Thermophysical properties : Towards a unified approach.

Samy Khennache

DOCTORAL THESIS

UNIVERSITE DE PAU ET DES PAYS DE L'ADOUR

École doctorale des sciences exactes et leurs applications

Thermophysical properties prediction: towards a unified approach

Dissertation defended on March 16, 2023

By

Samy KHENNACHE

In the fulfilment of the requirements for the degree of Doctor of Philosophy in

Process Engineering

Doctoral committee

Reviewers	
• Bernard ROUSSEAU	DR-CNRS, Université Paris Sud
• Patrice PARICAUD	Pr, ENSTA Paris
Examiners	
• Agilio PADUA	Pr, ENS Lyon
• Félix LLOVELL	Associate.Pr, Université Tarragone (Espagne)
• Romain PRIVAT	MdC, Université de Lorraine
• Céline HOURIEZ	CR, Mines ParisTech
• Olivier BAUDOUIN	ProSim
• Julien COLLELL	Dr. TotalEnergies
PhD Advisor	
• Guillaume GALLIERO	Pr, Université de Pau et des Pays de l'Adour

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR

THESE DE DOCTORAT DE L'UNIVERSITE DE PAU ET DES PAYS DE L'ADOUR

Ecole Doctorale n° 211 : Sciences Exactes et leurs Applications

Doctorat en Génie des Procédés

Par

Samy KHENNACHE

Prédiction des propriétés thermophysiques : vers une approche unifiée.

Bernard ROUSSEAU	DR-CNRS, Université Paris Sud	Rapporteur
Patrice PARICAUD	Pr, ENSTA Paris	Rapporteur
Agilio PADUA	Pr, ENS Lyon	Examinateur
Félix LLOVELL	Assoc.Pr, Université de Tarragone	Examinateur
Romain PRIVAT	MdC, Univesité de Lorraine	Examinateur
Céline HOURIEZ	CR, Mines ParisTech	Examinatrice
Olivier BAUDOUIN	ProSim	Examinateur
Julien COLLELL	Dr, TotalEnergies	Examinateur
Guillaume GALLIERO	Pr, Université de Pau et des Pays de l'Adour	Directeur

Laboratoire des Fluides Complexes et leurs Réservoirs,
UMR 5150 CNRS-TotalEnergies-UPPA

À mes chers parents

To my dear parents

Acknowledgements (in French)

Ce travail de thèse a été réalisé dans l'équipe de recherche « Propriétés Thermophysiques » du Laboratoire des Fluides Complexes et leurs Réservoirs (LFCR) à l'Université de Pau et des Pays de l'Adour.

Je remercie dans un premier temps Bernard ROUSSEAU et Patrice PARICAUD qui m'ont fait l'honneur d'accepter d'être rapporteurs de cette thèse. Aussi, je remercie Agilio PADUA, Félix LLOVELL et Céline HOURIEZ pour avoir accepté d'examiner ce travail et de faire partie du jury de soutenance. Un merci particulier à Julien COLLELL qui avait appuyé ma candidature pour cette thèse, ainsi qu'à Romain PRIVAT d'avoir toujours été disponible à chaque fois que je l'avais sollicité durant cette thèse.

Je tiens tout particulièrement à remercier mon Directeur de thèse, Guillaume GALLIERO, pour son soutien humain avant tout, son encadrement, son apport intellectuel, sa patience, sa totale disponibilité et les conseils, ô combien enrichissant qu'ils m'ont apportés tout au long de ces années de thèse. Mes remerciements s'adressent également au Dr.Hai Hoang, pour nos très nombreux échanges, mais aussi pour sa sympathie et ses encouragements tout au long de cette thèse. Je remercie énormément Dr.Stéphanie Delage-Santacreu pour son immense gentillesse, sa totale disponibilité et toute l'aide qu'elle m'a apporté concernant le code de simulation moléculaire.

C'est aussi l'occasion pour moi de remercier tous mes collègues du labo que j'ai côtoyés au cours de ces années de thèse et tout le personnel du LFCR. Un merci particulier à Catherine et Patricia pour leur gentillesse et bienveillance en abondance ! à Patrice pour sa sympathie et nos échanges au tour d'une tasse de café. Que Romain Vermorel trouve ici mes sincères remerciements pour ses conseils, son soutien et tous les échanges que j'avais eu le plaisir de partager avec lui.

Enfin, mes plus sincères remerciements vont à l'endroit de toute ma famille et en particulier mes Parents, pour les sacrifices consentis et toutes les valeurs qu'ils m'ont inculquées : ***ce travail de thèse leur est spécialement dédié*** *!*

Samy KHENNACHE

Abstract (in French)

Une des limites des approches couramment utilisées dans l'industrie est que les propriétés thermodynamiques et de transport d'un fluide donné sont décrites par des modèles distincts, ce qui augmente le nombre d'options et souligne le manque de cohérence de ce type d'approche. Cette thèse vise donc à relever ce défi en développant un nouveau type de modèle : une équation d'état " moléculaire " capable de fournir simultanément les propriétés thermodynamiques et de transport d'un fluide. L'approche proposée sera plus robuste que les approches actuelles car elle sera contrainte par un plus grand nombre de propriétés à décrire simultanément, et surtout, elle possédera une cohérence interne qui simplifiera son utilisation, car elle ne nécessitera qu'un seul jeu de paramètres par espèce (ou pseudo-espèce) impliquée.

Ainsi, dans une première étape des simulations moléculaires du modèle de chaine de Lennard-Jones (LJ) à rigidité variable ont été réalisées. Cela a permis d'obtenir les diagrammes d'équilibres de phases des corps purs et les courbes d'évolution de la viscosité et des coefficients de diffusion en phase liquide de ce modèle moléculaire. Les propriétés thermodynamiques (dont l'entropie résiduelle) ont été déduites d'une approche du type Monte Carlo. Les propriétés de transport ont été obtenues par des approches du type dynamique moléculaire classique.

Une fois les propriétés thermophysiques du modèle gros grains obtenus, une stratégie originale de type top-down basée sur l'approche des états correspondants a été développée afin de définir les paramètres moléculaires de composés réels. Les résultats de simulation sur ces modèles gros grains définis ont démontré la qualité du choix du champ de force et de ces paramétrisations relativement aux modèles existants, en particulier pour prédire les propriétés de transport. Il a été notamment montré que le modèle gros-gains ainsi développé permettait de prédire les propriétés thermophysiques des fluides de manière plus fidèle que les modèles moléculaires de la littérature avec un coût de simulation d'un à deux ordres de grandeur inférieur.

A partir de l'ensemble des données produites par simulation moléculaire, une équation d'état de type SAFT décrivant le comportement d'une chaine de sphères de LJ de rigidité variable a été développée pour la première fois dans la littérature. Ceci a été possible par l'utilisation d'une SAFT-dimère à laquelle a été intégrée la prise en compte de la rigidité interne au travers de la fonction de distribution radiale. Les résultats obtenus ont montré que cette équation d'état était capable de reproduire l'ensemble des propriétés thermodynamiques (en particulier équilibre de phase) de manière très fidèle avec ce qui avait été obtenu par simulation moléculaire. Cette équation d'état, avec la paramétrisation du modèle gros grains réalisée via les résultats de simulation moléculaire a ensuite été appliquée à des composés réels (purs). Les résultats sont très probants et ont permis de montrer la pertinence de l'approche ainsi que la qualité de l'équation d'état développée.

Dans une dernière étape, toujours en cours, une approche de type « entropy-scaling » a été développée. L'idée est d'utiliser l'équation d'état pour prédire la viscosité (et d'autres propriétés de transport) à partir de l'entropie résiduelle calculée. Même si le travail n'est pas

encore finalisé, l'approche est prometteuse. Cependant il est apparu que l'entropie résiduelle ne permettait pas de complétement intégré les effets des degrés de libertés internes sur la viscosité, notamment aux fortes densité (comportement super-Arrhenius).

Mots clefs : Propriétés Thermophysiques, Simulation moléculaire, modèle gros-gains, rgidité moléculaire, Etats correspondants, Equations d'état SAFT, Entropy-scaling

Abstract

One of the limitations of the approaches commonly used in industry is that the thermodynamic and transport properties of a given fluid are described by separate models, which increases the number of options and highlights the lack of consistency of this type of approach. Therefore, this PhD aims to address this challenge by developing a new type of model: a "molecular" equation of state capable of simultaneously providing the thermodynamic and transport properties of a fluid. This proposed approach will be more robust than current ones because it will be constrained by a greater of properties to be described simultaneously, and, above all, will possess an internal consistency that will simplify its use, as it will require only one set of parameters per species (or pseudo-species) involved.

In a first step, molecular simulations of the Lennard-Jones (LJ) chain model with variable rigidity were carried out. This allowed for the obtaining of phase equilibrium diagrams of pure substances and the viscosity and diffusion coefficients in the liquid phase of this molecular model. Thermodynamic properties (including residual entropy) were deduced from a Monte Carlo approach. Transport properties were obtained by using classical molecular dynamics approaches.

Once the thermophysical properties of the coarse-grained model were obtained, an original top-down strategy based on the corresponding states approach was developed to define the molecular parameters of real compounds. The simulation results on these defined coarse-grained models demonstrated the quality of the choice of the force field and these parameterizations compared to existing models, particularly for predicting transport properties. It was particularly shown that the coarse-grained model thus developed made it possible to predict the thermophysical properties of fluids more accurately than the molecular models of the literature with a simulation cost of one to two orders of magnitude lower.

From the set of data produced by molecular simulation, an equation of state of the SAFT type describing the behavior of a Lennard-Jones chain of variable rigidity was developed for the first time in the literature. This was possible through the use of a SAFT dimer to which the internal rigidity was integrated through the radial distribution function. The results obtained showed that this equation of state was able to reproduce all thermodynamic properties (in particular phase equilibrium) very accurately with what had been obtained by molecular simulation. This equation of state, with the coarse-grained model parameterization carried out via the results of molecular simulation was then applied to real compounds (pure). The results are very promising

and have shown the relevance of the approach as well as the quality of the developed equation of state.

In a final step, still ongoing, an "entropy-scaling" approach has been developed. The idea is to use the equation of state to predict viscosity (and other transport properties) from the calculated residual entropy. Although the work is not yet finalized, the approach is promising. However, it appeared that the residual entropy did not allow for completely integrating the effects of internal degrees of freedom on viscosity, especially at high densities (super-Arrhenius behavior)."

Keywords: Thermophysical properties, Molecular simulation, Coarse grained models, molecular rigidity, Corresponding states, SAFT Equations of state, Entropy-scaling

Table of contents

Acknowledgements (in French) i
Abstract (in French) ii
Abstract iii
Table of contents v
List of figures viii
List of tables xii
Chapter 1. Introduction 1
1.1 Global context 2
1.2 Motivations 3
1.3 Objectives 6
1.4 Thesis overview 7
1.5 References 9
Chapter 2. Statistical physics and molecular simulation 10
2.1 Ensembles 11
2.1.1 Canonical ensemble NVT 12
2.1.2 NpT ensemble 16
2.1.3 Microcanonical ensemble NVE 19
2.2 Interaction potentials 20
2.2.1 Intramolecular potentials 20
2.2.2 Intermolecular potentials 23
2.3 Molecular simulations 25
2.3.1 Molecular dynamics 25
2.3.2 Monte Carlo 32
2.3.3 Monte Carlo or Molecular dynamics? 36
2.3.4 Optimization of the simulation efficiency 37
2.3.5 Transport properties and structural properties calculation 40
2.4 References 44
Chapter 3. Development of an LJC coarse grained molecular model 46
3.1 Molecular models: up-scaling 47
3.2 Parametrization strategy: Bottom-up vs Top down 49
3.2.1 Bottom-up approach 49
3.2.2 Top-down approach 51
3.3 An overview of the predictive capabilities of some up-scaled force fields 53
3.3.1 United atoms 53

3.3.2 Coarse grained 55
3.3.3 SAFT based Coarse grained 56
3.4 Towards an improved coarse grained model 57
3.4.1 Semi Rigid Lennard Jones Chain Coarse Grained model 58
3.4.2 Simulations details 60
3.4.3 Rigidity effect on phase equilibria and transport properties of the fluid model: The Semi-Rigid LJCCG 64
3.4.4 Choice of the parametrization strategy 66
3.4.5 Optimization procedure 68
3.4.6 Application to pure real fluids 70
3.5 Summary 84
3.6 References 85
Chapter 4. Integration of a bending potential into SAFT model 92
4.1 Introduction 93
4.2 Wertheim perturbation theory: a global overview 95
4.2.1 Pair potential 95
4.2.2 Assumptions and Steric Incompatibilities (SI) 97
4.2.3 Multiple density formalism 100
4.2.4 Derived Equations 101
4.3 SAFT models 111
4.3.1 SAFT-HS and CK-SAFT 111
4.3.2 LJ-SAFT (soft-SAFT) 112
4.3.3 SAFT-VR 113
4.3.4 PC-SAFT 115
4.4 Effect of the approximations on the chain bond formation term 117
4.5 A simple alternative to higher orders of the TPTn … 122
4.6 Does Wertheim TPT includes hard intramolecular interactions (vibrations)? 129
4.7 The model 132
4.8 Model validation 134
4.8.1 Single phase 134
4.8.2 Liquid vapour equilibrium (LVE) 136
4.8.3 Internal energy & Entropy at saturation 138
4.9 Real fluid properties 140
4.9.1 Optimization of the molecular parameters 140
4.9.2 Results and discussions 144
4.9.3 Consistency check and transferability of the parameters 155

4.10 Summary 161
4.11 References 163
Chapter 5. Entropy scaling based SR-SAFTD for transport properties 169
5.1 Overview of the evolution and application of the entropy scaling approach 170
5.2 Towards a corresponding state entropy scaling for the LJC fluid model 172
5.3 Application of the entropy scaling for the real fluids 181
5.4 Summary 185
5.5 References 186
Chapter 6. Conclusion and perspectives 188
6.1 Conclusion 189
6.2 Perspectives 191
6.3 References 194
Chapter 7. Appendices 195
Appendix 7.A 196
7.A.1 Experimental data used for the parameterization. 196
7.A.1.1 Experimental data used in the optimization of the parameters of the FF-LJCCG, SR-LJCCG and SR-SAFTD models. 196
7.A.2 Fitting parameters: fitting constants and optimized molecular parameters 197
7.A.2.1 Correlations used in the parametrization of the SR-LJCCG and the SR-SAFTD models. 197
7.A.2.2 Optimized parameters for the SR-LJCCG model. 199
7.A.2.3 Optimized parameters for the FF-LJCCG model. 200
7.A.2.4 Optimized parameters for the SR-SAFTD model. 201
7.A.2.5 SR-SAFTD fitting parameters. 202
7.A.2.6 Fitting parameters of the original Johnson parameters for the SAFTD (TPT1-D) [1] and SAFTM (TPT1-M) [2] models. 205
7.A.2.7 Helmholtz free energy of the monomer segment. 205
7.A.2.8 Transferable parameters used for the prediction of the properties of the iso-alkanes obtained from the SR-SAFTD. 207
Appendix 7.B 209
7.B.1 Surface tension and associative molecules 209
7.B.1.1 Surface tension prediction. 209
7.B.1.2 Associative molecules 210
7.B.2 References (Appendix A & B) 213

List of figures

Figure 1.1: Different domains where thermophysical properties are required 2
Figure 1.2: Different approaches to obtain thermophysical properties. 4
Figure 1.3: Illustration of the general framework where the molecular simulation approach can be used. 5
Figure 2.1: Virtual representation of microstates ensemble 12
Figure 2.2: Schematic representation of the NVT ensemble 13
Figure 2.3: Schematic representation of the NpT ensemble 16
Figure 2.4: Schematic representation of the NVE ensemble 20
Figure 2.5: Schematic representation of the bond stretch and associated harmonic potential 21
Figure 2.6: Schematic representation of the bond bending and associated harmonic potential 22
Figure 2.7: Schematic representation of the bond torsion and associated polynomial potential 22
Figure 2.8: Different types of intermolecular potential interactions 25
The global workflow of the molecular dynamics simulations corresponds to the steps described in Figure 2.9. 26
Figure 2.10: Workflow of the molecular dynamics simulation 26
Figure 2.11: Boltzmann analysis of the Metropolis algorithm 34
Figure 2.12: Schematic representation of the Gibbs ensemble. left: Molecular dynamics, right: Monte Carlo 37
Figure 2.13: Schematic representation of the PBC 38
Figure 2.14: Schematic representation of the cut-off radius 39
Figure 2.15: Schematic representation of the Verlet List and Verlet radius 39
Figure 2.16: Schematic view of the momentum transfer process proposed by Müller-Plathe. The white arrows indicate the direction of the shear. The red colour indicates the hot slabs (& particles) and the blue indicate the cold slabs (& particles). The black arrows indicate the direction of the momentum exchange. 41
Figure 2.17: Left: schematic representation of the spatially divided system into shell volumes used for the calculation of the radial distribution function (RDF). Right: typical response of the RDF in the liquid phase. 42
Figure 3.1: Schematic representation of the up-scaling approach in molecular modelling. From Ab-initio scale to the continuum scale. AA stands for All Atom model, UA stands for United Atom model and CG stands for Coarse grained model. 48
Figure 3.2: Schematic representation of the Bottom-up and Top-down parameterization approaches. 50
Figure 3.3: Test of different objective functions for the parametrization of the fully-flexible LJCCG force field along the vapour-liquid coexistence line. 59
Figure 3.4: Evolution of the critical temperature, density and the acentric factor with respect to the number of data points included in the fit. 61
Figure 3.5: Effect of the shear rate on the viscosity of chain of m=5 at T_r=0.7 and ρ^*_{sat} corresponding to the different rigidities considered. 63
Figure 3.6: Effect of the shear rate on the normalized stress differences Nxy and Nyz for m=5 at the same thermodynamic conditions given in Figure 3.5. 63
Figure 3.7: Effect of the rigidity on the coexistence densities and saturation pressure of chain of m=5. (a) L-V Equilibrium densities, (b) Saturation pressure. 65
Figure 3.8: Effect of the rigidity on the saturated liquid viscosity of chain of m=5 66
Figure 3.9: Thermophysical properties of the semi-rigid LJCCG model used in the parametrization of the coarse grained force field proposed. 67
Figure 3.10: (a) flowchart of the general methodology, (b) flowchart of the optimization strategy. 69
Figure 3.11: Equilibrium liquid-vapour saturated densities and saturation pressures of iso-alkanes vs correlated experimental data from NIST & DIPPR databases. 70

Figure 3.12: Equilibrium liquid-vapour saturated densities and saturation pressures of iso-alkanes vs correlated experimental data from NIST database. Experimental liquid density for iC16 is taken from Klein et al.2020, and Psat from Ambrose et al.1988. 71
Figure 3.13: Equilibrium liquid-vapour saturated densities and saturation pressures of some polar components vs correlated experimental data from NIST database. 71
Figure 3.14: Temperature dependence of the predicted second order derivative properties of n-alkanes with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa. 73
Figure 3.15: Temperature dependence of the predicted second order derivative properties of iso-alkanes with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa. 74
Figure 3.16: Temperature dependence of the predicted second order derivative properties of some polar components with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa. 75
Figure 3.17: Temperature dependence of the single phase density predicted with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa. (a) n-alkanes, (b) iso-alkanes, (c) polar components. 76
Figure 3.18: Saturated and compressed liquid viscosity of normal alkanes, iso-alkanes and some polar components. (a, b, c) viscosity of n-alkanes, iso-alkanes and polar components at saturation conditions, (d) compressed liquid viscosity of n-alkanes at temperature of T=373K. 78
Figure 3.19: Relative errors for the compressed liquid viscosity and density of n-C12 at T=373K. Results of viscosity are those presented in previous figure and for density are not shown but correspond to those at which the viscosity is simulated. 79
Figure 3.20: Benchmark between different force fields proposed in the literature and experiments for the saturated liquid viscosity prediction. (a) n-C12, (b) n-C16. 80
Figure 3.21: (a) Temperature and chain length effect for various n-alkanes components at P=1atm. (b) Pressure effect for nC6 and nC8 at T=298K. 81
Figure 3.22: Center of mass intermolecular radial distribution function of n-hexane and n-nonane molecules in the subcritical liquid state (Tr=0.7) and supercritical fluid (Tr=1.20). 83
Figure 4.1: Wertheim's SI1, where two bonded hard cores prevent a third hard core from association to one of them due to the impossible overlapping between cores. 98
Figure 4.2: (a) Wertheim's SI2W, where a site from one molecule can only be bonded to one site from another molecule. (b) Wertheim's SI2S, where double bonding is not allowed as a consequence of SI1 and SI2W. 99
Figure 4.3: Wertheim's SI3, where, for small molecules, the angle between the two site and the center of mass of the molecule is large enough , so that the molecules can be considered as rigid and linear (a). (b) Another consequence is that rings are not allowed to form for small molecules. 100
Figure 4.4: Polymerization process to form a linear trimer molecule. It highlights the excluded volume ($V_{molecule\ (1)} + V_{grey\ zone}$) for molecule (3). Consequently, the bonding energy between {3,2} is much greater that between {1,2}. 118
Figure 4.5: This is the continuation of the previous figure, where here it highlights that the energy required to form a branch {2,4} is much greater than that needed to form a linear 4-mer chain due to the higher excluded volume for molecule (4). 119
Figure 4.6: In the logic of TPT1, in (a) if a molecule (1) is hold fixed, the center of mass of molecule (2) can freely move over the red circle to bond on molecule (1). In (b), molecule (3) can also move over the green circle which is confounded with the red circle without any restriction due to the fact that there is no information between molecules (1) and (3) as they can overlap. 120
Figure 4.7: Internal energy and saturation pressure of chain of m=4 for variable rigidities vs the prediction from TPT1. (a) internal energy at T*=4, (b) saturation pressure 121
Figure 4.8: Polymerization of monomers to form chains of 4-mer fluid. Path (a) follows the TPT1 scheme where the formation of chain depends only on the information included at the monomer level, thus, no information is included beyond nearest bond. Path (b and c) follows the TPT1-D scheme, where first, the dimer is formed through the TPT1-M. Then, the 4-mer chains fluid is formed by associating dimerization of these dimer molecules. 123

Figure 4.9: Comparison between the exact molecular simulations of the fully flexible chains model (m=1, 2, 3, 4, 8, 16) at saturation conditions vs the predictions of the TPT1-M (dotted lines) and the TPT1-D (solid lines) models. (a) LVE, (b) saturation pressures, (c) internal energy, (d) entropy..... 126
Figure 4.10: Zero-density limit internal energy of fully flexible chains up to m=50 at T*=2 obtained from monte carlo simulations vs the predictions using the SAFT-M, SAFT-D, SAFT-T and SAFT-Q. 128
Figure 4.11: Some of the possible configurations that two dimers in contact may hold during the path (c). 130
Figure 4.12: Zero-density limit internal energy at T*=2 predicted by the SR-SAFTD (solid lines) for variable rigidities (K*=0, 1, 3, 10, 50) and chains up to m=50 vs the exact monte carlo simulations (dots)..... 134
Figure 4.13: Residual internal energy for chains of m=4, 8, 12 at supercritical(T*=4) and subcritical(T*=2) conditions for variable rigidities (K*=0, 1, 3, 10) obtained from molecular simulations (symbols) and predictions using the SR-SAFTD (solid lines). 135
Figure 4.14: Pressure for chain of m=8 at supercritical(T*=4) and subcritical(T*=2) conditions for variable rigidities (K*=0, 1, 3, 10) obtained from molecular simulations (symbols) and predictions using the SR-SAFTD (solid lines). 136
Figure 4.15: Liquid-vapour equilibrium properties and saturation pressures predicted by the SR-SAFTD (solid lines) for chains of m=4, 8, 16 vs molecular simulations (symbols) for variable rigidities (K*=0, 3, 10)..... 137
Figure 4.16: Critical properties and acentric factor predicted by the SR-SAFTD (solid lines) and soft-SAFT (dashed lines) for chain of m=5 vs molecular simulations (symbols) for variable rigidities (K*=0 to 20). 138
Figure 4.17: Liquid-vapour internal energy and entropy predicted by the SR-SAFTD (solid lines) for chains of m=4, 8 vs molecular simulations (symbols) for variable rigidities (K*=0, 3, 5, 10). 139
Figure 4.18: Prediction of the phase equilibria properties of pure n-octane using a corresponding state strategy (CS). 140
Figure 4.19: (a) flowchart of the general methodology using the SR-SAFTD, (b) flowchart of the optimisation strategy using SR-SAFTD. 142
Figure 4.20 : Liquid-vapour saturated densities and saturation pressures predicted with the SR-SAFTD and the SR-LJCCG vs the correlated experimental data from NIST database. (a,b) n-alkanes, (c,d) iso-alkanes, (e,f) polar compounds. 145
Figure 4.21: Relative deviations of the critical properties predicted with SR-SAFTD vs the carbon number of series of n-alkanes. 146
Figure 4.22: Liquid-vapour saturated densities and saturation pressures predicted with the SR-SAFTD, the fully flexible SAFTD and the soft-SAFT vs the correlated experimental data from NIST database of n-C6 and n-C12. 146
Figure 4.23: Liquid-vapour enthalpy predicted with the SR-SAFTD vs the correlated experimental data from NIST. (a) n-alkanes, (b) iso-alkanes, (c) polar components. 148
Figure 4.24: Temperature dependence of the second order derivative properties of the n-alkanes components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa. 150
Figure 4.25: Temperature dependence of the second order derivative properties of the iso-alkanes components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa. 151
Figure 4.26: Temperature dependence of the second order derivative properties of the polar components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa. 152
Figure 4.27: Temperature dependence of the single phase enthalpy predicted with the SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.. (a) n-alkanes, (b) iso-alkanes, (c) polar components..... 153

Figure 4.28: Temperature dependence of the single phase density predicted with the SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.. (a) n-alkanes, (b) iso-alkanes, (c) polar components..... 154
Figure 4.29: Optimized molecular parameters of the SR-SAFTD EOS for the series of n-alkanes Vs That of the fully flexible SAFTD and the fully flexible soft-SAFT. (a) σ Vs Mw. (b) ε Vs Mw. (c) m Vs Mw. (d) K/ε Vs Mw..... 156
Figure 4.30: Reduced experimental viscosity of n-alkanes taken from NIST database Vs Mw 157
Figure 4.31: Transferable parameters of the SR-SAFTD EOS for the series of n-alkanes, correlated up to n-C_{22}. (a) m Vs Mw. (b) $m\sigma^3$ Vs Mw. (c) mε Vs Mw. (d) K/ε Vs Mw..... 158
Figure 4.32: Transferable parameters of the SR-SAFTD EoS for the series of iso-alkanes, correlated up to i-C_{16}. (a) m Vs Mw. (b) $m\sigma^3$ Vs Mw. (c) mε Vs Mw. (d) K/ε Vs Mw..... 159
Figure 4.33: Prediction of the thermodynamic properties of nC_{36} using the SR-SAFTD and soft-SAFT with parameters obtained from the extrapolated correlations. (a) LVE, (b) Psat 160
Figure 4.34: Liquid density of polyethylene at various pressures predicted using SR- SAFTD and soft-SAFT with parameters obtained from the extrapolation of the correlations. The prediction with SAFT-γ-Mie is also included taken from Rahman et al.2018 . The experimental (dots) data are also obtained from the same paper (see therein for therein exact origin)..... 161
Figure 5.1: Test of different scaling approaches on the fully flexible LJC fluid model..... 173
Figure 5.2: Test of the different approaches on the Semi-Rigid LJC fluid model. 174
Figure 5.3: Testing the different scaling approaches using the new variable suggested in this work for the abscissa..... 175
Figure 5.4: Arrhenius and super-Arrhenius transition analysis using the modified Dehlouz approach. (a) Entropy scaling for LJC3 at variable rigidities (the dashed lines represent the transition limits), (b) represents the evolution of the transition abscissa with respect to the rigidity. 177
Figure 5.5: Entropy scaling of fully flexible LJC3 showing the possibility of deducing the gas-liquid transition similarly to the transition Arrhenius-Super-Arrhenius..... 178
Figure 5.6: Testing different new approaches by dividing the viscosity in the ordinate by the segment number of the chains and using of the reduced quantity in the abscissa for the Galliéro, Novak and Bell approaches..... 179
Figure 5.7:Corresponding state entropy scaling approach for the fully flexible LJ chains. 180
Figure 5.8: Performance of the soft-SAFT EoS in predicting a corresponding state entropy scaling for the fully flexible LJC fluid model. 180
Figure 5.9: (a) prediction of the viscosity of the semi-rigid LJC at zero-density limit and T*=4, (b) The fitting parameter Alpha function of rigidity. 182
Figure 5.10: Test of the correlation (solide lines) for the entropy scaling against the data used in the fitting (symbols). (a) fully flexible chains K=0, (b) semi-rigid chains K=5. 183
Figure 5.11: Saturated viscosity prediction of some n-alkanes with different approaches, direct molecular simulations and entropy scaling using the SR-SAFTD. The FF-SAFTD is the SR-SAFTD with K*=0. Comparison is made between the two approaches and to the correlated experimental data taken from the NIST database. 184
Figure 7.1: (a) Temperature dependence of the zero-density limit g_{00} (T*, ρ*→0, K*, σ) at various rigidities. (b) density dependence of the term in g_{ee} (T*=4, ρ*, K*, σ) substracted out g_{00} (T*, ρ*→0, K*, σ) expressed by the matrix $C_{ij;}$ (c)Extended g_{00} (T*, ρ*→0, K*→0, σ) to infinite T*. 203
Figure 7.2: Rigidity dependency of the matrix C_{ij}(K*) coefficient used to calculate g_{ee} (T*, ρ*, K*, σ) 204
Figure 7.3: Temperature dependent parameter ai and bi given in $A^{seg.}$..... 206
Figure 7.4: Density dependent parameters of G_i coefficients given in A^{seg}..... 206
Figure 7.5: Predicted surface tension of normal alkanes and polar components using correlation from Galliéro (2010). (a) normal alkanes. (b) Polar components. 210
Figure 7.6: Test of the model and parameterization strategy on methanol and water molecules..... 211

List of tables

Table 3.1: Radius of gyration for the fully flexible and semi-rigid n-hexane and n-nonane molecules at Tr=0.7 and Tr=1.20 and densities indicated on figure 3.18. 82
Table 3.2: End-to-End distance for the fully flexible and semi-rigid n-hexane and n-nonane molecules at Tr=0.7 and Tr=1.20 and densities indicated on figure 3.18..... 83
Table 4.1: Constants of the transferable parameters correlations for the n-alkanes..... 158
Table 5.1: Parameters of the α(K*) function..... 181
Table 5.2: Numerical fitting parameters used in the correlation for the entropy scaling. 183
Table 7.1: Experimental data of the normal alkanes used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models..... 196
Table 7.2: Experimental data of the iso- alkanes used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models..... 196
Table 7.3: Experimental data of the some polar components used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models. 196
Table 7.4: Experimental data of methanol and water molecules used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models. 197
Table 7.5: Fitting parameters of the acentric factor..... 197
Table 7.6: Fitting parameters for the critical temperature. 198
Table 7.7: Fitting parameters of the saturated liquid density at Tr=0.7. 198
Table 7.8: Fitting parameters of the saturated liquid viscosity at Tr=0.7..... 199
Table 7.9: Optimized molecular parameters of the SR-LJCCG model for the n-alkanes. 199
Table 7.10: Optimized molecular parameters of the SR-LJCCG model for the iso-alkanes..... 199
Table 7.11: Optimized molecular parameters of the SR-LJCCG model for some polar components. 199
Table 7.12: Optimized molecular parameters of the SR-LJCCG model for methanol and water molecules..... 200
Table 7.13: Optimized molecular parameters of the FF-LJCCG model for the n-alkanes..... 200
Table 7.14: Optimized molecular parameters of the FF-LJCCG model for the iso-alkanes. 200
Table 7.15: Optimized molecular parameters of the FF-LJCCG model for some polar components. 200
Table 7.16: Optimized molecular parameters of the SR-SAFTD model for the n-alkanes. 201
Table 7.17: Optimized molecular parameters of the SR-SAFTD model for the iso-alkanes. 201
Table 7.18: Optimized molecular parameters of the SR-SAFTD model for some polar components. 201
Table 7.19: Optimized molecular parameters of the SR-SAFTD model for methanol and water molecules..... 201
Table 7.20: αlj coefficients used to calculate the cij matrix. 202
Table 7.21: q_i coefficients used to calculate the c_{ij} matrix. 202
Table 7.22: v_{lj} coefficients used in the $g_{ee}(T^*, \rho^*\rightarrow 0, K^*, \sigma)$.. 202
Table 7.23: Coefficients used in g_{00} $(T^*, \rho^*\rightarrow 0, K^*\rightarrow 0, \sigma)$. 203
Table 7.24: c_{ij} coefficients used in the g_{ee} (T^*, ρ^*, σ) given by Johnson(1996)..... 205
Table 7.25: : Coefficients used in the g_{ee} $(T^*, \rho^*\rightarrow 0, \sigma)$ given by Johnson(1996). 205
Table 7.26 c_{ij} coefficients used in the g^M (T^*, ρ^*, σ) given by Johnson et al.(1994). 205
Table 7.27: x[i] coefficients given in A^{seg} 207
Table 7.28: Constants of the transferable correlations of the iso-alkanes. 208

Chapter 1. Introduction

Contents

1.1 Global context: 2
1.2 Motivations 3
1.3 Objectives 6
1.4 Thesis overview 7
1.5 References 9

1.1 Global context

The idea for this PhD project arose from the recognition that in order to optimize an existing process or develop a new one, it is necessary to be able to quantify the properties of the fluid material present in the system [1, 2]. In the applicative context of this project, which deals with exploitation, storage and use of fluid resources, energy conversion, material transformation ... the targeted properties are those of the fluid phase(s) stored and/or circulating in the considered system. As shown in Figure 1.1, the crucial quantities in this perspective are the thermodynamic properties (such as phase equilibrium, quantities involved in energy and exergy balances such as enthalpy, entropy ...) and transport properties (like viscosity, mass diffusion, thermal conductivity ...), whose are collectively known as thermophysical properties.

Thus, optimizing or developing innovative processes requires, among other things, a thorough and reliable description of the properties of the fluids present in the system in question. This generally implies the use of macroscopic models that can quickly provide access to these physical quantities. However, there are many approaches available, often ad-hoc or correlative, that are tailored for specific types of system, properties, precise thermodynamic conditions, etc. These approaches have proven effective in the industrial world, but they require an expert user (to select from the many options) and may not be easily applied to new situations when experimental data (which could validate or adjust the models) are lacking or hard to obtain.

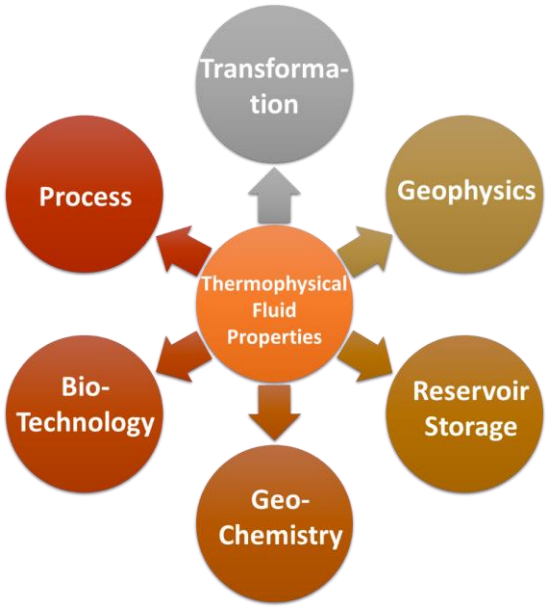


Figure 1.1: Different domains where thermophysical properties are required

Thus, one of the limitations of the approaches commonly used in industry is that the thermodynamic and transport properties of a given fluid are described by separate models, which increases the number of options and highlights the lack of consistency of this type of approach. Therefore, this PhD aims to address this challenge by developing a new type of model: a "molecular" equation of state capable of simultaneously providing the thermodynamic

and transport properties of a fluid. This approach will be more robust than current ones because it will be constrained by a greater number of properties to be described simultaneously, and, above all, will possess an internal consistency that will simplify its use, as it will require only one set of parameters per species (or pseudo-species) involved.

The model developed within this project will serve as a foundation for the prediction of the thermophysical properties of fluids under a wide range of thermodynamic conditions, making it applicable to issues related to georesources (such as exploitation, storage...) as well as processes (such as energy conversion, product design ...).

1.2 Motivations

Accurate estimates of thermophysical properties are crucial for the efficient design and optimization of many industrial processes as shown in Figure 1.1. There are several approaches that can be used to estimate these properties as shown in Figure 1.2, including experiments, ad-hoc correlations, machine learning, theories and molecular simulations. The use experiments would be the natural choice as they provide the real thermophysical properties. However, experiment cannot be conducted indefinitely to cover all thermodynamic states, as they can be costly and time-consuming. Moreover, there are many cases where experiments are unfeasible, tedious, and unsafe to conduct, such as when dealing with complex mixtures, toxic substances or studying systems under extreme pressure and temperature conditions. To optimize the number of experiments conducted, and also to use these properties in simulation software, ad-hoc correlations are often proposed. However, these correlations are typically only valid for one specific property of a specific fluid and at a specific range of condition, which is a major weakness as their extrapolation beyond the fitting range is then questionable. The same problem is also encountered when using a machine learning approach. However, the main limitation of this approach is the large amount of data required to build the model, which are often not available. When a robust theory is available, a more suitable approach is to fill in the lack of experiments data instead of using ad-hoc correlations or machine learning approach. It can also provide a fundamental explanation of the behaviour of fluids that are observed with experiments. However, in some cases, these explanations must be taken with caution because the development of a complex and complete theory is always based on a model which could be questionable and is sometimes difficult to interpret. As a result, most theories rely on assumptions to facilitate their development and resolution. Therefore, it is important to evaluate the extent to which these assumptions are valid to ensure the reliability of the explanations provided. It would be very useful to have a tool that allows the validation of theories before they are directly applied to real systems.

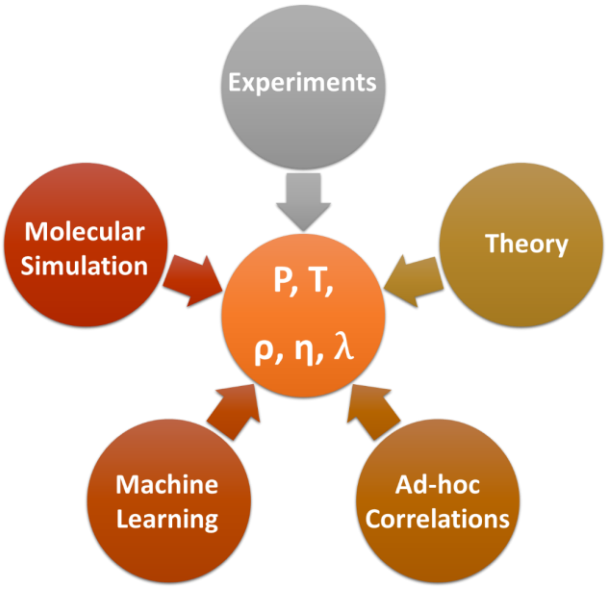


Figure 1.2: Different approaches to obtain thermophysical properties.

There are many theories that allows an accurate estimation of thermophysical properties, such as cubic equations of state (EoS) or SAFT equations of state [3] for thermodynamic properties, the Chapman-Enskog theory [4] or the friction theory [5] for viscosity… However, there is currently no generic theory that can provide both type of properties (thermodynamic and transport) simultaneously. A promising approach in that direction is to use a scaling method to connect these two types of properties, such as the density scaling [6, 7] or the entropy scaling [8, 9, 10]. The combination of a transport property and a thermodynamic property can exhibit a master curve that can be easily fitted with a simple correlation. This can be seen as an interesting tool to overcome the lack of transport properties data or to allow for a quality check of some data. However, the success of such scaling is closely linked to how the available experimental data are dispersed in the thermodynamic space. As mentioned previously, there are many cases where there is a lack of data, and when available, they are often limited to specific operational conditions. Furthermore, this scaling is often always constructed using experimental data from different sources, each with their own errors, resulting in a disparate data curve instead of a single master curve. A consistency check is then necessary, although sometimes it cannot be achieved. Finally, it is worth to underline that, some issues may arise from the equations of state used for instance to calculate the entropy for such scaling approaches, as to the best of our knowledge, they have never been thoroughly evaluated for their ability to provide correct entropy.

Among the approaches discussed above, molecular simulation is probably the most powerful alternative to experiments as it can provide quasi-experimental data on fluid properties. Moreover, this numerical approach can be used to obtain thermodynamic, transport interfacial and structural properties of fluids, where other approaches fall short. For example, the equations of states can only provide data on equilibrium properties. Although molecular simulation will never fully replace experiments, there are many situations where it is extremely useful. For example, it can reduce the number of experiments needed to deal with extreme thermodynamic conditions, complex mixtures and toxic substances. Thus, it can be considered as a

complementary tool to experiments. In addition, molecular simulation provides information at the molecular level, offering robust interpretations of the complex physical and/or chemical phenomena that govern fluid properties at the macroscopic level. In fact, these interpretations can also be given by theories when available. However, as mentioned earlier, these theories rely on assumptions and approximations, and therefore, the interpretations and theoretical predictions obtained should be valid to the same extent that these assumptions and approximations are valid. It is then common to rely on the molecular simulations as a tool for the validation of some theories in the outset of their development before they are applied to real systems. This is for example the case of the Statistical Associating Fluid Theory (SAFT) equations of state.

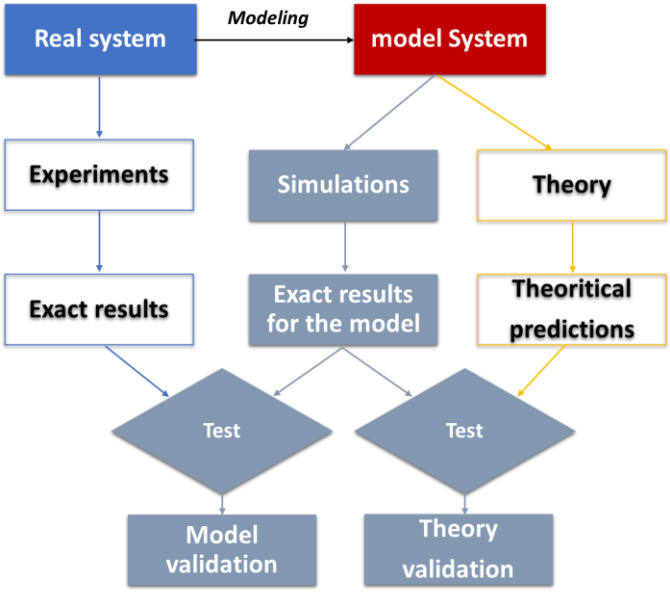


Figure 1.3: Illustration of the general framework where the molecular simulation approach can be used.

Despite the many advantages of the molecular simulation approach, there is a significant need to improve the computational efficiency when performing atomistic simulations. Simulations of systems at the mesoscale level using an atomistic model require the evaluation of billions of interactions at each step of the simulation, making the simulations extremely time-consuming. Therefore, the widespread use of molecular simulation approach in engineering applications require the development of new molecular models that would enable large time and length scale simulations. To address this limitation, coarse grained model has merged as a viable alternative to atomistic models. In a coarse-grained model, a group of atoms are gathered into a single center of force, called a pseudo-atom. This drastically reduces the number of interactions that needs to be evaluated at each simulation step, and can lead to computational time reductions up to five orders of magnitude [11]. However, the loss of the internal degrees of freedom due to the up-scaling makes coarse-grained models less capable of providing accurate thermal, structural and transport properties. This is not the only reason for their failure. Indeed, most of the coarse-grained models designed for the prediction of thermophysical properties are parametrized only on equilibrium properties such as phase equilibria, enthalpy of vaporization, heat capacity…and

do not include transport properties or structural properties as target properties to be reproduced during the optimization of the potential interaction parameters. This is because they are often based on SAFT equations of states, which do not allow for the inclusion of properties other than thermodynamic properties. The only exception is that proposed in the SAFT-γ-Mie coarse grained model [12], where it is proposed to include structural properties as a target property to optimize the stretching and bending molecular internal degrees of freedom. A more accurate coarse grained model is that proposed by Hoang et al. [13], which was optimized using extensive molecular simulations rather than using a SAFT equation of state. Hoang's model is a fully flexible Mie fluid model known as MCCG (Mie Chain Corase Grained), where the parameters of the interaction potential are optimized based on both equilibrium and transport properties by introducing a viscosity as a target property for the first time. The model showed excellent predictions for both type of properties, however, when dealing with large molecules such as normal decane and dodecane, a systematic underestimation of the liquid viscosity at low temperature was observed, which was attributed to the fully flexibility nature of the MCCG model.

1.3 Objectives

Currently, even with its limitation in predicting liquid viscosity, the MCCG model is probably among the best performing coarse grained model for thermodynamic and transport properties of simple fluids. To improve upon the performance of the MCCG model, a new model that integrates some internal degrees of freedom of the molecule needs to be developed. This is the main objective of this thesis. A new molecular model with a variable rigidity parameter of the bending type will be proposed. This molecular model is based on the Lennard-Jones interaction potential, and is referred to as the Semi-Rigid Lennard-Jones Chain Coarse-Grained molecular model (SR-LJCCG). It will be demonstrated that this new model is capable of providing accurate predictions of thermodynamic properties and transport properties, such as self-diffusion and viscosity, simultaneously. The limitation of the MCCG model in predicting liquid viscosity for large chain molecule will be addressed in this work. Additionally, a new parametrization strategy will be proposed, which is based on the corresponding states strategy inspired from the works of Mejia et al. [14] and Hoang et al. [13].

Despite the impressive computational time gains offered by the coarse-grained model over fine-grained models, it is more suitable for engineering applications to have a mathematical formulation of the model that allows for a rapid evaluation of thermophysical properties (within seconds!) and can be implemented in process simulators. Unfortunately, all available SAFT EoS versions are based on fully flexible molecular models. Therefore, the second objective of this thesis is to develop a new SAFT EoS that incorporates a rigidity parameter. The equation will be referred to as the SR-SAFTD (where D stands for dimer theory). This is the first time a SAFT version that takes into account the internal degrees of freedom of a molecule of the bending type will be proposed. This equation of state will first be validated using exact molecular simulation data of the fluid model and then applied to real fluids. The parametrization of the new SAFT model is based on the same strategy as the SR-LJCCG model. Additionally, the same parameters obtained for the SR-LJCCG will be used for the SR-SAFTD equation of state, highlighting the coherence and consistency of the proposed approach.

Last, to obtain a transport property such as viscosity, instead of relying on extensive molecular simulations, a new scaling based on the entropy scaling approach will be proposed as a third objective for this thesis. The predictions of the viscosity using the proposed scaling on real fluids will be obtained with the same parameters used for the equilibrium properties, without any further re-parametrization.

It is clear from this introduction that the goal of this PhD thesis is to propose a new, coherent, and consistent strategy where the same molecular model is used for the prediction of equilibrium and transport properties, all within a unified approach.

1.4 Thesis overview

The work in this thesis presents a coarse-grained molecular model developed for the predictions of the thermodynamic and transport properties of fluids. To do so, this work is divided into chapters, where in the second chapter, elements on the statistical physics and molecular simulations will be given. We first give some notions on the different ensembles used in this work, and then presents the techniques used to simulate the ensembles. Thus, both molecular dynamics and Monte Carlo simulations are presented, in addition to the different potential interactions. In the third chapter, it will be given insights on the up-scaling approach used in molecular modelling, in addition to the different parametrization strategy used in the literature. It will be followed by an overview of the different existing force fields, particularly those based on united atoms models and coarse grained force fields models, where their predictive capability is discussed. The new force field proposed in this work referred to as SR-LJCCG will then be presented, in addition to the general methodology and strategy of the parametrization used to optimize the SR-LJCCG force field. Results on the fluid model, as well as on the real fluid properties covering wide range of thermodynamic conditions, and applied over many properties and molecular species will be shown.

In the fourth chapter, an overview of the Wertheim thermodynamic perturbation theory is presented in addition to some SAFT models. Discussion on the strengths and weaknesses of the Wertheim theory is highlighted, particularly when the comparison between monomer based SAFT and dimer based SAFT is made. Then, the strategy proposed for the development of the new SR-SAFTD EoS that integrates a rigidity parameter of the bending type will be presented, in addition to the validation of this EoS model on exact molecular simulations of the fluid model. After that, will be present the new parameterization strategy used for the SR-SAFTD, where for the first time, a transport property (viscosity) is included in the parameter estimation of SAFT models. Then the optimised SR-SAFTD EoS is tested against molecular simulations and experimental data over a wide thermodynamic conditions, and covering many properties and molecular species. We will end-up by highlighting the transferability of the molecular parameters developed with the SR-SAFTD within the same family.

In the fifth chapter, will be discussed the coupling between the SAFT EoSs and molecular simulations used in the entropy scaling approach for the prediction of transport properties. First, a review is made for the LJC fluids using different scaling approaches given in the literature. It

will be shown that, while for the fully flexible LJC fluids a good scaling can be obtained with the different approaches, the case of semi-rigid is not trivial. Modified scaling is then presented for the generalization of the entropy scaling approach to semi-rigid LJ chains. Even though a corresponding state is not yet achieved, to illustrate the final goal of this work, it will be shown some applications on the prediction of the liquid viscosity using the SR-SAFTD in the modified entropy scaling. Very promising results will be shown, where the same parameters used for the prediction of the thermodynamic properties are used in the scaling for the viscosity prediction, highlighting the consistence and the coherence of the global methodology or approach proposed in this work.

Last, we finish this work with a general conclusion and some perspectives recommended for the future work.

1.5 References

1. Hendriks, E. *et al.* Industrial Requirements for Thermodynamics and Transport Properties. *Ind. Eng. Chem. Res.* **49**, 11131–11141 (2010).
2. Kontogeorgis, G. M. *et al.* Industrial Requirements for Thermodynamic and Transport Properties: 2020. *Ind. Eng. Chem. Res.* **60**, 4987–5013 (2021).
3. Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **29**, 1709–1721 (1990).
4. Chapman, S. & Cowling, T. G. *The Mathematical Theory of Non-uniform Gases: An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*. (Cambridge University Press, 1990).
5. Assael, M. J., Goodwin, A. R. H., Vesovic, V. & Wakeham, W. A. *Experimental Thermodynamics Volume IX: Advances in Transport Properties of Fluids*. (Royal Society of Chemistry, 2014).
6. Roland, C. M., Hensel-Bielowka, S., Paluch, M. & Casalini, R. Supercooled dynamics of glass-forming liquids and polymers under hydrostatic pressure. *Rep. Prog. Phys.* **68**, 1405 (2005).
7. Delage Santacreu, S., Hoang, H., Khennache, S. & Galliero, G. Thermodynamic Scaling of the Shear Viscosity of Lennard-Jones Chains of Variable Rigidity. *Liquids* **1**, 96–108 (2021).
8. Rosenfeld, Y. Relation between the transport coefficients and the internal entropy of simple systems. *Phys. Rev. A* **15**, 2545–2549 (1977).
9. Galliero, G., Boned, C. & Fernández, J. Scaling of the viscosity of the Lennard-Jones chain fluid model, argon, and some normal alkanes. *J. Chem. Phys.* **134**, 064505 (2011).
10. Dyre, J. C. Perspective: Excess-entropy scaling. *J. Chem. Phys.* **149**, 210901 (2018).
11. Ingólfsson, H. I. *et al.* The power of coarse graining in biomolecular simulations. *Wiley Interdiscip Rev Comput Mol Sci* **4**, 225–248 (2014).
12. Rahman, S. *et al.* SAFT-γ Force Field for the Simulation of Molecular Fluids. 5. Hetero-Group Coarse-Grained Models of Linear Alkanes and the Importance of Intramolecular Interactions. *J. Phys. Chem. B* **122**, 9161–9177 (2018).
13. Hoang, H., Delage-Santacreu, S. & Galliero, G. Simultaneous Description of Equilibrium, Interfacial, and Transport Properties of Fluids Using a Mie Chain Coarse-Grained Force Field. *Ind. Eng. Chem. Res.* **56**, 9213–9226 (2017).
14. Mejía, A., Herdes, C. & Müller, E. A. Force Fields for Coarse-Grained Molecular Simulations from a Corresponding States Correlation. *Ind. Eng. Chem. Res.* **53**, 4131–4141 (2014).

Chapter 2. Statistical physics and molecular simulation

Contents

2.1 Ensembles 11
2.1.1 Canonical ensemble NVT 12
2.1.2 NpT ensemble 16
2.1.3 Microcanonical ensemble NVE 19
2.2 Interaction potentials 20
2.2.1 Intramolecular potentials 20
2.2.2 Intermolecular potentials 23
2.3 Molecular simulations 25
2.3.1 Molecular dynamics 25
2.3.2 Monte Carlo 32
2.3.3 Monte Carlo or Molecular dynamics? 36
2.3.4 Optimization of the simulation efficiency 37
2.3.5 Transport properties and structural properties calculation 40
2.4 References 44

This chapter, which presents an overview of the theories and numerical methods used in this PhD thesis, is mostly inspired by statistical mechanic and molecular simulations textbooks, including Allen and Tildesley [1] , Smit and Frenkel [2], Hansen and McDonald [3], Sandler [4], Raabe [5], Ungerer et al. [6], Leach [7], McQuarrie [8], Rappaport [9] and Vlugt [10]. Additional references can be found therein.

2.1 Ensembles

As it will be described latter on, molecular simulations are a set of a numerical tools that can be run in various statistical ensembles. As defined by statistical mechanics, a statistical ensemble is a collection of an extremely large number of systems in thermodynamic equilibrium, where each system is in a different microstate *j* characterized by its total energy H_j (the Hamiltonian). Each microstate consists of *N* particles, each described by its position vector "r" defined in the 3D spatial coordinates, known as configurational space, and also by its momentum vector "*p*" defined as well in the 3D spatial coordinates. Thus, the system of *N* particles, is defined in a space of 6*N* dimensions, r^N and p^N, referred to as "phase space".

Such a representation is theoretical construction to depict the time evolution of a real macroscopic system. This is because, from a macroscopic point of view, a measure can be viewed as the time average of a successive frozen microstate of the molecules in the system. Therefore, if *X* is the macroscopic quantity that is observed, it can be obtained in the Boltzmann's sense by the time average over all the microstates in the phase space as,

$$X_{obs} =< X >_{time}= \lim_{t\to\infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} X\,(r^N(t), p^N(t))dt \tag{2.1}$$

where *t* is the time and $< X >_{time}$ denotes the time average of *X*.

This equation is valid under the assumption that the duration of the experiments is sufficiently large to generate an extremely large number of microstates (which is practically always verified).

Similarly, Gibbs provided a definition for the ensemble average, which allows to make a connection between the macroscopic properties and the ensemble averages. In this framework, the observable *X* is given by:

$$X_{obs} =< X >_{ensemble}= \int X(r^N, p^N) f(r^N, p^N) dr^N dp^N \tag{2.2}$$

where f is the probability density written as:

$$f(r^N, p^N) = \frac{Exp\left(\frac{-H(r^N, p^N)}{k_B T}\right)}{Q(r^N, p^N)} \tag{2.3}$$

where $Q(r^N, p^N)$ is the sum over all the microstates, called the partition function, which is a function of the macroscopic properties that defines the ensemble, and $Exp\left(\frac{-H(r^N, p^N)}{k_B T}\right)$ is the Boltzmann factor expressing the fact that low energy microstates are more favourable than higher ones. The Boltzmann factor acts as a weighting function for the occurrence of a certain microstate j. T is the temperature and k_B is the Boltzmann constant.

From a molecular simulation point of view, an ensemble can be seen as a set of microstates which can be thought as different snapshots of a system's configuration at different time or move steps. The properties of the system can be approximated as those of the ensemble if the number of configurations generated by the molecular simulation is sufficiently large. In reality, this can never be fully achieved as the molecular simulation will preferentially sample only the low-energy region of the phase space, and thus, it can only provide an approximation of the partition function. Therefore, the properties deduced are only approximate and not the true ones. However, if one can consider that the high energy levels contribute very little according to the Boltzmann probability, one can adopt the Ergodicity postulate, on which the statistical thermodynamics relies on. This postulate stipulates that the ensemble average and time average can be considered equivalent. Thus, one can write:

$$X_{obs} \approx < X >_{ensemble} \approx < X >_{time} \tag{2.4}$$

where X_{obs} is the experimental measurement, while $< X >_{ensemble}$ can be obtained using Monte Carlo simulations and $< X >_{time}$ can be obtained using molecular dynamics simulations. The two molecular simulations techniques will be detailed in the following.

Statistical thermodynamics and classical thermodynamics are connected through the bridge relation between the thermodynamic potentials of the ensemble (with its minimum defining the equilibrium condition) and the partition function of these microstates, which is given by:

$$\mathbb{\Gamma} = -k_B T ln(Q_{ensemble}) \tag{2.5}$$

Thus, if the partition function is known, all macroscopic thermodynamic properties can be derived.

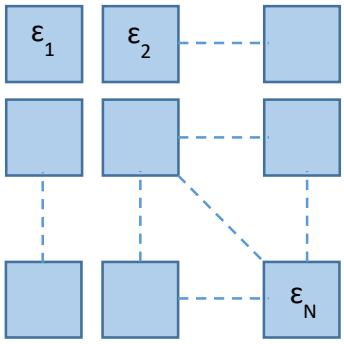


Figure 2.1: Virtual representation of microstates ensemble

2.1.1 Canonical ensemble NVT

In this ensemble, the number of particles, the volume and the temperature of the system are hold constant. In other words, it is the density and the temperature of the system which are fixed, and thus, the other variables of the system such as the total energy of the system as well as the pressure which are allowed to fluctuate.

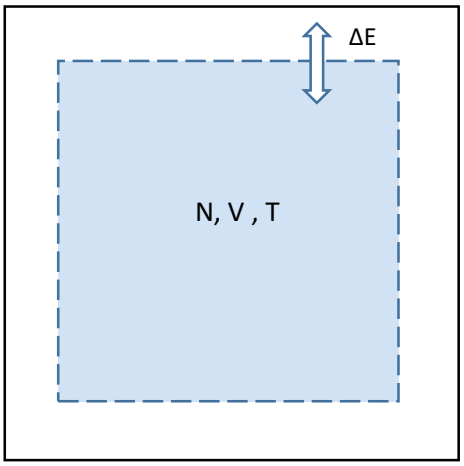


Figure 2.2: Schematic representation of the NVT ensemble

The probability density of a given microstate *j* is given by:

$$\rho_j^{NVT} = \frac{Exp\left(\frac{-H(r^N,p^N)}{k_BT}\right)}{Q_{NVT}} \tag{2.6}$$

where Q_{NVT} is the partition function in this ensemble given by:

$$Q_{ensemble} = Q_{NVT} = \sum_{j=1}^{n} Exp\left(\frac{-H(r^N,p^N)}{k_BT}\right) \tag{2.7}$$

If *n*, which is the number of microstates, is supposed sufficiently large, but, finite, the summation remains finite as well. Thus, the partition function can be written as a continuous integral as follow:

$$Q_{NVT} = \frac{1}{h^{3N}N!}\int_{r^N}\int_{p^N} Exp\left(\frac{-H(r^N,p^N)}{k_BT}\right) dr^N dp^N \tag{2.8}$$

where, $N!$ denotes the number of indistinguishable microstates, h^{3N} is the volume of an individual quantum microstate in the phase space, and h is the Planck constant.

$H_j(r^N, p^N)$ can be separated as a summation of two contributions, a kinetic part "*k*" and a potential part "*U*". Furthermore, if one assumes that the kinetic part is only dependent on p^N, and the potential part is only dependent on r^N, thus, the two integrals can be separated as:

$$Q_{NVT} = \frac{1}{h^{3N}N!}\int_{p^N} Exp\left(\frac{-k(p^N)}{k_BT}\right) dp^N \int_{r^N} Exp\left(\frac{-U(r^N)}{k_BT}\right) dr^N \tag{2.9}$$

After some calculations, one may show that:

$$\frac{1}{h^{3N}}\int_{p^N} Exp\left(\frac{-k(p^N)}{k_BT}\right) dp^N = \left(\frac{2\pi m}{k_BTh^2}\right)^{3N/2} = \Lambda^{-3N} \tag{2.10}$$

where Λ is the de Broglie wavelength:

$$\Lambda = \frac{h}{\sqrt{2\pi m k_BT}} \tag{2.11}$$

The partition function becomes:

$$Q_{NVT} = \frac{\Lambda^{-3N}}{N!} \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N \tag{2.12}$$

The integral in this equation is often referred as the configurational partition function and noted by the symbol Z_{NVT}:

$$Z_{NVT} = \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N \tag{2.13}$$

For a perfect gas system, $U_j(r^N) = 0$. Hence, the ideal partition function, which can only be approximated by the means of molecular dynamics, is:

$$\begin{aligned} Q_{ideal} &= \frac{\Lambda^{-3N}}{N!} V^N \\ &= \frac{q^N}{N!} = \frac{q_{translation}^N}{N!} \end{aligned} \tag{2.14}$$

where q^N stands for the molecular partition function, and since the molecules are approximated as a perfect gas, neither vibrational nor rotational kinetic energies contribute to the total kinetic energy. Only translational contribution can be considered. Therefore, it is possible write $q^N = q_{trans}^N$.

Thus one can write:

$$Q_{NVT} = Q_{ideal} Q_{residual} \tag{2.15}$$

where $Q_{residual}$ can be approximated by both molecular dynamics and Monte Carlo as it involves only the knowledge of the position of the atoms and the potential of interaction, given by:

$$Q_{residual} = \frac{Z_{NVT}}{V^N} \tag{2.16}$$

Various thermodynamics properties can be obtained if the partition function is known. Some of them are:

$$A = -k_B T ln(Q_{NVT}) \tag{2.17}$$

$$U = \frac{k_B T^2}{Q_{NVT}} \left(\frac{\partial Q_{NVT}}{\partial T}\right)_{V,N} \tag{2.18}$$

$$H = \frac{k_B T^2}{Q_{NVT}} \left(\frac{\partial Q_{NVT}}{\partial T}\right)_{V,N} + \frac{k_B T V}{Q_{NVT}} \left(\frac{\partial Q_{NVT}}{\partial V}\right)_{T,N} \tag{2.19}$$

$$G = -k_B T ln(Q_{NVT}) + \frac{k_B T V}{Q_{NVT}} \left(\frac{\partial Q_{NVT}}{\partial V}\right)_{T,N} \tag{2.20}$$

It is common to express these quantities as a combination of two parts: the ideal part, which is solely due to the kinetic energy of the system, and the residual part, which is due to the interaction between the molecules of the system. The total Helmholtz free energy in the *NVT* ensemble can be rewritten as:

$$\begin{aligned} A &= -k_B T ln(Q_{ideal} Q_{residual}) \\ &= -k_B T ln(Q_{ideal}) - \frac{k_B T}{V^N} \ln(Z_{NVT}) \\ &= A^{ideal} + A^{residual} \end{aligned} \tag{2.21}$$

The ideal gas law is then simply obtained by setting $A^{residual} = 0$.

$$\begin{aligned} A^{ideal} &= -k_B T ln(Q_{ideal}) \\ &= -k_B T[ln(V^N) - ln(N!) - \ln(\Lambda^{3N})] \\ &= -k_B T[Nln(V) - (Nln(N) - N) - N\ln(\Lambda^3)] \\ &= -k_B TN[ln(V/N) + 1 - \ln(\Lambda^3)] \\ &= -k_B TN\left[ln\left(\frac{1}{\rho\Lambda^3}\right) + 1\right] \\ &= k_B TN[ln(\rho\Lambda^3) - 1] \end{aligned} \tag{2.22}$$

If, furthermore, the molecules are considered as hard spheres (not punctual), the equation of state can be obtained as follow:

$$\begin{aligned} A &= k_B TN\left[ln\left(\frac{N}{V-Nb}\Lambda^3\right) - 1\right] \\ &= k_B TN\left[ln\left(\frac{N}{V-Nb}\right) + \ln(\Lambda^3) - 1 + ln\left(\frac{V-Nb}{N}\right) - ln\left(\frac{V-Nb}{N}\right)\right] \\ &= k_B TN\left[ln\left(\frac{N}{V}\right) + \ln(\Lambda^3) - 1 - ln\left(\frac{V-Nb}{N}\right)\right] \\ &= A^{ideal} - ln(1-\rho b) \end{aligned} \tag{2.23}$$

where $-ln(1-\rho b)$ is the excess free energy due to the consideration of hard sphere. Note that, if $b = 0$, one recovers the ideal gas law.

The pressure can also be deduced from the A quantity as follow:

$$\begin{aligned} p &= k_B T\left(\frac{\partial \ln(Q_{NVT})}{\partial V}\right)_{T,N} \\ &= k_B T\left(\frac{\partial \ln(Q_{ideal})}{\partial V}\right)_{T,N} + k_B T\left(\frac{\partial \ln(Q_{residual})}{\partial V}\right)_{T,N} \\ &= k_B TN + k_B T\left(\frac{\partial \ln(Q_{residual})}{\partial V}\right)_{T,N} \\ &= p_{ideal} + p_{residual} \end{aligned} \tag{2.24}$$

Second-order derivative properties such as the isochoric heat capacity can be deduced as well. It is given by:

$$\begin{aligned} C_v = \frac{\partial U}{\partial T} &= \frac{\partial\left(\frac{k_B T^2}{Q_{NVT}}\left(\frac{\partial Q_{NVT}}{\partial T}\right)_{V,N}\right)_{N,V}}{\partial T} = \\ &= \frac{k_B T^2}{Q_{NVT}}\left(\frac{\partial^2 Q_{NVT}^2}{\partial T^2}\right)_{V,N} + \frac{2k_B T}{Q_{NVT}}\left(\frac{\partial Q_{NVT}}{\partial T}\right)_{V,N} - \frac{k_B T^2}{Q_{NVT}^2}\left(\frac{\partial Q_{NVT}}{\partial T}\right)_{V,N}^2 \\ &\vdots \\ &= \frac{1}{k_B T^2}(< U^2 > - < U >^2) \end{aligned} \tag{2.25}$$

which is evaluated from the fluctuation theory using molecular simulations. For more details on how such formula can be derived, please refers to the next section regarding the *NpT* ensemble where many other derivative properties are considered.

The evaluation of $C_v$with the use of the fluctuation theory may sometimes be subject to high statistical uncertainties. Alternatively, this can be evaluated by performing simulation at different isotherms. Applying numerical derivation for the C_v, one obtains:

$$C_v = \frac{\partial U}{\partial T} = \frac{U_2 - U_1}{T_2 - T_1} \tag{2.26}$$

The interest of the canonical ensemble lies in its ease of implementation in addition to the fact that there are many thermodynamic quantities that are written in terms of density and

temperature as variables, i.e. $f(V,T)$, such as in the SAFT equations of states, allowing for a direct comparison between the simulation and model results.

2.1.2 NpT ensemble

Most experiments are typically performed at fixed temperature and pressure rather than fixed *NVT*. Thus, to make a direct comparison to experiments, it is more practical to perform molecular simulations at fixed *NpT*, allowing both volume and energy to fluctuate.

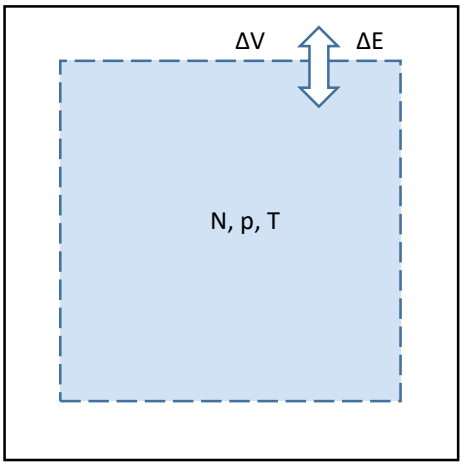


Figure 2.3: Schematic representation of the NpT ensemble

In this ensemble, the partition function is given by:

$$\begin{aligned}\Delta_{NpT} &= \frac{1}{N!}\int_{r^N}\int_{p^N}\int_V \left(Exp\left(\frac{-H(r^N,p^N)}{k_BT}-\frac{pV}{k_BT}\right)\right)dr^N dp^N dV \\ &= \frac{\Lambda^{-3N}}{N!}Q_{NVT}\int_V(-pV)dV\end{aligned} \tag{2.27}$$

Similar to the *NVT* ensemble, the positions, the momentum and the volume are independent, allowing for the integrals to be separated as follows:

$$\begin{aligned}\Delta_{NpT} &= \frac{1}{N!}\int_V Exp\left(\frac{-pV}{k_BT}\right)dV\int_{p^N}Exp\left(\frac{-k_j(p^N)}{k_BT}\right)dp^N\int_{r^N}Exp\left(\frac{-U(r^N)}{k_BT}\right)dr^N \\ &= \frac{\Lambda^{-3N}}{N!}\int_V Exp\left(\frac{-pV}{k_BT}\right)dV\int_{r^N}Exp\left(\frac{-U(r^N)}{k_BT}\right)dr^N \\ &= \frac{\Lambda^{-3N}}{N!}Z_{NpT}\end{aligned} \tag{2.28}$$

where the Z_{NpT} is the configurational partition function in the *NpT* ensemble. Using the bridge equation, the Gibbs free energy is given by:

$$G = -k_BT\ln(\Delta_{NpT}) \tag{2.29}$$

The variation of the Gibbs free energy is given by the Gibbs-Duhem equation:

$$dG = -SdT + Vdp + \mu dN \tag{2.30}$$

where S, V and μ are given by:

$$S = k_B \ln(\Delta_{NpT}) + \frac{k_B T}{\Delta_{NpT}} \left(\frac{\partial \Delta_{NpT}}{\partial T}\right)_{N,p} \tag{2.31}$$

$$V = -\frac{k_B T}{\Delta_{NpT}} \left(\frac{\partial \Delta_{NpT}}{\partial p}\right)_{N,T} \tag{2.32}$$

$$\mu = -\frac{k_B T}{\Delta_{NpT}} \left(\frac{\partial \Delta_{NpT}}{\partial N}\right)_{p,T} \tag{2.33}$$

Using the fluctuation theory, many derivative properties can be deduced from the *NpT* ensemble.

- Isothermal compressibility

$$k_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_{N,T} \tag{2.34}$$

First, the average volume V is:

$$\begin{aligned} V &= -\frac{k_B T}{\Delta_{NpT}} \left(\frac{\partial \Delta_{NpT}}{\partial p}\right)_{N,T} \\ &= -\frac{k_B T}{\Delta_{NpT}} \frac{\Lambda^{-3N}}{N!} \int_V \left(\frac{-V}{k_B T}\right) Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N \\ &= \frac{\int_V (V) Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N}{\int_V Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N} \\ &= \frac{\int_V (V) Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N}{Z_{NpT}} \\ &= < V >_{NpT} \end{aligned} \tag{2.35}$$

Thus, the compressibility is given by (using similar results than those obtained for *V*):

$$\begin{aligned} k_T &= -\frac{1}{<V>_{NpT}} \left(\frac{\partial \left(-\frac{k_B T}{\Delta_{NpT}} \left(\frac{\partial \Delta_{NpT}}{\partial p}\right)_{N,T} \right)}{\partial p} \right)_{N,T} \\ &= \frac{\int_V \left(\frac{V^2}{k_B T}\right) Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U_j(r^N)}{k_B T}\right) dr^N}{Z_{NpT}} + \frac{\left[\int_V (V^2) Exp\left(\frac{-pV}{k_B T}\right) dV \int_{r^N} Exp\left(\frac{-U(r^N)}{k_B T}\right) dr^N\right]^2}{k_B T Z_{NpT}^2} \\ &= \frac{1}{k_B T} \frac{\left(<V^2>_{NpT} - <V>_{NpT}^2\right)}{<V>_{NpT}} \end{aligned} \tag{2.36}$$

Similar to the isochoric heat capacity, the fluctuations of the volume may be subject to large uncertainties. Thus, alternatively, the compressibility can be obtained by numerical derivation between of two *NpT* runs as follow:

$$\begin{aligned} k_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_{N,T} &= \frac{1}{\rho} \left(\frac{\partial \rho}{\partial p}\right)_{N,T} \\ &= \left(\frac{\partial \ln(\rho)}{\partial p}\right)_{N,T} \\ &= \frac{\ln(\rho_2/\rho_1)}{(p_2 - p_1)} \end{aligned} \tag{2.37}$$

a) Thermal expansivity

$$\begin{aligned}\alpha_P &= \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{N,p} \\ &= \frac{1}{<V>_{NpT}}\left(\frac{\partial\left(-\frac{k_BT}{\Delta_{NpT}}\left(\frac{\partial \Delta_{NpT}}{\partial p}\right)_{N,T}\right)}{\partial T}\right)_{N,p} \\ &= \frac{1}{<V>_{NpT}}\left(\frac{\partial\left(-\frac{k_BT}{\Delta_{NpT}}\left(\frac{\partial \Delta_{NpT}}{\partial p}\right)_{N,T}\right)}{\partial T}\right)_{N,p} \\ &= \frac{1}{k_BT^2<V>_{NpT}}\left(\frac{-\int_V VExp\left(\frac{-pV}{k_BT}\right)dV\int_{r^N}(U+pV)Exp\left(\frac{-U\left(r^N\right)}{k_BT}\right)dr^N}{Z_{NpT}} + \right. \\ &\quad \left. \frac{\int_V VExp\left(\frac{-pV}{k_BT}\right)dV\int_{r^N}Exp\left(\frac{-U\left(r^N\right)}{k_BT}\right)dr^N}{Z_{NpT}}\frac{\int_V Exp\left(\frac{-pV}{k_BT}\right)dV\int_{r^N}(U+pV)Exp\left(\frac{-U\left(r^N\right)}{k_BT}\right)dr^N}{Z_{NpT}}\right)\end{aligned} \tag{2.38}$$

The quantity $U + pV$ is the configurational enthalpy noted $\hat{H}$. Thus, the formula becomes:

$$\alpha_P = \frac{<V\hat{H}>_{NpT} - <V>_{NpT}<\hat{H}>_{NpT}}{k_BT^2<V>_{NpT}} \tag{2.39}$$

The alternative approach can also be applied by performing isobaric simulations, using this formula:

$$\begin{aligned}\alpha_P &= \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{N,p} \\ &= -\frac{\ln(\rho_2/\rho_1)}{T_2 - T_1}\end{aligned} \tag{2.40}$$

b) Isobaric heat capacity

$$c_P = \left(\frac{\partial H}{\partial T}\right)_{N,p} \tag{2.41}$$

As for the isochoric heat capacity, it is common in molecular simulation to separate the kinetic from the residual part to express the isobaric heat capacity. In the formula above, *H* is the total enthalpy given by:

$$H = U^{inter} + U^{intra} + pV + U^{kinetic} \tag{2.42}$$

According to Lagache [11], the total enthalpy can be separated into two parts, the kinetic (ideal) and the configurational parts given as follow:

$$H^{res} = U^{inter} + pV - Nk_BT \tag{2.43}$$

$$H^{ideal} = U^{intra} + U^{kinetic} + Nk_BT \tag{2.44}$$

Thus, the residual heat capacity is:

$$\begin{aligned}c_P^{res} &= \left(\frac{\partial H^{res}}{\partial T}\right)_{N,p} \\ &= \left(\frac{\partial U^{inter}}{\partial T}\right)_{N,p} + p\left(\frac{\partial V}{\partial T}\right)_{N,p} - Nk_B\end{aligned} \tag{2.45}$$

The fluctuations formulas for the two first terms can be obtained by generalizing the results obtained for the thermal expansivity, as it is possible to show that for any quantity *X* one obtains:

$$\frac{\partial <X>_{NpT}}{\partial \beta} = -\left(< X\hat{H} >_{NpT} - < X >_{NpT} < \hat{H} >_{NpT}\right) \tag{2.46}$$

Thus, the residual

$$\begin{aligned} c_P^{res} &= \left(\frac{\partial U^{inter}}{\partial T}\right)_{N,p} + p\left(\frac{\partial V}{\partial T}\right)_{N,p} - Nk_B \\ &= \frac{1}{k_B T^2}\left(< U^{inter}\hat{H} >_{NpT} - < U^{inter} >_{NpT} < \hat{H} >_{NpT}\right)_{N,p} + \\ &\quad \frac{p}{k_B T^2}\left(< V\hat{H} >_{NpT} - < V >_{NpT} < \hat{H} >_{NpT}\right) - Nk_B \end{aligned} \tag{2.47}$$

Once the density, the isothermal compressibility, the thermal expansion and the isobaric heat capacity are deduced from the fluctuation theory, many other second-order derivative properties can be assessed through the use of the classical thermodynamics relations.

c) Isochoric heat capacity

$$c_V^{res} = c_P^{res} + \frac{Tk_P^2}{\rho k_T} \tag{2.48}$$

d) Speed of sounds

$$V_s = \sqrt{\frac{c_P^{res} + c_P^{ideal}}{c_V^{res} + c_V^{ideal}} \frac{1}{\rho k_T}} \tag{2.49}$$

e) Joule-Thomson coefficient

$$\mu_{JT} = \frac{\rho}{c_P^{res} + c_P^{ideal}}(Tk_P - 1) \tag{2.50}$$

2.1.3 Microcanonical ensemble NVE

To simulate a system with maximum entropy, i.e. exploring a maximum of microstates, the *NVE* ensemble is the most adapted one, particularly for molecular dynamics. This is because in molecular dynamics it is a conservative approach for a system of fixed number of particles and volume. Thus, solving the equations of motions will allow the system to evolve in a natural way, exploring all the microstates. However, the challenge with this ensemble is that some properties are difficult to obtain. Additionally, such ensemble is difficult to realize experimentally as energy is always a fluctuating quantity.

In such ensemble, the partition function is:

$$\Omega = \frac{1}{h^{3N} N!} \int_{p^N} \int_{N^N} dr^N dp^N \tag{2.51}$$

The bridge equation that relates the entropy to the partition function is given by:

$$S = k_B \ln(\Omega) \tag{2.52}$$

Fundamentally, it is one of the most powerful results in thermodynamics. It states that the more accessible microstates an isolated system has, the higher its entropy will be. Thus, the change in the entropy is given by:

$$\Delta S = k_B \ln(\Omega_2/\Omega_1) \tag{2.53}$$

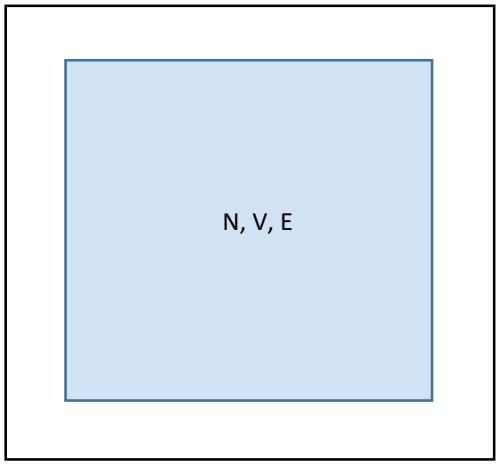


Figure 2.4: Schematic representation of the NVE ensemble

2.2 Interaction potentials

From the introduction to statistical ensembles, it is clear that the interaction energy between the molecule play an essential role. This is because, it is these interactions that allow molecules to explore the available microstates. As mentioned in the previous part, the total potential energy is given the Hamiltonian H, which is a combination of two parts, the kinetic energy and the configuration energy. Thus, one may write:

$$H = K + U \tag{2.54}$$

where *K* is the kinetic energy which only specific to molecular dynamics simulations (and will be discussed in the MD part), and U is the configurational energy coming from the way the atoms interact between them. For simplification purpose, this latter can also be split into two parts, intramolecular potential and intermolecular potentials (more precisely interatomic potentials). Thus,

$$U = U_{intra} + U_{inter} \tag{2.55}$$

2.2.1 Intramolecular potentials

The intramolecular potential is set of contributions due to the molecular internal degrees of freedom, such as the part of the potential energy due to the vibration of molecules, which is absent for monoatomic molecules. The molecular degrees of freedom are mainly of three types: molecular bonding, bending and torsion. The intramolecular potential is then a sum of all these contributions and can be expressed as follow:

$$U_{intra} = U_{Bond} + U_{Bend} + U_{Torsion} \tag{2.56}$$

Here it is supposed that these interactions do not interfere, otherwise, one should define additional cross interactions.

2.2.1.1 Stretching energy

This is the energy associated with the stretching of two bounded atoms around an equilibrium bond length l_0. A realistic mathematical function to model the stretching potential is provided by the Morse model, which takes into account the complete dissociation of the bonded atoms:

$$U_{Stretch} = D_e \left(1 - e^{-\sqrt{\frac{k}{2D_e}}(l-l_0)^2}\right)^2 \tag{2.57}$$

However, the Morse potential is not computationally efficient and it involves three parameters. Furthermore, the deviation far from the equilibrium is rarely the case as the covalent bonds are known to be of high energy. Alternatively, a simpler expression can be used to mimic this stretching energy covering the region around the equilibrium length. For this the harmonic potential type is found to be a good choice given by:

$$U_{Bond} = \frac{1}{2} K_{Stretch}(l - l_0)^2 \tag{2.58}$$

where K_{Bond} is the stiffness constant associated to the stretch. l_0 is the equilibrium length that minimises the stretching potential.

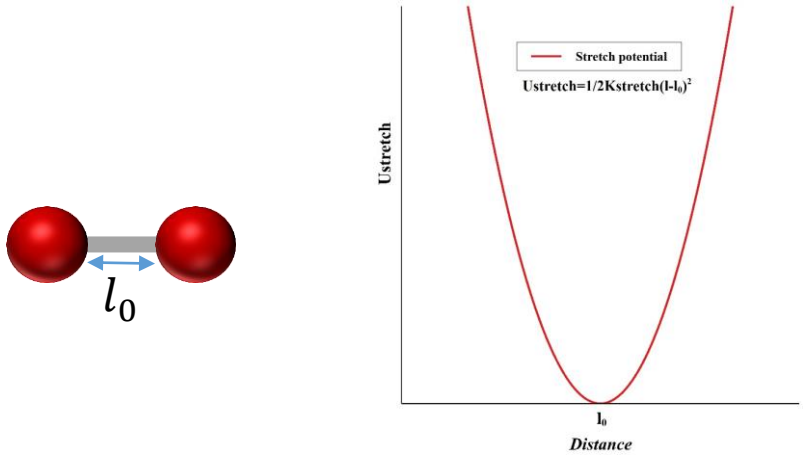


Figure 2.5: Schematic representation of the bond stretch and associated harmonic potential

2.2.1.2 Bending energy

This is the energy associated to the vibration around an equilibrium angle involving three successive bonded atoms. This energy is not as important as that of the stretching, thus, it is easier to bend the molecule rather than to stretch or compress a bond. The bending energy is also often modelled by a harmonic type given by:

$$U_{Bend} = \frac{1}{2} K_{Bend}(\theta - \theta_0)^2 \tag{2.59}$$

where K_{Bend} is the rigidity constant and θ_0 is the equilibrium angle.

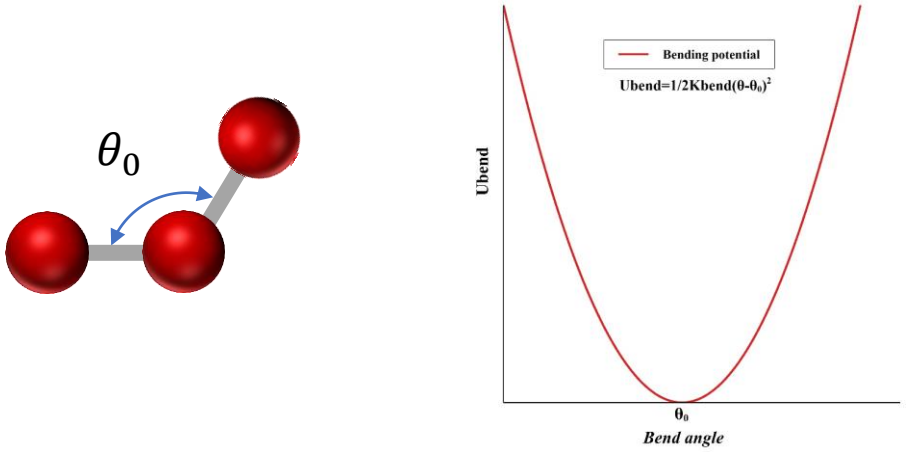


Figure 2.6: Schematic representation of the bond bending and associated harmonic potential

2.2.1.3 Torsion energy

This is the energy required for the rotation of four successive bonded atoms. Such an energy plays an essential role in the structural properties and conformational analysis. The torsion energy is often modelled with a cosine series expansion of order L given by:

$$U_{Torsion} = \sum_{n=0}^{L} \frac{1}{2} V_n (1 + \cos(n\varphi - \delta)) \quad (2.60)$$

where V_n is the barriers height energy, and φ is the torsion angle, δ the phase of the torsional term.

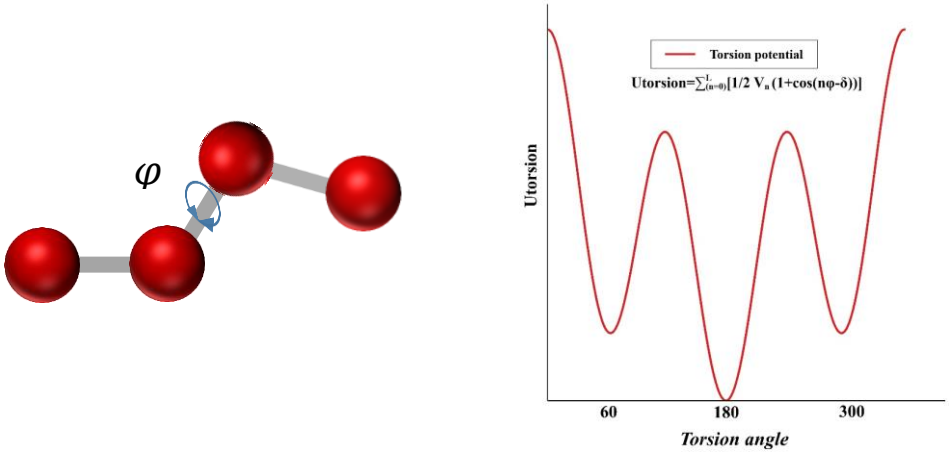


Figure 2.7: Schematic representation of the bond torsion and associated polynomial potential

2.2.2 Intermolecular potentials

Intermolecular interactions are the contribution to the total potential energy due to the interaction between atoms of different molecules. They are mainly of two types, the van der Waals interaction and the electrostatic interactions with the latter not being considered in this work.

The case of noble gases is probably the best example of van der Waals interactions, as it is well known that such molecules do not have any polar interactions, yet they can undergo phase transitions. If two molecules are placed sufficiently far apart, their interaction energy is null. However, as the distance between them, *r*, is reduced, the interaction energy becomes negative to reach a certain minimum at an equilibrium distance r_{eq}. This explains the presence of the attractive forces (dispersive forces) between the two molecules, responsible for the cohesion of the liquids (London's dispersive forces). If this distance is further reduced, the interaction energy increases and become positive, eventually diverging to infinity. This highlights the presence of a repulsive interaction that only acts at very short range distances. This is the consequence of the repulsive interactions between the two clouds of electron of the two molecules, prohibiting the total overlapping of these molecules which is known as the Pauli exclusion principle.

The exact intermolecular potential can be obtained by quantum calculations for small molecules, but it remains challenging for large molecules due to computational time and assumptions for the calculations. In the case of simple fluids like the noble gases or alkanes, the global behaviour of the long range dispersive interactions is seen to vary as r^{-6}, while that of the short range repulsive are seen to vary as r^{-n} with $n \geq 9$, or probably better as an exponential decay of the type e^{-br}. There are many models that are proposed in the literature which are often ad-hoc.

2.2.2.1 Hard sphere potential

Although it does not properly account for Van der Waals interactions, the hard sphere potential is widely used as a reference model in many theories, including statistical physics, due to its simplicity. This is the case in the kinetic theory of gases or the association theory of Wertheim and its derived theories like the SAFT type equations of state and integral equations of states as well as cubic equations of states. Similarly, many theories on transport properties also rely on this model, such as the Enskog-Chapman theory for predicting viscosity. In reality, the closest situation to this model is a gas at high temperature where only the interactions between atoms can be considered as infinitely elastic bouncing.

Mathematically, it is expressed as follow:

$$U_{HS}(r) = \begin{cases} \infty\ ; & |r| < \sigma \\ 0\ ; & |r| \geq \sigma \end{cases} \tag{2.61}$$

where σ is the impenetrable hard-sphere diameter.

2.2.2.2 Square-well potential

The previous potential is found to be sufficiently adequate to reproduce the structure of the fluids at high density. However, it only exists in a gas phase if not coupled with a perturbative approach to take into account the dispersive interaction responsible for the cohesion of liquids. The square-well potential is a good model for more realistic representation of the fluids with a capability of inducing the gas-liquid transition. The repulsive part is modelled with the HS model while the dispersive interactions are modelled with a square well given by this form:

$$U_{SW}(r) = \begin{cases} \infty\ ; & |r| < \sigma \\ -\epsilon\ ; & \lambda\sigma > |r| \geq \sigma \\ 0\ ; & |r| \geq \lambda\sigma \end{cases} \tag{2.62}$$

where ϵ is the depth of the well and λ is the range of the attractive interactions that generally takes a value of 1.5.

2.2.2.3 Lennard-Jones potential

The Lennard-Jones (LJ) potential, which requires only two parameters and exhibits very interesting features in terms of speeding out the calculations, is a widely used realistic potential. In the LJ potential, both the repulsive and attractive contribution are modelled with an inverse power law with respectively 12 and 6 as stiffness exponents. While the exponent of 6 is physically justified by the London dispersion forces, the exponent of 12 in the repulsive part is purely empirical and is used to speed the calculations by simply the squaring the repulsive part. The model is given by:

$$U_{LJ} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2.63}$$

here ϵ is the depth of the LJ-potential and σ is the collisional diameter for which the LJ-potential is zero.

A generalized LJ potential, often named Mie Potential, for any powers decay can be written as:

$$U_{Mie} = \left(\frac{n}{n-m}\right)\left(\frac{n}{m}\right)^{m/(n-m)}\epsilon\left[\left(\frac{\sigma}{r}\right)^{n} - \left(\frac{\sigma}{r}\right)^{m}\right] \tag{2.64}$$

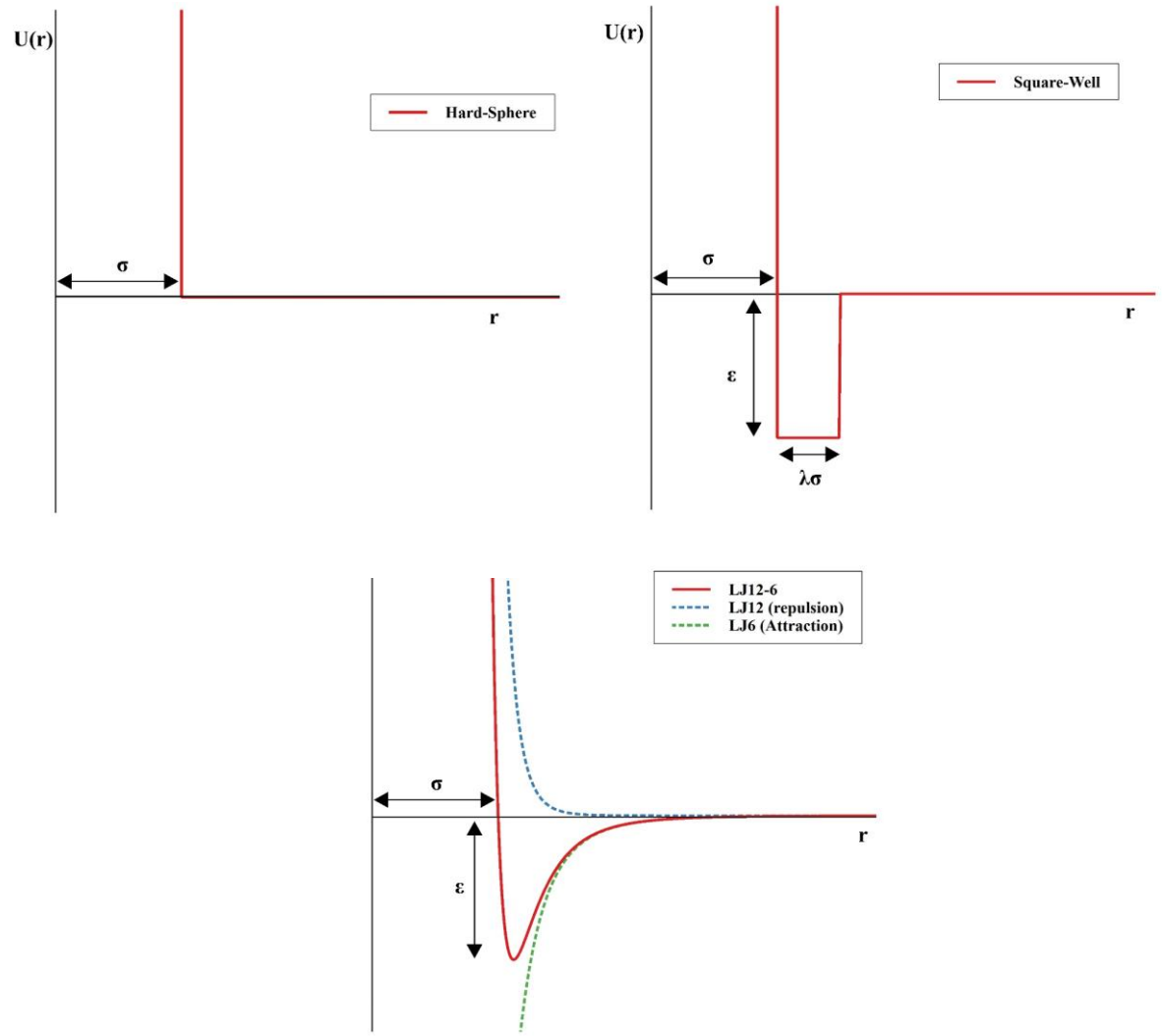


Figure 2.8: Different types of intermolecular potential interactions

2.3 Molecular simulations

There are two techniques that are used to perform molecular simulations in different ensembles, the Monte Carlo (MC) technique and the molecular dynamics technique (MD).

2.3.1 Molecular dynamics

In molecular dynamics simulations, the collection of the microstates is obtained through the time evolution of the system of particles. This dynamic evolution corresponds to different configurations of the molecules in the system that are generated by integrating the Newton's equations of motions in the phase space $(r^N(t), p^N(t))$. The properties of the material are obtained at each time step corresponding to a certain snapshot of the system, and averaged after many time steps to obtain a mean value:

$$< X >_{time} = \lim_{t \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} X\,(r^N(t), p^N(t))dt \qquad (2.65)$$

Of course, the longer the simulation, the better the estimated property.

The Newton's equations of motions are naturally conservative $\left(\frac{dH(t)}{dt} = \frac{dU^{kin}(t)}{dt} + \frac{dU^{conf}(t)}{dt} = 0\right)$. Thus, if one puts a number *N* of molecules in a certain volume *V*, and allows the system to evolve by solving the equations of motions, this will correspond to the *NVE* ensemble, with *E* being the total energy of the system. The Newton's second law is:

$$\begin{cases} F_i(t) = m_i a_i(r^N(t)) \\ F_i(t) = -\frac{\partial U_i(r^N(t))}{\partial r^N(t)} \end{cases} \quad (2.66)$$

where m_i and $a_i(t)$ are respectively the mass and the acceleration of the particle "i". $F_i(t)$ and $U_i(r^N(t))$ are respectively the force and the interaction potential acting on particle "i" by the remaining particles in the system. The force acting on particle "i" by particle " j" should be reciprocal, which leads to the Newton's third law:

$$F_i(t) = -F_j(t) \quad (2.67)$$

These equations are solved for all the particles in the system at each time step. Of course, solving them analytically is just impossible, thus, numerical methods are required. Many algorithms have been proposed to solve the equations of motions, including the Verlet algorithm [1212], the Leapfrog algorithm [13], the Verlet velocity algorithm [14], the Beeman algorithm [15], predictor-corrector [7], Runge-Kutta algorithm...

The global workflow of the molecular dynamics simulations corresponds to the steps described in Figure 2.9.

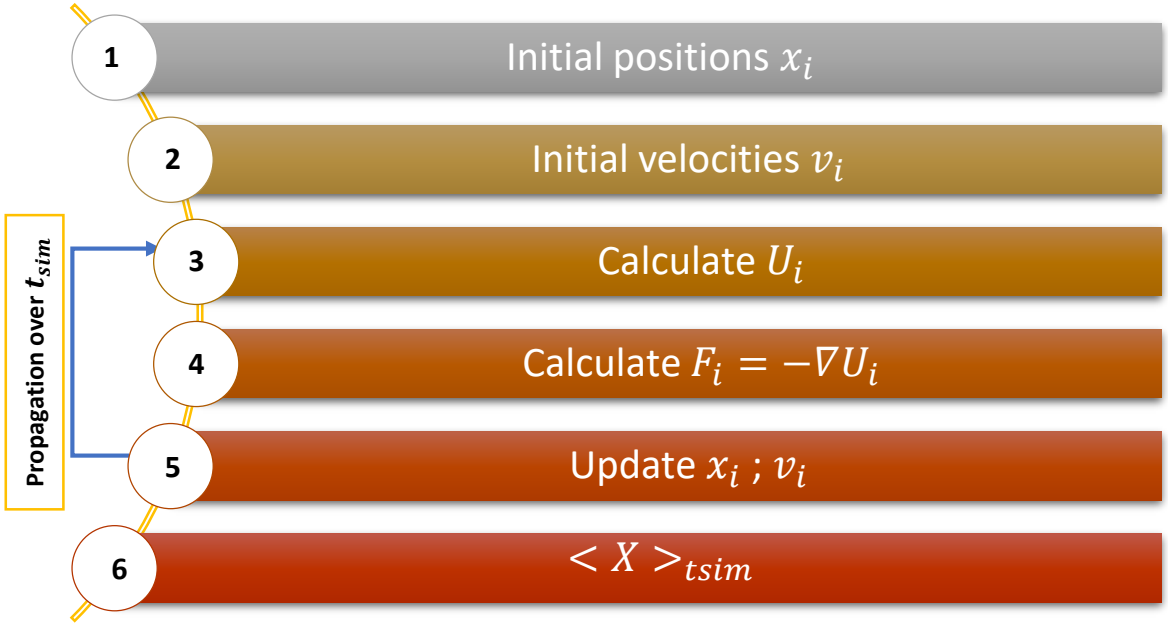


Figure 2.10: Workflow of the molecular dynamics simulation

To initialize the system, the positions of atoms can be generated in two ways. The first consists in placing the atoms inside the box randomly, while the second consists in placing them on a lattice. Concerning the initial velocities, they are obtained according to the Maxwell-Boltzmann distribution, where, at a given temperature, the velocity of the particles in the system should follow this distribution:

$$f(v_i) = \left[\frac{m_i}{2\pi k_B T}\right]^{1/2} e^{-\frac{m_i v_i^2}{2k_B T}} \tag{2.68}$$

However, at equilibrium, the centre of mass of the fluid should not move, thus, the velocity generated should satisfy the criterion on the total momentum of the system, given by:

$$p = \sum_i m_i v_i = 0 \tag{2.69}$$

Once the positions are known, and the force field is chosen, the interaction energy between the particles are easily obtained by evaluating this model. In fact, the model here refers to the sum of each contribution to the interaction potential, thus, the total energy is:

$$U = U_{vdw} + U_{strech} + U_{bend} + U_{Torsion} + \cdots \tag{2.70}$$

The forces, and so the accelerations, are then deduced by using the derivative of the potential interaction model. Then, updated velocities and positions are deduced from a time integrator.

2.3.1.1 Time integrator

As already pointed out, many integrators exist in the literature, but the Verlet velocity algorithm is probably the most used one in molecular dynamics simulations. The main advantage of this algorithm is that the positions, velocities and accelerations are all known for each step $(t + \delta t)$ of the simulation where δt is the time step. This algorithm is provided by the following equations:

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) \tag{2.71}$$

$$v\left(t + \frac{\delta t}{2}\right) = v(t) + \frac{\delta t}{2} a(t) \tag{2.72}$$

$$a(t + \delta t) = -\frac{1}{m_i} F(t + \delta t) \tag{2.73}$$

$$v(t + \delta t) = v(t) + \frac{\delta t}{2}[a(t) + a(t + \delta t)] \tag{2.74}$$

2.3.1.2 Time step

The choice of the time step is an important question as it depends on several factors. In fact, the finer is the time step, the more correct the simulated trajectory and the better the phase space covered, but this requires more *CPU* time to reach a certain run duration. In the case of chain molecules, the time step should be lower than the shortest period of vibration:

$$\delta t \ll 2/f_{max} \tag{2.75}$$

where f_{max} is the maximum vibrational frequency.

Generally, the highest frequency is that corresponding to stretching. However, in many cases, such vibrations have only a minimal effect on the overall behaviour of the system. Thus, in many force fields, they are frozen out, enabling to take larger time step.

In the case of too large time step, this may lead to some instabilities in the dynamics, such as when two particles overlap too much, resulting in their ejection from the box at the next time step if $\delta t v(t) > \frac{L_{box}}{2}$. A good balance is then needed to satisfy all the factors mentioned previously.

2.3.1.3 Constraints for non-spherical molecules

It is common in molecular dynamic simulations to deal with constrained molecules. There are many reasons for this, for example, to avoid the presence of high vibrational frequencies, which impose taking very short time step. An additional reason is that it simplifies the development of a theory for a particular model. A good example for this is the molecular model we are going to study in this work which consists in chains composed of spheres that are tangentially bonded with an infinite bonding energy. Thus, the simulation of this model requires the use of time-independent techniques to constrain these bonds. To do so, the two most popular algorithms are the Rattle algorithm [16] and the shake algorithm [17]. It is the Rattle method that is used in this work, and thus the only one that will be discussed in the following.

This algorithm consists in correcting the unconstrained dynamics though the coordinates, in a first step, and the velocities in a second step. More precisely, the corrected (new) positions are given by:

$$\begin{cases} r_i^{new} = r_i^{old} - \gamma r_{ij}^{old} \\ r_j^{new} = r_j^{old} + \gamma r_{ij}^{old} \end{cases} \tag{2.76}$$

where i and j denote the two atoms in consideration. r_{ij}^{old} is the unconstrained distance between the two atoms. γ is an adjustable parameter which should satisfy:

$$\left(r_{ij}^{new} \right)^2 = \left(b_{ij} \right)^2 \tag{2.77}$$

where b_{ij} is the length of the bond considered. From the system of equation and the condition above, if the quadratic term $\left(4\gamma^2 \left(r_{ij}^{old} \right)^2 \right)$ is neglected, one obtains:

$$\gamma = \frac{\left(r_{ij}^{old} \right)^2 - \left(b_{ij} \right)^2}{4\left(r_{ij}^{old} \right)^2} \tag{2.78}$$

The process is repeated iteratively until:

$$\left| \left(r_{ij}^{new} \right)^2 - \left(b_{ij} \right)^2 \right| < \varepsilon_r \left(b_{ij} \right)^2 \tag{2.79}$$

where ε_r is the specified tolerance. If the mass of the considered atoms are not the same, please refer to the general method given in the Rapaport's book [9].

Similarly, the velocities are corrected using these equations:

$$\begin{cases} v_i^{new} = v_i^{old} - \gamma v_{ij}^{old} \\ v_j^{new} = v_j^{old} + \gamma v_{ij}^{old} \end{cases} \quad (2.80)$$

$$2v_{ij}^{old} r_{ij} = 0 \quad (2.81)$$

where γ is given by:

$$\gamma = \frac{v_{ij}^{old} r_{ij}}{2(b_{ij})^2} \quad (2.82)$$

2.3.1.4 Molecular Dynamics in different ensembles

Normally the NVE ensemble does not require any external intervention as *N* and *V* are explicitly imposed when defining the system, and the energy is naturally conserved with Newton's equations of motions. However, imposing the energy of the system may imply change in temperature (and pressure). The NVT and NpT ensembles, on the other hand, require explicit constraints to maintain the imposed temperature (and pressure) in order to ensure the correct statistical ensemble and thus, correct properties collected. Therefore, techniques such as thermostating and barostating are introduced to control these conditions. However, the challenge is to choose a technique that has as little impact as possible on the natural dynamics of the molecules in the system while satisfying these external conditions.

2.3.1.4.1 Thermostats

Experimentally, a thermostat is achieved by making the system in contact with a bath that controls its average temperature. In MD simulations, the temperature is not directly controlled. It is inferred through the kinetic energy as they are connected by the equipartition energy principle:

$$\langle U^{kin} \rangle = \frac{N_f k_B T}{2} \quad (2.83)$$

where $N_f = 3N - N_{constraints}$ is the number of degrees of freedom (=3*N* for monoatomic molecules), and the instantaneous kinetic energy is given by:

$$U^{kin} = \sum_{i=1}^{N} \frac{m_i v_i^2}{2} \quad (2.84)$$

The instantaneous temperature is then obtained with the following formula:

$$T = \frac{1}{N_f k_B} \sum_{i=1}^{N} m_i v_i^2 \quad (2.85)$$

and thus, the average temperature writes:

$$\langle T \rangle = \langle \frac{1}{N_f k_B} \sum_{i=1}^{N} m_i v_i^2 \rangle \quad (2.86)$$

It is clear from this formula, that it is the velocity over which one intervenes to satisfy the temperature. However, the way this velocity is changed is extremely important and this depends on the algorithm technique that is used. Among them one can cite, the velocity rescaling [18], the Anderson thermostat [19], the Berendsen thermostat [20], the Nosé-Hoover thermostat [21], the Langevin thermostat [22]…

a) velocity rescaling

The velocities are rescaled every time step to satisfy the instantaneous temperature of the system. The temperature change for an unconstrained system corresponds to:

$$\Delta T = \frac{1}{2}\sum_{i=1}^{N}\frac{2}{3}\frac{m_i(\lambda v_i)^2}{Nk_B} - \frac{1}{2}\sum_{i=1}^{N}\frac{2}{3}\frac{m_i(v_i)^2}{Nk_B} \quad (2.87)$$

$$\Delta T = (\lambda^2 - 1)T_{current}(t) \quad (2.88)$$

$$\lambda = \sqrt{T_{new}(t)/T_{current}(t)} \quad (2.89)$$

where λ is the rescaling factor so that $v_i^n = \lambda v_i^{n-1}$**.** This method can be considered as the simplest, but is an abrupt way to control the temperature. The dynamics deviate from the correct trajectory as it removes the natural temperature fluctuation. Therefore, the time dependent properties (transport properties) are affected as the time correlation function is also affected. However, this can be used during some steps at the beginning of the simulation to accelerate its "equilibration" to a target temperature.

b) Berendsen

Instead of an abrupt rescaling which kills all natural temperature fluctuations, the Berendsen thermostat proposes a smoother way to do it by rescaling the velocities gradually at each time step. In this method, the system is coupled to an external heat bath which acts as a source of thermal energy, a closer way to mimic the real experiments. The coupling is given by the following equation whose solution is an exponential decay:

$$\frac{dT(t)}{dt} = \frac{1}{\tau_T}\left(T_{bath}(t) - T(t)\right) \quad (2.90)$$

where τ_T is the coupling parameter or the relaxation time which controls the strength of the coupling. It is generally set in the range $[500\delta t - 1000\delta t]$ to avoid a too strong perturbation of the dynamics. Note that if $\tau_T \to \infty$ no coupling is made, and if $\tau_T \to \delta t$ it is possible to recover the rescaling velocity algorithm. The scaling factor is hence given by:

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T}\left(T_{bath}(t)/T(t)\right)} \quad (2.91)$$

c) Nosé-Hoover Thermostat

The concept of thermostating is quite different in this deterministic approach, where the heat bath reservoir is considered as a part of the system, and thus, considered as an extra degree of freedom noted *s*. The variable *s* is the scaling factor of the momentum to adjust the temperature of the particles written as:

$$v_i^n = s v_i^{n-1} \quad (2.92)$$

Its potential and kinetic energy are given by:

$$U_s = (3N + 1)k_B T_{bath}\ln(s) \quad (2.93)$$

$$U_{kin,s} = 1/2\,Q_s(ds/dt)^2 \quad (2.94)$$

where "1" is the additional degree of freedom associated to the reservoir and Q_s is the coupling parameter between the reservoir and the real system expressed in [energy/time²].

The Hamiltonian of the extended system should read:

$$H = U \ + U_s + \ U_{kin} + U_{kin,s} \tag{2.95}$$

However, instead of scaling the velocities and the time explicitly, a friction factor is defined as:

$$\xi = p_s/Q_s \tag{2.96}$$

which can be integrated in the equations of motions as follow:

$$m_i a(t) = F_i(t) - \ \xi_i m_i v_i \tag{2.97}$$

where the evolution in time of the friction factor is given by:

$$\frac{d\xi}{dt} = (3N + 1)k_B/Q_s[T(t) - T_{bath}] \tag{2.98}$$

2.3.1.4.2 Barostat

The pressure average is obtained with the following Virial expression:

$$P = \rho k_B T + \frac{1}{3V}\langle\sum_{i<j} \vec{F}(r_{ij}) \,.\, \vec{r}_{ij}\rangle \tag{2.99}$$

There are probably as many methods for regulating temperature as there are for regulating pressure. For instance, the instantaneous barostat [23] where the instantaneous volume is scaled to match the desired instantaneous pressure similarly to the velocity rescaling thermostat. One also may cite the Anderson barostat [19], Berendsen barostat [20], Nosé-Hoover barostat [24], Hoover barostat [25], Parrinello-Rahman barostat [26], Langevin barostat [27], Gauss barostat…However, in this work only the Berendsen barostat has been used.

- **Berendsen barostat**

In this method, it is the mean volume which is scaled by a correction factor which depends on the variation between the current mean pressure and the desired pressure of the external reservoir in contact $\Delta P = P(t) - P_{desired}$. The coupling is given by the following equation whose solution is an exponential decay:

$$\frac{dP(t)}{dt} = \frac{1}{\tau_P}\left(P_{desired}(t) - P(t)\right) \tag{2.100}$$

where τ is the coupling parameter or the relaxation time which controls the strength of the coupling. It is generally set in the range $[500\delta t - 1000\delta t]$ to avoid strong perturbation of the dynamics. Note that if $\tau_P \to \infty$ no coupling is made, and if $\tau_P \to \delta t$ it is possible to recover an instantaneous barostat.

The volume is scaled with this factor given by:

$$\lambda = 1 - k_T \frac{\Delta t}{\tau_P}\left(P(t) - P_{desired}(t)\right) \tag{2.101}$$

where k_T is the isothermal compressibility.

Thus, the new volume reads $V^n = \lambda V^{n-1}$ which is also equivalent to scaling the positions with $\lambda^{1/3}$. Hence, the new positions are $r_i^n = \lambda^{1/3} r_i^{n-1}$.

2.3.2 Monte Carlo

In MD simulations, new configurations are generated by the forces between particle of the system according the Newton's equations of motion. However, in Monte Carlo simulations, these configurations, corresponding the microstates of a certain ensemble, are generated randomly. Once these configurations are obtained after N_{trial} of iterations, the average properties $\langle X \rangle$ are obtained in the canonical ensemble as follows:

$$\langle X \rangle_{NVT} = \sum_{i=1}^{N_{trial}} X_i \, \rho_i^{NVT} \tag{2.102}$$

where ρ_i^{NVT} is the probability density, and is given by:

$$\rho_i^{NVT} = \frac{Exp\left(\frac{-H\left(r^N, p^N\right)}{k_B T}\right)}{Q_{NVT}} \tag{2.103}$$

Taking the advantage that the partition function can be separated into an ideal contribution times the residual contribution due to interactions (cf. introduction to ensembles), the Monte Carlo simulation does not deal with kinetic (ideal) part, as this can be obtained analytically, but only with the configurational part. Therefore, the probability density corresponds only to the configurational one given as:

$$\rho_i^{NVT} = \frac{Exp\left(\frac{-U\left(r^N\right)}{k_B T}\right)}{Q_{NVT}^{conf}} \tag{2.104}$$

The main question now is how Monte Carlo samples the phase space? Analytically it is just impossible, and numerically using quadrature or the simple Monte Carlo method (not to be confused with the Monte Carlo methods used in molecular simulation) is just unfeasible because of the extremely large number of iterations needed. In fact, the problem with numerical quadrature-like methods is that they assign the same probability for each configuration, even though there are an extremely large number of configurations that are unphysical and correspond to overlapping particles, which leads to a vanishing Boltzmann factor. However, if computational resources were not an issue, there are not limit and the average of the quantity *X* is simply calculated by:

$$\langle X \rangle = \frac{1}{N_{trial}} \sum_{i=1}^{N_{trial}} X_i \tag{2.105}$$

2.3.2.1 Metropolis algorithm

Instead of generating configurations randomly with equal probability and then weighting them by a Boltzmann factor $Exp\left(\frac{-U(r^N)}{k_B T}\right)$, alternatively, a clever way it to choose configurations with a probability $Exp\left(\frac{-U(r^N)}{k_B T}\right)$ from the outset and then weight them equally. This is the heart of the N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller work [28] in 1953, often referred to as the Metropolis algorithm.

Once the configurations are chosen in the low energy region, the second step consists of determining whether these configurations will drive the system to a more stable state or to a higher energy state level. In other words, generating the right configurations does not mean that they are systematically accepted in order to move the system from the current microstate "1" to microstate "2". Thus, how is this decision made? Such a decision relies on the microscopic reversibility condition of the Markov chain, which states that at equilibrium the transition between states "1" and "2" occurs at the same rate. Thus if $\pi_{1\to2}$ is the transition matrix from state "1" to state "2" ($\pi_{2\to1}$ is the inverse), and ρ_1 is the probability density at state *"1"*(ρ_2 for state *"2"*), the condition of microscopic reversibility reads:

$$\pi_{1\to2}\rho_1 = \pi_{2\to1}\rho_2 \tag{2.106}$$

where the transition matrix can be separated as follow:

$$\pi_{1\to2} = \alpha_{1\to2} acc_{1\to2} \tag{2.107}$$

where $\alpha_{1\to2}$ is the probability of proposing the transition $\pi_{1\to2}$, and $acc_{1\to2}$ is the probability to be accepted. Thus, going from one state to another, goes through two steps; first is proposed a certain move $\alpha_{1\to2}$ (elements of the transition matrix), and second, it is decided whether to accept or reject it.

In the work of Metropolis, the probability of choosing to move from state *"1"* to state *"2"* is set the same as that of moving from *"2"* to *"1"*. Thus, the stochastic matrix α is said to be symmetrical and it is possible to write:

$$\alpha_{1\to2} = \alpha_{2\to1} \tag{2.108}$$

Hence, the ratio between these transition matrix elements gives:

$$\frac{\pi_{1\to2}}{\pi_{2\to1}} = \frac{acc_{1\to2}}{acc_{2\to1}} = \frac{\rho_2}{\rho_1} \tag{2.109}$$

There are choices for *acc* that obeys this condition, and the one chosen by Metropolis is:

$$\begin{cases} acc_{1\to2} = 1 \ ; & \rho_2 \geq \rho_1 \\ acc_{1\to2} = \frac{\rho_2}{\rho_1} \ ; & \rho_2 < \rho_1 \end{cases} \tag{2.110}$$

which can be summarized as:

$$acc_{1\to2} = min\left[1, \frac{\rho_2}{\rho_1}\right] \tag{2.111}$$

Thus, If the energy of state *"2"* is lower or equal to the energy of state *"1"*, the system is considered at equilibrium or moving towards a more stable equilibrium. Then, the transition is accepted. If the energy of state *"2"* is now higher than energy of state *"1"*, this move is not systematically rejected. However, it can be accepted with a probability given by $\frac{\rho_2}{\rho_1}$. The decision for this case is taken by comparing this ratio to a random number *Rand* obeying a uniform distribution as summarized in Figure 2.11. This leads to two possible situations:

$$\begin{cases} Rand \leq \frac{\rho_2}{\rho_1} \ \to \ accept \\ Rand > \frac{\rho_2}{\rho_1} \ \to \ reject \end{cases} \tag{2.112}$$

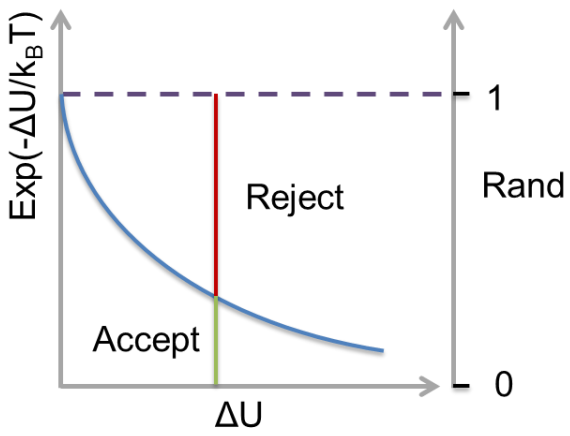


Figure 2.11: Boltzmann analysis of the Metropolis algorithm

2.3.2.2 Type of moves

There are many types of move that can be applied to molecules to allow the transition from one microstate to another. These moves ensure that the molecules will explore as much of the phase space as possible, thus ensuring that the ergodicity is respected to some extent.

a) Translational move

In the translational type of move, the whole molecule is translated to a new position by changing its Cartesian coordinates according to the following equations:

$$\begin{cases} x_{new} = x_{old} + (2\chi - 1)\delta r_{max} \\ y_{new} = y_{old} + (2\chi - 1)\delta r_{max} \\ z_{new} = z_{old} + (2\chi - 1)\delta r_{max} \end{cases} \quad (2.113)$$

where is χ a random number, and δr_{max} is the maximum displacement allowed.

Once this new configuration is generated, it is subject to the process of decision making according to the Metropolis algorithm. If the new configuration has a lower energy than the previous one, the translation move is accepted. Otherwise, one should compare a random number *Rand* to the probability ratio of the two states, which can be expressed in terms of the Boltzmann factor:

$$\alpha_{1\to 2} = min\left[1,\ Exp\left(\frac{-\left(U_2(r^N) - U_1(r^N)\right)}{k_B T}\right)\right] \quad (2.114)$$

The choice of the maximum displacement is very important as it affects the efficiency of the exploration of the phase space. If is this too large, many of the trial move will be rejected due to overlapping, and if it is too small, many of the moves will be accepted, but the phase space will be poorly explored. In practice, the maximum displacement is often set to change automatically in order to ensure that approximately 50% of the trial move are accepted.

a) Rotational move

The rotational moves are achieved in a similar way to the translation. First, a randomly chosen molecule is rotated in a random direction by applying a small change to the Euler angles $\emptyset, \theta, \psi$. For rigid chains this writes:

$$\begin{cases} \emptyset_{new} = \emptyset_{old} + (2\chi - 1)\delta\emptyset_{max} \\ \theta_{new} = \theta_{old} + (2\chi - 1)\delta\theta_{max} \\ \psi_{new} = \psi_{old} + (2\chi - 1)\delta\psi_{max} \end{cases} \quad (2.115)$$

where $\delta\emptyset_{max}, \delta\theta_{max}, \delta\psi_{max}$ are the maximum change in the Euler angles. The acceptance or the rejection of the rotation move decision is taken upon:

$$\alpha_{1\to2} = min\left[1, \frac{\sin(\theta_{new})}{\sin(\theta_{old})} Exp\left(\frac{-(U_2(r^N) - U_1(r^N) + p\Delta V)}{k_B T}\right)\right] \quad (2.116)$$

The sampling in θ may lead to a problem if θ_{old} is equal zero. Alternatively, the idea is to sample using $\cos(\theta)$, thus, leading to:

$$cos\,(\theta_{new}) = \cos(\theta_{old}) + (2\chi - 1)\cos(\theta)_{max} \quad (2.117)$$

b) Volume move

The volume move is characteristic of the *NpT* ensemble, where the volume of the system is allowed to vary during the simulations. Thus, the simulation box may be expanded or shrunk, in one, two or three directions with ΔV

$$V_{new} = V_{old} + \Delta V \quad (2.118)$$

The new coordinates of the volume are given :

$$\begin{cases} Vx_{new} = Vx_{old} + (2\chi - 1)\delta V_{max} \\ Vy_{new} = Vy_{old} + (2\chi - 1)\delta V_{max} \\ Vz_{new} = Vz_{old} + (2\chi - 1)\delta V_{max} \end{cases} \quad (2.119)$$

δV_{max} is the maximum volume change.

The decision criterion is given by:

$$\alpha_{1\to2} = min\left[1, \left(\frac{V + \Delta V}{V}\right)^N Exp\left(\frac{-(U_2(r^N) - U_1(r^N) + p\Delta V)}{k_B T}\right)\right] \quad (2.120)$$

Note that, the Cartesian coordinates should also be updated, however, the relative distance between the molecules should remain the same.

Here, have only been discussed some of the main moves. However, there are many other moves developed to improve the sampling when considering, for instance, cyclic molecules or long flexible chain molecules. Example include the reptation move, flip move, progressive regrowth move…

c) Insertion move

Some moves are specific to certain ensemble, such as the transfer of molecule from one box to another in the Gibbs ensemble, or the deletion move in the Grand Canonical ensemble. The

insertion move, on the other hand, can be used in all ensembles to compute the chemical potential using the Widom's insertion method. However, the insertion move is not done in a random way, as it is unlikely that a molecule would be inserted in a dense phase randomly due to the high probability of overlap. If the molecules are considered as spherical particles, this probability (X) is extremely small. However, for chain molecule composed of n spherical particles, this probability is even smaller (as it scales with X^n). Consequently, this requires an extremely large number of trials in order to succeed with an insertion, limiting the use of the original Gibbs ensemble to small molecules. To overcome this limitation, a Configurational Bias Monte Carlo (CBMC) [29, 30, 31] technique has been proposed, which takes advantage of the flexibility of the molecule. The chain molecule is inserted progressively, atom by atom, until the whole chain is inserted. For each atom, k attempts are considered corresponding to many possible orientations, each characterized by a certain probability, then one is selected. The probability to insert atom 'i', where 'i' indicates its position in the chain is:

$$P(i) = \frac{Exp\left(\frac{-U(i)}{k_B T}\right)}{\sum_{j=1}^{k} Exp\left(\frac{-U(j)}{k_B T}\right)} \tag{2.121}$$

where $U(j)$ is the total configurational energy of a jth trial direction.

However, by doing so, the insertion move is now biased, as the algorithm is designed to sample only favourable regions in the phase space (generate configurations that follow the Boltzmann distribution). Thus, the probability of proposing the transition α involved in the microscopic reversibility condition of the Markov chain, is no longer symmetrical, but, depends on its configurational energy:

$$\begin{cases} \alpha_{1\to 2} = f(U_2^{conf}) \\ \alpha_{2\to 1} = f(U_1^{conf}) \end{cases} \tag{2.122}$$

With this, the acceptance rule is modified, and takes into account the fact that a bias was introduced to enhance the acceptance. Thus, to remove this bias, the modified rule reads:

$$\frac{acc_{1\to 2}}{acc_{2\to 1}} = \frac{f(U_2^{conf})}{f(U_1^{conf})}\frac{\rho_2}{\rho_1} \tag{2.123}$$

In the Gibbs ensemble [30], this can be expressed as:

$$\alpha_{1\to 2} = min\left[1,\ \frac{(N-n_1+1)V_1}{n_1(V-V_1)}\frac{\rho_2}{\rho_1}\right] \tag{2.124}$$

where V and N are respectively the total volume and particle of the system. V_1 and $V_2 = (V - V_1)$ are the volumes of two boxes. n_1 and $n_2 = (N - n_1)$ are the number of particles in the two boxes.

Actually, the CBMC method is not only applied to the insertion move, but even for the displacement and rotations…

2.3.3 Monte Carlo or Molecular dynamics?

Normally, if there is no justification to use Monte Carlo, molecular dynamics is the natural choice. In this work, the time dependent properties, i.e. transport properties, were obtained from molecular dynamics as this is the only option. However, to study the phase equilibria of not too

large molecules, the Monte Carlo Gibbs ensemble [32] was the best choice even though this can also be obtained by MD simulations. In fact, the use of MD for phase equilibria where an interface is explicitly involved requires very large systems which is computationally not efficient. Moreover, the interface represents a high free energy barrier for the transfer of molecules from one phase to another, making it difficult to reach thermodynamic equilibrium.

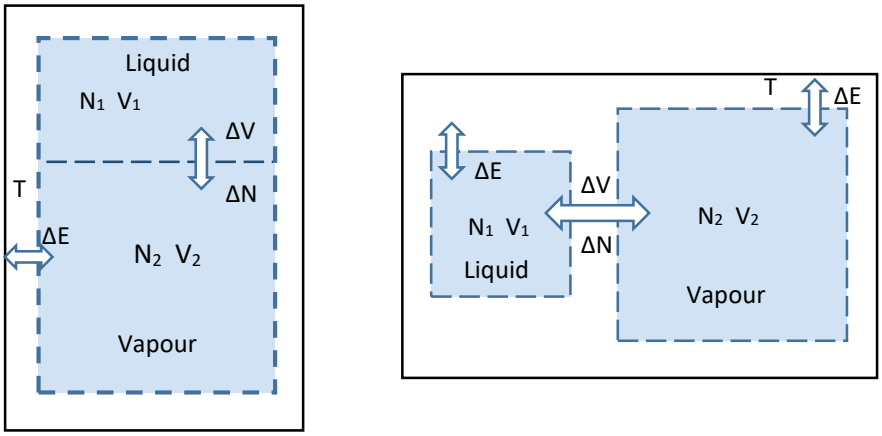


Figure 2.12: Schematic representation of the Gibbs ensemble. left: Molecular dynamics, right: Monte Carlo

The monophasic phase equilibrium properties in the NVT or NpT ensembles can be studied with both techniques, with preference in this work to MD techniques as the properties can be obtained in a very reasonable CPU time. However, when a single chain is considered, the Monte Carlo is the preferred choice, as it is well-known that many barriers exist in this case and Monte Carlo is able to "jump" them by generating configurations (it may explore the high energy regions of the phase space) that MD cannot access. As a final comment, the chemical potential and derived properties (entropy, free energy…) always requires exchange (insertion, swapping, deletion) of particles, which are not related to real dynamics, thus, Monte Carlo is the preferable option for such properties.

2.3.4 Optimization of the simulation efficiency

2.3.4.1 Periodic Boundary Conditions PBC

Periodic boundary conditions are used to overcome the fact that the simulation box size is finite due to a finite number of particles simulated. Thus, the molecules close to the boundaries will exhibit different behaviour due to the boundary effect or "surface effects" from that of the bulk. Consequently, their thermophysical properties will be different from those of the bulk. In fact, it is the fraction of the molecules that experience the boundaries which is important. Thus, to reduce this, one may tend to increase the size of the system, so that this fraction becomes very small and their influence on the total bulk thermophysical properties is negligible. However, such a solution is not efficient, as it requires high computational time ($time_{CPU} \propto N^2$). Alternatively, periodic boundary conditions can be applied, where the simulation box is replicated in all the space directions. This means that every time a molecule leaves the central box, it is immediately replaced by its image which enters from the opposite side with exactly

the same properties as the one that is replaced. The range of interaction for a certain molecule should not exceed the radius of $L/2$, to avoid self-interactions, known as the minimum image convention (MIC). Doing so, makes all the molecules of the system to always be in an infinite bulk. The molecules close to the boundaries now interact with the molecules in the image box.

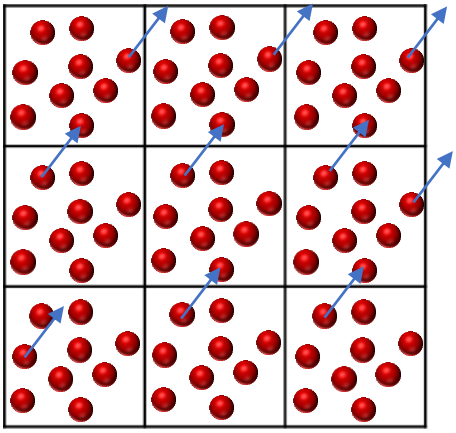


Figure 2.13: Schematic representation of the PBC

2.3.4.2 Cut-off radius & Long range corrections

Intermolecular forces calculation is the most demanding operation in terms of CPU time, thus, optimizing it turns out necessary. In fact, for the evaluation of the forces, ideally, a molecule should explore all the interactions with molecules within the simulation box. However, as already mentioned, the computational increases as N^2, whereas, for many intermolecular interactions models, the configurational energy due to the long range interactions is mostly coming from the immediate surrounding molecules, and only a small contribution comes from the molecules beyond a certain threshold radius called the cut-off distance. For instance, if one considers the Lennard-Jones potential, the contribution at distance $r = 2.5\sigma$ is only around 1.63% of its value at $r = r_{min}$ corresponding to the well-depth $U_{LJ}(r_{min}) = \epsilon$. Thus, it is reasonable to truncate the interaction at a distance radius around $r_{cutoff} \approx 2.5\sigma < L/2$ and neglecting the interactions beyond the cut-off radius by setting $U_{LJ}(r_{cutoff} < r) = 0$. The part of the energy neglected can then be easily obtained analytically using the following long range correction (where the fluid is considered isotropic, thus, g(r)≈1):

$$U^{correction} = 2\pi\rho N \int_{r_{cut}}^{\infty} g(r)U(r)r^2 dr = 0 \tag{2.125}$$

In the NVT ensemble, this correction is constant and so is calculated only once (at the end of the simulation), however, it should be recalculated each time the volume changes in the *NpT* and Gibbs MC ensembles. Similar long range corrections exist for other properties, like pressure.

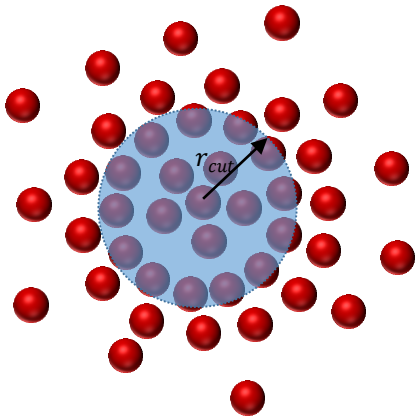


Figure 2.14: Schematic representation of the cut-off radius

2.3.4.3 Verlet Neighbor Lists

Even though the forces with the molecules beyond the cut-off radius are not calculated, the computational time is not fully optimized as at each time step or move, a check should be done on all molecules to determine whether they are inside or outside the cut-off distance. The Verlet list method, proposed by Loup Verlet [33], suggests creating lists of particles close to the one being processed by taking advantage of the fact that the molecules remain in the same region of space during a few integration steps. The new radius r_{verlet} which allows the creation of this Verlet list is chosen in such a way that a molecule located outside the sphere described by this radius does not have the time, between two updates of the list (usually 10 to 20 steps), to penetrate the interaction zone of the studied molecules.

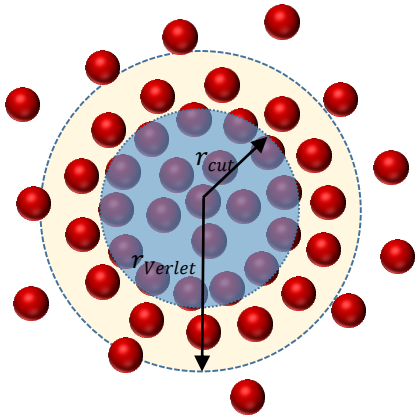


Figure 2.15: Schematic representation of the Verlet List and Verlet radius

2.3.5 Transport properties and structural properties calculation

2.3.5.1 Viscosity

The viscosity can be calculated during equilibrium molecular dynamics (EMD) simulations using the Green-Kubo formula or the Einstein formula [34, 35]. However, such method is computationally demanding, as multiple independent simulations are needed and long simulations are recommended. As an alternative, particularly for dense fluids, the Reverse Non-equilibrium molecular dynamics (R-NEMD) simulations method proposed by Muller-Plathe [36] is an interesting option. In this approach, the system is perturbed to be out of equilibrium by applying an external force. In this scheme, the cause and effect are reversed: the flux is imposed and the induced field is measured. This method takes the advantage over some other NEMD methods in cases where the flux is difficult to be defined microscopically or difficult to be converged. It is this method which is used in this work for the calculation of the liquid viscosity.

In terms of simulation details, as shown in Figure 2.16, this R-NEMD method consists in dividing the simulation box into $N_s = 24$ slabs in the z-direction. The momentum is exchanged in the x-direction by exchanging particle velocities, between the first and the $\frac{N_s}{2}$ slab, and between the $\frac{N_s}{2} + 1$ slab and the N_s slab, every N_{swap}time steps, inducing than a shear stress which is calculated using:

$$J_{xz} = \frac{\Delta p_x}{2A\delta t L_x L_y} \tag{2.126}$$

where, Δp_x is the total exchanged momentum during the time t, L_x and L_y are the lengths of the simulation box in the x and y directions, respectively.

At the stationary state, viscosity is calculated using the newton's law for the viscosity of fluids by:

$$\eta = -\frac{J_{xz}}{\frac{\partial v_x}{\partial z}} \tag{2.127}$$

where $\frac{\partial v_x}{\partial z}$ corresponds to the shear rate of the velocity profile in one half box (symmetrical) evaluated thanks to the local velocity computed in each slab. It is worthwhile to note that the choice of the N_{swap} is appropriately done depending on temperature, density, the stiffness coefficient and the chain length to avoid any shear thinning that may occur during the simulation. The shear thinning check is done by evaluating the first and second normal stresses differences: $N_1 = \tau_{xx} - \tau_{yy}$ and $N_2 = \tau_{yy} - \tau_{zz}$ which should be zero in the case of isotropic fluid [37].

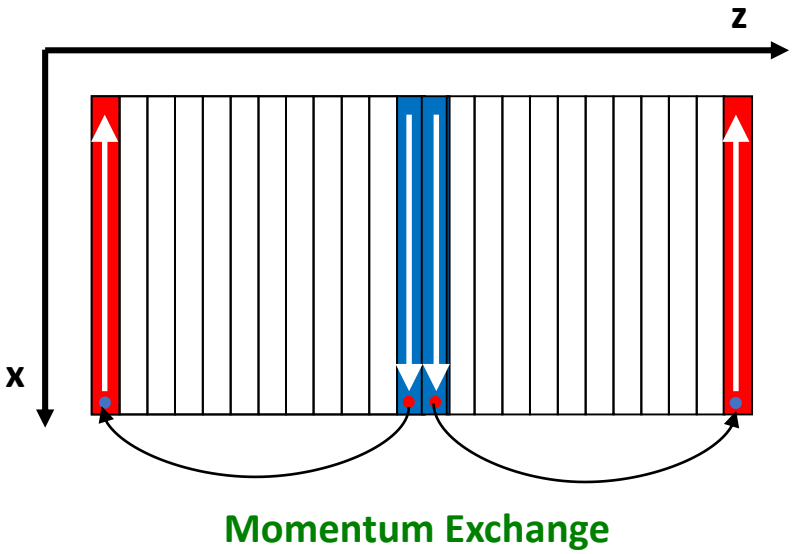


Figure 2.16: Schematic view of the momentum transfer process proposed by Müller-Plathe. The white arrows indicate the direction of the shear. The red colour indicates the hot slabs (& particles) and the blue indicate the cold slabs (& particles). The black arrows indicate the direction of the momentum exchange.

2.3.5.2 Self-diffusion coefficient

The self-diffusion coefficient is obtained from EMD [35] simulations using the Einstein relation given by the following formula:

$$D = \frac{1}{2d} \lim_{t \to \infty} \frac{<[x(t_0+t) - x(t_0)]^2>}{t} \quad (2.128)$$

Where $< [x(t_0 + t) - x(t_0)]^2 >$ is the mean square displacement (msd), $x(t_0)$ is the molecular centre of mass at the original position, $x(t_0 + t)$ is the molecular centre of mass position after time t, d is the number of dimensions (here is equal to 3) and D is the self-diffusion coefficient.

2.3.5.3 Radial distribution function

This radial distribution function (RDF) represents the probability that a particle has a neighbour at a given distance *r* from its center [3]. It indicates how the matter is locally structured and gives information on the state of the matter. It is related to the structure factor obtained in scattering experiments. It also plays an essential role in many theories, such as the thermodynamic

perturbation theory which is of particular interest in this work. Such quantity is easily deduced from molecular simulation technique. It consists in dividing the system spatially into shell volumes with increasing the radius r centred on particle "i" and counting the number of surrounding particles present in each of the concentric shell volumes. This operation is made for each particle in the system and propagated over the N time or move steps. They are then binned into a histogram and normalized with respect to the overall density, ρ, of the system and the volume of the shell, $4\pi r^2 dr$, where dr is the thickness of the shells. For a homogeneous fluid, the formula is given by:

$$g(r) = \frac{2V}{N^2}\left\langle \sum_i^N \sum_{i<j}^N \delta\left(r - r_{ij}\right)\right\rangle = \frac{V}{N}\frac{\langle \Delta N(r)\rangle}{4\pi r^2 \Delta r} \tag{2.129}$$

where N is the number of particles in the system volume V, is the Dirac distribution and $\langle \Delta N(r)\rangle$ is the mean number of particles in a shell volume of thickness Δr at distance r.

This idea is illustrated in Figure 2.17.

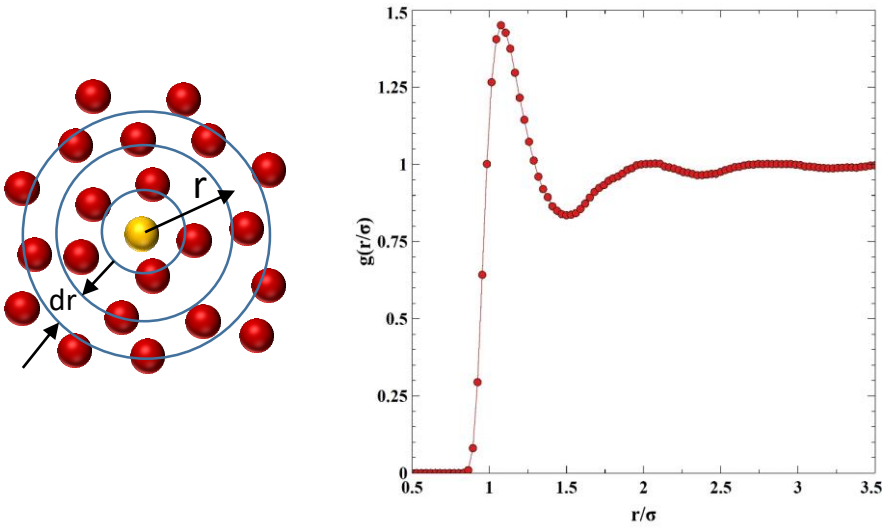


Figure 2.17: Left: schematic representation of the spatially divided system into shell volumes used for the calculation of the radial distribution function (RDF). Right: typical response of the RDF in the liquid phase.

2.3.5.4 Statistical uncertainties in molecular simulations

The properties deduced from molecular simulation technique are all subject to statistical errors. To estimate these errors, the block average method [1] is used to avoid estimations based on correlated data. The method consists in splitting the total simulation run into N_b blocks (minimum of 6 is suggested), then the standard error can be estimated using the following formula:

$$Err\langle X\rangle = \frac{std(X)}{\sqrt{N_b}} \tag{2.130}$$

where

$std(X)$ is the standard deviation of the blocks given by:

$$std(X) = \sqrt{\frac{1}{N_b}\sum_{i=1}^{N_b}(x_i - \langle X\rangle)^2} \tag{2.131}$$

where x_i are the averages of each block, and $\langle X\rangle$ is the mean of all the blocks given by:

$$\langle X\rangle = \frac{1}{N_b}\sum_{i=1}^{N_b} x_i \tag{2.132}$$

The standard error is often given in the confidence interval of 95%, and thus, in the limit where N_b is enough large, the standard error becomes:

$$Err\langle X\rangle = Z\frac{std(X)}{\sqrt{N_b}} \tag{2.133}$$

where Z~1.96 following the normal distribution. If N_b is <~200, a Student should be used instead to estimate Z.

2.4 References

1. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids: Second Edition*. (Oxford University Press, 2017).
2. Frenkel, D. & Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*. (Elsevier, 2001).
3. Hansen, J.-P. & McDonald, I. R. *Theory of Simple Liquids*. (Elsevier, 2006).
4. Sandler, S. I. *An Introduction to Applied Statistical Thermodynamics*. (John Wiley & Sons, 2010).
5. Raabe, G. *Molecular Simulation Studies on Thermophysical Properties*. (Springer, 2017).
6. Ungerer, P., Tavitian, B. & Boutin, A. *Applications of Molecular Simulation in the Oil and Gas Industry: Monte Carlo Methods*. (Editions TECHNIP, 2005).
7. Leach, A. R. & AR, L. *Molecular Modelling: Principles and Applications*. (Pearson Education, 2001).
8. McQuarrie, D. A. *Statistical Mechanics*. (University Science Books, 2000).
9. Rapaport, D. C. *The Art of Molecular Dynamics Simulation*. (Cambridge University Press, 2004).
10. Vlugt, T., Eerden, J. V., Dijkstra, M., Smit, B. & Frenkel, D. Introduction to Molecular Simulation and Statistical Thermodynamics. in.
11. Lagache, M., Ungerer, P., Boutin, A. & Fuchs, A. H. Prediction of thermodynamic derivative properties of fluids by Monte Carlo simulation. *Phys. Chem. Chem. Phys.* **3**, 4333–4339 (2001).
12. Verlet, L. Computer 'Experiments' on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **159**, 98–103 (1967).
13. R W Hockney, J W Eastwood. *Computer Simulation Using Particles*. (CRC Press, 2021).
14. Swope, W. C., Andersen, H. C., Berens, P. H. & Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *Journal of Chemical Physics* **76**, 637–649 (1982).
15. Beeman, D. Some Multistep Methods for Use in Molecular Dynamics Calculations. *Journal of Computational Physics* **20**, 130–139 (1976).
16. Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics* **52**, 24–34 (1983).
17. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *Journal of Computational Physics* **23**, 327–341 (1977).
18. Woodcock, L. V. Isothermal molecular dynamics calculations for liquid salts. *Chemical Physics Letters* **10**, 257–261 (1971).
19. Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **72**, 2384–2393 (1980).
20. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

21. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
22. Schneider, T. & Stoll, E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B* **17**, 1302–1322 (1978).
23. Evans, D. J. & Morriss, G. P. Isothermal-isobaric molecular dynamics. *Chemical Physics* **77**, 63–66 (1983).
24. Hoover, W. G. Constant-pressure equations of motion. *Phys. Rev. A* **34**, 2499–2500 (1986).
25. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177–4189 (1994).
26. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182–7190 (1981).
27. Kolb, A. & Dünweg, B. Optimized constant pressure stochastic dynamics. *J. Chem. Phys.* **111**, 4453–4459 (1999).
28. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **21**, 1087–1092 (1953).
29. Siepmann, J. I. & Frenkel, D. Configurational bias Monte Carlo: a new sampling scheme for flexible chains. *Molecular Physics* **75**, 59–70 (1992).
30. Mooij, G. C. A. M., Frenkel, D. & Smit, B. Direct simulation of phase equilibria of chain molecules. *J. Phys.: Condens. Matter* **4**, L255 (1992).
31. Smit, B., Karaborni, S. & Siepmann, J. I. Computer simulations of vapor–liquid phase equilibria of n-alkanes. *J. Chem. Phys.* **102**, 2126–2140 (1995).
32. Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. & Tildesley, D. J. Phase equilibria by simulation in the Gibbs ensemble: alternative derivation, generalization and application to mixture and membrane equilibria. *Molecular Physics* **63**, 527–545 (1988).
33. VERLET, L. Computer 'Experiments' on Classical Fluids. II. Equilibrium Correlation Functions. *Phys. Rev.* **165**, 201–214 (1968).
34. Zhang, Y., Otani, A. & Maginn, E. J. Reliable Viscosity Calculation from Equilibrium Molecular Dynamics Simulations: A Time Decomposition Method. *J. Chem. Theory Comput.* **11**, 3537–3546 (2015).
35. Maginn, E. J., Messerly, R. A., Carlson, D. J., Roe, D. R. & Elliot, J. R. Best Practices for Computing Transport Properties 1. Self-Diffusivity and Viscosity from Equilibrium Molecular Dynamics [Article v1.0]. *Living Journal of Computational Molecular Science* **1**, 6324–6324 (2019).
36. Bordat, P. & Müller-Plathe, F. The shear viscosity of molecular fluids: A calculation by reverse nonequilibrium molecular dynamics. *Journal of Chemical Physics, v.116, 3362-3369 (2002)* **116**, (2002).
37. Delage Santacreu, S., Hoang, H., Khennache, S. & Galliero, G. Thermodynamic Scaling of the Shear Viscosity of Lennard-Jones Chains of Variable Rigidity. *Liquids* **1**, 96–108 (2021).

Chapter 3. Development of an LJC coarse grained molecular model

Contents

3.1 Molecular models: up-scaling 47
3.2 Parametrization strategy: Bottom-up vs Top down 49
3.2.1 Bottom-up approach 49
3.2.2 Top-down approach 51
3.3 An overview of the predictive capabilities of some up-scaled force fields 53
3.3.1 United atoms 53
3.3.2 Coarse grained 55
3.3.3 SAFT based Coarse grained 56
3.4 Towards an improved coarse grained model 57
3.4.1 Semi Rigid Lennard Jones Chain Coarse Grained model 58
3.4.2 Simulations details 60
3.4.3 Rigidity effect on phase equilibria and transport properties of the fluid model: The Semi-Rigid LJCCG 64
3.4.4 Choice of the parametrization strategy 66
3.4.5 Optimization procedure 68
3.4.6 Application to pure real fluids 70
3.5 Summary 84
3.6 References 85

3.1 Molecular models: up-scaling

Computer simulations can be conducted at different spatiotemporal scales, depending on the level of details used to describe the molecular model. At the smallest scale, the molecule is described by its electrons and protons, and the properties of the system are obtained by solving equations such as the Schrodinger or the Dirac wave functions. This quantum, or ab-initio, level of detail allows for the determination of the "exact" properties of the matter at very small time and length scales, which is necessary for example necessary to understand many phenomena involved in chemical reactions. However, only small systems can be considered, as solving the many-body wave functions increases exponentially in terms of computational costs with the number of simulated particles. Consequently, accessing larger time and length scales requires some simplifications.

Up-scaling leads to an atomistic, i.e. classic, description of the molecule. At this level, electron motion is no longer described, but each atom is represented by an effective center of force. This model is known as the All Atom model (AA) and will sometimes be referred here as the high resolution model. The interactions between the atoms of the system are described by semi-empirical force fields, which are a set of mathematical equations (pair intermolecular interaction potentials and intramolecular interaction potentials, as described in the previous chapter) whose forms are generally defined to reproduce the effective potential interactions of quantum models. The parameters of the force fields are either obtained using a bottom-up approach by fitting these equations to ab-initio calculations, or using a top-down approach by fitting macroscopic experimental properties of the matter. At this level, the matter description is close to that of real molecules, and the model is then expected to give quasi-experimental thermodynamic, transport, interfacial and structural properties of the matter using molecular dynamics simulations or Monte Carlo simulations. Examples of models in this category includes the COMPASS model [1], the CHARMM model [2], the AMBER model [3], the TraPPE-EH model [4] and the OPLS-AA model [5].

The AA model may appear to be the natural choice for accurate predictions of fluid properties as it uses a representation very close to that of real molecules. However, it is very demanding in CPU time when dealing with long molecules and large systems (such as self-assembly and micro-phase separation structures). For example, even a relatively small molecule like normal decane contains 32 (22H + 10C) center forces and so would require the evaluation of 1024 intermolecular distances between tow molecules, not counting those coming from within the same molecule. Therefore, considering a large system makes the task very challenging in terms of computation time, knowing that the CPU time evolves with the square of the number of force centres in the system. Another limitation of the AA model is it parameterization. It requires a high number of parameters and requires an exact characterization of the real molecule to be modelled, which can be difficult in some cases, such as for asphaltenes. To overcome these limitations, further simplifications can be proposed.

At the next level of the computer simulations scale, instead of using a high resolution definition of the molecule, a center force can be represented by chemical groups. This model is known as the United Atom model (UA). In this model, hydrogens are not considered explicitly, but are

taken into account in the pseudo-atoms such as CH, CH2 and CH3 chemical groups. This reduces the number of internal degrees of freedom that define the molecule and, consequently, the number of parameters defining the force field are reduced as well, leading to a simpler parameterization compared to the AA model. For example, using a UA model, the decane molecule is defined by only 10 center forces (2CH3 + 8CH2) and the interactions between two molecules would require the evaluation of 100 intermolecular distances instead of 1024 for the AA model. As mentioned previously, the CPU time scales roughly with the square of the number of center forces in the system. Hence, it is expected to be reduced by 1/10th of that required for an AA model. Consequently, UA models are widely used in the literature for many applications as they make feasible simulations of large time and length scales. Among the most popular UA models, there are the TraPPE-UA [6], the OPLS-UA [7], the NERD [8] and the anisotropic ones such as that of Toxvaerd [9] and the AUA [10].

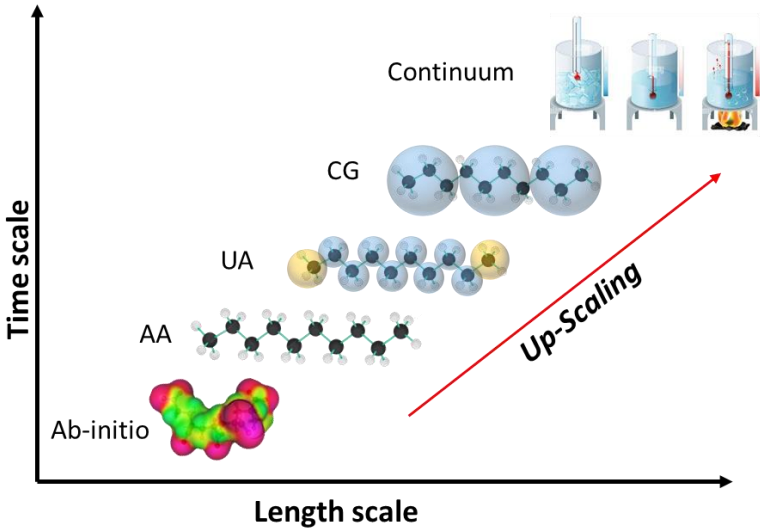


Figure 3.1: Schematic representation of the up-scaling approach in molecular modelling. From Ab-initio scale to the continuum scale. AA stands for All Atom model, UA stands for United Atom model and CG stands for Coarse grained model.

Despite the success of the UA model in accurately representing the properties of materials in a reasonable amount of computational time, more effort is needed to tackle problems at the mesoscale level, such as the study of dynamics of proteins and the behaviour of polymer molecules, among others. Even for less complex molecules such as n-alkanes, in terms of engineering requirements, it is always desirable to obtain data in a short time with fewer computer resources. Thus, in order to access to large time and length scales and to make molecular simulations compatible with engineering requirements by speeding up the simulations, it is necessary to scale-up the model by relaxing further internal degrees of freedom. A coarse grained (CG) model is then an appealing option, where a pseudo atom may contain two, three or four CH_2 or CH_3 groups, leading to a reduction in the dimensionality of the system. For example, if a decane molecule is modelled using a coarse grained approach, the whole decane molecule could be described by a three center forces model. The total intermolecular distance to evaluate for two molecules of decane is then reduced to 9 with the

coarse grained model! As a consequence, the CPU time is reduced by 1/10th of the amount need with a UA model and by 1/100th of the amount required for an AA model.

Compared to the AA or UA models, the energy landscape of the CG model becomes smoother due to the loss of many fluctuations caused by molecular vibrations. This makes it possible to perform simulations with a larger time step, thus better sample the phase space. Larger and more complex systems [11,12] can then be explored in a much more reasonable computational time, while still keeping most of the physics well described. Moreover, the number of parameters of the model is much lower than the AA and UA model, making their determination much easier. Another interesting feature of some of the CG model lies in their compatibility with the Statistical Association Fluid Theory (SAFT), making the transferability of the parameters of the force field between the two approaches possible. Additional advantage is that the equilibrium properties of the model could be obtained in seconds with the use of SAFT equations of state, instead of performing extensive molecular simulations. Such a CG representation is also widely used in the literature, particularly with the development of SAFT models. One may cite the works of Klein [11], Nielsen [13], Voth [14], Reith [15], Moore [16], Shell [17], Hoang [18], MARTINI [19], and those from SAFT equations of states such as the soft-SAFT [20] and SAFT-γ-Mie [21].

However, because of the lack of many internal degrees of freedom, the structural properties and dynamical properties are often only approximately predicted using CG models. This is a problem that we aim to address in this work by, first, restoring some internal degrees to the molecules and second, by proposing a new robust parametrization strategy for the determination of the parameters of the force field.

3.2 Parametrization strategy: Bottom-up vs Top down

Two main approaches are available in the literature for determining the parameters of a CG model. The first is the bottom-up approach [22], where, a priori, the number of atoms assigned to a single bead (sphere, monomer or segment) is determined and the parameters of the CG model are determined from the fine grained (FG) model. The second is the top-down approach, which is widely used to develop CG SAFT based models. It relies on constraining the parameters of the model to reproduce macroscopic properties of fluids that are experimentally measurable.

3.2.1 Bottom-up approach

The most widely used approach in a bottom-up approach is probably the iterative Boltzmann Inversion Method (BIM). In its original form, the non-bonded parameters of the low resolution model are updated until the radial distribution function (RDF) of the low and high resolution models, at a specific state point, match [15, 23] according to the following scheme:

$$U_{CG}^{n+1}(r) = U_{CG}^{n}(r) - k_B T ln \frac{g_{AA}(r)}{g_{CG}(r)} \tag{3.1}$$

The bonded parameters are determined in a non-iterative way using a direct BIM by inverting an analytical distribution function for bonds, angles, dihedrals [24]. This allows the model to

reproduce short ranged properties, as the local structure of fluids is more sensitive to the repulsive part of the potential [25]. However, it is less sensitive to the properties affected by long range interactions. As a result, the resulting CG model may fail to reproduce properties such as pressure or isothermal compressibility [15, 26], which are heavily influenced by the long range contributions. To address this issue, an additional post-optimisation process is done by adding pressure or Kirkwood-Buff integrals (KBI) as a target in the merit function in addition to the RDF [15, 26, 27, 28]. Actually, it is the Potential of the Mean Force (PMF) which is updated to meet the target function (RDF) and not the pair particles interaction potential. This means that the optimised parameters of the mesoscale model are state dependent and should, therefore, not be used at another state (weak transferability). To overcome this limitations, using a multistate iterative BIM [16] scheme can significantly reduce the dependence of the derived potentials on the thermodynamic state used in the optimisation, compared to the original version.

Moreover, the choice of the mapping scheme is not an easy task [29] as the structural fluid properties are highly sensitive to it. The choice of the FG model to fit is also crucial. In other words, the quality of the CG developed using this approach is strongly dependent on the quality of the high resolution model (such as an AA or UA force field), as some of them may exhibit inadequate behaviours at some conditions, such as what was observed for normal dodecane (nC_{12}) using a version of the OPLS-UA FF [30, 31]. These problems arise due to the parametrisation strategy used to fit the intermolecular parameters, where, in the case of the OPLS-UA, they are fitted on liquid densities and the heat of vaporisation at atmospheric pressure.

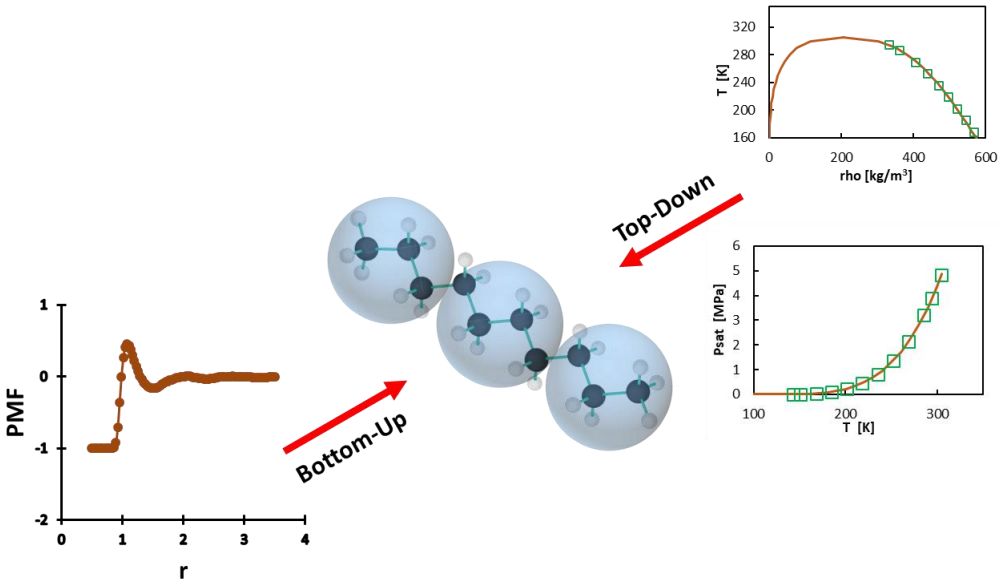


Figure 3.2: Schematic representation of the Bottom-up and Top-down parameterization approaches.

Alternatives to the IBM are available in the literature, such as the Force Matching (FM) method [32] proposed by Voth and co-workers, or the elegant one based on the relative entropy concept [33, 34] proposed by Shell and co-workers. The latter method measures the loss of information when moving from a high resolution model to a low resolution one. In other words, it measures the change in the configuration-space probabilities between the low resolution model and that

of the mapped coarse grained model. In the FM method, instead of looking at the structure of the fluid as the BIM does, it focuses on the forces and trajectories. The intermolecular interactions parameters are then obtained by minimising the difference between the forces acting on the coarse grained model and that of the atomistic model (or between the atomistic model and that from ab-initio). Unlike the iterative BIM, the FM method is non-iterative and is based on a least-square minimisation procedure.

3.2.2 Top-down approach

A top-down approach consists of parametrizing the force field by constraining the molecular parameters to reproduce experimental target properties. In practice, such an approach is only used to obtain the parameters of the intermolecular potentials of fully flexible models. This is the case of most models based on the Statistical Association Fluid Theory approach [35] developed by Chapman and co-workers. Among them, the one may cite the original-SAFT [35], the soft-SAFT [36], the SAFT-VR-Mie [37–39] and the PC-SAFT [40]. These equations of state describe a fluid consisting of freely and tangentially bonded segments interacting through different type of intermolecular potentials. Since all common SAFT models rely on the Thermodynamic Perturbation Theory (TPT) of Wertheim truncated at the first level [41, 42], they do not include any information regarding the intramolecular interactions and structure. This is a problem that will be addressed in the next chapter of this thesis. The intermolecular parameters can be fitted on liquid density, saturation pressure, heat of vaporization, heat capacity... The optimisation is done upon this following objective function:

$$min\, F(\theta) = \sum_{i=1}^{N^{prop}} \frac{w_{N^{prop}}}{N_{X^{exp}}} \sum_{n=1}^{N_{X^{exp}}} \left[\frac{X^{exp} - X^{SAFT}(\theta)}{X^{exp}}\right]^2 \quad (3.2)$$

Where N^{prop} is the number of different properties involved in the fit, $N_{X^{exp}}$ is the number of experimental data points per property used in the fit, $w_{N^{prop}}$ is the property associated weight, X^{exp} is the experimental property to optimise on, $X^{SAFT}(\theta)$ is the calculated property from the SAFT model, and finally θ is the set of the intermolecular parameters that defines the intermolecular potential used in the SAFT version model. Such models are known to provide accurate thermodynamic properties. However, the lack of most internal degrees of freedom of the molecule in the model, makes them limited for the prediction of the structural and transport properties as it will be shown latter.

The parametrization of a semi-rigid model is usually done by coupling two approaches, where the bottom-up approach serves for the determination of the parameters of the intramolecular interaction potentials such as stretching, bending or torsional potentials, and the top-down approach for the estimation of the parameters of the intermolecular interaction potentials. This approach is widely used to parametrize the united atom or coarse grained force field models, such as the TraPPE-UA and the TraPPE-CG developed by Siepmann et al. [6, 43], the AUA developed by Ungerer et al. [44], the OPLS-UA and OPLS-AA developed by Jorgensen et al. [5, 7, 45], that of Nielsen et al. [46], the MARTINI developed by Marrink et al. [19, 47]. The non-bonded intermolecular parameters are constrained to fit certain macroscopic measurable experimental data, while the bonded interaction parameters are usually fitted to ab-initio data.

The choice of the properties used to adjust the intermolecular parameters is often dependent on the application for which the force field is designed. In the case of LJ-based force fields such as the AUA(4) [48] model, it uses vapour pressure, liquid density and heat of vaporization as target properties. The TraPPE-UA [6] uses the critical temperature and the saturated liquid density and, in a recent version, a saturation pressure was included by Janecek and Paricaud [49]. The MARTINI force field, which is probably the most widely used one for studying biomolecular systems, is fitted to the water-oil partitioning free energy ($\Delta G_{Oil/water}$) for 18 CG selected particles [47]:

$$\Delta G_{Oil/water} \approx k_B T ln \frac{\rho_{oil}}{\rho_{water}} \tag{3.3}$$

The parameters are estimated in a similar way as for the coarse grained SAFT based models by minimising the objective function iteratively. In both approaches, the strategy relies on including a large amount of experimental data covering wide range of thermodynamics conditions in the objective function. Indeed, the larger the amount of data used, the better the transferability of the force field, but this is subject to some limitations, for example:

- it often involves a large amount of experimental data to cover a wide range of thermodynamics states that are not always available;
- many solutions can exist because of the presence of many local minima, due to the use of many data and different types of properties, making the convergence to the global minima challenging. The local minima make the parameters estimated physically less consistent;
- the critical points are usually overestimated, leading to poor prediction of supercritical properties;
- the transport properties predicted with this approach are, most of the time, inaccurate.

Alternatively, another strategy based on the top-down approach, called Corresponding State strategy (CS), mainly inspired from Cubic Equations of States (C-EoS) parametrization, is available. It was first applied by Mejia et al. [39] to parameterize a coarse grained model based on the SAFT-γ-Mie EoS, and later by Hoang et al. [18] to parametrize the Mie chain CG model (MCCG) and by Moine et al. to parametrize the I-PC-SAFT EoS. The CS strategy presents excellent features, among them, the number of experimental data required is the same as the number of the parameters held by the model. This reduces considerably the amount of macroscopic experimental data required for the parameters estimation when compared to the classical top-down strategy. Moreover, to ensure that the force field parameters are physically consistent, each parameter is independently obtained from one specific macroscopic property chosen from a dimensional analysis. For instance, the intermolecular energetic interaction parameter is constrained by a purely energetic macroscopic property, such as the critical temperature. The segment size diameter is constrained by a purely volumetric property, such as density. Other properties can also be considered, such as the critical pressure used by Moine et al. [50] or a transport property like viscosity considered by Hoang et al.[18]. The inclusion of this latter property in the optimisation showed extremely interesting results. First, if only equilibrium properties are used for the parametrization may lead to an indeterminacy in the set of parameters that minimises the objective function. The inclusion of a transport property, such

as viscosity, ensures the uniqueness and the consistency of the obtained set of parameters [18]. Second, the prediction of the liquid viscosity is strongly improved compared to force fields parametrized only on equilibrium properties, as shown in the works of Messerly et al. [51] and Papavasileiou et al. [30]. However, a systematic underestimation of the liquid viscosity at low temperature is observed when considering large chain molecules, which is attributed to the fully flexibility nature of the model. A problem that we are going to address in this work.

3.3 An overview of the predictive capabilities of some up-scaled force fields

3.3.1 United atoms

The first modern force field using extensive adjustment of the molecular parameters is the OPLS-UA (optimized potentials for liquid simulations) developed by Jorgensen et al. [7], where the intermolecular parameters were adjusted to reproduce liquid density and heat of vaporization of organic compounds at ambient pressure and temperature. The intramolecular parameters were adjusted on quantum mechanical calculations. The model was found relatively accurate in reproducing the experimental equilibrium properties of small molecules.

Few years later, based on the work of Ryckaert and Bellemans [52], who proposed the first version of united atom model, Toxvaerd [53] showed that the models based on the UA did not scale very well with respect to temperature and density because they did not properly account for the effect of hydrogens. He then proposed the first version of the so-called Anisotropic UA model (AUA(1)) which can be considered as a more realistic model of n-alkanes. Better agreement was obtained for the normal propane, pentane and decane molecules.

A year later, Padilla and Toxvaerd [9] showed that the model needed to be readjusted to better fit the structural and transport properties. They reparametrized the parameters of the torsional potential of the AUA(1) to better fit the self-diffusion coefficient. This was probably the first time where a transport property was included as a target property in the fitting procedure. Moreover, that work highlighted that transport properties are very sensitive to the intramolecular conformations. A third generation of the anisotropic was latter proposed by Toxvaerd [54] by refining the molecular parameters to high pressure data of pentane and decane in the range of [298K-673K].

In the meantime, Smit et al. [55, 56] noted that the OPLS-UA model only reproduces equilibrium properties correctly for short molecules. For instance, the critical temperature was found to be highly overestimated for molecules with more than 5 carbons for the normal alkanes. The authors proposed a new model often referred to as the SKS model (stands for the first letter of the authors: Smit, Karborni and Siepmann) addressing the issue of the critical temperature observed on the n-alkanes when using the OPLS-UA model. It was later refined and extended to branched alkanes by Siepmann et al. [57] and referred to as SMMK. A year later, Martin and Siepmann [6] published the final version of the SKS model referred to as ‘transferable potentials for phase equilibria’ (TraPPE) force field. The parameters were re-adjusted to yield correct critical temperature and coexistence liquid curve densities of ethane and n-octane. It was then tested on a series of n-alkanes up to nC_{48}, and the model was found to give better prediction

compared to all the previous force fields cited until now. At the same time, Nath et al. [8] proposed a slightly modified version of the TraPPE force field denoted as the NERD force field. It was fitted to orthobaric densities and second virial coefficient of short and long n-alkanes, and was found to yield globally better prediction of the equilibrium properties, particularly, for the saturation pressure, but still exceeded experimental measurements.

In 2000, Dysthe et al. [58] conducted the first extensive study on the predictive capability of the force fields cited until now (except for the NERD) regarding transport properties. The study showed that none of the force fields were capable of predicting the correct liquid viscosity which was underestimated, or the self-diffusion coefficient, which was overestimated, of the n-alkanes. The study showed that the mean deviations over all states, considering both viscosity and self-diffusion for both nC_4 and nC_{10} are: AUA(2) 24%, SKS 26%, AUA(3) 30%, SKS/2 42%, OPLS 48%, SMMK 54% and SMMK/2 99%. Unsurprisingly, it is the AUA(2) that yields the least worst predictions due to the fact that in the Padilla and Toxvaerd work, self-diffusion was included in the fit. The SKS/2 and the SMMK/2 are the same as the SKS and SMMK, except that their torsional potential (taken from the OPLS-UA) has been dived by factor of two. Therefore, the deviations are increased almost by a factor of two also, highlighting the high sensitivity of the transport properties on the molecular internal degrees of freedom.

At the same time, Errington and Panagiotopoulos [59] proposed a UA force field based on the Buckingham potential which performs better than the TraPPE or the NERD models regarding the equilibrium properties (on which it was fitted), and Ungerer et al. [44] proposed a fourth generation of the AUA(4) force field, where the parameters are fitted on saturation pressure, liquid density and heat of vaporization of three n-alkanes: ethane, pentane and dodecane. In this latter work, an approach based on the gradient method was proposed for the minimization of the objective function. The model has been compared to the AUA(3) over many properties and molecules as well as for the self-diffusion coefficient and CO_2-n-alkanes mixture. It was found that in all cases, that the AUA(4) yields better agreement with the experiments. Zhang and Ely [60] proposed an AUA model based on the Buckingham potential similarly to that of Errington and Panagiotopoulos. However, the parametrization was done on the viscosity. The model yields very good prediction on the viscosity mixture of alkanes-alcohol, but unfortunately, the authors did not report any data on its performance on the equilibrium properties.

To improve the prediction of the transport properties with the AUA(4), Nieto-Draghi et al. [10] have proposed to refit the torsional potential to reproduce the liquid viscosity of the n-heptane and the internal CH-group vector reorientation dynamics reported by the NMR $C^{13}T1$ spin lattice experiments. The modified model is denoted as AUA(4m) and yield better viscosity and self-diffusion properties compared to the previous one.

Other models based on the AUA representation have been proposed later by Potoff and Bernard-Brunel [61] denoted as Potoff force field, and by Hemmen and Gross [62] denoted as TAMie force field. The two models are based on the generalized LJ-potential (λ-6), where in the case of n-alkanes, λ is fixed to 16 for the Potoff and to 14 for the TAMie. The two models have been regressed on the equilibrium properties and saturation pressure of n-alkanes, and results have shown that while the Potoff is found better for the coexistence densities, the TAMie is found better on the saturation pressure prediction. The united atom TraPPE force field, has

recently be reparametrized to improve the vapour pressure prediction. This was done in the work of Van Westen et al. [63] who proposed a parametrization based on the PC-SAFT EoS, where the fitting has been made on liquid density, saturation pressure and heat of vaporization. The same target properties were also considered by Janecek and Paricaud [49] who proposed a re-parametrization using a strategy of minimization of the objective function based on the gradient method proposed by Bourasseau et al. [48]. More recently, Shah and Siepmann [64] proposed an improved version, called TraPPE-UA2, for the ethane and ethylene. Messerly et al. [51] studied the predictive capability on the liquid viscosity of some united atom force fields including the TraPPE–UA, the AUA(4m), the Pottof-AUA and the TAMie-AUA. The Pottof and the TAMie were found giving better predictions compared to the two LJ based models, indicating the superiority of the of Mie-6 potential. The Pottof exhibited some advantages at saturation conditions, while the TAMie was found to give a correct trend at high pressure. Globally for both models, results were found to degrade for larger molecules at low temperature and at high density.

3.3.2 Coarse grained

A large number of coarse-grained force fields have been proposed in the literature. Most of them rely on iterative molecular simulations for the optimization of the force field. The advantage of these force fields is that they are designed to provide some specific properties. However, few of them are dedicated to perform well on the thermophysical properties of simple fluids, as most of them are designed to study properties of biomolecules (proteins, lipids…).

Among those which consider thermophysical properties as target properties, one may cite the work of Shelley et al. [65]. In that work, for the n-alkanes, the direct BIM was used to obtain the intramolecular parameters, while the intermolecular parameters interaction of the LJ(9-6) potential were obtained by fitting them to reproduce the liquid density and vapour pressure of the n-dodecane. Nielsen et al. [46] proposed a new parametrization of the Shelley's work. While the same strategy was used for the intramolecular parameters by fitting them to structural properties of a CHARMM all atom force field, the intermolecular interaction potentials were adjusted to reproduce bulk density and surface tension of alkanes at 1 bar and T=303K. The model provided roughly good structural and surface tension properties at the same temperature of fitting. However, the temperature dependence of the density was not accurately captured and the self-diffusion coefficient prediction was only qualitative. Later, Shinoda et al.[66] proposed a multi-properties fitting. The authors considered the bulk density, surface/interfacial tension, structural and compressibility of water and free energy. Even though they consider thermodynamic properties as target properties, this force field was not widely used for the prediction of thermophysical properties of fluids. Thus, its predictive capability on the viscosity is unknown, but, one may expect some deviations as the self-diffusion is predicted only qualitatively for the n-alkanes [46].

The MARTINI force field[19, 47], which is widely used to study biomolecules such as proteins, lipid, amino acids…has been developed using a direct Boltzmann inversion method to reproduce the structural properties of an all atom model. The intermolecular parameters of the LJ(12-6) potential have been obtained by fitting them to the oil/water partitioning coefficients,

the free energy of hydration and the free energy of vaporization. Because no information on the vapour phase is included in the parametrization, the model is expected to give poor predictions of the vapour density and saturation pressure. Papavasileiou [30] showed that the MARTINI overestimates liquid densities, particularly in the case of n-C_{12}.

An et al. [67] proposed a coarse grained force field for hydrocarbons, where all the parameters (inter and intra) were optimized on experimental values of density, enthalpy of vaporization, self-diffusion coefficient, and surface tension of normal decane at 300 K and refined to map the trajectories of the CHARMM all atom force field. The model showed good transferability when tested at other temperatures and gives satisfactory results on all optimized properties, except self-diffusion which exhibits noticeable deviations increasing with the carbon number.

A coarse grained version of the TraPPE force field was developed by Maerzke and Siepmann [43], named as TraPPE-CG. The model was parametrized using the multistate iterative Boltzmann method at three distinct temperatures (300K, 400K and 500K). The model showed very good reproduction of the equilibrium and structural properties of the TraPPE-UA which was used as the reference high resolution model to fit upon. Knowing the relatively poor prediction of the TraPPE-UA for vapour density and saturation pressure, one may expect that the comparison of the TraPPE-CG to experiments will exhibit the same weakness. The transport properties predictions as well are expected to be relatively poor, even if not evaluated.

3.3.3 SAFT based Coarse grained

In the last decade of the 20th century, another type of force fields emerged, of the coarse grained type, that relied on the perturbation theory of Wertheim [41], and more precisely, on the SAFT type equations of state developed by Chapman [68]. These included the SAFT-HS developed by Huang and Radosz [69], the LJ-SAFT developed by Johnson et al. [70], the soft-SAFT developed by Blas and Vega [71], the SAFT-VR developed by Gil-Villegas et al. [72], the PC-SAFTD developed by Gross and Sadowski [40] and the SAFT-VR-Mie developed by Lafitte et al. [37]. These models became increasingly popular for a variety of applications due to their compatibility with molecular simulations, which allows for the transfer of force field parameters from the EoS to molecular simulations. Additionally, these models are able to produce results in seconds, whereas molecular simulations can require hours or days. However, molecular simulations also allow for the computation of other type of properties, such as the transport, interfacial and structural properties which not directly accessible form the EoS.

A good example of the transferability of the force field from EoS to molecular simulations can be found in the series of papers of the Jackson, Müller and co-workers using SAFT-γ-Mie group contribution EoS. They systematically transferred the parameters and performed molecular simulations for a variety of systems. Example of these articles include the work of Avendano et al. [73, 74], Lafitte et al. [75], Lobanova et al. [76, 77] and Rahman et al. [21]. These works showed that the thermodynamic properties predicted by either the equation of state or the molecular simulation were in very good agreement with experiments. The work of Rahman et al. is of particular interest, as it demonstrates for the first time that SAFT based force field approach are subject to a great weakness when considering transport properties. This is because all SAFT models are based on the TPT1 of Wertheim which do not take into account the internal degrees

of freedom of a molecule, and thus, the SAFT force fields are fully flexible. To alleviate this weakness, Rahman et al. [21] proposed adding some internal degrees of freedom (bonding and bending potentials) by mapping their force field to that of the TraPPE united atom and then obtained the parameters using a multistate direct Boltzmann inversion method. The results show very good predictions for viscosity and self-diffusion coefficient.

3.4 Towards an improved coarse grained model

The poor prediction of the viscosity with coarse grained models is particularly highlighted in the work of Galliéro [78]. The author used a fully flexible Lennard-Jones chain coarse grained (LJCCG) molecular model, where the molecular parameters were optimized on the liquid density and saturation pressure. The results showed that while accurate equilibrium properties could be obtained, less accurate surface tension was obtained and very poor prediction were observed for the transport properties such as viscosity and thermal conductivity. Galliéro pointed out that the lack in the internal degrees of freedom is responsible for the deterioration of the transport properties prediction when considering large alkanes. The author performed few simulations by adding a bending potential, and showed that it is possible to remedy the weakness of the viscosity prediction if an additional fourth parameter, molecular "rigidity", is included to the model while maintaining the same accuracy on equilibrium properties. Later, Hoang et al. [18] proposed a Mie chain coarse grained (MCCG) force field, similar to that of Mejia et al. [39], but including viscosity as a target property in addition to equilibrium properties, as viscosity in dense medium is very sensitive to the repulsive contribution of the interaction potential [79]. The results obtained by the model named MCCG were excellent for equilibrium properties, interfacial and transport properties, probably better than most existing force fields including fine grained models [6]. However, systematic underestimation of the liquid viscosity at low temperature for large alkanes was noted, attributed to the fully flexibility nature of the force field.

This overview showed that there is a lack of force fields that are capable to provide simultaneously accurate thermodynamic and transport properties of simple fluids such as normal alkanes. Most of the force field are designed to reproduce only equilibrium properties or structural properties, and only a few are capable of reproducing accurate transport properties. While accurate equilibrium properties are driven by a robust parametrization of the intermolecular potential interaction, accurate transport properties also require a robust parametrization of the intramolecular potential interactions. Indeed, the work of Nieto-Draghi [10] for the AUA(4m) and of Galliéro [78] for the LJCCG, highlighted that if these intramolecular potential are parametrized on a transport property as well, significant improvement could be achieved while keeping the equilibrium properties accurately predicted.

Thus, the development of a coarse grained force field that is capable of predicting both transport and equilibrium properties simultaneously with the same accuracy is still an open issue. If, in addition, the force field is also able to predict structural properties and interfacial properties, this would be of great interest as well. This is the aim of the proposed option described in this chapter.

3.4.1 Semi Rigid Lennard Jones Chain Coarse Grained model

The coarse grained molecular model chosen in this work consists in spherical beads, tangentially connected to form chain molecules, which interacts through a Lennard-Jones 12-6 potential:

$$U_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3.4}$$

where ε_{ij} is the potential well depth, σ_{ij} is the bead diameter and r_{ij} is the atom-atom distance.

This model is called the Fully-Flexible LJCCG (FF-LJCCG). It involves two molecular parameters (ε_{ij} ; σ_{ij}) and the number of segments (m) forming the chain molecule. This fully-flexible-LJCCG model has been widely studied in the literature using direct molecular simulations [15, 80–83], or through the use of the Statistical Association Fluid Theory (SAFT) [70, 71]. It showed very good capabilities to describe the equilibrium properties of real fluids and their mixtures over a wide range of thermodynamics conditions. However, this simple FF-LJCCG has some difficulties in predicting transport properties such as viscosity [78] at low temperature and as the chain length becomes longer. This weakness is mostly due to its fully flexible nature and so the bad representation of the internal degrees of freedom.

A question that may be raised is, instead of an additional molecular parameter, what would be the predictive capability of this fully-flexible model if a transport property is included as a target property to reproduce? To answer this question, there are three distinct options that could be evaluated. The first consists in parametrizing the intermolecular parameters on liquid density and saturation pressure only, the second option includes saturation liquid viscosity to the objective function, and the last option consists of using only saturated liquid viscosity as the target property. The results for the n-decane molecule are shown in Figure 3.3.

The results of the first option is already known from the work of Galliéro [78], but it was added for comparison. Results of the second and third options, are more interesting to analyze. In the case where parametrization is done upon both type of properties (equilibrium + viscosity), results show that a small improvement could be obtained on liquid viscosity. However, strong deviations are obtained on saturation pressure and critical temperature. In the third case, where the viscosity is set as the only target property, results show that a slight improvement on viscosity is obtained compared to the latter case, but, strong deviations are obtained on equilibrium densities and saturation pressure.

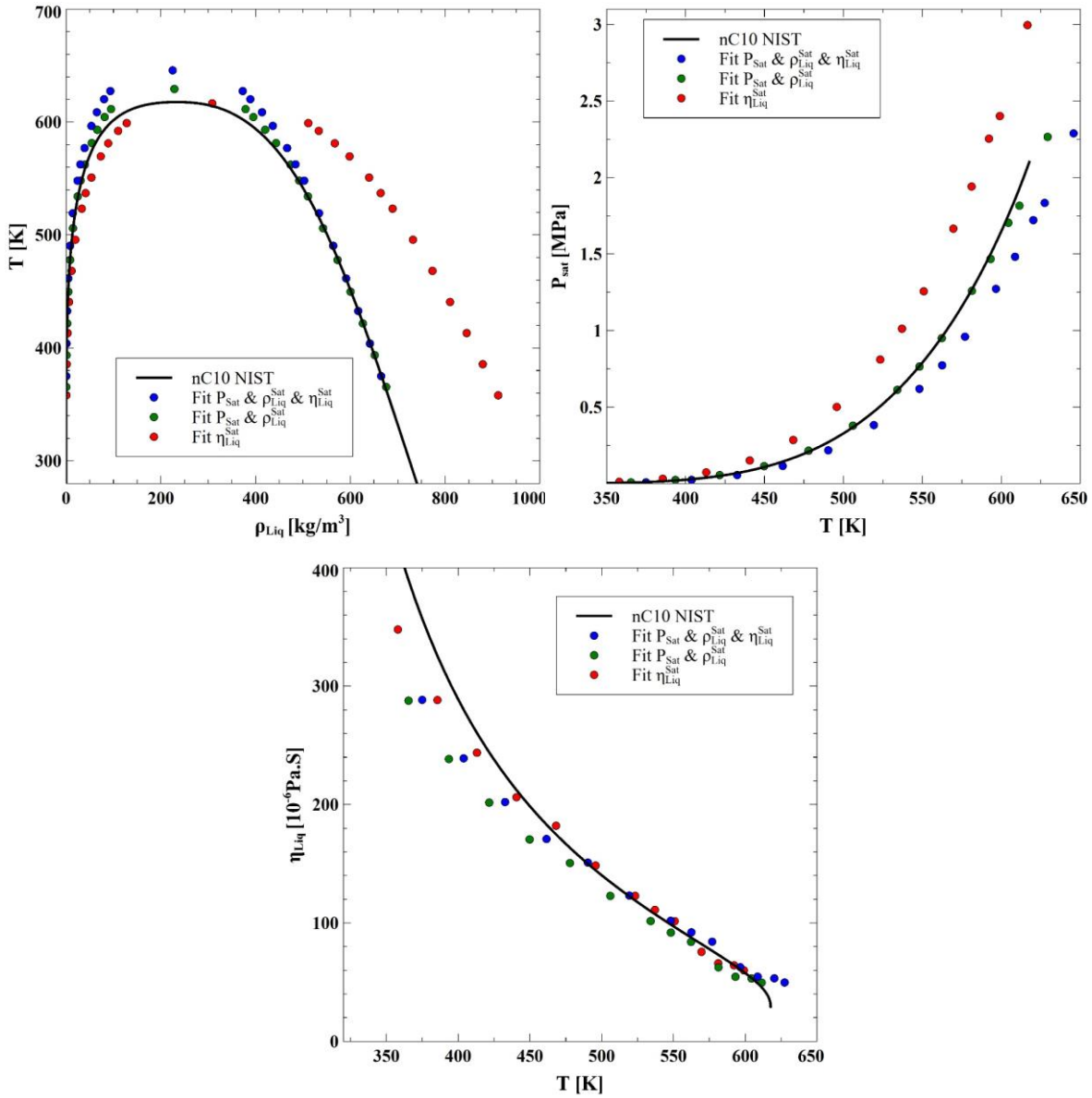


Figure 3.3: Test of different objective functions for the parametrization of the fully-flexible LJCCG force field along the vapour-liquid coexistence line.

From these results, it is clear that none of the options lead to a good compromise in the prediction of the two type of properties, due to the fact that three parameters are not enough to give flexibility to the model to represent these fluid properties simultaneously. Moreover, the incapability to reproduce the viscosity at low temperature, even in the case where this property was set as the only target property, indicates some intrinsic limitations of the chosen CG model.

Thus, to address the issue raised previously, it is proposed in this work to add an additional fourth molecular parameter to the FF-LJCCG related to the molecular internal degrees of freedom. More precisely, this additional parameter consists of modulating the molecular "stiffness" using a bending potential type:

$$U_{bend}(K_\theta) = \frac{1}{2}K_\theta(\theta - \bar{\theta})^2 \quad (3.5)$$

where K_θ is the stiffness coefficient, $\bar{\theta} = \pi$ is the equilibrium angle between three adjacent beads.

Thus, rather than being completely freely jointed, three successive segments of the same molecule, are now constrained through a harmonic bending potential around an equilibrium angle of 180°. The proposed model is so called Semi-Rigid LJCCG (SR-LJCCG).

3.4.2 Simulations details

3.4.2.1 Equilibrium properties calculations

Extensive Monte Carlo molecular simulations were performed to determine the vapour liquid phase equilibria of the of the Semi-Rigid LJCCG in the Gibbs Ensemble (GEMC) [84, 85] using MCCCS Towhee software [86]. The configurational-bias Monte Carlo method [87, 88] was used to improve molecular acceptance moves and transfer between the two boxes when dealing with long chains. A minimum of 1000 monomers were taken for the longest chains and a maximum of 2000 monomers were taken for the shortest ones. Each simulation was performed in two steps, an equilibration step consists in $40\,x\,10^6$ moves, followed by a production step consisting in $100\,x\,10^6$ moves. The cut-off distance was set equal to 4.2σ.

In the following most results are presented in classical LJ dimensionless units, which are given as follows:

$$\delta t^* = t\epsilon^{1/2}M^{-1/2}\sigma^{-1} \tag{3.6}$$

$$\rho^* = \rho_n\sigma^3 \tag{3.7}$$

$$P^* = \frac{P\sigma^3}{\epsilon} \tag{3.8}$$

$$T^* = \frac{k_B T}{\epsilon} \tag{3.9}$$

$$K^* = \frac{K_\theta}{\epsilon} \tag{3.10}$$

$$\eta^r = \eta\frac{\sigma^2\rho^{-2/3}}{M\sqrt{\epsilon k_B T}} \tag{3.11}$$

where t is the time, $\rho_n = N/V$ is the number density, N is the total number of spheres in the system (i.e., $N=mNc$, with m is the chain length and Nc is the total number of molecules), V is the volume of the simulation box, M is the mass of a sphere, P is the pressure, T is the temperature, η is the viscosity and η^r is the reduced viscosity.
The critical temperature and density were estimated using the following classical scaling laws:

$$\rho_L^* - \rho_V^* = A(T_C^* - T^*)^\beta \tag{3.12}$$

$$\frac{\rho_L^* - \rho_V^*}{2} = \rho_C^* + B(T_C^* - T^*) \tag{3.13}$$

The critical pressure is then deduced from a second order Antoine like relation as follow:

$$Ln(P_C^*) = \frac{A}{T_C^{*2}} + \frac{B}{T_C^*} + C \tag{3.14}$$

Where *A*, *B* and *C* are the fitted parameters on the saturation pressure up to $T/T_C \leq 0.95$.

More precisely, regarding the determination of T_C^* and ρ_C^*, the scaling have been fitted progressively by increasing the temperature range up to the critical point. Doing so, may provide critical properties which are function of the number of data points considered in the fitting. After reaching a stabilization, T_C^* is taken as the average of the stabilized plateau. A similar strategy has been applied to the calculation of the acentric factor, by including progressively low saturation pressures in the fitting. An example is given in Figure 3.4. This strategy helps to avoid including data points that may bias the fit.

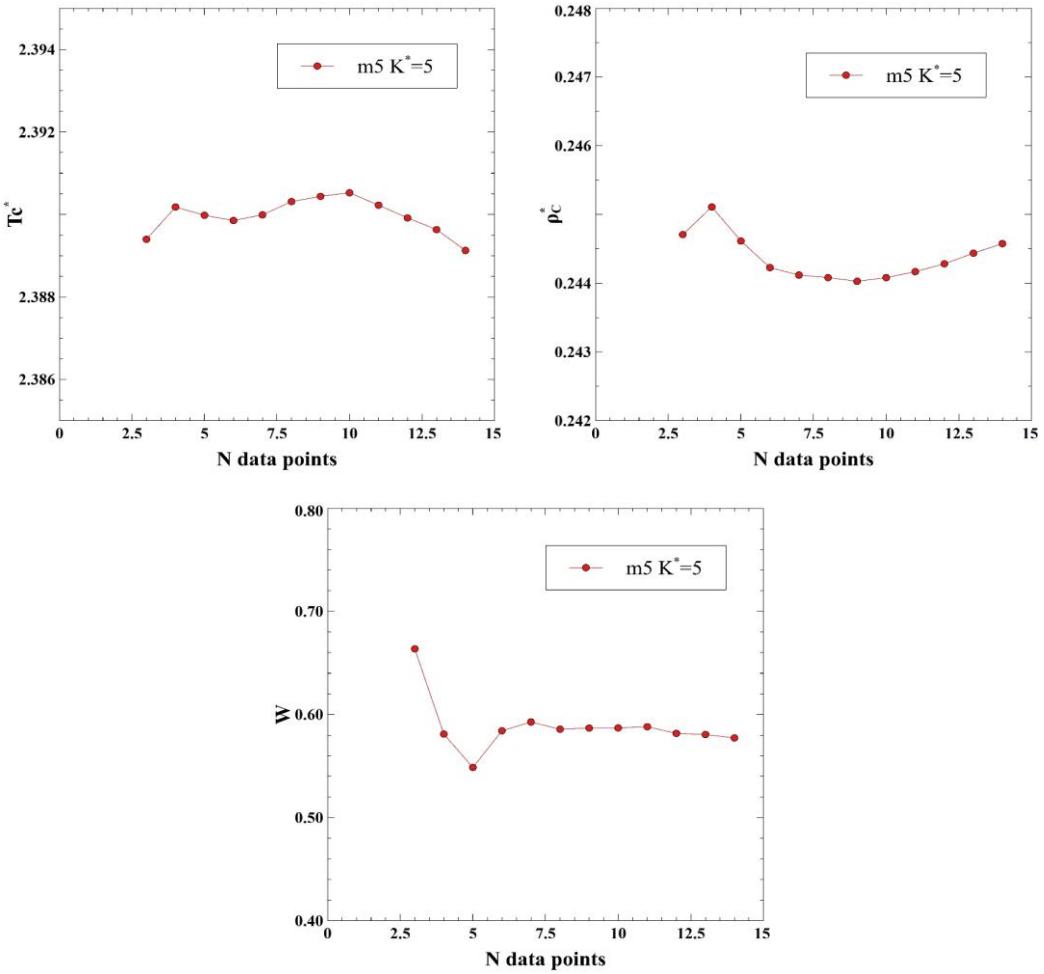


Figure 3.4: Evolution of the critical temperature, density and the acentric factor with respect to the number of data points included in the fit.

The derivative properties were calculated using Monte Carlo simulation in the NPT Ensemble by means of an in-house code [89, 90]. In all cases, an equilibration step consisting of a $50\ x\ 10^6$ moves followed by an equilibration step of a $200\ x\ 10^6$ steps were used. The cut-off radius was put equal to 5σ. These properties were deduced from the fluctuation theory. The reader is referred to chapter2 (NpT ensemble section) for the details of the equations.

3.4.2.2 Transport properties calculations

The viscosity was calculated using the Reverse Non-Equilibrium Molecular Dynamics (RNEMD) [91] simulation in the NVT ensemble by means of the homemade code "Transpore" validated elsewhere [18, 81, 92]. The system is equilibrated for a minimum of $1.5\ x\ 10^6$ time steps, followed by a transient regime of $2\ x\ 10^6$, and then a production step of a minimum $10\ x\ 10^6$ (for short molecules) to $16\ x\ 10^6$ (for the longest molecules) time steps. The self-diffusion coefficient is obtained using MD simulation run in NVT ensemble. The run was separated into an equilibration step of a $1.5\ x\ 10^6$ time steps and a production step of $18\ x\ 10^6$ time steps. The initial position was updated each $3\ x\ 10^6$ time steps.

The velocity Verlet algorithm is employed to integrate the equations of motion with time step $\delta t^* = [0.002 - 0.003]$. The bond length between successive beads in the chain molecule was maintained fixed by means of the RATTLE method [93]. To maintain the temperature, the Berendsen thermostat [94, 95] was applied. The classical periodic boundary conditions with Verlet neighbor's lists are applied. The simulation box, containing a minimum of 2000 monomers s is required to have sufficient molecules in each slab of the simulation domain, and were set up such that $Lz = 2Lx = 2Ly$. The cut-off radius of 3σ was employed to compute the LJ interactions.

The use of the RNEMD simulations for viscosity requires a careful choice of the particle exchange frequency F_{exch} (N_{swap}). This should be updated function of the thermodynamic conditions and the length of the chain molecules, to ensure that no shear thinning is occurring [92, 96, 97] and that the signal to noise ratio is high. Too high values of F_{exch} will lead to shear thinning, and too low values will result in high error bares as can be seen from Figure 3.5.

Figure 3.5 reveals that the shear thinning effect becomes more pronounced, and occurs at an earlier stage, as the rigidity increases. The trends obtained correspond typically to a Carreau like model [98] used for viscosity, where the plateau, corresponding to the Newtonian regime, is only achieved at very low frequencies for high rigid chains. This shear thinning is due to the local reorganization of the fluid, making the system anisotropic. This molecular organization is more likely to occur in linear chains. Thus, to limit this effect, we used the same strategy has the one already been presented in [92], by comparing the normal components of the stress tensor: $N_{xy} = \tau_{xx} - \tau_{yy}$ and $N_{yz} = \tau_{yy} - \tau_{zz}$, which should be zero in the case of an isotropic fluid. The normal stress has been calculated with the following formula:

$$\tau_{JJ} = -\frac{1}{V^*}\left(\sum_{i=1}^{N_T} v_{i,l}^* v_{i,l}^* + \sum_{i=1}^{N_T}\sum_{j>l}^{N_T} r_{ij,l}^* F_{ij,l}^*\right) \quad (3.15)$$

where $l = x, y$ or z, $v_{i,l}^*$ is the l component of the velocity of particle i, $r_{ij,l}^*$ is the l component of the center-to-center distance between particle i and j and is the l component of the interaction forces between particle i and j.

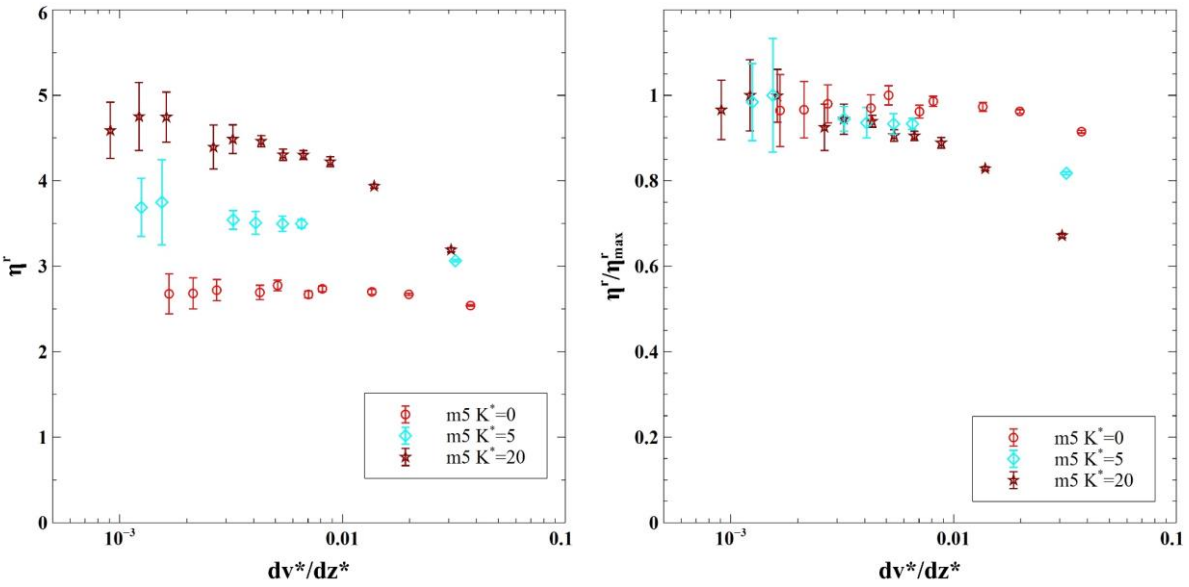


Figure 3.5: Effect of the shear rate on the viscosity of chain of m=5 at T_r=0.7 and ρ^*_{sat} corresponding to the different rigidities considered.

An example of the behaviour of the N_{xy} and N_{yz} with respect to the shear rate is given in Figure 3.6. It is clear that the differences of the normal stresses are decreasing with the shear rate, consistently with what observed in Figure 3.5. It confirms that the Newtonian plateau regime corresponds to the zero values (or almost) of the N_{xy} and N_{yz}, while the sear thinning regime corresponds to non-null values of the two components N_{xy} and N_{yz}. These quantities have then been used as quality check parameters to validate the viscosity data produced in this work.

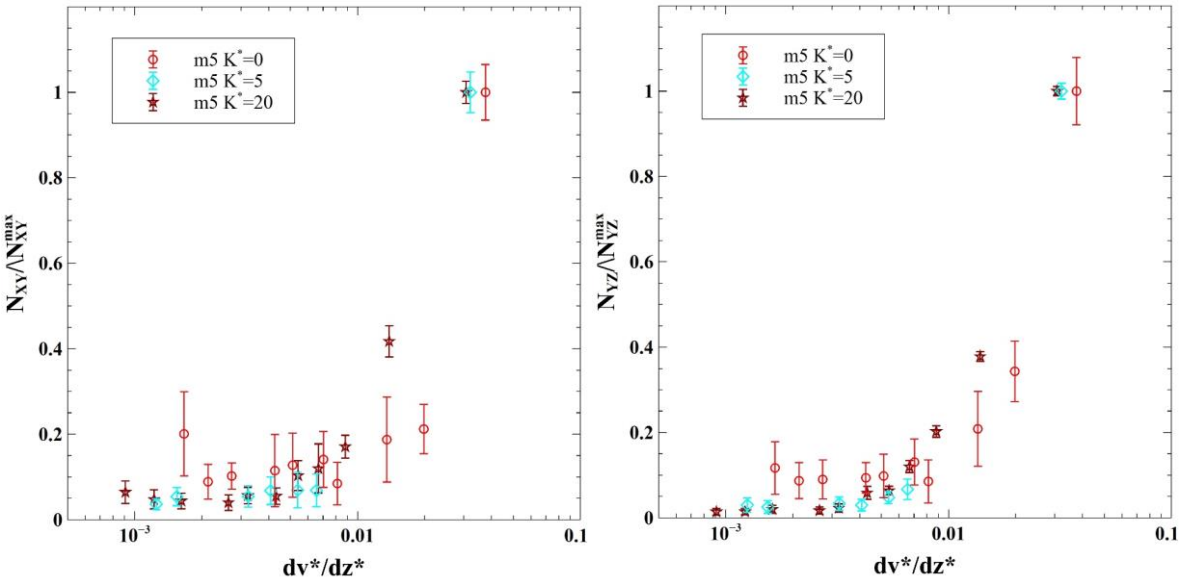


Figure 3.6: Effect of the shear rate on the normalized stress differences Nxy and Nyz for m=5 at the same thermodynamic conditions given in Figure 3.5.

3.4.3 Rigidity effect on phase equilibria and transport properties of the fluid model: The Semi-Rigid LJCCG

3.4.3.1 Phase equilibria

Studying the effect of molecular internal degrees of freedom on the phase equilibria of some coarse grained model was investigated years ago by several authors, [99, 102, 103, 104], in particular for hard sphere chains and Lennard-Jones chains.

In a fully flexible chain, the bending and torsional potentials are lacking, allowing the molecule to explore all possible conformations. Whereas, in a fully rigid chain, these potentials are considered and their corresponding stiffness constants are large enough to prevent the molecule from exploring conformations other than that close to the equilibrium angles (bending and torsion). In the case of a linear rigid model, the conformation corresponds to an angle of 180°. This linear fully rigid LJCCG exhibits a wider phase envelope compared to the fully flexible LJCCG, a higher critical temperature, but a lower critical density and pressure [99, 103]. It is expected that molecules consisting of more than six segments will not exhibit stable phase diagrams [99], but could exhibit liquid crystalline mesophases (nematic and smectic) at high temperature [82] as already been observed for hard sphere chain molecules [104, 105]. In the case where the equilibrium angle is different from that of a linear model (180°), the behaviour exhibited is completely different and is strongly dependent on the equilibrium angle [100, 101]. For small angles, the phase envelope becomes wider, and when the angle increases, it causes a shrinkage of the phase envelope to again becomes wider after certain threshold angle.

It worth mentioning that in all the studies mentioned above, none of them have investigated the effect of a semi-rigid model on the phase equilibria and particularly on transport properties. by gradually varying the bending stiffness constant. Hence, in this work, it is proposed to study in a first step the effect of the bending stiffness constant on the behaviour of fluid models, and, in a second step, the capabilities of such coarse grained model to describe the real fluid properties.

The effect of gradually varying the stiffness parameter from $K^*=0$, corresponding to the fully flexible model, to infinity $K^*=\infty$, corresponding to the linear fully rigid model, on the phase equilibria is presented in Figure 3.7 for chains consisting of five tangent spherical segments (LJC5). For small values of K^* (up to around 5), the SR-LJCCG model exhibits a phase envelope shifted to the low densities and the critical properties are all lowered as illustrated in Figure 3.7 when compared to the fully flexible model. At high values of K^*, the vapour side of the phase envelope is still shifted to the low densities, whereas, the liquid side is shifted towards the high densities. One may imagine that this increase in density could be assigned to the fact that the molecules become almost linear making them more favourable for a better arrangement while this is less favourable at small K^* values. The same non-monotonous behaviour is observed on critical temperature T_c^*, while the critical density exhibits a monotonous decrease. These non-monotonous behaviours are consistent with what was observed in another work [92] for the effective repulsive parameter γ involved in the thermodynamics scaling approach when considering semi-rigid chains molecules. Similar conclusions were obtained when varying the equilibrium angle [101] of the fully rigid LJ chains.

Regarding the saturation pressure, a monotonous decrease is obtained for all stiffness values as shown in the same figure This is consistent with the monotonous decrease behaviour obtained for the vapour phase density since both properties are nearly linearly connected (perfect gas law). The critical pressure P_c represented by the end points of the saturation pressure curves are globally showing a monotonous decrease with the chain stiffness as for critical density.

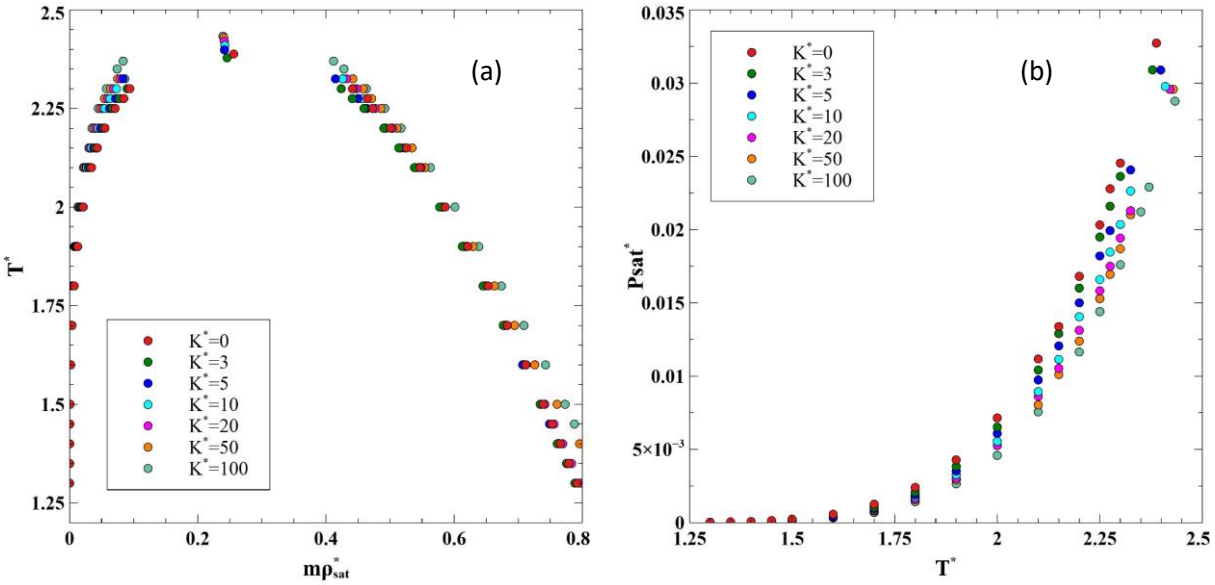


Figure 3.7: Effect of the rigidity on the coexistence densities and saturation pressure of chain of m=5. (a) L-V Equilibrium densities, (b) Saturation pressure.

3.4.3.2 Saturated liquid viscosity:

It is well known that the effect of the internal degrees of freedom on the viscosity are not negligible [10, 78, 106, 107]. Figure 3.8 shows that the liquid saturated viscosity is a lot more sensitive to the bending potential strength than what is observed on the equilibrium properties, especially at low temperature. For instance, at T^* =1.35, the liquid density for K^* =20 increases by 0.55% compared to the fully flexible case, while the liquid viscosity increases by 141.8%. This particular behaviour is particularly interesting and justifies once again our choice for this molecular model to overcome the limitation of the low temperature viscosity prediction of real fluids when dealing with fully flexible chains [78].

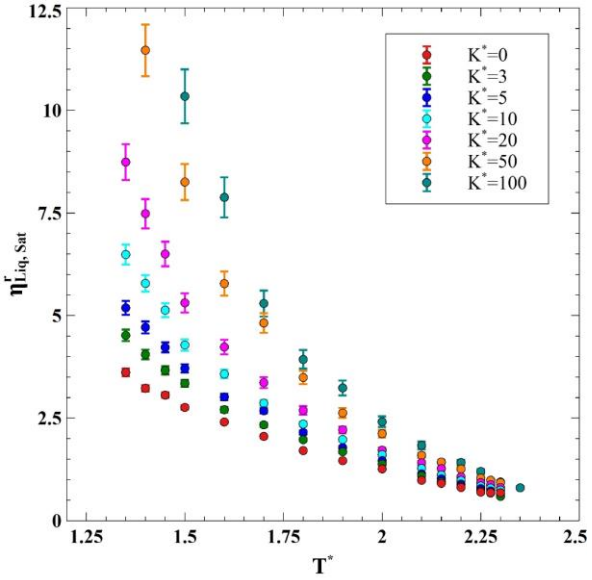


Figure 3.8: Effect of the rigidity on the saturated liquid viscosity of chain of m=5

3.4.4 Choice of the parametrization strategy

We have used the strategy of the corresponding states to parametrize the SR-LJCCG molecular model force field [18, 39]. The SR-LJCCG model contains four molecular parameters $(\varepsilon, \sigma, K^*, m)$ which will be determined using a top-down approach based on the CS strategy. With this approach the molecular parameters are constrained to reproduce some experimental macroscopic properties. Each parameter is optimised on a certain macroscopic property, as determined by a dimensional analysis.

The well depth ε, which is the energetic interaction parameter at the molecular scale, is directly related to an energetic macroscopic quantity which is the critical temperature. The segment collisional diameter σ, which represents the molecular volume at the molecular scale, is directly related to a purely volumetric macroscopic property, which is the density at a reduced temperature, $T_r = T/T_c$, equal to 0.7.

Now, regarding the stiffness parameter and the number of segments (K^*, m), this couple of parameters is optimized on both the acentric factor ω [108] and the liquid saturated viscosity at T_r=0.7, as both properties are affected by the two parameters. The acentric factor ω, which mainly quantifies the deviation of the molecular shape from being a sphere, is strongly dependent on the number of sphere segments m forming the molecule. The rigidity *K* has also an effect on the shape of the molecule and so on the acentric factor. Regarding the viscosity, it is well known that this property is affected by the number of segments [109] because of the increase of the strength of the intermolecular interactions involved and also due to the increase of the probability of molecular entanglements. The viscosity is also strongly sensitive to the molecular stiffness K^* parameter, particularly at low temperature as shown in Figure 3.8 .

The effect of the segment number and the rigidity parameter on the selected properties for the optimization are summarized in Figure 3.9. Note that, to reduce the uncertainties on these properties, a strategy adopted in this work for the acentric factor and critical temperature were discussed in the simulations details section, however, the case of $\eta^{r,Exp}_{LiqSat}$ or the ρ^{Exp}_{LiqSat} at $Tr =$

0.7, were not. Thus, it is worth to mention that, these properties are not determined by performing simulations at exact thermodynamic conditions corresponding to $Tr = 0.7$, but, instead, they were obtained by fitting the saturation curves for both properties approximately in the range $Tr = [0.6{\sim}0.85]$.

The proposition of using the $\eta_{LiqSat}^{r,Exp}$ or the ρ_{LiqSat}^{Exp} at $Tr = 0.7$, instead of using the corresponding critical quantities, is, first, to be in line with the acentric factor which is calculated at the same reduced temperature, and second, because the critical viscosity does not exist (it diverges) and the critical density is one of the most uncertain experimental thermodynamic property, as it is an extrapolated one.

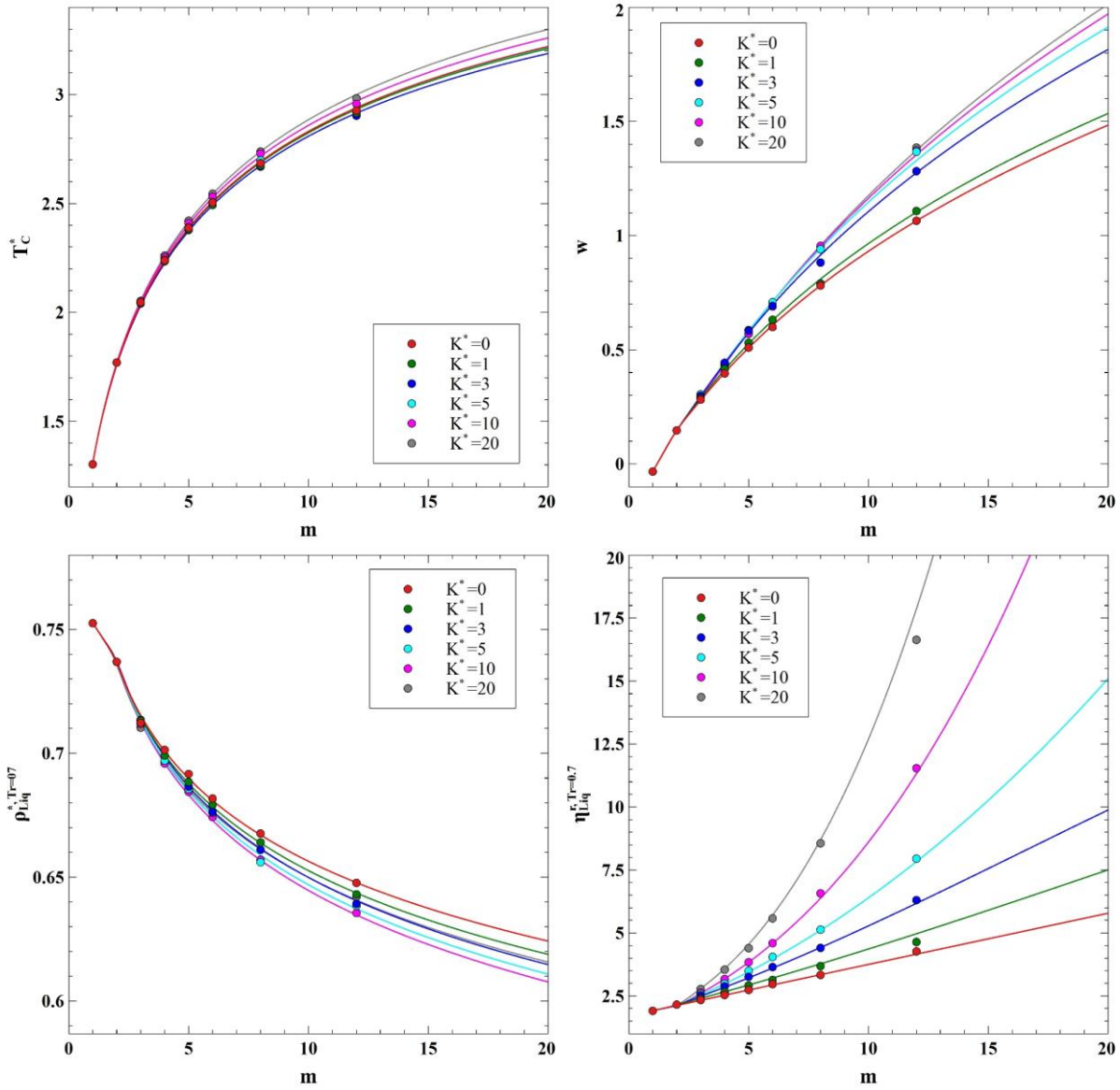


Figure 3.9: Thermophysical properties of the semi-rigid LJCCG model used in the parametrization of the coarse grained force field proposed.

The use of the CS offers many advantages. First, it ensures obtaining a unique set of parameters avoiding the overfitting problem that may arise when using many experimental data and many

different properties to simultaneously optimise the force field parameters. Second, the estimated parameters are physically consistent, since thanks to this strategy, a one to one relationship between the molecular scale parameters and their corresponding macroscopic properties is used. Third, it requires only four accessible experimental measurements per species.

3.4.5 Optimization procedure

Thus, to determine the set of parameters $(\varepsilon, \sigma, K^*, m)$ that define the SR-LJCCG parameters of a given fluid requires the knowledge of the macroscopic experimental data T_C^{Exp}, ω^{Exp}, *and* $\eta_{sat,L,Tr=0.7}^{r,Exp}$ and that from the Coarse Grained model $T_C^{CG}(K^*, m)$, $\rho_{sat,L,Tr=0.7}^{CG}(K, m)$, $\omega^{CG}(K^*, m)$, and $\eta_{sat,L,Tr=0.7}^{r,CG}(K^*, m)$.

Most of the time, the experimental data are available in the databases such as those of the NIST [110], DIPPR [111], DDB [112], DECHEMA [113], DETHERM [114], TDE [115], ETB [116], Cheméo [117]…or from some books like [118, 119]. If no experimental data are available for certain property or because it could not be measured, a group contribution method or existing correlations provided in the literature could be an alternative. For instance, the properties of CO2 at a reduced temperature of 0.7 do not exist because the triple point of this molecule is situated above T_r=0.7. For this specific case, an extrapolation of the equilibrium density using some scaling [120] give reasonable prediction. For the viscosity, the use of the entropy and density scaling is suggested [121, 122].

Regarding the coarse grained model properties, extensive molecular dynamics and Monte Carlo simulations were performed to obtain these properties. Then, they were correlated as shown in Figure 3.9 . The mathematical formulation and the fitting parameters of these correlations are given in the appendix 7.A.2.

In practice, the set of parameters (m, σ, ε, K^*) of a given fluid of interest are optimised as follow. The K^* and m are first obtained by minimising the following objective function:

$$minF(m, K^*) = \left(\frac{\omega^{CG}(m,K^*) - \omega^{Exp}}{1+\omega^{Exp}}\right)^2 + \left(\frac{\eta_{sat,L,Tr=0.7}^{r,CG}(m,K^*) - \eta_{LiqSat,\ Tr=0.7}^{Exp}}{\eta_{LiqSat,\ Tr=0.7}^{r,Exp}}\right)^2 \quad (3.16)$$

Then, the ε and σ, are directly deduced from the following relations:

$$\varepsilon = \frac{k_B T_C^{Exp}}{T_C^{CG}(m, K^*)} \quad (3.17)$$

$$\sigma = \sqrt[3]{\frac{M \rho_{sat,Liq,Tr=0.7}^{CG}}{\rho_{sat,Liq,Tr=0.7}^{Exp}}} \quad (3.18)$$

The minimization on the acentric factor is carried out on the quantity (1 + ω^{Exp}) to avoid any singularity when ω^{Exp} is approaching zero for some small molecules. It is also worth to mention that, the choice of the Euclidean norm for the minimisation is to force the objective function to converge to the global minimum. This can be easily handled for large molecule, but can be pronounced for small molecules. In fact, for such molecules, the rigidity has only a very small effect on the acentric factor favouring then the appearance of some local minima. It is recommended, for such small molecules, to try a set of initial guesses. Another

recommendation, regarding the initial guess when dealing with hydrocarbons, is to take an initial guess for *m* from the scheme *m→1+[(C-1)/3]* [123, 37], where C is the carbon number, and a value of *K**=5.

Another important point to consider is that, in the first step of the optimization procedure, the *F function* is minimized to obtain the number of segments forming the molecule. This number is, in most cases, a non-integer number and needs to be rounded to its nearest integer value. For certain species, this real number is exactly at halfway between two successive integer numbers. In this case, the question is: rounding to the closest high integer value or to its closest low integer value. In addition, one may then imagine that in this special case, the molecule will be less accurately represented, especially when *m*=2.5, as the molecule could be represented by two segments for which rigidity is not defined, or could be represented by three segments for which the rigidity is defined. In our experience, we found that, when rounding up, the transport properties are better predicted (because of the rigidity parameter). However, less accurate thermodynamics properties are obtained, particularly those highly sensitive to the long range interactions, such as the saturation pressure and the *Cp* as well. While when rounding down, it is the thermodynamics properties that are well predicted, whereas less accurate transport properties are obtained. A good illustration for this is found for the *iC*$_4$, for which the optimised number of segments is *m*=2.54, as it will be shown in the next section. It is worth to emphasize that if a SAFT Eos is used to predict the fluid properties instead of using a molecular simulation tool, this problem does not appear as the EoS allows using non integer number of segments.

For sake of clarity, two flow charts are given in Figure 3.10. The first shows the general methodology used for the parametrisation of the proposed force field, and the second shows the optimisation procedure followed to obtain the set of parameters of the target fluid.

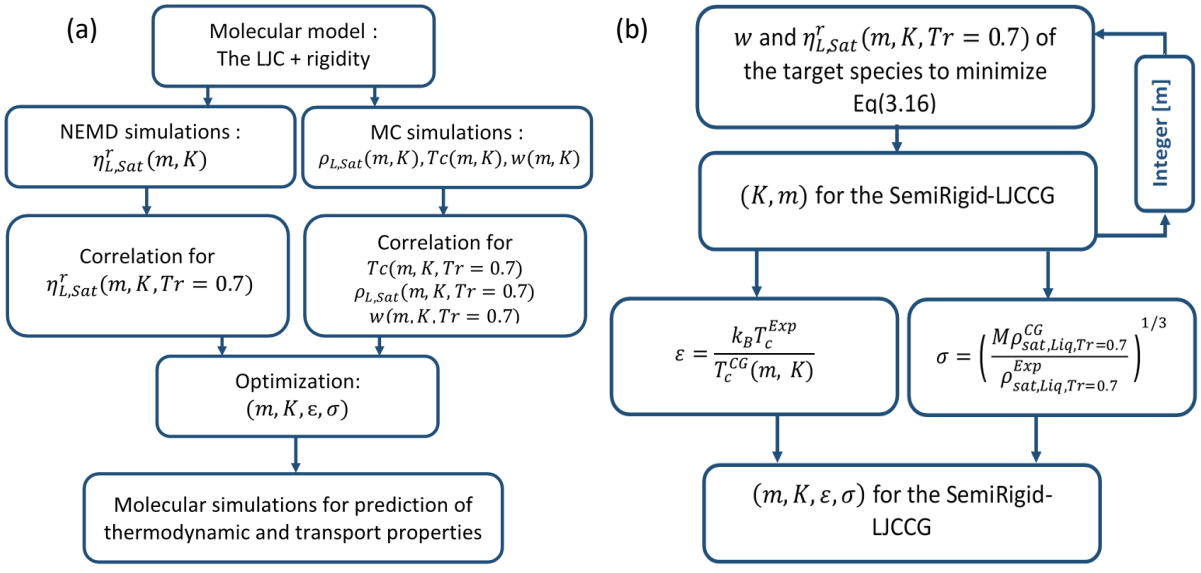


Figure 3.10: (a) flowchart of the general methodology, (b) flowchart of the optimization strategy.

3.4.6 Application to pure real fluids

In this part, the proposed model is applied to several real pure components, specifically the alkanes family and other common molecules. Various properties including phase equilibria, second order derivative, and transport properties are investigated. Note that, for a better clarity of the figures, error bars for the FF-LJCCG model are not shown. They have the same magnitude as for the SR-LJCCG model. Error bars for densities and pressures have not been included on the figures because they are most of the time smaller than the symbols.

3.4.6.1 Equilibrium properties

Extensive Monte Carlo simulations were performed to test the predictive capability of the SRLJCCG model on the vapour-liquid equilibrium properties for various polar and non-polar species as presented Figure 3.11 to Figure 3.13. As shown in Figure 3.11 excellent predictions are obtained for both equilibrium LV densities and saturation pressure for the normal alkanes up to the n-hexatriacontane (nC_{36}). Similarly, a good accuracy is also found for the iso-alkanes shown in Figure 3.12, except for the saturation pressure (and consequently, the vapour density) of iC_4, as expected from the discussion in the optimization part (open symbols correspond to a representation with two segments, while the full symbols correspond to a representation with three segments). Some polar components were also considered, as shown in Figure 3.13, and a very good accuracy is obtained for these molecules with the proposed SR-LJCCG, which is worth noting, does not include any polar parameters.

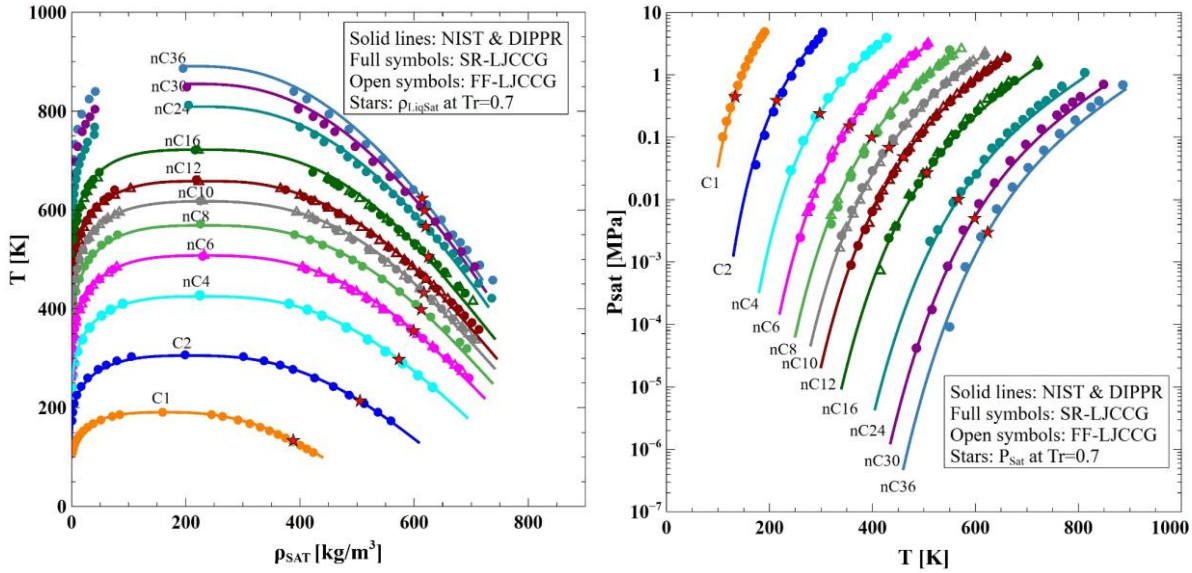


Figure 3.11: Equilibrium liquid-vapour saturated densities and saturation pressures of n-alkanes vs correlated experimental data from NIST & DIPPR databases.

These results demonstrate the robustness of the proposed molecular model and the importance of the parameterization strategy (CS) used to obtain consistent molecular parameters. It is worth mentioning that, only a few experimental data points were imposed (represented by a red star symbol in all the figures, with all other data points being pure predictions. Another interesting feature of this CS strategy is that the obtained interaction parameters can safely be transferred

to other states that are away from those at which the model was parametrized. Of course, if one is interested in very low temperatures close to the triple point T_t, the deviation will certainly be more pronounced. However, overall, the molecular parameters are less state dependent compared to what could be obtained using a BIM procedure, and it is assumed that the SR-LJCCG model will still yield reasonable results at these extreme conditions, as one would deduce from the trend of the results obtained.

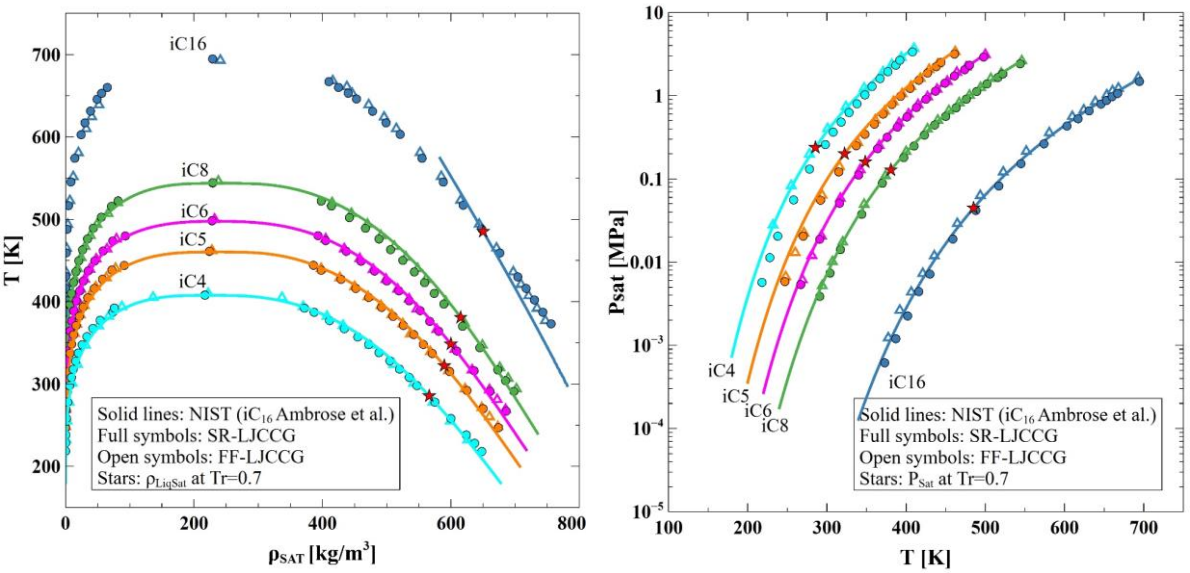


Figure 3.12: Equilibrium liquid-vapour saturated densities and saturation pressures of iso-alkanes vs correlated experimental data from NIST database. Experimental liquid density for iC16 is taken from Klein et al.2020, and Psat from Ambrose et al.1988.

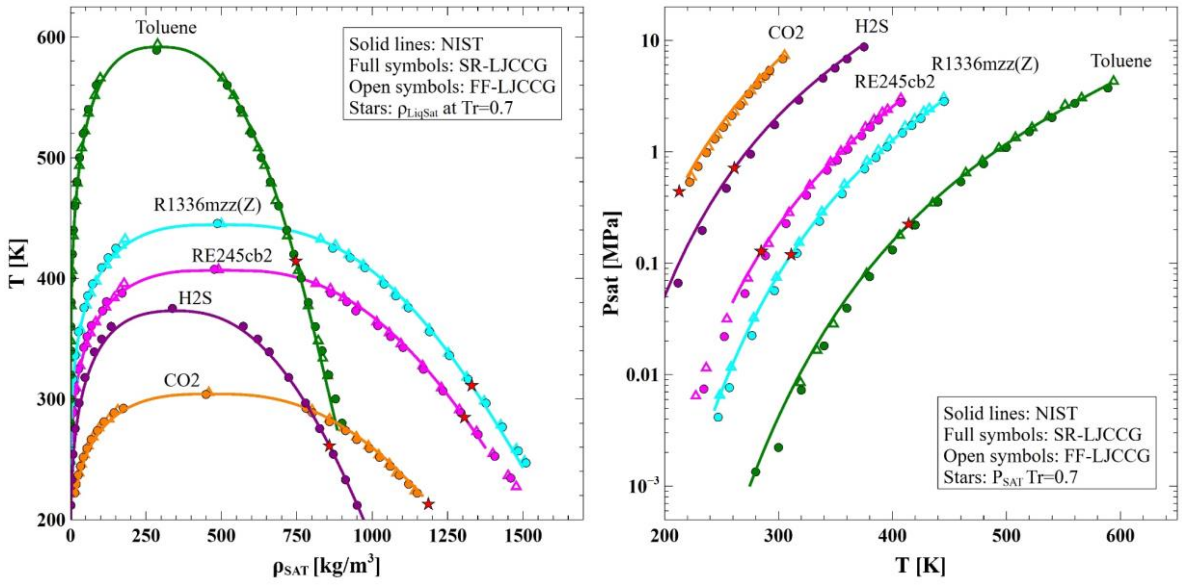


Figure 3.13: Equilibrium liquid-vapour saturated densities and saturation pressures of some polar components vs correlated experimental data from NIST database.

3.4.6.2 Second order derivative properties

Second order derivative properties are known to be one of the most difficult properties to obtain with either any equation of state (cubic or SAFT type model) or molecular models, even if Accurate predictions for these properties have already been obtained using some SAFT EoS [38, 124, 125] and some force fields [126] with intermolecular parameters being optimised on equilibrium properties. They are then an excellent type of property to test the representability of the proposed model and the parametrisation strategy (CS). The low resolution of any coarse grained model, makes them unable to accurately predict properties dictated by the molecular internal degrees of freedom (ideal part) such as the energetic and entropic second order derivative properties. To alleviate this problem, instead of using the incorrect ideal part predicted by the model, it is the correct ideal part taken from experimental correlations [118] that will be used in a similar way to what is done using cubic or SAFT EoS. The expressions used for the derivative properties are given in the chapter 2 (section NpT ensemble). Error bars are given only for the properties deduced from the direct Monte Carlo simulations. Thus, they are provided only for κ_T, α_P and the C_P. The cases of the C_V, μ_{JT} and V_s require propagation errors calculations.

Figure 3.14 to Figure 3.16 show the temperature dependency of many volumetric and energetic derivative properties at a pressure of 10MPa for the compounds studied previously. Overall, excellent results are obtained for most of the properties investigated in this work, except small deviations observed on the C_P and the V_s for large molecules. The case of iC_4 is an exception, similarly to what was observed for the saturation pressure and vapour density. Results show excellent agreement with experiments when the molecule is represented with three segments (open triangles), while the representation with two segments deviate from the experiments.

Now, concerning the speed of sound for the largest molecules (n-alkanes and iso-alkanes) at the lowest temperature, the deviations probably arise from the use of a semi-empirical LJ(12,6) potential. This was already pointed in the literature [125, 37] and a better prediction is obtained if this property is included in the fit [37, 38, 124] and when using a generalised LJ potential (Mie potential). This was at the heart of the motivation for the development of the SAFT-VR-Mie EoS version [37]. However, except at low temperature, results of the proposed model show very good agreement with experiments.

Moreover, as shown in Figure 14, the model shows good capabilities on the prediction of these derivative properties near the critical region. Indeed, at the critical point, the derivative properties exhibit a highly non regular behaviour (singularities), which then can be considered as a severe test for all models. The model proposed was then tested on some molecules such as the nC_4, iC_4 or more interestingly CO_2 and H_2S. Excellent illustration of the predictive capability of the proposed model are shown for these molecules near the critical conditions. Such an accuracy is the consequence of the robust CS parametrization used in this work.

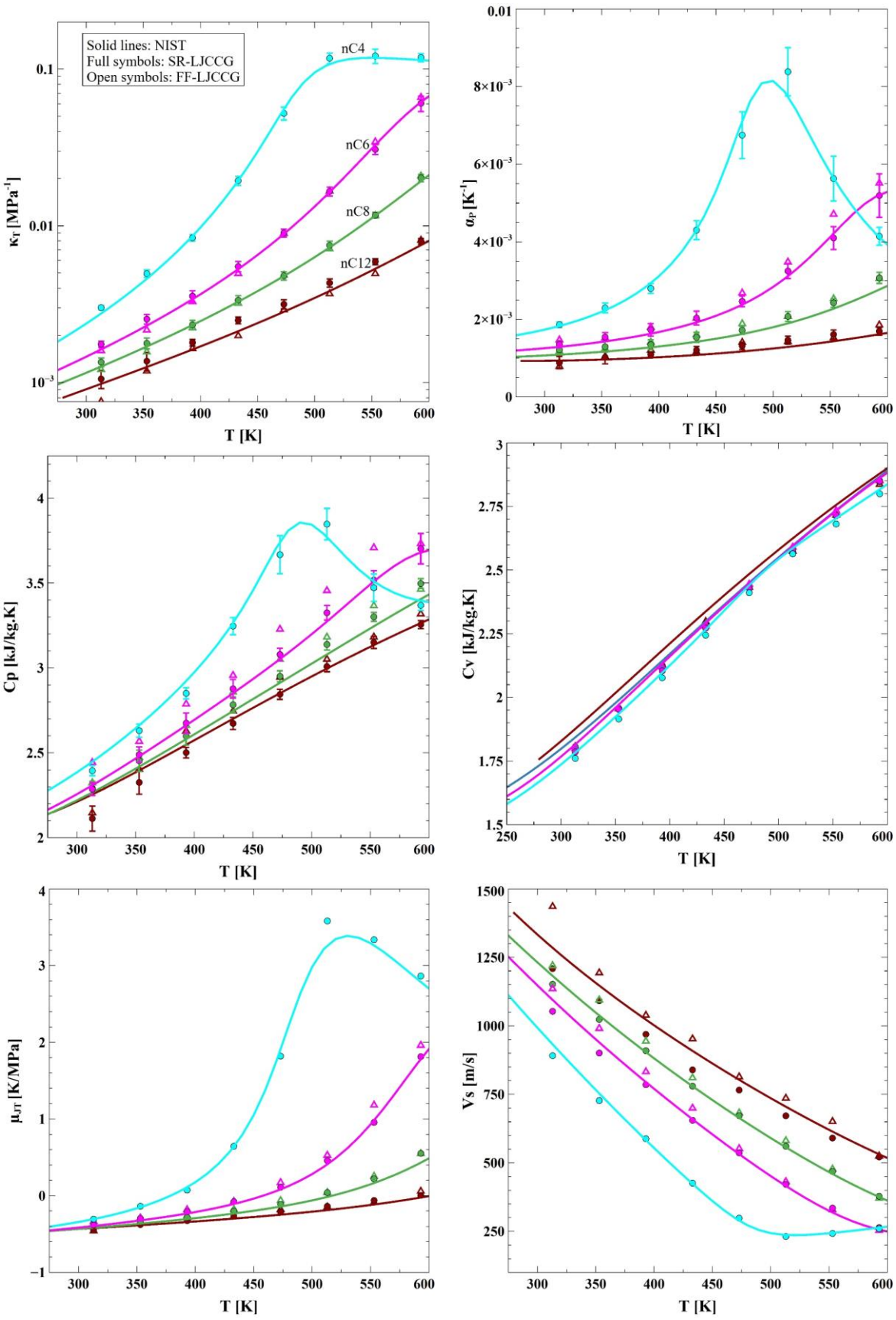


Figure 3.14: Temperature dependence of the predicted second order derivative properties of n-alkanes with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa.

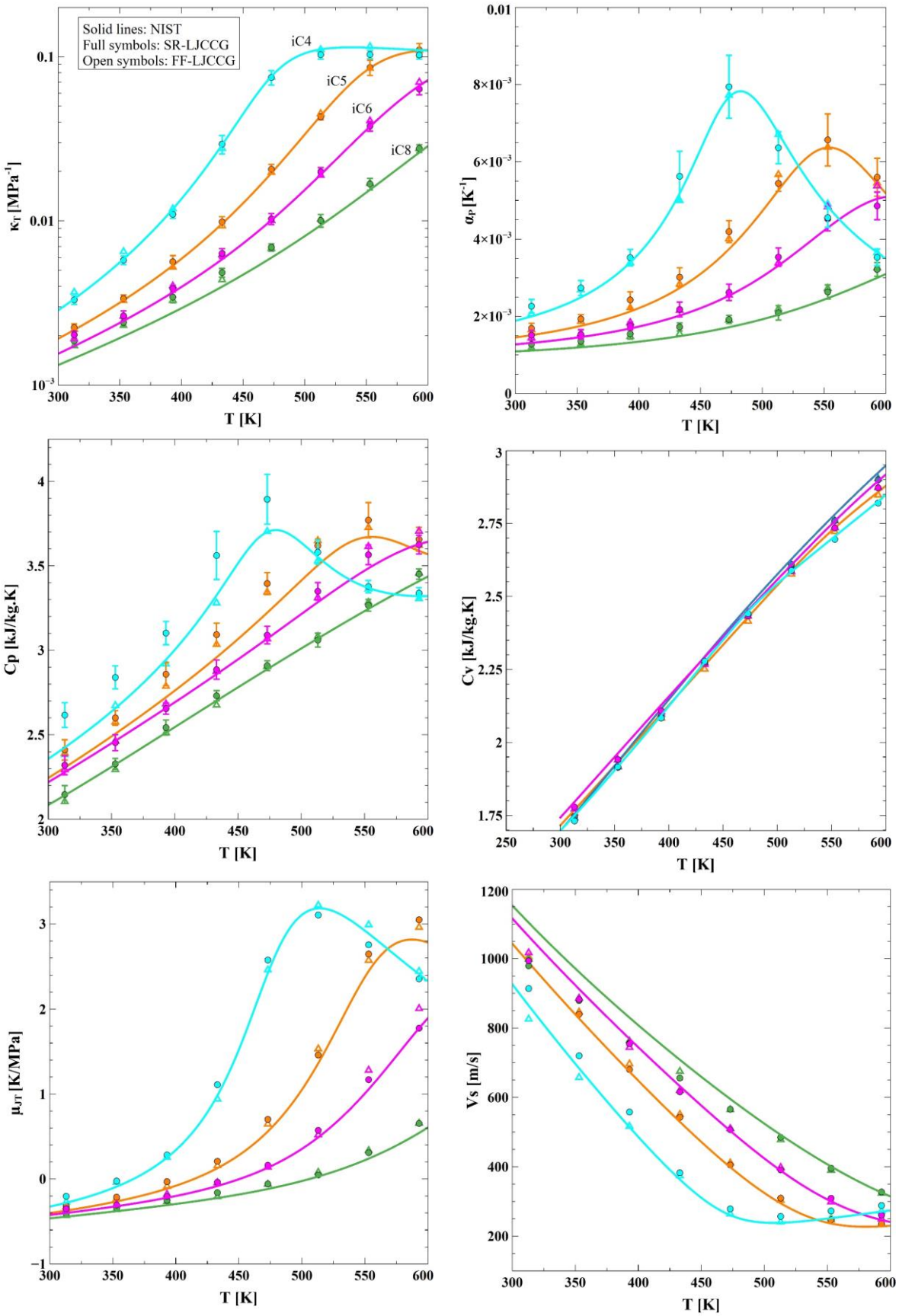


Figure 3.15: Temperature dependence of the predicted second order derivative properties of iso-alkanes with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa.

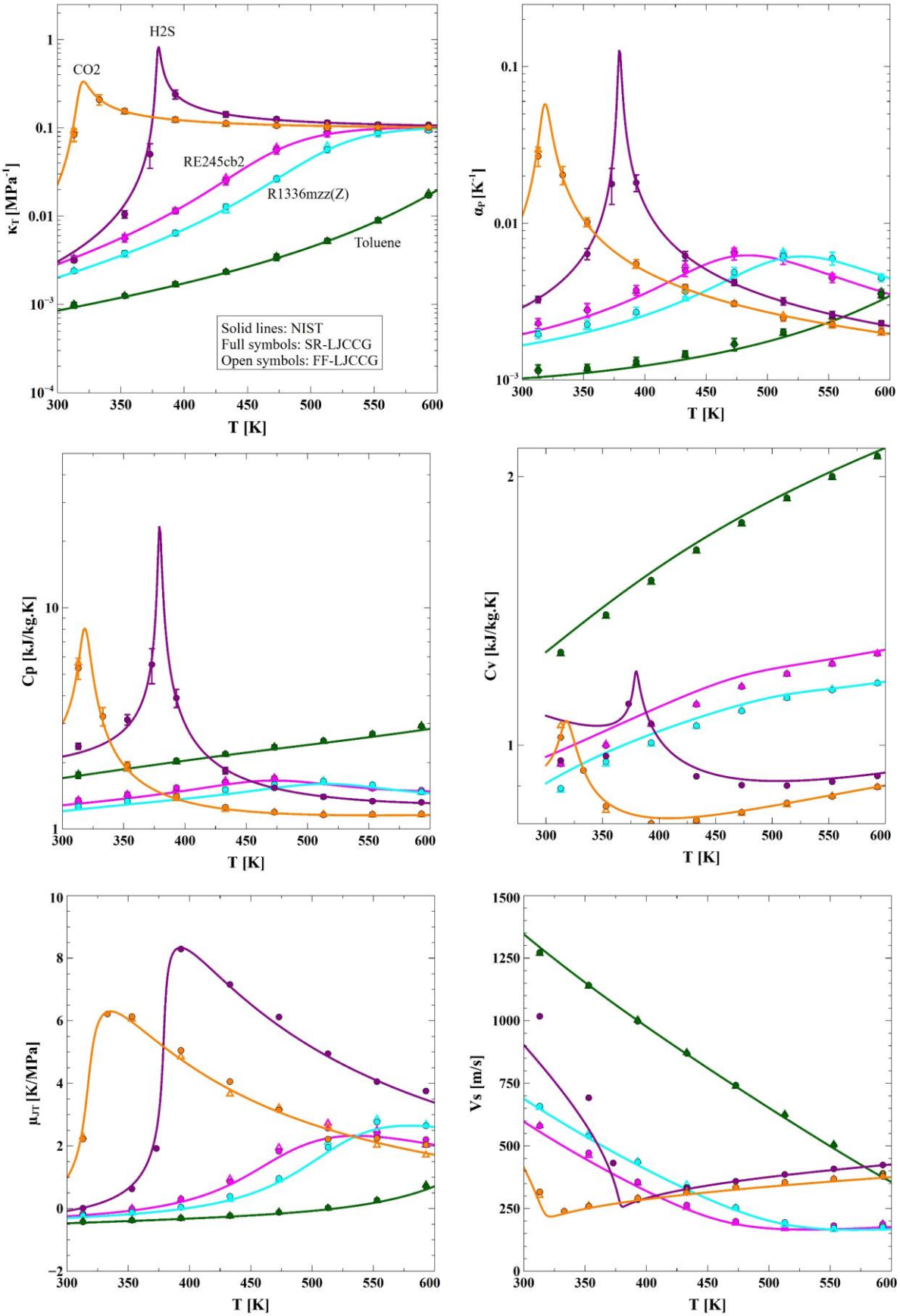


Figure 3.16: Temperature dependence of the predicted second order derivative properties of some polar components with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa.

Since the density is an output of the NpT simulations conducted to obtain the second order derivative properties shown above, we investigated its temperature dependency as illustrated in Figure 3.17. Excellent results are obtained for this property in the dense liquid phase, in the critical region and in the supercritical conditions.

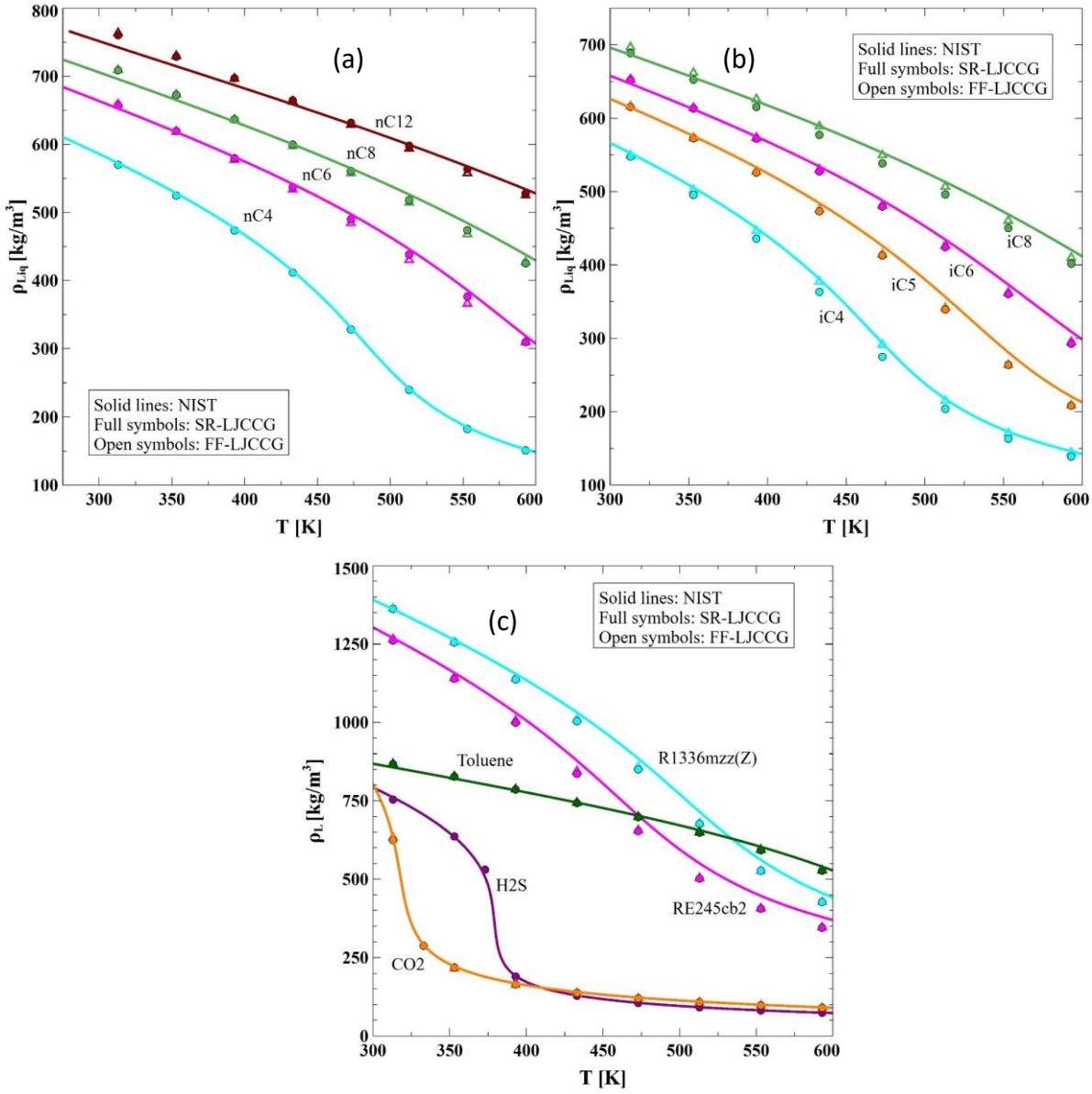


Figure 3.17: Temperature dependence of the single phase density predicted with the SR-LJCCG and FF-LJCCG vs experimental data from NIST at P=10MPa. (a) n-alkanes, (b) iso-alkanes, (c) polar components.

Last, results obtained with the FF-LJCCG are comparable to those obtained with the SR-LJCCG model regarding thermodynamic properties. Thus, the inclusion of the viscosity as a target property did not provide any improvement neither deterioration in the prediction of the thermodynamic properties investigated in this work compared to the fully flexible model. However, when considering transport properties, a noticeable difference appears between the two coarse grained models as it will be shown in the in the next point.

3.4.6.3 Viscosity

While accurate models exist for the viscosity predictions at dilute states [127], this becomes more complex where the mean free path becomes smaller than the atomic size (i.e. dense state). Many approaches have been proposed in the literature [128, 129], but they are still not fully satisfactory in terms of industrial requirements [130]. Among the alternatives to estimate shear viscosity, molecular dynamics simulation is one interesting option. However, the accuracy of the prediction obtained with MD simulations, is closely linked to the robustness of the force field used, as demonstrated in several papers for the most common existing force fields [51, 89, 78, 107, 18, 131, 132]. Indeed, a robust and simple (i.e. coarse grained) force field capable of simultaneously providing accurate transport and thermodynamic properties, is, to our knowledge, not yet available. Most of the coarse grained and united atom existing force fields tend to underestimate the viscosity, probably due to due to the over-smoothness in the molecular geometry [133]. A way to circumvent this, is to use a parametrization that takes into account viscosity as target property [10] as proposed here. In the following the performance of the SR-LJCCG model is tested on the prediction of the liquid viscosity of the compounds studied previously.

As shown in Figure 3.18, excellent estimates are obtained for the viscosity over the whole panel of molecules studied. The model shows an excellent transferability as only the red stars data ($\eta^{Exp}_{LiqSat,\,Tr=0.7}$) are imposed in the parametrization. Indeed, the prediction accuracy remains conserved even for molecules with high molecular weight (up to the nC_{36}), contrary to what was observed when using a MieCCG model [18, 89], or when using the fully flexible model as shown here.

The case of the iC_4 is as expected. Results obtained on the viscosity are excellent with the three segments model, while results deteriorate with the model with two segments. This support the discussion on this special case, where the optimal number of segments is found halfway between successive integer numbers. Results on nC_4 converges also to similar conclusion, as this molecule is represented by two segments, giving accurate thermodynamics properties, but less accurate viscosity.

To better assess the robustness of the proposed model, simulations have been performed at pressure conditions far from those used for the parameterization. To do so, simulations have been conducted on n-alkanes from atmospheric pressure (0.1MPa) up to 100MPa (about 1000 times the saturation pressure conditions). The predictions are shown in Figure 3.18 (d), where the model shows very good performance. Still, deviations start to be noticeable for the nC_{12} (and certainly for longer molecules than nC_{12}), but they remain reasonable compared to what is obtained with the fully flexible model. The deviations in terms of the compressed viscosity and density (at which the viscosity is simulated) are given for the nC_{12} molecule in Figure 3.19. The SR-LJCCG gives mean errors of ~2.31% and ~9.32% respectively for the density and viscosity, while this is around ~1.91% and 35.67% for the FF-LJCCG model. The maximum deviations are obtained for both models at pressure 100MPa, where the SR-LJCCG is underestimating the viscosity by ~12.87%, while it ~46.73% for the FF-LJCCG. Moreover, as can be seen from Figure 3.19, the deviations in terms density are well predicted by the FF-LJCCG, where the model is exhibiting a stabilization of the deviations when the pressure increases which is not the case for the SR-LJCCG model. Consequently, in terms of viscosity,

the FF-LJCCG model is exhibiting increasing deviations with increasing the pressure, while a stabilization is obtained for the SR-LJCCG model. This is because viscosity is highly sensitive to small changes in density, thus, only 2% on density may induce a change of 30% in viscosity. However, improvements are still needed to achieve high accuracy under extreme conditions, for example to meet the needs of the lubricant industry. This may possibly be done by including a torsional potential to the model similarly to what was observed in [10].

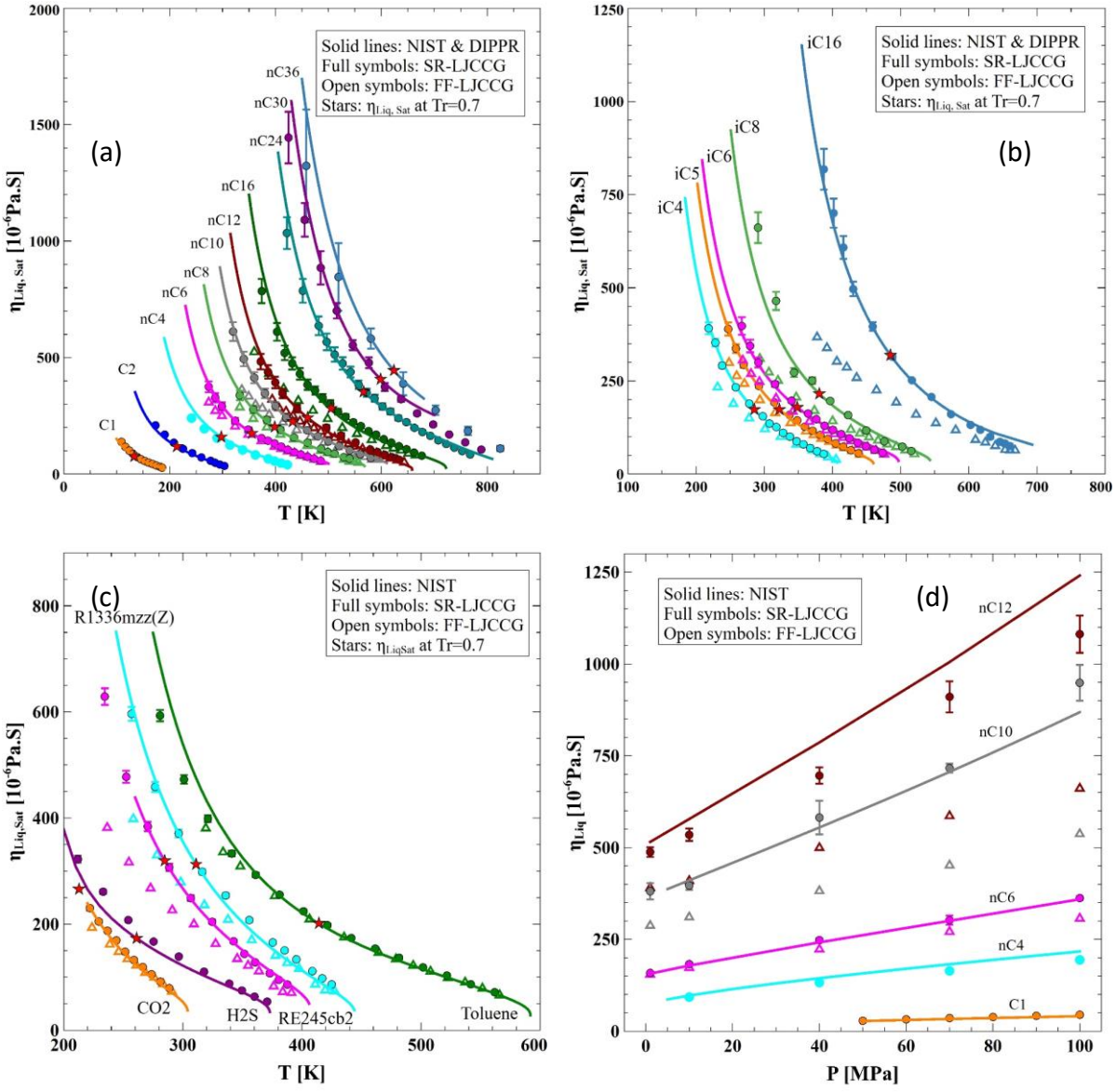


Figure 3.18: Saturated and compressed liquid viscosity of normal alkanes, iso-alkanes and some polar components. (a, b, c) viscosity of n-alkanes, iso-alkanes and polar components at saturation conditions, (d) compressed liquid viscosity of n-alkanes at temperature of T=373K

As the viscosity is the main target property that we want to improve in our prediction, an additional benchmark is conducted for this property. We compare the prediction of the proposed SR-LJCCG model to various existing force fields proposed in the literature for two normal alkanes, the n-dodecane nC_{12} and the n-hexadecane nC_{16}. Among these force fields, some are

based on fully atomistic descriptions of the molecules such as the Lipid14 [134], L-OPLS [45] and its modified version [31], while others are based on united atom models, such as the TraPPE-ua [6], POTOFF [61], TAMIE [62], SKS-AUA [135]. In addition, coarse grained models such as the Fully-Flexible LJCCG and the Mie-CCG [18] model, which were parametrized with the same strategy (corresponding states) as for the SR-LJCCG, are also included. In this latter category of CG models, the well-known MARTINI force field [19] is also included.

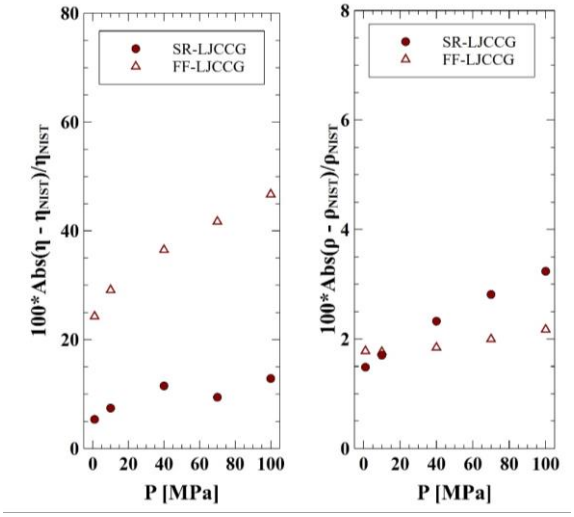


Figure 3.19: Relative errors for the compressed liquid viscosity and density of n-C12 at T=373K. Results of viscosity are those presented in previous figure and for density are not shown but correspond to those at which the viscosity is simulated.

The simulation data are taken from [51] for the TraPPE-ua, POTOFF, TAMIE. Data from [30] are used for L-OPLS, Lipid14 and MARTINI force fields. The modified-L-OPLS data are taken from [31], and those of SKS from [136]. Regarding the coarse grained models FF-LJCCG and MieCCG, data were computed in this work, while the experimental data are from NIST database [110].

It can be clearly seen from Figure 3.20 that, the Semi-Rigid LJCCG gives among the best predictions, overall, where very good agreement with experiments are obtained for both molecules at high temperature. The deviations obtained at low temperature remain reasonable, and, in relative, they are smaller than the other force fields.

Another interesting result highlighted by this benchmark is that, the success of the Semi-rigid LJCCG is not only due to the inclusion of the additional rigidity parameter, but also to the parametrization strategy adopted in this work. In fact, this latter point could be understood by benchmarking results of the Fully-Flexible LJCCG and the TraPPE-UA as both are based on the LJ intermolecular potential interaction. Comparable results are obtained with both models, whereas the TraPPE-UA integrates all the internal degrees of freedom. Thus, it shows that, the inclusion of all the internal degrees of freedom of the molecule do not systematically lead to improvement of the transport properties if not appropriately parametrized. A possible way to extend the TraPPE-UA for the transport properties could be achieved by re-parameterizing the torsional and/or the bending potentials to reproduce the viscosity, coupled to the re-parameterization proposed by Janecek and Paricaud [49] to improve the saturation pressure.

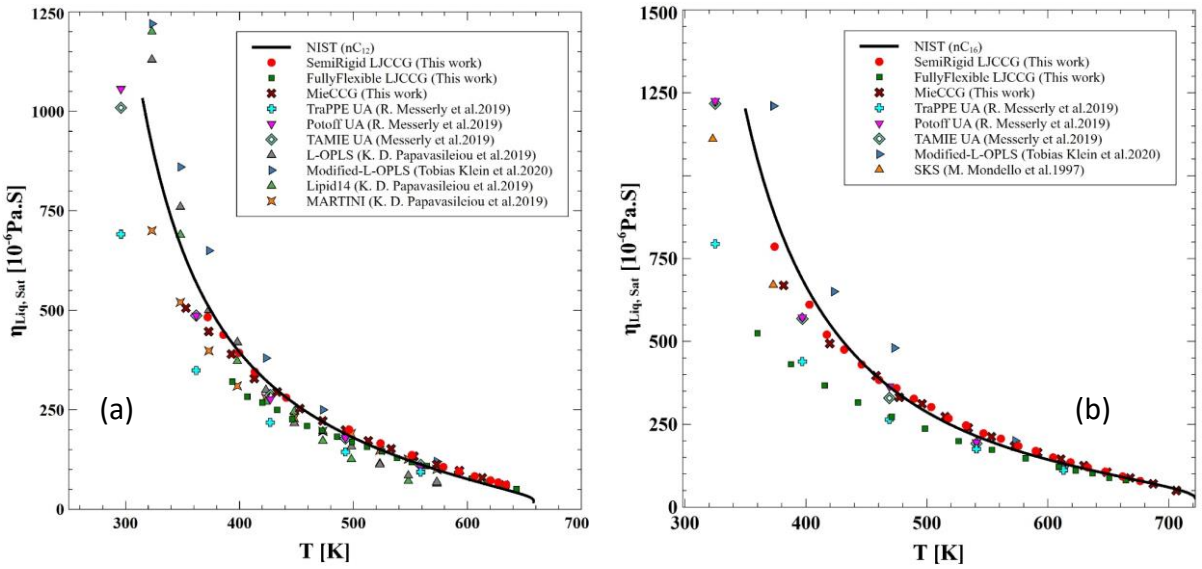


Figure 3.20: Benchmark between different force fields proposed in the literature and experiments for the saturated liquid viscosity prediction. (a) n-C12, (b) n-C16.

3.4.6.4 Diffusivity

The diffusivity is another important transport property to test the robustness of the SR-LJCCG force field. Only few experimental data are reported in the literature for this property, which explain the limited number of compounds evaluated here.

Results for the self-diffusion coefficient for different n-alkanes at temperatures T=303K, T=343K and T=383K and at densities corresponding to a pressure of 0.1MPa are reported in Figure 3.21. Good agreement with the experimental data [137] are found for the three temperatures studied here and for chain length up to nC_{28}. Slightly better results are obtained with the fully flexible LJCCG for low molecular weight species, but this model tends to largely overestimate the diffusivity compared to the SR-LJCCG as the molecular weight increases.

The pressure effect up to 150MPa was also checked for the nC_6 and nC_8 at a temperature of 298K, as shown in Figure 3.21 . The predictions are all in fair agreement with the experiments [138, 139] using both SR-LJCCG and FF-LJCCG models. It is interesting to point out that self-diffusion overestimation is often attributed to the smoothness in the molecular shape. However, results obtained in this work, indicate that the molecular shape is probably not the only explanation.

It is worth to mention that the system size correction for the self-diffusion coefficient proposed by Yeh and Hummer [140] has not been considered in this work. Thus, results are expected to vary a few (~10%) as shown in the work of Moultos et al. [141]. However, overall, the SR-LJCCG model yields predictions in the *T* and *P* spaces and for various chain lengths, which could be considered as very satisfactory for the self-diffusion coefficient, confirming thus the good representability of the proposed force field.

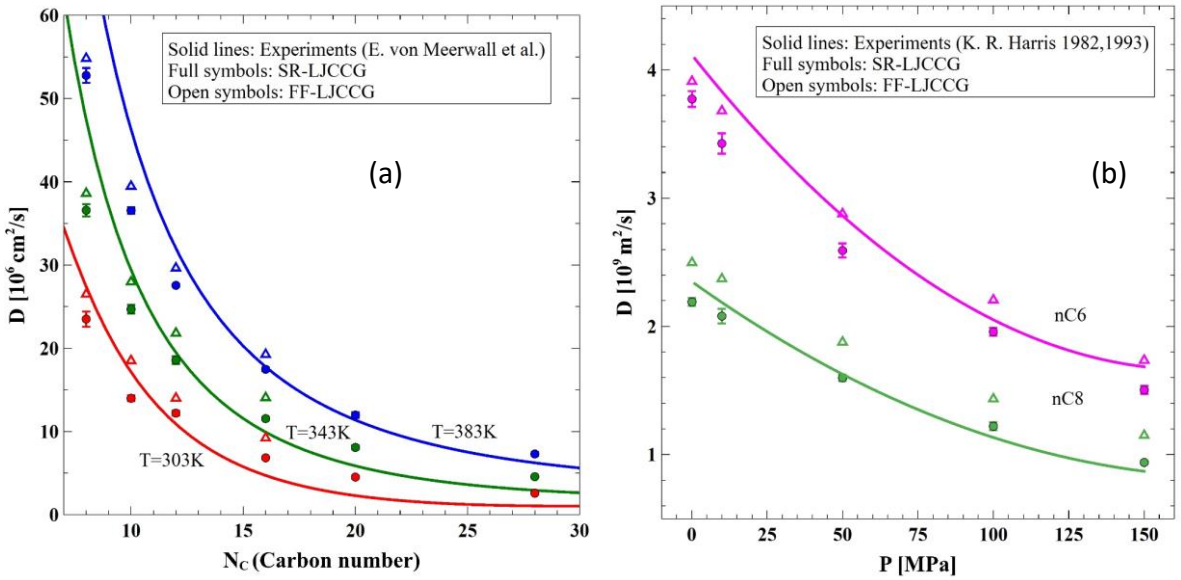


Figure 3.21: (a) Temperature and chain length effect for various n-alkanes components at P=1atm. (b) Pressure effect for nC6 and nC8 at T=298K.

3.4.6.5 Structural properties

As a final test, we have made a comparison between the Fully-Flexible and the Semi-Rigid LJCCG models regarding their respective predictions of structural properties. The first investigated property is the center of mass (COM) intermolecular radial distribution function (RDF) of two n-alkane molecules, the n-hexane (nC_6) and the n-nonane (nC_9). For both molecules, the RDF is computed in the liquid phase at reduced temperatures of T_r=0.75 and T_r=1.20 and at various densities as indicated in Figure 3.22.

Other properties related to the molecular conformations, such as the end-to-end distance (R_{EE}) and gyration radius (R_G) are also measured and are reported in Table 3.1 and Table 3.2. The tables indicate that for both metrics (R_{EE} and R_G), the fully flexible model is showing lower values compared to the semi-rigid model. This indicates that the rigidity tends to open the molecule, as expected. Note the values given in the table are the mean average values, as both metrics consist in distributions.

The increase in the thermal energy of the system lead to a decrease of the R_G and R_{EE} in the case of the semi-rigid model, as more non-linear configuration could be explored because of thermal motion. The behaviour is opposite in the case of the fully flexible model, because increasing the temperature allows to explore more configurations in which the molecules is les condensed (because of intramolecular interactions).

Regarding the RDF, for both molecules, and either in the subcritical dense liquid or in the supercritical fluid, the fully flexible model gives the same response, where it exhibits a single, but sharp pic expressing a well-structured fluid. However, in the case of the semi-rigid model, different responses are obtained depending on the thermodynamic conditions. In the dense liquid, the model exhibits a less structured fluid compared to the fully flexible model, with an attenuated first pic.

In fact, at low temperature, the semi rigid molecules are mostly "open" (rather linear) due to rigidity, whereas the fully flexible molecules are coiled up as pointed out previously. It is then this rigidity that perturbs the nearest molecules from being too close. The first solvation layer is then repelled to larger distances to appear in a small second pic. At high temperature, the semi-rigid model, exhibits practically the same response as for the fully flexible model, where a single sharper pic is obtained. This indicates that at high temperature, the effect of rigidity is reduced by the thermal effect, which is consistent with results obtained for the effect of rigidity on the viscosity shown in Figure 3.8.

At high distances, both models are converging towards a unity, but differently. The difference is more pronounced in the dense liquid state, where contrary to the fully flexible model, the semi-rigid model is exhibiting oscillations which is a typical response of liquids. Such a behaviour may have impact on properties deduced from the radial distribution function, such as the Kirkwood Buff Integrals.

It is interesting to mention that, the conditions at which we conducted the simulations for the RDF are exactly the same conditions studied by Rahman et al.[21] for the SAFT-γ-Mie model and the TraPPE-UA model. The results obtained with the SR-LJCCG model were found very close to those of the higher resolution model (the TraPPE-UA) similarly to the SAFT-γ-Mie. This indicates that, in addition to the improvements obtained on the transport properties when compared to the fully flexible model, the semi-rigid model provides also better performance in terms of structural properties predictions.

Radius of gyration (R_G)	$R_G^* = \frac{R_G}{\sigma}$ **SemiRigid nC_6**	$R_G^* = \frac{R_G}{\sigma}$ **FullyFlex nC_6**	$R_G^* = \frac{R_G}{\sigma}$ **SemiRigid nC_9**	$R_G^* = \frac{R_G}{\sigma}$ **FullyFlex nC_9**
Tr=0.7	1.282	1.191	1.557	1.343
Tr=1.20	1.265	1.194	1.352	1.525

Table 3.1: Radius of gyration for the fully flexible and semi-rigid n-hexane and n-nonane molecules at Tr=0.7 and Tr=1.20 and densities indicated on figure 3.22.

End-End distance (R_{EE})	$R^*_{EE} = \frac{R_{EE}}{\sigma}$ Semi-rigid nC_6	$R^*_{EE} = \frac{R_{EE}}{\sigma}$ Fully-flex nC_6	$R^*_{EE} = \frac{R_{EE}}{\sigma}$ Semi-rigid nC_9	$R^*_G = \frac{R_{EE}}{\sigma}$ Fully-flex nC_9
Tr=0.7	1.869	1.483	2.752	1.868
Tr=1.20	1.799	1.501	2.618	1.909

Table 3.2: End-to-End distance for the fully flexible and semi-rigid n-hexane and n-nonane molecules at Tr=0.7 and Tr=1.20 and densities indicated on figure 3.22.

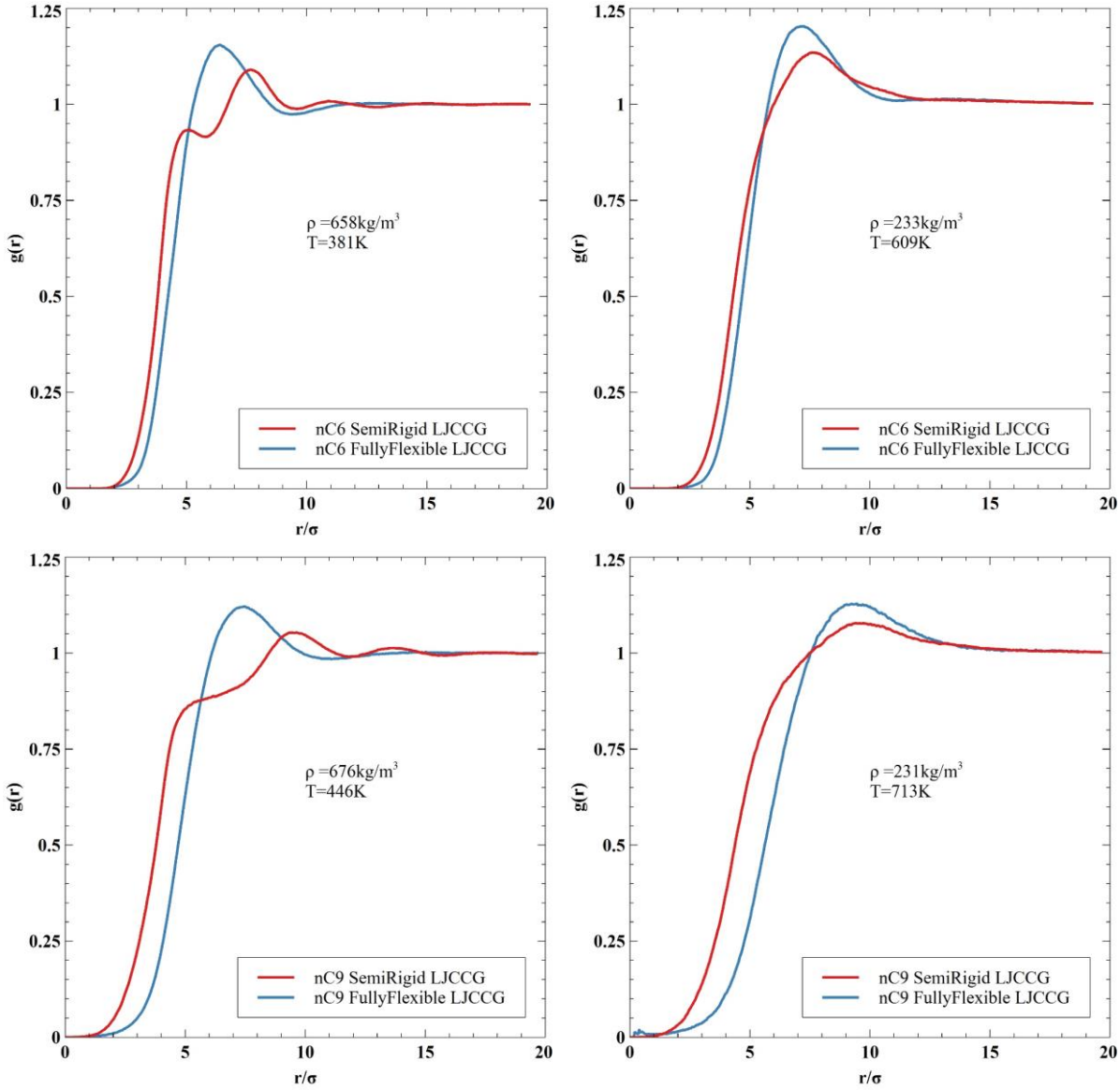


Figure 3.22: Center of mass intermolecular radial distribution function of n-hexane and n-nonane molecules in the subcritical liquid state (Tr=0.7) and supercritical fluid (Tr=1.20).

3.5 Summary

In this part we have proposed a semi-rigid coarse grained molecular referred to as SR-LJCCG model. The model contains only four molecular parameters which require, thanks to the corresponding strategy adopted in this work, only four macroscopic for their parametrization.

The model has been tested over many thermodynamic (phase equilibria and second order derivative properties), transport (viscosity and self-diffusion coefficient) and structural properties (radial distribution function), and found to yield excellent agreement with experiments. Moreover, the model has been tested on three different molecular families, including molecules with variable polarities which were all found fairly represented by the new model without any additional polar contribution. However, the case of strongly associative molecules is not considered, such as alcohols and water molecules. This is because the strategy we proposed in this work is not fully adequate to deal with such molecules, even though qualitative results can be achieved, the physical meaning of the parameter might be questionable. Additional efforts are needed as shown and discussed in the appendix 7.B.1 for methanol and water molecules.

The comparison has been systematically made against the reference model, the fully flexible model, and results showed that, while in terms of the thermodynamic properties, the models were found comparable, noticeable improvements were obtained with the SR-LJCCG model when considering transport properties, particularly the viscosity, and the structural properties as well.

As the viscosity was a target property for the development of this molecular model, further comparison was made by benchmarking the SR-LJCCG against many other force fields proposed in the literature. In this latter, were included high resolution models and united atoms models. The benchmark was made on the nC_{12} and nC_{16} at saturation conditions, and results showed that, only the SR-LJCCG model was among the best to predict correctly the viscosity over a wide range of temperature. This success of the SR-LJCCG model over other force fields is not only due to the rigidity parameter, but is also due to the robust parametrization strategy proposed in this work.

Elsewhere, to avoid performing extensive molecular simulations, an equation of state for the semi-rigid LJCCG model would be of great interest. However, for the time being, a SAFT like EoS including an internal rigidity term of the bending type is not yet available. This is the topic of the next chapter.

3.6 References

1. Sun, H. COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *J. Phys. Chem. B* **102**, 7338–7364 (1998).
2. MacKerell, A. D. Jr., Wiorkiewicz-Kuczera, J. & Karplus, M. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* **117**, 11946–11975 (1995).
3. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *ACS Publications* https://pubs.acs.org/ (2002).
4. Chen, B. & Siepmann, J. I. Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes. *J. Phys. Chem. B* **103**, 5370–5379 (1999).
5. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
6. Martin, M. G. & Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J. Phys. Chem. B* **102**, 2569–2577 (1998).
7. Jorgensen, W. L., Madura, J. D. & Swenson, C. J. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.* **106**, 6638–6646 (1984).
8. Nath, S. K., Escobedo, F. A. & de Pablo, J. J. On the simulation of vapor–liquid equilibria for alkanes. *J. Chem. Phys.* **108**, 9905–9911 (1998).
9. Padilla, P. & Toxvaerd, S. Structure and dynamical behavior of fluid n-alkanes. *J. Chem. Phys.* **95**, 509–519 (1991).
10. Nieto-Draghi, C., Ungerer, P. & Rousseau, B. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes: Improvement of transport properties. *J. Chem. Phys.* **125**, 044517 (2006).
11. Klein, M. & Shinoda, W. Large-Scale Molecular Dynamics Simulations of Self-Assembling Systems. *Science (New York, N.Y.)* **321**, 798–800 (2008).
12. Fayaz-Torshizi, M. & A. Müller, E. Coarse-grained molecular dynamics study of the self-assembly of polyphilic bolaamphiphiles using the SAFT-γ Mie force field. *Molecular Systems Design & Engineering* **6**, 594–608 (2021).
13. Nielsen, S. O., Lopez, C. F., Srinivas, G. & Klein, M. L. A coarse grain model for n-alkanes parameterized from surface tension data. *J. Chem. Phys.* **119**, 7043–7049 (2003).
14. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. (CRC Press, 2008).
15. Reith, D., Pütz, M. & Müller-Plathe, F. Deriving effective mesoscale potentials from atomistic simulations. *J Comput Chem* **24**, 1624–1636 (2003).
16. Moore, T. C., Iacovella, C. R. & McCabe, C. Derivation of coarse-grained potentials via multistate iterative Boltzmann inversion. *J. Chem. Phys.* **140**, 224104 (2014).
17. Shell, M. S. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *The Journal of Chemical Physics* **129**, 144108 (2008).
18. Hoang, H., Delage-Santacreu, S. & Galliero, G. Simultaneous Description of Equilibrium, Interfacial, and Transport Properties of Fluids Using a Mie Chain Coarse-Grained Force Field. *Ind. Eng. Chem. Res.* **56**, 9213–9226 (2017).
19. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
20. Pàmies, J. C. & Vega, L. F. Vapor−Liquid Equilibria and Critical Behavior of Heavy n-Alkanes Using Transferable Parameters from the Soft-SAFT Equation of State. *Ind. Eng. Chem. Res.* **40**, 2532–2543 (2001).

21. Rahman, S. *et al.* SAFT-γ Force Field for the Simulation of Molecular Fluids. 5. Hetero-Group Coarse-Grained Models of Linear Alkanes and the Importance of Intramolecular Interactions. *J. Phys. Chem. B* **122**, 9161–9177 (2018).
22. Brini, E. *et al.* Systematic coarse-graining methods for soft matter simulations – a review. *Soft Matter* **9**, 2108–2119 (2013).
23. Lyubartsev, A. P. & Laaksonen, A. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Phys. Rev. E* **52**, 3730–3737 (1995).
24. Milano, G., Goudeau, S. & Müller-Plathe, F. Multicentered Gaussian-based potentials for coarse-grained polymer simulations: Linking atomistic and mesoscopic scales. *Journal of Polymer Science Part B: Polymer Physics* **43**, 871–885 (2005).
25. Andersen, H. C., Chandler, D. & Weeks, J. D. Roles of Repulsive and Attractive Forces in Liquids : The Equilibrium Theory of Classical Fluids. in *Advances in Chemical Physics* 105–156 (John Wiley & Sons, Ltd, 1976).
26. Dunn, N. J. H. & Noid, W. G. Bottom-up coarse-grained models that accurately describe the structure, pressure, and compressibility of molecular liquids. *J. Chem. Phys.* **143**, 243148 (2015).
27. Rosenberger, D., Hanke, M. & van der Vegt, N. F. A. Comparison of iterative inverse coarse-graining methods. *Eur. Phys. J. Spec. Top.* **225**, 1323–1345 (2016).
28. Ganguly, P., Mukherji, D., Junghans, C. & van der Vegt, N. F. A. Kirkwood–Buff Coarse-Grained Force Fields for Aqueous Solutions. *J. Chem. Theory Comput.* **8**, 1802–1807 (2012).
29. Dallavalle, M. & Vegt, N. F. A. van der. Evaluation of mapping schemes for systematic coarse graining of higher alkanes. *Phys. Chem. Chem. Phys.* **19**, 23034–23042 (2017).
30. Papavasileiou, K. D., Peristeras, L. D., Bick, A. & Economou, I. G. Molecular Dynamics Simulation of Pure *n* -Alkanes and Their Mixtures at Elevated Temperatures Using Atomistic and Coarse-Grained Force Fields. *J. Phys. Chem. B* **123**, 6229–6243 (2019).
31. Klein, T. *et al.* Characterization of Long Linear and Branched Alkanes and Alcohols for Temperatures up to 573.15 K by Surface Light Scattering and Molecular Dynamics Simulations. *J. Phys. Chem. B* **124**, 4146–4163 (2020).
32. Izvekov, S., Parrinello, M., Burnham, C. J. & Voth, G. A. Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: a new method for force-matching. *J Chem Phys* **120**, 10896–10913 (2004).
33. Chaimovich, A. & Shell, M. S. Coarse-graining errors and numerical optimization using a relative entropy framework. *J. Chem. Phys.* **134**, 094112 (2011).
34. Rice, S. A. & Dinner, A. R. *Advances in Chemical Physics, Volume 161*. (John Wiley & Sons, 2016).
35. Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. SAFT: Equation-of-state solution model for associating fluids. *Fluid Phase Equilibria* **52**, 31–38 (1989).
36. Blas, F. J. & Vega, L. F. Prediction of Binary and Ternary Diagrams Using the Statistical Associating Fluid Theory (SAFT) Equation of State. *Ind. Eng. Chem. Res.* **37**, 660–674 (1998).
37. Lafitte, T., Bessieres, D., Piñeiro, M. M. & Daridon, J.-L. Simultaneous estimation of phase behavior and second-derivative properties using the statistical associating fluid theory with variable range approach. *J Chem Phys* **124**, 024509 (2006).
38. Lafitte, T. *et al.* Accurate statistical associating fluid theory for chain molecules formed from Mie segments. *J Chem Phys* **139**, 154504 (2013).
39. Mejía, A., Herdes, C. & Müller, E. A. Force Fields for Coarse-Grained Molecular Simulations from a Corresponding States Correlation. *Ind. Eng. Chem. Res.* **53**, 4131–4141 (2014).
40. Gross, J. & Sadowski, G. Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. *Ind. Eng. Chem. Res.* **40**, 1244–1260 (2001).

41. Wertheim, M. S. Fluids with highly directional attractive forces. I. Statistical thermodynamics. *J Stat Phys* **35**, 19–34 (1984).
42. Wertheim, M. S. Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations. *J Stat Phys* **35**, 35–47 (1984).
43. Maerzke, K. A. & Siepmann, J. I. Transferable Potentials for Phase Equilibria−Coarse-Grain Description for Linear Alkanes. *J. Phys. Chem. B* **115**, 3452–3465 (2011).
44. Ungerer, P. *et al.* Optimization of the anisotropic united atoms intermolecular potential for n-alkanes. *J. Chem. Phys.* **112**, 5499–5510 (2000).
45. Siu, S. W. I., Pluhackova, K. & Böckmann, R. A. Optimization of the OPLS-AA Force Field for Long Hydrocarbons. *J. Chem. Theory Comput.* **8**, 1459–1470 (2012).
46. Nielsen, S. O., Lopez, C. F., Srinivas, G. & Klein, M. L. A coarse grain model for n-alkanes parameterized from surface tension data. *J. Chem. Phys.* **119**, 7043–7049 (2003).
47. Periole, X. & Marrink, S.-J. The Martini coarse-grained force field. *Methods Mol Biol* **924**, 533–565 (2013).
48. Bourasseau, E., Ungerer, P., Boutin, A. & Fuchs, A. H. Monte Carlo simulation of branched alkanes and long chain n -alkanes with anisotropic united atoms intermolecular potential. *Molecular Simulation* **28**, 317–336 (2002).
49. Janeček, J. & Paricaud, P. A new optimization method for the determination of classical force fields. Application to the united atom force field of short alkanes. *Fluid Phase Equilibria* **429**, 27–36 (2016).
50. Moine, E., Piña-Martinez, A., Jaubert, J.-N., Sirjean, B. & Privat, R. I-PC-SAFT: An Industrialized Version of the Volume-Translated PC-SAFT Equation of State for Pure Components, Resulting from Experience Acquired All through the Years on the Parameterization of SAFT-Type and Cubic Models. *Ind. Eng. Chem. Res.* **58**, 20815–20827 (2019).
51. Messerly, R. A., Anderson, M. C., Razavi, S. M. & Elliott, J. R. Improvements and limitations of Mie λ-6 potential for prediction of saturated and compressed liquid viscosity. *Fluid Phase Equilibria* **483**, 101–115 (2019).
52. Ryckaert, J.-P. & Bellemans, A. Molecular dynamics of liquid n-butane near its boiling point. *Chemical Physics Letters* **30**, 123–125 (1975).
53. Toxvaerd, S. Molecular dynamics calculation of the equation of state of alkanes. *J. Chem. Phys.* **93**, 4290–4295 (1990).
54. Toxvaerd, S. Equation of state of alkanes II. *J. Chem. Phys.* **107**, 5197–5204 (1997).
55. Smit, B., Karaborni, S. & Siepmann, J. I. Computer simulations of vapor–liquid phase equilibria of n-alkanes. *J. Chem. Phys.* **102**, 2126–2140 (1995).
56. Smit, B., Karaborni, S. & Siepmann, J. I. Erratum: "Computer simulations of vapor–liquid phase equilibria of n-alkanes" [J. Chem. Phys. 102, 2126 (1995)]. *J. Chem. Phys.* **109**, 352–352 (1998).
57. ILJA SIEPMANN, B. J. & MARTIN, M. G. Intermolecular potentials for branched alkanes and the vapour-liquid phase equilibria of n-heptane, 2-methylhexane, and 3-ethylpentane. *Molecular Physics* **90**, 687–694 (1997).
58. Dysthe, D. K., Fuchs, A. H. & Rousseau, B. Fluid transport properties by equilibrium molecular dynamics. III. Evaluation of united atom interaction potential models for pure alkanes. *J. Chem. Phys.* **112**, 7581–7590 (2000).
59. Errington, J. R. & Panagiotopoulos, A. Z. A New Intermolecular Potential Model for the n-Alkane Homologous Series. *J. Phys. Chem. B* **103**, 6314–6322 (1999).
60. Zhang, H. & Ely, J. F. AUA model NEMD and EMD simulations of the shear viscosity of alkane and alcohol systems. *Fluid Phase Equilibria* **217**, 111–118 (2004).
61. Potoff, J. J. & Bernard-Brunel, D. A. Mie Potentials for Phase Equilibria Calculations: Application to Alkanes and Perfluoroalkanes. *J. Phys. Chem. B* **113**, 14725–14731 (2009).

62. Hemmen, A. & Gross, J. Transferable Anisotropic United-Atom Force Field Based on the Mie Potential for Phase Equilibrium Calculations: n-Alkanes and n-Olefins. *J. Phys. Chem. B* **119**, 11695–11707 (2015).
63. van Westen, T., Vlugt, T. J. H. & Gross, J. Determining Force Field Parameters Using a Physically Based Equation of State. *J. Phys. Chem. B* **115**, 7872–7880 (2011).
64. Shah, M. S., Siepmann, J. I. & Tsapatsis, M. Transferable potentials for phase equilibria. Improved united-atom description of ethane and ethylene. *AIChE Journal* **63**, 5098–5110 (2017).
65. Shelley, J. C., Shelley, M. Y., Reeder, R. C., Bandyopadhyay, S. & Klein, M. L. A Coarse Grain Model for Phospholipid Simulations. *J. Phys. Chem. B* **105**, 4464–4470 (2001).
66. Shinoda, W., DeVane, R. & Klein, M. L. Multi-property fitting and parameterization of a coarse grained model for aqueous surfactants. *Molecular Simulation* **33**, 27–36 (2007).
67. An, Y., Bejagam, K. K. & Deshmukh, S. A. Development of New Transferable Coarse-Grained Models of Hydrocarbons. *J. Phys. Chem. B* **122**, 7143–7153 (2018).
68. Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **29**, 1709–1721 (1990).
69. Huang, S. H. & Radosz, M. Equation of state for small, large, polydisperse, and associating molecules. *ACS Publications* https://pubs.acs.org/(2002).
70. Johnson, J. K., Mueller, E. A. & Gubbins, K. E. Equation of State for Lennard-Jones Chains. *J. Phys. Chem.* **98**, 6413–6419 (1994).
71. FELIPE J. BLAS and LOURDES F. VEGA. Thermodynamic behaviour of homonuclear and heteronuclear Lennard-Jones chains with association sites from simulation and theory. *Molecular Physics* **92**, 135–150 (1997).
72. Gil-Villegas, A. *et al.* Statistical associating fluid theory for chain molecules with attractive potentials of variable range. *The Journal of Chemical Physics* **106**, 4168–4186 (1997).
73. Avendaño, C. *et al.* SAFT-γ Force Field for the Simulation of Molecular Fluids. 1. A Single-Site Coarse Grained Model of Carbon Dioxide. *J. Phys. Chem. B* **115**, 11154–11169 (2011).
74. Avendaño, C. *et al.* SAFT-γ Force Field for the Simulation of Molecular Fluids: 2. Coarse-Grained Models of Greenhouse Gases, Refrigerants, and Long Alkanes. *J. Phys. Chem. B* **117**, 2717–2733 (2013).
75. Lafitte, T. *et al.* SAFT- *γ* force field for the simulation of molecular fluids: 3. Coarse-grained models of benzene and hetero-group models of *n* -decylbenzene. *Molecular Physics* **110**, 1189–1203 (2012).
76. Lobanova, O., Avendaño, C., Lafitte, T., Müller, E. A. & Jackson, G. SAFT-γ force field for the simulation of molecular fluids: 4. A single-site coarse-grained model of water applicable over a wide temperature range. *Molecular Physics* **113**, 1228–1249 (2015).
77. Lobanova, O., Mejía, A., Jackson, G. & Müller, E. A. SAFT-γ force field for the simulation of molecular fluids 6: Binary and ternary mixtures comprising water, carbon dioxide, and n-alkanes. *The Journal of Chemical Thermodynamics* **93**, 320–336 (2016).
78. Galliero, G. Equilibrium, interfacial and transport properties of n-alkanes: Towards the simplest coarse grained molecular model. *Chemical Engineering Research and Design* **92**, 3031–3037 (2014).
79. Galliero, G., Lafitte, T., Bessieres, D. & Boned, C. Thermodynamic properties of the Mie n-6 fluid: A comparison between statistical associating fluid theory of variable range approach and molecular dynamics results. *J. Chem. Phys.* **127**, 184506 (2007).
80. Galliero, G. Surface tension of short flexible Lennard-Jones chains: Corresponding states behavior. *The Journal of chemical physics* **133**, 074705 (2010).
81. Galliero, G., Boned, C. & Fernández, J. Scaling of the viscosity of the Lennard-Jones chain fluid model, argon, and some normal alkanes. *The Journal of chemical physics* **134**, 064505 (2011).

82. Blas, F. J., Ignacio Moreno-Ventas Bravo, A., Míguez, J. M., Piñeiro, M. M. & MacDowell, L. G. Vapor-liquid interfacial properties of rigid-linear Lennard-Jones chains. *J. Chem. Phys.* **137**, 084706 (2012).
83. Blas, F. J., Bravo, A. M.-V., Algaba, J., Martínez-Ruiz, F. J. & MacDowell, L. G. Effect of molecular flexibility of Lennard-Jones chains on vapor-liquid interfacial properties. *The Journal of Chemical Physics* **140**, 114705 (2014).
84. Panagiotopoulos, A. Z. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Molecular Physics* **61**, 813–826 (1987).
85. Panagiotopoulos, A. Z., Quirke, N., Stapleton, M. & Tildesley, D. J. Phase equilibria by simulation in the Gibbs ensemble: alternative derivation, generalization and application to mixture and membrane equilibria. *Molecular Physics* **63**, 527–545 (1988).
86. Martin, M. G. MCCCS Towhee: a tool for Monte Carlo molecular simulation. *Molecular Simulation* **39**, 1212–1222 (2013).
87. Siepmann, J. I. & Frenkel, D. Configurational bias Monte Carlo: a new sampling scheme for flexible chains. (1992).
88. Panagiotopoulos, A. Z. Monte Carlo methods for phase equilibria of fluids. *Journal of Physics: Condensed Matter* **12**, R25 (2000).
89. Hamani, A. W. S. *et al.* Thermophysical properties of simple molecular liquid mixtures: on the limitations of some force fields. *Journal of Molecular Liquids* **303**, 112663 (2020).
90. The, N. *et al.* Entropy Scaling for Viscosity of Pure Lennard-Jones Fluids and Their Binary Mixtures. *Communications in Physics* **32**, (2022).
91. Bordat, P. & Müller-Plathe, F. The shear viscosity of molecular fluids: A calculation by reverse nonequilibrium molecular dynamics. *Journal of Chemical Physics, v.116, 3362-3369 (2002)* **116**, (2002).
92. Delage Santacreu, S., Hoang, H., Khennache, S. & Galliero, G. Thermodynamic Scaling of the Shear Viscosity of Lennard-Jones Chains of Variable Rigidity. *Liquids* **1**, 96–108 (2021).
93. Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics* **52**, 24–34 (1983).
94. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
95. Understanding Molecular Simulation - 2nd Edition. https://www.elsevier.com/books/understanding-molecular-simulation/frenkel/978-0-12-267351-1.
96. Ramírez-González, P. V. & Escobar-Barrios, V. A. Viscosity and normal stress forces of Lennard-Jones chains using reverse non-equilibrium molecular dynamics. *Molecular Physics* **115**, 2970–2977 (2017).
97. Kröger, M. Simple models for complex nonequilibrium fluids. *Physics Reports* **390**, 453–551 (2004).
98. Carreau, P. J. Rheological Equations from Molecular Network Theories. *Transactions of the Society of Rheology* **16**, 99–127 (1972).
99. Galindo, A. *et al.* Computer simulation study of the global phase behavior of linear rigid Lennard-Jones chain molecules: Comparison with flexible models. *J. Chem. Phys.* **120**, 3957–3968 (2004).
100. Sheng, Y.-J., Panagiotopoulos, A. Z. & Kumar, S. K. Effect of Chain Stiffness on Polymer Phase Behavior. *Macromolecules* **29**, 4444–4446 (1996).
101. Laghaei, R. & Nasrabad, A. E. The Influence of Bond Angle on Thermophysical Properties of Three-Center Lennard-Jones Fluids: Computer Simulation and Theory. *Zeitschrift für Physikalische Chemie* **233**, 551–576 (2019).

102. Moller, E. A. & Gubbins, K. E. Simulation of hard triatomic and tetratomic molecules A test of associating fluid theories. 20.
103. van Westen, T., Vlugt, T. J. H. & Gross, J. On the vapor-liquid equilibrium of attractive chain fluids with variable degree of molecular flexibility. *J. Chem. Phys.* **142**, 224504 (2015).
104. Vega, C., McBride, C. & MacDowell, L. G. The effect of flexibility on the phase diagram of simple molecular models. *Phys. Chem. Chem. Phys.* **4**, 853–862 (2002).
105. Oyarzún, B., van Westen, T. & Vlugt, T. J. H. Isotropic-nematic phase equilibria of hard-sphere chain fluids—Pure components and binary mixtures. *J. Chem. Phys.* **142**, 064903 (2015).
106. Delage Santacreu, S., Galliero, G., Odunlami, M. & Boned, C. Low density shear viscosity of Lennard-Jones chains of variable rigidities. *Journal of Chemical Physics* **137**, 204306 (2012).
107. Jaeger, F., Matar, O. K. & Müller, E. A. Bulk viscosity of molecular fluids. *J. Chem. Phys.* **148**, 174504 (2018).
108. Pitzer, K. S., Lippmann, D. Z., Jr, R. F. C., Huggins, C. M. & Petersen, D. E. The Volumetric and Thermodynamic Properties of Fluids. II. Compressibility Factor, Vapor Pressure and Entropy of Vaporization1. *ACS Publications* https://pubs.acs.org/.
109. Galliero, G. & Boned, C. Shear viscosity of the Lennard-Jones chain fluid in its gaseous, supercritical, and liquid states. *Physical Review E : Statistical, Nonlinear, and Soft Matter Physics* **79**, 1–9 (2009).
110. Lemmon, E. W., Bell, I. H., Huber, M. L. & McLinden, M. O. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0, National Institute of Standards and Technology. *Standard Reference Data Program, Gaithersburg* (2018).
111. DIPPR | Design Institute for Physical Properties. https://www.aiche.org/dippr (2019).
112. DDB Search - DDBST GmbH. http://www.ddbst.com/ddb-search.html.
113. DECHEMA. Information Systems and Databases. https://i-systems.dechema.de/detherm/mixture.php (2013).
114. Westhaus, U., Dröge, T. & Sass, R. DETHERM®—a thermophysical property database. *Fluid Phase Equilibria* **158–160**, 429–435 (1999).
115. Frenkel, M. *et al.* ThermoData Engine (TDE): Software Implementation of the Dynamic Data Evaluation Concept. *J. Chem. Inf. Model.* **45**, 816–838 (2005).
116. Material Properties. https://www.engineeringtoolbox.com/material-properties-t_24.html.
117. Cheméo. *Cheméo* https://www.chemeo.com/.
118. Poling, B. E., Prausnitz, J. M. & O'Connell, J. P. *The properties of gases and liquids*. (McGraw-Hill, 2001).
119. Yaws, C. L. *Thermophysical Properties of Chemicals and Hydrocarbons*. (William Andrew, 2008).
120. Guggenheim, E. A. The Principle of Corresponding States. *J. Chem. Phys.* **13**, 253–261 (1945).
121. Bell, I. H., Fingerhut, R., Vrabec, J. & Costigliola, L. Connecting entropy scaling and density scaling. *J. Chem. Phys.* **157**, 074501 (2022).
122. Dehlouz, A., Privat, R., Galliero, G., Bonnissel, M. & Jaubert, J.-N. Revisiting the Entropy-Scaling Concept for Shear-Viscosity Estimation from Cubic and SAFT Equations of State: Application to Pure Fluids in Gas, Liquid and Supercritical States. *Ind. Eng. Chem. Res.* **60**, 12719–12739 (2021).
123. Jackson, G. & Gubbins, K. E. Mixtures of associating spherical and chain molecules. *Pure and Applied Chemistry* **61**, 1021–1026 (1989).
124. Oliveira, M. B., Llovell, F., Coutinho, J. A. P. & Vega, L. F. New Procedure for Enhancing the Transferability of Statistical Associating Fluid Theory (SAFT) Molecular Parameters: The Role of Derivative Properties. 14.
125. Llovell, F. & Vega, L. F. Prediction of Thermodynamic Derivative Properties of Pure Fluids through the Soft-SAFT Equation of State. *J. Phys. Chem. B* **110**, 11427–11437 (2006).

126. Lagache, M., Ungerer, P., Boutin, A. & Fuchs, A. H. Prediction of thermodynamic derivative properties of fluids by Monte Carlo simulation. *Phys. Chem. Chem. Phys.* **3**, 4333–4339 (2001).
127. The Mathematical Theory of Non-uniform Gases: An Account Of The Kinetic Theory Of Viscosity, Thermal Conduction And Diffusion In Gases - AbeBooks - Chapman, Sydney: 052140844X. https://www.abebooks.fr/9780521408448/Mathematical-Theory-Non-uniform-Gases-Account-052140844X/plp.
128. *Experimental Thermodynamics Volume IX*. (2014).
129. Baled, H. O., Gamwo, I. K., Enick, R. M. & McHugh, M. A. Viscosity models for pure hydrocarbons at extreme conditions: A review and comparative study. *Fuel* **218**, 89–111 (2018).
130. Kontogeorgis, G. M. *et al.* Industrial Requirements for Thermodynamic and Transport Properties: 2020. *Ind. Eng. Chem. Res.* **60**, 4987–5013 (2021).
131. Ewen, J. P. *et al.* A Comparison of Classical Force-Fields for Molecular Dynamics Simulations of Lubricants. *Materials* **9**, 651 (2016).
132. Payal, R. S. *et al.* Shear viscosity of linear alkanes through molecular simulations: quantitative tests for n-decane and n-hexadecane. *Molecular Simulation* **38**, 1234–1241 (2012).
133. Allen, W. & Rowley, R. L. Predicting the viscosity of alkanes using nonequilibrium molecular dynamics: Evaluation of intermolecular potential models. *J. Chem. Phys.* **106**, 10273–10281 (1997).
134. Dickson, C. J. *et al.* Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **10**, 865–879 (2014).
135. Smit, B., Karaborni, S. & Siepmann, J. I. Computer simulations of vapor–liquid phase equilibria of n-alkanes. *J. Chem. Phys.* **102**, 2126–2140 (1995).
136. Mondello, M. & Grest, G. S. Viscosity calculations of n-alkanes by equilibrium molecular dynamics. *J. Chem. Phys.* **106**, 9327–9336 (1997).
137. von Meerwall, E., Beckman, S., Jang, J. & Mattice, W. L. Diffusion of liquid n-alkanes: Free-volume and density effects. *J. Chem. Phys.* **108**, 4299–4304 (1998).
138. Harris, K. R. Temperature and density dependence of the self-diffusion coefficient of n-hexane from 223 to 333 K and up to 400 MPa. *J. Chem. Soc., Faraday Trans. 1* **78**, 2265–2274 (1982).
139. Harris, K. R. *et al.* Temperature and density dependence of the selfdiffusion coefficients of liquid n-octane and toluene. *Molecular Physics* **78**, 235–248 (1993).
140. Yeh, I.-C. & Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **108**, 15873–15879 (2004).
141. Moultos, O. A., Zhang, Y., Tsimpanogiannis, I. N., Economou, I. G. & Maginn, E. J. System-size corrections for self-diffusion coefficients calculated from molecular dynamics simulations: The case of CO2, n-alkanes, and poly(ethylene glycol) dimethyl ethers. *J Chem Phys* **145**, 074109 (2016).

Chapter 4. Integration of a bending potential into SAFT model

Contents

4.1 Introduction 93
4.2 Wertheim perturbation theory: a global overview 95
4.2.1 Pair potential 95
4.2.2 Assumptions and Steric Incompatibilities (SI) 97
4.2.3 Multiple density formalism 100
4.2.4 Derived Equations 101
4.3 SAFT models 111
4.3.1 SAFT-HS and CK-SAFT 111
4.3.2 LJ-SAFT (soft-SAFT) 112
4.3.3 SAFT-VR 113
4.3.4 PC-SAFT 115
4.4 Effect of the approximations on the chain bond formation term 117
4.5 A simple alternative to higher orders of the TPTn … 122
4.6 Does Wertheim TPT includes hard intramolecular interactions (vibrations)? 129
4.7 The model 132
4.8 Model validation 134
4.8.1 Single phase 134
4.8.2 Liquid vapour equilibrium (LVE) 136
4.8.3 Internal energy & Entropy at saturation 138
4.9 Real fluid properties 140
4.9.1 Optimization of the molecular parameters 140
4.9.2 Results and discussions 144
4.9.3 Consistency check and transferability of the parameters 155
4.10 Summary 161
4.11 References 163

4.1 Introduction

Mimicking the behaviour of simple fluids and polar fluids is still at the heart of the development of thermodynamics fluid theories. Simple models proposed at in the 1950s were found to be good for modelling the behaviour of simple fluids, such as the use of Hard sphere fluid model with some perturbation representing the attractive forces. However, extrapolating this idea to model polar fluids [1, 2] by considering higher order terms of the Helmholtz free energy only gave satisfactory results for weak polar components such as N_2, O_2. Systematic deviations with increasing polarity were obtained, and a multibody potential was suggested. This is due to the fact that the structure of simple fluids is not the same as that of polar fluids which exhibit hydrogen bonds. These interactions are very strong, act only at very short distances and are highly directional. These limitations of simple spherical and symmetrical interaction potentials have prompted researchers to develop new approaches to handle the specific complex interactions present in many substances such as water.

The thermodynamics perturbation theory (TPT) was developed by Wertheim in a series of papers [3, 4, 5, 6, 7, 8], and is considered as one of the most powerful existing theory that can model simple fluids and fluids with highly directional interactions such as hydrogen bonds. Wertheim was inspired by the work of Anderson on dilute gases [9] and dense liquids [10], who initiated the theory of association by proposing a molecular model composed of two interaction parts: a symmetrical hard sphere reference model, and a directional short ranged interactions modelled by a square well potential to mimic the effects of hydrogen bonds. The attractive hydrogen bonds were represented by a site at the edge of the hard cores. Anderson's approach is based on the single density graph Mayer's expansion [11]. However, this approach is known to be an infinite sum and can produce a catastrophe in some situations due to the large value of the Mayer function at low temperature when the ratio of bonding energy over thermal energy is highly negative, leading to an infinite Boltzmann factor. Anderson handles the first problem by introducing a steric incompatibility in which only one bond per site is permitted, allowing the use of the cancellation theorem. The second problem can be handled by working with renormalized Mayer function.

The Anderson approach was the first based on cluster graph expansion that allowed for the inclusion of directional interactions in a systematic way. However, the cancellations introduced had only a weak effect on the density expansion, and the approach was complicated to apply. Chandler and Pratt [12], instead of using density expansion, used the fugacity expansion by considering only physical clusters similarly to the Anderson's steric incompatibility. Based on the two previous works, Høye and Olaussen [13] showed that the fugacity expansion is simpler, more convenient and converge more rapidly than the density expansion. They also introduced the concept of multiple densities instead of singlet density ρ. Later, Wertheim, built upon these previous works and proposed a rather simple (compared to the previous ones) and convenient theory for hard sphere fluids with highly directional attractive interactions, called the thermodynamic perturbation theory (TPT). The TPT theory relies on the fugacity expansion introduced by Chandler and Pratt and integrates the concept of the two density theory by Høye and Olaussen. It considers simultaneously the presence of non-associating monomers characterized by a density ρ_0 and the presence of dimers in the system formed by the bonding

of two associated monomers, characterized by a density ρ_1. Furthermore, the theory builds on Anderson's concept of steric incompatibility and introduces more types of steric incompatibility, leading to many graph simplifications and cancellations.

In his first two papers [3, 4], Wertheim formulated his theory for associating fluids. The fluid model consists in hard spheres with a single association site at their edges. When the spheres are close to each other and in good orientation, the short ranged interactions are felt by the two sites giving them the possibility of being irreversibly dimerized. If this occurs for all the spheres in the system, this leads to a system of hard-dumbbells fluid. Between the two limits (only spheres or dimers), the system may exist as a mixture of hard spheres with single attractive sites and inseparable hard-dumbbells without any attractive interactions. Wertheim then derived an equation of state that is capable to model the behaviour of this mixture as well as the two limits by introducing a composition parameter. Comparisons against Monte Carlo simulations of the equation derived from TPT theory for the case of total system dimerization yielded excellent agreement [7]. The theory was extended to consider the possibility of multiple sites, where formation of chains with a number of spheres greater than two beads (dimers) [5, 6] is possible. An initial system consisting of hard spheres with two association sites in each sphere could evolve to form a mixture of chains with different chain lengths at equilibrium. Wertheim then derived two equations of states, one for the first-order TPT (TPT1) which works perfectly for the single site model and one for the second-order TPT (TPT2) to better model systems formed with two sites model.

In his last paper on TPT, Wertheim relaxed the constrain on the angle formed by the vectors between the hard core center and the position of the two sites. In fact, in his first work, this angle was fixed, corresponding to a formation of rigid chains. In this updated version, the sites are no more rigid, and are allowed to freely move on the surface of the hard core, but, in the limit where the angle is greater than $\frac{\pi}{3}$. Too small angles are not allowed as this is prohibited by some steric incompatibilities which will be discussed later.

Using TPT1, Wertheim compared his equation of state to molecular simulation data on the fully flexible hard chains of Dickman and Hall [14]**.** It is worth to mention that Dickman and Hall performed simulation with system of chains of identical chain length, and the comparison to Wertheim TPT was done after identifying the average chain length of the TPT to the finite fixed number of spheres forming the chain in [14]. Results for m=4 and m=8 were in very good agreement, but deviations were found when increasing the chain length, especially for m=16. Deviations were probably due to the fact that the interactions are only defined at the monomers level. In other words, the spheres forming the chain are not "aware" of the existence of the next-nearest sphere (no information between the relative site-site orientation) and so on. This is because, at TPT1 level, only pair correlation function is defined, and the introduction of the structural information in the theory needs, at least, three body correlation function. Wertheim then proposed in the same paper a version of TPT2, which includes the triplet distribution function for the three center-to-center distances (rather in terms of the bond angle). Comparison to simulation data of the fully flexible chains clearly showed that TPT2 yields better agreement compared to TPT1 as the chain length increases.

4.2 Wertheim perturbation theory: a global overview

In the following, a brief overview of the TPT of Wertheim will be provided. The theory is very complex, and requires strong knowledge on molecular physics as well as mathematics and graph theory as well. The following summary will only give some insights on the theory, inspired mostly from the original papers of Wertheim [3, 4, 5, 6, 7, 8], and also from some interesting text books [15, 16, 17], and from more recent thesis works of Marshall [18], Febra etal. [19] and that of Zimptas et al. [20].

4.2.1 Pair potential

The pair potential is defined by:

$$\emptyset(12) = \emptyset_R(12) + \sum_A \sum_B \emptyset_{AB}(|r_2 + d_B(\Omega_2) - r_1 - d_A(\Omega_1)|) \tag{4.1}$$

where $i = 1,2$ ... refers to the position r_i of the molecular center of mass and Ω_i to the molecular orientation. $j = A, B$ ... are the sites on molecules (1), (2) ...and d_j is the vector from the center of mass i to the center of site j. $\emptyset_{AB}$ is the short ranged attractive potential between site (A) of molecule (1) and site (B) of molecule (2). It is important to note that the interaction between the same site types is not allowed, leading to $\emptyset_{AA} = \emptyset_{BB} = 0$. $\emptyset_R$ is the short ranged repulsive interaction between hard molecules (1) and (2) ...

The repulsive potential takes the following form:

$$\emptyset_R(x) = \begin{cases} \infty, & if\ |x| < \sigma \\ 0, & if\ |x| \geq \sigma \end{cases} \tag{4.2}$$

where σ is the collisional hard sphere diameter.

The attractive potential is given by:

$$\emptyset_{AB}(x) = \begin{cases} < 0, & if\ |x| < a \\ 0, & if\ |x| \geq a \end{cases} \tag{4.3}$$

where a is the site diameter (= range of interaction)($<< \sigma$), and x the site-site distance:

$$x = |r_2 + d(\Omega_2) - r_1 - d(\Omega_1)| \tag{4.4}$$

note that d must satisfy:

$$\frac{1}{2}(\sigma - a) < d < \frac{1}{2}\sigma \tag{4.5}$$

For any potential, one may define the e-function and the Mayer f-function:

$$e(12) = Exp\big(-\beta\emptyset(12)\big) = Exp[-\beta(\emptyset_R(12) + \sum_A \sum_B \emptyset_{AB}(12))] \tag{4.6}$$

$$f(12) = e(12) - 1 \tag{4.7}$$

where $\beta = \frac{1}{k_B T}$ and k_B is the Boltzmann's constant and T is the temperature in K.

Using the product property of the exponential function, the two last equations give respectively the two following equations:

$$e(12) = e_R(12) \prod_{A\in\mathbb{I}} \prod_{B\in\mathbb{I}} e_{AB}(12) \tag{4.8}$$

$$f(12) = f_R(12) + e_R(12)(\prod_{A\in\mathbb{I}} \prod_{B\in\mathbb{I}}[1 + f_{AB}(12)] - 1) \tag{4.9}$$

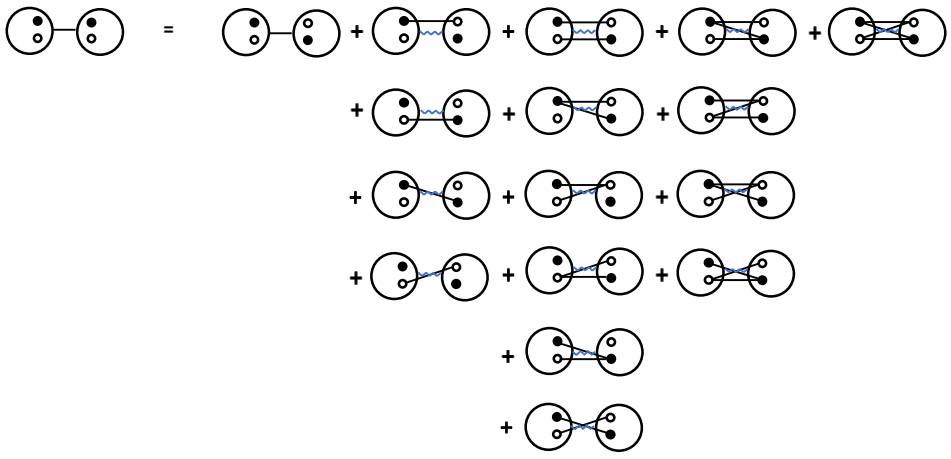
where

$$f_R(12) = \begin{cases} -1, & if\ |x| < \sigma \\ 0, & if\ |x| \geq \sigma \end{cases} \tag{4.10}$$

$$e_R(12) = \begin{cases} 0, & if\ |x| < \sigma \\ 1, & if\ |x| \geq \sigma \end{cases} \tag{4.11}$$

The $e_R(12)$ function was introduced take into account for steric incompatibilities when the rigid cores overlap, and $f_{AB}(12)$ accounts for the directional short ranged attractions. In the graph theory, the e-bonds will be represented by straight lines and the f-bonds by zigzag lines. The f_{AB}-bonds connects only two attractive sites, while f-bonds connects only hard cores.

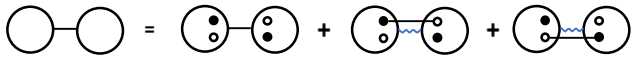
For illustration, let's take the same example as the one given in Wertheim's third paper [5]. Assuming the case of the two-site model, with identical particles {1, 2} possessing sites A (white) and B (black) each. All possible interactions between the two particles, expressed in terms of the Mayer function using the graph theory, are represented by the following graph:



The first simplification come from the assumption made in the potential interaction, where all interaction between the same type sites vanished as $f_{AA}(12) = f_{BB}(12) = 0$, due to the fact that $\emptyset_{AA} = \emptyset_{BB} = 0$. Thus, graph reduces to this:

Furthermore, Wertheim assumed that the double bonding is not allowed, as this will result in two overlapped cores because of the very short range of the attractive interactions. Mathematically, this vanishes due to the irrelevant value of the pre-factor $e_R(12)$. This is one of the steric incompatibilities introduced later by Wertheim which will be discussed just after. This leads then to the final f_function:

$$f(12) = f_R(12) + e_R(12)f_{AB}(12) + e_R(12)f_{BA}(12) \tag{4.12}$$



This simple example emphasizes the importance of the assumptions introduced by Wertheim which lead to many simplifications in the graph theory. In the same spirit of simplifications, as already mentioned, Wertheim introduced further approximations such as the single-chain approximation and other steric incompatibilities, leading to further simplifications in the graph theory.

4.2.2 Assumptions and Steric Incompatibilities (SI)

Among the keys to the success of the Wertheim theory, is the introduction of some assumptions and steric incompatibilities at an early stage. This allowed him to drastically reduce the number of clusters in the *z*-expansion, making this latter sum convergent. The first assumption is the prohibition made on the possible association between the same type sites of different molecules ($\emptyset_{AA} = \emptyset_{BB} = 0$). The second assumption is to account only graphs involving a single chain surrounded by a sea of monomers. This last assumption is called the single-chain approximation (or the ideal-chain approximation). With this approximation, there is no information on the simultaneous presence of at least two chains (dimers, trimers…) in the system. The interactions are then only defined at the monomeric level and not at the chain level.

Apart the fact that the aim is to simplify the theory to be easily applicable, there is no physical arguments to justify these assumptions. In reality, attraction may occur between any two distinct attractive sites (first assumption). Moreover, associated molecules could exist simultaneously in the system (second assumption). However, behind any theory development, there is always a trade-off between the degree of complexity and the degree of accuracy. For instance, without the single-chain approximation in TPT, a dimer-dimer interaction should be taken into account (note that this is the simplest case), which requires a four-body distribution function for which little is known. The effect of this approximation will be felt in extreme dilute system, especially for large chain molecules. The virial coefficients are then badly approximated, as well as phase equilibria, which requires the knowledge of accurate pressure and chemical potential of the vapour phase.

In addition to the two assumptions discussed above, some sensible approximations were introduced by Wertheim, called Steric Incompatibilities (SI). They are of three types, noted respectively SI1, SI2 and SI3. Their introduction allows much more simplifications and cancellations in the graph theory. Without them, the fugacity expansion used by Wertheim would be an infinite summation from a mathematical point of view (infinite possible configurations of clusters). Taking into account these SI from the outset is one of the keys to the success of the Wertheim TPT theory. Let's now introduce these three types of steric incompatibilities:

- **SI1**

The first steric incompatibility occurs when three molecules, with a single site each, are involved, as shown in Figure 4.1. It is taken from the Anderson's work [9] in which a third site {C} of molecule (3) could not be bonded to any of the already bonded sites {A, B} of molecules {1, 2}. In fact, this could happen if only site {C} in molecule (3) enters in the range of attraction of either site {A} or {B}. However, as the range of the attractions are very short, molecule (3) needs then to overlap molecules (1) and/or molecule (2) which is not possible because of the strong repulsion force induced by the hard-cores when overlapping, thus resulting in $e_R(13) = e_R(23) = 0$ and leading to a vanishing graph.

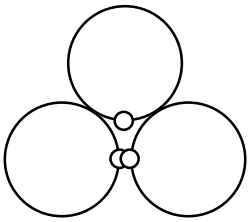


Figure 4.1: Wertheim's SI1, where two bonded hard cores prevent a third hard core from association to one of them due to the impossible overlapping between cores.

- **SI2**

This type of incompatibility occurs in a multiple sites model when only pairs of molecules are involved. This type of SI is physically justified by the fact that the angle formed between the two sites and the center of mass of the molecule is set to be sufficiently large, i.e. $\theta_{min} > \frac{\pi}{3}$, as too small angles can be considered as unrealistic. Thus, a bond between two molecules can only involve a maximum of one site from each molecules as the third site from either molecule *1* or *2* is situated at distance far enough from the bonding range. This limits the number of possible bonds per site, and is referred to as a weak SI, noted SI2W, as shown in Figure 4.2 (a).

Together with SI1, this results in SI2S, where S stands for strong, where multiple bonding between two molecules are excluded. A double bonding requires a high enough energy to be formed as this may happen only if the two cores overlap (SI1) due to the large enough angle between their sites and their centers of mass respectively (SI2W), as illustrated in Figure 4.2 (b).

Figure 4.2: (a) Wertheim's SI2W, where a site from one molecule can only be bonded to one site from another molecule. (b) Wertheim's SI2S, where double bonding is not allowed as a consequence of SI1 and SI2W.

- **SI3**

Consider the case of a chain formed by the association of spheres ($s-mers \geq 3$) in the multiple sites model. In the case of rigid linear or almost linear chains $\theta_{min} \approx \pi$, it is reasonable to consider that two spheres (with one free site each) of the same chain molecule, separated by at least one sphere, makes the distance between their relative free attractive sites larger than the site's attractive range, as shown in Figure 4.3. Hence, the association at one site can be considered independent of the association of the other sites within the same molecule. Thus, the chain self-hindrance occurs as the structure of the molecule is considered rigid. However, this can be true only if the range of attraction is very small as Wertheim wrote in the third paper "The extent to which it occurs depends critically on the range of the attractive interactions and the bond angles. For this reason, it is harder to incorporate in approximation theories than the other two types of SI".

Hence, the self-association, to permit ring formation, is then discarded with the use of SI3. However, in the case of a fully flexible chains (freely rotating), one may imagine that SI3 becomes less arguable! This is the reason for which Wertheim was somewhat hesitating to truncate on this SI3 type ("The problem of systematic exploitation of SI3 is a difficult one" [6]), and why he limited the application to very small chains. "However, the closest match between equilibrium TPT and the ringless DH model is obtained by again excluding physical rings via the simple means of dropping ring graphs" after Wertheim in [8].

It is worthwhile to note that most of the SI that simplifies the development of the Wertheim's TPT theory are a consequence of choosing the hard and spherically symmetrical model as a reference term (Wertheim's ingenuity!). Some of the SI would not be fully justified if a soft sphere was taken as a reference model, which would have made the development of the theory more complex.

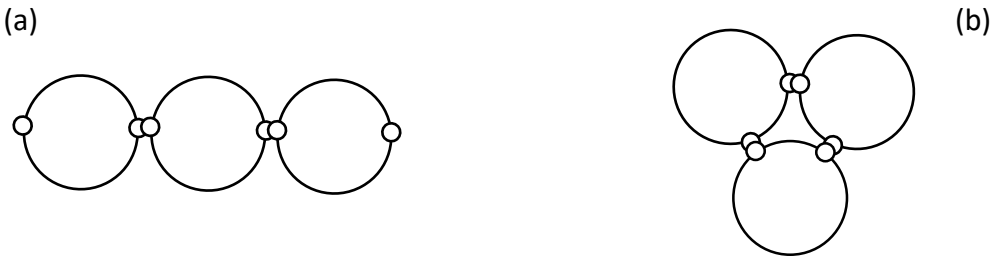


Figure 4.3: Wertheim's SI3, where, for small molecules, the angle between the two site and the center of mass of the molecule is large enough , so that the molecules can be considered as rigid and linear (a). (b) Another consequence is that rings are not allowed to form for small molecules.

4.2.3 Multiple density formalism

As already mentioned, the multiple density formalism is one of the keys of the success of the Wertheim. Unlike Anderson's work, where the overall density was used, Wertheim used another approach treating the molecules as different species in the system depending on their bonding state. Wertheim showed that, while the correct behaviour of a dimerizing gas in the low density limit can trivially be obtained with the two density formalism, it becomes heavy with the singlet-density expansion in the Mayer formalism.

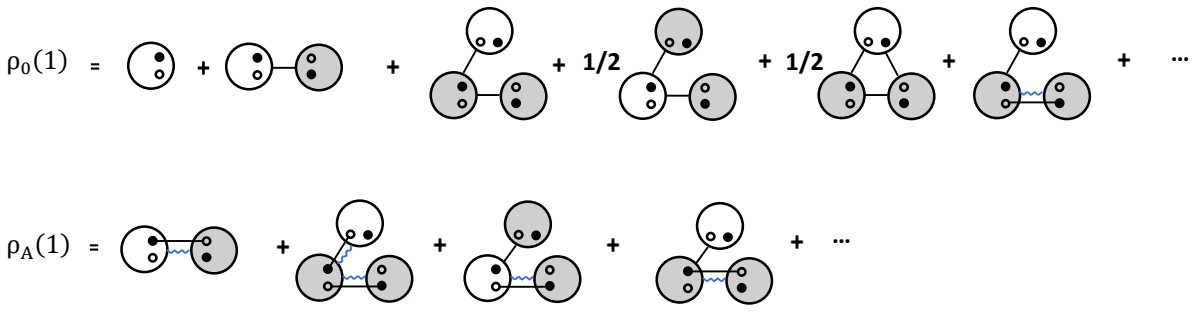
The overall number density of the system is then given as a sum of the number density of the bonded and non-bonded molecules. In the two site model, this is given by:

$$\rho(1) = \rho_0(1) + \rho_A(1) + \rho_B(1) + \rho_{AB}(1) \tag{4.13}$$

where $\rho_0(1)$ is the free monomer, $\rho_A(1)$ is the number density of the molecules bonded at site (A), $\rho_B(1)$ is the number density of the molecules bonded at site (B), and $\rho_{AB}(1)$ is the number density of the molecules bonded at site (A) and (B) simultaneously. In the one site model, this reduces to the two densities formalism:

$$\rho(1) = \rho_0(1) + \rho_A(1) \tag{4.14}$$

In terms of graph theory, an example on the different densities is represented by:



The white points are the fixed particles, and the grey ones are the system particles interacting with the fixed particles. The pre-factor corresponds to a symmetrical number.

4.2.4 Derived Equations

In this part, the derived equations of state of Wertheim using TPT theory are briefly provided. For details on the derivation, the reader is referred to the original articles of the Wertheim, or to the excellent works of Zmpitas [20] and Febra [19]. The derivation, presented in this work is rather heuristic and simple for understanding, mostly inspired from the original articles of Wertheim [3, 4, 5, 6, 7, 8], Chapman's and collaborators work [21, 22, 23, 24, 25], Galindo's work [26], Muller and Gubbins review [27], the work of Paricaud [28] and that of Zmpitas [20].

4.2.4.1 Single site model

In the single site model, the spheres possess only one attractive site. The interactions in the system may result in the formation of dimers at equilibrium. The system consists then in a mixture of hard-spheres and inseparable dispheres (hard-dumbbells) at certain composition. The number density is given by:

$$\rho = \rho_O + 2\rho_d \tag{4.15}$$

where ρ_d is dimers number density, and ρ_0 is the non-bonded spheres number density. In the ideal gas dimerization, the association (excess) Helmholtz free energy of a system is written as:

$$\beta A^{Asso} = \beta(A - A^{ideal}) \tag{4.16}$$

where $A = N\mu - PV$ is the total Helmholtz free energy of the system containing monomers and dimers, and $A^{ideal} = N\mu^{ideal} - P^{ideal}V$ is the ideal Helmholtz free energy of the non-associated system (only monomers). The chemical potential and the pressure for the ideal monomeric gas are:

$$\frac{\mu^{ideal}}{k_BT} = Ln(\rho(1)\Lambda^3) \tag{4.17}$$

$$\frac{P^{ideal}}{k_BT} = \int \rho(1)\, d(1) \tag{4.18}$$

The total Helmholtz free energy of the monomers ideal gas is:

$$\frac{A^{ideal}}{k_BT} = N\frac{\mu^{ideal}}{k_BT} - \frac{P^{ideal}}{k_BT}V = \int \rho(1)\, Ln(\rho(1)\Lambda^3) d(1) - \int \rho(1)\, d(1) \tag{4.19}$$

Note that $\int \rho(1)\, d(1) = N$.

Now, regarding the ideal gas mixture of monomers and dimers, the pressure is given by the Dalton's law:

$$\begin{aligned} \frac{P}{k_BT} &= \int \rho_{species}(1)\, d(1) \\ &= \int (\rho_0(1) + \rho_d(1)) d(1) \\ &= \int \tfrac{1}{2}(\rho(1) + \rho_0(1)) d(1) \end{aligned} \tag{4.20}$$

and the chemical potential of the dimers (in equilibrium with the monomers) is thus given by:

$$\frac{\mu}{k_BT} = Ln(\rho_0(1)\Lambda^3) \tag{4.21}$$

$$\begin{aligned}\beta A &= N\frac{\mu}{k_BT} - \frac{P}{k_BT}V \\ &= \int \rho(1)\, Ln(\rho_0(1)\Lambda^3)d(1) - \int \tfrac{1}{2}(\rho(1)+\rho_0(1))d(1)\end{aligned} \tag{4.22}$$

The association Helmholtz free energy due to the dimerization of the homogeneous ideal monomeric gas system at equilibrium is:

$$\begin{aligned}\beta A^{Asso} &= \beta(A - A^{ideal}) \\ &= \int \rho(1)\left(Ln\left(\frac{\rho_0(1)}{\rho(1)}\right) + \frac{1}{2} - \frac{1}{2}\left(\frac{\rho_0(1)}{\rho(1)}\right)\right)d(1) \\ &= N\left(Ln(X) + \frac{1}{2} - \frac{1}{2}X\right)\end{aligned} \tag{4.23}$$

Finally, the TPT derived equation of state for the single site model at certain temperature T, density ρ, is given in terms of the Helmholtz free energy A for N particles as follows:

$$\frac{\beta(A-A^{ideal})}{N} = Ln(X) - \frac{1}{2}X + \frac{1}{2} \tag{4.24}$$

This equation is the same as that given by Wertheim in Eq(9) of the Wertheim paper 2 [4].

Now, in order to use this equation to calculate any property of the system, one needs to determine the parameter X. This parameter is the fraction of monomers not bonded in the system which is a function of global density and non-bonded monomers density. The relation between both densities (monomer and dimer) is given by the mass action equation. However, this equation involves also the density of the dimers in the system, and thus, the parameter X is given by:

$$X = 1 - 2\frac{\rho_d}{\rho} \tag{4.25}$$

To obtain X, one needs to systematically know ρ_d, while ρ is, of course, already known. Thanks to the statistical mechanics applied to the chemical equilibrium, the density of the dimers present in the system under ideal conditions is given by:

$$\rho_d(1) = \frac{1}{2}\rho_0(1)\int \rho_0(2)e_R(12)f(12)d(2) \tag{4.26}$$

The equation for X then becomes when replacing ρ_d in it (for a homogeneous system) after rearrangements, as:

$$X = \frac{1}{1+\rho X\Delta} \tag{4.27}$$

which is a quadratic function that could be solved either analytically or numerically. Here Δ is given by:

$$\Delta = \int e_R f(12)d(2) \tag{4.28}$$

Remember that until now, this derivation is valid only for a dimerization occurring under ideal conditions. However, for an interacting system, Wertheim showed that the correct expression for the dimer density becomes [4]:

$$\rho_D(1) = \frac{1}{2}\rho_0(1)\int \rho_0(2)g_{00}(12)f(12)d(2) \tag{4.29}$$

Here $g_{00}(12)$ is the monomer-monomer pair correlation function given by this expression:

$$g_{00}(12) = \sum_{j=0}^{\infty} g_{00}^{j}(12) \tag{4.30}$$

where *j* here denotes the number of attractive Mayer function (bonds) involved, f_{AA}, in a certain cluster. Wertheim limited this to the lowest order approximation, where only one bond exist in a cluster. In other words, only a dimer surrounded by a sea of monomers is considered. In fact, in the one site model, this is not an approximation as the monomer has only one attractive site and then could only be bonded to one other monomer. However, if we anticipate the case of two sites model, the monomer now could be bonded to two other monomers (*j*=0 and *j*=1 are non-vanishing graphs if the one site model approximation is not made). With this approximation, $g_{00}(12)$ could then be simply approximated by the pair radial distribution function of the reference fluid $g_R(12)$. Furthermore, in the low density limit, one may approximate $g_R(12)$ by the Boltzmann factor $e_R(12)$.

The change in terms of pressure due to association is given by $P^{Asso} = \rho^2 \frac{\partial \frac{\beta A^{Asso}}{N}}{\partial \rho}|_{\beta}$ and that, in terms of the compressibility factor, is given by $Z^{asso} = \rho \frac{\partial A^{asso}}{\partial \rho}|_{\beta}$. Thus, the expression for Z^{Asso} is:

$$\begin{aligned} Z^{Asso} &= \rho \frac{\partial}{\partial \rho}\left(Ln(X) - \frac{1}{2}X + \frac{1}{2}\right) \\ &= \rho\left(\frac{\partial Ln(X)}{\partial X}\frac{\partial X}{\partial \rho} - \frac{1}{2}\frac{\partial X}{\partial X}\frac{\partial X}{\partial \rho}\right) \\ &= \rho \frac{\partial X}{\partial \rho}\left(\frac{1}{X} - \frac{1}{2}\right) \end{aligned} \tag{4.31}$$

where the term $\frac{\partial X}{\partial \rho}$ is obtained from the mass action equation and is given by:

$$\frac{\partial X}{\partial \rho} = \frac{-X^2(\Delta + \rho\frac{\partial \Delta}{\partial \rho})}{2\rho X\Delta + 1} \tag{4.32}$$

The equation for Z^{Asso} becomes then:

$$\begin{aligned} Z^{Asso} &= \frac{-\rho X^2}{2}\left(\frac{\partial \Delta}{\partial \rho}\rho + \Delta\right) \\ &= \frac{(X-1)}{2}\left(\frac{\partial Ln(\Delta)}{\partial Ln(\rho)} + 1\right) \end{aligned} \tag{4.33}$$

The pressure due to association is then simply obtained from:

$$P^{asso} = \rho Z^{Asso} \tag{4.34}$$

The total compressibility factor is:

$$Z = Z^{R} + Z^{Asso} \tag{4.35}$$

where Z^R is the compressibility factor of the reference term. In the case of hard sphere, its expression is given by the Carnahan-Starling equation [29]:

$$Z^{R} = \frac{(1+\eta+\eta^2-\eta^3)}{(1-\eta^3)} \tag{4.36}$$

Furthermore, because of the short range of the attractive potential, one may consider that dimerization occurs at contact only. Hence, the expression for Δ could be written as:

$$\begin{aligned}\Delta &= \int g_R(12) f(12) d(2) \\ &= g_R(\sigma,\rho) FK\end{aligned} \tag{4.37}$$

where $g_R(\sigma)$ is the radial distribution function of the hard sphere reference at contact, F is the Mayer function of the attractive site-site interaction, and K is the bonding volume. This leads to:

$$\begin{aligned}Z^{Asso} &= \frac{(X-1)}{2}\left(\frac{KF\rho}{\Delta}\frac{\partial(g_R)}{\partial(\rho)}+1\right) \\ &= \frac{(X-1)}{2}\left(\frac{\rho}{g_R}\frac{\partial(g_R)}{\partial(\rho)}+1\right)\end{aligned} \tag{4.38}$$

The expression for g_R is deduced from the equation of state of Carnahan-Starling [29], and is given by:

$$g_R(\sigma) = \frac{Z^R-1}{4\eta} \tag{4.39}$$

Therefore, the final TPT equation of state for the hard sphere with association is:

$$\begin{aligned}Z &= Z^R + Z^{Asso} \\ &= \frac{(1+\eta+\eta^2-\eta^3)}{(1-\eta^3)} - \frac{1}{2}\frac{(1+\eta-\frac{1}{2}\eta^2)(1-X)}{2(1-\eta)(1-\frac{1}{2}\eta)}\end{aligned} \tag{4.40}$$

It is interesting to check the validity of this TPT equation of state when approaching the limits. The case where only monomers are present in the system corresponds to $X = 1$, which, as expected, give a null association contribution. The second limit is when the whole system is dimerized, which corresponds to $X = 0$. In this latter case, the equation of state in terms of compressibility factor is given by:

$$Z_D^{Asso} = \frac{-1}{2}\left(\frac{\partial Ln(\Delta)}{\partial Ln(\rho)}+1\right) \tag{4.41}$$

Still, in this limit, it is more common to express the contribution in terms of the residual one:

$$Z_{D,res}^{Asso} = Z_D^{Asso} - Z_D^{ideal} \tag{4.42}$$

Thus:

$$\begin{aligned}Z_{D,res}^{Asso} &= Z_D^{Asso} - \lim_{\rho\to 0}(Z_D^{Asso}) \\ &= Z_D^{Asso} + \frac{1}{2} \\ &= -\frac{1}{2}\frac{\partial Ln(\Delta)}{\partial Ln(\rho)}\end{aligned} \tag{4.43}$$

The residual compressibility factor is then obtained by this expression:

$$Z_{D,res} = Z_{res}^R + Z_{D,res}^{Asso} \tag{4.44}$$

The residual Helmholtz free energy is obtained by integrating $Z_{D,res}$ over density, and the expression is given by:

$$\frac{\beta A_{D,res}}{N} = \frac{\beta A^{R}_{res}}{N} - \frac{1}{2} Ln(g_R(\sigma)) \tag{4.45}$$

where N being the number of spheres $N = m_s N_c$. m_s is the number of segments per chain and $N_c$is the number of chains in the system (dimers in this case). Thus, for a fully dimerized system, where $m_s = 2$, is written in terms of Helmholtz free energy as:

$$\frac{\beta A_{D,res}}{N_c} = m_s \frac{\beta A_{D,res}}{N} = 2\frac{\beta A^{ref}_{res}}{N} - Ln(g_R) \tag{4.46}$$

If we anticipate the progression in our review on the perturbation theory, one can note that this equation is exactly the same expression as that given by SAFT equation of state for a system of dimers. Moreover, if one assumes the additivity of the residual Helmholtz free energy due to association and that of the reference term, one obtains the complete scheme of the SAFT EoS. For instance, $A_{T,res}$ of a trimer, will be written as a contribution of three monomers + two bonds, … and so on, thus, one may write:

$$\begin{aligned}
\frac{\beta A_{D,res}}{N_c} &= 2\frac{\beta A^{ref}_{res}}{N} - 1Ln(g_R) \\
\frac{\beta A_{T,res}}{N_c} &= 3\frac{\beta A^{ref}_{res}}{N} - 2Ln(g_R) \\
\frac{\beta A_{Q,res}}{N_c} &= 4\frac{\beta A^{ref}_{res}}{N} - 3Ln(g_R) \\
&\vdots \\
\frac{\beta A_{m,res}}{N_c} &= m_s \frac{\beta A^{ref}_{res}}{N} - (m_s - 1)Ln(g_R)
\end{aligned} \tag{4.47}$$

This last equation is nothing more than the usual SAFT equation of state.

4.2.4.2 Two sites model

In the two sites model, a sphere contains two attractive sites. The sphere may then be bonded to one or two other sites of two different molecules. Hence, a polymerization may occur to form chains of dimers, trimers, or oligomers, while in the one-site model only dimers could be formed. Moreover, in this model, four number densities were defined by Wertheim instead of two in the one site model. The densities are as follow: ρ_0 is the non-bonded monomers, ρ_A is the density number of molecules bonded only on site A, ρ_B is the number of molecules bonded only on site B and ρ_{AB} is the number of molecules bonded simultaneously on sites A and B. For calculations simplicity, Wertheim proposed to work with the following density parameters related to the densities by:

$$\sigma_0 = \rho_0, \;\; \sigma_A = \rho_A + \rho_0, \;\; \sigma_B = \rho_B + \sigma_0 \tag{4.48}$$

$$\sigma = \rho_0 + \rho_A + \rho_B + \sigma_{AB} = \rho \tag{4.49}$$

As there is exactly the same number of sites of type A and B, this leads to $\sigma_A = \sigma_B = \sigma_G$, where σ_G will be used later to simplify the notation as done by Wertheim (note also that $\rho_A = \rho_B = \rho_G$).

In the following, the framework proposed for the one site model is used here to derive the Helmholtz free energy of the system in the two-sites model case.

If we assume the mixture as an ideal gas of monomers and dimers, the pressure is:

$$\frac{P}{k_BT} = \int \rho_{species}(1)\, d(1) \tag{4.50}$$

where the number of species are now multiple, and could be of type $0, A, B$ or AB. As sites of type A and type B are identical, one can consider them as the same species, noted G. Therefore, the number of species in the system is the total number of spheres minus those that are bonded which could be of type G or AB. The formula is then given by:

$$\begin{aligned} \rho_{species} &= \rho - \rho_G - \rho_{AB} \\ &= \sigma_G \end{aligned} \tag{4.51}$$

The pressure is then:

$$\begin{aligned} \frac{P}{k_BT} &= \int \rho_{species}(1)\, d(1) \\ &= \int \sigma_G(1)\, d(1) \end{aligned} \tag{4.52}$$

and the chemical potential of the dimers in equilibrium with monomers is thus given by:

$$\frac{\mu}{k_BT} = Ln(\rho_0(1)\Lambda^3) \tag{4.53}$$

The total Helmholtz free energy is:

$$\begin{aligned} \beta A &= N\frac{\mu}{k_BT} - \frac{P}{k_BT}V \\ &= \int \rho(1)\, Ln(\rho_0(1)\Lambda^3) d(1) - \int \rho_G(1)\, d(1) \end{aligned} \tag{4.54}$$

The residual Helmholtz free energy of the dimerization of the homogeneous gas system at equilibrium is:

$$\begin{aligned} \beta A^{Asso} &= \beta(A - A^{ideal}) \\ &= \int \sigma(1) \left(Ln\left(\frac{\sigma_0(1)}{\sigma(1)}\right) - \frac{\sigma_G(1)}{\sigma(1)} + 1 \right) d(1) \\ &= N\left(Ln\left(\frac{\sigma_0}{\sigma}\right) - \frac{\sigma_G}{\sigma} + 1 \right) \end{aligned} \tag{4.55}$$

Wertheim defined two other parameters that measure the degree of polymerization, given by $\nu = \frac{\sigma}{\sigma_G}$, $\tau = \frac{\sigma_G}{\sigma_0}$.

Hence,

$$\begin{aligned} \beta A^{Asso} &= \beta(A - A^{ideal}) \\ &= N\left(-Ln(\nu\tau) - \frac{1}{\nu} + 1 \right) \end{aligned} \tag{4.56}$$

and the compressibility factor is:

$$\begin{aligned} \beta Z^{Asso} &= -\rho\left(\frac{\partial Ln(\nu\tau)}{\partial \nu}\frac{\partial \nu}{\partial \rho} + \frac{\partial Ln(\nu\tau)}{\partial \tau}\frac{\partial \tau}{\partial \rho} + \frac{\partial\left(\frac{1}{\nu}\right)}{\partial \nu}\frac{\partial \nu}{\partial \rho} \right) \\ &= \rho\left[\frac{1}{\nu}\frac{\partial \nu}{\partial \rho}\left(\frac{1}{\nu} - 1\right) - \frac{1}{\tau}\frac{\partial \tau}{\partial \rho} \right] \end{aligned} \tag{4.57}$$

The self-consistent densities are obtained from the minimum principle of the Helmholtz free energy at chemical equilibrium. The generalized expression [8] is :

$$\beta(A - A_R) = \int \left(\sigma(1) ln \frac{\sigma_0(1)}{\sigma(r)} + \sigma(r) - \sigma_A(1) - \sigma_B(1) + \frac{\sigma_A(1)\sigma_B(1)}{\sigma_0(1)}\right) d(r) - \Delta c^{(0)} \quad (4.58)$$

where

$$\Delta c^{(0)} = \overline{\sigma_A}\,\overline{\sigma_B} \sum_{n=0}^{\infty} \overline{\sigma_0}\, I_{n+1} \quad (4.59)$$

where n stands for the TPT order. In TPT1, I_1 is given by:

$$I_1 = \int g_R(12) f_{BA}(12) d(2) \quad (4.60)$$

In the first order TPT, the stationary condition on densities gives:

$$\tau - 1 = \frac{\partial \Delta c^{(0)}}{\partial \sigma_G(1)} \quad (4.61)$$

$$\tau\nu - \tau\tau = \frac{\partial \Delta c^{(0)}}{\partial \sigma_0(1)} \quad (4.62)$$

The minimum of $\frac{\partial \beta(A - A_R)}{\partial \sigma_0(1)}$ is obtained when $\nu = \tau$. Therefore,

$$\frac{\beta A^{Asso}}{N} = -2Ln(\nu) - \frac{1}{\nu} + 1 \quad (4.63)$$

where ν is defined as the mean number of beads per chain (or simply the average chain length).

Now, the idea is to try to generalize this equation for multiple sites model from the two sites model. The following fraction: $X_A = \frac{\sigma_B}{\sigma} = \frac{1}{\nu_B}$ is the fraction of molecules with sites *A* being free and $X_B = \frac{\sigma_A}{\sigma} = \frac{1}{\nu_A}$ is the fraction of molecules with sites *B* being free, keeping in mind that the fraction of not bonded monomers is given by $X_0 = \frac{\sigma_0}{\sigma}$. From the self-consistent density equations, one obtains that $X_0 = X_A X_B = X_G^2$.

The association Helmholtz free energy can then be rewritten as:

$$\begin{aligned} \frac{\beta A^{Asso}}{N} &= 2Ln(X_A) - X_A + 1 \\ &= Ln(X_A) - \frac{1}{2}X_A + \frac{1}{2} + Ln(X_B) - \frac{1}{2}X_B + \frac{1}{2} \end{aligned} \quad (4.64)$$

where X_A and X_B are obtained using the first self-consistent equation:

$$X_G = \frac{1}{1 + \rho X_G \Delta_{AB}} \quad (4.65)$$

$$\Delta_{AB} = \int e_R f_{AB}(12) d(2) \quad (4.66)$$

This equation for $\frac{\beta A^{Asso}}{N}$ could be seen as the sum of two contributions of a single site model. This is due to the fact that in TPT1, the association at one site is independent on the association at all other sites. Hence, one may write the generalized scheme for *M* sites as follow:

$$\frac{\beta A^{Asso}}{N} = \sum_{i=1}^{M} \left(Ln(X_i) - \frac{X_i}{2}\right) + \frac{M}{2} \quad (4.67)$$

The compressibility factor is:

$$Z^{Asso} = \rho \sum_{i=1}^{M} \left(\frac{1}{X_i} - \frac{1}{2}\right) \frac{\partial X_i}{\partial \rho} \tag{4.68}$$

and X_i is obtained through:

$$X_i = \frac{1}{1 + \rho \sum_B X_B \Delta_{AB}} \tag{4.69}$$

The generalization for a multicomponent mixture where spheres may have different sites number writes,

$$\frac{\beta A^{Asso}}{N} = \sum_\alpha x_\alpha \left[\sum_i \left(Ln(X_i^\alpha) - \frac{X_i^\alpha}{2}\right) + \frac{M_\alpha}{2}\right] \tag{4.70}$$

where $x_\alpha = \frac{N_\alpha}{N}$ is the mole fraction of molecules α, and X_i^α is the fraction of molecules α which is not bonded at site i. The first sum is done over all the components, and the second one is done over all the sites of each component. Hence, the compressibility factor is obtained similarly to the one site model and is given by:

$$Z^{Asso} = \sum_\alpha x_\alpha \rho \left[\sum_i \left(\frac{1}{X_i^\alpha} - \frac{1}{2}\right) \frac{\partial X_i^\alpha}{\partial \rho}\right] \tag{4.71}$$

All the equations provided here, and derived from Wertheim TPT theory, have been successfully validated on molecular simulations of the fluid model [7, 8, 21, 22] as well as on real fluids. In the TPT theory of Wertheim, the system consists of a monodisperse fluid only in the case of non-associated spheres fluid or after a complete dimerization of a single site spheres fluid. Otherwise, the molecules of the system would be polydisperse due to the stoichiometry of the sites, meaning that polymerization would take place resulting in chains with different chain lengths. The average chain length is given by ν, and the comparison against molecular simulations of monodisperse chain like molecules with length m is made upon the condition $\nu = m$. However, it is obvious that such comparison is not fully correct, and thus, it is highly suitable to have a theory to describe the properties of a monodisperse associating or not associating fluids.

Chapman and collaborators have derived the expression for the Helmholtz free energy due to formation of chains in a mixture [25]. Their theory is fully based on the TPT of Wertheim, with some additional restrictions on the molecular model interactions. They have considered a mixture of spheres $(s_1, s_2, s_3 \ldots s_N)$ which will bond to form a mono-disperse chains fluid under the following restrictions. The site a, on sphere s_1, bonds only with site b on sphere s_2, site c of sphere s_2 bonds only to site d sphere $s_3 \ldots$site p of sphere s_{m-1} bonds only to site q of sphere s_m. Moreover, they have considered favourable stoichiometry, so that if the fraction of monomers is set to zero, this will result in mixture of r chains with length m ($s_N = r.m = \sum_{j=1}^{r} m_j$). The chain is then build-up with two types of spheres, single site ones allowed to bond for the end members, and $m - 2$ inner spheres with two sites each allowed to bond. The spheres may have other sites, but they are not allowed to interact with those chosen a priori to make bonds between them to build-up the chain molecule. This leads to:

$$Z^{Asso} = \rho_{chain} 2(m-1) \left[\frac{\partial(X)}{\partial(\rho)} \left(\frac{1}{X} - \frac{1}{2}\right)\right] \tag{4.72}$$

where $\frac{\partial(X)}{\partial(\rho)}$ is obtained as previously from the mass action equation:

$$\frac{\partial(X)}{\partial(\rho)}\left(\frac{1}{X}-\frac{1}{2}\right)=-\frac{1}{2}\frac{\rho_{chain}}{\rho}X^2\Delta\left(1+\frac{\partial Ln\Delta}{\partial Ln\rho}\right) \tag{4.73}$$

substitution of $\frac{\partial(X)}{\partial(\rho)}$ in the expression of Z^{Asso}, gives (note that $X^2\Delta=\frac{(1-X)}{\rho_{chain}}$ and $m=\frac{\rho_{chain}}{\rho}$)

$$\begin{aligned} Z^{Asso} &= -\rho_{chain}2(m-1)\left[\frac{1}{2}\frac{\rho_{chain}}{\rho}X^2\Delta\left(1+\frac{\partial Ln\Delta}{\partial Ln\rho}\right)\right] \\ &= -\rho_{chain}(m-1)\left[\frac{\rho_{chain}}{\rho}\frac{(1-X)}{\rho_{chain}}\left(1+\frac{\partial Ln\Delta}{\partial Ln\rho}\right)\right] \\ &= -\frac{(m-1)}{m}(1-X)\left[(1+\frac{\partial Ln\Delta}{\partial Ln\rho})\right] \end{aligned} \tag{4.74}$$

The complete polymerization is obtained when the fraction of monomers is set to zero, i.e. $X \to 0$, so that:

$$\begin{aligned} Z^{Chain} &= \lim_{X\to 0} Z^{Asso} \\ &= -\frac{(m-1)}{m}\left[\left(1+\frac{\partial Ln\Delta}{\partial Ln\rho}\right)\right] \end{aligned} \tag{4.75}$$

The ideal chain contribution is obtained when $\rho \to 0$:

$$\begin{aligned} Z_{res}^{chain} &= \lim_{\rho\to 0} Z^{Chain} \\ &= -\frac{(m-1)}{m} \end{aligned} \tag{4.76}$$

One may then obtain the residual contribution by subtracting this ideal part from the total contribution:

$$\begin{aligned} Z_{res}^{chain} &= Z^{Chain} - Z_{ideal}^{chain} \\ &= -\frac{(m-1)}{m}\frac{\partial Ln\Delta}{\partial Ln\rho} \\ &= -\frac{(m-1)}{m}\frac{\rho}{(g_R(\sigma)}\frac{\partial((g_R(\sigma))}{\partial(\rho)} \end{aligned} \tag{4.77}$$

The residual Helmholtz free energy is obtained by integrating this latter over density ρ and is then given by:

$$\frac{\beta A^{Asso}}{N} = -\frac{(m-1)}{m}Ln(g_R(\sigma)) \tag{4.78}$$

or could also be written as ($N = mN_{Chain}$):

$$\frac{\beta A^{Asso}}{N_{Chain}} = -(m-1)Ln(g_R(\sigma)) \tag{4.79}$$

and for a mixture, this becomes:

$$\frac{\beta A^{Asso}}{N_{Chain}} = -\sum_{j=1}^{r}(m_j-1)x_jLn(g_{R,jj}(\sigma_{jj})) \tag{4.80}$$

where the sum is done over the chains of type j, m_j is the length of the chains j, x_j is the mole fraction of molecule j, σ_{jj} is the center-to-center distance between two bonded spheres of molecule j.

With this expression, Chapman and collaborators gave birth to a new type of equation of state named SAFT [25], which stands for the Statistical Association Fluid Theory. In this form, the TPT of Wertheim is capable to control the geometry of a molecule, and becomes much more convenient to be compared to molecular simulations in a straightforward way for systems consisting of a monodisperse fluid.

Few years later, Y. Zhou and G. Stell [30] showed another alternative to derive the same expression given by Chapman using their theory on ionic systems. Their demonstration is based on the use of the cavity function, which is related to the change in the free energy of a system by this relation:

$$\beta\Delta A = -Ln(y(r_1, r_2, r_3, \dots r_{N_R},\ N_R, N_C)) \tag{4.81}$$

where, y is the m-body correlation function, r_i denote the position of the spheres, N_R is the number of spheres in the system and N_C is the number of chains formed due to bonding between the spheres.

A physical interpretation could be given as follow. In a system, two cavity particles do not interact with each other. Therefore, if *m* particles are present in the system, the dense medium will tend to drive them into each other to save space, costing a certain energy to the system. In other words, it is the energy required to bring two cavities from infinity to certain distance under thermodynamic conditions. In this case, it is interesting to bring them to $r = \sigma$ at which they are allowed to bond to form a dimer. This dimer can then bond to a third cavity…until a chain molecule is formed. This change in the free energy due to the formation of a chain in the system can be calculated by means of the cavity function. As already mentioned, this cavity function is multidimensional, and little is known about it. However, if one restricts only to its lowest order; two-body approximation; this becomes more convenient. Furthermore, the use of the linear approximation allows one to write:

$$y(r_1, r_2, r_3, \dots r_m) = y(r_1, r_2)y(r_2, r_3)y(r_3, r_4) \dots y(r_{m-1}, r_m) \tag{4.82}$$

Moreover, if one is interested only on bringing the cavities at a distance σ of each other, this equation becomes:

$$y(r_1, r_2, r_3, \dots r_m) = y(\sigma_{12})y(\sigma_{23})y(\sigma_{34}) \dots y(\sigma_{(m-1)m}) \tag{4.83}$$

Thus, the change in the free energy due to the formation of a chain molecule composed of m cavities is:

$$\beta\Delta A = -Ln(y(\sigma_{12})y(\sigma_{23})y(\sigma_{34}) \dots y(\sigma_{(m-1)m}), N_R - m, 1) \tag{4.84}$$

If all spheres are identical, $\sigma = \sigma_{ij}$, and the equation becomes:

$$\beta\Delta A = -Ln(y(\sigma)^{m-1}, N_R - m, 1) \tag{4.85}$$

This is the energy required to form a chain molecule composed of m spheres in a sea of non-connected spheres. Furthermore, as the approximation that the cavities do not interact between each other has been done, the formation of a chain is independent of the formation of another

chain (similarly to the single chain approximation in TPT theory). The energy required to bond the whole system into N_C chains is then additive, and is given by:

$$\begin{aligned}\beta\Delta A &= -N_C Ln(y(\sigma)^{m-1}, 0, N_C) \\ &= N_C(m-1)Ln(y(\sigma), 0, N_C) \\ &= N_C(m-1)Ln(g(\sigma))\end{aligned} \tag{4.86}$$

where $g(\sigma)$ is the radial distribution function of the reference fluid (hard spheres for example). This final expression is exactly the same as that given by Chapman et al. to account for the bond formation.

4.3 SAFT models

Unlike the equation of state of Wertheim, the SAFT equation of state developed by Chapman and co-workers [25] models a fluid consisting of a monodisperse chains with or without association. This EoS has a strong physical basis as it relies on the statistical mechanical perturbation theory compared to the empirical or semi-empirical ones. Moreover, as it relies on a well-defined molecular model, it is possible to compared it directly to molecular simulations of the same fluid type. This latter point is extremely advantageous, as the molecular simulation tool may provide "exact" properties (to the numerical uncertainties) of the well-defined molecular model Thus, the robustness or the deficiency of the SAFT EoS model can be assessed on the properties of the model fluid before applying it to real fluids.

The SAFT EoS is given in terms of the residual Helmholtz free energy in separate contributions:

$$\beta\frac{A^{res}}{N_C} = \beta m\frac{A^{seg}}{N_s} + \beta\frac{A^{chain}}{N_C} + \beta\frac{A^{asso}}{N_C} \tag{4.87}$$

Where A^{res} is the total residual Helmholtz free energy, A^{seg} is the contribution due to the intermolecular interaction of the reference fluid. A^{chain} accounts for the change in the energy due to the polymerization of segments, and A^{asso} is the association term taking into account the interactions due to association such as hydrogen bonds. N_s is the number of segments in the system, N_C is the number of chains composed of m segments present in the system, i.e. $m = \frac{N_s}{N_C}$.

Depending on the intermolecular interaction potential, different SAFT versions were developed in the literature. Among them, the one based on the reference hard sphere model developed by Champan [25] (SAFT-HS) and modified later by Huang and Radosz [31] (CK-SAFT), the LJ-SAFT [32] extended to mixtures by Blas and Vega [33] (soft-SAFT), the SAFT-VR [34, 35] developed initially by Gil-Villegas for different potential with variable range, and then extended for different potential interactions (Square Well, m-n potentials, Yukawa), or the PC-SAFT [36] with a slightly different formalism. It is the segment term and the associated radial distribution function that makes different SAFT versions, as the form of the chain and the association terms are roughly kept unchanged.

4.3.1 SAFT-HS and CK-SAFT

The original-SAFT (SAFT-HS) models Lennard-Jones chain fluids, where the reference term is composed of two parts, one accounting for the short range interaction using a hard sphere

(HS) potential and a perturbation part accounting for the dispersive interactions of the LJ potential. The Carnahan and Starling EoS [29] is used for the HS,

$$\frac{A^{HS}}{RT} = \frac{4\eta - 3\eta^2}{(1-\eta)^2} \tag{4.88}$$

where $\eta = \frac{\pi N_A}{6}\rho d^3 m$, and the dispersion term is taken from the Cotterman expression [37].

$$A^{Disp} = \frac{\varepsilon R}{k}\left(a_{01}^{Disp} + \frac{a_{02}^{Disp}}{T_R}\right) \tag{4.89}$$

where

$$a_{01}^{Disp} = \rho_R[-8.5959 - 4.5424\rho_R - 2.1268\rho_R^2 + 10.285\rho_R^3] \tag{4.90}$$

$$a_{02}^{Disp} = \rho_R[-1.9075 + 9.9724\rho_R - 22.216\rho_R^2 + 15.904\rho_R^3] \tag{4.91}$$

$T_R = \frac{kT}{\varepsilon}$ is the reduced temperature

$\rho_R = \frac{6}{2^{0.5}\pi}\eta$ is the reduced density

In the CK-SAFT of Chen and Kreglewski [38] the dispersive term is fitted on molecular simulation data (internal energy and PVT data and second virial coefficient) of particles interacting through a square-well potential and is given by:

$$\frac{A^{Disp}}{RT} = \sum_i \sum_j D_{ij}[\beta u]^i \ [\frac{\eta}{\tau}]^j \tag{4.92}$$

D_{ij} are universal constants given in the original paper [38]. η is the segment packing fraction and τ is the upper limit set to $\pi\frac{\sqrt{2}}{6}$. u is the segment-segment well depth temperature-dependent interaction parameter given by:

$$u = u^0(1 + \frac{e}{k_B T}) \tag{4.93}$$

where u^0 is the well depth temperature-independent energy parameter, and $\frac{e}{k_B} = 10$ for all molecules except for few small molecules (it is related to the Pitzer's acentric factor and to the critical temperature). The chain term is computed using the radial distribution function of the reference term corresponding to the Carnahan-Starling EoS.

This equation of sate was widely applied to model thermodynamic and phase equilibrium properties of complex pure fluids and their mixtures and can be found in [31, 39, 40, 41, 42, 43].

4.3.2 LJ-SAFT (soft-SAFT)

Instead of using a hard sphere fluid to describe the repulsive intermolecular interactions and a perturbation term to describe the dispersive interactions, the LJ potential possesses both type of interactions and can be used as a segment reference in SAFT EoS. The residual segment Helmholtz free energy of the LJ fluid is given by Johnson [44] using a modified BWD equation of state refitted on the internal energy and LVE properties. The expression is given by:

$$A^{seg} = \sum_{i=1}^{8} \frac{a_i \rho^{*i}}{i} + \sum_{i=1}^{6} b_i G_i \quad (4.94)$$

where a_i and b_i are temperature dependent parameters and G_i are density dependent parameters. Their expressions and their corresponding constant parameters can be found in [44]. This equation of state yields very accurate LVE and internal energy properties of the fluid in question, including the critical properties which were imposed during the optimization. The regression temperature range is $0.7 \leq T^* \leq 6$ and densities up to $\rho^* \leq 1.25$. Other equations of state are often used in the literature to represent the segment reference term of the LJ-SAFT. Among them, there is the one of Kolafa-Nezbeda [45] based on a perturbation virial expansion with temperature-dependent hard sphere diameter, or the more recent EoS of Thol et al. [46].

In its original form, the radial distribution function for the LJ reference fluid needed for the chain term is obtained by fitting the following functional form to molecular simulation data [44]:

$$y_R(\sigma) = g^{LJ}(T^*, \rho^*, \sigma) e^{\beta \emptyset^{LJ}(T^*, \rho^*, \sigma)} \quad (4.95)$$

where

$$g^{LJ}(T^*, \rho^*, \sigma) = 1 + \sum_{i=1}^{5} \sum_{j=1}^{5} a_{ij} \rho^{*i} T^{*(1-j)} \quad (4.96)$$

$$\emptyset^{LJ}(T^*, \rho^*, \sigma) = 0 \quad (4.97)$$

The regression parameters a_{ij} are given in [32], and covers the same range of temperature and density as the reference term of Johnson above.

The performance of the LJ-SAFT (sometimes named soft-SAFT) on the properties of the fluid model and real fluids as well as their mixtures can be found in [32, 47, 48, 49, 33, 50, 51, 52, 53, 54, 55, 56, 57].

4.3.3 SAFT-VR

The SAFT-VR developed by Gil-Villegas is based on the Barker and Henderson perturbation theory [58, 59], [60] which originates from the Zwanzig high-temperature expansion perturbation [61]. The expression for the Helmholtz free energy of the monomer is given by an expansion in the inverse of the temperature:

$$A^{seg} = A^{HS} + \beta A_1 + \beta^2 A_2 + \cdots \quad (4.98)$$

where A^{HS} is the Hard sphere reference term, and A_1, A_2 are the first and second orders perturbations contributions. At high temperatures, these high order terms should vanish to recover the fluid of hard spheres which is physically consistent (a van der Waals type fluid), while at low temperature they are no more negligible introducing the soft attractions. The expression of A_1 is:

$$A_1 = -2\pi \rho_s \epsilon \int_{\sigma}^{\infty} r^2 \emptyset(r) g^{HS}(r) dr \quad (4.99)$$

if $g^{HS}(r) = 1$ is taken, i.e. unstructured fluid, one obtains the classical van der Waals mean-field energy.

The second term becomes preponderant as the density increases and accounts for the fluctuations of the attractive energy due to the compression of the fluid. This requires the knowledge of high order correlation functions. However, this is complicated and thus, an approximation based on the Barker and Henderson approach was used instead. These fluctuations are correlated to the fluctuations in term of the number of particles inside the attractive well potential by means of the Local Compressibility Approximation (LCA). The expression for A_2 is then written as:

$$A_2 = \frac{1}{2}\rho_s \epsilon K^{HS} \frac{\partial A_1}{\partial \rho_s} \tag{4.100}$$

where K^{HS} is the hard sphere isothermal compressibility.

For a fluid interacting with a square-well potential, Gil-Villegas obtained the following expressions:

$$A_1^{sw} = -4\eta\epsilon(\lambda^3 - 1) g^{HS}(1; \eta_{eff}) \tag{4.101}$$

where $1.1 \leq \lambda \leq 1.8$ is the width of the potential.

$$g^{HS}(1; \eta_{eff}) = \frac{(1-0.5\eta_{eff})}{(1-\eta_{eff})} \tag{4.102}$$

and

$$\eta_{eff} = c_1\eta + c_2\eta^2 + c_3\eta^3 \tag{4.103}$$

the coefficient c_1, c_2, c_3, are given by the following matrix:

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 2.25855 & -1.50349 & 0.249434 \\ -0.669270 & 1.40049 & -0.827739 \\ 10.1576 & -15.0427 & 5.30827 \end{pmatrix} \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \end{pmatrix} \tag{4.104}$$

The second term is then obtained using the first one:

$$A_2^{sw} = \frac{1}{2}\rho_s \epsilon K^{HS} \frac{\partial A_1^{sw}}{\partial \rho_s} \tag{4.105}$$

and K^{HS} is obtained from the Percus-Yevick expression given by:

$$K^{HS} = \frac{(1-\eta)^4}{1+4\eta+4\eta^2} \tag{4.106}$$

Different expressions are obtained for Sutherland fluids, Yukawa fluids and for the Soft repulsive fluids in the original paper of Gil-Villegas [34]

Zhang [62] showed that significant improvement can be achieved for A_2^{sw} (particularly for $\lambda = 1.5$) using the Macroscopic Compressibility Approximation (MCA). He showed that at high densities, the number of molecules in successive shells are correlated. By assuming that the correlation is linear and is proportional to the particle number density and the volumes of the shells, he introduced a correction factor to A_2^{sw}, and thus, the expression becomes:

$$A_2^{sw} = \frac{1}{2}\rho_s \epsilon K^{HS}(1+\chi) \frac{\partial A_1^{sw}}{\partial \rho_s} \tag{4.107}$$

where $\chi = 8.23\eta^2$. If $\chi = 0$ we recover the original expression for A_2^{SW}. The improvement were also due to the introduction of a more accurate isothermal compressibility K^{HS} using the expression obtained from the Carnahan-Starling equation:

$$K^{HS} = \frac{(1-\eta)^4}{1+4\eta+4\eta^2-4\eta^3+\eta^4} \tag{4.108}$$

Paricaud [63] proposed a modification in the pre-factor correction to better predict properties of soft potentials where the correction of Zhang fails (it overestimates the saturation pressure and underestimates the liquid density).The pre-factor was optimized on the simulated LVE data for SW and Buckingham potential fluids, leading to a $\chi \propto \eta^2$ and η^5. The term in η^2 leads to decreasing the saturation pressure, while the term in η^5 leads to an increase in the liquid densities.

To account for the chain formation requires the knowledge of the radial distribution function of the square well model. This is given by the high-temperature expansion:

$$g^{Mono} = g^{HS} + \beta\epsilon g_1(r) + (\beta\epsilon)^2 g_2(r) + \cdots \tag{4.109}$$

Only $g_1(r)$ can be obtained as $g_2(r)$ requires the knowledge of the third order expansion of the free energy. $g_1(r)$ is then obtained using a self-consistent method, where identification is made between the compressibility factor obtained from the Clausius virial theorem $\mathrm{Z}^{Mono} = \frac{PV}{N_s k_B T}$ and that obtained from a density derivative of the Helmholtz free energy $\mathrm{Z}^{Mono} = \eta \frac{\partial A^M}{\partial \eta}$. The expression for the square well monomers at contact is then given by:

$$g^{SW} = g^{HS} + \frac{1}{4}\beta\left[\frac{\partial A_1^{SW}}{\partial \eta} - \frac{\lambda}{3\eta}\frac{\partial A_1^{SW}}{\partial \lambda}\right] \tag{4.110}$$

where

$$g^{HS} = \frac{(1-0.5\eta)}{(1-\eta)^3} \tag{4.111}$$

The SAFT-VR equation of state has shown successful applications over a wide range of industrial systems to describe their thermodynamic properties [35, 64, 65, 66, 67, 68, 69].

The Mie version of the SAFT-VR EoS developed by Lafitte [70] was found more accurate in describing both equilibrium and derivative properties of the liquid phase such as speed of sounds when compared to the SAFT-VR SW, SAFT-VR LJC, PC-SAFT, PR78 and SRK. Latter, the SAFT-VR Mie was improved by Lafitte [71] by using a perturbation expansion for the Helmholtz free energy of the Mie fluid up to the third order. This led to better agreement when compared to the previous version for equilibrium properties of the fluid model and as well as for the real fluids in particular close to the critical region.

4.3.4 PC-SAFT

Here instead of using hard spherical segment as a reference fluid, it is a hard chain which is used, similarly to the perturbed hard sphere chain theory (PHSC) [72] where the theory found its foundation. The PC-SAFT models a square-well fluids, and is written as follow:

$$\beta \frac{A^{res}}{N_C} = \beta \frac{A^{HC}}{N_C} + \beta \frac{A^{PC}}{N_C} + \beta \frac{A^{asso}}{N_C} \quad (4.112)$$

where A^{HC} is the reference hard sphere chain fluid and A^{PC} is the perturbed hard chain contribution.

The idea in PC-SAFT is first to consider a formation of chain of hard spheres and then to do the perturbation at the chain level instead of at the spherical segment level. Thus, the A^{HC} is obtained similarly to the original-SAFT and is given by :

$$\beta \frac{A^{HC}}{N_C} = m\beta \frac{A^{HS}}{N_s} + (1-m) ln(g^{HS}) \quad (4.113)$$

where A^{HS} and g^{HS} are obtained from the generalized Carnahan-Starling equation of state to deal with mixtures of hard spheres given by Boublik [73] and Mansoori [74] as follow:

$$\beta \frac{A^{HS}}{N_s} = \frac{1}{\xi_0}\left[\frac{3\xi_1\xi_2}{1-\xi_3} + \frac{{\xi_2}^3}{\xi_3(1-\xi_3)^2} + \left(\frac{{\xi_2}^3}{{\xi_3}^2} - \xi_0\right) ln(1-\xi_0)\right] \quad (4.114)$$

$$g^{HS} = \frac{1}{1-\xi_3} + \left(\frac{\sigma_i\sigma_j}{\sigma_i+\sigma_j}\right)\frac{3\xi_2}{(1-\xi_3)^2} + \left(\frac{\sigma_i\sigma_j}{\sigma_i+\sigma_j}\right)^2 \frac{2{\xi_2}^2}{(1-\xi_3)^3} \quad (4.115)$$

where

$$\xi_n = \frac{\pi}{6}\rho \sum_i x_i m_i \sigma_i^n \quad (4.116)$$

Note that for a mixture of hard spheres with the same size diameter σ, the Carnahan-Starling equation is recovered.

To take into account of the softness of the segments, an effective temperature dependent-diameter suggested by Barker and Henderson theory is used (it is the case as well in SAFT-VR), and is given by the following form:

$$\sigma(T) = \int_0^\sigma \left(1 - exp(-\frac{\epsilon}{k_B T})\right) \quad (4.117)$$

where ϵ is the potential well depth.

The effective diameter of the hard sphere d decreases when T increases. This is because molecules at high temperature have a higher kinetic energy and tend to enter each other more (overlapping).

The perturbation term is obtained using the Barker-Henderson expansion truncated at the second order. This term writes as follow:

$$\beta \frac{A^{PC}}{N_C} = \beta \frac{A_1}{N_c} + \beta \frac{A_2}{N_c} \quad (4.118)$$

where A_1 and A_2 are respectively the first and second order perturbation terms given by:

$$\beta \frac{A_1}{N_c} = -2\pi m^2 (\frac{\epsilon}{k_B T})\sigma^3 \int_1^\lambda g^{HC}(m; x, \rho) x^2 dx \quad (4.119)$$

$$\beta \frac{A_2}{N_c} = -\pi\rho m (Z^{HC} + \rho \frac{\partial Z^{HC}}{\partial \rho})^{-1} m^2 \sigma^3 (\frac{\epsilon}{k_B T})^2 \frac{\partial}{\partial \rho}\left(\rho \int_1^\lambda g^{HC}(m; x, \rho) x^2 dx\right) \quad (4.120)$$

where $g^{HC}(m;x,\rho)$ is used instead of $g^{HS}(x,\rho)$. For pure chains, the hard chain compressibility factor can be obtained from A^{HS} and is given by:

$$Z^{HC}+\rho\frac{\partial Z^{HC}}{\partial\rho}=1+m\frac{8\eta-2\eta^2}{(17-\eta)^4}+(1-m)\frac{20\eta-27\eta^2+12\eta^3-2\eta^4}{(1-\eta)(2-\eta)^2} \quad (4.121)$$

and

$$I_1(\eta,m)=\int_1^\lambda g^{HC}(m;x,\rho)x^2dx=\sum_{i=0}^{6}a_i(m)\eta^i \quad (4.122)$$

$$I_2(\eta,m)=\frac{\partial}{\partial\rho}\left(\rho\int_1^\lambda g^{HC}(m;x,\rho)x^2dx\right)=\frac{\partial(\eta I_1(\eta,m))}{\partial\eta}\sum_{i=0}^{6}b_i(m)\eta^i \quad (4.123)$$

$a_i(m)$ and $b_i(m)$ are given by the by Liu and Hu [75] correlation:

$$a_i(m)=a_{0i}+\frac{m-1}{m}a_{1i}+\frac{m-1}{m}\frac{m-2}{2}a_{2i} \quad (4.124)$$

$$b_i(m)=b_{0i}+\frac{m-1}{m}b_{1i}+\frac{m-1}{m}\frac{m-2}{2}b_{2i} \quad (4.125)$$

The PC-SAFT EoS is probably the one that has been the most widely used in the academic and industrial communities. Many parameters are available in the literature, including those for different type of fluids (hydrocarbons, associating fluids, polymers, pharmaceutical…). Some of the great success of PC-SAFT are described in [36, 76, 77, 78, 79, 80, 81, 82, 83].

4.4 Effect of the approximations on the chain bond formation term

The polymerization process consists in bonding monomers of a system between them to form chain molecules. A complete polymerization of system is achieved when all monomers are bonded resulting in no non-bonded monomers. This is the fundamental concept on which TPT relies on. However, they are some restrictions and approximations made in the development of the theory development, such as the single-chain approximation and the neglect of higher orders in the correlation function. The SAFT theory which is based on TPT, leads to molecular based EoS, allowing the theory to be assessed from molecular simulations before applying it to real fluids. This is of course one of the most valuable feature of these types of EoS constructions.

In TPT, the energy required to form a chain in a sea of monomers is the same as that of forming a chain in a mixture of monomers and chains when they are simultaneously present in the system. Of course, this is a crude approximation, as the monomers forming the chains already present in the system do not behave as free monomers because they are constrained by the bonds. Hence, the energy needed to form a second chain should certainly be different to that of the first chain since the external forces on the monomers of the second chain should not be the same as those acting on the monomers of the first chain and so on. This is the consequence of the single chain approximation where all clusters involving chain-chain interactions vanished, and thus, the single chain formed should behave as an ideal chain in a monomers system. The effect this approximation should be less pronounced when this polymerization process occurs at high density, as one can consider that the internal degrees of freedom of a molecule may be hidden by the effect of density. However, this is incorrect at low density limit. Hence, the TPT of Wertheim should fail in the prediction of the fluid behaviour for systems of chains with

internal degrees of freedom (s-mers>2) at the low-density limit, until this approximation is relaxed. Marshall et al. [84] went beyond this approximation by considering clusters involving dimer-dimer interactions in the graph theory and proposed a dual instead of single-chain approximation. Marshall showed that the inclusion of more interactions in the theory has an effect in lowering the compressibility factor of hard-sphere chains as well as for LJ-sphere chains, leading to better predictions when compared to Monte Carlo simulations, particularly at low densities and as the chains become larger.

The second approximation consists in removing all graphs involving more than two connected particles, leading to the first order perturbation theory (TPT1) on which most of the SAFT EoS are based. In this case all information beyond those of particle-particle interactions are lacking. It is then assumed that the energy required to link two monomers to form a dimer is exactly the same as that required to link a monomer to any other monomer belonging to a chain molecule, regardless of its position in the molecule. Hence, two main consequences emerge. The first is that the energy required to form a chain molecule is assumed to be simply the bonding energy of the dimer multiplied by the number of bonds in the molecule. The second is that, TPT1 makes no difference whether the molecule is flexible linear, rigid linear or even branched.

To better illustrate this second approximation, let's take two examples. The first consists in the formation of a trimer composed of monomers {1, 2, 3} successively bonded as illustrated in Figure 4.4. Suppose holding the first monomer fixed, and, that the second monomer can be bonded to the first regardless of the direction to form a dimer. Then, the third monomer has less accessible bonding volume on monomer (2), as there is an excluded volume due to the presence of monomer (1) (volume of monomer (1) + grey zone). Thus, the energy required to bond monomers {2,3} is different from that of bonding monomers {1, 2}. Nevertheless, one may expect that the bonding energy should converge to a plateau after certain number of bonds, as the excluded volume for the n^{th} monomer will roughly be the same as that for the $n\text{-}1^{th}$ monomer.

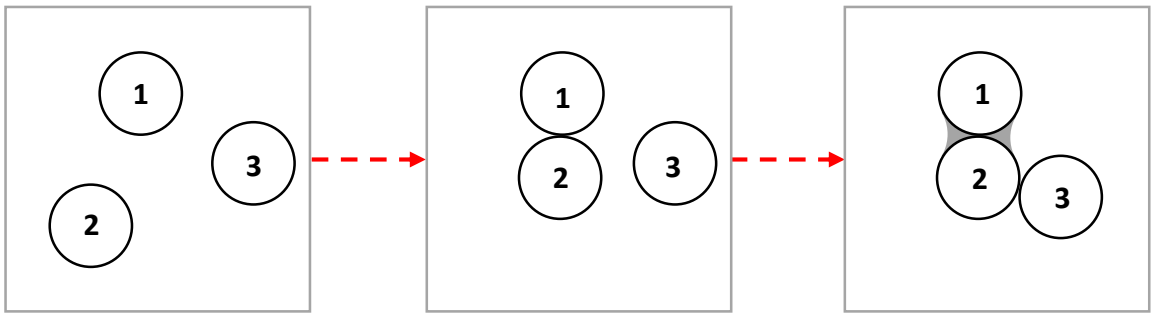


Figure 4.4: Polymerization process to form a linear trimer molecule. It highlights the excluded volume ($V_{molecule\ (1)} + V_{grey\ zone}$) for molecule (3). Consequently, the bonding energy between {3,2} is much greater that between {1,2}.

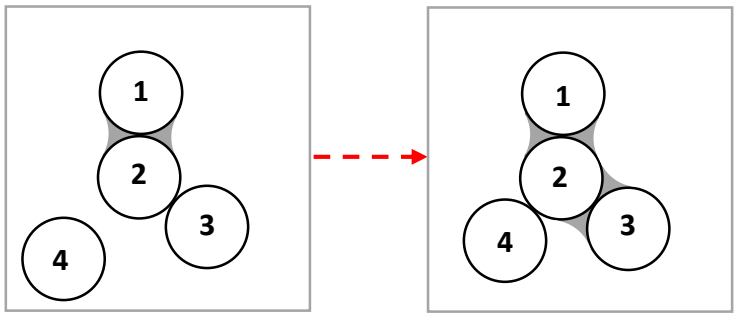


Figure 4.5: This is the continuation of the previous figure, where here it highlights that the energy required to form a branch {2,4} is much greater than that needed to form a linear 4-mer chain due to the higher excluded volume for molecule (4).

The second example is that consisting in the formation of branched molecule as illustrated in Figure 4.5. Suppose now, we want to bond a fourth monomer (4) to the second monomer (2) belonging a trimer. Here, the free volume offered to monomer (4) is less than that offered to bond {1,2} or even {2,3} due to the presence of the two monomers {1, 3} simultaneously. Consequently, forming a branch should require more energy than that needed to form a linear chain. In practice, the signature of these energy barriers in the TPT equation of state are expressed through the pair correlation function at contact which should decrease with increasing of the excluded volume.

For sake of clarity, TPT normally integrates the notion of excluded volume when introducing steric incompatibilities. However, it is the truncation at TPT1 level which leads somewhat to the omission of this excluded volume. Thus, in the spirit of TPT1, the surface of the volume offered to monomer (2) to bond on monomer (1), as shown in Figure 4.6 (a), is the same as that offered to monomer (3) to bond on monomer (1), as shown in Figure 4.6 (b). The bonded sites are fully free to move over the surface of the volume of the two monomers involved independently of the existence of other bonded sites in the same chain. Thus, when forming a chain, sites may overlap without any energetic penalty! To address this issue, the hole effect (excluded volume effect) should be included in TPT through the correlation function of the reference fluid, by including information on the next-nearest monomer, next-to-next-nearest monomer…so that the overlapping is avoided.

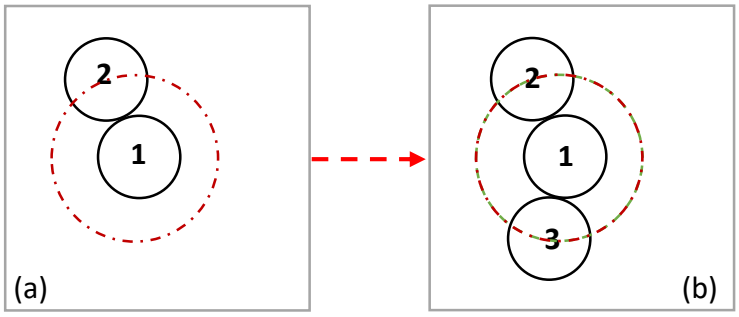


Figure 4.6: In the logic of TPT1, in (a) if a molecule (1) is hold fixed, the center of mass of molecule (2) can freely move over the red circle to bond on molecule (1). In (b), molecule (3) can also move over the green circle which is confounded with the red circle without any restriction due to the fact that there is no information between molecules (1) and (3) as they can overlap.

Thus, since TPT1 does not integrate information on the relative interactions between the segments of the same chain, a closed molecular model to TPT1 is, at some extent, that consisting in rigid chains. In the rigid chain system, intramolecular interactions account for a negligible part of the total contribution. This is because, in such molecular model, the information on the next-nearest monomer does not play an essential role, as the interactions between the first and the third monomer are screened by the second monomer (SI3), particularly for very short ranged potentials. To illustrate this, we have performed molecular simulation of a 4-mer LJ chains system with variable rigidities of the bending type. Results are shown in Figure 4.7 for the internal energy at a supercritical temperature ($T^* = 4$), and for the saturation pressure.

From Figure 4.7 (a), it is clear that for both the internal energy and the saturation pressure, the TPT1 model is closer to the chains model with the highest rigidities. Note that the computed internal energy considered here is that coming only from intermolecular interactions between monomer of different molecules. This to better mimic the TPT1-M model which does not account for the intramolecular interactions between monomers of the same chain. However, to better illustrate this latter, we have also added to the Figure 4.7 (a) the data of the total internal energy of the fully flexible chains. While the approximation of the TPT1 model could be considered satisfactory at high densities, as the intra contribution becomes less important due to the entanglement between molecules, deviations appear at medium densities and are more pronounced at low densities where the intramolecular interactions becomes important. The maximum error is expected be at the zero-density limit where the TPT1-M model predicts zero internal energy (no intramolecular interaction), which is incorrect.

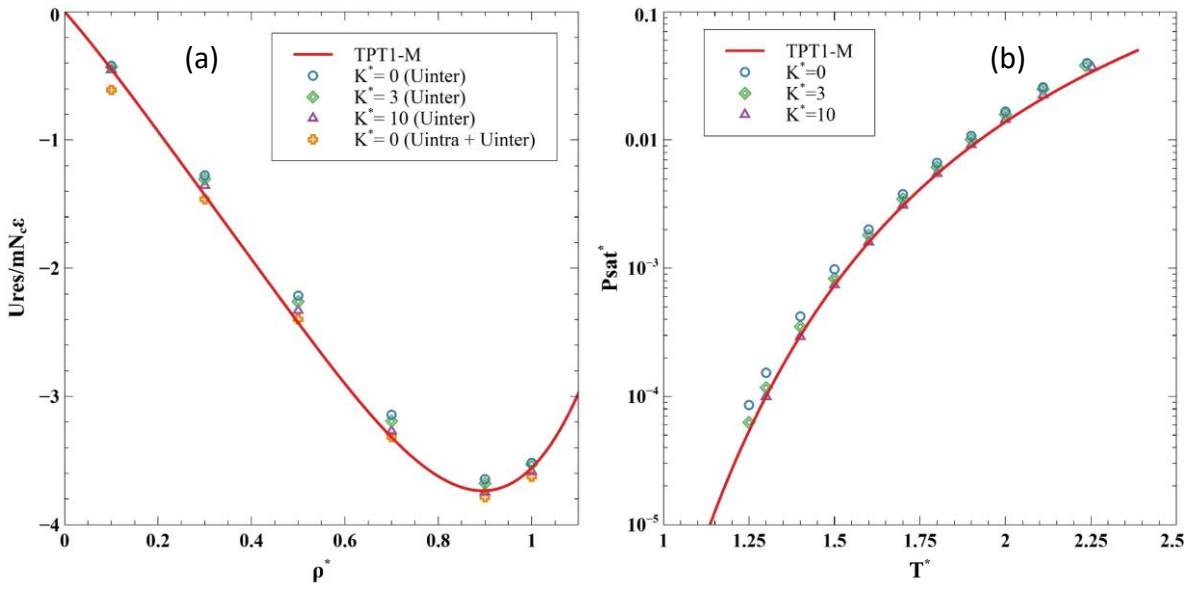


Figure 4.7: Internal energy and saturation pressure of chain of m=4 for variable rigidities vs the prediction from TPT1. (a) internal energy at T=4, (b) saturation pressure*

One way to address this issue is to go beyond the first level of TPT1 by including graphs involving more than two bonded monomers. However, high order levels are more challenging to develop theoretically as they require multi-body correlations functions. Therefore, it makes sense to limit the orders to the second order leading to TPT2. In TPT2, information on particle-particle-particle interactions is integrated, and thus, include information on the positions of the nearest and next-nearest monomers through their relative angle. The equation of state for the compressibility factor using the TPT2 is given by [8]:

$$\beta(Z - Z^{ref}) = \left(-1 + \frac{1}{\nu}\right)\left(\frac{\rho}{\Delta}\frac{\partial(\Delta)}{\partial(\rho)} + 1\right) - \frac{\nu-\tau}{\nu}\frac{\rho}{\lambda}\frac{\partial(\lambda)}{\partial(\rho)} \tag{4.126}$$

where $\frac{\nu-\tau}{\nu}\frac{\rho}{\lambda}\frac{\partial(\lambda)}{\partial(\rho)}$ defines the second order term, and the quantity $\frac{\nu-\tau}{\nu}$ comes from the self-consistent density relations for the TPT2 [8] given by:

$$\frac{\nu-\tau}{\nu} = \frac{1}{2} - \frac{(1+4\lambda-4\lambda\nu^{-2})^{\frac{1}{2}}}{2(1+4\lambda)} - \frac{2\lambda}{(1+4\lambda)\nu} \tag{4.127}$$

The evaluation of λ is given for the fully flexible chains by:

$$\lambda = \int_{\alpha_{min}}^{\pi}\left(\frac{g\left(\sigma,\sigma,2\sigma\,sin\left(\frac{1}{2}\alpha\right)\right)}{g^2(\sigma)} - 1\right)\xi(\alpha)sin(\alpha)d\alpha \tag{4.128}$$

The lower limit of α_{min} is set to $\frac{\pi}{3}$ due to steric incompatibilities in which the too small angles are prohibited. ξ is a distribution of angles between three successive bonded monomers given by:

$$\xi(\alpha) = \begin{cases} 0 & for\ 0 < \alpha < \pi/3 \\ 2/3 & for\ \pi/3 < \alpha < \pi \end{cases} \tag{4.129}$$

For a fixed rigid bond angle, characterizing a frozen rigid chain molecule, λ is given by:

$$\lambda = \frac{g\left(\sigma,\sigma,2\sigma\, sin\left(\frac{1}{2}\alpha\right)\right)}{g^2(\sigma)} - 1 \qquad (4.130)$$

By doing so, Wertheim [8] showed that TPT2 yields better predictions than TPT1 for the fully flexible chains, especially for large chains. Using TPT2, it is possible to differentiate between fully flexible chains and rigid ones by constraining the bond angle between three successive bonded monomers [85], or also between linear and branched chains [51]. More recently, Zmpitas developed a third order approximation TPT theory (TPT3). He found that only a small improvement was obtained when compared to TPT2, suggesting that considering higher orders beyond TPT2 does not seem efficient as the improvement obtained is negligible compared to the required efforts. An interesting comparison to molecular simulations can be found in the article of Marshall et al. [84] on the predictions of the compressibility factor between TPT1 and TPT2 and their respective versions when considering the dual-chain approximation.

4.5 A simple alternative to higher orders of the TPTn …

Since the theory of Wertheim does not impose the geometry of the reference fluid to be spherically symmetrical or even the isotropy of the potential interaction, it is possible to use any type of reference fluid unless the pair correlation function is known for this fluid. Thus, instead of a monomer, a dimer, trimer or any s-mer fluid can be used as a reference fluid. The use of this idea was motivated by the results obtained by Honnell and Hall [86] using a Generalized Flory Dimer theory (GFD). The authors showed that, the modelling of the thermodynamic behaviour of fully flexible chains is largely improved when a dimer reference fluid is used instead of a monomer used in the Generalized Flory theory (GF) [87]. In the same spirit, an alternative to TPT1, based on the monomer reference fluid often referred to as TPT1-M (M for Monomer), was developed, based on the dimer reference fluid, referred as TPT1-Dimer (TPT1-D). This theory was independently developed by both Chang and Sandler [88] and Ghonasgi and Chapman [89] in 1994 for hard chains. It was extended few years later by Johnson in 1996 for LJ-chains [90].

TPT1-D showed an improvement compared to TPT1-M for the fully flexible hard chains and the LJ chains over many properties. For example, the second virial coefficient and the compressibility factor of hard chains are better predicted with TPT1-D than with TPT1-M. More interestingly, TPT1-D theory provides also predictions that are superior to what could be obtained using TPT2-M or the GFD theory. However, despite the good capabilities of TPT1-D, surprisingly, this theory has not extensively been applied to real fluids whereas TPT1-M has been used extensively.

The secret of the TPT1-D success is that the loss of the structural information in the monomer based theory is now better recovered through the dimer theory. While the TPT1-M integrates information only on the nearest neighbours along the chain, the TPT1-D integrates information on the nearest, the next-nearest and the next-to-next-nearest neighbour along the chain. To illustrate the dimer perturbation theory, we follow the scheme proposed by Chang and Sandler [88]. Figure 4.8 shows two ways to transform a system composed of monomers to a system of chain molecules composed of 4-mers. The way (1) is the one using TPT1-M where chains are formed by assembling monomers one after another to form a chain of 4-mers and the process

is repeated until the whole system is polymerized into chains of 4-mers (keeping in mind the single chain approximation, which implies that the chains of 4-mers formed do not interact between them). With TPT1-D, this polymerization process is split into two steps. First, the system is completely dimerized using TPT1-M (remember that TPT1-M is perfectly adapted to model dimerization process). Then, the system is polymerized in a second step by assembling the dimers to form chains of 4-mers system.

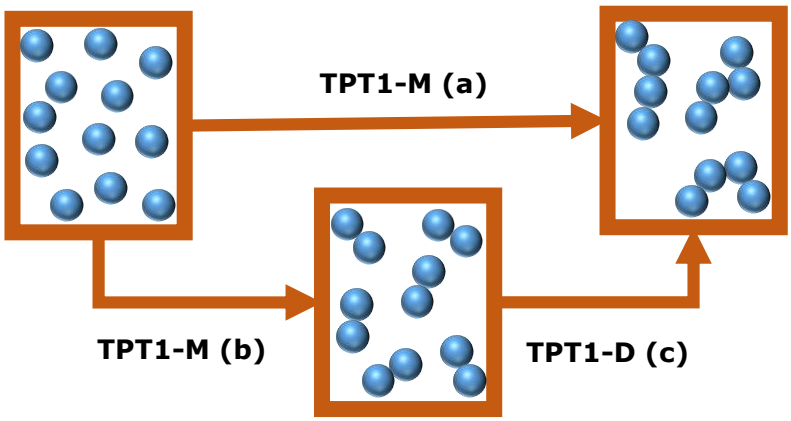


Figure 4.8: Polymerization of monomers to form chains of 4-mer fluid. Path (a) follows the TPT1 scheme where the formation of chain depends only on the information included at the monomer level, thus, no information is included beyond nearest bond. Path (b and c) follows the TPT1-D scheme, where first, the dimer is formed through the TPT1-M. Then, the 4-mer chains fluid is formed by associating dimerization of these dimer molecules.

It is the second step that makes the whole difference between TPT1-M or TPT2-M and TPT1-D. In fact, the bonding energy for associating two dimers, is now dependent on the presence of the attached monomers in both sides of the dimers, eliminating then all ghost (impossible) configurations, corresponding to a complete folding of a molecule. The Helmholtz free energy of irreversible association between dimers is given by

$$A_{bond}^{D} = -k_B T ln\left(g_{ee}^{D}(\sigma, T)\right) \tag{4.131}$$

where $g_{ee}^{D}(\sigma, \mathrm{T})$ is the end-end contact distance radial distribution function of two dimers in contact.

In practice, $g_{ee}^{D}(\sigma)$ can be computed using molecular simulation of dimer system. $g_{ee}^{D}(\sigma)$ is an averaged quantity over all possible configurations for two dimers at contact and over all positions and orientations of the surrounding molecules. During the simulation, the two adjacent monomers prevent the surrounding molecules of the system to be close enough to bond (the hole effect). As a consequence $g_{ee}^{D}(\sigma) < g_{ee}^{M}(\sigma)$. The best illustration of that, is the low density limit where $g_{ee}^{D}(\sigma, \rho \to 0) < 1$ leading to a positive free energy $A_{bond}^{D}(\sigma, \rho \to 0) > 0$ while this is zero in the case of monomers $A_{bond}^{M}(\sigma, \rho \to 0) = 0$. Therefore, an additional energy is required to overcome the hole effect.

Note that TPT1-D should normally be more accurate for systems of chains composed of an even number of monomers as the chains are built by assembling couple of dimers. However, one may imagine that accurate prediction for those with an odd number m could be obtained

by interpolation between that of $m-1$ and that of $m+1$ lengths, as their thermodynamic behaviour should lie in-between.

The generalization into any reference term composed of r-mers is given by:

$$\frac{A_{res}}{N_c k_B T} = \frac{m}{r}\frac{A_{res}^{r-mer}}{N_s k_B T} + \left(1-\frac{m}{r}\right) ln\left(g_{ee}^{r-mer}(\sigma)\right) \tag{4.132}$$

Introducing the reduced units $\mathrm{A}_{\mathrm{res}}^{*,\mathrm{chain}} = \frac{\mathrm{A_{res}}}{N_c\epsilon}$ and $\mathrm{A}_{\mathrm{res}}^{*,\mathrm{r-mer}} = \frac{\mathrm{A}_{\mathrm{res}}^{\mathrm{r-mer}}}{\mathrm{N_s}\epsilon}$ allows to rewrite it as:

$$A_{res}^{*,chain} = \frac{m}{r}A_{r-mer,res}^{*} + \left(1-\frac{m}{r}\right) T^{*} ln\left(g_{ee}^{r-mer}(\sigma)\right) \tag{4.133}$$

which, if a dimer is taken as a reference term, becomes:

$$A_{res}^{*,chain} = \frac{m}{2}A_{D,res}^{*} + \left(1-\frac{m}{2}\right) T^{*} ln\left(g_{ee}^{D}(\sigma)\right) \tag{4.134}$$

Here $A_{D,res}^{*}$ is obtained from TPT1-M and is given by:

$$A_{D,res}^{*,chain} = 2A_{M,res}^{*} - T^{*} ln\left(g_{ee}^{M}(\sigma)\right) \tag{4.135}$$

By injecting it into $\mathrm{A}_{\mathrm{res}}^{*,\mathrm{chain}}$ it gives:

$$A_{res}^{*,chain} = mA_{M,res}^{*} - \frac{m}{2}T^{*} ln\left(g_{ee}^{M}(\sigma)\right) + (1-\frac{m}{2})T^{*} ln\left(g_{ee}^{D}(\sigma)\right) \tag{4.136}$$

where $g_{ee}^{D}(\sigma)$, in the case of fully flexible LJ chains, is given by Johnson [90]. It is given by a correlation fitted to the end-end radial distribution function of two dimers in contact obtained from molecular simulations (the fitting coefficients are given in the appendix 7.A.2):

$$g_{ee}^{D}(T^{*},\rho^{*},\sigma) = g_{00}^{D}(T^{*},\sigma) + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij}\rho^{*i}\, T^{*(1-j)} \tag{4.137}$$

where $g(\mathrm{T}^{*},\sigma)$ is the low density limit given by

$$g_{00}^{D}(\sigma,T^{*},\rho^{*}\to 0) = \sum_{j=1}^{7} a_j T^{*(1-j)} \tag{4.138}$$

In terms of the residual internal energy, TPT1-D is expressed as follow (with $U_{res}^{*,chain} = \frac{U_{res}}{N_c\epsilon}$):

$$U_{res}^{*,chain} = mU_{M,res}^{*} + \frac{m}{2}T^{*2} ln\left(g_{ee}^{M}(\sigma)\right) - (1-\frac{m}{2})T^{*2} ln\left(g_{ee}^{D}(\sigma)\right) \tag{4.139}$$

where $U_{M,res}^{*}$ is the residual internal energy of the monomer. Note that in the results we will show, the residual internal energy is expressed per monomer, so $U^{*} = U_{res}^{*} = \frac{U_{res}^{*,chain}}{m} = \frac{U_{res}}{mN_c\epsilon}$.

The residual entropy per monomer is given by:

$$S^{*} = S_{res}^{*} = \frac{S_{res}^{*,chain}}{mN_c k_B} = \frac{U_{res}^{*} - A_{res}^{*}}{T^{*}} \tag{4.140}$$

The pressure is given by:

$$P^{*} = P_{M}^{*} + \rho^{*}\left(\frac{1-m}{m}\right) + \frac{T^{*}\rho^{*}}{2}\left(\left(\frac{2-m}{m}\right)\frac{\partial ln\left(g_{ee}^{D}(T^{*},\rho^{*},,\sigma)\right)}{\partial ln(\rho^{*})} - \frac{\partial ln\left(g^{M}(T^{*},\rho^{*},\sigma)\right)}{\partial ln(\rho^{*})}\right) \tag{4.141}$$

where P_{M}^{*} is the pressure of the monomer.

In the case of LJ based TPT, the performance of TPT1-D has shown very interesting improvement compared to TPT1-M in the monophasic phases by Johnson in the original article [90]. Few years later, Blas and Vega extended the comparison to the LVE properties [91] where the difference in terms of performance was even more noticeable. This is because the LVE properties consists in a severe test for any equation of state, as their prediction accuracy strongly depend on how good the model is in both vapour and liquid phases simultaneously. TPT1-D was found to describe better the saturated liquid phase, but, the highest improvement was found for the vapour phase. This is because of the integration of the intramolecular interactions for which the vapour property is highly sensitive to. However, the benchmark between TPT1-M and TPT1-D for the predictions of the internal energy as well as of the entropy has never been done at equilibrium conditions. Thus, we found interesting to check how good is the TPT1-D relatively to TPT1-M. For this, we performed extensive Monte Carlo simulations for chains consisting of different lengths including 1-mer, 2-mers, 3-mers, 4-mers, 8-mers and 16-mers. Entropy was obtained using Widom insertion method, by performing a test of at least 100 insertions per cycle. The comparison for the calculation of the LVE has already been done in [91], and thus, some of the LVE results we provide here are just for making the illustration more complete.

As shown in Figure 4.9 (a and b), for the short to medium chains (4-mers, 8-mers and 16-mers), the critical properties are clearly not well predicted using the TPT1-M, while better (even if not perfect) estimates are obtained with the TPT1-D. Regarding the liquid phase, the improvement is noticeable for the 8 and 16-mer chains. It is interesting to note that the inclusion of the structural information in the TPT1-D model leads to a slope derivative of the liquid phase in better agreement with the simulations data. Thus, one may expect also better performance on the prediction of the isothermal compressibility. Concerning the vapour phase, high improvement is obtained with the TPT1-D due to the structural information for which this phase is very sensitive. Consequently, as expected, the vapour pressure is better predicted by the TPT1-D theory.

Another interesting benchmark is that shown in Figure 4.9 (c and d), for the internal energy and entropy at saturation conditions. The internal energy is explored for chains from the monomer to the hexadecamer obtained from the same simulations for the equilibrium properties. The comparison on the entropy was limited to small chains up to the 4-mer, because this property is very demanding in terms of CPU. These figures show that excellent predictions are obtained for the internal energy and the entropy as well with the TPT1-D, with noticeable deviations appearing for the longest chains in the vapour phase.

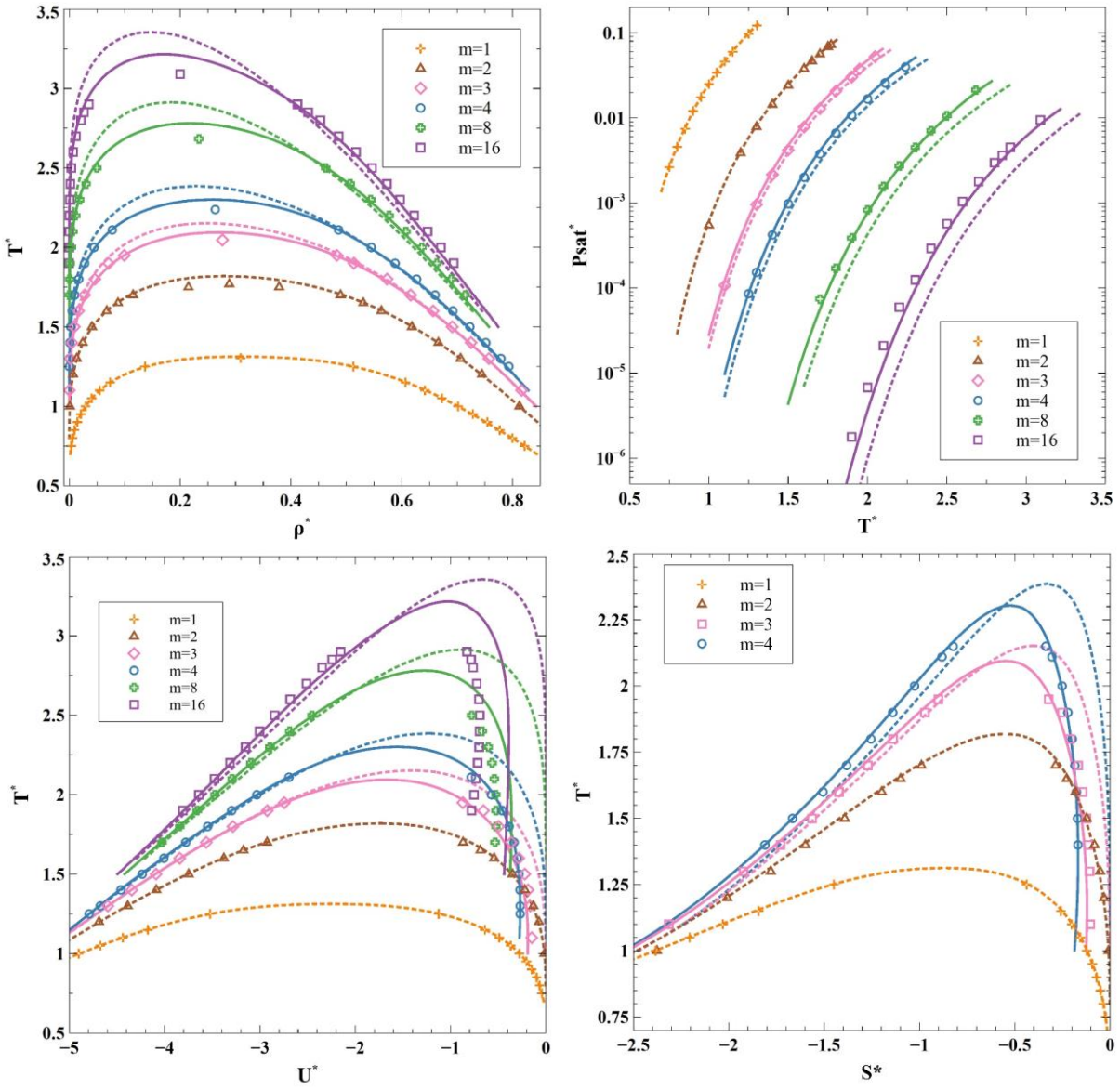


Figure 4.9: Comparison between the exact molecular simulations of the fully flexible chains model (m=1, 2, 3, 4, 8, 16) at saturation conditions vs the predictions of the TPT1-M (dotted lines) and the TPT1-D (solid lines) models. (a) LVE, (b) saturation pressures, (c) internal energy, (d) entropy.

The 4-mer chain is very accurately described for both properties, due to the fact that the model integrates exact zero-density limit pair correlation function. The TPT1-M is excellent for the monomer and the dimer which is not surprising because of the lack of any intramolecular interactions for these two molecules. However, deviations are more and more pronounced, particularly for the vapour phase, as the chain length increases. This is due to the fact that TPT1-M erroneously predicts zero internal energy for all chains. The liquid side for these properties is not badly predicted by the TPT1-M, due to the fact that the effect of the intramolecular interactions is reduced when the density increases. This is the consequence of the entanglement between the molecules in the system. Until this stage, we have shown that the TPT1-D provides better predictions of some equilibrium properties of the model fluid when compared to TPT1-M. The derivative properties were not included, but, one may expect also better results, as it

will be shown latter on. Mixtures as well would be interesting to be studied with both TPT1-M and TPT1-D.

Now, one may wonder whether the use of a dimer as a reference term is the best option? In other words, does the use of higher reference terms, such as a 3-mer (TPT1-T) or a 4-mer (TPT1-Q) would allow the model to yield better prediction than the TPT1-D? It is an interesting question, and to the best of our knowledge such high orders TPT1-r-mer versions do not exist yet in the literature. To answer this question, we have evaluated the capabilities of TPT1-T and TPT1-Q. However, we limited the evaluation to the zero-density limit conditions, as we think that this may be enough to draw a global conclusion, as done between TPT1-M and TPT1-D. The two models are given respectively using the TPT1-r-mer as follows:

$$A_{res}^{*,chain,TPT1-T} = \frac{m}{3} A_{T,res}^{*} + (1 - \frac{m}{3}) T^{*} ln(g_{ee}^{T}(\sigma)) \tag{4.142}$$

$$A_{res}^{*,chain,TPT1-Q} = \frac{m}{4} A_{Q,res}^{*} + (1 - \frac{m}{4}) T^{*} ln(g_{ee}^{Q}(\sigma) \tag{4.143}$$

The two reference terms are obtained using TPT1-M (one can use TPT1-D as well), which gives:

$$A_{T,res}^{*,chain} = 3A_{M,res}^{*} - 2T^{*} ln(g_{ee}^{M}(\sigma)) \tag{4.144}$$

$$A_{Q,res}^{*,chain} = 4A_{M,res}^{*} - 3T^{*} ln(g_{ee}^{M}(\sigma)) \tag{4.145}$$

Injecting them into $A_{res}^{*,chain,TPT1-T}$ and $A_{res}^{*,chain,TPT1-Q}$ respectively, one obtains the following final formulas:

$$A_{res}^{*,chain,TPT1-T} = mA_{M,res}^{*} - \frac{2m}{3} T^{*} ln(g_{ee}^{M}(\sigma)) + (1 - \frac{m}{3}) T^{*} ln(g_{ee}^{T}(\sigma)) \tag{4.146}$$

$$A_{res}^{*,chain,TPT1-Q} = mA_{M,res}^{*} - \frac{m}{4} T^{*} ln(g_{ee}^{M}(\sigma)) + (1 - \frac{m}{4}) T^{*} ln(g_{ee}^{Q}(\sigma)) \tag{4.147}$$

The zero-density limit for any r-mer is obtained from the averaged Boltzmann factor:

$$g_{ee}^{r-mer}(T^{*}, \sigma) = < exp(-\beta U_{12}(\sigma)) >_{\theta_1 \theta_2} \tag{4.148}$$

The brackets denote an unweighted average over all orientations keeping the end to end distance equal to σ. $U_{12}(\sigma)$ is the total potential interaction between the sites of the two r-mers in contact. This can be evaluated numerically by the use of three dimensional integration as in [90], by the use of any spherical mesh gridding method using any technique for generating many regular configurations, or by Monte Carlo simulation of a single molecule consisting of two bonded r-mers.

In our case we have chosen the latter method to obtain both $g_{ee}^{T}(\sigma, T^{*}, \rho^{*} \to 0)$ and $g_{ee}^{Q}(\sigma, T^{*}, \rho^{*} \to 0)$ using the relationship between the mean free energy and the correlation function:

$$A_{res}^{*,chain}(T^{*}, \rho^{*} \to 0) = -T^{*} ln\left(g_{ee}^{r-mer}(\sigma, T^{*})\right) \tag{4.149}$$

To evaluate the performance of the theories, we have performed molecular simulations for different LJ-chains at zero-density. A single chain consisting of two r-mers in contact is put in

the box and simulated in the NVT ensemble using Monte Carlo technique. $A_{res}^{*,chain}(T^*,\rho^* \to 0)$is obtained by the Widom particle test insertion method [92,93] given by:

$$\mu_i^* = -T^* ln < \frac{V}{(N_i+1)\,\Lambda_i^3} exp(-\beta U_{test}) >_{\theta_1\theta_2} \quad (4.150)$$

Where F is the partition function of the NVT ensemble, $Utest$ is the variation of the total potential energy by inserting the test particle, $\langle\ \rangle$ refers to the NVT ensemble average over all test insertions and Λ is the De Broglie wavelength already introduced in chapter2.

In the limit of zero-density, one may write:

$$A_{res}^{*,chain}(T^*,\rho^* \to 0) = \mu_{res}^{*,chain}(T^*,\rho^* \to 0) \quad (4.151)$$

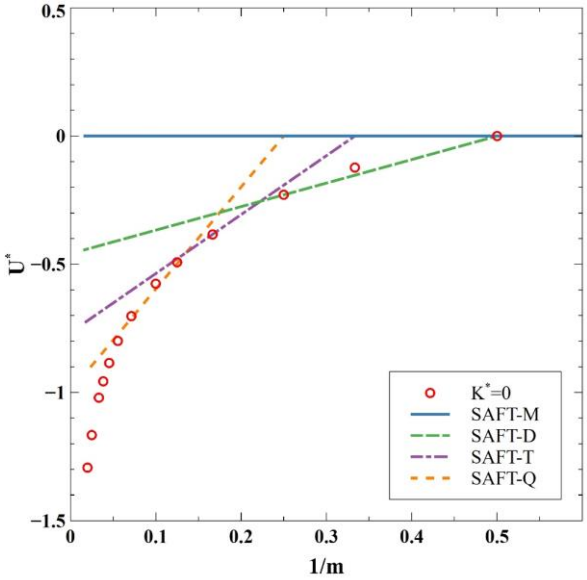


Figure 4.10: Zero-density limit internal energy of fully flexible chains up to m=50 at T=2 obtained from monte carlo simulations vs the predictions using the SAFT-M, SAFT-D, SAFT-T and SAFT-Q.*

As shown in Figure 4.10, while the TPT1-M predicts zero internal energy for all chain lengths, the TPT1-D, TPT1-T and TPT1-Q predicts non-zero values for chains lengths superior to 3-mers and 4-mers respectively. TPT1-D gives exact internal energy for a chain of 4-mers. Similar results are obtained for TPT1-T and TPT1-Q for chains of 6-mers and 8-mers respectively. This is because the models are forced to reproduce the exact zero-density limit through the use of the exact end-end correlation functions of two reference r-mers in contact. Thus, the longest the reference term introduced is, the better the prediction for longer chains. However, the deviations for chains with length below that of the reference term are very high, due to the use of TPT1-M which erroneously predicts zero internal energy at zero-density limit whatever the length of the chain. The use of a TPT1-D for the determination of the reference terms in TPT1-T and TPT1-Q would probably address this issue, but, care should be taken on the way this would be done. If not, a systematic shift of the internal energy will appear (towards low energies) due to the fact that the $g_{ee}^{D}(\sigma,T^*,\rho^* \to 0)$ involved in TPT1-D does not vanish at $\rho^* \to 0$.

An interesting point to note is regarding the linearity dependence of the internal energy with the inverse chain length $\frac{1}{m}$ exhibited by all approaches. It is in fact an intrinsic constraint of the model in itself, independently of which reference term is used. The expression for the internal energy for any r-mer reference term is given by:

$$U_{res}^{*,chain} = \frac{m}{r} U_{res}^{*r-mer} + \left(\frac{m}{r} - 1\right) T^{*2} \frac{\partial ln(g_{ee}^{r-mer}(\sigma))}{\partial T^*} \tag{4.152}$$

At the zero-density limit, the total Helmholtz free energy per monomer ($U_{res}^* = \frac{U_{res}}{mN_c\epsilon}$) gives:

$$U_{res}^*(\sigma, T^*, \rho^* \to 0) = \left(\frac{1}{r} - \frac{1}{m}\right) T^{*2} \frac{\partial ln(g_{ee}^{r-mer}(\sigma, T^*, \rho^* \to 0))}{\partial T^*} \tag{4.153}$$

from the equation, we can see that the model imposes the linear proportionality:

$$U_{res}^* \propto -\frac{1}{m} \tag{4.154}$$

4.6 Does Wertheim TPT includes hard intramolecular interactions (vibrations)?

Despite all improvements made by the TPT theory in modelling the behaviour of chains fluids by incorporating some structural information through the use of a dimer as a reference term (TPT1-D), or by using a second order perturbation theory (TPT2), there are still many open issues. One of these is the contribution to the molecular vibrations due to the presence of bending or torsional potential types, i.e. when dealing with chains of variable rigidity. Indeed, these internal degrees of freedom are known to strongly affect the structure of the fluid, with rigid chains exhibiting different behaviours than fully flexible ones.

The integration of these contributions requires the knowledge of the positions of three successive bonded monomers for a bending potential and the positions of four successive monomers for a torsional potential. TPT1-M only incorporates information on the position of the nearest bonded monomer, discarding any possibilities of defining structural information. On the other hand, TPT2 incorporates information on the positions of the nearest and next-nearest monomers through their relative angle, allowing for the definition of an angle potential. However, this corresponds to a certain molecular structure related to an equilibrium angle $\overline{\theta}_{eq}$ resulting in an infinite stiffness constant *K* for a completely rigid chain or a zero stiffness constant if this equilibrium angle is relaxed, and replaced by a distribution of angles $\theta \geq \frac{\pi}{3}$, corresponding then to a fully flexible chain. Hence, TPT2 can model only these two limits, a completely fully rigid chain or a fully flexible chain. Thus, the TPT2 scheme is not able to model semi-rigid chains, where the molecule has the possibility to vibrate around the equilibrium angle. Moreover, it is not possible to define a torsional potential as TPT2 does not incorporate information on the next-to-next-nearest neighbour monomer (the fourth monomer).

TPT1-D is an excellent candidate to describe semi-rigid chains as it allows to go beyond some of the limitations found with TPT1 and TPT2. In fact, as already explained previously, TPT1-D consists in assembling dimers instead of monomers, and so, information on the four successive bonded sites are known simultaneously, offering then the possibility of integrating bending and torsion potentials.

Indeed, as shown in the previous chapter, the semi-rigid LJ chain, when adequately parametrized, is a very interesting molecular model as it allows to obtain simultaneously accurate thermodynamic and transport properties of fluids. Thus, in this part of thesis, the aim is to develop an equation of state for the semi-rigid LJ chains based on the TPT1-D (SAFT-D). However, to do so, one needs to bring a modification into it, so as to integrate the contributions of the vibrational type. Here, only the case with contribution due to the bending potential of the harmonic type is dealt with, but, in principle, the approach could be extended to torsional contribution.

To model semi-rigid chains, it is useful to revisit the way SAFT-D incorporates structural information. It was previously mentioned that this is introduced through the correlation function by considering the excluded volume (hole effect) induced by the presence of existing bonds in the chain. Specifically, this hole effect is present in real molecules, but not in the TPT1-M, which limits its accuracy to some extent, and is taken into account by eliminating the ghost configurations (intramolecular overlaps) through the use of a dimer as a reference term instead of a monomer. As a result, the correlation function is lowered. In the same vein, the introduction of a bending potential will lead to an increase in ghost configurations. To illustrate this idea, consider the case of a fully rigid chain. In this case, the chain molecule needs to overcome an infinite energetic barrier to explore configurations other than the equilibrium making it highly unlikely for such configurations to occur and therefore should be discarded. The case of a semi-rigid chain is quite similar, but the barrier can be overcome depending on the stiffness constant and the thermal energy of the system. This allows more accessible configurations to occur compared to the fully rigid case, but fewer than that of a fully flexible chain, unless the stiffness constant is relaxed. In terms of the correlation function, one may then expect that it will decrease as the rigidity increases due to the elimination of additional ghost configurations compared to the fully flexible chains.

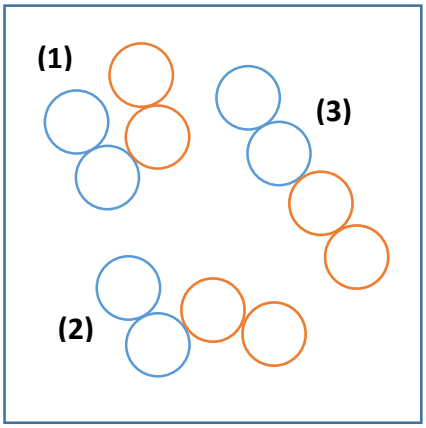


Figure 4.11: Some of the possible configurations that two dimers in contact may hold during the path (c).

Consider the case of chains given in Figure 4.11, consisting of the assembly of two dimers to form chains of 4-mer. During the simulation, the assembled molecule will sample the whole configurational space, where each configuration will occur with a certain probability corresponding to the Boltzmann factor. For instance, in the case of fully flexible chains,

configurations {1; 2; 3} will occur due to the possible self-folding of the molecule. However, in the case of fully rigid linear chains, it is clear that configuration {3} will have the greatest probability while configurations {1; 2} will see their occurring probability reduced to zero due to the high energetic barrier of the bending potential. Thus, during the simulation of dimer system each time a contact happens can be seen as a contribution to the energy required to form a chain molecule at certain temperature, density and configuration. Repeating this process during many time steps (or moves) will result in an averaged energy for making a bond at the thermodynamic conditions of the system, averaged over all the configurations occurring during the simulation. With this in mind, one may imagine that if the chain molecule formed possesses an intramolecular energy of the bending (or torsion) type, each configuration should be scaled by a probability taking into account this constrain. Note that, in the simulation, a bending or torsion cannot be defined for a system of dimers so that the constrain can be included intrinsically in the potential interaction. Thus, in order to scale the configurations, the Boltzmann factor has been used to weight each configuration so that its contribution to the correlation function will take into account a bending potential contribution to the formation of chains. This is given by:

$$Exp(-\beta U_{12}^{bend}(\theta_1, \theta_2)) \tag{4.155}$$

which is the probability that measures the ratio between the molecular bending energy induced by two dimers in contact relatively to the system thermal energy. θ_1, θ_2 are the bending angles of the 4-mer chain molecule formed.

For a fully flexible chain, the correlation function (in our case is the radial distribution function) is given by:

$$g_{ee}^{2-mer}(r = \sigma,\ T^*, \rho^*) = \frac{\langle n(r=\sigma)\rangle}{\rho V_{shell}} \tag{4.156}$$

where the brackets indicate an average over the simulation time (or moves) of the total number of particles found in the surrounding shell volume of radius $r = \sigma$. The total number of particles is obtained by (using the histogram method):

$$n(r) = n(r) + 1 \tag{4.157}$$

where "1" refers to each contact, thus, to each configuration.

Now, for the semi-rigid chains, each configuration is weighted by the Boltzmann factor, and thus, will not contribute by "1", but instead by:

$$n(r) = n(r) + 1Exp\left(-\beta U_{12}^{bend}(\theta_1, \theta_2)\right) \tag{4.158}$$

where the fully flexible case is obtained by taking:

$$U_{12}^{bend}(\theta_1, \theta_2) = 0 \tag{4.159}$$

i.e. a stiffness constant equal to zero.

Finally, the zero-density limit is obtained numerically by averaging the Boltzmann factor over all possible configurations of two dimers in contact:

$$g_{00}^{r-mer}(\sigma,\ T^*,K^*) = \langle exp\left(-\beta U_{12}^{Vdw+bend}(\theta_1,\theta_2)\right)\rangle_{\theta_1\theta_2} = \frac{\int exp\left(-\beta U_{12}^{Vdw+bend}(r,\theta_1,\theta_2)\right) sin(\theta_1) d\theta_1 d\theta_2}{\int sin(\theta_1) d\theta_1 d\theta_2} \quad (4.160)$$

4.7 The model

As done for the fully flexible TPT1-D, the radial distribution function has been obtained for different rigidities from molecular dynamics simulations of dimers system. The runs were performed at various temperatures and densities covering the following thermodynamic range $0.9 \leq T^* \leq 10$ and $0.1 \leq \rho^* \leq 1.0$. The simulations were run in the NVT ensemble using the Nosé-Hoover thermostat. The simulation is split into two steps, where the system is equilibrated for a minimum of $1.5x10^6$ time steps and followed with a production step for a minimum of $2.0x10^6$ time steps. The transient regime for the RDF is very small ($0.1x10^6$) allowing the rdf to converge very quickly (less than $0.5x10^6$ steps), and thus, the data obtained exhibited extremely small average error. The time step is taken equal to $\delta t^* = 2x10^{-3}$ for $T^* < 5$, and $\delta t^* = 10^{-3}$ for higher temperatures. The data have been fitted to the following correlation:

$$g_{ee}^{D}(T^*,\rho^*,K^*,\sigma) = g_{00}^{D}(T^*,K^*,\sigma) + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij}(K^*)\rho^{*i}\, Ln(1+T^*)^{(1-j)} \quad (4.161)$$

where

$$c_{ij}(K^*) = \sum_{i=1}^{5}\sum_{j=1}^{5}\sum_{l=5i-4}^{5i+1} \alpha_{lj}\, Ln(1+K^*)^{f(l-5(i-1))} \quad (4.162)$$

and

$$f(x) = \ q_1(x-1)^3 + \ q_2(x-1)^2 \ + \ q_3(x-1) \quad (4.163)$$

the coefficients α_{lj} as well as the q_{i} coefficients are given in the appendix 7.A.2 correlate the simulation data with an AAD of 0.15% over the temperature and density range $0.9<T^*<20$, $0<\rho^*<1.2$, $0<K^*<100$.
The $g_{00}^{D}(T^*,K^*,\sigma)$ is the low density limit given by:

$$g_{00}^{D}(T^*,K^*,\sigma) = g_{00}^{D}(T^*,K^*\to 0,\sigma)/Exp(\sum_{i=1}^{7}\sum_{j=1}^{7} v_{ij} Ln(1+K^*)^i\, Ln(1+T^*)^{(1-j)}) \quad (4.164)$$

where v_{ij} are given in the appendix 7.A.2 correlate the simulation data with an AAD of 0.098% over the temperature and density range $0.9<T^*<20$, $0<K^*<100$.
The function $g_{00}^{D}(T^*,K^*\to 0,\sigma)$ is similar to that provided by Johnson [90], but extended to the following temperature range [0.9, ∞[using:

$$g_{ee}^{D}(T^*,K^*\to 0,\sigma) = \frac{a+cLn(T)+eLn(T)^2+gLn(T)^3}{1+bLn(T)+dLn(T)^2+fLn(T)^3+hLn(T)^4} \quad (4.165)$$

The coefficients (*a,b, c, d, e, f, g, h*) are given in the appendix 7.A.2 and correlate the data with an AAD of 0.027% over the temperature range $0.9<T^*<\infty$.

The high number of fitting constants is due to the high non-linearity exhibited by the radial distribution function, which of course, is not suitable from the computational aspect. Thus, in practice, it is recommended to perform a pre-calculation of the matrix $c_{ij}(K^*)$ for the desired rigidity K^* of interest to avoid heavy calculation at each iteration. As a final comment, a hard

sphere model would be probably a better choice because the problem will be reduced to only two variables instead of three for the Lennard-Jones model.

To summarize, the SR-SAFTD model is written in terms of the Helmholtz free energy as:

$$A_{res}^{*,chain} = mA_{M,res}^{*} - \frac{m}{2}T^{*}ln\left(g^{M}(T^{*},\rho^{*},\ \sigma)\right) + (1-\frac{m}{2})T^{*}ln\left(g_{ee}^{D}(T^{*},\rho^{*},K^{*},\ \sigma)\right) \quad (4.166)$$

where the monomer contribution is obtained from the modified-BWR EoS of Johnson et al. [44] given by:

$$A_{M,res}^{*} = \sum_{i=1}^{8}\frac{a_i\rho^{*i}}{i} + \sum_{i=1}^{6}b_iG_i \quad (4.167)$$

where a_i and b_i are temperature dependent parameters and G_i are density dependent parameters, given in the appendix 7.A.2.

g_{ee}^{M} is the monomer-monomer radial distribution function obtained from the work of Johnson et al. [32] and given by:

$$g^{M}(T^{*},\rho^{*},\ \sigma) = 1 + \sum_{i=1}^{5}\sum_{j=1}^{5}a_{ij}\rho^{*i}\,T^{*(1-j)} \quad (4.168)$$

The regression parameters a_{ij} are given in the appendix 7.A.2.

The residual internal energy is obtained by:

$$U_{res}^{*,chain} = -T^{*2}\left(\frac{\partial\left(A_{res}^{*,chain}/T^{*}\right)}{\partial T^{*}}\right)_{\rho^{*},Nc} \quad (4.169)$$

which yields:

$$U_{res}^{*,chain} = mU_{M,res}^{*} + \frac{m}{2}T^{*2}\frac{\partial ln(g^{M}(T^{*},\rho^{*},\sigma))}{\partial T^{*}} + \left(\frac{m}{2}-1\right)T^{*2}\frac{\partial ln(g_{ee}^{D}(T^{*},\rho^{*},K^{*},\sigma))}{\partial T^{*}} \quad (4.170)$$

where $U_{M,res}^{*}$ is the residual internal of the monomer.

In this work, the results for the residual internal energy will be given per number of monomer, thus:

$$U_{res}^{*} = U_{M,res}^{*} + \frac{1}{2}T^{*2}\frac{\partial ln\left(g^{M}(T^{*},\rho^{*},\sigma)\right)}{\partial T^{*}} + \left(\frac{1}{2}-\frac{1}{m}\right)T^{*2}\frac{\partial ln(g_{ee}^{D}(T^{*},\rho^{*},K^{*},\sigma))}{\partial T^{*}} \quad (4.171)$$

The pressure P^{*}, is expressed by:

$$P^{*} = P_{ideal}^{*} + \rho^{*}\left(\frac{\partial\left(A_{res}^{*,chain}\right)}{\partial\rho^{*}}\right)_{T^{*},Nc} \quad (4.172)$$

where the full expression is:

$$P^{*} = P_{M}^{*} + \rho^{*}\left(\frac{1-m}{m}\right) + \frac{T^{*}\rho^{*}}{2}\left(\left(\frac{2-m}{m}\right)\frac{\partial ln\left(g_{ee}^{D}(T^{*},\rho^{*},K^{*},\sigma)\right)}{\partial ln(\rho^{*})} - \frac{\partial ln\left(g^{M}(T^{*},\rho^{*},\sigma)\right)}{\partial ln(\rho^{*})}\right) \quad (4.173)$$

where P_{M}^{*} is the pressure of the monomer.

4.8 Model validation

In the first step of this work, we have validated the developed equation of state on the properties of the fully flexible fluid model. The idea is now to extend the validation to the semi-rigid case using the full model developed in this work. To do so, extensive Monte Carlo simulations were performed to obtain the LVE properties of different chain lengths of LJ chains with variable rigidities. Results, shown in Figure 4.12 to Figure 4.17, have also been obtained in single phase domains including the dense liquid phase and the supercritical fluid from low density to high density. Note that, in what follows, new notation which will be used, where we refer to the generalized TPT1-D developed in this work as SR-SAFTD (SR stands for Semi-Rigid), and the TPT1-M as soft-SAFT.

4.8.1 Single phase

First, the low-density limit internal energy was investigated, as done with the fully flexible chains. From Figure 4.12, the results show that SR-SAFTD increasingly predicts the internal energy at zero density for different chain lengths with increasing stiffness. This interesting result highlights that the shielding effect, in which the interaction between two monomers of the same chain separated by at least one monomer, is cancelled by the presence of these intermediate monomers. This is all the more correct as the monomers are aligned and thus, as the rigidity is increased (SI3). In the case of a fully flexible chains, the SR-SAFTD model accounts perfectly for the interactions involving four successive monomers in the same chain, reason for which the internal energy for chain of 4-mers are accurately predicted. However, as the chain length increases, additional interactions are involved which may become dominant at a certain length due to the folding of the molecule on itself. The lack of these additional correlations in the SR-SAFTD model explains the deviations obtained in the Figure 4.12.

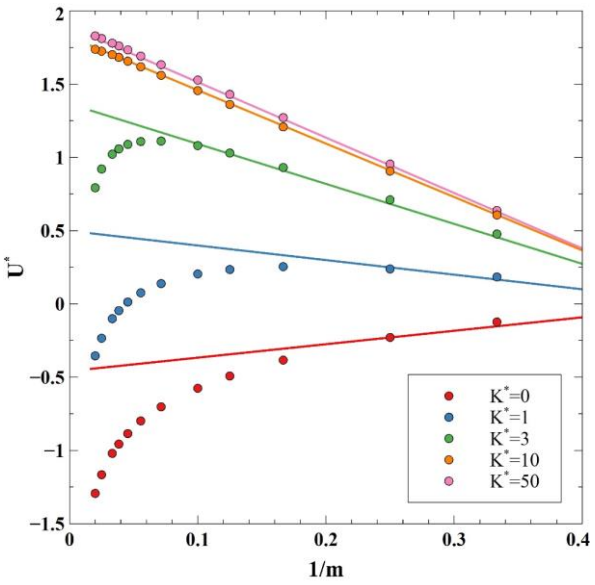


Figure 4.12: Zero-density limit internal energy at T=2 predicted by the SR-SAFTD (solid lines) for variable rigidities (K*=0, 1, 3, 10, 50) and chains up to m=50 vs the exact monte carlo simulations (dots).*

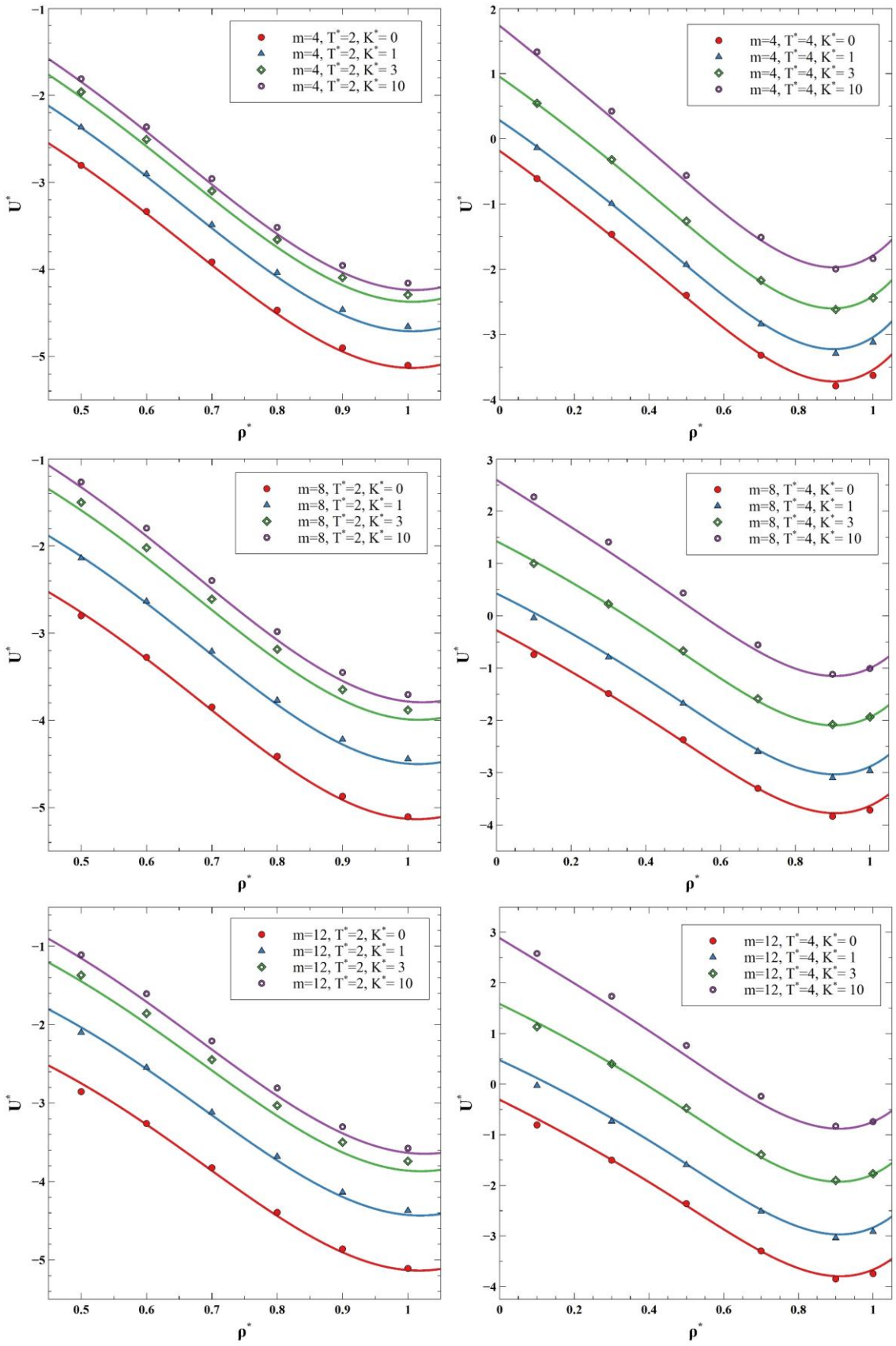


Figure 4.13: Residual internal energy for chains of m=4, 8, 12 at supercritical(T=4) and subcritical(T*=2) conditions for variable rigidities (K*=0, 1, 3, 10) obtained from molecular simulations (symbols) and predictions using the SR-SAFTD (solid lines).*

The results plotted in Figure 4.13 show that the SR-SAFTD gives excellent predictions of the internal energy in the subcritical conditions as well as in the supercritical conditions for the different chain lengths (m = 4, 8 and 12) investigated in this work. Still, slight deviations can be noticed when increasing the rigidity in the subcritical region, but the supercritical region is accurately predicted in all cases. The model is tested only at a maximum rigidity corresponding to K^*=10. However, we expect that the model will still providing reasonable results valid for rigidities around K^*=20, which is sufficient for the real fluid applications as it will be shown later.

Regarding the pressure, the rigidity has only a small effect as may be seen from the simulation data in Figure 4.14. However, the small deviation observed for the chain of 8-mers are correctly captured by the SR-SAFTD for both liquid and supercritical phases. The increase of pressure with increasing the rigidity seems to be consistent, as this indicates that it requires less energy to stack fully flexible chains which do not exhibit any resistance to fold, whereas the rigid chains require more energy due the high internal resistance.

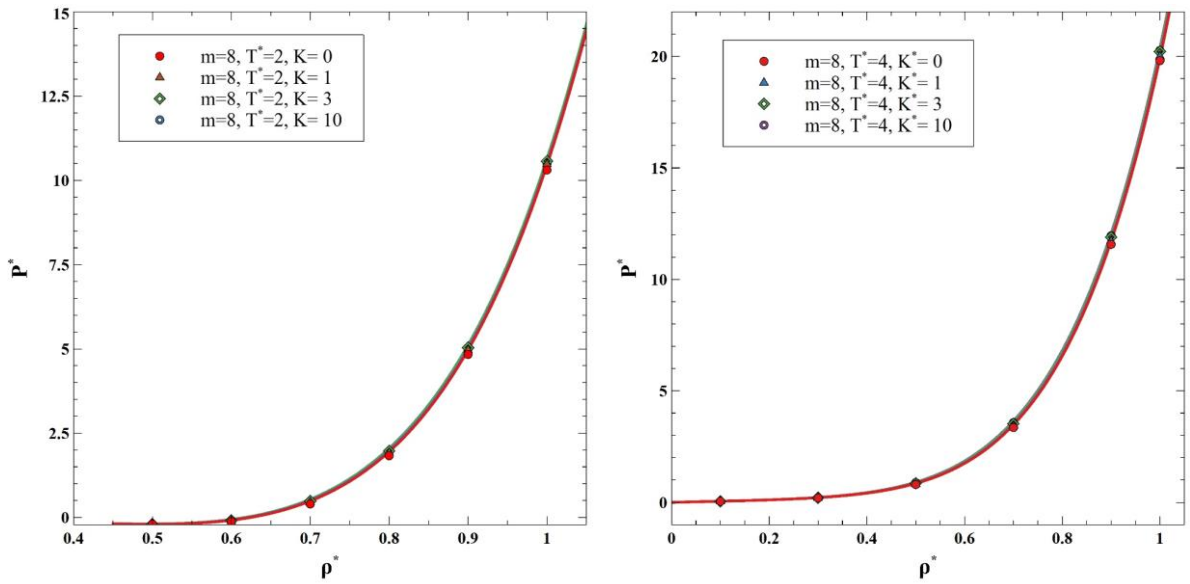


Figure 4.14: Pressure for chain of m=8 at supercritical(T=4) and subcritical(T*=2) conditions for variable rigidities (K*=0, 1, 3, 10) obtained from molecular simulations (symbols) and predictions using the SR-SAFTD (solid lines).*

4.8.2 Liquid vapour equilibrium (LVE)

In terms of liquid vapour equilibrium properties and saturation pressure, results obtained with the SR-SAFTD are all consistent with the computer simulations of the LJ chains with different lengths and variable rigidities as shown in Figure 4.15. While the predicted vapour phase density is always decreasing as the rigidity increases (not close to the critical point), which is consistent with the simulation data, the SR-SAFTD describes correctly the non-linear behaviour of the saturated liquid density, which tends to be lowered at small rigidities and then increased as the rigidity increases. Capturing this non-linearity is a very good indication that our model is physically consistent. Regarding the saturation pressure, the SR-SAFTD yields very good predictions for all rigidities and chain lengths. The behaviour of the critical point with rigidity

seems also very well predicted, as can be seen from the LVE curves. More precisely, the equation of state is predicting a lower value of the critical temperature for K^*=3 compared to the fully flexible chains, and a higher value for K^*=10 as found in the simulations.

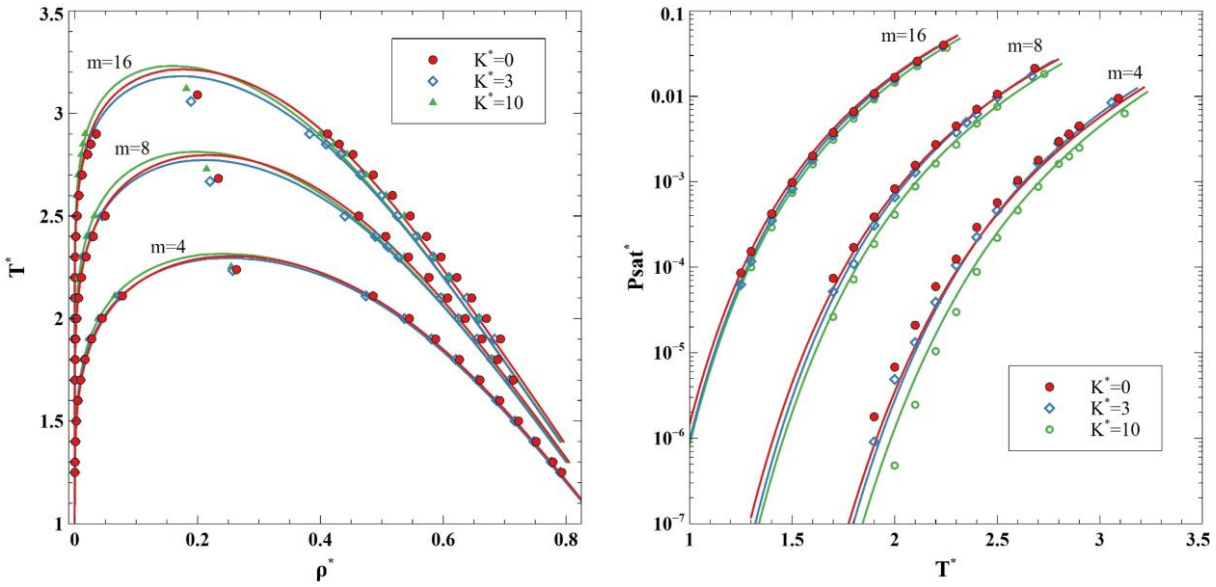


Figure 4.15: Liquid-vapour equilibrium properties and saturation pressures predicted by the SR-SAFTD (solid lines) for chains of m=4, 8, 16 vs molecular simulations (symbols) for variable rigidities (K=0, 3, 10).*

- **Critical properties**

To better show how correct is the prediction of the critical point location using the SR-SAFTD, all critical properties of the 5-mers chain have been plotted in Figure 4.16. The acentric factor is shown as well as it is very sensitive to the critical pressure and temperature. As can be seen from Figure 4.16, the SR-SAFTD captures qualitatively the behaviour of all critical properties and also the acentric factor. However, quantitatively, the critical temperature and the critical pressure are both overestimated by the model, while the critical density and the acentric factor are underestimated. When applied to real fluids, it is then expected that the SR-SAFTD will fail in reproducing accurately these critical properties. However, even if not perfect, these results are largely better than those that would be deduced from a soft-SAFT approach as shown in Figure 4.16. Furthermore, this is a classical weakness of any mean field model that needs a crossover contribution to be fully corrected.

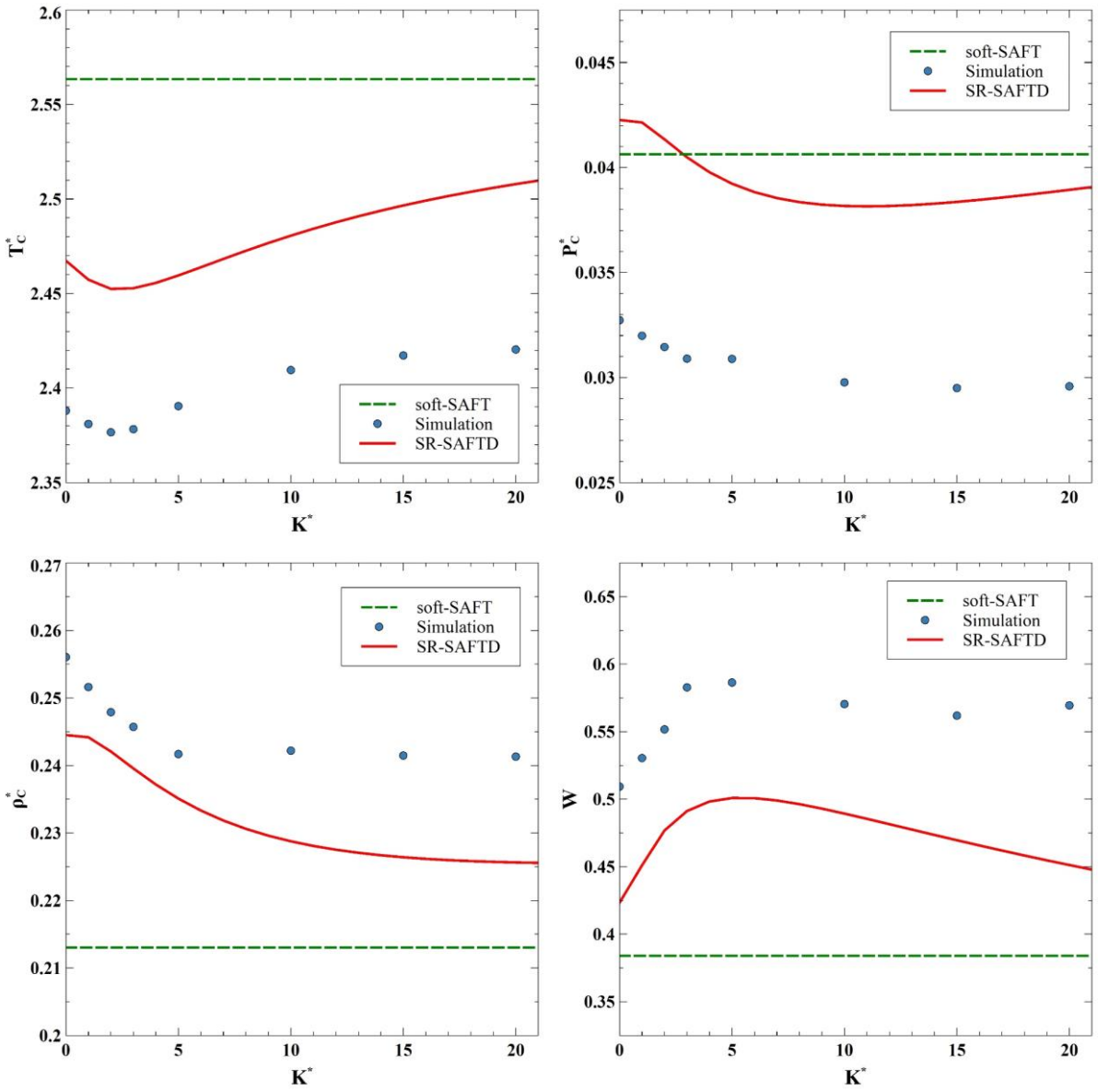


Figure 4.16: Critical properties and acentric factor predicted by the SR-SAFTD (solid lines) and soft-SAFT (dashed lines) for chain of m=5 vs molecular simulations (symbols) for variable rigidities (K=0 to 20).*

4.8.3 Internal energy & Entropy at saturation

In addition to the monophasic fluid and the LVE properties, excellent predictions are also obtained for the internal energy and the entropy at saturation conditions, as shown in Figure 4.17. Interestingly, the predictions become closer to the computer simulations results as the rigidity increases, consistently with what was observed for the zero-density limit of the internal energy. This is particularly obvious for the internal energy of 8-mers chain, due to the shielding effect. As already mentioned, the SR-SAFTD should be valid for chain length of any odd number, even if build to deal with an even number of mer. This assumption is confirmed by the excellent prediction for the entropy obtained for the trimer chain. In addition, the excellent entropy prediction with the SR-SAFTD is crucial, especially in our work, because this quantity

will be used to propose an entropy scaling approach to predict viscosity of the semi-rigid chains. This approach will be discussed in the next chapter of this thesis.

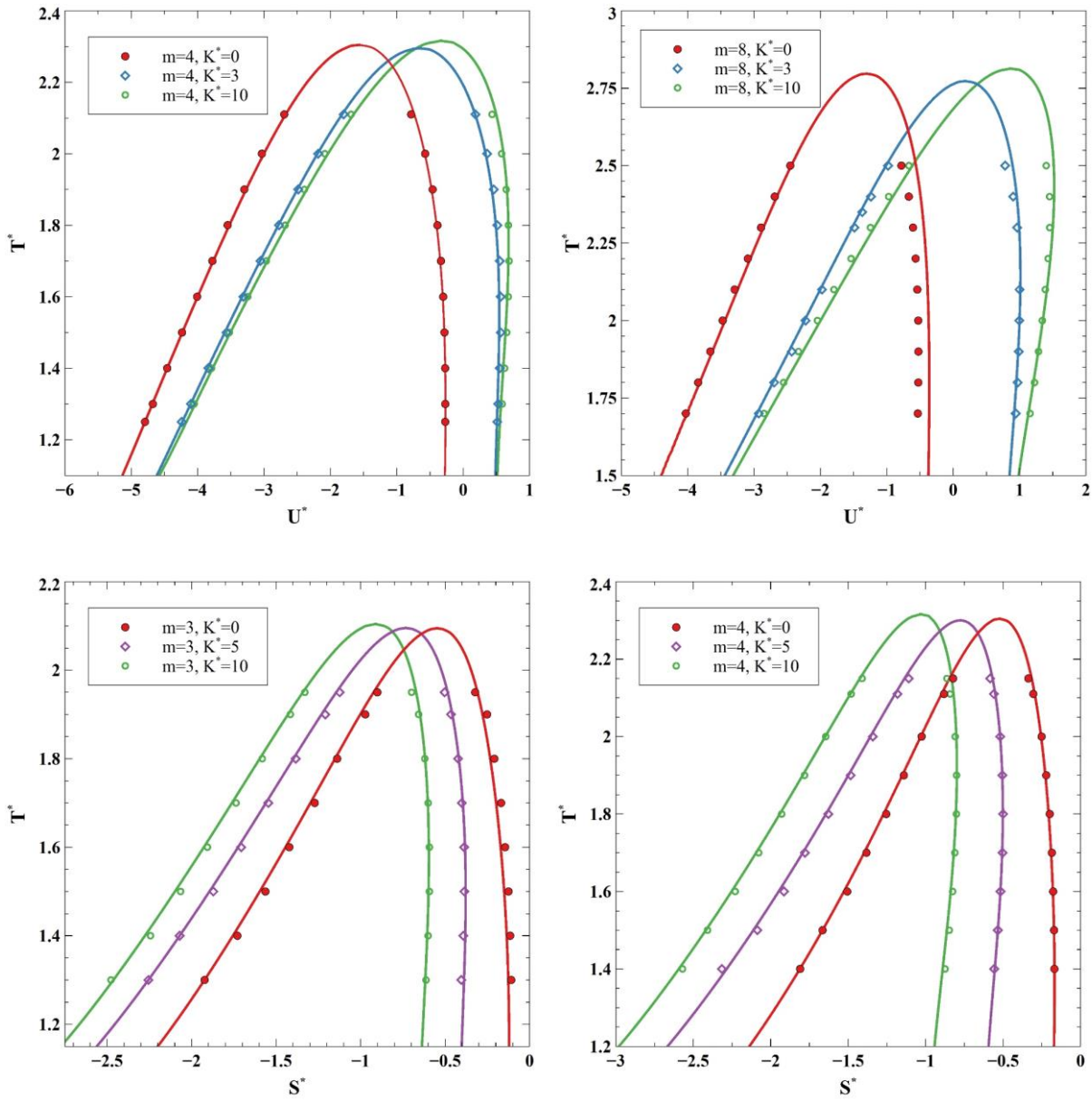


Figure 4.17: Liquid-vapour internal energy and entropy predicted by the SR-SAFTD (solid lines) for chains of m=4, 8 vs molecular simulations (symbols) for variable rigidities (K=0, 3, 5, 10).*

4.9 Real fluid properties

After having validated the SR-SAFTD on "exact" properties of the semi-rigid chains obtained from computer simulations, this equation of state will be applied to real fluids in this section. The application is limited to pure components only and mixtures will be considered in the future.

4.9.1 Optimization of the molecular parameters

Before showing any results, an important question is about the way the molecular parameters appearing in the SR-SAFTD are defined. In this work, relying on the results presented in the previous chapter, is proposed a parametrization based on the corresponding states similarly to what done for the semi-rigid LJCCG model. The corresponding state strategy has already been applied for the SAFT-γ-Mie [94], but including a transport property in the parametrization of a SAFT models has not yet been tested.

Thus, the parameterization strategy of the SR-SAFTD is similar to that for the LJCCG force field. However, instead of using an optimization based on the equation of state itself, and thus, using the theoretical prediction of the critical temperature, liquid density and acentric factor given by the SR-SAFTD EoS, it is the exact properties of the semi-rigid LJCCG model (the correlations for the Tc, ρ_{Tr07} and w, given in the appendix 7.A.2) that will be used. A priori, one may expect that a parametrization based on the corresponding state strategy will not give accurate predictions because of the wrong scaling exponent involved in the universal scaling laws when the thermodynamic properties are asymptotically approaching the critical point. This is a classical problem for any equation of state as it does not take into account the energy contribution due the density and composition fluctuations appearing in the vicinity of the critical point.

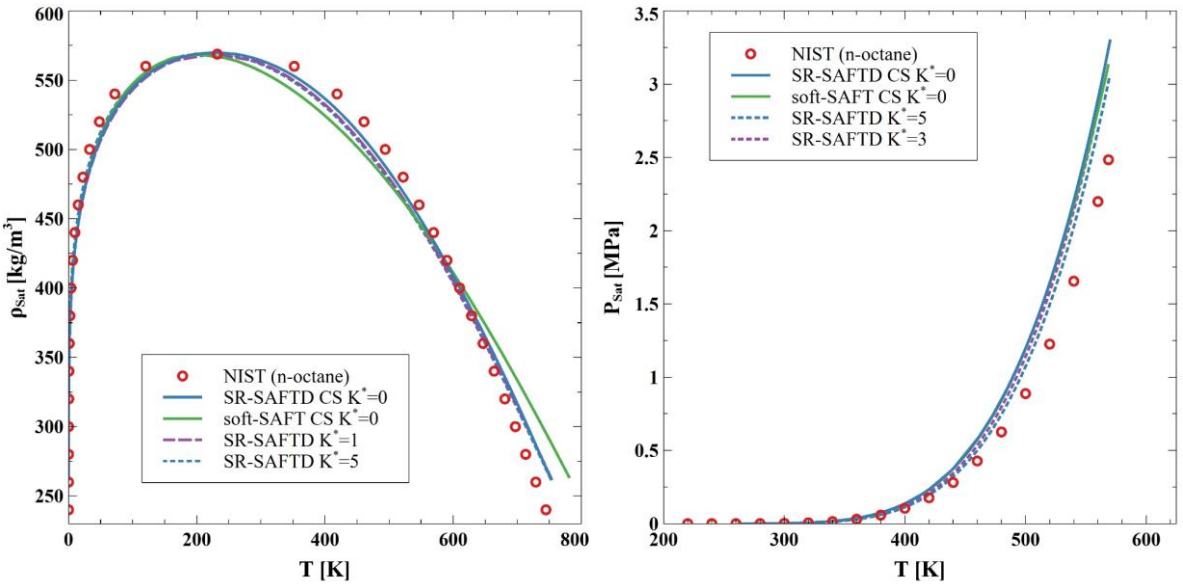


Figure 4.18: Prediction of the phase equilibria properties of pure n-octane using a corresponding state strategy (CS).

In order to illustrate the weakness of an optimization using a corresponding state based on the properties of the EoS, we took the case of the n-octane molecule using the SR-SAFTD and the soft-SAFT. In a first step, we did not consider the rigidity parameter for the SR-SAFTD, thus, the two EoS models are parametrized with exactly the same strategy. The number of segments is obtained via the acentric factor and the interaction and the diameter parameters are obtained from the critical temperature and saturated liquid density at Tr=0.7, similarly to what was done in chapter 3 for the SR-LJCCG model.

Results in Figure 4.18 show that neither the SR-SAFTD (with K^*=0) neither the soft-SAFT are able to provide simultaneously accurate coexistence densities and saturation pressure properties. However, because the SR-SAFTD performs better relatively to the soft-SAFT in the critical region, slightly better results are obtained with SR-SAFTD close to the critical point and for the liquid density. Quite similar strategy was applied on the soft-SAFT by Pàmies and Vega. [95] who tested the corresponding state strategy using only the critical properties similarly to what is done for the cubic EoS. Their results showed it is not possible to obtain accurate prediction of both LVE properties and saturation pressure. More recently, to overcome this limitation, Moine et al. [96] showed that a fourth molecular parameter is required. The authors proposed to use a volume translation correction parametrized on the saturated liquid density at T_r=0.8, and obtained very good results with the I-PC-SAFT. Another alternative consists in using the cross over treatment of the EoS to yield better scaling of the critical region such as its application on the soft-SAFT [97, 53, 98], SAFT-VR [99] or the PC-SAFT [100]. However, this latter option is somewhat complicated as it requires the rescaling of all the variables and additional parameters are required.

Since the SR-SAFTD contains four molecular parameters, one may then wonder whether the additional rigidity parameter can be used to remedy this issue instead of using the volume translation. In the Figure 4.18, we tested at which extent this rigidity parameter may play a role in improving the predictions. Two values of rigidity (K^*=1 and K^*=5) have been tested, and results showed that, only small improvement are obtained in the T-P space, but, in the detriment of a slightly deterioration of the liquid density close to the critical region. The phase diagram seems just to be slightly shifted to the low densities. In short, the rigidity parameter, does not give enough flexibility to the model in terms of equilibrium properties. Moreover, even if this strategy based on the CS using the EoS worked, the obtained parameters would not ensure obtaining accurate viscosity, and additional re-adjustment would certainly be needed. Finally, the set of parameters obtained would be largely different from those obtained for the SR-LJCCG model presented in chapter 3.

In the SR-SAFTD validation step, it was shown that this EoS gives accurate predictions of the set of properties tested when compared to the computer simulations of the fluid model. Therefore, if the parameters used for semi-rigid LJCCG model are injected into the SR-SAFTD EoS this should lead to very reasonable results on real fluids similarly to what we obtained in chapter 3, except very close to the critical point where the EoS would be not accurate.

As it will be confirmed in the following section, the parameters obtained for the coarse-grained force field model can be used for the SR-SAFTD model. However, when using the SR-SAFTD, there is no need to constrain the number of segments forming the chain to be an integer number,

as required when performing molecular simulations. This is because a SAFT model allows modelling fluids with a non-integer (even if sometimes their meaning can be questionable) contrary to molecular simulations. More in details, during the optimization procedure of the coarse grained model in the previous chapter, the parameters were obtained in two steps, where in the second step, the number of segments was forced to be an integer number. The optimization using a SAFT model does not require this second step, and it is the optimal parameters minimizing the F-function to the global minimum that will be used. The parameterization process is shown in Figure 4.19 (b). The global procedure for the optimization is similar to that of the coarse grained model as shown in Figure 4.19 (a), but the main difference appears in the last step to predict the real fluids properties. Instead of using molecular simulations, the thermodynamic properties will be predicted by the SR-SAFTD, and the viscosity will be predicted thanks to the entropy scaling approach as it will be discussed in the following chapter.

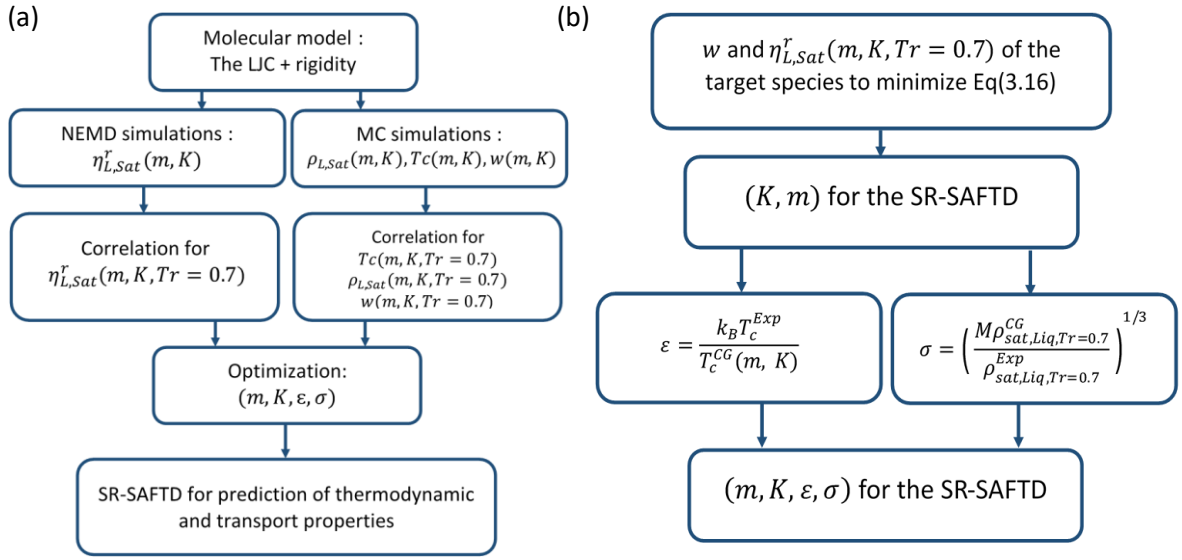


Figure 4.19: (a) flowchart of the general methodology using the SR-SAFTD, (b) flowchart of the optimisation strategy using SR-SAFTD.

It is worthwhile to mention it again, the optimization routine may be slightly dependent on the initial guess due to the presence of some local minima, but, can easily be overcome (see recommendation in the previous chapter) to converge to the global minima. This problem was mainly noted for short molecules with acentric factor around that of m~3. This is due to the non-linearity of the acentric factor with rigidity which is not compensated by the effect of the viscosity. In fact, the viscosity does not show any non-linearity with the rigidity and so tends to smooth the objective function when its contribution is found to be the dominating one, which is mainly the case for $m > 3$. The issue of the presence of multiple minima is a common problem for any equation of state when the parametrization is based on simultaneous optimization of a set of parameters using the classical fit over the saturation curves, or when derivative properties are included in the fit as target properties. Thus, in such strategies, high caution is required for the user to avoid obtaining unphysical parameters, particularly when four parameters equations of states are used [101, 102]. The SR-SAFTD possesses also four molecular parameters, but is not

subject to this problem because of the corresponding states framework which allows to decouple the fitting of the parameters. Indeed, only two parameters are fitted simultaneously (*m* and *K**) to only two experimental data and the two other ones (ε, σ) are simply deduced from the corresponding state relations.

4.9.2 Results and discussions

4.9.2.1 LVE and saturation pressure

The results obtained in the previous chapter using the optimized coarse grained force field by performing direct molecular simulations are reproduced in this part and compared to the prediction using the SR-SAFTD model. The predictions with the two approaches are then compared to the experimental measurements, and all results for LVE and saturation pressure are presented in Figure 4.20 for normal alkanes, iso-alkanes and some polar components.

The comparison between the simulations data and the SR-SAFTD predictions are in excellent agreement for the two properties and for all the species, except very close to the critical point which is not surprising, as already mentioned. The imposed saturated liquid density at T_r=0.7 is also accurately reproduced and the parameters used show very good transferability to other states not included in the fitting. Another interesting point to emphasize is the good prediction of the properties of the polar components without any use of polar or association terms. Indeed, it is very common in the SAFT community to add these contributions as soon as the molecules exhibits a dipole or a quadrupole moment, leading to more complexity in terms of parameters estimation. Here we show that this is perhaps not justified in some cases. Similar conclusion has been drawn in the recent study conducted by Ramirez-Vélez et al. [103] showing that there is no clear correlation between the association strength and the accuracy of the EoS. Thus, many polar molecules could probably be correctly modelled without any polar or association parameters using the corresponding state strategy proposed in this work. With this in mind, it is recommended that, for any compounds, one should first check whether the species could be modelled with SR-SAFTD before adding association or polar terms.

There is nevertheless exception for two cases, where the proposed strategy should fail. The first is that when it comes to mixtures, such contributions are most of the time asymmetric between the species. Thus, this requires additional fitting parameter to capture well the cross-interactions. The second is the situation corresponding to strongly associating molecules, such as alcohols and water. In fact, these molecules are characterized by a small molecular weight with high acentric factor and viscosity as well. Thus, applying this strategy for such molecules leads to a high number of segments which cannot be consistent (unphysical situation). For example, when trying to model methanol, the proposed approach gives a segment number around m=4.3, while it molecular weight is not far from that of ethane which is modelled by m=1.73. An illustration and a discussion of this particular case is given in the appendix 7.B.1.

As already pointed out and shown in Figure 4.20, the case of the critical properties are not accurately predicted for all molecules, particularly for large molecules. In the case of n- alkanes for example, the deviations noted are reasonable for the critical temperature (around 3%) and for the critical densities (around 7%) and seems to be relatively stable with increasing the chain length. However, the critical pressures are highly overestimated and are observed to increase with increasing the length of the chain (around 68% for the nC_{36}!) as shown in Figure 4.21. Note that, even though these deviations are high, we expect that they would be lower than what would be obtained with the soft-SAFT, as emphasized by the results shown in Figure 4.16

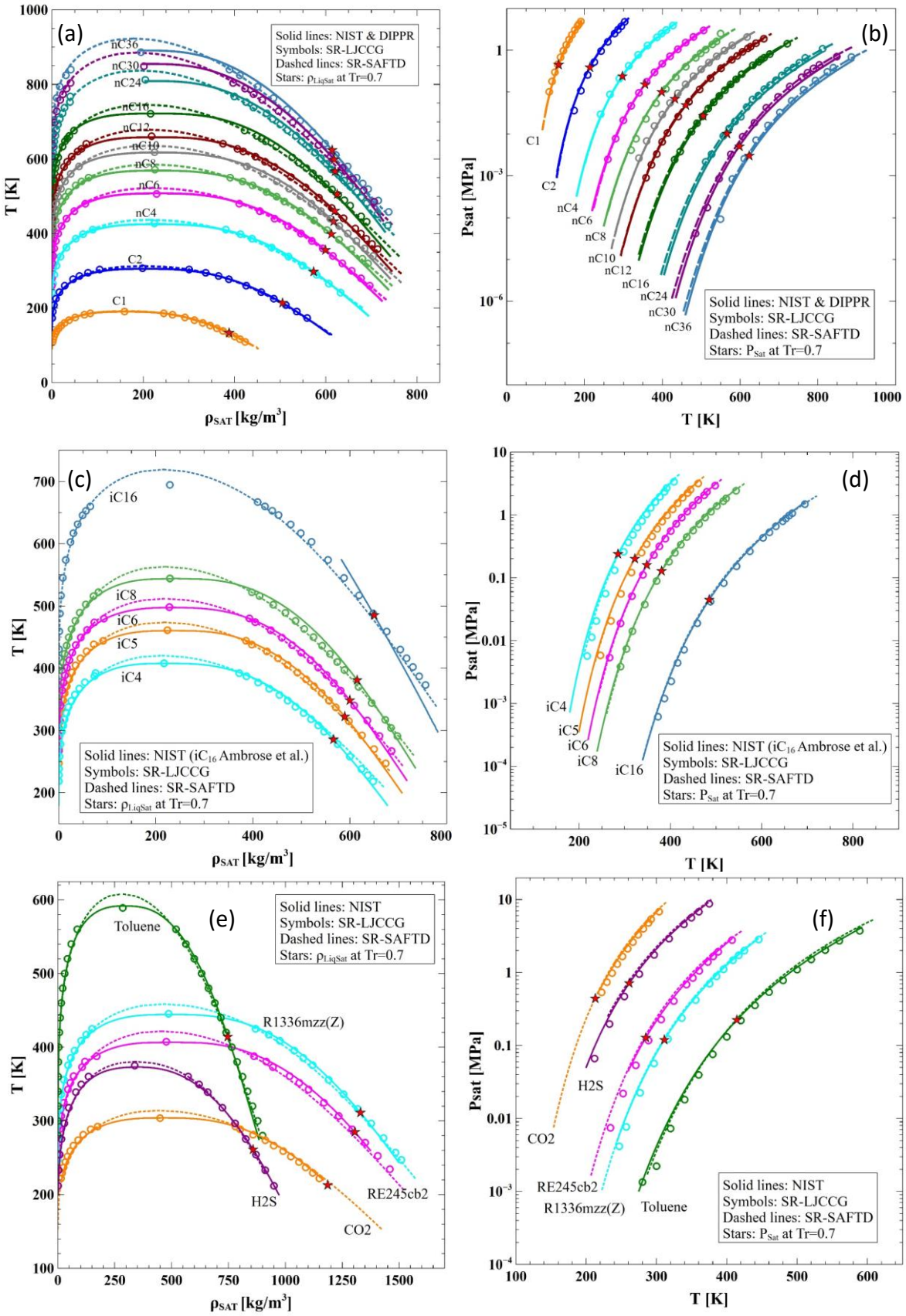


Figure 4.20 : Liquid-vapour saturated densities and saturation pressures predicted with the SR-SAFTD and the SR-LJCCG vs the correlated experimental data from NIST database. (a,b) n-alkanes, (c,d) iso-alkanes, (e,f) polar compounds.

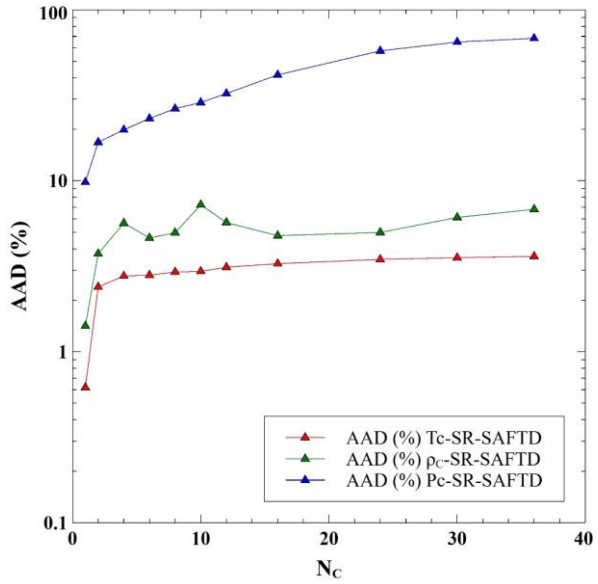


Figure 4.21: Relative deviations of the critical properties predicted with SR-SAFTD vs the carbon number of series of n-alkanes.

A comparison between the classical parametrization strategy based on the fitting of the saturated liquid density and the saturation pressure and that using a corresponding state strategy proposed in this work is shown in Figure 4.22. Two models are used for the classical fit strategy, the fully flexible SAFTD with the parameters obtained from the work of Zhang [104], and the fully flexible soft-SAFT with the parameter obtained from the work of Pàmies [52]. Two compounds are studied, the normal hexane and the normal dodecane.

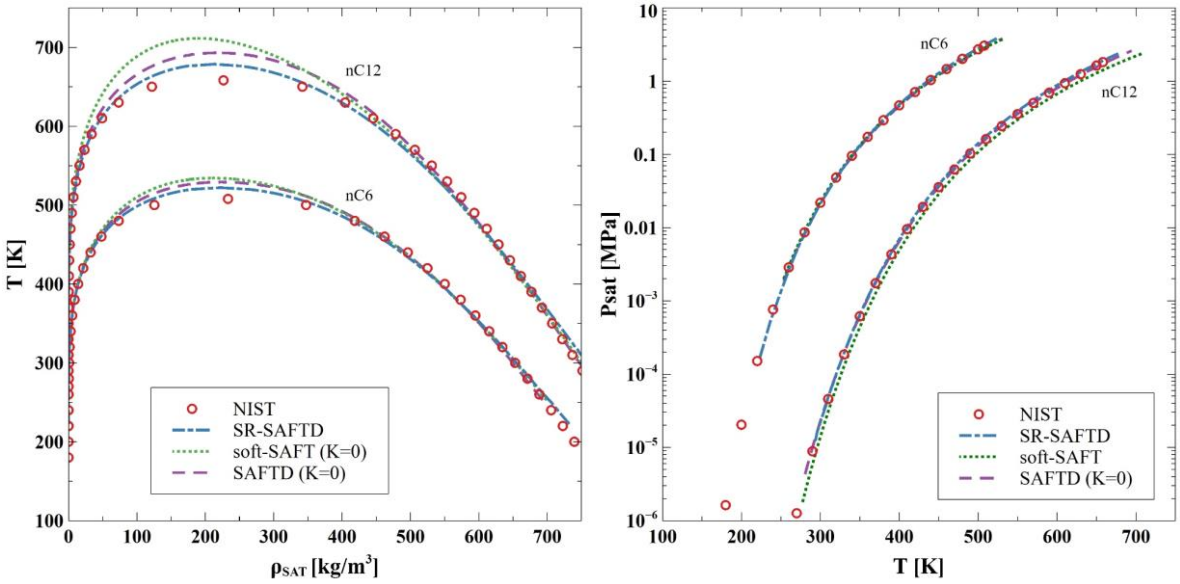


Figure 4.22: Liquid-vapour saturated densities and saturation pressures predicted with the SR-SAFTD, the fully flexible SAFTD and the soft-SAFT vs the correlated experimental data from NIST database of n-C6 and n-C12.

The fully flexible SAFTD and the SR-SAFTD exhibit practically the same accuracy for both LVE properties and saturation pressure, with a small advantage for the SAFTD at low temperatures (T_r<0.6) on the liquid side of the LVE, and a small advantage for the SR-SAFTD

when approaching the critical point. The soft-SAFT leads to greater deviations compared to the two previous cases for the LVE (on both sides: liquid and vapour) and the saturation pressure as well, with maximum deviations noted at the critical point. These results highlight that the use of a SAFTD improves considerably the predictions of the real fluids properties when compared to the soft-SAFT, which is not surprising if one refers to the improvement obtained on the model fluid.

All these results confirm once again the robustness of the corresponding state strategy as it is applied in this work. It provides predictions with the same accuracy as models (soft-SAFT and SAFTD) fitted only on equilibrium properties (which requires many data) and over large range of temperatures. Furthermore, the SR-SAFTD is expected to provide prediction of the viscosity with a similar accuracy as for the equilibrium properties when the optimized parameters are transferred to molecular simulations of the LJCCG model or when the entropy scaling approach is used, whereas those of soft-SAFT or SAFTD will largely fail to do so whatever the approach used.

4.9.2.2 Enthalpy at saturation

To obtain good prediction on the thermal properties of fluids, the enthalpy or its derivative properties are often included in the parameters estimation for the development of force fields such as the OPLS-UA [105] and the AUA(4) [106] and in the parametrization of equations of states [70, 71, 107, 108]. Such property is extremely important in many applications, and thus, it is important that the model is able to provide good predictions on it. Moreover, as in our case the enthalpy is not included in the optimization as a target property, this is a good test for the representability of the parameters obtained.

As shown in Figure 4.23, results obtained for this property are fully consistent with the experimental data for all studied fluids. The exception is at the critical point where the SR-SAFTD loses accuracy, similar to what was observed for the equilibrium properties. However, the overall fair behaviour obtained with the SR-SAFTD shows that it has a solid physical basis to capture not only the phase equilibria but also some thermal properties.

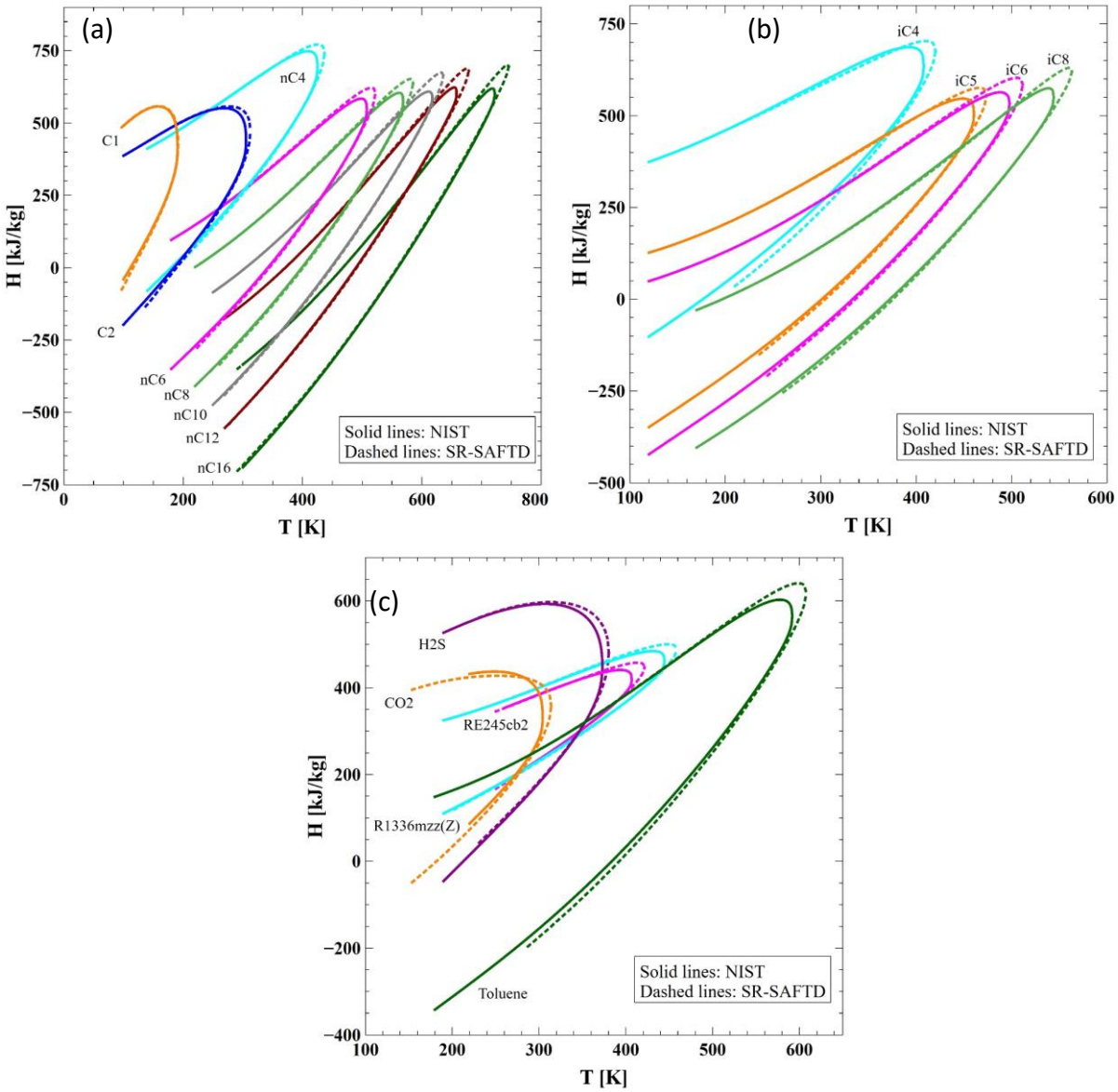


Figure 4.23: Liquid-vapour enthalpy predicted with the SR-SAFTD vs the correlated experimental data from NIST. (a) n-alkanes, (b) iso-alkanes, (c) polar components.

4.9.2.3 Second-order derivative properties

The good prediction of the first-order derivative properties does not always ensure a good prediction of the second-order derivatives properties as these latter are extremely sensitive to any variation in the former. In particular, these second-order derivative properties exhibit many anomalous behaviours particularly at the critical point, but, also relatively far from that peculiar point. This makes their prediction very challenging for any equation of state. Many efforts have been devoted to improve the prediction capability of many SAFT models (Soft-SAFT, SAFT-VR-Mie, PC-SAFT) [70, 71, 107, 108] on these properties. To do so, these properties were set as target properties in the objective function, with sometimes taking a weighting coefficient different from unity. However, the inclusion of derivative properties in the objective function for non-associating molecules was found to have only a marginal impact [107], while the

transferability of the parameter to other molecules leads to some complexities [101]. De Villiers [109] studied the predictive capability of the original-SAFT, PC-SAFT and CPA on the derivative properties without including them in the fitting of the parameters. It was concluded that these EoS are not capable to predict simultaneously accurate first-order properties and second-order ones, but advantages were found in many cases for the PC-SAFT. When tested in the near critical region, all these EoS fail to give even correct trend for the constant volume heat capacity (C_v), while it was shown that a correct behaviour could be obtained with the crossover treatment of the soft-SAFT [53]. Thus, as a strong test, the SR-SAFTD developed in this work has been assessed on its predictive capability on the derivative properties which were not included in the fit.

Results are shown in Figure 4.24 to Figure 4.26 for three families of molecules. The SR-SAFTD shows excellent prediction when compared to direct molecular simulations which is due to the consistent parametrization between the two approaches (the coarse grained force field and the SR-SAFTD). The comparison against experimental data is made over an extremely large range of temperatures at pressure corresponding to 100 bars. Many properties were assessed, including the isobaric heat capacity, the isochoric heat capacity, the thermal expansivity, the speed of sounds and the Joule-Thomson coefficient. As can be seen from these figures, the SR-SAFTD is showing excellent agreement with experimental data for all properties studied and for all the molecules considered in this work. The speed of sounds for example, which is known to be extremely challenging for many equations of states [109], is fairly well predicted by the SR-SAFTD. Still, small deviations are observed at low temperature when the chain length increases for the normal alkanes and for the iso-octane as well.

The SR-SAFTD is also capable to reproduce quantitatively the non-linearity exhibited by the derivatives properties for thermodynamic conditions not far from the critical point. This is a signature of the strong physical basis of the SR-SAFTD and confirms that the use of the LJ potential to account for the intermolecular interactions is a robust choice. Interestingly, carbon dioxide (CO_2) and hydrogen sulphide (H_2S) are tested under conditions extremely close to their critical points, which is a severe test for any equation of state. The results obtained with SR-SAFTD are in perfect agreement with the experimental data, reproducing all singularities of the derivative properties, except for small deviations observed for the isochoric heat capacity (C_v). The excellent trends predicted for the isochoric heat capacity confirms the overall superiority of the LJ-based SAFT, and thus the SR-SAFTD, when compared to some SAFT variants (CK-SAFT, PC-SAFT, CPA) as shown in [109]. To enhance quantitatively the prediction of the C_v would probably require further treatment by adding a kernel term that acts in favour of this property [110].

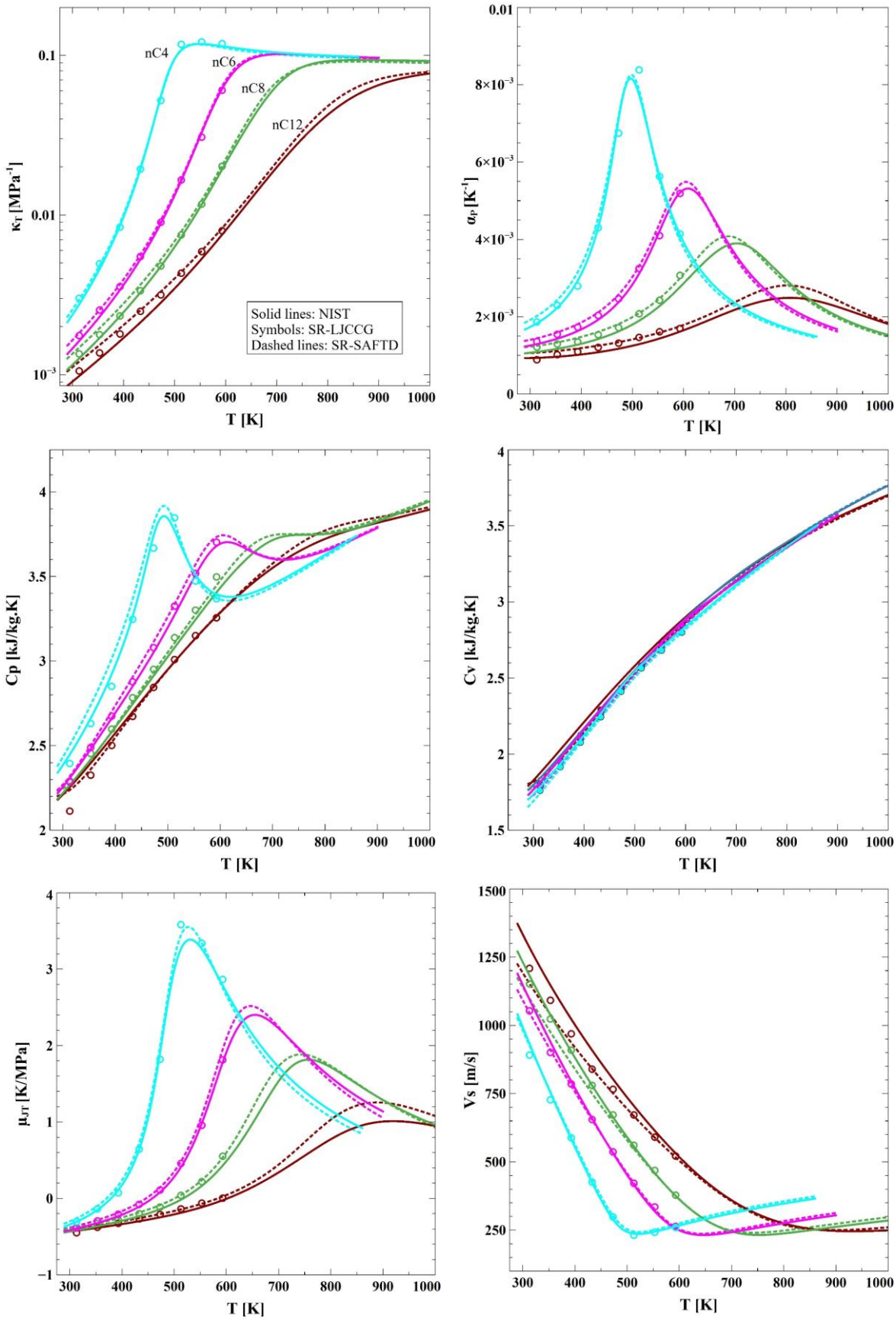


Figure 4.24: Temperature dependence of the second order derivative properties of the n-alkanes components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.

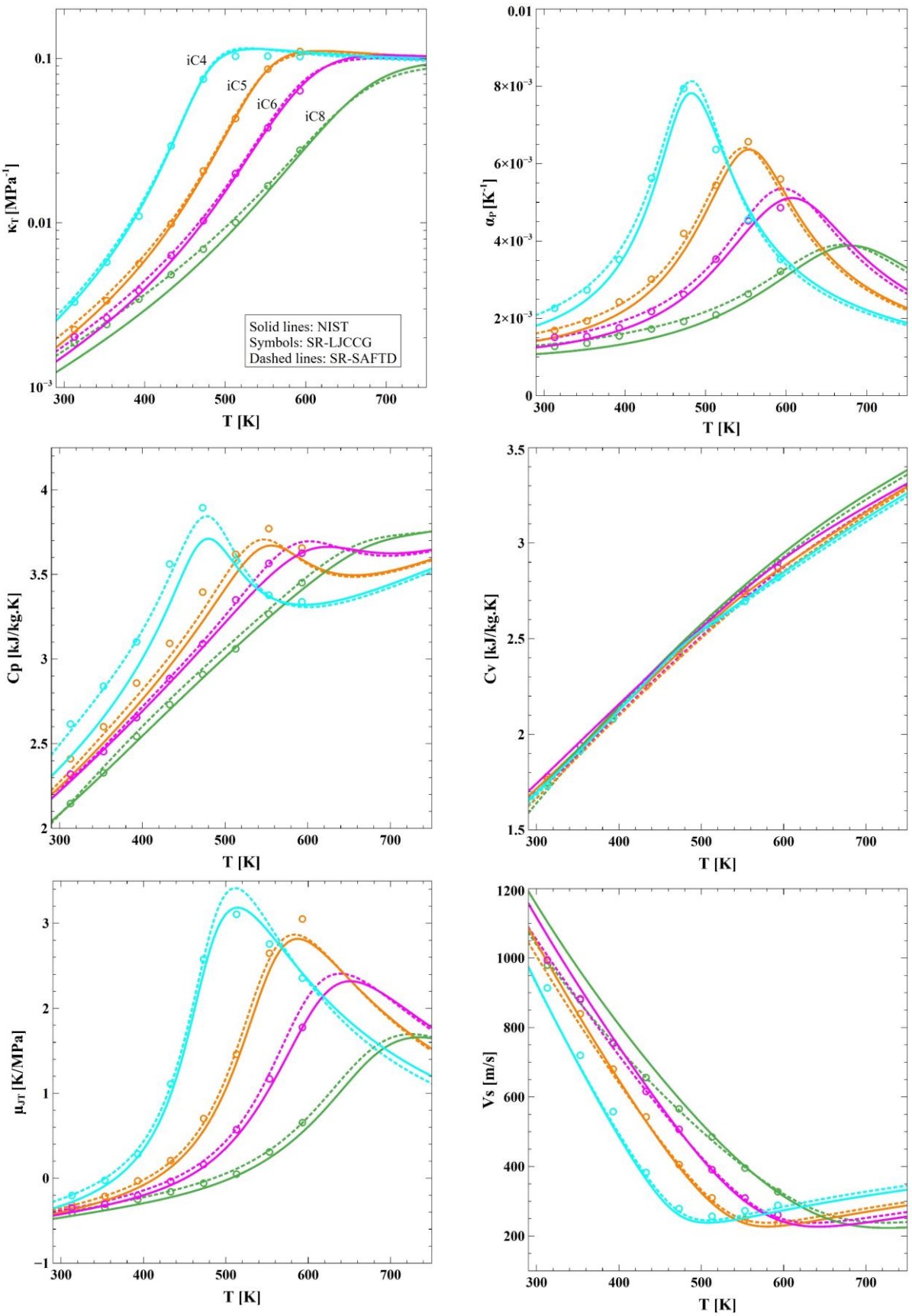


Figure 4.25: Temperature dependence of the second order derivative properties of the iso-alkanes components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.

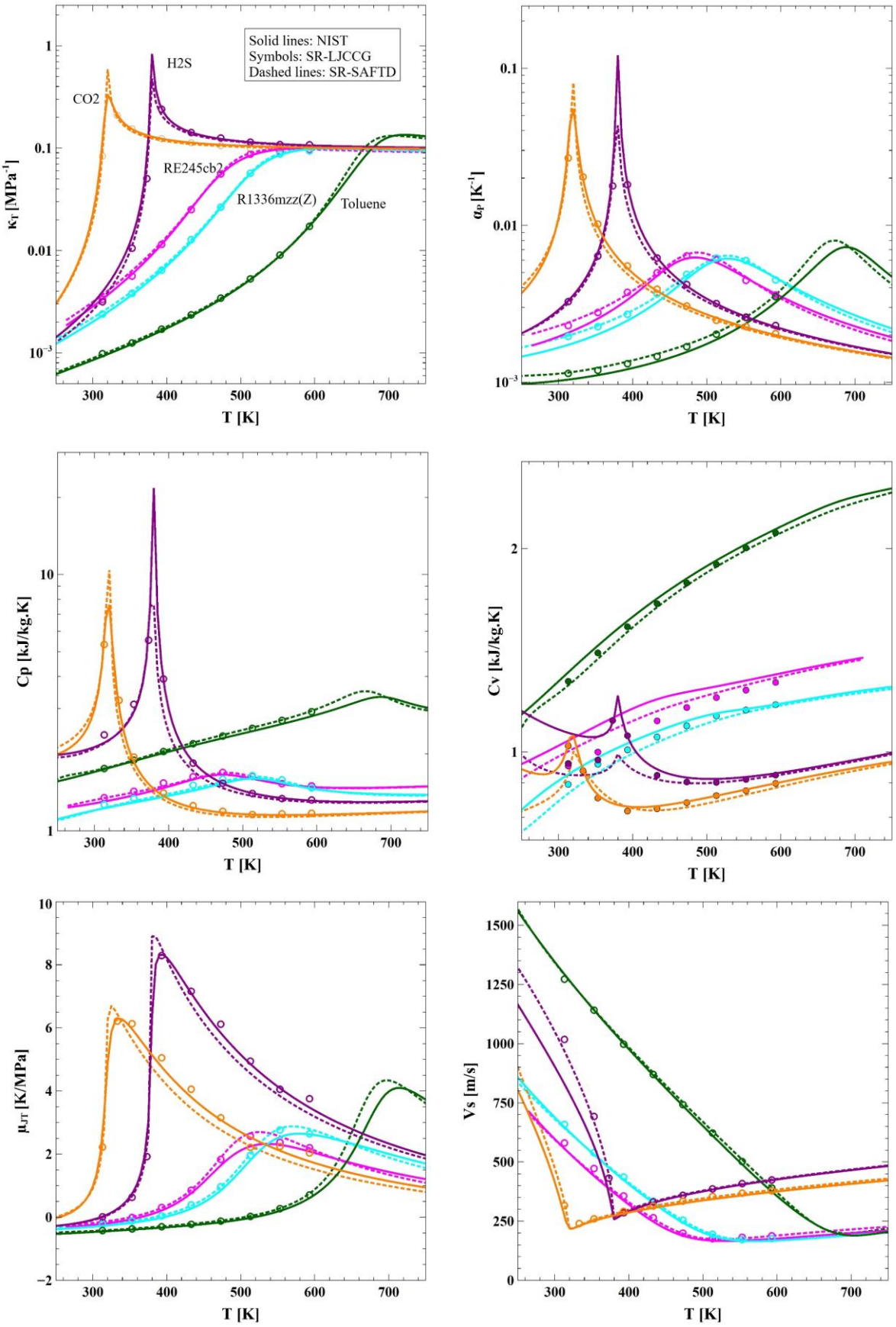


Figure 4.26: Temperature dependence of the second order derivative properties of the polar components predicted with the SR-LJCCG and SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.

4.9.2.4 Signe phase enthalpy

For the same conditions for which the second-order derivative properties were studied, the prediction for the single phase enthalpy was investigated. As shown in Figure 4.27, excellent predictions are obtained for the enthalpy which is consistent with the accurate isobaric heat capacity obtained. In most cases the experimental data and the predictions from the SR-SAFTD are not distinguishable.

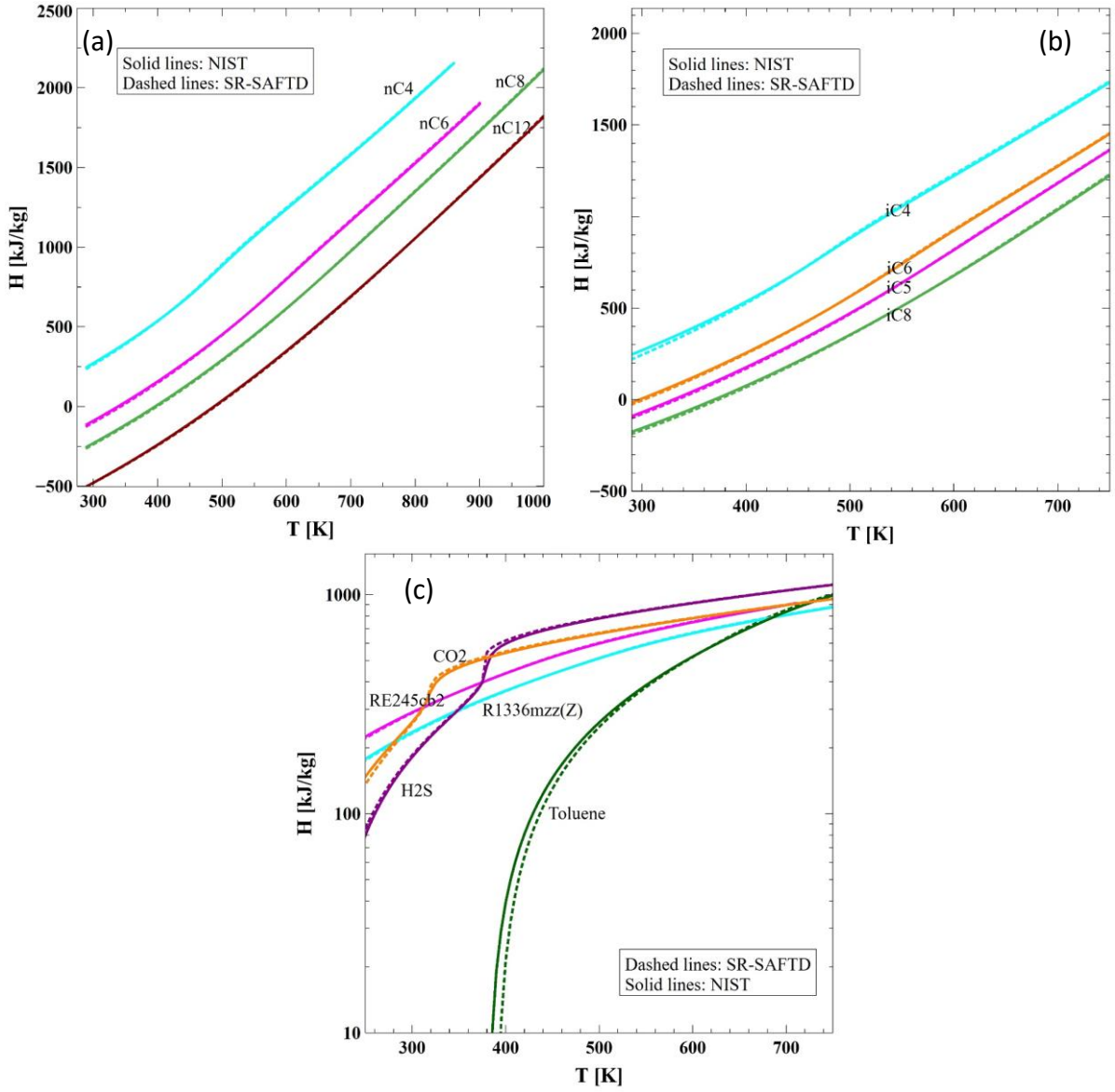


Figure 4.27: Temperature dependence of the single phase enthalpy predicted with the SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.. (a) n-alkanes, (b) iso-alkanes, (c) polar components.

4.9.2.5 Single phase density

As for the enthalpy, the single phase density behaviour at the same thermodynamic conditions has also been investigated. The predictions obtained with the SR-SAFTD are shown in Figure 4.28, and are in excellent agreement with those of the semi-rigid coarse grained model. When

compared to experimental data, the SR-SAFTD gives also very good predictions for all molecules. A slight overestimation is observed at low temperatures as the chain length increases for the iso and normal alkanes similarly to what was observed for the equilibrium properties, while a slight underestimation is obtained at high temperatures. The density behaviour close to the critical point for the H_2S and the CO_2 molecules is perfectly reproduced.

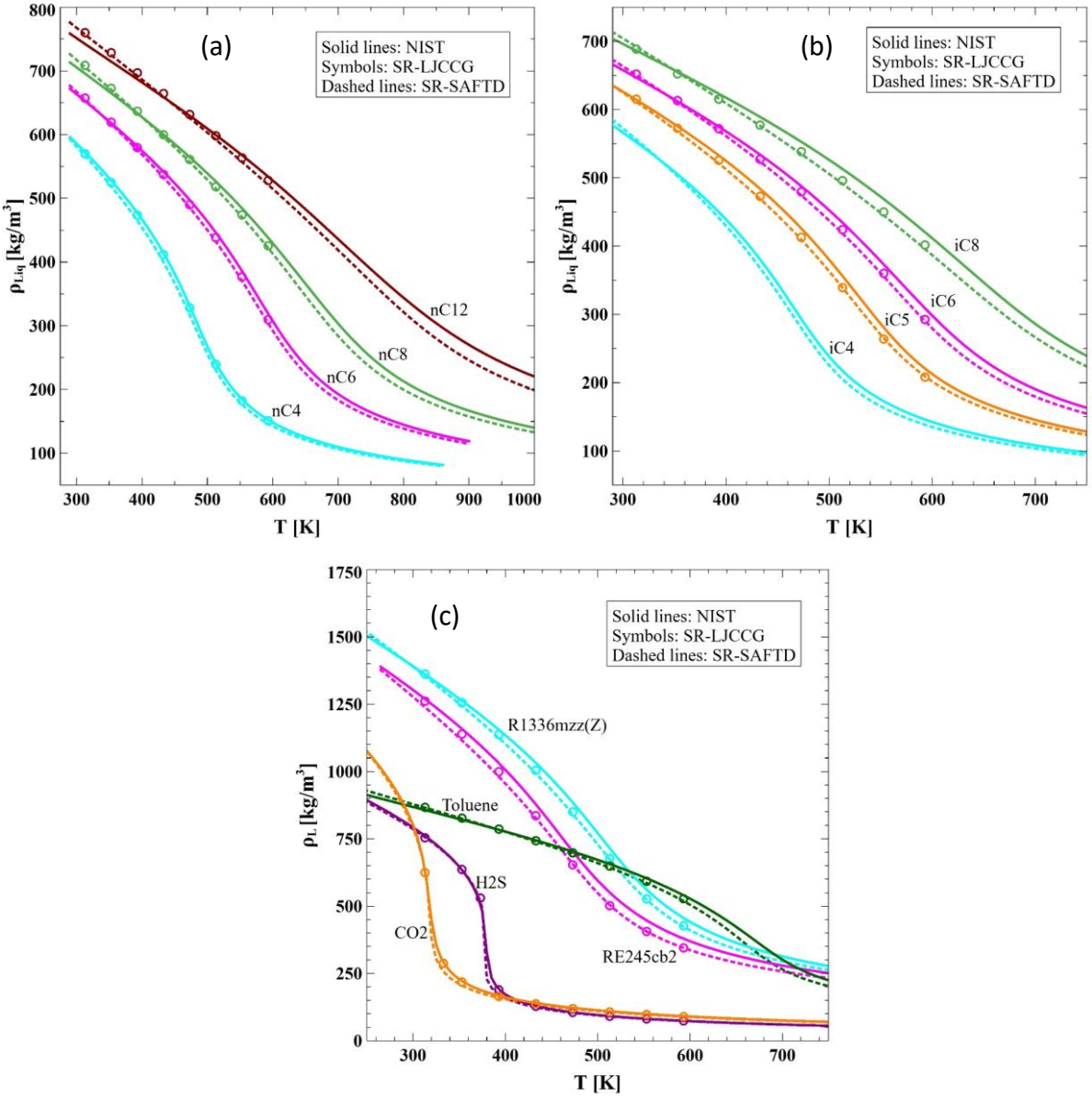


Figure 4.28: Temperature dependence of the single phase density predicted with the SR-SAFTD vs the correlated experimental data from NIST database at P=10MPa.. (a) n-alkanes, (b) iso-alkanes, (c) polar components.

From all these results, one may underline the robustness of the strategy of the parametrization as it allows obtaining consistent molecular parameters that can safely be transferred over a wide range of thermodynamic conditions. Moreover, the good results on the derivative properties and the enthalpy, which were not included in the fit, highlights the robustness and the representability of the obtained parameters.

4.9.3 Consistency check and transferability of the parameters

A major advantage of a SAFT equation of state resides in its solid physical basis. However, this latter can easily be lost if care has not been taken when coming to the parameter estimation due to the multiple possible set of parameters that may give the same predictions. Even though this risk becomes very low within the SR-SAFT framework due to the simplicity of the CS parametrization strategy, the user should always find a way to confirm the consistency of the estimated parameters. For instance, normal alkanes, or iso-alkanes, are modelled as a homonuclear chains, and thus, it is expected that the parameter within the same family should behave in a regular way as it is always the same functional group which is added to grow the chain. Such a consistency check has already been shown in the work of Pàmies [52] for the soft-SAFT and Lafitte [70] for the SAFT-VR-Mie first version, where the intermolecular interaction parameters and the number of segments forming the chain follow regular trends when plotted versus the molecular weight. Indeed, the parameters estimated from the SR-SAFTD in this work also follow the same trends as shown in Figure 4.29

More precisely, the energetic well-depth parameter increases rapidly for the small n-alkanes and then converges to a certain asymptotic limit, which is physically consistent, as the additional n^{th} functional group would not perturb the mean energy of the molecule which should remain finite for the polyethylene. The same reasoning also applies to the collisional diameter of the LJ potential which reflects the volume of the molecules. However, the number of segments is linear with molecular weight which an indication of the progressive addition of CH_2 functional groups.

In Figure 4.29, comparison is made against the parameters of the soft-SAFT taken from the work of Pàmies [52] and also against the fully flexible SAFTD of Johnson with the parameters taken from the work of Zhang [104]. While the results for the segments are practically superposed for the three LJ SAFT versions, the size and the energetic parameters shows small differences. Of course, such deviations reflect the differences in the optimization strategy used for the parameters estimation. The SR-SAFTD exhibits slightly lower σ and ϵ compared to the other versions, but the deviations for σ are more pronounced. Consequently, the SR-SAFTD models a less cohesive fluid but a denser one. As can be seen in Figure 4.22, the SR-SAFTD shows lower critical temperature of the n-hexane and the n-dodecane but slightly higher saturated liquid density at low temperatures. In other words, this deviation observed in sigma, reflects somewhat its slight dependence in temperature which is a common correction in the SAFT variants based on the hard sphere reference fluid to take into account the softness of the fluid model.

The rigidity parameter is specific to the SR-SAFT and it is thus this parameter that distinguishes this EoS from the other variants of SAFT. This behaviour of K^* exhibits a kind of hyperbolic decrease when plotted versus the molecular weight as shown inFigure 4.29 (d). For small molecules, the rigidity exhibits high values, which is consistent with the up-scaling approach on which the coarse grained model is based on. A monomer, which can represent a group of atoms in a real molecule is considered as a rigid unit. Two connected monomers are also considered as a fully rigid unit in the coarse graining approach. However, beyond two

monomers, the notion of molecular bending makes sense and so can be defined, which is the case of the n-butane. For the longest molecules, the rigidity is exhibiting a decreasing behaviour. This is also consistent as a polymer can locally have a certain stiffness (named Kuhn segment). However, globally this stiffness is reduced allowing the polymer to bend.

It is clear that for the case of C_6 and C_7, the optimized rigidity parameters seem to deviate from the regular decay behaviour exhibited by the other molecules of the family. It has already been shown that the viscosity is highly sensitive to the rigidity parameter, and thus, such deviations are nothing more than the signature of the deviation observed in the experimental data of the reduced saturated viscosity at T_r=0.7 for these two molecules when compared to their homologous series of alkanes as shown in Figure 4.30.

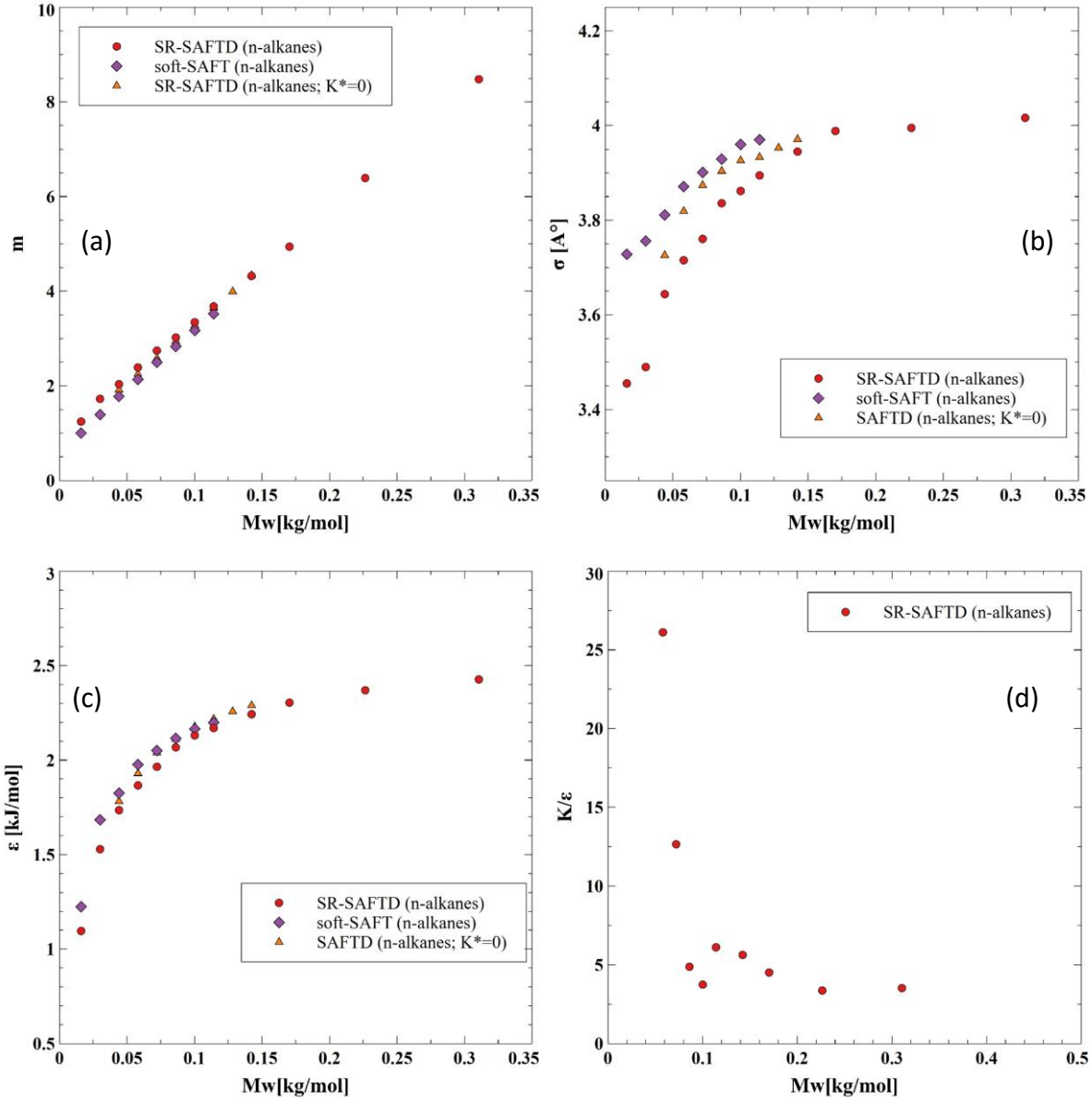


Figure 4.29: Optimized molecular parameters of the SR-SAFTD EOS for the series of n-alkanes Vs That of the fully flexible SAFTD and the fully flexible soft-SAFT. (a) σ Vs Mw. (b) ε Vs Mw. (c) m Vs Mw. (d) K/ε Vs Mw.

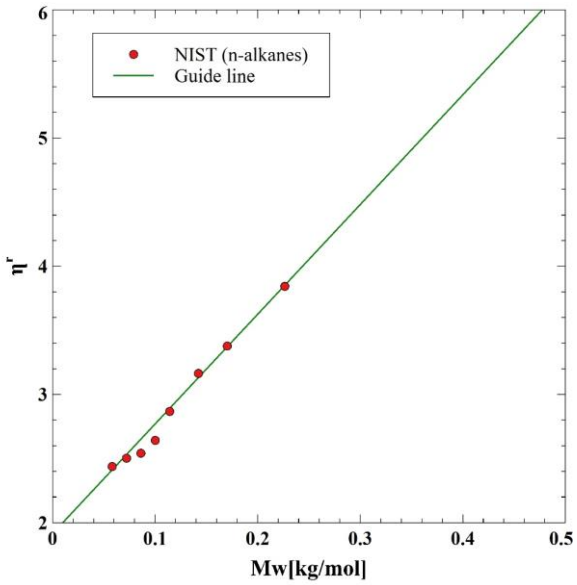


Figure 4.30: Reduced experimental viscosity of n-alkanes taken from NIST database Vs Mw

Now, regarding the transferability of the parameters to other molecules for which the experimental data needed for the optimization are not available, a classical way to obtain them is by exploiting their linear behaviour when plotted in terms of the molecular energy and volume instead of the atomic ones for the same family of molecules. The best example is that widely shared by most SAFT developers for n-alkanes molecules as shown by Blas [50] and Pàmies [52] for the soft-SAFT, McCabe [67] for the SAFT-VR, Tihic [111] for the PC-SAFT and Lafitte [70] for the first version of the SAFT-VR-Mie. Interestingly, the same behaviour is also found for the n-alkanes with the SR-SAFTD as shown in Figure 4.31.

Note that the case of methane is not included in the fit to not bias the fit as its parameter slightly deviate from the global trend. Regarding the rigidity, the n-hexane and n-heptane have been excluded from the fit due to their high deviation from the trend (as explained above). Thus, the fitting will provide higher value of stiffness constant for these two compounds. It is expected that this will only influence noticeably the viscosity at low temperatures for these two molecules.

The same work has been done for the iso-alkanes to obtain transferable parameters as shown in Figure 4.32. Similar to the case of the n-alkanes, deviations from the global trend for the stiffness constant is also observed for the iso-butane and iso-pentane and have then been excluded from the fit.

For the n-alkanes (for iso-alkanes see appendix 7.A.2), the correlations are given as follow:

$$m = a_m M_w + b_m \tag{4.174}$$

$$m\,\sigma^3[nm^3] = a_\sigma M_w + b_\sigma \tag{4.175}$$

$$m\,{}^{\epsilon}/_{k_B}[K] = a_\epsilon M_w + b_\epsilon \tag{4.176}$$

$${}^{K}/_{\epsilon} = \frac{a_K}{1+b_K Exp(c_K M_w)} \quad for\ Mw > Mw^{i-C3} \tag{4.177}$$

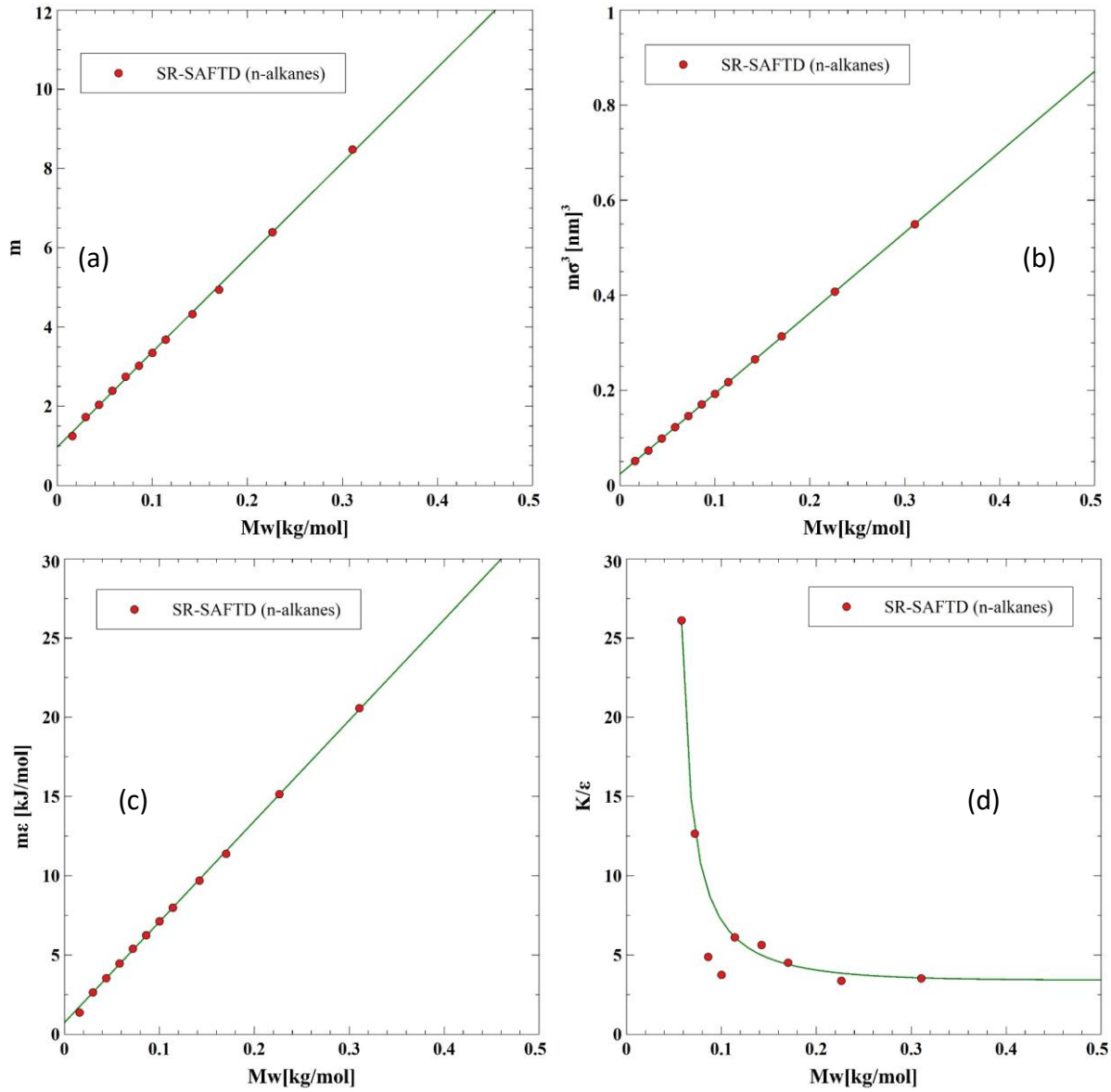


Figure 4.31: Transferable parameters of the SR-SAFTD EOS for the series of n-alkanes, correlated up to n-C_{22}. (a) m Vs Mw. (b) mσ³ Vs Mw. (c) mε Vs Mw. (d) K/ε Vs Mw.

where M_w is the molecular weight of the n-alkane expressed in $[\frac{kg}{mol}]$. The fitting constants are:

Normal alkanes	a_i	b_i	c_i
m	23.9475	0.9659	—
m σ^3	1.6949	0.02363	—
m ϵ/k_B	63.6103	0.7254	—
K/ϵ	3.4062	−1.7449	−11.9864

Table 4.1: Constants of the transferable parameters correlations for the n-alkanes

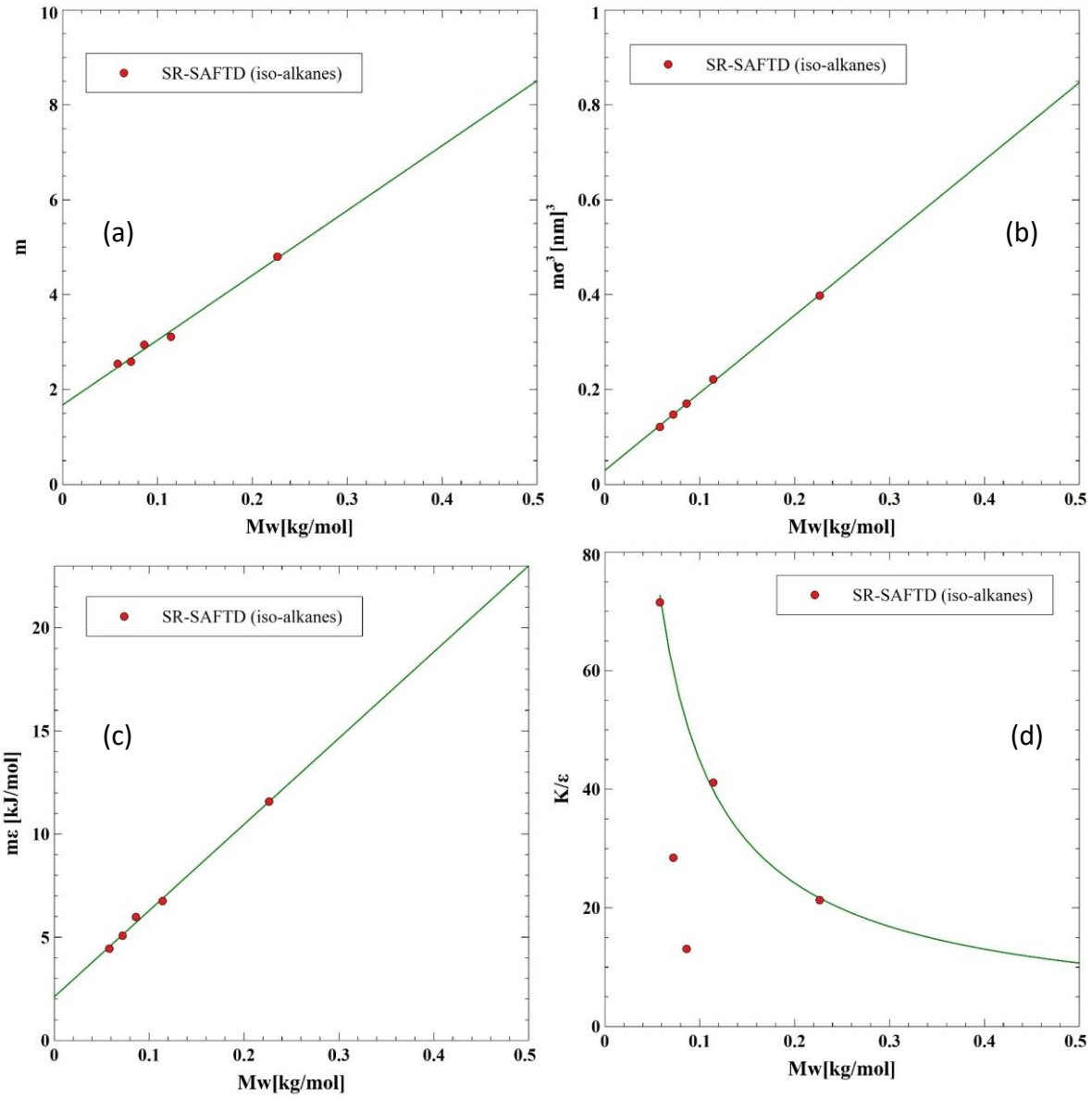


Figure 4.32: Transferable parameters of the SR-SAFTD EoS for the series of iso-alkanes, correlated up to i-C_{16}. (a) m Vs Mw. (b) mσ³ Vs Mw. (c) mε Vs Mw. (d) K/ε Vs Mw.

In order to check the validity of the correlations to predict properties of fluids not included in the fit, we have extrapolated them to obtain parameters of the normal-hexatriacontane (nC_{36}). Comparison is made between the exact parameters optimized on this molecule and those obtained from the correlations. In addition, we included data from molecular simulations and the predictions from the soft-SAFT with parameters obtained from similar correlations given by Pàmies and Vega [52]. Results are shown in Figure 4.34 for the LVE properties and saturation pressure. There are only small deviations on the saturation pressure at low temperature and a slight overestimation of the critical temperature compared to the predictions with the optimal parameters. The correlations can then be considered when experimental data needed for the optimization are not available. Higher deviations are obtained with the soft-SAFT, due probably to the fact that the fitting correlations for this EoS consider only small n-alkanes in the fit.

A second check is made upon the high-density linear polyethylene molecule constituted with 900 carbon atoms (molecular weight of 12602 g/mol). The molecule is modelled with around 303 segments for the SR-SAFTD and around 322 segments with the soft-SAFT. The experimental data are taken from the work of Rahman et al. [112] as well as the data for the SAFT-γ-Mie EoS which models the polyethylene with 300 segments. Figure 4.34 shows results for the three EoS models against experiments. The predictions of the SR-SAFTD and SAFT-γ-Mie are very close each other with slightly better prediction with this latter EoS at high pressures. However, both models overestimate the density for all the pressures. The soft-SAFT yields excellent predictions of the density compared to the two other EoSs. However, the slope (dρ/dT) seems to be better predicted by the SR-SAFTD and the SAFT-γ-Mie. The average deviations for the SR-SAFTD are of ~2.6%, ~3.8% and ~4.7% respectively from the lowest pressure to the highest pressure. The deviations for the SAFT-γ-Mie are ~2.35%, ~3.22, ~3.7% respectively. The soft-SAFT which yields the best predictions, thus, with much smaller deviations ~-0.62%, ~0.66%, ~1.49% respectively. Overall, all the EoSs tested here showed very reasonable results when extrapolated at the extreme limit.

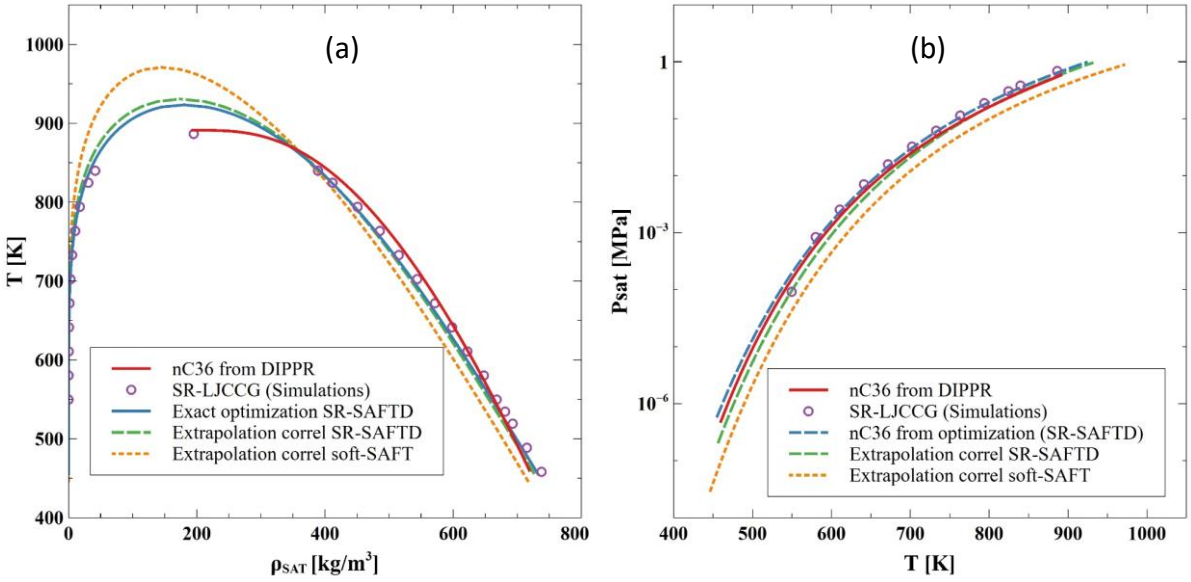


Figure 4.33: Prediction of the thermodynamic properties of nC_{36} using the SR-SAFTD and soft-SAFT with parameters obtained from the extrapolated correlations. (a) LVE, (b) Psat

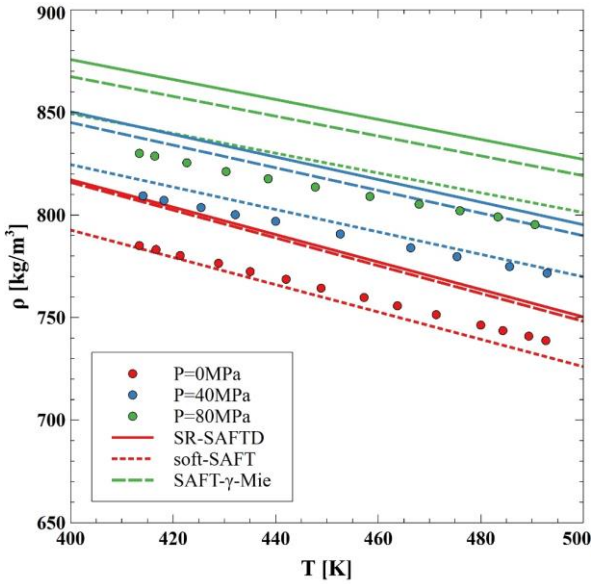


Figure 4.34: Liquid density of polyethylene at various pressures predicted using SR- SAFTD and soft-SAFT with parameters obtained from the extrapolation of the correlations. The prediction with SAFT-γ-Mie is also included taken from Rahman et al.2018 . The experimental (dots) data are also obtained from the same paper (see therein for therein exact origin).

4.10 Summary

Because of the interesting performance exhibited by the SR-LJCCG model presented in chapter 3, a SAFT equation of state providing the properties of this coarse grained molecular model would be of great interest to avoid performing extensive molecular simulations. However, before this work, such an EoS was not yet available. Thus, the goal of this part of the thesis was to develop a SAFT EoS for this fluid model.

In this chapter, we first provided an overview of the thermodynamic perturbation (TPT) theory of Wertheim, including the equations derived. Some of the strengths and weaknesses of the theory were discussed along this chapter, with a focus on the effect of the truncation at the first level. We found that TPT1-M yielded significant deviations when compared to molecular simulations, but TPT1-D performed much better due to the inclusion of structural information. Moreover, we determined that TPT1-D offers the ability to incorporate a bending or torsional potentials. Therefore, in a second step, a modification in the TPT1-D theory was proposed to include a bending potential. More precisely, thanks to molecular simulations, we proposed a scheme to calculate the radial distribution function that accounts for internal degrees of freedom, such as the bending potential. This new RDF in TPT1-D results in a more generalized TPT1-D referred to as SR-SAFTD (SR stands for Semi-Rigid) equation of state. This strategy proposed for the incorporation of the bending potential may be extended to other potentials related to the internal degrees of freedom. Moreover, this strategy is not specific to LJ-based EoS, but, can be transposed to other forms of intermolecular potentials, thus, used with any SAFT type model.

In the third step, the SR-SAFTD model developed has been validated on molecular simulations. Many properties were considered for the validation, such as the phase equilibria, the internal energy and the entropy over a large thermodynamic conditions. We showed that in all cases the model gives accurate predictions except very close to the critical conditions.

In the fourth step, the new equation of state was applied on real fluids with molecular parameters estimated using the corresponding state strategy developed for the coarse grained molecular model. This strategy was a novel aspect of this work, as it is based on the properties of the fluid model evaluated by molecular simulations, and it is for the first time that a transport property has been incorporated in the parametrization of a SAFT model. The SR-SAFTD has been tested on three different chemical families, including n-alkanes, iso-alkanes and some polar components. It was evaluated over a wide of thermodynamic conditions (saturation conditions, dense liquid and supercritical conditions) and on various properties such as phase equilibria, enthalpy and many derivatives properties. In all cases, the SR-SAFT was found to accurately reproduce the data of the SR-LJCCG model and also accurately predict experiments. Furthermore, we demonstrated that the parameters used in the SR-SAFTD are transferable within the same molecular family.

In addition to the success of the SR-SAFTD in predicting the thermodynamic properties of real fluids, it will be shown in the following chapter that this EoS stands out among others due to its ability in being used for predicting transport properties, specifically viscosity, when used in the entropy scaling approach.

4.11 References

1. Stell, G., Rasaiah, J. & Narang, H. Thermodynamic Perturbation Theory for Simple Polar Fluids. *Molecular Physics - MOL PHYS* **23**, 393–406 (1972).
2. Stell, G., Rasaiah, J. C. & Narang, H. Thermodynamic perturbation theory for simple polar fluids. II. *Molecular Physics* **27**, 1393–1414 (1974).
3. Wertheim, M. S. Fluids with highly directional attractive forces. I. Statistical thermodynamics. *J Stat Phys* **35**, 19–34 (1984).
4. Wertheim, M. S. Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations. *J Stat Phys* **35**, 35–47 (1984).
5. Wertheim, M. S. Fluids with highly directional attractive forces. III. Multiple attraction sites. *J Stat Phys* **42**, 459–476 (1986).
6. Wertheim, M. S. Fluids with highly directional attractive forces. IV. Equilibrium polymerization. *J Stat Phys* **42**, 477–492 (1986).
7. Wertheim, M. S. Fluids of dimerizing hard spheres, and fluid mixtures of hard spheres and dispheres. *J. Chem. Phys.* **85**, 2929–2936 (1986).
8. Wertheim, M. S. Thermodynamic perturbation theory of polymerization. *J. Chem. Phys.* **87**, 7323–7331 (1987).
9. Andersen, H. C. Cluster expansions for hydrogen-bonded fluids. I. Molecular association in dilute gases. *J. Chem. Phys.* **59**, 4714–4725 (1973).
10. Andersen, H. C. Cluster expansions for hydrogen bonded fluids. II. Dense liquids. *J. Chem. Phys.* **61**, 4985–4992 (1974).
11. Calais, J.-L. Statistical mechanics. J. E. Mayer and M. G. Mayer. 2nd ed. John Wiley & Sons, Inc., New York, 1977. Price: £17.50/$29.00. *International Journal of Quantum Chemistry* **15**, 259–259 (1979).
12. Chandler, D. & Pratt, L. R. Statistical mechanics of chemical equilibria and intramolecular structures of nonrigid molecules in condensed phases. *J. Chem. Phys.* **65**, 2925–2940 (1976).
13. HØye, J. S. & Olaussen, K. Statistical mechanical model with chemical reaction. *Physica A: Statistical Mechanics and its Applications* **104**, 435–446 (1980).
14. Dickman, R. & Hall, C. K. High density Monte Carlo simulations of chain molecules: Bulk equation of state and density profile near walls. *J. Chem. Phys.* **89**, 3168–3174 (1988).
15. Hill, T. L. *An Introduction to Statistical Thermodynamics*. (Courier Corporation, 1986).
16. Gray, C. G., Gubbins, K. E. & Joslin, C. G. *Theory of Molecular Fluids: Volume 2: Applications*. (OUP Oxford, 1984).
17. Hansen, J.-P. & McDonald, I. R. *Theory of Simple Liquids*. (Elsevier, 2006).
18. Marshall, B. D. Thermodynamic perturbation theory for associating fluids: Beyond first order. *arXiv preprint arXiv:1405.1775* (2014).
19. Febra, S. Ring formation in a statistical associating fluid theory framework. (2018).
20. Zmpitas, W. & Gross, J. Detailed pedagogical review and analysis of Wertheim's thermodynamic perturbation theory. *Fluid Phase Equilibria* **428**, 121–152 (2016).
21. Chapman, W. G., Gubbins, K. E., Joslin, C. G. & Gray, C. G. Theory and simulation of associating liquid mixtures. *Fluid Phase Equilibria* **29**, 337–346 (1986).
22. Joslin, C. G., Gray, C. G., Chapman, W. G. & Gubbins, K. E. Theory and simulation of associating liquid mixtures. II. *Molecular Physics* **62**, 843–860 (1987).
23. Jackson, G., Chapman, W. G. & Gubbins, K. E. Phase equilibria of associating fluids of spherical and chain molecules. *Int J Thermophys* **9**, 769–779 (1988).
24. Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. SAFT: Equation-of-state solution model for associating fluids. *Fluid Phase Equilibria* **52**, 31–38 (1989).

25. Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **29**, 1709–1721 (1990).
26. GALINDO, A., BURTON, S. J., JACKSON, G., VISCO, D. P. & KOFKE, D. A. Improved models for the phase behaviour of hydrogen fluoride: chain and ring aggregates in the SAFT approach and the AEOS model. *Molecular Physics* **100**, 2241–2259 (2002).
27. Müller, E. A. & Gubbins, K. E. Molecular-Based Equations of State for Associating Fluids: A Review of SAFT and Related Approaches. *Ind. Eng. Chem. Res.* **40**, 2193–2211 (2001).
28. Paricaud, P. Understanding the fluid phase behaviour of polymer systems with the SAFT theory. (2003).
29. Carnahan, N. F. & Starling, K. E. Equation of State for Nonattracting Rigid Spheres. *J. Chem. Phys.* **51**, 635–636 (1969).
30. Zhou, Y. & Stell, G. Chemical association in simple models of molecular and ionic fluids. III. The cavity function. *J. Chem. Phys.* **96**, 1507–1515 (1992).
31. Huang, S. H. & Radosz, M. Equation of state for small, large, polydisperse, and associating molecules. *ACS Publications* https://pubs.acs.org/doi/pdf/10.1021/ie00107a014 (2002) .
32. Johnson, J. K., Mueller, E. A. & Gubbins, K. E. Equation of State for Lennard-Jones Chains. *J. Phys. Chem.* **98**, 6413–6419 (1994).
33. FELIPE J. BLAS and LOURDES F. VEGA. Thermodynamic behaviour of homonuclear and heteronuclear Lennard-Jones chains with association sites from simulation and theory. *Molecular Physics* **92**, 135–150 (1997).
34. Gil-Villegas, A. *et al.* Statistical associating fluid theory for chain molecules with attractive potentials of variable range. *The Journal of Chemical Physics* **106**, 4168–4186 (1997).
35. Davies, L. A., Gil-Villegas, A. & Jackson, G. Describing the Properties of Chains of Segments Interacting Via Soft-Core Potentials of Variable Range with the SAFT-VR Approach. *International Journal of Thermophysics* **19**, 675–686 (1998).
36. Gross, J. & Sadowski, G. Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules. *Ind. Eng. Chem. Res.* **40**, 1244–1260 (2001).
37. Cotterman, R. L., Schwarz, B. J. & Prausnitz, J. M. Molecular thermodynamics for fluids at low and high densities. Part I: Pure fluids containing small or large molecules. *AIChE Journal* **32**, 1787–1798 (1986).
38. Chen, S. S. & Kreglewski, A. Applications of the Augmented van der Waals Theory of Fluids.: I. Pure Fluids. *Berichte der Bunsengesellschaft für physikalische Chemie* **81**, 1048–1052 (1977).
39. Gregg, C. J., Chen, S., Stein, F. P. & Radosz, M. Phase behavior of binary ethylene-propylene copolymer solutions in sub- and supercritical ethylene and propylene. *Fluid Phase Equilibria* **83**, 375–382 (1993).
40. Chen, S., Banaszak, M. & Radosz, M. Phase Behavior of Poly(ethylene-1-butene) in Subcritical and Supercritical Propane: Ethyl Branches Reduce Segment Energy and Enhance Miscibility. *Macromolecules* **28**, 1812–1817 (1995).
41. Chen, S. J. & Radosz, M. Density-tuned polyolefin phase equilibria. 1. Binary solutions of alternating poly(ethylene-propylene) in subcritical and supercritical propylene, 1-butene, and 1-hexene. Experiment and Flory-Patterson model. *ACS Publications* https://pubs.acs.org/(2002).
42. Jog, P. K., Sauer, S. G., Blaesing, J. & Chapman, W. G. Application of dipolar chain theory to the phase behavior of polar fluids and mixtures. *Industrial & engineering chemistry research* **40**, 4641–4648 (2001).
43. Economou, I. G. Statistical Associating Fluid Theory: A Successful Model for the Calculation of Thermodynamic and Phase Equilibrium Properties of Complex Fluid Mixtures. *Ind. Eng. Chem. Res.* **41**, 953–962 (2002).

44. Johnson, J. K., Zollweg, J. A. & Gubbins, K. E. The Lennard-Jones equation of state revisited. *Molecular Physics* **78**, 591–618 (1993).
45. Kolafa, J. & Nezbeda, I. The Lennard-Jones fluid: an accurate analytic and theoretically-based equation of state. *Fluid Phase Equilibria* **100**, 1–34 (1994).
46. Thol, M. *et al.* Equation of State for the Lennard-Jones Fluid. *Journal of Physical and Chemical Reference Data* **45**, 023101 (2016).
47. Müller, E. & Gubbins, K. An Equation of State for Water from a Simplified Intermolecular Potential. *Industrial & Engineering Chemistry Research - IND ENG CHEM RES* **34**, (1995).
48. Kraska, T. & Gubbins, K. E. Phase Equilibria Calculations with a Modified SAFT Equation of State. 1. Pure Alkanes, Alkanols, and Water. *Ind. Eng. Chem. Res.* **35**, 4727–4737 (1996).
49. Kraska, T. & Gubbins, K. E. Phase Equilibria Calculations with a Modified SAFT Equation of State. 2. Binary Mixtures of n-Alkanes, 1-Alkanols, and Water. *Ind. Eng. Chem. Res.* **35**, 4738–4746 (1996).
50. Blas, F. J. & Vega, L. F. Prediction of Binary and Ternary Diagrams Using the Statistical Associating Fluid Theory (SAFT) Equation of State. *Ind. Eng. Chem. Res.* **37**, 660–674 (1998).
51. Blas, F. J. & Vega, L. F. Thermodynamic properties and phase equilibria of branched chain fluids using first- and second-order Wertheim's thermodynamic perturbation theory. *J. Chem. Phys.* **115**, 3906–3915 (2001).
52. Pàmies, J. C. & Vega, L. F. Vapor−Liquid Equilibria and Critical Behavior of Heavy n-Alkanes Using Transferable Parameters from the Soft-SAFT Equation of State. *Ind. Eng. Chem. Res.* **40**, 2532–2543 (2001).
53. Llovell, F. & Vega, L. F. Prediction of Thermodynamic Derivative Properties of Pure Fluids through the Soft-SAFT Equation of State. *J. Phys. Chem. B* **110**, 11427–11437 (2006).
54. Dias, A. M. A., Pàmies, J. C., Coutinho, J. A. P., Marrucho, I. M. & Vega, L. F. SAFT Modeling of the Solubility of Gases in Perfluoroalkanes. *J. Phys. Chem. B* **108**, 1450–1457 (2004).
55. Ferreira, R., Pedrosa, N., Marrucho, I. M. & Rebelo, L. P. N. Biodegradable Polymer-Phase Behavior: Liquid–Liquid Equilibrium of Ethyl Lactate and Poly(Lactic Acid). *Journal of Chemical & Engineering Data* **53**, 588–590 (2008).
56. Belkadi, A., Llovell, F., Gerbaud, V. & Vega, L. F. Modeling the vapor–liquid equilibrium and association of nitrogen dioxide/dinitrogen tetroxide and its mixtures with carbon dioxide. *Fluid Phase Equilibria* **266**, 154–163 (2008).
57. Andreu, J. S. & Vega, L. F. Capturing the Solubility Behavior of CO2 in Ionic Liquids by a Simple Model. *J. Phys. Chem. C* **111**, 16028–16034 (2007).
58. Barker, J. A. & Henderson, D. Perturbation Theory and Equation of State for Fluids: The Square-Well Potential. *J. Chem. Phys.* **47**, 2856–2861 (1967).
59. Barker, J. A. & Henderson, D. Perturbation Theory and Equation of State for Fluids. II. A Successful Theory of Liquids. *J. Chem. Phys.* **47**, 4714–4721 (1967).
60. Barker, J. A. & Henderson, D. What is 'liquid'? Understanding the states of matter. *Rev. Mod. Phys.* **48**, 587–671 (1976).
61. Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **22**, 1420–1426 (1954).
62. Zhang, B.-J. Calculating thermodynamic properties from perturbation theory: I. An analytic representation of square-well potential hard-sphere perturbation theory. *Fluid Phase Equilibria* **154**, 1–10 (1999).
63. Paricaud, P. A general perturbation approach for equation of state development: Applications to simple fluids, ab initio potentials, and fullerenes. *J. Chem. Phys.* **124**, 154505 (2006).
64. McCabe, C., Gil-Villegas, A. & Jackson, G. Predicting the High-Pressure Phase Equilibria of Methane + n-Hexane Using the SAFT-VR Approach. *J. Phys. Chem. B* **102**, 4183–4188 (1998).

65. Bonifácio, R. P., Filipe, E. J. M., McCabe, C., Costa Gomes, M. F. & Pádua, A. A. H. Predicting the solubility of xenon in n-hexane and n-perfluorohexane: A simulation and theoretical study. *Molecular Physics* **100**, 2547–2553 (2002).
66. McCabe, C. & Kiselev, S. B. A crossover SAFT-VR equation of state for pure fluids: preliminary results for light hydrocarbons. *Fluid Phase Equilibria* **219**, 3–9 (2004).
67. McCabe, C. & Jackson, G. SAFT-VR modelling of the phase equilibrium of long-chain n-alkanes. *Physical Chemistry Chemical Physics* **1**, 2057–2064 (1999).
68. Peng, Y., Goff, K. D., dos Ramos, M. C. & McCabe, C. Developing a predictive group-contribution-based SAFT-VR equation of state. *Fluid Phase Equilibria* **277**, 131–144 (2009).
69. Haslam, A. J., Galindo, A. & Jackson, G. Prediction of binary intermolecular potential parameters for use in modelling fluid mixtures. *Fluid Phase Equilibria* **266**, 105–128 (2008).
70. Lafitte, T., Bessieres, D., Piñeiro, M. M. & Daridon, J.-L. Simultaneous estimation of phase behavior and second-derivative properties using the statistical associating fluid theory with variable range approach. *J Chem Phys* **124**, 024509 (2006).
71. Lafitte, T. *et al.* Accurate statistical associating fluid theory for chain molecules formed from Mie segments. *J Chem Phys* **139**, 154504 (2013).
72. Song, Y., Lambert, S. M. & Prausnitz, J. M. A Perturbed Hard-Sphere-Chain Equation of State for Normal Fluids and Polymers. *ACS Publications* https://pubs.acs.org/(2002).
73. Boublík, T. Hard-Sphere Equation of State. *J. Chem. Phys.* **53**, 471–472 (1970).
74. Mansoori, G. A., Carnahan, N. F., Starling, K. E. & Leland, T. W. Equilibrium Thermodynamic Properties of the Mixture of Hard Spheres. *J. Chem. Phys.* **54**, 1523–1525 (1971).
75. Liu, H. & Hu, Y. Molecular thermodynamic theory for polymer systems part II. Equation of state for chain fluids. *Fluid Phase Equilibria* **122**, 75–97 (1996).
76. Gross, J. & Sadowski, G. Modeling Polymer Systems Using the Perturbed-Chain Statistical Associating Fluid Theory Equation of State. *Ind. Eng. Chem. Res.* **41**, 1084–1093 (2002).
77. Gross, J. & Sadowski, G. Application of the Perturbed-Chain SAFT Equation of State to Associating Systems. *Ind. Eng. Chem. Res.* **41**, 5510–5515 (2002).
78. Gross, J., Spuhl, O., Tumakaka, F. & Sadowski, G. Modeling Copolymer Systems Using the Perturbed-Chain SAFT Equation of State. *Ind. Eng. Chem. Res.* **42**, 1266–1274 (2003).
79. Gonzalez, D. L., Ting, P. D., Hirasaki, G. J. & Chapman, W. G. Prediction of Asphaltene Instability under Gas Injection with the PC-SAFT Equation of State. *Energy Fuels* **19**, 1230–1234 (2005).
80. Yarrison, M. & Chapman, W. G. A systematic study of methanol+n-alkane vapor–liquid and liquid–liquid equilibria using the CK-SAFT and PC-SAFT equations of state. *Fluid Phase Equilibria* **226**, 195–205 (2004).
81. Grenner, A., Kontogeorgis, G. M., von Solms, N. & Michelsen, M. L. Application of PC-SAFT to glycol containing systems – PC-SAFT towards a predictive approach. *Fluid Phase Equilibria* **261**, 248–257 (2007).
82. Ruether, F. & Sadowski, G. Modeling the Solubility of Pharmaceuticals in Pure Solvents and Solvent Mixtures for Drug Process Design. *Journal of Pharmaceutical Sciences* **98**, 4205–4215 (2009).
83. Tihic, A., Kontogeorgis, G. M., von Solms, N., Michelsen, M. L. & Constantinou, L. A Predictive Group-Contribution Simplified PC-SAFT Equation of State: Application to Polymer Systems. *Ind. Eng. Chem. Res.* **47**, 5092–5101 (2008).
84. Marshall, B. D. Dual chain perturbation theory: A new equation of state for polyatomic molecules. *J. Chem. Phys.* **144**, 164104 (2016).
85. Moller, E. A. & Gubbins, K. E. Simulation of hard triatomic and tetratomic molecules A test of associating fluid theories. 20.

86. Honnell, K. G. & Hall, C. K. A new equation of state for athermal chains. *J. Chem. Phys.* **90**, 1841–1855 (1989).
87. Flory, P. J. Thermodynamics of High Polymer Solutions. *J. Chem. Phys.* **10**, 51–61 (1942).
88. Chang, J. & Sandler, S. I. An equation of state for the hard-sphere chain fluid: theory and Monte Carlo simulation. *Chemical Engineering Science* **49**, 2777–2791 (1994).
89. Ghonasgi, D. & Chapman, W. G. A new equation of state for hard chain molecules. *J. Chem. Phys.* **100**, 6633–6639 (1994).
90. Johnson, J. K. Perturbation theory and computer simulations for linear and ring model polymers. *J. Chem. Phys.* **104**, 1729–1742 (1996).
91. Blas, F. J. & Vega, L. F. Improved vapor–liquid equilibria predictions for Lennard-Jones chains from the statistical associating fluid dimer theory: Comparison with Monte Carlo simulations. *The Journal of Chemical Physics* **115**, 4355–4358 (2001).
92. Widom, B. Some Topics in the Theory of Fluids. *J. Chem. Phys.* **39**, 2808–2812 (1963).
93. Widom, B. Potential-distribution theory and the statistical mechanics of fluids. *J. Phys. Chem.* **86**, 869–872 (1982).
94. Mejía, A., Herdes, C. & Müller, E. A. Force Fields for Coarse-Grained Molecular Simulations from a Corresponding States Correlation. *Ind. Eng. Chem. Res.* **53**, 4131–4141 (2014).
95. PAMIES, J. C. & VEGA, L. F. Critical properties of homopolymer fluids studied by a Lennard-Jones statistical associating fluid theory. *Molecular Physics* **100**, 2519–2529 (2002).
96. Moine, E., Piña-Martinez, A., Jaubert, J.-N., Sirjean, B. & Privat, R. I-PC-SAFT: An Industrialized Version of the Volume-Translated PC-SAFT Equation of State for Pure Components, Resulting from Experience Acquired All through the Years on the Parameterization of SAFT-Type and Cubic Models. *Ind. Eng. Chem. Res.* **58**, 20815–20827 (2019).
97. Llovell, F., Pàmies, J. C. & Vega, L. F. Thermodynamic properties of Lennard-Jones chain molecules: Renormalization-group corrections to a modified statistical associating fluid theory. *J. Chem. Phys.* **121**, 10715–10724 (2004).
98. Llovell, F., Peters, C. J. & Vega, L. F. Second-order thermodynamic derivative properties of selected mixtures by the soft-SAFT equation of state. *Fluid Phase Equilibria* **248**, 115–122 (2006).
99. Forte, E., Llovell, F., Vega, L. F., Trusler, J. P. M. & Galindo, A. Application of a renormalization-group treatment to the statistical associating fluid theory for potentials of variable range (SAFT-VR). *J. Chem. Phys.* **134**, 154102 (2011).
100. Bymaster, A., Emborsky, C., Dominik, A. & Chapman, W. G. Renormalization-Group Corrections to a Perturbed-Chain Statistical Associating Fluid Theory for Pure Fluids Near to and Far from the Critical Region. *Ind. Eng. Chem. Res.* **47**, 6264–6274 (2008).
101. Dufal, S., Lafitte, T., Galindo, A., Jackson, G. & Haslam, A. J. Developing intermolecular-potential models for use with the SAFT-VR Mie equation of state. *AIChE Journal* **61**, 2891–2912 (2015).
102. dos Ramos, M. C., Docherty, H., Blas, F. J. & Galindo, A. Application of the generalised SAFT-VR approach for long-ranged square-well potentials to model the phase behaviour of real fluids. *Fluid Phase Equilibria* **276**, 116–126 (2009).
103. Ramírez-Vélez, N., Privat, R., Piña-Martinez, A. & Jaubert, J.-N. Assessing the performance of non-associating SAFT-type equations of state to reproduce vapor pressure, liquid density, enthalpy of vaporization, and liquid heat capacity data of 1800 pure fluids. *AIChE Journal* **68**, e17722 (2022).
104. Zhang, Y. & Chapman, W. G. Modeling Thermodynamic Properties of Isomeric Alkanes with a New Branched Equation of State. *Ind. Eng. Chem. Res.* **57**, 1679–1688 (2018).
105. Jorgensen, W. L., Madura, J. D. & Swenson, C. J. Optimized intermolecular potential functions for liquid hydrocarbons. *J. Am. Chem. Soc.* **106**, 6638–6646 (1984).

106. Bourasseau, E., Ungerer, P., Boutin, A. & Fuchs, A. H. Monte Carlo simulation of branched alkanes and long chain n -alkanes with anisotropic united atoms intermolecular potential. *Molecular Simulation* **28**, 317–336 (2002).
107. Oliveira, M. B., Llovell, F., Coutinho, J. A. P. & Vega, L. F. New Procedure for Enhancing the Transferability of Statistical Associating Fluid Theory (SAFT) Molecular Parameters: The Role of Derivative Properties. 14.
108. Ramírez-Vélez, N., Piña-Martinez, A., Jaubert, J.-N. & Privat, R. Parameterization of SAFT Models: Analysis of Different Parameter Estimation Strategies and Application to the Development of a Comprehensive Database of PC-SAFT Molecular Parameters. *J. Chem. Eng. Data* **65**, 5920–5932 (2020).
109. de Villiers, A. J., Schwarz, C. E., Burger, A. J. & Kontogeorgis, G. M. Evaluation of the PC-SAFT, SAFT and CPA equations of state in predicting derivative properties of selected non-polar and hydrogen-bonding compounds. *Fluid Phase Equilibria* **338**, 1–15 (2013).
110. Llovell, F., Vega, L. F., Anisimov, M. A. & Sengers, J. V. Incorporating critical divergence of isochoric heat capacity into the soft-SAFT equation of state. *AIChE Journal* **61**, 3073–3080 (2015).
111. Tihic, A., Kontogeorgis, G. M., von Solms, N. & Michelsen, M. L. Applications of the simplified perturbed-chain SAFT equation of state using an extended parameter table. *Fluid Phase Equilibria* **248**, 29–43 (2006).
112. Rahman, S. *et al.* SAFT-γ Force Field for the Simulation of Molecular Fluids. 5. Hetero-Group Coarse-Grained Models of Linear Alkanes and the Importance of Intramolecular Interactions. *J. Phys. Chem. B* **122**, 9161–9177 (2018).

Chapter 5. Entropy scaling based SR-SAFTD for transport properties

Contents

5.1 Overview of the evolution and application of the entropy scaling approach 170
5.2 Towards a corresponding state entropy scaling for the LJC fluid model 172
5.3 Application of the entropy scaling for the real fluids 181
5.4 Summary 185
5.5 References 186

In the previous chapter, we developed the SR-SAFTD equation of state which integrates a rigidity parameter of the bending type. This EoS accurately reproduced the thermodynamic properties of the SR-LJCCG model developed in chapter 3, making it a reliable alternative to extensive molecular simulations. In this chapter, to avoid performing these simulations to obtain transport properties, an entropy scaling approach will be developed for for the semi-rigid fluid model using the new SR-SAFTD EoS. Goel et al. [1] and later by Galliéro [2] previously applied this approach succefully to LJC fluids. However, to the best of our knowledge, it has never been applied to molecular models with internal rigidity even if of potential interest [3]. The combination of the SR-SAFTD equation of state with an entropy-based scaling approach should lead to good viscosity prediction provided that a simple entropy scaling would work as for the fully flexible chains. More important, and this is one of the originalities of this work, using this entropy-based approach would allow the SR-SAFTD to be capable to simultaneously allows the prediction of thermodynamic and transport properties in a single framework.

5.1 Overview of the evolution and application of the entropy scaling approach

The entropy scaling approach was originally proposed by Rosenfeld [4] in 1977. This approach postulates that, there is a universal scaling between the entropy and transport properties such as viscosity. More specifically, it is the reduced transport property which is related to the residual (sometimes called "excess") entropy, given in the case of viscosity by this relation:

$$\eta^r = \frac{\eta^*}{\eta_{ref}} = f(S^*_{res}) \tag{5.1}$$

where η^* is the viscosity in the reduced units, S^*_{res} is the residual entropy in reduced units, η_{ref} is the reference viscosity obtained thanks to a dimensional analysis, and is given by:

$$\eta_{ref} = \rho^{*2/3}(Mk_BT^*)^{1/2} \tag{5.2}$$

where ρ^* is the number density, M is the molecular mass and k_B is the Boltzmann constant.

Rosenfeld showed that the entropy scaling relation can be expressed by the following equation:

$$\eta^r = aExp(-bS^*_{res}) \tag{5.3}$$

where a = 0.2 and b = 0.8 are the fitting constants he deduced from simulations on a Lennard-Jones fluid.

Such a relation is valid for simple fluids in the dense states only. This is because at extremely low density, the scaling diverges due to the fact that, while the numerator tends to a certain finite non-null value provided by the kinetic theory of gases, the denominator tends to zero. In 1999, Rosenfeld [5] noted that, for different inverse power law model interaction, the reduced transport properties scale with the residual entropy at power -2/3. The relation for the dilute systems obeys then the following in the case of viscosity:

$$\eta^r = {S^*_{res}}^{-2/3} \tag{5.4}$$

Because of this divergence in the scaling proposed by Rosenfeld, several other authors have proposed alternative approaches. One approach is that proposed by Novak [6], who suggested to use the low density limit obtained from Chapman-Enskog theory to reduce viscosity instead of using macroscopic properties. This method, known as "microscopic reduction" uses molecular parameters to eliminate the diverging behaviour, extending then the approach to low density limit. The dimensionless viscosity writes then:

$$\eta^r = \frac{\eta^*}{\eta_{CE}} \tag{5.5}$$

where η_{CE} is the Chapman-Enskog zero density limit viscosity given by:

$$\eta_{CE} = \frac{5}{16\Omega^{(2,2)}(\sigma^2)}\sqrt{M\frac{k_B T^*}{\pi}} \tag{5.6}$$

where $\Omega^{(2,2)}$ is the collisional integral that can be obtained by the correlation of Neufeld et al. [7], and *M* is the molecular mass.

Another scaling was proposed by Galliéro et al. [2], which relies on the residual viscosity instead of the total reduced viscosity used in the Rosenfeld approach. The scaling is written as:

$$\eta^r_{res} = \frac{\eta^* - \eta^*_0}{\eta_{ref}} = f(S^*_{res}) \tag{5.7}$$

where η^*_0 is the zero density limit of the viscosity. The approach of Galliéro did not fully solve the incorrect scaling issue of the low density-limit, but, it extends the linear regime observed for the dense medium in the semi-log scale to lower residual entropies.

A third option is that proposed by Bell et al. [8–10] who used both the modifications proposed by Rosenfeld for the low density limit and that of Galliéro by using a residual viscosity instead of total viscosity.

$$\eta^r_{res} = \frac{\eta^* - \eta^*_0}{\eta_{ref}} S^{2/3}_{res} = f(S^*_{res}) \tag{5.8}$$

The Bell's approach was found to yield better scaling at low density limit compared to that of Galliéro, but, with data more dispersed than with the approach of Novak. In the high density regime, both Galliéro's and Bell's approaches perform better than the Novak approach as shown in our previous study [11] for the monomer.

The last approach, is that proposed by Dehlouz et al. [12], where they proposed to use a reduced residual entropy (with respect to the critical entropy "S^*_C") combined with the logarithm of the same quantity to alleviate the incorrect scaling at the low density limit. The expression is given by the following relation:

$$\eta^r = \frac{\eta^*}{\eta_{ref}} = f\left(\frac{S^*_{res}}{S^*_C} + ln\left(\frac{S^*_{res}}{S^*_C}\right)\right) \tag{5.9}$$

While the first term is preponderant in liquids, the second term in logarithm is dominant in gases. This approach allows obtaining a better scaling (less dispersed data) and covers the whole space of thermodynamic conditions.

During the last five years, the entropy scaling framework has been successfully applied to real pure fluids as well as their mixtures, including alkanes, iso-alkanes, refrigerants, alcohols, amines, esters, ethers as shown, among others, by the works of Lötgering-Lin et al. [13], that of Bell et al. [9, 10, 14] and the recent work of Dehlouz et al. [12]

5.2 Towards a corresponding state entropy scaling for the LJC fluid model

As pointed out previously, the main challenge is to find the best variables that induces the best scaling between viscosity and entropy. The chosen variables should imply as low as possible scattered data, extend the linear regime to cover a broad range of thermodynamic conditions in particular by eliminating the incorrect scaling at low densities and some complex behaviour at very high densities (that will be shown latter).

As a first step, it is interesting to test the existing approaches discussed previously on the LJ chains model, using MD data for the viscosities and the SR-SAFTD EoS for the residual entropies. Note, that the dilute/zero-density viscosity has been obtained from MD simulations and not from the Chapman-Enskog theory. The case of fully flexible LJ chains has been already treated by Goel et al. [1] using the original approach of Rosenfeld, and later by Galliéro et al. [2] using the Galliéro's approach. The case of the Novak, Bell and Dehlouz approaches have not yet been tested, thus, it has been decided to benchmark all of them together in this work. The results are reported in Figure 5.1 for the four different approaches. Note that, the values for the entropy are given per molecule/unit noted as $S^{*,chain}$. In the case of the Dehlouz approach, the abscissa consists in a ratio of two entropy quantities, thus, both molecule or atomic units are equivalent, so for simplicity of the notation we used are that of the atomic unit (S^* instead of $S^{*,chain}$).

The approach of Galliéro show very good scaling over the whole thermodynamic conditions, but its incorrect scaling at low densities makes it uses difficult for the prediction of the viscosity of gases. On the contrary, the approach of Novak shows excellent scaling at low densities, but shows scattered data at intermediate densities which is more pronounced for small molecules. The approach of Dehlouz is exhibiting a peculiar behaviour. While at high and low densities, all curves tend to be superposed, except for the dimer, at the intermediate densities they are all separated. However, when the curves are analysed separately, globally a good scaling is observed. The approach of Bell is found to be the best to deal with the fully flexible LJ chains, showing good trends over the whole thermodynamic space.

In a second step, it is interesting to test how these approaches could be applied to the case of the semi-rigid LJ chains model. Two distinct values of rigidities are considered, K^*=5 and K^*=10, in addition to the fully flexible case. The results obtained in Figure 5.2 show that none of the approaches is fully adequate for the semi-rigid chains. However, more in details, it appears that the approaches of Galliéro, Novak and Bell show that, the semi-rigid chains are exhibiting similar trends than for the fully flexible case, and are found close each other

particularly in the high density region. Still, when approaching the intermediate and low density regions, the data are very scattered. In the case of the Dehlouz approach, in addition to the fact that the data are very scattered similarly to the previous approaches, the behaviour exhibited for the semi-rigid chains is found completely different from that of the fully flexible chains.

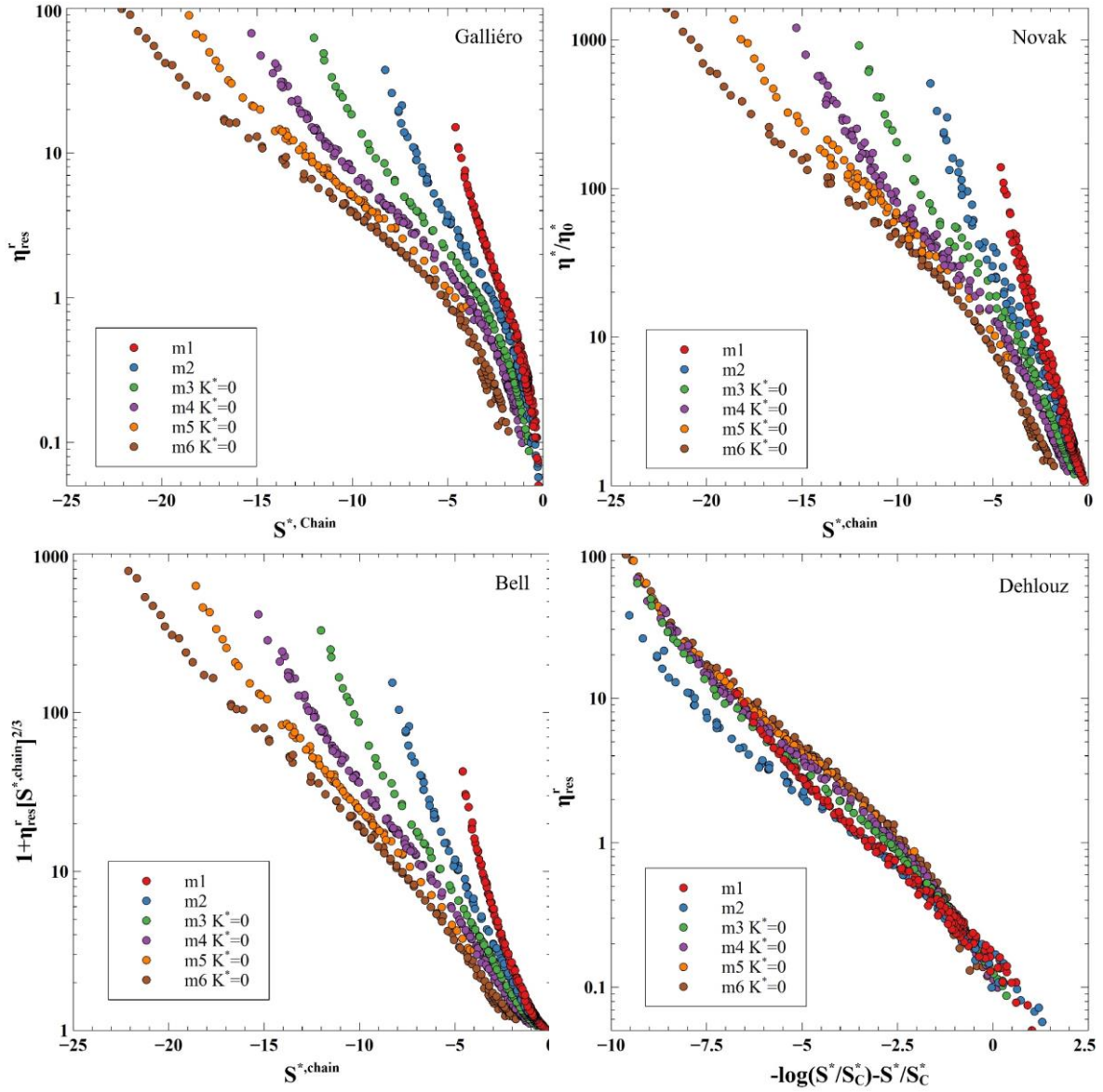


Figure 5.1: Test of different scaling approaches on the fully flexible LJC fluid model.

From these results, it appears that there is a need to develop a new scaling to deal with the additional rigidity parameter and that the residual entropy is not sufficient enough to capture the influence of the rigidity on the viscosity. Since the data are found mostly scattered in the diluted systems, this likely come from the choice of the quantity put on the abscissa. This is because, regarding ordinate, both total and residual viscosities used in the different approaches led to similar results. However, in the case of the abscissa, all the approaches used the total residual (or configurational) entropy. This latter is the total entropy of the system from which is subtracted out the ideal gas entropy contribution.

Thus, the residual entropy, as defined previously, still contains contributions from intramolecular interactions, such as LJ interactions between segments of the same molecule and interaction due to the bending potential. Hence, instead of using the correct residual entropy, we used so far only the part of the entropy contribution due to intermolecular interactions. By making the assumption that the intramolecular contribution is density-independent, it is possible within the SR-SAFTD scheme to calculate such contribution using this relation:

$$S^*_{intra} = S^*(T^*, \rho^* = 0, K^*, m) = \frac{U^*(T^*,\rho^*=0,K^*,m) - A^*(T^*,\rho^*=0,K^*,m)}{T^*} \quad (5.10)$$

and the residual intermolecular entropy contribution is then obtained from the following:

$$S^*_{inter} = S^*_{res} - S^*_{intra} \quad (5.11)$$

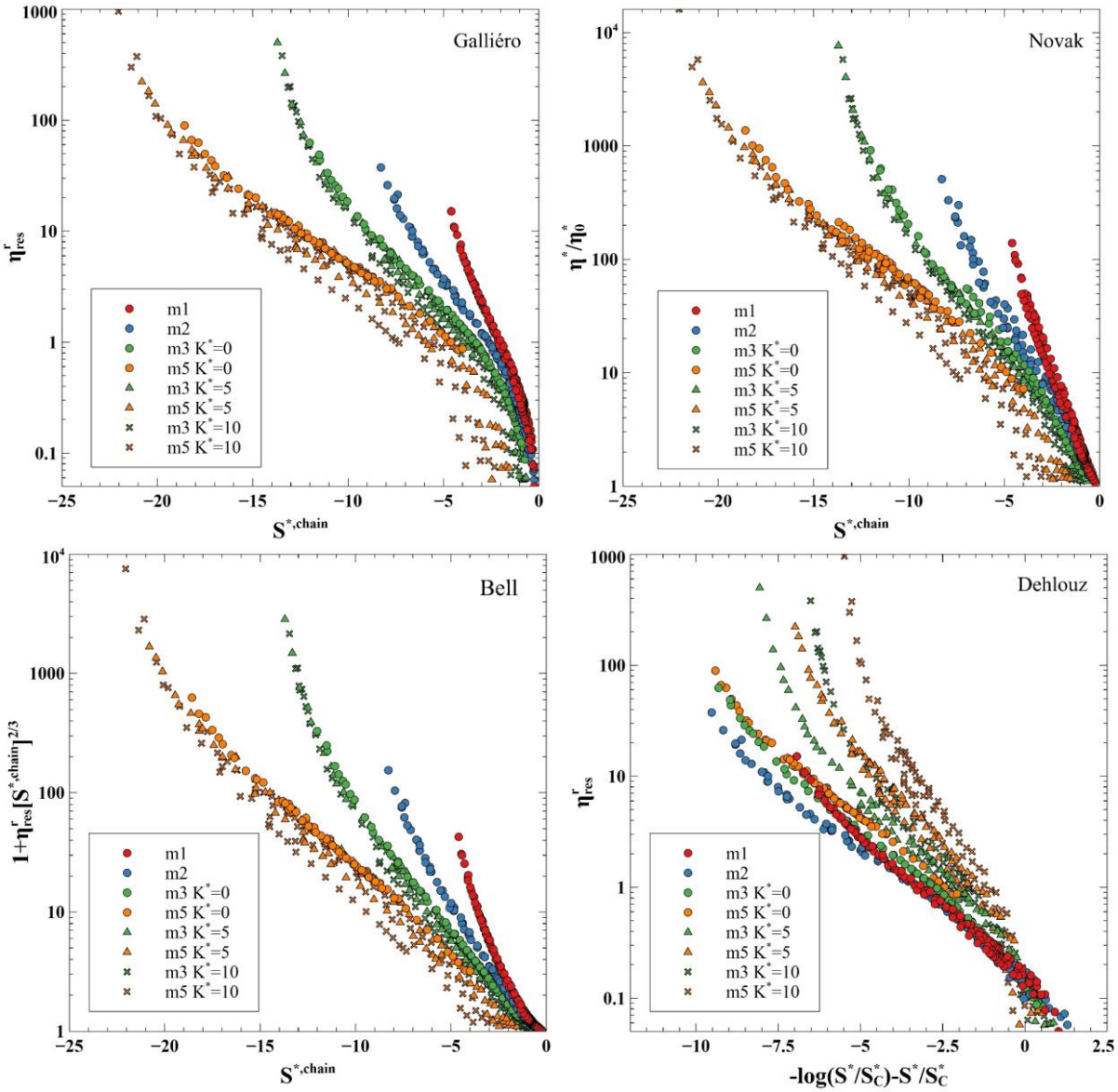


Figure 5.2: Test of the different approaches on the Semi-Rigid LJC fluid model.

Note that, in what follows, S^*_{intra} will be referred to as $S_0^{*,chain}$ as in the figures. Replacing the residual entropy by this quantity "S^*_{inter}" in the four scaling approaches considered in this work gives the results shown in Figure 5.3. Very interesting improvement are obtained compared to the previous case for the semi-rigid chains. The approaches of Galliéro and Bell, exhibit similar trends, where for the same chain length, the curves for all the rigidities are almost stacked on one curve, showing a corresponding state like behaviour. However, when the rigidity increases, particularly in the case of K^*=10, the curves are smoothly separated from each other in the Arrhenius region (at the intermediate densities where the curves exhibit a linear dependence between the ordinate and the abscissa in a semi-log scale), and more pronounced at very high densities leading to a super-Arrhenius (the dependence is larger than exponential) like behaviour, especially in the case of the Galliéro's approach. The notion of Arrhenius and super-Arrhenius regimes were introduced by Bell [9] and has no direct relationship with the conventional definition of the Arrhenius law for viscosity [15, 16].

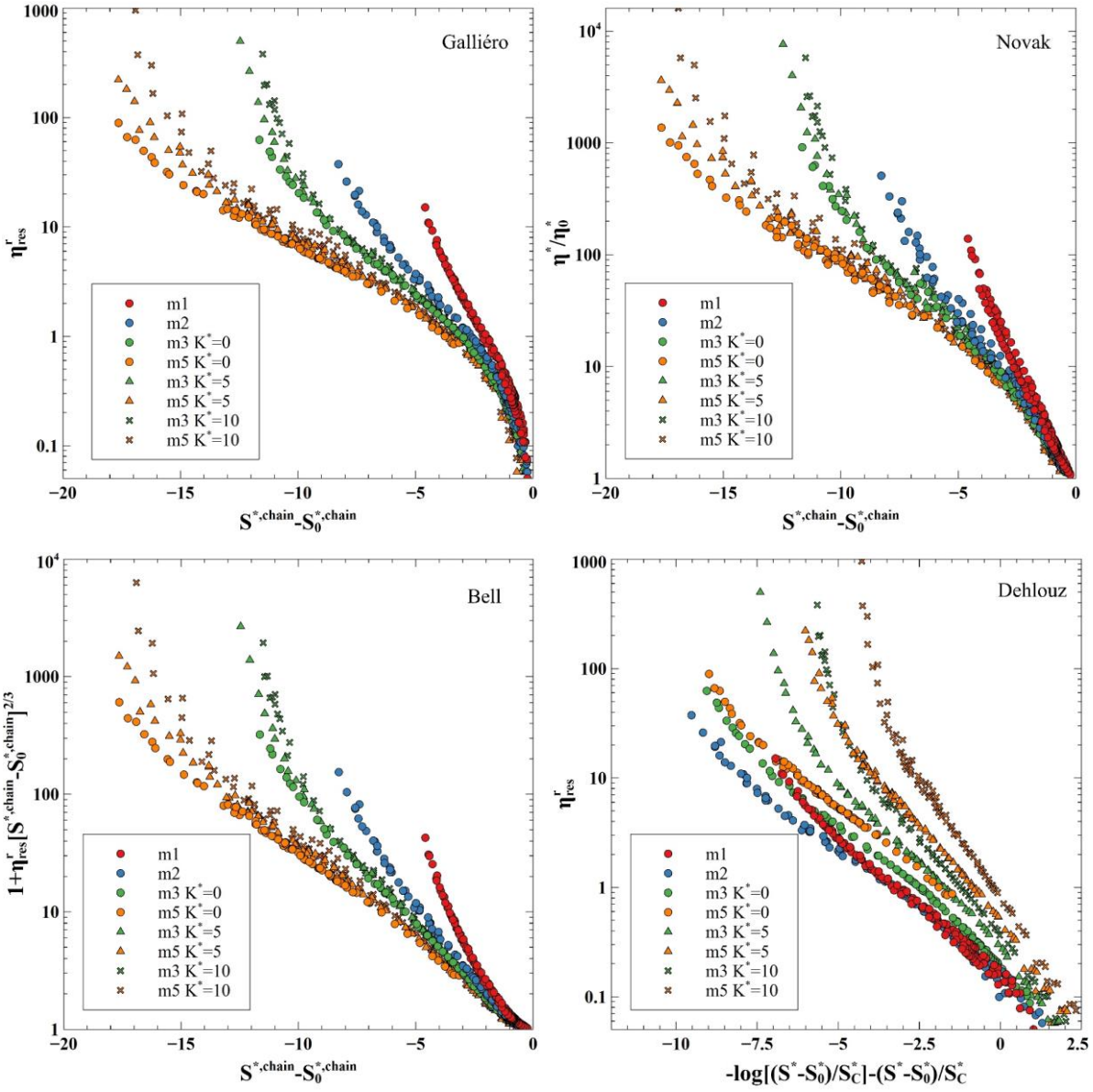


Figure 5.3: Testing the different scaling approaches using the new variable suggested in this work for the abscissa.

Interestingly, in the case of the fully flexible chains, this super-Arrhenius regime is observed only at very high density and as the chain length increases. In the case of the semi-rigid chains, this super-Arrhenius regime is found to be more pronounced and to appear earlier in terms of density and chain length. This departure of the curves of the semi-rigid chains from the fully flexible chains is interesting to analyse from the approach of Galliéro, since the trends exhibited by the most rigid chains in the high-density region tends to varies in the same way than what is observed in the very low density region. According to Bell et al. [17], the change from the Arrhenius to the super-Arrhenius regime in the dense states is occurring at half the distance between the entropy at the critical and triple points. Such an observation was only done for simple molecules, and results obtained in this work highlight that this observation should probably be not extrapolated to more complex molecules. This is because, even though not quantified, the change in regimes seems to evolve rapidly for the same chain length with a small increase of the rigidity.

The reason of the divergence in the dilute case has already been discussed and can be qualified as a "numerical divergence" (due to the density in the denominator). However, the divergence observed at high densities is "physical". In fact, as already pointed out by Galindo et al. [18], rigid chains exhibit a particular behaviour where the distance between the triple point and the critical point is observed to decrease drastically with increasing the rigidity of the molecule. An effect which is more pronounced when the chain length increases. It is, in fact, the triple point temperature which increases dramatically with the rigidity contrary to the critical point which is only very slightly affected. Moreover, in this work of Galindo et al. [18], it was shown that the increase in the triple point temperature affects also the liquid-solid transition location, which is observed to be shifted to the liquid side even until destabilization of the liquid-vapour phase diagram at high rigidities for chains of more than six segments. Thus, from the results of Galindo et al. one would expect that it is because of the displacement of the liquid-solid transition region with increasing the rigidity that there is this dramatic increase in the viscosity for the same temperature and density compared to the fully flexible case. Thus, the rate at which the viscosity increases is more important than that of the exponential of the residual entropy. However, these results, based on the entropy deduced from the SR-SAFTD EoS, will have to be confirmed by comparing them to those obtained using thermodynamic integration for the entropy. Checking the performance of the SR-SAFTD in capturing this transition effect with the rigidity would also be interesting, since the soft-SAFT tested in the work of Galindo et al. was found to be unable to capture all changes in the behaviour of the fluid observed from the molecular simulations.

Similar conclusion can be drawn from the approach of Novak, but, regarding the approach of Dehlouz, an interesting improvement is obtained. The modification brought by this approach, induced an interesting behaviour for all curves. The scaling showed less scattered data compared to all other approaches, particularly in the super-Arrhenius regime where the other approaches leads to less regular trends when increasing the rigidity. The transition from the Arrhenius regime to the super-Arrhenius regime is also well highlighted by this approach, where a clear distinction between the two regimes is observed. Indeed, it seems possible to identify the location of the transition which would correspond to the intersection of the two slopes fitted on each regime. In order to bring more light on this point, we have considered the

case of the LJ chains constituted of three tangentially bonded segments. We have considered variable rigidities from the fully flexible case K^*=0 to the case of K^*=50. Results are shown in Figure 5.4 (a) where have been plotted different entropy scaling for seven distinct rigidities. The transition is then deduced from the intersection of the slopes as shown for the case of K^*=50. Doing so for all cases, we have obtained the abscissa intersection named "$X_{transition}$". Values of this $X_{transition}$ with respect to the rigidity are reported in Figure 5.4 (b). The data are showing a clear trend, where there is a rapid variation of $X_{transition}$ for small rigidities reaching a plateau at high rigidities.

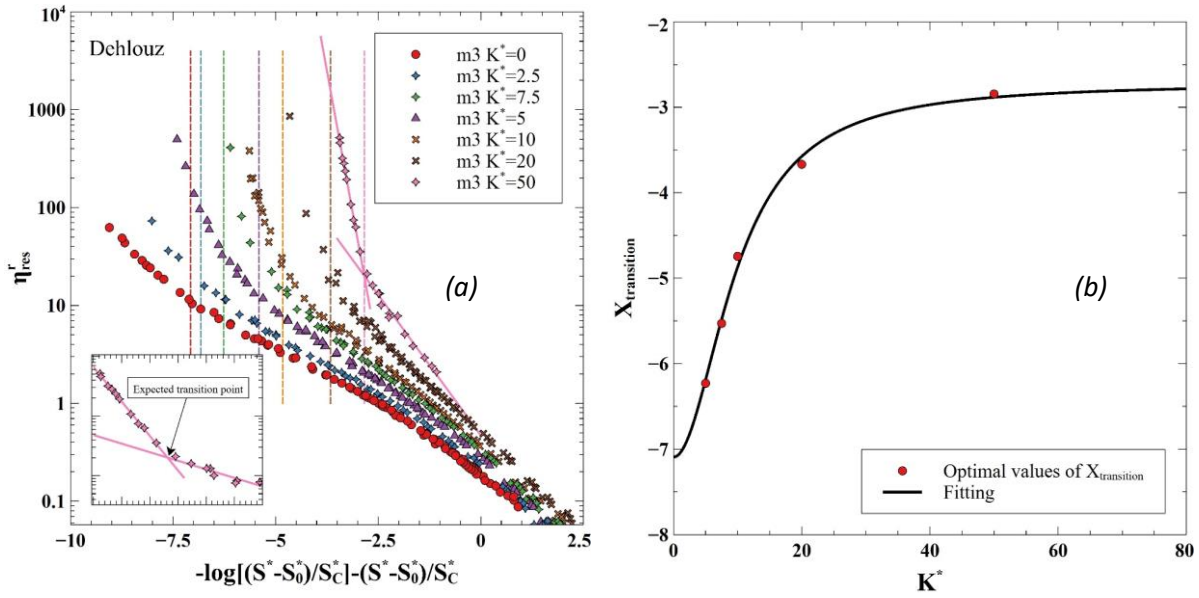


Figure 5.4: Arrhenius and super-Arrhenius transition analysis using the modified Dehlouz approach. (a) Entropy scaling for LJC3 at variable rigidities (the dashed lines represent the transition limits), (b) represents the evolution of the transition abscissa with respect to the rigidity.

Another interesting feature exhibited by this approach is that the transition from the gas region to the liquid region is also well highlighted as shown in Figure 5.5. Using the tangent method proposed, it is possible to determine the abscissa corresponding to this transition. However, contrary to the previous case considered, here the transition is more pronounced for the weakly rigid chains compared to the highly rigid chains. The case of highly rigid chains could probably be obtained by extrapolating the model fitted on the values obtained at low rigidities.

There remains an interesting point to discuss. The chains with the same rigidities are almost parallel to each other, which would be extremely useful for the development of a corresponding state like scaling. One could then rescale the ordinates by their respective segments number. Moreover, as for the Dehlouz approach, it has been decided to rescale the abscissa of the other approaches by the value of the entropy at the critical point. Doing so, as shown in Figure 5.6, leads to many interesting conclusions.

The new scaling proposed using the Novak approach leads to scattered data, and thus, the scaling is unusable. Another point is that, in the case of the Dehlouz approach, improvement is

remarkable compared to the previous case. Here, curves are clearly separated into a set of curves having the same rigidity similarly to what was obtained previously with the three other approaches. Regarding the approaches of Galliéro and Bell, the separation between curves with respect to the rigidities is still valid with the new variables introduced. Still, the same problem regarding the departure of the curves for the semi-rigid chains is observed in the Arrhenius regime and particularly in the super-Arrhenius regime. Further improvement is needed if one would obtain a corresponding state over each rigidity.

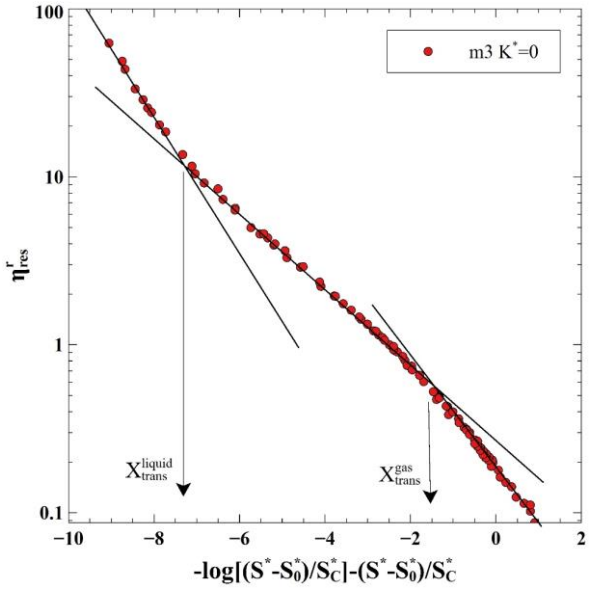


Figure 5.5: Entropy scaling of fully flexible LJC3 showing the possibility of deducing the gas-liquid transition similarly to the transition Arrhenius-Super-Arrhenius.

Contrary to the semi-rigid chains, in the case of the fully flexible chains the curves are exhibiting an extremely interesting behaviour shared by the approaches of Galliéro, Dehlouz and Bell. The introduction of the new variables leads to a master curve, except for the monomer. The separation of the monomer from the rest of the chains can be attributed to the lack of internal degrees of freedom. In fact, while the monomer, is characterized by only translational degrees of freedom, the other molecules includes in addition other internal degrees of freedom such as those related to the rotation of the molecule and that associated to the bonding. Thus, the scaling for the fully flexible chains, could probably be expressed as a contribution of two parts, a translational contribution given by the monomer, and a residual contribution due to the additional internal degrees of freedom, such as:

$$\chi_{total} = \frac{\eta_{res}^{red}}{m} = (m-1)\chi_{residual}^{chains} + m\chi_{translational}^{monomer} \tag{5.12}$$

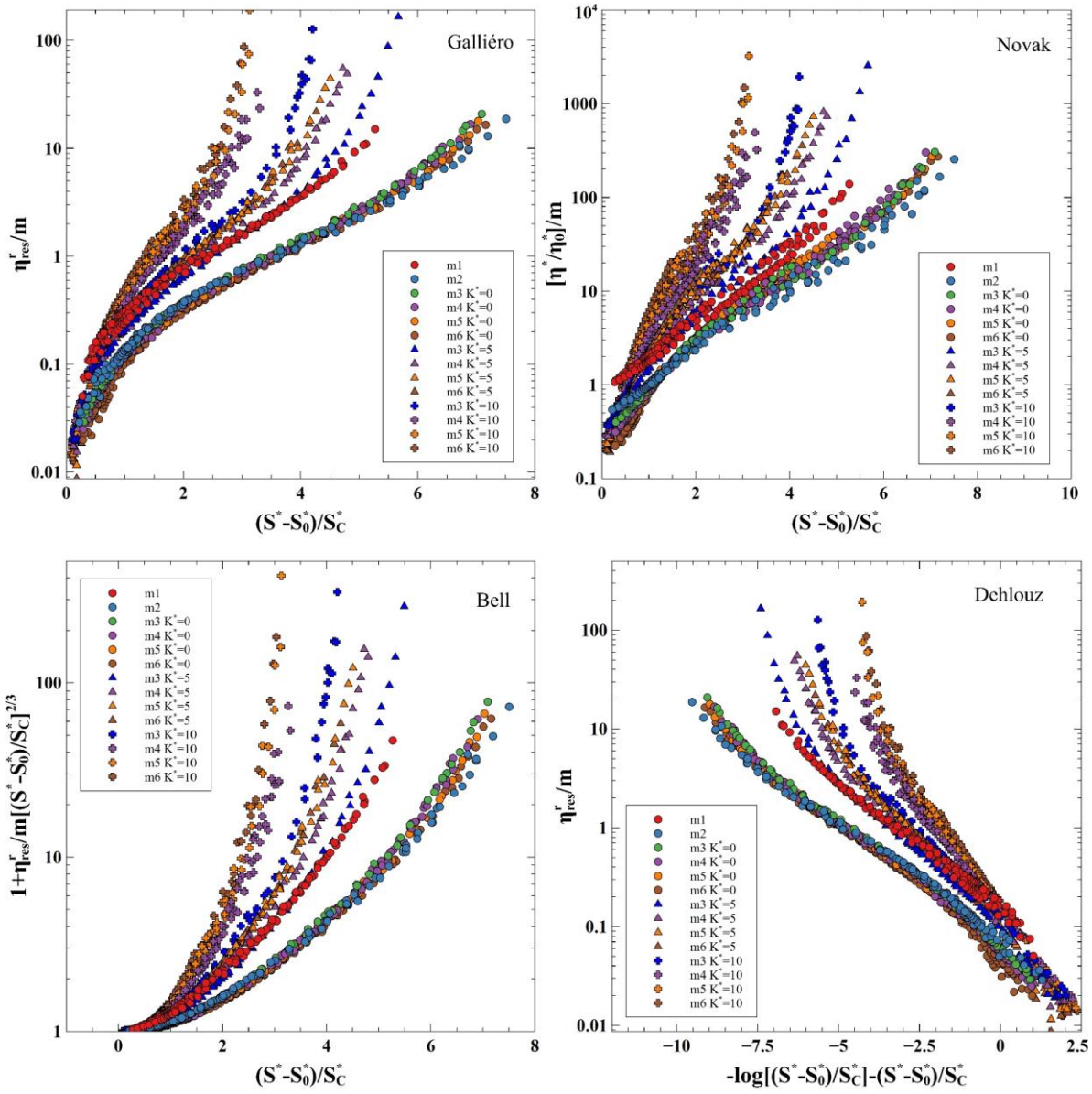


Figure 5.6: Testing different new approaches by dividing the viscosity in the ordinate by the segment number of the chains and using of the reduced quantity in the abscissa for the Galliéro, Novak and Bell approaches

To better illustrate this particular case, and to emphasize the needs of having an equation of state capable of estimating accurately residual entropy, results for the fully flexible chains are plotted alone in Figure 5.7 for the approaches of Galliéro and Dehlouz,. In addition, we have tested whether it is possible to obtain the same result using the soft-SAFT EoS. This latter case is shown in Figure 5.8 for the two approaches of Galliéro and Dehlouz.

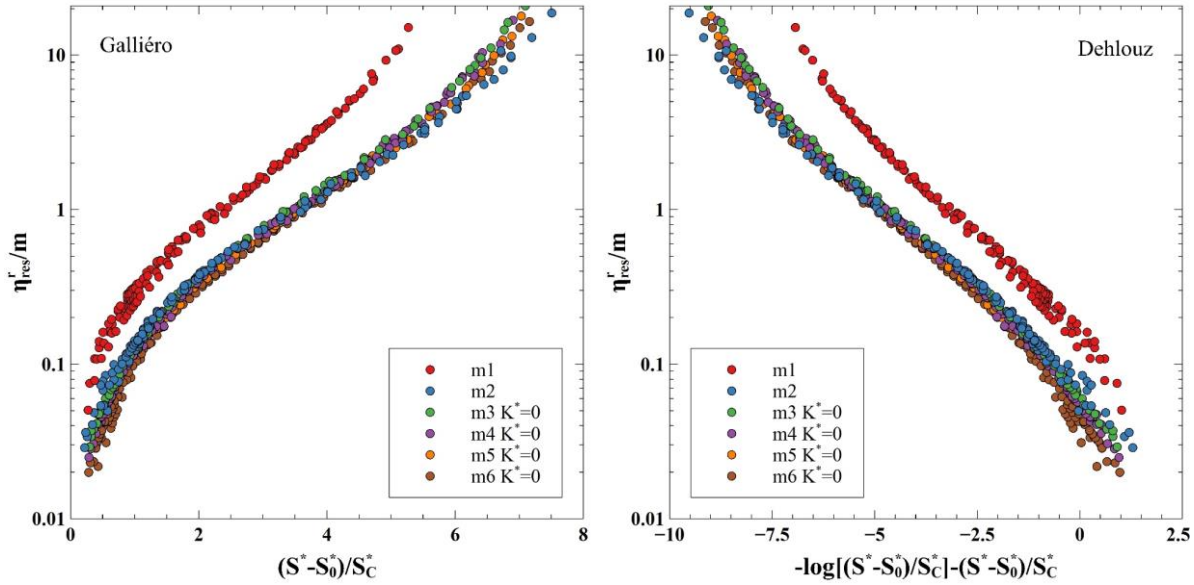


Figure 5.7:Corresponding state entropy scaling approach for the fully flexible LJ chains.

Results show clearly that no corresponding state is obtained when using a soft-SAFT EoS. Note that, in the case of the soft-SAFT, the intramolecular contribution is not taken into account in the model, and thus, the intramolecular entropy contribution is $S^{*,chain}_{intra} = 0$. This result highlights again the benefit of having a consistent equation of state. This weakness is not specific to the soft-SAFT, but, is present for all SAFT models based on a monomer representation (Original-SAFT, CK-SAFT, SAFT-VR, PC-SAFT, SAFT-VR-Mie…).

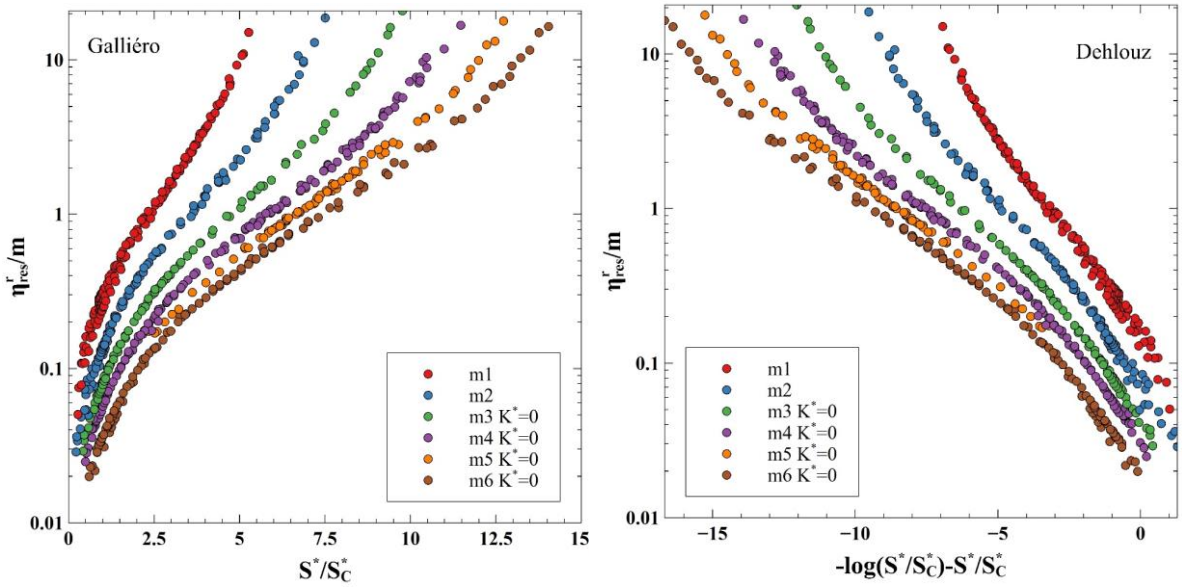


Figure 5.8: Performance of the soft-SAFT EoS in predicting a corresponding state entropy scaling for the fully flexible LJC fluid model.

While a corresponding state can be obtained for simple flexible LJC molecules, it is expected that the case of the semi-rigid chains could also be obtained with additional efforts.

5.3 Application of the entropy scaling for the real fluids

We have not yet developed a generic approach treating both fully flexible and semi-rigid chains across the entire thermodynamic space. However, to illustrate the goal of this work, we have developed an entropy scaling for the fully flexible and semi-rigid chains in the liquid phase. Please note that the correlation is not valid outside the fitting range.

We have chosen the classical approach of Galliéro for which the Arrhenius regime is simple to fit with an exponential function. However, the approach is modified by introducing a new variable on the abscissa ($S_{res}^{*,chain} - S_{intra}^{*,chain}$). In order to be applied to real fluids, a correlation for the zero-density viscosity is required. The case of the fully flexible chains has already been published by Galliéro et al. [19], and revisited by Delage-Santacreu et al. [20] later providing a more accurate correlations. In the latter work, a correlation for the fully rigid chains is also provided. However, a correlation for the case of chains with variable rigidities is not yet available. Thus, we have developed a correlation to treat this case by generalizing the correlation of Delage-Santacreu et al. For this, the translational viscosity for semi-rigid chains of lengths m=1 to 8, have been obtained from molecular dynamic simulations at a density of ρ^*=0.1 and temperatures T^*=2.5 to 6. The new correlation writes:

$$\begin{cases} \eta_{0,LJC}^{*} = \dfrac{\eta_{0,LJC2}^{*}}{\left(\frac{m}{2}\right)^{\alpha(K^*)}} & m \geq 2 \\ \eta_{0,LJC}^{*} = \left(-\dfrac{1}{3}\eta_{0}^{*} + \dfrac{4}{3}\right) & m \leq 2 \end{cases} \tag{5.13}$$

where $\eta_{0,LJC2}^{*}$ is the kinetic viscosity of the dimer, η_0^* is the kinetic viscosity of the monomer and α is a fitting parameter which is function of the rigidity given by:

$$\alpha(K^*) = \frac{ab + cK^{*d}}{b + K^{*d}} \tag{5.14}$$

a, b, c and d are fitting parameters given in Table 5.1.

a	b	c	d
0.5691	7.9860	0.8467	0.8196

Table 5.1: Parameters of the α(K) function*

The dimer zero-density viscosity is given by:

$$\eta_{0,LJC2}^{*} = \frac{2}{3}\eta_{0}^{*} = \frac{5}{24\Omega^{(2,2)}}\sqrt{\frac{T^*}{\pi}} \tag{5.15}$$

where η_0^* is the monomer zero-density viscosity.

The fitted parameter $\alpha(K)$ is given in Figure 5.9 (b), and the prediction of the viscosity at the zero-density limit and at T^*=4 for the semi-rigid chains are plotted in Figure 5.1 (a). Since, the SR-SAFTD allows using non-integer segment number, the case where this number is between m=1 and m=2 is obtained by a simple linear interpolation.

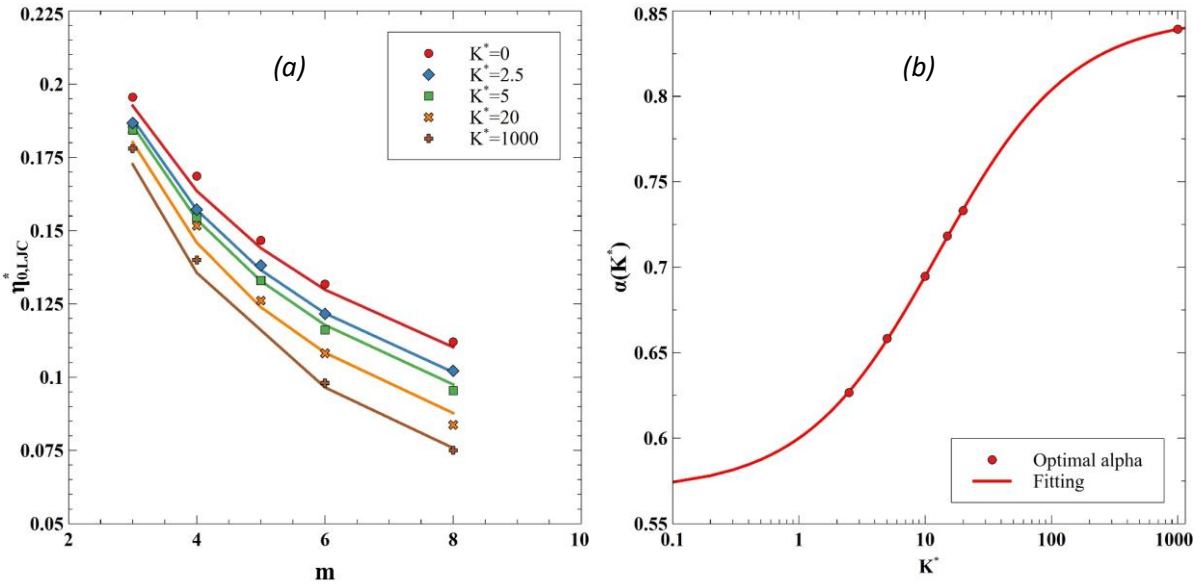


Figure 5.9: (a) prediction of the viscosity of the semi-rigid LJC at zero-density limit and T=4, (b) The fitting parameter Alpha function of rigidity.*

Regarding the correlation developed for the entropy scaling, it is given by the following expression:

$$\eta^*_{LJC}(\mathrm{m}, K^*) = A_0 Exp\big(B_0(S_{res} - S_0)\big)F + A_1 Exp\big(B_1(S_{res} - S_0)\big)(1 - F) \tag{5.16}$$

where F is a switching function dependent of K^*, given by:

$$F(K^*) = 2/(1 + Exp(10K^*)) \tag{5.17}$$

and A_0, B_0, A_1, B_1 are function of m and K^*, given by the following equations:

$$A_0(\mathrm{m}) = c_0 m^{c_1} \tag{5.18}$$

$$B_0(\mathrm{m}) = d_0 m^{d_1} \tag{5.19}$$

$$A_1(\mathrm{m}, K^*) = e_0(K) m^{e_1(K^*)} \tag{5.20}$$

$$B_1(\mathrm{m}, K^*) = f_0(K) m^{f_1(K^*)} \tag{5.21}$$

where

$$e_0(K^*) = g_0/(1 + g_1 Exp(-g_2 K^*) \tag{5.22}$$

$$e_1(K^*) = 1/(h_0 + h_1 K^* + h_2 K^{*2}) \tag{5.23}$$

$$f_0(K^*) = l_0/(1 + l_1 Exp(-l_2 K^*) \tag{5.24}$$

$$f_1(K^*) = n_0(K^* - n_1)^{n_2} \tag{5.25}$$

c_0	c_1	g_0	g_1	g_2
0.1242	0.6639	0.1249	-0.04269	0.2476
d_0	d_1	h_0	h_1	h_2
-0.9899	-0.7616	1.7107	-0.01337	0.0005423
		l_0	l_1	l_2
		-0.8072	-0.2784	0.1774
		n_0	n_1	n_2
		-1.2507	-3.7941	-0.3029

Table 5.2: Numerical fitting parameters used in the correlation for the entropy scaling.

We reiterate it again, this correlation is valid in the Arrhenius regime only, where the reduced residual viscosity scale linearly with the exponential of the residual entropy for which we have subtracted out the intramolecular entropy contribution.

We have considered chains from the monomer to the hexamer, and rigidities from K^*=0 to K^*=20. Results for this correlation are given in Figure 5.10 for two cases K^*=0 and K^*=5, in which this latter case was not included in the fit. Clearly, better predictions are obtained for the fully flexible chains, for which the range of validity of the correlation is extended compared to the semi-rigid chains. This is due to the fact that the super-Arrhenius regime is only slightly noticeable for the fully flexible chains, while it is more pronounced for the semi-rigid chains, limiting then the range of applicability of the correlation. The gas region is also not captured by the correlation as the scaling is incorrect at low densities.

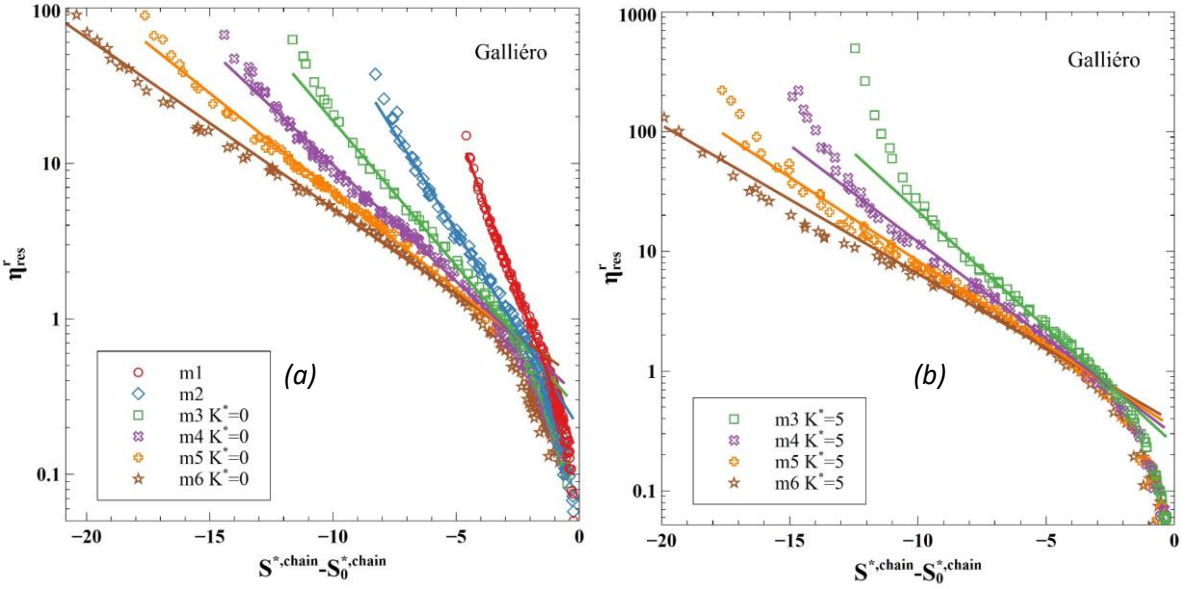


Figure 5.10: Test of the correlation (solide lines) for the entropy scaling against the data used in the fitting (symbols). (a) fully flexible chains K=0, (b) semi-rigid chains K=5.

The entropy scaling correlation developed with the SR-SAFTD EoS has finally been tested on real fluids as shown in Figure 5.11. In this figure we have considered some n-alkanes from the methane to the n-octane. A systematic comparison between the fully flexible model and the semi-rigid model is made for all the molecules with segment number m>2 and for both approaches, entropy scaling and direct molecular simulations.

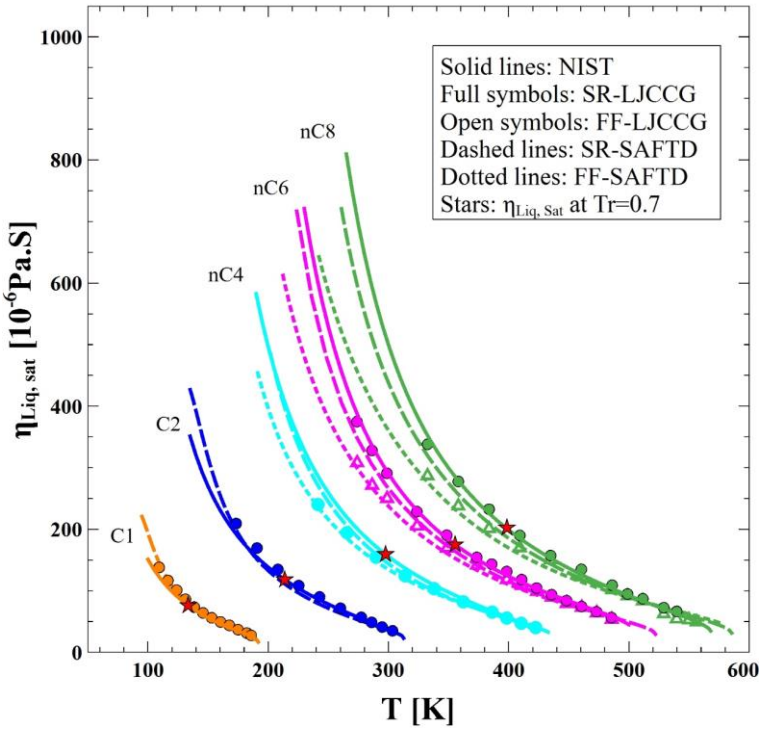


Figure 5.11: Saturated viscosity prediction of some n-alkanes with different approaches, direct molecular simulations and entropy scaling using the SR-SAFTD. The FF-SAFTD is the SR-SAFTD with K=0. Comparison is made between the two approaches and to the correlated experimental data taken from the NIST database.*

Results obtained with the entropy scaling for the viscosity are in agreement with those of molecular simulations for the same molecular fluid. This highlights the power of the entropy scaling approach, as it provides data with the same accuracy as those obtained from molecular simulations, but in only few seconds while molecular simulations may take hours or days.

The comparison between the results obtained with the scaling and the experiments highlights that the integration of the rigidity parameter in the SR-SAFT model is crucial to obtain accurate viscosity prediction. Very noticeable improvements are obtained with the semi-rigid model compared to the fully flexible model. However, small deviations are noticed for the semi-rigid model at low temperatures, which is probably due to the imprecision of the correlation, as at those conditions the density is high and may corresponds to the super-Arrhenius regime. It is worthwhile to mention that, the same parameters used for the prediction of the thermodynamic properties are used in the entropy scaling for the viscosity prediction. This also highlights that the SR-SAFTD developed in this work is able to provide simultaneously accurate thermodynamic and transport properties.

Although the work is not yet finalized, the preliminary results showed that the approach is very promising. In future works, we will try to extend this correlation to large chain lengths and to consider a larger ranger of thermodynamic conditions. Regarding this last point, it appeared that the residual entropy did not allow for completely integrating the effects of internal degrees of freedom on viscosity, especially at high densities (super-Arrhenius behaviour), when using the available entropy scaling approaches. However, results obtained with some modifications

brought in this work on these approaches, especially on that of Dehlouz, showed that a corresponding state entropy scaling could probably be obtained for the semi-rigid chains, similarly to what was obtained for the fully flexible chains, by considering further modifications.

5.4 Summary

In this chapter, we have revisited the entropy scaling approaches proposed in the literature. We first made a benchmark between all the approaches for the fully flexible fluid model. We tested the approaches on the semi-rigid fluid model, and results showed that none of the approaches were presenting a good scaling, indicating that the residual entropy is not sufficient to capture the influence of the rigidity of the viscosity. A modification of the variable in the abscissa was then proposed, by subtracting out the intramolecular entropy that contributes to the total residual entropy. Doing so, results showed that a better scaling can be achieved, extending then the entropy scaling approach to deal with semi-rigid chains. With this new variable, we showed that, except for the Dehlouz approach, the semi rigid chains and the fully flexible chains curves are superposed over a large range of thermodynamic conditions for the same chain length. However, in the super-Arrhenius regime, the curves are separated, and tends towards high viscosities as the rigidity increases.

We also showed that, even though the approach of Dehlouz could not be exploited directly, information on the transitions between the different regimes can be obtained. Such information may be of great usefulness for the construction of a corresponding state entropy scaling for the semi rigid chains. We showed that such a scaling could be written in terms of the internal degrees of freedom. In other words, this indicates that the entropy scaling is very sensitive to the internal degrees of freedom of the molecule, highlighting then the complexity exhibited by the semi-rigid chains.

Last, even though the full scaling was not yet obtained, in order to provide some insights on the goal of the approach developed in this work, we proposed an entropy scaling for the semi-rigid fluid model limited to the liquid (Arrhenius like) region. When applied in the fitting range, we showed that, this entropy scaling based on the SR-SAFTD is very promising as it was capable to yield accurately the liquid viscosity of some n-alkanes up to n-octane. It is worthwhile to underline that, the same molecular parameters used for the prediction of the thermodynamic properties are used for the prediction in the entropy scaling approach to predict the viscosity. It highlights the coherence of the approach proposed in this work.

5.5 References

1. Goel, T., Patra, C. N., Mukherjee, T. & Chakravarty, C. Excess entropy scaling of transport properties of Lennard-Jones chains. *J. Chem. Phys.* **129**, 164904 (2008).
2. Galliero, G., Boned, C. & Fernández, J. Scaling of the viscosity of the Lennard-Jones chain fluid model, argon, and some normal alkanes. *The Journal of chemical physics* **134**, 064505 (2011).
3. Galliero, G. Equilibrium, interfacial and transport properties of n-alkanes: Towards the simplest coarse grained molecular model. *Chemical Engineering Research and Design* **92**, 3031–3037 (2014).
4. Rosenfeld, Y. Relation between the transport coefficients and the internal entropy of simple systems. *Phys. Rev. A* **15**, 2545–2549 (1977).
5. Rosenfeld, Y. A quasi-universal scaling law for atomic transport in simple fluids. *Journal of Physics Condensed Matter* **11**, 5415–5427 (1999).
6. Novak, L. Self-Diffusion Coefficient and Viscosity in Fluids. *International Journal of Chemical Reactor Engineering* **9**, (2011).
7. Neufeld, P. D., Janzen, A. R. & Aziz, R. A. Empirical Equations to Calculate 16 of the Transport Collision Integrals Ω(l, s)* for the Lennard-Jones (12–6) Potential. *J. Chem. Phys.* **57**, 1100–1102 (1972).
8. Bell, I. H. Probing the link between residual entropy and viscosity of molecular fluids and model potentials. *Proceedings of the National Academy of Sciences* **116**, 4070–4079 (2019).
9. Bell, I. H. Entropy Scaling of Viscosity—I: A Case Study of Propane. *J. Chem. Eng. Data* **65**, 3203–3215 (2020).
10. Bell, I. H. Entropy Scaling of Viscosity—II: Predictive Scheme for Normal Alkanes. *J. Chem. Eng. Data* **65**, 5606–5616 (2020).
11. Nguyen, T. *et al.* Entropy Scaling for Viscosity of Pure Lennard-Jones Fluids and Their Binary Mixtures. *Communications in Physics* **32**, 187–187 (2022).
12. Dehlouz, A., Privat, R., Galliero, G., Bonnissel, M. & Jaubert, J.-N. Revisiting the Entropy-Scaling Concept for Shear-Viscosity Estimation from Cubic and SAFT Equations of State: Application to Pure Fluids in Gas, Liquid and Supercritical States. *Ind. Eng. Chem. Res.* **60**, 12719–12739 (2021).
13. Lötgering-Lin, O., Fischer, M., Hopp, M. & Gross, J. Pure Substance and Mixture Viscosities Based on Entropy Scaling and an Analytic Equation of State. *Ind. Eng. Chem. Res.* **57**, 4095–4114 (2018).
14. Yang, X., Xiao, X., May, E. F. & Bell, I. H. Entropy Scaling of Viscosity—III: Application to Refrigerants and Their Mixtures. *J. Chem. Eng. Data* **66**, 1385–1398 (2021).
15. Bair, S., Martinie, L. & Vergne, P. Classical EHL versus Quantitative EHL: A Perspective Part II - Super-Arrhenius Piezoviscosity, An Essential Component of Elastohydrodynamic Friction Missing from Classical EHL. *Tribology Letters* **63**, 37 (2016).
16. Messerly, R. A., Anderson, M. C., Razavi, S. M. & Elliott, J. R. Mie 16–6 force field predicts viscosity with faster-than-exponential pressure dependence for 2,2,4-trimethylhexane. *Fluid Phase Equilibria* **495**, 76–85 (2019).

17. Bell, I., Galliero, G., Santacreu, S. D. & Costigliola, L. An entropy scaling demarcation of gas- and liquid-like fluid behaviors. *Journal of Chemical Physics* **152**, 191102 (2020).
18. Galindo, A. *et al.* Computer simulation study of the global phase behavior of linear rigid Lennard-Jones chain molecules: Comparison with flexible models. *J. Chem. Phys.* **120**, 3957–3968 (2004).
19. Galliero, G. & Boned, C. Shear viscosity of the Lennard-Jones chain fluid in its gaseous, supercritical, and liquid states. *Physical Review E : Statistical, Nonlinear, and Soft Matter Physics* **79**, 1–9 (2009).
20. Delage Santacreu, S., Galliero, G., Odunlami, M. & Boned, C. Low density shear viscosity of Lennard-Jones chains of variable rigidities. *Journal of Chemical Physics* **137**, 204306 (2012).

Chapter 6. Conclusion and perspectives

Contents

6.1 Conclusion 189
6.2 Perspectives 191
6.3 References 194

6.1 Conclusion

The objective of this work was to present a new coarse grained molecular model capable to simultaneously predict accurate equilibrium and transport properties. The objective was also to develop a consistent and coherent global approach that links molecular simulation and equation of state, specifically for new coarse grained molecular model.

In the third chapter, we developed a new force field referred to as the Semi-Rigid Lennard-Jones Chains Coarse Grained (SR-LJCCG) molecular model, which is simply based on a Lennard Jones chain including a rigidity parameter of the bending type. This force field was parametrized using a top-down approach, utilizing a corresponding state strategy, which incorporates viscosity as a target property to constrain the molecular parameters (segment number and stiffness constant of the bending potential). This method only requires four accessible macroscopic properties and is easy to use. The optimization of the force field typically converges to the global minimum. This strategy presents many advantages compared to the classical top-down approach where many experimental data are needed, in addition to the difficulty to converge to the global minimum leading to sometimes inconsistent parameters.

The SR-LCCG molecular model was tested against experimental data over a wide range of thermodynamic conditions and chemical families. Many properties have been considered including those used in the fit such as viscosity and phase equilibria, but, also, other properties not included in the optimization such as derivative properties and self-diffusion coefficient. The model was found to provide excellent predictions. It was also compared to the fully-flexible LJCCG model, with the SR-LJCCG showing comparable results on thermodynamic properties and improved results on the liquid viscosity and the structural properties. The SR-LJCCG was also benchmarked against other force fields proposed in the literature, including high resolution models (L-OPLS-AA) and united atoms (TraPPE-UA force field), and found to provide the best prediction on nC_{12} and nC_{16} viscosities at saturation conditions.

To avoid the need for extensive molecular simulations, the idea was then to develop an equation of state for the semi-rigid LJCCG model. However, for the time being, a SAFT like EoS including an internal rigidity term of the bending type is not available. Thus, in the fourth chapter, we first revisited the thermodynamic perturbation theory and pointing out some strengths and weaknesses of the theory. We deeply discussed some approximations made in the theory and illustrating how these approximations made in the outset of the theory affect the predictions using the SAFT models. We found that the TPT1-Dimer was one of the available SAFT EoS that gives the closest predictions to the exact molecular simulations data of the fluid model. In addition, we determined that TPT1-D offers the ability to incorporate a bending or torsional potentials. Therefore, in a second step, a modification in the TPT1-D theory was proposed to include a bending potential. More precisely, thanks to molecular simulations, we proposed a scheme to calculate the radial distribution function that accounts for internal degrees of freedom, such as the bending potential. This new RDF in TPT1-D results in a more generalized TPT1-D referred to as SR-SAFTD equation of state. The SR-SAFTD has then been in a first step validated on molecular simulations. We showed that accurate predictions of the molecular simulations data were obtained over a wide range of thermodynamic conditions for all the properties considered and for chain lengths up to the 16-mer chain.

The new equation of state was then applied to real fluids. To do so, the molecular parameters optimized for SR-LJCCG force field were directly used, reinforcing the coherence between the two approaches. It is worth mentioning that, such a strategy was a novel aspect of this work, as it is based on the properties of the model fluid evaluated by molecular simulations, and it is for the first time that a transport property is incorporated in the parametrization of a SAFT model. The SR-SAFTD has then been tested on n-alkanes, iso-alkanes and some polar components over a wide range of thermodynamic conditions (saturation conditions, dense liquid and supercritical conditions). We considered various properties such as phase equilibria, enthalpy, single phase density and many derivatives properties. In all cases, the SR-SAFTD was found to accurately reproduce the data of the SR-LJCCG model and was also in very good agreement with experimental data. Furthermore, we demonstrated that the parameters used in the SR-SAFTD are transferable within the same molecular family.

The success of the SR-SAFTD model allows now to avoid performing extensive molecular simulations of the SR-LJCCG force field model to obtain thermodynamic properties of real fluids. However, the case of transport properties still requires performing molecular simulations as the such properties are not accessible via an equation of state. Thus, it was proposed in the fifth chapter to extend the predictive capability of the SR-SAFTD model to deal with transport properties thanks to the entropy scaling approach. In the first part of this chapter, we revisited some scaling approaches proposed in the literature and found that none of them was capable to yield a good scaling between the viscosity and the entropy of the semi-rigid chains. We found that, the semi-rigid model exhibit more complex behaviour, where the super-Arrhenius regime is strongly pronounced as the rigidity increases. In the case of the fully-flexible LJC fluid model, we proposed a robust corresponding state entropy scaling taking into account the internal degrees of freedom. Such a scaling is a demonstration of the strong sensitivity of the entropy scaling approach to the internal degrees of freedom of the molecule, highlighting then the complexity exhibited by the semi-rigid chains. Extending this corresponding state to the semi-rigid chains was not achieved in this work and would require more efforts.

Last, even though the full scaling was not yet obtained, in order to provide some insights on the goal of the approach developed in this work, we proposed an entropy scaling for the semi-rigid fluid model limited to the liquid (Arrhenius like) region. When applied in the fitting range, we showed that, this entropy scaling based on the SR-SAFTD is very promising as it was capable to yield accurately the liquid viscosity of some n-alkanes up to n-octane. Furthermore, and this is one of the most interesting features highlighted in this work, the same molecular parameters used for the prediction of the thermodynamic properties, were used in the entropy scaling without any additional re-parameterization, to predict the viscosity. This is highlighting once again the consistence and the coherence of the global approach proposed in this work.

6.2 Perspectives

Finally, this work sheds light on some perspectives which are provided below:

- It is important to mention that that the proposed coarse grained model will likely not perform well for strongly associating molecules such as water or alcohols. This is due to the fact that these molecules present highly directional hydrogen bond interactions. An example for water and methanol molecules are illustrated in the appendix 7.B.1. Such a limitation could probably be overcome by including association sites in the model, similar to what has been proposed in the SAFT EoS approach. Another option would be to include surface tension in the parameterization, as this property is, first, very sensitive to hydrogen bonds effect, and second, is compatible with the CS strategy proposed in this work, as shown in [1].

- The surface tension is one of the most important property that would be interesting to be evaluated with the SR-LJCCG model to confirm the representability of the molecular parameters. Further simulations are then needed for this property, however, a quick check can be made using the corresponding state for the surface tension developed by Galliéro [2]. Such a test is reported in the appendix 7.B.1 for some n-alkanes and polar components, and results showed very promising predictions can be achieved with the proposed force field.

- The application of the SR-LJCCG force field and the SR-SAFTD have not been extended to mixtures. Thus, it would be interesting to perform calculations with both approaches on mixtures to completely validate their capabilities. It is interesting to note that the use of a stiffness as a fourth parameter does not impose any additional combining rules when dealing with mixtures or systems involving different force fields. However, in this latter case, one should keep in mind that additional re-parametrization would be needed if this force field is combined with polarizable force fields.

- At very low temperature or at high pressures, the SR-LJCCG force field losses some of its performance (but, in relative, as it is still being better than many other force fields) when considering high molecular weights compounds. We expect that, such a weakness is related to the lack of other internal degrees of freedom of the molecule in the model, such as the torsional angle potential. Additional tests by considering a variable rigidity of the torsional potential would then be quite interesting to check how this could improve the predictions. In this regard, we expect that, similarly to what was done for the AUA(4m) [3], a re-parameterization of the torsional and/or the bending potentials of the TraPEE-UA force field to reproduce the viscosity, coupled to the re-parameterization proposed by Janecek and Paricaud [4] to improve the saturation pressure, could address its issue on transport properties.

- The strategy of the parametrization together with the pseudoization technique (coarse graining) can be combined to study complex mixtures such as petroleum fluids, where the petroleum fractions can be approximated by a pseudo-species. The CS strategy of the parametrization is flexible enough to use other properties than those used in this work. This is because in some cases, for instance petroleum fractions, the accessible thermophysical properties measured may be different from those used in this work (T_c, ω, μ, ρ).

- The main limitation of the proposed SR-SAFTD resides in the computational aspect when compared to other SAFT EoS. This is due to the large number of parameters used to fit the radial distribution function. Such a number of parameters is justified but the high non-linearity exhibited by the RDF when varying the rigidity parameter. This non-linearity is a consequence of the non-regular evolution with rigidity of the properties of the semi-rigid LJCCG fluid, such as what was observed on the critical properties and the saturated liquid density. However, when compared to molecular simulations the computational aspect is totally negligible and, by taking into account our recommendation in pre-calculating the matrix of the parameters before the main calculations allows to further reduce the computational time to be compatible with engineering requirements in a simulation software. A way to handle this issue would probably to limiting the range of rigidity, or to split this range. This last recommendation is because we found that the coefficients C_{ij} were monotonously decreasing in the rigidity range [0~5] and then a monotonous increase from [5 – ∞[. An illustration of the behaviour of these coefficients with rigidity is given in the appendix 7.A.2.

- An additional step towards the computational efficiency of the proposed SR-SAFTD, would consist in using an equation of state for the dimer reference fluid instead of using the TPT1-M of Johnson et al.[5]. In fact, the use of this latter requires 32 parameters for the monomer terms (modified-BWR EoS) and 25 other parameters for the $g^M(T^*, \rho^*, \sigma)$. Thus, for example, using an EoS of a Modified-BWR fitted on dimer properties, may avoid the need of g^M. Moreover, it is expected that a global improvement could also be achieved, particularly in the critical region. This is because the TPT1-M strongly overestimates the critical temperature, and thus, when used in TPT1-D, it has a high impact on the determination of the critical temperature.

- In its current version, the SR-SAFTD EoS does not ensure that it is free from some of the pitfalls observed in many other SAFT models due to their polynomial functions. Only a rigorous check following the recommendation analysis for the SAFT EoS of Polishuk et al. [6] may confirm it.

- As shown in the last chapter of this thesis, the SR-SAFTD coupled with the entropy scaling approach showed that very promising results could be achieved with this approach. However, some difficulties have been encountered in developing the scaling to cover a large thermodynamic conditions. In this regard, a new and simple scaling is needed to write a correlation which is function of the number of segments and the rigidity parameters. As done for the fully flexible model, the dimensionality of the problem may be reduced to only one parameter i.e. m, if a similar corresponding state entropy scaling approach could be developed.

- As soon as this entropy scaling for the semi-rigid chains will be obtained, it will be interesting to applied it also on mixtures. In our previous work of Quoc Viet et al. [7], we investigated the mass effect on viscosity of mixtures in the entropy scaling framework with application to Lennard-Jones mixtures. The study revealed that the choice of the mixing rule for the mass may have high impact. The case of LJ chains is not yet considered, thus, it will be interesting to extend such work with the SR-LJCCG model.

- There exist other approaches that might be used with the proposed model to predict transport properties. Among them, the thermodynamic scaling or sometimes referred to as density scaling. In our recent work [8], we showed that real fluids viscosity, such as n-alkanes, are exhibiting similar behaviour as the semi-rigid LJCCG model. Such an approach is then very promising. However, similarly to the entropy scaling, a limitation for the time being to build such an approach is related to the high non-linearity exhibited by the effective repulsive stiffness parameters with the rigidity and chain length.

- Despite the success of the proposed strategy for the parametrization of the SR-SAFTD, two other schemes based on the CS strategy would be interesting to investigate. The first, consists in a specific parametrization on the critical properties using the SR-SAFTD model, similarly to the work of Pàmies and Vega [9] using the soft-SAFT, where the idea is to develop a scaling parameters specially dedicated for the prediction of these properties. The idea is motivated by the fact that, some differences are observed between the soft-SAFTD and the SR-SAFTD, which then would lead to different predictions. Moreover, a theoretical model to predict such properties would be very suitable, as they are difficult to be measured experimentally, particularly for large chain molecules. such properties are used as input parameters in many models such as in cubic EoS to build the PVT tables that model petroleum fractions. A second strategy would be, instead of using the current strategy proposed, where the rigidity parameter K^* is constrained by the viscosity η, to use the ratio between the critical temperature T_c and that of the triple point T_t. This is also motivated by the fact that, it was already demonstrated from the work of Galindo et al. [10] that the rigidity largely affects the triple point location. The ratio T_c/T_t is so probably a good metric to determine K^*. Doing so, may lead to a simpler strategy in terms of experimental data to use, as large experimental database for the *Tt* is available for many fluids and also many correlations to estimate the T_t are proposed, compared to the viscosity. In addition, a probably better transferability may be achieved at very low temperatures. However, only a real test may confirm or infirm our expectations, particularly whether the viscosity will still be well represented, reason for which we recommend this test.

- Last, this strategy proposed in this work for the incorporation of the bending potential may be extended to other potentials related to the internal degrees of freedom, such as torsional potential. Moreover, this strategy is not specific to LJ-based EoS, but, can be transposed to other forms of intermolecular potentials, thus, used with any SAFT type model.

6.3 References

1. Galliero, G. Surface tension of short flexible Lennard-Jones chains: Corresponding states behavior. *The Journal of chemical physics* **133**, 074705 (2010).
2. Galliero, G. Surface tension of short flexible Lennard-Jones chains: Corresponding states behavior. *The Journal of Chemical Physics* **133**, 074705 (2010).
3. Nieto-Draghi, C., Ungerer, P. & Rousseau, B. Optimization of the anisotropic united atoms intermolecular potential for n-alkanes: Improvement of transport properties. *J. Chem. Phys.* **125**, 044517 (2006).
4. Janeček, J. & Paricaud, P. A new optimization method for the determination of classical force fields. Application to the united atom force field of short alkanes. *Fluid Phase Equilibria* **429**, 27–36 (2016).
5. Johnson, J. K., Mueller, E. A. & Gubbins, K. E. Equation of State for Lennard-Jones Chains. *J. Phys. Chem.* **98**, 6413–6419 (1994).
6. Polishuk, I., Privat, R. & Jaubert, J.-N. Novel Methodology for Analysis and Evaluation of SAFT-Type Equations of State. *Ind. Eng. Chem. Res.* **52**, 13875–13885 (2013).
7. Viet, T. Q. Q. *et al.* Mass effect on viscosity of mixtures in entropy scaling framework: Application to Lennard-Jones mixtures. *Fluid Phase Equilibria* **558**, 113459 (2022).
8. Delage Santacreu, S., Hoang, H., Khennache, S. & Galliero, G. Thermodynamic Scaling of the Shear Viscosity of Lennard-Jones Chains of Variable Rigidity. *Liquids* **1**, 96–108 (2021).
9. PAMIES, J. C. & VEGA, L. F. Critical properties of homopolymer fluids studied by a Lennard-Jones statistical associating fluid theory. *Molecular Physics* **100**, 2519–2529 (2002).
10. Galindo, A. *et al.* Computer simulation study of the global phase behavior of linear rigid Lennard-Jones chain molecules: Comparison with flexible models. *J. Chem. Phys.* **120**, 3957–3968 (2004).

Chapter 7. Appendices

Contents

Appendix 7.A **196**
7.A.1 Experimental data used for the parameterization. 196
7.A.2 Fitting parameters: fitting constants and optimized molecular parameters 197
Appendix 7.B..... **209**
7.B.1 Surface tension and associative molecules 209
7.B.2 References (Appendix A & B)..... 213

Appendix 7.A

7.A.1 Experimental data used for the parameterization.

7.A.1.1 Experimental data used in the optimization of the parameters of the FF-LJCCG, SR-LJCCG and SR-SAFTD models.

a. Normal alkanes

Molecule	Tc [K]	$\rho_{liq,\ sat\ Tr=0.7}$	w	η_{red}	Source
n-C36	891	614.049	1.288	6.447	DIPPR
n-C30	855.1	618.447	1.127	5.819	DIPPR
n-C24	809.3	620.445	0.943	4.991	DIPPR
n-C16	722	625.060	0.708	3.931	DIPPR
n-C12	658.1	619.770	0.574	3.378	NIST
n-C10	617.7	617.350	0.488	3.164	NIST
n-C8	569.32	612.450	0.395	2.868	NIST
n-C6	507.82	599.250	0.299	2.541	NIST
n-C4	425.13	573.480	0.201	2.438	NIST
C2	305.32	505.250	0.100	2.080	NIST
C1	190.56	388.380	0.011	1.823	NIST

Table 7.1: Experimental data of the normal alkanes used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models.

b. Iso-alkanes

Molecule	Tc [K]	$\rho_{liq,\ sat\ Tr=0.7}$	w	η_{red}	Source
i-C4	407.81	566.17	0.184	2.729	NIST
i-C5	460.35	589.6	0.227	2.594	NIST
i-C6	497.7	599.95	0.280	2.624	NIST
i-C8	544	615.13	0.303	3.117	NIST
i-C16	693	650.203	0.548	4.412	NIST

Table 7.2: Experimental data of the iso- alkanes used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models.

c. Some polar components

Molecule	Tc [K]	$\rho_{liq,\ sat\ Tr=0.7}$	w	η_{red}	Source
CO2	304.13	1186.18	0.224	2.832	NIST
Toluene	591.75	746.86	0.266	2.367	NIST
H2S	373.1	858.25	0.1005	1.979	NIST
RE245cb2	406.81	1305.6	0.354	3.384	NIST
R1336mzz(Z)	444.5	1330.2	0.386	3.179	NIST

Table 7.3: Experimental data of the some polar components used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models.

d. Associative molecules

Molecule	Tc [K]	$\rho_{liq,\ sat\ Tr=0.7}$	w	η_{red}	Source
Methanol	512.6	726.62	0.5625	2.7829664	NIST
Water	647.1	887.19	0.3443	1.14873826	NIST

Table 7.4: Experimental data of methanol and water molecules used for the optimization of the FF-LJCCG, SR-LJCCG and SR-SAFTD models.

7.A.2 Fitting parameters: fitting constants and optimized molecular parameters

7.A.2.1 Correlations used in the parametrization of the SR-LJCCG and the SR-SAFTD models.

a. Acentric factor

$$a_w = a_{w1}Asinh\big((a_{w2} + K)(K^2 + a_{w3})\big)$$

$$b_w = \Big(b_{w1} + Abs\big(b_{w2} + b_{w3}Abs(b_{w4} + K)\big)\Big)^{bw_5} - (b_{w6} + b_{w7}K)$$

$$c_w = c_{w1} + c_{w2}Asinh\Big(c_{w3} + Abs\big(c_{w4}(c_{w5} + K)\big)\Big)$$

$$w(m,K) = (0.1803m - 0.214)f_1 + \Big(a_w + b_wLn(m) + c_w\big(Ln(m)\big)^2\Big)f_2$$

$$f_1 = 1/\Big(1 + Exp\big(5(m-2)\big)\Big)$$

$$f_2 = 1/\Big(1 + Exp\big(-5(m-2)\big)\Big)$$

a_{wi}	b_{wi}	c_{wi}
0.01051	0.09735	0.16734
-1.65794	-0.60567	0.01948
-2.70019	0.56255	-3.38914
	-0.41008	-1.75709
	-0.06502	-0.43947
	0.96498	
	-0.00017	

Table 7.5: Fitting parameters of the acentric factor.

b. Critical temperature

$$a_{Tc} = atan(abs(a_{Tc1} + K) + a_{Tc2})/(a_{Tc3} + a_{Tc4}K)$$

$$b_{Tc} = b_{Tc1}/\big(b_{Tc2}K + b_{Tc3} + b_{Tc4} * abs(K + b_{Tc5})\big) + b_{Tc6}$$

$$T_C(m,K) = \frac{1}{a_{Tc}\left(\frac{1}{m^{0.5}} + \frac{1}{2m}\right) + b_{Tc}}$$

a_{Tci}	b_{Tci}
-2.75057	0.02474
13.61730	0.01125
4.15653	1.13867
-0.00018	0.09957
	-2.95074
	0.20299

Table 7.6: Fitting parameters for the critical temperature.

c. Saturated liquid density

$$a_\rho = a_{\rho1} + a_{\rho2}cos\left(a_{\rho3} - K^{0.5}\right) + a_{\rho4}\left(K + tan\left(K\left(a_{\rho5}K + a_{\rho6}\right)\right)\right)$$

$$b_\rho = b_{\rho1}Cos\left(\left(b_{\rho2}\left(b_{\rho3} + K + erf(K)\right)\right)^{0.5}\right)\left(b_{\rho4} + K\right) + b_{\rho5}$$

$$\rho_{Tr=0.7}^{Liq,Sat}(m,K) = (-0.0156m + 0.7681)f_1 + \left(a_\rho + m^{b_\rho}\right)f_2$$

$a_{\rho i}$	$b_{\rho i}$
-0.22344	0.00020
-0.00413	1.04394
-0.20208	0.12155
-0.00010	18.98966
-0.99859	-0.05724
-0.99859	
-995.00260	

Table 7.7: Fitting parameters of the saturated liquid density at Tr=0.7.

d. Saturated liquid viscosity

$$a_\mu = a_{\mu1}tanh\left(a_{\mu2}\left(a_{\mu3} + \left(a_{\mu4}cos\left((a_{\mu5}(a_{\mu6} + K))^{0.5}\right) + K\right)\right)erf\left(a_{\mu7}sinh(K)\right)\right)$$

$$b_\mu = K\left(b_{\mu1} + b_{\mu2}K^{b_{\mu3}}\right)^{\left(-1/b_{\mu3}\right)}$$

$$c_\mu = c_{\mu1}{}^{atan\left(\left(C_{\mu2}+K\right)C_{\mu3}K\right)}$$

$$\eta_{Tr=0.7}^{Liq,Sat} = (0.20346523m + 1.71486499)\left(a_\mu m^2 + b_\mu m + c_\mu\right) + 50\left(1 - exp(-10K)\right)/\left(1 + exp\left(10(m-1)\right)\right)$$

$a_{\mu i}$	$b_{\mu i}$	$c_{\mu i}$
0.03332	55677.12767	0.81887
0.03014	660.77650	5.38634
1.74329	2.88551	0.02548
1.26883		
0.40713		

1.26883		
0.40713		
0.11966		

Table 7.8: Fitting parameters of the saturated liquid viscosity at Tr=0.7.

7.A.2.2 Optimized parameters for the SR-LJCCG model.

a. Normal alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
n-C36	12	3.339	4.182	2538.061
n-C30	10	3.923	4.188	2522.341
n-C24	8	4.621	4.213	2503.175
n-C16	6	4.752	4.076	2396.719
n-C12	5	4.243	3.971	2295.615
n-C10	4	9.542	4.054	2284.894
n-C8	4	3.072	3.782	2126.740
n-C6	3	5.323	3.845	2070.560
n-C4	3	0.428	3.423	1734.347
C2	2	0.000	3.310	1441.630
C1	1	0.000	3.725	1212.071

Table 7.9: Optimized molecular parameters of the SR-LJCCG model for the n-alkanes.

b. Iso-alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
i-C4	3	17.020	3.434	1651.188
i-C5	3	10.238	3.641	1868.895
i-C6	3	11.351	3.841	2019.301
i-C8	3	137.494	4.213	2198.669
i-C16	5	17.878	4.295	2385.961

Table 7.10: Optimized molecular parameters of the SR-LJCCG model for the iso-alkanes.

c. Some polar components

Molecule	m	K*	sigma (Ang)	eps (J/mol)
CO2	3	22.553	2.451	1230.047
Toluene	3	0.725	3.655	2414.729
H2S	2	0.000	2.896	1761.620
RE245cb2	4	15.053	3.214	1499.589
R1336mzz(Z)	4	10.174	3.291	1643.354

Table 7.11: Optimized molecular parameters of the SR-LJCCG model for some polar components.

d. Associative molecules

Molecule	m	K*	sigma (Ang)	eps (J/mol)
Methanol	4	1.487	2.339	1909.772
Water	4	0	1.808	2406.339

Table 7.12: Optimized molecular parameters of the SR-LJCCG model for methanol and water molecules

7.A.2.3 Optimized parameters for the FF-LJCCG model.

a. Normal alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
n-C36	16	0.000	3.788	2388.384
n-C30	13	0.000	3.831	2381.400
n-C24	10	0.000	3.904	2377.067
n-C16	7	0.000	3.868	2301.980
n-C12	6	0.000	3.728	2182.513
n-C10	5	0.000	3.752	2151.909
n-C8	4	0.000	3.786	2117.102
n-C6	3	0.000	3.847	2070.925
n-C4	2	0.000	3.957	2007.283
C2	2	0.000	3.310	1441.630
C1	1	0.000	3.725	1212.071

Table 7.13: Optimized molecular parameters of the FF-LJCCG model for the n-alkanes.

b. Iso-alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
i-C4	2	0.000	3.974	1925.505
I-C5	3	0.000	3.646	1877.339
i-C6	3	0.000	3.846	2029.655
i-C8	3	0.000	4.189	2218.469
i-C16	5	0.000	4.305	2414.235

Table 7.14: Optimized molecular parameters of the FF-LJCCG model for the iso-alkanes.

c. Some polar components

Molecule	m	K*	sigma (Ang)	eps (J/mol)
CO2	3	0.000	2.449	1240.263
Toluene	3	0.000	3.656	2413.197
H2S	2	0.000	2.896	1761.620
RE245cb2	4	0.000	3.222	1512.784
R1336mzz(Z)	4	0.000	3.298	1652.940

Table 7.15: Optimized molecular parameters of the FF-LJCCG model for some polar components.

7.A.2.4 Optimized parameters for the SR-SAFTD model.

a. Normal alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
n-C36	11.911	3.393	4.193	2541.181
n-C30	9.906	4.018	4.206	2526.562
n-C24	8.023	4.579	4.209	2501.813
n-C22	8.478	3.522	4.016	2426.378
n-C16	5.978	4.828	4.084	2398.619
n-C12	4.940	4.512	3.989	2303.928
n-C10	4.322	5.630	3.945	2242.610
n-C8	3.679	6.112	3.895	2169.760
n-C6	3.020	4.876	3.836	2067.267
n-C4	2.389	26.110	3.715	1866.222
C2	1.725	0.000	3.490	1528.445
C1	1.246	0.000	3.455	1095.604

Table 7.16: Optimized molecular parameters of the SR-SAFTD model for the n-alkanes.

b. Iso-alkanes

Molecule	m	K*	sigma (Ang)	eps (J/mol)
i-C4	2.540	71.514	3.625	1748.929
i-C5	2.586	28.451	3.847	1962.073
i-C6	2.943	13.078	3.868	2031.162
i-C8	3.113	41.115	4.144	2169.076
i-C16	4.800	21.298	4.360	2411.532

Table 7.17: Optimized molecular parameters of the SR-SAFTD model for the iso-alkanes.

c. Some polar components

Molecule	m	K*	sigma (Ang)	eps (J/mol)
CO2	2.678	73.056	2.530	1279.106
Toluene	3.095	0.250	3.615	2388.609
H2S	1.611	0.000	3.123	1920.826
RE245cb2	3.463	33.832	3.399	1564.935
R1336mzz(Z)	3.671	16.018	3.395	1683.107

Table 7.18: Optimized molecular parameters of the SR-SAFTD model for some polar components.

d. Associative molecules

Molecule	m	K*	sigma (Ang)	eps (J/mol)
Methanol	4.233	1.446	2.292	1877.561
Water	3.530	0	1.890	2501.780

Table 7.19: Optimized molecular parameters of the SR-SAFTD model for methanol and water molecules.

7.A.2.5 SR-SAFTD fitting parameters.

a. α_{lj} coefficients used to calculate the c_{ij} matrix

l\j	*1*	*2*	*3*	*4*	*5*
1	-0.58801849	8.93225587	-17.16107315	5.85525327	3.92782695
2	-13.07890819	86.88731712	-191.89846330	170.97117100	-53.03154320
3	13.13557664	-78.49722038	139.50630931	-80.20303501	6.53526447
4	-5.30880509	28.25448854	-35.64721975	-2.94402296	15.36598713
5	0.20240413	-0.63937475	-1.21763487	4.46776181	-2.79327811
6	3.83081242	-19.56823476	31.53240140	0.10112261	-23.38485467
7	26.25619506	-203.29310385	502.68793284	-499.87664352	175.24988251
8	-9.84340292	72.35946434	-100.01669572	-19.23219732	56.01554169
9	-1.98340524	19.15499324	-110.55802870	204.00883254	-110.52363414
10	0.75278287	-6.36676352	22.10984463	-31.25184826	14.78055871
11	-8.44207147	38.57642265	-47.21206915	-9.87321449	42.47727198
12	-19.84615242	218.57447611	-645.38526640	727.98241848	-283.22968406
13	-7.94179017	-8.96344182	18.81635182	70.61732051	-72.26774120
14	12.80696068	-70.61683404	220.10564487	-323.85740492	162.21157956
15	-1.50293395	10.85744733	-34.33469904	46.72424890	-21.87343485
16					
17	6.66655719	-30.87648765	34.94070012	12.56943771	-33.80971330
18	12.17366812	-160.02203581	507.88956689	-603.90498357	244.63374560
19	-0.00636818	60.88054434	-174.94081197	142.57387203	-27.20198564
20	-3.61536062	7.17952195	-43.82553103	107.49179493	-68.69754904
21	0.42363454	-3.23850142	13.71387165	-22.48104121	11.76624811
22	-1.80814917	8.35180122	-9.14437080	-5.20177168	10.12049684
23	-6.00350224	63.00131114	-187.70183897	218.16308591	-87.35589323
24	6.42453547	-59.42692330	148.77476268	-141.16616465	44.38336952
25	-3.09447947	24.70227845	-49.66922886	31.65327665	-2.77356734
26	0.22968144	-1.33183265	1.28203944	1.27542451	-1.54156070

Table 7.20: αlj coefficients used to calculate the cij matrix.

q1	q2	q3
-0.63089354	2.25607667	-0.22702254

Table 7.21: q_i coefficients used to calculate the c_{ij} matrix.

b. ν_{lj} coefficients used in the $g_{ee}(T^*, \rho^* \rightarrow 0, K^*, \sigma)$.

i\j	*1*	*2*	*3*	*4*	*5*
1	-0.63089354	2.25607667	-0.22702254	-1.30073433	1.38530097
2	1.35107203	-4.02973091	-1.08379600	11.35304245	-8.77586355
3	-0.39774620	-3.52590997	21.75755370	-33.20343099	18.94425929
4	-0.31369245	4.64059564	-15.88218095	20.15666803	-10.63865063
5	0.11944725	-1.20595685	3.58032678	-4.25208977	2.17288077
6	-0.01032958	0.09193378	-0.25615124	0.29431764	-0.14785685

Table 7.22: v_{lj} coefficients used in the $g_{ee}(T^*, \rho^* \rightarrow 0, K^*, \sigma)$..

c. Coefficients used in g_{00} (T*, ρ*→0, K*→0, σ).

a	1.47734517
b	2.08369891
c	1.39653716
d	1.21571582
e	0.43088421
f	0.07453163
g	0.13440669
h	0.00095443

Table 7.23: Coefficients used in g_{00} (T, ρ*→0, K*→0, σ).*

d. Illustration of the quality of the fit of g_{ee}(T*, ρ*, K*, σ).

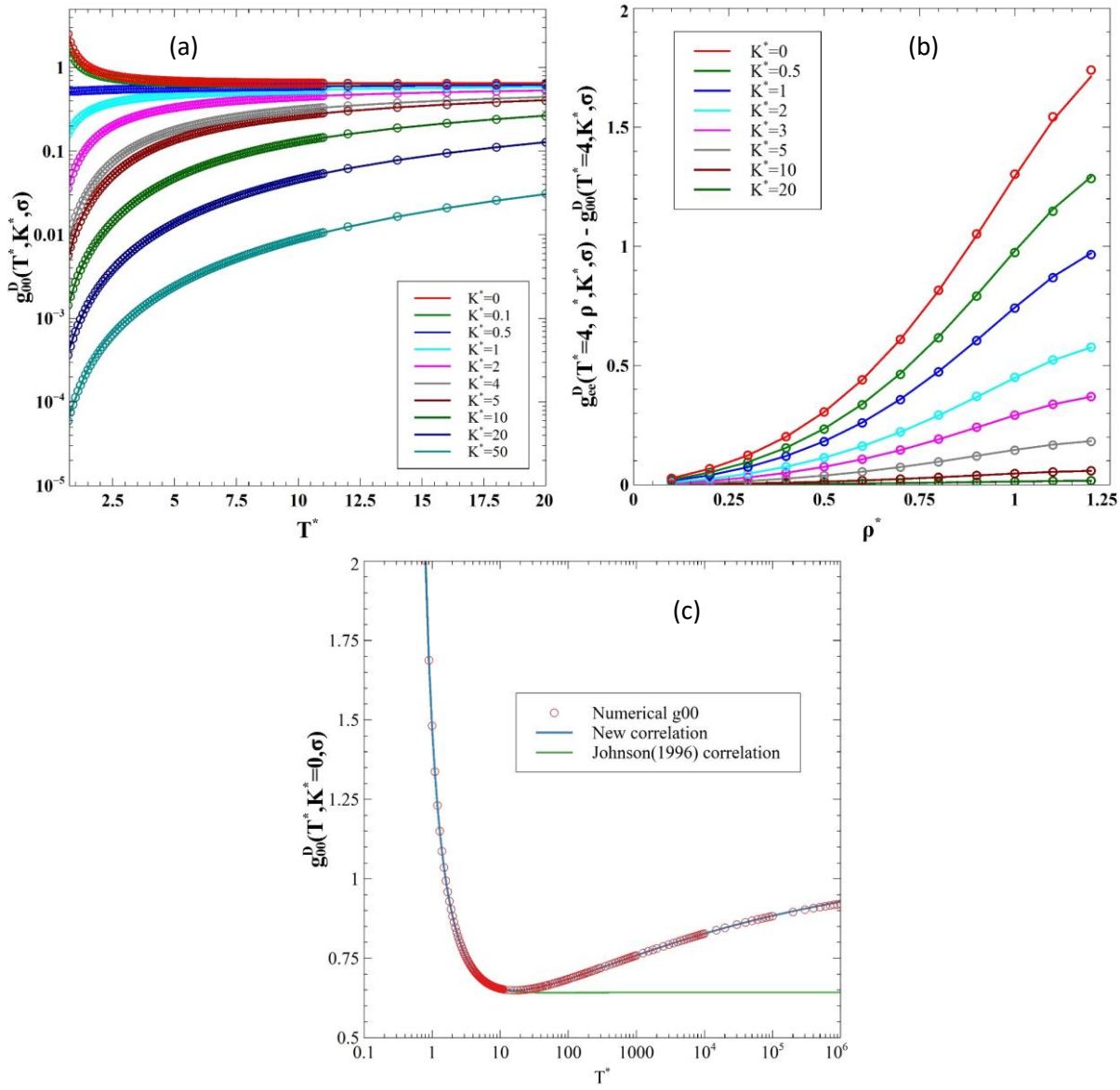


Figure 7.1: (a) Temperature dependence of the zero-density limit g_{00} (T, ρ*→0, K*, σ) at various rigidities. (b) density dependence of the term in g_{ee} (T*=4, ρ*, K*, σ) substracted out g_{00} (T*, ρ*→0, K*, σ) expressed by the matrix C_{ij}; (c)Extended g_{00} (T*, ρ*→0, K*→0, σ) to infinite T*.*

e. Illustration of the evolution of the parameters C_{ij}(K*).

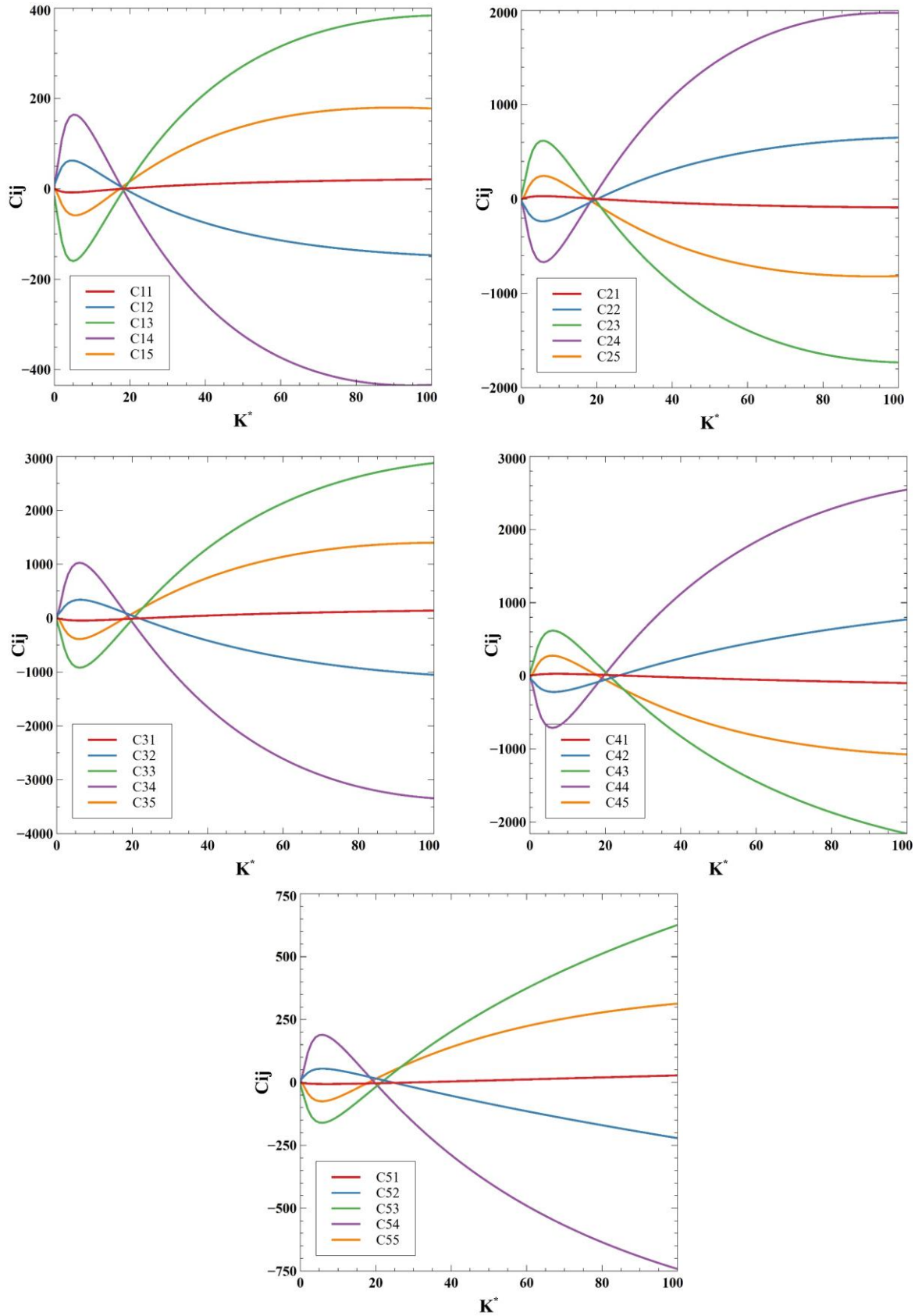


Figure 7.2: Rigidity dependency of the matrix $C_{ij}(K^*)$ coefficient used to calculate g_{ee} (T^*, ρ^*, K^*, σ)

7.A.2.6 Fitting parameters of the original Johnson parameters for the SAFTD (TPT1-D) [1] and SAFTM (TPT1-M) [2] models.

a. C_{ij} coefficients used in $g_{ee}(T^*, \rho^*, \sigma)$ (Eq(4.137))

i\j	1	2	3	4	5
1	1.92261089	-18.3410702	73.108684	-137.940285	100.181165
2	-15.3921034	202.930258	-875.941677	1580.79021	-1056.37502
3	45.4344615	-574.833828	2474.62424	-4449.62984	2920.48646
4	-49.920765	636.914217	-2710.64383	4849.76263	-3155.28520
5	18.7183896	-243.507214	1034.88816	-1844.93845	1194.01241

Table 7.24: c_{ij} coefficients used in the g_{ee} (T^*, ρ^*, σ) given by Johnson(1996).

b. α_i coefficients used in g_{00} (T^*, $\rho^* \rightarrow 0$, σ) (Eq(4.138))

a1	0.642180417
a2	-0.069376926
a3	2.21199344
a4	-4.2425198
a5	5.65047326
a6	-3.75078373
a7	1.0398651

Table 7.25: : Coefficients used in the g_{ee} (T^*, $\rho^* \rightarrow 0$, σ) given by Johnson(1996).

c. α_{ij} coefficients used in $g_{LJ}(T^*, \rho^*, \sigma)$ (Eq(4.96) & Eq(4.168))

i\j	1	2	3	4	5
1	0.493043466	2.152834989	-15.955682	24.03599967	-8.643795851
2	-0.470319831	1.147164749	37.889828	-84.66712149	39.64391411
3	5.032548624	-25.91539923	-18.862251	107.6370738	-66.60264974
4	-7.363315043	51.55356534	-40.519369	-38.79669265	44.6051392
5	2.90436073	-24.47881287	31.5001868	-5.336892037	-9.518344018

Table 7.26 c_{ij} coefficients used in the g^M (T^*, ρ^*, σ) given by Johnson et al.(1994).

7.A.2.7 Helmholtz free energy of the monomer segment.

The residual segment Helmholtz free energy of the LJ fluid is given by Johnson [3] using a modified BWD equation of state refitted on the internal energy and LVE properties. The expression is given by:

$$A^{seg} = \sum_{i=1}^{8} \frac{a_i \rho^{*i}}{i} + \sum_{i=1}^{6} b_i G_i$$

where a_i and b_i are temperature dependent parameters and G_i are density dependent parameters. This equation is used also in the for the segment monomer term given in SAFTM (TPT1-M) [2] (Eq(4.94), SAFTD (TPT1-D) [2] (Eq(4.136)) and in SR-SAFTD (generalized TPT1-D) of this work (Eq(4.167)). Their expressions and their corresponding constant parameters are:

```
def a(T):
    a1=x[1]*T + x[2]*T**0.5 + x[3] + x[4]*T**-1 + x[5]*T**-2
    a2=x[6]*T + x[7] + x[8]*T**-1 + x[9]*T**-2
    a3=x[10]*T + x[11] + x[12]*T**-1
    a4=x[13]
    a5=x[14]*T**-1 + x[15]*T**-2
    a6= x[16]*T**-1
    a7=x[17]*T**-1 + x[18]*T**-2
    a8= x[19]*T**-2
    return a1, a2, a3, a4, a5, a6, a7, a8

def b(T):
    b1= x[20]*T**-2 +x[21]*T**-3
    b2= x[22]*T**-2 +x[23]*T**-4
    b3= x[24]*T**-2 +x[25]*T**-3
    b4= x[26]*T**-2 +x[27]*T**-4
    b5= x[28]*T**-2 +x[29]*T**-3
    b6= x[30]*T**-2 +x[31]*T**-3 +x[32]*T**-4
    return b1, b2, b3, b4, b5, b6
```

Figure 7.3: Temperature dependent parameter ai and bi given in A^{seg}.

```
def F(d):
    return np.exp(-gamma*(d**2))

def G(d):
    G1= (1-F(d))/(2*gamma)
    G2= -(F(d)*d**2 -2*G1)/(2*gamma)
    G3= -(F(d)*d**4 -4*G2)/(2*gamma)
    G4= -(F(d)*d**6 -6*G3)/(2*gamma)
    G5= -(F(d)*d**8 -8*G4)/(2*gamma)
    G6= -(F(d)*d**10 -10*G5)/(2*gamma)
    return G1, G2, G3, G4, G5, G6
```

Figure 7.4: Density dependent parameters of G_i coefficients given in A^{seg}.

where the x[i] coefficients are given by the following table:

i	x[i]
1	0.86230851
2	2.976218765822090
3	-8.402230115796030
4	0.105413662920355
5	-0.856458382817459
6	1.582759470107600
7	0.763942194830545
8	1.753173414312040
9	2798.291772190370000
10	-0.048394220260858
11	0.996326519772193
12	-36.980002912724900
13	20.840122994346400
14	83.054021247172800
15	-957.479971520306000
16	-147.774622923499000
17	63.986078524715000
18	16.039936732948300
19	68.059166158643700
20	-2791.293578795940000
21	-6.245128304568450
22	-8116.836104958410000
23	14.887355595612200
24	-10593.467546550800000
25	-113.160763280282000
26	-8867.771540418820000
27	-39.869828444505400
28	-4689.270299917260000
29	259.353527743871000
30	-2694.523589434900000
31	-721.848763155021000
32	172.18020638632600

Table 7.27: x[i] coefficients given in A^{seg}

7.A.2.8 Transferable parameters used for the prediction of the properties of the iso-alkanes obtained from the SR-SAFTD.

For the iso-alkanes, the correlations are given by:

$$m = a_m M_w + b_m$$

$$m\,\sigma^3[nm^3] = a_\sigma M_w + b_\sigma$$

$$m\,{}^{\epsilon}/_{k_B}[K] = a_\epsilon M_w + b_\epsilon$$

$${}^{K}/_{\epsilon} = a_K\, {M_w}^{c_K}$$

where M_w is the molecular weight of the iso-alkane expressed in $[\frac{kg}{mol}]$, and the fitting constants are:

Iso alkanes	a_i	b_i
m	13.6705119366	1.6731353775
$m\,\sigma^3$	1.6356766033	0.0293422020
$m\epsilon/k_B$	41.8127799636	2.1052935109
K/ϵ	5.7628226402	−0.8908529915

Table 7.28: Constants of the transferable correlations of the iso-alkanes.

Appendix 7.B

7.B.1 Surface tension and associative molecules

7.B.1.1 Surface tension prediction.

In this subsection, the SR-LJCCG force field is evaluated on its capability to predict surface tension. Such property is known to be very challenging to be computed via direct molecular simulation tool due to some problems raised by the use of an appropriate r_{cut} , the system size, low temperature …[4, 5]. Thus, it is proposed here to use the correlation (and not simulations, reason for which results are shown in the appendix) based on the corresponding states approach for the surface tension given in [4, 6, 7] for fluids interacting through Mie or LJ potentials. The correlation relies on extensive molecular simulations run with $r_{cut} = 10\sigma$ and has been largely validated.

As shown in Figure 7.5 (a), very good agreement with the experimental data are found for the n-alkanes studied in with the SR-LJCCG model. The FF-LJCCG model tends to overestimates the surface tension at low temperature compared to the semi-rigid model, while both models predict accurately the high temperature region up to the critical point. Such an improvement close to the critical region is explained by the use of the CS strategy where the critical temperature is imposed, contrary to what could be obtained by using the classical fit over saturation pressure and liquid density [8] which exhibits systematic overestimation of the critical temperature, reflected then on the surface tension. Concerning polar components, good agreement is also obtained between the SR-LJCCG and the experimental data as shown in Figure 7.5 (b).

All these results on the surface tension confirm again the good representability of the proposed force field parameters, since no data on surface tension were included in the optimisation. Finally, even though results obtained using a correlation are very interesting, a direct simulation of this property would be required to confirm these results.

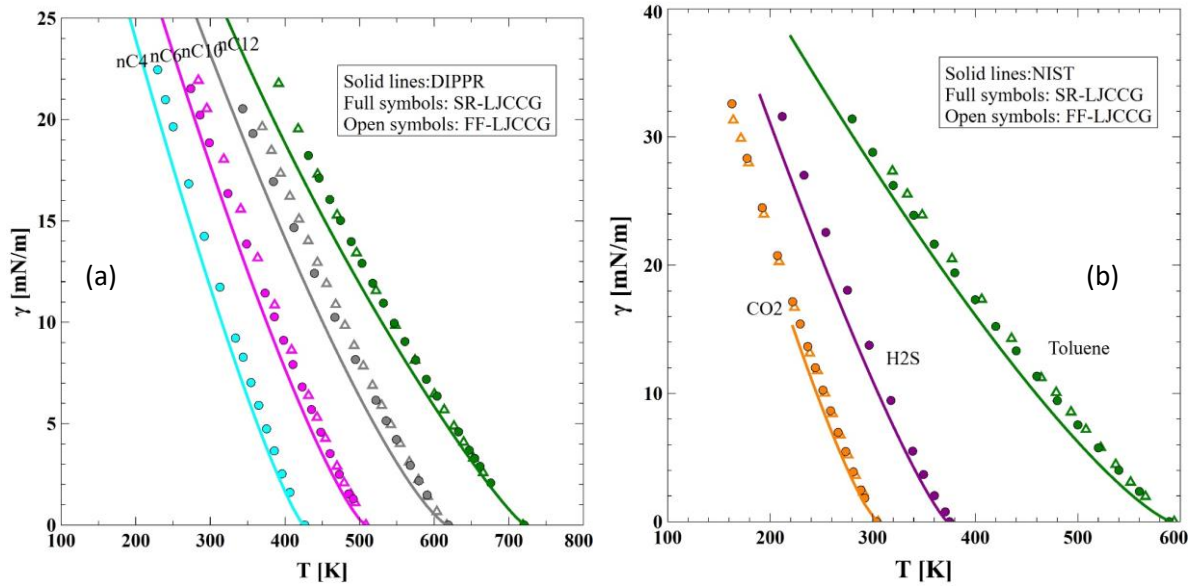


Figure 7.5: Predicted surface tension of normal alkanes and polar components using correlation from Galliéro (2010). (a) normal alkanes. (b) Polar components.

7.B.1.2 Associative molecules

As pointed it out in the discussions in chapters 3 and 4, the model and the strategy of the parameterization proposed in this work are not fully adequate for strongly associating molecules. This is illustrated in Figure 7.6 for methanol and water molecules. In fact, it appears that qualitative results could be obtained, depending on the property of interest. For example, the LVE properties are roughly captured by the model for the two studied molecules. Of course, the maximum in liquid density of water could not be reproduced by any non-associating model, unless in the case of our model temperature dependent parameters are chosen, similarly to what is done with the SAFT-γ-Mie force field for water [9]. The saturation pressure for both molecules are only qualitatively reproduced, which probably can be improved if a saturation pressure is imposed instead of the acentric factor. Regarding the viscosity, while quite good predictions are obtained for methanol, the model fails completely in the case of water.

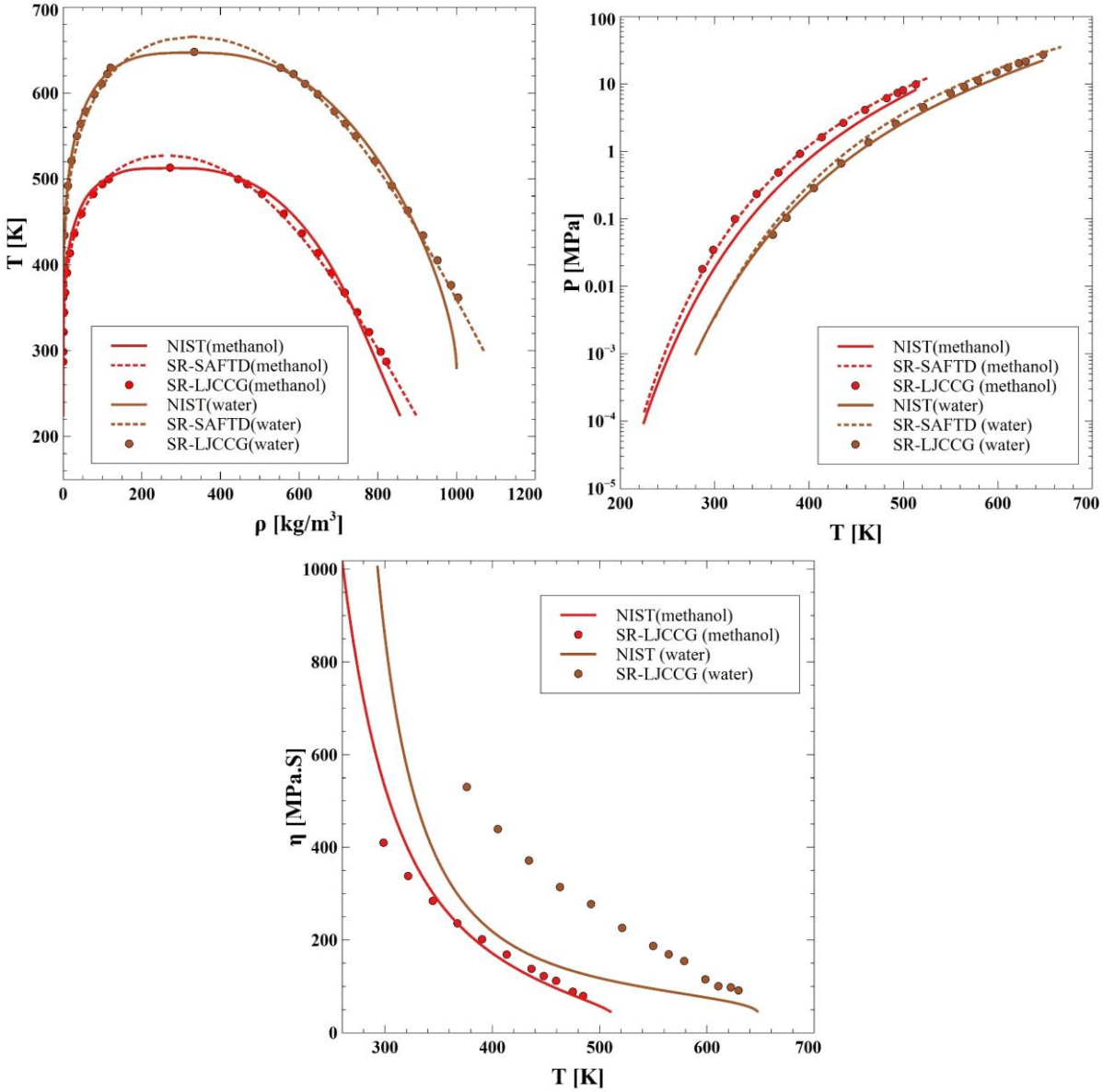


Figure 7.6: Test of the model and parameterization strategy on methanol and water molecules.

More precisely, the viscosity of water is not included in the fitting parameters due to the non-convergence of the F function (Min(W+η)), because, while the acentric factor of water is high, its viscosity is very low at T_r=0.7. If we refer to an equivalent of water for a normal alkane in terms of acentric factor, water and n-heptane have quite similar acentric factor, however, the viscosity of water at T_r =0.7 is two times lower.

Even though qualitative results could be obtained for some associating molecules, another question that may be raised is about the physical interpretation of the molecular optimized parameters. More precisely, these molecules exhibit high acentric factor due to their polarity while their molecular weight is not high when compared to their equivalent for normal alkanes molecules having the same acentric factor. Thus, applying the parametrization strategy proposed in this work for such molecules, would lead to high segment number "*m*" which may be interpreted as corresponding to large molecule, while it is not. This is a limitation of the

proposed strategy which relies on the assumptions that the acentric factor is purely an indication on how the molecular shape deviates from the sphere. This is relatively valid only for non-polar molecules as the shape of the molecule is roughly the only contribution to the acentric factor, whereas the polarity contributes strongly to the acentric factor, as the saturation pressure is strongly affected.

Last, it is worthwhile to underline that, beyond the parametrization strategy, the interaction potential used in this work is spherical and symmetrical, while such molecules are known to exhibit highly directional potential interactions. Thus, the modelling of these molecules without introducing the hydrogen sites explicitly, can only be considered as a crude approximation.

7.B.2 References (Appendix A & B)

1. Johnson, J. K. Perturbation theory and computer simulations for linear and ring model polymers. *J. Chem. Phys.* **104**, 1729–1742 (1996).
2. Johnson, J. K., Mueller, E. A. & Gubbins, K. E. Equation of State for Lennard-Jones Chains. *J. Phys. Chem.* **98**, 6413–6419 (1994).
3. Johnson, J. K., Zollweg, J. A. & Gubbins, K. E. The Lennard-Jones equation of state revisited. *Molecular Physics* **78**, 591–618 (1993).
4. Galliero, G. *et al.* Interfacial properties of the Mie n−6 fluid: Molecular simulations and gradient theory results. *J. Chem. Phys.* **130**, 104704 (2009).
5. Müller, E. A., Ervik, Å. & Mejìa, A. A guide to computing interfacial properties of fluids from molecular simulations. *A guide to computing interfacial properties of fluids from molecular simulations* (2021).
6. Galliero, G. Surface tension of short flexible Lennard-Jones chains: Corresponding states behavior. *The Journal of chemical physics* **133**, 074705 (2010).
7. Blas, F. J., Martínez-Ruiz, F. J., Moreno-Ventas Bravo, A. I. & MacDowell, L. G. Universal scaling behaviour of surface tension of molecular chains. *J. Chem. Phys.* **137**, 024702 (2012).
8. Galliero, G. Equilibrium, interfacial and transport properties of n-alkanes: Towards the simplest coarse grained molecular model. *Chemical Engineering Research and Design* **92**, 3031–3037 (2014).
9. Lobanova, O., Avendaño, C., Lafitte, T., Müller, E. A. & Jackson, G. SAFT-γ force field for the simulation of molecular fluids: 4. A single-site coarse-grained model of water applicable over a wide temperature range. *Molecular Physics* **113**, 1228–1249 (2015).

ECOLE DOCTORALE :

École doctorale des sciences exactes et leurs applications

LABORATOIRE :

Laboratoire des Fluides Complexes et leurs Réservoirs,
UMR 5150 CNRS-TOTAL-UPPA

CONTACT

samy.khennache@univ-pau.fr
Université de Pau et des Pays de l'Adour
Avenue de l'Université
BP 576
64012 Pau Cedex