Doctoral Dissertation
Doctoral Program in Chemical Engineering (31st cycle)

Multiscale Modelling of Polymer Self-Assembly in Binary Solvent Mixtures

Alessio Domenico Lavino

* * * * * *

Supervisors
Prof. Daniele Marchisio
Prof. Paola Carbone

Doctoral Examination Committee:
Prof. Jan Sefcik, Referee, University of Strathclyde
Prof. Alberto Giacomello, Referee, Università degli studi di Roma 'La Sapienza'
Prof. Giancarlo Cicero, Politecnico di Torino
Prof. Guido Raos, Politecnico di Milano
Prof. Eliodoro Chiavazzo, Politecnico di Torino

Politecnico di Torino
May 2, 2019

This thesis is licensed under a Creative Commons License, Attribution - Noncommercial- NoDerivative Works 4.0 International: see www.creativecommons.org. The text may be reproduced for non-commercial purposes, provided that credit is given to the original author.

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

. . . .
Alessio Domenico Lavino
Turin, May 2, 2019

Summary

This thesis focuses on the multiscale modelling and simulation of the formation process of particulate colloidal systems. In particular, the focus is on the self-assembly of polymer molecules in solution, leading to the formation of polymer nanoparticles (NP). The process is also known as flash nano-precipitation (FNP). In FNP, NP formation is induced by solvent displacement: by mixing a good solvent (i.e., acetone) containing polymer molecules (poly-ε-caprolactone, PCL, in this study) with a bad solvent (i.e., water) their self-assembly in molecular clusters, or NP, is induced. It usually takes place in very small devices and here the confined impinging jets mixer (CIJM) is considered. It consists of a cylindrical axial symmetric chamber with two inlets on the sides and one outlet at the bottom. NP production finds a wide range of applications; among the others it is worthwhile to mention drug delivery systems, in which NP are responsible for carrying a drug inside the human body, with the specific aim of attacking an ill site (e.g., cancer cells). As drug nano-carriers, control of mean NP size, together with particle size distribution (or, equally, the cluster mass distribution), becomes a crucial aspect in modelling and simulating such complex systems.

Due the complexity of the phenomena involved in this process, three different scales are then investigated: the atomistic scale, by means of molecular dynamics (MD) and coarse-grained molecular dynamics (CGMD); the cluster scale, by means of a population balance model (PBM); the vessel scale, in which fluid dynamics effects on NP formation are investigated through computational fluid dynamics (CFD). These three scales are interconnected with each other: the PBM is implemented and coupled into a CFD code and it is, in turn, built up from MD simulations. One of the several advantages in using a multiscale method is that, sometimes, one scale turns out to be useful in understanding or further investigating some peculiar behaviours that cannot be captured at just one level of description.

More specifically, in this work, full-atom MD is used to study the mixing behaviour at the atomistic scale of acetone-water mixtures. MD calculations of acetone-water mixtures exhibit a strong phase separation (with most of the standard all-atom force fields), despite the well-known experimental fact that acetone is miscible with water in any proportion at room temperature. The strong de-mixing behaviour (particularly stressed at low acetone molar fractions) is demonstrated to be caused by a bad polarization response of the classical all-atom force fields, usually employed to

simulate such liquid systems. Therefore, the de-mixing of acetone-water mixtures is here overcome at the molecular scale by means of a Charge-on-Particle model (COP), in which a pseudo-polarizable approach is used. A charge displacement is introduced on the carbonyl group of the acetone molecule, and suitably modified in function of the acetone molar fraction in order to get the desired polarization response. Thanks to the COP model, the dipole moment is shown to increase together with the content of water and de-mixing vanishes. Other dynamical properties, such as viscosity or diffusivity, are also shown to have a better profile with respect to the original all-atom force field, in terms of agreement with experimental trends.

Modelling of long PCL chains in solution at molecular level is also investigated by means of coarse-grained molecular dynamics (CGMD). More specifically, the MARTINI CG force field is used, in which both thermodynamics and structural parametrizations are employed to optimise the CG model. It is a four-to-one mapping method, namely four heavy atoms are grouped in one CG bead. The main advantage in using a CG procedure is the possibility of simulating much larger systems than the one explored with all-atom simulations, saving computational time. Non-bonded Lennard-Jones parameters are chosen as tuning functions, for the optimization procedure. The single CG bead types are set up by matching the solvation free energies of the corresponding atomistic compounds, by means of the Bennett's Acceptance Ratio method. The non-bonded intrachain interactions are then optimised by matching the radius of gyration of the corresponding atomic chains. Although the CG model developed here shows the typical limitations of the CG approaches (e.g., the sharp globule-to-coil transition at medium mixture concentrations range), results are quite satisfactory and self-assembly dynamics of several PCL CG chains in solution can be investigated by using this CG model. Both the all-atom and CGMD simulations are carried out by means of the GROMACS simulation package.

Going to the cluster scale, a PBM is developed, in order to describe the evolution of the cluster mass distribution (CMD, i.e., the number of polymer molecules that belong to a NP) thanks to a suitable governing equation (i.e., population balance equation) and, at the same time, to predict the final mean NP size at the mixer outlet. Particle size is described in terms of the mean radius of gyration, by using the Flory's theory. The whole PBM is treated with quadrature-based moments methods (QBMM). The advantage of solving the problem in terms of the moments of the CMD is twofold: on one hand, moments are scalars that correspond to known physical properties; on the other hand, much less equations need to be solved and so the computational cost is lowered. Being FNP an extremely fast process, the effect of turbulent fluctuations (also known as micro-mixing) must be accounted for. Turbulent fluctuation effects on NP formation is then considered in the PBM in the context of the Favre-averaged Navier-Stokes equation approach, thanks to the direct quadrature method of moments coupled together with the interaction-and-exchange with the mean (DQMOM-IEM) method. The rate at which two polymer clusters collide and aggregate is modelled by means of aggregation kernels, built up from MD calculations, and suitably implemented in the PBM. Finally, the effect of fluid dynamics on NP formation (in terms of influence on both CMD and final radius of gyration), is studied by CFD, proving that kinetics effects are one of the key parameters in controlling the mean NP size. The model referring to the latter two scales are coupled together in a simple computational tool, by means of proper user-defined function (UDF) of a CFD code.

The multiscale model is eventually exploited to simulate the effect of different good solvents on the NP formation process with a particular focus on the effect on particle size. Acetonitrile and THF have been studied and the key parameters, responsible for the different NP size have been successfully identified, thanks to both empirical and theoretical (i.e., Flory-Huggins solubility theory) approaches. This fully-predictive model is validated against experiments. Modelling predictions are in a good agreement with experiments, especially for medium and high initial supersaturation ratios (i.e., local PCL to equilibrium PCL concentration ratio). At low supersaturations the model still shows a gap with experiments. Further investigations are required to figure out the key phenomena at low solute concentration, but the multiscale model yields very interesting insights into the solvent effects.

As alternative mathematical framework for the population dynamics scale, the conditional quadrature method of moments is developed here and implemented in an open-source tool, OpenQBMM. PBM and CFD simulations are performed with both commercial (ANSYS Fluent) and open-source (OpenFOAM) packages. Concluding, the achievements of this work can be summarised as follows:

- De-mixing issues in MD simulations of acetone-water mixtures was overcome thanks to a charge-on-particle model, at the atomistic scale;

- A coarse-grained model was developed to properly describe the thermodynamics and the structural behaviour of long polymer chains in binary mixtures, by means of the MARTINI force field.

- A novel population balance model (with different modelling approaches) for FNP is employed in which the number of the polymer molecules that belong to a cluster was used for the first time as internal coordinate of the PBE. This allowed also a better understanding on the different good solvents effects in FNP.

Future investigations can focus on:

- Considering the presence of the active principle or drug (together with the carrier/polymer);

- Using the CG model to achieve a better insight into the key phenomena of the dynamics of the self-assembly at the molecular scale, to be eventually passed to the PBM scale;

- Investigating the different good solvent effects at the atomistic scale.

Acknowledgements

I would like to acknowledge the University of Manchester and the IOWA State University that hosted me during my PhD. In particular, Dr Paola Carbone, Dr Alberto Passalacqua, prof. Rodney Fox, as well as all the guys and girls that belong to their research groups and that helped me a lot to get used to the new cities to which I moved. Let me mention, in particular, my former flatmate Giuliana Giunta with whom I have faced the incomparable experience of being homeless in Manchester for one day!

I know that my supervisors are paid to do their job and, therefore, they should not be cited here; however, I would like to acknowledge their laudable effort in their job, every day, in a seamless manner, not such a taken for granted and common thing at all.

That said, I am going to acknowledge all the other people in Italian. Un primo pensiero non può che andare ai miei compagni di avventura, Hermes e Luca, con i quali ho vissuto questi tre anni molto intensamente, a stretto contatto, anche quando eravamo distanti, addirittura in continenti diversi, partite a FIFA sempre con un solo risultato, il tavolo da biliardo del Borsellino con un fisica tutta sua, sagre e giornate passate a mangiare e bere in modo spensierato! La loro costante presenza mi ha dato la forza per superare molte difficoltà, non solo professionali; i discorsi fino a notte tarda, chiamate su skype di cinque ore da continenti diversi, insomma che dire.. BBC forever! A questa categoria si aggiunge anche Salvatore, sempre presente e disponibile, una persona sincera ed un vero e proprio amico!

Un ringraziamento va anche a tutti gli altri colleghi che hanno condiviso questo mio cammino pieno di pause caffè e di discorsi, a volte interminabili!

Come non menzionare i miei compagni di avventura universitaria di triennale e magistrale. Due di loro, tuttora presenti nella mia vita, Stefano e Riccardo, con i quali condividiamo ed abbiamo condiviso esperienze di ogni tipo.

Ringrazio inoltre tutti i miei amici anche quelli che sono lontani, ma che, nonostante la lontananza, mi hanno fatto sentire il loro affetto quotidianamente. Una particolare menzione va a Donato, Sara, Paolo ed Antonio.

Un pensiero mi va anche a tutte quelle belle persone conosciute in questi tre anni di Torino, Manchester ed Ames (IOWA). Le porterò per sempre nei miei ricordi e nel mio cuore!

Vorrei ringraziare in aggiunta tutti i miei amici ed amiche con i quali ho condiviso questi anni sportivi: i miei vecchi compagni di squadra del primo anno di dottorato

(Alfredo, Pia, Daria, Cinzia, Pietro.. sono tantissimi!) con i quali ho vissuto a stretto contatto persino nei mesi estivi, giocando fino a tarda notte a beach volley e condividendo tanti tanti bei momenti insieme, comprese le mie ustioni durante i tornei sotto il sole cocente (Oz, Simon, Cri, Lilia, Alice)! The University of Manchester Volleyball Club that engaged me and gave me so many moments full of happiness, friendships and, above all, the opportunity to meet such nice and kind people! La mia attuale squadra di pallavolo dell' ASD Kolbe, formata da un gruppo di ragazzi assolutamente fantastici, quasi una seconda famiglia!

In ultimo, ma non per importanza, ringrazio la mia famiglia che mi ha sempre sostenuto nel corso di tutta la mia carriera universitaria, supportando tutte le mie scelte. In particolare, i miei zii e i miei nonni (mia nonna Tina soprattutto), i miei genitori Antonio e Maria Luce e mia sorella Ilaria, ed Amelia che più di ogni altro mi ha non solo supportato ma anche sopportato nei momenti di crisi e nelle mie isterie! Un ultimo pensiero va a tutti i miei cari che ora non ci sono più e che continueranno a sostenermi da lassù.

Un grazie davvero di cuore a tutti.

To my family and Amelia
who have been constantly
supporting me in all the
difficulties throughout
this long, exhausting and
fulfilling pathway.

Contents

List of Tables xvi

List of Figures xviii

Notation xxvi

I Introduction and Theoretical Backgrounds 1

1 Introduction 3

2 Statistical Mechanics 13
- 2.1 Introduction to Statistical Mechanics for Atomistic Modelling 13
- 2.2 Statistical Ensembles 15
  - 2.2.1 Microcanonical Ensemble 15
  - 2.2.2 Canonical Ensemble 15
  - 2.2.3 NPT Ensemble 16
  - 2.2.4 Grand Canonical Ensemble 16

3 Theory of Atomistic Modeling 19
- 3.1 Generalised Lagrangian and Hamiltonian Equations of Motion 19
- 3.2 Liouville's Theorem and Ergodicity 21

4 Theory of Coarse-Graining Modeling 25
- 4.1 Relevant degrees of freedom and characteristic equations 27
- 4.2 Generalised Langevin Equation of Motion 29
- 4.3 Potential of mean force and superatom approach 30
  - 4.3.1 Superatom technique 31
- 4.4 Fluctuation-dissipation theorem 31
- 4.5 Langevin equation in generalised coordinates 33
- 4.6 Brownian dynamics 35
- 4.7 Fokker-Planck equation 36

5 **Theory of Continuum Modeling** **37**
5.1 From Atomistic to Continuum Modelling 37
5.1.1 From Liouville to Boltzmann Equation 37
5.1.2 From Boltzmann to Continuity, Navier-Stokes and Energy equations 39
5.2 Turbulence Modelling 42
5.2.1 Reynolds decomposition and RANS equations 42
5.2.2 $k - \epsilon$ standard turbulence model 43
5.3 Population Balance Modelling 44
5.3.1 PBE and GPBE 44
5.3.2 Laminar and Turbulent PBE 45
5.3.3 Second-order point processes and aggregation kernels for fine particles 46

II Computational Details, Algorithms and Numerical Techniques **49**

6 **All-Atom Molecular Dynamics** **51**
6.1 MD Algorithms 51
6.2 Integration Schemes 52
6.2.1 Euler Algorithm 53
6.2.2 Verlet Algorithm 53
6.2.3 Leap-frog Algorithm 54
6.2.4 Velocity-Verlet Algorithm 54
6.2.5 Predictor-Corrector Schemes 54
6.2.6 Stochastic and Brownian Dynamics Algorithms 55
6.3 Trotter expansion 56
6.4 Temperature Control 57
6.4.1 Berendsen Thermostat 57
6.4.2 Velocity-Rescale Thermostat 57
6.5 Pressure Control 58
6.5.1 Berendsen Barostat 58
6.5.2 Parrinello-Rahman Barostat 58
6.6 Force Fields and Interaction Potentials 59
6.6.1 Interaction Potentials 59
6.7 Constraints, Periodic Boundary Conditions and Cut-off 64
6.8 Polarization Effect 66
6.8.1 Virtual Interaction Site 66

7 **Coarse-Grained MARTINI Force Field** **71**

8 Population Balance Model and Computational Fluid Dynamics 75
8.1 Univariate Distributions and Gaussian Quadrature 75
8.2 Computing Nodes and Weights of the Quadrature Approximation . . . 79
8.2.1 Product-Difference Algorithm 79
8.2.2 Wheeler Algorithm 79
8.3 Realizability of a Moments Set and Correction Algorithms 80
8.4 Multivariate Distributions 81
8.4.1 Conditional Quadrature Method of Moments 83
8.5 Direct Quadrature Method of Moments 85
8.6 Turbulent Mixing 86
8.7 Transport Equation of the Moments of the NDF 87

III Results, Discussions and Future Developments 89

9 Full atom molecular dynamics: the demixing problem in acetone-water mixtures 91
9.1 Introduction to the problem 91
9.2 The Charge-on-Particle Approach and Numerical Details 93
9.2.1 Numerical Details about the COP model 97
9.3 Results and discussion 98
9.3.1 Polarization response 98
9.3.2 Dynamical properties and de-mixing analysis 101
9.4 Conclusions 111

10 MARTINI coarse grained model in PCL self-assembly 115
10.1 Introduction 115
10.2 Atomistic Simulations 116
10.3 Coarse-Grained MARTINI Model 120
10.4 Bennett's Acceptance Ratio Method 121
10.5 Operating Conditions 123
10.6 Results and Discussions 125
10.6.1 Estimation of the solvation free energy of a single bead 126
10.6.2 Structural investigation of the polymer chain 128
10.7 Conclusions 135

11 Population balance model and computational fluid dynamics in NP production 139
11.1 Introduction 139
11.2 Governing Equations 141
11.2.1 Population Balance Model for FNP 141
11.2.2 QMOM and DQMOM-IEM for Flash Nano-Precipitation 145

11.2.3 CQMOM as alternative model for FNP 146
11.3 Operating conditions and computational details 150
11.4 Results and discussion 159
11.4.1 Effect of different good solvents on FNP 171
11.5 Conclusions 180

12 Conclusions **183**

A Industrial and Modeling Overview on Number Density Functions **187**
A.1 Number density function and population balance equation 187
A.2 Properties of number density functions 189

B Derivation of Source Terms for Second-Order Point Processes in PBE **191**
B.1 Source term derivation 191

C From PBE to moments transport equation - derivation and closure of the micro-mixing term **193**

D Topology of Charge-on-Particle Acetone model **195**

Bibliography **199**

List of Tables

6.1 MD general algorithm 52

9.1 Variation of the density of the mixture and the acetone dipole moment with the VS parameter, a, at acetone molar fraction, x_A, equal to 0.25. . 98

9.2 VS parameter, a, at different acetone molar fractions, x_A, obtained by the fitting against the experimental density of the mixture. 103

10.1 Acetone charge (e) distribution used at different acetone molar fractions, xA. C2-O2 corresponds to the carbonyl group; all the other atoms correspond to the two methyl groups. For each charge distribution, the corresponding acetone dipole moment, $\boldsymbol{\mu}$, is reported. 117

10.2 Lennard-Jones interactions and parameter ε, between the repeat unit bead types P_2, C_α and the solvent bead types P_4 and N_a. 129

10.3 Mean radius of gyration, $\langle R_g \rangle$, of coarse-grained (CG) and atomistic (AA) PCL-10 and PCL-20, at different acetone molar fractions. The LJ interaction levels are II for all the possible beads interactions, namely the set of LJ parameters that best fit the atomistic radii of gyration for the solely PCL-10 (preliminary analysis). 131

10.4 Mean radius of gyration, $\langle R_g \rangle$, of coarse-grained (CG) and atomistic (AA) PCL-10, PCL-20 and PCL-30, at different acetone molar fractions, x_A. The LJ interaction levels are of type II ($\varepsilon = 4.5$ kJ mol^{-1}) for P_2- P_2, III ($\varepsilon = 4.0$ kJ mol^{-1}) for C_α- C_α and IV ($\varepsilon = 3.5$ kJ mol^{-1}) for P_2- C_α. . 132

10.5 Mean radius of gyration of the CG PCL-40 and PCL-60 at different acetone molar fractions. 134

11.1 Operating conditions adopted in the simulations related to PCL-14000 and to PCL-80000, with acetone as good solvent in the CIJM. The first and second columns refer to the initial PCL concentration in the acetone stream in mg mL^{-1} and mol m^{-3}. The third column refers to the final PCL concentration in the outlet stream (after mixing of acetone and water streams). The fourth column quantifies the supersaturation ratio ($S(\xi)$) at the outlet of the mixer. 152

11.2 Involved physical variables and relative values adopted in the simulations. 157

11.3 Boundary conditions related to additional scalars employed in the CFD simulations. 159

11.4 Main physical properties related to the three good solvents investigated here: density, viscosity and molar volume. 171
11.5 HSP and distance from PCL (Eq. 11.39) for all the components of the investigated systems. 173
11.6 Flory-Huggins parameter, $\chi_{s,p}$, for acetone, acetonitrile and THF as good solvents with the relative scaling factors obtained from Eq. (11.44), at mean good solvent molar fraction, $\bar{x}_s$. 174
11.7 Flory's parameters functional forms for acetone, already reported in Eq. (11.34) and (11.35), acetonitrile and THF. 175

List of Figures

1.1 Sketch of CIJM. Acetone, with PCL dissolved, and water flow respectively from left and right. Black arrows show the flux lines. There may be also organic molecules (e.g. drug, light blue beads in the figure). The mixing leads to supersaturation and to precipitation of polymeric clusters in micellar structures that may contain the organic molecules. In this work the presence of the drug is neglected. 6
1.2 Multiscale framework adopted in this work. From left to right all the investigated scales are reported: atomistic scale for a range up to 100 nm of characteristic length scale, clusters scale in the range of 100 nm - 1 μm, and the vessel scale that considers phenomena regarding length scales up to millimeters order of magnitude. 9
1.3 Models used in this work: atomistic models (i.e., all-atom molecular dynamics), mesoscopic models (e.g., coarse-grained molecular dynamics) and continuum models (population balance model and computational fluid dynamics). For each model, the corresponding physics representation, the main physics equations and the materials relation (e.g., the set of information needed to close the mathematical model) are reported. 10
1.4 Fully predictive modelling approach adopted in this work. 11
4.1 Sketch of a CG mapping (red beads on right) on a generic atomistic system (brown on left). 26
6.1 Lennard-Jones potential. 60
6.2 Representation of a bond between two generic atoms i and j. 61
6.3 Representation of an angle between a triplet of atoms i,j,k. 62
6.4 Representation of an improper dihedral (torsional) angle ξ_{ijkl} between the planes $i-j-k$ and $j-k-l$. 63
6.5 Representation of a proper dihedral (torsional) angle ϕ_{ijkl} between the planes $i-j-k$ and $j-k-l$. 63
6.6 Periodic boundary conditions in MD simulations, using cubic simulation boxes. Image taken and modified from Allen and Tildesley (1987) 65

6.7 Generic scheme of a VS built up with respect to N surrounding atoms $i, i+1, ..., i+N-1$. The force $\mathbf{F}_s$ acting on the VS is redistributed on the N surrounding atoms with respect to which the VS itself is constrained, contributing to the total force $\mathbf{F}_i$ acting on them (Eq. 6.41). 67
6.8 Scheme of VS built up with respect to $N = 2$ surrounding atoms. The three particles (VS plus the two atoms i and $i+1$) lie on the axis x. The relative distance from the two surrounding atoms is governed by the weights $w_i = 1 - a$ and $w_{i+1} = a$. 68
6.9 Numerical example of a VS built up with respect to $N = 2$ surrounding atoms and $a = 0.75$. The three particles (VS plus the two atoms i and $i+1$) lie on the axis x. The relative distance from the two surrounding atoms is governed by the weights $w_i = 1 - a = 0.25$ and $w_{i+1} = a = 0.75$. 69
7.1 Sketch of the MARTINI force field mapping. This image has been taken and modified from Marrink, Risselada, et al. (2007). 72
7.2 Levels of non-bonded LJ interaction parameters used in the MARTINI model. 73
7.3 Bead types non-bonded interactions in the MARTINI force field. This image has been taken and modified from Marrink, Risselada, et al. (2007). 74
9.1 Schematic representation of the charge-on-particle (COP) model. The VS bead is attached to the carbon of the carbonyl group through a harmonic spring and a represents its normalized position, depicted here as in the general scheme shown in section 6.8.1. The positive charge, $\delta^+ = q_c = -q_o$ is placed on the C; the negative charge, $\delta^- = q_o = q_s$ (initial oxygen charge, without COP model), is totally shifted on the VS ($q_{o,l} = 0$). The charge displacement contributes to the dipole moment, $\boldsymbol{\mu}$. An electric field acting on the VS makes it vibrate around its equilibrium position, $a \cdot r_0$, creating an induced pseudo-polarization effect. Being a greater than unity, the VS lies on the other side of the oxygen of the carbonyl group. 95
9.2 Carbon-oxygen (bottom) and carbon-VS (top) distances over simulation time at acetone molar fraction equal to 0.25. The first 30 ps are represented here, in which it is clear how the carbon-VS bond is independent on the carbon-oxygen one, as stated by the red and blue flags, enlarged in the insets. Indeed, when the carbon-VS distance reaches a local maximum (red flag, top inset), at the same time instant the carbonyl-oxygen bond is still increasing (red flag, bottom inset) and vice-versa for the blue flag. The same behaviour is observed for the whole simulation time (20 ns). 100

9.3 Carbon-oxygen (black solid line) and carbon-VS (green dashed line) distances over simulation time at acetone molar fraction equal to 1.00. The first 30 ps are represented here, proving that when there is no water, no polarization effect is detected. Consequently, the carbon-VS curve collapses on the carbon-oxygen one, fluctuating in phase as stated by the red and blue flags, enlarged in the insets. 101

9.4 Experimental density of the acetone-water mixture (filled blue diamonds) versus the acetone molar fraction, x_A, fitted by the COP model (open purple squares). The original SPC/E-OPLS model predictions (open black squares) are also reported. 102

9.5 Average dipole moment, $\boldsymbol{\mu}$, of acetone molecules versus the acetone molar fraction, x_A, obtained from the COP model (open purple squares) and by Pereyra et al. (2011) with the modified CHARMM27 (open black circles). The experimental value of the acetone dipole moment is 2.88 $\boldsymbol{D}$ (gas phase, Marcus, 1985), slightly lower than the QM/MM calculations (2.98 $\boldsymbol{D}$) to neat liquid (3.33 $\boldsymbol{D}$, Pereyra et al., 2011), to infinite dilution (4.8 $\boldsymbol{D}$, Georg et al., 2006). These last values are omitted in the picture, for a sake a clarity. 104

9.6 Comparison between the experimental acetone diffusion coefficient (Toryanik and Taranenko, 1987) versus the acetone molar fraction, x_A, (filled blue diamonds), with the COP model predictions (open purple squares), together with the original SPC/E-OPLS (Perera and Sokolić, 2004) model predictions (open black squares). 105

9.7 Viscosity dependence on acceleration amplitude at acetone molar fraction equal to 0.10 calculated by NEMD. Error bars are reported. . . 106

9.8 Comparison between the experimental viscosity coefficient (filled blue diamonds) of the acetone-water mixture (K. Howard and McAllister, 1958; Noda et al., 1982) as a function of the acetone molar fraction, x_A, with the COP model predictions (open purple squares) Error bars referred to the average viscosity estimation are reported; at $x_A = 0.75$ and 1.00 they are too small (10^{-5} order of magnitude) to be graphically detected. 107

9.9 Snapshots of the acetone-water mixture referred to the COP model. From left to right and from up to bottom the acetone molar fractions are 0.10, 0.25, 0.50, 0.75. The oxygen is represented in red, the hydrogen in white, carbon in green and the VS is shown in pink. 109

9.10 Radial distribution functions related to acetone-acetone, $g_{AA}(r)$, water-acetone, $g_{WA}(r)$, and water-water, $g_{WW}(r)$. The acetone molar fractions shown here are equal to 0.10 (black), 0.25 (red), 0.50 (green) and 0.75 (blue). 110

9.11 Density profiles of the COP model for acetone-water mixture along the box. From left to right and from up to bottom the acetone molar fractions are 0.10, 0.25, 0.50, 0.75. The three space directions are reported, namely density along x (black), y (red), and z axes (green). 111

10.1 From left to right there are reported the three snapshots related to acetone molar fractions equal to 0.25, 0.50 and 0.75 respectively in cubic simulation boxes of 3 nm length. It is clear how no de-mixing occurs. 117

10.2 Atomistic radial distribution functions for acetone-acetone, $g_{AA}(r)$, water-acetone, $g_{WA}(r)$ and water-water, $g_{WW}(r)$, at different acetone molar fractions, x_A. The red, green and black lines correspond respectively to x_A= 0.25, 0.50 and 0.75. 119

10.3 Initial mapping of both the PCL chain and the solvents used in this work. The repeat unit is described by N_{am} (a new bead introduced here) and C_1 bead types respectively for the ester and the alkyl part. Water and acetone are mapped respectively by P_4 and N_a bead types. 120

10.4 Thermodynamic pathways that lead to the determination of the ΔG^{solv} from a state A to a state B. On the left side, there is a schematic representation of the definition of ΔG^{solv}; on the right side, a sketch of the BAR method is reported, going from λ=0 (cavity production, non-interacting system) to λ=1 (fully interacting system). 122

10.5 (a) Snapshot of the N_a-P_4 CG acetone(blue)-water(pink) mixture for x_A = 0.25 and box length equal to 15 nm, with 27000 particles and after 10 ns. (b) Detail of the N_a-P_4 box in which a little cluster (blue, N_a) it seems to form; however, no evident phase separation is detected at the CG level. Densities profiles along x (black), y (red) and z (green) are reported at acetone molar fraction equal to 0.25 (c), 0.50 (d) and 0.75 (e). 126

10.6 Free energy of solvation of butane (squares) and methyl formate (triangles) in mixture from the atomistic simulations (black symbols) carried out in this work. The red results are referred to the MARTINI model, C_1 (squares) and N_{am} (triangles), without any modifications of the LJ parameter, ε. The purple curves refer to the CG model developed here, in terms of single beads (squares for the alkyl part and triangles for the ester part of the PCL repeat unit) varying suitably the LJ parameter, ε. The green symbols stand for the experimental ΔG^{solv} values respectively of butane and methyl formate in pure water. 127

10.7 Snapshots referred to the preliminary analysis conducted in this work, in which CG PCL-10 (left) and PCL-20 (right) are compared and reported here in pure acetone with the LJ intrachain interaction levels equal to II. This scheme refers to the numerical results reported in Table 10.3. . 130

10.8 Snapshots referred to the final configuration of the CG PCL, after the optimization procedure conducted on the atomistic radius of gyration. The LJ intrachain interaction levels of type II ($\varepsilon = 4.5$ kJ mol^{-1}) for P$_2$-P$_2$, III ($\varepsilon = 4.0$ kJ mol^{-1}) for C$_\alpha$- C$_\alpha$ and IV ($\varepsilon = 3.5$ kJ mol^{-1}) for P$_2$-C$_\alpha$. This scheme refers to the numerical results reported in Table 10.4. . 132
10.9 Flory's theory predictions (dashed lines) of the CG model developed here (discrete symbols) in pure good solvent (acetone, red) and pure bad solvent (water, blue) for all the PCL chain lengths investigated in this work. The dashed lines represent the Flory law reported in Eq. (10.9). 135
11.1 Sketch of the quadrature procedure with two nodes for the mixing ξ and the aggregation number n. 147
11.2 Sketch of CIJM. Acetone, with PCL dissolved, and water flow respectively from left and right. Black arrows show the flux lines. There may be also organic molecules (e.g. drug, light blue beads in the figure). The mixing leads to supersaturation and to precipitation of polymeric clusters in micellar structures that may contain the organic molecules. In this work the presence of the drug is neglected. 151
11.3 Solubility curve for PCL-14000 in acetone-water mixtures, reported as acetone molar fraction, x_A, versus the equilibrium concentration, $c_{eq}^{PCL}(x_A)$ (continuous line). Points above the curve correspond to stable mixture where PCL molecules do not self assemble, otherwise for the points below. The symbols (empty triangles) identify the polymer concentrations at the outlet, c_{out}^{PCL}, for equal acetone and water flow rates, investigated in this work. The dashed curve corresponds to the supersaturation value equal to 200. The reason for choosing this latter value will be explained later on. 155
11.4 Solubility curve for PCL-80000 in acetone-water mixtures, reported as acetone molar fraction, x_A, versus the equilibrium concentration, $c_{eq}^{PCL}(x_A)$. Points above the curve correspond to stable mixture where PCL molecules do not self assemble, otherwise for the points below. The symbols (empty triangles) identify the six initial polymer concentrations, c_{out}^{PCL}, for equal acetone and water flow rates, investigated in this work. The dashed curve corresponds to the supersaturation value equal to 200. The reason for choosing this latter value will be explained later on. 156
11.5 Sketch of geometry adopted for the CFD simulation of the CIJM (corresponding to half of the actual CIJM geometry cut across the symmetry plane into two identical parts). The characteristic geometric parameters assume the following values: $d = 1$ mm, $D = 4.76$ mm, $H = 9.52$ mm and $W = 2$ m. 158

11.6 Contour plots for the magnitude of the velocity vector (m s^{-1}) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}. 161
11.7 Contour plots for the turbulent kinetic energy, k (m^2 s^{-2}) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}. 162
11.8 Contour plots for the turbulence dissipation rate, ε (m^2 s^{-3}) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}. 163
11.9 Contour plots for the moments of the CMD on the symmetry plane of the CIJM for an initial PCL-14000 concentration of 0.5 mg mL^{-1} and for, from left to right, inlet flow rates of 10, 40 and 120 mL min^{-1}; moments are reported according to their order (i.e. $j = 0, 1, 2, 3$) from top to bottom. 164
11.10 Contour plots for the mean radius of gyration (nm) on the symmetry plane of the CIJM for PCL-14000 with an initial concentration of 0.5 mg mL^{-1} and for, from left to right and top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}. 166
11.11 Reconstructed CMD at the CIJM outlet, $f(n)$, as a function of the number of molecules per cluster, n, at PCL-14000 concentration of 0.5 mg mL^{-1} and for, from left to right and top to bottom, inlet flow rates of 10, 20, 40, 60, 80 and 120 mL min^{-1}. The red vertical line indicates the mean value of n of the distribution, namely $m^{(1)}/m^{(0)}$. 167
11.12 Mean radius of gyration of the cluster exiting the CIJM versus the inlet flow rate as measured in experiments (black triangles) and as predicted by the model with only Brownian aggregation (continuous line) and Brownian and turbulent aggregation (dashed line) for PCL-14000 initial concentrations of, from left to right and top to bottom, of 0.5, 2.5, 5.0, 10.0, 15.0 and 25.0 mg mL^{-1}. 169
11.13 Mean radius of gyration of the clusters exiting the CIJM versus the inlet flow rate as measured in experiments (black triangles) and as predicted by the model with Brownian and turbulent aggregation (dashed line) for PCL-80000 initial concentrations of, from left to right and top to bottom, of 0.02, 0.2, 1.0, 3.0, 4.0 and 6.0 mg/mL. 170
11.14 Mean radius of gyration of the NP exiting the CIJM versus the inlet flow rate as measured in experiments (black symbols) and as predicted by the purely-aggregative model (dashed line, empty symbols) for PCL-14000 initial concentrations equal to 3.0 (top panel) and 5.0 (bottom panel) mg/mL in THF as the good solvent. 177

11.15 Mean radius of gyration of the NP exiting the CIJM versus the inlet flow rate as measured in experiments (black symbols) and as predicted by the purely-aggregative model (dashed line, empty symbols) for PCL-14000 initial concentrations equal to 3.0 (top), 6.0 (middle) and 9.0 (bottom) mg/mL in acetone (triangles) and acetonitrile (squares) as the good solvents. 178
11.16 Flory's exponent ν^* profiles in function of good solvent molar fraction for acetone (red), acetonitrile (green) and THF (blue). The discrete symbols correspond to the conditions of bad and good solvent (1/3 and 3/5, respectively) and at the outlet mean good solvent molar fraction (inset). 179
11.17 Mean squared radius of gyration in function of the number of molecules that form a cluster (aggregation number, n) for acetone (red), acetonitrile (green) and THF (blue). 180
A.1 Sketch of a probability density function, $f(x)$. The quantity $f(x)dx$ represents the probability of the system to lie in the configurational interval $[x, x + dx]$ 188

Notation

Symbols

$\langle A \rangle$	ensemble-averaged generic property A
$\mathbb{C}$	collision integral
$c_{eq,i}^{PCL}$	equilibrium PCL concentration in the quadrature environment i
$c_{loc,i}^{PCL}$	local PCL concentration in the quadrature environment i
d_p	particle diameter
$\mathbf{F}^C$	conservative (systematic) forces
$\mathbf{F}^D$	dissipative (frictional) forces
$\mathbf{F}^R$	stochastic (random) forces
$f(n), f_N$	number density function
$f(\xi), f_\phi(\xi), f_{N_p}$	probability density function
G^{solv}	free energy of solvation
$\mathscr{H}$	Hamiltonian
iL	Liouville operator
$\mathscr{K}$	kinetic energy
k	turbulent kinetic energy
k_B	Boltzmann constant
$k(x_A)$	Flory's coefficient
$\mathscr{L}$	Lagrangian
$\mathbf{M}$	mass tensor
M_w	molecular weight
m_i	$i-th$ particle mass
m^k, m_k	generic k^{th}-order moment (see also Greeks Letters)
m_i^j	generic j^{th}-order moment in the quadrature environment i
n_α	generic nodes of the quadrature formula/aggregation quadrature
P	pressure
p_α	generic weights of the quadrature formula/mixing quadrature

$\mathbf{p}$	generalized momenta coordinates vector
$\mathbf{q}$	generalized position coordinates vector
Q_{ens}	partition function referred to the ensemble *ens*
$\langle R_g \rangle$	mean radius of gyration
$\mathbf{r}$	generic position vector
S	entropy
S_ξ	supersaturation ratio
T	temperature
$\mathbf{U}$	number-averaged fluid velocity
$\mathscr{V}$	potential energy
$\mathbf{v}$	Cartesian velocity vector
w_α	generic weights of the quadrature formula/aggregation quadrature
$\mathbf{x}$	Cartesian position vector
x_A	acetone molar fraction
x_S	generic good solvent molar fraction

Greek Letters

β	aggregation kernel
Γ	molecular diffusion coefficient
Γ_t	turbulent diffusion coefficient
$\boldsymbol{\Gamma}$	phase space coordinates (positions and momenta)
γ_M	micro-mixing rate
δ	Dirac delta function
ε	turbulence dissipation rate
ζ	friction kernel in the CG dissipative term
$\boldsymbol{\eta}$	generic stochastic fluctuations
$\boldsymbol{\mu}$	dipole moment
μ	viscosity
μ_t	turbulent viscosity
ν	kinematic viscosity
$\nu(x_A)$	Flory's exponent
ξ	internal coordinate variable
ξ_α	generic nodes of the quadrature formula/mixing quadrature
$\langle \xi^k \rangle$	generic k^{th}-order moment (see also Symbols)
$\boldsymbol{\xi}$	internal coordinates variable vector
ρ	density
$\boldsymbol{\Sigma}$	viscous stress tensor
τ	generic characteristic time
χ_{sp}	Flory-Huggins parameter
Ω_ξ	internal coordinate phase space domain
Ω_v	velocity phase space domain

Operators

Δ	generic difference operator between a final and initial state
$\partial/\partial\mathbf{x}$	gradient operator in physical space
$\nabla_{\mathbf{x}}$	gradient operator in physical space
$\partial/\partial\mathbf{v}$	gradient operator in velocity space
$\nabla_{\mathbf{v}}$	gradient operator in velocity space
$\partial/\partial\boldsymbol{\xi}$	gradient operator in internal-coordinate phase space
$\nabla_{\boldsymbol{\xi}}$	gradient operator in internal-coordinate phase space
$\cdot$	scalar product

Abbreviations

AA	All-Atom
BAR	Bennett's Acceptance Ratio
CFD	Computational Fluid Dynamics
CGMD	Coarse-Grained Molecular Dynamics
CIJM	Confined Impinging Jets Mixer
CMD	Cluster Mass Distribution
COP	Charge-on-Particle
CQMOM	Conditional Quadrature Method of Moments
DOF	Degrees of Freedom
DQMOM	Direct Quadrature Method of Moments
FNP	Flash Nano-Precipitation
FVM	Finite Volume Method
GPBE	Generalized Population Balance Equation
IEM	Interaction and Exchange with the Mean
MD	Molecular Dynamics
NDF	Number Density Function
NEMD	Non-Equilibrium Molecular Dynamics
NP	Nanoparticles
PBE	Population Balance Equation
PBM	Population Balance Model
PCL	Poly-ε-caprolactone
PDF	Probability Density Function
QBMM	Quadrature-Based Moments Methods
QMOM	Quadrature Method of Moments
RDF	Radial Distribution Functions
UDF	User-Defined Function

PART I

Introduction and Theoretical Backgrounds

Chapter 1

Introduction

Research efforts in nanotechnology are becoming more and more popular in the last years, thanks to the wide range of applications that characterize the nano-world. Among the others, one of the most famous is the production of nanoparticles (NP). Macromolecule self-assembly into NP or molecular clusters is being widely investigated with both experimental and modelling approaches in the last decade (Celasco et al., 2014; Valente, Stella, et al., 2012) and due to their wide range of applications, polymer NP are becoming more and more important. The application fields spread out on a wide spectrum, such as medicine, cosmetics, pharmaceuticals and textile industry. In the latter one, for example, NP are produced in order to carry an active agent responsible of antimicrobial or antibacterial actions (Ferri et al., 2017); in the pharmaceutical field, the most common and investigated area is the NP production as a method for drug delivery systems inside the human body, in which NP behave as drug nano-carriers. Polymer NP can in fact carry a drug entrapped in them throughout the blood stream and, thanks to their degradation time, the polymer degrades releasing the drug and making it attack a specific ill site inside the human body (e.g. cancer cells). Drug releasing action is not the only technique employed nowadays in drug delivery systems; other techniques are also followed and tested as well, such as the UV-light activated metal NP (Ancona et al., 2018), but polymer NP represent surely a well-established and one of the most investigated approaches, pointing itself out as a solid alternative to the more aggressive and dangerous chemotherapy for the cure of cancer.

Polymer NP can be classified in nanospheres and nanocapsules (Valente, Celasco, et al., 2012): nanospheres have a monolithic matrix in which drugs are dispersed or absorbed on the surfaces or in the particles; on the other hand, nanocapsules consist of a layer-based structure, resulting in a vesicular system with an inner solid or liquid core surrounded by a polymeric membrane (the drug in this case is dissolved in the liquid core). In this work, nanospheres will be considered, referring to them with the generic term nanoparticles (NP). As drug delivery systems, mean NP size and particle size distribution are of paramount importance, being both of them related to the bead circulation time. The circulation time is, in turn, related to the reachability of the targeted area without escaping too soon or too late from the blood stream (Moghimi et al., 2001). First of all, NP size cannot exceed 1μm in order to freely circulate in the human blood stream and, at the same time, even if below this upmost limit, it has been proven that NP dimensions below 300 nm can increase the circulation time. Therefore, it is clear how the control of NP dimension and their size distribution play a key role in drug delivery systems (Maeda, 2001).

One of the numerous advantages in using drug delivery nano-carriers consists not only in the ability of reaching a targeted area (ill site inside the human body, e.g. cancer cells) but also the ability to carry water-insoluble drugs (hydrophobic molecules) that otherwise would face several problems in their use. For instance, the low-water insolubility causes the poor absorption and low bioavailability of the drugs (Lipinski et al., 2001); moreover, the attitude of this kind of drugs to lead aggregation phenomena might represent a serious risk for human health, since they may lead to embolism (Torchilin, 2007). Also these issues can be easily overcome by using polymer NP as pharmaceutical nano-carriers.

It is therefore straightforward that a biocompatible polymer must be used in order to avoid undesirable side effects for human health. In this work the interesting case of poly-ε-caprolactone (PCL) is considered. PCL is a saturated aliphatic polyester with hexanoate repeated units employed in a wide range of applications, ranging from biomedical devices to controlled drug-delivery systems. Its popularity is due to its biocompatibility with the human body (Maeda, 2001), as well as to the harmlessness of the degradation products, the main of which is the 6-hydroxycaproic acid that can be expelled out of the human body by metabolism (Who et al., 2000). The methodology developed and tested in this work has however a wider range of applications, as it can be used also to describe the self-assembly and aggregation of other colloidal and “soft systems”, such as surfactants, liquid crystals, proteins, DNA, and many more (Bockstaller et al., 2005), involved in the production of cosmetic, electronic, pharmaceutical and food products (Cohen et al., 2011; Collins et al., 2004; Ferrone et al., 1985; Knowles et al., 2009; Oosawa and Asakura, 1975; Oosawa and Kasai, 1962; Zhou and Ferrone, 1990).

Several techniques can be employed in order to induce self-assembly in solution. One of the most used techniques is represented by flash nano-precipitation (FNP) via solvent displacement. It usually takes place in very small mixers (order of magnitude of millimeters) and consists in mixing two or more inlet streams, of which at least one is made by a solvent in which a solute (the polymer in this case) is completely dissolved (for this reason called good solvent) and at least another one is made by another solvent completely miscible with the previous one, but incompatible with the solute (for this reason called antisolvent or non-solvent or bad solvent). As soon as the streams of good solvent/solute and antisolvent mix together, the antisolvent destabilises the mixture inducing the solute molecules (Jones, 2002) to aggregate and form NP. This phenomenon is called solvent displacement and is driven by the excess of concentration of the solute (i.e., polymer) in the final mixture. This excess is labelled as supersaturation.

Since supersaturation is crucial in NP formation by solvent displacement, a key role is played by mixing at all scales. Mixing is then a critical point in controlling NP mean size, as well as their distribution, since NP formation is a very rapid process (Johnson and Prud'homme, 2003b). Special mixers must be used to achieve the desired degree of mixing (mixing efficiency). In the last years several devices have been employed, tested and simulated and they can be classified mainly in two categories. The first one is the family of the so-called active mixers, namely a complete and rapid mixing obtained through an external energy input (e.g. electrospraying, Chan and Kwok, 2011). The second one consists of the so-called passive mixers that has the advantage to not require an energy input to induce mixing, and, on the other hand, the only power consumption needed is for the pumps that feed the inlet streams. Among them, the most used in the NP production field are: the vortex mixer (VM, Marchisio, Omegna, and Barresi, 2009; Marchisio, Omegna, Barresi, and Bowen, 2008), the multi-inlet vortex mixer (MIVM, Liu, C. Cheng, et al., 2008), the T-mixer (Gradl et al., 2006), the Y-mixer (Choi et al., 2005) and the confined impinging jets mixer (CIJM, Johnson and Prud'homme, 2003a).

This work focuses on FNP inside the CIJM. It is a mixer made by a cylindrical chamber with two inlets on the opposite sides and one outlet in the bottom. A sketch of the CIJM is reported in Figure 1.1. Looking at Figure 1.1, on the left side there is the good solvent inlet (acetone in this work), with the PCL dissolved in it. The presence of the drug has been neglected. On the right side, the antisolvent inlet is made by water.

Due to the several phenomena involved in FNP, a multiscale modeling approach is here presented and validated in which three different scales are investigated and interconnected between themselves:

- *Atomistic-scale*: by means of molecular dynamics (MD, all-atom and coarse-grained);

- *Clusters/NP-scale*: by means of a population balance model (PBM);

- *Vessel-scale*: by means of computational fluid dynamics (CFD).

The multiscale framework is depicted in Figure 1.2. Each scale is, in turn, investigated by means of different models and methods. In this work, the atomistic-scale is studied by both atomistic models such as all-atom molecular dynamics (MD), concerning small systems, and by mesoscopic models (i. e., coarse-grained molecular dynamics, CGMD), in order to characterise larger systems, such as long polymer chains in solution. Continuum models are employed to investigate the clusters- and vessel-scales. This is schematically reported in Figure 1.3. For each model, it is also reported the physics which the model itself is based on and represents (atoms, beads,

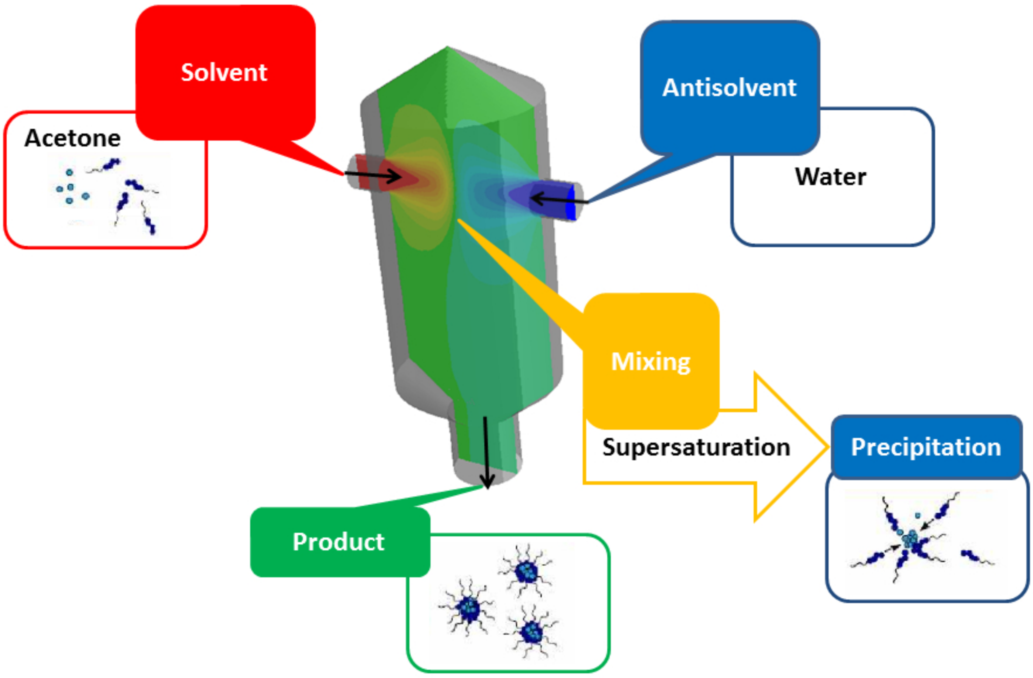

Solvent
Acetone
Antisolvent
Water
Mixing
Supersaturation
Precipitation
Product

Figure 1.1: Sketch of CIJM. Acetone, with PCL dissolved, and water flow respectively from left and right. Black arrows show the flux lines. There may be also organic molecules (e.g. drug, light blue beads in the figure). The mixing leads to supersaturation and to precipitation of polymeric clusters in micellar structures that may contain the organic molecules. In this work the presence of the drug is neglected.

etc.) together with the main equations and the materials relations, namely the set of information necessary to close the mathematical model.

More specifically, all-atom MD is here employed to investigate the intimate mixing properties of good and bad solvent in mixture (acetone and water in this work), addressing and overcoming the well-known issue of de-mixing of acetone-water mixtures, by means of a charge-on-particle model. Also the interactions and behaviour of short polymer chains with the solvents are analysed, by using the most common all-atom force fields for liquid systems (e.g., OPLS-AA, SPC/E). Coarse-grained molecular dynamics (CGMD) is used instead to investigate the behaviour of longer polymer chains in solution, developing a coarse-grained model based on the MARTINI force field. This CG model can be used in the future to get a better insight into the self-assembly phenomenon of multiple polymer chains in mixture, that otherwise would not be possible to evaluate by all-atom MD, due to the too high associated computational cost. All-atom MD and CGMD simulations are carried out by using the GROMACS simulation package (open-source tool).

All the information achieved via MD (all-atom and coarse-grained) at the atomistic-scale can be passed to the upper scale, namely the clusters-scale, investigated by a suitable PBM. The aim of the PBM is to study the evolution of the cluster mass distribution (CMD, i.e., the number of polymer molecules that form a NP) and, at the same time, the mean NP size at the CIJM outlet in terms of mean radius of gyration, thanks to the Flory's theory of real polymer chains (Flory, 1953). A novel approach is here proposed, in which the number of molecules that form a cluster (or, NP) is used as internal coordinate of the PBM, overcoming in this way the usual distinction between nucleation, molecular growth and aggregation. Thanks to this modelling strategy, for the first time the aggregation kernels are built upon MD simulations, in which the functional forms of Flory's parameters are fitted directly from such atomistic models, as function of the good solvent molar fraction. Turbulent fluctuations on NP formation is taken into account by means of the direct quadrature method of moments coupled together with the interaction and exchange with the mean method (DQMOM-IEM) and by solving the Favre-averaged Navier-Stokes equation. The $k - \varepsilon$ standard turbulence model is used and the quantities transported in it are directly implemented into the aggregation kernels of the PBM, representing another important link between the clusters and vessel scales. In the flow field context, all the properties are Favre-averaged, in order to take into account the density fluctuations in CFD analysis. The whole PBM is treated with a quadrature-based method of moments approach (QBMM, Marchisio and Fox, 2013). Two nodes in the quadrature approximation are considered for both mixing and aggregation, since it is a good compromise in terms of accuracy of the method and computational cost. Moreover, two different modelling approaches are presented, following the QMOM procedure: the first one consists in keeping mixing and aggregation separated into two different univariate PBEs, in which mixing is treated with the DQMOM-IEM formalism; the second one is a new approach in which aggregation and mixing are solved together in a unique bivariate PBE, applying the conditional quadrature method of moments (CQMOM) and closing the micro-mixing term with the IEM model. The main novelty of this new procedure is the use of pure moments in aggregation as internal coordinates of the bivariate PBE. Finally, at these length scales, the effect of different good solvents in NP production (CMD and mean NP size) was evaluated, investigating acetonitrile and THF besides acetone. By using the continuum models presented here, together with a computational fitting procedure, new functional forms for the Flory's law were extrapolated and validated against experiments.

The latter two scales (clusters- and vessel scales) are coupled together in a unique tool in the CFD code. Both commercial (ANSYS Fluent) and open-source (OpenFOAM) packages are employed to carry out the simulations related to clusters- and vessels scales. A fully predictive model is then developed and presented in this work and the way in which the different models interact with each other is schematically summarized in Figure 1.4.

This work is structured in three main parts: in Part I there is a general description of the main theoretical backgrounds behind the three scales investigated here: a general overview on statistical mechanics is reported in Chapt. 2; theoretical backgrounds of all-atom MD is presented in Chapt. 3, the coarse-grained general theory in Chapt. 4 and the theory of continuum modeling is reported in Chapt. 5. Part II is focused on the computational details concerning the three different scales always in the same order: a closer overview on the main algorithms and computational tools are presented in Chapt. 6 for all-atom MD, in Chapt. 7 to depict the specific technique used in this work (MARTINI force field) concerning the CGMD and in Chapt. 8 the QBMM used to solve PBM-CFD scale is presented. Part III deals with the results of this multiscale modelling approach, in which different topics and goals have been analyzed and hugely discussed: Chapt. 9 deals with the de-mixing issue of acetone-water mixtures in MD simulations, in which a new pseudo-polarizable model is presented and validated; a MARTINI CG model is developed and validated for PCL in acetone/water mixtures in Chapt. 10 and PBM-CFD approach is used to study FNP in Chapt. 11. Final considerations and future developments are finally reported in Chapt. 12.

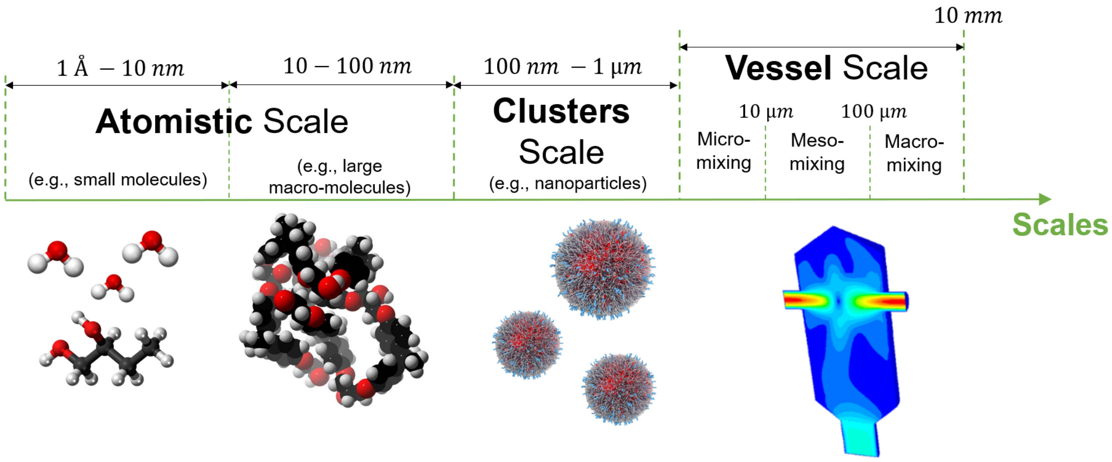
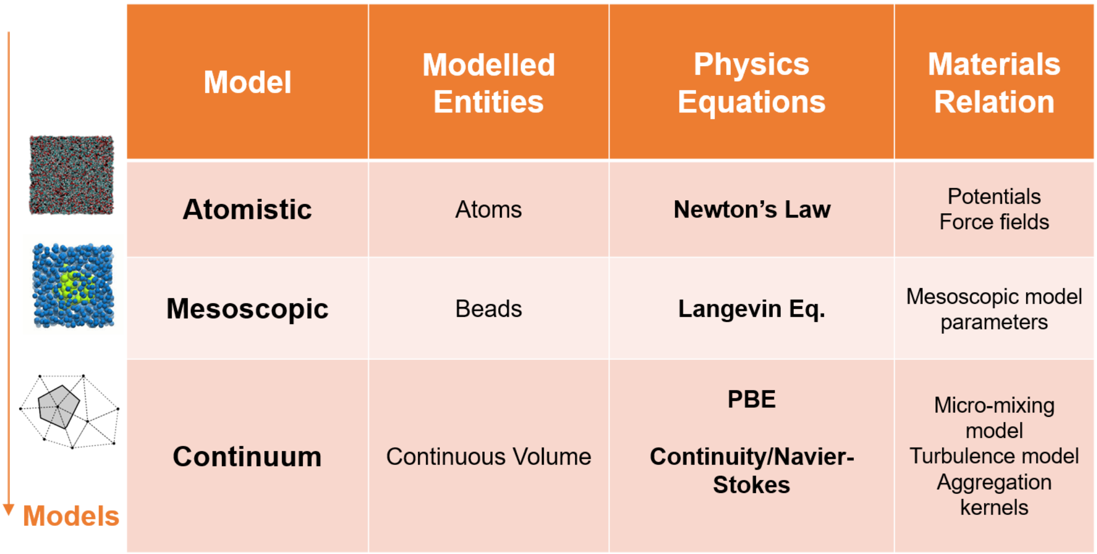

10 mm
1 Å – 10 nm
10 – 100 nm
100 nm – 1 μm
Vessel Scale
Atomistic Scale
Clusters Scale
10 μm
100 μm
Micro-mixing
Meso-mixing
Macro-mixing
(e.g., small molecules)
(e.g., large macro-molecules)
(e.g., nanoparticles)
Scales

Figure 1.2: Multiscale framework adopted in this work. From left to right all the investigated scales are reported: atomistic scale for a range up to 100 nm of characteristic length scale, clusters scale in the range of 100 nm - 1 μm, and the vessel scale that considers phenomena regarding length scales up to millimeters order of magnitude.



Model	Modelled Entities	Physics Equations	Materials Relation
Atomistic	Atoms	**Newton’s Law**	Potentials Force fields
Mesoscopic	Beads	**Langevin Eq.**	Mesoscopic model parameters
Continuum	Continuous Volume	**PBE** **Continuity/Navier-Stokes**	Micro-mixing model Turbulence model Aggregation kernels

Figure 1.3: Models used in this work: atomistic models (i.e., all-atom molecular dynamics), mesoscopic models (e.g., coarse-grained molecular dynamics) and continuum models (population balance model and computational fluid dynamics). For each model, the corresponding physics representation, the main physics equations and the materials relation (e.g., the set of information needed to close the mathematical model) are reported.

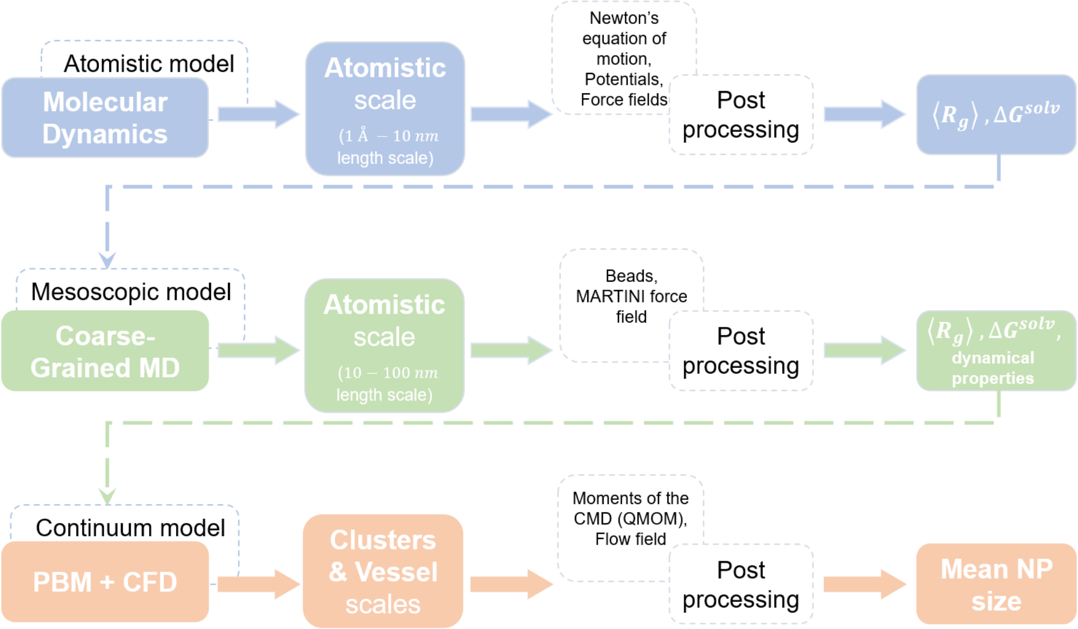

Atomistic model
Molecular Dynamics
Atomistic scale
(1 Å – 10 nm length scale)
Newton's equation of motion, Potentials, Force fields
Post processing
$\langle R_g \rangle, \Delta G^{solv}$
Mesoscopic model
Coarse-Grained MD
Atomistic scale
(10 – 100 nm length scale)
Beads, MARTINI force field
Post processing
$\langle R_g \rangle, \Delta G^{solv}$, dynamical properties
Continuum model
PBM + CFD
Clusters & Vessel scales
Moments of the CMD (QMOM), Flow field
Post processing
Mean NP size

Figure 1.4: Fully predictive modelling approach adopted in this work.

Chapter 2

Statistical Mechanics

This second chapter is meant to give a theoretical background, necessary for a deeper insight into atomistic modelling. A brief introduction concerning statistical mechanics is here reported, jumping over the quantum derivation and showing the results of its classical limit, being this one the basis of classical molecular simulations.

2.1 Introduction to Statistical Mechanics for Atomistic Modelling

In this first section a general overview on statistical mechanics is presented, being this area the base of molecular simulations. Molecular simulation was developed in order to analyse aspects that with a practical experiment are not possible to discern. For instance, in MD simulations of liquid water it is possible to measure instantaneous velocities and positions of each atom and molecule; on the other hand, these kinds of information cannot be directly tracked during experiments and, therefore, no direct comparison is possible with them. Experiments are then able to measure averaged properties, namely properties measured in a given volume that contains a given number of particles in some given physical conditions and, more importantly, in a given time interval. Basically, it is possible to infer that experiments give back a measurement of both ensemble- and time-averaged properties.

How can molecular simulations be used then to obtain results directly comparable to experiments? Or, in other terms, how can microscopic properties (directly evaluated by molecular simulation) be related to the macroscopic state (properties measured experimentally) of a system? The answer to this question is the statistical mechanics, which can be interpreted as the language needed to link molecular simulations results in terms of particles trajectories to averaged results. In fact, since a system is made by a huge number of molecules, it would be impossible to track the motion of each particle and, therefore, statistical methods turn out to be more effective. From an historical point of view, statistical mechanics appears for the first time in the scientific panorama during the last decades of 1800 thanks to the kinetic theory of gases developed by Maxwell and Boltzmann, even if at the time quantum mechanics was not discovered yet and, consequently, the first results were not too much accurate in some cases. Once quantum theory was developed, the suitable modifications have been made.

Let us start with some important definitions, passing then to the concept of ensemble average and the classical limit of statistical mechanics as powerful tool on which all MD is built upon. Let us define *thermodynamic state* of a given system as the values of macroscopic parameters that characterise a given system. For example, we might have a system made by 28 g of nitrogen at 40 ℃ and 3.5 atm. A *quantum state* of a system is all the set of equations needed to solve all the possible energy states of the Schrödinger equation related to a system of $6x10^{23}$ particles that are at 40 ℃ and 3.5 atm. For each time instant the system is, in a univocally defined quantum state, characterized by a given wave function (result of Schröedinger equation) and a given energy level characterized, in turn, by specific quantum numbers. Another important distinction to make is between micro- and macro-state. The latter is the thermodynamic state of a system. The former is instead the quantum state of a system. Different micro-states may correspond to the same macro-state. The macro-state is experimentally measurable and can be defined by giving thermodynamic variables such as temperature, pressure and number of particles (or other combination with three thermodynamic properties such as temperature, volume and number of particles, or even the energy instead of the temperature). The micro-state, on the other hand, is univocally specified by giving the quantum numbers that, in turn, define the wave function related to that specific particle. Now the question is: what is the link between micro- and macro-state? The answer is the statistical ensemble. In a classical scenario, which classical MD is based on, the microscopic state of the system is expressed in its classical limit. More specifically, a classical system made by N particles that evolves in space $\mathbf{x}$ and time t, each particle with its own trajectory, is characterised by a probability density function (PDF), $f_{N_p}(t, \mathbf{x}, \mathbf{v})$, that expresses the probability of finding the system in a specific configuration, namely in a specific infinitesimal interval of phase space $[\mathbf{x}, \mathbf{x} + d\mathbf{x}]$, where $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_N)$, and $[\mathbf{v}, \mathbf{v} + d\mathbf{v}]$, where $\mathbf{v} = (\mathbf{v}_1, ..., \mathbf{v}_i, ..., \mathbf{v}_N)$ are the particles velocity, to which a specific infinitesimal energy configuration corresponds. This continuous energy configuration is the Hamiltonian of the system, that corresponds to the total energy (summation of kinetic and potential energies) and it is formally defined in the following chapter, in Eq. (3.8).

Therefore, each generic macroscopic property A can be averaged on the ensemble behaviour (ensemble-average) by using the PDF as weight of the average itself:

$$\langle A \rangle = \int A f_{N_p}(t, \mathbf{x}, \mathbf{v}) d\mathbf{x} d\mathbf{v}. \tag{2.1}$$

As in the quantum approach the average behaviour is weighted on the discrete quantum states and the probability of finding the system in a quantum state is a discrete probability, in the classical limit that probability becomes a continuous function, namely the PDF, f_{N_p}, and the weight of the average is not related anymore to the discrete energy corresponding to that specific quantum state, but to the Hamiltonian of the system, that in this case is a continuous function. In either cases (discrete and continuous), this probability (the PDF in the classical limit) is proportional to a function that accounts for all the possible configurations (states in quantum approach) and expresses how these configurations are distributed or ”partitioned” in the system: this is the so-called partition function. It takes different forms depending on the ensemble, and represents the way to give ensemble-averages of macroscopic properties, starting from the microscopic (molecular) state of the system. In the next section the different statistical ensembles will be presented, together with the corresponding partition functions in their classical limit forms.

2.2 Statistical Ensembles

In this section the different statistical ensembles will be described, namely the different ways to relate the macrostate to all the possible corresponding microstates for a given system, showing their final classical-limit expression. The most used ensembles are: the microcanonical one, that is at constant number of particles N, volume V and total energy E (NVE), the canonical ensemble (NVT), the isothermal-isobaric ensemble (NPT), and the grand-canonical ensemble μVT.

2.2.1 Microcanonical Ensemble

The PDF for a microcanonical ensemble NVE is proportional to $\delta(\mathscr{H}(\mathbf{\Gamma}) - E)$, with $\mathbf{\Gamma}$ representing the phase space coordinates (positions and momenta), $\mathscr{H}$ is the Hamiltonian of the system. δ is the Dirac delta function for continuous representation (Kronecker delta for finite discrete representation). The NVE quasi-classical partition function reads as follows:

$$Q_{NVE} = \frac{1}{N!h^{3N}} \int \delta(\mathscr{H}(\mathbf{q}, \mathbf{p}) - E)\mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}. \tag{2.2}$$

The relationship to macroscopic thermodynamic variables is usually summarized in the following expression:

$$S/k_B = \ln Q_{NVE} \, , \tag{2.3}$$

where S indicates the entropy of the system and k_B is the Boltzmann constant.

2.2.2 Canonical Ensemble

Concerning the canonical ensemble NVT, its PDF is proportional to $\exp(-\mathscr{H}(\mathbf{\Gamma})/k_B T)$, and the classical limit of its partition function in the continuous form is

$$Q_{NVT} = \frac{1}{N!h^{3N}} \int \exp\left(-\mathscr{H}(\mathbf{q},\mathbf{p})/k_B T\right) \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}. \tag{2.4}$$

The correspondence to thermodynamic macroscopic variables is generally formulated in terms of the Helmholtz free energy A:

$$A/k_B T = -\ln Q_{NVT}. \tag{2.5}$$

Remembering the definition of Hamiltonian Eq. (3.8) of a system, then Eq. (2.5) can be factorized into two terms, the first one corresponding to the ideal contribution (kinetic), the second one corresponding to the excess contribution, or configurational contribution (potential):

$$\begin{aligned} Q_{NVT} &= \frac{1}{N!h^{3N}} \int \exp\left(-\mathscr{H}(\mathbf{q},\mathbf{p})/k_B T\right) \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p} = \\ &= \frac{1}{N!h^{3N}} \int \exp\left(-\mathscr{K}(\mathbf{p})/k_B T\right) \mathrm{d}\mathbf{p} \int \exp\left(-\mathscr{V}(\mathbf{q})/k_B T\right) \mathrm{d}\mathbf{q}. \end{aligned} \tag{2.6}$$

This possibility to split Eq. (2.4) in two different contributions turns out to be a crucial feature in molecular simulations, in particular for Monte Carlo methods (Allen and Tildesley, 1987).

2.2.3 NPT Ensemble

The isothermal-isobaric NPT ensemble PDF is proportional to $\exp(-(\mathscr{H}(\mathbf{q},\mathbf{p}) + PV)/k_B T)$, where the averaged quantity appearing in the exponent numerator results to be the thermodynamic enthalpy ($H = \langle \mathscr{H} \rangle + P\langle V \rangle$). The classical limit of NPT partition function is

$$Q_{NPT} = \frac{1}{N!h^{3N}} \frac{1}{V_0} \int \mathrm{d}V \int \exp\left(-(\mathscr{H}(\mathbf{q},\mathbf{p}) + PV)/k_B T\right) \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}. \tag{2.7}$$

Notice now that also the volume V has become a microscopic quantity that, therefore, needs to be integrated. Obviously, the dimensions of the basic volume unit V_0 must be chosen in order to guarantee Q_{NPT} to be dimensionless. The related thermodynamic function is correlated to the Gibbs free energy G through

$$G/k_B T = -\ln Q_{NPT}. \tag{2.8}$$

2.2.4 Grand Canonical Ensemble

The PDF corresponding to the grand canonical system μPT is $\exp(-(\mathscr{H}(\mathbf{q},\mathbf{p}) - \mu N)/k_B T)$ in which the chemical potential μ appears. The quasi-classical μPT partition function reads as follows

$$Q_{\mu PT} = \sum_N \frac{1}{N!h^{3N}} \exp\left(\mu N / k_B T\right) \int \exp\left(-\mathscr{H}(\mathbf{q}, \mathbf{p}) / k_B T\right) \mathrm{d}\mathbf{q} \mathrm{d}\mathbf{p}. \tag{2.9}$$

Note that now also the number of particles N is a variable of the partition function. Despite the continuous form of the latter, in many applications it turns out to be more useful to keep the discrete form, that is a finite summation over N. Its appropriate thermodynamic function is

$$-PV/k_B T = -\ln Q_{\mu PT}. \tag{2.10}$$

Chapter 3

Theory of Atomistic Modeling

This chapter describes a general overview on the theory of atomistic modelling. The Liouville's theorem, pillar of classical atomistic molecular dynamics (MD), together with the corresponding equations of motion are here presented.

3.1 Generalised Lagrangian and Hamiltonian Equations of Motion

A set of N particles (atoms, molecules) that form a system can be described by their generalized coordinates, namely positions $\mathbf{q}_i$ and momenta $\mathbf{p}_i$, with $i = 1, ..., N$:

$$\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N), \tag{3.1}$$

$$\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N). \tag{3.2}$$

In a classical system, particles can interact through positions-dependent potential $\mathscr{V}(\mathbf{q})$ and, due to their momenta, they also have a kinetic energy $\mathscr{K}(\mathbf{p})$, that can be defined as follows:

$$\mathscr{K}(\mathbf{p}) = \sum_{i=1}^{N} \sum_{\alpha} \frac{{p_{i\alpha}}^2}{2m_i}, \tag{3.3}$$

where m_i is the mass of the particle i and the index α runs over all the three spatial coordinates (q_{i1}, q_{i2}, q_{i3}) of the particle i. The potential energy $\mathscr{V}(\mathbf{q})$ contains information related to the interactions between all the particles in the system.

Let us define the Lagrangian, $\mathscr{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$, of such a system as

$$\mathscr{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathscr{K}(\mathbf{p}) - \mathscr{V}(\mathbf{q}) = \sum_{i=1}^{N} \sum_{\alpha} \frac{{p_{i\alpha}}^2}{2m_i} - \mathscr{V}(\mathbf{q}), \tag{3.4}$$

being the momenta related to the time derivative of the generalised positions, $\mathbf{p}_i = m_i \cdot \dot{\mathbf{q}}_i$. The generalised coordinates (Eq.s 3.1 and 3.2) evolve in time and, by using the Lagrangian function, the action S can be defined:

$$S = \int_{t_1}^{t_2} \mathscr{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt, \tag{3.5}$$

at times t_1 and t_2, and must be minimal. The principle of minimal action leads to the Lagrangian equation of motions (Berendsen, 2007):

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{\mathbf{q}}_i}\right) - \frac{\partial \mathscr{L}}{\partial \mathbf{q}_i} = 0, \tag{3.6}$$

for $i = 1, 2, ..., N$. By combining the definition of Lagrangian given in Eq. (3.4) with Eq. (3.6), the Newton's equation of motion is easily obtained:

$$m\ddot{\mathbf{q}} = -\frac{\partial \mathscr{V}(\mathbf{q})}{\partial \mathbf{q}}. \tag{3.7}$$

Another way to describe classical systems is represented by the Hamiltonian formalism. The Hamiltonian, $\mathscr{H}(\mathbf{q}, \mathbf{p})$, of the system is defined as the summation of kinetic $\mathscr{K}(\mathbf{p})$ and potential energy $\mathscr{V}(\mathbf{q})$, in turn, function of the generalized coordinates:

$$\mathscr{H}(\mathbf{q}, \mathbf{p}) = \mathscr{K}(\mathbf{p}) + \mathscr{V}(\mathbf{q}) = \sum_{i=1}^{N} \sum_{\alpha} \frac{{p_{i\alpha}}^2}{2m_i} + \mathscr{V}(\mathbf{q}), \tag{3.8}$$

representing the total energy of the system. It is possible to show (Berendsen, 2007) that, starting from the definition reported in Eq. (3.8), the Hamilton's equations of motion can be derived:

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\partial \mathscr{H}(\boldsymbol{\Gamma})}{\partial \mathbf{p}_i}, \\ \dot{\mathbf{p}}_i &= -\frac{\partial \mathscr{H}(\boldsymbol{\Gamma})}{\partial \mathbf{q}_i}, \end{aligned} \tag{3.9}$$

and their correspondence to Newton's equation of motion easily achieved, by just applying Eq. (3.8) to Eq. (3.9). Starting from the Hamilton formalism, it is then possible to derive all governing equations for a classical system and, by knowing initial positions and initial velocities of a set of molecules, then future positions and velocities can be computed. This is the starting point of MD simulations.

3.2 Liouville's Theorem and Ergodicity

In this section, the concept of average of a generic observable property is introduced, culminating in the formal statement of the Liouville's equation and the ergodic theorem. When we think about an average trend of a system, we tend to think about an average behaviour of such system over time. Furthermore, from an experimental point of view, it happens that a given quantity is measured in a given time interval, sufficiently long to determine a trustable measurement of that property itself leading to the so-called *time average*. However, the time average is not the only useful average in atomistic modelling, since, in terms of statistical ensembles introduced in Chapt. 2, also the average behaviour over all the possible micro-states is somehow interesting to measure. This is what is formally called *ensemble average*, defined in Eq. (2.1).

Let us call, then, the observable property in an experiment A_{obs} and let us relate it to an average over time, the observation time τ_{obs}. The last element we need to introduce is the generic phase space point, $\mathbf{\Gamma}(t)$ evolving over time, t, of the macroscopic system Γ considered in the experimental evaluation (it might be an NPT, or NVT, NVE, etc. ensemble). The relationship between the observable property A_{obs} and its time average $\langle A \rangle_{time}$ is

$$A_{obs} = \langle A \rangle_{time} = \langle A(\mathbf{\Gamma}(t)) \rangle_{time} = \lim_{t_{obs} \to \infty} \frac{1}{t_{obs}} \int_0^{t_{obs}} A(\mathbf{\Gamma}(t)) \mathrm{d}t. \tag{3.10}$$

Obviously, in a real classical molecular simulation, Newton's equations are solved for the time evolution of the trajectory $\mathbf{\Gamma}(t)$ for each particle, but the time related to this time evolution cannot be infinity. Therefore, a sufficiently long finite time t_{obs} will be chosen and the Newton's equations of motion will be integrated in a step-by-step procedure, by choosing a suitable number of time steps, $n_{timstep}$, so that the time interval obeys: $\Delta t = t_{obs}/n_{timstep}$. Consequently, Eq. (3.10) becomes

$$A_{obs} = \langle A \rangle_{time} = \frac{1}{n_{timstep}} \sum_{n=1}^{n_{timstep}} A(\mathbf{\Gamma}(n)). \tag{3.11}$$

Since the system is a group of N particles evolving in time and space and following the Newton's law of motion, then each particle will represent a point of the phase space $\mathbf{\Gamma}(t)$. This point in the phase space will change in time and physical space and, therefore, the N particles will be distributed at each time instant (or, more correctly, at each time step) following a probability density function (PDF) $f_{N_p}(\mathbf{\Gamma}(t)) = f_{N_p}(t; \mathbf{q}_i, \mathbf{p}_i)$, already introduced in Chapt. 2. Being the PDF an ensemble dependent property of the system, therefore dependent on the time t, on particle position $\mathbf{q}_i$ and on particle momenta $\mathbf{p}_i$, it will assume a different functional form depending on the statistical ensemble which it is referred to. For this reason the generic notation $f_{N_p}^{ens}(\mathbf{\Gamma}(t))$ should be used, in which $'ens' = NVT, NPT$, etc, but for a sake of brevity the solely notation f_{N_p}

will be generically used from now on, keeping in mind that the PDF assumes different functional forms by varying statistical ensembles. The conservation of a Hamiltonian system is guaranteed by the *Liouville's theorem*, which states that $df_{N_p}/dt = 0$, where d/dt is the total time derivative, that, in turn, can be decomposed in all the variables of the phase space, namely

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \dot{\mathbf{q}} \cdot \nabla_{\mathbf{q}} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}} \tag{3.12}$$

where $\partial/\partial t$ represents the partial differentiation with respect to time, $\dot{\mathbf{q}}$ and $\dot{\mathbf{p}}$ are the time derivative of position and momenta and $\nabla_{\mathbf{q}}$ and $\nabla_{\mathbf{p}}$ are respectively the derivatives with respect to positions and momenta. The notation $\nabla_{\mathbf{q}}$ stands then for $\nabla_{\mathbf{q}} = \partial/\partial\mathbf{q} = (\partial/\partial q_{i1}, \partial/\partial q_{i2}, \partial/\partial q_{i3})$, for the generic particle i. This notation may be used interchangeably from now on. By introducing the Liouville operator $iL = \left(\dot{\mathbf{q}} \cdot \nabla_{\mathbf{q}} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}}\right)$ and by using the Liouville's theorem, the Liouville's equation reads as follows:

$$\frac{\partial f_{N_p}(\mathbf{\Gamma}(t))}{\partial t} = -iLf_{N_p}(\mathbf{\Gamma}(t)) \tag{3.13}$$

whose formal solution is $f_{N_p}(\mathbf{\Gamma}(t)) = \exp(-iLt)f_{N_p}(\mathbf{\Gamma}(0))$. Considering, moreover, that $\exp(-iLt) \approx 1 - iLt + \mathcal{O}(t^2)$, the equation of motion referred to a generic property $A(\mathbf{\Gamma})$ then can be written in the form corresponding to $\dot{A}(\mathbf{\Gamma}(t)) = iL\ A(\mathbf{\Gamma}(t))$ or $A(\mathbf{\Gamma}(t)) = \exp(-iLt)A(\mathbf{\Gamma}(0))$.

Liouville's equation (Eq. 3.13) and its formal solution are the pillar of molecular simulations; particular attention deserves the exponential term $\exp(-iLt)$, called "propagator", which is the starting point to obtain all the algorithms useful to solve the equations of motion for classical MD simulations (Eq. 3.9), as it will be explained more clearly in Chap. 6. By using Eq. (3.9), the Liouville operator becomes $iL = \sum_i \left(\frac{\partial\mathcal{H}}{\partial\mathbf{p}_i} \cdot \nabla_{\mathbf{q}} - \frac{\partial\mathcal{H}}{\partial\mathbf{q}_i} \cdot \nabla_{\mathbf{p}}\right)$.

Let us now introduce the concept of ergodicity. Considering the system Γ, made by N classical particles, at equilibrium the particles PDF will be time independent, namely $\partial f_{N_p}/\partial t = 0$. The system will evolve over time t, occupying a point of the phase space at each time interval n_t, $\mathbf{\Gamma}(n_t)$, moving to the next point of the phase space $\mathbf{\Gamma}(n_{t+1})$ at the following time interval n_{t+1} and being replaced by another system that arrives from a state $\mathbf{\Gamma}(n_{t-1})$. All these movements over the points of the phase space will generate different trajectories in the phase space itself. If there is just one trajectory for which f_{N_p} is non-zero, it means that all the systems that are evolving over time will be experiencing all the points of the phase space (more clearly, all the points of the phase space will see all the possible evolutions of the system Γ in a long closed circuit, a trajectory $\mathbf{\Gamma}(t)$). This system is called *ergodic* and this property just described above is named *ergodicity*. The most important consequence of such a property is the fact that, in order to achieve an average behaviour of the system there are two choices: the first one consists in waiting for the entire time evolution of the system to evaluate how it behaves over time; the second one, straight consequence of ergodicity, is that one can freeze the system in a given status and evaluate an instant picture of the whole group of n_t systems in those frozen conditions, instead of following a specific one over the total observation time, since all the possible systems will experience all the possible points in the phase space. The two average evaluations (over time, $\langle A\rangle_{time}$, and ensemble average, $\langle A\rangle_{ens}$) will give the same result; in more mathematical terms, the ergodicity allows us to infer that

$$A_{obs} = \langle A\rangle_{ens} = \sum_{\mathbf{\Gamma}} A(\mathbf{\Gamma}) f_{N_p}(\mathbf{\Gamma}) = \langle A\rangle_{time} = \frac{1}{n_{timstep}} \sum_{n_t=1}^{n_{timstep}} A(\mathbf{\Gamma}(n_t)). \tag{3.14}$$

This configuration is the key feature of any molecular simulation and satisfies the following conditions:

- $f_{N_p}(\mathbf{\Gamma})$ for a given statistical ensemble does not change as the system evolves;
- any 'reasonable' starting $f_{N_p}(\mathbf{\Gamma})$ tends to its stationary solution for any initial conditions;
- the system is ergodic.

Chapter 4

Theory of Coarse-Graining Modeling

Due to the large time and length scales involved in many applications, the full-atom description of classical MD turns out to be not sufficient to fulfil a complete insight into such complex systems. Coarse-grained molecular dynamics (CGMD) is then needed, in order to reduce the number of degrees of freedom (DOF, Carbone and Avendaño, 2014; Cheung and Carbone, 2013; Karimi-Varzaneh et al., 2012; Taddese, Cheung, et al., 2015), by grouping a given number of atoms or even entire molecules together in a unique particle denominated “bead”. This action of grouping atoms altogether in a unique bead is called “mapping”. The mapping varies dependently on the kind of CG that one is using, but a general scheme is depicted in Figure 4.1. Thanks to the possibility of simulating larger systems for very long times, CGMD is becoming more and more popular in many biological and complex fields, such as proteins (Bonomi et al., 2017), DNA (Dans et al., 2010) and lipids (Marrink, de Vries, et al., 2004).

Different classifications may be done among the several existing CG techniques; basically, it is possible to identify two main categories: the bottom-up (systematic) and the top-down CG models. The latter are built up upon experimental data observed at the length scale at which the CG model is targeted to. The former are systematically derived from first principles or atomistic detailed-level MD. Examples of bottom-up CG models are the Iterative Boltzmann Inversion (IBI, Muller-Plathe, 2002), the Inverse Monte Carlo (this one by means of Monte Carlo simulations, Lyubartsev and Laaksonen, 1995), the force-matching models (Izvekov, Parrinello, et al., 2004; Izvekov and Voth, 2005; Noid et al., 2008) and more rigorous derived bottom-up techniques such as the conditional reversible work (CRW, Brini and van der Vegt, 2012), which uses thermodynamic cycles to calculate non-bonded interaction potentials, and the Mori-Zwanzig framework (Di Pasquale, Hudson, et al., 2018; Hijon et al., 2010), capable to derive a closed set of equations for the dynamics of CG systems.

At higher coarse-graining levels of resolution, mesoscopic CG models have been developed in order to capture the hydrodynamics of the system, such as the multiparticle collision dynamics (M. Howard et al., 2018) and the well-known dissipative particle dynamics (DPD, Groot and Warren, 1997). Developed at the

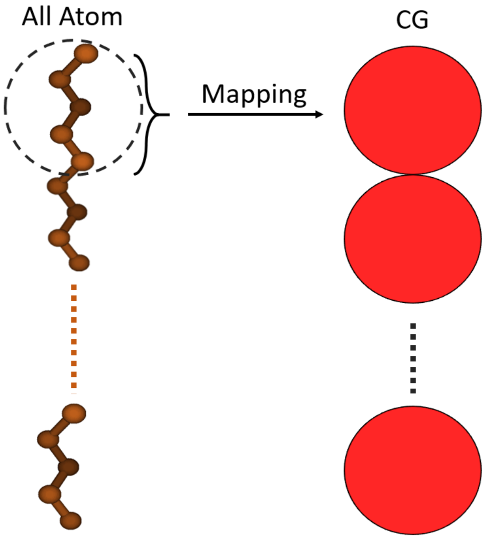

All Atom
CG
Mapping

Figure 4.1: Sketch of a CG mapping (red beads on right) on a generic atomistic system (brown on left).

beginning as a top-down model, recent works showed that bottom-up approaches may be used in order to achieve more accurate information about the range of validity of the DPD assumptions (especially concerning the Markovian friction term, Kauzlaric et al., 2012). Furthermore, DPD can be systematically derived via bottom-up techniques, such as the CRW-DPD (Deichmann et al., 2014). It is worth noticing also that the rigorous Mori-Zwanzig derivation can lead to the mathematical framework of DPD, despite the two approaches are completely different. Other CG methods can be built up as a combination of both top-down and bottom-up techniques. This is the case of the MARTINI force field, developed by Marrink, Risselada, et al. (2007). Here, the basic theory concerning the CGMD will be presented and has to be intended as a general theoretical approach to coarse-graining, independently on what kind of technique the user wants to adopt.

4.1 Relevant degrees of freedom and characteristic equations

Reducing the number of DOF of a given system leads to a new concept that governs the mesoscale world: the relevant variables (or, equally, relevant DOF). It means that the final simplified approach must be able to reproduce the "interesting" behaviour of a given system, neglecting the "uninteresting" details (some "uninteresting" variables are not considered anymore). This action of reducing the number of DOF to catch the relevant behaviour of a dynamic system is the core of the mesoscale, and every technique based on that falls under the label of CG. The relevant DOF can be chosen as spatial coordinates averaged of a given number of atoms or may belong to the densities grid type. The former leads to a discrete CG that is named "superatom approach" (treated later on in this chapter); the latter to the so-called mesoscopic continuum dynamics. Therefore, the first question is necessarily about what can be considered relevant and, on the other side, what can be considered irrelevant. In a preliminary way, it is mandatory to stress that a dynamic system is characterised by many DOF and may be described at various levels. For each level of description some variables are considered relevant in order to specify univocally the state of that system at that specific level.

This new concept of "level" suggests that the multiscale pathway, from atomistic to larger length scales, has a hierarchical structure, and each level of this "hierarchy" can capture a given amount of information at a given time scale, by considering the evolution of the characteristic equations related to those relevant variables. For example, at low length-ranges of atomistic scale, full-atomistic models capture the dynamics of the system considering each atom positions and momenta as relevant variables, and by solving then the Hamilton's equations of motion at the time scale corresponding to atoms collisions and vibrations (as explained in Chapt. 3). On the other hand, in continuum modelling, the corresponding relevant variables are dynamical invariant (mass, momentum and energy). At this level no equation of motion is considered, as the relevant variables are constant in time and the corresponding time scale is infinite (continuum).

The main concepts which the CG world is based on have been now introduced: the concept of “relevant/interesting” and “irrelevant/uninteresting” variables; the concept of “levels” and their hierarchical structure, as well as the very important concept of characteristic equations related to the specific level. It is noteworthy to stress here that the term “irrelevant” does not imply that those variables are neglected or forgotten, but, more correctly, are taken into account in different ways by using proper mathematical expressions (closures), that at a higher-coarser level allow to suitably reproduce the correct dynamics of the system. Another important definition regards the so-called Markovian systems. A system is defined Markovian if its future state is determined by the present, but not the past, state of the relevant variables. This reflects the main feature that the time scales of the relevant variables can be completely decoupled from the time scales of the irrelevant ones, and such a system is described by differential equations. When this time scales decoupling is not possible, then, a system is defined as non-Markovian and its characteristic equations are integro-differential.

The loss of information due to the elimination of some irrelevant DOF brings to two main terms that are considered in coarser levels, in order to reproduce the correct dynamics of the system. The first one is a friction term and it comes out from the principle that, by neglecting DOF in the coarse-graining procedure, the final CG system would be artificially “accelerated” by a fictious inertia that obviously is not real. The second term is a stochastic effect that aims to mimic the Brownian interactions that the solute molecules would have with the surrounding solvent. The principle on which this stochastic description is based on is that even if the system is Markovian, knowing the present CG state is not sufficient to correctly predict the future, since many microstates may correspond to the same macrostate. This uncertainty brings also to a statistical description of the equation of the relevant variables in terms of probability density function. Under the Markovian assumption it has the form of the so-called Fokker-Planck equation. The eliminated DOF in Fokker-Planck equation assume the form of both dissipative and thermal fluctuations effects. Dissipation and fluctuations have then the same cause which is the CG procedure of eliminating irrelevant variables and, more importantly, they are correlated by the fluctuation-dissipation theorem. The main goal of the coarse-graining is to develop a suitable form for the Fokker-Planck equation, depending on the specific CG level at which one is aiming to study a given system. For example, in MARTINI CG force field the coarse level is so fine that both dissipation and fluctuation terms are neglected; on the contrary, in much coarser techniques such as dissipative particle dynamics (DPD, Groot and Warren, 1997), they assume a specific expression that allows to account for the relevant phenomena at that coarse scale. The Fokker-Planck equation decouples in a set of equations that describes the motion in a CG system. This leads to the so-called generalised Langevin equation, which, in turn, will assume different forms depending on the particular CG technique (DPD, Brownian dynamics, etc ...).

4.2 Generalised Langevin Equation of Motion

Let us consider a system with N' relevant particles (coarse level) with positions $\mathbf{r}_i'(t)$ and velocities $\mathbf{v}_i'(t)$, with $i = 1,2,\dots,N'$ and N'' irrelevant particles characterised by positions $\mathbf{r}''{}_j(t)$ and velocities $\mathbf{v}''{}_j(t)$ with $j = 1,2,\dots,N''$.

The forces acting on the primed particles can be split in the following terms, in line with the theory explained in Sect. 4.1:

- *systematic (conservative) forces*, $\mathbf{F}_i^C(\mathbf{r}')$, function of primed positions that accounts for pairwise interactions between primed particles, as well as the interactions with the double-primed particles as long as these are related to the primed positions;

- *frictional forces*, $\mathbf{F}_i^D(\mathbf{r}')$, functions of the primed velocities. As for the systematic forces, the frictional ones may include also the dependence on the double-primed velocities, as far as they are related to the primed ones;

- *stochastic (random) forces*, $\mathbf{F}_i^R(\mathbf{r}')$, that accounts for the interactions with the double-primed particles. They are characterised by statistical distributions and may parametrically depend on primed coordinates.

Despite this formalism is general and elegant, two more assumptions are necessary to solve practical problems:

a) $\mathbf{F}_i^C(\mathbf{r}')$ can be written equal to the gradient of the potential in primed coordinates, that is equivalent to infer that systematic forces have no curl. The potential is the *potential of mean force* and it will be explained in the next section;

b) $\mathbf{F}_i^D(\mathbf{r}')$ have linear dependence on primed velocities at earlier times (first order truncation in terms of velocities). "Earlier times" lead to take into account the causality of the motion.

The resulting equation of motion turns out to be the generalised Langevin equation:

$$m_i \frac{d\mathbf{v}_i}{dt} = -\underbrace{\frac{\partial V^{mf}}{\partial \mathbf{r}_i}}_{\mathbf{F}_i^C} - \underbrace{\sum_j \int_0^t \zeta_{ij}(\tau)\mathbf{v}_j(t-\tau)d\tau}_{\mathbf{F}_i^D} + \underbrace{\boldsymbol{\eta}_i(t)}_{\mathbf{F}_i^R} \ , \tag{4.1}$$

where V^{mf} is the potential of mean force (extensively explained in the next section) and $\zeta_{ij}(\tau)$ corresponds to the friction kernel defined only for $\tau \geq 0$. The integration interval in the second term of the right side in Eq. (4.1) can be extended from [0 to finite time, t] to $[0, \infty)$, if t is much greater than the correlation time of the friction kernel.

Therefore, $\mathbf{F}_i^D$ can be seen as a linear description of the velocities derivative, that takes into account the past trajectory of the system.

The last term on the right side of Eq. (4.1) is a random term that considers the stochastic fluctuations in the CG system, which obeys to the following mathematical properties:

$$\langle \boldsymbol{\eta}(t) \rangle = 0 \tag{4.2}$$

$$\langle \mathbf{v}_i(t) \cdot \boldsymbol{\eta}_j(t') \rangle = 0, \text{for any } i, j \text{ and } t' \geq \tau. \tag{4.3}$$

The two last terms of Eq. (4.1) (dissipation and fluctuations) are correlated with each other through the so-called fluctuation-dissipation theorem (shown in sect. 4.4). The generalised Langevin equation presented in Eq. (4.1) is expressed in Cartesian coordinates. Concerning the generalised coordinates $\{\mathbf{q}, \mathbf{p}\}$, the mass corresponds to a tensor and further details are shown in sect. 4.5.

In the following sections there will be presented the potential of mean force (sect. 4.3) together with the fluctuation-dissipation theorem (sect. 4.4), as well as all the approximations that bring to simple Langevin equations (sect. 4.5) and to Brownian dynamics (sect. 4.6). Finally, the Fokker-Planck equation will be shown in sect. 4.7.

4.3 Potential of mean force and superatom approach

In order to reproduce the thermodynamics of the CG system at equilibrium with a suitable accuracy, the partition function of the coarser system must be proportional to the partition function referred to the finer system. This means that, in the canonical ensemble, defining $\omega(\mathbf{q'})$ as the probability density function referred to primed particles, the following proportionality may be established:

$$\omega(\mathbf{q'}) \propto \int_{\Omega_{q''}} e^{-\beta V(\mathbf{q'},\mathbf{q''})} \mathrm{d}\mathbf{q''}, \tag{4.4}$$

where $\Omega_{q''}$ is the integration domain referred to $\mathbf{q''}$ and, as already mentioned, $\beta = (k_B T)^{-1}$.

The potential of mean force (PMF) can be defined as:

$$V^{mf}(\mathbf{q'}) = -k_B T \ln \int_{\Omega_{q''}} e^{-\beta V(\mathbf{q'},\mathbf{q''})} \mathrm{d}\mathbf{q''}. \tag{4.5}$$

Combining the last two equations, the following identity is obtained:

$$\omega(\mathbf{q'}) \mathrm{d}\mathbf{q'} \propto e^{-\beta V(\mathbf{q'},\mathbf{q''})} \mathrm{d}\mathbf{q'}. \tag{4.6}$$

By differentiating the expression reported in Eq. (4.5) with respect to $\mathbf{q'}$, the average force acting on the ensemble of primed particles can be obtained:

$$\nabla_{q'} V^{mf}(\mathbf{q'}) = \frac{\int_{\Omega_{q''}} \nabla_{q'} V(\mathbf{q'}, \mathbf{q''}) \cdot e^{-\beta V(\mathbf{q'}, \mathbf{q''})} \mathrm{d}\mathbf{q''}}{\int_{\Omega_{q''}} e^{-\beta V(\mathbf{q'}, \mathbf{q''})} \mathrm{d}\mathbf{q''}}. \tag{4.7}$$

More specifically, in a system of N particles, the PMF is the potential that gives the mean force acting on a j particle integrated over the $n+1, \ldots, N$ configurations, when the first $1, \ldots, n$ particles are kept fixed (Kirkwood, 1935).

4.3.1 Superatom technique

One of the several ways to coarse-grain a system is represented by the so-called superatom approach. This means that a given number of atoms are grouped together in a unique particle called superatom (or bead). This technique turns out to be useful in long-chain systems like lipids, proteins or polymers, representing a reasonable trade-off between computational cost and accuracy. Following the superatom approach, few techniques have been developed (as already discussed in the introduction of this chapter), but some of them may neglect the two additional terms (friction and stochastic fluctuations) of the Langevin equation. This leads to an artificially accelerated system, and this is why this exclusion works only for low number of particles mapping (namely, low number of atoms grouped together in a superatom or bead).

The interactions between two close beads in a chain are usually described through soft harmonic potentials, which corresponds to describe the distance evolution of the beads as a Gaussian distribution. Several approaches have been developed following this technique; the most common ones are represented by the Finitely Extendable Nonlinear Elastic (FENE) chain model, which limits the chain extension to a finite upper limit. Some stochastic applications of FENE models have also been developed (Fan et al., 2003).

Another model is the one developed by Nielsen et al. (2003) that specifically reproduces targeted compounds, by using a three non-hydrogen atoms mapping together with soft harmonic bond and angle potentials. This model is based on a parametrization of surface tension and density of the compounds by reproducing the end-to-end distance distributions from atomistic simulations. A last example of a very well-established and employed tool is the MARTINI model (Marrink, Risselada, et al., 2007), which lies on a four-to-one mapping and that will be extensively described in the second part of this work.

4.4 Fluctuation-dissipation theorem

In this section the correlation between the last two terms (friction and noise) of the Langevin equation is presented. Let us consider the long-time range contribution of the kinetic energy $\mathscr{K} = \sum_i 1/2 m_i {v_i}^2$. While the conservative force, $\mathbf{F}^C$, keeps the total energy (potential + kinetic) constant, the friction (or dissipative) force, $\mathbf{F}^D$, lowers $\mathscr{K}$

and the noise (or stochastic) term, $\mathbf{F}^R$, increases it. Let us consider now the velocity correlation function, expressed as $\langle \mathbf{v}_i(t)\mathbf{v}_j(t+\tau)\rangle$. For large values of t it becomes independent on t itself. The thermodynamic quantity $\langle \mathbf{v}_i(t)\mathbf{v}_j(t)\rangle$ must obey the so-called equipartition theorem:

$$\langle v_{i\alpha} v_{j\beta}\rangle = \frac{k_B T}{m_i}\delta_{ij}\delta_{\alpha\beta}. \tag{4.8}$$

This is equivalent to infer that if a process always starts with the identical initial conditions at $t = 0$, and several simulations are performed, after a suitably long simulation time, the average velocities correlation functions must fulfil the equipartition theorem.

A more general formulation for the fluctuation-dissipation theorem is the one proposed by Kubo (1966), making a distinction between a first and second fluctuation-dissipation theorem. In order to show them, let us define a single velocity $v(t)$ for a pure Langevin equation (neglecting the systematic forces):

$$m\dot{v}(t) = -\int_0^t \zeta(t)v(t-\tau)\mathrm{d}\tau + \eta(t) + F^{ex}(t), \tag{4.9}$$

where $F^{ex}(t)$ is an external force necessary to measure the linear response of the system. Applying an external impulse-force: $F^{ex}(t) = mv_0\delta(t)$ where $\delta(t)$ is the Dirac delta function, the ensemble-averaged velocity $\langle v(t)\rangle$ will correspond to the response: $v_0\phi(t)$. This means that starting from random initial conditions and unperturbed equilibrium distribution, it is true that $\langle v(t)\rangle = v_0\phi(t)$ and, for independent realizations of the noise, the first fluctuation-dissipation theorem corresponds to:

$$\phi(t) = \frac{\langle v(t_0)v(t_0+t)\rangle}{\langle v^2\rangle}, \tag{4.10}$$

where the correlation functions are stationary (independent on t_0) and ensemble averaged. The proof of the correlation reported in Eq. (4.10) is extensively demonstrated in Berendsen (2007).

The second fluctuation-dissipation theorem states that:

$$\langle \eta(t_0)\eta(t_0+t)\rangle = m\langle v^2\rangle\zeta(t) = k_B T\zeta(t). \tag{4.11}$$

In the case of Langevin equation without considering the systematic forces (pure Langevin equation), the expression reported in Eq. (4.11) can be rigorously demonstrated (Kubo et al., 1985). As it happens in all the realistic MD simulations, systematic forces are considered and the theorem cannot be generally proven anymore. Several particular cases have been considered and validated in the latter scenario (Bossis et al., 1982; Ciccotti and Ryckaert, 1981; McDowell, 2000). In a more general way, Berendsen (2007) recommends to use time-independent friction kernels when the systematic forces relaxation time is of the same order as the friction characteristic times. In this case, the system is called Markovian and the characteristic equation is labelled as Markovian-Langevin equation. In a Markovian process, the changes in $\mathscr{K}$ due to friction (decrease) and noise (increase) are shown to be:

$$\left(\frac{d\mathscr{K}}{dt}\right)_{friction} = mv\dot{v} = -\zeta v^2 = -\frac{2\zeta}{m}\mathscr{K}, \tag{4.12}$$

$$\left\langle\frac{d\mathscr{K}}{dt}\right\rangle_{noise} = \frac{A_\eta}{2m}, \tag{4.13}$$

where A_η is the intensity if a δ-noise-force, $A_\eta\delta(t)$. The balance of Eq. (4.12) and Eq. (4.13) leads to $\langle\mathscr{K}\rangle = A_\eta/(4\zeta)$. The stationary equilibrium at temperature T_0 is shown to satisfy

$$A_\eta = 2\zeta k_B T_0, \tag{4.14}$$

for one degree of freedom (one dimension system). If the temperature deviates from the reference value T_0, then the decay will be exponential with respect to T_0 itself. In other words,

$$\frac{dT}{dt} = -\frac{2\zeta}{m}(T - T_0), \tag{4.15}$$

where $m/2\zeta$ is the characteristic time constant of the decay. The physical meaning of all the theory just presented above is the following: velocity fluctuations are stabilised thanks to the presence of friction and noise terms; moreover, the increase and decrease of kinetic energy, due to noise and friction respectively, can be thought as heat exchange with a bath at a reference temperature T_0 (independent on both systematic forces and time, as well as time-dependence of velocity autocorrelation functions).

In the case of non-Markovian systems, the friction decrease reads as follows

$$\left\langle\frac{d\mathscr{K}}{dt}\right\rangle_{friction} = m\langle v(t)\dot{v}(t)\rangle = -\int_0^\infty \zeta(t)\langle v(t)v(t-\tau)\rangle d\tau, \tag{4.16}$$

depending on both velocity autocorrelation function and the systematic (conservative) forces.

4.5 Langevin equation in generalised coordinates

Let us define a system with N particles identified in the cartesian coordinates by $\mathbf{r}$. The generalised coordinates $\mathbf{q}$ are defined in such a way that $\mathbf{r}_i = \mathbf{r}_i(q_1, .., q_n)$, with $i = 1, \ldots, N$ and $n = 3N$. Considering also the generalised momenta $\mathbf{p}$, the set of generalised variables is $\mathbf{z} = \{\mathbf{q}, \mathbf{p}\}$. The mass is expressed by the so-called mass tensor $\mathbf{M}$. In order to find it out, let us consider the expression of kinetic energy $\mathscr{K}$, in terms of generalised coordinates $\mathbf{q}$:

$$\mathscr{K} = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\mathbf{r}}^2 = \sum_{k,l=1}^{n} \sum_{i=1}^{N} m_i \frac{\partial \mathbf{r}_i}{\partial q_k} \frac{\partial \mathbf{r}_i}{\partial q_l} \dot{q}_k \dot{q}_l, \tag{4.17}$$

where:

$$\mathbf{M} = M_{kl} = \sum_{i=1}^{N} m_i \frac{\partial \mathbf{r}_i}{\partial q_k} \frac{\partial \mathbf{r}_i}{\partial q_l}. \tag{4.18}$$

The mass tensor $\mathbf{M}$ is symmetric and invertible (det $\mathbf{M} \neq 0$) and its eigenvalues are the masses m_i. It is correlated to the generalised coordinates through the following relation:

$$\mathbf{p} = \mathbf{M}\dot{\mathbf{q}}, \tag{4.19}$$

or, equally

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{p}. \tag{4.20}$$

Labelling the relevant coordinates as $\mathbf{q}'$ and the irrelevant ones with $\mathbf{q}''$, the mass tensor can be decomposed as

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y}^T & \mathbf{Z} \end{pmatrix}. \tag{4.21}$$

Consequently,

$$(\dot{\mathbf{q}}' \;\; \dot{\mathbf{q}}'') = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y}^T & \mathbf{Z} \end{pmatrix} \begin{pmatrix} \mathbf{p}' \\ \mathbf{p}'' \end{pmatrix}. \tag{4.22}$$

From Eq. (4.22) the two following relations hold:

$$\dot{\mathbf{q}}' = \langle \mathbf{X} \rangle_{''} \mathbf{p}' + \langle \mathbf{Y}\mathbf{p}'' \rangle_{''}, \tag{4.23}$$

$$\dot{\mathbf{p}}' = -\left\langle \frac{\partial V}{\partial \mathbf{q}'} \right\rangle_{''} + \text{friction} + \text{noise}, \tag{4.24}$$

where $\langle\rangle_{''}$ denotes a canonical distribution average. On the right side of Eq. (4.24) the three terms are similar to the ones seen in the Langevin equation in cartesian coordinates. However, on the left hand side of Eq. (4.24) the expression is different, in line with the matrix notation, yielding:

$$\frac{d}{dt} \langle \mathbf{X} \rangle_{''}^{-1} \mathbf{v}(t) = -\frac{\partial V^{mf}}{\partial \mathbf{q}} - \int_0^t \boldsymbol{\zeta}(\tau) \mathbf{v}(t-\tau) d\tau + \boldsymbol{\eta}(t). \tag{4.25}$$

Berendsen (2007) and Fixman (1974) extensively show that, in some specific cases, if the mass tensor is independent on time-dependent variables, then the generalised Langevin equations can be written down as follows:

$$\mathbf{M}\dot{\mathbf{v}} = \mathbf{F}^C(\mathbf{q}') - \int_0^t \boldsymbol{\zeta}(\tau)\mathbf{v}(t-\tau)d\tau + \boldsymbol{\eta}(t), \tag{4.26}$$

$$\mathbf{v} = \dot{\mathbf{q}}, \tag{4.27}$$

$$\mathbf{M} = \mathbf{X}^{-1}, \tag{4.28}$$

where $\mathbf{M}$ is the mass tensor, $X_{kl} = \sum_i \left(1/m_i \cdot \partial q''_k/\partial \mathbf{r}_i \cdot \partial q''_l/\partial \mathbf{r}_i\right)$ and $\mathbf{F}^C(\mathbf{q}')$ are the systematic (conservative) forces.

4.6 Brownian dynamics

When the conservative forces are slow, namely when they are pretty much constant with the time scale $\tau_C = m/\zeta$ related to the velocity correlation function, the Langevin equation can be averaged on a time step $\Delta t > \tau_C$. The average over the left hand side term of Eq. 4.26 yields a small quantity that can be neglected. Therefore, a non-inertial Langevin dynamics is obtained:

$$0 \approx F_i[\mathbf{q}(t)] - \sum_j \zeta_{ij} v_j(t) + \eta_i(t), \tag{4.29}$$

which corresponds to the following matrix notation:

$$\boldsymbol{\zeta}\mathbf{v} = \mathbf{F} + \boldsymbol{\eta}(t). \tag{4.30}$$

The Brownian equation related to the velocities becomes:

$$\mathbf{v} = \dot{\mathbf{q}} = \boldsymbol{\zeta}^{-1}\mathbf{F} + \boldsymbol{\zeta}^{-1}\mathbf{B}\boldsymbol{\eta}^0(t), \tag{4.31}$$

where $\mathbf{B}$ and $\boldsymbol{\eta}^0(t)$ are quantities that are defined so that:

$$\mathbf{B}\mathbf{B}^T = 2\boldsymbol{\zeta} k_B T, \tag{4.32}$$

$$\langle \boldsymbol{\eta}^0(t) \rangle = 0, \tag{4.33}$$

$$\langle \boldsymbol{\eta}^0(t_0)\boldsymbol{\eta}^0(t_0+t) \rangle = \mathbf{I}\delta(t). \tag{4.34}$$

In the context of MD simulations, starting from Eq. (4.31), the positions $\mathbf{q}$ can be updated through the Euler step and, in components notation, the equation becomes:

$$q_i(t+\Delta t) = q_i(t) + \frac{\Gamma_i}{k_B T} F_i(t)\Delta t + \xi \tag{4.35}$$

where $\Gamma_i = k_B T/\zeta_{ii}$ and ξ is a random number (from a probability distribution) so that:

$$\langle \xi \rangle = 0, \quad \langle \xi^2 \rangle = 2\Gamma\Delta t. \tag{4.36}$$

4.7 Fokker-Planck equation

As in classical atomistic MD the Liouville equation is used to describe the evolution of the density function in the phase space, here in CG systems the equivalent approach can be expressed by the so-called Fokker-Planck equation (FPE). The two approaches are similar in terms of statistical mechanics, both of them describing the evolution of density functions and directly related to the corresponding dynamics equations of motion (Hamiltonian equation for MD and Langevin equation for CG systems). The main difference lies on the stochastic nature that Fokker-Planck equations take into account.

The general FPE proposed (Risken and Frank, 1989; van Kampen, 1981) in matrix notation reads as follows:

$$\frac{\partial f_{N_p}}{\partial t} = -\nabla_x^T(\mathbf{a} f_{N_p}) + \frac{1}{2}\mathrm{tr}\,(\nabla_x \nabla_x^T \mathbf{B}\mathbf{B}^T f_{N_p}), \tag{4.37}$$

where the first and the second terms on the right hand side correspond respectively to a drift and a diffusion contribution. Applying Eq. (4.37) to the generalised Langevin dynamics, assuming **B** constant, the following relation holds:

$$\begin{aligned}\frac{\partial f_{N_p}}{\partial t} = -\mathbf{v}^T\nabla_q f_{N_p} - \mathbf{F}^T\mathbf{M}^{-1}\nabla_v f_{N_p} + \mathrm{tr}\,(\mathbf{M}^{-1}\boldsymbol{\zeta}) f_{N_p} + \\ + \mathbf{v}^T\boldsymbol{\zeta}\mathbf{M}^{-1}\nabla_v f_{N_p} + \frac{1}{2}\mathrm{tr}\,(\mathbf{M}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{M}^{-1}\nabla_q\nabla_q f_{N_p}).\end{aligned} \tag{4.38}$$

Reducing Eq. (4.38) to the one-dimensional case, the so-called Kramers equation (Kramers, 1940) is obtained:

$$\frac{\partial \rho}{\partial t} = -v\frac{\partial \rho}{\partial q} - \frac{F}{m}\frac{\partial \rho}{\partial v} + \frac{\rho v}{m}\frac{\partial \rho}{\partial v} + \frac{\zeta}{m}\rho + \frac{\zeta k_B T}{m^2}\frac{\partial^2 \rho}{\partial v^2}. \tag{4.39}$$

The corresponding equilibrium distribution is shown to be:

$$f_{N_p}^{eq}(q, v) \propto \exp\left[-\frac{mv^2}{2k_B T}\right] \exp\left[-\frac{V(q)}{k_B T}\right], \tag{4.40}$$

where $V(q)$ is the potential related to the force, $F = -dV(q)/dq$.

Chapter 5

Theory of Continuum Modeling

5.1 From Atomistic to Continuum Modelling

In this section, continuum models are derived starting from one of the main results shown in atomistic modelling, namely the Liouville equation. The aim of continuum modelling is to describe the evolution of continuous properties, such as number density functions (NDFs) and other macroscopic variables (density, momentum, etc.). In the continuum modelling context, NDFs represent group of molecules that have in common the same infinitesimal physical and phase space that evolves over time, without tracking the trajectory of single particles, as it is done in atomistic modelling. Passing then through Boltzmann equation and by using the moments of the NDF, continuity, Navier-Stokes and energy transport equations are derived. Further details concerning importance and modelling properties of number density functions are reported in Appendix A. Turbulence modelling is exploited, with particular attention to Reynolds- and Favre-averaged Navier-Stokes equation. Population balance modelling is finally introduced in both laminar and turbulent conditions, focusing on the functional form of PBEs source term and, more specifically, about aggregation kernels.

5.1.1 From Liouville to Boltzmann Equation

Starting from the Liouville equation reported in Eq. (3.13), referred to the probability density function f_{N_p} for a system of N_p particles, and developing explicitly the Liouville operator iLf_{N_p}, Liouville equation in generalised coordinates can be rearranged as follows:

$$\frac{\mathrm{d}f_{N_p}}{\mathrm{d}t} = \frac{\partial f_{N_p}}{\partial t} + \dot{\mathbf{q}} \cdot \nabla_{\mathbf{q}} f_{N_p} + \dot{\mathbf{p}} \cdot \nabla_{\mathbf{p}} f_{N_p} = 0. \tag{5.1}$$

Let pass from generalised coordinates $\mathbf{q}$ and $\mathbf{p}$ to Cartesian coordinates $\mathbf{x} = (x, y, z)$. Each i^{th} particle is characterised by a position vector $\mathbf{x}_i = (x_i, y_i, z_i)$ and velocity $\mathbf{v}_i =$

(v_{ix}, v_{iy}, v_{iz}). The momenta can be therefore expressed by the relation $\mathbf{p} = m\mathbf{v}$. Eq. (5.1) becomes:

$$\frac{\partial f_{N_p}}{\partial t} + \dot{\mathbf{x}} \cdot \nabla_{\mathbf{x}} f_{N_p} + \dot{\mathbf{v}} \cdot \nabla_{\mathbf{v}} f_{N_p} = 0, \tag{5.2}$$

where the notation $\nabla_{\mathbf{x}}$ stands for the partial derivative vector with respect to position $\mathbf{x}$, that is $\nabla_{\mathbf{x}} = \partial/\partial\mathbf{x} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$. The same notation is obviously valid for $\nabla_{\mathbf{v}} = \partial/\partial\mathbf{v} = (\partial/\partial v_x, \partial/\partial v_y, \partial/\partial v_z)$ and from now on these two equivalent expressions will be used interchangeably.

The acceleration term $\dot{\mathbf{v}}$ can be furthermore decomposed into two different contributions

$$\dot{\mathbf{p}} = m\dot{\mathbf{v}} = \mathbf{F} + m\mathbf{g}, \tag{5.3}$$

that are molecular accelerations due to force fields ($\mathbf{g}$) and to discrete events of molecular collisions ($\mathbf{F}$). The Liouville equation can be therefore written down accounting explicitly for all the i^{th} particles contribution (summation over i) in the following way:

$$\frac{\partial f_{N_p}}{\partial t} + \sum_{i=1}^{N_p} \left(\mathbf{v}_i \cdot \frac{\partial f_{N_p}}{\partial \mathbf{x}_i} \right) + \mathbf{g} \cdot \left(\sum_{i=1}^{N_p} \frac{\partial f_{N_p}}{\partial \mathbf{v}_i} \right) = -\sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \frac{\mathbf{F}(\mathbf{r}_{i,j})}{m} \cdot \frac{\partial f_{N_p}}{\partial \mathbf{v}_i}, \tag{5.4}$$

where the third term on left hand side is a molecular acceleration term in velocity space and the term on right hand side corresponds to discrete events due to collisions/molecular interactions, between generic particles i and j that have a distance vector $\mathbf{r}_{i,j}$. Eq. (5.4) represents an alternative form of Eq. (3.13), always fully deterministic and time reversible but expressed in terms of Cartesian coordinates.

Now, let us introduce the number density function, $f_N(t, \mathbf{x}, \mathbf{v}) = \sum_i \langle \delta(\mathbf{x}_i - \mathbf{x}) \delta(\mathbf{v}_i - \mathbf{v}) \rangle$ which accounts for the number of molecules that are in the same infinitesimal configuration, in terms of position $\mathbf{x}$ and velocity $\mathbf{v}$. Therefore, from now on, the NDF will identify a group of molecules, of which the single trajectories have been averaged. The brackets $\langle \cdot \rangle$ identify this average over all the possibles states of the N_p particles system, namely $\langle \cdot \rangle = \int \cdot f_{N_p} d\mathbf{x} d\mathbf{v}$, where $\mathbf{x} = \{\mathbf{x}_1, ..., \mathbf{x}_{N_p}\}$ and $\mathbf{v} = \{\mathbf{v}_1, ..., \mathbf{v}_{N_p}\}$. By integrating Eq. (5.4) over all the molecular variables except for i^{th} molecule and summing over all possible i, the Boltzmann equation comes out (indistinguishable molecules with equal mass m):

$$\frac{\partial f_N}{\partial t} + \mathbf{v} \cdot \frac{\partial f_N}{\partial \mathbf{x}} + \mathbf{g} \cdot \frac{\partial f_N}{\partial \mathbf{v}} = \mathbb{C}. \tag{5.5}$$

The term $\mathbb{C}$ is the collision integral and accounts for all molecular interactions (collisions, discrete events). It usually has a very complicated functional form, especially due to the highly non-linear intermolecular potentials. It is worth mentioning some interesting properties about the term $\mathbb{C}$: the collision integral applied to mass, momentum and energy is always null, stating the conservation of those properties. More generally, it is possible to infer that at equilibrium the collision term is null and the distribution of the NDF tends to a Maxwell-Boltzmann distribution (Gaussian). The Maxwell-Boltzmann distribution is valid in case of Knudsen number much smaller than unity (many collisions occurs) and especially when the number of states available is much greater than the number of molecules; otherwise, a quantum approach is needed. Independently on the complexity of the interaction potential employed in the collision term, every system at equilibrium will always evolve following the Maxwell-Boltzmann distribution.

However, there are many cases in which, for a given range of intermediate Knudsen number ($0.1 < Kn < 100$), the system is out of equilibrium, it cannot be treated as a continuum, and the collision integral is therefore non-zero. In this case, suitable closures are needed; one of the most famous example is represented by the hard-spheres potential, in which the so-called Boltzmann Stosszahlansatz approximation is done (indistinguishable particles, two-particles PDF equal to the product of one-particles PDF, molecular chaos) in order to close the collision term.

5.1.2 From Boltzmann to Continuity, Navier-Stokes and Energy equations

In this section the main equations of change for continuum systems will be presented, starting from Boltzmann equation (5.5). In a preliminary way, let us take a NDF $f_N(t, \mathbf{x}, \mathbf{v})$ and define the generic $\mathbf{k}^{th}$ order moment of a distribution:

$$m^{(\mathbf{k})} = \int_{\Omega_v} \mathbf{v}^{\mathbf{k}} f_N(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \tag{5.6}$$

where Ω_v is the velocity phase space domain.

The 0^{th}-order moment corresponds to the total number, N_t of entities (disperse phase, e.g., bubbles, particles, molecules, ...) in the mesoscale system, per unit volume:

$$m^{(0,0,0)} = \int_{\Omega_v} f_N(t, \mathbf{x}, \mathbf{v}) d\mathbf{v} = N_t. \tag{5.7}$$

The 1^{th}-order moment divided by 0^{th} one corresponds to the mean velocity in number. Therefore, the following equality holds:

$$m^{(1,0,0)} = \int_{\Omega_v} \mathbf{v} f_N(t, \mathbf{x}, \mathbf{v}) d\mathbf{v} = N_t(t, \mathbf{x}) \cdot U_x(t, \mathbf{x}), \tag{5.8}$$

where $\mathbf{U}(t, \mathbf{x}) = (U_x, U_y, U_z)$ is the number-averaged velocity. It is straightforward that $U_x = m^{(1,0,0)}/N_t$, $U_y = m^{(0,1,0)}/N_t$ and $U_z = m^{(0,0,1)}/N_t$.

The mass density (total mass of molecules per unit volume) is

$$\rho(t, \mathbf{x}) = m \int_{\Omega_v} f_N(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \tag{5.9}$$

where m is the molecule mass.

The momentum density is defined as follows:

$$\rho \mathbf{U} = m \int_{\Omega_v} \mathbf{v} f_N(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}. \tag{5.10}$$

Let us take now the Boltzmann equation (5.5). Let us multiply it by $m d\mathbf{v}$ and integrate it out in Ω_v (0^{th}-order moment),

$$\int_{\Omega_v} \left(\frac{\partial f_N}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{v} f_N) + \frac{\partial}{\partial \mathbf{v}} \cdot (\mathbf{g} f_N) = \mathbb{C} \right) m d\mathbf{v}. \tag{5.11}$$

By applying the definitions reported in Eq. (5.7) and Eq. (5.9), the continuity equation comes out:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{U}) = 0, \tag{5.12}$$

where the collisional term $\mathbb{C}$ is null if integrated over all the velocity phase space domain at equilibrium (Maxwell-Boltzmann distribution) and the acceleration term results zero by integrating it out by parts (the NDF is null at phase space domain border $\partial \Omega_v$).

Following the same procedure, let us apply the 1^{st}-order moment (Eq. 5.8) to Eq. (5.5):

$$\int_{\Omega_v} \left(\frac{\partial f_N}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{v} f_N) + \frac{\partial}{\partial \mathbf{v}} \cdot (\mathbf{g} f_N) = \mathbb{C} \right) m \mathbf{v} d\mathbf{v}. \tag{5.13}$$

Introducing the celerity $\mathbf{c} = \mathbf{v} - \mathbf{U}$ the following equality trivially holds:

$$\int f_N(\mathbf{v}) \mathbf{c} d\mathbf{v} = \int f_N(\mathbf{c}) \mathbf{c} d\mathbf{c} = 0. \tag{5.14}$$

After some trivial calculations and by using the definitions given in Eq. (5.8) and Eq. (5.10), Eq. (5.13) results in:

$$\frac{\partial}{\partial t}(\rho \mathbf{U}) + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{U}\mathbf{U}) + \frac{\partial}{\partial \mathbf{x}} \cdot \left[m \int_{\Omega_v} \mathbf{c}\mathbf{c} f_N(\mathbf{c}) d\mathbf{c} \right] - \mathbf{g} m \int_{\Omega_v} f_N(\mathbf{v}) d\mathbf{v} = 0, \tag{5.15}$$

where the quantity $\mathbf{aa}$, with $\mathbf{a}$ generic tensor, represents a generic tensor product. By defining the momentum flux tensor:

$$\mathbf{\Pi} = m \int_{\Omega_v} \mathbf{c}\mathbf{c} f_N(\mathbf{c}) d\mathbf{c} = -\mathbf{\Sigma}, \tag{5.16}$$

opposite of the viscous stress tensor $\mathbf{\Sigma}$, the following equation is obtained:

$$\frac{\partial}{\partial t}(\rho \mathbf{U}) + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{U}\mathbf{U}) = \rho \mathbf{g} - \frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{\Pi}. \tag{5.17}$$

By decomposing now the quantity $\mathbf{\Pi} = \boldsymbol{\tau} + p\mathbf{I}$, where $p = 1/3tr(\mathbf{\Pi})$ and $\boldsymbol{\tau}$ is a traceless stress tensor, the equation of motion becomes:

$$\frac{\partial}{\partial t}(\rho \mathbf{U}) + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{U}\mathbf{U}) = -\frac{\partial}{\partial \mathbf{x}} p + \rho \mathbf{g} - \frac{\partial}{\partial \mathbf{x}} \cdot \boldsymbol{\tau}. \tag{5.18}$$

Having a closer look at Eq. (5.23), it turns out that Eq. (5.23) is unclosed due to the viscous term $\boldsymbol{\tau}$. For Newtonian (constant viscosity μ) and incompressible fluids (ρ constant), it can be closed with the expression $\boldsymbol{\tau} = -\mu(\nabla \mathbf{v} + \nabla \mathbf{v}^T)$ (∇ coincides with the expression $\partial/\partial \mathbf{x}$), yielding the very famous so-called Navier-Stokes equation:

$$\rho \frac{\partial}{\partial t}(\mathbf{U}) + \mathbf{U} \cdot \frac{\partial}{\partial \mathbf{x}}(\rho \mathbf{U}) = -\frac{\partial}{\partial \mathbf{x}} p + \rho \mathbf{g} + \mu \frac{\partial^2}{\partial \mathbf{x}^2} \mathbf{v} \tag{5.19}$$

or, equally:

$$\rho \frac{\partial}{\partial t}(\mathbf{U}) + \mathbf{U} \cdot \nabla(\rho \mathbf{U}) = -\nabla p + \rho \mathbf{g} + \mu \nabla^2 \mathbf{v}. \tag{5.20}$$

Following again the same approach, by applying the 2^{nd}-order moment transform to Boltzmann equation, the mechanical energy equation can be obtained. Let us define the internal energy, $\rho\hat{U}$, and kinetic energy, $\rho\hat{K}$:

$$\rho\hat{U} = \int_{\Omega_v} \left(\frac{1}{2} m \mathbf{c}^2\right) f_N(\mathbf{v}) d\mathbf{v} = \frac{m}{2} \int_{\Omega_v} \mathbf{c}^2 d\mathbf{c}, \tag{5.21}$$

$$\rho\hat{K} = \int_{\Omega_v} \left(\frac{1}{2} m \mathbf{U}^2\right) f_N(\mathbf{v}) d\mathbf{v} = \frac{m\mathbf{U}^2}{2} \int_{\Omega_c} d\mathbf{c} = \rho \frac{\mathbf{U}^2}{2}. \tag{5.22}$$

By multiplying Eq. (5.5) by the quantity $1/2m\mathbf{v}^2 d\mathbf{v}$ and integrating out over $\mathbf{\Omega}_v$ (remember the equality $\mathbf{v} = \mathbf{U} + \mathbf{c}$) the energy balance equation referred to the energy per unit mass $\hat{E} = \hat{K} + \hat{U}$ is obtained:

$$\frac{\partial}{\partial t}(\rho \hat{E}) + \frac{\partial}{\partial \mathbf{x}} \cdot (\rho \mathbf{U} \hat{E}) = -\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{J}_q + \rho \mathbf{g} \cdot \mathbf{U} - \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{\Pi} \cdot \mathbf{U}), \tag{5.23}$$

where the three terms on right hand side correspond respectively to conduction, work due to gravity and work due to viscous stresses.

5.2 Turbulence Modelling

5.2.1 Reynolds decomposition and RANS equations

When the flow becomes turbulent, a proper mathematical description must be introduced since turbulence affects heat and mass transfer, as well as aggregation and breakage phenomena. Being turbulent flows chaotic, random and irregular, a statistical description of the evolution of instant variables is needed. More than 100 years ago, Reynolds introduced a statistical averaging method in order to simplify the description of the fluctuating properties in turbulent flows. He proposed to split the instantaneous property $\phi(t, \mathbf{x})$ that fluctuates in time t and physical space $\mathbf{x}$ into two contributions: an average contribution plus the fluctuation of the given property. This is the very well-known Reynolds decomposition:

$$\phi = \bar{\phi} + \phi', \tag{5.24}$$

where ϕ' is the fluctuating component and $\bar{\phi}$ represents the average, also known as Reynolds average:

$$\bar{\phi}(t, \mathbf{x}) = \frac{1}{\tau} \int_{t}^{t+\tau} \phi(\mathbf{x}, \tilde{t}) d\tilde{t}, \tag{5.25}$$

representing the average of ϕ in a time interval τ. Thanks to Reynolds decomposition, shown here in Eq. (5.24), a statistical description of the flow field can be obtained together with the evolution of turbulent properties. More specifically, let us decompose the instantaneous fluid velocity in the following way: $\mathbf{U} = \overline{\mathbf{U}} + \mathbf{u}'$. Applying Reynolds decomposition approach to continuity and Navier-Stokes equations (the body forces $\mathbf{g}$ are here neglected) already presented in this chapter respectively in Eq. (5.12) and Eq. (5.19), the so-called Reynolds-averaged Navier-Stokes (RANS) equations are obtained. The whole mathematical derivation is here omitted (the reader can find it in Andersson et al., 2012) and only the final result is shown. RANS equations read as follows:

$$\frac{\partial \overline{\mathbf{U}}}{\partial t} + \overline{\mathbf{U}} \cdot \frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}} = -\frac{1}{\rho} \frac{\partial}{\partial \mathbf{x}} \cdot \left[\bar{P} \mathbf{I} + \mu \left(\frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}} + \frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}}^{T} \right) - \rho \overline{\mathbf{u}'\mathbf{u}'} \right], \tag{5.26}$$

where the term $-\rho \overline{\mathbf{u}'\mathbf{u}'}$ is labelled as Reynolds stresses, which represents an unclosed term. The Boussinesq approximation is therefore adopted in order to close the Reynolds stress tensor, in terms of fluid mean velocity field. It is based on the hypothesis that the Reynolds stress tensor is proportional to the mean velocity gradients:

$$-\rho \overline{\mathbf{u}'\mathbf{u}'} = \mu_t \left(\frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}} + \frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}}^{T} \right) - \frac{2}{3} \rho k \mathbf{I}, \tag{5.27}$$

where μ_t is the turbulent eddy viscosity (it is not a physical quantity, but rather a numerical term that comes out from the closure) and the turbulent kinetic energy $k = \frac{1}{2}\overline{\mathbf{u}'\mathbf{u}'}$ is introduced.

When density fluctuations are important (e.g., compressible gas flows, two-liquids flows), a density-average, or Favre-average (Favre, 1965), must be introduced, strictly related to Reynolds average:

$$\langle \phi(t,\mathbf{x}) \rangle = \frac{\int \rho(t,\mathbf{x})\phi(t,\mathbf{x})\mathrm{d}t}{\int \rho(t,\mathbf{x})\mathrm{d}t} = \frac{\overline{\rho\phi}}{\bar{\rho}}, \tag{5.28}$$

where $\langle\cdot\rangle$ represents the Favre average. Following the same procedure and the same closure as for the RANS equations, the FANS equation is easily obtained:

$$\begin{aligned} \frac{\partial \bar{\rho}\langle \mathbf{U}\rangle}{\partial t} + \frac{\partial \bar{\rho}\langle \mathbf{U}\rangle\langle \mathbf{U}\rangle}{\partial \mathbf{x}} = -\frac{\partial}{\partial \mathbf{x}} \cdot \Bigg[\langle P\rangle\ \mathbf{I} + \\ + \left(\mu + \mu_t\right)\left(\frac{\partial\langle \mathbf{U}\rangle}{\partial \mathbf{x}} + \frac{\partial\langle \mathbf{U}\rangle}{\partial \mathbf{x}}^T - \frac{2}{3}\frac{\partial}{\partial \mathbf{x}}\cdot\langle \mathbf{U}\rangle\mathbf{I}\right) - \frac{2}{3}\bar{\rho}k\mathbf{I}\Bigg], \end{aligned} \tag{5.29}$$

where the divergence of the velocity $\partial/\partial\mathbf{x}\cdot\mathbf{U}$ is also taken into account, showing the general case of compressible fluid.

As said above, the turbulent viscosity μ_t, in contrast with molecular viscosity, is not a physical property, but a mathematical artifact of the turbulence model. The mathematical closure procedure is now moved to this term, that is still unclosed. In general, turbulent eddy viscosity ν_t is proportional to the large-scale eddy velocity u and its characteristic length scale l. Depending on the kind of models adopted to close it, several turbulence models have been developed and classified in three main categories: zero-, one- and two-equation models. Here one of the two-equation models will be presented, being the turbulence model adopted in the CFD simulations in the current work. This is the $k-\varepsilon$ standard and will be presented in the next section.

5.2.2 $k-\epsilon$ standard turbulence model

The $k-\varepsilon$ standard is one of the most employed turbulence models in CFD calculations, thanks to its simplicity and to its direct link to key turbulent properties, such as the turbulent dissipation rate ε, also necessary to close k-equation itself. Therefore, two transport equations for k and ε are solved:

$$\frac{\partial k}{\partial t} + \overline{\mathbf{U}}\cdot\frac{\partial k}{\partial \mathbf{x}} = \nu_t\left[\left(\frac{\partial\overline{\mathbf{U}}}{\partial \mathbf{x}} + \frac{\partial\overline{\mathbf{U}}}{\partial \mathbf{x}}^T\right)\cdot\frac{\partial\overline{\mathbf{U}}}{\partial \mathbf{x}}\right] - \varepsilon + \frac{\partial}{\partial \mathbf{x}}\cdot\left[\left(\nu + \frac{\nu_t}{\sigma_k}\right)\frac{\partial k}{\partial \mathbf{x}}\right], \tag{5.30}$$

$$\begin{aligned}\frac{\partial \varepsilon}{\partial t}+\overline{\mathbf{U}}\cdot\frac{\partial \varepsilon}{\partial \mathbf{x}}=C_{\varepsilon 1}\nu_t\frac{\varepsilon}{k}\left[\left(\frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}}+\frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}}^T\right)\cdot\frac{\partial \overline{\mathbf{U}}}{\partial \mathbf{x}}\right] \\ -C_{\varepsilon 2}\frac{\varepsilon^2}{k}+\frac{\partial}{\partial \mathbf{x}}\cdot\left[\left(\nu+\frac{\nu_t}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial \mathbf{x}}\right],\end{aligned} \tag{5.31}$$

where $\nu_t=\frac{\mu_t}{\rho}=C_\mu\frac{k^2}{\varepsilon}$, $C_\mu=0.09$, $C_{\varepsilon 1}=1.44$, $C_{\varepsilon 2}=1.92$, $\sigma_k=1.00$ and $\sigma_\varepsilon=1.30$.

5.3 Population Balance Modelling

5.3.1 PBE and GPBE

The aim of this section is to write down the kinetic equation related to NDF in terms of PBE. In a preliminary way, it is worthwhile to stress again that a NDF, as defined above, is an average quantity of the disperse phase based on the atomistic possible states (look up in section 5.1.1 at the definition of NDF). However, at the clusters scale, this information (stochastic behaviour and trajectories of the single particles) is lost and the evolution of the entire population is considered (in an Eulerian approach), by trasporting a fully deterministic equation (PBE or GPBE) as long as the population is made by a large number of particles.

The last definition that is noteworthy to introduce is the distinction between internal and external coordinates. The former consist in something that can be classified as an intimate property (internal) of the disperse phase (particle velocity, concentration, particle size). On the other hand, the external coordinates are those that do not belong directly to the particles; for example, physical space represents a set of external coordinates.

Adopting the same approach already seen for the Boltzmann and Liouville equations, the PBE referred to the NDF $f_N(t,\mathbf{x},\boldsymbol{\xi})$ evolving in time, t, physical space $\mathbf{x}$ and with respect to M internal coordinates $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_M)$ reads as follows (Marchisio and Fox, 2013; Ramkrishna, 2000):

$$\frac{\partial f_N}{\partial t}+\frac{\partial}{\partial \mathbf{x}}\cdot(\mathbf{v}f_N)+\frac{\partial}{\partial \boldsymbol{\xi}}\cdot(\dot{\boldsymbol{\xi}}f_N)=h_\xi, \tag{5.32}$$

where the quantity $\dot{\boldsymbol{\xi}}$ corresponds to the time derivative of the internal coordinate vector $\boldsymbol{\xi}$. For the numerical solution of Eq. (5.32), suitable expressions (closures) for the phase space coordinates rates (last term on left hand side of Eq. 5.32) are needed and vary from case to case.

Solving numerically Eq. (5.32) means achieving the knowledge of the NDF at every physical point in space and time, and every phase space point. If the internal coordinates vector $\boldsymbol{\xi}$ is made by just one element (e.g., $\xi=L$ particle size) then the PBE is said to be univariate. If the elements of the vector $\boldsymbol{\xi}$ are two, then the PBE is defined bivariate and, finally, multivariate PBE for higher dimensional internal coordinates vector $\boldsymbol{\xi}$. It is worthwhile to mention here that in Eq. (5.32) together with the advection term in physical space (second term on left hand side), there might be a diffusive-flux term when the disperse phase consists of fine particles (characteristic dimension less than 1 micron) or low-inertia particles (e.g., bubbles), due to Brownian and turbulent fluctuations.

Let us consider now the transport equation for the NDF $f_N(t, \mathbf{x}, \mathbf{v}, \boldsymbol{\xi})$, in which the particle velocity $\mathbf{v}$ is considered as internal coordinate vector. This is the case of non negligible inertia particles (very high particle Stokes number), and the disperse phase has its own velocity distribution. The kinetic equation referred to $f_N(t, \mathbf{x}, \mathbf{v}, \boldsymbol{\xi})$ is the GPBE and reads as follows:

$$\frac{\partial f_N}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{v} f_N) + \frac{\partial}{\partial \mathbf{v}} \cdot (\mathbf{A}_p f_N) + \frac{\partial}{\partial \boldsymbol{\xi}} \cdot (\dot{\boldsymbol{\xi}} f_N) = h_{\xi}, \tag{5.33}$$

where $\mathbf{A}_p$ is a continuous term that states the rate of change of particle velocity (acceleration term), or the force per unit mass acting on the particle (e.g., drag force, gravity). The term h_{ξ} is a discontinuous jump term that considers discrete events (e.g., particle collisions, nucleation, aggregation, breakage). In GPBE there are never diffusive terms in the physical space; however, diffusion in velocity space is quite often taken into consideration, to account for Brownian motion and turbulent fluctuations of the continuous phase effect on the inertial particles movement.

5.3.2 Laminar and Turbulent PBE

Referring to PBE (Eq. 5.32), if the flow is laminar, then the NDF changes due to advection and Brownian motion (fine particles); on the other side, if the flow is turbulent, then turbulent fluctuations must be taken into account because they affect the micromixing (turbulent mixing) of the system. In this section, laminar and turbulent PBE for fine particles (negligible inertia, very small Stokes number) are presented.

In an Eulerian approach, for fine particles, the Brownian motion is expressed by a size-dependent diffusion coefficient multiplied by the gradient of the NDF. The laminar PBE reads then as follows:

$$\frac{\partial f_N}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{v} f_N) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\Gamma(\boldsymbol{\xi}) \frac{\partial f_N}{\partial \mathbf{x}} \right) + \frac{\partial}{\partial \boldsymbol{\xi}} \cdot (\dot{\boldsymbol{\xi}} f_N) = h_{\xi}, \tag{5.34}$$

where $\Gamma(\boldsymbol{\xi})$ is the diffusion coefficient that can be estimated by Nernst-Einstein theory (Bird et al., 1960).

When the flow field is turbulent, the NDF fluctuates due to turbulence itself, due to non-linear convection term in the equation of motion related to the continuous phase. When the system is turbulent a Reynolds decomposition is introduce, in order to consider the fluctuating component of the turbulent flow.

By expressing the turbulent NDF f_N and the turbulent velocity $\mathbf{v}$ through the Reynolds decomposition (Eq. 5.24), a Reynolds-average transport equation for the NDF can be written down:

$$\frac{\partial\langle f_N\rangle}{\partial t}+\frac{\partial}{\partial\mathbf{x}}\cdot(\langle\mathbf{v}f_N\rangle)-\frac{\partial}{\partial\mathbf{x}}\cdot\left(\Gamma(\xi)\frac{\partial\langle f_N\rangle}{\partial\mathbf{x}}\right)+\frac{\partial}{\partial\xi}\cdot(\dot{\xi}\langle f_N\rangle)=\langle h_\xi\rangle, \tag{5.35}$$

where the $\langle\cdot\rangle$ indicates the Reynolds average (Eq. 5.25). The argument of the convection in physical space $\langle\mathbf{v}f_N\rangle$ represents though an unclosed term. The equality $\langle\mathbf{v}f_N\rangle=\langle\mathbf{v}\rangle\langle f_N\rangle+\langle\mathbf{v}'f_N'\rangle$ holds by definition (Eq. 5.24). The unclosed term $\langle\mathbf{v}'f_N'\rangle$ is called turbulent flux of the NDF and is closed thanks to the gradient-diffusion model:

$$\langle\mathbf{v}'f_N'\rangle\approx-\Gamma_T\nabla\langle f_N\rangle, \tag{5.36}$$

where Γ_T is the turbulent diffusion coefficient, calculated by proper turbulence models for the fluid phase (extensively explained in the Part II of this work).

The turbulent PBE reads then as follows:

$$\frac{\partial\langle f_N\rangle}{\partial t}+\frac{\partial}{\partial\mathbf{x}}\cdot(\langle\mathbf{v}\rangle\langle f_N\rangle)-\frac{\partial}{\partial\mathbf{x}}\cdot\left((\Gamma(\xi)+\Gamma_T)\frac{\partial\langle f_N\rangle}{\partial\mathbf{x}}\right)+\frac{\partial}{\partial\xi}\cdot(\dot{\xi}\langle f_N\rangle)=\langle h_\xi\rangle. \tag{5.37}$$

The source term $\langle h_\xi\rangle$ is closed if it can be expressed in terms of $\langle f_N\rangle$ and the phase space advection $\partial/\partial\xi\cdot(\dot{\xi}\langle f_N\rangle)$ needs to be closed with a suitable micromixing model (see Part II, sect. 8.6).

5.3.3 Second-order point processes and aggregation kernels for fine particles

So far, the source term of the PBE shown in Eq. (5.34) has been indicated with the generic expression h_ξ . Here the explicit form of the h_ξ is shown for the particular case of second-order point processes, being these strictly related to the modelling case studied in this work. They are discrete events involving the discontinuous interactions of two particles. This is the case of aggregation between nanoparticles or coalescence of bubbles, very often treated as second-order processes. More generally, the aim of this section is to give a mathematical form for events involving two particles that collide, sticking together in a solely particle as result.

The rigorous derivation of the source term functional form is reported in Appendix B. Here, only the specific case of aggregation of fine particles is explicitly shown, being this of paramount importance for the test case analysed in the following chapters. The nomenclature fine particles refers to a system in which the disperse phase is characterised by Stokes number much smaller than unity, so that their inertia is negligible. This assumption simplifies the derivation of the source term, since particle velocities are not accounted in it and are therefore considered equal to the fluid mean velocity.

Let us consider the evolution of the NDF $f_N(t, \mathbf{x}, \xi_p)$, in which the generic phase space variable referred to the fine particle is indicated with the symbol ξ_p (it might correspond to the mass or the dimension of the particle). Let us define ξ_p' and $\tilde{\xi}_p$ the phase space variables referred to the two particles that are colliding; let us suppose it represents the mass or the dimension, so that the particle that is going to form will have a phase space variable $\xi_p = \xi_p' + \tilde{\xi}_p$. The aggregation source term reads as follows (Marchisio and Fox, 2013):

$$h_\xi(\xi_p) = \frac{1}{2}\int_0^{\xi_p} \beta(\xi_p', \tilde{\xi}_p) f_N(\xi_p') f_N(\tilde{\xi}_p) d\xi_p' - \int_0^{\infty} \beta(\xi_p', \xi_p) f_N(\xi_p') f_N(\xi_p) d\xi_p', \tag{5.38}$$

that is very often written explicitly in terms of the final phase space variable ξ_p:

$$h_\xi(\xi_p) = \frac{1}{2}\int_0^{\xi_p} \beta(\xi_p', \xi_p - \xi_p') f_N(\xi_p') f_N(\xi_p - \xi_p') d\xi_p' - \int_0^{\infty} \beta(\xi_p', \xi_p) f_N(\xi_p') f_N(\xi_p) d\xi_p', \tag{5.39}$$

where $\beta(\xi_p', \tilde{\xi}_p)$ is the aggregation kernel and represents the rate at which two particles undergo second-order point process, e.g., the rate at which two particles, with dimension ξ_p' and $\tilde{\xi}_p$ respectively, aggregate forming a particle with dimension $\xi_p' = \xi_p' + \tilde{\xi}_p$. Depending on the physical meaning of the internal coordinate ξ_p, suitable closures are needed in order to calculate the source term.

Concerning the aggregation kernel $\beta(\xi_p', \tilde{\xi}_p)$, its mathematical expression strongly depends on transport phenomena that make particles come close to each other and collide. The first very important phenomenon is represented by Brownian motion. Assuming that the internal coordinate ξ_p corresponds to the particle diameter, d_p, the rate at which two particles with dimension d_p' and $\tilde{d}_p$ are brought closely to each other, collide and aggregate due to Brownian phenomena is (Elimelech et al., 1998):

$$\beta_{Brown}(d_p', \tilde{d}_p) = 4\pi \left(\Gamma' + \tilde{\Gamma}\right) \frac{(d_p' + \tilde{d}_p)}{2}, \tag{5.40}$$

where, $(d_p' + \tilde{d}_p)/2$ is the collision radius of the two particles, Γ is the diffusion coefficient that by using the Stokes-Einstein equation is:

$$\Gamma = \frac{k_B T}{3\pi \mu d_p}, \tag{5.41}$$

where μ is the fluid phase viscosity, k_B is the Boltzmann constant and T is the fluid phase temperature. By using Eq. (5.41) in Eq. (5.40) and under the assumption of Knudsen number (free molecular path to its dimension ratio) much smaller than unity, the Brownian aggregation kernel reads as follows:

$$\beta_{Brown}(d'_p, \tilde{d}_p) = \frac{2k_B T}{3\mu}\frac{(d'_p + \tilde{d}_p)^2}{d'_p \tilde{d}_p}. \tag{5.42}$$

It is very important to stress again that the Brownian kernel reported in Eq. (5.42) is valid only in continuous regime, namely Knudsen number much smaller than unity. When the Knudsen number is higher, then different and more complex functional forms such as the Fuchs equation (Fuchs, 1964) must be employed. Being not this one the case of the current work, they are omitted here and the reader can found them in Elimelech et al. (1998) and Marchisio and Fox (2013).

Another important phenomenon that brings to fine particles aggregation is the spatial gradients fluid velocity. In fact, without any spatial inhomogeneity in fluid flow field, two particles will never collide, following the same velocity line in parallel motion. Thanks to the presence of local shear rates, the flow field turns out to be locally inhomogeneous and so the orthokinetic aggregation may occur. The corresponding aggregation kernel proposed by von Smoluchowski (1917) is:

$$\beta_{per}(d'_p, \tilde{d}_p) = \frac{4}{3}G\left(\frac{d'_p}{2} + \frac{\tilde{d}_p}{2}\right)^3, \tag{5.43}$$

where $\boldsymbol{G}$ is the fluid shear rate. For laminar flows it corresponds to the second invariant of the rate-of-strain tensor $\boldsymbol{G} = 1/2(s_{ij}s_{ij} - s^2_{kk})$ where $s_{ij} = 1/2(\partial U_i/\partial x_j + \partial U_j/\partial x_i)$ and repeated indices imply summation (Einstein notation).

For turbulent flows G = $\sqrt{\varepsilon/\nu}$, where ε is the turbulent dissipation rate and ν is the kinematic viscosity, both of them referred to the continuous phase. This assumption is valid only in case of particle size smaller than Kolmogorov length scale $\eta_k = (\nu^3/\varepsilon)^{1/4}$, resulting in the following turbulent aggregation kernel (Saffman and Turner, 1956):

$$\beta_{Turb}(d'_p, \tilde{d}_p) = \left(\frac{8\pi}{15}\right)^{0.5}\left(\frac{\varepsilon}{\nu}\right)^{0.5}\left(\frac{d'_p}{2} + \frac{\tilde{d}_p}{2}\right)^3 \approx 1.2944\left(\frac{\varepsilon}{\nu}\right)^{0.5}\left(\frac{d'_p}{2} + \frac{\tilde{d}_p}{2}\right)^3. \tag{5.44}$$

For particle size greater than Kolmogorov length scale, different functional forms of the turbulent aggregation kernels must be used. It is worthwhile stressing one more time that these expressions reported in Eq. (5.42) and Eq. (5.44) are valid only in case of particles Stokes number that approaches to zero. When the Stokes number is greater than zero, than the particles inertia must be taken into account and suitable corrections to the aggregation kernels must be introduced (Ammar and Reeks, 2009). The total aggregation kernel is usually calculated as the summation of the single contributions reported in Eq. (5.42) and Eq. (5.44), $\beta = \beta_{Brown} + \beta_{Turb}$.

PART II

Computational Details, Algorithms and Numerical Techniques

Chapter 6

All-Atom Molecular Dynamics

6.1 MD Algorithms

In this chapter, technical details about the algorithms used for all-atom MD simulations will be provided. In order to achieve a complete description of all the most important algorithms, let us start from the dynamic equations typical of MD, namely the Hamilton's equations of motion for a classical system reported in Eq. (3.9) that in terms of second Newton's law read as follows:

$$\begin{aligned} m_i \ddot{\mathbf{r}}_i &= \mathbf{F}_i, \\ \mathbf{F}_i &= -\frac{\partial \mathscr{V}(\mathbf{r})}{\partial \mathbf{r}_i}, \end{aligned} \tag{6.1}$$

where $\mathbf{F}_i$ is the force acting on the atom i, at position $\mathbf{r}_i$, with mass m_i, and interatomic potential $\mathscr{V}(\mathbf{r})$, with $\mathbf{r} = (\mathbf{r}_1, ..., \mathbf{r}_i, ..., \mathbf{r}_N)$ are the coordinates of the N particles. Note that $\mathbf{r}_i$ is a vector in spatial Cartesian coordinates.

Introducing the momenta vector $\mathbf{p} = (\mathbf{p}_1, ..., \mathbf{p}_i, ..., \mathbf{p}_N)$ the classical equations of motion in Cartesian coordinates become:

$$\begin{aligned} \dot{\mathbf{r}}_i &= \mathbf{p}_i / m_i, \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i. \end{aligned} \tag{6.2}$$

It is noteworthy to stress here that whereas the system of equations reported in Eq. (6.1) is a set of $3N$ second-order differential equations, the latter is a set of $6N$ first-order differential equations. Eq. (6.1) can be solved with the Verlet-based algorithm; Eq. (6.2) is the base of the Euler algorithm. All of them will be schematically presented in the following sections. In a preliminary way, the general scheme which MD simulations are based on is presented in Table 6.1.

THE GENERAL MD ALGORITHM
1. Input initial conditions Potential interaction $\mathcal{V}$ as a function of atom positions Positions $\mathbf{r}_i$ of all atoms in the system Momenta $\mathbf{p}_i$ of all atoms in the system $\Downarrow$
repeat 2,3,4 for the required number of steps:
2. Compute forces The force on each atom $\mathbf{F}_i = -\frac{\partial \mathcal{V}(\mathbf{r})}{\partial \mathbf{r}_i}$ is computed (summation of bonded, non-bonded, restrains, external forces) Potential and kinetic energies and the pressure tensor are computed $\Downarrow$
3. Update configuration The movement of the atoms is simulated by numerically solving Newton's equations of motion $m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i$, or Eq. (6.2) $\Downarrow$
4. if required: **Output step** positions, velocities, energies, temperature, pressure, etc.

Table 6.1: MD general algorithm

6.2 Integration Schemes

A common method to solve ordinary differential equations such as Eq. (6.1) and (6.2) is represented by the finite difference method. Starting from initial conditions (atoms positions and momenta) and following the general scheme reported in Table 6.1, numerical frameworks are necessary to update positions, velocities, etc.. from a simulation time t to the following simulation time $t + \Delta t$. The equations are then solved by a step-by-step method, and the choice of the time-step Δt depends on the other numerical conditions, as well as the kind of algorithm used for the integration scheme. Before going into the specific algorithms, assuming the classical trajectories to be continuous, let us apply a Taylor expansion to the positions and velocities at the instant time $t + \Delta t$:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \Delta t\, \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i(t) + \frac{\Delta t^3}{3!}\dddot{\mathbf{r}}_i(t) + \mathcal{O}(\Delta t^4), \tag{6.3}$$

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{\Delta t}{m_i}\,\mathbf{F}_i(t) + \frac{\Delta t^2}{2}\ddot{\mathbf{v}}_i(t) + \frac{\Delta t^3}{3!}\dddot{\mathbf{v}}_i(t) + \mathcal{O}(\Delta t^4). \tag{6.4}$$

6.2.1 Euler Algorithm

The simplest integration scheme is the Euler integration scheme which reads as follows:

$$\begin{aligned} \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \Delta t\, \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i(t) + \mathcal{O}(\Delta t^3), \\ \mathbf{v}_i(t+\Delta t) &= \mathbf{v}_i(t) + \frac{\Delta t}{m_i}\,\mathbf{F}_i(t) + \mathcal{O}(\Delta t^2). \end{aligned} \tag{6.5}$$

Despite its simplicity, it has the disadvantages that it is no-time reversible and does not preserve the phase-space, becoming then not very used in the computational field.

6.2.2 Verlet Algorithm

One of the most used algorithms in MD simulations is the Verlet algorithm. By solving the second order system reported in Eq. (6.1), considering the position and forces at the instant time t, and considering the previous time $t - \Delta t$, the Taylor expansion becomes:

$$\mathbf{r}_i(t-\Delta t) = \mathbf{r}_i(t) - \Delta t\, \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i(t) - \frac{\Delta t^3}{3!}\dddot{\mathbf{r}}_i(t) + \mathcal{O}(\Delta t^4). \tag{6.6}$$

By summing Eq. (6.3) to Eq. (6.6), and by subtracting the same equations for the velocities, the obtained set is:

$$\begin{aligned} \mathbf{r}_i(t+\Delta t) &= 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t) + \frac{\Delta t^2}{m_i}\mathbf{F}_i(t) + \mathcal{O}(\Delta t^4), \\ \mathbf{v}_i(t) &= \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^3). \end{aligned} \tag{6.7}$$

A closer look at Eq. (6.7) reveals that velocities are not needed anymore to compute the trajectories; however the code needs to calculate them in order to compute the kinetic energy and other observable properties. The algorithm gives some numerical imprecision due to the fact that a small quantity (order Δt^2) is added to a difference of order $\mathcal{O}(1)$-factors.

6.2.3 Leap-frog Algorithm

The leap-frog algorithm reads as follows:

$$\begin{aligned} \mathbf{v}_i(t+\frac{\Delta t}{2}) &= \mathbf{v}_i(t-\frac{\Delta t}{2}) + \frac{\Delta t}{m_i}\mathbf{F}_i(t), \\ \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \Delta t\ \mathbf{v}_i(t+\frac{\Delta t}{2}). \end{aligned} \tag{6.8}$$

The velocity is calculated at half time-steps, 'leaping' the positions. The update velocity is then estimated as:

$$\mathbf{v}_i(t) = \frac{\mathbf{v}_i(t+\frac{\Delta t}{2}) + \mathbf{v}_i(t-\frac{\Delta t}{2})}{2}. \tag{6.9}$$

6.2.4 Velocity-Verlet Algorithm

The velocity-Verlet algorithm gives the advantage of having positions, velocities and forces at the same simulation time. It reads as follows:

$$\begin{aligned} \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \Delta t\ \mathbf{v}_i(t) + \frac{\Delta t^2}{m_i}\mathbf{F}_i(t) + \mathcal{O}(\Delta t^3), \\ \mathbf{v}_i(t+\Delta t) &= \mathbf{v}_i(t) + \frac{\Delta t}{2m_i}(\mathbf{F}_i(t) + \mathbf{F}_i(t+\Delta t)) + \mathcal{O}(\Delta t^3). \end{aligned} \tag{6.10}$$

It preserves the phase space and it is very stable, being therefore one of the most used algorithms in MD. The normal Verlet scheme (Eq. 6.7) can be derived from Eq. (6.10), by eliminating the velocities during the positions calculation.

6.2.5 Predictor-Corrector Schemes

Unlike the methods presented so far, the predictor-corrector schemes approach to the calculation of the trajectories in a different, optimised way. The general idea is to predict positions, velocities, accelerations, and third time derivative of positions (all the terms appearing in Eq. 6.3) and starting from them, to correct the value at the next step in a proper way. Let us define the predicted quantities in the following system:

$$
\begin{aligned}
\mathbf{r}^p(t+\Delta t) &= \mathbf{r}(t) + \Delta t\,\mathbf{v}_i(t) + \frac{\Delta t^2}{2}\mathbf{a}(t) + \frac{\Delta t^3}{3!}\dddot{\mathbf{r}}(t) + \ldots \\
\mathbf{v}^p(t+\Delta t) &= \mathbf{v}(t) + \Delta t\,\mathbf{a}(t) + \frac{\Delta t^2}{2}\dddot{\mathbf{r}}(t)\ldots \\
\mathbf{a}^p(t+\Delta t) &= \mathbf{a}(t) + \Delta t\,\dddot{\mathbf{r}}(t)\ldots \\
\dddot{\mathbf{r}}^p(t+\Delta t) &= \dddot{\mathbf{r}}(t) + \ldots
\end{aligned}
\tag{6.11}
$$

Once the predicted positions, $\mathbf{r}^p(t+\Delta t)$, are known, the correct accelerations, $\mathbf{a}^c(t+\Delta t)$, (passing through the force) can be computed. Therefore, an error in the prediction step is:

$$\Delta\mathbf{a}(t+\Delta t) = \mathbf{a}^c(t+\Delta t) - \mathbf{a}^p(t+\Delta t). \tag{6.12}$$

By using Eq. (6.12), the new correct quantities can be evaluated:

$$
\begin{aligned}
\mathbf{r}^c(t+\Delta t) &= \mathbf{r}^p(t+\Delta t) + c_0\Delta\mathbf{a}(t+\Delta t) \\
\mathbf{v}^c(t+\Delta t) &= \mathbf{v}^p(t+\Delta t) + c_1\Delta\mathbf{a}(t+\Delta t) \\
\mathbf{a}^c(t+\Delta t) &= \mathbf{a}^p(t+\Delta t) + c_2\Delta\mathbf{a}(t+\Delta t) \\
\dddot{\mathbf{r}}^c(t+\Delta t) &= \dddot{\mathbf{r}}^p(t+\Delta t) + c_3\Delta\mathbf{a}(t+\Delta t)
\end{aligned}
\tag{6.13}
$$

Equations (6.13) represent the corrector step of the MD simulation. The choice of the coefficients c_0, c_1, c_2, c_3 depends on the specific chosen predictor-corrector technique, as discussed by Gear (1967).

6.2.6 Stochastic and Brownian Dynamics Algorithms

Considering the stochastic dynamics, the generalised Langevin equation (Eq. 4.1) is solved by either simple integrators (Hess, Kutzner, et al., 2008) or a more complex leap-frog one. Details about the last one can be found in van Gunsteren and Berendsen (1988).

Concerning instead Brownian dynamics (sect. 4.6), namely when the system has a negligible inertia (non-inertial Langevin dynamics) a simple Euler scheme is adopted:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \frac{\Delta t}{\zeta_i}\mathbf{F}_i^C(\mathbf{r}(t)) + \sqrt{2k_B T\,\frac{\Delta t}{\zeta_i}}\,\mathbf{r}_i^G, \tag{6.14}$$

where $\mathbf{r}_i^G$ is a Gaussian distribution to describe the noise effect.

6.3 Trotter expansion

As already mentioned in sect. 3.2, the solution of the equation of motion referred to a generic property $A(\mathbf{\Gamma}, t)$ is represented by $A(\mathbf{\Gamma}(t)) = e^{-iLt}A(\mathbf{\Gamma}(0))$, introducing also the propagator e^{-iLt}. Starting from the propagator, the aim of this section is to investigate Trotter dynamics, in order to better understand the main properties of the time-reversible algorithms (such as the velocity-Verlet, Eq. 6.10).

The Liouville operator expressed in Cartesian coordinates $(\mathbf{r}, \mathbf{p})$, for positions and momenta respectively, can be split into two contributions:

$$iL = iL_{\mathbf{r}} + iL_{\mathbf{p}}, \tag{6.15}$$

with

$$iL_{\mathbf{r}} = \sum_i \dot{\mathbf{r}}_i \frac{\partial}{\partial \mathbf{r}_i} \tag{6.16}$$

and

$$iL_{\mathbf{p}} = \sum_i \dot{\mathbf{p}}_i \frac{\partial}{\partial \mathbf{p}_i}. \tag{6.17}$$

The Trotter expansions reads as follows:

$$e^{i(L_{\mathbf{r}}+L_{\mathbf{p}})\Delta t} = e^{iL_{\mathbf{p}}\frac{\Delta t}{2}} e^{iL_{\mathbf{r}}\Delta t} e^{iL_{\mathbf{p}}\frac{\Delta t}{2}} + \mathcal{O}(\Delta t^3). \tag{6.18}$$

Applying the Trotter expansion to positions $\mathbf{r}_i$ and momenta $\mathbf{p}_i$,

$$\begin{aligned} e^{iL_{\mathbf{p}}\frac{\Delta t}{2}}\mathbf{r}_i &= \mathbf{r}_i \\ e^{iL_{\mathbf{r}}\Delta t}\mathbf{r}_i &= \mathbf{r}_i + \Delta t\, \dot{\mathbf{r}}_i \\ e^{iL_{\mathbf{p}}\frac{\Delta t}{2}}\mathbf{p}_i &= \mathbf{p}_i + \frac{\Delta t}{2}\dot{\mathbf{p}}_i \\ e^{iL_{\mathbf{r}}\Delta t}\mathbf{p}_i &= \mathbf{p}_i, \end{aligned} \tag{6.19}$$

the following set of equations can be rewritten as follows:

$$\begin{aligned} e^{iL_{\mathbf{p}}\frac{\Delta t}{2}} e^{iL_{\mathbf{r}}\Delta t} e^{iL_{\mathbf{p}}\frac{\Delta t}{2}}\mathbf{r}_i(t) &= \mathbf{r}_i(t) + \Delta t\, \dot{\mathbf{r}}_i(\frac{\Delta t}{2}) = \\ &= \mathbf{r}_i(t) + \Delta t\, \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i}\mathbf{F}_i(t) \\ e^{iL_{\mathbf{p}}\frac{\Delta t}{2}} e^{iL_{\mathbf{r}}\Delta t} e^{iL_{\mathbf{p}}\frac{\Delta t}{2}}\mathbf{p}_i(t) &= \mathbf{p}_i(t) + \frac{\Delta t}{2}\,(\dot{\mathbf{p}}_i(t) + \dot{\mathbf{p}}_i(t+\Delta t)) = \\ &= \mathbf{p}_i(t) + \frac{\Delta t}{2}\,(\mathbf{F}_i(t) + \mathbf{F}_i(t+\Delta t)), \end{aligned} \tag{6.20}$$

yielding exactly the equations of the velocity-Verlet algorithm. The Trotter operators and the analysis reported in Eq. (6.16) and (6.17) show that these methods based on the velocity-Verlet algorithm conserve the volume in phase-space and guarantee time reversibility (due to the coordinates symmetry of the equations).

6.4 Temperature Control

In MD simulations at constant temperature (NVT, NPT ensembles) the temperature is controlled by suitable thermostats that are coupled to the system. Several thermostats are available in MD codes and can be found in literature, such as the stochastic Andersen thermostat (Andersen, 1980), Nosè-Hoover (Hoover, 1985; Nose, 1984), Berendsen (Berendsen, Postma, DiNola, et al., 1984) and velocity rescale thermostats (Bussi et al., 2007). Here, only the last two will be presented since they have been used in this work.

6.4.1 Berendsen Thermostat

The Berendsen thermostat is a weak-coupling of first-order kinetics with respect to the reference temperature T_0:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau}, \tag{6.21}$$

implying an exponential decay with time constant τ. It is a very robust method, but it suppresses kinetic energy fluctuations, impeding the proper generation of a canonical ensemble. This error decreases as the number of particles increases, therefore for large system it might be successfully employed.

6.4.2 Velocity-Rescale Thermostat

The above-mentioned issue can be overcome by using a velocity-rescale thermostat. It simply corresponds to a classical Berendsen thermostat plus a stochastic term that allows the development of the correct kinetic energy distribution:

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}} \cdot \frac{dW}{\sqrt{\tau_T}}, \tag{6.22}$$

where N_f corresponds to the total number of degrees of freedom, K is the kinetic energy and dW is a Wiener process (i.e., a stochastic continuous Gaussian process). This approach has the advantage of reproducing the correct ensemble, together with those already listed in the Berendsen section.

6.5 Pressure Control

In line with the theory, also a pressure control is applied to some MD simulations. Different barostats can be used, some of them more suitable for specific thermostats. The most used ones are the Berendsen (Berendsen, Postma, DiNola, et al., 1984), the Parrinello-Rahman algorithm (Nose and Klein, 1983; Parrinello and Rahman, 1981) and the Martyna-Tuckerman-Tobias-Klein (MTTK) one (Martyna et al., 1996). The first two can be used with any thermostat; on the other hand, the last one can only be coupled with the Nosè-Hoover thermostat.

6.5.1 Berendsen Barostat

Following the Berendsen thermostat, the barostat follows the same approach:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P}. \tag{6.23}$$

A scaling matrix μ is implemented in order to rescale positions and box vectors at every number of time steps, n_{step}, defined as follows:

$$\mu_i = \delta_{ij} - \frac{n_{step}\Delta t}{3\tau_p}\beta_{ij}\{P_{0ij} - P_{ij}(t)\}, \tag{6.24}$$

where β is the isothermal compressibility.

6.5.2 Parrinello-Rahman Barostat

When pressure fluctuations are too large and the simulation boxes are quite small, then a weak-coupling might not suffice to reproduce the correct canonical ensemble. Parrinello-Rahman barostat is then needed. The box vectors are represented by a matrix b, following the matrix equation of motion:

$$\frac{d^2b}{dt^2} = V\mathbf{W}^{-1}b'^{-1}(\mathbf{P} - \mathbf{P}_{ref}), \tag{6.25}$$

where the $\mathbf{P}_{ref}$ and $\mathbf{P}$ are the reference and current pressure, V is the box volume and $\mathbf{W}$ is a matrix parameter that governs the strength of the coupling.

The pressure matrix $\mathbf{P}$ is calculated from the virial $\mathbf{\Xi}$ and the kinetic energy $\mathbf{K}$:

$$\mathbf{P} = \frac{2}{V}(\mathbf{K} - \mathbf{\Xi}). \tag{6.26}$$

The scalar $P = 1/3tr(\mathbf{P})$ and the virial is defined as:

$$\mathbf{\Xi} = -\frac{1}{2}\sum_{i<j}\mathbf{r}_{ij} \otimes \mathbf{F}_{ij}, \tag{6.27}$$

where the symbol $\otimes$ identifies the direct product between two vectors which gives a second-order tensor.

6.6 Force Fields and Interaction Potentials

Before dealing with the interactions between particles inside a classical MD system, the concept of force field must be introduced. In MD, a force field is a set of analytical functions (potentials, parameters, etc.) that fully describes a given system at the atomistic scale level. In line with the multiscale modeling approach and, above all, with the concept of relevant degrees of freedom going from one scale to another (sect. 4.1), these functions are calculated at a lower scale (ab initio methods, i. e., quantum mechanical simulations, Maple et al., 1988) and, therefore, ”closed” by eliminating the irrelevant DOF (e.g., electrons, etc.). This closure leads to a set of parameters that characterise the specific force field. Different force fields are available for MD simulations; they differ from case to case and their choice depends on the kind of system that has to be simulated (i.e., proteins, gas, liquid-liquid, etc.).

6.6.1 Interaction Potentials

In classical MD, the interaction potentials taken into account in the force fields consist of the summation of bonded and non-bonded interactions between atoms. Different functional forms exist to describe them and are schematically reported in the following subsections.

Non-bonded Potentials

Pairwise potentials are employed to account for the non-bonded interactions (particle attractions and repulsions). The most used is the Lennard-Jones (LJ) potential, $V_{LJ}(r_{ij})$, that for two atoms i and j in Cartesian coordinates reads in the following forms:

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}}, \tag{6.28}$$

where the parameters $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ depend on the kind of atoms considered and are stored in the force fields files of the MD simulation package. An alternative form is:

$$V_{LJ}(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{6.29}$$

where the LJ parameters ϵ_{ij} and σ_{ij} represent respectively the well depth of the potential minimum point and the distance at which the interactions are null, as shown in Figure 10.1.

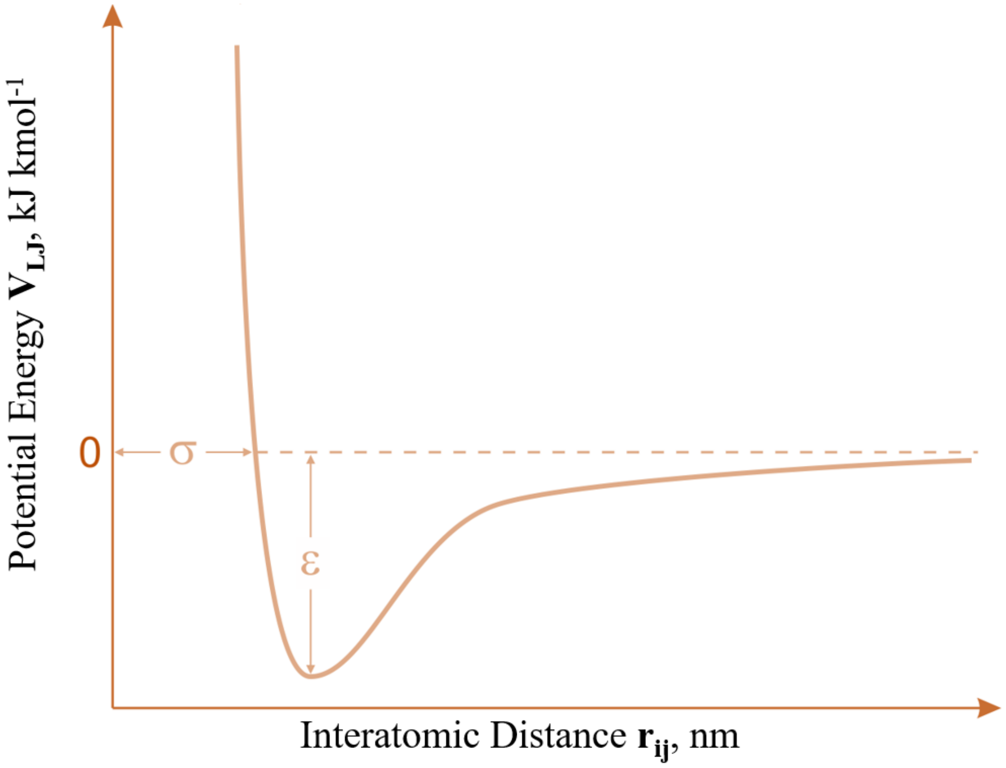


Figure 6.1: Lennard-Jones potential.

The way in which the LJ parameters, appearing in both equations Eq. (6.28) and Eq. (6.29), are calculated is named combination rules. It simply consists in the way in which the single atom i parameter C_i, for instance, is combined with the atom j parameter C_j, obtaining C_{ij}. The combination rules can be either geometric or arithmetic averages or a mix of them (Lorentz-Berthelot rules).

Alternatively to the LJ potential, the Buckingham potential can be used:

$$V_{bh}(r_{ij}) = A_{ij}exp(-B_{ij}r_{ij}) - \frac{C_{ij}}{r_{ij}^6}, \tag{6.30}$$

where a more advantageous repulsion term is used, rising though the computational cost.

Concerning electrostatic interactions, the Coulomb potential is typically used:

$$V_C(r_{ij}) = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{\epsilon_r r_{ij}}, \tag{6.31}$$

where ϵ_0 and ϵ_r are the dielectric constants respectively in vacuum and in the medium, r_{ij} is the distance between atoms i and j which carry the charges q_i and q_j.

Bonded Potentials

Bonded potentials are functional forms that account for bonded atoms interactions and are divided in bonds (2-atoms interaction), angles (3-atoms interaction), proper and improper dihedrals (4-atoms interaction).

The bond stretching between two atoms i and j is modelled by a harmonic spring:

$$V_b(r_{ij}) = \frac{1}{2}k^b_{ij}(r_{ij} - b_{ij})^2, \tag{6.32}$$

where k^b_{ij} and b_{ij} corresponds respectively to the bond force constant and the equilibrium bond distance between two generic bonded atoms i and j. The situation is depicted in Figure 6.2.

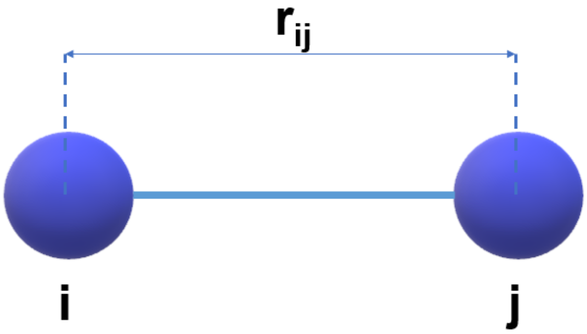


Figure 6.2: Representation of a bond between two generic atoms i and j.

The harmonic angle potential between three atoms i, j and k reads as follows:

$$V_\theta(\theta_{ijk}) = \frac{1}{2}k^\theta_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2, \tag{6.33}$$

where k^θ_{ijk} and θ^0_{ijk} corresponds respectively to the angle force constant and the equilibrium angle value between a generic triplet of atoms i, j, k consecutively bonded. The situation is shown in Figure 6.3.

The dihedrals are group of four atoms bonded to each other. If they are bonded in a continuous way, one by the other one, following a line, then they are defined proper dihedral and describe the torsion of the molecule. Sometimes, four atoms are connected to each other in a non-continuous way, but belonging to a plane (carboxyl group). In this case the dihedral is said to be improper and means to keep that plane planar, as well as to prevent a chiral molecule to change its conformation (by mirroring itself).

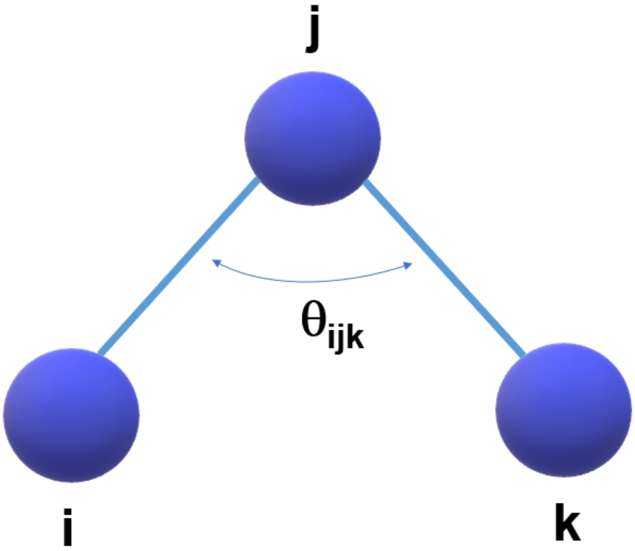


Figure 6.3: Representation of an angle between a triplet of atoms i,j,k.

The improper dihedral potential is still a harmonic spring:

$$V_{\xi}(\xi_{ijkl}) = \frac{1}{2}k^{\xi}_{ijkl}(\xi_{ijkl} - \xi^{0}_{ijk})^{2}, \tag{6.34}$$

where the improper dihedral angle ξ_{ijkl} is defined as the angle between the planes $i - j - k$ and $j - k - l$. This scenario is reported in Figure 6.4.

The proper dihedrals potential follows a continous periodic function:

$$V_{d}(\phi_{ijkl}) = k_{\phi}(1 + \cos(n\phi - \phi_{s})), \tag{6.35}$$

where ϕ is the angle between the plane $i - j - k$ and $j - k - l$. This situation is reported in Figure 6.5. An alternative form is represented by the Ryckaert-Bellemans function, namely an expansion in powers of cosines:

$$V_{rb}(\phi_{ijkl}) = \sum_{n=0}^{5} C_{n}(cos(\psi))^{n}, \tag{6.36}$$

where $\psi = \phi - 180°$.

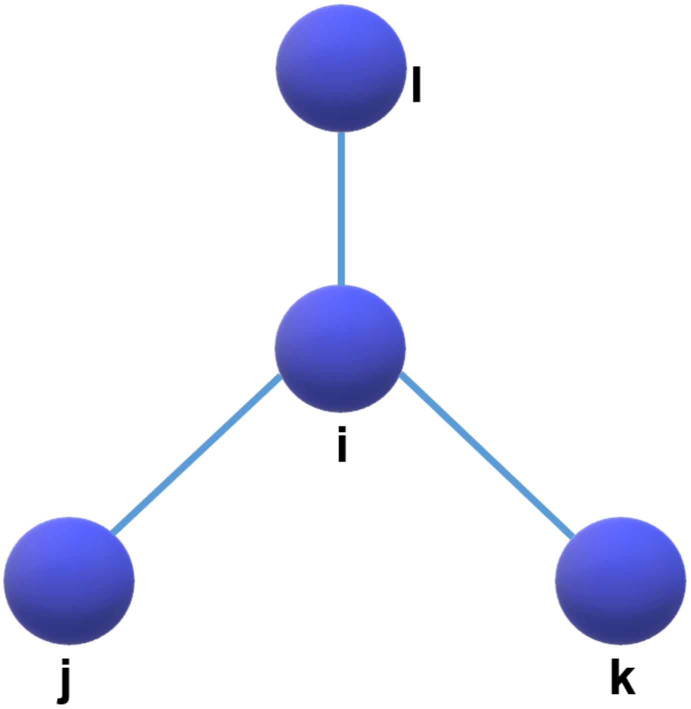


Figure 6.4: Representation of an improper dihedral (torsional) angle ξ_{ijkl} between the planes $i - j - k$ and $j - k - l$.

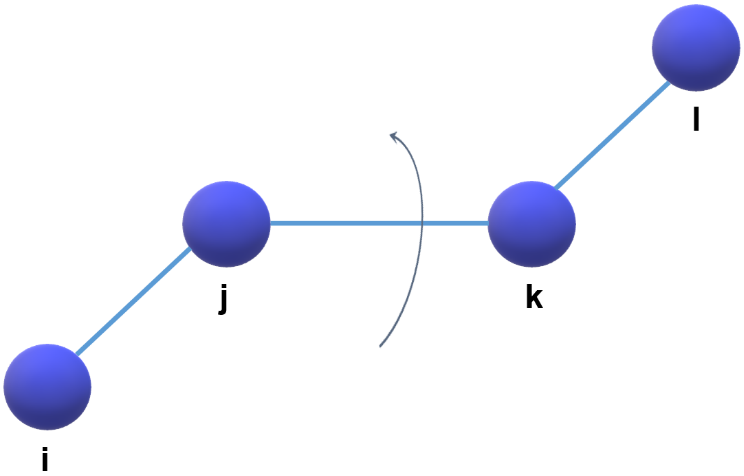


Figure 6.5: Representation of a proper dihedral (torsional) angle ϕ_{ijkl} between the planes $i - j - k$ and $j - k - l$.

6.7 Constraints, Periodic Boundary Conditions and Cut-off

In classical MD simulations there is a very common practice to consider the bond lengths constrained at a fixed bond length, instead of considering their intramolecular potential. The advantage is twofold: first, the torsional potentials (proper and improper dihedrals) vibrate at a much lower frequency than the chemical bonds and, secondly, a practical reason consists in the possibility of using a four-five times larger time-step (van Gunsteren and Berendsen, 1977) in the simulation run, giving higher stability and lowering the computational cost. This approximations turns out to be very useful in treating the hydrogen vibrations, being a proton very light and therefore vibrating at very high frequencies. Being interested in the long-term conformation/evolution of the molecule dynamics, the vibrations that happens at a lower time-scale are then approximated with good results in the perspective of the classical limit of the MD.

Several techniques have been developed to include this constrained dynamics in the equation of motion and they are based on the so-called Lagrange multipliers (Allen and Tildesley, 1987; Ryckaert, 1985). The most used algorithms are SHAKE (Ryckaert et al., 1977) for the Verlet algorithm and the more stable and faster LINCS (Hess, Bekker, et al., 1997).

MD simulations take place in a finite control volume that is denominated simulation box. Due to the limited, finite size periodic boundary conditions (PBC) are employed to minimise the edge effect on the evolution of the system. It simply consists in surrounding the simulation box with its replicas on each box face. In this way each particle interacts with the surrounding particles, including the PBC images thereof. This is schematically reported in Figure 6.6.

For the sake of computational cost the molecular interactions are limited to a given distance from the specific particle. This distance is the so-called cutoff radius. Each particle can interact with all the others that are placed at a distance less than the cutoff including the replica images due to PBC. In certain cases, the non-bonded interactions are also calculated in a way to include also the effect of long-range interactions,. This is the case of the Ewald summation techinque or the Particle Mesh Ewald (PME, Darden et al. (1993) and Essmann et al. (1995)) algorithms for long-range Coulomb interactions, which avoid the accumulation of charges in the cutoff sphere that would be caused by using the cutoff method. The cutoff radius must be shorter than the half shortest box length.

Different simulation boxes may be used, such as cubic, rhombic dodecahedron, truncated octahedron, etc. From the point of view of the computational cost, it is more convenient to store the pair atoms interactions of the neighbour atoms around a given particle. These neighbours pairs are store in the so-called neighbour list, identified by a radius as well. Only the neighbour pairs stored in the neighbour list are considered for the force calculation (cutoff LJ potential, etc.), saving computing time. The neighbour

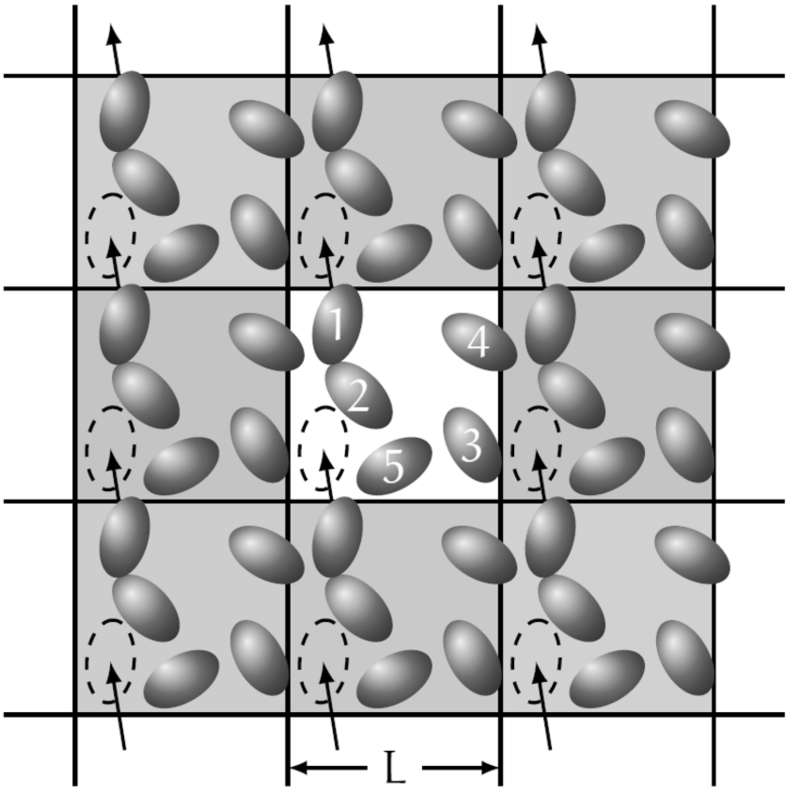


Figure 6.6: Periodic boundary conditions in MD simulations, using cubic simulation boxes. Image taken and modified from Allen and Tildesley (1987)

list radius must be greater or equal to the cutoff one.

6.8 Polarization Effect

Polarization effect can be introduced in classical MD by implementing some artificial particles responsible of carrying the charge on themselves, following the Born-Oppenheimer approximation, i.e. decoupling the electronic and nuclear motions degrees of freedom. These are divided into two groups: the shell (drude) particles and the virtual interaction sites. The former are particles with a given mass and charge, whereas the latter are massless by definition. Also a mix of them can be used in order to improve the complexity of the polarization effect on a given molecule. A very famous application of this technique is represented by the so-called Drude Oscillator Model (Lemkul et al., 2016).

When more than one shell particle is implemented on a molecule, the intra-molecular electrostatic interaction must be taken into account, thanks to the Thole polarization effect (Thole, 1981), and more specifically by introducing the Thole potential between two shell particles i and j:

$$V_{th} = \frac{q_i q_j}{r_{ij}} \left[1 - \left(1 + \frac{\hat{r}_{ij}}{2} \right) e^{-\hat{r}_{ij}} \right], \tag{6.37}$$

where

$$\hat{r}_{ij} = a \frac{r_{ij}}{(\alpha_i \alpha_j)^{1/6}}, \tag{6.38}$$

with the a parameter equal to 2.6 (Noskov et al., 2005); α_i and α_j are the polarizabilities related to the two shell particles i and j respectively.

Being the second technique (virtual interaction site) employed in part of this work (see Chapt. 9), a more detailed discussion is included in the following subsection.

6.8.1 Virtual Interaction Site

A virtual interaction site (VS) is used in MD simulations in several ways. Basically, it is “virtual” since it is massless, but, an ”interaction site” to interact with other atoms thanks to a charge placed on it.

In MD simulations, VSs are built up by means of position constrains with respect to a given number of N surrounding atoms. Therefore, the VS will be position-constrained to the chosen N surrounding atoms, as shown in Figure 6.7 which depicts a very generic scenario. Let us consider a system of N surrounding particles with respect to which the VS is built up; the simplest choice is to write the VS position vector, $\mathbf{r}_s = (x_s, y_s, z_s)$, as a linear combination of the overall surrounding atoms positions, $\mathbf{r}_i = (x_i, y_i, z_i)$, with $i = 1, ..., N$:

$$\mathbf{r}_s = \sum_{i=1}^{N} w_i \mathbf{r}_i, \tag{6.39}$$

where w_i are the weights of the linear combination, suitably chosen depending on the VS type, and that will be defined later on in this section. Since a VS is massless, the system takes it into account by redistributing the force $\mathbf{F}_s$ (depicted in brown in Figure 6.7) acting on the VS (before the integration of the equation of motion) over the other N particles (with mass) with respect to which the VS was constrained, in a consistent way (Berendsen and van Gunsteren, 1984). The force contribution due to the VS presence is then redistributed according to the weights w_i in the same proportional way, namely

$$\mathbf{F}'_i = \sum_{i=1}^{N} w_i \mathbf{F}_s. \tag{6.40}$$

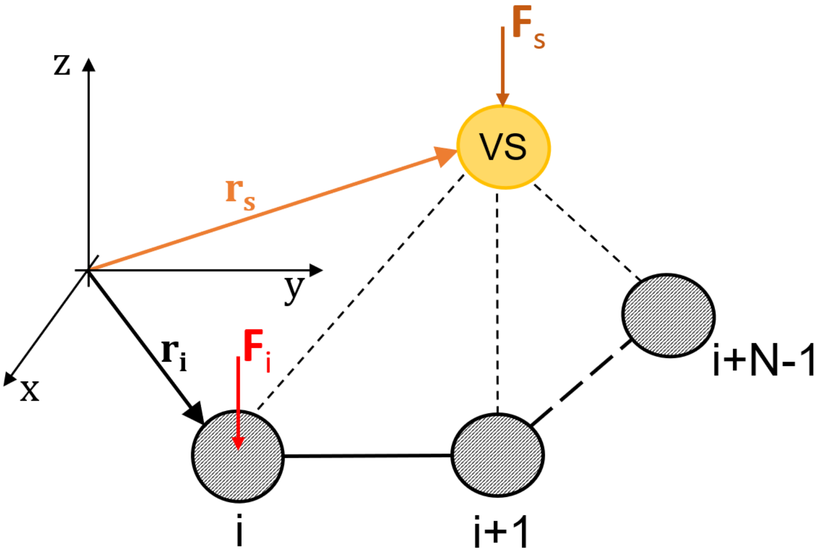


Figure 6.7: Generic scheme of a VS built up with respect to N surrounding atoms $i, i+1, ..., i+N-1$. The force $\mathbf{F}_s$ acting on the VS is redistributed on the N surrounding atoms with respect to which the VS itself is constrained, contributing to the total force $\mathbf{F}_i$ acting on them (Eq. 6.41).

The total force $\mathbf{F}_i$ (red arrow in Figure 6.7), acting on the generic surrounding atom i, is:

$$\mathbf{F}_i = \mathbf{F}_i^{direct} + \mathbf{F}_i' = \mathbf{F}_i^{direct} + \frac{\partial \mathbf{r}_s}{\partial \mathbf{r}_i}\mathbf{F}_s = \mathbf{F}_i^{direct} + \begin{bmatrix} \frac{\partial x_s}{\partial x_i} & \frac{\partial y_s}{\partial x_i} & \frac{\partial z_s}{\partial x_i} \\ \frac{\partial x_s}{\partial y_i} & \frac{\partial y_s}{\partial y_i} & \frac{\partial z_s}{\partial y_i} \\ \frac{\partial x_s}{\partial z_i} & \frac{\partial y_s}{\partial z_i} & \frac{\partial z_s}{\partial z_i} \end{bmatrix}\mathbf{F}_s, \tag{6.41}$$

where $\mathbf{F}_i^{direct}$ is the normal force acting on the particle i (in absence of VS), the second term is the force due to the VS, strictly dependent on the force $\mathbf{F}_s$, and also a general definition of the weights w_i is therefore given (matrix of partial position derivatives).

Thanks to this modelling approach, the total potential of the system is able to consider the presence of the new particle (VS) in the conservation laws, so that the total force and total torque are conserved.

Depending on the number of surrounding atoms N, different kinds of VS can be set up. In the simplest scenario (adopted also in this work to develop the charge-on-particle model, Chapt. 9), $N = 2$ and, consequently, the generic scheme shown in Figure 6.7 reduces to:

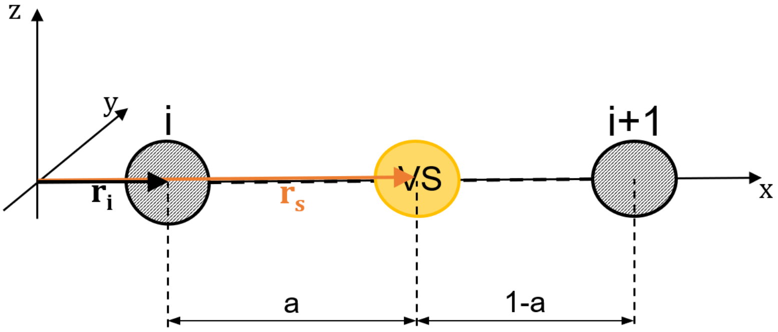


Figure 6.8: Scheme of VS built up with respect to $N = 2$ surrounding atoms. The three particles (VS plus the two atoms i and $i + 1$) lie on the axis x. The relative distance from the two surrounding atoms is governed by the weights $w_i = 1 - a$ and $w_{i+1} = a$.

In this very simple scenario, the weights easily read as:

$$\begin{aligned} w_i &= 1 - a \\ w_{i+1} &= a, \end{aligned} \tag{6.42}$$

where a is a geometric parameter suitably chosen by the user, depending on the physics that one wants to reproduce (e.g., electrons pair around an atom, electron cloud deformation and so on). Being this one the VS type implemented in the current work, in order to make it more clear, a numerical example is here proposed and depicted in Figure 6.9. Let us suppose to be in the following situation:

$$\begin{aligned} \mathbf{r}_i = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{r}_{i+1} &= \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \\ a &= 0.75, \\ w_i &= 1 - a = 0.25, \\ w_{i+1} &= a = 0.75. \end{aligned} \tag{6.43}$$

By applying Eq. (6.39), the VS position $\mathbf{r}_s$ can be found out:

$$\mathbf{r}_s = \sum_{i=1}^{2} w_i \mathbf{r}_i = w_i \mathbf{r}_i + w_{i+1} \mathbf{r}_{i+1} = 0.25 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0.75 \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.75 \\ 0 \\ 0 \end{pmatrix}, \tag{6.44}$$

as schematically reported in Figure 6.9.

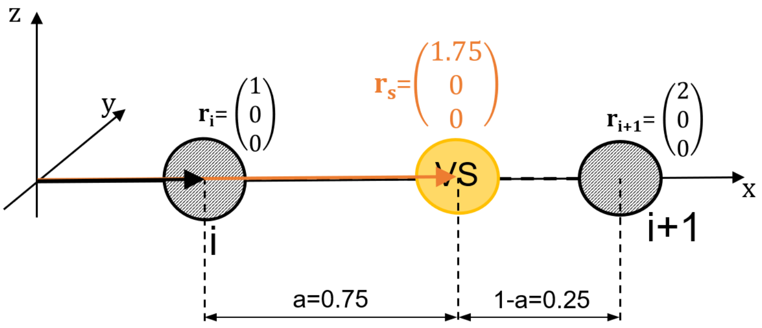

z
y
$\mathbf{r}_i = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$
$\mathbf{r}_s = \begin{pmatrix} 1.75 \\ 0 \\ 0 \end{pmatrix}$
$\mathbf{r}_{i+1} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$
VS
x
i
i+1
a=0.75
1-a=0.25

Figure 6.9: Numerical example of a VS built up with respect to $N = 2$ surrounding atoms and $a = 0.75$. The three particles (VS plus the two atoms i and $i + 1$) lie on the axis x. The relative distance from the two surrounding atoms is governed by the weights $w_i = 1 - a = 0.25$ and $w_{i+1} = a = 0.75$.

What happens if $a > 1$? By following the simple calculation reported in Eq. (6.43) and Eq. (6.44), the reader can easily find out that in this case the VS will fall on the other side (right side, looking at Figure 6.9) with respect to atom $(i + 1)$. For example, by employing $a = 1.2$ in the example presented in Eq. (6.43), it is easy to find out that $\mathbf{r}_s = (2.2, 0, 0)$, namely on the right side of atom $(i + 1)$ in Figure 6.9.

Chapter 7

Coarse-Grained MARTINI Force Field

In this chapter a general overview about the MARTINI Coarse-Grained force-field is given. Most of the information provided here are extensively reported in a more detailed way in Marrink, Risselada, et al. (2007). Realised at the beginning only for lipids systems (Marrink, de Vries, et al., 2004), it has been extended throughout the years to other biological systems, such as proteins (Monticelli et al., 2008), carbohydrates (Lopez et al., 2009), DNA (Uusitalo et al., 2015) and also employed in polymer systems (Bochicchio and Pavan, 2017). Recently, also a polarizable version of the MARTINI force field has been developed (Michalowsky et al., 2018).

It consists of a four-to-one mapping, namely four heavy atoms per CG bead. A qualitative example is depicted in Figure 7.1. Being the level of coarse-graining very fine, the MARTINI force field accounts only for the conservative forces, neglecting therefore the dissipative and stochastic terms of the Langevin equation, reported in Eq. (4.1). The CG bead types are classified into four general main classes: polar (P), non-polar (N), apolar (C) and charged (Q). For each type, a subtype is present, in order to lead to a more accurate representation of the chemical nature of the atomic compounds. Subtypes can be either letters or numbers, indicating whether that compound is a electron donor (d), acceptor (a), both of them (da) or none of them (0), or indicating the polarity, from 1 (low) to 5 (high). Bead types interact among themselves thorough bonded and non-bonded interactions. The bonded interactions have been parametrised from atomistic simulations (bottom-up), whereas the non-bonded interactions have been evaluated by reproducing the experimental partitioning free energies referred to a wide range of chemical compounds (top-down).

Non-bonded interactions are described through shifted potentials (LJ and Coulomb interactions), and all the algorithms used for MARTINI CG simulations are the same as the ones described in Chapt. 6 for the full-atom molecular dynamics. Obviously, this is another direct consequence of the very fine level of coarsening that characterises this method. Concerning the LJ potentials, discrete levels of non-bonded interactions are available and, more specifically, there are ten levels of LJ interactions from the strongest (0) to the weakest (IX), briefly summarised in Figure 7.2, in which the LJ parameters σ

and ε are reported. A complete overview of the all bead types present in the MARTINI model, as well as all the possible non-bonded levels of interactions, are reported in Figure 7.3.

Looking at the algorithms details, larger time steps can be used compared to the full-atom MD (almost one order of magnitude grater), together with larger cutoffs for the neighbour list searching. The use of shifted potentials guarantees smooth and continuous functions during the integration of the equations of motion, leading to higher stability and solidity of this method (the energy conservation is facilitated).

Further technical details concerning the application of the MARTINI force field to the specific test case investigated in this work is extensively reported in Chapt. 10.

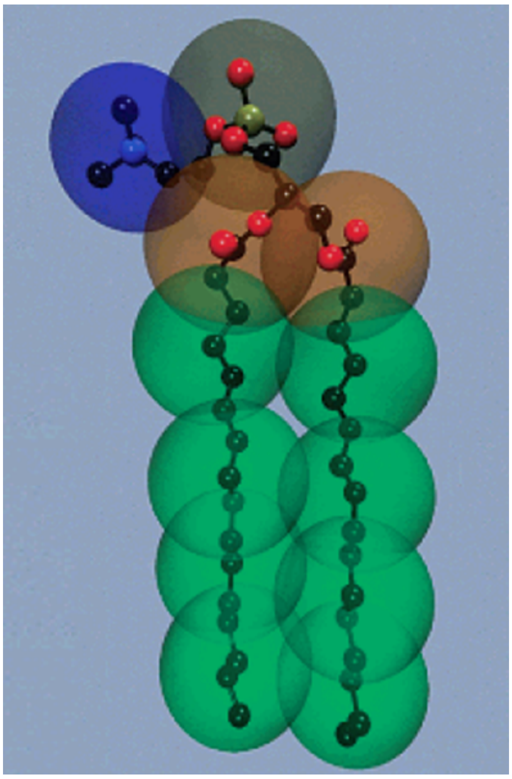
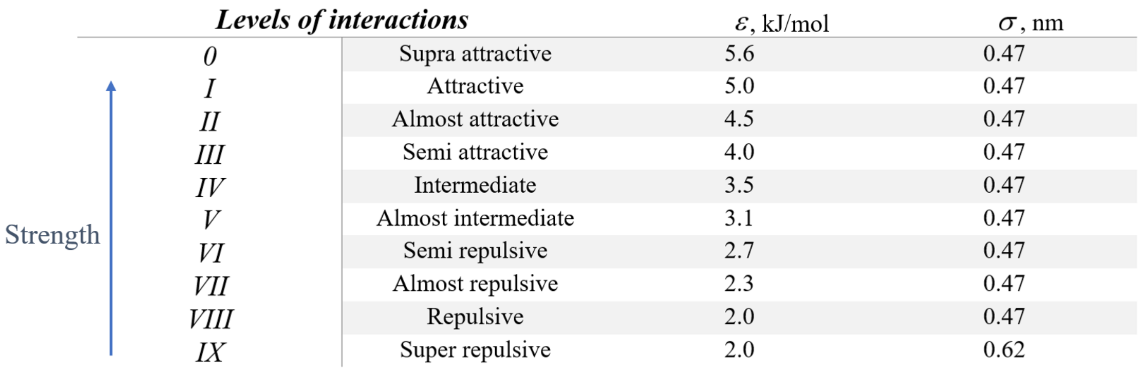


Figure 7.1: Sketch of the MARTINI force field mapping. This image has been taken and modified from Marrink, Risselada, et al. (2007).



Levels of interactions		ε, kJ/mol	σ, nm
0	Supra attractive	5.6	0.47
I	Attractive	5.0	0.47
II	Almost attractive	4.5	0.47
III	Semi attractive	4.0	0.47
IV	Intermediate	3.5	0.47
V	Almost intermediate	3.1	0.47
VI	Semi repulsive	2.7	0.47
VII	Almost repulsive	2.3	0.47
VIII	Repulsive	2.0	0.47
IX	Super repulsive	2.0	0.62

Figure 7.2: Levels of non-bonded LJ interaction parameters used in the MARTINI model.

		Q				P					N				C				
	sub	da	d	a	0	5	4	3	2	1	da	d	a	0	5	4	3	2	1
Q	da	O	O	O	II	O	O	O	I	I	I	I	I	IV	V	VI	VII	IX	IX
	d	O	I	O	II	O	O	O	I	I	I	III	I	IV	V	VI	VII	IX	IX
	a	O	O	I	II	O	O	O	I	I	I	I	III	IV	V	VI	VII	IX	IX
	0	II	II	II	IV	I	O	I	II	III	III	III	III	IV	V	VI	VII	IX	IX
P	5	O	O	O	I	O	O	O	O	O	I	I	I	IV	V	VI	VI	VII	VIII
	4	O	O	O	O	O	I	I	II	II	III	III	III	IV	V	VI	VI	VII	VIII
	3	O	O	O	I	O	I	I	II	II	II	II	II	IV	IV	V	V	VI	VII
	2	I	I	I	II	O	II	II	II	II	II	II	II	III	IV	IV	V	VI	VII
	1	I	I	I	III	O	II	II	II	II	II	II	II	III	IV	IV	IV	V	VI
N	da	I	I	I	III	I	III	II	II	II	II	II	II	IV	IV	V	VI	VI	VI
	d	I	III	I	III	I	III	II	II	II	II	III	II	IV	IV	V	VI	VI	VI
	a	I	I	III	III	I	III	II	II	II	II	II	III	IV	IV	V	VI	VI	VI
	0	IV	IV	IV	IV	IV	IV	IV	III	III	IV	IV	IV	IV	IV	IV	IV	V	VI
C	5	V	V	V	V	V	V	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	V	V
	4	VI	VI	VI	VI	VI	VI	V	IV	IV	V	V	V	IV	IV	IV	IV	V	V
	3	VII	VII	VII	VII	VI	VI	V	V	IV	VI	VI	VI	IV	IV	IV	IV	IV	IV
	2	IX	IX	IX	IX	VII	VII	VI	VI	V	VI	VI	VI	V	V	V	IV	IV	IV
	1	IX	IX	IX	IX	VIII	VIII	VII	VII	VI	VI	VI	VI	VI	V	V	IV	IV	IV

Figure 7.3: Bead types non-bonded interactions in the MARTINI force field. This image has been taken and modified from Marrink, Risselada, et al. (2007).

Chapter 8

Population Balance Model and Computational Fluid Dynamics

In this chapter Gaussian quadrature and quadrature-based method of moments (QBMM) are presented as powerful tool to solve the PBE introduced in Chapt. 5. Particular attention is dedicated to univariate and multivariate distributions, stressing two specific methods of moments: the direct quadrature method of moments (DQMOM) and the conditional quadrature method of moments (CQMOM). Turbulent and micro-mixing models are also discussed, being one of the key aspects of the test case analysed in this work. All the governing equations presented in Chapt. 5 together with the modelling algorithms shown in this chapter are implemented in finite volume CFD codes. In line with the finite volume method (FVM), the computational domain is discretised in volume elements denominated *cells*. Each cell has a number of faces (the edges of the cell domain). Governing equations (continuity, Navier-Stokes, turbulence modelling, PBE) are discretised and the partial derivatives appearing in their original form are transformed into algebraic quantities that must be integrated in each cell. Some of the terms appearing in the governing equations are discretised in FVM as fluxes through the cell domain (entering into and going out of the cell faces). Being the flux of properties entering in a finite volume cell equal to the flux that goes out from the adjacent cell, FVM turns out to be strictly conservative and, for this reason, it represents one of the most employed methods in CFD simulation codes.

8.1 Univariate Distributions and Gaussian Quadrature

As already discussed in Chapt. 5, the population balance equation (PBE) is affected by some closure problems. In general, a closure can be represented by the following integral:

$$I = \int_{\Omega_\xi} f_N(\xi) g(\xi) d\xi, \tag{8.1}$$

where $f_N(\xi)$ is a generic univariate NDF that is transported in the PBE, function of the internal coordinate ξ in the phase space domain Ω_ξ and space and time dependence is omitted for brevity; $g(\xi)$ is a function that encloses all the terms appearing in the unclosed integral I, except the NDF. It is straightforward that the integration domain changes, depending on the physical meaning of the quantities that are considered in the PBE.

In the context of real applications, it would be almost impossible to solve directly the transport equation of the NDF. It turns out to be more strategic therefore to transport the moments of the NDF instead of the NDF itself. The advantage is twofold: first, while the NDF is unknown, its moments correspond to very well-known physical quantities; second, by transporting the moments of the NDF much less equations (just the numbers of moments belonging to the moments set) have to be solved and so the computational cost decreases.

The general idea is to transport the moments of the NDF and to reconstruct it directly from its moments set by applying the quadrature method, as it will extensively explained later on. It is important to stress that the reconstruction does not want to reconstruct the actual unknown NDF, but to identify a possible reconstructed NDF, that shares the same moments with the unknown NDF, and that can be used to overcome the closure problem.

Let us define the generic k^{th}-order moment of the NDF:

$$m_k(\xi) = \int_{\Omega_\xi} \xi^k f_N(\xi) d\xi. \tag{8.2}$$

Let us also give the following fundamental definition: a set of polynomials $\{P_\alpha(\xi)\} = \{P_0(\xi), P_1(\xi), ..., P_\alpha(\xi), ...\}$, with $P_\alpha(\xi) = \sum_{l=0}^{\alpha} k_{\alpha,l} x^{\alpha-l}$, is defined as orthogonal with respect to the NDF, also known as weight or measure, if (Marchisio and Fox, 2013):

$$\int_{\Omega_\xi} P_\alpha(\xi) P_\beta(\xi) f_N(\xi) d\xi = \begin{cases} 0, (\alpha \neq \beta) \\ > 0, (\alpha = \beta), \ \ = 1 \text{ orthonormal.} \end{cases} \tag{8.3}$$

A polynomial is said to be monic if its term $k_{\alpha,0}$ is unitary. Any set of orthogonal polynomials can be written with the following recurrence formula:

$$P_{\alpha+1}(\xi) = (\xi - a_\alpha) P_\alpha(\xi) - b_\alpha P_{\alpha-1}(\xi), \ \alpha = 0, 1, 2, ..., \tag{8.4}$$

where $P_{-1}(\xi) = 0$ and $P_0(\xi) = 1$, and the coefficients are defined as follows:

$$
\begin{aligned}
a_\alpha &= \frac{\int_{\Omega_\xi} \xi f_N(\xi) P_\alpha(\xi) P_\alpha(\xi) d\xi}{\int_{\Omega_\xi} f_N(\xi) P_\alpha(\xi) P_\alpha(\xi) d\xi}, \quad \alpha = 0, 1, 2, ..., \\
b_\alpha &= \frac{\int_{\Omega_\xi} f_N(\xi) P_\alpha(\xi) P_\alpha(\xi) d\xi}{\int_{\Omega_\xi} f_N(\xi) P_{\alpha-1}(\xi) P_{\alpha-1}(\xi) d\xi} \quad \alpha = 0, 1, 2, ...
\end{aligned}
\tag{8.5}
$$

Knowledge of these coefficients is the pillar of the quadrature method itself, paving the way to the calculation of the zeros of the orthogonal polynomials. The formula reported in Eq. (8.4) leads to a set of monic polynomials that are orthogonal to the weight function in the phase space domain. Starting from Eq. (8.4) and (8.5) a set of coefficients can be written down, as a function of the moments of the NDF. To reconstruct a N-order polynomial, a set of $2N - 1$ moments must be used.

Each orthogonal polynomials set expressed as in Eq. (8.4) has distinct, real roots belonging to the phase space integration domain. It is noteworthy to stress also that each root ξ_α of the generic polynomial P_α lies in between the two roots corresponding to the polynomials of order $\alpha - 1$ and $\alpha + 1$. Finally, the key feature that comes from the theory just presented here is that the roots of the orthogonal polynomials correspond to nodes of the Gaussian quadrature method.

The generic unclosed integral shown in Eq. (8.1) can be approximate as follows:

$$
I = \int_{\Omega_\xi} f_N(\xi) g(\xi) d\xi \approx \sum_{\alpha=1}^{N} w_\alpha g(\xi_\alpha), \tag{8.6}
$$

where w_α and ξ_α correspond, respectively, to nodes and weights of the quadrature formula by using N nodes. A quadrature formula of degree of accuracy d means that the interpolation formula is exact (Eq. 8.6) by using a d(or less-)-order polynomial, and gives an error if at least one $d+1$-order polynomial is employed. The advantage of using the Gaussian quadrature method consists in the fact that for N distinct nodes a degree of accuracy equal to $2N-1$ is achieved. This means that by using for example only two nodes, the first four moments ($m_0, ..., m_3$) of the weight function are known exactly:

$$\begin{aligned} m_0 &= \sum_{\alpha=1}^{N} w_\alpha \\ m_1 &= \sum_{\alpha=1}^{N} w_\alpha \xi_\alpha \\ m_2 &= \sum_{\alpha=1}^{N} w_\alpha \xi_\alpha^2 \\ &\vdots \\ m_{2N-1} &= \sum_{\alpha=1}^{N} w_\alpha \xi_\alpha^{2N-1}. \end{aligned} \tag{8.7}$$

Solving the nonlinear system shown here (Eq. 8.7) is computationally too expensive; consequently, other more efficient algorithms must be employed to ensure a stable convergence of the method.

Therefore, a more efficient method consists in using the recurrence formula seen in Eq. (8.4) (with $\alpha = 0, 1, ..., N$) and solving the following system in matrix notation:

$$\xi \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_{N-2} \\ P_{N-1} \end{bmatrix} = \begin{bmatrix} a_0 & 1 & & & & & \\ b_1 & a_1 & 1 & & & & \\ & b_2 & a_2 & 1 & & & \\ & & b_3 & a_3 & 1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & a_{N-2} & 1 \\ & & & & & b_{N-1} & a_{N-1} \end{bmatrix} \begin{bmatrix} P_0(\xi) \\ P_1(\xi) \\ P_2(\xi) \\ P_3(\xi) \\ \vdots \\ P_{N-2}(\xi) \\ P_{N-1}(\xi) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ P_N(\xi) \end{bmatrix}, \tag{8.8}$$

which clearly corresponds to an eigenvalues problem. More specifically, the nodes $\{\xi_\alpha\}$ of the quadrature formula are the eigenvalues of a tridiagonal matrix. The latter can be transformed in a symmetric one, leading to the Jacobi matrix:

$$\mathbf{J} = \begin{bmatrix} a_0 & \sqrt{b_1} & & & & & \\ \sqrt{b_1} & a_1 & \sqrt{b_2} & & & & \\ & \sqrt{b_2} & a_2 & \sqrt{b_3} & & & \\ & & \sqrt{b_3} & a_3 & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & a_{N-2} & \sqrt{b_{N-1}} \\ & & & & & \sqrt{b_{N-1}} & a_{N-1} \end{bmatrix}. \tag{8.9}$$

This approach stabilises the numerical solution of the search of the orthogonal polynomials roots $\{\xi_\alpha\}$ by solving a well-conditioned eigenvalues problem. The weights are calculated through the relation $w_\alpha = m_0 \phi^2_{\alpha,1}$ (Wilf, 1962), where $\phi_{\alpha,1}$ is the first component of the eigenvector $\boldsymbol{\phi}_\alpha$ related to **J**. The coefficients reported in Eq. (8.5) and appearing in **J** can be calculated by means of suitable algorithms. In the next section, two of them (that use the moments of the NDF) will be briefly presented and discussed.

8.2 Computing Nodes and Weights of the Quadrature Approximation

In this section two main algorithms are presented and discussed: the Product-Difference (PD) and the Wheeler algorithm. Both of them take advantage of the orthogonality properties and the knowledge of the moments of the weight functions (NDF).

8.2.1 Product-Difference Algorithm

The Product-Difference (PD) Algorithm, proposed by Gordon (1968), consists in constructing a **P** matrix in which the components $P_{\alpha,\beta}$ are a combination of moments of the NDF. The first column is $P_{\alpha,1} = \delta_{\alpha,1}$ with $\alpha = 1, 2, ..., 2N + 1$ and $\delta_{\alpha,1}$ is the Kronecker delta. The second column is $P_{\alpha,2} = (-1)^{\alpha-1} m_{\alpha-1}$, with $\alpha = 1, 2, ..., 2N$. The other components are $P_{\alpha,\beta} = P_{1,\beta-1} P_{\alpha+1,\beta-2} - P_{1,\beta-2} P_{\alpha+1,\beta-1}$, with $\beta = 3, ..., 2N + 1$ and $\alpha = 1, ..., 2N + 2 - j$ for the j columns of **P**.

Furthermore, a set of coefficients ζ_α is evaluated:

$$\zeta_\alpha = \frac{P_{1,\alpha+1}}{P_{1,\alpha} P_{1,\alpha-1}}, \; \alpha = 2, ..., 2N. \tag{8.10}$$

The coefficients of the matrix **J** are then calculated as:

$$\begin{aligned} a_\alpha &= \zeta_{2\alpha} + \zeta_{2\alpha-1}, \; \alpha = 1, ..., N \\ b_\alpha &= -\sqrt{\zeta_{2\alpha+1}\zeta_{2\alpha}} \; \alpha = 1, ..., N-1. \end{aligned} \tag{8.11}$$

8.2.2 Wheeler Algorithm

The Wheeler Algorithm was proposed by Sack and Donovan (1971), in order to rearrange the coefficients calculation of Eq. (8.4) in a different way. The main novelty lies on the introduction of new basis functions $\{\pi_\alpha(\xi)\}$ to represent the orthogonal polynomials, taking over the classical ξ^k. This strategy implies a better and more stable numerical solution. The modified moments are then:

$$\nu_k(\xi) = \int_{\Omega_\xi} \pi_k(\xi) f_N(\xi) d\xi, \; k = 0, 1, ..., 2N - 1, \tag{8.12}$$

assuming that:

$$\begin{aligned}\pi_{-1}(\xi) &= 0\\ \pi_0(\xi) &= 1\\ \pi_\alpha(\xi) &= (\xi - a'_\alpha)\pi_\alpha(\xi) - b'_\alpha\pi_{\alpha-1}(\xi).\end{aligned} \tag{8.13}$$

By using this new approach it is straightforward that the coefficients (a'_α, b'_α) are different from the ones reported in Eq. (8.5) and, obviously, must be calculated. Wheeler (1974) proposed an algorithm to do that, calculating some intermediate quantities:

$$\sigma_{\alpha,\beta} = \int_{\Omega_\xi} \pi_\alpha(\xi)\pi_\beta(\xi)f_N(\xi)d\xi,\ \ \alpha, \beta > 1. \tag{8.14}$$

The initial values are:

$$\begin{aligned}\sigma(-1, \alpha) &= 0,\ \ \alpha = 1{,}2, ..., 2N - 2\\ \sigma(0, \alpha) &= \nu_\alpha,\ \ \alpha = 1{,}2, ..., 2N - 1\\ a_0 &= a'_0 + \frac{\nu_1}{\nu_0}\\ b_0 &= 0.\end{aligned} \tag{8.15}$$

Evaluating then the quantities $\sigma_{\alpha,\beta} = \sigma_{\alpha-1,\beta+1} - (a_{\beta-1} - a'_\beta)\sigma_{\alpha-1,\beta} - b_{\beta-1}\sigma_{\alpha-2,\beta} + b'_\beta\sigma_{\alpha-1,\beta-1}$ for $\alpha = 1, 2, ..., N - 1$ and $\beta = \alpha, \alpha + 1, ..., 2N - \alpha - 1$, the Jacobi matrix coefficients are:

$$\begin{aligned}a_\alpha &= a'_\alpha - \frac{\sigma_{\alpha-1,\alpha}}{\sigma_{\alpha-1,\alpha-1}} + \frac{\sigma_{\alpha,\alpha+1}}{\sigma_{\alpha,\alpha}}\\ b_\alpha &= \frac{\sigma_{\alpha,\alpha}}{\sigma_{\alpha-1,\alpha-1}}.\end{aligned} \tag{8.16}$$

Compared to the PD algorithm, the Wheeler one turns out to be more robust, especially at very high-order quadrature approximations.

8.3 Realizability of a Moments Set and Correction Algorithms

All the properties concerning the quadrature approximation, reported so far, are valid only if the moments set is realizable. The concept of realizability comes from physical intuitions based on the physical meaning of the low-order moments themselves. From a mathematical and rigorous point of view, a moment set is realizable if the Hankel-Hadamard determinants are non-negative (Gautschi, 2004):

$$\Delta_{k,l} = \det(\mathbf{H}_{k,l}) = \begin{vmatrix} m_k & m_{k+1} & \cdots & m_{k+l} \\ m_{k+1} & m_{k+2} & \cdots & m_{k+l+1} \\ \vdots & \vdots & \vdots & \vdots \\ m_{k+l} & m_{k+l+1} & \cdots & m_{k+l+l} \end{vmatrix} \geq 0, \quad k = 0{,}1 \ \ l \geq 0. \tag{8.17}$$

The relationship reported in Eq. (8.17) represents a necessary and sufficient condition for the realizability. A weaker condition, only necessary, is the convexity of $\ln(m_k)$ with respect to k:

$$\frac{\ln(m_k) - \ln(m_{k-2})}{2} \geq \ln(m_{k-1}), \ \ k = 2{,}3, \ldots, \tag{8.18}$$

which corresponds to the physical intuition that, being the moments the representation of physical quantities, m_0, m_1 must be positive and also the variance $m_2 - m_1^2/m_0 \geq 0$ (is null only in case of delta-function distributions). In fact, the condition reported in Eq. (8.18) is valid only for the first four-order moments, after that strongest conditions must be employed.

From a computational cost point of view, computing the Hankel determinants at each iteration would be very expensive, especially for a number of nodes greater than two. Very often, this is overcome by taking advantage of other properties of the quadrature approximation, namely evaluating the sign of the coefficients appearing in the recurrence formula, since they are strictly correlated to the positiveness of the Hankel determinants (Dette and Studden, 2002).

It might happen that, even starting from a realizable moments set, at a certain iteration the numerical solution gives back a non-realizable moments set in which at least one of the transported moments is unrealiazable. This is due to the numerical scheme and the discretization adopted in the simulation, and is named moment corruption. When this happens, a correction algorithm can be applied to adjust that corrupted moment with one that guarantees the realizability. In this sense, two different correction algorithms have been proposed by McGraw (2006) and Wright (2007). The main aim of these two algorithm is to minimize the number of corrections artificially (so, arbitrary) introduced in the numerical scheme and, at the same time, to maximize the smoothness of the moments convexity (Eq. 8.18).

8.4 Multivariate Distributions

When the NDF depends on more than one internal coordinate (e.g., size and particle velocity, or size and mixture fraction), the distribution that one aims to reconstruct is said to be multivariate. In this case, the unclosed term can be generalised as follows:

$$I = \int_{\Omega_\xi} f_N(\boldsymbol{\xi}) g(\boldsymbol{\xi}) d\boldsymbol{\xi}, \tag{8.19}$$

where $\boldsymbol{\xi} = \{\xi_1, \xi_2, ..., \xi_M\}$ is the generic internal coordinate vector that considers the M internal coordinates, $f_N(\boldsymbol{\xi})$ is the multivariate NDF (unknown) and $g(\boldsymbol{\xi})$ accounts for all the other terms appearing in the integral I.

Now the quadrature approximation has the aim of finding N weights for the N nodes of the M internal coordinates. The quadrature approximation for a multivariate problem reads as follows:

$$\int_{\Omega_\xi} f_N(\boldsymbol{\xi}) g(\boldsymbol{\xi}) d\boldsymbol{\xi} \approx \sum_{\alpha=1}^{N} w_\alpha g(\boldsymbol{\xi}_\alpha), \tag{8.20}$$

and the assumption of the functional form for the NDF is:

$$f_N(\boldsymbol{\xi}) = \sum_{\alpha=1}^{N} w_\alpha \delta(\boldsymbol{\xi} - \boldsymbol{\xi}_\alpha) = \sum_{\alpha=1}^{N} w_\alpha \prod_{\beta=1}^{M} \delta(\xi_\beta - \xi_{\alpha,\beta}), \tag{8.21}$$

where each node $\boldsymbol{\xi}_\alpha$ of the quadrature approximation is now a vector of length M. It is straightforward that the case $M = 1$ corresponds to the univariate case.

The moments of the multivariate NDF are now a set of mixed moments of the single internal coordinates:

$$m_{k_1,k_2,...,k_M}(\boldsymbol{\xi}) = \int_{\Omega_\xi} \xi_1^{k_1} \xi_2^{k_2} ... \xi_M^{k_M} f_N(\boldsymbol{\xi}) d\boldsymbol{\xi} = \langle \xi_1^{k_1} \xi_2^{k_2} ... \xi_M^{k_M} \rangle, \tag{8.22}$$

in which the new notation, $\langle \xi_\beta^{k_\beta} \rangle$ has been introduced to identify the generic k^{th}-order moment. The corresponding multivariate approximation reads as follows:

$$m_{k_1,k_2,...,k_M}(\boldsymbol{\xi}) = \sum_{\alpha=1}^{N} w_\alpha \prod_{\beta=1}^{M} \xi_{\alpha,\beta}^{k_\beta}. \tag{8.23}$$

Due to the multivariate nature of the unclosed term I, the normal Gaussian quadrature cannot be applied to this new mathematical framework. New algorithms are needed, in order to reconstruct the original multivariate NDF, by applying different assumptions. One of them will be treated in detail in the next section and it is called conditional quadrature method of moments (CQMOM), taking advantage of the conditional probability theory. Alternatives to the CQMOM are represented by the brute-force QMOM (Wright et al., 2001) and the tensor-product QMOM (Yoon and McGraw, 2004).

8.4.1 Conditional Quadrature Method of Moments

Based on the conditional probability theory, CQMOM starts from considering the so-called conditional density function. It represents the probability of finding an internal coordinate in a given infinitesimal interval when all the other internal coordinates are equal to a fixed value. In mathematical terms:

$$h(\xi_M|\xi_1,\xi_2,...,\xi_{M-1}) = \frac{f_N(\boldsymbol{\xi})}{h(\xi_1,\xi_2,...,\xi_{M-1})}, \tag{8.24}$$

which states the probability of having the internal coordinate ξ_M in a range interval ξ_M and $\xi_M + d\xi_M$ when all the others $\xi_1,...,\xi_{M-1}$ are fixed to constant values. The conditional density function integrates to one, that is:

$$h(\xi_1,\xi_2,...,\xi_{M-1}) = \int_{\Omega_{\xi_M}} f_N(\boldsymbol{\xi})d\xi_M. \tag{8.25}$$

By applying the definition reported in Eq. (8.24), the following equality holds:

$$h(\xi_{M-1}|\xi_1,\xi_2,...,\xi_{M-2}) = \frac{h(\xi_1,\xi_2,...,\xi_{M-1})}{h(\xi_1,\xi_2,...,\xi_{M-2})}, \tag{8.26}$$

and, by repeating the recurrent relation, the following equality is valid and represents the starting point of the CQMOM:

$$f_N(\boldsymbol{\xi}) = h(\xi_M|\xi_1,\xi_2,...,\xi_{M-1})h(\xi_{M-1}|\xi_1,\xi_2,...,\xi_{M-2})\cdots h(\xi_3|\xi_1,\xi_2)h(\xi_2|\xi_1)h(\xi_1). \tag{8.27}$$

By using Eq. (8.25), the following equality is trivial to show:

$$m_{\mathbf{0}} = \int_{\Omega_{\xi_1}} h(\xi_1)d\xi_1. \tag{8.28}$$

Thanks to all the properties and equations reported so far, it is now possible to build up a multivariate quadrature approach. More specifically, the definition of the conditional density function itself allows to transform the multivariate quadrature in a series of single univariate quadratures.

Therefore, the first step consists in constructing the first univariate quadrature in the direction of the first internal coordinate ξ_1. The moments set of ξ_1 is known and, by applying the Wheeler algorithm (sect. 8.2.2), it can be inverted obtaining the weights and nodes referred to the first internal coordinate. In mathematical notation:

$$\{m_{k_1,0,0,...,0}\},\ k_1 = 0,1,...,2N_1-1 \Rightarrow \text{Wheeler} \Rightarrow \{\xi_{1,\alpha_1}, w_{\alpha_1}\},\ \ \alpha_1 = 1,2,...,N_1, \tag{8.29}$$

where the first quadrature inversion has been done on N_1 nodes. The last term of the right hand side of Eq. (8.27) is then reconstructed:

$$h(\xi_1) = \sum_{\alpha_1=1}^{N_1} w_{\alpha_1} \delta(\xi_1 - \xi_{1,\alpha_1}). \tag{8.30}$$

The generic mixed moment can be expressed as:

$$\begin{aligned} m_{k_1,k_2,\dots,k_M} = & \sum_{\alpha_1=1}^{N_1} w_{\alpha_1} \xi_{1,\alpha_1}^{k_1} \\ & \int_{\Omega_{\xi_M}} \xi_2^{k_2} \cdots \xi_M^{k_M} h(\xi_M | \xi_1, \xi_2, \dots, \xi_{M-1}) h(\xi_{M-1} | \xi_1, \xi_2, \dots, \xi_{M-2}) \\ & \cdots h(\xi_3 | \xi_1, \xi_2) h(\xi_2 | \xi_{1,\alpha_1}) d\xi_2 \cdots d\xi_{M-1} d\xi_M. \end{aligned} \tag{8.31}$$

The mixed moment of order k_1, k_2 reads then as follows:

$$m_{k_1,k_2,0,\dots,0} = \sum_{\alpha_1=1}^{N_1} w_{\alpha_1} \xi_{1,\alpha_1}^{k_1} \int_{\Omega_{\xi_2}} \xi_2^{k_2} h(\xi_2 | \xi_{1,\alpha_1}) d\xi_2. \tag{8.32}$$

The integral on the right hand side of the last equation is the conditional moment of order k_2 with respect to the second internal coordinate ξ_2 conditioned on a fixed value of the first internal coordinate ξ_{1,α_1}:

$$m_{k_2}(\xi_2 | \xi_{1,\alpha_1}) = \langle \xi_2^{k_2} \rangle (\xi_{1,\alpha_1}) = \int_{\Omega_{\xi_2}} \xi_2^{k_2} h(\xi_2 | \xi_{1,\alpha_1}) d\xi_2, \tag{8.33}$$

that from now on will be identified with the shorter notation $\langle \xi_2^{k_2} \rangle_{\alpha_1}$.

Following again the same procedure as in Eq. (8.29) the second set of nodes and weights must be calculated at this step:

$$\langle \xi_2^{k_2} \rangle_{\alpha_1},\ k_2 = 0, 1, \dots, 2N_2 - 1 \Rightarrow \text{Wheeler} \Rightarrow \{\xi_{2,\alpha_1,\beta}, w_{\alpha_1,\beta}\},\quad \beta = 1, 2, \dots, N_2, \forall \alpha_1. \tag{8.34}$$

Unlike the first quadrature inversion reported in Eq. (8.29), the conditional moments $\langle \xi_2^{k_2} \rangle_{\alpha_1}$ reported in Eq. (8.34) are not known a priori, and can be calculated from the following Vandermonde linear system:

$$\mathbf{V}_1 \mathbf{R}_1 \begin{bmatrix} \langle \xi_2^{k_2} \rangle_1 \\ \langle \xi_2^{k_2} \rangle_2 \\ \vdots \\ \langle \xi_2^{k_2} \rangle_{N_1} \end{bmatrix} = \begin{bmatrix} m_{0,k_2,0,\dots,0} \\ m_{1,k_2,0,\dots,0} \\ \vdots \\ m_{N_1-1,k_2,0,\dots,0} \end{bmatrix},\ k_2 = 0, 1, \dots, 2N_2 - 1, \tag{8.35}$$

with

$$\mathbf{V}_1 = \begin{bmatrix} 1 & \cdots & 1 \\ \xi_{1,1} & \cdots & \xi_{1,N_1} \\ \vdots & \vdots & \vdots \\ \xi_{1,1}^{N_1-1} & \cdots & \xi_{1,N_1}^{N_1-1} \end{bmatrix} \tag{8.36}$$

and

$$\mathbf{R}_1 = \begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_{N_1} \end{bmatrix} \tag{8.37}$$

solved by using the algorithm proposed by Rybicki (Press et al., 1992).
The NDF referred to the first two internal coordinates is reconstructed:

$$h(\xi_1,\xi_2) = h(\xi_2|\xi_1)h(\xi_1) = \sum_{\alpha_1=1}^{N_1}\sum_{\alpha_2=1}^{N_2} w_{\alpha_1} w_{\alpha_1,\alpha_2} \delta(\xi_1 - \xi_{1,\alpha_1})\delta(\xi_2 - \xi_{2,\alpha_1,\alpha_2}). \tag{8.38}$$

This procedure can be repeated until the M^{th} internal coordinate, so that the final $f_N(\xi_1, \xi_2, ..., \xi_M)$ is reconstructed.

8.5 Direct Quadrature Method of Moments

Unlike all the other inversion algorithms, the direct quadrature method of moments (DQMOM) is not a proper inversion technique. The only initial values of weights and abscissas are inverted and, after that, weights and weighted abscissas are solved through suitable transport equations (Marchisio and Fox, 2005):

$$\begin{gathered} \frac{\partial w_\alpha}{\partial t} + \frac{\partial (u w_\alpha)}{\partial x} = \Gamma \frac{\partial^2 w_\alpha}{\partial x^2} + a_\alpha, \\ \frac{\partial w_\alpha \xi_\alpha}{\partial t} + \frac{\partial (u w_\alpha \xi_\alpha)}{\partial x} = \Gamma \frac{\partial^2 w_\alpha \xi_\alpha}{\partial x^2} + w_\alpha S_\alpha + b_\alpha, \end{gathered} \tag{8.39}$$

where $w_\alpha(t, \mathbf{x})$ and $w_\alpha(t, \mathbf{x})\xi_\alpha(t, \mathbf{x})$ are, respectively, the weights and the weighted abscissas (the dependence on the independent variable is omitted in the last set of equations for a sake of compactness), S_α represents a generic source term, Γ is a diffusion coefficient, and u is a known velocity. Also the initial distribution is assumed to be known. The two terms a_α and b_α are additive source terms necessary to be taken into account for a coherent description of the moments transport equations.

A practical application of the DQMOM in describing the turbulent fluctuation effect on a nanoparticles formation process (Lavino, Di Pasquale, et al., 2017) is extensively reported Chapt. 11.

8.6 Turbulent Mixing

In many industrial applications (nanoparticles formation, crystallization processes, etc.) the mixing at molecular level plays a key role in determining the final desired properties (e.g., particle size). This effect is taken into account by implementing proper transport equations for a passive scalar ϕ (Madadi-Kandjani et al., 2017). The mixture fraction is described by using a probability density function (PDF), $f_\phi(t, \mathbf{x}, \boldsymbol{\psi})$, whose the transport equation reads as follows (Fox, 2003; Pope, 2000):

$$\frac{\partial f_\phi}{\partial t} + \langle U_i \rangle \frac{\partial f_\phi}{\partial x_i} + \frac{\partial}{\partial x_i}(\langle u_i | \boldsymbol{\psi} \rangle f_\phi) = -\frac{\partial}{\partial \psi_i}(\langle \Gamma_i \nabla^2 \phi_i' | \boldsymbol{\psi} \rangle f_\phi) - \frac{\partial}{\partial \psi_i}(\Gamma_i \nabla^2 \langle \phi_i \rangle f_\phi), \quad (8.40)$$

where Γ_i is the molecular diffusivity of th especies i, repeated indices imply summation and the Reynolds decomposition of $\phi = \langle \phi \rangle + \phi'$ and $v = \langle U \rangle + u$ for the composition variable ϕ and the velocity v has been applied. On the left hand side, the physical meaning of Eq. (8.40) is: rate of change of the PDF plus the convection in physical space due to mean velocity (macromixing) plus the convection in physical space due to the fluctuations of the velocity (mesomixing) conditioned on the phase space composition variables $\boldsymbol{\psi}$ (note that ϕ is the composition variable, $\boldsymbol{\psi}$ is its representation in the phase space). On the right hand side, transport in phase space due to molecular mixing (micro-mixing) is represented.

For statistical homogeneous systems, $\nabla^2 \langle \phi_i \rangle$ and the convection in physical space are null. Therefore the only first term on right hand side of Eq. (8.40) must be closed. Two of the most used techniques are represented by the interaction-and-exchange with the mean (IEM) and the Fokker-Planck (FP) models.

The IEM model is a linear relaxation of the passive scalar with its mean (Fox, 2003):

$$-\frac{\partial}{\partial \psi}(\langle \Gamma \nabla^2 \phi' | \boldsymbol{\psi} \rangle f_\phi) = -\frac{\partial}{\partial \psi}\left[\frac{\varepsilon_\phi}{2\langle \phi'^2 \rangle}(\langle \phi \rangle - \psi) f_\phi\right], \quad (8.41)$$

where ε_ϕ is the scalar covariance dissipation rate, $\langle \phi'^2 \rangle$ is the covariance of the scalars. Furthermore, their ratio is related to the time scale of the turbulent micro-mixing interactions $1/\tau_T = \varepsilon_\phi / \langle \phi'^2 \rangle$. This term is usually closed by means of the mechanical-to-scalar ratio parameter C_ϕ, evaluated from the turbulent energy spectrum (Liu and Fox, 2006):

$$\frac{\varepsilon_\phi}{2\langle \phi'^2 \rangle} = \frac{1}{2\tau_T} = \frac{C_\phi \varepsilon}{2k} = \gamma_M, \quad (8.42)$$

where the γ_M is the micro-mixing rate, expressed in terms of turbulence dissipation rate ε and turbulent kinetic energy k. The last two quantities are closed thanks to a two-equation model, such as the $k - \varepsilon$ standard model (Ferziger and Peric, 2001). C_ϕ is a function of the local Reynolds number and shows different profiles, depending on the Schmidt number (Liu and Fox, 2006).

The FP model is able to overcome some limitations of the IEM method (differential diffusion, control of the shape of the PDF). Introduced by Fox (1994, 1999) it closes Eq. (8.40) as follows:

$$-\frac{\partial}{\partial\psi}(\langle\Gamma\nabla^2\phi'|\boldsymbol{\psi}\rangle f_\phi) = -\frac{c_{FP}+1}{2}\frac{\partial}{\partial\psi}\left[\frac{\varepsilon_\phi}{\langle\phi'^2\rangle}(\langle\phi\rangle-\psi)f_\phi\right] + \frac{c_{FP}}{2}\frac{\partial^2}{\partial\psi^2}(\langle\epsilon_\phi|\psi\rangle f_\phi), \tag{8.43}$$

where c_{FP} is a positive constant, $\langle\epsilon_\phi|\psi\rangle$ is a term that closes the conditional diffusion and is provided by the user to close the model (Fox, 2003). If c_{FP} is equal to zero, then the IEM model is obtained.

8.7 Transport Equation of the Moments of the NDF

The aim of this section is to start from the general transport equation of a multivariate NDF, in order to obtain the transport equation of its generic k^{th}-order moments. Considering the general PBE presented in Eq. (5.32) applied to the multivariate NDF $f_N(\boldsymbol{\xi})$ presented in section 8.4 (time and space dependence is omitted for brevity), where $\boldsymbol{\xi} = \{\xi_1, \xi_2, ..., \xi_M\}$ is the internal coordinates vector, the following PBE is obtained:

$$\frac{\partial f_N(\boldsymbol{\xi})}{\partial t} + \frac{\partial}{\partial \mathbf{x}}\cdot(\mathbf{v}f_N(\boldsymbol{\xi})) + \frac{\partial}{\partial\boldsymbol{\xi}}\cdot(\dot{\boldsymbol{\xi}}f_N(\boldsymbol{\xi})) = h_\xi(\boldsymbol{\xi}), \tag{8.44}$$

where $h_\xi(\boldsymbol{\xi})$ represents the generic source term, extensively described in section 5.3.3.

By using the multivariate moment definition reported in Eq. (8.22), multiplying Eq. (8.44) by $\xi^{k_1}...\xi^{k_M}\cdot \mathrm{d}\boldsymbol{\xi}$ and integrating out all the terms in the phase space domain of the internal coordinates Ω_ξ, the trasport equation related to the moments of the NDF is obtained:

$$\frac{\partial m_{\mathbf{k}}(\boldsymbol{\xi})}{\partial t} + \frac{\partial}{\partial\mathbf{x}}\cdot(\mathbf{v}m_{\mathbf{k}}(\boldsymbol{\xi})) - \mathbf{k}\gamma_M\cdot[m_{\mathbf{k}-1}(\boldsymbol{\xi})m_1(\boldsymbol{\xi}) - m_{\mathbf{k}}(\boldsymbol{\xi})] = h_\xi(m_{\mathbf{k}}(\boldsymbol{\xi})), \tag{8.45}$$

where the micro-mixing term (third term on left hand side) is closed thanks to the IEM model together with the micro-mixing rate parameter, γ_M, introduced in the previous section, followed by integration by parts. All this mathematical procedure is here omitted and reported in Appendix C. $\mathbf{k} = \{k_1, k_2, ..., k_M\}$ corresponds to the M^{th}-order of the moments and, as already said, $\boldsymbol{\xi} = \{\xi_1, \xi_2, ..., \xi_M\}$. It is straightforward that if $M = 1$, then the univariate moments transport equation is trivially obtained.

PART III

Results, Discussions and Future Developments

Chapter 9

Full atom molecular dynamics: the demixing problem in acetone-water mixtures

Most of this chapter has been published in Alessio D. Lavino, Luca Banetta, Paola Carbone, and Daniele L. Marchisio (2018). Extended Charge-On-Particle Optimized Potentials for Liquid Simulation Acetone Model: The Case of Acetone–Water Mixtures, *J. Phys. Chem. B*, 122 (20), pp 5234–5241.

9.1 Introduction to the problem

Molecular simulations of complex fluids, such as aqueous mixtures, and their micro-heterogeneities have received a lot of attention during the last decades (Perera, Mazighi, et al., 2012). One of the most common complex fluids is represented by water (Guillot, 2002), due to its network of hydrogen bonds (HB Angell, 1982), as well as its polarization effects. For this reason it has been widely investigated via computer simulations and several models have been developed (Berendsen, Postma, Von Gusteren, et al., 1981; Jorgensen, Chandrasekhar, et al., 1983; Mahoney and Jorgensen, 2000). Complex liquids and their mixtures are subjected to density or concentration fluctuations which, if not properly modeled, can lead to phase separation phenomena, resulting in a wrong evaluation of both thermodynamic correlations and transport properties, such as the Kirkwood-Buff integrals (KBI), strictly related to concentration fluctuations (Allison et al., 2005; Gupta, 2012; Mountain, 2010), and, consequently, the radial distribution functions (RDF).

It is surprising to find out how computer simulations of common aqueous mixtures, such as acetone-water mixtures, can be non-trivial. More specifically, in classical molecular dynamics (MD) simulations (Ferrario et al., 1990; Freitas et al., 1999; Jorgensen, Briggs, et al., 1990) numerous acetone force fields showed an evident de-mixing and phase separation in acetone-water mixtures (Di Pasquale, Marchisio, Barresi, and Carbone, 2014; Jedlovszky et al., 2009; Perera and Sokolić, 2004; Pereyra et al., 2011; Weerasinghe and Smith, 2003), despite water and acetone are perfectly miscible at any proportion at room temperature.

Phase separation depends on the stability of a given mixture and the stability is, in turn, strictly related to the local order of the molecules. The fact that water undergoes clustering effects in the simulation of acetone-water mixtures is due to the wrong description of acetone molecules, which have a low dipole moment in aqueous mixtures with respect to the expected trend, not being affected by the presence of a polar solvent. In other words, the polarization effect of water on acetone molecules is usually not properly taken into account by all-atom acetone force fields. Similarly, water molecules are not affected by such low polar acetone interactions and tend to self-organise together, driven by a self-induced polarization effect due to their own dipole moment. Consequently, in MD simulations the local order increases (together with the stability of the system) leading to unphysical de-mixing and phase separation.

Molecular mechanics/quantum mechanics (QM/MM) iterative calculations (Georg et al., 2006) showed that the electronic polarization on the acetone molecule is strongly influenced by the aqueous environment and that the acetone dipole moment spans from 3.3 D (neat liquid acetone) to 4.8 D (one molecule of acetone surrounded by 200 water molecules). These results are not in accordance with the original OPLS acetone force field that predicts a constant dipole moment of 2.35 D for the acetone molecule, as reported in a previous work (Perera and Sokolić, 2004), and, consequently, leads to phase separation at low acetone concentrations in the mixture.

Since the classical force fields turn out not to be adequate for describing these solvents mixtures, research efforts have focused on the development of new ones specifically addressing the miscibility issue of acetone-water mixtures. Pereyra et al. (2011) developed a modified version of the CHARMM27 acetone force field, thanks to a variable charge distribution on the acetone molecule and validated in terms of addressing the de-mixing issue from a thermodynamic point of view (Pinke and Jedlovszky, 2012). The method introduced by Pereyra et al. (2011) consisted in varying the charge distribution on the acetone molecule with the acetone molar fraction, in order to fit the excess mixing enthalpy. This approach relied on the idea that, in order to gain a better description of the several features related to the micro-heterogeneity of acetone-water mixtures, it is mandatory to introduce a polarization effect in the acetone force field. However, this approach requires changing the charge distribution when the composition of the mixture is modified. Here we propose an alternative way, namely to implement a pseudo-polarizable model for the acetone molecule which adjust the acetone polarizability without the need to modify the atomic charge distribution, but just acting on the charge displacement along the carbonyl group, as explained later on.

Numerous techniques have been recently developed to introduce polarizability effects into classical MD simulations, such as the point polarizable dipole model (PPD)

(Van Belle et al., 1987; Vesely, 1977; Warshel and Levitt, 1976), the charge-on-spring (COS, Straatsma and McCammon, 1990) (also called Drude oscillator, Lemkul et al., 2016) model and the fluctuating charge (FQ, Rick et al., 1994) model. The approach proposed here consists in using a charge-on-particle (COP) model by means of a Virtual interaction Site (VS), which lies on the basic idea of the COS approach, in which molecules polarization is modelled by charged particles attached to the nuclei of their core atoms by harmonic springs. Applications of polarizable force fields spread out on a wide range, especially in biological systems, but they have never been applied to acetone-water mixtures. Lying on the idea of the COS model, the COP model is based on the concept of induced polarizability by means of a charged particle (the virtual site, VS). In the case of the acetone molecule, the charged particle is placed along the carbonyl group and attached to the carbonyl carbon.

The aim of this work is therefore to reproduce the induced polarizability of acetone, optimizing the charge equilibrium position. This allows us to model the variable stretching of the electron cloud around the carbonyl group, depending on the environment in which the acetone molecule is placed. It merges the idea of polarization which the polarizable force fields lie on, preserving at the same time the simplicity and originality of the classical OPLS force field. Consequently, the novelty of this approach consists in acting on the charge position, instead of modifying the charge distribution, and without modifying the force field itself like previous efforts did (Pereyra et al., 2011), stating therefore an improvement with respect to them. It is worthwhile to mention that the model proposed here cannot be defined as pure-polarizable, since it does not adjust the polarizability of the acetone on-the-fly at different acetone molar fractions, and it is therefore labelled as charge-on-particle (COP) model.

9.2 The Charge-on-Particle Approach and Numerical Details

The model developed here is based on the general idea of the self-consistent field (SCF) procedure, in line with the Born-Oppenheimer approximation, i.e. keeping the electronic and nuclear degrees of freedom separated. A charge-on-spring (COS) approach is adopted by placing a charge along the carbonyl axis, attached to a core atom by a harmonic spring. Being the charge carried on a virtual particle, that is a virtual interaction site (VS), this approach will be named as charge-on-particle (COP) model from now on. VS is used in MD simulations in several ways. It is defined "virtual" because it is massless, but, "interaction site" because it interacts with other atoms, by means of a charge placed on it. As explained in section 6.8.1, in general, the VS is a position-constrained massless particle, built upon a given number of surrounding atoms N, as reference positions to constrain the VS itself. In this case, $N = 2$ and they correspond to the carbon and the oxygen of the carbonyl group. Therefore, the position of the VS is uniquely defined with respect to the carbonyl group as reference atoms, and, more specifically, the carbon represents its core atom. A harmonic non-chemical bond is introduced between the VS and its core atom. All the computational details concerning the implementation of the VS in acetone topology are reported in Appendix D. The force acting on the VS is redistributed by the code on these reference atoms in a consistent way (Berendsen and van Gunsteren, 1984), following the procedure presented in section 6.8.1 and also explained in the GROMACS simulation package (Hess, Kutzner, et al., 2008) manual.

Thanks to this modelling approach, the total potential of the system is able to consider the presence of the new particle (VS) in the conservation laws, so that the linear and angular momenta are conserved. Both experiments (Applequist et al., 1972; Thole, 1981) and solvent accessible area (Connolly, 1983; Eisenhaber et al., 1995) calculations of a single acetone molecule in water environment show how polarizability of acetone is enhanced on the carbonyl axis. Therefore, the COP model aims to reproduce this polarization effect around the carbonyl group, which stretches its electron cloud depending on the polarity of the environment (induced electronic polarization, Israelachvili, 2011). Hence, the stretching of the electron cloud is modelled by the equilibrium position of the charge (carried on the VS), which vibrates via a harmonic spring around it.

In the case of $N = 2$, the equilibrium position of the VS is characterised by the geometrical parameter introduced in section 6.8.1, which we denote with the symbol a. The parameter a represents the normalized distance over the carbon-oxygen bond of the VS from its core atom (in this case, the carbonyl carbon). If $a = 0$ the VS is placed on the carbon; for $a = 1$ the VS lies on the oxygen, for $0 < a < 1$ the VS lies in between the carbon and the oxygen, and finally if $a > 1$ then the VS is placed on the other side of the oxygen with respect to the carbon. A numerical example is also reported in section 6.8.1. This scheme is depicted in Figure 9.1. By following this procedure, the VS is always constrained to lie on the carbonyl axis. The choice of this VS type, built upon only two atoms, assures the best trade off in terms of both computational cost and physical interpretation of the charge movement.

As stated above, the VS is attached to the carbonyl carbon (Figure 9.1) and placed at an equilibrium distance from it equal to $a \cdot r_0$ (where r_0 is the equilibrium carbon-oxygen distance). The VS charge, q_s, must be chosen in order to respect the equality $q_s + q_{o,l} = q_o$, where $q_{o,l}$ is the charge left on the oxygen and q_o is the original oxygen atom charge, before the introduction of the VS (original OPLS force field). This new approach consists in shifting the totality of the oxygen charge (q_o in this case) on the VS, such that $q_s = q_o$; therefore $q_{o,l}$ is equal to zero, and no charge is left on the oxygen after the introduction of the VS. At the carbon-oxygen-VS group the following isotropic polarizability is associated:

$$\alpha_s = \frac{q_s^2}{k_s}, \tag{9.1}$$

where k_s is the force constant of the carbon-VS harmonic spring, in line with

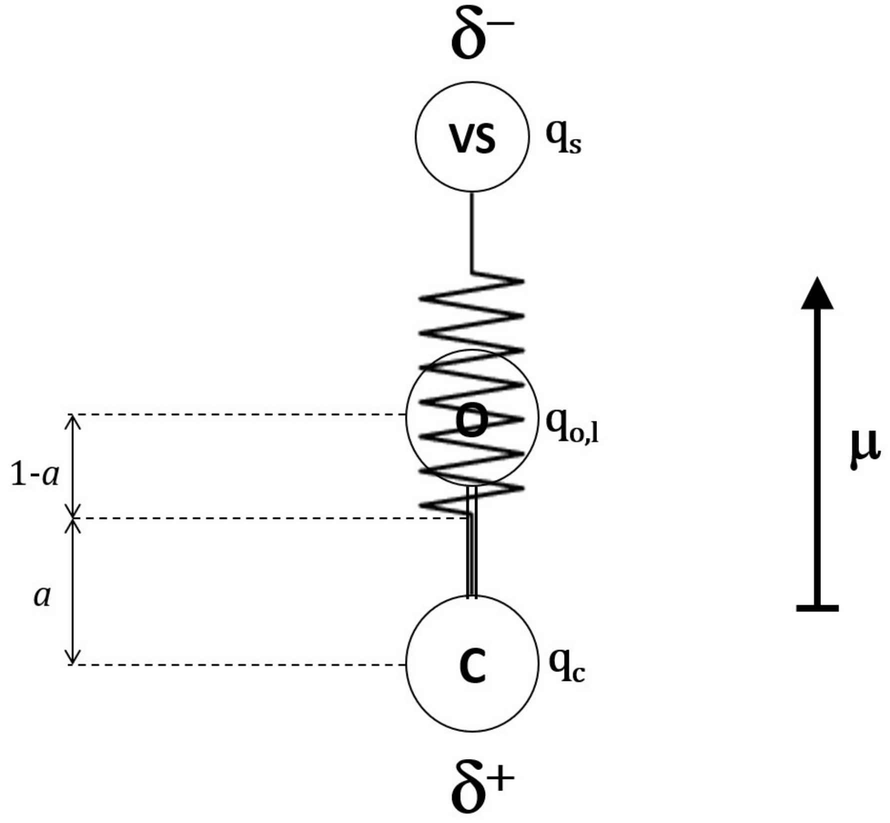


Figure 9.1: Schematic representation of the charge-on-particle (COP) model. The VS bead is attached to the carbon of the carbonyl group through a harmonic spring and a represents its normalized position, depicted here as in the general scheme shown in section 6.8.1. The positive charge, $\delta^+ = q_c = -q_o$ is placed on the C; the negative charge, $\delta^- = q_o = q_s$ (initial oxygen charge, without COP model), is totally shifted on the VS ($q_{o,l} = 0$). The charge displacement contributes to the dipole moment, $\boldsymbol{\mu}$. An electric field acting on the VS makes it vibrate around its equilibrium position, $a \cdot r_0$, creating an induced pseudo-polarization effect. Being a greater than unity, the VS lies on the other side of the oxygen of the carbonyl group.

the isotropic polarizable force fields theory (Lamoureux, MacKerell Jr., et al., 2003; Lamoureux and Roux, 2003), q_s is the charge carried on the VS and $\alpha_s = 6.39 \mathring{A}^3$ is the experimental polarizability referred to the acetone molecule (Israelachvili, 2011; Thole, 1981). Being the polarizability defined as the induced dipole moment to the electric field ratio, it is usually normalised to a factor of $4\pi\varepsilon_0$, where ε_0 is the vacuum dielectric constant, and instead of being expressed in $C^2 m^2 J^{-1}$ units, is more often indicated with volume units (e.g., in $\mathring{A}^3$). The energy associated to the carbon-VS harmonic spring is $V_{self} = \frac{1}{2} k_s (r - a \cdot r_0)^2$, where a is the characteristic normalized and dimensionless VS parameter, $r_0 = 0.1229\ nm$ (equilibrium distance of the carbonyl bond in the OPLS force field), and $k_s = 4804$ kJ mol^{-1}nm^{-2} has been calculated achieving the appropriate polarization response (Applequist et al., 1972), according to Eq. (9.1).

By shifting the oxygen partial charge onto the VS, a relative motion is created between the oxygen charge (displaced now on the VS) and the oxygen atom, thanks to both the VS parameter a (greater than one) and the spring constant of the VS harmonic potential, k_s which is different from the one of the carbonyl bond. This is more extensively explained in the sect. 9.3.1, where a sensitivity analysis with respect to both a and k_s is also reported. Since k_s is derived by a physical measurable quantity (i.e. acetone polarizability) and seems not to affect much the final predictions, more attention is paid to the value of a, upon which model predictions are extremely sensitive. The parameter a is therefore varied in order to fit a target property of the mixture, namely the density of the mixture (Figure 9.4), and, consequently a set of a values is obtained, depending on the acetone molar fraction (Table 9.2).

Moreover, the necessity to change the value of the parameter a lies on the aim of creating a variable charge displacement, in order to reproduce higher acetone dipole moments in more polar environment (lower acetone molar fractions). As already mentioned, it turns out that the spring constant, k_s, plays a secondary role when a polar molecule interacts with the VS and, on the other side, it is the equilibrium position, $a \cdot r_0$, that mainly determines the polarization response. This is related to the following fact: all the chemical bonds are constrained to their equilibrium positions (as usually happens in MD simulations) in order to lower the computational cost; the VS is a position-constrained particle by definition and, in this case, it is constrained to the C-O bond. Being the bonds vibrations damped by the constraint algorithms (LINCS, in this work, as reported in the operating conditions, sect. 9.2.1) it turns out that such constraints damp the displaced-charge oscillations as well, leading to very small oscillations for both C-O (atom-atom) and C-VS (atom-charge) bonds. These small fluctuations, damped by the constraints, make the charge oscillations and, consequently, the effect of the spring constant k_s negligible, with respect to the charge equilibrium position $a \cdot r_0$.

9.2.1 Numerical Details about the COP model

Simulations are performed in this study by using GROMACS 4.5.6 MD package (Hess, Kutzner, et al., 2008). The VMD program is used to produce the graphical images of the molecular systems. The OPLS (Jorgensen, Maxwell, et al., 1996) acetone force field is employed, coupled with SPC/E (Berendsen, Grigera, et al., 1987) force field for water. All the simulations are conducted in a NPT ensemble, with a total of $N = 1500$ molecules in a cubic box with periodic boundary conditions and box length equal to 4 nm; this leads to a sufficiently large system to guarantee the uniformity condition for the RDF. Weak harmonic potential functions are used for bonds, angles and improper dihedrals, whereas the Ryckaert-Bellemans potential is implemented for the proper dihedrals.

The whole range of acetone molar fractions has been investigated, with values of x_A considered equal to $0.10, 0.25, 0.50, 0.75, 1$. The system is equilibrated by using a leap-frog algorithm for 100 ps with a 1 fs timestep. The simulations are performed for 20 ns with a 1 fs timestep, and bonds are constrained using the LINCS algorithm. Both simulation time and box length are large enough to detect the de-mixing phenomenon. The electrostatic interactions are evaluated by using the Particle-mesh Ewald (PME) summation, with Van der Waals and electrostatic cut-off radii equals to 0.9 nm; long-range dispersion corrections are taken into account for both energy and pressure. The non-bonded interactions are settled by Lennard-Jones and Coulomb potential functions. In particular, the Lennard-Jones interactions are taken into account by the combination rule defined as follows:

$$\sigma_{ij} = (\sigma_i \sigma_j)^{1/2}, \tag{9.2}$$

$$\varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2}, \tag{9.3}$$

where σ and ε are the Lennard-Jones parameters and i and j represent two different non-bonded atoms.

The temperature is fixed at 300 K by means of a velocity-rescale algorithm with a time constant equal to 0.1 ps. The Berendsen barostat is implemented with a coupling constant equal to 1.0 ps to keep the system at the reference pressure of 1 bar; the compressibility is set to $4.5 \cdot 10^{-5}\text{bar}^{-1}$. Diffusion coefficients are calculated from the mean square displacement, as it is customary.

The viscosity of the acetone-water mixture has been calculated via non-equilibrium molecular dynamics (NEMD, Allen and Tildesley, 1987; Song and Dai, 2010) simulations, since it has been shown (Hess, 2002) that the periodic perturbation method represents the best choice for the evaluation of the viscosity of liquids, in terms of accuracy and computational cost. It briefly consists in perturbating an equilibrated system (it is crucial to not bring the system too far from an equilibrium condition to not affect the final results) by applying a periodic force. This force is characterized by an acceleration amplitude that in a MD simulation is considered by a term called cos-acceleration. After the system was equilibrated in a NPT ensemble for few hundred picoseconds (as much as it needs to reach the equilibrium), it has been perturbed by setting a cos-acceleration value equal to 0.01 nm ps^{-2}. This value represents a good compromise between too low values which could not guarantee a reasonable perturbation and, on the other hand, too high values which would lead to very high shear rates affecting the structure of the liquid.

9.3 Results and discussion

Before going into the modelling results details in terms of dynamical properties and de-mixing behaviour, a deeper look at the polarization response introduced by the COP model is done in a preliminary way in the following subsection, in order to better clarify limits and advantages of this kind of approach.

9.3.1 Polarization response

In this section, further details regarding the polarizability effect implemented by the COP model, as well as the roles of the VS characteristic parameter a and of the associated harmonic spring will be shown and discussed. First of all, it is interesting to evaluate the dependence of the mixture density and the acetone dipole moment on the VS parameter a. This is done at acetone molar fraction equal to 0.25 and the relative results are shown in Table 9.1.

Virtual Site parameter, a	**Density ($kg \cdot m^{-3}$)**	**Dipole Moment (D)**
1.00	895.89	3.10
1.10	902.91	3.37
1.16	913.89	3.54
1.20	921.37	3.65
1.25	931.45	3.79

Table 9.1: Variation of the density of the mixture and the acetone dipole moment with the VS parameter, a, at acetone molar fraction, x_A, equal to 0.25.

By looking at Table 9.1, it is clear that both density and dipole moment increase proportionally with a. This is explained by the following reason: increasing the characteristic parameter a, the charge is displaced at a greater equilibrium distance from the carbonyl carbon, equal to $a \cdot r_0$; this implies a higher acetone dipole moment (third column in Table 9.1) and, consequently, stronger electrostatic interactions with the surrounding environment. This brings the molecules to interact more among themselves leading to a lower volume and, therefore, the density increases, as shown in Table 9.1.

As already explained in the final part of the previous section, constraining all chemical bonds makes the bonds oscillation almost vanish and this reflects directly to the charge oscillations as well. Consequently, model predictions were found to be less sensitive to the value of the VS harmonic constant k_s. Even changes of orders of magnitude did not significantly changed the predictions. Attention was therefore paid to the value of a only.

In order to assess the polarization effect induced by the presence of water in the mixture, both carbon-oxygen and carbon-VS harmonic bonds have been monitored over the whole simulation time (20 ns) for this mixture composition, i.e. x_A, equal to 0.25, and results are reported in Figure 9.2. As it can be seen, the decoupling between the two harmonic bonds is evident, demonstrating that the VS vibrates around the position $a \cdot r_0 = 1.16 \cdot 0.1229\ nm = 0.1426\ nm$ (top panel), while the carbon-oxygen bond is vibrating around its equilibrium position given by $r_0 = 0.1229\ nm$, despite the corresponding fluctuations are very small. Although the oscillations are limited and do not influence the final dipole moment values, it is noteworthy to stress that the carbon-VS bond is independent on the carbon-oxygen one, as proven by the insets of Figure 9.2 and pointed out by the red and blue flags. For instance, enlarging the interval around the red flag, at 0.35 ps, the carbon-VS distance reaches a local maximum, whereas the carbon-oxygen one is still increasing. The same behaviour is detected around the blue flag (look at the insets respectively in the bottom and top panels). This univocally proves that the two fluctuations are out of phase, thanks to the polarization response caused by the presence of water. For the sake of clarity in the results depicted in both Figure 9.2 and Figure 9.3, only the first 30 ps are represented; however, this behaviour has been observed for the whole simulation time (20 ns).

This polarization effect vanishes in pure acetone ($x_A = 1$, no water in the box and $a = 1$), where the carbon-oxygen and carbon-VS behave exactly in the same way. Therefore, the bonds fluctuations are in phase at any time instant (Figure 9.3) and the VS model is consistent with the original OPLS force field in pure acetone.

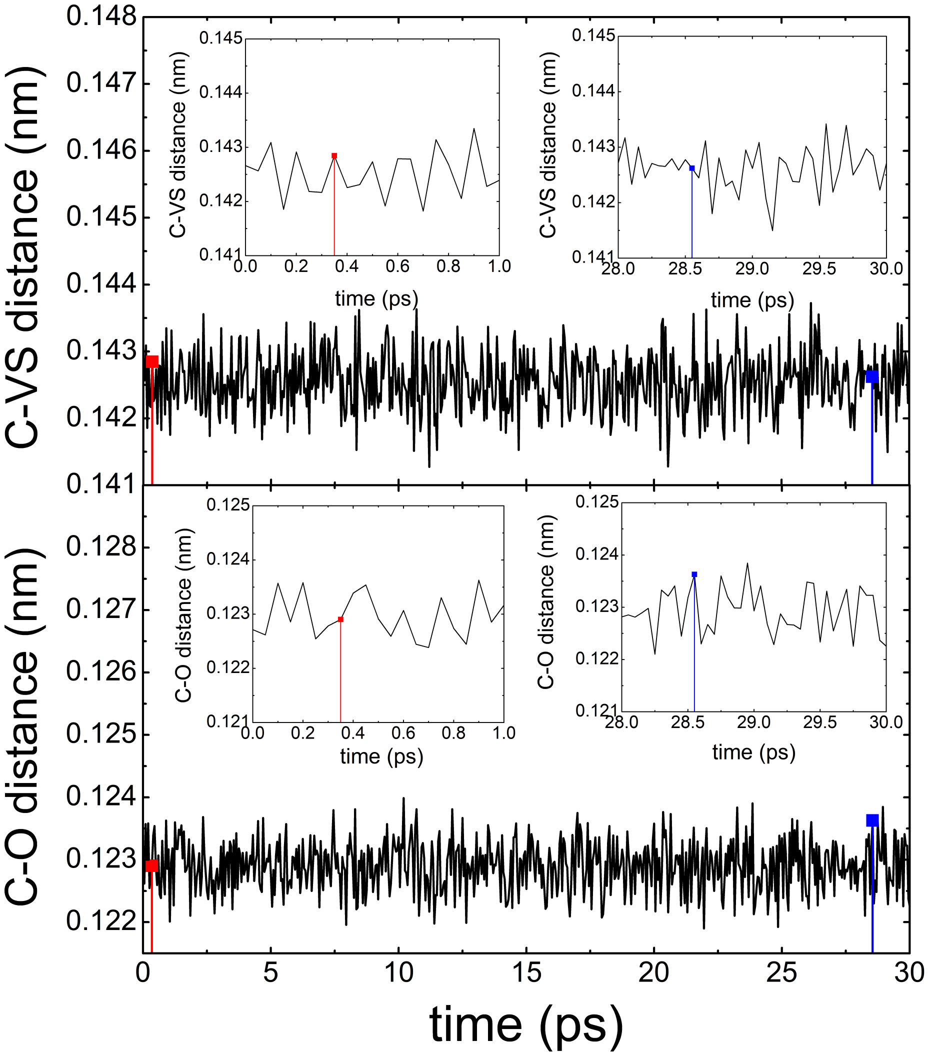


Figure 9.2: Carbon-oxygen (bottom) and carbon-VS (top) distances over simulation time at acetone molar fraction equal to 0.25. The first 30 ps are represented here, in which it is clear how the carbon-VS bond is independent on the carbon-oxygen one, as stated by the red and blue flags, enlarged in the insets. Indeed, when the carbon-VS distance reaches a local maximum (red flag, top inset), at the same time instant the carbonyl-oxygen bond is still increasing (red flag, bottom inset) and vice-versa for the blue flag. The same behaviour is observed for the whole simulation time (20 ns).

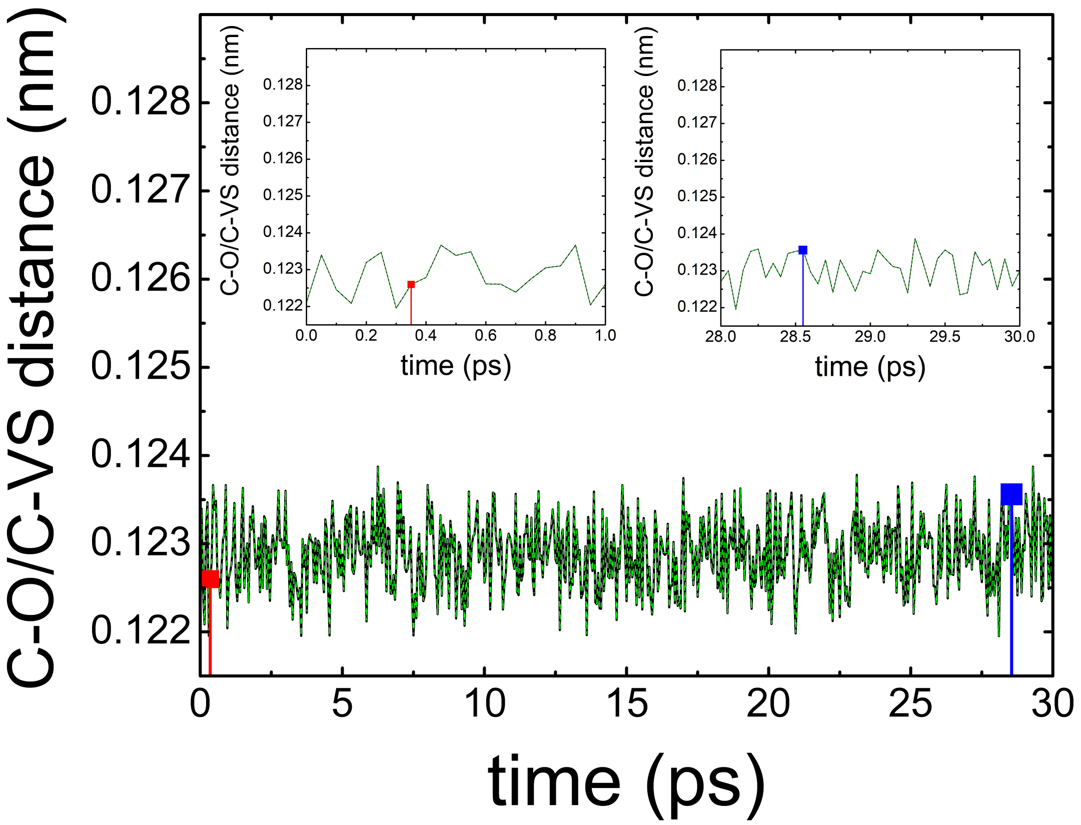


Figure 9.3: Carbon-oxygen (black solid line) and carbon-VS (green dashed line) distances over simulation time at acetone molar fraction equal to 1.00. The first 30 ps are represented here, proving that when there is no water, no polarization effect is detected. Consequently, the carbon-VS curve collapses on the carbon-oxygen one, fluctuating in phase as stated by the red and blue flags, enlarged in the insets.

9.3.2 Dynamical properties and de-mixing analysis

In line with the adopted modelling approach, several target functions can be chosen in order to fit the VS parameter *a*. In this case, the density was identified to be the most suitable target property due to its general relevance and role in the de-mixing issue. The result of this fitting is depicted in Figure 9.4, in which the experimental density (blue filled symbols, Thomas and McAllister, 1957) is fitted by the VS parameter (open purple squares) at different acetone molar fractions, x_A. The density profile corresponding to the original SPC/E-OPLS force field (open black squares) is also reported. Looking at Figure 9.4, it is evident that the original SPC/E-OPLS force field is inadequate to quantitatively reproduce the density of the mixture, especially far away from the pure components; this gap is overcome by fitting the parameter *a* of the COP model with the experimental data. More specifically, the fitting procedure consists in a very straightforward trials and errors approach, by taking advantage of the monotonic profile of the density with respect to a at constant acetone molar fraction (look at Table 9.1).

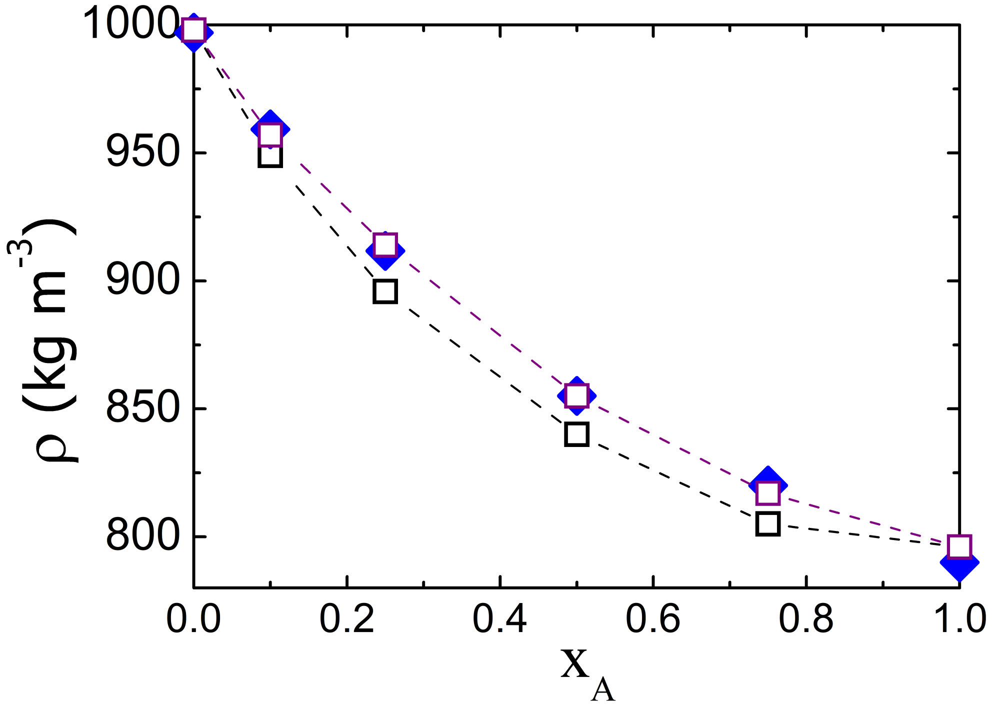


Figure 9.4: Experimental density of the acetone-water mixture (filled blue diamonds) versus the acetone molar fraction, x_A, fitted by the COP model (open purple squares). The original SPC/E-OPLS model predictions (open black squares) are also reported.

The a parameter values, corresponding to the different acetone molar fractions investigated in this work, are summarized in Table 9.2. Increasing the acetone molar fraction, x_A, the parameter a clearly decreases, indicating that the acetone molecule is more polarized when the water concentration is higher. At $x_A = 1$ (pure acetone) a is equal to 1, namely the mobile charge is on the oxygen, relaxing into the original OPLS force field. As explained in more details in sect. 9.3.1, the decoupling between the charge and its reference atom (i.e. the oxygen) brings out a more stretched electron cloud. The stretching of the electron cloud around the carbonyl group is enhanced by acting on the equilibrium position $a \cdot r_0$, and, more specifically, increasing the value of the VS parameter a leads the molecules to interact among themselves in a more attractive way resulting in higher density values in better agreement with experiments.

x_A	0.10	0.25	0.50	0.75	1.00
a	1.30	1.16	1.08	1.03	1.00

Table 9.2: VS parameter, a, at different acetone molar fractions, x_A, obtained by the fitting against the experimental density of the mixture.

The polarizability response of the mixture in function of acetone molar fractions is also demonstrated by the profile of the acetone dipole moment obtained from the COP model (empty purple squares), as shown in Figure 9.5. The acetone dipole moment spans in the range of 3-4 D. This is qualitatively in line with the results obtained by Pereyra et al. (2011) through a modified charge distribution on acetone molecules by means of a CHARMM27 acetone force field (black empty circles) but slightly in contrast with the value of acetone dipole moment reported by Perera and Sokolić (2004) that predicts a value of 2.35 D for pure OPLS acetone as also reported in the introduction of the current chapter; this might be addressed to the different operating conditions as well as the different versions of the tools used. Moreover, what we obtained here is confirmed by the trends experimentally observed (Marcus, 1985), as well as by quantum mechanical calculations (Georg et al., 2006) of the induced polarizability of the acetone molecule surrounded by water molecules. Both experiments and quantum mechanical calculations state indeed that the acetone dipole moment must increase as the acetone molar fraction decreases in acetone-water mixtures.

Figure 9.6 reports the acetone diffusion coefficient as a function of the acetone molar fraction. As shown in Figure 9.6, the COP model (empty purple squares) can reproduce qualitatively better the experimental values (blue filled diamonds, Toryanik and Taranenko, 1987), with respect to the original SPC/E-OPLS force field (empty black squares), which is shown to overestimate the acetone diffusion coefficient except for $x_A = 1$. On the contrary, the COP model underestimates the diffusion coefficient in all the acetone molar fraction range, although the agreement is very good, especially at low acetone molar fractions. This is due to the fact that the COP model takes into consideration stronger electrostatic interactions between molecules than the original one, so that the mobility of the acetone molecules is hindered, resulting in smaller diffusion coefficients. Furthermore, it is qualitatively in good accordance with the experimental trend (the minimum is correctly reached), similarly to the technique proposed by Pereyra et al., 2011, which however leads to a crossover between simulations results and experiments, resulting in a flatter diffusivity profile.

Regarding the viscosity coefficient η, it was calculated by means of non-equilibrium molecular dynamics (NEMD) calculations. In line with the NEMD theory, a trade off value for the cos-acceleration has been identified thanks to a sensitivity analysis. The viscosity profile versus cos-acceleration values is here reported in Figure 9.7 at acetone molar fraction equal to 0.10; the same qualitative results are obtained for all the others acetone molar fractions investigated in this work. By looking at Figure 9.7, it is visible

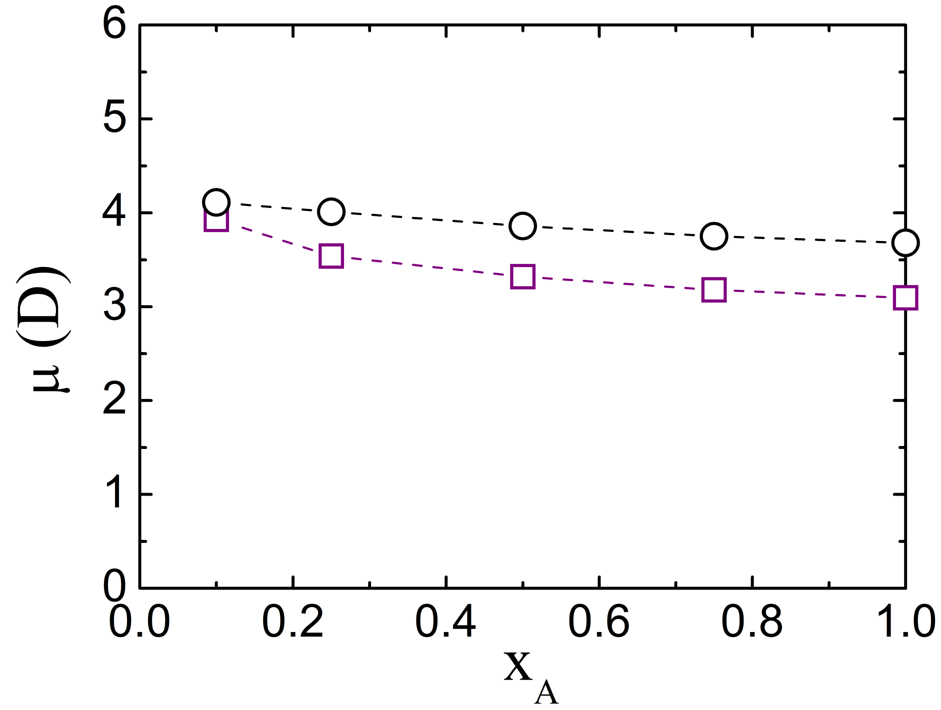


Figure 9.5: Average dipole moment, $\boldsymbol{\mu}$, of acetone molecules versus the acetone molar fraction, x_A, obtained from the COP model (open purple squares) and by Pereyra et al. (2011) with the modified CHARMM27 (open black circles). The experimental value of the acetone dipole moment is 2.88 D (gas phase, Marcus, 1985), slightly lower than the QM/MM calculations (2.98 D) to neat liquid (3.33 D, Pereyra et al., 2011), to infinite dilution (4.8 D, Georg et al., 2006). These last values are omitted in the picture, for a sake a clarity.

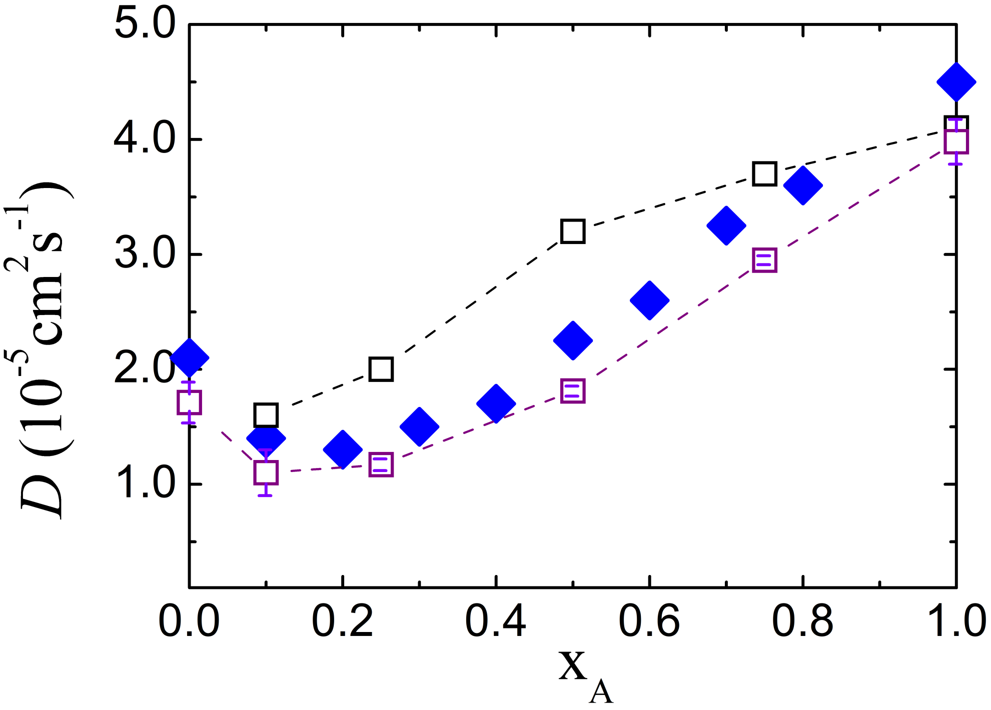


Figure 9.6: Comparison between the experimental acetone diffusion coefficient (Toryanik and Taranenko, 1987) versus the acetone molar fraction, x_A, (filled blue diamonds), with the COP model predictions (open purple squares), together with the original SPC/E-OPLS (Perera and Sokolić, 2004) model predictions (open black squares).

that a value of 0.01 nm·ps^{-2} represents a good trade off, since smaller values of cos-acceleration bring to greater errors of viscosity estimation. On the other hand, the error estimation decreases with cos-acceleration, leading however to a wrong evaluation of viscosity for cos-acceleration values greater than 0.1 nm·ps^{-2}.

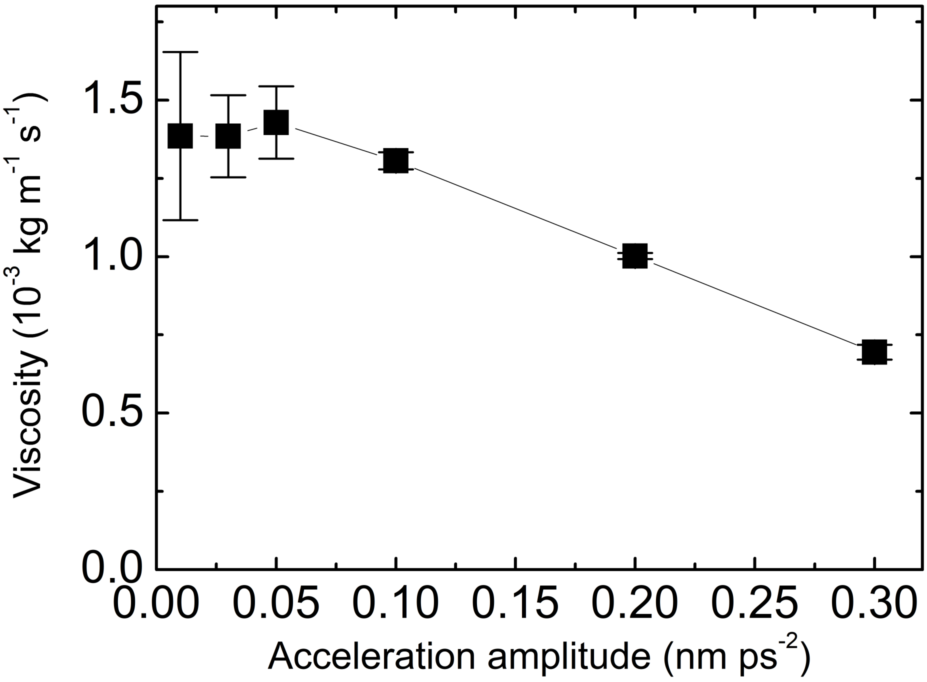


Figure 9.7: Viscosity dependence on acceleration amplitude at acetone molar fraction equal to 0.10 calculated by NEMD. Error bars are reported.

In Figure 9.8 the COP model results (empty purple squares) are compared against the experimental data (K. Howard and McAllister, 1958; Noda et al., 1982), showing excellent agreement. The largest mismatch is detected at $x_A = 0.50$, but it still can be considered an acceptable agreement. This is a very encouraging result, which demonstrates that the COP model proposed here reproduces very well not only the density of the mixture (target property), but also, the molecular diffusion and the rheological behaviour of acetone-water mixtures.

The COP model not only results in better predictions for density, diffusion and viscosity coefficients, but it also fixes the unphysical de-mixing of the mixture observed with the original SPC/E-OPLS model. This is shown here by looking at the snapshots of the MD simulations, the radial distribution functions, as well as the density profiles along the simulation box. The snapshots related to the four acetone molar fractions investigated here are reported in Figure 9.9, taken at very long simulation times (20 ns). It is evident that no de-mixing occurs anymore, by looking at the acetone molecules (the green atoms univocally refer to the acetone carbons) which are uniformly distributed in the simulation box after 20 ns. It is also possible to see the VS (pink bead) attached to its core nuclei on the carbonyl group of the acetone. The affinity with water is enhanced thanks to the VS approach since more HB between water and acetone are clearly visible (O red bead - H white bead — VS pink bead – O red bead) in the snapshots, overcoming the characteristic hydrophobicity of the original OPLS force-field.

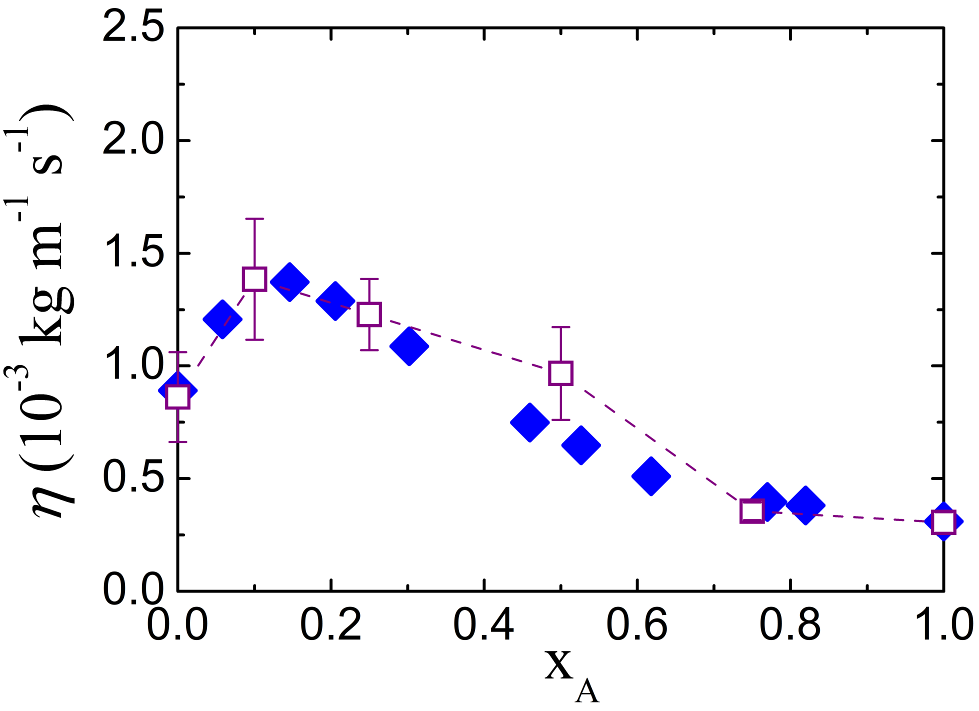


Figure 9.8: Comparison between the experimental viscosity coefficient (filled blue diamonds) of the acetone-water mixture (K. Howard and McAllister, 1958; Noda et al., 1982) as a function of the acetone molar fraction, x_A, with the COP model predictions (open purple squares) Error bars referred to the average viscosity estimation are reported; at x_A = 0.75 and 1.00 they are too small (10^{-5} order of magnitude) to be graphically detected.

This is also confirmed by the RDF referring to acetone-acetone molecules, $g_{AA}(r)$, water-acetone molecules, $g_{WA}(r)$, and water-water molecules, $g_{WW}(r)$, as it can be seen in Figure 9.10. Looking at $g_{AA}(r)$, no first peak is present anymore, unlike previous works (Perera and Sokolić, 2004), observed especially at low acetone molar fractions, as well as no long-range tail effect is visible, confirming that acetone-acetone clustering is clearly milder with the COP model, and, therefore pointing out that the hydrophilicity of the OPLS acetone model has been enhanced, as shown in the snapshots before (Figure 9.9). As far as the $g_{WA}(r)$ is concerned, although the RDFs in the middle panel do not show an evident first peak, it is noteworthy to stress that shallow first peaks can be observed, especially increasing the acetone concentration. This is explained by the so-called effective neighbour interactions, which increase with the acetone molar fraction, since acetone can form HB to water. This fact states that the interactions with acetone are improved, thanks to the introduction of the polarizability response in the model. The bottom panel of Figure 9.10 shows $g_{WW}(r)$, in which the first peak increases with x_A. This is in line with previous works, and states that the HB between water molecules prevail over the H-bonds that a molecule of water can build with a molecule of acetone, especially when the system is mostly made of acetone. This is because the water concentration is too low and it is not able to induce a sufficient polarization of acetone, preferring to interact with itself, since the affinity with acetone lowers. Despite the first peak observed, it is much less enhanced with respect to the original SPC/E-OPLS model; moreover, no long-range tail effect can be detected despite the $g_{WW}(r)$ goes to unity from the top, at increasing acetone concentrations, but this effect is immediately damped.

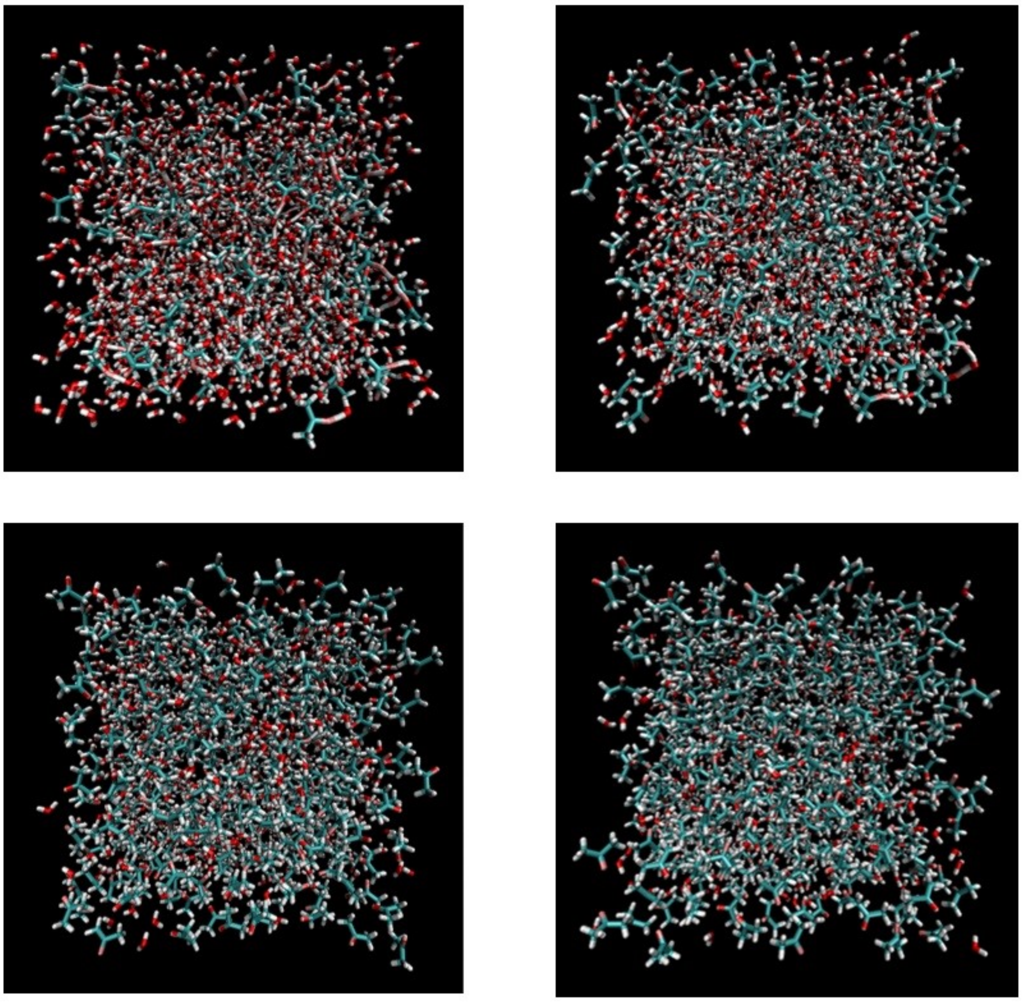


Figure 9.9: Snapshots of the acetone-water mixture referred to the COP model. From left to right and from up to bottom the acetone molar fractions are 0.10, 0.25, 0.50, 0.75. The oxygen is represented in red, the hydrogen in white, carbon in green and the VS is shown in pink.

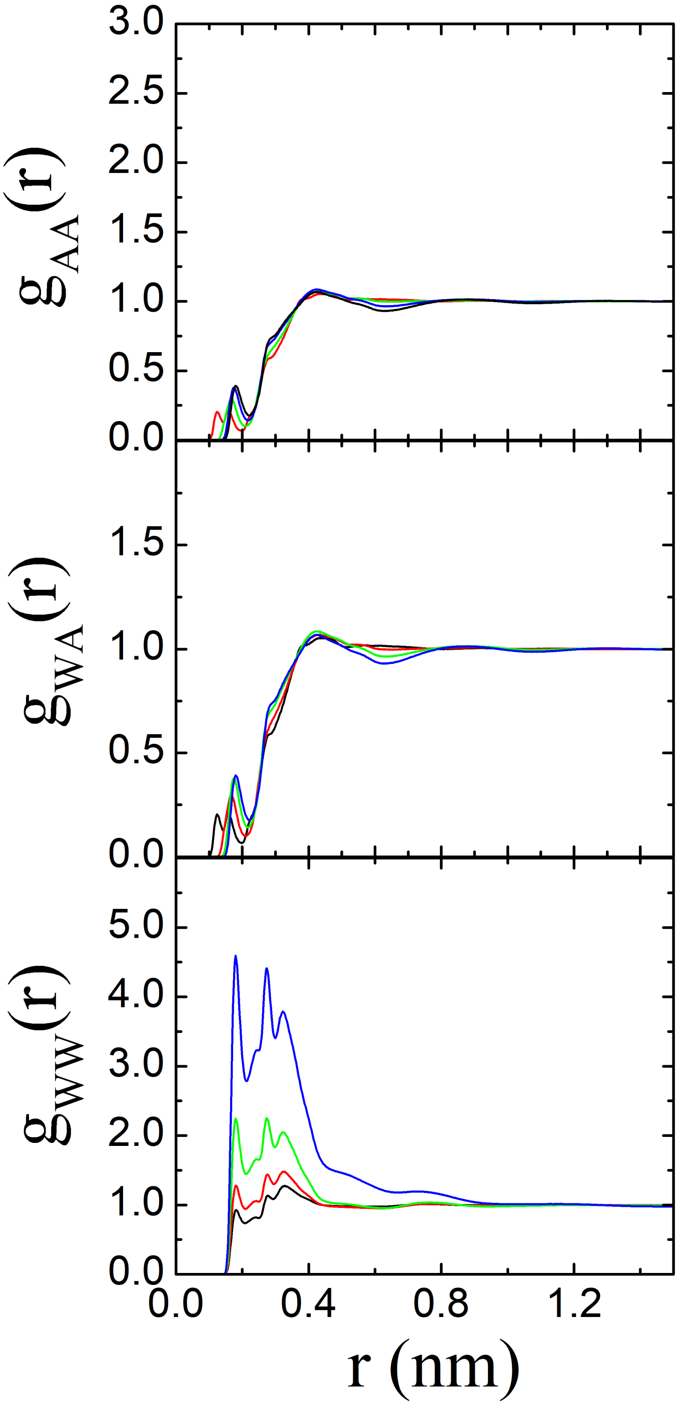


Figure 9.10: Radial distribution functions related to acetone-acetone, $g_{AA}(r)$, water-acetone, $g_{WA}(r)$, and water-water, $g_{WW}(r)$. The acetone molar fractions shown here are equal to 0.10 (black), 0.25 (red), 0.50 (green) and 0.75 (blue).

Density profiles averaged over the entire simulation time (20 ns) along the box are shown in Figure 9.11, proving that the COP model developed here can definitely solve the de-mixing issue, typical of MD simulations of acetone-water mixtures. Looking at Figure 9.11, no density gradients are observed at any acetone molar fraction and at any spatial direction investigated. On the other hand, if the system underwent phase separation, evident density oscillations should be detected along the box, proving that one of the two phases prefers to segregate in a specific region of the simulation box.

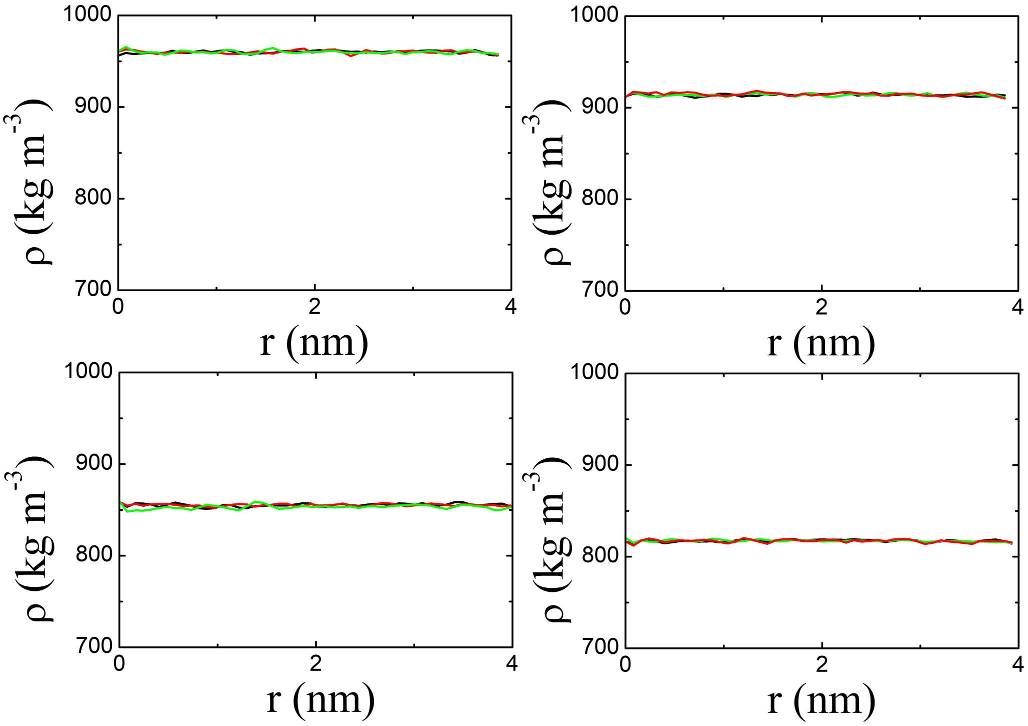


Figure 9.11: Density profiles of the COP model for acetone-water mixture along the box. From left to right and from up to bottom the acetone molar fractions are 0.10, 0.25, 0.50, 0.75. The three space directions are reported, namely density along x (black), y (red), and z axes (green).

9.4 Conclusions

In this chapter, the well-known issue of de-mixing in MD simulations of acetone-water mixtures (Di Pasquale, Marchisio, Barresi, and Carbone, 2014; Jedlovszky et al., 2009; Perera and Sokolić, 2004; Pereyra et al., 2011; Weerasinghe and Smith, 2003) has been investigated, stating the difficulty in reproducing perfect miscibility at all concentrations with classical all-atom force fields. Both quantum mechanical calculations (Georg et al., 2006) and experiments (Applequist et al., 1972; Thole, 1981) show that the effect of induced polarization must be taken into account. Here, lying on the idea of the charge-on-spring approach, a charge-on-particle (COP) model that uses a virtual site (VS) bead for carrying the charge together with the standard OPLS acetone force field is proposed and applied. The VS is attached to its core atom (carbonyl carbon) via a harmonic spring. The COP model is able to account for the effect of the induced polarization due to the interaction with a more polar solvent such as water. This effect is reached by fitting the VS characteristic parameter, a, on the mixture density as target function. The polarization response is reproduced then by varying the equilibrium charge displacement, namely the value of the VS parameter, a, reaching a higher dipole moment value around the carbonyl group of the acetone molecule.

The advantage of the adopted approach consists in merging the simplicity of the well-known force fields (e.g. OPLS) with the main features of the more complex polarizable models such as the Drude oscillator force fields (Lemkul et al., 2016) or the fluctuating charge model (Rick et al., 1994). Regarding the latter ones, we believe that they are more suitable for more complex biological systems, like proteins or DNA, representing a far too complex approach for this simple binary system, since further degrees of freedom would have been added (e.g. introducing a small mass particle to carry the charge, in line with the extended Lagrangian approach (Rick et al., 1994)). Here, we aim to preserve the originality of the adopted force fields together with the simplicity of the analysed binary mixture as well as the ease of implementing the COP model in classical MD simulations, which consequently turns out to be also computational cost saving. The only limitation of this kind of approach is the necessity of adjusting the charge displacement, namely the equilibrium position of the VS, at different mixture molar fractions.

The explanation of this limitation lies on the fact that, being all the chemical bonds constrained, and considering the nature of the VS that is a massless position-constrained particle by definition, it turns out the the charge fluctuations are considerably damped and, therefore, the harmonic potential is negligible in solving the equation of motion during the MD simulations. The detected bond and charge fluctuations are then too small (as shown in the results section) to guarantee a polarizable response by the system, without suitably modifying the charge displacement. However, it surely represents an improvement with respect to previous efforts (Pereyra et al., 2011) that modified the entire charge distribution to reach the polarization effect. The results are very encouraging, showing an evident improvement with respect to the original SPC/E-OPLS force field in all the properties investigated, namely in terms of the density of the mixture (target property), as well as diffusion and viscosity coefficients of the mixture.

Furthermore, the COP model is able to solve the de-mixing issue of acetone-water mixtures, since no more phase separation is detected, leading to well micro-mixed systems. It is noteworthy to stress that another advantage of this COP model lies on its universality and transferability of the general approach adopted here in case of simple binary systems, namely on the possibility from the user to fit suitably the VS parameter a on a given chemical-physical target property which plays a key role in the considered system, depending on its own purpose that one aims to reach. Future work could be done to carry out thermodynamics analysis of the COP model here presented and validated only in terms of dynamic and structural properties, in order to assess the miscibility behaviour also on a thermodynamic point of view and, at the same time, the evaluation of more mixture properties like dielectric constant, heat capacitance and other spectroscopic properties of the acetone-water solution system.

Chapter 10

MARTINI coarse grained model in PCL self-assembly

Most of this chapter has been already published in Alessio D. Lavino, Paola Carbone, and Daniele L. Marchisio (2019), submitted in Soft Matter.

Here, we present the coarse-graining of the poly-ε-caprolactone (PCL) by using the MARTINI force field. In particular, we combine atomistic and coarse-grained (CG) molecular dynamics to model the structural behaviour of long PCL chains in acetone, water and their mixtures. A thermodynamic/structural procedure is adopted to build up the CG model, based on the classical MARTINI force field. The single CG bead is parametrised upon solvation free energy calculations, by means of the Bennett's Acceptance Ratio (BAR) method, whereas the whole polymer chain is optimized in terms of radius of gyration at different chain lengths.

10.1 Introduction

As already explained in Chapt. 4, simulations of realistic polymer chain length by the solely classical full-atom MD is complicated by the associated computational cost. In order to catch interesting and fundamental phenomena that take place at large time and length scales, the use of alternative methods, such as CGMD, turns out to be particularly interesting. One of the most used CG force field is represented by the MARTINI model (Marrink, Risselada, et al., 2007), introduced and extensively described in Chapt. 7.

Despite the MARTINI CG force field is able to catch the partition properties of several compounds in different mixtures, recent works by Rossi et al. (2011), Lee, de Vries, et al. (2009), Lee and Larson (2006) and Taddese and Carbone (2017) have shown how the MARTINI force field can be suitably adjusted and improved for polymer systems, introducing new bead types compatible with the existing ones. These new beads can be parametrised through the tuning of several properties, such as the solvation free energy. Polymer systems still represent a challenging area, because of the several properties concerning the polymer chain lengths, as well as the different conformations they assume depending on the environment. More specifically, polymer self-assembly in solution is still an open topic, due to the numerous applications (Ancona et al., 2018; Valente, Celasco, et al., 2012; Zelenková et al., 2018) it spreads on. Particularly important is the understanding of the polymer chain behaviour in ”good” and ”bad” solvents and a mixture of them.

The aim of this chapter is to develop and validate a CG model for the poly-ε-caprolactone (PCL) in solution, based on the MARTINI CG force field and suitably modified to account for the thermodynamics of a mixed system containing polymers chains and two solvents in different concentration. The parameter space optimized using the standard ”MARTINI approach” is expanded including solvation free energies of the monomer (calculated by means of the Bennett's Acceptance Ratio method Bennett, 1976), and the structure of the polymer chains in terms of radius of gyration. The validation of the model has been done by comparing it against the atomistic simulations, some of which are performed in this work, others in a previous one (Di Pasquale, Marchisio, Barresi, and Carbone, 2014). Once all the parameters related to a single CG polymer chain are set up, it is possible to use this CGMD model to simulate larger systems and systems involving multiple polymer chains in the simulation box.

10.2 Atomistic Simulations

As reported in Chapt. 9, most of the classical full-atom force fields are not able to describe the miscibility of acetone-water mixtures, because of the bad polarization response by the acetone molecule. Two different approaches are possible to consider the polarization of acetone molecules at different acetone molar fractions in aqueous mixtures. Here we followed the approach proposed by Pereyra et al. (2011), in order to carry out the full-atom simulations needed to develop the CG model. The reason of this choice is simply due to the fact that the other way, proposed by Lavino, Banetta, et al. (2018), was not fully tested and complete when we developed this CG model.

Following the Pereyra approach, different charge distributions (depending on mixture molar fraction) on the acetone molecule have been settled, as shown in Table 10.1 and in which C2-O2 corresponds to the acetone carbonyl group. The charge distributions adopted in this context lead to an acetone dipole moment sufficiently high (Table 10.1) to guarantee a good affinity with the water model, therefore no de-mixing occurs during the MD simulations.

The OPLS/AA force field has been implemented in the atomistic simulations, whereas the TIP4P model has been used for water. The last row in Table 10.1 corresponds to the original OPLS/AA force field (pure acetone), properly modified at the other acetone molar fractions, in order to avoid phase separation, in line with the Pereyra et al. (2011) and Lavino, Banetta, et al. (2018) approaches. By means of the acetone charge distributions reported in Table 10.1 no de-mixing takes place, as

$\mathbf{x_A}$	**C2**	**O2**	**C**	**C**	**H**	**H**	**H**	**H**	**H**	**H**	$\boldsymbol{\mu}$, *D*
0.25	0.79	-0.73	-0.36	-0.36	0.11	0.11	0.11	0.11	0.11	0.11	4.60
0.50	0.79	-0.73	-0.36	-0.36	0.11	0.11	0.11	0.11	0.11	0.11	4.60
0.75	0.55	-0.55	-0.27	-0.27	0.09	0.09	0.09	0.09	0.09	0.09	3.69
1.00	0.47	-0.47	-0.18	-0.18	0.06	0.06	0.06	0.06	0.06	0.06	3.07

Table 10.1: Acetone charge (e) distribution used at different acetone molar fractions, xA. C2-O2 corresponds to the carbonyl group; all the other atoms correspond to the two methyl groups. For each charge distribution, the corresponding acetone dipole moment, $\boldsymbol{\mu}$, is reported.

qualitatively shown in Figure 10.1.

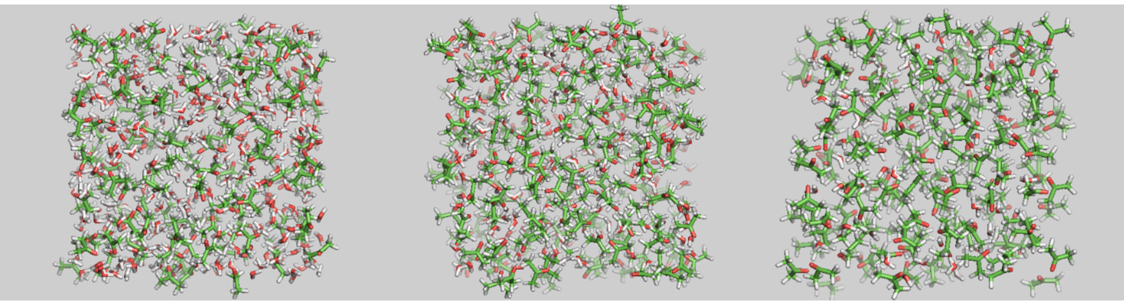


Figure 10.1: From left to right there are reported the three snapshots related to acetone molar fractions equal to 0.25, 0.50 and 0.75 respectively in cubic simulation boxes of 3 nm length. It is clear how no de-mixing occurs.

From a more quantitative point of view, in order to detect possible de-mixing behaviours, it is noteworthy to evaluate also the radial distribution functions (RDF), $g_{\alpha\beta}$. Figure 10.2 (top panel) shows the acetone-acetone RDF, $g_{AA}(r)$, and it is clear that acetone clustering is not detected at all acetone molar fractions, since no first peak is observed. Figure 10.2, in the middle and bottom panels, shows a different trend for $g_{WA}(r)$ and $g_{WW}(r)$ in terms of first peak. In particular, a small peak is shown by $g_{WA}(r)$ highlighting the so-called acetone-water "effective-neighbour" effect. Nevertheless, also $g_{WW}(r)$ shows an overshoot and therefore the water-water clustering seems to be still quite strong but only at $x_A = 0.75$. Results are in line with those obtained by Lavino, Banetta, et al. (2018), in which clustering effect vanishes at all acetone molar fractions, thanks to a pseudo-polarizable model (as also shown in Chapt. 9). Indeed, by comparing Figure 10.2 with Figure 9.10 (RDF obtained with the COP model), it is interesting to note that the profiles are similar, as well as the trends at different acetone molar fractions. The only notable difference is detected for the acetone-water RDF profile (middle panel) that in the latter approach (Figure 10.2) shows a slight first overshoot, leading to stronger acetone-water interactions than in the COP model. Acetone-acetone and water-water RDFs are instead very similar in the two cases. It is then possible to infer that under these operating conditions the system can be considered well micro-mixed enough to perform all the sets of atomistic simulations which the single CG beads will be characterised on. All the atomistic details concerning the latter point will be presented accurately in section 10.5.

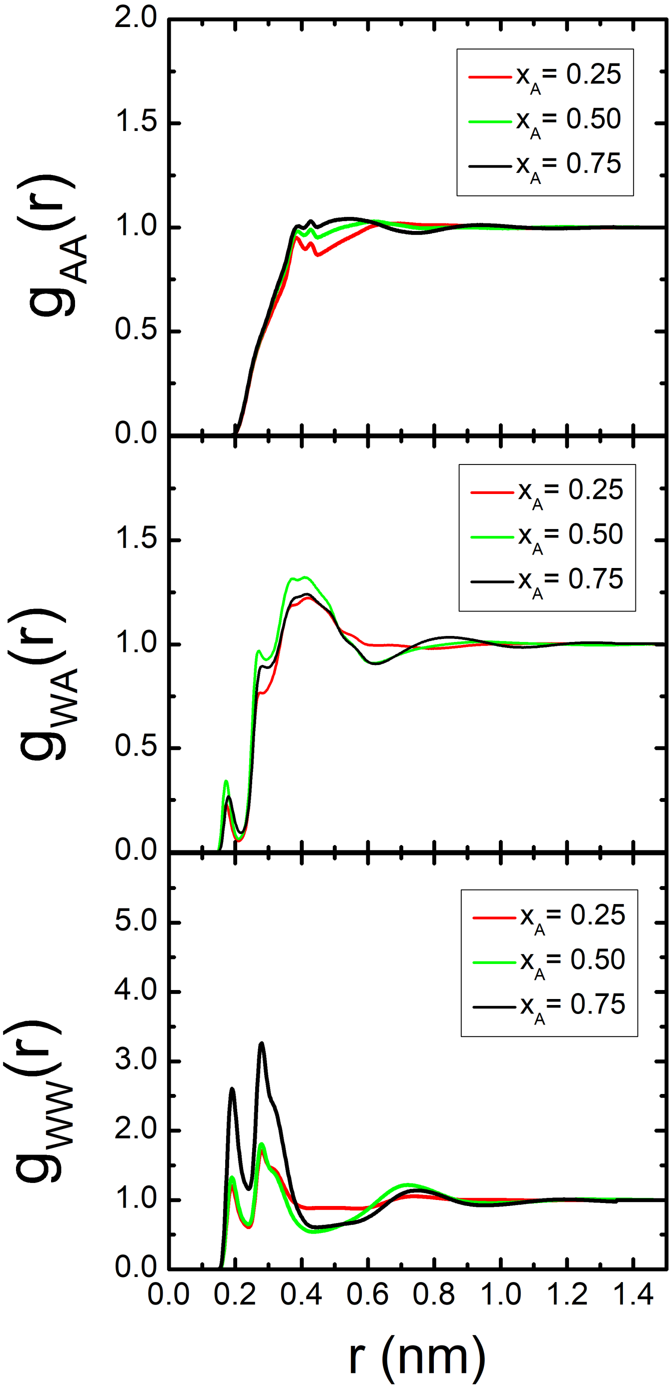


Figure 10.2: Atomistic radial distribution functions for acetone-acetone, $g_{AA}(r)$, water-acetone, $g_{WA}(r)$ and water-water, $g_{WW}(r)$, at different acetone molar fractions, x_A. The red, green and black lines correspond respectively to x_A= 0.25, 0.50 and 0.75.

10.3 Coarse-Grained MARTINI Model

In line with the MARTINI force field (Marrink, Risselada, et al., 2007), a four-to-one mapping has been considered, which means four heavy atoms per CG bead. The two beads of the PCL repeat unit have been chosen in order to account for both the alkyl and the ester part of the polymer. Therefore, at the beginning a C_1 bead type (corresponding to butane) was used for the alkyl part, whereas a bead type labelled N_{am} (corresponding to methyl formate) was used for the ester part, based on the existing one N_a and suitably modified in order to differentiate it from the latter, which corresponds to acetone. The mapping of the system is schematically depicted in Figure 10.3. Water is modelled by a P_4 bead type (corresponding to four water molecules), in line with the MARTINI force field. It is crucial to stress here that this mapping (N_{am} - C_1) represents just a starting choice, since the outcome of the CG optimization procedure we adopt here may lead to different bead types (for more suitably describing the behaviour of PCL in acetone/water mixtures) as it will be shown in the results section.

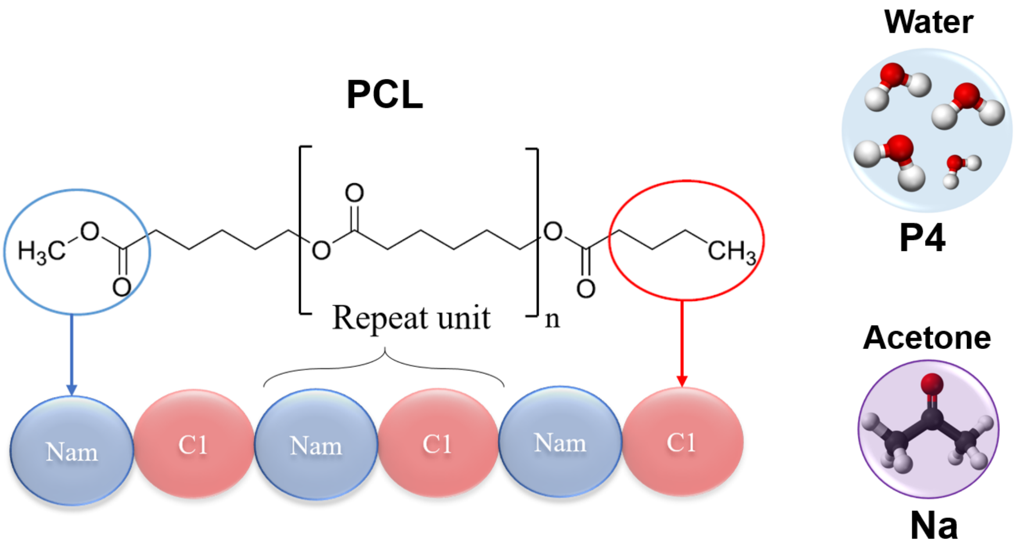


Figure 10.3: Initial mapping of both the PCL chain and the solvents used in this work. The repeat unit is described by N_{am} (a new bead introduced here) and C_1 bead types respectively for the ester and the alkyl part. Water and acetone are mapped respectively by P_4 and N_a bead types.

The approach proposed here consists in parametrising the conservative potentials of the single CG beads by matching the atomistic results in terms of solvation free energy, evaluated by using the BAR method (extensively explained in the next section). More specifically, the Lennard-Jones (LJ) parameter ε has been varied, keeping σ

constant. The default MARTINI non-bonded LJ interaction levels have been used to build up the model, in line with the MARTINI force field. Every interaction between bead types is biunique as reported in the interaction matrix of the MARTINI work (Marrink, Risselada, et al., 2007). A pairwise shifted LJ 12-6 potential energy function:

$$U_{LJ}(r) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \tag{10.1}$$

coupled with a Coulombic potential energy function:

$$U_{el}(r) = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r}, \tag{10.2}$$

are used to describe non-bonded interactions, where σ_{ij} represents the closest distance between two particles and ε_{ij} the strength of their interaction; in the Coulombic potential, $\varepsilon_r = 15$ is the dielectric constant (explicit screening). Concerning the bonded interactions, weak harmonic potentials have been used for both bonds and angles:

$$V_b(r) = \frac{1}{2}K_b(r - r_b)^2, \tag{10.3}$$

$$V_{angle}(r) = \frac{1}{2}K_{angle}(\theta - \theta_0)^2, \tag{10.4}$$

where K_b and r_b are respectively the bond stretching force constant and the equilibrium bond length, while K_{angle} and θ_0 are respectively the angle force constant and the equilibrium angle. Their values will be reported in the simulation details section.

10.4 Bennett's Acceptance Ratio Method

The free energy of solvation, ΔG^{solv}, is evaluated by means of the Bennett's Acceptance Ratio (BAR) method (Bennett, 1976), since it has been shown that it represents a more efficient method compared to other thermodynamic integration methods (Bruckner and Boresch, 2011; Shirts and Pande, 2005). The BAR method allows to calculate the energy difference in a thermodynamic transformation from a state A to a state B, evaluating the energy differences between several intermediate states. The idea is to sample several ensembles in between the two reference ones that one wants to estimate. These sampled ensembles must be close enough to each other in terms of configuration space (potential functions), and the BAR method uses a probability estimator to efficiently compute the energy difference between two consecutive ensembles.

ΔG^{solv} is defined as the reversible work required to transfer a molecule from the ideal gas state to ideal solution, at 298.15 K, in the pure solvent. However, from a computational point of view, it is more convenient to define the ΔG^{solv} as the balance between the work required for the production of a cavity in the bulk solvent, plus the gain reached to solvate a given solute which gradually compares in that cavity and interacts with the surrounding solvent molecules, in terms of Van der Waals, Coulomb and hydrogen interactions. Since ΔG^{solv} is a state function, it is independent on the specific thermodynamic pathway, therefore the two interpretations are thermodynamically equivalent (Figure 10.4).

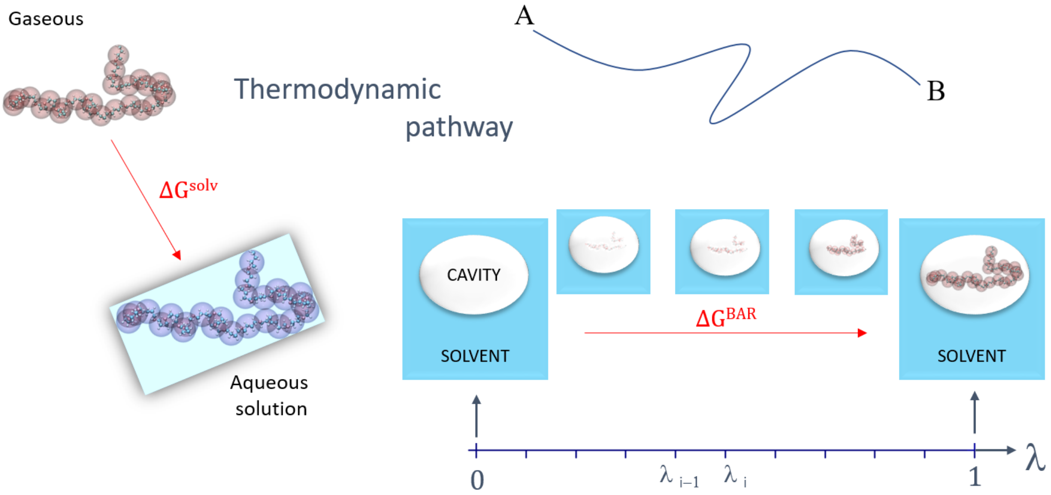


Figure 10.4: Thermodynamic pathways that lead to the determination of the ΔG^{solv} from a state A to a state B. On the left side, there is a schematic representation of the definition of ΔG^{solv}; on the right side, a sketch of the BAR method is reported, going from λ=0 (cavity production, non-interacting system) to λ=1 (fully interacting system).

Although the work related to the cavity production is always an unfavourable process, the gain in terms of enthalpic and entropic solvation can be more or less favourable, depending on the solute hydrophobicity (if the solvent is pure water) or, in general, on the affinity with the solvent. ΔG^{solv} can be interpreted as the balance between these two factors (Skyner et al., 2015).

The BAR method is based on a switching/interacting parameter, λ ,which states the coupling between the solute and the solvent, and spans the range [0,1]. Stating $\lambda = 0$ means non-interacting state, a cavity is created in the bulk solvent (vacuum); on the other hand, $\lambda = 1$ stands for a fully interacting state between the solute and the solvent. All the other values in between [0,1] state a gradual appearance of the solute in the cavity. The chosen number of λ points represents a crucial aspect of this method. In fact, on one side, the computational cost decreases together with $[\lambda_{i-1}, \lambda_i]$ intervals (a single simulation is required for each interval); on the other side, it must be high enough to guarantee a suitable phase space overlap (Wu and Kofke, 2005) along the thermodynamic pathway. A single simulation is run for each $[\lambda_{i-1}, \lambda_i]$ interval and, then, the ΔG^{solv} from a generic state A to a state B is evaluated by means of the variation of the Hamiltonian of the system with respect to λ; therefore, the total ΔG^{solv} can be expressed by

$$\Delta G^{solv} = \sum_{i=1}^{N} \Delta G_i^{BAR}[\lambda_{i-1}, \lambda_i], \tag{10.5}$$

where $\Delta G_i^{BAR}[\lambda_{i-1}, \lambda_i]$ is the solvation free energy related to the single interval $[\lambda_{i-1}, \lambda_i]$, evaluated by means of the BAR method on the N intervals.

Computational details about ΔG^{solv} simulations will be reported in the next section; potential functions deserve particular attention. More specifically, soft-core potentials have been used for the BAR simulations, in order to avoid singularities when nonbonded interactions are turned off. Soft-core potentials V_{sc} are defined as a shifted version of the regular potentials:

$$V_{sc} = (1 - \lambda)V_A(r_A) + \lambda V_B(r_B) \tag{10.6}$$

$$r_A = (\alpha\sigma_A^6 \lambda^p + r^6)^{1/6} \tag{10.7}$$

$$r_B = (\alpha\sigma_B^6 (1 - \lambda)^p + r^6)^{1/6}, \tag{10.8}$$

where $V_A(r_A)$ and $V_B(r_B)$ are the normal ”hard-core” potentials referred to the states A and B, α is the soft-core parameter (it controls the value of the potentials when r is approaching to zero), p is the soft-core λ power and σ is the radius of the interaction.

10.5 Operating Conditions

All simulations were performed by using GROMACS 4.5.6 molecular dynamics package (Hess, Kutzner, et al., 2008). The VMD program was used to produce graphical images of the molecular systems.

Atomistic Simulations

The atomistic simulations were carried out in a cubic box with box length equal to 3 nm, in an NPT ensemble. The OPLS-AA force-fields was employed for acetone (suitably modified according to the charge distributions shown in sect. 10.2) and TIP4P force field for water. The weak harmonic potential functions were used for bonds, angles and improper dihedrals, whereas the Ryckaert-Bellemans potential was implemented for the proper dihedrals. It was simulated a single butane and a single methyl formate (corresponding to the atomistic part of the single CG beads) molecule in mixture at acetone molar fractions equal to 0, 0.25, 0.50, 0.75, 1. The energy minimization of the system was carried out with a steepest descent algorithm. The system was then equilibrated by using a leap-frog algorithm for 1 ns with a 2 fs timestep. Berendsen thermostat and Berendsen barostat (coupled with Parrinello-Rahman) fixed the temperature and the pressure respectively at 300 K and 1 bar during the equilibration, with a coupling time constant respectively equal to 0.2 ps and 5 ps. The simulations were performed for 10 ns with a 1 fs timestep, and bonds were constrained using the LINCS algorithm. 3D periodic boundary conditions were employed and electrostatic interactions were evaluated using the Particle-Mesh Ewald (PME) summation. A temperature of 300 K temperature was maintained by means of a velocity-rescale algorithm with a time constant equal to 0.1 ps. An isotropic pressure was set to 1 bar by using a Parrinello-Rahman scheme with a coupling constant equal 1.0 ps and a compressibility set to 4.5 10-5 bar-1. To keep away from singularities and numerical instabilities, a soft-core LJ and Coulomb potentials were used50 as reported in sect. 10.4. A stochastic dynamics integrator was implemented in order to avoid singularities when λ approaches to zero, and the number of lambda points has been chosen equal to 10 for the butane and 20 for the methyl formate, since it represents a good trade-off between the computational cost and the phase space overlap needed to guarantee a correct thermodynamic pathway. Concerning the soft-core potential parameters (Eq. 10.6) used in the atomistic BAR simulations, p was set equal to 1, α equal to 1 and σ equal to 0.3 nm.

Coarse-Grained Simulations

The CG simulations were performed in a cubic box, whose box length varied depending on the simulated system. It spans from 4 nm for the simulation of the single CG bead in mixture to 15 nm for the longest PCL chains. Different PCL chains lengths have been simulated. In particular, we focused on the PCL-10, PCL-20 and PCL-30, namely a PCL chain with a number of repeat units respectively equal to 10, 20 and 30. Bond stretching and angle bending motions were treated with harmonic potentials, reported in Eq. (10.3) and (10.4) where r_b=0.415 nm, θ_0=170 ° (N_{am}-C_1-N_{am}) θ_0=129 ° (C_1-N_{am}-C_1) and the force constants K_b=5000 kJ mol^{-1} nm^{-2} and K_{angle}=50 kJ mol^{-1}, since it has been demonstrated that this set of values can reproduce the atomistic distributions (Raman et al., 2017). The energy minimization was carried out by means of the steepest descent algorithm. The system was equilibrated with a leap-frog algorithm for 1-2 ns by using a Berendsen thermostat and a combination of Berendsen and Parrinello-Rahman barostat (T = 300 K and p = 1 bar). The temperature time constant for the Berendsen scheme is equal to 1 ps, Parrinello-Rahman coupling time constant is equal to 4-12 ps (higher constants values ensure more stability).

The simulations were performed for 50 ns with a 20 fs timestep. A velocity-rescale thermostat scheme was used with a temperature time constant of 1.0 ps, whereas a Parrinello-Rahman barostat was employed with a time constant equal to 4 ps and a compressibility set to 4.5 10-5 bar^{-1}. A soft-core LJ and Coulomb potentials were used for BAR calculations, and the number of lambda points has been chosen equal to 10 for both of the CG solutes. Concerning the soft-core potential parameters (Eq. 10.6) used in the CG BAR simulations, the soft-core parameters have been chosen to be p equal to 1, α equal to 0.5 and σ equal to 0.47 nm.

10.6 Results and Discussions

In this section, the main results concerning the developed CG model and the adopted methodology will be discussed. First of all, it is noteworthy to underline that, unlike the atomistic system, CG acetone-water mixtures do not lead to evident phase separation or de-mixing. Little clusters seem to appear, as shown in Figure 10.5 in which a simulation box of N_a-P_4 (CG acetone-water) at $x_A = 0.25$ and box length equal to 15 nm, with 27000 particles and after 10 ns is reported. Since the system does not undergo an evident phase separation, by looking at the snapshots reported in Figure 10.5, it is reasonable to infer that no de-mixing occurs at the CG level. This is also confirmed from a more quantitative analysis by looking at the density profiles along the three spatial directions x, y, z (Figure 10.5c, d, e) that are shown to be flat. This is far too different from the all-atom MD simulations of acetone-water mixtures, in which evident phase separation takes place with most of the classical MD force fields (Lavino, Banetta, et al., 2018; Perera and Sokolić, 2004). It was already explained in Chapt. 9 that the cause of such de-mixing at molecular level lies on a bad polarization response of acetone molecule in water. The reason why this does not happen at CG level is probably due to the CG beads framework, in which polarization effects are less important since CG water and acetone bead types do not carry any charge on themselves, showing therefore a better affinity (Figure 10.5).

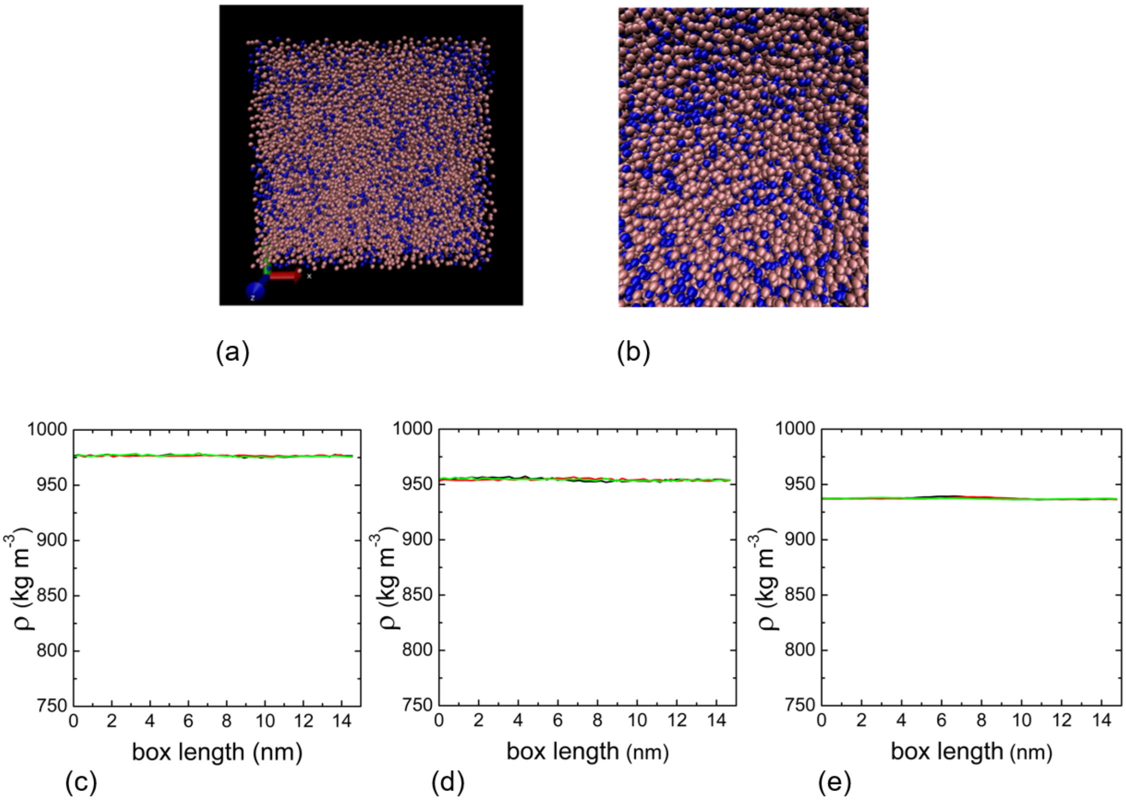


Figure 10.5: (a) Snapshot of the N_a-P_4 CG acetone(blue)-water(pink) mixture for x_A = 0.25 and box length equal to 15 nm, with 27000 particles and after 10 ns. (b) Detail of the N_a-P_4 box in which a little cluster (blue, N_a) it seems to form; however, no evident phase separation is detected at the CG level. Densities profiles along x (black), y (red) and z (green) are reported at acetone molar fraction equal to 0.25 (c), 0.50 (d) and 0.75 (e).

10.6.1 Estimation of the solvation free energy of a single bead

Firstly, full-atom and CG simulations of a single bead in solution were carried out, in order to match the ΔG^{solv} between the CG beads and the atomistic part of the polymer chain that it represents (mapping). Therefore, atomistic models for butane, methyl formate, and the beads C_1 and N_{am} are simulated in acetone-water mixtures and the relative ΔG^{solv} are evaluated. The main results are shown in Figure 10.6.

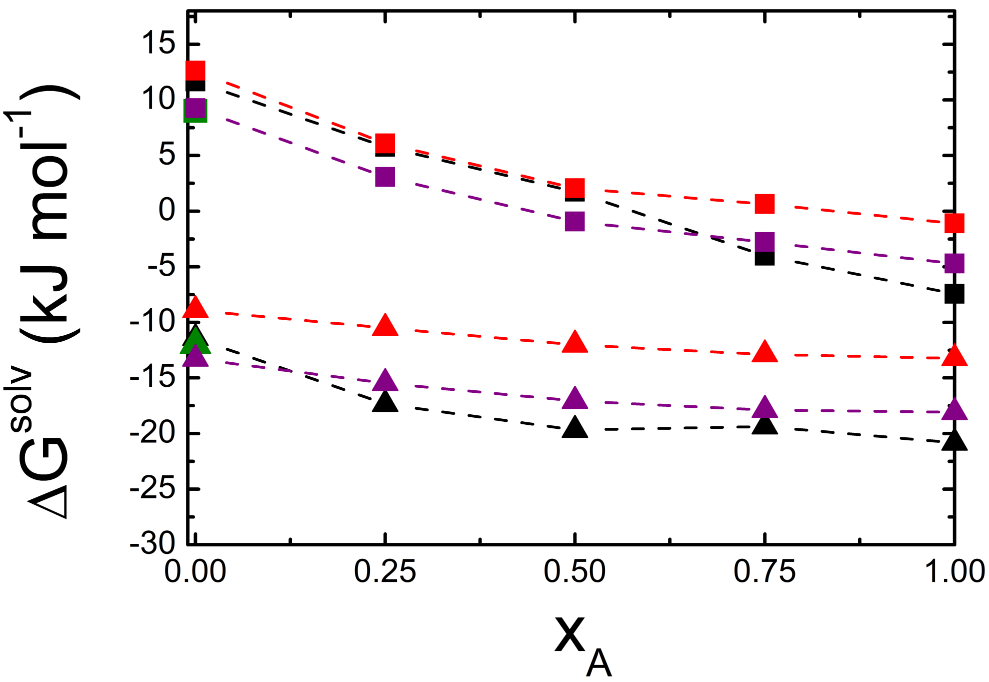


Figure 10.6: Free energy of solvation of butane (squares) and methyl formate (triangles) in mixture from the atomistic simulations (black symbols) carried out in this work. The red results are referred to the MARTINI model, C_1 (squares) and N_{am} (triangles), without any modifications of the LJ parameter, ε. The purple curves refer to the CG model developed here, in terms of single beads (squares for the alkyl part and triangles for the ester part of the PCL repeat unit) varying suitably the LJ parameter, ε. The green symbols stand for the experimental ΔG^{solv} values respectively of butane and methyl formate in pure water.

Figure 10.6 clearly points out that the MARTINI beads C_1 and N_{am} (red symbols) show a mismatch with the atomistic results, in terms of ΔG^{solv} for the whole composition of the mixture. The mismatch is quite evident at high acetone molar fractions, as far as the butane is concerned, and for the whole acetone molar fraction range, for the methyl formate. By modifying properly the LJ parameter ε, namely, by moving upwards and downwards along the MARTINI LJ interaction levels, a better agreement in terms of ΔG^{solv} between atomistic and CG model is here achieved (purple profiles in Figure 10.6). The only experimental values available in literature are those referred to pure water (Ben-Naim and Marcus, 1984). It is evident that the agreement between atomistic, experimental and CG values is improved also in pure water (Figure 10.6, $x_A = 0$).

The new set of LJ parameters (between CG solvents and CG polymer bead types) that best fit the ΔG^{solv} (purple profiles in Figure 10.6) is reported in Table 10.2. Looking at Table 10.2, it is worthwhile to stress that the hydrophilicity of the ester part of the PCL is enhanced in acetone-water mixture, behaving at CG level of resolution as a P_2 bead type (polar), instead of a N bead type (non-polar) as considered in the initial mapping. This implies that, in order to better reproduce the atomistic and experimental ΔG^{solv} in acetone-water mixtures, a more polar bead type is necessary to suitably describe the affinity of such a group in a mixture of solvents with which polarizable effects may take place. This can be justified by looking at the atomistic scale, in which water can H-bond with the carboxyl group of the PCL. This may lead to a negative charge dislocation on it, which, in turn, dislocates partial positive charge on the rest of the molecule increasing therefore also the affinity with the acetone carbonyl group. In a CG framework, this fact can be taken into account by increasing the interaction level between polymer bead and solvents from III (weaker, ε= 4.0 kJ/mol) to II (stronger, ε = 4.5 kJ/mol) in the LJ levels of the MARTINI force field. Therefore, in the mapping procedure, the bead N_{am} is substituted by a P_2 bead type.

Concerning instead the alkyl part of the PCL repeat unit, it is visible, by looking at Figure 10.6, that the MARTINI (red symbols) model does not match the experimental (green symbol) ΔG^{solv} in pure water, overestimating this thermodynamic quantity. In order to better reproduce the experimental value, a stronger affinity (higher values of ε) between the CG alkyl group and CG water (P_4), is necessary. Therefore, a C_1 bead type is clearly not suitable to reach this purpose. In order to get it and, at the same time, to have a better qualitative agreement with the atomistic results, this alkyl bead type has to relate to water (P_4) and acetone (N_a) as a C bead type, but the subscript has to be greater than 1, belonging to the range [2,4], namely being slightly more polar. However, regarding the alkyl bead type there is not a unique bead belonging to the interaction matrix of the MARTINI model that can catch up on those variations just introduced above (with respect to both water and acetone), since it behaves as a C_4 with acetone and as a C_2 with water. This makes us conclude that only ten different LJ interaction levels are not sufficient to catch all the variability that may takes place in binary mixtures, at CG level of description, and more bead types must be introduced in this CG approach. Therefore, in order to reproduce the best profile in terms of ΔG^{solv}, the CG alkyl group must be mapped by a generic C_α bead type, whose LJ interactions with the solvent beads are reported in Table 10.2 and still belong to the range of the non-polar C bead types family of the MARTINI force field.

10.6.2 Structural investigation of the polymer chain

Once the ”polymer single bead – solvent” non-bonded interactions are optimized upon mixture thermodynamics, an investigation on the polymer chain structural properties can be carried out. The aim here is to evaluate how different non-bonded intrachain interactions can affect the final structure of the PCL chain, looking for the optimal set of polymer intrachain LJ parameter ε that brings to the best agreement

	Polymer bead, P_2		Polymer bead, C_α	
	LJ interaction	ε, kJ mol^{-1}	LJ interaction	ε, kJ mol^{-1}
Water bead, P_4	**II**	**4.5**	**VII**	**2.3**
Acetone bead, N_a	**II**	**4.5**	**V**	**3.1**

Table 10.2: Lennard-Jones interactions and parameter ε, between the repeat unit bead types P_2, C_α and the solvent bead types P_4 and N_a.

with the atomistic radii of gyration.

Three different PCL chain lengths are used here to develop the CG model: PCL-10 (i.e. 10 monomers), PCL-20 (20 monomers) and PCL-30 (30 monomers), being these the only ones investigated in the literature at the atomistic level, at all the acetone mixture fractions, in a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014). However, before showing the final set of LJ intrachain parameters that best reproduces the atomistic radius of gyration for these three chain lengths (PCL-10, PCL-20 and PCL-30, Table 10.4), it is worthwhile to do a preliminary analysis by just considering PCL-10 and PCL-20, since interesting properties and limitations of these kind of approach can be pointed out.

Let us start therefore in a preliminary way with the comparison between PCL-10 and PCL-20. After optimising the LJ intrachain interactions for PCL-10 in terms of matching atomistic and CG radii of gyration at all acetone molar fractions investigated in this work, it turns out that this set of LJ parameters underestimates the radius of gyration of the CG system for PCL-20. The mean radius of gyration even decreases with respect to the PCL-10 at acetone molar fractions $x_A > 0.25$. This behaviour is quantitative and qualitative reported respectively in Table 10.3 and Figure 10.7 and can be summarised as follows: the best configuration in terms of non-bonded interactions for a given chain length N turns out to be not the optimal configuration in the same operating conditions, for a chain with a double length $2N$. This peculiar aspect makes the used CG procedure less trivial than one can expect and, above all, allows us to infer that the final set up in terms of non-bonded parameters will be a compromise between the three different chain lengths investigated in this work (PCL-10, PCL-20 and PCL-30). This is clearly demonstrated in Table 10.3 in which the two PCL chain lengths used for this preliminary analysis (PCL-10 and PCL-20) are simulated with the best configuration found out initially for solely PCL-10. This configuration corresponds to LJ interaction levels equal to II ($\varepsilon = 4.5$ kJ mol^{-1}) for all the three PCL intrachain interactions. As Table 10.3 shows, the agreement is always good in pure water, but if the atomistic trend is well reproduced for PCL-10, concerning the PCL-20 the mismatch is evident at all the other acetone molar fractions. At constant acetone molar fraction, say equal to one, the radius of gyration even decreases substantially, going from 1.38 nm (PCL-10) to 0.99 nm (PCL-20). These LJ intrachain interactions are therefore too strong and, then, increasing the molecular weight, namely the chain length, the polymer collapses on itself rather than exposing to the solvent, since the intrachain bead-bead attractions are preferred on the bead-solvent ones (a globule form is then preferred).

At the end of this preliminary analysis we can conclude that the best configuration (in terms of LJ polymer intrachain interactions) at a given acetone molar fraction is not said to be the best one for the full range of acetone molar fractions. Furthermore, the best configuration for a given PCL chain length does not guarantee the best match at higher PCL chain lengths, since the polymer beads will prefer to interact more favourably with themselves (higher LJ interaction parameter, ε), rather than exposing to the solvent (poor solvent condition). This is quite evident for the case of the PCL-10 and -20, in which the length doubles (it smooths going to higher molecular weights) as shown in the snapshots reported in Figure 10.7, which represents PCL-10 and PCL-20 in pure acetone, after 50 ns of simulation with all the intrachain LJ parameters set to the level II ($\varepsilon = 4.5$ kJ mol^{-1}). It is evident that PCL-20 collapses on itself showing a globule conformation, instead of the PCL-10 which shows a more stretched (coil) conformation. Being this set of intrachain LJ parameters considered so far suitable to represent the PCL-10, but not good at all to represent the PCL-20, it is not worth to continue with higher molecular weights.

In line with the strategy adopted here, a trade-off set of LJ intrachain parameters has to be established in order to correctly reproduce the atomistic results for PCL-10, PCL-20 and PCL-30. After this has been done, by using this set of LJ parameters, higher molecular weights are investigated (PCL-40 and PCL-60) and reported at the end of this chapter.

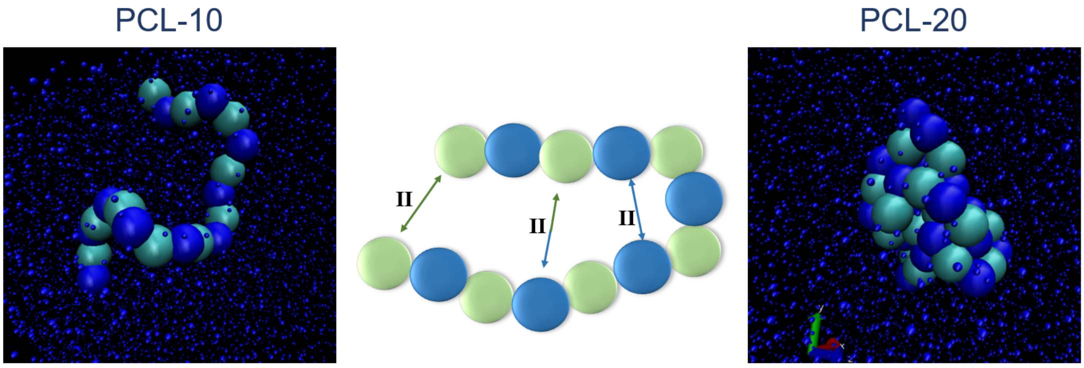


Figure 10.7: Snapshots referred to the preliminary analysis conducted in this work, in which CG PCL-10 (left) and PCL-20 (right) are compared and reported here in pure acetone with the LJ intrachain interaction levels equal to II. This scheme refers to the numerical results reported in Table 10.3.

At the end of the optimization procedure, the best fitting (in terms of accordance with atomistic radii of gyration among PCL-10, PCL-20 and PCL-30) is finally

		Acetone molar fraction, x_A				
Preliminary analysis	$\langle R_g \rangle$	0	0.25	0.50	0.75	1.00
PCL-10	CG, *nm*	0.64	0.73	0.94	1.37	1.38
	AA, *nm*	0.64	-	1.19	1.23	1.25
PCL-20	CG, *nm*	0.76	0.83	0.85	0.86	0.99
	AA, *nm*	0.77	-	1.66	1.73	1.85

Table 10.3: Mean radius of gyration, $\langle R_g \rangle$, of coarse-grained (CG) and atomistic (AA) PCL-10 and PCL-20, at different acetone molar fractions. The LJ interaction levels are II for all the possible beads interactions, namely the set of LJ parameters that best fit the atomistic radii of gyration for the solely PCL-10 (preliminary analysis).

represented by the LJ intrachain interaction levels of type II ($\varepsilon = 4.5$ kJ mol^{-1}) for P_2- P_2, III ($\varepsilon = 4.0$ kJ mol^{-1}) for C_α- C_α and IV ($\varepsilon = 3.5$ kJ mol^{-1}) for P_2- C_α. The results are schematically shown in the Table 10.4 and depicted in Figure 10.8. As it can be seen, the agreement is very good in pure water and at high acetone molar fractions. However, the mismatch is quite viewable at medium acetone molar fraction in mixture (i.e., $x_A = 0.50$). Let us stress here that despite the intrachain P_2- P_2 and P_2-C_α interactions are in accordance with those tabulated in the MARTINI force field, the optimal C_α-C_α LJ interaction level turns out to be one level stronger (level III, ε = 4.0 kJ/mol) than the ones (level IV, ε = 3.5 kJ/mol) reported in the interaction matrix of the MARTINI force field. This stresses again what was already pointed out for the single-bead characterization in this work: in order to capture more chemical/structural details of the atomistic systems, the corresponding CG framework with the MARTINI force field needs more LJ interaction levels, namely, a finer discretization of them.

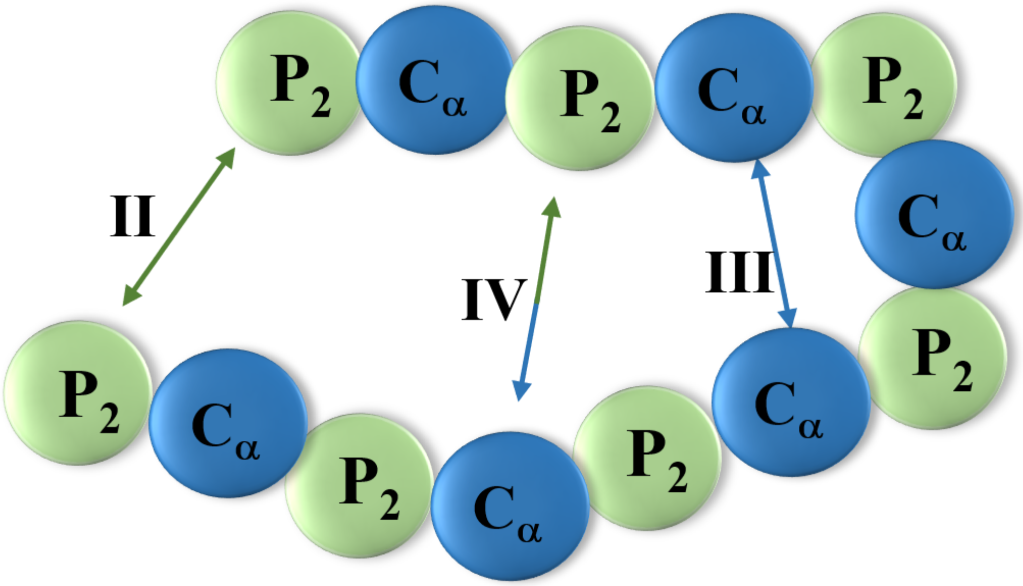


Figure 10.8: Snapshots referred to the final configuration of the CG PCL, after the optimization procedure conducted on the atomistic radius of gyration. The LJ intrachain interaction levels of type II (ε = 4.5 kJ mol^{-1}) for P_2- P_2, III (ε = 4.0 kJ mol^{-1}) for C_α- C_α and IV (ε = 3.5 kJ mol^{-1}) for P_2- C_α. This scheme refers to the numerical results reported in Table 10.4.

		Acetone molar fraction, x_A				
Final setup	$\langle R_g \rangle$	0	0.25	0.50	0.75	1.00
PCL-10	CG, *nm*	0.68	1.00	1.39	1.47	1.48
	AA, *nm*	0.64	-	1.19	1.23	1.25
PCL-20	CG, *nm*	0.81	1.02	1.51	1.75	2.10
	AA, *nm*	0.77	-	1.66	1.73	1.85
PCL-30	CG, *nm*	0.89	1.05	1.58	1.75	2.70
	AA, *nm*	0.88	-	2.00	2.17	2.42

Table 10.4: Mean radius of gyration, $\langle R_g \rangle$, of coarse-grained (CG) and atomistic (AA) PCL-10, PCL-20 and PCL-30, at different acetone molar fractions, x_A. The LJ interaction levels are of type II (ε = 4.5 kJ mol^{-1}) for P_2- P_2, III (ε = 4.0 kJ mol^{-1}) for C_α- C_α and IV (ε = 3.5 kJ mol^{-1}) for P_2- C_α.

The behaviour observed at x_A = 0.50 and, in general, the whole trend reported in Table 10.4, deserves further considerations. The most important one is about what happens at $x_A = 0.50$. This seems to be a threshold value at which a stepwise globule-to-coil transition takes place, instead of being smoother at medium acetone molar fractions, as it happens in the atomistic framework. Recent works have investigated this interesting behaviour, finding out that this is one of the limits of the CG models for long single chains in solution. More specifically, J. Cheng et al. (2014), demonstrated that a stepwise transition from globule-to-coil configuration is detected at medium acetone molar fraction in binary mixture ($x_A \approx 0.55$) for a single polymer tethered chain in solution, in a DPD-based analysis. They also detected another interesting property of the CG polymer systems, namely that this stepwise transition observed for a single CG chain becomes less sharp for polymer brushes. This means that the stepwise transition observed in CG single macromolecule systems fades out when more CG macromolecules are present in solution. Furthermore, the CG approach on one single macromolecule may be limited by the very low degrees of freedom of the system, especially when the mixture is at medium acetone molar fraction (around 0.50) and, above all, when the solvents mixture itself is made by a good and bad solvent. In this scenario, in fact, the number of ”good-solvent-beads” are equal to the number of ”bad-solvent-beads”. Consequently, the probability of interacting with a ”good-solvent-bead” (stronger, favourable interactions) is the same as the one of interacting with a ”bad-solvent-bead” (weaker, unfavourable interactions). Being the level of discretization much coarser than an atomistic scenario (much less number of degrees of freedom and only ten levels of LJ interactions in the MARTINI model), one single isolated CG chain in mixture is not able to guarantee a smooth transition in its conformation when it is experiencing a statistically equivalent number of interactions with both ”good” (higher ε) and ”bad” (lower ε) solvents beads. Indeed, it is expectable that going through intermediate good solvent molar fractions, from bad to good solvent conditions, a smooth polymer structural transition may be detected, by gradually changing its configuration from globule-to-coil conformation as long as the good solvent molar fraction increases. However, the CG results do not show this smooth conformational change, leading to a sharp globule-to-coil transition, at least sharper than the atomistic results.

Raman et al. (2017) observed the same behaviour at $x_A = 0.50$ for single CG PCL chains in acetone-water mixtures, by using the MARTINI force filed with a polarizable water model. This allows us to conclude that the anomalous behaviour detected at $x_A = 0.50$ (medium good solvent molar fraction) represents an intrinsic limit of CG approaches for a single polymer chain in binary ”good-bad” solvent mixtures.

However, it is still under debate the validity of the atomistic simulations reported in Table 10.4, since they are referred to a work in which acetone/water mixtures were affected by de-mixing issues. Although a complete phase separation between acetone and water is detected at acetone molar fraction, x_A, equal to 0.25 with the most common classical acetone force fields, the authors of the reference work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014) observed the formation of an acetone pocket around the PCL molecule at $x_A = 0.50$. This solvent behavior may be enhanced by a slight de-mixing that keeps taking place also at the other intermediate molar fractions (not only $x_A = 0.25$), forming a little acetone pocket around the PCL molecule and, therefore, leading to a more stretched final structure (higher mean radius of gyration). This could pave the way to further atomistic investigations, maybe employing the charge-on-particle model presented in the previous chapter, in order to better clarify the uncertainty related to the atomistic calculations.

By using the set of LJ parameters that come from the previous optimisation procedure on PCL-10, PCL-20 and PCL-30, the mean radius of gyration of longer CG single polymer chains in solution was calculated, despite no atomistic data are available in literature. The case of PCL-40 and PCL-60 was evaluated and is reported in Table 10.5. It is noteworthy to stress that also for longer polymer chains (PCL-40 and PCL-60) the trend is the same as the one observed at lower molecular weights (Table 10.4). A stepwise globule-to-coil conformation is detected also here in the range of acetone molar fractions, $x_A = [0.50, 0.75]$. However, in pure solvents and far away from medium acetone molar fractions the CG chain evolves more qualitatively in accordance with the trend predicted by atomistic simulations at lower chain lengths (Di Pasquale, Marchisio, Barresi, and Carbone, 2014; Raman et al., 2017). By looking at Table 10.4 and Table 10.5, it is worth noticing that the trend obtained by this new CG model of PCL in acetone-water mixture follows the Flory's theory of polymers, namely

$$\langle R_g \rangle \propto M_w^{\nu} \tag{10.9}$$

where M_w is the polymer chain molecular weight and ν corresponds to the Flory's exponent that is equal to 1/3 in pure anti-solvent (water) and 3/5 in pure good solvent (acetone), in line with the Flory's theory. The trend in pure water and pure acetone for all the M_w investigated here (PCL-10 / -60) is depicted in Figure 10.9, where the simulation results are reported with discrete symbols and the theory of Flory's predictions (Eq. 10.9) with dashed lines (an arbitrary constant is employed, being the Flory trend a proportionality relation).

		Acetone molar fraction, x_A				
Further investigation	$\langle R_g \rangle$	0	0.25	0.50	0.75	1.00
PCL-40	CG, *nm*	0.99	1.18	1.60	1.84	3.10
PCL-60	CG, *nm*	1.13	1.28	1.80	2.10	4.09

Table 10.5: Mean radius of gyration of the CG PCL-40 and PCL-60 at different acetone molar fractions.

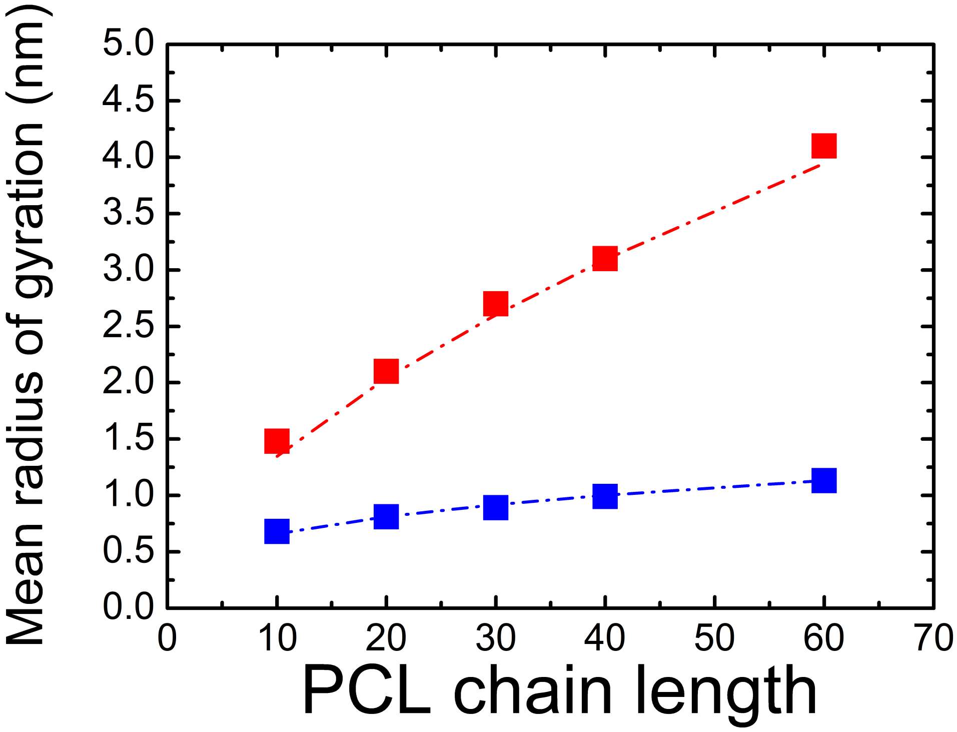


Figure 10.9: Flory's theory predictions (dashed lines) of the CG model developed here (discrete symbols) in pure good solvent (acetone, red) and pure bad solvent (water, blue) for all the PCL chain lengths investigated in this work. The dashed lines represent the Flory law reported in Eq. (10.9).

In pure solvents the CG model predictions are in an excellent accordance with the Flory's theory (dashed lines, Eq. 10.9), namely the mean radius of gyration scales with the Flory's exponent equal to 1/3 (blue dashed line) for bad solvent and to 3/5 (red dashed line) for good solvent. The accordance is shown to be slightly better in water, in which also the atomistic agreement was better than in pure acetone.

10.7 Conclusions

In this chapter a multiscale molecular dynamics approach for a polymer self-assembly in solution is proposed, in which classical all-atom molecular dynamics (MD) is combined together with the coarse-grained molecular dynamics (CGMD). The model developed here is based on the MARTINI force field (Marrink, Risselada, et al., 2007) and aims to build up a CG model for the poly-ε-caprolactone (PCL) in acetone, water and their mixtures. The adopted strategy consists in optimising both polymer - solvent and intrachain polymer - polymer non-bonded interactions, in terms of accordance respectively with thermodynamic and structural properties. Concerning the thermodynamics, the solvation free energy of single polymer beads was evaluated by means of the Bennett's Acceptance Ratio (BAR) method (Bennett, 1976). Regarding the structure of the polymer chains, the mean radius of gyration was chosen as target property, whose atomistic data are taken from a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014). Three different PCL chain lengths have been employed in this optimization procedure, that are PCL-10, -20 and -30 (i.e., 10, 20 and 30 repeat units). The model optimised on these three chain lengths was also validated for higher molecular weights (PCL-40 and -60).

Atomistic simulations related to the solvation free energy were carried out here, by means of different acetone charge distributions, in order to avoid the very well-known issue of the de-mixing of acetone-water mixtures (Lavino, Banetta, et al., 2018; Pereyra et al., 2011).

The CG mapping of the PCL was done by selecting at the beginning a non-polar bead type (N) for the ester part and an apolar (C) bead type for the alkyl part of the polymer repeat unit, as a starting choice. Thermodynamics results showed however that, in order to better reproduce the atomistic and experimental (at $x_A = 0$) free energy of solvation, a different mapping in terms of polarity degree of the PCL bead types must be chosen. This optimization procedure led to a bead P_2 (polar type) for the ester part and yet to a C (apolar) for the alkyl part of the PCL. Concerning the apolar bead type of the PCL, an interesting limit of this approach turns out from this work: a set of solely ten different levels of LJ interactions (which this version of the MARTINI force field lies on) are not sufficient to properly describe all the interactions and to reproduce the correct thermodynamics in binary mixtures made by good and bad solvents. For this reason, the bead type corresponding to the alkyl part was chosen to be labelled as a generic C_α, where the subscript α belongs to the range of degree of polarity [2,4] in the MARTINI interaction matrix.

Results in terms of accordance with atomistic mean radii of gyration depict also an interesting scenario. First of all, the optimization procedure is not trivial to afford since effects due to the chain length are important especially for single short CG chain lengths. It is therefore a sake of trade-off in finding the best set of intrachain non-bonded interactions parameters.

The agreement with atomistic results is very good in pure solvents and, more generally, far away from medium acetone molar fraction. At this threshold acetone concentration in mixture, a stepwise globule-to-coil transition is detected for the single PCL chain in solution. This trend was also observed by previous attempts (J. Cheng et al., 2014; Raman et al., 2017) and it turns out to be an intrinsic limitation of single CG macromolecules in binary mixtures, especially when the mixture itself is made by a good and a bad solvent. Investigation of higher PCL molecular weights shows the same trend observed for smaller chains, with the same limitations and advantages. The CG model shows also an excellent behaviour in reproducing the Flory's theory predictions in both pure good and bad solvents.

We can therefore conclude that the CG model developed here, based on the MARTINI force field, shows good results in terms of both thermodynamic and structural properties in binary mixtures. Interesting developments can be done in the future, by simulating multiple longer PCL chains in larger simulation boxes, in order to achieve a better insight into the dynamics of the polymer self-assembly.

Chapter 11

Population balance model and computational fluid dynamics in NP production

Most of this chapter has been published in Alessio D. Lavino, Nicodemo Di Pasquale, Paola Carbone, and Daniele L. Marchisio (2017). A novel multiscale model for the simulation of polymer flash nanoprecipitation, *Chemical Engineering Science* 171, 485-494.

11.1 Introduction

In this chapter the population dynamics combined with the effect of fluid dynamics on poly-ε-caprolactone (PCL) molecules self-assembly in acetone-water mixtures is studied.

Macromolecule self-assembly into molecular clusters or nanoparticles (NP) is investigated both with experimental and modeling approaches. Self-assembly is often described by using thermodynamics models designed to describe what happens at the equilibrium (Jones, 2002), whereas kinetics effects are well-known to play an important role (Johnson and Prud'homme, 2003b; Lince et al., 2008), in determining, for example, the size and structure of the final molecular clusters or nanoparticles (Celasco et al., 2014; Valente, Celasco, et al., 2012; Valente, Stella, et al., 2012).

As already explained in the Introduction of this thesis (Chapt. 1), in order to induce self-assembly, a particular technique named solvent-displacement is employed: PCL molecules are dissolved in a so-called 'good' solvent (e.g. acetone) and the solution is then mixed with a 'bad' solvent or anti-solvent (e.g. water) for the polymer, but with which the 'good' solvent is fully miscible. The presence of water induces the destabilization of the macromolecules, leading to their aggregation in molecular clusters (or nanoparticles) (Jones, 2002).

As it happens, the system (reaching the lowest energy state) can either undergo a pure spinodal separation in which macromolecules aggregate together without any resistance, or a metastable separation characterized by an energy barrier. In this latter case, the formation of clusters or NP is referred to as nucleation. The former instead consists of both a global and local instabilities defined with respect to small fluctuations in composition, and therefore the system undergoes a phase transition without any energy barrier to overcome. On the other hand, the latter is characterized by a global instability and local stability with respect to small fluctuations and thus it becomes necessary to consider an energy barrier. It is also worth mentioning here that often this solvent displacement process takes place in very small continuous mixers (look at Chapt. 1 for a complete list of them), such as the confined impinging jets mixer (Johnson and Prud'homme, 2003a) (CIJM) where solvent and anti-solvent are mixed together in a controlled way (Johnson and Prud'homme, 2003b).

In this chapter, a multiscale model, accounting for the most important phenomena involved, is validated. The model describes the macromolecule self-assembly as a purely aggregative phenomenon (spinodal decomposition), as also proposed by other authors (J.C. Cheng et al., 2010), but for the first time it employs kernels (to describe the rate with which macromolecules self-assemble) directly derived from molecular dynamics (MD) simulations (Di Pasquale, Marchisio, Barresi, and Carbone, 2014), resulting in a fully-predictive multiscale model, without adjusting parameters. Also for the first time this approach is validated by extensive comparison with experimental data derived under different operating conditions. This multiscale approach seems to be more effective than other alternatives, based for example on cluster prenucleation by using the classical nucleation theory (CNT) (Di Pasquale, Marchisio, and Barresi, 2012; Garcia et al., 2013) or other more detailed but similar theories (Di Pasquale, Marchisio, Carbone, et al., 2013).

The model implicitly assumes that aggregation is irreversible and is characterized by a unitary efficiency, as it will be explained better in section 11.2. The degree of aggregation and the resulting form of the cluster mass distribution (CMD) is quantified by the number of macromolecules belonging to a cluster (i.e., aggregation number) and by solving a population balance model (PBM). Moreover, in order to account for the inhomogeneous mixing between the acetone and water streams and for the influence of turbulence on the process, the PBM is solved with the quadrature method of moments (QMOM) (Marchisio, Vigil, et al., 2003) inside a computational fluid dynamics (CFD) code. Including the mixing in the description of the self-assembly is particularly important here, as the self-assembly process is very fast, almost instantaneous, and irreversible, and is therefore strongly influenced by spatial inhomogeneities (Johnson and Prud'homme, 2003a,b). Being, for the test cases considered in this work, the flow turbulent, it is also necessary to consider the effect of turbulent fluctuations on the aggregation process. This is done in the context of the Favre-averaged Navier-Stokes (FANS) equation approach, by using the direct quadrature method of moments coupled with the interaction and exchange with the mean (DQMOM-IEM) model (Fox, 2003).

The main novelty of the results collected and the methodology employed in this chapter stands in the relationship employed to calculate the radius of gyration of the single PCL molecule, based on the Flory law (Flory, 1953). The use of Flory's theory is particular important in this context as it allows to predict molecule conformation taking into account the solvent nanoscopic mixing effect on the chain entropy (Rubinstein and Colby, 2003). The coefficient and exponent appearing in the Flory law are calculated from MD simulations, performed in a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014) and reported in the operating conditions, section 11.3. These MD simulations results also show that the Flory law derived for a single macromolecule, can be safely extended to clusters of macromolecules. The use of MD to estimate the coefficient and exponent of the Flory law, together with the use of PBM and of CFD, make the model fully-predictive and without adjustable parameters. The model is in this chapter validated against experimental data referring to different mixing conditions in a CIJM, different initial polymer concentrations and two different molecular weights for the PCL. Comparison with experimental data allows to assess the range of validity of the approach, in line with the theory. A particular section is here dedicated also to the effect of other good solvents on mean NP size, showing a method to suitably adjust the Flory law, at this scale, when acetone is replaced by acetonitrile and THF.

11.2 Governing Equations

As already mentioned, turbulent mixing between acetone and water is described here with CFD. Turbulence is treated with the Favre-averaged Navier-Stokes equation (FANS) approach, introduced in Chapt. 5, Eq. (5.29), since two fluids with different densities are involved. The fluid can still be considered incompressible, but density fluctuations are accounted for in this way. The continuity equation for the acetone-water mixture is solved together with the Navier-Stokes equations and the equations for the turbulent kinetic energy, k, and the turbulent dissipation rate, ε, as common practice with the standard $k-\varepsilon$ model (section 5.2.2). All these equations are reported in Chapt. 5. In the following subsection, the detailed PBM for FNP will be shown. Here, we will be referring to acetone as good solvent, but the model is valid for any choice of good solvents, as it will be shown in section 11.4.1.

11.2.1 Population Balance Model for FNP

In order to efficiently describe the evolution of the population of polymer clusters, a PBM is employed. The PBM operates on the CMD, $f(\mathbf{x}, n)$, defined in such a way that the quantity, $f(\mathbf{x}, n)\mathrm{d}n$, represents the number density of molecular clusters containing n macromolecules at position $\mathbf{x}$. The variable n, referred to in what follows as dimensionless cluster mass or aggregation number, is discrete, but since it varies between one and very large numbers, is treated here as a continuous variable. It is also useful to highlight here that, in the acetone inlet, when the macromolecules are dissolved in the solvent, and do not show any tendency to self-assemble, the CMD is always equal to zero, except for $n = 1$, where it assumes the value corresponding to the initial macromolecule concentration (or number density) in acetone. Moreover, as the values assumed by the CMD would inevitably be very large, the CMD is normalized by dividing it to the Avogadro number. By using this normalization the CMD in the acetone inlet corresponds to the initial polymer molar concentration.

In line with the theory of PBM reported in Chapt. 5, section 5.3.2, the evolution of the CMD, for a stationary problem and before Favre average, reads as follows:

$$\frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{U} f(n)) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\Gamma(n) \frac{\partial f(n)}{\partial \mathbf{x}} \right) = \\ = \frac{1}{2} \int_0^n \beta(n-n', n') f(n-n') f(n') \mathrm{d}n' - \int_0^\infty \beta(n, n') f(n) f(n') \mathrm{d}n', \quad (11.1)$$

where space dependencies are omitted for brevity, $\Gamma(n)$ is the diffusion coefficient due to Brownian motions of a cluster of size n, $\mathbf{U}$ is the cluster velocity and $\beta(n, n')$ is the second-order point aggregation kernel, previously introduced in section 5.3.3. It is assumed here that $\mathbf{U}$ is identical to the fluid velocity, as clusters are characterized by very small particle Stokes number (Baldyga and Orciuch, 2001) (i.e. pseudo-homogeneous system hypothesis).

As already seen in section 5.3.3, the aggregation kernel quantifies the rate at which clusters of different size aggregate together forming larger clusters and its specific functional form, in the context of aggregation number as internal coordinate of the PBE, will be introduced later on. It is important to stress here that aggregation results in two source terms, one positive related to the formation of a cluster containing n macromolecules and one negative due to the disappearance of clusters formed by n' macromolecules. Consequently, the term $\beta(n = 1, n' = 1)$ corresponds to the rate with which individual molecules disappear forming dimers.

Since the solution of the PBM with discretized methods (Marchisio and Fox, 2013) leads to very high computational costs, it is more convenient to solve the problem in terms of the moments of the CMD. Let us recall the defintion of the generic j^{th}-order moment:

$$m^{(j)} = \int_0^\infty f(n) n^j dn. \quad (11.2)$$

An additional advantage of solving the PBM in terms of the moments of the CMD is that they represent physical measurable quantities; in fact, $m^{(0)}$ represents the total cluster number density, $m^{(1)}$ is a conserved quantity and represents the total number density of macromolecules, whereas the ratio between $m^{(1)}$ and $m^{(0)}$ results in the average number of macromolecules per cluster. By applying the moment transform and the Favre average the steady-state transport equation for the moment of order j

assumes the following form:

$$\frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\langle \mathbf{U}\rangle \left\langle m^{(j)}\right\rangle\right) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\Gamma_t \frac{\partial \left\langle m^{(j)}\right\rangle}{\partial \mathbf{x}}\right) = \\ = \frac{\bar{\rho}}{2}\left\langle \int_0^\infty \int_0^\infty \left[(n+n')^j - n^j - n'^j\right] \beta(n,n') f(n) f(n') dn dn' \right\rangle, \tag{11.3}$$

where during derivation it was assumed that the turbulent diffusivity, $\Gamma_t \approx \nu_t/0.7$, which does not depend on the moment order, is much larger than the Brownian diffusivity and is calculated from the turbulent viscosity, ν_t, by assuming a turbulent Schmidt number of 0.7 (Andersson et al., 2012), $\bar{\rho}$ is the time-averaged fluid density and $\langle \mathbf{U}\rangle$ is the Favre-averaged fluid velocity.

Let us now define the functional form of the aggregation kernel, $\beta(n, n')$, stating the rate of aggregation of clusters of size n with clusters of size n', appearing in Eq. (11.3). To do that it is useful to know that our MD simulations, performed in a previous work (Di Pasquale, Marchisio, and Barresi, 2012), have shown that the PCL molecule behaves as a freely-jointed chain (Rubinstein and Colby, 2003). This observation allows to assume that the behavior of a molecular cluster with n PCL molecules of molecular weigth M_w, is similar to the behavior of a single PCL molecule with molecular weight nM_w. This assumption allows to extend the Flory law, valid for a single PCL molecule:

$$[\![{R_g}^2]\!] = k {M_w}^{2\nu}, \tag{11.4}$$

to a cluster of n PCL molecules:

$$[\![{R_g}^2(n)]\!] = k (n M_w)^{2\nu}, \tag{11.5}$$

where $[\![{R_g}^2]\!] = [\![{R_g}^2(n=1)]\!]$ is the ensemble-averaged squared radius of gyration of a single PCL molecule, $[\![{R_g}^2(n)]\!]$ is the ensemble-averaged squared radius of gyration of a PCL cluster with n molecules, M_w is the molecular weight of a single PCL chain and k and ν are the Flory parameters, that depend on the acetone molar fraction x_A in the acetone-water mixture. Their functional dependences on x_A were determined in a previous work via MD simulations (Di Pasquale, Marchisio, and Barresi, 2012) and are reported in the next section, in Eq. (11.34) and Eq. (11.35).

The aggregation kernel accounts for two mechanisms: collisions due to Brownian motions and those due to turbulent fluctuations. By using the following approximation: $[\![R_g]\!] \simeq \sqrt{[\![{R_g}^2]\!]}$ and by assuming that the diffusion coefficient due to Brownian motions can be calculated with the Stokes-Einstein equation (look at section 5.3.3), the following expression is obtained (J.C. Cheng et al., 2010):

$$\begin{aligned}
\beta(n,n') &= \\
&= \eta\left(\frac{2k_BT}{3\mu}\frac{\left(\llbracket R_g(n)\rrbracket+\llbracket R_g(n')\rrbracket\right)^2}{\llbracket R_g(n)\rrbracket\llbracket R_g(n')\rrbracket}+1.2944\sqrt{\frac{\varepsilon}{\nu}}\left(\llbracket R_g(n)\rrbracket+\llbracket R_g(n')\rrbracket\right)^3\right)= \\
&= \eta\left(\frac{2k_BT}{3\mu}\frac{\left(\sqrt{k\left(nM_w\right)^{2\nu}}+\sqrt{k\left(n'M_w\right)^{2\nu}}\right)^2}{\sqrt{k\left(nM_w\right)^{2\nu}}\sqrt{k\left(n'M_w\right)^{2\nu}}}+\right. \\
&\left.\quad+1.2944\sqrt{\frac{\varepsilon}{\nu}}\left(\sqrt{k\left(nM_w\right)^{2\nu}}+\sqrt{k\left(n'M_w\right)^{2\nu}}\right)^3\right), \qquad (11.6)
\end{aligned}$$

where η is the aggregation efficiency that will be defined later on, together with μ, the molecular viscosity, both exhibiting a strong dependence on the acetone molar fraction x_A, where k_B is the Boltzmann constant, ε is the turbulent dissipation rate and ν is the mixture kinematic viscosity. Derivation of Eq. (11.6) also made use of the assumption that the two aggregation mechanisms are simply additive: at low PCL initial concentrations and low turbulence intensity the Brownian aggregation mechanism dominates on the turbulent one and viceversa at high PCL initial concentrations and high turbulence intensity. The whole procedure to derive Eq. (11.6) is extensively explained in section 5.3.3.

It is worth showing that Eq. (11.6), after substitution of Eq. (11.5), can be semplified as follows:

$$\beta(n,n')=\eta\left(\frac{2k_BT}{3\mu}\frac{\left(n^{\nu}+n'^{\nu}\right)^2}{n^{\nu}n'^{\nu}}+1.2944\sqrt{\frac{\varepsilon}{\nu}}\left(k^{\frac{3}{2}}M_w^{3\nu}\right)\left(n^{\nu}+n'^{\nu}\right)^3\right). \qquad (11.7)$$

As it can be seen, the Brownian term is in line with the theory since it does not depend on the PCL molecular weight. In fact, thanks to the Stokes-Einstein theory, the diffusion coefficient is inversely proportionally to the molecular weight, whereas the radius of gyration increases over the molecular weight, and so the two effects cancel out. If Brownian aggregation were the only acting process, outcome would be the same in terms of n but not in terms of radius of gyration, because this latter depends on molecular weight.

The integrals that appear on the right-hand side of Eq. (11.3) are affected by a twofold closure problem, one due to the impossibility of writing the integrals in terms of the moments of the CMD and one related to the fact that turbulent fluctuations, that affect the CMD, are characterized by time-scales comparable with that of aggregation. To overcome these two closure problems a synergic strategy is employed: by using QMOM to overcome the first and DQMOM-IEM to overcome the second. In the next sections some information concerning the use of these two methods will be given.

11.2.2 QMOM and DQMOM-IEM for Flash Nano-Precipitation

In a turbulent flow all its properties, including the moments of the CMD, $m^{(j)}$, fluctuate. In this work these fluctuations are characterized in terms of the probability density function (PDF) for the acetone mass fraction, ξ, in the acetone-water mixture. Following the DQMOM-IEM (sections 8.5 and 8.6) formalism, this PDF is approximated as a summation of N_e delta functions: $p(\mathbf{x},\xi) \approx \sum_{i=1}^{N_e} p_i(\mathbf{x})\delta\left[\xi - \xi_i(\mathbf{x})\right]$, where p_i are known as weights and ξ_i as nodes of the underlying quadrature approximation. By setting $N_e = 2$ one has that: $i = 1{,}2$, and the following transport equation is solved:

$$\frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\langle \mathbf{U}\rangle p_1\right) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\Gamma_t \frac{\partial p_1}{\partial \mathbf{x}}\right) = 0, \tag{11.8}$$

whereas no equation for p_2 is solved, as $p_2 = 1 - p_1$ (i.e. the PDF integrates to unity), together with:

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\langle \mathbf{U}\rangle p_1\xi_1\right) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\Gamma_t \frac{\partial}{\partial \mathbf{x}}(p_1\xi_1)\right) = \\ = \bar{\rho}\gamma_M p_1 p_2\left(\xi_2 - \xi_1\right) + \frac{\bar{\rho}\Gamma_t}{\xi_1 - \xi_2}\left(p_1\frac{\partial \xi_1}{\partial \mathbf{x}} \cdot \frac{\partial \xi_1}{\partial \mathbf{x}} + p_2\frac{\partial \xi_2}{\partial \mathbf{x}} \cdot \frac{\partial \xi_2}{\partial \mathbf{x}}\right).\end{aligned} \tag{11.9}$$

In Eq. (11.9) the right-hand side represents molecular mixing closed with the IEM model and $\gamma_M = \frac{C_\phi}{2}\frac{\varepsilon}{k}$ is the micromixing rate, where C_ϕ is a constant. More details can be found in section 8.6. Inverting subscripts 1 and 2, the transport equation related to $p_2\xi_2$, can be readily obtained. It is important to stress here that ξ_1 and ξ_2 can be thought of as local acetone mass fractions in two 'environments', representing turbulent composition fluctuations. Following the DQMOM-IEM formalism the Favre-averaged acetone mass fraction can be calculated as follows:

$$\langle \xi \rangle = \int_0^1 p(\xi)\xi \mathrm{d}\xi \approx p_1\xi_1 + p_2\xi_2. \tag{11.10}$$

By using this formalism, together with the modeling details presented in section 8.7, the moments of the CMD are transported in the following equation:

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\langle \mathbf{U}\rangle p_1 m_1^{(j)}\right) - \frac{\partial}{\partial \mathbf{x}} \cdot \left(\bar{\rho}\Gamma_t \frac{\partial}{\partial \mathbf{x}}\left(p_1 m_1^{(j)}\right)\right) = \\ = \bar{\rho}\gamma_M p_1 p_2\left(m_2^{(j)} - m_1^{(j)}\right) + \frac{\bar{\rho}\Gamma_t}{m_1^{(j)} - m_2^{(j)}}\left(p_1\frac{\partial m_1^{(j)}}{\partial \mathbf{x}} \cdot \frac{\partial m_1^{(j)}}{\partial \mathbf{x}} + p_2\frac{\partial m_2^{(j)}}{\partial \mathbf{x}} \cdot \frac{\partial m_2^{(j)}}{\partial \mathbf{x}}\right) + \\ + \frac{p_1\bar{\rho}}{2}\sum_{\alpha=1}^{N}\sum_{\gamma=1}^{N} w_{\alpha,1}w_{\gamma,1}\beta_{\alpha,\gamma,1}\left[\left(n_{\alpha,1} + n_{\gamma,1}\right)^j - n_{\alpha,1}^j - n_{\gamma,1}^j\right],\end{aligned} \tag{11.11}$$

where again inverting the subscripts 1 and 2, the equation for $m_2^{(j)}$ is readily obtained. In Eq. (11.11) the source term due to aggregation is approximated by using QMOM. The N nodes and weights of the quadrature approximation $n_{\alpha,1}$, $w_{\alpha,1}$ are calculated from the first $2N$ moments in the first environment, $m_1^{(j)}$, by using the Product-Difference (PD) algorithm (and similarly for environment 2). Details concerning QMOM can be found elsewhere (Marchisio and Fox, 2013). In this work N is taken equal to two, transporting therefore four moments (for each of the two environments), since this value represents a good trade off between stability, computational cost and accuracy (Barrett and Webb, 1998). The Favre-averaged moment of order k, $\langle m \rangle^{(j)}$, can be calculated again from the DQMOM-IEM formalism, resulting in the following equation:

$$\left\langle m^{(j)} \right\rangle = p_1 m_1^{(j)} + p_2 m_2^{(j)}. \tag{11.12}$$

By using the same approach, also the final Favre-averaged mean radius of gyration, $\langle R_g \rangle$, for the entire population of clusters can be calculated as the summation weighted over the two environments, as shown in the following equation:

$$\langle R_g \rangle = \sum_{i=1}^{N_e} p_i \sum_{\alpha=1}^{N} \left[\frac{w_{\alpha,i} [\![R_{g,\alpha,i}]\!]}{\sum_{\alpha=1}^{N} w_{\alpha,i}} \right] = \sum_{i=1}^{N_e} p_i \sum_{\alpha=1}^{N} \left[\frac{w_{\alpha,i} \sqrt{k(x_{A,i}) \left(n_i M_w\right)^{2\nu(x_{A,i})}}}{\sum_{\alpha=1}^{N} w_{\alpha,i}} \right] \tag{11.13}$$

where $R_{g,\alpha,i}$, $k(x_{A,i})$ and $\nu(x_{A,i})$ are respectevely the radius of gyration, the Flory coefficient and the Flory exponent, functions of the acetone molar fraction x_A in the environment i, with $i = 1{,}2$, and in the quadrature node α, with $\alpha = 1{,}2$ calculated by means of QMOM. Functional forms of Flory's parameters $k(x_{A,i})$ and $\nu(x_{A,i})$ are directly evaluated from MD simulations (Di Pasquale, Marchisio, Barresi, and Carbone, 2014) and their expressions are reported in Eq. (11.34) and Eq. (11.35). The value obtained from Eq. (11.13) will be then compared against experiments.

11.2.3 CQMOM as alternative model for FNP

In this section, the conditional quadrature method of moments (CQMOM) applied to FNP is developed, in line with the theory presented in section 8.4.1. Let us define the bivariate NDF $g(t, \mathbf{x}, \xi, n)$, in which two internal coordinates are considered, namely the mixture fraction ξ and the number of molecules n that form a cluster, also identified as aggregation number. The quantity $g(t, \mathbf{x}, \xi, n)\mathrm{d}\xi\mathrm{d}n$ corresponds therefore to the number of molecules per unit volume which form a molecular cluster, at a given position $\mathbf{x}$, at the time instant t in a blob of liquid with composition ξ. For a sake of brevity, from now on time and space coordinates will be omitted in the NDF, resulting in $g(\xi, n)$, but the two forms are totally equivalent. In line with the conditional quadrature method of moments (CQMOM) theory, the NDF $g(\xi, n)$ is defined as

$$g(\xi, n) = \sum_{\alpha=1}^{N_1} p_\alpha \delta(\xi - \xi_\alpha) \sum_{\gamma=1}^{N_2} w_{\alpha,\gamma} \delta(n - n_{\alpha,\gamma}) \tag{11.14}$$

where N_1 and N_2 are the number of nodes of the quadrature referred respectively to the first and second internal coordinate. Choosing $N_1 = N_2 = 2$ represents a good compromise between computational cost and accuracy, leading to the quadrature scheme depicted in Figure 11.1.

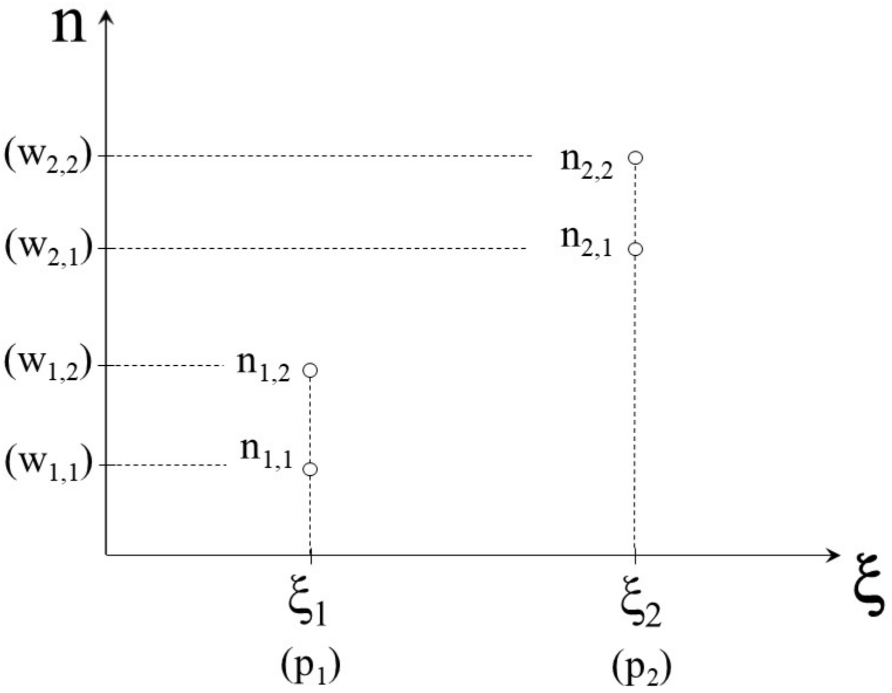


Figure 11.1: Sketch of the quadrature procedure with two nodes for the mixing ξ and the aggregation number n.

The mixed moment of order k_1 in ξ and k_2 in n is defined as follows:

$$m_{k_1,k_2} = \langle \xi^{k_1} n^{k_2} \rangle = \int_0^1 \int_1^\infty \xi^{k_1} n^{k_2} g(\xi, n) \mathrm{d}\xi \mathrm{d}n = \\ = \int_0^1 \xi^{k_1} f(\xi) \mathrm{d}\xi \int_1^\infty n^{k_2} f(n|\xi) \mathrm{d}n, \tag{11.15}$$

where $f(n|\xi)$ is the marginal NDF, defined as:

$$f(n|\xi) = \frac{g(\xi, n)}{f(\xi)}, \tag{11.16}$$

and where

$$f(\xi) = \frac{\int_1^\infty g(\xi, n)\mathrm{d}n}{N_t}, \tag{11.17}$$

where N_t is the total clusters number density, namely the 0^{th}-order moment in the variable n.

The interchangeable notation $m_{k_1,k_2} = \langle \xi^{k_1} n^{k_2} \rangle$ is used again and from now on will be used very frequently.

Instead of transporting and solving the NDF $g(\xi, n)$, it turns out to be more strategic and interesting to consider the NDF $h(\xi, \mathbf{m})$, where

$$\mathbf{m} = \{m_{k_2}\} = \int_{\Omega_n} n^{k_2} f(n)\mathrm{d}n, \quad k_2 = 0, 1, 2, 3 \quad = \begin{pmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{pmatrix} \tag{11.18}$$

is the pure (in n variable) moments vector, considered now as an internal variable of the PBM. Being the number of nodes for the quadrature in n equal to two, the first $2N_2 - 1$ order moments are considered, namely $m_0, ..., m_3$.

Coherently with the multivariate distribution theory, the new NDF is defined as follows:

$$h(\xi, \mathbf{m}) = \sum_{\alpha=1}^{N_1} p_\alpha \delta(\xi - \xi_\alpha)\delta(\mathbf{m} - \mathbf{m}_\alpha). \tag{11.19}$$

The transport equation of the joint NDF $h(\xi, \mathbf{m})$ reads as follows:

$$\begin{aligned} \frac{\partial h(\xi, \mathbf{m})}{\partial t} + \langle \mathbf{U} \rangle \cdot \frac{\partial h(\xi, \mathbf{m})}{\partial \mathbf{x}} - \frac{\partial}{\partial \mathbf{x}} \cdot \left[\Gamma_t \frac{\partial}{\partial \mathbf{x}} h(\xi, \mathbf{m}) \right] &= \\ = -\frac{\partial}{\partial \xi} \left(\frac{\epsilon_\phi}{2\langle \xi'^2 \rangle} (\langle \xi \rangle - \xi) h(\xi, \mathbf{m}) \right) &+ \\ -\frac{\partial}{\partial \mathbf{m}} \left(\frac{\epsilon_\phi}{2\langle \xi'^2 \rangle} (\langle \mathbf{m} \rangle - \mathbf{m}) h(\xi, \mathbf{m}) \right) &+ \\ -\frac{\partial}{\partial \mathbf{m}} \left[R(\xi, \mathbf{m}) h(\xi, \mathbf{m}) \right], & \end{aligned} \tag{11.20}$$

where the two transport terms in mesoscale phase space (first two terms on the right hand side) have been closed by means of the IEM model.

Let us introduce the new mixed moments for this alternative mathematical framework:

$$\hat{m}_{k_1,k_2} = \langle \xi^{k_1}\mathbf{m}\rangle = \int_{\Omega_\xi}\int_{\Omega_n} \xi^{k_1}\mathbf{m}h(\xi,\mathbf{m})\mathrm{d}\xi\mathrm{d}\mathbf{m}. \tag{11.21}$$

By multiplying Eq. (11.20) by $\xi^{k_1}\mathbf{m}d\xi d\mathbf{m}$ and by integrating over all the phase space domain, from Eq. (11.21) the mixed moments transport equation is obtained:

$$\begin{aligned}
&\frac{\partial\langle\xi^{k_1}\mathbf{m}\rangle}{\partial t} + \langle\mathbf{U}\rangle\cdot\frac{\partial\langle\xi^{k_1}\mathbf{m}\rangle}{\partial\mathbf{x}} - \frac{\partial}{\partial\mathbf{x}}\cdot\left[\Gamma_t\frac{\partial}{\partial\mathbf{x}}\langle\xi^{k_1}\mathbf{m}\rangle\right] = \\
&\qquad = -\int_{\Omega_\xi}\int_{\Omega_n}\xi^{k_1}\mathbf{m}\frac{\partial}{\partial\xi}\left(\frac{\epsilon_\phi}{2\langle\xi'^2\rangle}(\langle\xi\rangle-\xi)h(\xi,\mathbf{m})\right)\mathrm{d}\xi\mathrm{d}\mathbf{m}+ \\
&\qquad -\int_{\Omega_\xi}\int_{\Omega_n}\xi^{k_1}\mathbf{m}\frac{\partial}{\partial\mathbf{m}}\left(\frac{\epsilon_\phi}{2\langle\xi'^2\rangle}(\langle\mathbf{m}\rangle-\mathbf{m})h(\xi,\mathbf{m})\right)\mathrm{d}\xi\mathrm{d}\mathbf{m}+ \\
&\qquad -\int_{\Omega_\xi}\int_{\Omega_n}\xi^{k_1}\mathbf{m}\frac{\partial}{\partial\mathbf{m}}\left[R(\xi,\mathbf{m})h(\xi,\mathbf{m})\right]\mathrm{d}\xi\mathrm{d}\mathbf{m}.
\end{aligned} \tag{11.22}$$

By using the micro-mixing rate introduced in Chapt. 8, by integrating by parts the left hand side of Eq. (11.22) (in order to make the phase space derivatives vanish off, see also Appendix C), and by means of the definition reported in Eq. (11.21), the general transport equation of the mixed moments of order k_1 (note that $k_2 = 0, 1, 2, 3$ is implicitly present in the internal coordinate vector $\mathbf{m}$, as reported in Eq. (11.18)) becomes:

$$\begin{aligned}
&\frac{\partial\langle\xi^{k_1}\mathbf{m}\rangle}{\partial t} + \langle\mathbf{U}\rangle\cdot\frac{\partial\langle\xi^{k_1}\mathbf{m}\rangle}{\partial\mathbf{x}} - \frac{\partial}{\partial\mathbf{x}}\cdot\left[\Gamma_t\frac{\partial}{\partial\mathbf{x}}\langle\xi^{k_1}\mathbf{m}\rangle\right] = \\
&\qquad = +\,k_1\gamma_M(\langle\xi\rangle\langle\xi^{k_1-1}\mathbf{m}\rangle - \langle\xi^{k_1}\mathbf{m}\rangle) + \\
&\qquad\quad +\,\gamma_M(\langle\xi^{k_1}\rangle\langle\mathbf{m}\rangle - \langle\xi^{k_1}\mathbf{m}\rangle) + \langle\xi^{k_1}R(\xi,\mathbf{m})\rangle.
\end{aligned} \tag{11.23}$$

In order to close the quadrature scheme of this particular problem, only two equations, corresponding to $k_1 = 0$ and $k_1 = 1$, must be solved leading to:

$$\begin{aligned}
&\frac{\partial\langle\mathbf{m}\rangle}{\partial t} + \langle\mathbf{U}\rangle\cdot\frac{\partial\langle\mathbf{m}\rangle}{\partial\mathbf{x}} - \frac{\partial}{\partial\mathbf{x}}\cdot\left[\Gamma_t\frac{\partial}{\partial\mathbf{x}}\langle\mathbf{m}\rangle\right] = \langle R(\xi,\mathbf{m})\rangle, \\
&\frac{\partial\langle\xi\mathbf{m}\rangle}{\partial t} + \langle\mathbf{U}\rangle\cdot\frac{\partial\langle\xi\mathbf{m}\rangle}{\partial\mathbf{x}} - \frac{\partial}{\partial\mathbf{x}}\cdot\left[\Gamma_t\frac{\partial}{\partial\mathbf{x}}\langle\xi\mathbf{m}\rangle\right] = \\
&\qquad =2\gamma_M(\langle\xi\rangle\langle\mathbf{m}\rangle - \langle\xi\mathbf{m}\rangle) + \langle\xi R(\xi,\mathbf{m})\rangle,
\end{aligned} \tag{11.24}$$

where $R(\xi, \mathbf{m})$ is the source term (containing the aggregation kernels) already defined in the previous section, and

$$\langle R(\xi, \mathbf{m}) \rangle = \sum_{\alpha=1}^{2} p_\alpha R(\xi_\alpha, \mathbf{m}_\alpha), \tag{11.25}$$

$$\langle \xi R(\xi, \mathbf{m}) \rangle = \sum_{\alpha=1}^{2} p_\alpha \xi_\alpha R(\xi_\alpha, \mathbf{m}_\alpha), \tag{11.26}$$

closed with the CQMOM.
The solving scheme for this approach consists in the same iteration in:

- Calculating the quadrature in the first phase space variable ξ ; thanks to the pure moments in ξ and a suitable inversion algorithm (e.g., Wheeler), the nodes and weights $\{p_\alpha, \xi_\alpha\}$ are obtained;
- The two mixed moments reported in Eq.s (11.24) are solved and the mixed moments $\langle \mathbf{m} \rangle$ and $\langle \xi \mathbf{m} \rangle$ are then known;
- The linear system based on the mixed moments definition is solved:

  $$\begin{cases} \langle \mathbf{m} \rangle = p_1 \mathbf{m}_1 + p_2 \mathbf{m}_2 \\ \langle \xi \mathbf{m} \rangle = p_1 \xi_1 \mathbf{m}_1 + p_2 \xi_2 \mathbf{m}_2 \end{cases} \tag{11.27}$$

  and the pure moments in the two environments of the mixing quadrature $\mathbf{m}_1$ and $\mathbf{m}_2$ are obtained.
- By applying the Wheeler algorithm to the moments set vector $\mathbf{m}_1$ and $\mathbf{m}_2$ (remember the definition reported in Eq. 11.18), the weights and nodes of the quadrature of the pure moments in n, $\{w_{\alpha,\gamma}, n_{\alpha,\gamma}\}$ (look at Figure 11.1), are found and can be employed in the calculation of the mean radius of gyration.

This alternative mathematical framework, presented in this section, has been implemented in an open-source tool, OpenQBMM, an OpenFOAM-based package. Being still under numerical and computational optimizations, all the results here presented are obtained by means of the previous model shown in section 11.2.2.

11.3 Operating conditions and computational details

In order to validate the model developed, predictions are compared with experiments obtained under different operating conditions. In the experiments the water and acetone streams are mixed in a CIJM continuosly, as shown in Figure 11.2. First of all, two molecular weights for PCL are simulated: the first one has $M_w = 14000$

g mol^{-1} (indicated in what follows as PCL-14000), whereas the second one has M_w = 80000 g mol^{-1} (indicated as PCL-80000). For both of them, different operating conditions have been considered, in terms of initial PCL concentrations (in the acetone stream) and inlet acetone and water flow rates, whereas the flow rate ratio between water and acetone was kept equal to one. Therefore, the final acetone concentration at the CIJM outlet in terms of volume, molar and mass fraction is respectively equal to 0.5, 0.2 and 0.44.

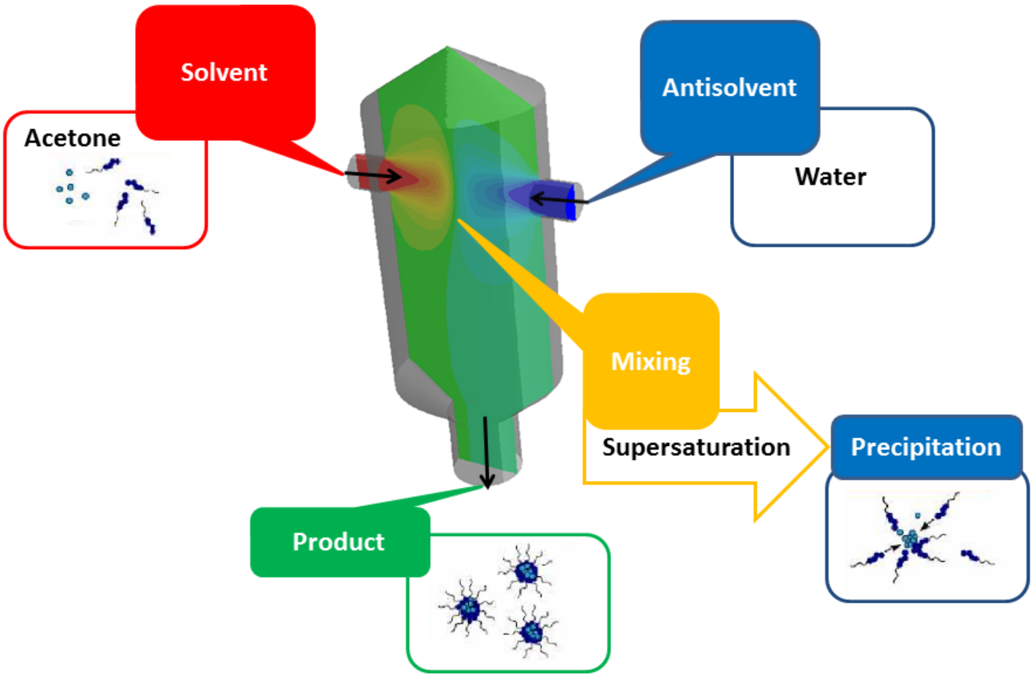


Figure 11.2: Sketch of CIJM. Acetone, with PCL dissolved, and water flow respectively from left and right. Black arrows show the flux lines. There may be also organic molecules (e.g. drug, light blue beads in the figure). The mixing leads to supersaturation and to precipitation of polymeric clusters in micellar structures that may contain the organic molecules. In this work the presence of the drug is neglected.

The operating conditions related to PCL-14000 and PCL-80000 are summarized in Tab. 11.1. In order to clearly explain the operating conditions considered, a double-entry table is employed. The supersaturation ratio $S(\xi)$ (fourth column in Tab. 11.1) is function of the local good solvent mixture fraction ξ and is calculated as

$$S(\xi) = \frac{c_{loc}^{PCL}(\xi)}{c_{eq}^{PCL}(\xi)}, \tag{11.28}$$

where c_{loc}^{PCL} is the local PCL concentration, after mixing of water and acetone streams, and c_{eq}^{PCL} is the equilibrium PCL concentration, both of them defined later on respectively in Eq. (11.29) and Eq. (11.30). In Tab. 11.1 the $S(\xi)$ is calculated at the outlet of the mixer, therefore by taking the PCL concentrations reported in the third column of the table. This value is very meaningful since it represents the driving force for self-assembly. Moderate $S(\xi)$ values probably corrispond to nucleation, whereas large $S(\xi)$ values probably corrispond to spinodal decomposition.

$c_{in,PCL}$, $mg\ mL^{-1}$	$c_{in,PCL}$, $mol\ m^{-3}$	$c_{out,PCL}$, $mol\ m^{-3}$	$S(\xi)$	**Flow rate,** $mL\ min^{-1}$					
				10	**20**	**40**	**60**	**80**	**120**
0.5	0.0357	0.0179	23.3	PCL-14000					
2.5	0.1786	0.0893	116.7						
5.0	0.3571	0.1786	233.4						
10.0	0.7143	0.3571	466.7						
15.0	1.0714	0.5357	700.1						
25.0	1.7857	0.8929	1166.8						
0.02	0.00025	0.000125	0.9	PCL-80000					
0.2	0.0025	0.00125	9.3						
1.0	0.0125	0.00625	46.7						
3.0	0.0375	0.01875	140.0						
4.0	0.0500	0.025	186.7						
6.0	0.0750	0.0375	280.0						

Table 11.1: Operating conditions adopted in the simulations related to PCL-14000 and to PCL-80000, with acetone as good solvent in the CIJM. The first and second columns refer to the initial PCL concentration in the acetone stream in mg mL^{-1} and mol m^{-3}. The third column refers to the final PCL concentration in the outlet stream (after mixing of acetone and water streams). The fourth column quantifies the supersaturation ratio ($S(\xi)$) at the outlet of the mixer.

The case referring to PCL with the smaller molecular weight (i.e. PCL-14000) is first investigated only under the effect of Brownian-induced aggregation and then subsequently under the effect of both Brownian-induced and turbulent-induced aggregation, in order to assess the relative importance of the two mechanisms. For the other molecular weight (i.e. PCL-80000) only predictions accounting for both effects will be discussed.

As already mentioned, PCL molecule self-assembly is induced by a change in the solvent. The aggregation efficiency, η, introduced in Eq. (11.6) is assumed to be zero if the local PCL molecule concentration in environment i, $c_{loc,i}^{PCL}$, is smaller than the equilibrium concentration and equal to one otherwise. The physical meaning behind this modelling choice deserves a further explanation. The aggregation efficiency is represented by a stepwise function, zero for undersaturated solutions and one for supersaturated solutions. This means that if locally the supersaturation ratio overcomes its solubility threshold of unity, and if two clusters/NP collide, then they stick together undergoing an aggregation event (second-order point process). This is physically justified by the fact that, locally, a PCL molecule (or cluster) is interacting with a poor solvent condition, since the mixing between the good and bad solvent is extremely fast and characterised by a very low variance (the system tends to a perfect micro-mixed). Therefore, the PCL molecule (or cluster) once the solubility limit is overcome (supersaturation ratio greater than unity) will always face a twofold choice: either interacting with a poor solvent molecule, or with another PCL molecule (cluster). It is a matter of fact that due to the low compatibility with water, the PCL chains would always prefer to interact with themselves, sticking together rather than exposing to the poor solvent (much lower energy interaction). Besides the energetic effect just introduced here, there is also an entropic effect that is worthwhile to mention. At the beginning single PCL macromolecules are completely dissolved in pure acetone stream and this represents the best possible configuration. As soon as water and acetone streams collide and mix together, the single PCL molecules can be either in their mono-macromolecular status (the cluster distribution is a delta centred on unity, namely each cluster can be thought as formed by just one single PCL macromolecule), or can aggregate together forming two-macromolecule cluster, a three-macromolecule cluster and so on. The single PCL chains system will correspond to a given number of possible configurations, but the clusters that would form from a possible aggregation will have many more configurations that describe the same aggregate and consequently, the aggregated system turns out to be more probable, from an entropic point of view. That is why the system is entropically and energetically favourable to spontaneous self-assemble, driven by the gradient of the chemical potential (that decreases as long as the aggregation takes place reaching a plateau), which is simply another form to express the entropic and enthalpic contributions described above. All of these phenomena are local phenomena, strictly dependent on a macroscopic quantity that is the PCL initial concentration in acetone stream.

The local PCL concentration is calculated in the environment i from the local mixture fraction, ξ_i, via the following equation:

$$c_{loc,i}^{PCL}(\xi_i) = \frac{c_{in}^{PCL}}{M_w}\left[1 + \left(\frac{1}{\xi_i} - 1\right)\left(\frac{\rho_A}{\rho_W}\right)\right]^{-1}, \tag{11.29}$$

where M_w is the PCL molecular weight in g mol^{-1}, c_{in}^{PCL} is the initial PCL concentration expressed in mg mL^{-1}, ρ_A and ρ_W are respectively the density of acetone and water, and ξ_i is the mass fraction of acetone in the environment i; whereas the equilibrium concentration in environment i, $c_{eq,i}^{PCL}(\xi_i)$, was determined experimentally in a previous work (Lince et al., 2008) and fitted (Di Pasquale, Marchisio, and Barresi, 2012) obtaining the exponential correlation:

$$c_{eq,i}^{PCL}(\xi_i) = \frac{1200}{M_w} \cdot \exp\left[-14.533\left(1 - x_{A,i}(\xi_i)\right)\right], \tag{11.30}$$

where M_w is the PCL molecular weight in g mol^{-1}, $x_{A,i}(\xi_i)$ is the good solvent (i.e., acetone) molar fraction in the environment i and its dependence on the mixture fraction is reported in Eq. (11.33). The PCL equilibrium concentration profiles are reported in Figure 11.3 and 11.4 respectively for PCL-14000 and PCL-80000. The figures report the equilibrium condition in terms of the molar acetone fraction, x_A, as a function of the PCL equilibrium concentration together with the operating conditions investigated in the experiments. The symbols report the final PCL concentrations in the acetone-water mixture with the final solution composition, after complete mixing of the acetone and water solutions.

The time-averaged acetone-water mixture density is calculated from the following equation:

$$\bar{\rho} = \left(\frac{\langle\xi\rangle}{\rho_A} + \frac{1-\langle\xi\rangle}{\rho_W}\right)^{-1}, \tag{11.31}$$

where ρ_A and ρ_W are the acetone and water densities, respectively, and $\langle\xi\rangle$ is the Favre-averaged mixture fraction, namely the acetone mass fraction in the mixture. The viscosity of the acetone-water mixture is instead calculated with the following empirical correlation:

$$\mu(x_{A,i}) = \exp\left[x_{A,i}\ln\left(\mu_A\right) + (1 - x_{A,i})\ln\left(\mu_W\right)\right], \tag{11.32}$$

where μ_A and μ_W are the acetone and water viscosities, respectively, and $x_{A,i}$ is acetone molar fraction in the environment i calculated as follows:

$$x_{A,i} = 1 - \left[\frac{\frac{(1-\xi_{v,i})}{V_W}}{\frac{\xi_{v,i}}{V_A} + \frac{(1-\xi_{v,i})}{V_W}}\right], \tag{11.33}$$

where $\xi_{v,i}$ is the acetone volume fraction, V_A and V_W are respectively the molar volume of acetone and water.

The Flory parameters, k and ν, are obtained from MD simulations performed in a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014) and depend on the acetone molar fraction as follows:

$$k(x_{A,i}) = 0.0064\exp\left(-3.15x_{A,i}\right), \tag{11.34}$$

$$\nu(x_{A,i}) = 0.30 + 0.45x_{A,i} - 0.15x_{A,i}^2. \tag{11.35}$$

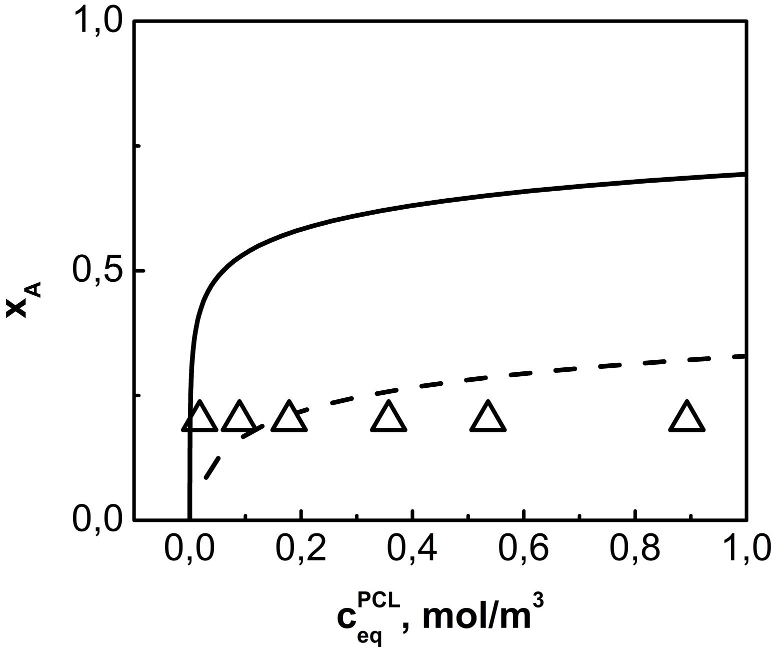


Figure 11.3: Solubility curve for PCL-14000 in acetone-water mixtures, reported as acetone molar fraction, x_A, versus the equilibrium concentration, $c_{eq}^{PCL}(x_A)$ (continuous line). Points above the curve correspond to stable mixture where PCL molecules do not self assemble, otherwise for the points below. The symbols (empty triangles) identify the polymer concentrations at the outlet, c_{out}^{PCL}, for equal acetone and water flow rates, investigated in this work. The dashed curve corresponds to the supersaturation value equal to 200. The reason for choosing this latter value will be explained later on.

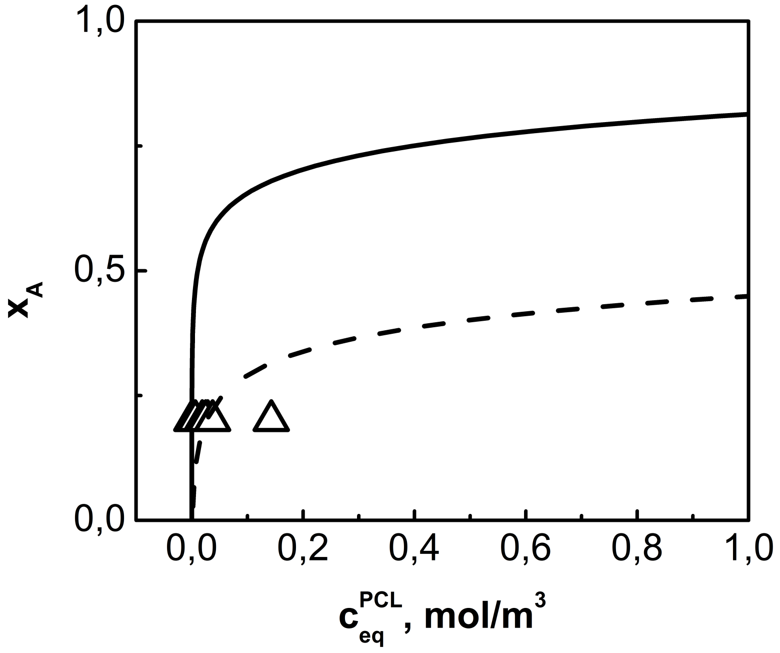


Figure 11.4: Solubility curve for PCL-80000 in acetone-water mixtures, reported as acetone molar fraction, x_A, versus the equilibrium concentration, $c_{eq}^{PCL}(x_A)$. Points above the curve correspond to stable mixture where PCL molecules do not self assemble, otherwise for the points below. The symbols (empty triangles) identify the six initial polymer concentrations, c_{out}^{PCL}, for equal acetone and water flow rates, investigated in this work. The dashed curve corresponds to the supersaturation value equal to 200. The reason for choosing this latter value will be explained later on.

Physical variable	**Value**
Temperature, K	303
Pressure, Pa	101325
Water density, kg m^{-3}	994
Acetone density, kg m^{-3}	780
Water viscosity, kg $m^{-1}s^{-1}$	8.0×10^{-4}
Acetone viscosity (Kestin et al., 1978), kg $m^{-1}s^{-1}$	3.1×10^{-4}

Table 11.2: Involved physical variables and relative values adopted in the simulations.

It is interesting to notice that Eq. (11.35) returns for the exponent ν the theoretical values of 1/3 and 3/5 for bad (i.e. pure water) and good solvents (i.e. pure acetone), respectively. The values assumed by all the other relevant variables is reported in Tab. 11.2.

As mentioned simulations were performed in a CIJM, which consists of two small (millimeter size) jets impinging at the center of a cylindrical chamber. The two jets, being confined in a small volume, after mixing exit from the bottom. Both top and bottom have conical ends. More details concerning this particular mixer and its mixing features can be found in the literature (Icardi, Gavi, Marchisio, Barresi, et al., 2011; Icardi, Gavi, Marchisio, Olsen, et al., 2011; Lince et al., 2011, 2009). By making use of its symmetry, one can simulate half of the geometry, as depicted in Figure 11.5.

Ansys Fluent 15 was used in this work. The computational grid consists of about 120.000 uniform hexahedral cells (for half of the actual geometry), with a refinement near the impinging plane and in the region around the two inlet flows. The pressure-velocity coupling has been done by using the SIMPLE algorithm, whereas the employed numerical scheme for spatial dicretization is the first-order upwind. In order to efficiently study turbulence inside the mixer the standard $k - \varepsilon$ model is employed with enhanced wall treatment near the wall, as this can be considered an adequate approach for this system (Gavi et al., 2007). Outlet boundary conditions are set to zero normal gradients for all flow variables, except for pressure; in this way, outflow boundary values are not imposed but are calculated from the interior.

The DQMOM-IEM and QMOM model were implemented via user-defined functions (UDF) and scalars (UDS) in Fluent 15.0.

The boundary conditions referred to the latter are reported in Tab. 11.3; as it is seen, in the pure acetone flow, inlet 1 has a unitary weight (i.e. $p_1 = 1$), because the whole fluid is constituted by acetone only. Similarly, being the pure water inlet consisting of pure water, the weight p_1 is necessarly zero. For the same reason, it is easy to understand the reported values referring to $p_1\xi_1$ and $p_2\xi_2$: in the first inlet both p_1 and ξ_1 are unitary (pure acetone), whereas in second inlet p_2 is equal to one, but ξ_2 is null because the nodes always refer to acetone (pure acetone).

As far as the boundary conditions for the j-th order moments, $m^{(j)}$, is concerned they represents the initial PCL concentration (expressed in mol m^{-3}), as the PCL enter

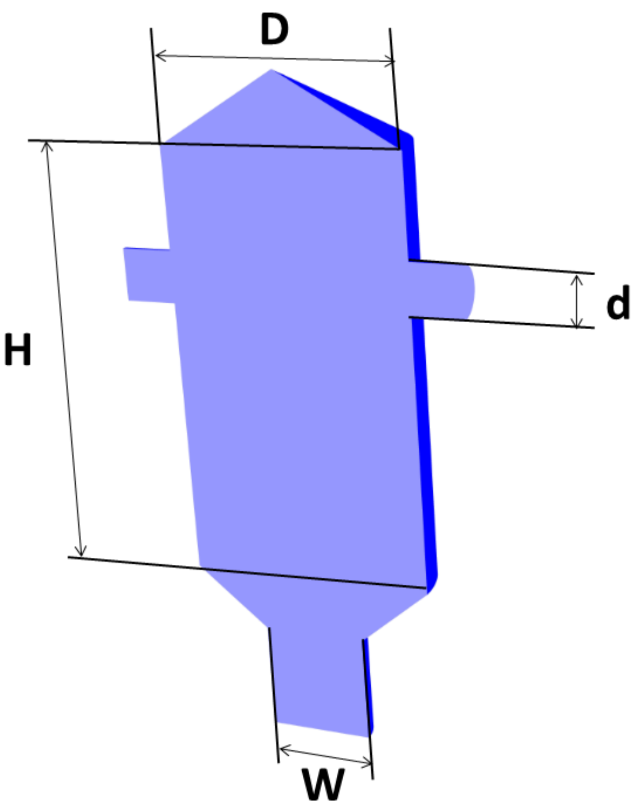


Figure 11.5: Sketch of geometry adopted for the CFD simulation of the CIJM (corresponding to half of the actual CIJM geometry cut across the symmetry plane into two identical parts). The characteristic geometric parameters assume the following values: $d = 1$ mm, $D = 4.76$ mm, $H = 9.52$ mm and $W = 2$ m.

in the CIJM in a stable molecular solution, corresponding to a CMD centered on n=1 (i.e. clusters made up of only one PCL molecule). For this reason, this value is the same for the all four moments. It is noteworthy to highlight that the moments boundary conditions are conveniently corrected taking into account for the Favre-averaged density of the mixture and acetone density, since also the moments are Favre-averaged scalar values, obeying the following equation:

$$m^{(j)} = \frac{c_{in}^{PCL}}{M_w \rho_A} \bar{\rho}, \tag{11.36}$$

where c_{in}^{PCL} represents the initial PCL concentration expressed in mg mL^{-1} and reported in Tab. 11.1, M_w, ρ_A and $\bar{\rho}$ are respectively PCL molecular weight, acetone mixture and Favre-averaged mixture density. Since the PCL is dissolved in the acetone inlet stream, the boundary conditions related to the moments in the water inlet are equal to zero.

Variable	**Boundary conditions**	
	Pure acetone inlet	Pure water inlet
p_1	1	0
ξ_1	1	1
ξ_2	0	0
$p_1\xi_1$	1	0
$p_2\xi_2$	0	0
$m^{(j)}, j = 0{,}1{,}2{,}3$	Eq. (11.36)	0

Table 11.3: Boundary conditions related to additional scalars employed in the CFD simulations.

11.4 Results and discussion

In this section the main results are reported and discussed. Figure 11.6 reports contour plots of the velocity magnitude in the symmetry plane of the CIJM for increasing flow rates of the two inlet streams of water and acetone. As it can be seen, the two inlet jets collide in an impinging plane, confined in the central part of the mixing chamber. This impinging plane is where most of the kinetic energy is produced and dissipated and is where mixing takes place. It is also interesting to notice that the impinging plane is slightly shifted towards the acetone inlet due to the difference of density between acetone and water. Closer observation of Figure 11.6 reveals that the flow is laminar in the two inlets, as evident from the parabolic velocity profile, and is instead turbulent in other regions. It is also interesting to observe that some small stagnant zones are generated near the walls. This leads to a higher residence time of the fluid in these regions that, as we will see, influences directly the final PCL molecule self-assembly.

The contour plots, on the symmetry plane of the CIJM, for the turbulent kinetic energy, k, and the turbulence dissipation rate, ε, at different inlet flow rates, are instead reported in Figure 11.7 and 11.8. As it can be seen, as the low flow rate values are increased (from 10 to 120 mL min^{-1}) the region where most of the turbulent kinetic energy is produced and dissipated becomes more and more confined. Moreover, by increasing the inlet flow rates the characteristic values of both k and ε increase, indicating faster and faster mixing.

The turbulence dissipation rate is particular important as it determines the rate with which mixing at the molecular level, or micromixing, occurs. In other words the turbulent dissipation rate dictates the rate with which acetone (and PCL) and water molecules enter into contact, dictating therefore also the rate with which the self-assembly process occurs. The turbulent dissipation rate is also important because it defines the rate of cluster aggregation due to turbulent fluctuations, as clearly evidenced by Eq. (11.6). In general, the effect of increasing ε is twofold: on one hand it improves micromixing, leading to the formation of smaller molecular clusters, but on the other hand it increments cluster aggregation, leading to their subsequent aggregation. Its final effect of the final size of the aggregates is therefore difficult to predict a priori.

Figure 11.9 reports the contour plots on the symmetry plane of the CIJM of the moments of the CMD ranging from order zero to three at different flow rates. Results are reported here only for PCL-14000 and for an initial PCL concentration in the acetone stream of 0.5 mg mL^{-1}, as for the other molecular weight and for the other PCL initial concentrations very similar trends are observed.

As already mentioned, $m^{(0)}$ represents the total cluster number density (see top row of Figure 11.9). When PCL molecules enter the CIJM through the acetone inlet, they are in a stable molecular solution, and therefore each PCL molecule is an independent cluster characterized by a cluster dimensionless mass, or aggregation number, of one (i.e. $n = 1$). As the acetone stream mixes with the water stream it is diluted, and therefore the total cluster number density decreases. Moreover, by mixing the acetone stream with water, the PCL equilibrium concentration is overcome, supersaturation is built up, and aggregation triggered, as depicted in Figure 11.3 and 11.4. As soon as aggregation takes place, the original molecular clusters entering the CIJM, with the acetone stream, and containing one single PCL molecule ($n = 1$), self-assemble forming larger clusters, increasing n and further reducing the total cluster number density. In fact, if four molecules self-assemble into one single cluster, the total cluster number density is reduced from four to one.

The second row of Figure 11.9 reports instead $m^{(1)}$ that represents the total molecule number density. This quantity is invariant during aggregation. In fact, if four PCL molecules self-assemble into a single cluster, $m^{(1)}$ does not change, as the total molecule number density quantifies the total number of molecules distributed over the different clusters. The behavior observed in the contour plots of Figure 11.9 is therefore solely

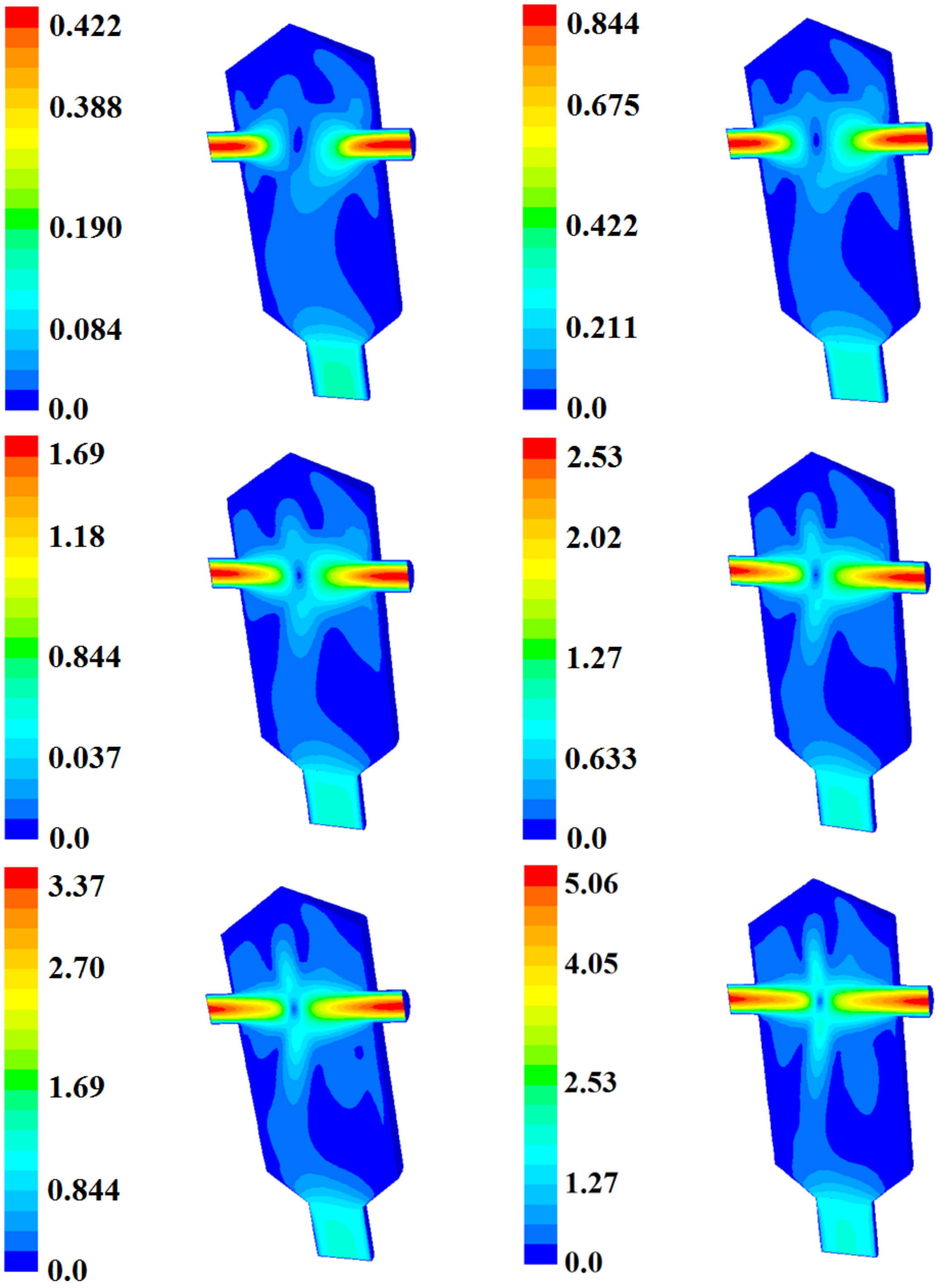


Figure 11.6: Contour plots for the magnitude of the velocity vector (m s^{-1}) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}.

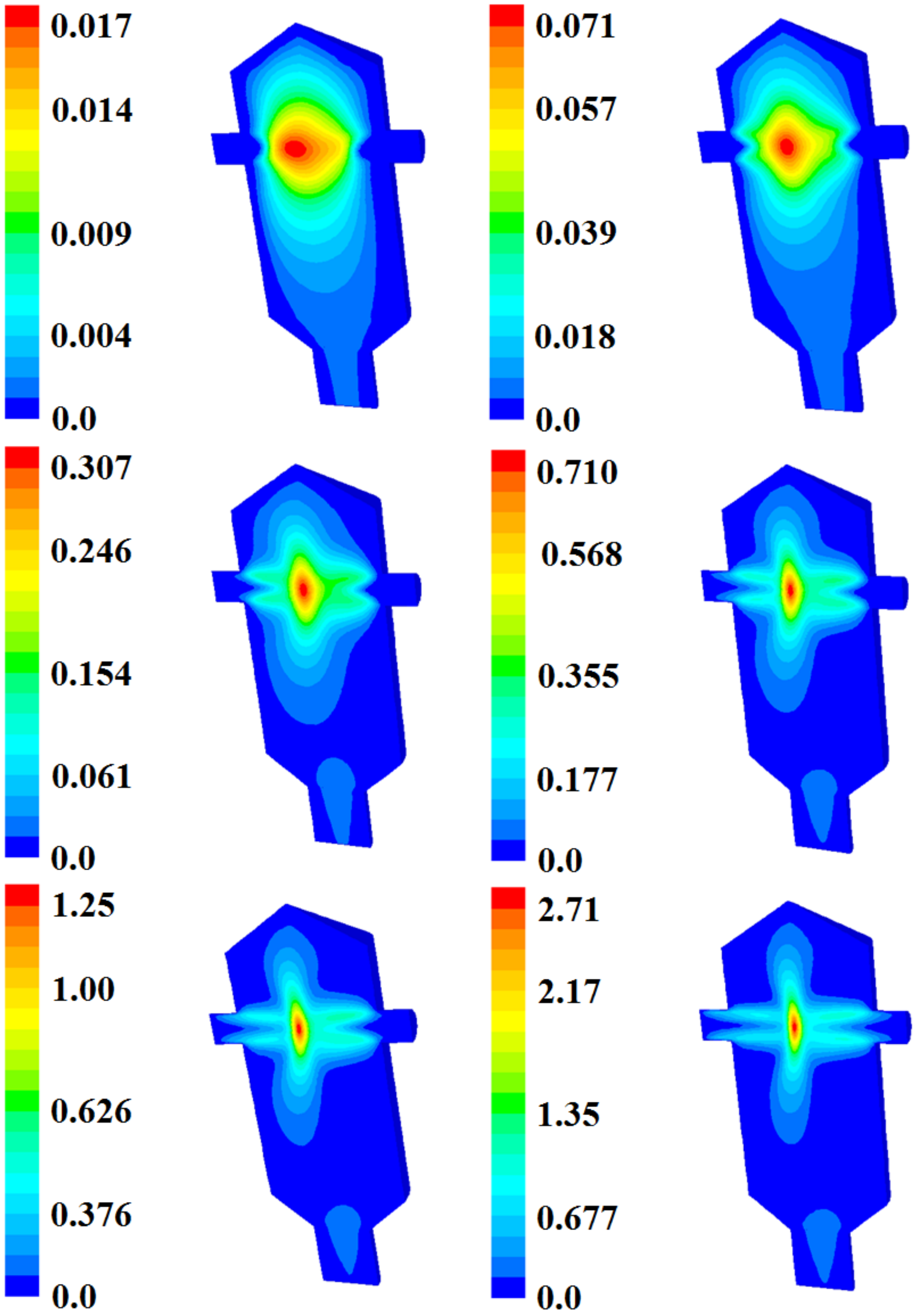


Figure 11.7: Contour plots for the turbulent kinetic energy, k ($m^2\ s^{-2}$) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}.

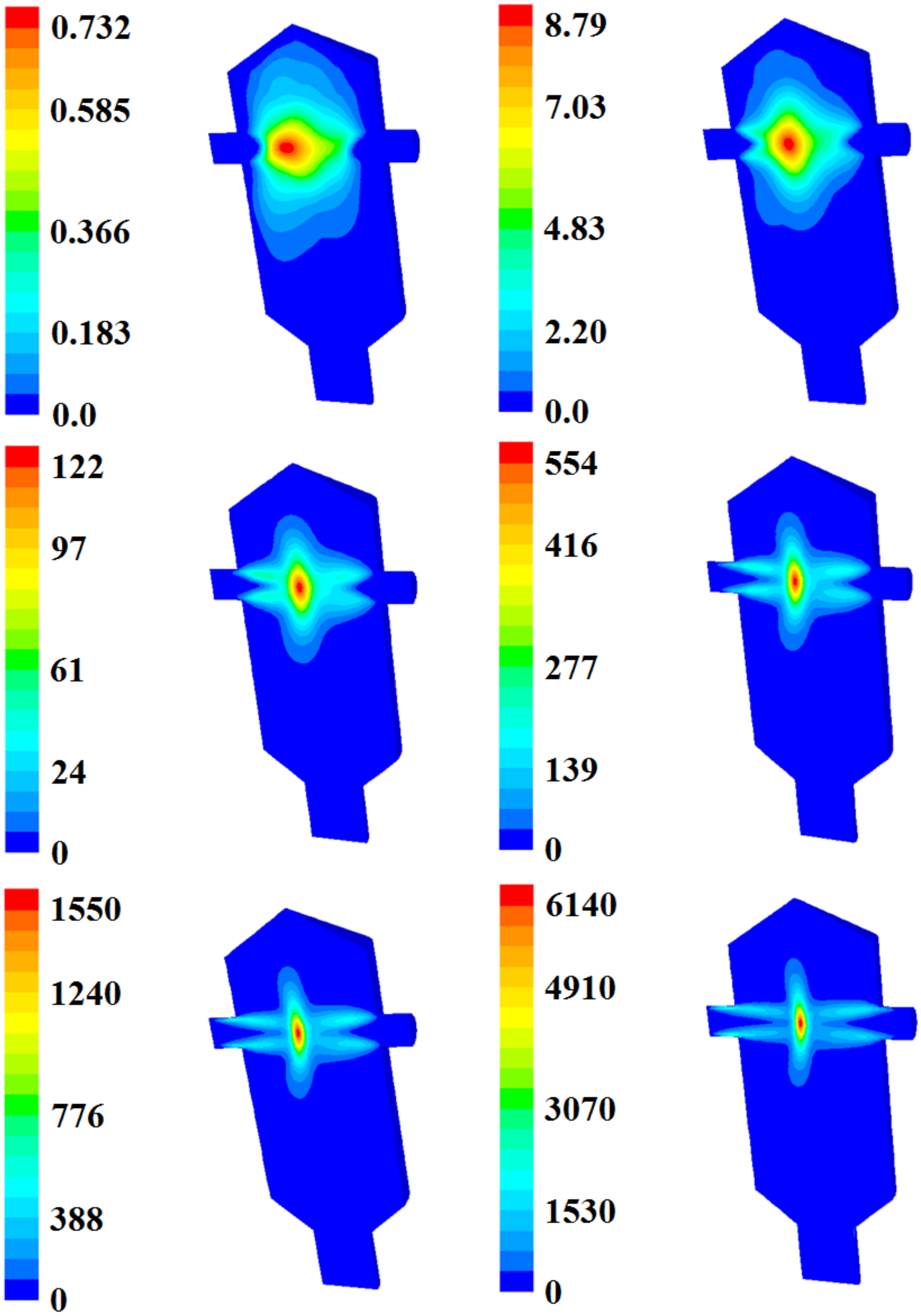


Figure 11.8: Contour plots for the turbulence dissipation rate, ε (m^2 s^{-3}) on the symmetry plane of the CIJM for, from left to right and from top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}.

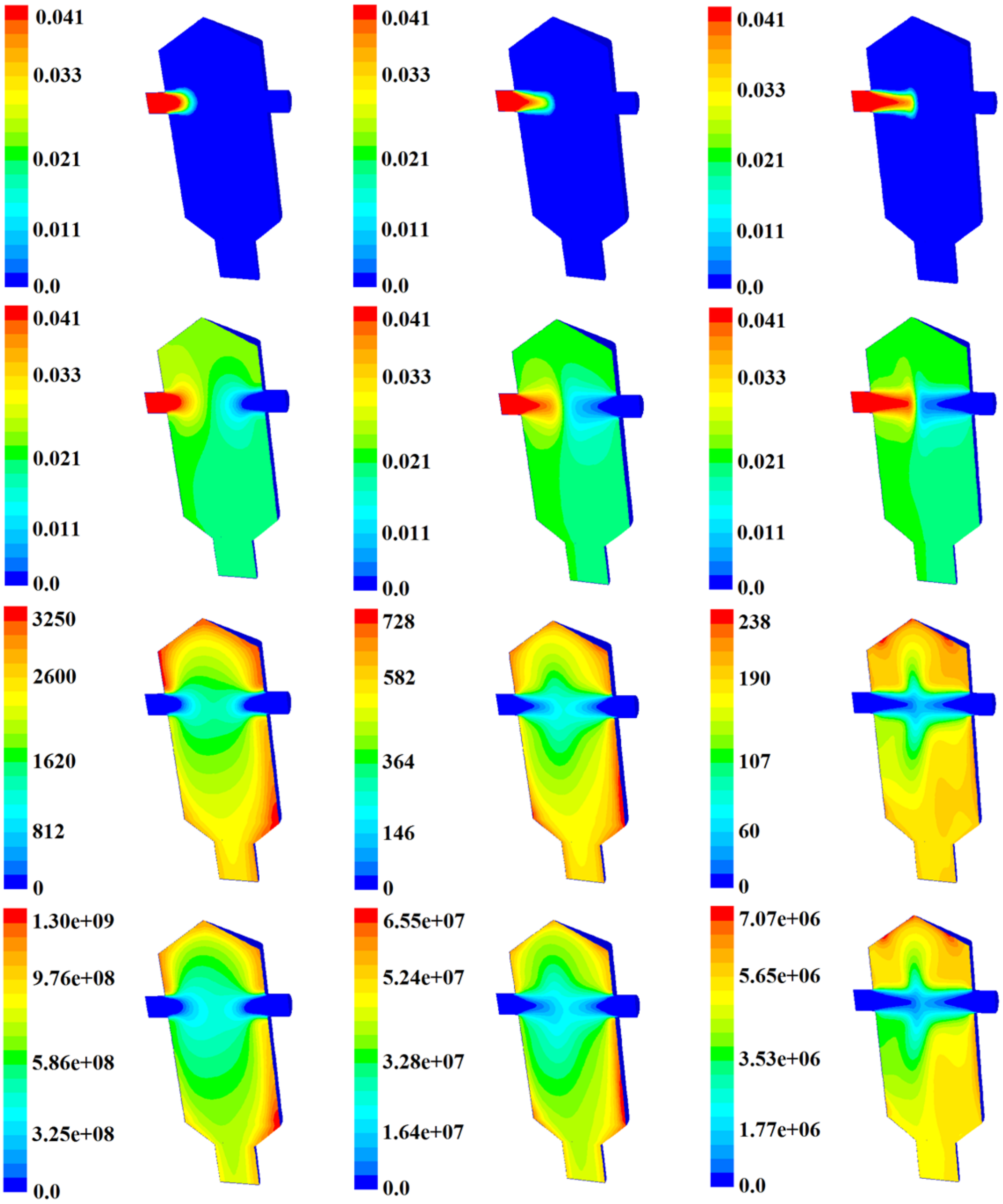


Figure 11.9: Contour plots for the moments of the CMD on the symmetry plane of the CIJM for an initial PCL-14000 concentration of 0.5 mg mL^{-1} and for, from left to right, inlet flow rates of 10, 40 and 120 mL min^{-1}; moments are reported according to their order (i.e. $j = 0, 1, 2, 3$) from top to bottom.

due to dilution.

The third and fourth rows of Figure 11.9 report finally $m^{(2)}$ and $m^{(3)}$ quantifying the variance and skewness of the CMD. The higher $m^{(2)}$ and $m^{(3)}$ are, the more spread and skewed the CMD is. It is interesting to observe the effect of the inlet flow rate on these two quantities. As seen by increasing the inlet flow rate both $m^{(2)}$ and $m^{(3)}$ are sensibly reduced, indicating that the corresponding CMD become narrower and less skewed.

As discussed in the previous sections each cluster, depending on the value of n and the composition of the surrounding solution (i.e. x_A), is characterized by a different radius of gyration. However, as anticipated in Eq. (11.13), a mean radius of gyration for the entire population of clusters can be calculated by making use of the quadrature approximations of QMOM and DQMOM-IEM. The contour plots of this quantity on the symmetry plane of the CIJM for different inlet flow rates is reported in Figure 11.10. As it can be noted, at the acetone inlet the mean radius of gyration is very small (i.e. a few nanometers) and corresponds to the size of a single PCL-14000 molecule in pure acetone. As soon as water and acetone mix, molecules self-assemble and the mean radius of gyration of the aggregates increases, reaching the final value at the outlet. It is interesting to observe that the maximum values for the radius of gyration are reached near the wall, where, as highlighted by the analysis of the flow field, stagnant zones are detected. In these stagnant zones the fluid, and the clusters, spend more time and have therefore more time to self-assemble. It is also interesting to observe that the mean radius of gyration sensibly decreases with the increase of the inlet flow rate. The reduction of the value of the mean radius of gyration is due to the fact that, by increasing the inlet flow rate, faster mixing between the two inlet streams is achieved, resulting in the formation of more clusters, then that have less PCL molecules available for subsequent growth via further aggregation.

By reconstructing the CMD as a lognormal distribution from the tracked four moments (Marchisio and Fox, 2013) at the CIJM outlet, the results reported in Figure 11.11 are obtained. The data are reported here for different inlet flow rates, but one single PCL-14000 concentration (i.e. 0.5 mg mL^{-1}), and are plotted together with a line indicating the mean cluster mass at the reactor outlet. As already discussed, by increasing the inlet flow rates the CMD becomes narrower (i.e. the variance decreases), less skewed, whereas the mean cluster mass (reported in the figure as red lines and calculated as the ratio of the moments of order one and zero) decreases. This confirms once again that manipulating the inlet flow rate allows to control the CMD. In particular, an increase in the inlet flow rates results in the formation of smaller clusters with narrower CMD.

The mean radius of gyration at the reactor outlet can now be compared with experimental data collected in another work (Lince et al., 2008). In fact, the mean cluster size measured experimentally for clusters collected at the outlet of CIJM with Dynamic Light Scattering (DLS) corresponds to the mean radius of gyration calculated with Eq. (11.13). It is noteworthy to stress that the DLS measures the mean hydrodynamic radius, which is by definition different from the mean radius of gyration

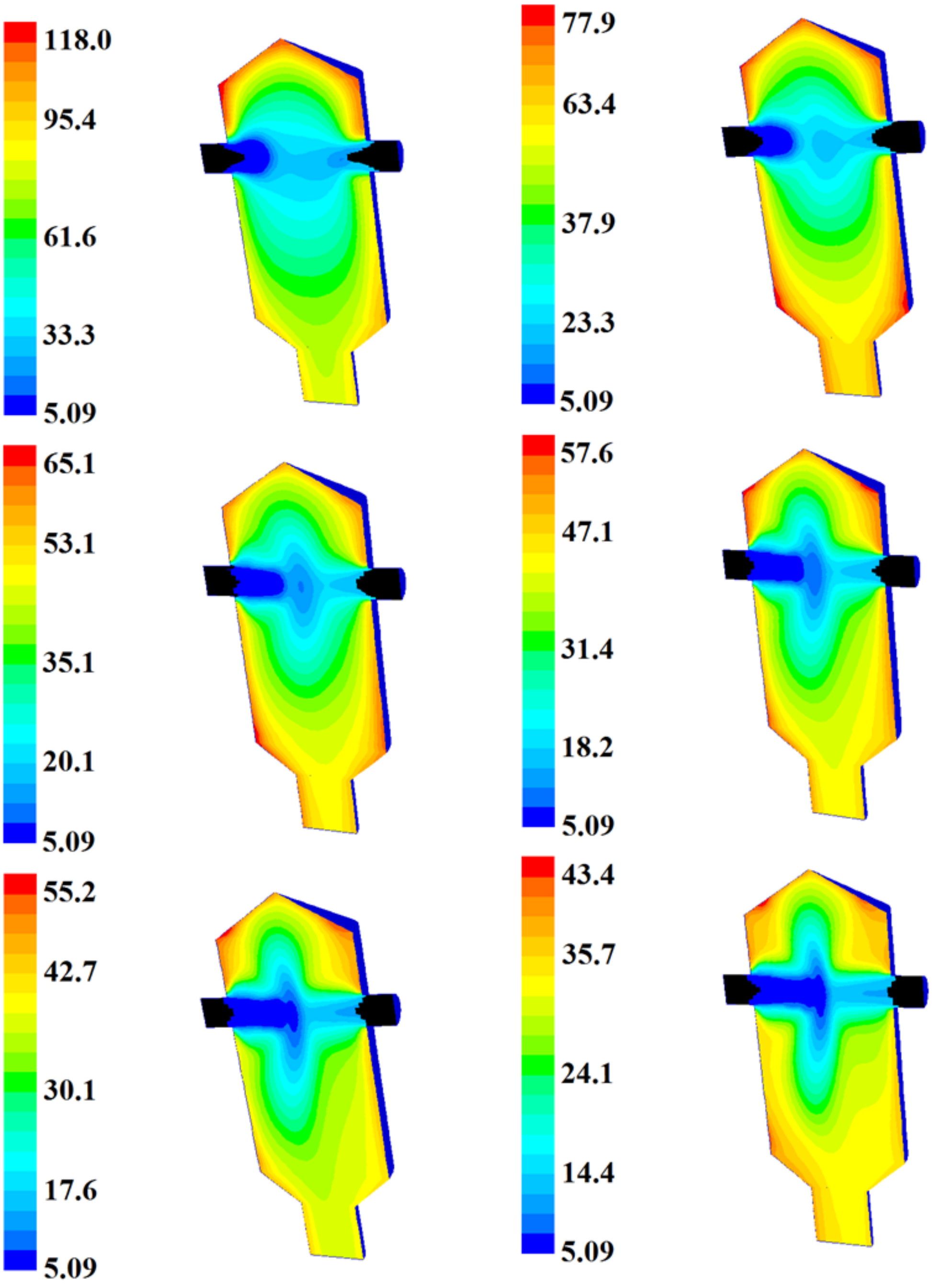


Figure 11.10: Contour plots for the mean radius of gyration (nm) on the symmetry plane of the CIJM for PCL-14000 with an initial concentration of 0.5 mg mL^{-1} and for, from left to right and top to bottom, flow rate values of 10, 20, 40, 60, 80 and 120 mL min^{-1}.

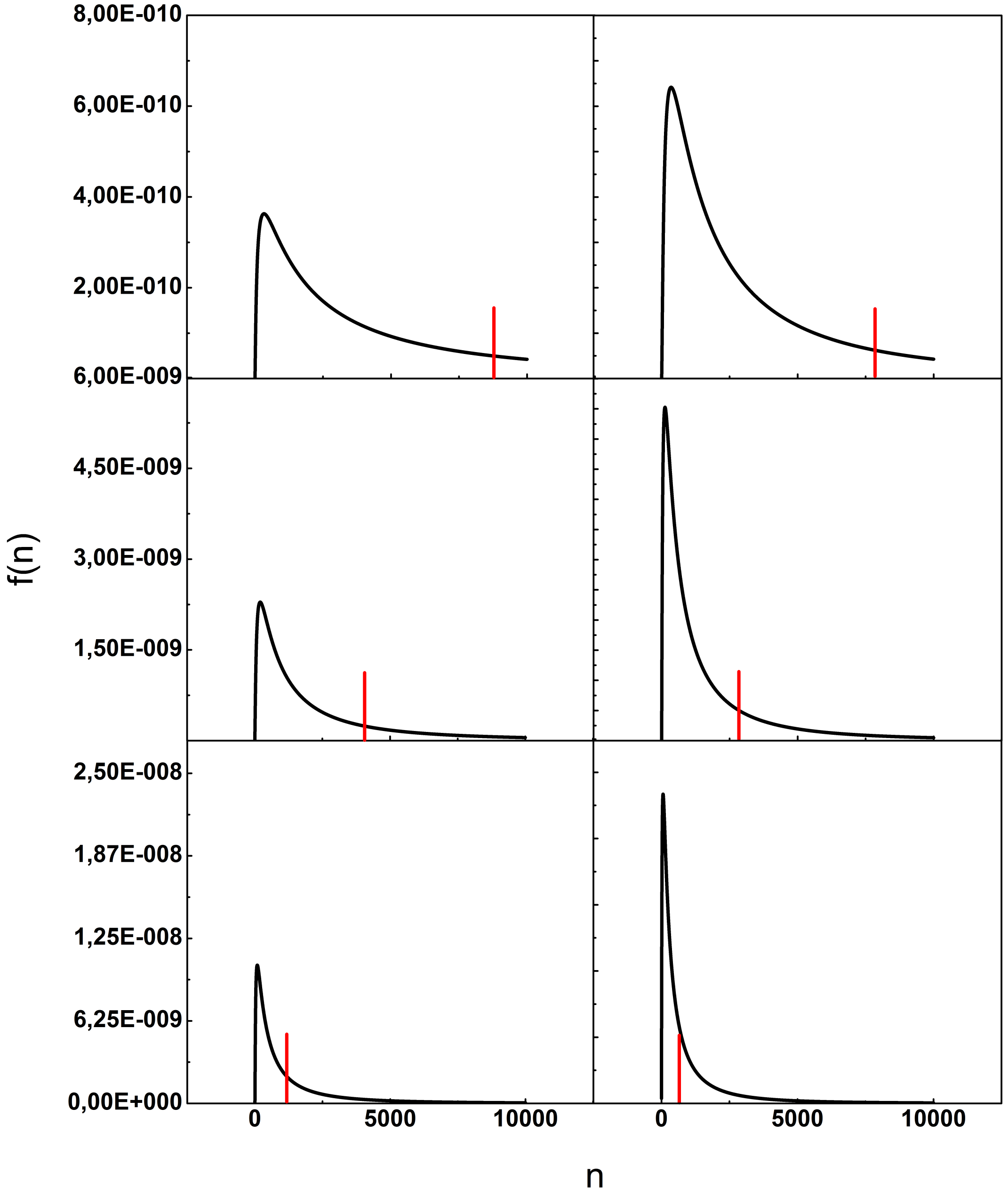


Figure 11.11: Reconstructed CMD at the CIJM outlet, $f(n)$, as a function of the number of molecules per cluster, n, at PCL-14000 concentration of 0.5 mg mL^{-1} and for, from left to right and top to bottom, inlet flow rates of 10, 20, 40, 60, 80 and 120 mL min^{-1}. The red vertical line indicates the mean value of n of the distribution, namely $m^{(1)}/m^{(0)}$.

predicted by simulations. For spherical particles the hydrodynamic to gyration radius ratio is close to unity and, being the polymer NP experimentally proven to have a spherical shape, this turns out to be a reasonable approximation.

The comparison for the different PCL-14000 initial concentrations and for the different inlet flow rates investigated in this work is reported in Figure 11.12. In general acceptable agreement between experiments and simulations (with Brownian and turbulent aggregation) is observed at all initial PCL-14000 concentrations, but the agreement improves as the concentration increases and becomes very good for concentrations equal to or higher than 2.5 mg mL^{-1}). This corresponds to supersaturation ratio ($S(\xi)$) greater than about 200.

For these simulations predictions are evaluated by considering the effect of Brownian aggregation only and of both Brownian and turbulent aggregation. The comparison of predictions in the two cases reveals that at low initial PCL-14000 concentrations the effect of turbulent motions is negligible on cluster aggregation. This is probably due to the fact that at low initial PCL-14000 concentrations the cluster size obtained from Brownian aggregation is too small to have some effect on turbulent fluctuations, as the turbulent aggregation kernel strongly depends on the cluster size. Under the operating conditions investigated in this work, the local turbulence intensities are such that turbulent aggregation is relevant only for cluster larger than 200 nm.

As it can be seen, generally both experimental data and model predictions indicate a decrease of the mean radius of gyration with the inlet flow rate, as already observed and explained. The only exception is at the highest initial polymer concentration, where after a first decrease the mean radius of gyration increases at large inlet flow rates. This behavior, observed also in the experiments, is due again to the strong dependency of the aggregation kernel on the cluster size. Only at this high initial polymer concentration the clusters become so large to be extremely sensitive to turbulent fluctuations. At this initial polymer concentration a first increase of the inlet flow rate results in the formation of more clusters, that then tend to growth smaller, but a further increase of the flow rate, results in an increase of the turbulent dissipation rate and of the resulting turbulent aggregation rate.

The same trend is observed in Figure 11.13 where model predictions are compared with experiments for PCL-80000. In this case model predictions due to both Brownian and turbulent aggregation are reported. The similarity of the trends confirms the generality of the model that can be therefore applied for both molecular weights. Also in this case predictions at low concentrations result in poorer agreement, when compared with the agreement obtained at high concentrations. In this case, the agreement starts being acceptable for an initial PCL concentration of 6.0 mg mL^{-1}, corresponding to an supersaturation ratio larger than 200.

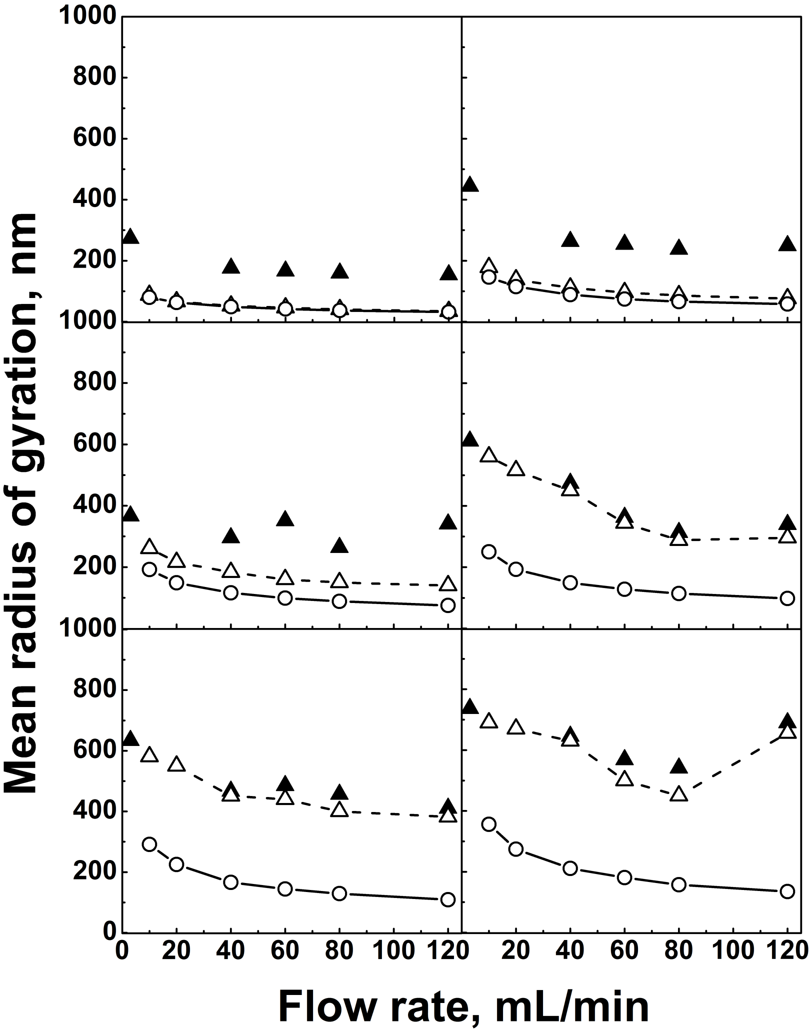


Figure 11.12: Mean radius of gyration of the cluster exiting the CIJM versus the inlet flow rate as measured in experiments (black triangles) and as predicted by the model with only Brownian aggregation (continuous line) and Brownian and turbulent aggregation (dashed line) for PCL-14000 initial concentrations of, from left to right and top to bottom, of 0.5, 2.5, 5.0, 10.0, 15.0 and 25.0 mg mL^{-1}.

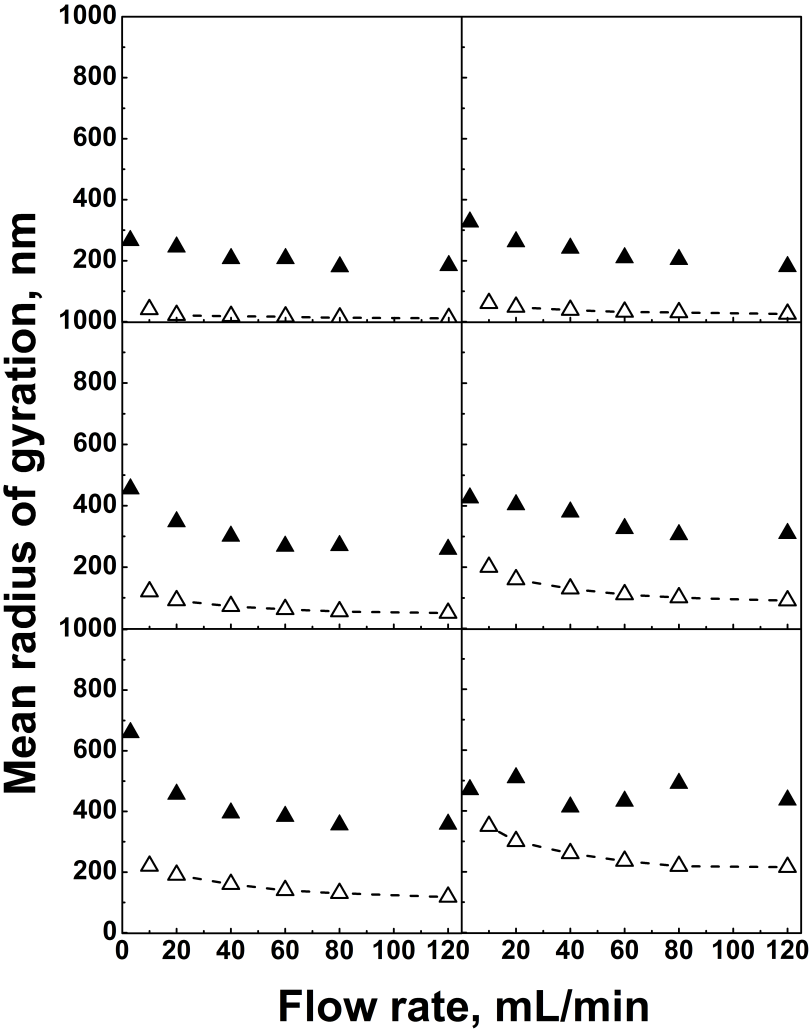


Figure 11.13: Mean radius of gyration of the clusters exiting the CIJM versus the inlet flow rate as measured in experiments (black triangles) and as predicted by the model with Brownian and turbulent aggregation (dashed line) for PCL-80000 initial concentrations, from left to right and top to bottom, of 0.02, 0.2, 1.0, 3.0, 4.0 and 6.0 *mg*/*mL*.

Results seems to confirm that there exists a concentration range under which molecular clusters formation is controlled by aggregation and can be described as a purely-aggregative process. This happens for initial superaturation greater than 200. Then, $S(\xi)$ over 200 allows to describe the process as a spinodal decomposition. On the other hand, for supersaturation ratio smaller than 200 the purely-aggregative model does not result in good agreement, proving that a nucleative mechanism controls the process. Therefore, it is reasonable to infer that the spinodal line (i.e. the border line between nucleation and unstable zone) could be placed at this value of $S(\xi)$, as shown in Figure 11.3 and 11.4.

11.4.1 Effect of different good solvents on FNP

In this section the effect of different good solvents in FNP is investigated, by means of the QMOM and DQMOM-IEM model presented in section 11.2.2. Experiments show, in fact, that CMD and mean NP size at the outlet of the mixer are affected by the nature of the good solvent employed (Ferri et al., 2017). Here, this effect is investigated at the clusters and vessel scales, with the aim to suitably adjust the Flory's law, developing a transferable PBM-CFD model when different good solvents are used. A computational fitting together with a theoretical approach is employed, in order to identify the key parameters that lead to different mean NP size at the outlet of the mixer at the same operating process conditions. Three different good solvents are compared: acetone, acetonitrile and tetrahydrofuran (THF).

Being the acetone investigated throughout this thesis, also at smaller length scales (atomistic scale), and having already all the results and functional forms related to the Flory's law presented in the previous sections and chapters, it is strategic to choose acetone as reference solvent in this analysis.

In the modeling framework extensively exploited in this chapter, three main physical properties must be taken into account when the good solvent is changed: density, viscosity and molar volume. These physical properties referred to acetone, acetonitrile and THF are schematically reported in Table 11.4.

Good solvent	**Density, kg m^{-3}**	**Viscosity, Pa·s**	**Molar volume, cm^{3}mol^{-1}**
Acetone	780.85	$3.10 \cdot 10^{-4}$	74.38
Acetonitrile	771.45	$3.26 \cdot 10^{-4}$	53.21
THF	874.78	$4.34 \cdot 10^{-4}$	82.43

Table 11.4: Main physical properties related to the three good solvents investigated here: density, viscosity and molar volume.

Looking at Table 11.4, it turns out that acetone and acetonitrile have very similar values of density and viscosity, but acetonitrile has a smaller molar volume; on the contrary, THF exhibits higher density, viscosity and molar volume with respect to acetone.

It is worthwhile mentioning now that molar volume will be the most important physical parameter, affecting the final NP size, since it is related to the good solvent molar fraction, x_s, by means of Eq. (11.33) and recalling that x_s represents the independent variable of Flory's parameters functional forms.

The theoretical approach used in this analysis consists in correlating the NP size, in terms of radius of gyration, to the Flory-Huggins parameter χ that states the interaction between two substances (a solvent and a solute, for instance), in turn, strictly dependent on the solubility parameters of the components of the system.

The Hildebrand solubility parameter (Hildebrand and Scott, 1950) of a solute (e.g., polymer) reads as follows:

$$\delta = \left(\frac{\Delta H_v - RT}{V_m} \right), \tag{11.37}$$

where ΔH_v is the vaporization enthalpy and the whole numerator corresponds to the cohesion energy per unit volume; R is the gas constant, T is the absolute temperature and V_m is the molar volume of the solute.

Hansen (2007) proposed a decomposition of the Hildebrand parameter into three different contributions, leading to the so-called Hansen solubility parameters (HSP):

$$\delta = \left(\delta_D^2 + \delta_P^2 + \delta_H^2 \right)^{1/2}, \tag{11.38}$$

where δ_D considers the dispersion attractive forces (non-polar), δ_P accounts for the permanent dipole-permanent dipole interactions and δ_H for the hydrogen bonds. The advantage of the HSP approach lies on the fact that polar effects are explicitly considered, differently from the Hildebrand parameter that is more suitable for apolar systems. The ability of a given solvent to solubilize a solute (e.g., polymer) is expressed in terms of solubility 'distance' from the solute itself and takes the form of the radius of a sphere in Hansen solubility space $\{\delta_D, \delta_P, \delta_H\}$:

$$R_a = 4(\delta_{D,s} - \delta_{D,p})^2 + (\delta_{P,s} - \delta_{P,p})^2 + (\delta_{H,s} - \delta_{H,p})^2, \tag{11.39}$$

where R_a is the solubility 'distance', in terms of sphere radius in the Hansen space, and the subscripts s and p refer respectively to the solvent and the polymer. The term '4' is added in order to make the shape of this functional form more spherical and less elliptical. The whole derivation can be found in Hansen, 2007. The HSP and the corresponding distance from PCL is summarized in Table 11.5.

In several contexts, it turns out to be more practical to rescale the distance R_a in the following way:

$$RA = (\delta_{D,s} - \delta_{D,p})^2 + (\delta_{P,s} - \delta_{P,p})^2/4 + (\delta_{H,s} - \delta_{H,p})^2/4, \tag{11.40}$$

where the simply scaling relation holds: $RA = R_a/2$.

The Flory-Huggins parameter χ_{sp} can be expressed in function of the HSP (Hansen, 2007) through:

	δ_D, MPa$^{1/2}$	δ_P, MPa$^{1/2}$	δ_H, MPa$^{1/2}$	R_a, MPa$^{1/2}$
PCL	17.0	4.8	8.3	-
THF	16.8	5.7	8.0	1.0
Acetone	15.5	10.4	7.0	6.5
Acetonitrile	15.3	18.0	6.1	13.8
Water	15.6	16.0	42.3	35.9

Table 11.5: HSP and distance from PCL (Eq. 11.39) for all the components of the investigated systems.

$$\chi_{sp} = \frac{(RA)^2}{2R_m^2}, \tag{11.41}$$

where R_m represents the radius of the solubility sphere, namely the maximum solubility distance (in Hansen solubility space) that allows a solvent to dissolve the solute. Solvents characterised by $RA > R_m$ are classified as non-solvents (or poor solvents) with respect to that specific solute. On the contrary, all the solvents that belong to the Hansen solubility sphere (RA/R_m smaller than unity) are classified as good solvents for that solute. In this analysis, the solute is the PCL and the solvent is the 'good solvent'-water mixture. Being the second phase made by a binary mixture, all the solubility parameters (e.g., HSP) and physical properties (e.g., molar volume) involved in the following calculations are weighted on the good-bad solvents volume and molar fractions, and it will be generically labelled as 'solvent'.

The parameter R_m is able to consider the dependence of the solubility parameters on polymer molecular weight, thanks to the following expression:

$$R_m^2 = 0.5(1 + 1/r^{1/2})RT/V_m, \tag{11.42}$$

where r is the degree of polymerization and it is calculated in this approach as the PCL macromolecule molecular weight $M_w = 14000$ g mol^{-1} to the PCL monomer molecular weight $M_o = 114$ g mol^{-1} ratio.

With the assumption of considering the polymer structure at the mean good solvent molar fraction $\bar{x}_s$, after mixing occurs, in the range [0.15,0.25] for all the three good solvents considered, it is reasonable to infer that in the three-component phase diagram (water, good solvent and PCL) the system is in the condition of poor solvent, for which the radius of gyration depends on $\chi_{s,p}$ by means of the following equation (Rubinstein and Colby, 2003):

$$R_g \approx \frac{b^2}{|v|^{1/3}} N^{1/3} = \frac{bN^{1/3}}{(2\chi_{s,p} - 1)^{1/3}}, \tag{11.43}$$

where b is the Kuhn length, v the excluded volume and N is the number of monomers that form the polymer chain.

By using the relation reported in Eq. (11.43) for two different solvents $s1$ and $s2$ (e.g., $s1$ = acetonitrile-water and $s2$ = acetone-water), and considering that the Kuhn length b is a solute property, therefore the same for any solvent used, the following ratio can be readily obtained:

$$\frac{\langle R_g(n=1)\rangle_{s1}}{\langle R_g(n=1)\rangle_{s2}} = \frac{(2\chi_{s2,p}-1)^{1/3}}{(2\chi_{s1,p}-1)^{1/3}} = S_f, \tag{11.44}$$

where S_f is a scaling ratio, function of the solely Flory-Huggins parameters $\chi_{s1,p}$ and $\chi_{s2,p}$, calculated starting from the HSP through Eq. (11.41). The adopted strategy consists in setting $s2$ as the reference solvent, i.e., acetone-water in this case, for which the functional forms of Flory's parameters are known from MD. The following equality holds:

$$\sqrt{k_{s1}(x_{s1})M_w^{2\nu_{s1}(x_{s1})}} = S_f \cdot \sqrt{k_{s,ref}(x_{s,ref})M_w^{2\nu_{s,ref}(x_{s,ref})}}, \tag{11.45}$$

where, as stated above, the subscript ref refers to acetone-water mixture. Scaling factors S_f and Flory-Huggins parameters $\chi_{s,p}$ referred to the three solvent-polymer systems, namely acetone-water, acetonitrile-water and THF-water, evaluated at mean good solvent molar fraction $\bar{x}_s$, are reported in Table 11.6.

Solvent	**Flory-Huggins parameter, $\chi_{s,p}$**	**Scaling factor, S_f**
Acetone-water	0.923	1.00
Acetonitrile-water	1.003	0.94
THF-water	0.868	1.05

Table 11.6: Flory-Huggins parameter, $\chi_{s,p}$, for acetone, acetonitrile and THF as good solvents with the relative scaling factors obtained from Eq. (11.44), at mean good solvent molar fraction, $\bar{x}_s$.

Looking at Table 11.6, it is worth noticing that for all the systems the Flory-Huggins parameter is greater than 0.5 (θ-condition), confirming the initial assumption of poor solvent condition (the solvent is made by a good-solvent-to-water ratio in volume equal to unity).

At this point, an iterative procedure can be conducted, based on Eq. (11.45) and the results reported in Table 11.6. More specifically, Eq. (11.45) still presents two degrees of freedom, assuming to set the reference solvent as acetone-water mixtures for which everything is known. Therefore, by putting $\nu_{s1}(x_{s1}) = \nu_{s,ref}(x_{s,ref})$ as beginning value, $k_{s1}(x_{s1})$ is univocally determined and with this new set of Flory's parameters, say $k^*_{s1}(x_{s1})$ and $\nu^*_{s1}(x_{s1})$, simulations can be performed and an error against the experimental data will be detected. Based on the error with the experiments, the Flory's exponent can be suitably adjust to a new functional form, $\nu^{**}_{s1}(x_{s1})$ and, by means of Eq. (11.45), a new functional form for $k_{s1}^{**}(x_{s1})$ is obtained. With this new set of Flory's parameters simulations are performed until this iterative procedure will lead to the minimum error. This minimum error corresponds to the functional forms of the Flory's parameters reported in Table 11.7, thanks to which the agreement with experiments is very good and reported in Figure 11.14, as far as the THF-water system is concerned, and in Figure 11.15, as far as acetone-water and acetonitrile-water are concerned, in a range of PCL initial concentration in good solvent stream that spans from 3.0 to 9.0 mg mL^{-1}. The agreement with experiments is line with the results obtained in the previous section, therefore no further discussion is needed.

The adjustment of the functional forms of Flory's parameters deserves further explanations. Regarding the Flory's constant, only the proportionality constant of the exponential is adjusted with respect to the function referred to acetone, Eq. (11.34). As far as the Flory's exponent is concerned, it corresponds to a parabolic profile (Eq. 11.35); therefore, three conditions are needed: two of them are represented by the exponent value in pure good and pure bad solvent, that are respectively 3/5 and 1/3. The third condition is the only degree of freedom that the user needs to fulfill, and it might correspond to the exponent value at the mean good solvent molar fraction, $\bar{x}_s$, that can be suitably adjusted during the iterative procedure presented above.

	Acetone	**Acetonitrile**	**THF**
$k_s(x_s)$	$0.0064\exp(-3.15x_s)$	$0.0055\exp(-3.15x_s)$	$0.0047\exp(-3.15x_s)$
$\nu_s(x_s)$	$0.30+0.45x_s-0.15x_s^2$	$0.30+0.40x_s-0.10x_s^2$	$0.30+0.62x_s-0.32x_s^2$

Table 11.7: Flory's parameters functional forms for acetone, already reported in Eq. (11.34) and (11.35), acetonitrile and THF.

The Flory's exponent profiles are depicted in Figure 11.16. The inset shows the ν^* values corresponding to the different mean good solvent molar fractions $\bar{x}_s$ (through the discrete symbols) for the three good solvents compared in this analysis. The related value of ν^* justifies the fact PCL aggregates more in THF (blue line), then acetonitrile (green line) and finally in acetone, in line with experiments. In fact, at mean good solvent molar fraction $\bar{x}_s$ the inset shows the following relation: $\nu^*(\bar{x}_{THF}) > \nu^*(\bar{x}_{acetonitrile}) > \nu^*(\bar{x}_{acetone})$. Figure 11.16 also points out the affinity order found out in terms of HSP and solubility distance from PCL shown in Table 11.5, namely THF>acetone>acetonitrile for which the solubility distance from PCL is respectively 1.0, 6.5 and 13.8 MPa$^{1/2}$, as reported in Table 11.5. This means that, operating at constant good solvent molar fraction, the ν^* profiles, or the solubility affinity with PCL, implies that the single PCL macromolecule dimension follows the order THF>acetone>acetonitrile. As usually happens in industrial contexts, however, the operating conditions consist in keeping constant the good-solvent-to-water ratio in volume, which leads to a mean good solvent molar fraction in the order $\bar{x}_{acetonitrile} > \bar{x}_{acetone} > \bar{x}_{THF}$, due to the different molar volume of the good solvents, affecting the Flory's parameters profiles and, consequently, aggregation. The combination of molar volume and the aggregation kernels (dependent on Flory's parameters) lead to the aggregation order THF>acetonitrile>acetone. This is visible, by looking at Figure 11.17, in which the mean squared radius of gyration is reported for acetone (red), acetonitrile (green) and THF (blue). The crossover between the red and green profiles (acetone and acetonitrile) shows the two different tendency and contributions in case of single molecule ($\langle R_g \rangle (n = 1)$ higher in acetone) and at high aggregation number, n, in which $\langle R_g \rangle$ is higher in acetonitrile than in acetone. This means that, despite the single PCL macromolecule occupies a smaller volume, the PCL cluster is bigger in acetonitrile than in acetone, namely more macromolecules belong to the same cluster (or NP) in acetonitrile (m^1/m^0 in acetonitrile is greater than in acetone).

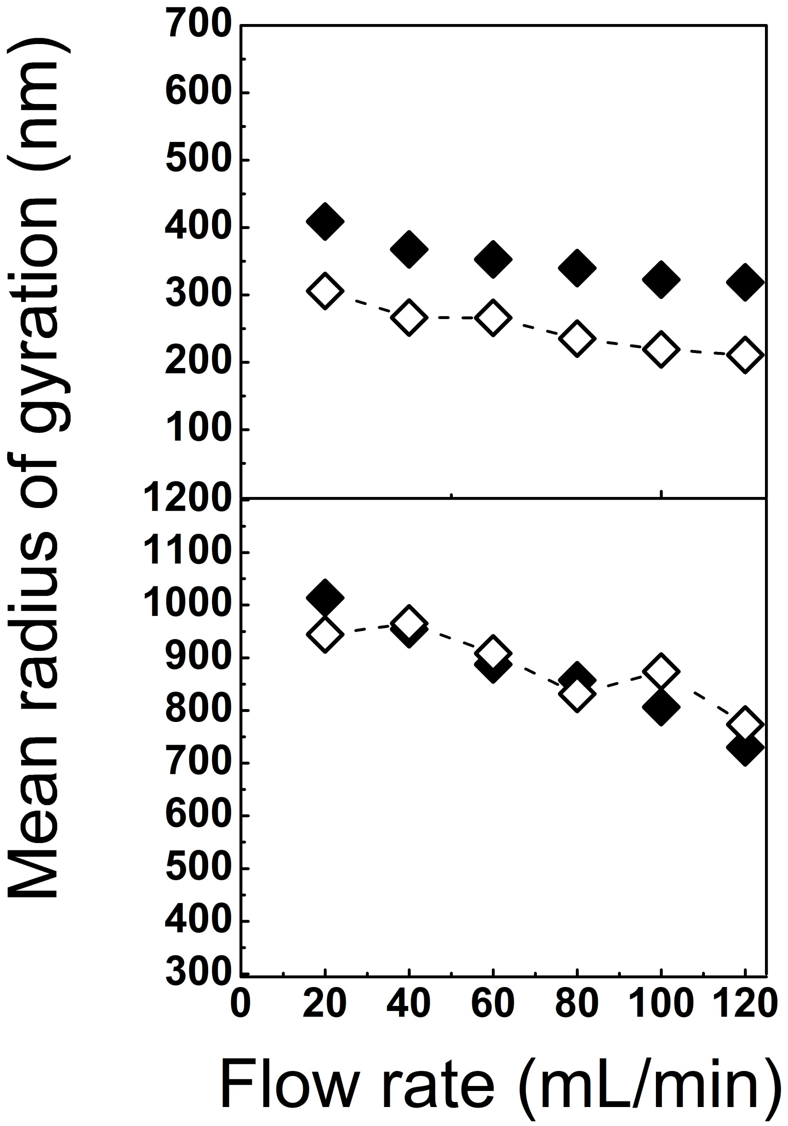


Figure 11.14: Mean radius of gyration of the NP exiting the CIJM versus the inlet flow rate as measured in experiments (black symbols) and as predicted by the purely-aggregative model (dashed line, empty symbols) for PCL-14000 initial concentrations equal to 3.0 (top panel) and 5.0 (bottom panel) mg/mL in THF as the good solvent.

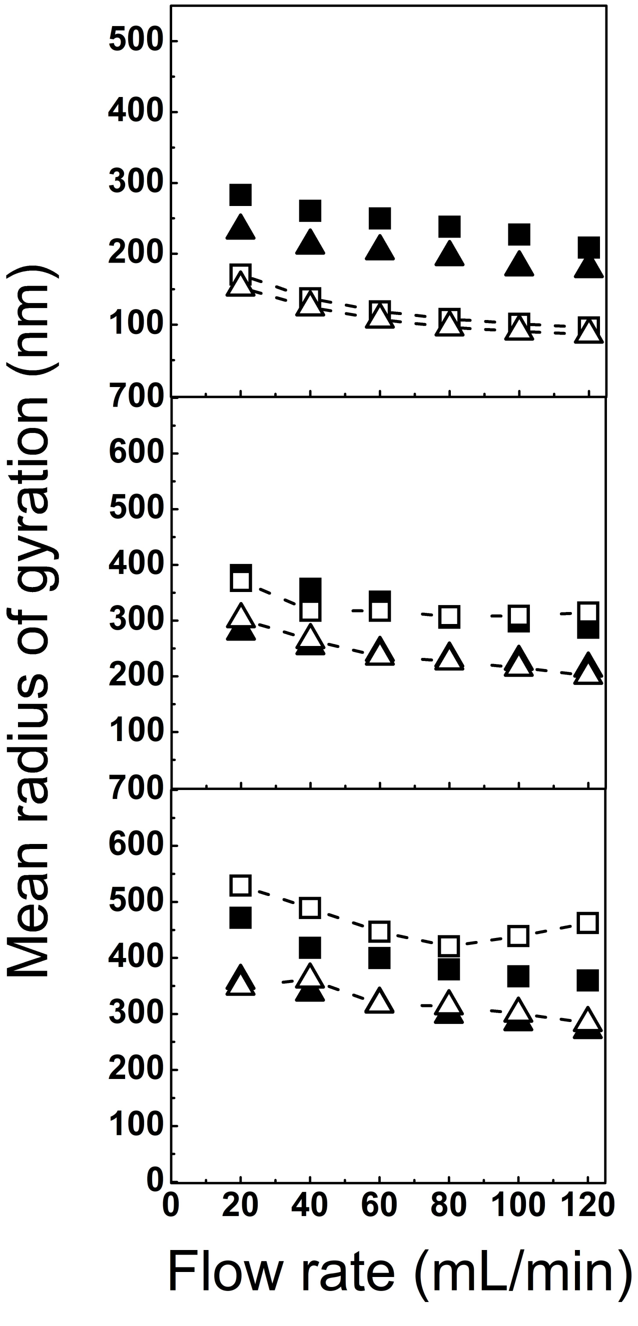


Figure 11.15: Mean radius of gyration of the NP exiting the CIJM versus the inlet flow rate as measured in experiments (black symbols) and as predicted by the purely-aggregative model (dashed line, empty symbols) for PCL-14000 initial concentrations equal to 3.0 (top), 6.0 (middle) and 9.0 (bottom) *mg/mL* in acetone (triangles) and acetonitrile (squares) as the good solvents.

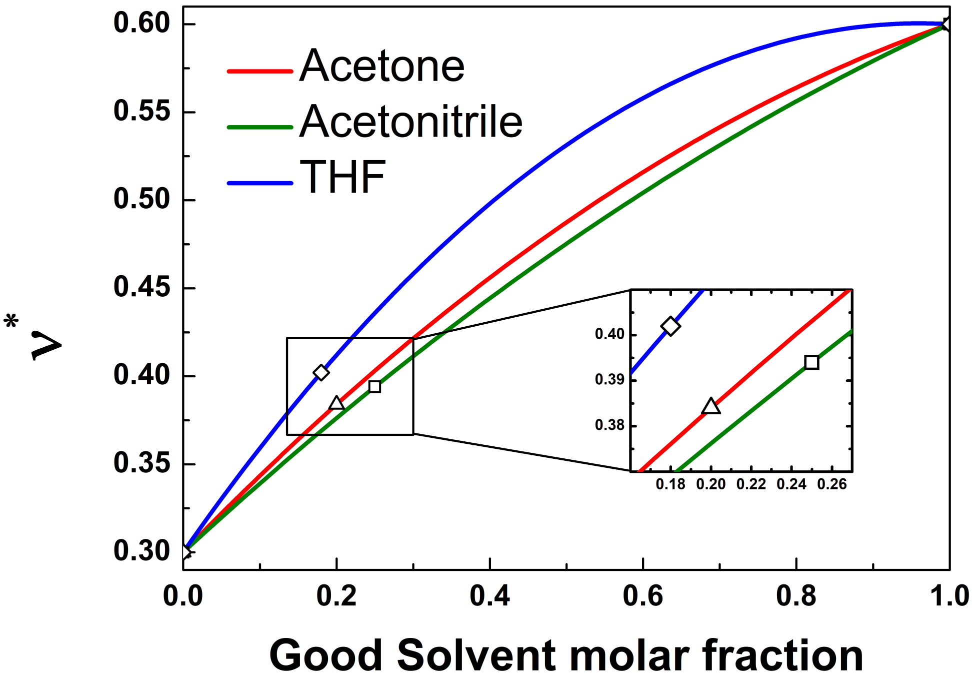


Figure 11.16: Flory's exponent ν^* profiles in function of good solvent molar fraction for acetone (red), acetonitrile (green) and THF (blue). The discrete symbols correspond to the conditions of bad and good solvent (1/3 and 3/5, respectively) and at the outlet mean good solvent molar fraction (inset).

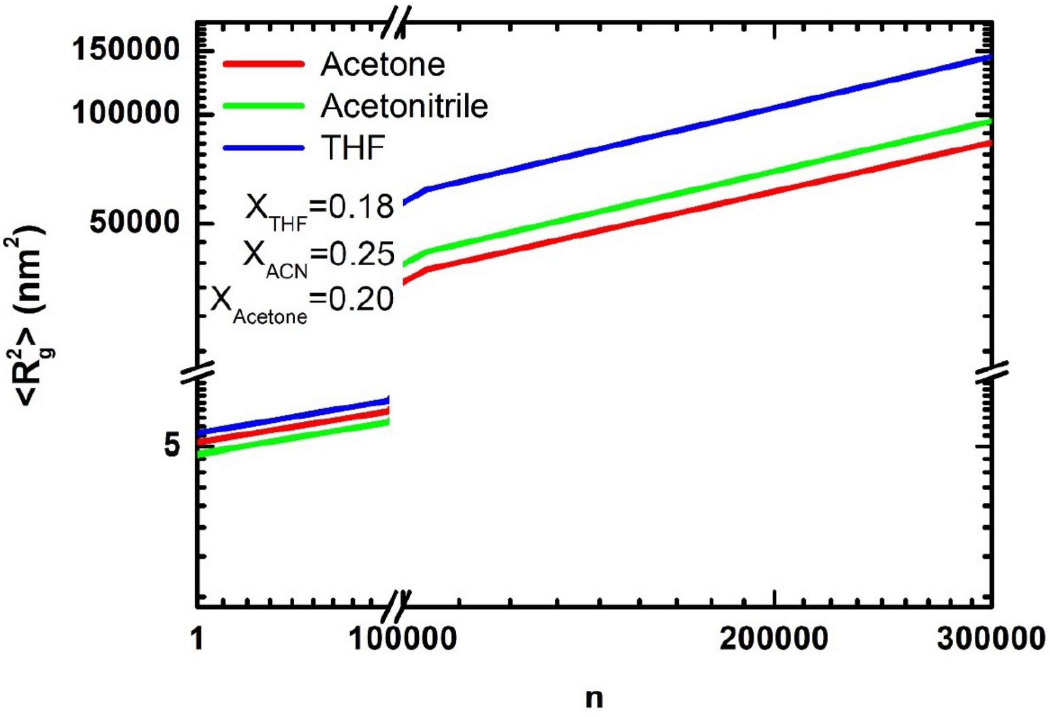


Figure 11.17: Mean squared radius of gyration in function of the number of molecules that form a cluster (aggregation number, *n*) for acetone (red), acetonitrile (green) and THF (blue).

11.5 Conclusions

In this chapter the self-assembly of PCL in acetone-water mixtures is modeled, for FNP conducted inside a CIJM, with the aim of developing a fully-predictive model capable of describing NP formation. The modeling approach is based on CFD to simulate the mixing between the two streams of acetone and water, but it uses also a PBM to describe the evolution of the CMD. The main novelty consists in considering for the first time the number *n* of polymer molecules (i.e., aggregation number) that form a cluster (or NP) as internal coordinate of the univariate population balance equation (PBE), overcoming the usual distinction between nucleation, molecular growth and aggregation. The resulting model is therefore defined as purely aggregative process without any form of energy barrier. In the theory of phase separation this corresponds to modeling the phenomenon as a spinodal decomposition (instead of a nucleation process). This assumption is valid especially at high initial supersaturation ratios of the polymer, which characterise FNP processes. The aggregation number as internal coordinate of the PBE allows to have a more direct link with the atomistic scale, building the aggregation kernels directly from MD simulations (in terms of Flory's law) performed in a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014).

In order to assess the validity of the model, several operating conditions are evaluated in terms of initial PCL concentrations in the acetone inlet stream and in terms of inlet flow rate values. Moreover, two molecular weights for PCL are considered: 14000 and 80000 g mol^{-1}. Simulation results are validated against experiments, resulting in good agreement, especially at high polymer concentrations. The agreement deteriorates as the polymer concentration is decreased, indicating that, in line with the corresponding theory, at low polymer concentrations and therefore low driving force for phase separation, the role of nucleation starts being important.

In the final part of this chapter the effect of different good solvents is evaluated, showing a transferable procedure to assess new functional forms for the Flory's parameters when the good solvent changes and, furthermore, assessing the key physical parameters (molar volume, solubility parameters) that are responsible of the different NP size at the outlet of the mixer, when acetone is substituted by acetonitrile and THF. This approach is conducted at the clusters and vessel scales, it has a general validity and can be applied to other good solvents. Future developments may be done, by investigating the validity of the results obtained here at the atomistic scale, through MD simulations.

Moreover, future work may include also the investigation of other polymers, experimentally investigated in the past, together with considering the presence of the drug inside the NP.

Chapter 12

Conclusions

In this work a multiscale modelling approach for soft matter self-assembly in binary mixtures is developed and tested on the case of Flash Nano-Precipitation (FNP) of poly-ε-caprolactone (PCL) in acetone-water mixtures. Considering the complexity of the phenomena involved in this process, three different scales are investigated: the atomistic scale, in which molecular phenomena are modelled thanks to Molecular Dynamics (MD) and Coarse-Grained Molecular Dynamics (CGMD); the clusters (or, nanoparticles, NP) scale, namely the scale of the population of polymer clusters dynamics, by means of a population balance model (PBM); the vessel scale, that is the scale of continuum, in which fluid dynamics effects on NP formation are evaluated through Computational Fluid Dynamics (CFD).

The PBM is implemented and coupled into a CFD code and it is, in turn, built up from MD simulations performed in a previous work (Di Pasquale, Marchisio, Barresi, and Carbone, 2014).

Full-atom MD is used to investigate the intimate mixing properties of good and bad solvent (acetone and water). Despite the well-known experimental fact that acetone is miscible with water in any proportion at room temperature, MD calculations showed to be affected by a strong phase separation when using most of the standard all-atom force fields. The strong de-mixing behaviour (that is particularly stressed at low acetone molar fractions) was shown to be due to the bad polarization response of the classical all-atom force fields, usually employed to simulate such liquid systems (Perera and Sokolić, 2004; Pereyra et al., 2011). The de-mixing of acetone-water mixtures is overcome at the atomistic scale by means of a Charge-on-Particle model (COP) (Lavino, Banetta, et al., 2018), in which a pseudo-polarizable approach is adopted. A charge displacement is introduced on the carbonyl group of the acetone molecule, and suitably modified in function of the acetone molar fraction in order to get the desired polarizable response. The dipole moment increases in this way together with the content of water and the de-mixing is shown to be disappeared. Beyond the de-mixing that is demonstrated to vanish, also dynamical properties are shown to be in better agreement with experiments, with respect to the standard all-atom force fields.

Coarse-grained molecular dynamics is here employed in order to investigate the behaviour of long polymer chains in solution. The MARTINI CG force field is used (Marrink, Risselada, et al., 2007), in which both thermodynamics and structural investigations are adopted to build up the model. Non-bonded interactions referred to the single CG bead types are set up by matching the solvation free energies of the corresponding atomistic structures, by means of the Bennett's Acceptance Ratio method. The CG polymer chain is then optimised always in terms of non-bonded interactions, by matching the radius of gyration of the atomic chains. Although the CG model developed here shows the typical limitations of the CG approaches (e.g., the sharp globule-to-coil transition at medium good solvent molar fraction in mixture), results are quite satisfactory in terms of both thermodynamics and structural behaviour of the system, paving the way towards future simulations of multiple chains in the simulation box, in order to get a better insight into the PCL self-assembly in solution. Molecular dynamics simulations are performed in an open-source tool, the GROMACS simulation package (Hess, Kutzner, et al., 2008).

Moving to the clusters scale, the aim of the PBM is to describe the evolution of the cluster mass distribution (CMD, i.e., the number of polymer molecules that form a NP) thanks to suitable transport equations (population balance equations) and, at the same time, to predict the final mean NP size at the mixer outlet, in terms of mean radius of gyration, by using the Flory's theory (Flory, 1953) of real polymers. The whole PBM is treated with a quadrature-based method of moments approach (QBMM, Marchisio and Fox, 2013). A novelty of this PBM approach is represented by setting the number of polymer molecules n (i.e., the aggregation number) that form a cluster as internal coordinate of the univariate PBE, overcoming the usual distinction between nucleation, molecular growth and aggregation steps. The model can be therefore defined as fully predictive aggregative. Turbulence fluctuations on NP formation are taken into account thanks to the direct quadrature method of moments coupled together with the interaction-and-exchange with the mean (DQMOM-IEM) method and by solving the Favre-averaged continuity and Navier-Stokes equations. The $k - \varepsilon$ standard model is employed. The rate at which two polymer clusters collide and aggregate is modelled by means of the aggregation kernels (built up from MD calculations) and suitably implemented in the PBM. Finally, the effect of fluid dynamics on NP formation (in terms of influence on both CMD and final radius of gyration), is studied by CFD. The latter two scales (PBM and CFD) are coupled together in a suitable user defined function in a commercial CFD code (ANSYS Fluent 15.0).

An alternative mathematical approach is also proposed, in which mixing and aggregation are considered together in a bivariate PBE. Mixture fraction and the pure moments in aggregation are used as internal coordinates of the bivariate PBE, resulting in the transport of the mixed moments of order zero and one. The conditional quadrature method of moments is employed to close the mathematical framwork, and the entire PBM is developed and implemented in an open-source tool, OpenQBMM (Passalacqua et al., 2018), based on OpenFOAM package. Being currently under numerical optimization, simulations results are still not present here and future investigations can be made by comparing the results obtained with the last technique to the first approach proposed and validated in this work (Lavino, Di Pasquale, et al., 2017).

Furthermore, the effect of different good solvents on NP formation is evaluated at this scale. Acetonitrile and THF are studied and the key parameters, responsible for the different NP size have been successfully identified, thanks to both tuning techniques and theoretical (Flory-Huggins solubility theory) approaches. Further investigations on the different solvents effect may be carried out at the atomistic scale, by means of MD analysis.

Future developments regarding this multiscale approach can be done, by validating the current models on different soft matter systems, as well as by considering the presence of the organic drug (together with the PCL) inside this fully-predictive pure-aggregative model.

Appendix A

Industrial and Modeling Overview on Number Density Functions

A.1 Number density function and population balance equation

In several industrial applications, chemical processes are characterised by multiphase flows, namely a flow in which continuous and disperse phases can be identified (crystallizers, fluidized beds, bubble columns). The dispersed phase, in turn, is characterised by some properties such as mass, velocity, particle size, etc. When at least one of these properties varies from particle to particle, then the disperse phase is characterised by a distribution function of that specific property and the corresponding system is labelled as polydisperse system. For example, in a crystallisation process, the crystals and the quality of the product strongly depend on the crystal size distribution (CSD) at the outlet of the crystallizer, in turn, influenced by the evolution of the CSD inside the reactor. The evolution of the CSD inside the control volume (e.g., the crystallizer) is governed by a kinetic equation, named population balance equation (PBE).

In general, a PBE is a continuity statement written in terms of partial integro-differential equations that describes the evolution of density functions referred to the disperse phase. Moreover, also the effect of the fluid dynamics of the continuous phase may affect the behaviour of the disperse one, in terms of evolution of density functions. This means that a theoretical framework together with a suitable computational modelling approach are necessary, in order to describe such complex systems.

Regarding the density functions, they have been introduced so far as functions able to consider one or more peculiar features of the disperse phase. For instance, in a crystallisation process it might be the distribution of the crystal sizes; in a carrier drug delivery system it might correspond to the number of molecules that form a delivery nano-cluster. It might be also the gas volume fraction of a given component in a gas-liquid system (bubble columns). In the first two cases the density function corresponds respectively to the crystal-size distribution (or, equally, the particle size distribution, PSD) and cluster mass distribution (CMD) in the case of drug delivery systems. In the latter case it corresponds instead to a distribution function that surely belongs to the range [0,1] (volume fraction). Therefore, in the case of CMD and PSD (or CSD) the density function is generically defined as number density function (NDF); in the case of bubble columns it corresponds to a probability density function (PDF) since the integral domain is limited to [0, 1], or in other words the NDF is normalised to one. Another example of PDF is represented by the mixture fraction, which has a key role in many industrial processes. From a more mathematical point of view the PDF, $f(x)$, is defined such that

$$f(x)dx = Pr[x, x + dx] \tag{A.1}$$

namely, the quantity $f(x)dx$ represents the probability of finding the system in a configuration between x and $x + dx$. This is also schematically depicted in Figure A.1.

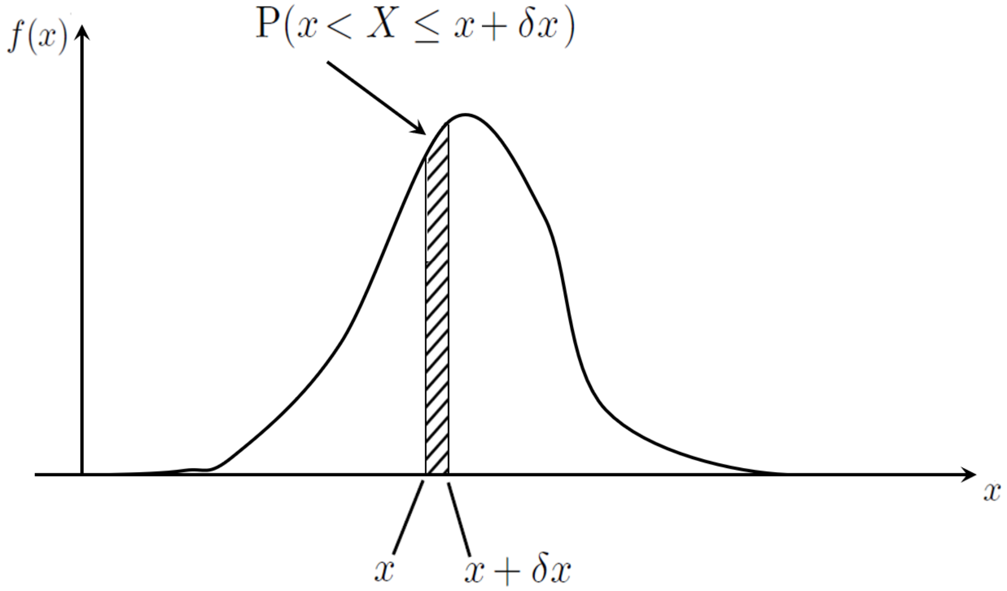


Figure A.1: Sketch of a probability density function, $f(x)$. The quantity $f(x)dx$ represents the probability of the system to lie in the configurational interval $[x, x + dx]$

.

Being a probability, another property of the PDF is that the integral over the whole domain yields the unity:

$$\int_0^1 f(x)dx = 1. \tag{A.2}$$

If the integral over the phase space domain yields to a number $A \neq 1$, then the PDF is a NDF. When the particles in a system (disperse phase) have non-negligible inertia, then also the velocity associated to each particle becomes important. This is the case of bubbles column systems. Each particle (bubble, for instance) has different velocity, and the whole system is represented by a velocity distribution function. The evolution of the velocity distribution function is taken into account by a suitable kinetic equation. Taking into consideration the velocity distribution function (namely, considering the inertia of the single particles) leads to the so-called generalised population balance equation (GPBE). In the following section, a closer look at further properties of NDF is given.

A.2 Properties of number density functions

The NDF can be defined as the number of entities (particles) per unit volume with a given range of cluster scale variables. The cluster scale variables are those variables that play a key role in the cluster scale system (mixture fraction, volume fraction, particle size ...). Cluster scale variables are said to be realisable if they can be reproduced and observed at the atomistic scale level; otherwise, they are defined unrealisable. Cluster scale variables are very often correlated to each other; thus, the cluster scale model (the NDF) must take this into account. This means that two correlated cluster scale variables cannot be reconstructed by two separate NDF referred to each variable separately. For example, in a binary collisions system (e.g., spheres that collide pairwise) two particles are correlated to each other in terms of velocities before and after a collision. Consequently, the NDF referred to one-particle (so-called *one-particle NDF*) for velocity is influenced by *two-particle NDF* due to the collisions. A cluster scale model framework is then needed, in order to capture the complexity of the correlations between cluster scale variables. Let us give some other useful definitions.

The evolution of NDF is governed by the so-called kinetic equation. Kinetic equations for NDF that do not take into account the particle velocity are labelled population balance equation (PBE); on the other side, if the kinetic equation accounts for the velocity of the disperse phase, then it is defined as generalised population balance equation (GPBE), in both cases together with all the other cluster scale variables.

In general, the variables that appear in a kinetic equation (PBE or, more generally, in a GPBE) can be classified in physical variables (time, space) if they represent a point in the real space, and cluster scale variables (cluster scale point) represented by a point in the so-called phase space (mixture fraction, particle size, ...). By integrating the kinetic equation over the phase space domain, a continuum model is thus obtained (only time and space as independent variables). The last very crucial concept that is mandatory to be introduced is the *closure*. A kinetic equation is said to be closed if all the terms appearing in it are formulated in terms of cluster scale variables. The main goal of cluster scale modelling is to get to closed kinetic equations that are able to account for the atomistic scale physics effects on the cluster scale.

Appendix B

Derivation of Source Terms for Second-Order Point Processes in PBE

B.1 Source term derivation

In this appendix a deeper insight into the derivation of the source term for second-order point processes is provided. Let us recall the definition of second-order point process: these are processes involving the interaction of two particles, during which velocities, mass and other internal properties are instantaneously changed Therefore, aggregation, coalescence and agglomeration represent an example of second-order point processes. The assumption which these mathematical modeling lies on is that, during the interacting event (e.g., collision), only two particles interact at the time.

For a sake of simplicity, let us consider a granular system (i.e., without any solvent), described by the number density function (NDF) $f_N(t, \mathbf{x}, \boldsymbol{\xi})$ evolving in time t, physical space $\mathbf{x}$ and that is characterised by the internal coordinate vector $\boldsymbol{\xi}$. The generic source term, in a second-order point process, reads as follows:

$$h_{\xi} = h^{+} - h^{-}, \tag{B.1}$$

where h^{+} is the production rate and h^{-} represents the particle loss rate due to the second-order event.

Let us identify one of the two particles with its center of mass position vector $\mathbf{x}'$ and its phase space vector $\boldsymbol{\xi}'$; another particle is instead placed at $\tilde{\mathbf{x}}$ with its phase space vector $\tilde{\boldsymbol{\xi}}$. The frequency of this second-order process is

$$a(\mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}), \tag{B.2}$$

and the quantity $a(\mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})\mathrm{d}t$ represents the number of particles undergoing second-order events in the time interval $\mathrm{d}t$. The NDF related to this second order process is $f_N^{(2)}(t, \mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})$ and it represents the expected number of particle pairs with $(\mathbf{x}', \boldsymbol{\xi}')$

and $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})$. Using the pairs NDF would lead to a too high computational cost due to the solution of an equation for $f_N^{(3)}$, therefore the following closure is adopted:

$$f_N^{(2)}(t, \mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \approx f_N(t, \mathbf{x}', \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}). \tag{B.3}$$

The particles located at $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})$ are labelled as *test* particles, whereas particles placed at $(\mathbf{x}', \boldsymbol{\xi}')$ are known as *field* particles. The number of second-order events, N_e occurring between particles located near $(\tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})$, per unit volume and per unit time is:

$$N_e = \int a(\mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) f_N(t, \mathbf{x}', \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \mathrm{d}\mathbf{x}' \mathrm{d}\boldsymbol{\xi}'. \tag{B.4}$$

Being very often the test and field particles distances much greater than particles dimensions, f_N does not change significantly in physical space within these distances. Consequently, the closure $f_N(t, \mathbf{x}', \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \approx f_N(t, \tilde{\mathbf{x}}, \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}})$ can be done. This allows us to rearrange Eq. (B.4) in the following way:

$$N_e = \int \left(\int a(\mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \mathrm{d}\mathbf{x}' \right) f_N(t, \tilde{\mathbf{x}}, \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \mathrm{d}\boldsymbol{\xi}'. \tag{B.5}$$

The integral appearing in Eq. (B.5) is the kernel of the second-order point process, already introduced for the particular case of aggregation of fine particles in section 5.3.3:

$$\beta(\tilde{\mathbf{x}}; \boldsymbol{\xi}', \tilde{\boldsymbol{\xi}}) = \int a(\mathbf{x}', \boldsymbol{\xi}'; \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \mathrm{d}\mathbf{x}'. \tag{B.6}$$

The dimensions of the kernel usually correspond to physical volume per unit time, stating the volumetric rate at which these processes occur.

Let us consider a system in which the two internal coordinates vectors $\boldsymbol{\xi}'$ and $\tilde{\boldsymbol{\xi}}$ result in $\boldsymbol{\xi}$ after the interaction takes place. Recalling Eq. (B.1), its positive term h^+ reads as follows (Ramkrishna, 2000):

$$h^+ = \frac{1}{\delta} \int \beta(\boldsymbol{\xi}', \tilde{\boldsymbol{\xi}}) f_N(t, \tilde{\mathbf{x}}, \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) J(\tilde{\boldsymbol{\xi}}, \boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}', \tag{B.7}$$

where δ is a symmetry factor avoiding multiple counting (it is equal to two for fine particles aggregation) and $J(\tilde{\boldsymbol{\xi}}, \boldsymbol{\xi})$ is the determinant of the partial derivatives matrix $\partial \tilde{\xi}_i / \partial \xi_j$ corresponding to the transformation from phase space variable $\tilde{\xi}$, before collision, to ξ, after collision.

The particles loss rate, h^- is:

$$h^- = \int \beta(\boldsymbol{\xi}', \tilde{\boldsymbol{\xi}}) f_N(t, \tilde{\mathbf{x}}, \boldsymbol{\xi}') f_N(t, \tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) \mathrm{d}\boldsymbol{\xi}'. \tag{B.8}$$

Appendix C

From PBE to moments transport equation - derivation and closure of the micro-mixing term

In this section a more extensive derivation of the transport equations of moments of CMD is shown, underlying the key mathematical details. Let us start with a PBE referred to the generic NDF $f_N(t, \mathbf{x}, \xi)$ presented in Eq. (8.44) and, for a sake of simplicity, by butting the number of internal variables equal to one.

$$\frac{\partial f_N(\xi)}{\partial t} + \frac{\partial}{\partial \mathbf{x}} \cdot (\mathbf{v} f_N(\xi)) + \frac{\partial}{\partial \xi} \cdot (\dot{\xi} f_N(\xi)) = h_\xi(\xi), \tag{C.1}$$

where $h_\xi(\xi)$ is the source term and only the internal variable ξ is indicated for brevity. The third term on the left hand side needs a closure with a mixing model; the interaction and exchange with the mean method (IEM) is chosen and already shown in this work in section 8.6. Recalling the IEM model, it is a linear relaxation of the passive scalar with its mean (Fox, 2003):

$$-\frac{\partial}{\partial \psi}(\langle \Gamma \nabla^2 \phi' | \boldsymbol{\psi} \rangle f_\phi) = -\frac{\partial}{\partial \psi}\left[\frac{\varepsilon_\phi}{2\langle \phi'^2 \rangle}(\langle \phi \rangle - \psi) f_\phi\right], \tag{C.2}$$

where ε_ϕ is called scalar covariance dissipation rate and $\langle \phi'^2 \rangle$ is the inverse covariance matrix of the scalars. This term is closed, as already seen, by employing the micro-mixing rate γ_M:

$$\frac{\varepsilon_\phi}{2\langle \phi'^2 \rangle} = \frac{1}{2\tau_T} = \frac{C_\phi \varepsilon}{2k} = \gamma_M, \tag{C.3}$$

Let us also recall the definition of generic k^{th}-order moment:

$$m_k(\xi) = \langle \xi^k \rangle = \int_{\Omega_\xi} \xi^k f_N(\xi) d\xi, \tag{C.4}$$

Multiplying Eq. (C.1) by ξ^k dξ and integrating out over the internal variable phase space domain Ω_ξ, the following equation is obtained:

$$\frac{\partial\left(\int_{\Omega_\xi}\xi^k f_N(\xi)\mathrm{d}\xi\right)}{\partial t}+\frac{\partial}{\partial\mathbf{x}}\cdot\left[\mathbf{v}\left(\int_{\Omega_\xi}\xi^k f_N(\xi)\mathrm{d}\xi\right)\right]+$$

$$+\int_{\Omega_\xi}\xi^k\frac{\partial}{\partial\xi}\cdot\left[\gamma_M\left(\langle\xi\rangle-\xi\right)f_N(\xi)\right]\mathrm{d}\xi=\int_{\Omega_\xi}\xi^k h_\xi(\xi)\mathrm{d}\xi, \quad \text{(C.5)}$$

where, in the first two terms on left hand side, the big round parenthesis correspond to the k^{th}-order moment, defined in Eq. (C.4). The third term on left hand side needs further developments in order to eliminate the partial derivative, by integrating it by parts:

$$\gamma_M\int_{\Omega_\xi}\xi^k\frac{\partial}{\partial\xi}\cdot\left[\left(\langle\xi\rangle-\xi\right)f_N(\xi)\right]\mathrm{d}\xi=-k\gamma_M\int_{\Omega_\xi}\xi^{k-1}\left(\langle\xi\rangle-\xi\right)f_N(\xi)\mathrm{d}\xi, \quad \text{(C.6)}$$

and, by applying the moment definition given in Eq. (C.4) ($\langle\xi\rangle=m_1$):

$$-k\gamma_M\int_{\Omega_\xi}\left(\xi^{k-1}\langle\xi\rangle-\xi^{k-1}\xi\right)f_N(\xi)\mathrm{d}\xi=$$

$$=-k\gamma_M\langle\xi\rangle\int_{\Omega_\xi}\xi^{k-1}f_N(\xi)\mathrm{d}\xi+k\gamma_M\int_{\Omega_\xi}\xi^k f_N(\xi)\mathrm{d}\xi=$$

$$=-k\gamma_M\left(m_{k-1}m_1-m_k\right). \quad \text{(C.7)}$$

The final transport equation of the moments of the NDF reads then as follows:

$$\frac{\partial m_k(\xi)}{\partial t}+\frac{\partial}{\partial\mathbf{x}}\cdot\left(\mathbf{v}m_k(\xi)\right)-k\gamma_M\cdot\left[m_{k-1}(\xi)m_1(\xi)-m_k(\xi)\right]=h_\xi(m_k(\xi)), \quad \text{(C.8)}$$

where the source term on right hand side may need a proper closure, depending on the physical meaning of the internal coordinate ξ.

Appendix D

Topology of Charge-on-Particle Acetone model

In this appendix, the topology of charge-on-particle acetone model is shown, implemented at the atomistic scale in MD simulations.

```
  ;Charge-on-particle OPLSAA topology for acetone
;
#include ”oplsaaff.itp”
#include ”oplsaa.ff/spc.itp”
[ moleculetype ]
; Name            nrexcl
acetone            3

[ atoms ]
;nr   type   resnr    residue    atom  cgnr   charge  mass
1    opls_280   1        LIG        C     1      0.47    12.011
2    opls_135   1        LIG        C     2     -0.18    12.011
3    opls_135   1        LIG        C     3     -0.18    12.011
4    opls_281   1        LIG        O     4      0.0    15.9994
5    opls_282   1        LIG        H     5      0.06     1.008
6    opls_282   1        LIG        H     6      0.06     1.008
7    opls_282   1        LIG        H     7      0.06     1.008
8    opls_282   1        LIG        H     8      0.06     1.008
9    opls_282   1        LIG        H     9      0.06     1.008
10  opls_282    1        LIG        H     10     0.06     1.008
11      VS      1        LIG       VS    11     -0.47      0.0


[ bonds ]
```

```
;   ai     aj  funct
     1      2      1
     1      3      1
     1      4      1
     1     11      6
     2      5      1
     2      6      1
     2      7      1
     3      8      1
     3      9      1
     3     10      1

[ pairs ]
;   ai     aj  funct
     2      8      1
     2      9      1
     2     10      1
     3      5      1
     3      6      1
     3      7      1
     4      5      1
     4      6      1
     4      7      1
     4      8      1
     4      9      1
     4     10      1

[ angles ]
;   ai     aj     ak  funct
     2      1      3      1
     2      1      4      1
     3      1      4      1
     1      2      5      1
     1      2      6      1
     1      2      7      1
     5      2      6      1
     5      2      7      1
     6      2      7      1
     1      3      8      1
     1      3      9      1
     1      3     10      1
     8      3      9      1
```

```
    8     3    10     1
    9     3    10     1

[ dihedrals ]
;  ai    aj    ak    al funct
    3     1     2     5     3
    3     1     2     6     3
    3     1     2     7     3
    4     1     2     5     3
    4     1     2     6     3
    4     1     2     7     3
    2     1     3     8     3
    2     1     3     9     3
    2     1     3    10     3
    4     1     3     8     3
    4     1     3     9     3
    4     1     3    10     3

[ virtual_sites2 ]
; site  ai  aj  funct   a
    11   1   4    1     1.16

[ system ]
mixture
[ molecules ]
acetone 301
SOL 899
```

Bibliography

Allen, M. P. and D. J. Tildesley (1987), *Computer Simulation of Liquids*, Oxford Science Publication.

Allison, S.K., J.P. Fox, R. Hargreaves, and S.P. Bates (2005), “Clustering and Microimmiscibility in Alcohol-Water Mixtures: Evidence from Molecular-Dynamics Simulations”, *Phys. Rev. B*, 71, p. 024201.

Ammar, Y. and M. Reeks (2009), “Agglomeration of inertial particles in a random rotating symmetric straining flow”, *International Journal of Multiphase Flow*, 35, pp. 840-853.

Ancona, A., B. Dumontel, N. Garino, B. Demarco, D. Chatzitheodoridou, W. Fazzini, H. Engelke, and V. Cauda (2018), “Lipid-Coated Zinc Oxide Nanoparticles as Innovative ROS-Generators for Photodynamic Therapy in Cancer Cells”, *Nanomaterials*, 8, 3, p. 143.

Andersen, H. C. (1980), “Molecular dynamics simulations at constant pressure and/or temperature”, *J. Chem. Phys.* 72, p. 2384.

Andersson, B, R Andersson, L Hakansson, M Mortensen, R Sudiyo, and B van Wachem (2012), *Computational fluid dynamics for engineering*, Cambridge University Press.

Angell, C.A. (1982), *Water: A Comprehensive Treatise*, Vol. 7, p. 1, Plenum: New York, USA.

Applequist, J., J.R. Carl, and K.K. Fung (1972), “Atom Dipole Interaction Model for Molecular Polarizability. Application to Polyatomic Molecules and Determination of Atom Polarizabilities”, *J. Am. Chem. Soc.* 94, pp. 2952-2960.

Baldyga, J and W Orciuch (2001), “Some hydrodynamic aspects of precipitation”, *Powder Technol.* 121, p. 9.

Barrett, J C and N A Webb (1998), “A comparison of some approximate methods for solving the aerosol general dynamic equation”, *J. Aerosol Sci.* 29, p. 31.

Ben-Naim, A. and Y. Marcus (1984), “Solvation thermodynamics of nonionic solutes”, *The Journal of Chemical Physics*, 81, 4, pp. 2016-2027.

Bennett, C.H. (1976), “Efficient estimation of free energy differences from Monte Carlo data”, *Journal of Computational Physics*, 22, 2, pp. 245-268.

Berendsen, H.J.C. (2007), *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*, Cambridge University Press.

Berendsen, H.J.C., J.R. Grigera, and T.P. Straatsma (1987), “The Missing Term in Effective Pair Potentials”, *J. Phys. Chem.* 91, 6269–6271.

Berendsen, H.J.C., J.P.M. Postma, A. DiNola, and J.R. Haak (1984), “Molecular dynamics with coupling to an external bath”, *J. Chem. Phys.* 81, pp. 3684-3690.

Berendsen, H.J.C., J.P.M. Postma, W.F. Von Gusteren, and J. Hermans (1981), “Interaction Models for Water in Relation to Protein Hydration”, *In: Intermolecular forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry, B., Pullman, Editor, Reidel Publ. Company, Dordrecht, Holland*, pp. 331-342.

Berendsen, H.J.C. and W.F. van Gunsteren (1984), “Molecular dynamics simulations: Techniques and approaches”, *in ”Molecular Liquids: Dynamics and Interactions”, eds A.J. Barnes et al., NATO ASI Series*, C 135, pp. 475-500.

Bird, R. B., W. E. Stewart, and E. N. Lightfoot (1960), *Transport Phenomena*, John Wiley and Sons.

Bochicchio, D. and G. M. Pavan (2017), “From Cooperative Self-Assembly to Water-Soluble Supramolecular Polymers Using Coarse-Grained Simulations”, *ACS Nano*, 11, pp. 1000-1011.

Bockstaller, M. R., R. A. Mickiewicz, and E. Thomas (2005), “Block Copolymer Nanocomposites: Perspectives for Tailored Functional Materials”, *Adv. Mater.* 17, pp. 1331-1349.

Bonomi, M., G. T. Heller, C. Camilloni, and M. Vendruscolo (2017), “Principles of protein structural ensemble determination”, *Current Opinion in Structural Biology*, 42, pp. 106-116.

Bossis, G., B. Quentrec, and J. P. Boon (1982), “Brownian dynamics and the fluctuation-dissipation theorem”, *Mol. Phys.* 45, pp. 191-6.

Brini, E. and N. F. van der Vegt (2012), “Chemically transferable coarse-grained potentials from conditional reversible work calculations”, *J. Chem. Phys.* 137, pp. 15411310.1063/1.4758936.

Bruckner, S. and S. Boresch (2011), “Efficiency of alchemical free energy simulations. I. A practical comparison of the exponential formula, thermodynamic integration, and Bennett’s acceptance ratio method”, *Journal of Computational Chemistry*, 32, 7, pp. 1303-1319.

Bussi, G., D. Donadio, and M. Parrinello (2007), “Canonical sampling through velocity rescaling”, *J. Chem. Phys.* 126, p. 014101.

Carbone, P. and C. Avendaño (2014), “Coarse-Grained Methods for Polymeric Materials: Enthalpy- and Entropy-Driven Models”, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4, 1, pp. 62-70.

Celasco, E., I. Valente, D.L. Marchisio, and A.A. Barresi (2014), “Dynamic light scattering and X-ray photoelectron spectroscopy characterization of PEGylated polymer nanocarriers: internal structure and surface properties”, *Langmuir*, 30, pp. 8326-8335.

Chan, H. K. and P. C. L. Kwok (2011), "Production methods for nanodrug particles using the bottom-up approach", *Advanced Drug Delivery Reviews*, 63, p. 406.

Cheng, J., A. Vishnyakov, and A.V. Neimark (2014), "Morphological Transformations in Polymer Brushes in Binary Mixtures: DPD Study", *Langmuir*, 30, 43, pp. 12932-12940.

Cheng, J.C., R.D. Vigil, and R.O. Fox (2010), "A competitive aggregation model for Flash NanoPrecipitation", *J. Colloid Interface Sci.* 351, 2, pp. 330-342.

Cheung, D. L. and P. Carbone (2013), "How Stable Are Amphiphilic Dendrimers at the Liquid–Liquid Interface?", *Soft Matter*, 9, 6841–6850.

Choi, Y. J., S. T. Chung, M. Oh, and H. S. Kim (2005), "Investigation of cristallization in a Jet Y-Mixer by a hybrid computational fluid dynamics and process simulation approach", *Cristal Growth & Design*, 5, p. 959.

Ciccotti, G. and J.P. Ryckaert (1981), "On the derivation of the generalized Langevin equation for interacting Brownian particles", *J. Stat. Phys.* 26, pp. 73-82.

Cohen, S. I. A., M. Vendruscolo, M. E. Welland, C. M. Dobson, E. M. Terentjev, and T. P. J. Knowles (2011), "Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments", *J. Chem. Phys.* 135, p. 065105.

Collins, S. R., A. Douglass, R. D. Vale, and J. S. Weissman (2004), "Mechanism of Prion Propagation: Amyloid Growth Occurs by Monomer Addition", *PLoS Biol.* 2, e321.

Connolly, M.L. (1983), "Solvent-Accessible Surfaces of Proteins and Nucleic Acids", *Science*, 221, pp. 709-713.

Dans, P. D., A. Zeida, M. R. Machado, and S. Pantano (2010), "A Coarse Grained Model for Atomic-Detailed DNA Simulations with Explicit Electrostatics", *J. Chem. Theory Comput.* 6, 1711–1725.

Darden, T., D. York, and L. Pedersen (1993), "Particle mesh Ewald: An N· log(N) method for Ewald sums in large systems", *J. Chem. Phys.* 98, pp. 10089-10092.

Deichmann, G., V. Marcon, and N. van der Vegt (2014), "Bottom-up derivation of conservative and dissipative interactions for coarse-grained molecular liquids with the conditional reversible work method", *J. Chem. Phys.* 141, p. 224109.

Dette, Holger and William J. Studden (2002), "Matrix measures, moment spaces and Favard's theorem for the interval [0,1] and [0,∞)", *Linear Algebra and its Applications*, 345, pp. 169-193.

Di Pasquale, N., T. Hudson, and M. Icardi (2018), "Systematic derivation of hybrid coarse-grained models", *ArXiv e-prints*, -, pp. -.

Di Pasquale, N., D.L. Marchisio, and A.A. Barresi (2012), "Model validation for precipitation in solvent-displacement processes", *Chem. Eng. Sci.* 84, p. 671.

Di Pasquale, N., D.L. Marchisio, A.A. Barresi, and P. Carbone (2014), "Solvent Structuring and its Effect on the Polymer Structure and Processability: the Case of Water-Acetone Poly--Caprolactone Mixtures", *J. Phys. Chem. B*, 118, p. 13258.

Di Pasquale, N., D.L. Marchisio, P. Carbone, and A.A. Barresi (2013), "Identification of nucleation rate parameters with MD and validation of the CFD model for polymer particle precipitation", *Chem. Eng. Res. and Des.* 91, p. 2275.

Eisenhaber, F., P. Lijnzaad, P. Argos, C. Sander, and M. Scharf (1995), "The Double Cubic Lattice Method: Efficient Approaches to Numerical Integration of Surface Area and Volume and to Dot Surface Contouring of Molecular Assemblies", *J. Comput. Chem.* 16, pp. 273-284.

Elimelech, M., J. Gregory, X. Jia, and R. Williams (1998), *Particle Deposition and Aggregation: Measurement, Modelling and Simulation*, Woburn (MA): Butterworth-Heinemann.

Essmann, U., L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen (1995), "A smooth particle mesh ewald potential", *J. Chem. Phys.* 103, pp. 8577-8592.

Fan, X., N. Phan-Thien, N. T. Yong, X. Wu, and D. Xu (2003), "Microchannel flow of a macromolecular suspension", *Physics of Fluids*, 15, pp. 11-21.

Favre, A. (1965), "Equations des gaz turbulents compressibles I. Formes générales", *Journal de Mecanique*, 4, pp. 361-390.

Ferrario, M., M. Haughney, I.R. McDonald, and M.L. Klein (1990), "Molecular-Dynamics Simulation of Aqueous Mixtures: Methanol, Acetone, and Ammonia", *J. Chem. Phys.* 93, pp. 5156-5166.

Ferri, A., N. Kumari, R. Peila, and A.A. Barresi (2017), "Production of menthol-loaded nanoparticles by solvent displacement", *The Canadian Journal of Chemical Engineering*, 95, 9, pp. 1690-1706.

Ferrone, F. A., J. Hofrichter, and W. A. Eaton (1985), "Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism", *J. Mol. Biol.* 183, p. 611.

Ferziger, J.H. and M. Peric (2001), *Computational Methods for Fluid Dynamics*, Springer Berlin Heidelberg.

Fixman, M. (1974), "Classical statistical mechanics of constraints: a theorem and application to polymers", *Proc. Natl. Acad. Sci.* 71, pp. 3050-3.

Flory, P. (1953), *Principles of Polymer Chemistry*, Cornell Univ. Press.

Fox, R.O. (2003), *Computational models for turbulent reacting flows*, Cambridge University Press.

— (1994), "Improved Fokker–Planck model for the joint scalar, scalar gradient PDF", *Phys. Fluids*, 6, p. 334.

— (1999), "The Lagrangian spectral relaxation model for differential diffusion in homogeneous turbulence", *Phys. Fluids*, 11, p. 1550.

Freitas, L.C.G., J.M.M. Cordeiro, and F.L.L. Garbujo (1999), "Theoretical Studies of Liquids by Computer Simulations: The Water-Acetone Mixture", *J. Mol. Liq.* 79, pp. 1-15.

Fuchs, N. (1964), *Mechanics of Aerosols*, New York (NY): Pergamon.

Garcia, G A, S I A Cohen, C M Dobson, and T P J Knowles (2013), "Nucleation-conversion-polymerization reactions of biological macromolecules with prenucleation cluster", *Phys. Rew. E*, 88.

Gautschi, W. (2004), *Orthogonal Polynomials: Computation and Approximation*, Oxford: Oxford University Press.

Gavi, E., L. Rivautella, D.L. Marchisio, M. Vanni, A.A. Barresi, and G. Baldi (2007), “CFD modelling of nano-particle precipitation in confined impinging jet reactor”, *Chem. Eng. Sci.* 85, p. 753.

Gear, C. W. (1967), “The numerical integration of ordinary differential equations”, *Math. Comp.* 21, pp. 146-156.

Georg, H.C., K. Coutinho, and S. Canuto (2006), “Converged Electronic Polarization of Acetone in Liquid Water and the Role in the n−π^* Transition”, *Chem. Phys. Letters*, 429, pp. 119-123.

Gordon, R. G. (1968), “Error bounds in equilibrium statistical mechanics”, *Journal of Mathematical Physics*, 9, pp. 655-663.

Gradl, J., H. C. Schwarzer, F. Schwertfirm, M. Manhart, and W. Peukert (2006), “Precipitation of nanoparticles in a T-mixer: coupling the particle population dynamics with hydrodynamics through direct numerical simulation”, *Chem. Eng. Process.* 45, p. 908.

Groot, R. D. and P. B. Warren (1997), “Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation”, *Journal of Chemical Physics*, 107, p. 4423.

Guillot, B.A. (2002), “Reappraisal of What We Have Learnt During Three Decades of Computer Simulations on Water”, *J. Mol. Liq.* 101, pp. 219-245.

Gupta R.; Patey, G.N. (2012), “Aggregation in Dilute Aqueous Tert-Butyl Alcohol Solutions: Insights from Large-Scale Simulations”, *J. Chem. Phys.* 137, p. 034509.

Hansen, C.M. (2007), *Hansen Solubility Parameters: A User’s Handbook, 2nd ed.* CRC Press, Boca Raton.

Hess, B. (2002), “Determining the Shear Viscosity of Model Liquids from Molecular Dynamics Simulations”, *J. Chem. Phys.* 116, 209–217.

Hess, B., H. Bekker, H.J.C. Berendsen, and J. G. E. M. Fraaije (1997), “LINCS: A linear constraint solver for molecular simulations”, *J. Comp. Chem.* 18, pp. 1463-1472.

Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl (2008), “GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation”, *J. Chem. Theory Comput.* 4, 435–447.

Hijon, C., P. Espanol, E. Vanden-Eijnden, and R. Delgado-Buscalioni (2010), “Mori–Zwanzig formalism as a practical computational tool”, *Faraday Discussions*, 144, p. 301.

Hildebrand, J. and R.L. Scott (1950), *Solubility of Nonelectrolytes*, 3rd ed. Reinhold, New York.

Hoover, W. G. (1985), “Canonical dynamics: equilibrium phase-space distributions”, *Phys. Rev. A*, 31, pp. 1695-1697.

Howard, K.S. and R.A. McAllister (1958), “The Viscosity of Acetone-Water Solutions up to Their Normal Boiling Points”, *A.I.Ch.E. Journal*, 4, 362–366.

Howard, M.P., A.Z. Panagiotopoulos, and A. Nikoubashman (2018), “Efficient mesoscale hydrodynamics: Multiparticle collision dynamics with massively parallel GPU acceleration”, *Computer Physics Communications*, 230, pp. 10-20.

Icardi, M., E. Gavi, D.L. Marchisio, A.A. Barresi, M.G. Olsen, R.O. Fox, and D. Lakehal (2011), "Investigation of the flow field in a three-dimensional Confined Impinging Jets Reactor by means of microPIV and DNS", *Chem. Eng. J.* 166, 1, pp. 294-305.

Icardi, M., E. Gavi, D.L. Marchisio, M.G. Olsen, R.O. Fox, and D. Lakehal (2011), "Validation of LES predictions for turbulent flow in a Confined Impinging Jets Reactor", *App. Math. Mod.* 35, 4, pp. 1591-1602.

Israelachvili, J.N. (2011), "5 - Interactions Involving the Polarization of Molecules", in *Intermolecular and Surface Forces (Third Edition)*, ed. by Jacob N. Israelachvili, Third Edition, Academic Press, San Diego, pp. 91-106, ISBN: 978-0-12-375182-9.

Izvekov, S., M. Parrinello, C. J. Burnham, and G. A. Voth (2004), "Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching", *Journal of Chemical Physics*, 120, p. 10896.

Izvekov, S. and G. A. Voth (2005), "Multiscale coarse graining of liquid-state systems", *Journal of Chemical Physics*, 123, p. 134105.

Jedlovszky, P., A. Idrissi, and G. Jancsó (2009), "Can Existing Models Qualitatively Describe the Mixing Behavior of Acetone with Water?", *J. Chem. Phys.* 130, p. 124516.

Johnson, B. K. and R. K. Prud'homme (2003a), "Chemical processing and micro-mixing in confined impinging jets", *AIChE J.* 49, p. 2264.

— (2003b), "Flash nano-precipitation of organic actives and block copolymers using a confined impinging jets mixer", *Aust. J. Chem*, 56, p. 1021.

Jones, R. A. L. (2002), *Soft condensed matter*, Oxford Univ. Press.

Jorgensen, W.L., J.M. Briggs, and M.L. Contreras (1990), "Relative Partition Coefficients for Organic Solutes from Fluid Simulations", *J. Phys. Chem.* 94, pp. 1683-1686.

Jorgensen, W.L., J. Chandrasekhar, J.D. Madura, R.W. Impey, and M.L. Klein (1983), "Comparison of Simple Potential Functions for Simulating Liquid Water", *J. Chem. Phys.* 79, pp. 926-935.

Jorgensen, W.L., D.S. Maxwell, and J. Tirado-Rives (1996), "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.

Karimi-Varzaneh, H. A., N. F. A. van der Vegt, F. Muller-Plathe, and P. Carbone (2012), "How Good Are Coarse-Grained Polymer Models? A Comparison for Atactic Polystyrene", *ChemPhysChem*, 13, 3428–3439.

Kauzlaric, D., O. Liba, Y. Hanein, P. Espanol, A. Greiner, S. Succi, and J. G. Korvink (2012), "Top-down vs. bottom-up coarse-graining of graphene and CNTs for nanodevice simulation", *7th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Kyoto,* pp. 298-303.

Kestin, J, M Sokolov, and W A Wakeham (1978), "viscosity of Liquid Water in the Range -8℃ To 150 ℃", *J. Phys. Chem.* 7, p. 941.

Kirkwood, John G. (1935), "Statistical Mechanics of Fluid Mixtures", *The Journal of Chemical Physics*, 3, pp. 300-313.

Knowles, T. P. J., C. A. Waudby, G. L. Devlin, S. I. A. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson (2009), "An analytical solution to the kinetics of breakable filament assembly", *Science*, 326, p. 1533.

Kramers, H. A. (1940), "Brownian motion in a field of force and the diffusion model of chemical reactions", *Physica*, 7, pp. 284-304.

Kubo, R. (1966), *The fluctuation–dissipation theorem and Brownian motion*, In References 571 Many-Body Theory, ed. R. Kubo. Tokyo–New York, Syokabo–Benjamin,

Kubo, R., M. Toda, and N. Hashitsume (1985), *Statistical Physics. II. Nonequilibrium Statistical Mechanics*, Springer–Verlag.

Lamoureux, G., A.D. MacKerell Jr., and B. Roux (2003), "A Simple Polarizable Model of Water Based on Classical Drude Oscillators", *J. Chem. Phys.* 119, 5185–5197.

Lamoureux, G. and B. Roux (2003), "Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm", *J. Chem. Phys.* 119, pp. 3025-3039.

Lavino, A. D., L. Banetta, P. Carbone, and D.L. Marchisio (2018), "Extended Charge-On-Particle Optimized Potentials for Liquid Simulation Acetone Model: The Case of Acetone–Water Mixtures", *The Journal of Physical Chemistry B*, 122, 20, pp. 5234-5241.

Lavino, A. D., N. Di Pasquale, P. Carbone, and D.L. Marchisio (2017), "A novel multiscale model for the simulation of polymer flash nano-precipitation", *Chemical Engineering Science*, 171, pp. 485-494.

Lee, H., A.H. de Vries, S.J. Marrink, and R.W. Pastor (2009), "A Coarse-Grained Model for Polyethylene Oxide and Polyethylene Glycol: Conformation and Hydrodynamics", *The Journal of Physical Chemistry B*, 113, 40, pp. 13186-13194.

Lee, H. and R.G. Larson (2006), "Molecular Dynamics Simulations of PAMAM Dendrimer-Induced Pore Formation in DPPC Bilayers with a Coarse-Grained Model", *The Journal of Physical Chemistry B*, 110, 37, pp. 18204-18211.

Lemkul, J. A., J. Huang, B. Roux, and Jr. MacKerell A. D. (2016), "An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications", *Chem. Rev.* 116, 4983–5013.

Lince, F., D.L. Marchisio, and A.A. Barresi (2011), "A comparative study for nanoparticle production with passive mixers via solvent-displacement: Use of CFD models for optimization and design", *Chem. Eng. Proc.* 50, 4, pp. 356-368.

— (2009), "Smart mixers and reactors for the production of pharmaceutical nanoparticles: Proof of concept", *Chem. Eng. Res. and Des.* 87, 4, pp. 543-549.

— (2008), "Strategies to control the particle size distribution of poly--caprolactone nanoparticles for pharmaceutical applications", *J. Colloid Interface Sci.* 322, p. 505.

Lipinski, C. A., F. Lombardo, B. W. Dominy, and P. J. Feeney (2001), "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings", *Advanced Drug Delivery*, 46, pp. 3-26.

Liu, Y., C. Cheng, R. K. Prud'homme, and R.O. Fox (2008), "Mixing in a multi-inlet vortex mixer (MIVM) for flash nano-precipitation", *Chem. Eng. Sci.* 63, p. 2829.

Liu, Y. and R.O. Fox (2006), "CFD predictions for chemical processing in a confined impinging-jets reactor", *AIChE Journal*, 52, 2, pp. 731-744.

Lopez, C.A., A.J. Rzepiela, A.H. de Vries, L. Dijkhuizen, P.H. Hunenberger, and S.J. Marrink (2009), "Martini Coarse-Grained. Force Field: Extension to Carbohydrates", *J. Chem. Theory Comput.* 5, 3195–3210.

Lyubartsev, A. and A. Laaksonen (1995), "Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach", *Physical Review E*, 52, p. 3730.

Madadi-Kandjani, E., R.O. Fox, and A. Passalacqua (2017), "Application of the Fokker-Planck molecular mixing model to turbulent scalar mixing using moment methods", *Physics of Fluids*, 29, 6, p. 065109.

Maeda, H. (2001), "SMANCS and polymer-conjugated macromolecular drugs: advantages in cancer chemotherapy", *Adv. Drug Delivery Rev.* 46, pp. 169-185.

Mahoney, M.W. and W.L. Jorgensen (2000), "A Five-Site Model for Liquid Water and The Reproduction of The Density Anomaly by Rigid, Nonpolarizable Potential Functions", *J. Chem. Phys.* 112, pp. 8910-8922.

Maple, J. R., U. Dinur, and A. T. Hagler (1988), "Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces", *Proceedings of the National Academy of Sciences*, 85, 15, pp. 5350-5354.

Marchisio, D.L. and R.O. Fox (2013), *Computational models for polydisperse particulate and multiphase flows*, Cambridge University Press.

— (2005), "Solution of population balance equations using the direct quadrature method of moments", *Journal of Aerosol Science*, 36, pp. 43-73.

Marchisio, D.L., F. Omegna, and A.A. Barresi (2009), "Production of TiO2 nanoparticles with controlled characteristics by means of a vortex reactor", *Chem. Eng. J.* 146, p. 456.

Marchisio, D.L., F. Omegna, A.A. Barresi, and P. Bowen (2008), "Effect of mixing and other operating parameters in sol-gel processes", *Industrial & Engineering Chemistry Research*, 47, p. 7202.

Marchisio, D.L., R. D. Vigil, and R.O. Fox (2003), "Quadrature method of moments for aggregation-breakage processes", *J. Colloid Int. Sci.* 258, pp. 322-334.

Marcus, Y. (1985), *Ion Solvation*, John Wiley & Sons Limited, New York.

Marrink, S.J., A.H. de Vries, and A.E. Mark (2004), "Coarse grained model for semiquantitative lipid simulations", *J. Phys. Chem. B*, 108, 750–760.

Marrink, S.J., H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. de Vries (2007), "The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations", *J. Phys. Chem. B*, 111, pp. 7812-7824.

Martyna, G. J., M. E. Tuckerman, D. J. Tobias, and M. L. Klein (1996), "Explicit reversible integrators for extended systems dynamics", *Mol. Phys.* 87, pp. 1117-1157.

McDowell, H. K. (2000), "Quantum generalized Langevin equation: explicit inclusion of nonlinear system dynamics", *J. Chem. Phys.* 112, pp. 6971-82.

McGraw, R. (2006), *Correcting moment sequences for errors associated with advective transport*, http://www.ecd.bnl.gov/pubs/momentcorrection mcgraw2006.pdf.

Michalowsky, J., J. Zeman, C. Holm, and J. Smiatek (2018), “A polarizable MARTINI model for monovalent ions in aqueous solution”, *The Journal of Chemical Physics*, 149, 16, p. 163319.

Moghimi, S. M., A. C. Hunter, and J. C. Murray (2001), “Long-circulating and target-specific nanoparticles: theory to practice”, *Pharm. Rev.* 53, pp. 283-318.

Monticelli, L., S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and S.J. Marrink (2008), “The MARTINI coarse-grained force field: Extension to proteins”, *J. Chem. Theory Comput.* 4, 819–834.

Mountain, R.D. (2010), “Microstructure and Hydrogen Bonding in Water–Acetonitrile Mixtures”, *J. Phys. Chem. B*, 114, pp. 16460-16464.

Muller-Plathe, F. (2002), “Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back”, *Chem. Phys. Chem.* 3, pp. 754-769.

Nielsen, S. O., C. F. Lopez, G. Srinivas, and M. L. Klein (2003), “A coarse grain model for n-alkanes parameterized from surface tension data”, *The Journal of Chemical Physics*, 119, pp. 7043-7049.

Noda, K., M. Ohashi, and K. Ishida (1982), “Viscosities and Densities at 298.15 K for Mixtures of Methanol, Acetone, and Water”, *J. Chem. Eng. Data*, 27, 326–328.

Noid, W. G., J. Chu, G. A. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen (2008), “The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models”, *Journal of Chemical Physics*, 128, p. 244114.

Nose, S. (1984), “A molecular dynamics method for simulations in the canonical ensemble”, *Mol. Phys.* 52, pp. 255-268.

Nose, S. and M. L. Klein (1983), “Constant pressure molecular dynamics for molecular systems”, *Mol. Phys.* 50, pp. 1055-1076.

Noskov, S. Y., G. Lamoureux, and B. Roux (2005), “Molecular dynamics study of hydration in ethanol-water mixtures using a polarizable force field”, *J. Phys. Chem. B*, 109, pp. 6705-6713.

Oosawa, F. and S. Asakura (1975), *Thermodynamics of the Polymerization of Protein*, Academic Press.

Oosawa, F. and M. Kasai (1962), “A theory of linear and helical aggregations of macromolecules”, *J. Mol. Biol.* 4, p. 10.

Parrinello, M. and A. Rahman (1981), “Polymorphic transitions in single crystals: A new molecular dynamics method”, *J. Appl. Phys.* 52, pp. 7182-7190.

Passalacqua, A., F. Laurent, E. Madadi-Kandjani, J. C. Heylmun, and R.O. Fox (2018), “An open-source quadrature-based population balance solver for OpenFOAM”, *Chemical Engineering Science*, 176, pp. 306-318.

Perera, A., R. Mazighi, and B. Kezic (2012), “Fluctuations and Micro-Heterogeneity in Aqueous Mixtures”, *J. Chem. Phys.* 136, p. 174516.

Perera, A. and F. Sokolić (2004), "Modeling Nonionic Aqueous Solutions: The Acetone-Water Mixture", *J. Chem. Phys.* 121, 11272–11282.

Pereyra, R.G., M.L. Asar, and M.A. Carignano (2011), "The Role of Acetone Dipole Moment in Acetone–Water Mixture", *Chem. Phys. Lett.* 507, 240–243.

Pinke, A. and P. Jedlovszky (2012), "Modeling of Mixing Acetone and Water: How Can Their Full Miscibility Be Reproduced in Computer Simulations?", *J. Phys. Chem. B*, 116, 5977–5984.

Pope, B. (2000), *Turbulent Flows*, Cambridge University Press.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992), *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 2nd edn., Cambridge: Cambridge University Press.

Raman, A.S., A. Vishnyakov, and Y. C. Chiew (2017), "A coarse-grained model for PCL: conformation, self-assembly of MePEG-b-PCL amphiphilic diblock copolymers", *Molecular Simulation*, 43, 2, pp. 92-101.

Ramkrishna, D. (2000), *Population Balances*, Academic Press.

Rick, S.W., S.J. Stuart, and B.J. Berne (1994), "Dynamical Fluctuating Charge Force Fields: Application to Liquid Water", *J. Chem. Phys.* 101, pp. 6141-6156.

Risken, H. and T. Frank (1989), *The Fokker-Planck Equation*, Springer-Verlag Berlin Heidelberg.

Rossi, Giulia, Luca Monticelli, Sakari R. Puisto, Ilpo Vattulainen, and Tapio Ala-Nissila (2011), "Coarse-graining polymers with the MARTINI force-field: polystyrene as a benchmark case", *Soft Matter*, 7 (2 2011), pp. 698-708.

Rubinstein, M and R H Colby (2003), *Polymer Physics*, Oxford Univ. Press.

Ryckaert, J.P. (1985), "Special geometrical constraints in the molecular dynamics of chain molecules", *Molecular Physics*, 55, 3, pp. 549-556.

Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen (1977), "Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes", *Journal of Computational Physics*, 23, 3, pp. 327-341.

Sack, R. A. and A. F. Donovan (1971), "An algorithm for Gaussian quadrature given modified moments", *Numerische Mathematik*, 18, pp. 465-478.

Saffman, P.G. and J.S. Turner (1956), "On the collision of drops in turbulent clouds", *Journal of Fluid Mechanics*, 1, pp. 16-30.

Shirts, M.R. and V.S. Pande (2005), "Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration", *The Journal of Chemical Physics*, 122, 14, p. 144107.

Skyner, R. E., J. L. McDonagh, C. R. Groom, T. van Mourik, and J. B. O. Mitchell (2015), "A review of methods for the calculation of solution free energies and the modelling of systems in solution", *Phys. Chem. Chem. Phys.* 17 (9 2015), pp. 6174-6191.

Song, Y. and L.L. Dai (2010), "The Shear Viscosities of Common Water Models by Non-Equilibrium Molecular Dynamics Simulations", *Mol. Simul.* 36, 560–567.

Straatsma, T.P. and J.A. McCammon (1990), "Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water", *Mol. Simul.* 5, pp. 181-192.

Taddese, T. and P. Carbone (2017), "Effect of Chain Length on the Partition Properties of Poly(ethylene oxide): Comparison between MARTINI Coarse-Grained and Atomistic Models", *The Journal of Physical Chemistry B*, 121, 7, pp. 1601-1609.

Taddese, T., D. L. Cheung, and P. Carbone (2015), "Scaling Behavior of Polymers at Liquid/Liquid Interfaces", *ACS Macro Lett.* 4, 1089–1093.

Thole, B. T. (1981), "Molecular Polarizabilities Calculated with a Modified Dipole Interaction", *Chem. Phys.* 59, 341–350.

Thomas, K.T. and R.A. McAllister (1957), "Densities of Liquid-Acetone-Water Solutions up to Their Normal Boiling Points", *A.I.Ch.E. Journal*, 3, 161–164.

Torchilin, V. P. (2007), "Targeted Pharmaceutical Nanocarriers for Cancer Therapy and Imaging", *The AAPS Journal,* 9, E128-47.

Toryanik, A.I. and V.N. Taranenko (1987), "Molecular Mobility and Structure in Water-Acetone Mixtures", *J. Struct. Chem.* 28, 714–719.

Uusitalo, J.J., H.I. Ingolfsson, P. Akhshi, D.P. Tieleman, and S.J. Marrink (2015), "Martini Coarse-Grained Force Field: Extension to DNA", *J. Chem. Theory Comput.* 11, 3932–3945.

Valente, I., E. Celasco, D.L. Marchisio, and A.A. Barresi (2012), "Nanoprecipitation in confined impinging jets mixers: Production, characterization and scale-up of pegylated nanospheres and nanocapsules for pharmaceutical use", *Chemical Engineering Science*, 77, pp. 217-227.

Valente, I., B. Stella, D.L. Marchisio, F. Dosio, and A.A. Barresi (2012), "Production of PEGylated nanocapsules through solvent-displacement in confined impinging jets mixers", *J. Pharm. Science*, 101, pp. 2490-2501.

Van Belle, D., I. Couplet, M. Prevost, and S.J. Wodak (1987), "Calculations of Electrostatic Properties in Proteins. Analysis of Contributions from Induced Protein Dipoles", *J. Mol. Biol.* 198, pp. 721-735.

Van Gunsteren, W.F. and H.J.C. Berendsen (1988), "A leap-frog algorithm for stochastic dynamics", *Mol. Sim.* 1, pp. 173-185.

— (1977), "Algorithms for macromolecular dynamics and constraint dynamics", *Mol. Phys.* 34, pp. 1311-1327.

Van Kampen, N. G. (1981), *Stochastic Processes in Physics and Chemistry*, North Holland.

Vesely, F.J. (1977), "N-particle Dynamics of Polarizable Stockmayer-type Molecules", *J. Comput. Phys.* 24, pp. 361-371.

Von Smoluchowski, M. (1917), "Versuch einer mathematischen Theorie der Koagulationskinetik kolloidaler Losungen", *Zeitschrift fur Physikalische Chemie*, 92, pp. 129-168.

Warshel, A. and M. Levitt (1976), "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme", *J. Mol. Biol.* 103, pp. 227-249.

Weerasinghe, S. and P.E. Smith (2003), “Kirkwood–Buff Derived Force Field for Mixtures of Acetone and Water”, *J. Chem. Phys.* 118, pp. 10663-10670.

Wheeler, J. C. (1974), “Modified moments and Gaussian quadratures”, *Rocky Mountain Journal of Mathematics*, 4, pp. 287-296.

Who, C., T. Jim, Z. Gan, Y. Zhao, and S. Wang (2000), “A heterogeneous catalytic kinetics for enzymatic biodegradation of poly(ε-caprolactone) nanoparticles in aqueous solution”, *Polymer*, 41, pp. 3593-3597.

Wilf, H. S. (1962), *Mathematics for the Physical Sciences*, New York (NY): John Wiley and Sons.

Wright, D. L. (2007), “Numerical advection of moments of the particle size distribution in Eulerian models”, *Journal of Aerosol Science*, 38, pp. 352-369.

Wright, D. L., R. McGraw, and D. E. Rosner (2001), “Bivariate extension of the quadrature method of moments for modeling simultaneous coagulation and sintering particle populations”, *Journal of Colloid and Interface Science*, 236, pp. 242-251.

Wu, D. and D.A. Kofke (2005), “Phase-space overlap measures. II. Design and implementation of staging methods for free-energy calculations”, *The Journal of Chemical Physics*, 123, 8, p. 084109.

Yoon, C. and R. McGraw (2004), “Representation of generally mixed multivariate aerosols by the quadrature method of moments: I. Statistical foundation”, *Journal of Aerosol Science*, 35, pp. 561-576.

Zelenková, T., M.J. Mora, A.A. Barresi, G.E. Granero, and D. Fissore (2018), “On the Production of Chitosan-Coated Polycaprolactone Nanoparticles in a Confined Impinging Jet Reactor”, *Journal of Pharmaceutical Sciences*, 107, 4, pp. 1157-1166.

Zhou, H. X. and F. A. Ferrone (1990), “Theoretical description of the spatial dependence of sickle hemoglobin polymerization”, *Biophys. J.* 58, p. 695.

This Ph.D. thesis has been typeset by means of the TeX-system facilities. The typesetting engine was LuaLaTeX. The document class was `toptesi`, by Claudio Beccari, with option `tipotesi=scudo`. This class is available in every up-to-date and complete TeX-system installation.