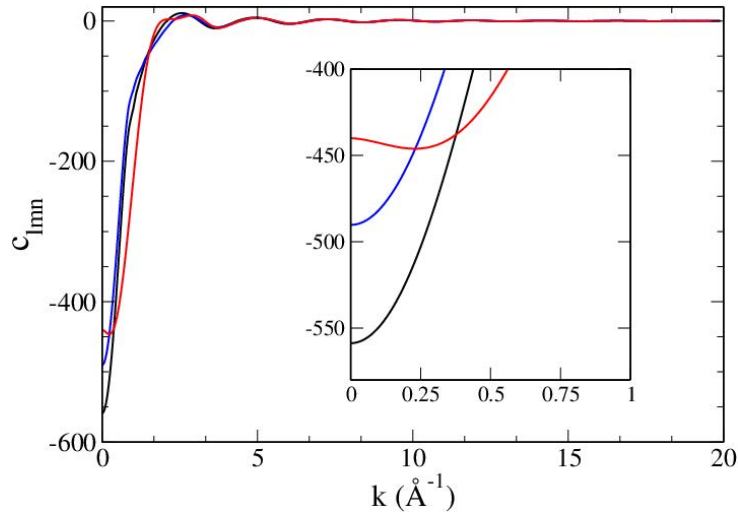
où  $B$  est appelée fonction bridge. Si cette fonction est connue on a une expression exacte de la fonction de corrélation totale. Cependant, cette fonction bridge n'a pas d'expression connue dans le cas général. Une approximation de l'éq. 3.2.11 s'appelle une relation de fermeture et en résolvant de manière conjointe cette équation de fermeture et l'équation de Ornstein-Zernike on peut calculer les différentes fonctions de corrélation. Une approximation classique, l'approximation HNC, consiste à prendre cette fonction bridge nulle.

Si on suppose que  $\mathcal{F}_{\text{cor}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]$  est nulle dans l'éq. 3.2.6, on retrouve la fonctionnelle d'excès dans l'approximation HRF qui est équivalente à l'approximation HNC. On voit le rôle similaire que joue la fonctionnelle  $\mathcal{F}_{\text{cor}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]$  dans la cDFT et la fonction bridge dans la théorie des liquides. C'est le terme à ajouter à l'approximation HNC pour avoir une solution exacte.

En première approximation, on peut négliger les termes de couplage entre la polarisation et la densité, ce qui revient à omettre les termes qui dépendent de  $c_{011}$  et  $c_{101}$  dans l'éq. 3.2.6. On peut remarquer que les termes en  $c_{110}$  et  $c_{112}$  ont une forme ressemblant au potentiel d'interaction dipôle-dipôle.

Les parties idéales et extérieures peuvent être calculées relativement facilement. La partie d'excès, telle qu'elle est écrite en éq. 3.2.6, est désormais beaucoup moins coûteuse à utiliser numériquement : contrairement à l'éq. 3.2.4, elle ne requiert plus d'intégration sur les angles. La double

intégration sur l'espace est calculée efficacement par transformées de Fourier rapides (FFT), voir la sec. 6.2. Pour calculer la partie d'excès, il est nécessaire de connaître les 3 projections des fonctions de corrélation directe utilisées  $c_{000}$ ,  $c_{110}$ ,  $c_{112}$ , au lieu de la fonction de corrélation  $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{\Omega}_1, \mathbf{\Omega}_2)$ . Cette fonction, qui dépend de douze variables est particulièrement compliquée à déterminer par simulation numérique. Le problème est donc considérablement simplifié puisqu'on le ramène à la détermination de trois fonctions à une seule variable. La détermination de ces fonctions peut se faire à partir de simulations de dynamique moléculaire de deux manières, soit en résolvant l'équation de Ornstein-Zernike dans l'espace de Fourier[56], soit dans l'espace direct en résolvant les équations de Baxter[50, 57] ; ce qui a été fait pour l'eau au laboratoire avant mon arrivée. Ces fonctions peuvent également être déterminées par des mesures expérimentales.



**FIGURE 3.2.2.:** Composantes des projections de la fonction de corrélation directe de l'eau SPC/E,  $c_{000}$  est en rouge,  $c_{110}$  en noir et  $c_{112}$  en bleu. Dans l'encart est représenté un zoom entre 0 et 1  $\text{\AA}^{-1}$ .

Dans le cas où on néglige les couplages entre polarisation et densité, c'est-à-dire si on suppose que les termes  $c_{011}$  et  $c_{101}$  sont nuls, les transformées de Fourier de la fonction  $c_{000}$  et celle du facteur de structure sont directement liées[4],

$$\hat{S}(\mathbf{k}) = \frac{1}{1 - n_b \hat{c}_{000}(\mathbf{k})}. \quad (3.2.12)$$

Les transformées de Fourier de  $c_{110}$  et  $c_{112}$  sont quant à elles reliées aux composantes transverse et longitudinale des permittivités diélectriques[58, 59] :

$$\frac{n_b}{3} (\hat{c}_{110}(\mathbf{k}) + 2\hat{c}_{112}(\mathbf{k})) = 1 - \frac{4\beta\mu^2 n_b}{3\epsilon_0 (1 - \hat{\epsilon}_L^{-1}(\mathbf{k}))} \quad (3.2.13)$$

et

$$\frac{n_b}{3} (\hat{c}_{110}(\mathbf{k}) - \hat{c}_{112}(\mathbf{k})) = 1 - \frac{4\beta\mu^2 n_b}{3\epsilon_0 (\hat{\epsilon}_T(\mathbf{k}) - 1)}. \quad (3.2.14)$$

En déterminant expérimentalement les deux composantes des permittivités diélectriques et le facteur de structure, on peut ensuite calculer ces fonctions de corrélation.

Les fonctions de corrélation ne doivent être déterminées qu'une seule fois pour chaque solvant. On peut ensuite étudier la solvation de n'importe quel soluté dans ce solvant puisque celui-ci n'intervient que comme une perturbation, au travers du champ extérieur.

On va maintenant tester la qualité de cette fonctionnelle dipolaire en étudiant les propriétés de solvation de divers solutés en milieux aqueux. Pour l'instant, on fait l'hypothèse simplificatrice que le terme de bridge  $\mathcal{F}_{\text{cor}}$  est nul. On généralisera par la suite au cas d'une fonctionnelle multipolaire tenant compte de la distribution de charge complète des molécules d'eau.

### 3.2.2. Structure de solvation

Nous allons tester la fonctionnelle précédemment introduite pour évaluer sa précision pour la prédiction des propriétés structurales de solvation d'ions et de petites molécules en milieu aqueux. Les fonctions de distribution radiale calculées à partir des densités d'équilibre obtenues par minimisation de la fonctionnelle seront comparées aux résultats exacts obtenus par dynamique moléculaire. Les études ont été réalisées en utilisant les mêmes modèles de solutés et de solvants, *i.e.*, les mêmes potentiels d'interaction dans les simulations MD et dans les calculs MDFT.

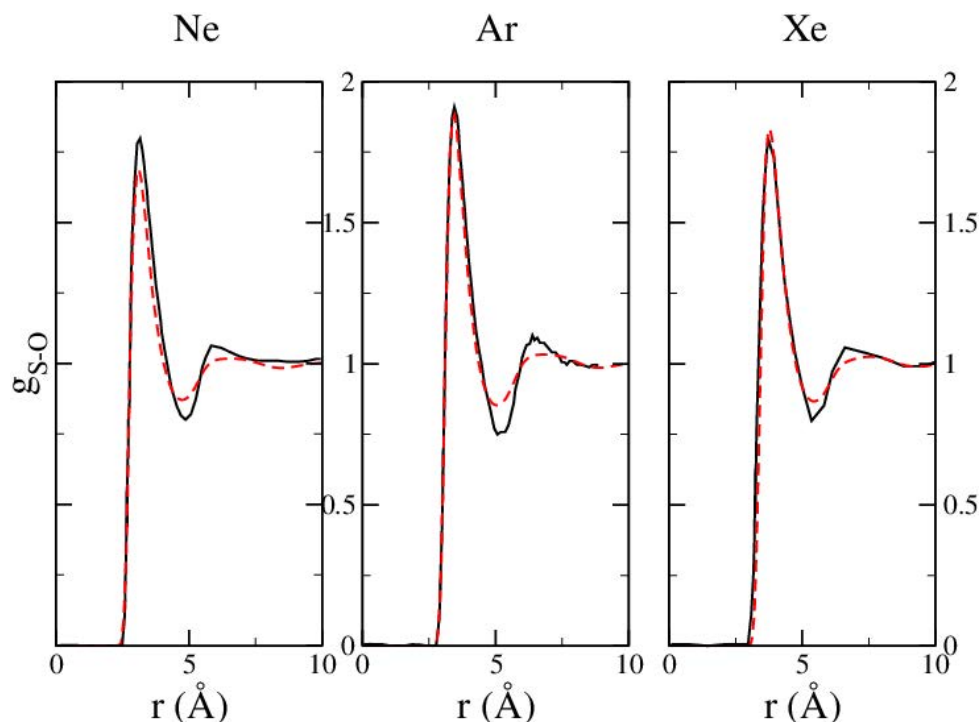
#### 3.2.2.1. Application aux solutés apolaires

On étudie la solvation en milieu aqueux supposé infiniment dilué des atomes d'argon, néon et xénon. Ces atomes sont neutres et les paramètres du champ de force Lennard-Jones[60] représentant l'interaction entre atomes de gaz rares et solvant sont donnés dans le tab. 3.2. Les fonctions de distribution radiale obtenues par minimisation de la fonctionnelle ainsi que celles obtenues par dynamique moléculaire[60, 61] sont présentées en fig. 3.2.3. Dans les trois cas, l'accord entre MDFT et MD est satisfaisant puisque les positions des extrema de ces fonctions sont très bien reproduites, ainsi que la hauteur du pic correspondant à la première couche de solvation. On sous-estime légèrement la hauteur du pic dû à la seconde couche de solvation, ainsi que la déplétion entre les deux maxima. On peut remarquer que l'accord est meilleur à mesure que la taille des solutés augmente.

Atome	$\sigma$ (Å)	$\epsilon$ (kJ.mol <sup>-1</sup> )	$q$ (e)
Ne	3.035	0.15432	0.0
Ar	3.415	1.03891	0.0
Xe	3.975	1.78510	0.0

**TABLE 3.2.:** Champ de force pour les modèles de gaz rares étudiés.

On s'intéresse ensuite à des solutés apolaires plus complexes car non-sphériques en examinant la série des six premiers alcanes linéaires. Les paramètres de potentiel d'interactions sont ceux du



**FIGURE 3.2.3.:** Fonctions de distribution radiale entre l'oxygène de l'eau et les atomes de néon, d'argon et de xéon. Les résultats obtenus par dynamique moléculaire[60, 61] sont représentés par une ligne noire continue, tandis que les résultats obtenus par MDFT sont présentés par une ligne rouge discontinue.

champ de force OPLS (Optimized Intermolecular Potential Functions for Liquid Hydrocarbons)[62] et sont donnés dans le tab. 3.3. Dans ce modèle les groupements  $\text{CH}_n$  sont représentés par une seule sphère de Lennard-Jones. La comparaison entre les fonctions de distribution radiale obtenues par MD et MDFT est donnée en fig. 3.2.4. Seuls les résultats pour les trois premiers alcanes sont présentés. Là encore, l'accord entre MD et MDFT est satisfaisant. On reproduit correctement la position des première et deuxième sphères de solvation (maxima), ainsi que la hauteur de ces pics. La déplétion entre les deux pics est encore une fois légèrement sous-estimée.

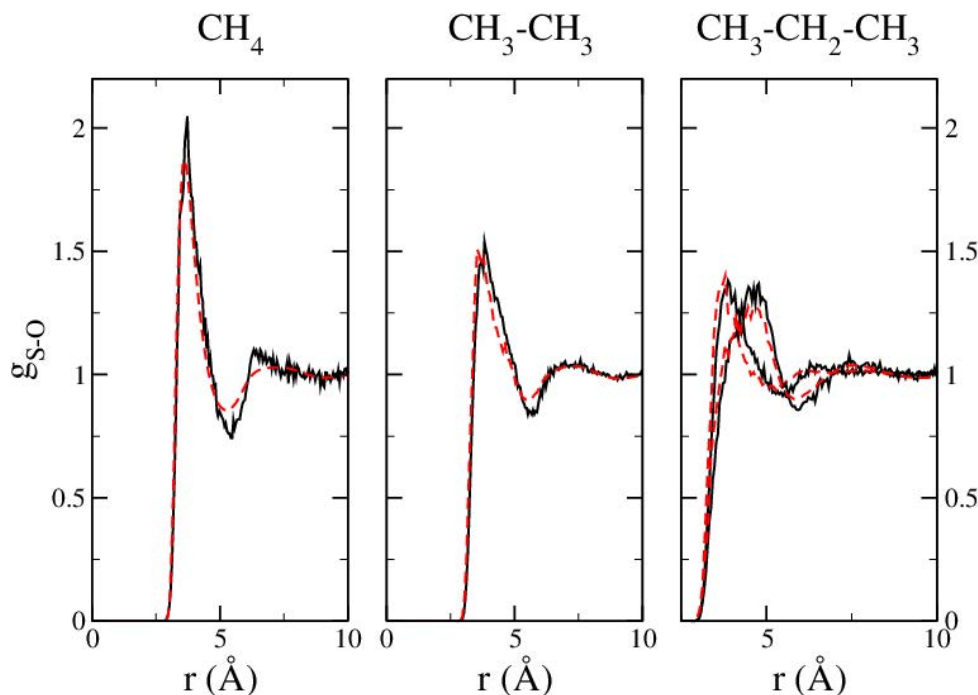
Site	$\sigma$ (Å)	$\epsilon$ (kJ.mol <sup>-1</sup> )	$q$ (e)
CH <sub>4</sub> (méthane)	3.73	1.23	0.0
CH <sub>3</sub> (éthane)	3.775	0.8661	0.0
CH <sub>3</sub>	3.905	0.732	0.0
CH <sub>2</sub>	3.905	0.494	0.0

**TABLE 3.3.:** Champ de force OPLS pour les alcanes étudiés. Les valeurs données dans les deux dernières lignes sont communes au propane, au butane, au pentane et à l'hexane.

### 3.2.2.2. Application à des solutés hydrophiles

La fonctionnelle dipolaire utilisée donne de bons résultats pour des solutés neutres. On s'intéresse maintenant à la solvation de molécules chargées : l'eau et la N-méthylacétamide qui est un





**FIGURE 3.2.4.:** Fonctions de distribution radiale entre l'oxygène de l'eau et les groupements  $\text{CH}_n$  des trois premiers alcanes. Les résultats obtenus par dynamique moléculaire[62] sont représentés par une ligne noire continue, tandis que les résultats obtenus par MDFT sont présentés par une ligne rouge discontinue.

modèle de liaison peptidique.

Le modèle utilisé pour le soluté eau est le même que celui utilisé pour le solvant, SPC/E. Les fonctions de distribution radiale entre l'oxygène du solvant et les oxygènes et hydrogènes de la molécule d'eau soluté, sont présentées en fig. 3.2.5. La minimisation de la fonctionnelle ne permet pas de reproduire des résultats similaires à ceux obtenus par MD.

En effet, dans la fonction de distribution radiale oxygène-oxygène calculée par MD, le premier pic, correspondant à la première couche de solvation est plus haut et plus fin que pour un fluide de Lennard-Jones (par exemple ceux de la fig. 3.2.3).

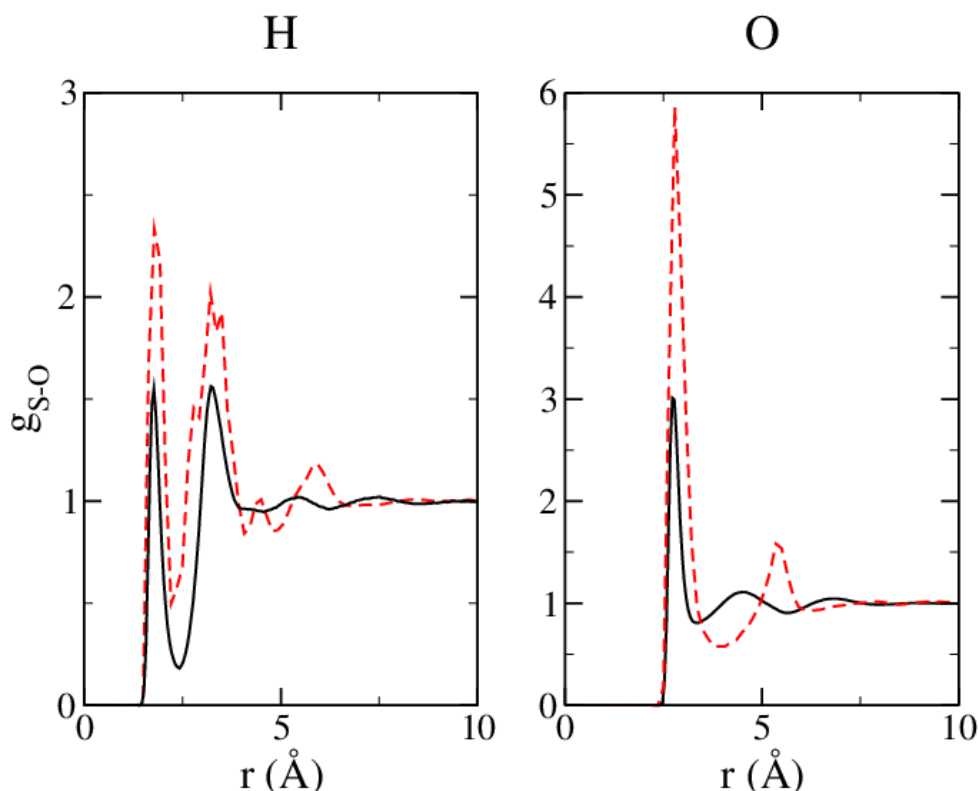
Le second pic, qui correspond à la seconde couche de solvation, est plus proche du premier pic que dans le fluide de Lennard-Jones, où il se trouve à peu près à deux fois la distance du premier pic.

Ce profil particulier peut s'expliquer par la géométrie tétraédrique de l'eau liquide. Cet arrangement spatial particulier se voit aussi sur la fonction de distribution radiale hydrogène-oxygène qui présente deux pics marqués qui proviennent des deux hydrogènes des molécules d'eau situées dans la première couche de solvation. Si l'eau ne possédait pas de structure spatiale localement tétraédrique on ne distinguerait pas ces deux pics mais un seul, dû à une contribution moyenne des hydrogènes.

Pour les résultats obtenus par minimisation fonctionnelle, la position des pics de la première

couche de solvation est partiellement reproduite mais leurs largeurs et leurs hauteurs ne sont pas satisfaisantes. Les pics décrivant la seconde couche de solvation sont déplacés et surestimés.

La fonction de distribution radiale oxygène-oxygène obtenue par MDFT pour l'eau est en fait similaire à celle obtenue pour une sphère Lennard-Jones. La théorie de la fonctionnelle de la densité présentée ici échoue donc à reproduire l'ordre tétraédrique. On peut espérer que ce problème



**FIGURE 3.2.5.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et l'oxygène et l'hydrogène de l'eau soluté. MD en trait noir plein, MDFT en tirets rouges.

soit inhérent à « l'eau dans l'eau ». On étudie donc la solvation d'autres petites molécules polaires. On donne ici l'exemple de la N-méthylacétamide (NMA) avec les paramètres de potentiel d'interaction donnés dans le tab.3.4 Les fonctions de distribution radiale obtenues par MD et

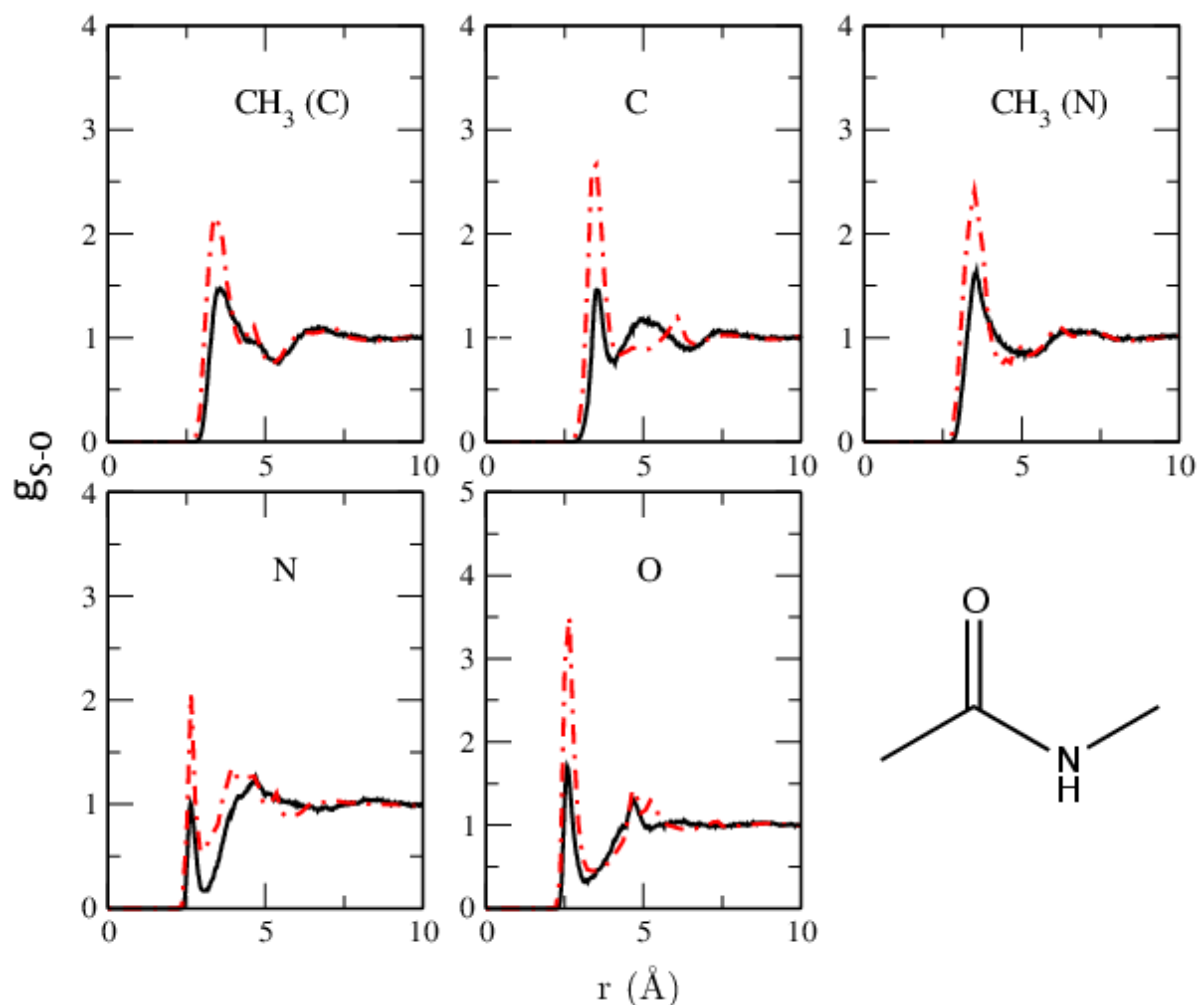
Site	$\sigma$ (Å)	$\epsilon$ (kJ.mol <sup>-1</sup> )	$q$ (e)
CH <sub>3</sub> (C)	3.91	0.160	0.00
C	3.75	0.105	0.50
O	2.96	0.210	-0.50
N	3.25	0.170	-0.57
H	0.00	0.000	0.37
CH <sub>3</sub> (N)	3.80	0.170	0.20

**TABLE 3.4.:** Paramètres Lennard-Jones et charges partielles de la N-Méthylacétamide.

minimisation fonctionnelle sont données en fig.3.2.6. Les résultats obtenus sont à rapprocher des précédents. Pour les sites peu ou non chargés, c'est-à-dire les groupements méthyles, l'accord

entre MD et MDFT est correct quoique moins bon que dans le cas des solutés apolaires, alors que pour les sites chargés, l'azote, l'oxygène et le carbone de la fonction amide, les premiers pics sont surestimés et les seconds sont situés à une distance trop grande par rapport aux résultats de la MD, comme dans le cas de l'eau.

Ces sites sont engagés dans des liaisons hydrogène avec le solvant ce qui crée également un arrangement spatial tétraédrique local autour du soluté. Il est donc probable que le mauvais accord entre MD et MDFT ait la même origine que pour le soluté eau. Pour tester cette hypothèse,



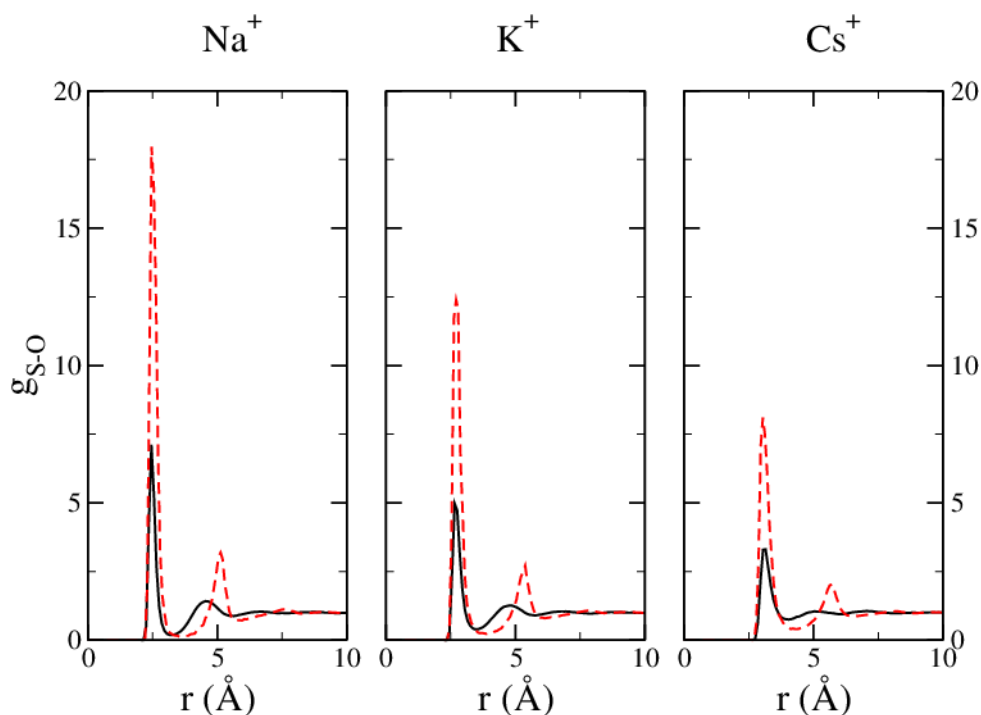
**FIGURE 3.2.6.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et les différents sites de la molécule NMA. La légende est la même que précédemment.

nous avons étudié des solutés fortement hydrophiles : des cations alcalins et des anions halogénures. Nous avons utilisé les potentiels d'interaction indiqués dans le tab.3.5. Les fonctions de distribution radiale, obtenues par dynamique moléculaire et par minimisation fonctionnelle sont montrées en fig. 3.2.7 et fig. 3.2.8. L'écart entre la MDFT et les résultats exacts obtenus par MD est encore plus flagrant que dans le cas des molécules polaires, puisque la MDFT surestime grandement le premier pic de solvation. Les pics dus aux deuxième et troisième couches de

Site	$\sigma$ (Å)	$\epsilon$ (kJ.mol <sup>-1</sup> )	$q$ (e)
Na <sup>+</sup>	2.581	0.129	1.0
K <sup>+</sup>	2.931	0.760	1.0
Cs <sup>+</sup>	3.531	1.500	1.0
Cl <sup>-</sup>	4.035	0.410	-1.0
Br <sup>-</sup>	4.581	0.499	-1.0
I <sup>-</sup>	4.921	0.670	-1.0

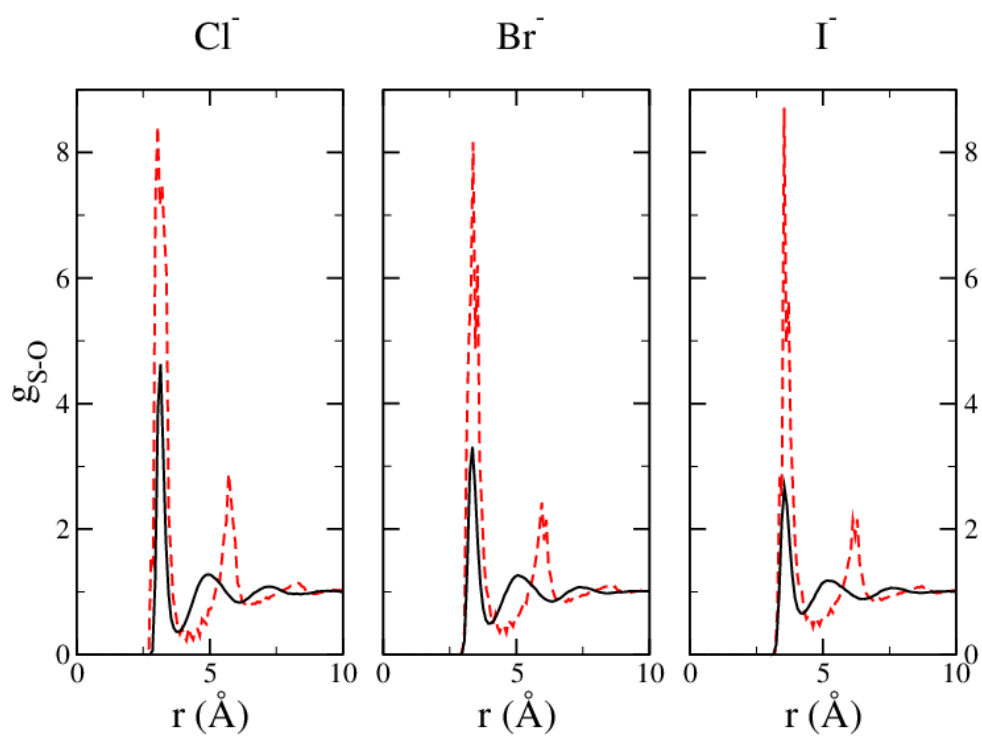
**TABLE 3.5.:** Paramètres Lennard-Jones et charges partielles des cations alcalins et anions halogénures.

solvatation sont eux aussi surestimés mais également situés à une distance trop grande du soluté.



**FIGURE 3.2.7.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et trois cations alcalins : sodium, potassium et césium. Les résultats des simulations MD sont en trait noir plein et ceux des calculs MFT en tirets rouges.

Pour résumer, une fonctionnelle dans l'approximation du fluide homogène de référence avec un traitement de la polarisation strictement dipolaire, et sans terme de bridge, donne d'excellents résultats quant à la structure de solvatation des solutés apolaires. Cependant, elle ne permet pas de reproduire correctement les structures de solvatation lorsque les solutés sont constitués de sites portant une charge électrique importante.



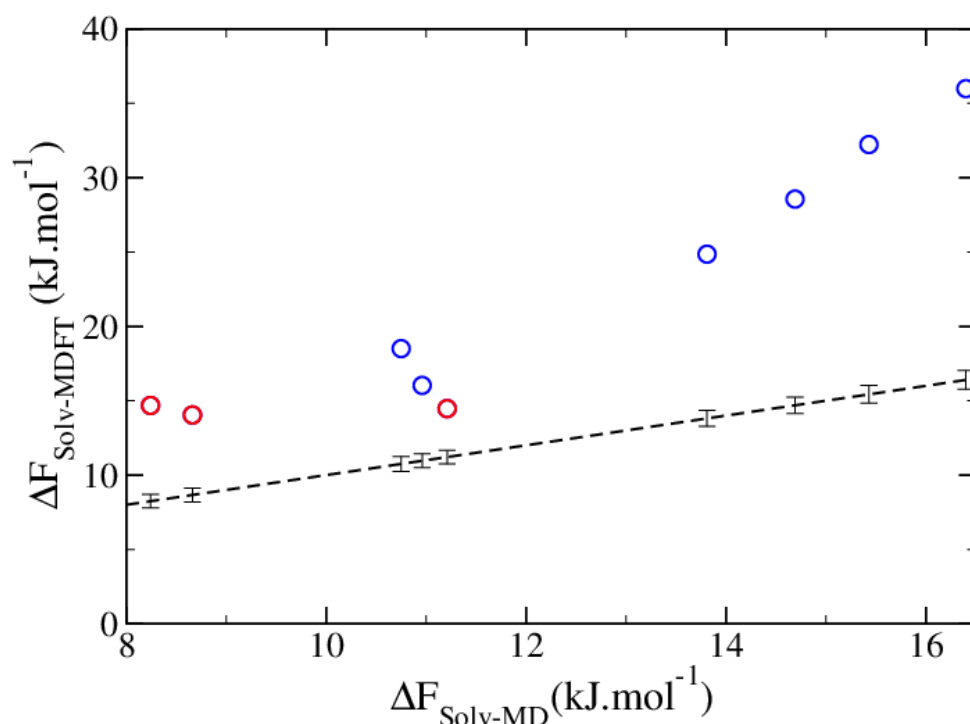
**FIGURE 3.2.8.:** Même figure que la fig. 3.2.7 avec des solutés halogénures : chlorure, bromure, iodure.

### 3.2.3. Énergies de solvation

Nous allons désormais examiner les prédictions des énergies libres de Helmholtz de solvation en milieux aqueux par minimisation de la fonctionnelle. Ces résultats seront cette fois encore comparés à ceux obtenus par dynamique moléculaire. On rappelle qu'en dynamique moléculaire il est nécessaire d'utiliser des techniques d'intégration thermodynamique pour accéder aux énergies libres.

#### 3.2.3.1. Application aux solutés apolaires

La comparaison des résultats obtenus par MD et MDFT est présentée sur la fig. 3.2.9. L'accord est loin d'être satisfaisant puisque les énergies calculées par minimisation fonctionnelle valent pratiquement le double de celles calculées par MD. On peut s'interroger sur l'origine de ces mauvais résultats alors que la structure de solvation est très bien reproduite (voir fig. 3.2.3 et fig. 3.2.4). On peut penser que l'interaction entre soluté et solvant est correctement estimée par la fonctionnelle, ce qui explique notamment l'excellente reproduction de la première couche de solvation, mais que l'interaction solvant-solvant est moins bien représentée. Ceci explique la légère différence au niveau des couches de solvation suivantes. Cette mauvaise estimation de



**FIGURE 3.2.9.:** Comparaison des énergies libres de solvation obtenues par MD et MDFT pour les gaz rares et les six premiers alcanes linéaires. La droite en tirets noirs est le résultat obtenu par MD [63, 60]. Les cercles sont les valeurs obtenues par minimisation de la fonctionnelle, en rouge pour les gaz rares, en bleu pour les alcanes. Les barres d'erreurs sont celles obtenues par MD.

l'énergie libre de solvation provient donc des approximations réalisées sur la partie d'excès,

c'est-à-dire l'approximation du fluide homogène de référence et le développement à l'ordre strictement dipolaire de la fonction de corrélation directe. En restant dans l'HRF, on peut essayer d'améliorer la fonctionnelle en proposant une approximation du terme de bridge  $\mathcal{F}_{\text{cor}}[\rho(\mathbf{r}, \boldsymbol{\Omega})]$  qui avait été supposé nul jusqu'à présent. Ce terme inconnu peut s'écrire en développant de manière systématique la fonction de corrélation directe du solvant pur en fonction des fonctions de corrélation à  $N$  corps, avec  $N \geq 3$ . Cette approche n'a pas été réalisée car elle nécessite la connaissance de ces fonctions de corrélation d'ordre supérieur pour le liquide pur. En revanche, une approximation de ce terme a été proposée au laboratoire[53] : l'approximation du bridge de sphères dures (HSB pour hard spheres bridge). Cette approximation [64, 65, 66, 67, 68] consiste à remplacer le terme de bridge, inconnu pour l'eau, par le terme de bridge d'un fluide de sphères dures de même densité, qui lui est connu exactement,

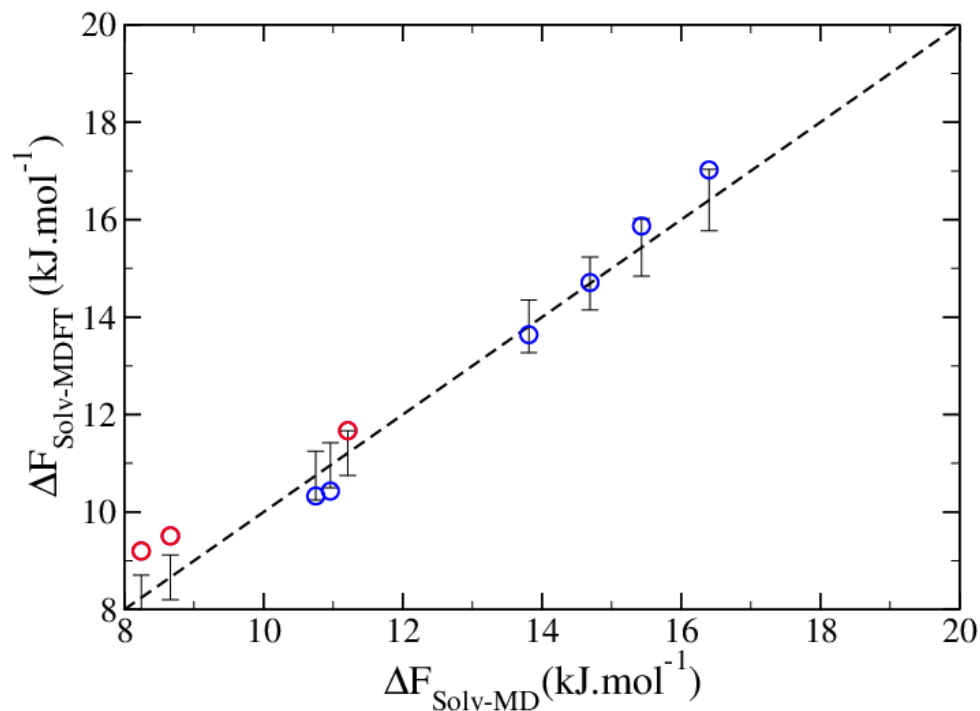
$$\begin{aligned} \mathcal{F}_{\text{HSB}}[n(\mathbf{r})] &= \mathcal{F}^{\text{HS}}[n(\mathbf{r})] - \mathcal{F}^{\text{HS}}[n_b] - \mu^{\text{HS}} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) d\mathbf{r} \\ &+ \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} n\rho(\mathbf{r}) c_{\text{HS}}^{(2)}(\|\mathbf{r} - \mathbf{r}'\|; \rho_b) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \end{aligned} \quad (3.2.15)$$

L'écriture de ce terme fait intervenir une fonctionnelle du fluide de sphères dures  $\mathcal{F}^{\text{HS}}$ , dont des expressions de très bonnes qualités sont connues [69, 70, 71] (l'Appendice C traite de ces fonctionnelles de sphères dures). Intervient également la fonction de corrélation directe du fluide de sphères dures  $c_{\text{HS}}^{(2)}$ , que l'on peut calculer comme la dérivée fonctionnelle seconde de la fonctionnelle de sphères dures par rapport à la densité. On voit que cette approximation revient à remplacer tous les termes en  $\Delta\rho^n$  avec  $n \geq 3$ , supposés nuls dans l'approximation du fluide homogène de référence, par ceux du fluide de sphères dures.

On peut voir cette approximation du bridge de sphères dures de deux points de vue différents, selon qu'on corrige les approximations du modèle approché développé ici (MDFT-HRF) ou au contraire qu'on modifie un modèle de sphères dures pour le faire ressembler à un liquide réel :

- Dans le cas du fluide de sphères dures, la fonctionnelle est connue avec une excellente approximation et il est possible de calculer la fonction de corrélation directe d'ordre deux en utilisant l'éq. 2.3.2. On peut donc modifier la fonctionnelle de sphères dures en lui ajoutant un « léger goût » d'eau. Pour cela, on « remplace » la fonction de corrélation directe d'ordre deux par celle connue de l'eau dans l'approximation du fluide de référence.
- Dans le cas d'un développement de Taylor de la fonctionnelle dans le cadre de l'approximation du fluide de référence, on a « supprimé » tous les termes d'ordre supérieur à deux en densité, et ce terme de bridge sphères dures sert à corriger cette approximation. Ces termes d'ordres supérieurs sont à courte-portée donc essentiellement des termes d'empilement, il ne semble pas absurde de les remplacer par des termes sphères dures.

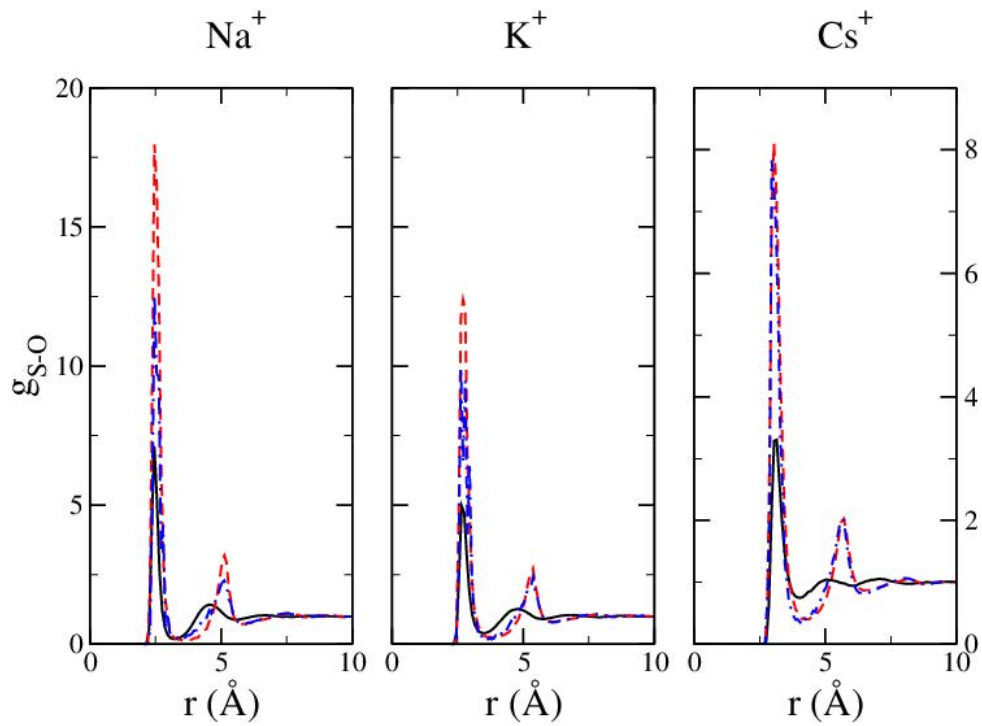
Une figure analogue à la fig. 3.2.9 mais où l'on a inclus ce bridge sphères dures est présentée en fig. 3.2.10. On constate que cette correction améliore grandement les valeurs des énergies libres de solvation obtenues sans modifier les bons résultats des fig. 3.2.3 et fig. 3.2.4. Sur la fig. 3.2.11 sont présentées les fonctions de distribution radiale pour les cations alcalins obtenues par MD,



**FIGURE 3.2.10.:** Comparaison des énergies libres de solvation obtenues par MD et MDFT avec la correction de bridge sphères dures pour les gaz rares et les six premiers alcanes linéaires. La droite en tirets noirs est le résultat obtenu par MD [63, 60]. Les cercles sont les valeurs obtenues par minimisation de la fonctionnelle, en rouge pour les gaz rares, en bleu pour les alcanes. Les barres d'erreurs sont celles obtenues par MD.

MDFT et MDFT avec la correction bridge sphères dures. Cette correction améliore légèrement les fonctions de distribution radiale en réduisant la hauteur du premier pic, mais l'accord reste néanmoins insuffisant. Elle ne résout pas le manque d'ordre tétraédrique dans le modèle.





**FIGURE 3.2.11.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et trois cations alcalins, sodium, potassium et césium. Les résultats des simulations MD sont en trait noir plein, ceux des calculs MDFT en tirets rouges et ceux obtenus avec MDFT et la correction HSB en tirets bleus.

### 3.3. Au delà de l'ordre dipolaire

On s'est limité dans la partie précédente à un traitement strictement dipolaire de la partie d'excès pour l'eau. Lors de ma thèse, j'ai développé une écriture de cette partie d'excès ne restreignant plus le traitement de la polarisation à l'ordre dipolaire. Cette formulation est présentée ici.

#### 3.3.1. Écriture de la polarisation multipolaire

La polarisation décrite dans l'éq. 3.2.8 est d'ordre purement dipolaire. C'est-à-dire que l'on a restreint la symétrie de l'interaction effective contenue dans la fonction  $c$  à l'ordre 2. Une telle approximation est valable pour un fluide purement dipolaire. Cependant, si on considère le modèle d'eau SPC/E, on se rend compte que la distribution de charge n'est pas purement dipolaire puisque le modèle est constitué de trois charges partielles, voir fig. 3.1.1. L'approximation réalisée précédemment n'est pas justifiée a priori.

On peut définir une densité de charge  $\sigma(\mathbf{r}, \mathbf{\Omega})$  et de polarisation  $\boldsymbol{\mu}(\mathbf{r}, \mathbf{\Omega})$  [58, 59] pour une molécule d'eau SPC/E prise à l'origine du repère et ayant une orientation  $\mathbf{\Omega}$ ,

$$\begin{aligned}\sigma(\mathbf{r}, \mathbf{\Omega}) &= \sum_{\mathbf{m}} q_{\mathbf{m}} \delta(\mathbf{r} - \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})) \\ &= q_O \delta(\mathbf{r}) + q_H [\delta(\mathbf{r} - \mathbf{s}_{H_1}(\mathbf{\Omega})) + \delta(\mathbf{r} - \mathbf{s}_{H_2}(\mathbf{\Omega}))],\end{aligned}\tag{3.3.1}$$

et

$$\boldsymbol{\mu}(\mathbf{r}, \mathbf{\Omega}) = \sum_{\mathbf{m}} q_{\mathbf{m}} \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega}) \int_0^1 \delta(\mathbf{r} - u \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})) du.\tag{3.3.2}$$

$\mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})$  désigne la position du  $\mathbf{m}^{\text{ième}}$  site du solvant quand la molécule possède l'orientation  $\mathbf{\Omega}$ , et  $q_{\mathbf{m}}$  désigne la charge de ce site. On peut réécrire la polarisation microscopique dans l'espace de Fourier,

$$\begin{aligned}\boldsymbol{\mu}(\mathbf{k}, \mathbf{\Omega}) &= \sum_{\mathbf{m}} q_{\mathbf{m}} \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega}) \int_0^1 e^{i u \mathbf{k} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})} du \\ &= -i \sum_{\mathbf{m}} q_{\mathbf{m}} \frac{\mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})}{\mathbf{k} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})} (e^{i \mathbf{k} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})} - 1) \\ &= \sum_{\mathbf{m}} q_{\mathbf{m}} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega}) + i \frac{q_{\mathbf{m}}}{2} (\mathbf{k} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})) \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega}) + \mathcal{O}(\mathbf{k}^2) \\ &= \boldsymbol{\mu}(\mathbf{\Omega}) + i \frac{q_{\mathbf{m}}}{2} (\mathbf{k} \cdot \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega})) \mathbf{s}_{\mathbf{m}}(\mathbf{\Omega}) + \mathcal{O}(\mathbf{k}^2).\end{aligned}\tag{3.3.3}$$

On reconnaît le premier terme, qui est le terme purement dipolaire. Les termes d'ordres supérieurs (c'est-à-dire le terme d'ordre un et les termes d'ordres supérieurs qui ont été écrits, pour des

raisons de compacité, en  $\mathcal{O}(\mathbf{k}^2)$ ) proviennent de la distribution de charge non dipolaire de la molécule. On peut vérifier que c'est en fait la densité de polarisation associée à la distribution de charge d'une molécule de solvant centrée à l'origine via l'équation de Poisson locale :

$$\sigma(\mathbf{r}, \mathbf{\Omega}) = -\nabla \cdot \boldsymbol{\mu}(\mathbf{r}, \mathbf{\Omega}), \quad (3.3.4)$$

soit dans l'espace de Fourier,

$$\sigma(\mathbf{k}, \mathbf{\Omega}) = -i\boldsymbol{\mu}(\mathbf{k}, \mathbf{\Omega}). \quad (3.3.5)$$

Dans le cas d'un solvant constitué d'un seul site Lennard-Jones (comme l'eau SPC/E), le Hamiltonien qui décrit le système de  $N$  molécules d'eau en présence d'un soluté s'écrit

$$\mathcal{H} = \mathcal{K} + \mathcal{U} + \iiint_{\mathbb{R}^3} \tilde{n}(\mathbf{r}) v_{\text{LJ}}(\mathbf{r}) d\mathbf{r} - \iiint_{\mathbb{R}^3} \tilde{\mathbf{P}}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}) d\mathbf{r} \quad (3.3.6)$$

$v_{\text{LJ}}$  désigne le potentiel extérieur Lennard-Jones dû à l'interaction soluté-solvant et  $\mathbf{E}$  le champ électrostatique créé par le soluté en un point  $\mathbf{r}$ .  $\tilde{\rho}(\mathbf{r}, \mathbf{\Omega})$  et  $\tilde{n}(\mathbf{r})$  sont les densités de solvant microscopiques et  $\tilde{\mathbf{P}}(\mathbf{r})$  la densité de polarisation microscopique, c'est-à-dire calculées pour une configuration des molécules donnée. Ces grandeurs s'écrivent :

$$\tilde{\rho}(\mathbf{r}, \mathbf{\Omega}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{\Omega} - \mathbf{\Omega}_i) \quad (3.3.7)$$

$$\tilde{n}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) = \iiint_{8\pi^2} \tilde{\rho}(\mathbf{r}, \mathbf{\Omega}) d\mathbf{\Omega} \quad (3.3.8)$$

$$\tilde{\mathbf{P}}(\mathbf{r}) = \sum_{i=1}^N \boldsymbol{\mu}(\mathbf{r} - \mathbf{r}_i, \mathbf{\Omega}_i) = \iiint_{\mathbb{R}^3} \int_{8\pi^2} \boldsymbol{\mu}(\mathbf{r} - \mathbf{r}', \mathbf{\Omega}) \tilde{\rho}(\mathbf{r}', \mathbf{\Omega}) d\mathbf{r}' d\mathbf{\Omega}. \quad (3.3.9)$$

On peut alors définir des densités de particule et de polarisation  $n(\mathbf{r}) = \langle \tilde{n}(\mathbf{r}) \rangle$  et  $\mathbf{P}(\mathbf{r}) = \langle \tilde{\mathbf{P}}(\mathbf{r}) \rangle$ . On peut montrer, en suivant une démonstration analogue à celle d'Evans [46], que le grand-potentiel, que l'on notera désormais  $\Theta$  pour éviter la confusion avec les orientations, peut s'exprimer comme une fonctionnelle de ces densités de particule et de polarisation. Ce grand potentiel atteint son minimum pour les densités de particule et polarisation à l'équilibre thermodynamique. La démonstration est donnée en Appendice B.

On va utiliser l'approximation du fluide homogène de référence décrite en sec. 2.3.1. On définit alors de manière identique la fonctionnelle d'énergie libre  $\mathcal{F}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \Theta[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] - \Theta_0$ , qui est également une fonctionnelle des densités de particule et de polarisation.

On introduit la fonctionnelle intrinsèque, qui est la partie ne dépendant pas des champs extérieurs :

$$\mathcal{F}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \mathcal{F}_{\text{int}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] + \iiint_{\mathbb{R}^3} n(\mathbf{r}) v_{\text{LJ}}(\mathbf{r}) d\mathbf{r} - \iiint_{\mathbb{R}^3} \mathbf{P}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}) d\mathbf{r} \quad (3.3.10)$$

Cette partie intrinsèque est celle due au solvant seul. Elle contient comme précédemment une partie entropique et une partie d'interaction entre molécules de solvant. On aimerait pouvoir scinder ce terme en un terme idéal et un terme d'excès comme dans l'éq. 2.2.26. Cependant, on ne connaît pas d'expression de l'énergie idéale en fonction des densités de particule et de polarisation. On va donc se ramener à une expression connue et écrire la contribution idéale de la partie intrinsèque comme dépendant de la densité de particule « totale »  $\rho(\mathbf{r}, \mathbf{\Omega})$ , c'est-à-dire utiliser l'expression de l'éq. 2.3.14. On peut montrer, dans le cadre de la théorie de la réponse linéaire, que la fonctionnelle intrinsèque peut s'écrire, sous l'hypothèse d'un développement quadratique, en fonction des corrélations entre densité de particule et densité de polarisation. La démonstration de cette écriture, dans le cas plus simple d'une fonctionnelle dépendant d'une seule variable, est proposée dans le prochain encart. Une démonstration parfaitement similaire est réalisable dans le cas présent. On peut donc écrire la fonctionnelle d'excès comme,

$$\beta \mathcal{F}_{\text{exc}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \beta \mathcal{F}_{\text{int}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})] - \beta \mathcal{F}_{\text{id}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})] \quad (3.3.11)$$

$$\begin{aligned} \beta \mathcal{F}_{\text{exc}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})] &= \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} S^{-1}(r_{12}) \Delta n(\mathbf{r}_1) \Delta n(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ \frac{1}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \chi_L^{-1}(r_{12}) \mathbf{P}_L(\mathbf{r}_1) \mathbf{P}_L(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ \frac{1}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \chi_T^{-1}(r_{12}) \mathbf{P}_T(\mathbf{r}_1) \mathbf{P}_T(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ &- k_B T \iiint_{\mathbb{R}^3} \frac{\Delta n(\mathbf{r})^2}{2n_0} d\mathbf{r} - k_B T \iiint_{\mathbb{R}^3} \frac{3}{2\mu_0 n_0^2} \mathbf{P}(\mathbf{r})^2 d\mathbf{r} \\ &+ \beta \mathcal{F}_{\text{cor}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})]. \end{aligned} \quad (3.3.12)$$

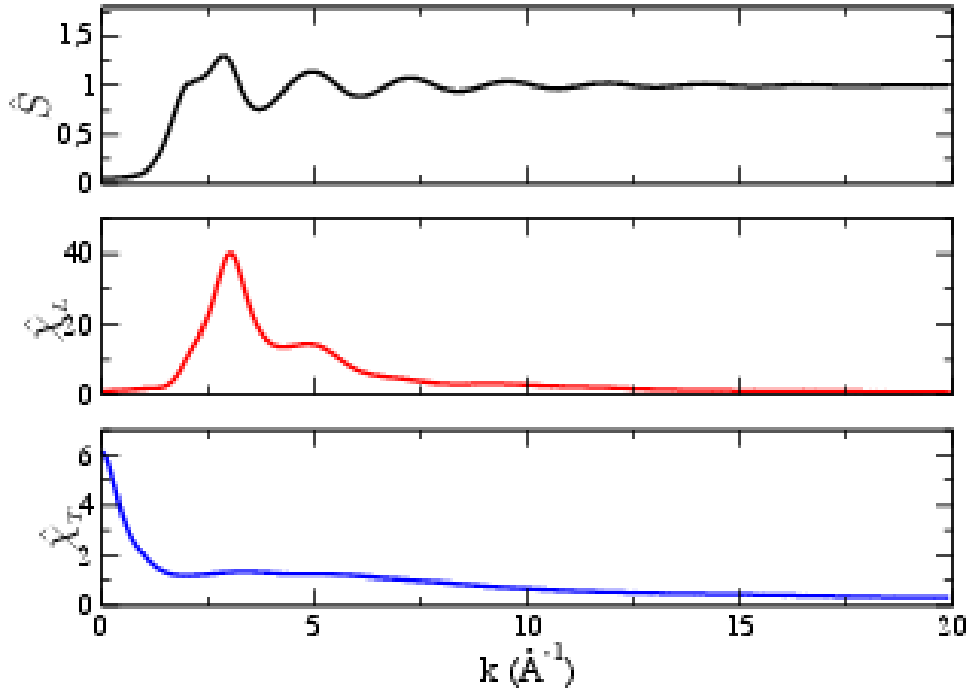
Dans cette expression, on a négligé les termes de couplage entre densité et polarisation. Les trois premiers termes viennent du développement quadratique de la partie intrinsèque. On a séparé ici les composantes longitudinale et transverse de la polarisation, qui sont obtenues dans l'espace de Fourier,

$$\hat{\mathbf{P}}_L(\mathbf{k}) = \frac{(\hat{\mathbf{P}}(\mathbf{k}) \cdot \mathbf{k}) \mathbf{k}}{\|\mathbf{k}\|^2}, \quad \hat{\mathbf{P}}_T(\mathbf{k}) = \hat{\mathbf{P}}(\mathbf{k}) - \hat{\mathbf{P}}_L(\mathbf{k}) \quad (3.3.13)$$

$S^{-1}$ ,  $\chi_L^{-1}$  et  $\chi_T^{-1}$  sont définies comme les transformées de Fourier inverses des inverses du facteur de structure  $\hat{S}$  et des susceptibilités diélectriques longitudinale et transverse  $\hat{\chi}_L$  et  $\hat{\chi}_T$ . Soulignons que ces fonctions sont liées aux fluctuations de densité et de polarisation du solvant pur. Ces fonctions ont été calculées par dynamique moléculaire dans l'espace de Fourier, grâce à la méthode introduite par Bopp et collaborateurs [72, 73]. Ces fonctions sont présentées en fig. 3.3.1.

On mentionne que les susceptibilités diélectriques longitudinale et transverse peuvent être reliées aux composantes longitudinale et transverse des constantes diélectriques par les relations,

$$\hat{\chi}_L(k) = \frac{1}{1 - \hat{\epsilon}_L(k)}, \quad \hat{\chi}_T(k) = \frac{\hat{\epsilon}_T(k) - 1}{4\pi}, \quad (3.3.14)$$



**FIGURE 3.3.1.:** Facteur de structure et composantes longitudinale et transverse des susceptibilités de l'eau SPC/E, calculées avec les formules de Bopp et collaborateurs[73].

de telle sorte que leurs valeurs aux petites valeurs de  $k$ , qui sont mal estimées par les simulations à cause d'effet de tailles finies, peuvent être extrapolées puisque l'on connaît la valeur macroscopique des constantes diélectriques  $\epsilon_T(0) = \epsilon_L(0) = 71$  pour l'eau SPC/E[74].

Les deux termes suivants de l'éq. 3.3.12 sont les termes linéarisés de la partie idéale, que l'on retire pour ne pas les compter deux fois. Le dernier terme est un terme correctif inconnu qui contient à priori tous les ordres supérieurs à deux en  $\Delta n$  et  $\mathbf{P}$ . Ce terme est l'équivalent du terme de bridge dans le développement dipolaire de la fonctionnelle présenté plus haut.

Soulignons que dans tout ce développement, les termes de couplage entre densité de particule et polarisation ont été négligés pour être cohérent avec l'approche dipolaire réalisée précédemment. Cependant une expression de la fonctionnelle d'excès incluant ces termes est proposée en Appendice E.

On peut donner une expression alternative de la fonctionnelle en fonction de la densité de charge de molécules de solvant  $\rho_c$ . La densité de charge du solvant s'exprime également en fonction de la densité de charge microscopique :

$$\rho_c(\mathbf{r}) = \langle \tilde{\rho}_c(\mathbf{r}) \rangle, \quad (3.3.15)$$

avec

$$\tilde{\rho}_c(\mathbf{r}) = \iiint_{8\pi^2} \iiint_{\mathbb{R}^3} \sigma(\mathbf{r} - \mathbf{r}', \mathbf{\Omega}) \tilde{\rho}(\mathbf{r}', \mathbf{\Omega}) d\mathbf{r} d\mathbf{\Omega}. \quad (3.3.16)$$

La densité de charge est également reliée à la polarisation via une équation de Poisson,

$$\rho_c(\mathbf{r}) = -\nabla \cdot \mathbf{P}(\mathbf{r}). \quad (3.3.17)$$

La fonctionnelle se réécrit en fonction de  $\rho_c$ ,

$$\begin{aligned} \beta \mathcal{F}_{\text{exc}}[\Delta n(\mathbf{r}), \rho_c(\mathbf{r})] = & \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} S^{-1}(r_{12}) \Delta n(\mathbf{r}_1) \Delta n(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ & + \frac{1}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} S_{cc}^{-1}(r_{12}) \rho_c(\mathbf{r}_1) \rho_c(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ & + \frac{1}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \chi_T^{-1}(r_{12}) \mathbf{P}_T(\mathbf{r}_1) \mathbf{P}_T(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ & - k_B T \iiint_{\mathbb{R}^3} \frac{\Delta n(\mathbf{r})^2}{2n_0} d\mathbf{r} - k_B T \iiint_{\mathbb{R}^3} \frac{3}{2\mu_0 n_0^2} \mathbf{P}(\mathbf{r})^2 d\mathbf{r} \\ & + \beta \mathcal{F}_{\text{cor}}[\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})]. \end{aligned} \quad (3.3.18)$$

On a introduit le facteur de structure charge-charge défini comme

$$k^2 S_{cc}(k) = \chi_L(k). \quad (3.3.19)$$

On montre ici, dans le cas d'une fonctionnelle ne dépendant que de la densité à une particule  $n(\mathbf{r})$ , comment on peut proposer une expression de la fonctionnelle d'excès.

On suppose qu'on décrit le système dans un état proche de celui du fluide de référence à la densité homogène  $n_b$ . On étudie une petite modulation de la densité de particule  $\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_b$ . On suppose que cette modulation provient de la présence d'un petit potentiel extérieur  $\delta\phi$ .

Ceci revient à supposer que le hamiltonien décrivant le système est égal au hamiltonien du fluide homogène plus cette petite perturbation,

$$\mathcal{H} = \mathcal{H}_b + \delta\phi(\mathbf{r}). \quad (3.3.20)$$

Cette perturbation étant faible, on peut supposer que la réponse de la densité peut être décrite, dans le cadre de la théorie de la réponse linéaire, en fonction de la fonction de réponse  $\chi(\mathbf{r}, \mathbf{r}')$ ,

$$\Delta n(\mathbf{r}) = \iiint_{\mathbb{R}^3} \chi(\mathbf{r}, \mathbf{r}') \delta\phi(\mathbf{r}') d\mathbf{r}' \quad (3.3.21)$$

La fonction de réponse vaut par définition,

$$\chi(\mathbf{r}, \mathbf{r}') = \left. \frac{\delta n(\mathbf{r})}{\delta\phi(\mathbf{r}')} \right|_{\delta\phi=0} = - \left. \frac{\delta n(\mathbf{r})}{\delta\psi(\mathbf{r}')} \right|_{\delta\phi=0} = -\beta \langle \Delta \tilde{n}(\mathbf{r}) \Delta \tilde{n}(\mathbf{r}') \rangle. \quad (3.3.22)$$

L'égalité de gauche se trouve par dérivée fonctionnelle de l'éq. 3.3.21. Le lien avec les fluctuations provient du théorème de fluctuation-dissipation[75].

La fonction de réponse est donc égale aux corrélations densité-densité du système non perturbé. Si l'on prend la transformée de Fourier de l'éq. 3.3.21, on trouve la relation suivante qui relie fonction de réponse, densité et perturbation

$$\Delta\hat{n}(\mathbf{k}) = \hat{\chi}(\mathbf{k})\delta\hat{\phi}(\mathbf{k}). \quad (3.3.23)$$

Si on suppose que  $\mathcal{F}_{\text{int}}$  est quadratique en la modulation de densité, alors,

$$\mathcal{F}_{\text{int}}[\Delta n] = \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) X_0(\|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' + \mathcal{O}(\Delta n^3). \quad (3.3.24)$$

Puisque pour un potentiel extérieur constant la fonctionnelle doit être minimale pour une densité uniforme, il n'y a pas de terme d'ordre 1 en  $\Delta n$ .

La fonction  $X_0(\mathbf{r}, \mathbf{r}')$  est une grandeur caractéristique du fluide de référence et ne dépend donc que de  $\|\mathbf{r} - \mathbf{r}'\|$ .

Si on réécrit l'éq. 3.3.24 dans l'espace réciproque, en transformées de Fourier discrète, celle-ci devient,

$$\mathcal{F}_{\text{int}}[\Delta n] = \frac{1}{2V} \sum_{\mathbf{k}} \Delta\hat{n}(\mathbf{k}) \hat{X}_0(\mathbf{k}) \Delta\hat{n}(-\mathbf{k}) d\mathbf{k} + \mathcal{O}(\Delta\rho^3). \quad (3.3.25)$$

Si on applique le principe variationnel à cette équation on trouve, en utilisant éq. 2.2.27,

$$\Delta\hat{n}(\mathbf{k}) \hat{X}_0(\mathbf{k}) = -\Delta\phi(\mathbf{k}) \quad (3.3.26)$$

Ce qui, en comparant l'éq. 3.3.23 et l'éq. 3.3.26, donne

$$\hat{X}_0(\mathbf{k}) = -\hat{\chi}^{-1}(\mathbf{k}) \quad (3.3.27)$$

Or, on peut relier la fonction de corrélation  $\chi$  au facteur de structure  $S[4]$  :

$$\hat{\chi}(\mathbf{k}) = -\beta n_b \hat{S}(\mathbf{k}) = \frac{-\beta n_b}{1 - n_b \hat{c}(\mathbf{k})} \quad (3.3.28)$$

On a donc dans le cas présent :

$$\begin{aligned} \mathcal{F}_{\text{exc}}[\Delta n] &= \mathcal{F}_{\text{int}}[n] - \mathcal{F}_{\text{id}}[n] \\ &= \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) \chi^{-1}(\|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' - \mathcal{F}_{\text{id}}[n] \\ &= -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) S(\|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' - \mathcal{F}_{\text{id}}[n] \end{aligned} \quad (3.3.29)$$

Si on linéarise le terme idéal, on obtient :

$$\mathcal{F}_{\text{exc}}[\Delta n] = -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) c(\|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \quad (3.3.30)$$

On retrouve bien l'expression de la partie d'excès dans le cadre de l'approximation du fluide homogène de référence.

### 3.3.2. Structures de solvation

Comme attendu, les structures de solvation des solutés neutres, alcanes et gaz rares, sont inchangées à l'ordre multipolaire puisque ceux-ci ne créent pas de champ électrique.

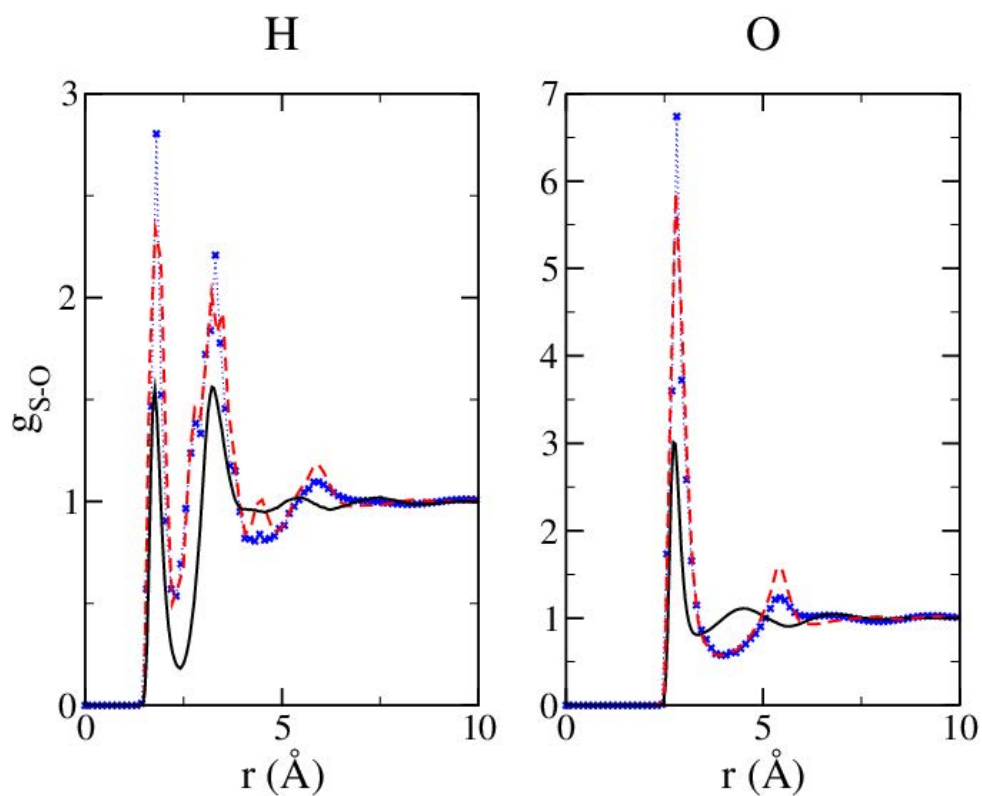
#### 3.3.2.1. Les solutés polaires

On présente, sur les fig. 3.3.2 et fig. 3.3.3, les fonctions de distribution radiale obtenues par simulation de dynamique moléculaire et par minimisation des fonctionnelles dipolaires et multipolaires des sec. 3.2 et sec. 3.3, pour l'eau et la NMA, avec les champs de force des tab. 3.1 et tab. 3.4. On remarque que l'approche multipolaire ne modifie que peu l'allure des fonctions de distribution radiale obtenues par MDFT dipolaire pour les molécules polaires. Il y a une très légère amélioration de la hauteur des pics dus à la seconde couche de solvation dont l'intensité diminue pour l'eau et la NMA.

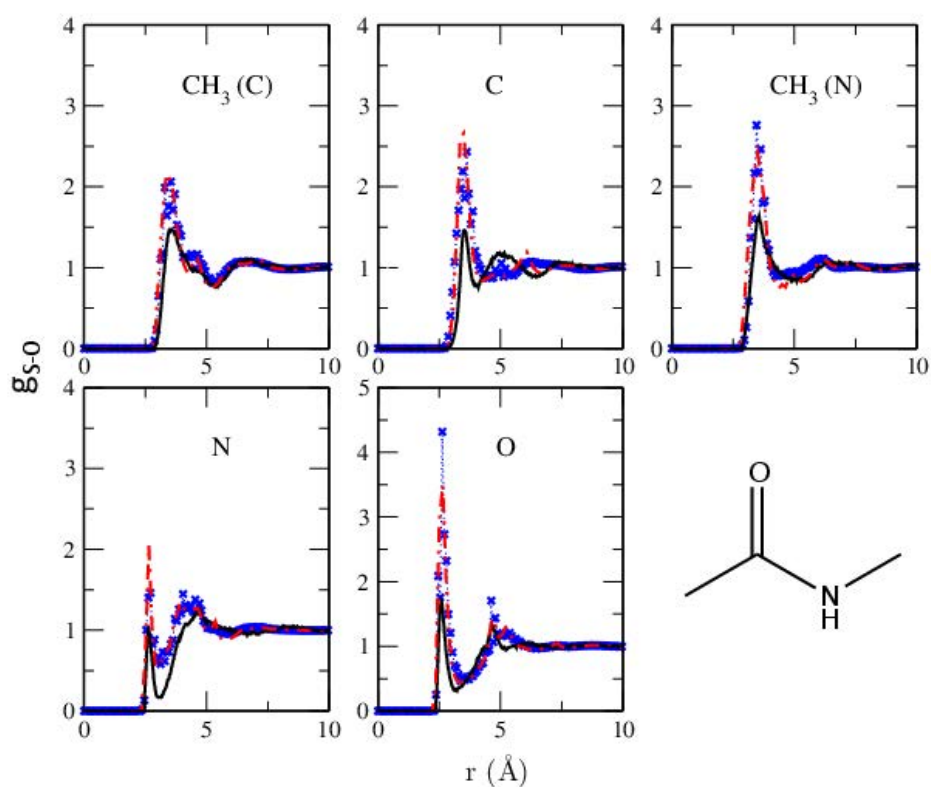
Cette tendance est confirmée pour les ions. On étudie par exemple les cations alcalins dans la fig. 3.3.4, où la hauteur du premier pic et la déplétion entre le premier et le second pic sont nettement diminués.

On arrive donc à une conclusion un peu décevante : un traitement rigoureux de l'électrostatique au delà d'une approximation dipolaire n'améliore pas sensiblement les résultats. Une théorie bilinéaire en  $n(\mathbf{r})$  et  $\mathbf{P}(\mathbf{r})$  conduit à une mésestimation de l'ordre tétraédrique. Pallier ce défaut est l'objet du chapitre suivant.

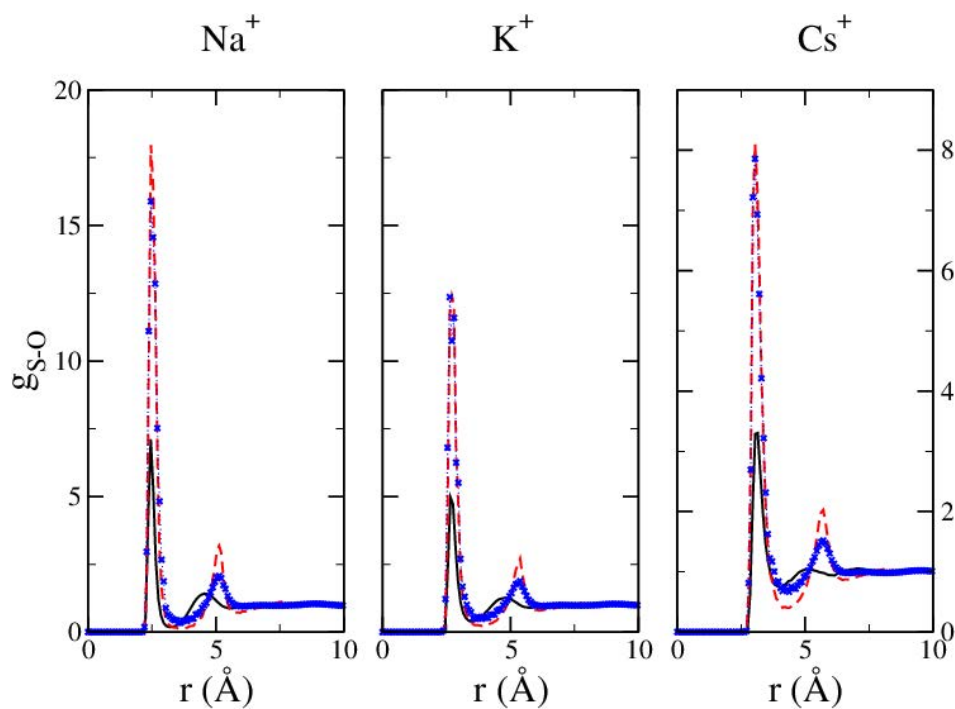




**FIGURE 3.3.2.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et l'oxygène et l'hydrogène de l'eau soluté. Les résultats calculés par MD sont en trait noir plein, tandis que ceux obtenus par minimisation de la fonctionnelle multipolaire sont représentés par des croix bleues reliées par des pointillés. Pour mémoire, on a laissé les résultats obtenus avec la fonctionnelle dipolaire en traits rouges discontinus.



**FIGURE 3.3.3.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et les différents sites de la molécule NMA. La légende est la même que sur la fig. 3.3.2.



**FIGURE 3.3.4.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et trois cations alcalins. La légende est la même que sur la fig. 3.3.2.

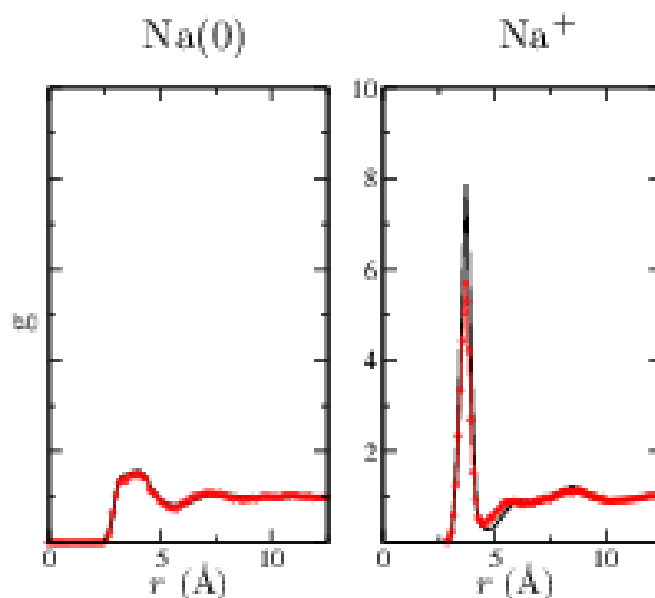
## 4. Correction à trois corps

### 4.1. Retour sur les solutés chargés

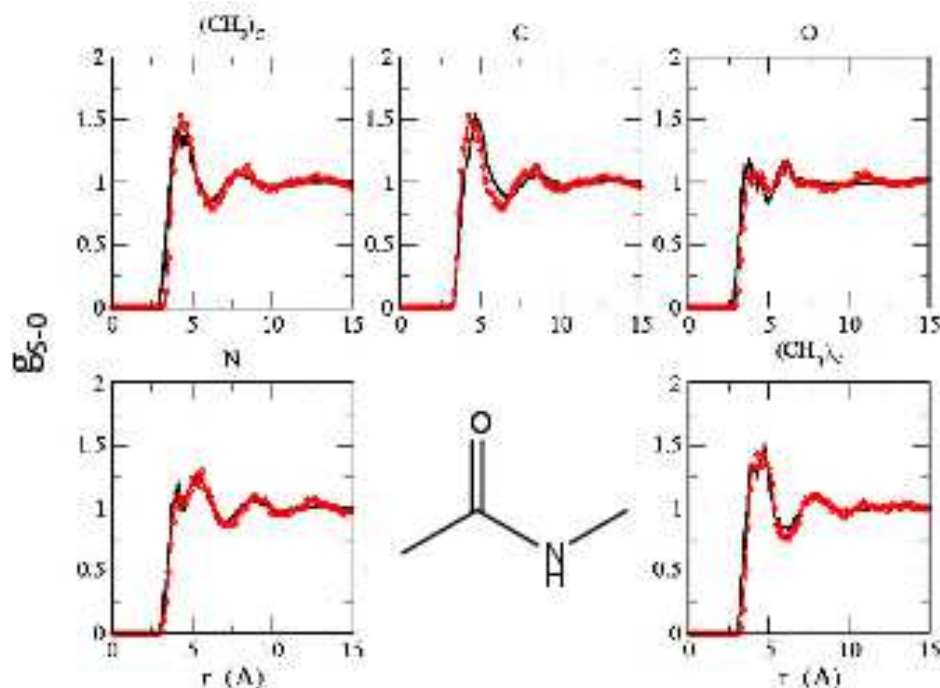
Nous avons vu dans le chapitre 3 que l'on peut écrire une théorie de la fonctionnelle de la densité pour l'eau dans l'approximation du fluide homogène de référence. Cette fonctionnelle donne des résultats satisfaisants pour les solutés apolaires mais pas pour les solutés polaires qui ne sont pas en accord avec la MD.

Cette théorie avait pourtant déjà été utilisée avec succès pour l'étude de molécules polaires et d'ions dans le solvant acétonitrile[51, 52] : par exemple, les fonctions de distribution radiale entre l'acétonitrile et un atome neutre ou un cation sont présentées en fig. 4.1.1, ou avec la molécule de N-méthylacétamide (NMA) en fig. 4.1.2. Les pics des fonctions de distribution radiale dus aux première et deuxième couches de solvation sont très bien reproduits par la MDFT. De plus, les énergies de solvation pour la série des anions halogénures sont en accord avec les mesures expérimentales[51].

Pourquoi les résultats pour la solvation des solutés chargés dans l'eau diffèrent autant de la MD,



**FIGURE 4.1.1.:** Fonctions de distribution radiale entre le centre de masse de la molécule d'acétonitrile et un atome (à gauche) et un ion sodium (à droite) [51]. MD en rouge, MDFT en noir.



**FIGURE 4.1.2.:** Fonctions de distribution radiale entre le centre de masse de la molécule d'acétonitrile et les différents sites de la molécule de NMA[51]. MD en rouge, MDFT en noir.

alors que l'accord est quasi-parfait pour l'acétonitrile ? Ces deux solvants sont polaires (1.85 D pour l'eau, 3.9 D pour l'acétonitrile). L'eau est un solvant protique ce qui en fait un bon donneur de liaisons hydrogènes. De plus, l'eau possède un caractère accepteur de liaisons hydrogènes. L'acétonitrile en revanche, ne développe pas de liaisons hydrogènes. Sur la fig. 3.2.11, le premier pic de la fonction de distribution radiale est à la bonne position mais il est surestimé. Le second pic est lui-aussi surestimé et situé à une distance trop éloignée par rapport au soluté. Ce second pic est en réalité très similaire à celui obtenu avec le solvant acétonitrile sur la fig. 4.1.1. Au contraire lorsque les sites de soluté sont neutres comme sur la fig. 3.2.11, ils ne peuvent interagir avec le solvant par liaisons hydrogène, dans ce cas l'accord entre MD et MDFT est bon. C'est donc l'interaction par des liaisons hydrogènes de l'eau avec les solutés chargés qui explique le désaccord entre MD et MDFT.

Les liaisons hydrogènes, dans l'eau, provoquent un arrangement tétraédrique local autour des solutés chargés. Cet arrangement particulier, illustré sur la fig. 4.1.3, a deux conséquences sur les fonctions de distribution radiale. Le nombre de premiers voisins est réduit à quatre : le pic dû à la première sphère de solvation est donc plus faible que dans un fluide non-associé. De plus, les molécules de la seconde sphère de solvation peuvent se rapprocher plus du soluté car l'encombrement dû aux molécules de la première couche de solvation est plus faible : le pic dû à la seconde couche de solvation est plus proche soluté.

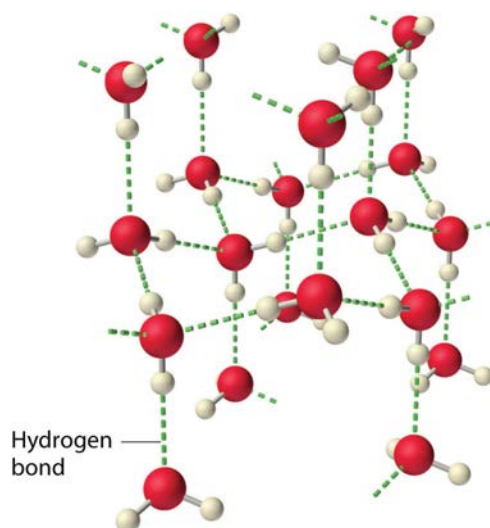


FIGURE 4.1.3.: Illustration de l'arrangement tétraédrique de l'eau liquide, d'après [76].

## 4.2. Terme de correction à 3 corps

Les résultats présentés dans la partie précédente pour la solvation dans l'eau montrent que les fonctionnelles d'excès utilisées ne parviennent pas à reproduire la structure tétraédrique du solvant autour des solutés chargés. Ceci peut en partie s'expliquer par un développement systématique de la fonction de corrélation directe du solvant pur qui a été limité à l'ordre deux. On pourrait imaginer améliorer la fonctionnelle en incluant les corrélations d'ordre supérieur, mais cette idée n'a pas été retenue pour des raisons numériques. De plus, il est compliqué d'obtenir avec précision les termes supérieurs du développement de la fonction de corrélation directe. Une des premières tentatives pour corriger ce problème a été l'utilisation du « bridge » de sphères dures décrit à l'éq. 3.2.15. L'inclusion de ce terme n'a que peu amélioré les résultats, comme on peut le voir sur la fig. 3.2.11. Le seul ion pour lequel l'ajout de ce terme HSB a un effet notable est le cation sodium. Ceci peut s'expliquer par le fait que dû à son importante densité de charge (qu'on peut définir ici comme le rapport entre la charge et le volume de la sphère Lennard-Jones), il attire particulièrement les molécules de solvant dans sa première sphère de solvation. Il en découle un premier pic très intense. L'ajout d'un bridge de sphères dures empêche un empilement très élevé des molécules de solvant, ce qui tend à faire diminuer ce premier pic. Cet effet est déjà moins visible pour le cation potassium et est inexistant pour le cation césium. Cette correction rajoute des termes de corrélation d'ordre supérieur d'un fluide de sphères dures. Ce sont des termes d'empilement isotropes qui ne permettent pas de reproduire un ordre tétraédrique local. Cette correction ne peut donc pas modifier la position du second pic de la fonction de distribution radiale.

En conséquence, nous avons décidé d'introduire un terme correctif qui renforce l'ordre tétraédrique local entre molécules d'eau et solutés chargés. Ce terme correctif s'inspire d'un modèle d'eau gros-grains (MW) développé par Molinero et Moore[77]. C'est une adaptation d'un potentiel pour des fluides tétracoordinés comme le silicium, qui a été originellement proposé par

Stillinger et Weber[78].

Dans ce modèle, l'interaction entre molécules de solvant se fait à l'aide de potentiels d'interaction de deux types : un potentiel Lennard-Jones à deux corps et un potentiel à deux corps qui pénalise les arrangements non tétraédrique de la forme :

$$\phi_3(\mathbf{r}_{ij}, \mathbf{r}_{ik}, \theta_{ijk}) = \lambda_{ijk} \epsilon (\cos \theta_{ijk} - \cos \theta_0)^2 f(\|\mathbf{r}_{ij}\|) f(\|\mathbf{r}_{ik}\|). \quad (4.2.1)$$

C'est un potentiel d'interaction entre des sites positionnés en  $\mathbf{r}_i, \mathbf{r}_j$  et  $\mathbf{r}_k$ , avec  $\theta_{ijk}$  l'angle entre ces trois sites et  $\|\mathbf{r}_{ij}\|$  la distance entre les sites  $i$  et  $j$ . Les fonctions  $f$  sont des fonctions qui limitent la portée de ce potentiel.  $\theta_0 = 109.4^\circ$  désigne l'angle tétraédrique.  $\epsilon$  est la profondeur du potentiel Lennard-Jones et les paramètres  $\lambda_{ijk}$  servent à moduler l'importance relative de ce potentiel d'interaction à trois corps par rapport au potentiel à deux corps. Ce potentiel crée une pénalité énergétique d'autant plus forte que trois molécules de solvant suffisamment proches forment un angle différent de l'angle tétraédrique  $\theta_0$ .

Grâce à ce modèle d'eau simplifié, Molinero et Moore parviennent à reproduire de nombreuses propriétés de l'eau et notamment les propriétés structurales et de densité.

Le terme de correction ajouté à la fonctionnelle pour renforcer l'ordre tétraédrique local inspiré de ce potentiel de Molinero et Moore est donc :

$$\mathcal{F}_{\text{cor}}^{3\text{B}}[n(\mathbf{r})] = \mathcal{F}_{\text{cor}}^{3\text{B}-1\text{S}}[n(\mathbf{r})] + \mathcal{F}_{\text{cor}}^{3\text{B}-2\text{S}}[n(\mathbf{r})] + \mathcal{F}_{\text{cor}}^{3\text{B}-\text{ww}}[n(\mathbf{r})]. \quad (4.2.2)$$

Avec :

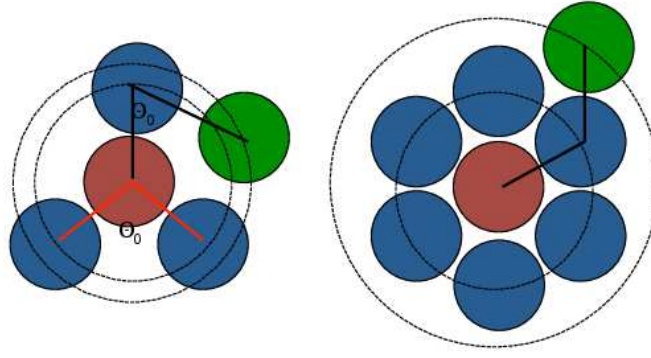
$$\beta \mathcal{F}_{\text{cor}}^{3\text{B}-1\text{S}}[n(\mathbf{r})] = \frac{1}{2} \sum_m \lambda_m^{1\text{S}} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} f_m(r_{m2}) f_m(r_{m3}) \left( \frac{\mathbf{r}_{m2} \cdot \mathbf{r}_{m3}}{r_{m2} r_{m3}} - \cos \theta_0 \right)^2 n(\mathbf{r}_2) n(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3, \quad (4.2.3)$$

$$\beta \mathcal{F}_{\text{cor}}^{3\text{B}-2\text{S}}[n(\mathbf{r})] = \frac{1}{2} \sum_m \lambda_m^{2\text{S}} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} f_m(r_{m2}) f_w(r_{23}) \left( \frac{\mathbf{r}_{m2} \cdot \mathbf{r}_{23}}{r_{m2} r_{23}} - \cos \theta_0 \right)^2 n(\mathbf{r}_2) n(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3, \quad (4.2.4)$$

$$\beta \mathcal{F}_{\text{cor}}^{3\text{B}-\text{ww}}[n(\mathbf{r})] = \frac{1}{2} \lambda_w \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} n(\mathbf{r}_1) d\mathbf{r}_1 \iiint_{\mathbb{R}^3} f_w(r_{12}) f_w(r_{13}) \left( \frac{\mathbf{r}_{12} \cdot \mathbf{r}_{13}}{r_{12} r_{13}} - \cos \theta_0 \right)^2 n(\mathbf{r}_2) n(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3, \quad (4.2.5)$$

où  $n$  désigne la densité de solvant moyennée sur les angles. Les expressions des dérivés fonctionnelles sont données en Appendice D.

La correction se décompose en la somme de trois termes car il existe un arrangement tétraédrique entre les molécules de solvant mais aussi entre le soluté et le solvant. Le terme de l'éq. 4.2.5 correspond à un terme d'interaction entre molécules de solvant, c'est donc le terme équivalent au potentiel du modèle d'eau MW de l'éq. 4.2.1. Le terme de l'éq. 4.2.3 renforce l'arrangement tétraédrique entre les sites des solutés et deux molécules d'eau se trouvant dans la première couche de solvation. Celui de l'éq. 4.2.4 renforce l'ordre tétraédrique entre les sites de solutés



**FIGURE 4.2.1.:** Schéma de la structure de solvation autour d'un soluté (en rouge) dans un fluide associé (à gauche) et non-associé (à droite). Les molécules de la première couche de solvation sont représentées en bleu, celles de la deuxième en vert. L'angle tétraédrique  $\theta_0$  favorisé par l'utilisation de l'éq. 4.2.3 est en rouge, celui favorisé par l'éq. 4.2.4 est en noir.

et deux molécules d'eau, l'une se trouvant dans la première couche de solvation et l'autre dans la seconde. Les arrangements spatiaux favorisés par ces deux derniers termes sont illustrés sur la fig. 4.2.1.

Une sommation est réalisée sur les  $m$  sites du soluté.  $\lambda_m^{1S}$  et  $\lambda_m^{2S}$  permettent, pour chaque site du soluté, de moduler l'importance des termes des première et deuxième couches de solvation. Typiquement, pour un site très chargé les valeurs de  $\lambda_m^{1S}$  et  $\lambda_m^{2S}$  sont élevées, pour un site peu chargé elles sont faibles. Les fonctions  $f_m$  et  $f_w$  sont, ici aussi, des fonctions à courte portée, elles sont définies comme,

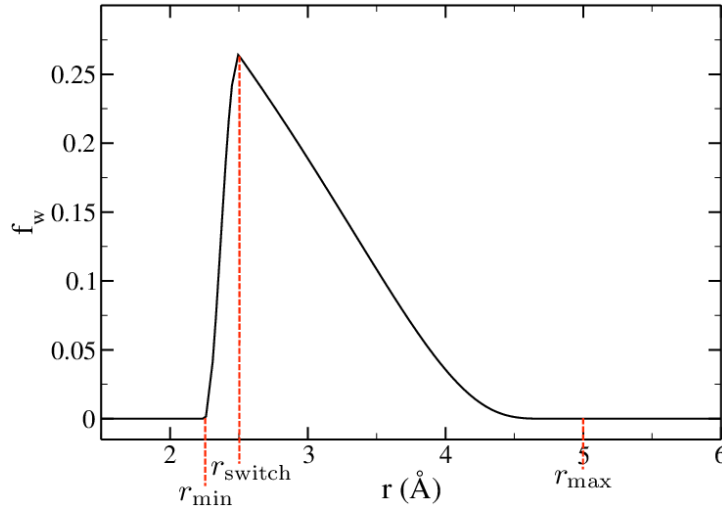
$$f(r; r_{\min}, r_{\max}, r_{\text{swap}}) = S(r; r_{\min}, r_{\text{switch}}) \exp \left[ \frac{2}{3} \frac{r_{\max}}{r - r_{\max}} \right] \text{ si } r < r_{\max} \text{ et } 0 \text{ sinon.} \quad (4.2.6)$$

$S(r; r_{\min}, r_{\text{switch}})$  est une fonction cubique, construite de telle sorte qu'elle soit nulle pour  $r < r_{\min}$  et égale à 1 pour  $r > r_{\text{switch}}$ . Cette fonction a pour but de couper l'interaction à trois corps à l'intérieur des particules, de la rendre maximale en  $r_{\text{switch}}$  et de la faire décroître jusqu'à ce qu'elle soit nulle en  $r > r_{\max}$ . L'allure de la fonction  $f$  est donnée en fig. 4.2.2, avec les paramètres utilisés dans le code.

Les termes relatifs aux interactions entre molécules de solvant et ions ont déjà été proposés dans la ref. [52] mais ces termes étaient évalués dans l'espace direct ce qui rendait leur calcul coûteux. J'ai développé le calcul des différents termes de l'éq. 4.2.2 dans l'espace de Fourier, pour un coût en  $N \log N$  plutôt que  $N^3$ .

Si on introduit les produits de convolution suivants, calculables par le produit de leurs transformées de Fourier,

$$F_{\alpha\beta}(\mathbf{r}_1) = \iiint_{\mathbb{R}^3} f_w(r_{12}) \frac{\alpha_{12} \beta_{12}}{r_{12}^2} n(\mathbf{r}_2) d\mathbf{r}_2, \text{ avec } \alpha \text{ et } \beta = x \text{ ou } y \text{ ou } z, \quad (4.2.7)$$



**FIGURE 4.2.2.:** Allure de la fonction  $f_w$  utilisée dans la correction à trois corps en fonction de la distance  $r$  entre deux sites. Les paramètres utilisés sont  $r_{\min} = 2.25 \text{ \AA}$ ,  $r_{\max} = 5.0 \text{ \AA}$  et  $r_{\text{switch}} = 2.5 \text{ \AA}$ . Cette fonction est maximale pour  $r = r_{\text{switch}}$  qui est choisi à une valeur typique de la distance entre oxygènes de deux molécules d'eau partageant une liaison hydrogène dans l'eau bulk.

$$\mathbf{F}(\mathbf{r}_1) = \iiint_{\mathbb{R}^3} f_w(r_{12}) \frac{\mathbf{r}_{12}}{r_{12}} n(\mathbf{r}_2) d\mathbf{r}_2, \quad (4.2.8)$$

$$F_0(\mathbf{r}_1) = \iiint_{\mathbb{R}^3} f_w(r_{12}) n(\mathbf{r}_2) d\mathbf{r}_2, \quad (4.2.9)$$

l'éq. 4.2.5 se réécrit en fonction de ces produits de convolution,

$$\beta \mathcal{F}_{\text{cor}}^{3\text{B-ww}}[n(\mathbf{r})] = \frac{1}{2} \lambda_w \iiint_{\mathbb{R}^3} n(\mathbf{r}_1) \left[ \sum_{(\alpha, \beta) \in \{x, y, z\}^2} F_{\alpha\beta}(\mathbf{r}_1)^2 + \cos^2 \theta_0 F_0(\mathbf{r}_1)^2 - 2 \cos \theta_0 \mathbf{F}(\mathbf{r}_1) \cdot \mathbf{F}(\mathbf{r}_1) \right] d\mathbf{r}_1. \quad (4.2.10)$$

On définit également plusieurs intégrales permettant de calculer les termes d'interaction à trois corps entre soluté et solvant. On souligne la présence de l'exposant  $m$  qui rappelle que ces produits de convolution sont définis pour chaque site de soluté,

$$\mathbf{H}_{\alpha\beta}^m = \iiint_{\mathbb{R}^3} f_m(r_{m2}) \frac{\alpha_{m2} \beta_{m2}}{r_{m2}^2} n(\mathbf{r}_2) d\mathbf{r}_2, \quad , \text{ avec } \alpha \text{ et } \beta = x \text{ ou } y \text{ ou } z, \quad (4.2.11)$$

$$\mathbf{H}^m = \iiint_{\mathbb{R}^3} f_m(r_{m2}) \frac{\mathbf{r}_{m2}}{r_{m2}} n(\mathbf{r}_2) d\mathbf{r}_2, \quad (4.2.12)$$

$$H_0^m = \iiint_{\mathbb{R}^3} f_w(r_{m2}) n(\mathbf{r}_2) d\mathbf{r}_2. \quad (4.2.13)$$

On calcule ces intégrales dans l'espace direct, sur un petit volume autour des sites  $m$  car les



fonctions  $f$  sont à courte portée.

Le terme de l'éq. 4.2.3, jouant sur la première couche de solvation, s'écrit alors,

$$\beta \mathcal{F}_{\text{cor}}^{3\text{B}-1\text{S}}[n(\mathbf{r})] = \sum_m \lambda_m^{1\text{S}} \left[ \sum_{(\alpha,\beta) \in \{x,y,z\}^2} (\mathbf{H}_{\alpha\beta}^m)^2 + (\cos \theta_0 \mathbf{H}_0^m)^2 - 2 \cos \theta_0 \mathbf{H}^m \cdot \mathbf{H}^m \right]. \quad (4.2.14)$$

Le terme de l'éq. 4.2.4, jouant sur la seconde couche de solvation, peut se réécrire en utilisant les produit de convolution  $F_0$ ,  $\mathbf{F}$  et  $F_{\alpha\beta}$  que l'on vient d'introduire,

$$\begin{aligned} \beta \mathcal{F}_{\text{cor}}^{3\text{B}-2\text{S}}[n(\mathbf{r})] = & \sum_{(\alpha,\beta) \in \{x,y,z\}^2} \iiint_{\mathbb{R}^3} F_{\alpha\beta}(\mathbf{r}_1) G_{\alpha\beta}(\mathbf{r}_1) d\mathbf{r}_1 + \cos^2 \theta_0 \iiint_{\mathbb{R}^3} F_0(\mathbf{r}_1) G_0(\mathbf{r}_1) d\mathbf{r}_1 \\ & - 2 \cos \theta_0 \iiint_{\mathbb{R}^3} \mathbf{F}(\mathbf{r}_1) \cdot \mathbf{G}(\mathbf{r}_1) d\mathbf{r}_1, \end{aligned} \quad (4.2.15)$$

où on a introduit les notations suivantes,

$$G_{\alpha\beta}(\mathbf{r}_1) = \sum_m \lambda_m^{2\text{S}} f_m(r_{m1}) \frac{\alpha_{m1} \beta_{m1}}{r_{m1}^2} n(\mathbf{r}_1), \quad (4.2.16)$$

$$\mathbf{G}(\mathbf{r}_1) = \sum_m \lambda_m^{2\text{S}} f_m(r_{m1}) \frac{\mathbf{r}_{m1}}{r_{m1}} n(\mathbf{r}_1), \quad (4.2.17)$$

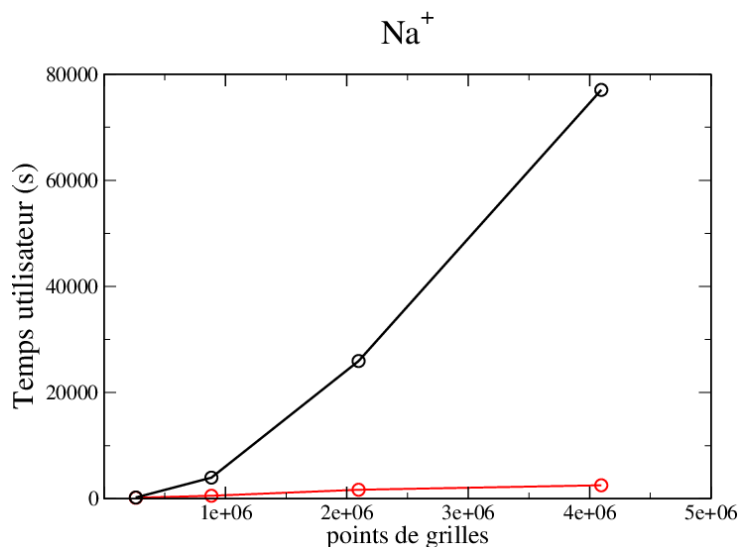
$$G_0(\mathbf{r}_1) = \sum_m \lambda_m^{2\text{S}} f_m(r_{m1}) n(\mathbf{r}_1). \quad (4.2.18)$$

Ces huit fonctions  $G$  sont calculées dans l'espace direct.

On insiste sur le fait qu'au lieu d'avoir à calculer des doubles intégrales sur le volume dans l'espace direct, on se ramène au calcul des dix produits de convolution  $F_0$ ,  $F_{\alpha\beta}$  et  $\mathbf{F}$ , soit vingt transformées de Fourier (directes et inverses). Il suffit ensuite de calculer des intégrales simples sur le volume. Sur la fig. 4.2.3, on compare l'efficacité des deux routines pour le cation alcalin sodium. L'une calcule ce terme à trois corps dans l'espace direct, l'autre en utilisant le nouvel algorithme. On constate que l'utilisation de transformées de Fourier réduit considérablement le temps de calcul du terme à trois corps et, a fortiori, le temps de calcul global de la minimisation fonctionnelle. La formulation utilisant les transformées de Fourier introduite est donc parfaitement adaptée à l'étude d'un soluté possédant un ou plusieurs sites chargés.

### 4.3. Application à la solvation

Dans l'éq. 4.2.3 et l'éq. 4.2.4, les paramètres  $\lambda_m^{1\text{S}}$  et  $\lambda_m^{2\text{S}}$  permettent de moduler l'importance du terme de correction à trois corps. Lorsque les sites de solvant ne sont pas des sites susceptibles d'être engagés dans une liaison hydrogène, ces deux paramètres sont choisis nuls :  $\lambda_m^{1\text{S}} = \lambda_m^{2\text{S}} = 0$ . Il n'y a donc aucune modification des résultats structuraux ou énergétiques pour ces solutés,



**FIGURE 4.2.3.:** Temps de calcul utilisateur nécessaire à la minimisation d'une fonctionnelle utilisant le terme à trois corps, soit calculé dans l'espace direct (en noir) soit à l'aide de transformées des Fourier discrètes (en rouge), dans le cas de  $\text{Na}^+$ . On compare les résultats pour une boîte cubique d'arête 25 Å, avec un maillage de  $64^3$ ,  $96^3$ ,  $128^3$ ,  $160^3$  points.

notamment pour les alcanes et les gaz rares présentés dans la sec.1.1.1. Pour les sites chargés, il est évident qu'il faudrait choisir ces termes en fonction de la force des liaisons hydrogènes mettant en jeu le site considéré.

Cependant, ce travail de thèse a pour but de prouver que le principe de l'utilisation de la théorie de la fonctionnelle de la densité au cas de l'eau est viable. L'optimisation de tous les paramètres mis en jeu, notamment dans ce cas particulier, est repoussée à plus tard, quand on sera satisfait de la fonctionnelle utilisée.

Par simplicité, et pour ne pas faire un ajustement des paramètres sur les résultats que l'on cherche à obtenir, l'hypothèse simplificatrice suivante a été faite : pour les sites ayant une charge entière, typiquement les ions monovalents, on choisit  $k_B T \lambda_m^{1S} = k_B T \lambda_m^{2S} = 100 \text{ kJ.mol}^{-1}$ . Pour les sites impliqués dans des liaisons hydrogènes portant des charges partielles (par exemple l'oxygène d'une molécule d'eau), on choisit  $k_B T \lambda_m^{1S} = k_B T \lambda_m^{2S} = 75 \text{ kJ.mol}^{-1}$ . Le choix des valeurs pour les ions a été fait pour obtenir le meilleur profil de distribution radiale possible pour l'ion chlorure tandis que les paramètres pour les molécules neutres permettent d'obtenir le meilleur profil de distribution radiale possible pour l'eau.

### 4.3.1. Structure

#### 4.3.1.1. Les ions

On donne les fonctions de distribution radiale obtenues par dynamique moléculaire et par minimisation fonctionnelle pour des cations alcalins sur la fig.4.3.1, et pour des halogénures sur la fig.4.3.2. On utilise la fonctionnelle multipolaire et la correction à trois corps. Ces résultats sont

donc à comparer avec la fig. 3.3.4 pour les alcalins. Pour les halogénures, on peut comparer avec la fig. 3.2.8 puisque l'utilisation de la fonctionnelle multipolaire modifie peu les fonctions de distribution radiale. On constate que l'introduction de ce terme à trois corps améliore notablement l'allure des fonctions de distribution radiale. La hauteur du premier pic est désormais correcte, la troisième couche de solvation est très bien reproduite en position et en intensité. Le second pic, dû à la seconde couche de solvation est également amélioré même si on surestime toujours légèrement l'intensité pour des distances juste inférieures au maximum. On peut donc conclure que c'est bien l'arrangement local tétraédrique des molécules d'eau autour des solutés chargés qui est responsable de l'allure caractéristique de leurs fonctions de distribution radiale.

L'écart par rapport à la MD pour le second pic empire à mesure que l'on descend dans le tableau périodique mais les paramètres  $\lambda_m^{1S}$  et  $\lambda_m^{2S}$  ont été choisis pour le sodium et le chlorure. Ces ions se lient fortement avec l'eau par liaisons hydrogènes. Quand on descend dans le tableau périodique, les rayons ioniques augmentent et la densité de charge diminue. Le potassium et le bromure et à fortiori l'iodure et le césium, vont donc faire des liaisons hydrogènes moins fortes avec l'eau. On a vérifié que des paramètres  $\lambda_m^{1S}$  et  $\lambda_m^{2S}$  plus faibles sont mieux adaptés pour ces ions plus gros. À ce stade, il est cependant satisfaisant de pouvoir utiliser les mêmes paramètres pour tous les ions monovalents.

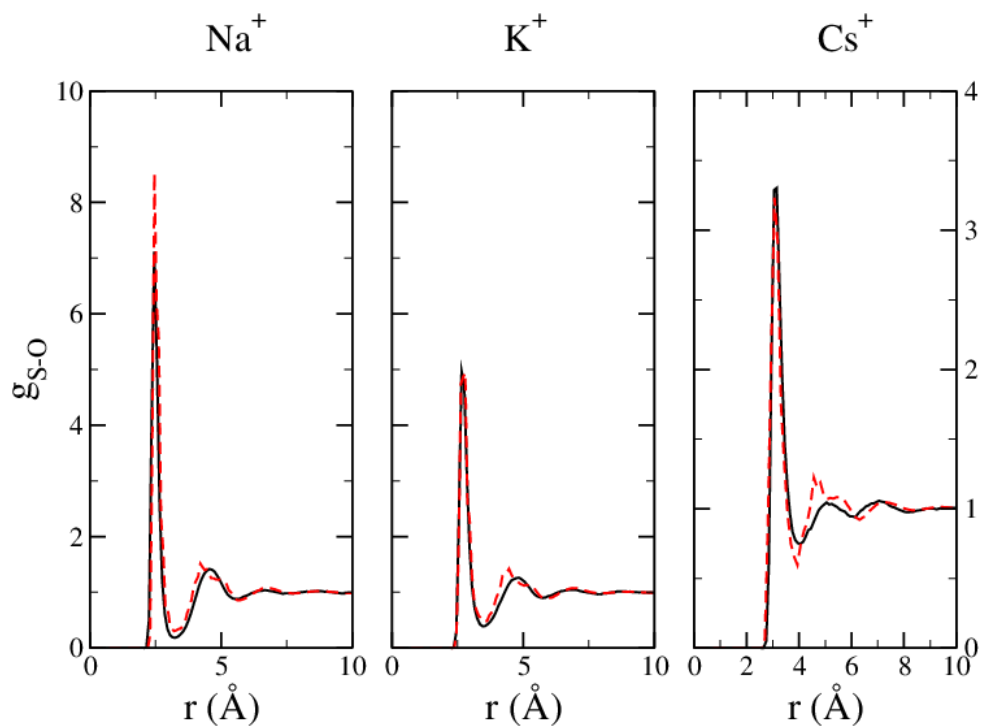
En trouvant une façon de choisir ces paramètres à priori en fonction de la force des liaisons hydrogènes mises en jeu, on peut s'attendre à reproduire correctement l'évolution de l'allure des fonctions de distribution radiale dans un groupe du tableau périodique.

### 4.3.1.2. Les solutés neutres

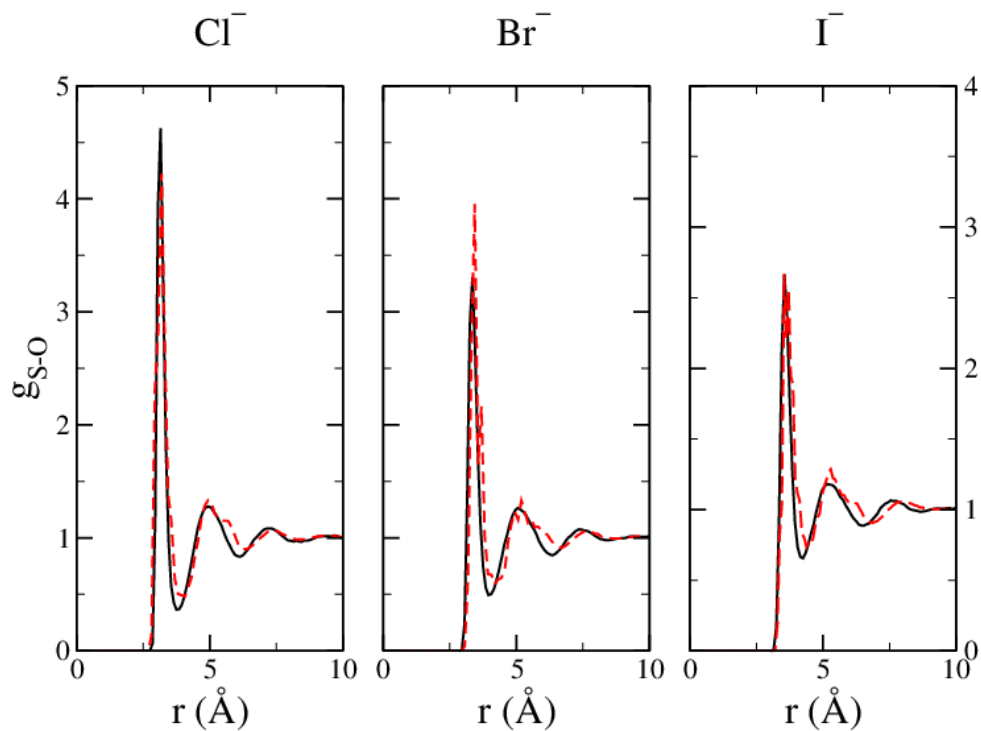
En ce qui concerne les solutés neutres, on a vu en sec. 3.2.2.2 que les profils de solvation des sites peu chargés sont bien reproduits par minimisation de la fonctionnelle avec un traitement multipolaire de la polarisation. Ce n'est pas le cas pour les sites chargés.

On observe sur la fig. 4.3.3 que l'ajout du terme correctif à trois corps améliore considérablement l'accord entre les fonctions de distribution radiale obtenues par dynamique moléculaire et celles obtenues avec le code mdft pour le soluté eau SPC/E. La première couche de solvation de l'eau est responsable des deux premiers pics de la fonction de distribution radiale entre l'hydrogène du soluté et l'oxygène du solvant, et du premier pic de celle entre l'oxygène du soluté et celui du solvant. Ces pics sont désormais très bien reproduits par minimisation fonctionnelle. Leurs positions et leurs hauteurs sont désormais semblables pour les deux méthodes, MD et MDFT. On surestime toujours, mais moins, la déplétion entre la première et la deuxième couche de solvation pour  $g_{OO}$ .

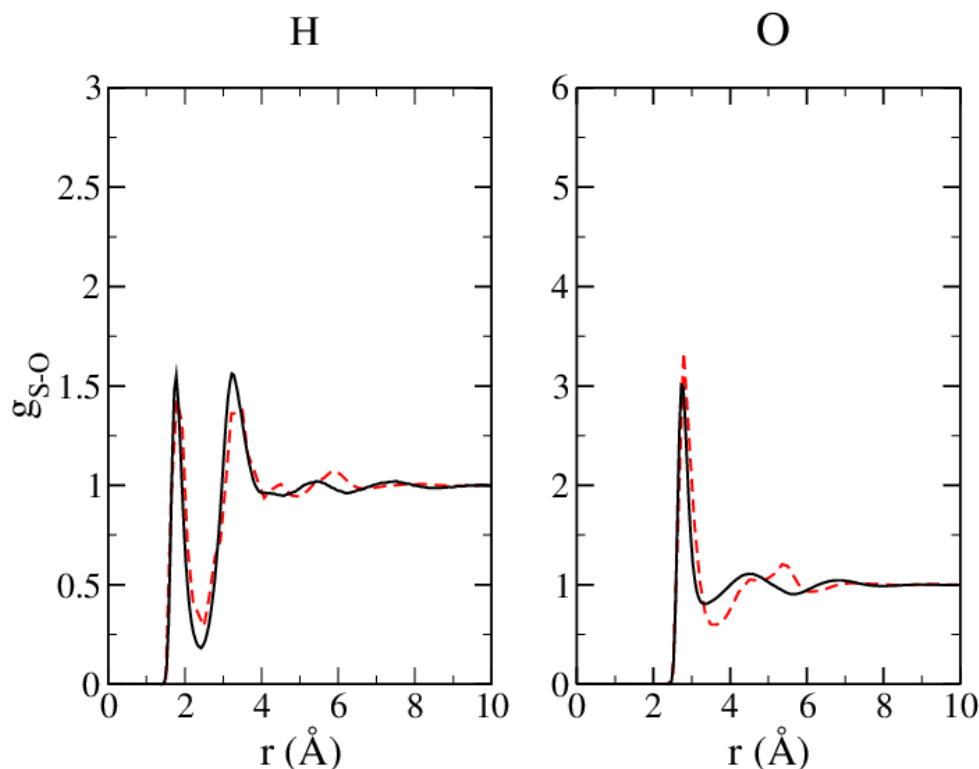
Dans le cas de  $g_{OO}$  tout comme dans celui des ions, on constate une légère amélioration de la prédiction de la seconde couche de solvation, avec une réduction du pic situé vers 5.5 Å. Ce pic était situé trop loin de l'atome d'oxygène du soluté. Un nouveau pic apparaît vers 4.5 Å, en accord avec la MD. L'effet du terme à trois corps est clair : il renforce l'ordre tétraédrique, réduit la distance entre soluté et molécules d'eau de la deuxième couche de solvation, augmentant ainsi



**FIGURE 4.3.1.:** Fonctions de distribution radiale entre l'oxygène de l'eau et trois cations alcalins, sodium, potassium et césium. Les résultats des simulations MD sont en trait noir plein et ceux des calculs MDFT, incluant la correction trois corps, en tirets rouges.



**FIGURE 4.3.2.:** Fonctions de distribution radiale entre le solvant et trois anions halogénures, chlorure, bromure et iodure. La légende est la même que sur la fig. 4.3.1.



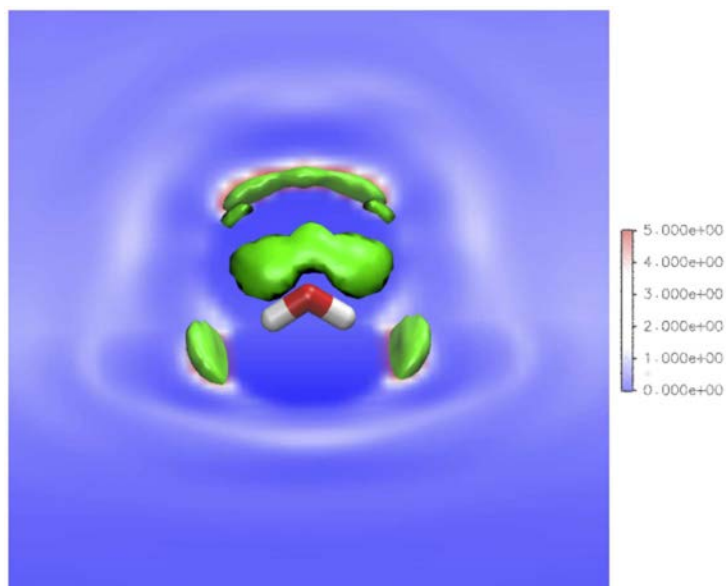
**FIGURE 4.3.3.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et l'oxygène et l'hydrogène de l'eau soluté. MD en traits noirs pleins et MDFT incluant le terme correctif à trois corps en traits rouges discontinus.

la densité pour des zones de l'espace situées vers 4.5 Å. C'est ce phénomène qui est responsable de ce nouveau pic. La répulsion entre molécules d'eau a pour conséquence de réduire les densités au voisinage de cette zone de plus haute densité à 4.5 Å. En conséquence, la densité diminue là où se situaient les zones de hautes densités dues à la seconde couche de solvation précédemment, et le pic situé vers 5.5 Å diminue.

Le troisième pic, dû la troisième couche de solvation est lui aussi grandement amélioré par l'introduction du terme à trois corps. C'est une conséquence directe de la modification des zones de haute densité de la seconde couche de solvation. Ces progrès pour les deuxième et troisième couches de solvation sont moins visibles sur la fonction de distribution radiale entre l'oxygène du solvant et l'hydrogène du soluté.

Sur la fig. 4.3.4, on représente des cartes de densité autour du soluté eau, ainsi que des isosurfaces de haute densité,  $(n(\mathbf{r})/n_b)=3$ . On constate que les régions de l'espace où on s'attend à observer des liaisons hydrogènes coïncident avec les zones de haute densité. On remarque aussi que les régions qui jouxtent ces zones de haute densité subissent une déplétion. De telles cartes de densité sont compliquées à obtenir par MD ou MC car elles requièrent de faire des histogrammes de densité. Elles sont au contraire obtenues directement par le code mdft, la densité étant la variable naturelle de la théorie.

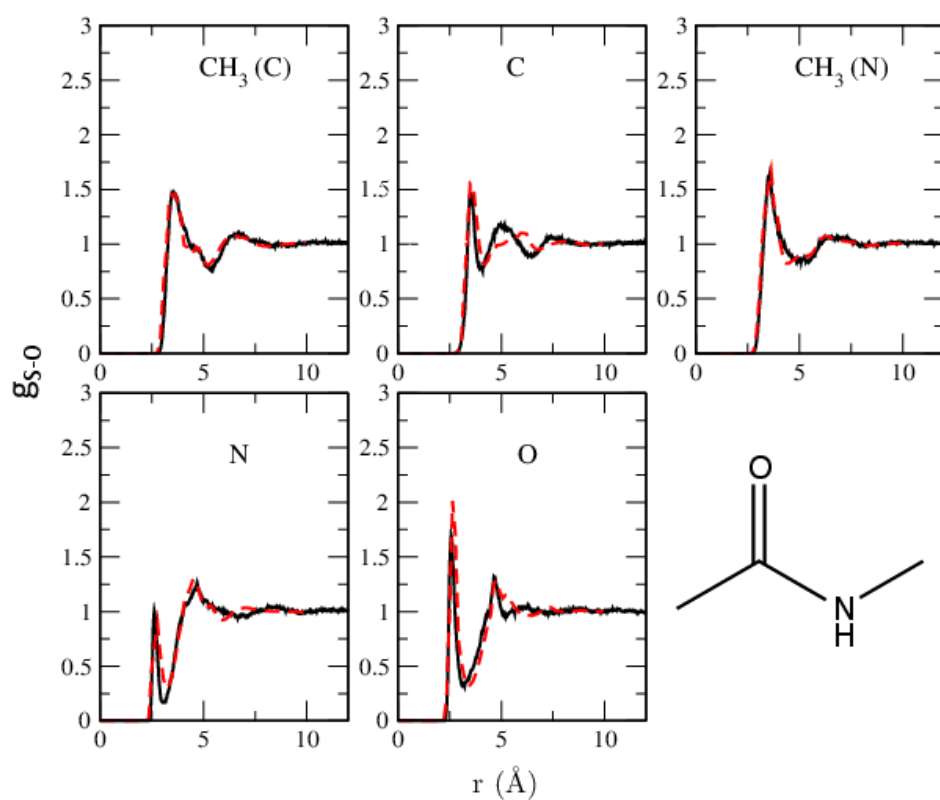
Cette amélioration réalisée grâce à l'inclusion du terme à trois corps est confirmée par les fonc-



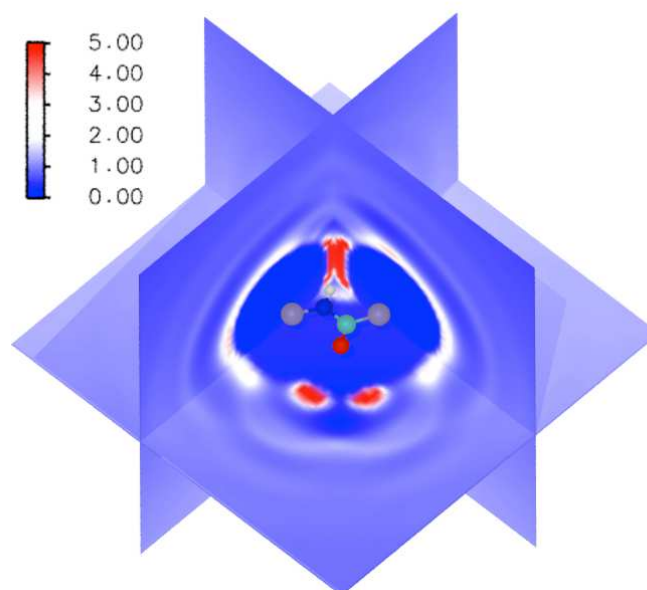
**FIGURE 4.3.4.:** Carte de densité autour de la molécule d'eau obtenue en incluant le terme correctif à trois corps. Des isosurfaces de haute densité sont représentées en vert ( $n(\mathbf{r})/n_b=3$ ) pour aider à la visualisation.

tions de distribution radiale obtenues pour la N-méthylacétamide présentées en fig. 4.3.5. Les pics des première et deuxième couches de solvation des sites impliqués dans des liaisons hydrogènes, l'oxygène et l'azote, sont améliorés en terme de hauteur et de position. Les fonctions de distribution radiale des sites ne formant pas de liaisons hydrogènes sont elles-aussi considérablement améliorées. Ceci s'explique par le fait que les modèles de molécule utilisés sont rigides. En conséquence, la modification de l'environnement proche d'un site modifie directement l'allure de la fonction de distribution radiale calculée depuis un autre site.

L'utilisation du terme à trois corps permet de renforcer la coordination tétraédrique sur les sites engagés dans des liaisons hydrogènes sur la fig. 4.3.6. Des zones de plus haute densité, en rouge, apparaissent clairement autour du groupe carbonyle et de la liaison azote-hydrogène de la molécule NMA.



**FIGURE 4.3.5.:** Fonctions de distribution radiale entre l'oxygène de l'eau solvant et les différents sites de la molécule NMA. La légende est la même que sur la figure fig. 4.3.3.



**FIGURE 4.3.6.:** Densité 3D du solvant eau autour de la molécule fixe de NMA obtenue par MDFT.

L'inclusion d'un terme *ad-hoc* qui renforce l'ordre tétraédrique entre un site donneur ou accepteur de liaisons hydrogènes et les molécules d'eau du solvant améliore considérablement les structures de solvation obtenues par minimisation fonctionnelle des solutés polaires. En réalité ce terme modifie le potentiel d'interaction entre soluté et solvant et peut être considéré comme un terme extérieur plutôt que d'excès. *Stricto sensu* les potentiels d'interaction utilisés pour les calculs MD et MDFT ne sont alors plus les mêmes.



## 5. Hydrophobicité et couplage multi-échelle

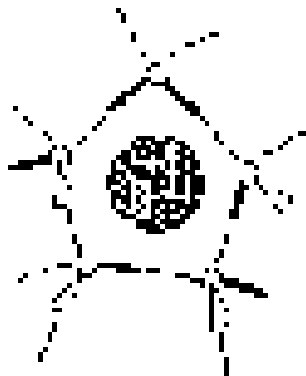
La théorie de la fonctionnelle de la densité classique MDFT présentée dans les chapitres précédents est construite à un niveau de description moléculaire. Elle donne des résultats sensiblement comparables à ceux des simulations moléculaires pour des petits solutés hydrophobes. Les propriétés de solvation dans l'eau des petits solutés hydrophobes (quelques angströms) sont très différentes de celles des solutés de taille plus importante (nanométrique)[79]. Cette différence de comportement s'explique par une modification du phénomène physique gouvernant la mise en solution lorsque la taille du soluté augmente. On va d'abord décrire les propriétés de solvation de solutés hydrophobes à ces deux échelles, avant de proposer un moyen d'introduire dans la fonctionnelle la solvation des solutés hydrophobes aux échelles microscopiques et mésoscopiques. Cette modification de la fonctionnelle constitue une illustration de la possibilité d'utiliser la MDFT pour étudier des problèmes multi-échelles. L'objectif est de pouvoir coupler la MDFT, qui permet une description rapide du solvant au niveau moléculaire, avec des théories mésoscopiques utilisant aussi la densité comme l'hydrodynamique.

### 5.1. La solvation des solutés hydrophobes aux échelles microscopique et mésoscopique

On nomme hydrophobes les espèces peu solubles dans l'eau. Les solutés hydrophobes apolaires et aprotiques ne peuvent pas créer de liaisons hydrogènes avec les molécules d'eau. De ce fait, la réorganisation des molécules de solvant autour des solutés vise à conserver au maximum le nombre de liaisons hydrogènes entre molécules d'eau.

Comme schématisé sur la fig. 5.1.1, l'ordre spatial des molécules d'eau autour des petits solutés permet de ne pas perdre de liaisons hydrogènes entre molécules de solvant. Le coût de cette réorganisation est principalement entropique et explique la faible solubilité des petits solutés apolaires dans l'eau. On peut montrer que l'énergie libre de solvation d'un tel soluté hydrophobe est proportionnel au volume exclu, c'est-à-dire au volume effectif du soluté[80, 79].

La mise en solution des solutés de plus grande taille peut causer un démouillage, c'est-à-dire qu'il peut se former à la surface du soluté une interface liquide-gaz[12]. Les molécules d'eau en surface du soluté perdent une partie de leurs liaisons hydrogènes, le coût énergétique de la solvation est essentiellement enthalpique. On peut montrer que l'énergie de solvation de ces grands solutés est proportionnelle à leur surface.



**FIGURE 5.1.1.:** Schéma de la structure des molécules d'eau autour d'un petit soluté hydrophobe. Les tirets indiquent des liaisons hydrogènes (figure extraite de [79]).

Dans la référence [81], Chandler et ses collaborateurs illustrent ce changement de comportement en calculant l'énergie de solvation par unité surface d'un soluté sphère dure dans de l'eau SPC/E, ces résultats sont donnés en fig. 5.1.2. Pour des solutés de faible rayon, l'énergie libre de solvation est proportionnelle au volume du soluté. Pour des solutés plus gros l'énergie libre de solvation est proportionnelle à la surface du soluté.

La solvation des solutés hydrophobes est gouvernée par des phénomènes physiques différents à des échelles différentes. Nous allons voir que si le comportement aux grandes échelles n'est pas inclus à priori dans le formalisme MDFT développé précédemment, il est néanmoins possible de l'inclure par une approche multi-échelle.

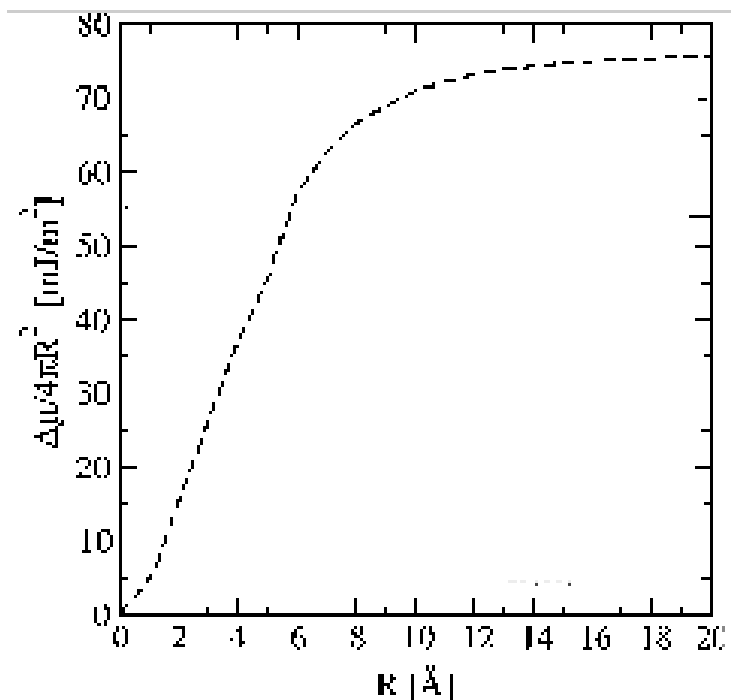
## 5.2. La théorie MDFT/HRF avec HSB

On s'intéresse à deux systèmes hydrophobes déjà étudiés par simulation numérique (MD ou MC). Chandler, dans les réf. [81, 82], a étudié la solvation d'une sphère dure dans l'eau SPC/E, tandis que Hansen et Dzubiella dans la ref. [83] ont étudié la solvation d'une sphère molle purement répulsive. Cette sphère interagit avec le solvant avec le potentiel répulsif de l'éq. 6.2.9, en  $r^{-12}$ .

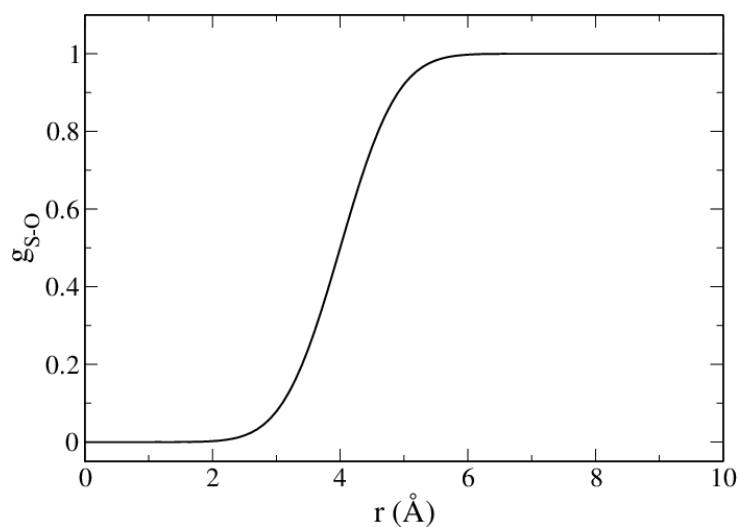
Ces deux systèmes ont été étudiés avec la même fonctionnelle que celle utilisée pour les alcanes dans la sec. 1.1.1, avec le terme de bridge de sphères dures. Dans la fig. 5.2.2, on compare les fonctions de distribution radiale obtenues par MDFT à celles obtenues par MD pour ces deux systèmes, en fonction de la taille des sphères.

En l'absence de démouillage les fonctions de distribution radiale ont une allure similaire à celles rencontrées pour les gaz rares. Lorsqu'il y a démouillage, les fonctions de distribution radiale ont une allure sigmoïdale, voir la fig. 5.2.1.

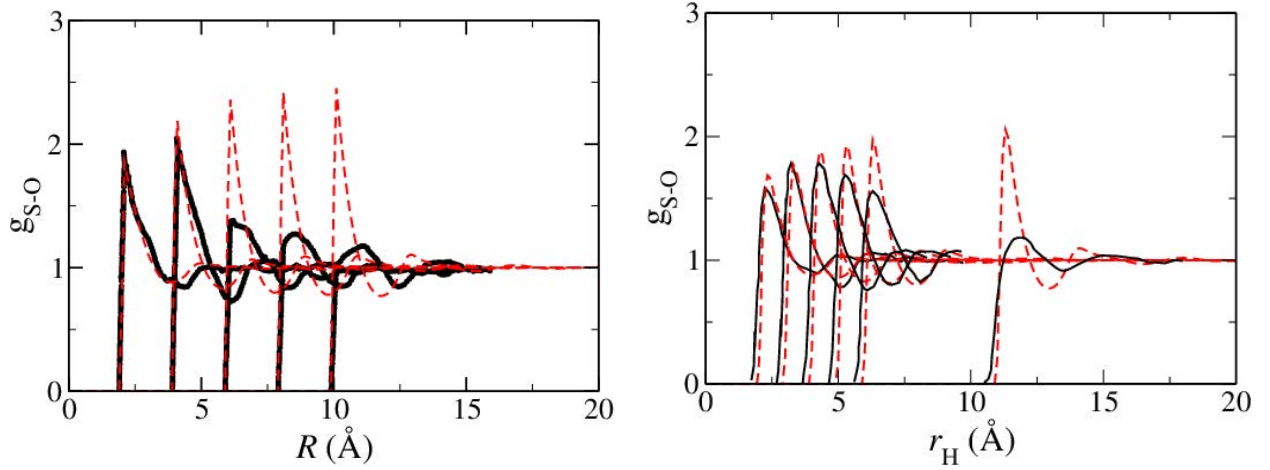
Le changement de régime apparaît clairement sur les fonctions de distribution radiale obtenues par simulation numérique pour les deux systèmes étudiés. On observe d'abord une augmentation de la hauteur du pic de la première couche de solvation pour des rayons inférieurs à 5 Å, puis



**FIGURE 5.1.2.:** Energie libre de solvation par unité de surface, notée  $\Delta\mu$  pour un soluté sphère dure de rayon  $R$ , à 298 K dans l'eau SPC/E d'après [81].



**FIGURE 5.2.1.:** Allure de la fonction de distribution radiale entre soluté et solvant dans le cas d'un démouillage.



**FIGURE 5.2.2.:** Comparaison des fonctions de distribution radiale :

- entre un soluté sphère dure de rayon  $R$  et l'eau à gauche ;
- entre un soluté sphère molle de rayon  $r_H$  et l'eau à droite.

Les résultats obtenus par MDFT avec le bridge de sphères dures sont en tirets rouges, ceux obtenus par simulation Monte-Carlo[81] ou par dynamique moléculaire[83] sont en noir.

la hauteur de ce pic diminue pour des rayons supérieurs. Dans les deux cas, les fonctions de distribution radiale obtenues par minimisation fonctionnelle ne présentent pas ce changement de comportement : les pics de la première couche de solvation augmentent avec le rayon de la sphère pour toutes les valeurs considérées.

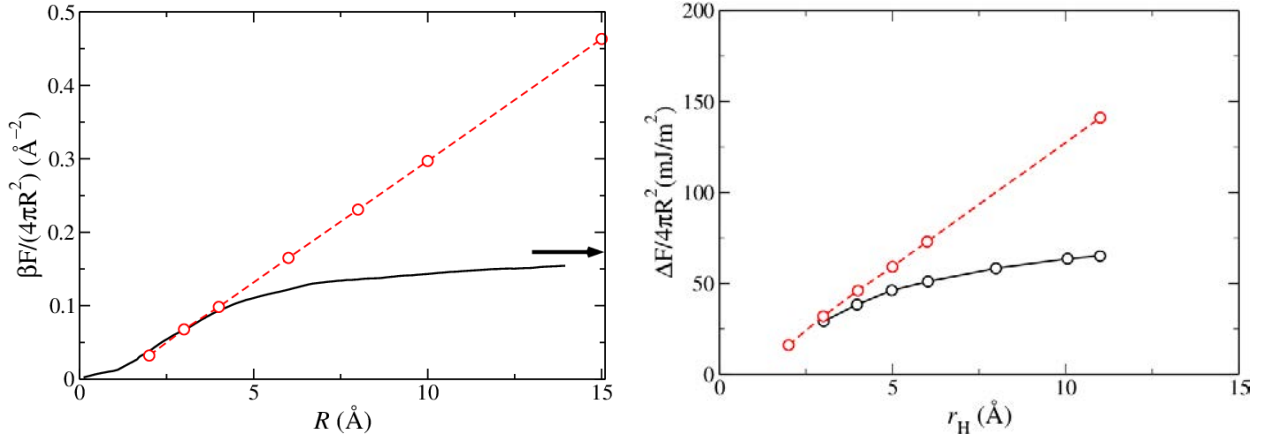
Les énergies libres de solvation sont présentées en fig. 5.2.3. On observe à nouveau un changement de régime sur les courbes obtenues par simulations numériques. Avant 5 Å l'énergie de solvation est proportionnelle au volume. Pour des rayons plus grands la pente diminue pour finalement tendre vers un palier égal à la tension de surface liquide-gaz. L'énergie libre est alors proportionnelle à la surface. La limite indiquée par la flèche noire correspond à la valeur théorique de la tension de surface liquide-gaz de l'eau. Les énergies libres obtenues par MDFT sont en revanche proportionnelles au volume sur toute la plage des rayons considérés.

Il apparaît clair que la théorie de la fonctionnelle de la densité utilisée est adaptée à l'étude de solutés hydrophobes de petite taille mais ne permet pas de reproduire le comportement à échelle mésoscopique de ces solutés.

## 5.3. Description de l'hydrophobicité à différentes échelles dans MDFT

### 5.3.1. Théorie

On introduit ici une nouvelle physique dans la fonctionnelle qui permet de reproduire les propriétés structurale et énergétique de solvation des solutés hydrophobes[84] de dimensions microscopique et mésoscopique.



**FIGURE 5.2.3.:** Energie libre de solvatation par unité de surface pour les systèmes sphères dures (à gauche) et sphères molles (à droite). Les résultats obtenus avec le code MDFT sont les cercles rouges, ceux obtenus par MC et MD sont en noir.

Hughes a proposé [85] une théorie de la fonctionnelle de la densité classique basée sur la fonctionnelle sphères dures dans l'approche FMT (fundamental measure theory) [69] additionnée d'interactions attractives basées sur la théorie statistique des fluides associés (SAFT). SAFT est une équation d'état construite pour décrire les fluides associés. Par rapport à cela, notre théorie à l'avantage d'inclure en plus le traitement de la polarisation et donc l'interaction avec des solutés chargés.

On part de la fonctionnelle multipolaire décrite dans la sec. 3.3. Comme les solutés étudiés ici sont hydrophobes et que l'on néglige un éventuel couplage entre polarisation et densité, les champs de polarisation longitudinale et transverse sont nuls. La fonctionnelle se réduit alors à,

$$\begin{aligned} \mathcal{F}_{\text{exc}}[n(\mathbf{r})] = & -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r})^2 d\mathbf{r} + \mathcal{F}_{\text{cor}}[n(\mathbf{r})] \\ & + \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} S^{-1}(n_b; \|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \end{aligned} \quad (5.3.1)$$

ou bien,

$$\mathcal{F}_{\text{exc}}[n(\mathbf{r})] = -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}(n_b; \|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}' + \mathcal{F}_{\text{cor}}[n(\mathbf{r})], \quad (5.3.2)$$

où on rappelle que  $\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_b$ . On retrouve les termes quadratiques qui proviennent du développement autour du fluide de référence, que l'on peut écrire, de manière équivalente, en fonction de l'inverse du facteur de structure, ou de la fonction de corrélation directe de l'eau<sup>1</sup>. Tous les termes d'ordre supérieur à deux sont rassemblés dans la fonctionnelle inconnue  $\mathcal{F}_{\text{cor}}$ . On approxime ce terme, comme dans d'autres articles [64, 65, 66, 67, 68] et comme dans le chapitre 3,

1. Facteur de structure et fonction de corrélation sont liés dans l'espace de Fourier en l'absence de polarisation,  $n_b \hat{c}_{000}(n_b; k) = 1 - \hat{S}^{-1}(n_b; k)$

par un bridge pour un fluide de référence sphères dures de rayon  $R_0$  et de même densité  $n_b$  que le fluide étudié.

$$\begin{aligned}\mathcal{F}_{\text{cor}}[n(\mathbf{r})] &= \mathcal{F}_{\text{exc}}^{\text{HS}}[n(\mathbf{r})] - \mathcal{F}_{\text{exc}}^{\text{HS}}[n_b] - \mu_{\text{exc}}^{\text{HS}} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) d\mathbf{r} \\ &\quad + \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}^{\text{HS}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}'.\end{aligned}\quad (5.3.3)$$

On rappelle que ce terme revient à remplacer les termes de corrélation d'ordre supérieur ou égaux à trois par ceux d'un fluide de sphères dures.

Puisque l'on veut décrire des phénomènes physiques à deux échelles différentes, microscopique et mésoscopique, on veut pouvoir séparer la fonctionnelle en une partie courte-distance et une partie longue-distance. Pour cela, on définit une densité gros grains  $\bar{n}(\mathbf{r})$ ,

$$\bar{n}(\mathbf{r}) = \iiint_{\mathbb{R}^3} G(\|\mathbf{r} - \mathbf{r}'\|) n(\mathbf{r}') d\mathbf{r}', \quad (5.3.4)$$

où  $G$  est une fonction de convolution.

Le produit de convolution de l'éq. 5.3.4 se réécrit plus simplement dans l'espace de Fourier,

$$\bar{\hat{n}}(\mathbf{k}) = \hat{G}(k) \hat{n}(\mathbf{k}). \quad (5.3.5)$$

Pour simplifier l'explication qui suit supposons que cette fonction  $G$  est la fonction de Heaviside  $\Theta(k_c - k)$ ; c'est-à-dire une fonction égale à 1 si  $k < k_c$  et 0 sinon;  $k_c^{-1}$  est de l'ordre de quelques dimensions caractéristiques de la molécule d'eau, c'est à dire quelques angströms.

Dans ce cas,

$$\bar{\hat{n}}(\mathbf{k}) (\hat{n}(\mathbf{k}) - \bar{\hat{n}}(\mathbf{k})) = 0, \quad \forall \mathbf{k}. \quad (5.3.6)$$

On peut alors rassembler les termes quadratiques en  $\Delta n$  des éq. 5.3.2 et éq. 5.3.3 en un seul terme définissant une énergie libre attractive.

$$\mathcal{F}_{\text{exc}}^{\text{VdW}}[n(\mathbf{r})] = -\frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}^{\text{VdW}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) \Delta n(\mathbf{r}) \Delta n(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (5.3.7)$$

avec  $c_{000}^{\text{VdW}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) = c_{000}(n_b; \|\mathbf{r} - \mathbf{r}'\|) - c_{000}^{\text{HS}}(n_b; \|\mathbf{r} - \mathbf{r}'\|)$ . On a alors un terme qui rappelle l'équation de Van-der-Waals puisque le fluide est traité comme un fluide de sphères dures auquel on ajoute ce terme globalement attractif, c'est pourquoi on appellera par la suite la fonctionnelle utilisée ici correction de Van-der-Waals.

En utilisant l'éq. 5.3.6, on peut décomposer cette fonctionnelle en une partie courte portée et une

partie longue portée :

$$\begin{aligned}\mathcal{F}_{\text{exc}}^{\text{VdW}}[n(\mathbf{r})] &= - \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}^{\text{VdW}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) (n(\mathbf{r}) - \bar{n}(\mathbf{r})) (n(\mathbf{r}') - \bar{n}(\mathbf{r}')) d\mathbf{r} d\mathbf{r}' \\ &\quad - \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}^{\text{VdW}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) \Delta \bar{n}(\mathbf{r}) \Delta \bar{n}(\mathbf{r}') d\mathbf{r} d\mathbf{r}'.\end{aligned}\quad (5.3.8)$$

Comme la fonction  $c_{000}^{\text{VdW}}$  est à courte portée et que la densité gros grains  $\bar{n}$  varie lentement, la seconde intégrale, à longue portée, de l'éq. 5.3.8 peut s'exprimer dans l'espace de Fourier comme,

$$\begin{aligned}&\iiint_{\mathbb{R}^3} c_{000}^{\text{VdW}}(n_b; k) \Delta \bar{n}(\mathbf{k}) \Delta \bar{n}(-\mathbf{k}) d\mathbf{k} = \\ &\iiint_{\mathbb{R}^3} \left( c_{000}^{\text{VdW}}(k=0) + k^2 \frac{d^2 c_{000}^{\text{VdW}}}{dk^2} \Big|_{k=0} + \dots \right) \Delta \bar{n}(\mathbf{k}) \Delta \bar{n}(-\mathbf{k}) d\mathbf{k} = \\ &a \iiint_{\mathbb{R}^3} \Delta \bar{n}(\mathbf{k}) \Delta \bar{n}(-\mathbf{k}) d\mathbf{k} - m \iiint_{\mathbb{R}^3} k^2 \Delta \bar{n}(\mathbf{k}) \Delta \bar{n}(-\mathbf{k}) d\mathbf{k} + \dots = \\ &a \iiint_{\mathbb{R}^3} \Delta \bar{n}(\mathbf{r})^2 d\mathbf{r} - m \iiint_{\mathbb{R}^3} (\nabla \bar{n}(\mathbf{r}))^2 d\mathbf{r} + \dots,\end{aligned}\quad (5.3.9)$$

où  $a$  et  $m$  sont des réels positifs. Le terme en gradient est un terme similaire à celui de l'équation de Cahn-Hilliard. Il pénalise la création d'inhomogénéité, en particulier d'interface.

On peut vérifier le signe de ces paramètres sur la figure fig. 5.3.1 où sont représentées les fonctions de corrélation pour l'eau et le fluide de sphères dures ainsi que la fonction  $c_{000}^{\text{VdW}}$ . Bien que ces deux termes soient fixés par la théorie FMT et par la fonction de corrélation directe de l'eau utilisée comme input, on considérera ci-après que ce sont deux paramètres phénoménologiques que l'on s'autorise à faire varier.

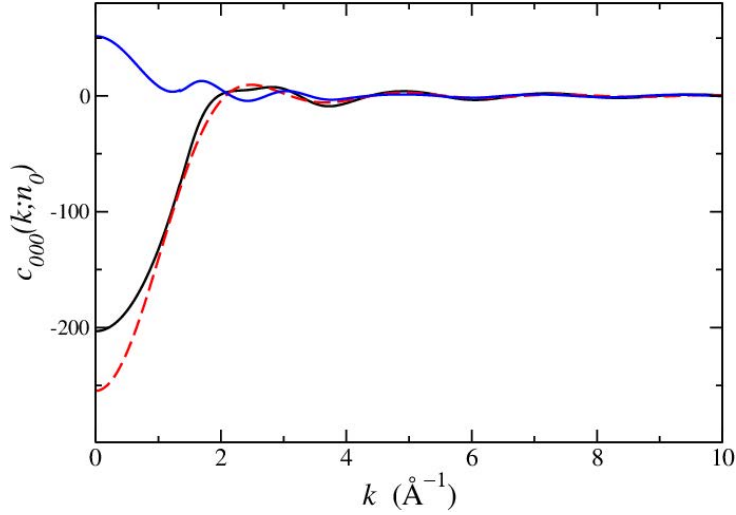
En rassemblant les développements réalisés précédemment on aboutit à la fonctionnelle suivante

$$\begin{aligned}\mathcal{F}_{\text{exc}}^{\text{VdW}}[n(\mathbf{r})] &= - \frac{k_B T}{2} \left[ a \iiint_{\mathbb{R}^3} \Delta \bar{n}(\mathbf{r})^2 d\mathbf{r} - m \iiint_{\mathbb{R}^3} (\nabla \bar{n}(\mathbf{r}))^2 d\mathbf{r} \right] \\ &\quad + \frac{k_B T}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} c_{000}^{\text{VdW}}(n_b; \|\mathbf{r} - \mathbf{r}'\|) (n(\mathbf{r}) - \bar{n}(\mathbf{r})) (n(\mathbf{r}') - \bar{n}(\mathbf{r}')) d\mathbf{r} d\mathbf{r}' \\ &\quad + \mathcal{F}_{\text{exc}}^{\text{HS}}[n(\mathbf{r})] - \mathcal{F}_{\text{exc}}^{\text{HS}}[n_b] - \mu_{\text{exc}}^{\text{HS}} \iiint_{\mathbb{R}^3} \Delta n(\mathbf{r}) d\mathbf{r}.\end{aligned}\quad (5.3.10)$$

Pour la fonctionnelle sphères dures on utilise la fonctionnelle FMT scalaire de Kierlik et Rosinberg (KR-FMT)[70, 71] décrite dans l'Appendice C.

Lorsque  $n = \bar{n}$ , le terme courte portée (celui en  $n - \bar{n}$ ) s'annule et on peut approximer la fonctionnelle sphères dures par sa forme locale, soit celle de Percus-Yevick (PY), soit celle de Carnahan-Starling (CS) selon la forme de fonctionnelle KR-FMT choisie[53].

$$\mathcal{F}_{\text{exc}}^{\text{HS}}[\bar{n}(\mathbf{r})] = \iiint_{\mathbb{R}^3} f_{\text{exc}}^{\text{HS}}(\bar{n}(\mathbf{r})) \bar{n}(\mathbf{r}) d\mathbf{r}, \quad (5.3.11)$$



**FIGURE 5.3.1.:** Comparaison des fonctions de corrélation directe pour l'eau (en noir) et pour un fluide de sphères dures de rayon  $R_0 = 1.27 \text{ \AA}$  (en rouge) à la même densité  $n_b = 0.0333 \text{ \AA}^{-3}$ . La courbe bleue est la différence entre ces deux fonctions, c'est-à-dire la fonction de corrélation « attractive »  $c_{000}^{\text{VdW}}$ . On peut vérifier que  $a = c_{000}^{\text{VdW}}(k=0)$  et  $m = -k^2 \frac{d^2 c_{000}^{\text{VdW}}}{dk^2} \Big|_{k=0}$  sont positifs.

avec si on choisit par exemple celle de Carnahan-Starling,

$$f_{\text{exc}}^{\text{HS}}(n(\mathbf{r})) = \frac{\eta(4 - 3\eta)}{(1 - \eta)^2}, \quad (5.3.12)$$

où  $\eta = 4\pi R_0^3 n/3$  est la fraction d'empilement. Il est montré dans la ref. [53] que le choix de PY ou CS donne des résultats très proches pour l'état fluide. On peut alors écrire la partie intrinsèque (terme idéal et terme d'excès) de  $\mathcal{F}[\bar{n}(\mathbf{r})]$  comme,

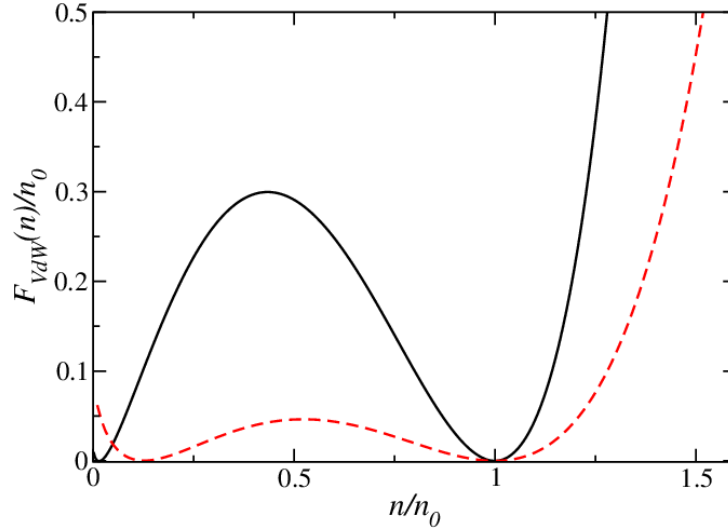
$$\mathcal{F}_{\text{int}}[\bar{n}(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \left[ F_{\text{VdW}}(\bar{n}(\mathbf{r})) + \frac{1}{2} m (\nabla \bar{n}(\mathbf{r}))^2 \right] d\mathbf{r}, \quad (5.3.13)$$

avec,

$$\begin{aligned} F_{\text{VdW}}(\bar{n}(\mathbf{r})) &= \bar{n}(\mathbf{r}) \ln \left( \frac{\bar{n}(\mathbf{r})}{n_b} \right) - \Delta \bar{n}(\mathbf{r}) + \bar{n}(\mathbf{r}) f_{\text{exc}}^{\text{HS}}(\bar{n}(\mathbf{r})) - n_b f_{\text{exc}}^{\text{HS}}(n_b) \\ &\quad - \left( f_{\text{exc}}^{\text{HS}}(n_b) + n_b \frac{df_{\text{exc}}^{\text{HS}}}{dn} \Big|_{n=n_b} \right) \Delta \bar{n}(\mathbf{r}) - \frac{1}{2} a \Delta \bar{n}(\mathbf{r})^2. \end{aligned} \quad (5.3.14)$$

L'eau liquide dans les conditions ambiantes est proche de la coexistence liquide-gaz. Il est donc important de choisir un paramètre  $a$  tel que les énergies libres associées au gaz et au liquide soient égales, ce qui revient à imposer que  $F_{\text{VdW}}(n(\mathbf{r}))$  ait deux minima équivalents, l'un à la densité du liquide  $n_b$ , l'autre à une densité très faible, celle du gaz. On représente une telle courbe sur la fig. 5.3.2 pour un liquide de densité  $n_b = 0.0333 \text{ \AA}^{-3}$  et pour des valeurs de rayons de sphères





**FIGURE 5.3.2.:** Fonctions de Van-der-Waals pour les densités gros grains, décrites par l'éq. 5.3.14, proches de la coexistence liquide gaz. La courbe noire correspond à un rayon de sphères dures du fluide de référence de 1.42 Å et la courbe rouge à un rayon de 1.27 Å.

dures  $R_0 = 1.27$  Å et  $R_0 = 1.42$  Å. Ces deux valeurs ont déjà été utilisées dans la littérature pour étudier un fluide de sphères dures mimant le cœur dur de l'eau[86, 87]. Pour ces deux rayons, les minima ont des valeurs similaires pour des paramètres valant respectivement  $n_{ba} = 6.7$  et  $n_{ba} = 12.5$ . Avec cette forme simple de l'équation de Van-der-Waals, on ne peut pas choisir indépendamment la hauteur de la barrière et la densité de la phase gaz, c'est cette version de la correction hydrophobe qui a été utilisée dans la ref. [84]. Le paramètre  $m$  semble lui n'avoir que peu d'influence. Dans le cadre du modèle de Chandler[79, 81, 82] de la théorie de Cahn-Hilliard pour une interface liquide-gaz, le profil de densité d'un fluide près d'un soluté hydrophobe évolue comme une tangente hyperbolique de la distance, et l'énergie libre de Van-der-Waals est,

$$F_{VdW}(n(\mathbf{r})) = \frac{6\gamma}{d}(n(\mathbf{r}) - n_l)^2(n(\mathbf{r}) - n_g)^2, \quad (5.3.15)$$

où  $\gamma$  est la tension de surface liquide-gaz,  $n_l$  et  $n_g$  sont les densités d'équilibre du liquide et du gaz respectivement. Si on compare le profil de l'énergie de Van-de-Waals de notre théorie avec celle issue du modèle de Chandler on se rend compte que pour pouvoir fixer la position du second minimum ainsi que la hauteur de la barrière, il est nécessaire d'ajouter un terme cubique et un terme d'ordre quatre en  $\Delta\bar{n}$ . Ce développement a été réalisé en collaboration avec Stojanović dans un article à paraître[88].

Il suffit ensuite de procéder à une minimisation fonctionnelle comme on l'a fait précédemment.

Soulignons ici que cette théorie a l'avantage d'être auto-cohérente : même si l'énergie libre est écrite comme une fonctionnelle de  $n$  et de  $\bar{n}$ , la minimisation est toujours conduite uniquement sur  $n$ .  $\bar{n}$  est elle-même une fonctionnelle de  $n$  dans l'éq. 5.3.4.

Si la dérivation a été faite avec une fonction de convolution de Heaviside, nous avons utilisé ici

une fonction gaussienne avec une largeur à mi-hauteur de 4 Å.

### 5.3.2. Résultats

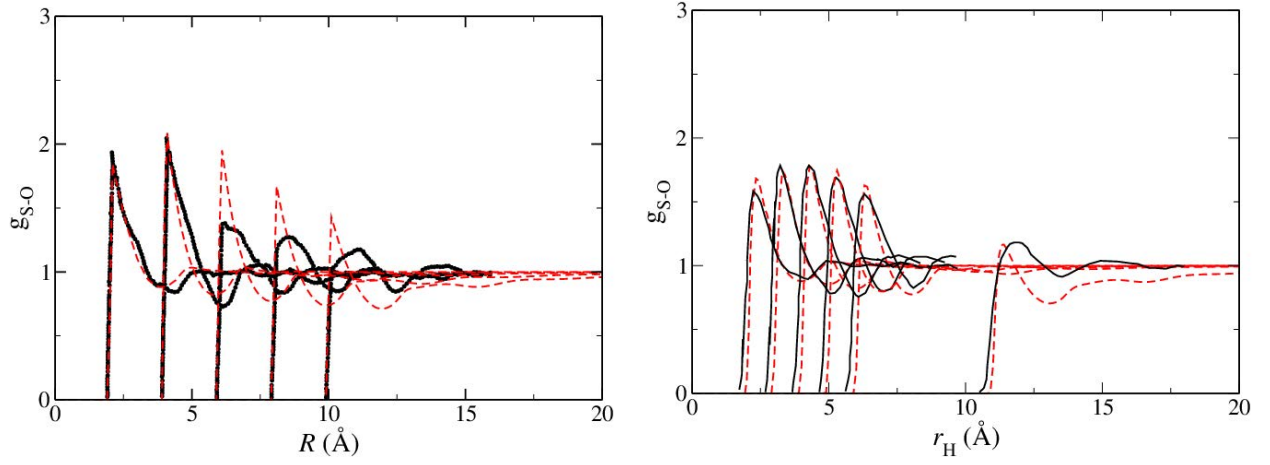
Les résultats discutés ici ont été obtenus avec une énergie libre de Van-der-Waals limitée à l'ordre deux.

L'effet de l'ajout de cette correction de l'hydrophobicité sur les fonctions de distribution radiale pour les systèmes de sphères dures et de sphères molles est donné en fig. 5.3.3.

Cette correction hydrophobe améliore les résultats par rapport à la simple inclusion du bridge sphères dures. Avec cette description de l'hydrophobicité à longue portée, on observe bien la diminution du pic de la première couche de solvation à partir d'un rayon de 5 Å, pour les deux systèmes considérés. Cependant, la correction n'est pas parfaite puisque pour les sphères dures de rayon  $R > 4$  Å, les pics apparaissent trop grands et trop piqués. De plus, pour les solutés de rayon supérieur à 6 Å, la déplétion après le premier maximum est surestimée. Cette observation est confirmée sur les sphères molles où les pics dus à la première couche de solvation sont mieux reproduits mais où on surestime toujours la déplétion après le premier pic pour des solutés de rayon supérieur à 6 Å. On donne aussi sur la fig. 5.3.4 l'effet de cette correction hydrophobe sur les énergies libres de solvation. Là encore la phénoménologie du changement de comportement avec l'augmentation en taille des solutés est bien reproduite. L'énergie libre de solvation est d'abord proportionnelle au volume du soluté jusqu'à 6 Å, la pente de la courbe diminue ensuite. Cependant l'énergie libre n'atteint pas de plateau pour les rayons considérés et est surestimée par rapport aux simulations. Le doublet de paramètres utilisé a été choisi pour obtenir le meilleur compromis entre structures et énergies libres de solvation et permet d'étudier l'ensemble des solutés.

On peut vérifier que cette correction, qui contient toujours le bridge de sphères dures, donne toujours des bons résultats pour les petits solutés. À titre d'exemple, on présente la fonction de distribution radiale entre le solvant et le centre de masse du néopentane ainsi que la carte de densité obtenue autour de ce solvant. Dans la ref. [84] est également donnée l'énergie libre de solvation pour les six premiers alcanes en incluant la correction hydrophobe, qui est en tout point semblable à la fig. 3.2.10.

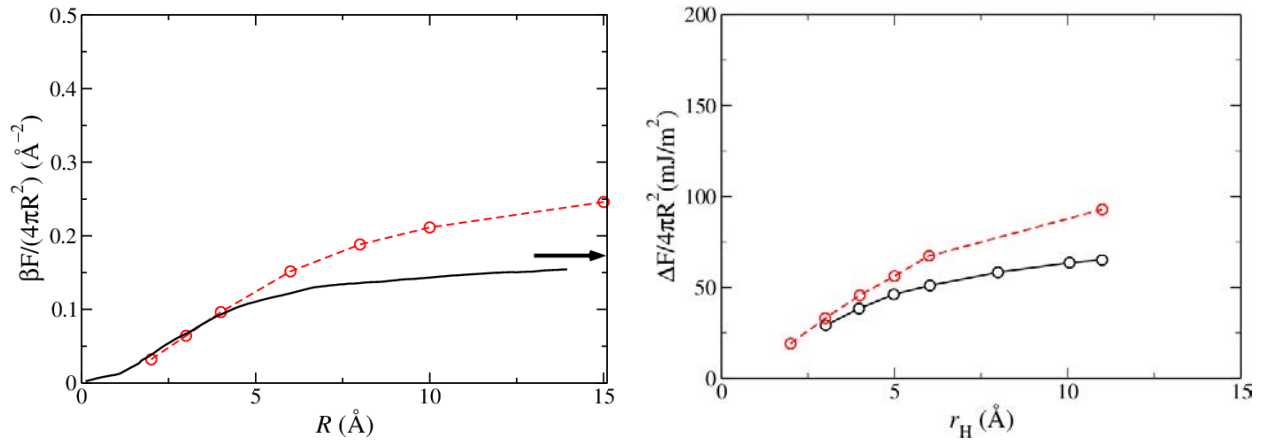
On a introduit dans la fonctionnelle une description phénoménologique qui reproduit correctement la solvation des solutés hydrophobes de petite et de grande taille. Cette correction contient des termes physiques similaires à la théorie de Cahn-Hilliard utilisée pour décrire les séparations de phases. Il est possible d'utiliser une densité gros grains pour traiter des phénomènes physiques se passant à des échelles mésoscopiques. Ceci peut permettre de coupler la MDFT, qui décrit la solvation au niveau moléculaire, avec des méthodes permettant d'étudier des plus grandes échelles comme les méthodes Lattice-Boltzmann.



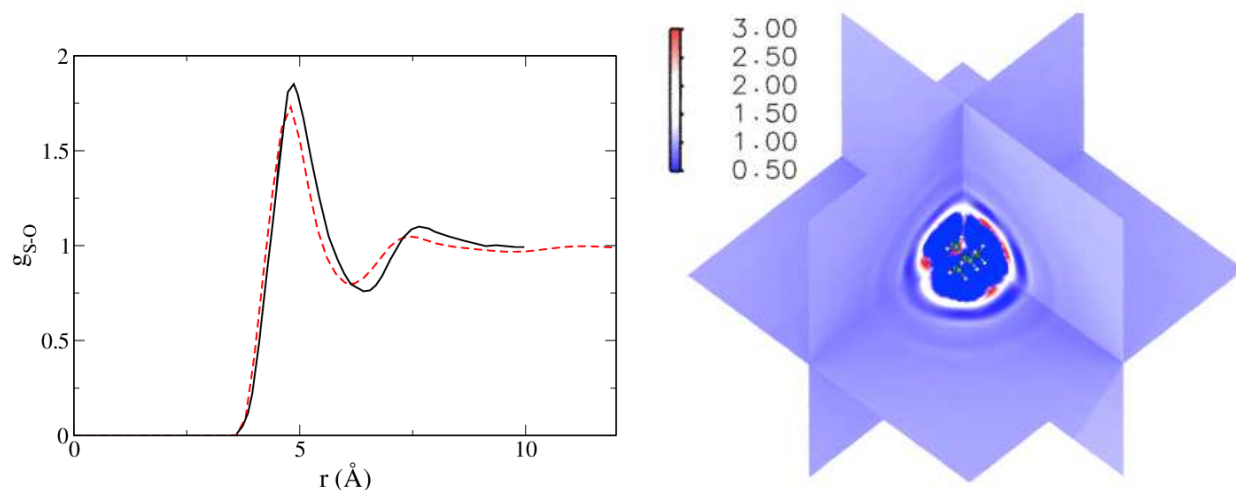
**FIGURE 5.3.3.:** Comparaison des fonctions de distribution radiale :

- entre un soluté sphère dure de rayon  $R$  et l'eau à gauche ;
- entre sphère molle de rayon  $r_H$  et l'eau à droite.

Les résultats obtenus par MDFT avec la correction hydrophobe sont en tirets rouges, ceux obtenus par simulation Monte-Carlo[81] ou par dynamique moléculaire[83] sont en noir.



**FIGURE 5.3.4.:** Energie libre de solvation par unité de surface pour les systèmes sphères dures (à gauche) et sphères molles (à droite). Les résultats obtenus avec le code mdft avec la correction hydrophobe sont en cercles rouges, ceux obtenus par MC et MD sont en noir.



**FIGURE 5.3.5.:** À gauche, la fonction de distribution radiale entre le centre de masse du néopentane et l'eau obtenue par MDFT avec la correction hydrophobe en tirets rouges, et par MD tirés de la ref. [89] en noir. À droite, une carte de densité du solvant eau autour du soluté néopentane.

## 6. Implémentation numérique de la MDFT

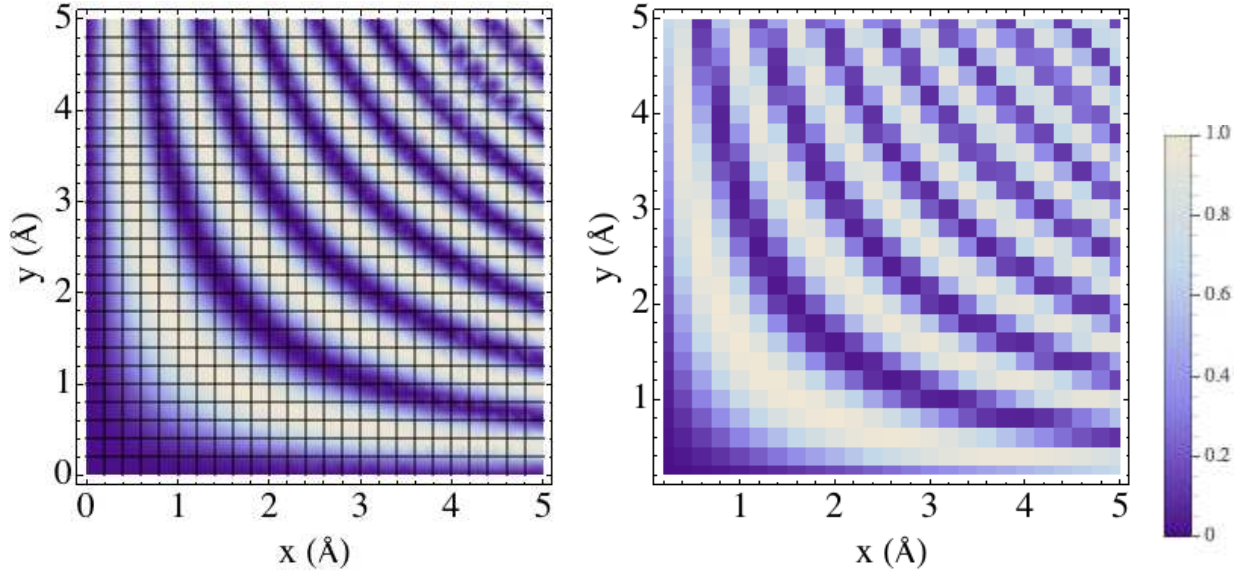
Dans ce chapitre, on décrit le fonctionnement du code mdft, basé sur la théorie MDFT présentée dans cette thèse. Ce code vise à trouver la valeur de la densité de solvant qui minimise la fonctionnelle d'énergie libre, et la valeur que prend la fonctionnelle pour cette densité de solvant.

### 6.1. Discrétisation de la densité sur une double grille spatiale et angulaire

Dans le cadre de la théorie MDFT, la densité est un champ scalaire  $\rho(\mathbf{r}, \mathbf{\Omega})$  qui dépend de six coordonnées  $(\mathbf{r}, \mathbf{\Omega}) = (x, y, z, \theta, \phi, \psi)$ , avec  $x, y$  et  $z$  les coordonnées cartésiennes dans le référentiel du laboratoire et  $\theta, \phi$  et  $\psi$  les trois angles d'Euler. Comme une minimisation analytique n'est pas réalisable, un code, mdft (molecular density functional theory), est écrit en Fortran moderne. Celui-ci permet de minimiser numériquement une fonctionnelle de la densité. Comme on ne peut pas travailler numériquement avec un espace de variable continu, l'espace et les orientations sont discrétisés sur deux grilles tridimensionnelles.

Cette discrétisation est illustrée sur une densité à deux dimensions sur la fig.6.1.1. Au lieu de chercher la valeur de la densité en tout point  $(x, y)$  du plan, on se contente de chercher cette densité sur les nœuds d'une grille d'un pas régulier de 0.1 Å. Dans l'implémentation actuelle, la grille spatiale est orthorhombique (un prisme rectangulaire à base rectangulaire). On utilise typiquement 3 ou 4 nœuds par angström. On appellera  $L_x$ ,  $L_y$  et  $L_z$  les dimensions de la grille spatiale selon les axes  $\mathbf{O}_x$ ,  $\mathbf{O}_y$  et  $\mathbf{O}_z$  du repère cartésien. L'espace angulaire est lui aussi discrétisé et on utilise généralement 6 nœuds pour les angles  $(\theta, \phi)$  et 4 nœuds pour l'angle  $\psi$ , soit environ 1200 nœuds par angström cube et un nombre total de l'ordre de  $10^7$ - $10^8$  variables de minimisation  $\{\mathbf{r}_i, \mathbf{\Omega}_j\}$ . Pour que la densité en chaque point ne puisse être que positive, ce qui est une contrainte physique évidente, la minimisation est en réalité conduite sur la variable  $\Psi(\{\mathbf{r}_i, \mathbf{\Omega}_j\}) = \sqrt{\rho(\{\mathbf{r}_i, \mathbf{\Omega}_j\})}$  qui est en quelque sorte un équivalent classique fictif de la fonction d'onde électronique. Cependant, hormis pour le calcul des gradients, on travaille toujours avec la densité recalculée à partir de  $\Psi$ . Dans la suite du texte, par soucis de simplicité, on continuera donc à utiliser la densité comme variable de minimisation.

La minimisation réalisée sur la variable  $\Psi$  peut être évitée en réalisant une minimisation sous contrainte de la variable  $\rho$ , en imposant comme contrainte que la densité soit positive.



**FIGURE 6.1.1.:** Illustration de la discrétisation de l'espace. À gauche on présente une fonction définie sur  $\mathbb{R}^2$  où l'on a représenté le grillage utilisé. À droite, la même fonction telle qu'elle est approchée par la discrétisation dans le code. Elle prend une valeur constante en chaque point de grille, égale à la valeur moyenne de la fonction continue sur le pixel contenant le point de grille.

## 6.2. Fonctionnement du code mdft

Le but du code mdft est de minimiser une fonctionnelle de la forme présentée en éq. 3.2.1 dans l'approximation du fluide homogène de référence. Cela consiste à trouver en chaque point  $\{\mathbf{r}_i, \mathbf{\Omega}_j\}$  des deux grilles tridimensionnelles spatiale et angulaire, la valeur de la densité à l'équilibre thermodynamique  $\rho_{\text{eq}}(\mathbf{r}, \mathbf{\Omega})$ , telle que le grand potentiel soit minimal. L'algorithme de ce programme est schématisé sur la fig. 6.2.1.

### 6.2.1. Description moléculaire du soluté et du solvant

Le soluté est décrit par un fichier ayant un format présenté dans le tab. 6.1. Ici le soluté est décrit par  $N^u=12$  sites différents. Chacun de ces sites est caractérisé par une charge et des paramètres Lennard-Jones  $\sigma_i^u$  et  $\epsilon_i^u$ . Les cinquième et sixième colonnes contiennent les paramètres  $\lambda^{1S}$  et  $\lambda^{2S}$  associés à la correction à trois corps de ce site.

Les trois colonnes suivantes décrivent les coordonnées cartésiennes des sites du soluté  $\mathbf{r}_i^u = (x_i^u, y_i^u, z_i^u)$ . Les trois dernières colonnes donnent des informations sur le site mais ne sont pas utilisées directement dans le code. Pour des raisons de mémoire, on dit que deux sites sont équivalents s'ils possèdent les mêmes charges et paramètres Lennard-Jones. Ainsi, dans l'exemple donné, tous les hydrogènes et tous les carbones sont équivalents entre eux.

On étudie un solvant perturbé par un soluté, le fichier d'input du solvant est similaire à celui du soluté.

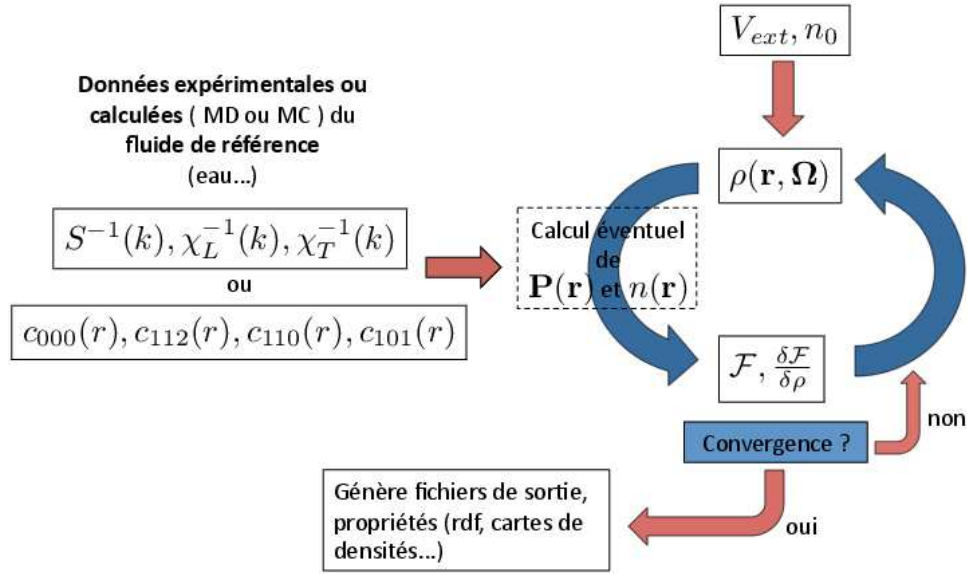


FIGURE 6.2.1.: Représentation schématique de l'algorithme du code mdft.

Le soluté intervient dans le calcul de la fonctionnelle à travers le potentiel extérieur, qui est la somme de plusieurs contributions possibles. Les potentiels extérieurs actuellement utilisables dans le programme sont celles présentées ci-dessous.

### 6.2.2. Calcul du potentiel extérieur

#### Calcul du potentiel extérieur électrostatique

Les valeurs et les positions des charges du soluté définissent une distribution de charge  $\sigma_c^u(\mathbf{r})$ . Le potentiel électrostatique généré par cette distribution de charge s'exprime à l'aide de l'équation

Benzène 12 2 #	charge	$\sigma$	$\epsilon$	$\lambda^{1S}$	$\lambda^{2S}$	$x$	$y$	$z$	Z	atome	surnom
1	-0.115	3.55	0.292669	0.0	0.0	1.386	0.000	0.0	6	C	C
1	-0.115	3.55	0.292669	0.0	0.0	0.693	-1.200	0.0	6	C	C
1	-0.115	3.55	0.292669	0.0	0.0	-0.693	-1.200	0.0	6	C	C
1	-0.115	3.55	0.292669	0.0	0.0	-1.386	0.000	0.0	6	C	C
1	-0.115	3.55	0.292669	0.0	0.0	-0.693	1.200	0.0	6	C	C
1	-0.115	3.55	0.292669	0.0	0.0	0.693	1.200	0.0	6	C	C
2	0.115	2.42	0.125248	0.0	0.0	2.462	0.000	0.0	1	H	H
2	0.115	2.42	0.125248	0.0	0.0	1.231	-2.132	0.0	1	H	H
2	0.115	2.42	0.125248	0.0	0.0	-1.231	-2.132	0.0	1	H	H
2	0.115	2.42	0.125248	0.0	0.0	-2.462	0.000	0.0	1	H	H
2	0.115	2.42	0.125248	0.0	0.0	-1.231	2.132	0.0	1	H	H
2	0.115	2.42	0.125248	0.0	0.0	1.231	2.132	0.0	1	H	H

TABLE 6.1.: Exemple d'un fichier d'input de soluté

de Poisson, en unités SI,

$$\Delta V_c = -\frac{\sigma_c^u}{\epsilon_0}, \quad (6.2.1)$$

où  $\Delta$  est l'opérateur laplacien et  $\epsilon_0$  est la permittivité du vide. Dans l'espace de Fourier cette équation de Poisson se réécrit,

$$\hat{V}_c(\mathbf{k}) = -\frac{\hat{\sigma}_c^u(\mathbf{k})}{\epsilon_0 k^2}, \quad (6.2.2)$$

où  $k = \|\mathbf{k}\|$  est la norme du vecteur réciproque et  $\hat{\sigma}_c^u$  est la transformée de Fourier (voir la sec. 6.4) de la densité de charge du soluté. Le programme calcule la transformée de Fourier du potentiel électrostatique à partir de celle de la densité de charge du soluté et de l'éq. 6.2.2. La solution,  $V_c$ , est donnée par la transformée de Fourier inverse de  $\hat{V}_c$ .

Le potentiel extérieur électrostatique est ensuite calculé à l'aide de la densité de charge d'une molécule de solvant,  $\sigma_c^v(\mathbf{r}, \mathbf{\Omega})$ ,<sup>1</sup>

$$v_c(\mathbf{r}, \mathbf{\Omega}) = \frac{1}{4\pi} V_c(\mathbf{r}) \sigma_c^v(\mathbf{r}, \mathbf{\Omega}). \quad (6.2.3)$$

### Calcul du potentiel extérieur Lennard-Jones

À partir des paramètres de Lennard-Jones des sites du soluté et de ceux du solvant, le programme calcule des paramètres Lennard-Jones mixtes  $\sigma_i^{uv}$  et  $\epsilon_i^{uv}$  via les règles de Lorentz-Berthelot, qui sont celles utilisées le plus généralement,

$$\sigma_i^{uv} = \frac{\sigma_i^u + \sigma^v}{2}, \quad (6.2.4)$$

$$\epsilon_i^{uv} = \sqrt{\epsilon^v \epsilon_i^u}. \quad (6.2.5)$$

Grâce à ces paramètres mixtes le programme calcule ensuite le potentiel extérieur Lennard-Jones

$$v_{\text{LJ}}(\mathbf{r}) = 4 \sum_{i=1}^{N^u} \epsilon_i^{uv} \left[ \left( \frac{\sigma_i^{uv}}{\|\mathbf{r}\|} \right)^{12} - \left( \frac{\sigma_i^{uv}}{\|\mathbf{r}\|} \right)^6 \right]. \quad (6.2.6)$$

Pour des raisons numériques, on borne la densité d'énergie totale à une valeur maximale, typiquement 100 kJ.mol<sup>-1</sup>.

### Calcul du potentiel extérieur sphères dures

Il est également possible de rajouter un rayon de sphère dure  $R_i^u$  sur les différents sites du soluté. On définit également un rayon de sphère dure pour le solvant  $R^v$ . Le potentiel extérieur

---

1. On rappelle que cette densité de charge moléculaire est utilisée pour calculer la densité de charge du solvant  $\rho_c(\mathbf{r}) = \iiint_{\mathbb{R}^3} \sigma_c^v(\mathbf{r} - \mathbf{r}', \mathbf{\Omega}) \rho(\mathbf{r}', \mathbf{\Omega}) d\mathbf{r}'$



sphères dures vaut alors,

$$v_{SD}(r) = \sum_{i=1}^{N^u} v_{iSD}(r), \quad (6.2.7)$$

avec

$$\begin{cases} v_{iSD}(r) = 0 & \text{si } |\mathbf{r} - \mathbf{r}_i^u| \leq R_i^u + R^v \\ v_{iSD}(r) = \infty & \text{si } |\mathbf{r} - \mathbf{r}_i^u| > R_i^u + R^v \end{cases}. \quad (6.2.8)$$

### Calcul d'un potentiel extérieur sphères molles

Un potentiel de sphère molle purement répulsive en  $r^{-p}$  où  $p$  est un entier positif, peut aussi être rajouté sur les sites de soluté en précisant la valeur des rayons de sphère molle  $r_{iH}$ ,

$$V_p(r) = k_B T \sum_{i=1}^{N^u} \|\mathbf{r} - \mathbf{r}_{iH}\|^{-p} \text{ si } \|\mathbf{r}\| > r_{iH}, \text{ et } \infty \text{ sinon} \quad (6.2.9)$$

On note que si  $p=12$ , ce potentiel correspond à celui utilisé dans la référence [83].

### 6.2.3. Initialisation de la densité

Une fois la densité d'énergie extérieure connue, on peut initier la densité à

$$\rho_{\text{init}}(\mathbf{r}, \boldsymbol{\Omega}) = \rho_b \exp[-\beta (v_{\text{ext}}(\mathbf{r}) - v_c(\mathbf{r}))], \quad (6.2.10)$$

où  $v_{\text{ext}}$  est la densité d'énergie extérieure totale, on rappelle que  $\rho_b = n_b/(8\pi^2)$ . On retire la contribution due à l'énergie électrostatique, pathologique, qui peut diverger. Cette initialisation permet d'avoir des densités raisonnables, c'est-à-dire élevées (respectivement faibles) là où le potentiel extérieur est faible (respectivement élevé). La plupart du temps on initie donc la densité à  $\rho_{\text{init}}(\mathbf{r}, \boldsymbol{\Omega}) = \rho_b \exp[-\beta v_{\text{LJ}}(\mathbf{r})]$ .

On peut remarquer que l'éq. 6.2.10 correspond à la solution exacte pour un fluide idéal (dont le terme d'excès est nul) et un potentiel extérieur  $v_{\text{ext}}(\mathbf{r}) - v_c(\mathbf{r})$ .

### 6.2.4. Fonctions de corrélation directe

Pour le calcul du terme d'excès, il est nécessaire d'introduire des fonctions de corrélation obtenues expérimentalement ou calculées par simulations numériques. Celles-ci dépendent de la forme de la fonctionnelle que l'on veut utiliser. Ce sont soit les fonctions de corrélation directe,  $c_{000}$ ,  $c_{112}$ ,  $c_{110}$  et  $c_{101}$  si on utilise la forme dipolaire de la fonctionnelle présentée en sec. 3.2, soit le facteur de structure  $S$  et les inverses des constantes diélectriques,  $\chi_L^{-1}$  et  $\chi_T^{-1}$ , si on utilise la forme multipolaire de la sec. 3.3. On souligne une fois encore que la force de la théorie réside dans le fait que ces fonctions de corrélation ne doivent être déterminées qu'une seule fois pour chaque

solvant bulk de densité donnée, et que l'on peut ensuite étudier la solvation de n'importe quel soluté dans ce solvant. Les moments multipolaires de la molécule d'eau sont également nécessaires et il sont fournis de manière implicite par le choix du modèle moléculaire.

### 6.2.5. Minimiseur et minimisation

#### Description du minimiseur

Pour réaliser la minimisation fonctionnelle, le code mdft utilise le minimiseur L-BFGS (Limited-memory BFGS) [90] qui est une méthode quasi-Newton pour trouver les minima d'une fonction. Ce minimiseur est très efficace pour les problèmes impliquant un grand nombre de variables. Cette routine trouve les points stationnaires d'une fonction, c'est-à-dire les points où le gradient est nul. Au lieu d'utiliser le gradient et la matrice Hessienne (la matrice des dérivés secondes) de la fonction à minimiser, comme un algorithme gradient-conjugué classique, le minimiseur utilise les  $m$  valeurs obtenues aux  $m$  pas précédents de la densité et du gradient de la fonctionnelle (dans notre code on utilise généralement  $m = 4$ ). Ces valeurs stockées servent à réaliser implicitement les opérations requérant normalement l'utilisation de la matrice Hessienne. Ceci permet d'éviter le calcul de la matrice Hessienne et de limiter la quantité de mémoire utilisée. Ce minimiseur nécessite donc qu'on lui fournisse, à chaque pas, les valeurs de l'énergie libre et des gradients en chaque nœud de la double grille. Il y a deux critères de convergence permettant de considérer que le minimum de la fonctionnelle est atteint. Le premier porte sur la norme du gradient, qui doit être inférieure à un critère arbitraire (généralement  $10^{-7}$ ). Le second porte sur la variation relative de l'énergie libre entre deux pas successifs  $((F_n - F_{n-1})/F)$  qui doit elle aussi être inférieure à un critère arbitraire (généralement  $10^{-6}$ ).

#### Processus de minimisation

À partir de la valeur initiale de la densité, du potentiel extérieur et des fonctions de corrélation, le code calcule les différentes parties de la fonctionnelle de l'éq. 3.2.1 ainsi que leur gradient. Le minimiseur calcule alors une nouvelle densité notée  $\rho_1$ . Le code recalcule les différentes parties de la fonctionnelle et leur gradient à partir de la densité  $\rho_1$ . Si l'un des deux critères de convergence décrits en sec. 6.2.5 est atteint, la procédure s'arrête et les dernières densités et énergie libre de solvation calculées sont considérées comme les grandeurs physiques à l'équilibre thermodynamique. Si aucun des critères de convergence n'est vérifié alors le minimiseur trouve une nouvelle densité  $\rho_2$  à partir de sa valeur précédente  $\rho_1$  et de celles des gradients. À partir de cette densité, les nouvelles valeurs des différentes composantes de l'énergie libre (de la fonctionnelle) et des gradients sont calculées. Le procédé est itéré jusqu'à ce qu'un des critères de convergence soit atteint. Une fois sorti de cette boucle de minimisation, le code calcule différentes observables d'intérêt à partir de la densité d'équilibre obtenue : fonctions de distribution radiale, densité, densité de polarisation à l'équilibre, etc.

### 6.3. Calculs des différentes intégrales rencontrées

Les différents termes de la fonctionnelle qui doivent être calculés lors de la minimisation contiennent plusieurs types d'intégrale sur les différentes variables mises en jeu. On distingue essentiellement trois types d'intégrale : des intégrales sur les angles, des intégrales sur le volume et des intégrales doubles sur le volume au carré.

#### 6.3.1. Intégrales sur le volume

Pour les termes qui contiennent une intégrale simple sur les coordonnées d'espace comme la partie idéale de l'éq. 2.2.6, on réalise directement le calcul dans l'espace direct. L'intégrale est simplement approchée par une somme finie,

$$\mathcal{F}[\rho(\mathbf{r})] = \iiint_{\mathbb{R}^3} A(\rho(\mathbf{r}))d\mathbf{r} \equiv \mathcal{F}[\{\rho(\mathbf{r}_i)\}] \equiv \sum_{i=1}^N A(\rho(\mathbf{r}_i))\Delta_V, \quad (6.3.1)$$

où  $N = N_x \cdot N_y \cdot N_z$  avec  $N_x$  (resp.  $N_y$ , resp.  $N_z$ ) le nombre de nœuds de la grille selon la direction  $\mathbf{O}_x$  (resp.  $\mathbf{O}_y$ , resp.  $\mathbf{O}_z$ ) et  $\Delta_V = L_x \cdot L_y \cdot L_z / (N_x \cdot N_y \cdot N_z)$  le volume d'un voxel élémentaire. On a fait ici une approximation implicite, qui consiste en l'utilisation d'une quadrature simple pour évaluer l'intégrale.

Les méthodes de quadrature sont des approximations pour évaluer la valeur numérique d'une intégrale. On remplace l'intégrale par une somme pondérée en un certain nombre de points appartenant au domaine d'intégration. Par exemple, on a à l'éq. 6.3.1 des points régulièrement espacés avec comme poids le volume du voxel élémentaire.

Pour évaluer l'intégrale sur le volume d'une fonction, on pourrait utiliser une quadrature plus adaptée, c'est à dire qui donnerait une bonne estimation numérique de l'intégrale avec un nombre de point plus faible. Il est évident, par exemple, que la densité varie peu loin de la perturbation, on pourrait ainsi utiliser un maillage plus lâche loin du soluté, plutôt qu'une grille régulière.

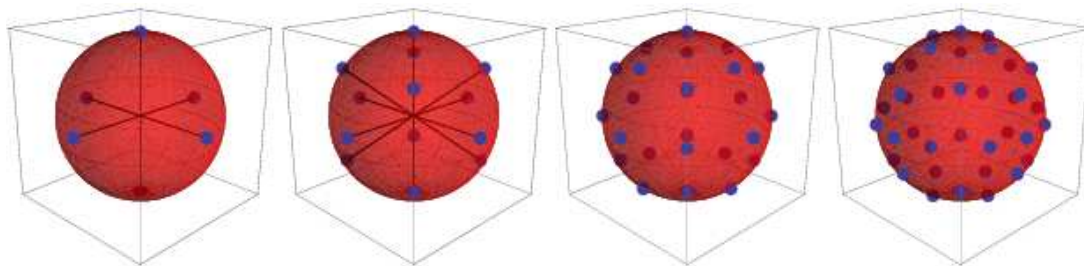
Cependant, comme on le verra par la suite, pour calculer efficacement les produits de convolution, nous faisons un usage intensif de transformées de Fourier rapides (FFT) et l'algorithme choisi pour faire ces FFT requiert l'utilisation d'une grille régulière.

#### 6.3.2. Intégrales sur la grille angulaire

Ce sont des intégrales de la forme,

$$n(\mathbf{r}) = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \rho(\mathbf{r}, \mathbf{\Omega}) d\mathbf{\Omega}. \quad (6.3.2)$$

On rappelle que  $\mathbf{\Omega}$  est la notation compacte des trois angles d'Euler.



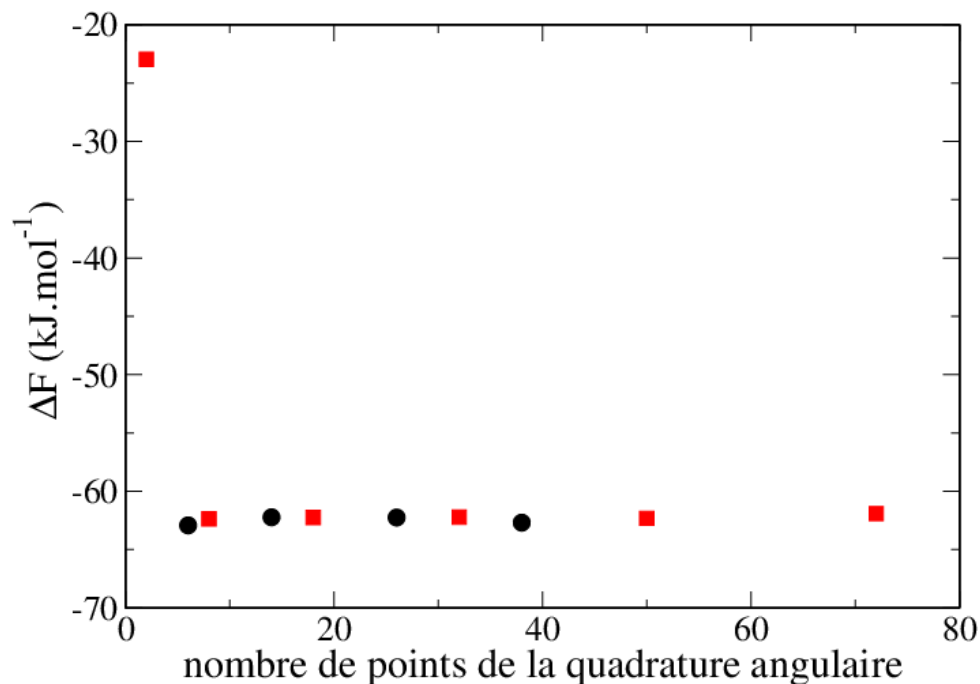
**FIGURE 6.3.1.:** Visualisation des orientations considérées dans la quadrature de Lebedev, avec de gauche à droite 6, 14, 26 et 38 points. La sphère unité est en rouge et les orientations considérées sont représentées par des points bleus.

### Quadrature pour les deux premiers angles d'Euler $(\theta, \phi)$

Les intégrations sur la grille angulaire sont réalisées dans l'espace direct exactement de la même manière que pour le volume. Cependant, on n'utilise plus ici de quadrature régulière. À mon arrivée au laboratoire une quadrature du type Gauss-Legendre[91] était utilisée pour les angles  $(\theta, \phi)$ . Cette quadrature est conçue pour des intégrales dont le domaine d'intégration est  $[-1, 1]$ . Elle donne les  $n$  points où la fonction doit être évaluée comme étant les racines du polynôme de Legendre d'ordre  $n$ , ainsi que les poids associés à ces points. Cette quadrature est classique et donne des résultats exacts pour les fonctions polynomiales de degré  $2n - 1$ , et donc les harmoniques sphériques d'ordre  $n - 1$ . Pour l'intégration sur la sphère unité, avec une quadrature de Gauss-Legendre d'ordre  $n$ , comme les nœuds  $x_i$  tels que  $i \in [1, n]$ , varie entre  $-1$  et  $1$ , on définit  $\theta_i$  tel que  $\cos(\theta_i) = x_i$ , de telle sorte que  $\theta$  varie entre  $0$  et  $\pi$ . Les angles  $\phi_i$  doivent varier entre  $0$  et  $2\pi$ , on les choisit régulièrement espacés sur cette intervalle. On a donc pour une quadrature de Gauss-Legendre d'ordre  $n$ ,  $2n^2$  points pour décrire les deux premiers angles d'Euler.

Au cours de ma thèse j'ai implémenté dans le code mdft une autre quadrature que celle-ci, la quadrature de Lebedev [92]. Cette quadrature est spécialement développée pour réaliser une intégration sur la sphère unité. Elle est construite sur une symétrie octaédrique. Ainsi, contrairement à la quadrature de Gauss-Legendre, les contraintes sur le nombre de points pour calculer numériquement l'intégrale sont fortes. En conséquence, dans la version actuelle du code, on peut utiliser 6, 14, 26 ou 38 points pour décrire les orientations sur la sphère unité. Par exemple, les 6 premières orientations sont dans les directions des centres des faces du cube unité contenant la sphère, les 8 points suivants sont dans les directions des sommets, les 12 points suivants sont orientés dans la direction du milieu des arêtes du cube comme illustré sur la fig. 6.3.1. La quadrature de Lebedev est plus efficace qu'une quadrature de Gauss-Legendre d'un facteur d'environ  $2/3$ . On peut voir sur la fig. 6.3.2 que les deux quadratures donnent des résultats similaires pour un soluté et un solvant simple.

L'introduction de cette quadrature a été une bonne façon de « mettre les mains » dans un code existant. De plus, il se peut que pour étudier des solutés plus complexe, il soit nécessaire d'avoir une meilleure résolution angulaire, ce qui sera facilité par l'utilisation de la quadrature de Lebedev.



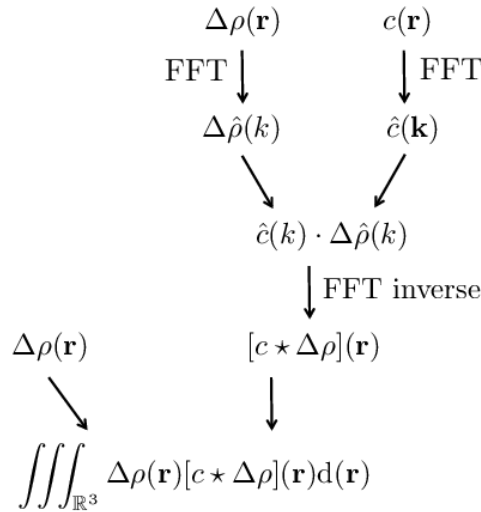
**FIGURE 6.3.2.:** Valeurs des énergies libres de solvation obtenues en utilisant une grille de Lebedev (cercles noirs) ou de Gauss Legendre (carrés rouges) en fonction du nombre de points de la grille. Le soluté utilisé est un sodium portant une demi charge élémentaire et le solvant un fluide de Stockmayer. La boîte fait  $25 \text{ \AA}^3$ , avec une grille de  $96^3$  points.

#### Quadrature pour l'angle de rotation propre $\psi$

Au début de ma thèse, le code ne fonctionnait que pour des solvants strictement linéaires (acétonitrile, fluide de Stockmayer, etc). Pour pouvoir l'utiliser pour le solvant eau il a fallu introduire un nouvel angle d'Euler, celui correspondant à la rotation propre d'une molécule de solvant sur son axe de symétrie principal  $C_n$ . J'ai donc introduit l'angle  $\psi$  comme une nouvelle variable dont dépendait la densité. Dans le but d'augmenter l'efficacité numérique du code, cet angle a été introduit comme variant entre 0 et  $2\pi/n$ , où  $n$  est l'ordre de l'axe de symétrie principal (soit 2 pour l'eau). La quadrature choisie pour cette nouvelle variable est une quadrature régulière, comparable à celle pour l'angle  $\phi$  dans la quadrature de Gauss-Legendre.

#### 6.3.3. Intégrales doubles sur le volume

Pour les termes contenant une intégrale double sur le volume, comme par exemple la partie d'excès de l'éq. 2.3.16, le calcul dans l'espace direct apparaît comme trop coûteux. Cependant,



**FIGURE 6.3.3.:** Représentation schématique de la façon dont sont calculées, dans le code mdft, les intégrales contenant un produit de convolution comme dans l'éq. 6.3.3.

ces intégrales sont des convolutions de la forme,

$$\mathcal{F}[\rho(\mathbf{r})] = \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}_1) c(\mathbf{r}_{12}) \Delta\rho(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (6.3.3)$$

$$= \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}_1) \left[ \iiint_{\mathbb{R}^3} c(\mathbf{r}_{12}) \Delta\rho(\mathbf{r}_2) d\mathbf{r}_2 \right] d\mathbf{r}_1 \quad (6.3.4)$$

$$= \iiint_{\mathbb{R}^3} \Delta\rho(\mathbf{r}_1) [c \star \Delta\rho](\mathbf{r}_1) d\mathbf{r}_1, \quad (6.3.5)$$

où on a séparé les termes dépendant de  $\mathbf{r}_2$ , et  $\star$  désigne le produit de convolution entre les deux fonctions  $\Delta\rho$  et  $c$ . On calcule ce produit de convolution en se servant de transformées de Fourier. En effet, une des propriétés intéressantes des transformées de Fourier (voir la sec. 6.4) est que le produit des transformées de deux fonctions est égal à la transformée de Fourier du produit de convolution de ces deux fonctions. Par exemple, la procédure pour calculer l'intégrale de l'éq. 6.3.3 est schématisée sur la fig. 6.3.3 :

- On calcule, si on ne les connaît pas déjà, les champs scalaires  $\Delta\rho$  et  $c$  en chaque point de l'espace, c'est-à-dire pour tous les points de notre grille.
- On calcule leurs transformées de Fourier respectives, que l'on note  $\Delta\hat{\rho}$  et  $\hat{c}$ .
- On multiplie ces deux transformées de Fourier pour calculer la transformée de Fourier du produit de convolution.
- On effectue la transformée de Fourier inverse de ce produit pour obtenir la fonction produit de convolution  $c \star \Delta\rho$  dans l'espace direct.
- On peut alors calculer l'intégrale de l'éq. 6.3.3 comme on l'a fait avec l'intégrale simple sur l'espace de l'éq. 6.3.1.

### 6.3.4. Calcul des densités de charges et de polarisation microscopique

La densité de charge microscopique du solvant est calculée dans l'espace de Fourier,

$$\sigma(\mathbf{k}, \boldsymbol{\Omega}) = \sum_m z_m e \exp(-i\mathbf{k} \cdot \mathbf{s}_m(\boldsymbol{\Omega})) \quad (6.3.6)$$

où  $z_m e$  désigne la charge partielle et  $\mathbf{s}_m(\boldsymbol{\Omega})$  la position du  $m^{\text{ième}}$  site de solvant d'une molécule d'eau située à l'origine et d'orientation  $\boldsymbol{\Omega}$ . La densité de polarisation microscopique est également calculée dans l'espace réciproque en utilisant l'éq. 3.3.3. Ces deux densités sont calculées à l'initialisation du programme et stockées pour être utilisées par la suite.

Comme les densités de charge et de polarisation sont des produits de convolution, entre ces densités microscopiques et la densité macroscopique  $\rho(\mathbf{r}, \boldsymbol{\Omega})$ , on les calcule également dans l'espace de Fourier :

$$\mathbf{P}(\mathbf{k}) = \iiint_{8\pi^2} \mu(\mathbf{k}, \boldsymbol{\Omega}) \rho(\mathbf{k}, \boldsymbol{\Omega}) d\boldsymbol{\Omega} \quad (6.3.7)$$

et

$$\rho_c(\mathbf{k}) = \iiint_{8\pi^2} \sigma(\mathbf{k}, \boldsymbol{\Omega}) \rho(\mathbf{k}, \boldsymbol{\Omega}) d\boldsymbol{\Omega}. \quad (6.3.8)$$

Les composantes longitudinale et transverse de la polarisation sont ensuite calculées directement dans l'espace réciproque grâce à l'éq. 3.3.13. On peut ensuite calculer la fonctionnelle multipolaire.

On souligne qu'il faut donc réaliser un nombre de transformées de Fourier directe et inverse égale au nombre total d'angles utilisés.

## 6.4. Transformées de Fourier rapides

Comme on l'a vu dans le calcul de la densité d'énergie électrostatique et le calcul des intégrales doubles sur le volume, le code mdft requiert le calcul de transformées de Fourier.

Ceci est fait numériquement par l'utilisation de transformées de Fourier discrètes efficaces. Celles-ci sont réalisées par des transformées de Fourier rapides (FFT), qui sont calculées grâce à l'algorithme FFTW3 (Fastest Fourier Transform in the West)[93, 94, 95]. On ne décrira pas, ici, comment fonctionne cet algorithme en particulier, ni les principes des transformées de Fourier discrètes et rapides puisqu'aucun travail n'a été réalisé sur cette partie. On se limite à l'utilisation de ces techniques numériques. On peut néanmoins souligner que l'on approxime ce qui est en réalité des intégrales continues (les transformées de Fourier) par des sommes finies (les transformées de Fourier discrètes),

$$\hat{f}(k) = \iiint_{\mathbb{R}^3} f(\mathbf{x}) e^{-2i\pi\mathbf{k} \cdot \mathbf{x}} d\mathbf{x} \approx \hat{f}(\mathbf{k}_p) = \sum_{n=0}^{N-1} f(\mathbf{x}_n) e^{-2i\pi\mathbf{k}_p \cdot \mathbf{x}_n} \Delta_V. \quad (6.4.1)$$

La transformée de Fourier discrète transforme alors une suite de  $N$  nombres (possiblement complexes, mais dans notre cas réels) en une suite  $N$ -périodique de nombres (complexes). Notons que la transformée de Fourier discrète inverse se construit de manière analogue,

$$f(\mathbf{x}_n) = \frac{1}{(2\pi)^3} \sum_{p=0}^{N-1} \hat{f}(\mathbf{k}_p) e^{2i\pi \mathbf{k}_p \cdot \mathbf{x}_n} \Delta V_k. \quad (6.4.2)$$

Cette équation implique également la  $N$ -périodicité des  $f(\mathbf{x}_n)$  et  $\hat{f}(\mathbf{k}_p)$  et donc des conditions aux bords périodiques.

On souligne également que les définitions données ici sont celles de FFTW3 qui ne sont pas normalisées : si on fait une transformée de Fourier discrète suivi d'une transformée de Fourier discrète inverse de la même fonction, on voit que le résultat est la série de départ multipliée par  $N$ , il faut donc à posteriori renormaliser ces résultats.

L'utilisation de transformées de Fourier rapides et l'efficacité du code FFTW3 sont en grande partie à l'origine de la rapidité du code mdft. On peut par exemple constater, pour la correction à trois corps, la différence flagrante de temps de calcul entre la routine qui calcule directement les intégrales doubles dans l'espace direct et celle qui calcule les produits de convolution à l'aide de transformées de Fourier, voir la fig. 4.2.3.

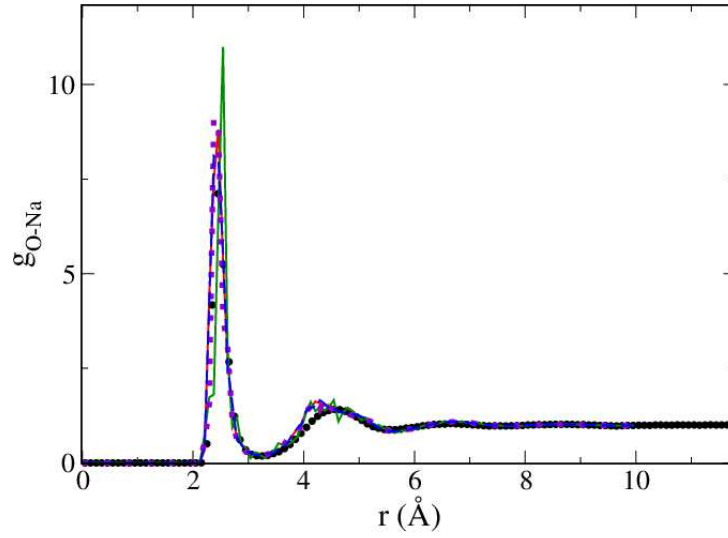
## 6.5. Influence du maillage sur la convergence

Après avoir présenté comment fonctionne le code mdft, on s'intéresse à l'influence sur les résultats obtenus de la résolution (la finesse du maillage) des grilles angulaire et spatiale, ainsi qu'à l'efficacité algorithmique du code.

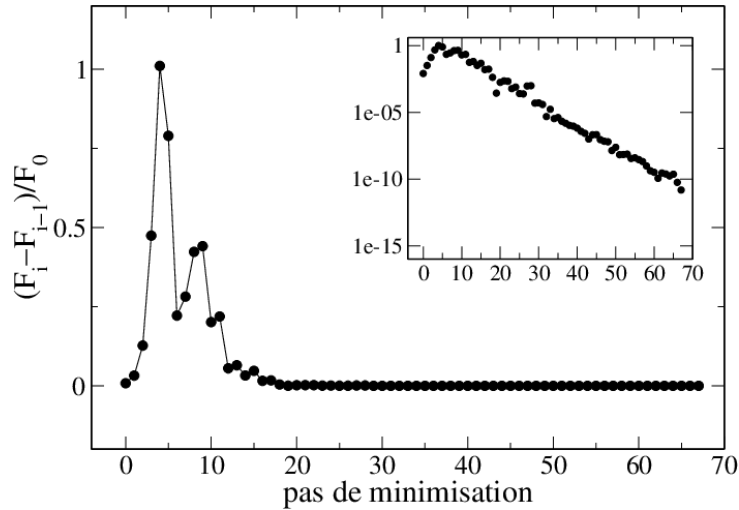
On montre sur la fig. 6.5.1 l'influence de la résolution de la grille spatiale sur les fonctions de distribution radiale. On a choisi comme exemple l'ion sodium, que l'on étudie avec la fonctionnelle multipolaire et la correction à trois corps décrite au chapitre 4. On obtient un profil correct à partir d'une résolution de 3 points par Å. Une augmentation de cette résolution n'a pas de conséquence notable sur la structure. En ce qui concerne l'énergie libre de solvation, un profil de convergence typique est présenté en fig. 6.5.2 pour un ion sodium, un critère de convergence de  $10^{-4}$  est atteint après une vingtaine de pas. Le temps de CPU (Central Processing Unit) nécessaire à la minimisation de la fonctionnelle augmente comme le nombre de points de grille angulaire et spatiale. Ceci est illustré sur la fig. 6.5.3.

Le code est donc linéaire en fonction du nombre de points de grille utilisé pour la grille angulaire, et pour la grille spatiale. Le temps de calcul est donc proportionnel à la taille du système étudié.

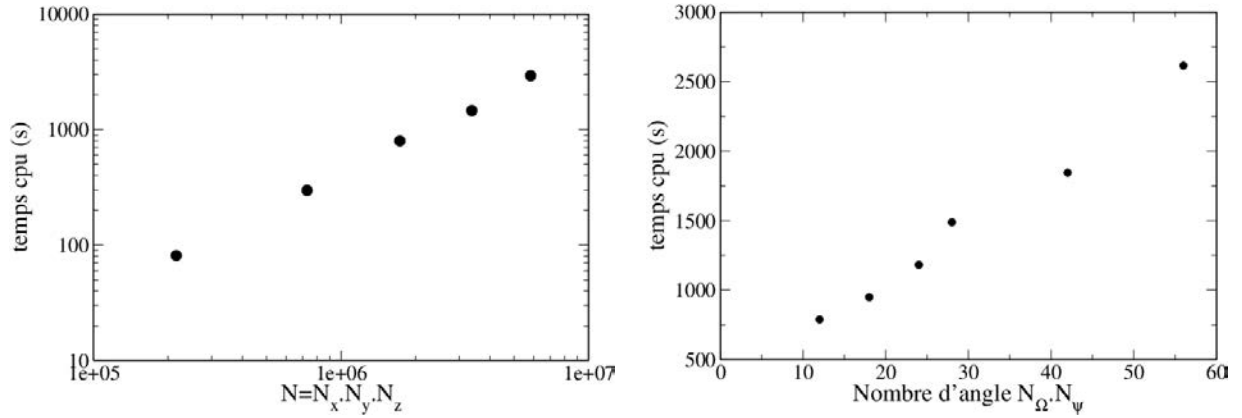




**FIGURE 6.5.1.:** Fonctions de distribution radiale entre le solvant eau et l'ion sodium. La référence calculée par dynamique moléculaire est donnée par les points noirs. Les calculs MDFT ont été réalisés avec une boîte de 30 Å, 12 points pour la grille angulaire et 2 (ligne verte), 3 (points violets), 4 (tirets bleus), et 5 (ligne rouge) points par Å pour la grille spatiale.



**FIGURE 6.5.2.:** Différences d'énergies libre de solvation entre deux pas successifs, normalisées par l'énergie libre de solvation initiale, pour un sodium avec une boîte de 30 Å, 5 points par Å et 12 points pour la grille angulaire en fonction du pas d'itération. L'encart représente la même courbe avec une échelle logarithmique pour l'axe des ordonnées.



**FIGURE 6.5.3.:** Temps CPU pour réaliser la minimisation totale. Le système est le même que pour la courbe de la fig. 6.5.2. À gauche, ce temps est présenté pour une grille angulaire de 12 points en fonction du nombre total de points de la grille spatiale. À droite, ce temps est présenté pour une grille spatiale de  $120^3$  points en fonction du nombre de points de la grille angulaire.

## 6.6. Conclusion

Le code mdft est une implémentation en Fortran moderne de la théorie MDFT présentée dans ce manuscrit. Il permet la minimisation fonctionnelle de l'énergie libre sur une double grille angulaire et spatiale (soit 6 dimensions). Une grande partie de l'efficacité numérique de ce programme provient de l'utilisation de transformées de Fourier rapides[93]. Le calcul de la partie d'excès nécessite l'injection de fonctions de corrélation obtenues préalablement par simulations numériques ou par l'expérience. Ce code permet d'obtenir des informations sur la structure et sur l'énergie libre de solvation comparables à celles obtenues par simulations numériques en étant  $10^2 - 10^3$  fois plus rapide. Un calcul typique de solvation d'une petite molécule prend moins de dix minutes sur un ordinateur portable. Le code n'a pas été parallélisé pour l'instant et, étant encore en développement, n'est pas encore mis à disposition de la communauté.

Troisième partie .

## Applications

## 7. Applications

Après avoir introduit la théorie de la fonctionnelle de la densité classique moléculaire et validé son utilisation sur un ensemble de petites molécules tests, nous allons montrer dans ce chapitre qu'elle est un outil adapté à l'étude de systèmes complexes.

Nous présenterons d'abord des calculs de potentiels de force moyenne. Nous nous intéressons ensuite à l'étude d'une argile, la pyrophyllite puis à celle d'un système biologique, la protéine lysozyme du blanc d'œuf de poule.

## 7.1. Calculs de potentiels de force moyenne (PMF)

Nous avons utilisé la théorie MDFT pour calculer des potentiels de force moyenne entre des petites molécules hydrophobes dans l'eau. Un potentiel de force moyenne (PMF) représente l'évolution de l'énergie libre d'un système en fonction d'une coordonnée de réaction (CR), tous les autres degrés de liberté étant moyennés thermodynamiquement. C'est une grandeur couramment utilisée en chimie numérique, par exemple pour étudier les interactions entre médicaments et récepteurs, protéines et membranes cellulaires. Dans un solvant, l'interaction entre solutés se fait à travers les molécules de solvant et le PMF contient l'effet du solvant sur cette l'interaction. Les PMF sont aussi utilisés pour la construction de modèles gros-grains pour les simulations moléculaires[96].

Pour ces raisons, des méthodes basées sur des simulations MD ou MC ont été développées pour calculer des PMF. C'est un problème particulièrement complexe car il nécessite un échantillonnage des configurations du solvant pour chaque valeur de la coordonnée réactionnelle. Certaines valeurs de la CR sont particulièrement difficiles à échantillonner car elles possèdent une énergie d'activation élevée. Il faut donc biaiser les simulations moléculaires pour forcer l'exploration le long de la coordonnée réactionnelle[8]. Ces méthodes fonctionnent mais nécessitent des simulations longues pour avoir une statistique suffisante.

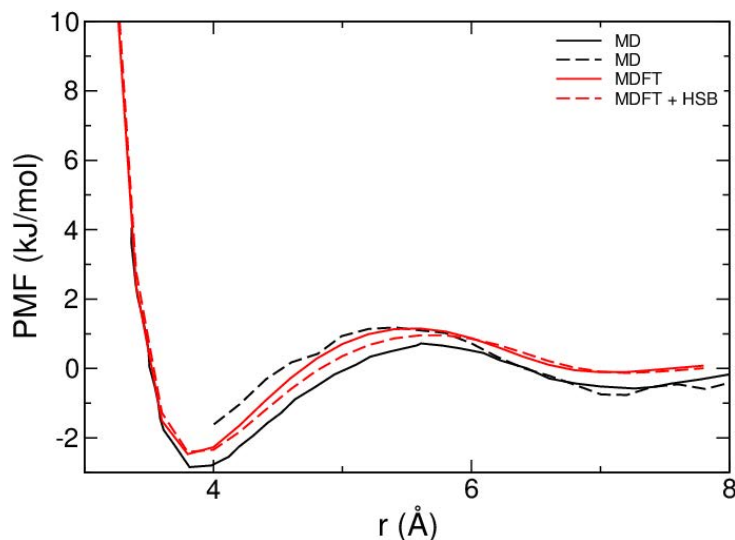
L'utilisation d'une méthode implicite peut faciliter l'accès aux PMF. Par exemple, 3D-RISM avec une fermeture HNC ou Kovalenko-Hirata a été utilisé pour calculer les PMF d'ions polyatomiques dans des solvants polaires[31] et d'ions simples dans l'eau[30]. On se propose d'utiliser la MDFT pour calculer les PMF[88] entre deux solutés sphériques. On considérera deux petits solutés (méthanes), deux gros solutés (néopentanes), deux solutés de tailles différentes (un méthane et un néopentane). On s'intéresse aussi à un PMF de torsion dans une molécule (butane). Pour le méthane et le néopentane, on utilise les modèles unifiés de OPLS/AA. Ces molécules sont donc représentées par des sphères Lennard-Jones.

### 7.1.1. Fonctionnelle

Comme les solutés étudiés ici ne sont pas chargés, on utilise une fonctionnelle où le terme d'excès se limite au terme fonction de  $c_{000}(r)$  de l'éq. 3.2.6 et au terme correctif. On a donc une fonctionnelle de la forme,

$$\mathcal{F}[n(\mathbf{r})] = \mathcal{F}_{\text{id}}[n(\mathbf{r})] + \mathcal{F}_{\text{ext}}[n(\mathbf{r})] - \frac{1}{2}k_B T \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} [c_{000}(\mathbf{r} - \mathbf{r}') \Delta n(\mathbf{r}) \Delta n(\mathbf{r}')] d\mathbf{r} d\mathbf{r}' + \mathcal{F}_{\text{cor}}[n(\mathbf{r})]. \quad (7.1.1)$$

Le terme correctif choisi est le bridge de sphères dures décrit à l'éq. 5.3.3 car il permet d'obtenir des valeurs des énergies libres de solvation correctes pour les solutés neutres. Le terme extérieur est dû à la présence des deux solutés à une distance fixée.



**FIGURE 7.1.1.:** PMF méthane-méthane calculé par MDFT en rouge, sans bridge de sphères dures (traits pleins) et avec bridge de sphères dures (tirets) et par MD en noir (les résultats de Asthagiri[97] sont en traits pleins et ceux de Shimizu en tirets[98]).

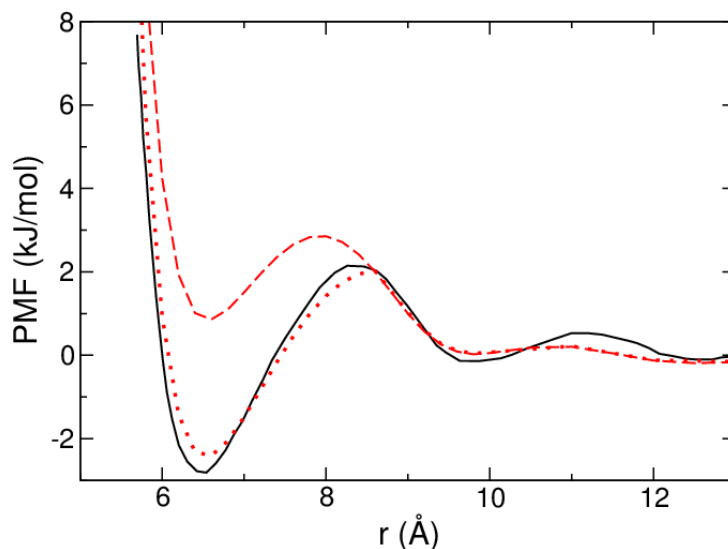
### 7.1.2. Résultats

**Méthane-méthane :** Le modèle de méthane utilisé est celui de Asthagiri et collaborateurs[97]. On utilise le même modèle d'eau, SPC. Sur la fig. 7.1.1, on présente le PMF obtenu entre deux molécules de méthane par MDFT et par deux dynamiques moléculaires différentes[97, 98]. Le PMF obtenu est en accord quantitatif avec ceux obtenus par les deux simulations. L'ajout du bridge de sphères dures ne modifie pas sensiblement le résultat, ce qui est attendu pour les petits solutés.

Les potentiels de force moyenne dans le cas de particules sphériques présentent généralement deux minima séparés par un maximum. C'est le cas ici. Le premier minimum, appelé *minimum de contact* (CM), est généralement plus faible que le second minimum, appelé *minimum séparé par le solvant* (SSM). Le CM correspond au cas où les deux solutés sont en contact et ne sont donc pas séparés par des molécules de solvant. Le SSM correspond à un état métastable où les deux solutés sont séparés par une couche unique de solvant. Dans le cas du PMF présenté en fig. 7.1.1, le premier minimum se situe à  $\approx 3.9$  Å et correspond à la configuration la plus stable de deux molécules de méthane dans l'eau. Dans le cas de particules hydrophobes, le CM est stabilisé pour des raisons entropiques tandis que le SSM l'est pour des raisons enthalpiques[99, 89].

Comme l'obtention des PMF entre petits solutés sphériques est possible, on s'intéresse maintenant à des solutés plus gros.

**Néopentane-néopentane :** Le comportement hydrophobe des molécules de néopentane a fait l'objet de plusieurs études théoriques[89]. Huang et collaborateurs ont étudié l'argon, le méthane et le néopentane dans l'eau et ont trouvé que, contrairement aux deux autres solutés, le néopentane a un comportement similaire aux solutés hydrophobes de grande taille. Une analyse

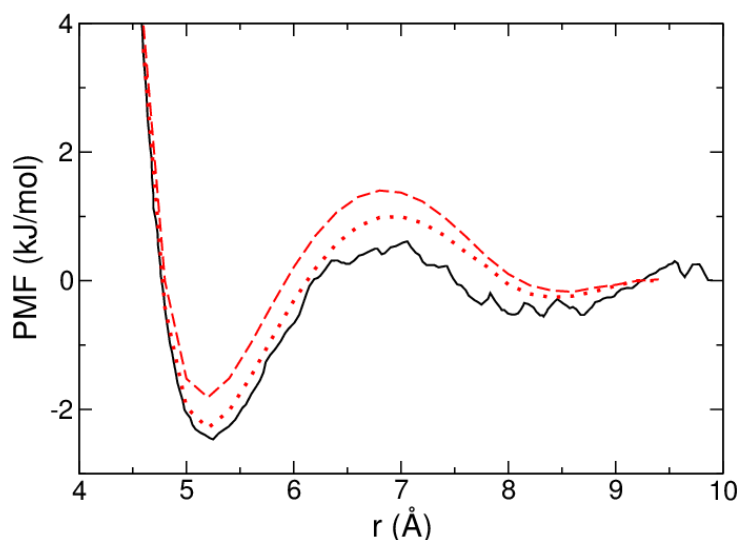


**FIGURE 7.1.2.:** PMF néopentane-néopentane calculé par MDFT en rouge, avec bridge de sphères dures (tirets) et avec la correction hydrophobe à longue distance (pointillés) et par MD[89] en noir.

orientationnelle de la molécule montre notamment que les molécules d'eau de la première couche de solvation ne peuvent former quatre liaisons hydrogènes avec les molécules d'eau voisines ce qui est typique des gros hydrophobes. Les potentiels de force moyenne entre deux molécules de néopentane, représentées par un modèle unifié (une seule sphère Lennard-Jones), sont présentés sur la fig. 7.1.2. On voit que les résultats obtenus par minimisation fonctionnelle avec un bridge de sphères dures sont très différents des résultats de MD. En effet, les positions des CM et SSM sont légèrement décalées vers les petites distances entre néopentane et les valeurs des deux extrema sont surestimées. Ces résultats mitigés ne sont pas surprenants car la solvation du néopentane, qui est un gros soluté, est gouvernée par des effets hydrophobes à longue portée. Pour améliorer la description du néopentane, on ajoute la correction hydrophobe à longue portée du chapitre 5, ce qui améliore l'accord entre MD et MDFT. Il est donc possible d'obtenir les PMF entre deux petits solutés et entre deux gros solutés. Qu'en est-il des PMF entre solutés de tailles différentes où ces deux échelles sont mélangées ?

**Méthane-néopentane :** La comparaison des PMF entre méthane et néopentane obtenus avec et sans la correction hydrophobe, présentée en fig. 7.1.3, prouve que l'inclusion de la correction hydrophobe à longue portée est nécessaire pour reproduire correctement les PMF impliquant des gros hydrophobes. L'accord entre les résultats obtenus par dynamique moléculaire[100] et par minimisation fonctionnelle est encore une fois très bon.

La MDFT permet donc d'obtenir des PMF corrects entre solutés sphériques, le prochain paragraphe s'intéresse à un PMF intramoléculaire.



**FIGURE 7.1.3.:** PMF méthane-néopentane calculé par MDFT en rouge, avec bridge de sphères dures (tirets) et avec la correction hydrophobe à longue distance (pointillés) et par MD[100] en noir .

**Torsion du butane :** Nous avons calculé le potentiel de force moyenne dû à la torsion d’une molécule de butane selon son angle dièdre. Cette molécule est représentée par un modèle à quatre sites. Les paramètres Lennard-Jones utilisés sont les mêmes que dans une étude de MD calculant ce PMF[101]. Ce sont les paramètres OPLS présentés dans le tab. 3.3. Les longueurs de liaisons ainsi que les angles entre sites Lennard-Jones sont fixés à leurs valeurs idéales 1.54 Å et 111°. Les résultats obtenus par minimisation fonctionnelle avec la correction bridge de sphères dures et par MD [101] sont présentés en fig. 7.1.4. L’accord entre les deux méthodes est une fois de plus très bon.

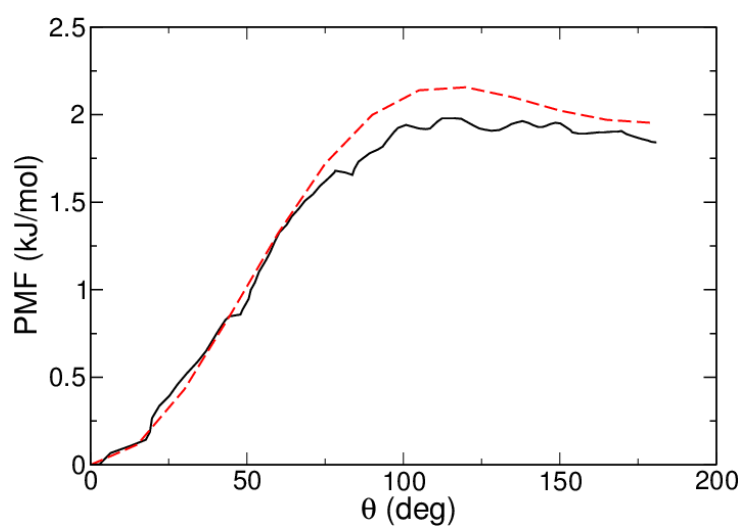
Le calcul de PMF en utilisant la théorie MDFT est possible pour des molécules neutres. Nous avons pu reproduire très rapidement des PMF intermoléculaires pour des solutés sphériques et un PMF intramoléculaire.

La rapidité de la méthode permet de réaliser des études systématiques qui sont hors de portée des techniques de simulations moléculaires.

Ces résultats sont encourageants, et ouvrent la voie à l’utilisation de la MDFT pour calculer des PMF pour des solutés plus complexes, notamment des solutés chargés.

Dans les prochaines parties nous présentons l’étude de solutés complexes par MDFT.





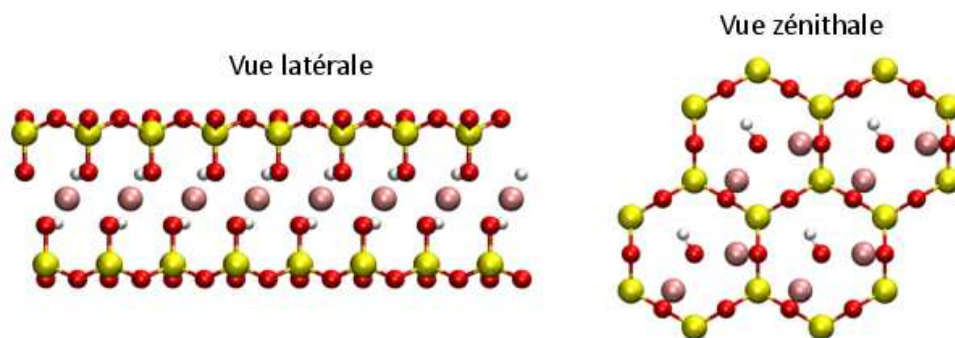
**FIGURE 7.1.4.:** PMF selon la torsion de l'angle dièdre du butane, calculé par MDFT avec bridge sphères dures en tirets rouges et par MD[101] en traits noirs continus.

## 7.2. Étude de la solvation d'une argile

Les phénomènes aux interfaces jouent un rôle important dans de nombreuses applications, notamment en catalyse hétérogène et dans les propriétés d'absorption des matériaux poreux. Ces phénomènes peuvent être étudiés expérimentalement, par spectroscopie ou diffraction[102, 103]. Ces expériences donnent des renseignements sur les quelques couches moléculaires de liquide adsorbées à la surface d'un matériaux. L'utilisation de méthodes numériques peut permettre de clarifier ou de rationaliser des observations expérimentales en donnant une explication moléculaire à ces phénomènes. Les méthodes théoriques couramment utilisées reposent généralement sur des simulations moléculaires (MD ou MC)[104]. L'étude de l'adsorption dans les matériaux poreux est un problème multi-échelle puisqu'on s'intéresse à l'adsorption de molécules nanométriques dans des pores de taille mésoscopique. Il est nécessaire de réaliser des simulations comportant des milliers ou des millions d'atomes pour étudier ces systèmes, ce qui requiert l'utilisation de supercalculateurs. Il est donc difficile de réaliser des études systématiques de ces systèmes ou de systèmes de plus grande taille étant donné les ressources computationnelles nécessaires. L'utilisation de méthodes implicites est une alternative pour éviter ce coût numérique et énergétique. La théorie MDFT a le double avantage d'être bien plus rapide et moins coûteuse numériquement que les simulations moléculaires, tout en conservant une description moléculaire du solvant.

On se propose d'utiliser la MDFT dans l'approximation du fluide homogène de référence pour étudier la solvation d'une argile, la pyrophyllite. Les argiles sont des matériaux d'intérêt pour diverses applications. Par exemple, l'Andra envisage de stocker sur des longues durées des déchets nucléaires dans les couches géologiques profondes constituées d'argile du site de Bure. En effet, les argiles sont imperméables, ce qui empêcherait la diffusion des radionucléotides dans les couches inférieures si les matériaux d'enrobage des déchets radioactifs venaient à se fissurer. Les argiles sont des matériaux multi-échelle constitués d'un empilement de feuillets atomiques séparés par des pores nanométriques. L'étude de ces pores de petite taille est tout à fait possible par MD. Les matrices d'argile sont également constituées de pores à l'échelle mésoscopique dont l'étude n'est pas réalisable par MD. On peut néanmoins les étudier par des simulations n'ayant pas une description moléculaire, comme la méthode de Lattice-Boltzmann. Pour avoir une description complète du matériau, il faudrait pouvoir étudier la solvation dans ces pores d'échelles différentes, tout en conservant une description moléculaire. La MDFT peut permettre d'étudier ces matériaux à l'échelle mésoscopique, avec une description moléculaire. De plus, il pourrait être possible de coupler la MDFT avec des modélisations macroscopiques, par exemple de dynamique des fluides, pour avoir un traitement multi-échelle complet de ces matériaux.

On utilisera tout d'abord un solvant simple, le fluide de Stockmayer, pour tester la faisabilité de l'approche[105], avant d'étudier la solvation par l'eau[106] qui est un solvant complexe nécessitant l'utilisation d'une fonctionnelle adaptée. Les résultats obtenus par MDFT seront dans les deux cas comparés à des résultats de dynamique moléculaire.



**FIGURE 7.2.1.:** Vues latérale et zénithale des feuillets de pyrophyllite. Les atomes de silicium sont en jaune, les atomes d'aluminium en rose, les oxygènes en rouge et les hydrogènes en blanc.

Atome	O	Si	O	H	Al	H	O	Si	O
z (Å)	39.03	39.62	41.21	41.30	0.00	1.00	1.09	2.68	3.27

**TABLE 7.1.:** Coordonnées des couches atomiques de l'argile pyrophyllite selon l'axe  $O_z$  perpendiculaire au plan de l'argile.

### 7.2.1. Solvation par le fluide de Stockmayer

On utilise comme solvant le fluide de Stockmayer constitué d'une sphère Lennard-Jones centrale et de deux sites portant une charge de  $\pm 1.91$  e situés à une distance de  $0.1$  Å de part et d'autre de ce site, résultant en un dipôle de  $1.85$  D. Les paramètres de Lennard-Jones donnés en tab. 7.2 et le dipôle ont été ajustés pour que le fluide mime certaines propriétés de l'eau (rayon moléculaire, dipôle, constante diélectrique). Cependant, ce solvant ne permet pas de reproduire toutes les propriétés de l'eau, comme la polarisation au-delà de l'ordre dipolaire, la création de liaisons hydrogènes intermoléculaires et par conséquent la structure locale tétraédrique de l'eau.

#### 7.2.1.1. Description du système

La pyrophyllite est une argile neutre qui cristallise dans le groupe monoclinique  $2/m$ . Elle est clivée selon le plan  $\{011\}$ . Elle est constituée d'un empilement de neuf couches atomiques. La couche supérieure est constituée d'atomes d'oxygène (O) qui se trouvent juste au-dessus d'une couche d'atomes de silicium (Si). Ces deux couches forment une structure hexagonale en nids d'abeilles. La couche centrale est composée d'atomes d'aluminium (Al). Entre cette couche d'aluminium centrale et les deux couches Si-O se trouvent deux couches constituées de groupements hydroxyles, dont l'oxygène se trouve au centre des hexagones formés par les deux couches supérieures. Les vues latérale et zénithale des feuillets d'argiles sont données en fig. 7.2.1. Les coordonnées des couches atomiques selon l'axe  $O_z$  perpendiculaire aux feuillets sont données dans le tab. 7.1, la couche d'aluminium centrale a été choisie comme l'origine de cet axe.

Molécule	Atome	$\epsilon$ (kJ.mol <sup>-1</sup> )	$\sigma$ (Å)	q (e)
Pyrophyllite	Al	5.56388e-6	4.27120	1.575
	Si	7.7005e-6	3.30203	2.100
	O <sub>G</sub>	0.650190	3.16554	-1.050
	O <sub>H</sub>	0.650190	3.16554	-0.950
	H <sub>G</sub>	0.0	0.0	0.425
Stockmayer	Central	1.847	3.024	0.0
	Site 1	0.0	0.0	1.91
	Site 2	0.0	0.0	-1.91

**TABLE 7.2.:** Champ de force utilisé pour le soluté pyrophyllite et le fluide de Stockmayer dans les simulations de dynamique moléculaire et les calculs MDFT. Les paramètres de la pyrophyllite sont tirés du champ de force CLAYFF.

### 7.2.1.2. Aspects numériques

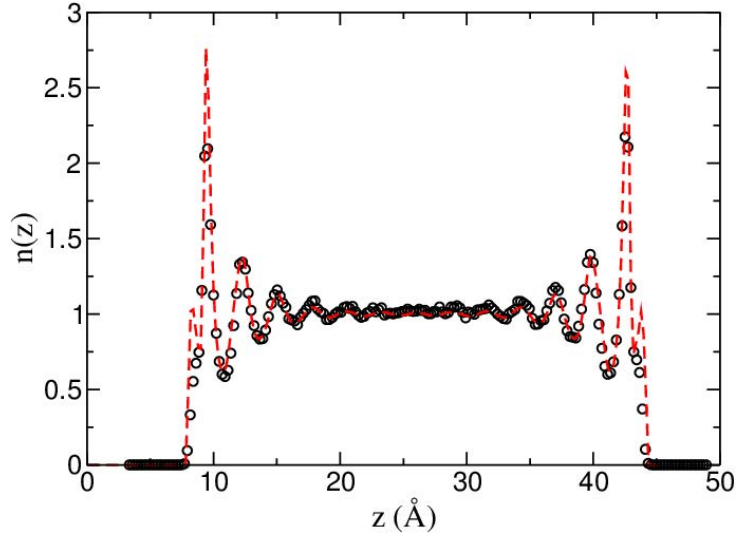
**Champ de force :** À la fois dans les simulations de dynamique moléculaire et dans les calculs MDFT, les atomes de la surface interagissent avec le solvant au travers du champ de force CLAYFF[107], constitué de sites Lennard-Jones et de charges ponctuelles. C'est un champ de force générique pour les systèmes minéraux et leurs interfaces avec les solutions. Les paramètres de champ de force pour l'argile et le solvant sont donnés dans le tab. 7.2. La boîte de simulation contient deux demi-couches d'argile de surface  $L_x \times L_y = 41.44 \times 35.88 \text{ Å}^2$ . La distance entre deux surfaces est de  $L_z = 45.57 \text{ Å}$ .

**Dynamique moléculaire :** Les simulations de dynamique moléculaire ont été réalisées par Virginie Marry au laboratoire Phenix de l'UPMC. Les simulations sont menées dans l'ensemble canonique avec un thermostat de Nosé-Hoover. Après une étape d'équilibration, les simulations ont été menées sur une trajectoire de 5 ns. La densité de solvant locale  $n(\mathbf{r})$  est calculée par moyenne temporelle dans des voxels de volume  $0.1^3 \text{ Å}^3$ . Les densités dans la direction perpendiculaire au plan des feuillets sont calculées à partir de la densité moyenne précédente que l'on moyenne dans les plans  $(x, y)$  parallèles aux feuillets. Les simulations de dynamique moléculaire ont été réalisées avec le logiciel DLPOLY[108, 109].

**MDFT :** Comme le solvant est strictement dipolaire, on utilisera la MDFT avec une fonctionnelle d'excès dipolaire, telle qu'elle a été présentée dans la sec. 3.2. La MDFT est environ mille fois plus rapide que la MD.

### 7.2.1.3. Résultats

On compare d'abord les profils de densité et d'orientation obtenus par MDFT aux résultats de MD. On analyse ensuite l'influence des interactions électrostatiques sur ces grandeurs.



**FIGURE 7.2.2.:** Profils de la densité planaire selon l'axe perpendiculaire au plan de l'argile. Les résultats calculés par dynamique moléculaire sont les cercles noirs. Ceux calculés par MDFT sont en tirets rouges.

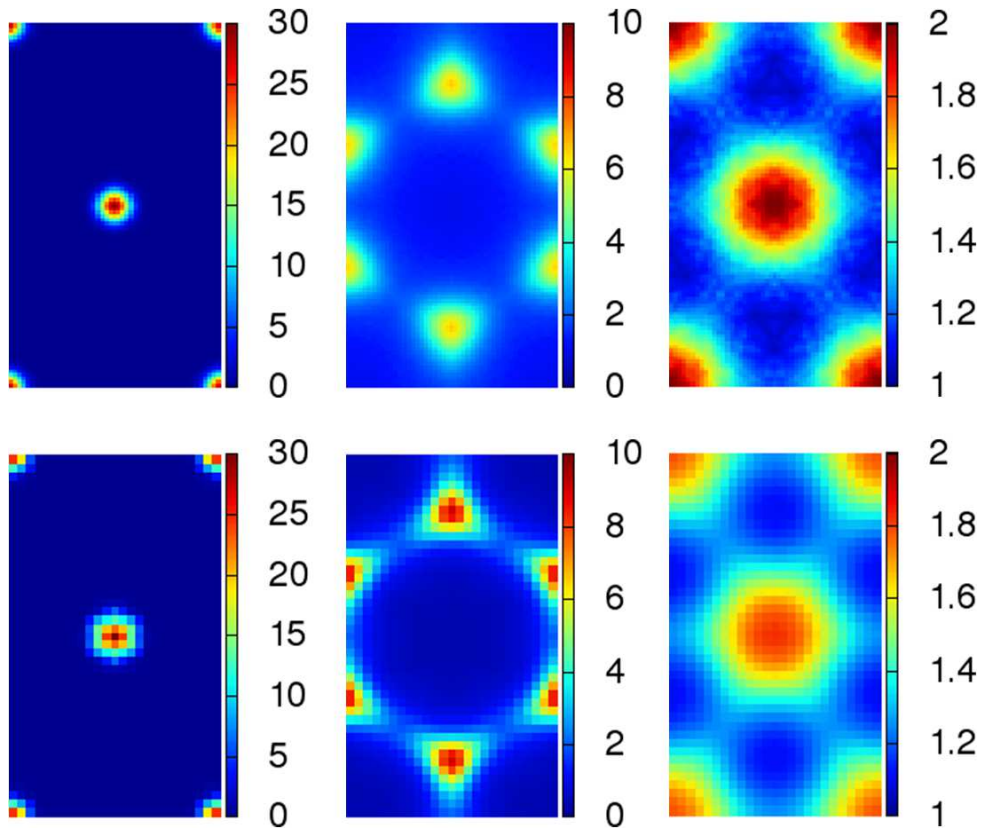
**Profil de densité :** On analyse l'interaction du solvant avec l'interface grâce à la densité planaire normalisée selon l'axe  $z$ , définie comme,

$$n(z) = \int_{x=0}^{L_x} \int_{y=0}^{L_y} \frac{n(\mathbf{r})}{n_b} dx dy. \quad (7.2.1)$$

Les profils de densité obtenus par dynamique moléculaire et par minimisation fonctionnelle sont donnés en fig. 7.2.2. On trouve un bon accord global entre les deux méthodes. En MD, le pic principal observé présente un épaulement alors qu'en MDFT on observe un « prépic » au lieu de cet épaulement. Le premier pic est suivi par un second pic plus faible et par des oscillations à longue distance. L'épaulement ou « prépic » se situe à  $z=8.3$  Å c'est-à-dire à seulement 1.76 Å de la couche de surface, son intensité est de  $\approx 0.7$  en MD et de  $\approx 1$  en MDFT. Le pic principal, qu'on appelle premier pic, se situe à  $z=9.5$  Å soit à 2.96 Å de la surface. Le second pic se trouve à 12.3 Å, soit à 5.76 Å de la surface. Au centre de la cellule on retrouve une densité égale à celle de l'eau bulk,  $n = n_b$ . Le solvant se trouve alors dans les mêmes conditions que dans la phase homogène.

La seule différence dans les résultats obtenus par les deux méthodes réside donc dans l'intensité du prépic, qui est plus marquée pour la MDFT. Ce prépic est très localisé. Cette observation est confirmée par la fig. 7.2.3 où sont présentées les cartes de densité, calculées par MD et par MDFT. Ces cartes présentent les densités dans les plans parallèles au plan de l'argile situés à une distance correspondant aux maxima de densité planaire de la fig. 7.2.2.

On obtient des cartes globalement similaires avec les deux méthodes. Pour MDFT, dans le plan du prépic, le pic de densité est légèrement plus large, avec une intensité similaire. Ceci explique la plus grande valeur lorsque l'on fait la moyenne dans le plan. Le prépic se trouve au centre



**FIGURE 7.2.3.:** Cartes de densité normalisée,  $n(\mathbf{r})/n_b$ , dans les plans du prépic (gauche), du pic principal (centre) et du second pic (droite), calculées par MD en haut et par MDFT en bas.

des hexagones Si-O. On peut souligner la valeur élevée de la densité à cet endroit ( $n/n_b \approx 30$ ). L'intégrale de ce pic correspond à la présence d'une molécule unique dans cette zone.

Sur la fig. 7.2.3 sont également présentées les cartes de densité dans les plans du premier et du second pic. Là encore, les cartes obtenues par les deux méthodes sont très similaires. Le premier pic est localisé essentiellement au-dessus des atomes de silicium alors que l'on observe une déplétion au-dessus des atomes d'oxygène. Le second pic est globalement localisé au centre des hexagones.

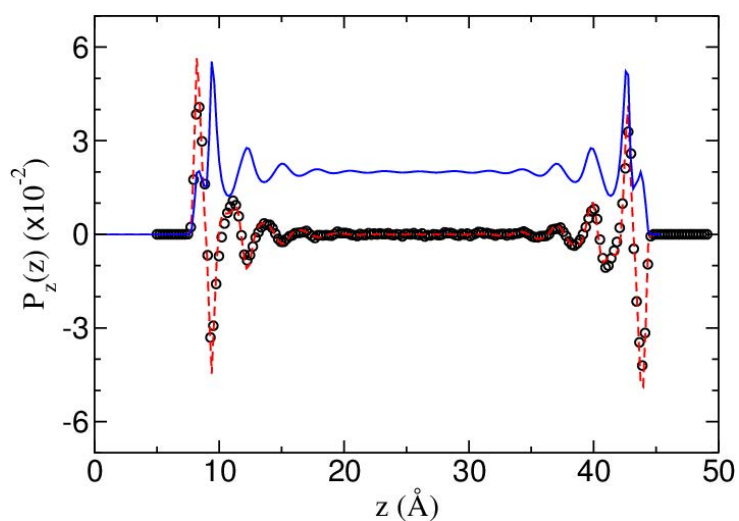
On a désormais une idée claire de la structure tridimensionnelle du solvant à la surface de la pyrophyllite. Quelques molécules de solvant sont adsorbées très près de la surface, au centre des hexagones. Au-dessus de ces molécules se trouve une couche très structurée de molécules de solvant localisées au-dessus des atomes de Si. Enfin, une troisième couche de molécules se trouve au centre des hexagones avec une structure moins bien définie. La distance relativement courte entre les couches est signe d'une interaction et donc d'une cohésion importante entre ces couches. Cette conclusion est en accord avec Rotenberg et collaborateurs qui ont montré récemment que la compétition entre adhésion et cohésion détermine l'hydrophobicité de ces surfaces[104].

**Propriétés orientationnelles :** On s'intéresse désormais aux propriétés orientationnelles des molécules solvatées à la surface de la pyrophyllite. Même si cela est fait de manière routinière, on peut souligner que les propriétés orientationnelles sont particulièrement lentes à sonder par des simulations moléculaires, MD ou MC. En effet, chaque élément de volume doit être échantillonné pour un grand nombre d'angles. La MDFT donne directement accès à ces propriétés puisque la densité spatialement et angulairement dépendante  $\rho(\mathbf{r}, \boldsymbol{\Omega})$  et la densité de polarisation  $\mathbf{P}(\mathbf{r})$  définie dans l'éq. 3.2.8 sont les variables sur lesquelles est conduite la minimisation.

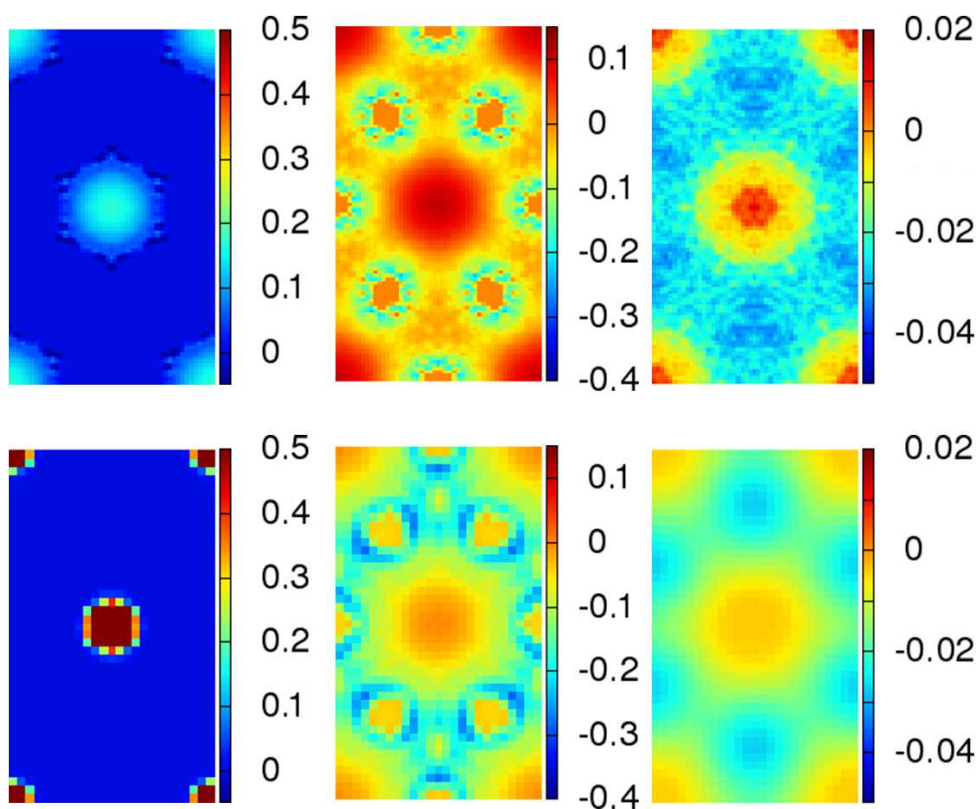
On constate, avec les deux méthodes utilisées, que la densité de polarisation est alignée selon l'axe Oz. Sur la fig. 7.2.4 est présentée la projection  $P_z$  de la polarisation  $\mathbf{P}$  selon l'axe Oz en fonction de la coordonnée z (c'est-à-dire moyennée dans le plan  $(x, y)$ ). Là encore, on trouve un accord quantitatif entre MD et MDFT. Un maxima est trouvé là où se trouve le prépic de densité, un second là où se trouve le premier pic. Entre ces deux maxima, le signe de  $P_z$  change.

On présente sur la fig. 7.2.5 les cartes de la composante selon z de la polarisation dans les plans des trois premiers pics de densité. La densité de polarisation est surestimée par MDFT dans le prépic. Comme l'axe  $O_z$  est toujours orienté vers l'extérieur de la surface, la projection  $P_z$  est positive quand le dipôle de l'eau, qui est orienté de l'oxygène vers le centre de masse des hydrogènes, pointe vers la surface. Dans le premier et le second pic, la polarisation est trouvée légèrement plus grande par MD. Pour essayer d'obtenir plus d'informations sur les propriétés orientationnelles des molécules, on définit deux paramètres d'ordre. Le premier est l'orientation locale de la polarisation définie par le cosinus de l'angle entre le champ de polarisation et la normale à la surface,

$$\cos(\theta_P(\mathbf{r})) = \frac{\mathbf{P}(\mathbf{r}) \cdot \mathbf{z}}{\|\mathbf{P}(\mathbf{r})\|} = \frac{P_z(\mathbf{r})}{\|\mathbf{P}(\mathbf{r})\|}. \quad (7.2.2)$$

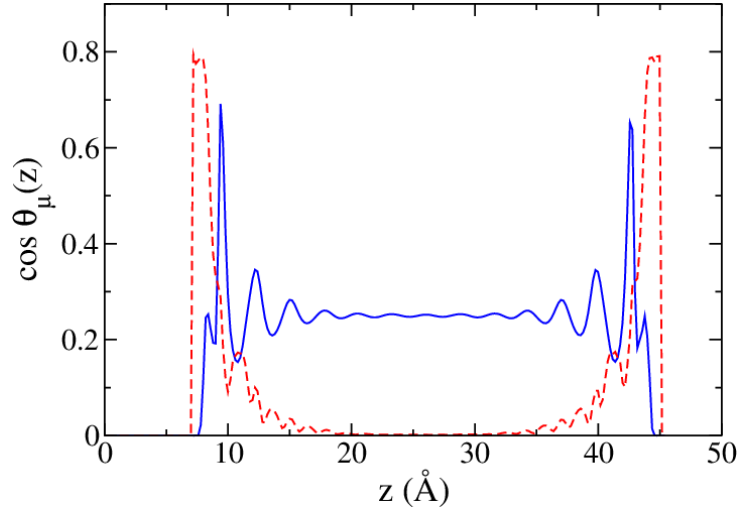


**FIGURE 7.2.4.:** Composante normale au plan de l'argile de la polarisation adimensionnée, entre deux couches de pyrophyllite calculée par MD (cercles noirs) et MDFT (tirets rouges). On rappelle le profil de densité en trait bleu pour repère. Les maxima de polarisation correspondent à des maxima de densité.

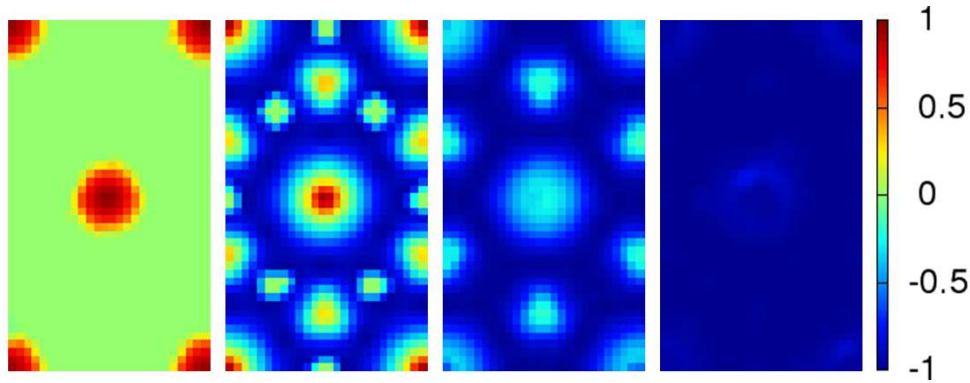


**FIGURE 7.2.5.:** Cartes de la polarisation  $P_z$  dans les plans du prépic (gauche), du pic principal (centre) et du second pic (droite), calculées par MD en haut et par MDFT en bas.





**FIGURE 7.2.6.:** Profil de l'orientation moléculaire moyenne  $\cos(\theta_\mu(z))$  selon la normale au plan de l'argile obtenu par MDFT en tirets rouges. Comme sur la fig. 7.2.5, on a rappelé le profil de densité en bleu.

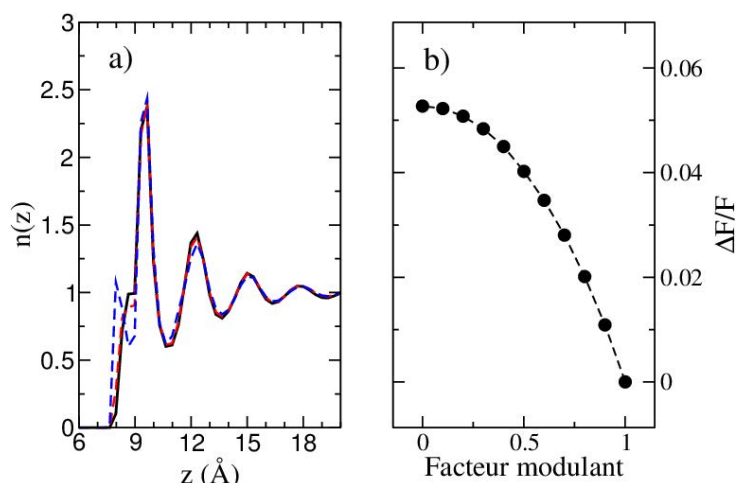


**FIGURE 7.2.7.:** De gauche à droite, paramètre d'ordre  $\cos(\theta_P(\mathbf{r}))$  dans le plan du prépic, dans le plan médian entre prépic et premier pic, dans le plan du premier pic et celui du second pic.

Le second est l'orientation moléculaire moyenne,

$$\cos(\theta_\mu(\mathbf{r})) = \frac{\|P_z(\mathbf{r})\|}{n(\mathbf{r})}. \quad (7.2.3)$$

Quand  $\cos(\theta_P(\mathbf{r}))$  est positif (resp. négatif) le dipôle pointe vers l'extérieur (resp. l'intérieur) de la surface. On représente sur la fig. 7.2.6 la valeur moyenne dans le plan  $(x, y)$  de l'orientation moléculaire moyenne  $\cos(\theta_\mu(z)) = \langle \cos(\theta_\mu(\mathbf{r})) \rangle_{x,y}$ , obtenue par MDFT. Les molécules de solvant sont fortement orientées dans le plan du prépic, avec une valeur élevée de l'orientation moyenne. Cette orientation préférentielle est moins marquée dans le plan du premier pic de solvation. Dans le plan du second, le pic de polarisation correspond à un grand nombre de molécules ayant une faible orientation préférentielle selon l'axe Oz. On peut confirmer cette analyse en regardant des cartes d'orientation locale obtenue par MDFT dans le plan des trois pics, définie par l'éq. 7.2.2. Ces cartes sont présentées sur la fig. 7.2.7. On observe que les molécules dans le plan



**FIGURE 7.2.8.:** À gauche : densités moyennes le long de l'axe perpendiculaire pour des charges correspondant à celles du champ de force CLAYFF (bleu), à moitié éteintes (rouge) et complètement éteintes (noir). À droite, le changement relatif de l'énergie libre de solvation en fonction du facteur modulant les charges, à 0 les charges sont éteintes. Ce changement relatif reste inférieur à 6%.

du prépic sont très orientées, et les dipôles, orientés du site négatif vers le site positif, pointent vers l'extérieur de la surface. L'orientation est plus faible dans le pic principal et les molécules situées au sommet des atomes de silicium sont orientées le dipôle pointant vers l'intérieur de la surface. Dans le plan du second pic, on trouve une orientation préférentielle faible avec un dipôle pointant vers l'extérieur de la surface.

On a maintenant une vision claire des propriétés de solvation de l'argile par le fluide de Stockmayer. Dans le plan du prépic, une molécule de solvant se trouve au centre des hexagones, avec un dipôle figé pointant vers l'extérieur de la surface. Un peu au-dessus, dans le plan du premier pic, les molécules de solvant sont essentiellement situées au dessus des atomes de silicium, avec une orientation préférentielle moins marquée des dipôles vers l'extérieur de la surface. Enfin, la seconde couche de solvation est globalement située au centre des hexagones avec une faible orientation des dipôles vers l'intérieur de la surface.

**Rôle de l'électrostatique :** L'efficacité numérique de la MDFT permet une étude systématique de la structure du solvant et du rôle des interactions de Van-der-Waals et électrostatique dans le champ de force CLAYFF, qui serait trop coûteuse par MD. Ces interactions sont modélisées par des interactions Lennard-Jones et électrostatique dont les paramètres sont données dans le tab. 7.2. On reporte les profils de densité moyenne le long l'axe perpendiculaire à la surface sur la fig. 7.2.8, pour des charges modulées par un facteur variant de 0 (charges nulles) à 1 (charges habituelles du champ de force). Sans surprise, seule la forme du prépic, où les molécules sont les plus fortement polarisées, est modifiée quand on supprime les charges. Sur la fig. 7.2.8 on présente les changements relatifs de l'énergie libre de solvation en fonction du facteur modulant les charges. L'énergie libre de solvation n'est affectée que de 6% quand on éteint les charges. La

contribution principale à l'énergie libre de solvation provient donc des forces de Van-der-Waals. Ce résultat n'est pas évident à prévoir, même pour une surface globalement non chargée.

On a montré que la MDFT est une méthode viable pour étudier des systèmes complexes constitués de plusieurs milliers d'atomes comme les argiles. Elle permet de reproduire quantitativement les prédictions MD concernant les propriétés structurales et orientationnelles du solvant, tout en réduisant le temps CPU de trois ordres de grandeur. La MDFT permet une étude systématique de ces systèmes. En particulier, il devient possible d'étudier l'influence des paramètres de champ de force sur les propriétés de solvation. Les profils de densité et les énergies libres de solvation sont ici peu sensibles à la valeur des charges de l'argile. Le développement de champ de force pourrait ainsi être grandement facilité par l'étude rapide de l'influence des paramètres.

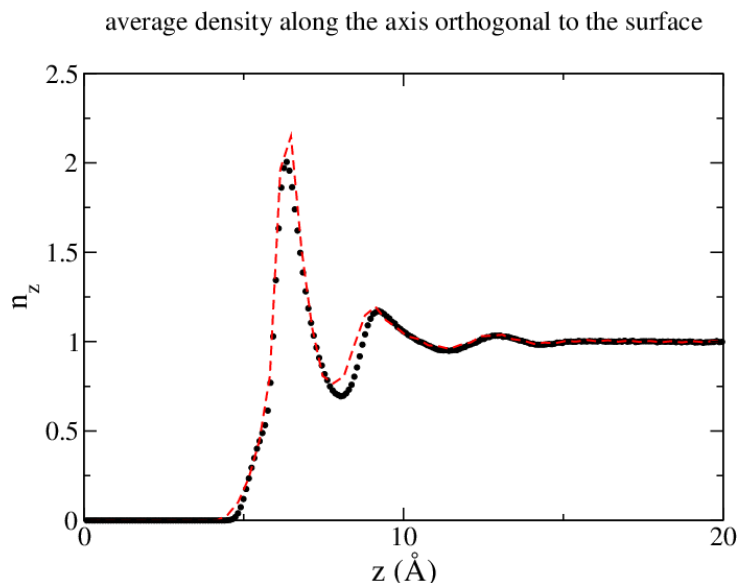
### 7.2.2. Solvation par l'eau

Nous avons étudié la solvation de la même argile, la pyrophyllite, avec le solvant eau SPC/E. On utilise donc la fonctionnelle dont la partie d'excès contient un traitement multipolaire de la polarisation décrite dans la sec.3.3. Pour une grille spatiale avec un maillage comportant 64 points par  $\text{\AA}^3$ , et 30 points pour décrire les orientations, la convergence est atteinte en une quinzaine d'itérations. Le calcul dure environ 25 minutes. C'est environ trois ordres de grandeur plus rapide que la dynamique moléculaire.

#### 7.2.2.1. Résultats

**Profils de densité :** Sur la fig. 7.2.9 est présentée la densité planaire. On observe un accord quantitatif entre MD et MDFT. Ce profil de densité comporte trois pics. Le premier, qui correspond à la première couche de solvation de l'argile, se trouve à 6.3  $\text{\AA}$  de la couche centrale de l'argile et comporte un épaulement à 5.7  $\text{\AA}$ . Il est intéressant de remarquer qu'avec une description de l'eau dipolaire, se limitant aux premiers invariants, on trouve, comme dans le cas du fluide de Stockmayer, que l'on surestime le premier épaulement. Une meilleure description du solvant corrige cet effet. Le second pic, dû à la seconde couche de solvation, se trouve à 2.9  $\text{\AA}$  du premier. Il y a une déplétion entre ces deux premiers pics. Enfin, un troisième pic se situe à  $z = 13.0 \text{\AA}$ . Son l'intensité est pratiquement négligeable mais on le distingue néanmoins en raison des deux zones de déplétion qui l'encadrent. À une distance plus grande de la surface, la densité tend vers celle du solvant homogène  $n_b = 0.033 \text{\AA}^{-3}$ .

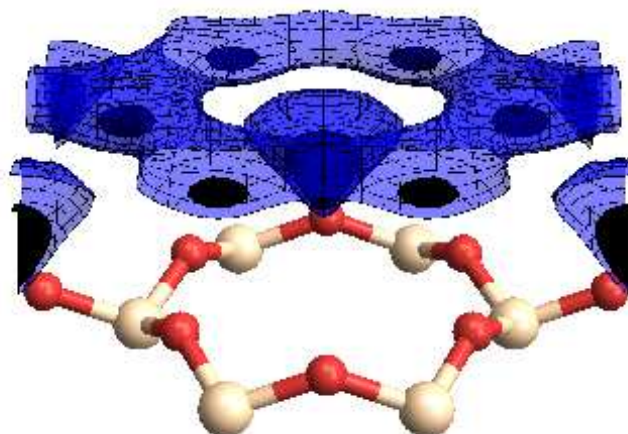
Le calcul de la densité locale  $n(\mathbf{r})$  ou de la densité moyenne  $n(z)$  par MD n'est pas fait de manière exacte. On mesure plutôt  $n(z) \star \Delta z$  et  $n(\mathbf{r}) \star \Delta V$ , où  $\star$  désigne une convolution et  $\Delta z$  et  $\Delta V$  sont des dimensions caractéristiques choisies arbitrairement. Ce sont respectivement la distance et le volume dans lesquels est réalisée la moyenne autour du point d'intérêt. Plus ces dimensions caractéristiques sont petites plus le temps nécessaire pour obtenir une bonne statistique augmente. Ce problème n'existe pas en MDFT puisque l'on n'effectue pas d'échantillonnage. Il est aisé de tracer des cartes d'isodensité, comme celles montrées en fig. 7.2.10, car la densité est la variable



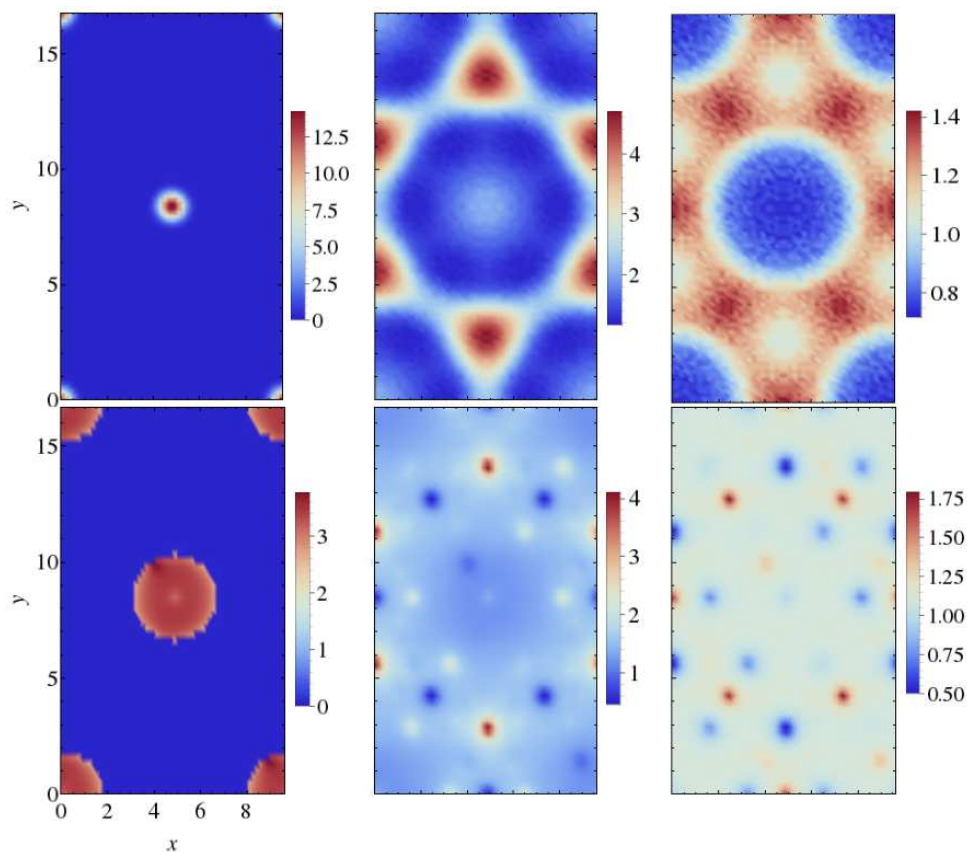
**FIGURE 7.2.9.:** Profil de la densité planaire selon l'axe perpendiculaire aux feuillets, définie à l'éq. 7.2.1. Les résultats calculés par dynamique moléculaire sont les cercles noirs. Ceux calculés par MDFT sont en tirets rouges.

naturelle. Cette carte présente deux surfaces d'isodensité ( $n(\mathbf{r})=2, 4$ ) au dessus d'un hexagone Si-O de la surface. Les zones de plus haute densité se trouvent au centre des hexagones et des atomes de silicium (ces zones sont respectivement responsables de l'épaulement et du premier pic dans le profil de densité de la fig. 7.2.9). Une zone de densité moins forte se trouve au dessus des hexagones, en forme d'anneau (cette zone est responsable du second pic dans le profil de densité).

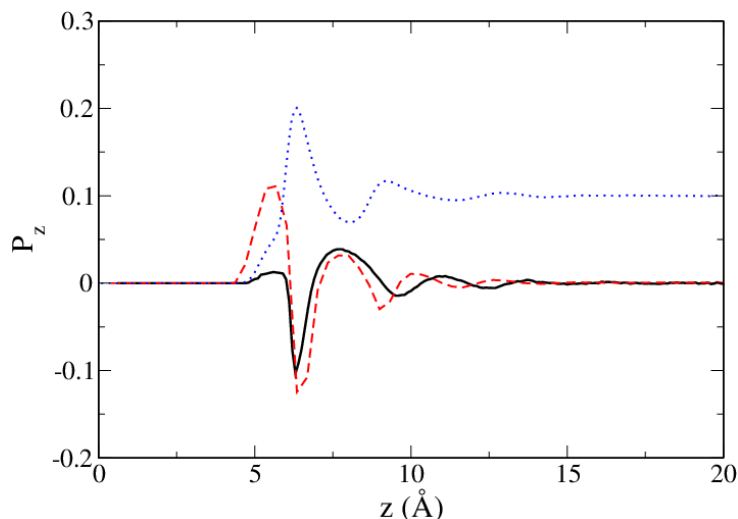
On représente des cartes de densité dans les plans de l'épaulement, du pic principal et du second pic sur la fig. 7.2.11. Dans le plan du prépic, la densité est fortement localisée au centre des hexagones, très proche de la surface hydrophobe. À cette distance l'eau est repoussée partout ailleurs que dans cette petite zone, ce qui explique que lorsqu'on moyenne la densité dans le plan, comme sur la fig. 7.2.9, on observe qu'un épaulement. Dans le plan du premier pic de la fig. 7.2.9, c'est-à-dire dans la première couche de solvation, on trouve une densité très localisée au dessus des atomes de silicium. Pour la seconde couche de solvation, qui correspond au deuxième pic du profil de densité, on observe une densité importante au sommet des atomes d'oxygène et une déplétion au sommet des atomes de silicium. On constate sur la fig. 7.2.11 que l'accord entre MD et MDFT est seulement qualitatif : on observe les mêmes tendances mais les cartes obtenues par MD sont plus diffuses. On peut faire l'hypothèse que les convolutions évoquées plus tôt jouent un rôle ici. On a vérifié que convoluer les profils de densité obtenus par MDFT avec un volume  $\Delta V$  de  $1 \text{ \AA}^3$  permettait d'obtenir des cartes en bien meilleur accord avec celles obtenues par MD. Cette différence illustre l'avantage de la méthode MDFT qui évite un échantillonnage d'autant plus coûteux qu'on veut des cartes précises.



**FIGURE 7.2.10.:** Surfaces d'isodensité d'eau, avec  $n(\mathbf{r})=2$  (en bleu) et 4 (en noir). On représente la structure en hexagone de la pyrophyllite pour aider à la visualisation, les atomes d'oxygène sont en rouge, les atomes de silicium en blanc.



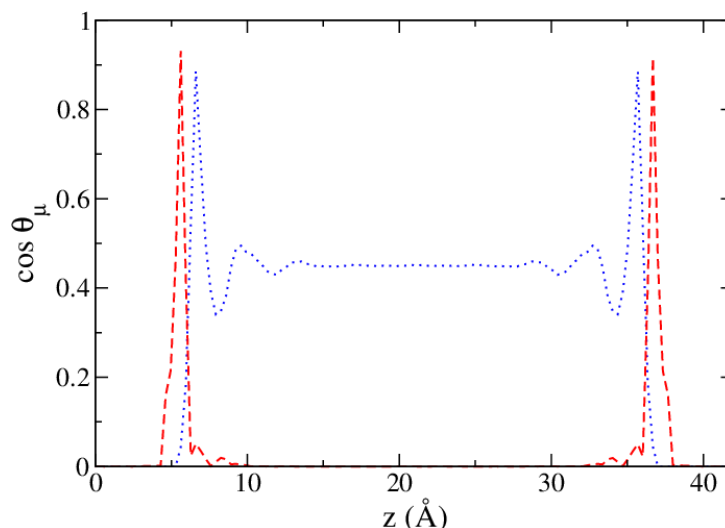
**FIGURE 7.2.11.:** Cartes de densité normalisée  $n(\mathbf{r})/n_b$  de l'eau dans les plans de l'épaulement (gauche), du pic principal (centre) et du second pic (droite). Les résultats calculés par MD sont en haut, ceux calculés par MDFT en bas.



**FIGURE 7.2.12.:** Projections de la polarisation selon l'axe  $z$ , moyennée dans chaque plan  $(x, y)$ . Le profil obtenu par MD est en trait noir plein, celui obtenu par MDFT en tirets rouges. On rappelle le profil de la densité  $n_z$  en pointillés bleu, en unité arbitraire.

**Propriétés orientationnelles :** On s'intéresse maintenant aux propriétés orientationnelles. On présente sur la fig. 7.2.12 la projection  $P_z$  de la polarisation sur l'axe perpendiculaire à la surface de l'argile, moyennée dans chaque plan. Les composantes transverses sont trouvées négligeables ( $P_x/P_z$  et  $P_y/P_z$  de l'ordre de  $10^{-3}$ ), la polarisation est donc alignée dans la direction perpendiculaire à la surface. Les extrema dans le profil de polarisation correspondent à des extrema dans le profil de densité. Un premier maximum de polarisation se trouve à la même abscisse que l'épaulement. Celui-ci est beaucoup plus marqué dans les résultats obtenus par MDFT. Le profil obtenu par MDFT est très semblable à celui obtenu pour le fluide de Stockmayer. Un second extremum, négatif, se trouve dans le plan du premier pic de densité. Entre ces deux extrema le signe de  $P_z$  change. L'intensité de la polarisation macroscopique est liée d'une part à la densité de molécule et d'autre part à leur orientation préférentielle dans la région de l'espace considérée. Ainsi, dans le plan du prépic, la polarisation est positive, c'est donc les atomes d'oxygène des molécules d'eau qui sont globalement les plus proches de la surface. Dans le plan du pic principal, les molécules d'eau sont essentiellement situées au-dessus des silicium. Leurs hydrogènes sont les plus proches de la surface, puisque le pic de polarisation correspondant est négatif. Enfin le second pic de densité est pratiquement non polarisé, il n'y a presque pas d'orientation préférentielle dans cette région de l'espace.

L'orientation moléculaire locale  $\cos \theta_\mu(\mathbf{r})$ , définie à l'éq. 7.2.3 obtenue par MDFT, moyennée dans chaque plan  $(x, y)$  est donnée en fig. 7.2.13. Dans le plan de l'épaulement, les molécules d'eau sont particulièrement orientées. L'atome d'oxygène de la molécule d'eau est le plus proche possible de la surface et de la charge positive des atomes de silicium. Au contraire, dans le plan du pic principal, les molécules d'eau sont globalement orientées dans la direction opposée. La faible valeur de  $\cos \theta_\mu(z)$  correspondant à ce pic principal suggère une orientation préférentielle faible. La seconde couche de solvatation ne possède pas d'orientation préférentielle.



**FIGURE 7.2.13.:** Valeur moyenne de l'orientation moléculaire locale  $\langle \cos \theta_\mu(\mathbf{r}) \rangle_{xy}$  obtenue par MDFT en tirets rouges. Pour faciliter la lecture on rappelle le profil de densité en pointillés bleus en unités arbitraires.

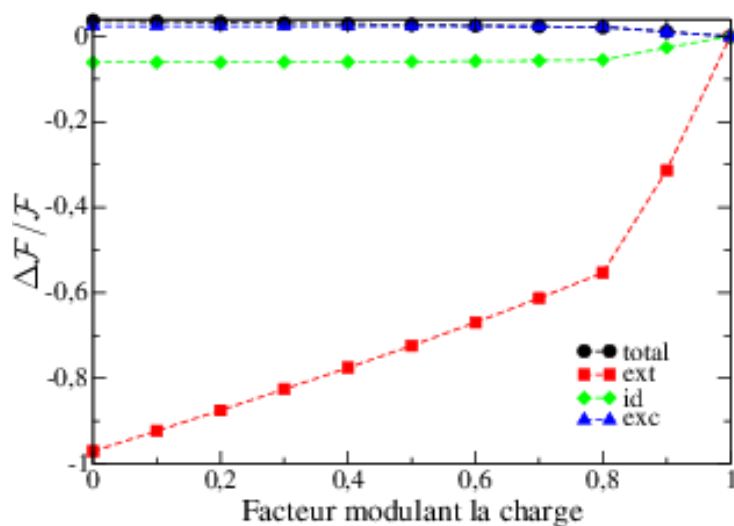
	$\mathcal{F}$	$\mathcal{F}_{\text{ext}}$	$\mathcal{F}_{\text{id}}$	$\mathcal{F}_{\text{exc}(\text{rad})}$	$\mathcal{F}_{\text{exc}(\text{ori}^\circ)}$
$\gamma_{\text{solv}} \text{ (mJ.m}^{-2}\text{)}$	332.4	-12.8	100.4	244.5	0.3

**TABLE 7.3.:** Les différentes composantes de l'énergie libre de solvation par unité de surface de l'argile.

**Propriétés énergétiques :** On peut analyser l'influence des différentes composantes énergétiques en regardant la valeur de chacun des termes de la fonctionnelle utilisée. La partie d'excès est décomposée en une partie radiale et une partie de polarisation. L'énergie libre de solvation par unité de surface de la pyrophyllite  $\gamma_{\text{solv}}$  est donnée au tab.7.3. Elle est de 332.4 mJ.m<sup>-2</sup>. La valeur positive confirme la nature hydrophobe du matériau. La contribution énergétique principale provient de la partie radiale des interactions solvant-solvant. La partie d'interaction soluté-solvant et la partie de polarisation ont une contribution très limitée à l'énergie totale. Ceci ne signifie pas pour autant que la polarisation et la perturbation créées par le soluté n'ont qu'une influence limitée dans le processus de solvation.

En effet, on peut imaginer la solvation d'un mur dur dans un gaz parfait. Le potentiel extérieur infini créé par le mur va évidemment entraîner une densité nulle à l'intérieur du mur. Il en résulte un terme extérieur  $\mathcal{F}_{\text{ext}}$  nul pour la densité d'équilibre. Cependant, c'est la minimisation de ce terme extérieur qui impose un profil de densité différent de celui obtenu en l'absence du mur dur.

**Le rôle de l'électrostatique :** L'évolution de l'énergie libre et de ses composantes avec le facteur modulant les charges de la surface de la pyrophyllite est donnée en fig.7.2.14. L'énergie libre de solvation totale ne change pas de plus de 3% quand on éteint totalement les charges. De plus, les profils de densité ne sont pas modifiés par ce procédé. Quand on éteint les charges, les énergies libres de solvation totale, idéale et la partie radiale du terme d'excès sont peu



**FIGURE 7.2.14.:** Évolution relative de l'énergie libre par unité de surface et de ses composantes, en fonction du facteur modulant la charge. Ces grandeurs sont normalisées par rapport aux charges non modifiées. Le changement relatif de l'énergie de solvation est de 3%.

modifiées en valeur absolue. La partie extérieure change beaucoup relativement mais sa valeur reste faible. Ceci confirme que l'électrostatique a peu d'importance pour étudier la solvation d'argiles globalement neutres.

Nous avons montré que la MDFT est adaptée à l'étude d'un matériaux constitué de plus de mille atomes, même dans un solvant complexe comme l'eau. Ceci peut être réalisé à un coût numérique mille fois plus faible que celui de la MD.

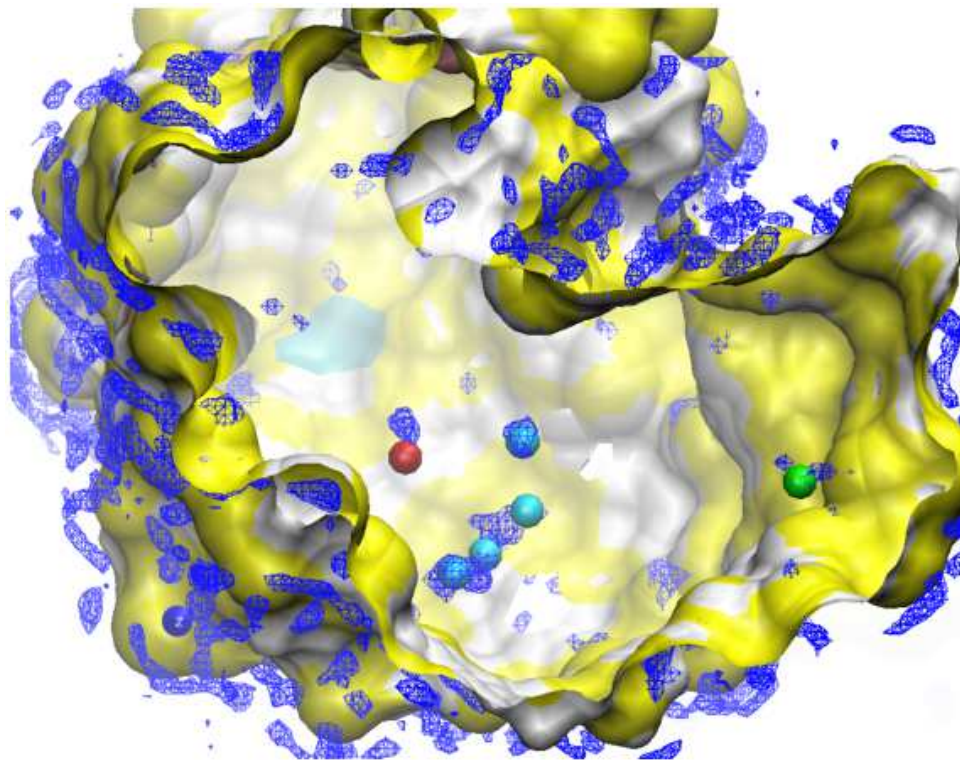


### 7.3. Étude d'une molécule d'intérêt biologique : le lysozyme

Les molécules d'eau confinées dans des cavités jouent un rôle important dans la compréhension de la structure, de la stabilité et des fonctions des biomolécules, notamment des protéines[110, 111]. Ces molécules d'eau internes peuvent être observées par cristallographie[112]. Il est difficile de sonder la présence de molécule d'eau dans des cavités internes de protéines par dynamique moléculaire. Elles sont généralement piégées grâce à des processus impliquant de larges fluctuations de la structure des protéines (repliement et dépliement) sur des échelles de temps trop longues (ms) pour être facilement simulées par MD. Des simulations sur des trajectoires de l'ordre de la microseconde permettant de voir des échanges de molécules d'eau entre l'intérieur de la protéine et le solvant bulk ont néanmoins été publiées récemment[113, 114]. Cependant, ces simulations nécessitent l'utilisation de supercalculateurs et restent très coûteuses numériquement à mettre en œuvre. Hirata et collaborateurs ont montré que l'utilisation de la méthode de solvant implicite 3D-RISM permet de détecter ces molécules piégées dans les protéines. Ces simulations ont un coût numérique beaucoup plus faible puisqu'elles évitent un traitement dynamique de la protéine[32, 115, 33]. Nous avons donc testé la théorie MDFT sur le même exemple que ces auteurs : le lysozyme du blanc d'œuf de poule. La structure 3D de cette protéine (code 1HEL) a été extraite de la protein data bank[116]. Cette protéine est connue pour avoir trois cavités (notées W1, W2 et W3) dans lesquelles sont respectivement détectées par cristallographie quatre, une et aucune molécules d'eau. Il existe d'autres structures cristallographiques[117] de la même protéine qui détectent une molécule d'eau dans la cavité W3 mais avec un facteur B très grand, signe d'une précision très faible quant à la position de cette molécule.

Toutes les molécules d'eau présentes dans la structure cristallographique ont été retirées pour ne conserver comme soluté que la protéine en elle-même. Le potentiel d'interaction de ce soluté a été généré en utilisant les paramètres de champ de force Amber (parm99)[118] en utilisant le programme GROMACS[119] pour ajouter les atomes d'hydrogène qui ne sont pas visibles en cristallographie et absents dans le fichier PDB.

Nous avons utilisé la fonctionnelle multipolaire, sans terme à trois corps, avec le modèle d'eau SPC/E. La minimisation s'est faite avec une boîte de  $64 \text{ \AA}^3$  pour une grille spatiale de  $192^3$  points et une grille angulaire de 24 angles. La minimisation se fait en 38 pas et 3270 secondes sur un seul CPU. On obtient la carte de densité du solvant à l'équilibre. Sur la fig. 7.3.1 est présentée une vue de l'intérieure de la protéine en représentation SURF où les positions des cavités sont représentées par des sphères colorées situées aux coordonnées des molécules d'eau détectées par cristallographie. On a également représenté des isosurfaces de haute densité. On observe que la densité n'est pas uniforme selon que les groupements situés à la surface de la protéine sont plutôt hydrophobes ou hydrophiles. À l'intérieur de la cavité, les seules zones où on rencontre une densité notable sont les zones où on détecte des molécules d'eau internes par cristallographie. Ces densités se trouvent exactement à l'endroit où les molécules d'eau sont détectées pour les cavités W1 et W2. La densité est légèrement décalée pour la cavité W3. Le facteur B pour cette molécule étant très faible, sa position cristallographique est incertaine. La carte de densité tend



**FIGURE 7.3.1.:** Vue de l'intérieur de la protéine lysozyme dans la représentation SURF. Les molécules d'eau visibles par cristallographie sont représentées par des sphères cyans pour la cavité W1, vertes pour la cavité W2 et rouges pour la cavité W3. Les surfaces de haute isodensité  $n(\mathbf{r})/n_b = 10$  obtenues par MDFT sont représentées avec un maillage bleu.

donc à confirmer la présence d'une molécule d'eau interne dans cette cavité.

La MDFT est donc une méthode de choix pour vérifier la présence de molécule de solvant à l'intérieur de cavité de molécules biologiques. Ceci permet de compléter les résultats expérimentaux de cristallographie, à un coût numérique très raisonnable, alors que ces systèmes sont pratiquement hors de portée pour les simulations numériques.

## 8. Conclusions et perspectives

### 8.1. Conclusions

L'objectif de cette thèse est le développement d'une méthode de solvant implicite, la théorie fonctionnelle de la densité moléculaire, MDFT, pour son application à l'étude de la solvation en milieu aqueux. C'est un problème physico-chimique important, généralement appréhendé théoriquement par des simulations moléculaires de type Monte Carlo ou dynamique moléculaire, couplées à des techniques d'intégration thermodynamique. Si ces méthodes donnent en principe des résultats avec une « précision chimique », leur utilisation pour l'étude de solutés complexes présente un coût numérique important. Ceci limite leur utilisation systématique, les systèmes les plus complexes restant à ce jour hors de portée. L'utilisation de méthodes de solvant implicite, basées sur des théories des liquides, comme la MDFT, permet de réduire considérablement ce coût. La méthode MDFT est exacte théoriquement et conserve un niveau de description moléculaire du solvant.

La théorie de la fonctionnelle de la densité classique est basée sur la preuve que le grand potentiel d'un solvant, en présence d'une perturbation créant un champ extérieur, peut s'écrire comme une fonctionnelle unique de la densité de solvant. Cette fonctionnelle étant inconnue, la clé de l'utilisation de cette méthode réside dans la formulation d'une approximation correcte de la fonctionnelle, et plus particulièrement du terme d'excès relatif aux interactions solvant-solvant. Dans cette thèse, nous nous sommes placés dans l'approximation du fluide homogène de référence qui postule que la fonctionnelle d'excès peut s'écrire en fonction de la fonction de corrélation directe du solvant pur.

Pendant cette thèse, j'ai développé la théorie de la fonctionnelle de la densité dans l'approximation du fluide de référence pour étudier la solvation d'un soluté quelconque dans l'eau. Cette théorie existait déjà pour des solvants plus simples, tels que l'acétonitrile ou le fluide de Stockmayer, mais son extension à l'eau a posé des problèmes particuliers, notamment dus au fait que l'eau (représentée par le modèle à trois sites SPC/E) ne possède pas une distribution de charge strictement dipolaire. Ceci a nécessité la formulation d'une fonctionnelle d'excès dépendant de la densité de particule et de la densité de polarisation. L'écriture de cette fonctionnelle fait intervenir le facteur de structure et les composantes transverse et longitudinale des susceptibilités diélectriques. L'utilisation de cette fonctionnelle permet de calculer la structure du solvant à l'équilibre thermodynamique autour d'un soluté quelconque ainsi que l'énergie libre de solvation du soluté. Ces fonctions de corrélation peuvent être obtenues expérimentalement ou calculées par des simulations de dynamique moléculaire. On insiste sur le fait qu'il est nécessaire

de déterminer ces fonctions de corrélation pour le solvant pur, à une densité donnée. Il suffit donc de connaître un jeu de trois fonctions pour chaque solvant pour pouvoir étudier la solvation de n'importe quel soluté dans ce solvant.

Cette théorie a néanmoins montré ses limites, notamment pour l'étude des solutés chargés. La cause de cette limitation provient des approximations réalisées lors du développement de la théorie. La fonctionnelle développée est incapable de reproduire l'ordre tétraédrique local de l'eau dû aux liaisons hydrogènes. Nous avons donc introduit un terme de correction à trois corps qui renforce l'ordre tétraédrique entre le solvant et les solutés. Une implémentation numériquement efficace de cette correction a aussi été réalisée. En incorporant cette correction, la structure du solvant autour des solutés chargés obtenue est très proche des résultats obtenus par des simulations numériques.

Une autre limite de la théorie développée est son incapacité à reproduire correctement le changement de comportement à petite et grande échelle de la structure et de l'énergie de solvation de solutés hydrophobes. Cet effet peut être corrigé par la séparation de la fonctionnelle en une partie courte et longue échelle, et par l'introduction d'une correction longue échelle de la fonctionnelle. Le développement de cette correction longue portée et son implémentation ont constitué une autre partie importante des travaux effectués au cours de cette thèse. Cette formulation prouve qu'il est possible de traiter des problèmes multi-échelles avec la théorie de la fonctionnelle de la densité dans le cadre de l'approximation du fluide homogène de référence.

Une autre partie importante de ce travail de thèse a consisté en l'implémentation de cette théorie dans un code de minimisation numérique de la fonctionnelle. Ce code, nommé `mdft`, utilise une discrétisation de l'espace et des orientations sur une double grille tridimensionnelle. Il est écrit en Fortran moderne.

Nous avons appliqué notre théorie sur un certain nombre de systèmes. Les résultats obtenus sur des solutés simples, atomes, ions et petites molécules sont en accord avec des résultats de référence obtenus par MD qui utilisent les mêmes champs de force. Nous avons également étudié des systèmes complexes. L'étude de la solvation d'une argile hydrophobe par le solvant de Stockmayer et l'eau par la théorie MDFT donne accès aux propriétés structurales et orientationnelles du solvant, en bon accord avec les simulations moléculaires menées sur le même système, tout en étant plus rapide d'environ trois ordres de grandeurs. La MDFT est donc une bonne alternative pour l'étude de ces systèmes. Le gain numérique vis-à-vis de la MD permet de plus une étude énergétique systématique en fonction de paramètres des champs de forces, notamment les charges utilisées pour décrire ces systèmes. La MDFT peut donc aussi être un outil permettant d'optimiser des champs de force moléculaires avec un coût raisonnable.

Nous nous sommes également servi de notre théorie pour étudier un soluté d'intérêt biologique, le lysozyme, et nous avons pu confirmer l'existence de cavités contenant des molécules d'eau à l'intérieur de la protéine. Une telle étude est impossible par des méthodes de simulation car l'entrée et la sortie des molécules d'eau à l'intérieur de la protéine requiert des dépliements et des repliements de la protéine qui ne sont pas observables par MD. L'échelle de temps de ces mouve-

ments étant à ce jour encore hors de portée pour ces techniques. La force de la méthode MDFT réside ici dans le fait que l'équilibre thermodynamique est atteint par minimisation fonctionnelle et non par échantillonnage.

Enfin, la théorie a été utilisée pour calculer des potentiels de force moyenne. La précision des résultats obtenus est comparable à celle obtenue par MD. L'important gain de temps de calcul par rapport à la MD laisse à penser que la MDFT pourrait devenir un outil standard pour calculer des PMF, qui sont des observables très utiles notamment pour la paramétrisation de champs de force gros-grains.

La construction de la fonctionnelle proposée est également intéressante pour la compréhension des phénomènes physiques mis en jeu dans la solvation. En effet, on peut rajouter différentes corrections et voir leurs effets sur la solvation. Ceci permet de voir quels phénomènes physiques, modélisés par ces corrections gouvernent les propriétés de solvation des systèmes étudiés. Il est plus compliqué de rationaliser ce lien entre théorie et résultats pour les simulations moléculaires puisque les seules choses que l'on peut modifier sont les paramètres des champs de force utilisés dont l'effet sur les propriétés d'intérêt est très sensible.

## 8.2. Perspectives

Si les résultats obtenus pendant cette thèse sont encourageants, il reste néanmoins de nombreuses tâches à accomplir avant de faire de la MDFT, et du code correspondant, une méthode standard. Tout d'abord, il est prévu d'utiliser une fonction de corrélation du solvant utilisant l'ensemble des projections sur les invariants rotationnels. Pour cela, une collaboration avec Luc Bellonni du CEA, qui est un expert du calcul de ces projections, a commencé en octobre 2013, sous forme de la thèse de Lu Ding. Ceci permettra d'avoir une bien meilleure description des interactions entre molécules de solvant et permettra peut-être de s'affranchir des différentes corrections : terme à trois corps, bridge de sphères dures et hydrophobicité à longue portée. Un point théoriquement capital, soulevé par Bob Evans, est de vérifier en quel endroit du diagramme de phase de l'eau se trouve le solvant dans notre théorie. Ceci peut se faire en vérifiant un certain nombre de règles thermodynamiques, telle que la règle d'adsorption de Gibbs et les règles de transition de phase. Un bon point de départ serait la construction d'un modèle d'eau spécifique à la MDFT, au lieu de prendre ceux couramment utilisés en MD. On pense, par exemple, partir d'un fluide de sphères dures, auquel serait ajouté une terme de polarisation, un terme à trois corps et une perturbation Lennard-Jones. Il est également prévu d'introduire un terme de polarisabilité électrostatique du solvant.

Enfin, l'objectif ultime, qui était à l'origine le but de cette thèse et qui s'est révélé irréalisable pendant ces trois ans faute de temps, est un couplage entre DFT électronique et MDFT. Ceci permettrait d'utiliser la théorie MDFT pour traiter le solvant tout en ayant une description quantique du soluté à un coût bien plus faible que les méthodes QM/MM et avec une précision supérieure aux méthodes utilisant un continuum comme PCM. La partie la plus technique résidera dans l'écriture d'un potentiel de couplage entre la partie classique et la partie quantique.

On se propose comme point de départ de réaliser des minimisations successives avec un couplage nul. D'abord d'eDFT pour obtenir une densité électronique, puis de se servir de cette densité électronique comme potentiel extérieur pour trouver la densité de solvant d'équilibre par MDFT. On pourra se servir du champ électrique créé par cette densité de solvant comme potentiel extérieur pour mener une nouvelle minimisation eDFT, et itérer le processus jusqu'à atteindre une convergence des densités moléculaire et électronique calculées par les deux méthodes.

Enfin, une perspective primordiale est la mise à disposition à moyen terme du programme mdft à l'ensemble de la communauté.

---

Ci-dessous, l'ensemble des articles publiés ou acceptés sur le sujet de ce manuscrit avant sa rédaction, le premier juin 2014.

M. Levesque, V. Marry, B. Rotenberg, G. Jeanmairet, R. Vuilleumier, D. Borgis. (2012) Solvation of complex surfaces via molecular density functional theory. *J. Chem. Phys.* **137**, 224107.

G. Jeanmairet, M. Levesque, R. Vuilleumier, D. Borgis. (2013) Molecular Density Functional Theory of Water, *J. Phys. Chem. Lett.* **4**, 619-624 .

G. Jeanmairet, M. Levesque, D. Borgis. (2013) Molecular Density Functional Theory of Water describing Hydrophobicity at Short and Long Length Scales, *J. Chem. Phys.* **139**, 154101.

G. Jeanmairet, V. Marry, M. Levesque, B. Rotenberg, D. Borgis. (2014) Hydration of Clays at the Molecular Scale : The Promising Perspective of Classical Density Functional Theory, *Mol. Phys.* **112**, 1320-1329.

V. P. Sergiievskyi, G. Jeanmairet, M. Levesque, D. Borgis. (2014) Fast Computation of Solvation Free Energies with Molecular Density Functional Theory : Thermodynamic-Ensemble Partial Molar Volume Corrections, *J. Phys. Chem. Lett.* **5**, 1935-1942.

G. Jeanmairet, N. Levy, M. Levesque, D. Borgis. (2014) Introduction to Classical Density Functional Theory by Computational Experiment, *arXiv preprint* arXiv :1401.1679.

Plusieurs autres articles sont soumis ou en cours de rédaction.

Quatrième partie .

Appendices



# A. Notions de fonctionnelle et de dérivation fonctionnelle

## A.1. Fonctionnelle

On appellera fonctionnelle  $\mathcal{F}$  une fonction prenant comme argument une autre fonction  $f$ . La fonctionnelle renvoie un scalaire (on se limitera aux scalaires réels) en fonction de la valeur prise par la fonction  $f$ , sur un intervalle  $I$ . Par exemple,

$$\mathcal{F}_{\text{id}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \left( \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \rho(\mathbf{r}) + \rho_b \right) d\mathbf{r}, \quad (\text{A.1.1})$$

où  $\mathcal{F}_{\text{id}}$  est une fonctionnelle de la densité  $\rho$ , elle-même fonction des coordonnées cartésiennes. La valeur de la fonctionnelle est liée à celle de la densité sur l'ensemble de son domaine de définition, ici  $\mathbb{R}^3$ .

## A.2. Dérivée fonctionnelle

Soit la fonction  $f(\mathbf{x})$  où  $\mathbf{x} = x_1, x_2, \dots, x_n$  est un vecteur à  $n$  dimensions. La variation de  $f$  due à une variation infinitésimale de  $\mathbf{x}$  s'écrit

$$df = f(\mathbf{x} + d\mathbf{x}) - f(\mathbf{x}) = \sum_{i=1}^n f'_i(\mathbf{x}) dx_i, \quad (\text{A.2.1})$$

avec,

$$f'_i(\mathbf{x}) = \frac{\partial f}{\partial x_i}. \quad (\text{A.2.2})$$

De même, si  $\mathcal{A}$  est une fonctionnelle de  $f(\mathbf{x})$ , c'est-à-dire qui dépend de la valeur de  $f$  sur un intervalle  $[a; b]$ , la variation de la fonctionnelle s'exprime comme

$$\delta \mathcal{A} = \mathcal{A}[f(\mathbf{x}) + \delta f(\mathbf{x})] - \mathcal{A}[f(\mathbf{x})] = \int_a^b \mathcal{A}'[f(\mathbf{x})] \delta f(\mathbf{x}) d\mathbf{x} \quad (\text{A.2.3})$$

avec la dérivée fonctionnelle

$$\mathcal{A}'[f(\mathbf{x})] = \frac{\delta \mathcal{A}}{\delta f(\mathbf{x})}. \quad (\text{A.2.4})$$

Illustrons ces notions sur un cas utile en DFT classique, par exemple la contribution idéale à la fonctionnelle d'énergie libre de l'éq. A.1.1.

$$\begin{aligned} \beta \delta \mathcal{F}_{\text{id}} &= \beta \mathcal{F}_{\text{id}}[\rho(\mathbf{r}) + \delta \rho(\mathbf{r})] - \beta \mathcal{F}_{\text{id}}[\rho(\mathbf{r})] \\ &= \iiint_{\mathbb{R}^3} \left[ (\rho(\mathbf{r}) + \delta \rho(\mathbf{r})) \ln \left( \frac{\rho(\mathbf{r}) + \delta \rho(\mathbf{r})}{\rho_b} \right) - \rho(\mathbf{r}) + \delta \rho(\mathbf{r}) + \rho_b \right] d\mathbf{r} \\ &\quad - \iiint_{\mathbb{R}^3} \left( \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \rho(\mathbf{r}) + \rho_b \right) d\mathbf{r} \\ &= \iiint_{\mathbb{R}^3} \left[ (\rho(\mathbf{r}) + \delta \rho(\mathbf{r})) \left[ \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) + \frac{\delta \rho(\mathbf{r})}{\rho(\mathbf{r})} + \mathcal{O}(\delta \rho(\mathbf{r})^2) \right] - \rho(\mathbf{r}) - \delta \rho(\mathbf{r}) + \rho_b \right] d\mathbf{r} \\ &\quad - \iiint_{\mathbb{R}^3} \left( \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) - \rho(\mathbf{r}) + \rho_b \right) d\mathbf{r} \\ &= \iiint_{\mathbb{R}^3} \left( \delta \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) + \delta \rho(\mathbf{r}) + \mathcal{O}(\delta \rho(\mathbf{r})^2) - \delta \rho(\mathbf{r}) \right) d\mathbf{r} \\ &= \iiint_{\mathbb{R}^3} \left( \delta \rho(\mathbf{r}) \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right) + \mathcal{O}(\delta \rho(\mathbf{r})^2) \right) d\mathbf{r} \end{aligned}$$

qui, par identification avec l'éq. A.2.4, donne la dérivée fonctionnelle de la partie idéale,

$$\beta \frac{\delta \mathcal{F}_{\text{id}}}{\delta \rho(\mathbf{r})} = \ln \left( \frac{\rho(\mathbf{r})}{\rho_b} \right). \quad (\text{A.2.5})$$

## B. Démonstration : Le grand potentiel est une fonctionnelle des densités de particule et de polarisation

### B.1. La densité qui minimise le grand potentiel est la densité d'équilibre

Cette démonstration a déjà été réalisée en sec. 2.2.2 et ne supposait pas d'écriture particulière du hamiltonien. Les résultats obtenus sont toujours vrais pour le hamiltonien de l'éq. 3.3.6. On a donc :<sup>1</sup>

$$\Theta[f] = \text{Tr}_{\text{cl}} [f(\mathcal{H}_N - \mu N + \beta^{-1} \ln f)], \quad (\text{B.1.1})$$

$$\Theta[f_0] = -k_B T \ln \Xi = \Theta, \quad (\text{B.1.2})$$

$$\Theta[f] > \Theta[f_0], \text{ si } f \neq f_0. \quad (\text{B.1.3})$$

### B.2. L'énergie libre intrinsèque est une fonctionnelle unique des densités de particule et de polarisation

On veut introduire les deux nouvelles fonctionnelles,

$$\mathcal{F}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \text{Tr}_{\text{cl}} [f_0 (\mathcal{K} + \mathcal{U} + \beta^{-1} \ln f_0)], \quad (\text{B.2.1})$$

et

$$\Theta_V[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \mathcal{F}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})] + \iiint_{\mathbb{R}^3} \Psi_n(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + - \iiint_{\mathbb{R}^3} \mathbf{P}(\mathbf{r}) \mathbf{E}(\mathbf{r}) d\mathbf{r}, \quad (\text{B.2.2})$$

---

1. On rappelle l'expression de la trace classique :

$$\text{Tr}_{\text{cl}} = \sum_{N=0}^{\infty} \frac{1}{h^{3N} N!} \int \cdots \int_{\mathbb{R}^{3N}} \int \cdots \int_{\mathbb{R}^{3N}} d\mathbf{r} \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N$$

où  $\Psi_n(\mathbf{r}) = \Phi_n(\mathbf{r}) - \mu$ . L'indice  $\mathcal{V}$  dans l'éq. B.2.2 rappelle que cette fonctionnelle dépend de la forme choisie pour le potentiel extérieur, c'est donc une notation compacte du couple  $(\phi_n, \mathbf{E})$ . Nous allons montrer qu'il est possible d'écrire ces deux fonctionnelles.

### B.3. Démonstration : $\mathcal{F}$ est une fonctionnelle unique de $n(\mathbf{r})$ et $\mathbf{P}(\mathbf{r})$

La densité de probabilité d'équilibre est une fonction des champs extérieurs  $\Phi_n$  et  $\mathbf{E}$ . Puisque

$$n_0(\mathbf{r}) = \text{Tr}_{\text{cl}} [f_0 \tilde{n}(\mathbf{r})] \quad (\text{B.3.1})$$

et

$$\mathbf{P}_0(\mathbf{r}) = \text{Tr}_{\text{cl}} [f_0 \tilde{\mathbf{P}}(\mathbf{r})], \quad (\text{B.3.2})$$

$n_0(\mathbf{r})$  et  $\mathbf{P}_0(\mathbf{r})$  à l'équilibre sont eux aussi des fonctionnelles de ces champs extérieurs<sup>2</sup>. Supposons qu'il existe deux couples de champs extérieurs différents qui engendrent les mêmes densités de particule et de polarisation à l'équilibre  $(n(\mathbf{r}), \mathbf{P}(\mathbf{r}))$ . Au hamiltonien  $\mathcal{H}_N$  on associe la distribution de probabilité à l'équilibre  $f_0$  et le grand potentiel  $\Theta$ . Au hamiltonien  $\mathcal{H}'_N$  on associe la distribution de probabilité à l'équilibre  $f'_0$  et le grand potentiel  $\Theta'$ . D'après l'éq. B.1.3, on a

$$\Theta' = \text{Tr}_{\text{cl}} f'_0 \left[ (\mathcal{H}'_N - \mu N + \beta^{-1} \ln f'_0) \right] < \Theta [f_0], \quad (\text{B.3.3})$$

$$\Theta' = \text{Tr}_{\text{cl}} \left[ f'_0 (\mathcal{H}'_N - \mu N + \beta^{-1} \ln f'_0) \right] < \text{Tr}_{\text{cl}} \left[ f_0 (\mathcal{H}'_N - \mu N + \beta^{-1} \ln f_0) \right], \quad (\text{B.3.4})$$

$$\Theta' < \Theta + \iiint_{\mathbb{R}^3} n_0(\mathbf{r}) (\phi'_N(\mathbf{r}) - \phi_N(\mathbf{r})) d\mathbf{r} - \iiint_{\mathbb{R}^3} \mathbf{P}_0(\mathbf{r}) \cdot (\mathbf{E}'(\mathbf{r}) - \mathbf{E}(\mathbf{r})) d\mathbf{r}. \quad (\text{B.3.5})$$

En échangeant les rôles de  $\Theta$  et  $\Theta'$ , on trouve

$$\Theta < \Theta' + \iiint_{\mathbb{R}^3} n_0(\mathbf{r}) (\phi_N(\mathbf{r}) - \phi'_N(\mathbf{r})) d\mathbf{r} - \iiint_{\mathbb{R}^3} \mathbf{P}_0(\mathbf{r}) \cdot (\mathbf{E}(\mathbf{r}) - \mathbf{E}'(\mathbf{r})) d\mathbf{r}. \quad (\text{B.3.6})$$

En sommant l'éq. B.3.5 et l'éq. B.3.6 on obtient  $\Theta + \Theta' < \Theta' + \Theta$ , ce qui est absurde. À un couple de densité de particule et de polarisation n'est donc associé qu'un unique potentiel extérieur. Le potentiel extérieur est déterminé de manière unique par les densités d'équilibre et réciproquement. Il en suit que la distribution de probabilité est une fonctionnelle unique des densités d'équilibre. La fonctionnelle donnée à l'éq. B.2.1 est donc une unique fonctionnelle des densités d'équilibre  $n$  et  $\mathbf{P}$ . Elle est dite universelle car elle possède la même expression pour tous les potentiels extérieurs.

Montrons que la fonctionnelle introduite en éq. B.2.2 est égale au grand potentiel à son minimum.

---

2. On rappelle que les  $\sim$  désignent les fonctions microscopiques.

Pour les densités d'équilibre  $n_0(\mathbf{r})$  et  $\mathbf{P}_0(\mathbf{r})$ , cette fonctionnelle est égale au grand potentiel,

$$\Theta_{\mathcal{V}}[n_0(\mathbf{r}), \mathbf{P}(\mathbf{r})] = \Theta. \quad (\text{B.3.7})$$

Supposons qu'il existe des densités de particule et polarisation  $n'(\mathbf{r})$  et  $\mathbf{P}'(\mathbf{r})$  associées à une distribution de probabilité  $f'$ .

$$\begin{aligned} \Theta[f'] &= \text{Tr}_{\text{cl}} [f'(\mathcal{H}_N - \mu N + \beta^{-1} \ln f')] \\ &= \mathcal{F}[n'(\mathbf{r}), \mathbf{P}'(\mathbf{r})] + \iiint_{\mathbb{R}^3} \Psi_n(\mathbf{r}) n'(\mathbf{r}) d\mathbf{r} + \iiint_{\mathbb{R}^3} \mathbf{P}'(\mathbf{r}) \mathbf{E}(\mathbf{r}) d\mathbf{r} \\ &= \Theta_{\mathcal{V}}[n'(\mathbf{r}), \mathbf{P}'(\mathbf{r})]. \end{aligned}$$

En utilisant l'éq. B.1.3, on a alors

$$\Theta_{\mathcal{V}}[n'(\mathbf{r}), \mathbf{P}'(\mathbf{r})] > \Theta_{\mathcal{V}}[n_0(\mathbf{r}), \mathbf{P}_0(\mathbf{r})]. \quad (\text{B.3.8})$$

Ainsi, les densités d'équilibre  $n_0(\mathbf{r})$  et  $\mathbf{P}_0(\mathbf{r})$  minimisent la fonctionnelle  $\Theta_{\mathcal{V}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]$ , ce que l'on peut noter :

$$\left. \frac{\delta \Theta_{\mathcal{V}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]}{\delta n(\mathbf{r})} \right|_{n_0} = 0 \text{ et } \left. \frac{\delta \Theta_{\mathcal{V}}[n(\mathbf{r}), \mathbf{P}(\mathbf{r})]}{\delta \mathbf{P}(\mathbf{r})} \right|_{\mathbf{P}_0} = 0. \quad (\text{B.3.9})$$

## C. Fonctionnelle de la densité du fluide de sphères dures

Nous avons implémenté une fonctionnelle pour le fluide de sphères dures. La théorie de la mesure fondamentale (FMT) a été introduite par Rosenfeld, nous la notons R-FMT[69]. Cette fonctionnelle est une excellente approximation de la fonctionnelle d'excès pour un fluide composé de sphères dures de rayon  $R_0$ . Elle nécessite le calcul de densités convoluées avec des fonctions poids :

$$\mathcal{F}_{\text{exc}}[\rho(\mathbf{r})] = k_B T \iiint_{\mathbb{R}^3} \Phi(\{n_\alpha(\mathbf{r})\}) d\mathbf{r}, \quad (\text{C.0.1})$$

avec,

$$n_\alpha(\mathbf{r}) = \iiint_{\mathbb{R}^3} \rho(\mathbf{r}') \omega_\alpha(\mathbf{r} - \mathbf{r}') d\mathbf{r}' = \rho(\mathbf{r}) \star \omega_\alpha(\mathbf{r}), \quad (\text{C.0.2})$$

où les  $\omega_\alpha$  désignent des fonctions poids géométriques que l'on définit plus bas,  $\star$  désigne le produit de convolution et  $\Phi$  est une densité d'énergie libre.

On donne la dérivée de cette énergie libre par rapport à la densité :

$$\begin{aligned} \frac{\delta \mathcal{F}_{\text{exc}}}{\delta \rho(\mathbf{r})} &= k_B T \sum_{\alpha=1}^{N_w} \iiint_{\mathbb{R}^3} \frac{\delta \Phi}{\delta n_\alpha(\mathbf{r}')} \frac{\delta n_\alpha(\mathbf{r}')}{\delta \rho(\mathbf{r})} d\mathbf{r}' = k_B T \sum_{\alpha=1}^{N_w} \iiint_{\mathbb{R}^3} \frac{\delta \Phi}{\delta n_\alpha(\mathbf{r}')} \omega_\alpha(\mathbf{r} - \mathbf{r}') d\mathbf{r}' \\ &= k_B T \sum_{\alpha=1}^{N_w} \frac{\delta \Phi}{\delta n_\alpha(\mathbf{r}')} \star \omega_\alpha(\mathbf{r}) \end{aligned} \quad (\text{C.0.3})$$

qui s'exprime comme une somme des dérivées partielles de  $\Phi$  par rapport aux fonctions poids.

Kierlik et Rosinberg ont dérivé une version alternative plus simple de FMT, que nous notons KR-FMT. Phan[120] et collaborateurs ont démontré que la formulation KR-FMT est équivalente à la formulation vectorielle R-FMT. Cette formulation fait intervenir moins de fonctions poids que la formulation de Rosenfeld. On calcule ces fonctions poids, qui sont des produits de convolution, par FFT. La formulation de KR diminue donc le nombre de FFT à réaliser. C'est donc la version que nous avons implémenté dans le code mdfft[53].

Elle nécessite quatre fonctions poids scalaires  $\omega_\alpha$  avec  $\alpha = 0, 1, 2, 3$ .

$$\omega_0(r) = -\frac{1}{8\pi}\delta''(R_0 - r) + \frac{1}{2\pi r}\delta'(R_0 - r), \quad (\text{C.0.4})$$

$$\omega_1(r) = \frac{1}{8\pi}\delta'(R_0 - r), \quad (\text{C.0.5})$$

$$\omega_2(\mathbf{r}) = 4\pi R_0 \omega_1(\mathbf{r}) = 4\pi R_0^2 \omega_0(\mathbf{r}) = \delta(R_0 - r), \quad (\text{C.0.6})$$

$$\omega_3(\mathbf{r}) = \Theta(R_0 - r), \quad (\text{C.0.7})$$

où  $\Theta$  désigne la fonction de Heaviside et  $\delta$  la distribution de Dirac. Cette formulation peut être utilisée avec une densité d'énergie libre  $\Phi$  dérivant de la théorie de Percus-Yevick ou de Carnahan-Starling. Les deux densités d'énergie libre, celle de Percus-Yevick (PY)

$$\Phi^{\text{PY}}[n_\alpha] = -n_0 \ln(1 - n_3) + \frac{n_1 n_2}{1 - n_3} + \frac{1}{24\pi} \frac{n_2^3}{(1 - n_3)^2} \quad (\text{C.0.8})$$

et celle de Carnahan-Starling (CS)

$$\Phi^{\text{CS}}[n_\alpha] = \left( \frac{1}{36\pi} \frac{n_2^3}{n_3^2} - n_0 \right) \ln(1 - n_3) + \frac{n_1 n_2}{1 - n_3} + \frac{1}{36\pi} \frac{n_2^3}{(1 - n_3)^2 n_3}, \quad (\text{C.0.9})$$

sont implémentées dans le code. Kierlik et Rosinberg ont montré que CS donne des résultats plus précis mais que son utilisation sans modification des fonctions poids cause une inconsistance thermodynamique. Dans la pratique, nous utilisons toujours la relation de CS.

Le potentiel chimique du fluide de sphères dures et la fonction de corrélation directe peuvent être calculés par dérivation fonctionnelle,

$$\mu_{\text{exc}}^{\text{HS}} = \left. \frac{\delta \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right|_{\rho(\mathbf{r})=\rho_b}, \quad (\text{C.0.10})$$

et

$$c_{000}^{\text{HS}}(r; \rho_b) = - \left. \frac{\delta^2 \mathcal{F}_{\text{exc}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')} \right|_{\rho(\mathbf{r})=\rho_b}. \quad (\text{C.0.11})$$

Ces fonctions sont à calculer si on veut utiliser le bridge de sphères dures décrit au chapitre 3. En pratique, la fonction de corrélation directe est calculée efficacement dans l'espace de Fourier,

$$c_{000}^{\text{HS}}(k; \rho_b) = - \sum_{\alpha, \beta} \frac{\partial^2 \Phi}{\partial n_\alpha \partial n_\beta}(\{n_\gamma^b\}) \hat{\omega}_\alpha(k) \hat{\omega}_\beta(k), \quad (\text{C.0.12})$$

où  $\{n_\gamma^b\}$  représente l'ensemble des fonctions pondérées pour le fluide de sphères dures homogène de densité  $\rho_b$ ; et  $\hat{\omega}_\alpha(k)$  les transformées de Fourier des fonctions poids.

## D. Dérivées des fonctionnelles

Les dérivées des différentes parties des fonctionnelles utilisées dans ce manuscrit sont rassemblées ici.

La dérivée de la partie idéale de l'éq. 3.2.2 est

$$\frac{\delta \mathcal{F}_{\text{id}} [\rho(\mathbf{r}, \boldsymbol{\Omega})]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} = k_B T \ln \left( \frac{\rho(\mathbf{r}, \boldsymbol{\Omega})}{\rho_b} \right). \quad (\text{D.0.1})$$

Celle de la partie extérieure donnée en éq. 3.2.3 est

$$\frac{\delta \mathcal{F}_{\text{ext}} [\rho(\mathbf{r}, \boldsymbol{\Omega})]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} = v_{\text{ext}}(\mathbf{r}, \boldsymbol{\Omega}). \quad (\text{D.0.2})$$

Ces deux termes ne posent aucun problème particulier pour être calculés. La dérivée de la partie d'excès dipolaire de l'équation éq. 3.2.6 est

$$\begin{aligned} \frac{\delta \mathcal{F}_{\text{exc}} [\rho(\mathbf{r}, \boldsymbol{\Omega})]}{\delta \rho(\mathbf{r}_2, \boldsymbol{\Omega})} &= k_B T \iiint_{\mathbb{R}^3} [c_{000}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_1)] d\mathbf{r}_1 \\ &- \frac{1}{2\mu} k_B T \iiint_{\mathbb{R}^3} [c_{101}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_1)] d\mathbf{r}_1 \\ &- \frac{1}{2} k_B T \iiint_{\mathbb{R}^3} [c_{101}(\|\mathbf{r}_{12}\|) \mathbf{P}(\mathbf{r}_2) \cdot \tilde{\mathbf{u}}_{12}] d\mathbf{r}_1 \\ &+ \frac{1}{2\mu} k_B T \iiint_{\mathbb{R}^3} [c_{011}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_1)] d\mathbf{r}_1 \\ &+ \frac{1}{2} k_B T \iiint_{\mathbb{R}^3} [c_{011}(\|\mathbf{r}_{12}\|) \mathbf{P}(\mathbf{r}_2) \cdot \tilde{\mathbf{u}}_{12}] d\mathbf{r}_1 \\ &- \frac{1}{\mu} k_B T \iiint_{\mathbb{R}^3} [c_{112}(\|\mathbf{r}_{12}\|) 3(\mathbf{P}(\mathbf{r}_1) \cdot \tilde{\mathbf{u}}_{12}) - \mathbf{P}(\mathbf{r}_1)] d\mathbf{r}_1 \\ &+ \frac{\delta \mathcal{F}_{\text{cor}} [\rho(\mathbf{r}, \boldsymbol{\Omega})]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} \end{aligned} \quad (\text{D.0.3})$$

$$+ \frac{\delta \mathcal{F}_{\text{cor}} [\rho(\mathbf{r}, \boldsymbol{\Omega})]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} \quad (\text{D.0.4})$$

Ces gradients sont des produits de convolution faciles à calculer par FFT, méthode employée de manière systématique dans le code.

La dérivée de la partie d'excès multipolaire de l'éq. 3.3.12 peut se décomposer en la somme de sa dérivée par rapport à la densité et celle par rapport à la polarisation,

$$\beta \frac{\delta \mathcal{F}_{\text{exc}} [\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})]}{\delta \Delta n(\mathbf{r}_2)} = \frac{1}{2} \iiint_{\mathbb{R}^3} S^{-1}(\|\mathbf{r}_{12}\|) \Delta n(\mathbf{r}_2) d\mathbf{r}_2 - \iiint_{\mathbb{R}^3} \frac{\Delta n(\mathbf{r}_2)}{n_0} d\mathbf{r}_2 + \frac{\delta \beta \mathcal{F}_{\text{corr}} [\Delta n(\mathbf{r}), \mathbf{P}(\mathbf{r})]}{\delta n(\mathbf{r}_2)}.$$



(D.0.5)

La dérivée par rapport à la polarisation est beaucoup plus facile à exprimer directement dans l'espace de Fourier,

$$\begin{aligned} \beta \frac{\delta \mathcal{F}_{\text{exc}}[\Delta \hat{n}(\mathbf{k}), \hat{\mathbf{P}}(\mathbf{k})]}{\delta \hat{\mathbf{P}}(\mathbf{k})} &= \frac{3}{\mu_0 n_0^2} \hat{\mathbf{P}}(\mathbf{k}) \hat{\mu}(\mathbf{k}, \Omega) + \frac{1}{8\pi\epsilon_0} \hat{\chi}_L^{-1}(k) \hat{\mathbf{P}}_L(\mathbf{k}) \left( \hat{\mathbf{P}}(\mathbf{k}) \cdot \mathbf{k} \right) \mathbf{k} \\ &+ \frac{1}{8\pi\epsilon_0} \hat{\chi}_T^{-1}(k) \hat{\mathbf{P}}_T(\mathbf{k}) \left[ \hat{\mu}(\mathbf{k}, \Omega) - \left( \hat{\mathbf{P}}(\mathbf{k}) \cdot \mathbf{k} \right) \mathbf{k} \right]. \end{aligned} \quad (\text{D.0.6})$$

Il faut ensuite calculer les composantes de cette dérivée dans l'espace direct, donc en effectuer la transformée de Fourier inverse.

La dérivée du terme de correction à trois corps de l'éq. 4.2.2 peut se faire composante par composante :

$$\left. \frac{\delta \mathcal{F}_{\text{cor}}^{\text{3B-1S}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right|_{\mathbf{r}=\mathbf{r}_2} = \sum_m \lambda_m^{\text{1S}} \iiint_{\mathbb{R}^3} f_m(r_{m2}) f_m(r_{m3}) \left( \frac{\mathbf{r}_{m2} \cdot \mathbf{r}_{m3}}{r_{m2} r_{m3}} - \cos \theta_0 \right)^2 n(\mathbf{r}_3) d\mathbf{r}_3 \quad (\text{D.0.7})$$

$$\begin{aligned} \left. \frac{\delta \mathcal{F}_{\text{cor}}^{\text{3B-2S}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right|_{\mathbf{r}=\mathbf{r}_2} &= \frac{1}{2} \sum_m \lambda_m^{\text{2S}} \iiint_{\mathbb{R}^3} f_m(r_{m3}) f_w(r_{23}) \left( \frac{\mathbf{r}_{m3} \cdot \mathbf{r}_{23}}{r_{m3} r_{23}} - \cos \theta_0 \right)^2 n(\mathbf{r}_3) d\mathbf{r}_3 \\ &+ \frac{1}{2} \sum_m \lambda_m^{\text{2S}} \iiint_{\mathbb{R}^3} f_m(r_{m2}) f_w(r_{23}) \left( \frac{\mathbf{r}_{m2} \cdot \mathbf{r}_{23}}{r_{m2} r_{23}} - \cos \theta_0 \right)^2 n(\mathbf{r}_3) d\mathbf{r}_3 \end{aligned} \quad (\text{D.0.8})$$

$$\begin{aligned} \left. \frac{\delta \mathcal{F}_{\text{cor}}^{\text{3B-ww}}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right|_{\mathbf{r}=\mathbf{r}_1} &= \\ \frac{1}{2} \lambda_w \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} f_w(r_{12}) f_w(r_{13}) \left( \frac{\mathbf{r}_{12} \cdot \mathbf{r}_{13}}{r_{12} r_{13}} - \cos \theta_0 \right)^2 n(\mathbf{r}_2) n(\mathbf{r}_3) d\mathbf{r}_2 d\mathbf{r}_3 \\ &+ \lambda_w \iiint_{\mathbb{R}^3} n(\mathbf{r}_2) \left[ \iiint_{\mathbb{R}^3} f_w(r_{12}) f_w(r_{23}) \left( \frac{\mathbf{r}_{12} \cdot \mathbf{r}_{23}}{r_{12} r_{23}} - \cos \theta_0 \right)^2 n(\mathbf{r}_3) d\mathbf{r}_3 \right] d\mathbf{r}_2 \end{aligned} \quad (\text{D.0.9})$$

## E. Couplage densité-polarisation dans la fonctionnelle d'excès multipolaire

Si, contrairement à la sec. 3.3, on ne néglige pas les couplages entre densité et polarisation, on peut écrire la fonctionnelle d'excès :

$$\begin{aligned}
\beta\mathcal{F}_{\text{exc}}[\Delta n(\mathbf{r}), \mathbf{P}_L(\mathbf{r}), \mathbf{P}_T(\mathbf{r})] &= \beta\mathcal{F}_{\text{int}}[\Delta n(\mathbf{r}), \mathbf{P}_L(\mathbf{r}), \mathbf{P}_T(\mathbf{r})] - \beta\mathcal{F}_{\text{id}}[\Delta n(\mathbf{r}), \mathbf{P}_L(\mathbf{r}), \mathbf{P}_T(\mathbf{r})] \\
&= \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} C_{\text{nn}}(r_{12}) \Delta n(\mathbf{r}_1) \Delta n(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 + \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} C_{\text{nP}}(r_{12}) \Delta n(\mathbf{r}_1) P_L(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\
&+ \frac{1}{2} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} C_{\text{PP}}(r_{12}) P_L(\mathbf{r}_1) P_L(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\
&+ \frac{\beta}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \iiint_{\mathbb{R}^3} \chi_T^{-1}(r_{12}) \mathbf{P}_T(\mathbf{r}_1) \mathbf{P}_T(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2,
\end{aligned} \tag{E.0.1}$$

où on a défini un champ de polarisation longitudinal, qui est la transformée inverse de  $\hat{P}_L(\mathbf{k})$ , défini par

$$\hat{\mathbf{P}}_L(\mathbf{k}) = \hat{P}_L(\mathbf{k}) \tilde{\mathbf{k}} \tag{E.0.2}$$

avec

$$\hat{P}_L(\mathbf{k}) = \tilde{\mathbf{k}} \cdot \hat{\mathbf{P}}(\mathbf{k}), \tag{E.0.3}$$

où  $\tilde{\mathbf{k}}$  désigne le vecteur unitaire de l'espace réciproque.

L'éq. E.0.1 se réécrit dans l'espace de Fourier :

$$\begin{aligned}
\beta\mathcal{F}_{\text{exc}}[\Delta \hat{n}(\mathbf{k}), \hat{\mathbf{P}}_L(\mathbf{k}), \hat{\mathbf{P}}_T(\mathbf{k})] &= \frac{1}{2} \iiint_{\mathbb{R}^3} \hat{C}_{\text{nn}}(k) \Delta \hat{n}(\mathbf{k}) \Delta \hat{n}(-\mathbf{k}) d\mathbf{k} \\
&+ \frac{1}{2} \iiint_{\mathbb{R}^3} \hat{C}_{\text{PP}}(k) \hat{P}_L(\mathbf{k}) \hat{P}_L(-\mathbf{k}) d\mathbf{k} + \iiint_{\mathbb{R}^3} \hat{C}_{\text{nP}}(k) \Delta \hat{n}(\mathbf{k}) \hat{P}_L(-\mathbf{k}) d\mathbf{k} \\
&\frac{1}{8\pi\epsilon_0} \iiint_{\mathbb{R}^3} \hat{\chi}_T^{-1}(k) \hat{\mathbf{P}}_T(\mathbf{k}) \cdot \hat{\mathbf{P}}_T(-\mathbf{k}) d\mathbf{k}
\end{aligned} \tag{E.0.4}$$

On reconnaît des expressions couplant les densités et les polarisations similaires à celles de l'éq. 3.3.12. Cependant, les fonctions de corrélation intervenant dans ces expressions sont modifiées. Il y a un nouveau terme qui couple densité et polarisation.

On définit les facteurs de structure densité-densité, densité-polarisation longitudinale et polarisation longitudinale-polarisation longitudinale :

$$\hat{S}_{nn}(k) = \langle \Delta \hat{n}(\mathbf{k}) \Delta \hat{n}(-\mathbf{k}) \rangle = \hat{S}(\mathbf{k}) \quad (\text{E.0.5})$$

$$\hat{S}_{nP}(k) = \langle \Delta \hat{n}(\mathbf{k}) \hat{P}_L(-\mathbf{k}) \rangle = -ik \langle \Delta \hat{n}(\mathbf{k}) \hat{\rho}_c(-\mathbf{k}) \rangle = -ik S_{nc}(k) \quad (\text{E.0.6})$$

$$\hat{S}_{PP}(k) = \langle \hat{P}_L(\mathbf{k}) \hat{P}_L(-\mathbf{k}) \rangle = -4\pi\epsilon_0 k_B T \hat{\chi}_L(k) \quad (\text{E.0.7})$$

$$= -\langle \hat{\rho}_c(\mathbf{k}) \hat{\rho}_c(-\mathbf{k}) \rangle / k^2 = -S_{cc}(k) / k^2 \quad (\text{E.0.8})$$

Les expressions des fonctions de corrélation intervenant dans l'éq. E.0.4 sont les suivantes :

$$\hat{C}_{nn}(k) = \frac{\hat{S}_{PP}(k)}{\hat{S}_{PP}(k)\hat{S}_{nn}(k) - \hat{S}_{nP}(k)^2}, \quad (\text{E.0.9})$$

$$\hat{C}_{nP}(k) = \frac{\hat{S}_{nP}(k)}{\hat{S}_{PP}(k)\hat{S}_{nn}(k) - \hat{S}_{nP}(k)^2}, \quad (\text{E.0.10})$$

$$\hat{C}_{cc}(k) = \frac{\hat{S}_{nn}(k)}{\hat{S}_{PP}(k)\hat{S}_{nn}(k) - \hat{S}_{nP}(k)^2}, \quad (\text{E.0.11})$$

Ainsi, si on néglige le couplage densité-polarisation, on retrouve l'éq. 3.3.12 avec,  $\hat{C}_{cc} = -\frac{1}{4\pi\epsilon_0} \hat{\chi}_L^{-1}$ , et  $\hat{C}_{nn} = \hat{S}^{-1}$ .

L'inclusion d'un terme de couplage entre densité et polarisation modifie donc aussi les fonctions de corrélation couplant les densités et les polarisations longitudinales. Néanmoins, l'effet sur ces fonctions est faible, comme on peut le voir sur la fig. E.0.1, où on voit que le terme  $(\hat{S}_{nP})^2$  au dénominateur des fonctions  $\hat{C}$  est négligeable par rapport au terme  $\hat{S}_{PP}\hat{S}_{nn}$  et qu'il est petit devant le terme  $\hat{S}_{PP}$ .

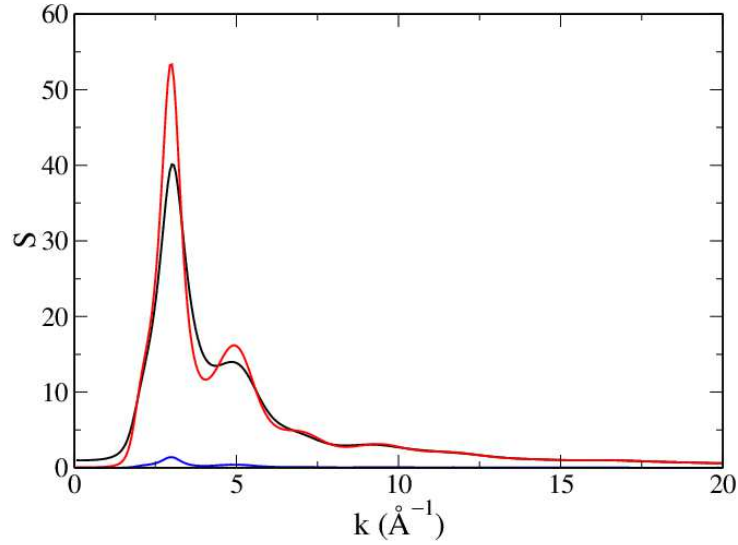
Cette fonctionnelle a été implémentée de manière à être calculée dans l'espace de Fourier.

Sa dérivée est également calculée dans l'espace de Fourier,

$$\frac{\delta \beta \mathcal{F}_{\text{exc}}[\Delta \hat{n}(\mathbf{k}), \hat{\mathbf{P}}_L(\mathbf{k}), \hat{\mathbf{P}}_T(\mathbf{k})]}{\delta \Delta \hat{n}(\mathbf{k})} = \hat{C}_{nn}(k) \Delta \hat{n}(\mathbf{k}) + \hat{C}_{nP}(k) \hat{P}_L(-\mathbf{k}),$$

$$\frac{\delta \beta \mathcal{F}_{\text{exc}}[\Delta \hat{n}(\mathbf{k}), \hat{\mathbf{P}}_L(\mathbf{k}), \hat{\mathbf{P}}_T(\mathbf{k})]}{\delta \Delta \hat{P}_L(\mathbf{k})} = \hat{C}_{PP}(k) \hat{P}_L(\mathbf{k}) + \hat{C}_{nP}(k) \Delta \hat{n}(\mathbf{k}), \quad (\text{E.0.12})$$

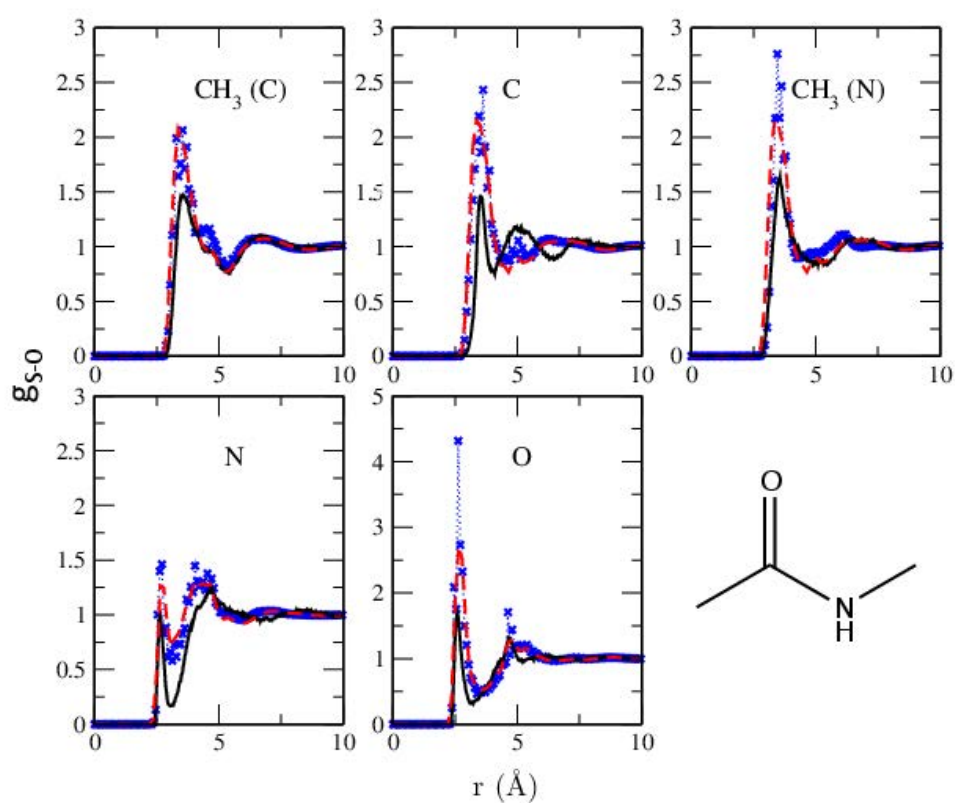
$$\frac{\delta \beta \mathcal{F}_{\text{exc}}[\Delta \hat{n}(\mathbf{k}), \hat{\mathbf{P}}_L(\mathbf{k}), \hat{\mathbf{P}}_T(\mathbf{k})]}{\delta \Delta \hat{\mathbf{P}}_T(\mathbf{k})} = \frac{1}{4\pi\epsilon_0} \hat{\chi}_T^{-1}(k) \hat{\mathbf{P}}_T(\mathbf{k}) \tilde{\mathbf{k}}, \quad (\text{E.0.13})$$



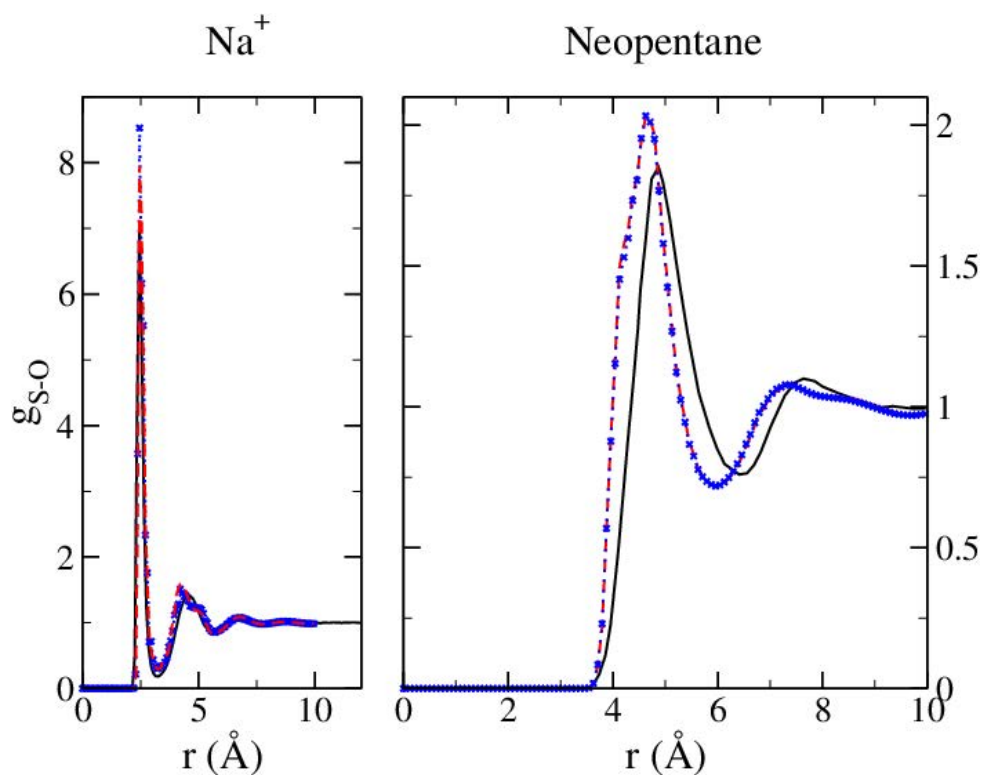
**FIGURE E.0.1.:** Comparaison des facteurs de structure  $\hat{S}_{PP}$  (noir),  $(\hat{S}_{nP})^2$  (bleu) et du produit  $\hat{S}_{PP}\hat{S}_{nn}$  (rouge) adimensionnés.

L'inclusion de ce terme de couplage n'a aucune influence visible sur les fonctions de distribution radiale obtenues par minimisation fonctionnelle pour les molécules hydrophobes comme le néopentane présenté en fig. E.0.3. Pour les molécules hydrophiles comme la N-méthylacétamide présentée en fig. E.0.2 ou les ions comme le sodium présenté en fig. E.0.3, l'ajout du couplage améliore légèrement les fonctions de distribution radiale : la hauteur des pics dus à la première couche de solvation diminue mais reste surestimée. Le second pic est peu affecté.

En ce qui concerne l'énergie libre de solvation, l'ajout du terme de couplage modifie l'énergie libre de solvation d'environ 5%. Parce que ce terme augmente le temps de calcul et que son influence sur les structures de solvation et les énergies libres de solvation obtenues est limitée, nous l'avons négligé dans le manuscrit.



**FIGURE E.0.2.:** Fonctions de distribution radiale entre le solvant et les différents sites de la N-méthylacétamide obtenus par minimisation fonctionnelle, sans terme de couplage (en croix bleues) et avec (en tirets rouges). En noir, les références de MD.



**FIGURE E.0.3.:** Fonctions de distribution radiale entre le solvant et l'ion sodium (à gauche) et le solvant et le centre de masse du néopentane (à droite). La légende est la même que celle de la fig. E.0.2.

# Bibliographie

- [1] Nič, M, Jirát, J, Košata, B, Jenkins, A, & McNaught, A, eds. (2009) *IUPAC Compendium of Chemical Terminology : Gold Book*. (IUPAC, Research Triangle Park, NC), 2.1.0 edition.
- [2] Fischer, H. E, Barnes, A. C, & Salmon, P. S. (2006) Neutron and x-ray diffraction studies of liquids and glasses. *Rep. Prog. Phys.* **69**, 233.
- [3] de Graaf, L. A & Mozer, B. (2003) Structure study of liquid neon by neutron diffraction. *J. Chem. Phys.* **55**, 4967–4973.
- [4] Hansen, J.-P & McDonald, I. (2013) *Theory of Simple Liquids*. (Academic Press), 4 edition.
- [5] Sorenson, J. M, Hura, G, Glaeser, R. M, & Head-Gordon, T. (2000) What can x-ray scattering tell us about the radial distribution functions of water? *J. Chem. Phys.* **113**, 9149.
- [6] Frenkel, D & Smit, B. (2002) *Understanding molecular simulation : from algorithms to applications*. (Academic Press).
- [7] Diu, B, Guthmann, C, Lederer, D, & Roulet, B. (1995) *Physique Statistique*. (Hermann, éditeurs des sciences et des arts), 1 edition.
- [8] Torrie, G. M & Valleau, J. P. (1977) Nonphysical sampling distributions in monte carlo free-energy estimation : Umbrella sampling. *J. Comput. Phys.* **23**, 187–199.
- [9] Pierotti, R. A. (1976) A scaled particle theory of aqueous and nonaqueous solutions. *Chem. Rev.* **76**, 717–726.
- [10] Reiss, H. (1965) in *Adv. Chem. Phys.*, ed. Prigogine, I. (John Wiley & Sons, Inc.), pp. 1–84.
- [11] Reiss, H, Frisch, H. L, & Lebowitz, J. L. (1959) Statistical mechanics of rigid spheres. *J. Chem. Phys.* **31**, 369–380.
- [12] Stillinger, F. H. (1973) Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J. Solution. Chem.* **2**, 141–158.
- [13] Tanford, C. (1979) Interfacial free energy and the hydrophobic effect. *P. Natl. Acad. Sci. USA* **76**, 4175–4176.
- [14] Roux, B & Simonson, T. (1999) Implicit solvent models. *Biophys. Chem.* **78**, 1–20.
- [15] Baker, N. A, Sept, D, Joseph, S, Holst, M. J, & McCammon, J. A. (2001) Electrostatics of nanosystems : Application to microtubules and the ribosome. *P. Natl. Acad. Sci. USA* **98**, 10037–10041.

- [16] Warwicker, J & Watson, H. C. (1982) Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* **157**, 671–679.
- [17] Honig, B & Nicholls, A. (1995) Classical electrostatics in biology and chemistry. *Science* **268**, 1144–1149.
- [18] Honig, B, Sharp, K, & Yang, A. S. (1993) Macroscopic models of aqueous solutions : biological and chemical applications. *J. Phys. Chem.* **97**, 1101–1109.
- [19] Bruccoleri, R. E, Novotny, J, Davis, M. E, & Sharp, K. A. (1997) Finite difference poisson-boltzmann electrostatic calculations : Increased accuracy achieved by harmonic dielectric smoothing and charge antialiasing. *J. Comput. Chem.* **18**, 268–276.
- [20] Ray Luo, L. D. (2002) Accelerated poisson-boltzmann calculations for static and dynamic systems. *J. Comput. Chem.* **23**, 1244–53.
- [21] Gilson, M. K, Sharp, K. A, & Honig, B. H. (1988) Calculating the electrostatic potential of molecules in solution : Method and error assessment. *J. Comput. Chem.* **9**, 327–335.
- [22] Nicholls, A & Honig, B. (1991) A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. *J. Comput. Chem.* **12**, 435–445.
- [23] Still, W. C, Tempczyk, A, Hawley, R. C, & Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127–6129.
- [24] Blum, L & Torruella, A. J. (1972) Invariant expansion for Two-Body correlations : Thermodynamic functions, scattering, and the Ornstein—Zernike equation. *J. Chem. Phys.* **56**, 303–310.
- [25] Blum, L. (1972) Invariant expansion. II. the Ornstein-Zernike equation for nonspherical molecules and an extended solution to the mean spherical model. *J. Chem. Phys.* **57**, 1862–1869.
- [26] Messiah, A. (1970) *Quantum Mechanics*. (North-Holland Publishing Company).
- [27] Puibasset, J & Belloni, L. (2012) Bridge function for the dipolar fluid from simulation. *J. Chem. Phys.* **136**, 154503–154503–15.
- [28] Beglov, D & Roux, B. (1997) An integral equation to describe the solvation of polar molecules in liquid water. *J. Phys. Chem. B* **101**, 7821–7826.
- [29] Kovalenko, A & Hirata, F. (1999) Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.* **110**, 10095–10112.
- [30] Kovalenko, A & Hirata, F. (2000) Potentials of mean force of simple ions in ambient aqueous solution. i. three-dimensional reference interaction site model approach. *J. Chem. Phys.* **112**, 10391–10402.
- [31] Kovalenko, A & Hirata, F. (1999) Potential of mean force between two molecular ions in a polar molecular solvent : A study by the three-dimensional reference interaction site model. *J. Phys. Chem. B* **103**, 7942–7957.



- [32] Imai, T, Hiraoka, R, Kovalenko, A, & Hirata, F. (2005) Water molecules in a protein cavity detected by a Statistical-Mechanical theory. *J. Am. Chem. Soc.* **127**, 15334–15335.
- [33] Imai, T, Hiraoka, R, Kovalenko, A, & Hirata, F. (2007) Locating missing water molecules in protein cavities by the three-dimensional reference interaction site model theory of molecular solvation. *Proteins : Struct., Funct., Bioinf.* **66**, 804–813.
- [34] Imai, T, Kovalenko, A, & Hirata, F. (2004) Solvation thermodynamics of protein studied by the 3D-RISM theory. *Chem. Phys. Lett.* **395**, 1–6.
- [35] Casanova, D, Gusarov, S, Kovalenko, A, & Ziegler, T. (2007) Evaluation of the SCF combination of KS-DFT and 3D-RISM-KH ; solvation effect on conformational equilibria, tautomerization energies, and activation barriers. *J. Chem. Theory Comput.* **3**, 458–476.
- [36] Palmer, D. S, Sergiievskiy, V. P, Jensen, F, & Fedorov, M. V. (2010) Accurate calculations of the hydration free energies of druglike molecules using the reference interaction site model. *J. Chem. Phys.* **133**, 044104.
- [37] Palmer, D. S, Frolov, A. I, Ratkova, E. L, & Fedorov, M. V. (2010) Towards a universal method for calculating hydration free energies : a 3D reference interaction site model with partial molar volume correction. *J. Phys. : Condens. Matter* **22**, 492101–1–492101–9.
- [38] Sergiievskiy, V. P & Fedorov, M. V. (2012) 3DRISM multigrid algorithm for fast solvation free energy calculations. *J. Chem. Theory Comput.* **8**, 2062–2070.
- [39] Szabo, A & Ostlund, N. S. (1996) *Modern Quantum Chemistry*. (Mineola, New York : Dover Publishing.).
- [40] Hohenberg, P & Kohn, W. (1964) Inhomogeneous electron gas. *Phys. Rev.* **136**, B864.
- [41] Kohn, W & Sham, L. J. (1965) Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133.
- [42] Blum, V, Gehrke, R, Hanke, F, Havu, P, Havu, V, Ren, X, Reuter, K, & Scheffler, M. (2009) Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175–2196.
- [43] Ziegler, T. (1991) Approximate density functional theory as a practical tool in molecular energetics and dynamics. *Chem. Rev.* **91**, 651–667.
- [44] Koch, W & Holthausen, M. C. (2000) *A chemist's guide to density functional theory*. (Wiley-VCH).
- [45] Mermin, N. D. (1965) Thermal properties of the inhomogeneous electron gas. *Phys. Rev.* **137**, A1441–A1443.
- [46] Evans, R. (1979) The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.* **28**, 143.
- [47] Berendsen, H. J. C, Grigera, J. R, & Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem.* **91**, 6269–6271.
- [48] Ramirez, R, Gebauer, R, Mareschal, M, & Borgis, D. (2002) Density functional theory of solvation in a polar solvent : Extracting the functional from homogeneous solvent simulations. *Phys. Rev. E* **66**, 031206–031206–8.

- [49] Ramirez, R & Borgis, D. (2005) Density functional theory of solvation and its relation to implicit solvent models. *J. Phys. Chem. B* **109**, 6754–6763.
- [50] Ramirez, R, Mareschal, M, & Borgis, D. (2005) Direct correlation functions and the density functional theory of polar solvents. *Chem. Phys.* **319**, 261–272.
- [51] Gendre, L, Ramirez, R, & Borgis, D. (2009) Classical density functional theory of solvation in molecular solvents : Angular grid implementation. *Chem. Phys. Lett.* **474**, 366–370.
- [52] Zhao, S, Ramirez, R, Vuilleumier, R, & Borgis, D. (2011) Molecular density functional theory of solvation : From polar solvents to water. *J. Chem. Phys.* **134**, 194102.
- [53] Levesque, M, Vuilleumier, R, & Borgis, D. (2012) Scalar fundamental measure theory for hard spheres in three dimensions : Application to hydrophobic solvation. *J. Chem. Phys.* **137**, 034115–034115–9.
- [54] Patey, G. (1977) An integral equation theory for the dense dipolar hard-sphere fluid. *Mol. Phys.* **34**, 427–440.
- [55] Duh, D.-M & Haymet, A. D. J. (1995) Integral equation theory for uncharged liquids : The Lennard-Jones fluid and the bridge function. *J. Chem. Phys.* **103**, 2625–2633.
- [56] Gendre, L. (2008) Ph.D. thesis (Université d’Evry-Val d’Essonne, France).
- [57] Zhao, S, Liu, H, Ramirez, R, & Borgis, D. (2013) Accurate evaluation of the angular-dependent direct correlation function of water. *J. Chem. Phys.* **139**, 034503.
- [58] Raineri, F. O, Resat, H, & Friedman, H. L. (1992) Static longitudinal dielectric function of model molecular fluids. *J. Chem. Phys.* **96**, 3068–3084.
- [59] Raineri, F. O & Friedman, H. L. (1993) Static transverse dielectric function of model molecular fluids. *J. Chem. Phys.* **98**, 8910–8918.
- [60] Guillot, B & Guissani, Y. (1993) A computer simulation study of the temperature dependence of the hydrophobic hydration. *J. Chem. Phys.* **99**, 8075.
- [61] Guillot, B, Guissani, Y, & Bratos, S. (1991) A computer-simulation study of hydrophobic hydration of rare gases and of methane. i. thermodynamic and structural properties. *J. Chem. Phys.* **95**, 3643–3648.
- [62] Jorgensen, W. L, Madura, J. D, & Swenson, C. J. (1984) Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.* **106**, 6638–6646.
- [63] Ashbaugh, H, Kaler, E, & Paulaitis, M. (1998) Hydration and conformational equilibria of simple hydrophobic and amphiphilic solutes. *Biophys. J.* **75**, 755–768.
- [64] Rosenfeld, Y. (1993) Free energy model for inhomogeneous fluid mixtures : Yukawa-charged hard spheres, general interactions, and plasmas. *J. Chem. Phys.* **98**, 8126–8148.
- [65] Liu, Y, Zhao, S, & Wu, J. (2013) A site density functional theory for water : Application to solvation of amino acid side chains. *J. Chem. Theory Comput.* **9**, 1896–1908.
- [66] Zhao, S, Jin, Z, & Wu, J. (2011) New theoretical method for rapid prediction of solvation free energy in water. *J. Phys. Chem. B* **115**, 6971.

- [67] Zhao, S, Jin, Z, & Wu, J. (2011) Correction to theoretical method for rapid prediction of solvation free energy in water. *J. Phys. Chem. B* **115**, 15445–15445.
- [68] Oettel, M. (2005) Integral equations for simple fluids in a general reference functional approach. *J. Phys. : Condens. Matter* **17**, 429.
- [69] Rosenfeld, Y. (1989) Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing. *Phys. Rev. Lett.* **63**, 980–983.
- [70] Kierlik, E & Rosinberg, M. L. (1990) Free-energy density functional for the inhomogeneous hard-sphere fluid : Application to interfacial adsorption. *Phys. Rev. A* **42**, 3382–3387.
- [71] Kierlik, E & Rosinberg, M. L. (1991) Density-functional theory for inhomogeneous fluids : Adsorption of binary mixtures. *Phys. Rev. A* **44**, 5025–5037.
- [72] Bopp, P. A, Kornyshev, A. A, & Sutmann, G. (1996) Static nonlocal dielectric function of liquid water. *Phys. Rev. Lett.* **76**, 1280–1283.
- [73] Bopp, P. A, Kornyshev, A. A, & Sutmann, G. (1998) Frequency and wave-vector dependent dielectric function of water : Collective modes and relaxation spectra. *J. Chem. Phys.* **109**, 1939–1958.
- [74] Rami Reddy, M & Berkowitz, M. (1989) The dielectric constant of SPC/E water. *Chem. Phys. Lett.* **155**, 173–176.
- [75] Kubo, R & Hashitsume, N. (2012) *Statistical Physics II : Nonequilibrium Statistical Mechanics*, Springer Series in Solid-State Sciences. (Springer).
- [76] Averill, B & Eldredge, P. (2007) *Chemistry : Principles, Patterns, and Applications*. (Benjamin Cummings).
- [77] Molinero, V & Moore, E. B. (2009) Water modeled as an intermediate element between carbon and silicon. *J. Phys. Chem. B* **113**, 4008–4016.
- [78] Stillinger, F. H & Weber, T. A. (1985) Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **31**, 5262–5271.
- [79] Lum, K, Chandler, D, & Weeks, J. (1999) Hydrophobicity and large length scales. *J. Phys. Chem.* **103**, 4570–4577.
- [80] Chandler, D & Varilly, P. (2011) Lectures on molecular- and nano-scale fluctuations in water. *arXiv 1101.2235*.
- [81] Huang, D. M & Chandler, D. (2002) The hydrophobic effect and the influence of Solute-Solvent attractions. *J. Phys. Chem. B* **106**, 2047–2053.
- [82] Varilly, P, Patel, A. J, & Chandler, D. (2011) An improved coarse-grained model of solvation and the hydrophobic effect. *J. Chem. Phys.* **134**, 074109.
- [83] Dzubiella, J & Hansen, J.-P. (2004) Competition of hydrophobic and coulombic interactions between nanosized solutes. *J. Chem. Phys.* **121**, 5514.
- [84] Jeanmairet, G, Levesque, M, & Borgis, D. (2013) Molecular density functional theory of water describing hydrophobicity at short and long length scales. *J. Chem. Phys.* **139**, 154101.

- [85] Hughes, J, Krebs, E. J, & Roundy, D. (2013) A classical density-functional theory for describing water interfaces. *J. Chem. Phys.* **138**, 024509–024509–8.
- [86] Zhao, S, Jin, Z, & Wu, J. (2011) A new theoretical method for rapid prediction of solvation free energy in water. *J. Phys. Chem. B* **115**, 6971–6975.
- [87] Oleksy, A & Hansen, J.-P. (2010) Wetting of a solid substrate by a civilized model of ionic solutions. *J. Chem. Phys.* **132**, 204702.
- [88] Stojanović, L, Jeanmairet, G, Levesque, M, & Borgis, D. (2014) A mdft approach for fast computation of potential of mean force between small hydrophobic solutes in water. *To be Published*.
- [89] Huang, X, Margulis, C. J, & Berne, B. J. (2003) Do molecules as small as neopentane induce a hydrophobic response similar to that of large hydrophobic surfaces? *J. Phys. Chem. B* **107**, 11742–11748.
- [90] Zhu, C, Byrd, R. H, Lu, P, & Nocedal, J. (1997) Algorithm 778 : L-BFGS-B : fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. Math. Soft.* **23**, 550–560.
- [91] Press, W. (1996) *Numerical recipes in FORTRAN 90 : the art of parallel scientific computing*. (Cambridge University Press, Cambridge), 2nd ed. edition.
- [92] Lebedev, V. I & Laikov, D. N. (1999) A quadrature formula for the sphere of the 131st algebraic order of accuracy. *Dokl. Math.* **59**, 477–481.
- [93] Frigo, M & Johnson, S. (1998) FFTW : an adaptive software architecture for the FFT. *CASSP Conference proceedings* **3**, 1381–1384.
- [94] (accessed 2011-13-10) Fftw3 (<http://www.fftw.org>).
- [95] Frigo, M & Johnson, S. (2005) The design and implementation of FFTW3. *Proc. IEEE* **93**, 216–231.
- [96] Marrink, S. J, Risselada, H. J, Yefimov, S, Tieleman, D. P, & de Vries, A. H. (2007) The MARTINI force field : Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824.
- [97] Asthagiri, D, Merchant, S, & Pratt, L. R. (2008) Role of attractive methane-water interactions in the potential of mean force between methane molecules in water. *J. Chem. Phys.* **128**, 244512.
- [98] Shimizu, S & Chan, H. S. (2000) Temperature dependence of hydrophobic interactions : A mean force perspective, effects of water density, and nonadditivity of thermodynamic signatures. *J. Chem. Phys.* **113**, 4683–4700.
- [99] Smith, D. E, Zhang, L, & Haymet, A. D. J. (1992) Entropy of association of methane in water : a new molecular dynamics computer simulation. *J. Am. Chem. Soc.* **114**, 5875–5876.
- [100] Shimizu, S & Chan, H. S. (2002) Origins of protein denatured state compactness and hydrophobic clustering in aqueous urea : Inferences from nonpolar potentials of mean force. *Proteins* **49**, 560–566.

- [101] Beglov, D & Roux, B. (1994) Finite representation of an infinite bulk system : Solvent boundary potential for computer simulations. *J. Chem. Phys.* **100**, 9050–9063.
- [102] Richmond, G. L. (2002) Molecular bonding and interactions at aqueous surfaces as probed by vibrational sum frequency spectroscopy. *Chem. Rev.* **102**, 2693–2724.
- [103] Marry, V, Dubois, E, Malikova, N, Durand-Vidal, S, Longeville, S, & Breu, J. (2011) Water dynamics in hectorite clays : Influence of temperature studied by coupling neutron spin echo and molecular dynamics. *Env. Sci. & Tech.* **45**, 2850–2855.
- [104] Rotenberg, B, Patel, A. J, & Chandler, D. (2011) Molecular explanation for why talc surfaces can be both hydrophilic and hydrophobic. *J. Am. Chem. Soc.* **133**, 20521–20527.
- [105] Levesque, M, Marry, V, Rotenberg, B, Jeanmairet, G, Vuilleumier, R, & Borgis, D. (2012) Solvation of complex surfaces via molecular density functional theory. *J. Chem. Phys.* **137**, 224107–224107–8.
- [106] Jeanmairet, G, Marry, V, Levesque, M, Rotenberg, B, & Borgis, D. (2014) Hydration of clays at the molecular scale : the promising perspective of classical density functional theory. *Mol. Phys.* **112**, 1320–1329.
- [107] Cygan, R. T, Liang, J.-J, & Kalinichev, A. G. (2004) Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field. *J. Phys. Chem. B* **108**, 1255–1266.
- [108] Todorov, I. T, Smith, W, Trachenko, K, & Dove, M. T. (2006) DL\_poly\_3 : new dimensions in molecular dynamics simulations via massive parallelism. *J. Mater. Chem.* **16**, 1911.
- [109] (Accessed : 2012-09-30) DL\_POLY\_3 ([http://www.ccp5.ac.uk/DL\\_POLY](http://www.ccp5.ac.uk/DL_POLY)).
- [110] Eriksson, A. E, Baase, W. A, Zhang, X. J, Heinz, D. W, Blaber, M, Baldwin, E. P, & Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **255**, 178–183. PMID : 1553543.
- [111] Zhou, Y, Morais-Cabral, J. H, Kaufman, A, & MacKinnon, R. (2001) Chemistry of ion coordination and hydration revealed by a  $K^+$  channel–Fab complex at 2.0 angstroms resolution. *Nature* **414**, 43–48.
- [112] Nakasako, M. (2004) Water–protein interactions from high–resolution protein crystallography. *Phil. Trans. R. Soc. Lond. B* **359**, 1191–1206.
- [113] Grossfield, A, Pitman, M. C, Feller, S. E, Soubias, O, & Gawrisch, K. (2008) Internal hydration increases during activation of the g-protein-coupled receptor rhodopsin. *J. Mol. Biol.* **381**, 478–486.
- [114] Shaw, D. E, Maragakis, P, Lindorff-Larsen, K, Piana, S, Dror, R. O, Eastwood, M. P, Bank, J. A, Jumper, J. M, Salmon, J. K, Shan, Y, & Wriggers, W. (2010) Atomic-level characterization of the structural dynamics of proteins. *Science* **330**, 341–346.
- [115] Yoshida, N, Imai, T, Phongphanphanee, S, Kovalenko, A, & Hirata, F. (2009) Molecular recognition in biomolecules studied by statistical-mechanical integral-equation theory of liquids. *J. Phys. Chem. B* **113**, 873–886.

- [116] Wilson, K. P, Malcolm, B. A, & Matthews, B. W. (1992) Structural and thermodynamic analysis of compensating mutations within the core of chicken egg white lysozyme. *J. Biol. Chem.* **267**, 10842–10849.
- [117] Kundrot, C. E & Richards, F. M. (1987) Crystal structure of hen egg-white lysozyme at a hydrostatic pressure of 1000 atmospheres. *J. Mol. Biol.* **193**, 157–170.
- [118] Wang, J, Cieplak, P, & Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049–1074.
- [119] Berendsen, H. J. C, van der Spoel, D, & van Drunen, R. (1995) GROMACS : a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56.
- [120] Phan, S, Kierlik, E, Rosinberg, M. L, Bildstein, B, & Kahl, G. (1993) Equivalence of two free-energy models for the inhomogeneous hard-sphere fluid. *Phys. Rev. E* **48**, 618–620.