



TECHNISCHE  
UNIVERSITÄT  
WIEN  
Vienna University of Technology

## Dissertation

---

# In- and out-of-equilibrium simulations of polymeric molecules at the monomeric and the coarse-grained level

---

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin der  
Naturwissenschaften unter der Leitung von

Ao. Univ. Prof. Dr. Gerhard Kahl  
Institut für Theoretische Physik (E136)  
Technische Universität Wien

eingereicht an der Technischen Universität Wien  
Fakultät für Physik

von

Marta Montes Saralegui  
Matrikelnummer 1027395  
Franzengasse 10/6, 1050 Wien

Wien, im November 2014

---

Marta Montes Saralegui



## Kurzfassung

Mittelt man bei polymerischen Makromolekülen mit Hilfe geeigneter Methoden der Statistischen Physik über die Freiheitsgrade der monomerischen Einheiten, so erhält man sogenannte effektive Potentiale, die bei kleinen Molekülabständen endliche Werte (in der Größenordnung thermischer Energien) annehmen. Dies bedeutet, dass die Schwerpunkte dieser Teilchen mit einem relativ geringem Energiepönale einander beliebig nahe kommen können, ohne dass es zu einem Überlapp der Monomere kommt. In der Literatur werden derartige Systeme als ultraweich bezeichnet. Unter gewissen Bedingungen können derartige Polymersysteme sogenannte Cluster-Kristalle bilden: die Positionen eines periodischen Gitters sind hier von Clustern besetzt, die aus einander überlappenden Makromolekülen bestehen. Die Stabilisierung dieser Aggregate von einander abstoßenden Teilchen erfolgt dadurch, dass die benachbarten Cluster auf die Moleküle eine noch viel stärkere Repulsion ausüben und somit diese Teilchen in ihren Cluster zurückdrängen. Bisherige Untersuchungen haben gezeigt, dass diese Cluster-Kristalle sehr bemerkenswerte Eigenschaften aufweisen, wie etwa einen Gitterabstand, der trotz einer Erhöhung der Dichte konstant bleibt oder charakteristische "hopping"-Prozesse, über die sich die Teilchen von einem Cluster zum Nächsten durch den Kristall bewegen.

Diese Arbeit ist Untersuchungen von dynamischen Eigenschaften dieser Clusterbildenden Systeme gewidmet. Auf dem Niveau *effektiver* Wechselwirkungen untersuchen wir den Einfluss hydrodynamischer Kräfte, die durch das (mikroskopische) Lösungsmittel verursacht werden, auf die "hopping"-Prozesse in Cluster-Kristallen. Weiters haben wir mit Hilfe von Computersimulationen die Teilcentrajektorien berechnet, wenn man einen Cluster-Kristall einem externen Druck aussetzt: individuelle Teilchenbewegungen und Cluster-Bildung konnten damit eingehend untersucht werden. Auf dem Niveau einer *monomerischen* Beschreibung der Makromoleküle haben wir *in silico* Modelle für amphiphile Polymerketten entwickelt, die in der kondensierten Phase Cluster-Kristalle bilden können: diese Ketten bestehen aus einer solvophoben Hauptkette, die mit solvophilen Monomeren dekoriert ist. Den Abschluss dieser Arbeit bildet die Implementierung einer numerischen Methode, die die Berechnungen effektiver Wechselwirkungen polymerischer Makromoleküle ermöglicht: da es diese Methode erlaubt, die effektiven (Paar-)Potentiale bei endlichen Dichten zu berechnen, können die Eigenschaften der Systeme über diese Wechselwirkungen verlässlicher beschrieben werden.



## Abstract

Ultrasoft polymeric macromolecules show bounded effective interactions, which arise as the internal degrees of freedom of the molecule are averaged out. This means that they can fully overlap their centres of mass without violation of the excluded volume at the monomer level. Under certain conditions, these systems form *cluster crystals*: crystals whose lattice sites are occupied by clusters of partially or fully overlapping particles. Even though the macromolecules within a cluster repel each other, the cluster as a whole is stabilised by the repulsion of the neighbouring clusters. This counter-intuitive, but meanwhile well understood, behaviour leads to a novel class of materials which shows remarkable properties, as for example a density independent lattice constant or hopping mechanisms, where particles diffuse hopping from one cluster to another.

In this thesis we study certain dynamic properties of these materials on the coarse-grained level: firstly we examine the effect of hydrodynamic interactions (due to the presence of a solvent) on the diffusion and hopping mechanism in pure and binary cluster crystals. Secondly we study the response of a pure cluster crystal to an external compression, paying special attention to the mechanisms through which the crystal can accommodate a volume reduction while keeping the spacing of the lattice unchanged. We also present an *in silico* design of a polymer amphiphilic chain, whose effective interaction shows the necessary properties for the formation of stable cluster crystals. The amphiphilic chains are composed by a solvophobic backbone decorated by solvophilic side groups. We perform monomer resolved simulations of a bulk crystal of these chains and verify the stability of the system. Finally we focus our attention on the computation of effective interactions of polymeric macromolecules. We make use of a method, which allows for the computation of the effective potentials at finite density leading to a more reliable description of the behaviour of the system.



# Contents

<b>1. Introduction</b>	<b>3</b>
<b>2. Model</b>	<b>9</b>
2.1. Introduction . . . . .	9
2.2. Ultrasoft potentials . . . . .	11
2.2.1. $Q^+$ and $Q^\pm$ potentials . . . . .	11
2.3. Cluster crystals . . . . .	14
2.3.1. Bulk modulus . . . . .	18
2.3.2. Dynamics in cluster crystals . . . . .	19
2.4. Generalised exponential model . . . . .	20
2.5. Dendrimers . . . . .	22
2.5.1. History of dendrimers . . . . .	22
2.5.2. Modelling dendrimers . . . . .	23
2.5.3. Cluster forming amphiphilic dendrimers . . . . .	25
<b>3. Methods</b>	<b>29</b>
3.1. Simulation techniques . . . . .	29
3.1.1. Molecular dynamics . . . . .	29
3.1.2. Multi particle collision dynamics . . . . .	31
3.2. Computation of effective interactions:	
Multi-scale coarse graining method . . . . .	35
3.2.1. General formulation of the MSCG method . . . . .	35
3.2.2. Effective pair interactions . . . . .	41
<b>4. Diffusion in cluster crystals in the explicit presence of a solvent</b>	<b>43</b>
4.1. Introduction . . . . .	43
4.2. Hopping processes in a pure system . . . . .	43
4.2.1. Mean-squared displacement . . . . .	45
4.2.2. Jump length distribution . . . . .	47
4.2.3. Distribution of the angle enclosed by successive jumps . . . . .	49
4.2.4. Self-part of the van Hove correlation function . . . . .	53
4.3. Hopping processes in a binary system . . . . .	54
4.3.1. Mean-squared displacement . . . . .	55
4.3.2. Self-part of the van Hove correlation function . . . . .	56
<b>5. Compression experiments on a two dimensional GEM-4 system</b>	<b>59</b>
5.1. Introduction . . . . .	59
5.2. Method . . . . .	62
5.2.1. Compression/expansion or heating/cooling protocol . . . . .	65
5.3. Results . . . . .	67
5.3.1. About the compression rate . . . . .	67
5.3.2. Compression experiments . . . . .	73
5.3.3. Equation of state . . . . .	87

<b>6. Amphiphilic, cluster-forming polymer chains</b>	<b>95</b>
6.1. Introduction . . . . .	95
6.2. Model . . . . .	97
6.3. Effective interactions . . . . .	100
6.3.1. ABC chains . . . . .	100
6.3.2. AB chains . . . . .	105
6.4. Monomer resolved and coarse grained simulations of AB chains . . . . .	107
6.4.1. Simulations in the fluid state . . . . .	107
6.4.2. Simulations in the crystal state . . . . .	109
<b>7. Effective interactions of polymeric macromolecules computed at finite densities</b>	<b>113</b>
7.1. Introduction . . . . .	113
7.2. Effective interactions at finite densities . . . . .	115
7.2.1. Non-clustering systems . . . . .	115
7.2.2. Clustering systems . . . . .	120
7.3. Effective interaction of a single dendrimer with a cluster . . . . .	129
<b>8. Conclusions &amp; Outlook</b>	<b>137</b>
<b>Appendices</b>	<b>143</b>
<b>A. Reduced units</b>	<b>143</b>
<b>B. Cluster Analysis</b>	<b>145</b>
<b>Bibliography</b>	<b>147</b>

# 1. Introduction

Soft Matter science is an interdisciplinary scientific field where physics, chemistry and biology meet. Materials which are *soft* exhibit, in contrast to their hard counterparts, such as atomic crystals, large responses to external mechanical forces. This effect can for instance be quantified by the shear modulus  $G$ , which measures the resistance of an object to a mechanical deformation.  $G$  is proportional to the ratio between the system's characteristic interaction energy,  $\epsilon$ , and the cube of the characteristic length,  $\sigma^3$ , *i.e.*  $G \sim \epsilon/\sigma^3$  [1]. In hard materials,  $\epsilon$  is of the order of  $10^{-18}$ J (corresponding to the bonding energy) [2] and  $\sigma$  is of the order of  $10^{-10}$ m (corresponding to the typical bond length) [2]; therefore  $G$  is of the order of  $10^{12}$ N/m<sup>2</sup>. In contrast, soft materials are composed of mesoscopic particles ( $\sigma \sim 10^{-9}$  to  $10^{-6}$ m) embedded in a microscopic solvent and the interactions governing the system are usually van der Waals interactions or hydrogen bonds; consequently  $\epsilon \sim 10^{-20}$ J [2, 3], and  $G \sim 10^7$  to  $10^{-2}$ N/m<sup>2</sup>. Thus the difference between the value of  $G$  in soft and hard materials is of 5 to 14 orders of magnitude. Soft materials are not only easily deformable, but their characteristic interaction energy is of the same order of magnitude as  $k_B T \simeq 4 \times 10^{-21}$ J, thus even thermal fluctuations can induce structural changes in these systems. Soft materials can be found in our daily life: blood, milk, ink, paint and protein solutions are only a few examples.

In this thesis we focus on the study of a particular class of soft matter systems, the so-called *ultrasoft* systems. The interactions between the ultrasoft particles are described by potentials which are bounded at a particle separation of  $r = 0$ , meaning that ultrasoft particles can overlap and the penalty for such a configuration is typically of the order of a few  $k_B T$ . These potentials arise naturally as effective interactions between the centres of mass of polymeric macromolecules, which are composed of hundreds to thousands of monomers. The large number of degrees of freedom of the monomers can be averaged out by suitable methods; these macromolecules can then be described as coarse grained points particles placed in their centres of mass, interacting *via* the effective particles. If these polymers can overlap their centres of mass without violating the excluded volume in the monomer level, then the effective interaction will be bounded, *i.e.* the system can be classified as ultrasoft. Polymer chains [4], dendrimers [5], ring polymers [6] and polyelectrolytes [7] are a few examples whose effective interactions are bounded.

The first ultrasoft systems to be studied theoretically were the Gaussian core model (GCM) by Stillinger *et al.* [8] and the penetrable sphere model (PSM) [9, 10]. Stillinger *et al.* found that the phase diagram of the CGM exhibits *re-entrant* melting [8]: below a certain freezing temperature  $T_f$ , the system freezes on increasing the density at constant temperature into a crystal; upon further compression it re-melts back to the fluid state. At temperatures above  $T_f$ , the system remains fluid for all densities. On the contrary, the PSM shows a very

different behaviour [10]: it shows *clustering*. At low and intermediate densities we observe a fluid phase of disordered clusters of overlapping particles; upon further compression, the system freezes for all temperatures into a cluster crystal, where the lattice positions of a regular bcc or fcc lattice are occupied by clusters of fully or partially overlapping particles. Although the particles inside a cluster repel each other, the mutual repulsion of neighbouring clusters stabilises the cluster crystal. The interest in the field of ultrasoft systems rapidly grew as Likos *et al.* presented in reference [11] a criterion that determines whether an ultrasoft system will show re-entrant melting or clustering: if the Fourier transform of the effective potential has negative components, as in the PSM, the system will freeze into a cluster; if the Fourier transform of the potential decays monotonically to zero, as in the GCM, the system will show re-entrant melting.

Part of the work presented in this thesis is dedicated to the computation of reliable effective interaction potentials of polymeric macromolecules which make it possible to predict if the system will freeze into a cluster crystal or not. Methods used up to now [12, 13] compute the effective pair interactions in the zero-density limit: by considering two isolated macromolecules. However, in states of high density many-body effects, which are not included in the zero-density effective pair interactions, become relevant. An example of a system where these effects play a crucial role are flexible polymer rings [14]: the zero-density effective interaction has negative Fourier components, however due to many-body effects related to the deformation of the rings, clustering is not observed at high densities.

In our work we apply a technique for the computation of effective interactions, the so-called multi-scale coarse graining method (MSCG) [15], which to our knowledge has yet not been used in the field of ultrasoft systems and allows for the calculation of effective pair interactions at finite densities. We show that the obtained potentials offer a much more reliable description of the high density states than the zero-density effective potentials. The agreement between pair correlation functions obtained in monomer resolved (MR) and coarse-grained (CG) simulations is significantly better when the CG simulations are carried out with the effective potentials computed at finite densities instead of those computed at zero-density. However, we find that finite density effective potentials are not able to properly describe the system if it forms clusters of overlapping particles; this is due to the fact that in such a case the local density is strongly inhomogeneous.

Re-entrant melting has been theoretically and experimentally observed: studies of re-entrant melting have been carried out in simulations at the coarse grained description level of the GCM [8], and at the monomer resolved description level of star polymers [16], microgels [7] and diblock copolymers [17]; and a series of small angle X-ray scattering (SAXS) experiments have been performed with copolymer micelles [18–20]. In contrast, cluster forming systems have not been identified in experiments so far. Theoretical studies of these systems have been carried out in coarse grained simulations of the generalised exponential model, GEM, [21–26] and the PSM [9, 10] and only very recently in monomer resolved simulations of amphiphilic dendrimers [27, 28].

---

We therefore dedicate a second part of our work to the *in silico* design of a better polymeric structure which freezes into a cluster crystal and aims to be easy to synthesise in a lab. We were inspired by the work of Meijer *et al.* [29], who synthesised a series of polymer chains built up with blocks of hydrophilic poly(ethylene glycol) and hydrophobic chiral benzene-1,3,5-tricarboxamide substituents which carry separated spatially catalytic units. We build on the computer a model of amphiphilic linear chains composed of solvophilic and solvophobic monomers and come up with a combination of the parameters in the model which results in negative components of the Fourier transform of the effective potential. We perform monomer resolved simulations in the cluster crystal phase and verify the stability of the system.

The rest of the work is dedicated to the study of certain properties of cluster crystals, carried out in coarse grained simulations of GEM systems. Previous studies have shown that cluster crystals have very novel and remarkable properties in comparison to atomic crystals:

- in 2006 Mladek *et al.* performed the first simulations of GEM cluster crystals at various densities. The authors found that the number of particles participating in a cluster crystal scales linearly with density, therefore cluster crystals feature **density-independent lattice constants** constants [21].
- In 2007 Likos *et al.* performed a series of density functional theory (DFT) calculations in the mean field approximation (MFA) for cluster-forming systems. The authors derived expressions for the location of the **freezing line** in the temperature and density plane and the value of the density independent **lattice constant in terms of position and value of the minimum of the Fourier transform of the interaction potential** [23].
- In the same year Mladek *et al.* performed calculations of the free energy of a cluster-forming GEM-4 system where the number of lattice sites ( $N_c$ ) was treated as a constraining thermodynamic variable [22]. The authors computed the bulk modulus ( $B$ ) as a derivative of the free energy, which now has two contributions: the term which accounts for the response that one intuitively expects: the affine shrinking of  $a$  when a crystal is compressed and a term arising from the fact that  $N_c$  is now a thermodynamic variable. The additional term accounts for the extra degree of freedom unique to cluster crystals: upon compression some lattice positions can be deleted and the particles which were populating the clusters in these lattice positions will re-arrange in other clusters of the system. It was shown that the additional term results in a **reduction of about 40% of the term due to affine shrinking**.

The phase diagrams obtained from the studies mentioned above in Monte Carlo simulations [21], DFT calculations in the MFA [23] and free energy calculations [22] show a very good agreement. Further studies showed that:

- another remarkable property of cluster crystals which was studied by Moreno *et al.* in 2007 [24] and Coslovich *et al.* in 2010 [25] is the presence **hopping mechanisms** in

these systems: clusters in the crystal show an incessant exchange of particles, which migrate from one cluster to another, realising often very long trajectories, and visiting thereby many intermediate clusters.

- The study of hopping processes in cluster crystals was extended by Camargo *et al.* in 2010 [30] to binary mixtures of cluster-forming GEM particles and non-clustering Gaussian particles.
- In 2010 Nikoubashman *et al.* [31] discovered a **unique response mechanism of cluster crystals to shear**: the crystal forms under external shear an ordered pattern of long strings oriented parallel to the flow and subsequently melts into a fluid at high shear rates.

We have contributed to these investigations in two directions: we have studied the influence of hydrodynamic interactions (HI), induced by the presence of a solvent, on the hopping activity of a pure and on a binary cluster crystal. The microscopic solvent, in which soft mesoscopic particles are embedded, is included in the coarse grained effective interaction between the solute particles and is not simulated explicitly. This approach is valid for studies of static quantities, such as static correlation functions. However, in investigations of the dynamic properties the solvent often plays a significant role and should not (or even must not) be neglected in a faithful description of the system. The relevance of the solvent for the dynamics has been shown for charged colloidal suspensions [32] and for short polymer chains [33]. To address the effect of HI we perform extensive simulations using the multi-particle collision dynamics method [34, 35], which explicitly includes the solvent and takes hydrodynamic interactions into account as faithfully as possible. We tuned the coupling between the solute and the solvent covering the range from no coupling (neglect of the solvent) up to strong coupling. We observe that the presence of the solvent hinders the hopping activity, and that the amount of particles involved in jumps and the lengths of their trajectories dramatically decrease with stronger coupling between the solute and the solvent.

Secondly, we study the response of a pure cluster crystal to an external compression. To address this question, we surround a system of cluster-forming GEM particles with an ensemble of ideal gas particles which exert a strong repulsion on the former. By adjusting the number of ideal gas particles and the distribution of their velocities we can tune the temperature and pressure in the GEM system [36]. We first perform a series of compression experiments at a target temperature: we observe that the cluster occupation increases linearly with the density and the spacing of the lattice,  $a$ , is almost unaffected by the compression, however, the value of  $a$  is below the equilibrium value,  $a_{\text{eq}}$ . The necessary particle and cluster rearrangements to attain  $a_{\text{eq}}$  entail a big energetic barrier between the current compressed configuration and a configuration with  $a = a_{\text{eq}}$ , which the system is not able to overcome at constant temperature. We observe that a cluster crystal reacts to compression by a increased hopping activity, which leads to a heterogeneous occupation of the clusters. Then, smaller clusters, which feel a strong repulsion of their bigger neighbours, are pushed together

---

to merge; thanks to the volume released the lattice constant can remain unchanged. In a second series of experiments we combine compression runs with annealing processes, where the temperature of the GEM system is raised at constant pressure and then lowered again to the original starting value. *Via* these experiments we are able to understand that there is indeed an optimal equilibrium value for the spacing of the lattice, which the system is not able to attain during a simple compression experiment. The introduction of the annealing processes helps the cluster crystal overcome the energetic barrier separating the compressed and the equilibrium state. With this set-up we are able to extract the equation of state of the system. Finally, we combine compression and expansion experiments and we observe that the first reaction of the cluster crystal is to contract or expand the lattice constant within a range of 10% of  $a_{\text{eq}}$  around  $a_{\text{eq}}$ . Only once this lower or upper value of the lattice constant is achieved will the cluster merging (in a compression run) or cluster splitting (in an expansion run) take place.

This thesis is organised as follows: in **chapter 2** we present the different model systems studied throughout the thesis and in **chapter 3** we present the used simulation techniques and the MSCG method, used for the computation of effective interactions. In **chapter 4** we present the results on the effect of hydrodynamic interactions on the hopping activity in a pure and a binary cluster crystal. In **chapter 5** we address the reaction of a pure cluster crystal to compression. In **chapter 6** we present our designed model of amphiphilic cluster-forming polymer chains. In **chapter 7** we present our results on the computation of effective interactions of polymeric macromolecules at finite densities. Finally, in **chapter 8** we present the conclusions and outlook of our work. In appendix A we present the reduced units used throughout the thesis and in appendix B we give a detailed explanation of the cluster analysis, a tool used in chapters 4 to 7.



# 2. Model

## 2.1. Introduction

This thesis is focused on the *cluster crystal* phase [21]: for certain ultrasoft particles (like very hollow polymeric colloids) the most stable phase at high densities is a crystal whose lattice sites are occupied by clusters of fully or partially overlapping mesoscopic particles. Even though these particles might repel each other, the repulsion of the neighbouring clusters stabilises the aggregate of overlapping particles [37]. The hollow structure of these colloids allows an overlap of the centres of mass of the polymers without violation of the excluded volume at the monomeric level (see figure 2.1) [38].

This exotic kind of crystals have very interesting properties. The spacing of the lattice has a density independent value [23]: when the density of the crystal increases, the occupation of the clusters increases, leaving the spacing of the lattice unchanged. The multiple occupancy of the crystal positions allows for activated hopping processes[24]: when the temperature of the crystal is high enough, thermal fluctuations enable some particles to leave their host cluster and diffuse from cluster to cluster without breaking the lattice structure.

Systems forming these kinds of crystals can be studied in computer simulations at two different levels of description: within a *monomer resolved* (MR) or within a *coarse grained* (CG) representation. In the MR representation each monomer is considered as an individual microscopic particle. The monomers interact *via* excluded volume (*e. g.* Morse, Lennard-Jones, etc ...) and bonding interactions (*e.g.* FENE, harmonic potential, etc ...). The MR description can easily be related to the real macromolecule [39]: parameters of the excluded volume interaction potential represent the size, softness and interaction of the monomers with the solvent; parameters of the bonding potential represent the bond length and stiffness. A MR representation gives a very detailed and precise description of the behaviour of polymers. However, the number of particles to be simulated might grow too quickly. Polymers can be composed of up to thousands of monomers ( $N_m \sim 10^2 - 10^3$ , being  $N_m$  the number of monomers per polymer). For a reliable study one might need to simulate an ensemble of hundreds or even thousands of polymers. The computing time needed to simulate such a system will then be unaffordable.

In such situations one might rather use a CG representation. The degrees of freedom of the monomers are in general of little relevance for our study: for example bond vibrations have a characteristic time scale which is orders of magnitude smaller than the relaxation time of the system. Such degrees of freedom are averaged out and each polymer is then represented by an effective particle positioned on its centre of mass (COM). The number of

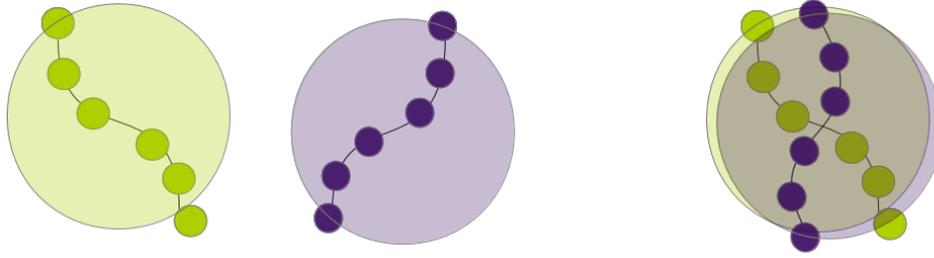


Figure 2.1.: Overlapping polymers. The centres of mass overlap without violating excluded volume in the monomer level.

particles to be simulated is now  $N_m$  times smaller. Such a representation is of course very advantageous: one can simulate bigger systems over longer time windows. However, it is not trivial to compute the functional form of an effective potential which *properly* describes the interaction between the effective particles (the so-called *effective interaction*). Both representations should yield faithful results: the thermodynamics (or partition function) of the system in the MR description level have to be reproduced by the CG representation. The computation of these interactions involves a tedious procedure and often transferability problems arise [40]: effective potentials computed at certain  $P$ ,  $T$  and  $\rho$  conditions are not valid at different conditions.

In this work we approach the study of cluster crystals in two manners: first we focus on details about certain properties of the crystals in the coarse grained level. In chapter 4 we study how the the presence of a solvent (and the consequent hydrodynamic interactions) affect the hopping processes in a cluster crystal and in chapter 5 we examine the reaction of a cluster nanocrystal to an external compression. Secondly we investigate in chapter 6 polymeric structures which would crystallise into a cluster crystal and finally in chapter 7 we exhaustively examine the validity of coarse grained representations by studying the state-dependence of effective interactions in different kinds of polymeric colloids. In this second part we combine MR and CG representations and validate their consistency.

## 2.2. Ultrasoft potentials

We define the energy scale  $\epsilon$  and the length scale  $\sigma$  of a potential as the quantities which allow us to express the potential as:

$$\phi(r) = \epsilon v(r/\sigma) \quad (2.1)$$

where  $v(r)$  is a dimensionless function of a dimensionless variable and  $r$  denotes the distance between two particles. *Ultrasoft potentials* are potentials whose energy scale  $\epsilon$  is of the order of  $k_B T$  ( $k_B$  is the Boltzmann constant) and the value of  $\phi(r)$  at a particle separation  $r = 0$  is bounded and then decays to zero for  $r \rightarrow \infty$ . The finite value of  $\phi(r)$  at  $r = 0$  is in strong contrast to diverging atomic potentials (like Lennard-Jones). This property might seem unphysical upon first sight, but such potentials have to be understood as effective interactions between polymeric macromolecules and then one easily understands that the value of the potential at  $r = 0$  corresponds to the (finite) energy penalty of a full overlap as shown in figure 2.1.

### 2.2.1. $Q^+$ and $Q^\pm$ potentials

Imagine a fluid with a spatially modulated density profile  $\rho(\mathbf{r})$  which does not vary too rapidly on the scale of  $\sigma$  set by the interaction (see equation (2.1)). At high densities,  $\rho\sigma^3 > 1$ , the average interparticle separation becomes vanishingly small, and it holds that the potential is extremely long ranged [11]. Particles interact with an enormous number of neighbours and in the absence of short range excluded volume interactions a mean field approximation (MFA) framework is justified [11]. The structure factor of the uniform fluid takes then the form [41]:

$$S(k) = \frac{1}{1 + \beta\rho\tilde{\phi}(k)} \quad (2.2)$$

where  $\beta = 1/(k_B T)$  is the inverse temperature and  $\tilde{\phi}(k)$  is the Fourier transform of the interaction potential  $\phi(r)$ .

If the Fourier transform of the potential attains negative values for some finite  $k$  there will be a line in the  $(\rho, T)$ -plane where the structure factor diverges, the so-called ‘‘spinodal line’’. This leads to a Kirkwood instability [42] of the uniform liquid, implying a crystallization of the system. Crystallization will occur at all temperatures and the spacing of the lattice will be dictated by the  $k$  value at which  $\tilde{\phi}(k)$  attains its minimum value [23], independently of the density. For the lattice constant to be independent of  $\rho$  there is only one solution: particles have to form clusters on every lattice site, *i.e.* a cluster crystal has to be formed.

Based on this ideas Likos *et al.* [11] established a criterion, to classify ultrasoft potentials into two classes ( $Q^+$ ,  $Q^\pm$ ), which determines the behaviour of the phase diagram, *i.e.* it

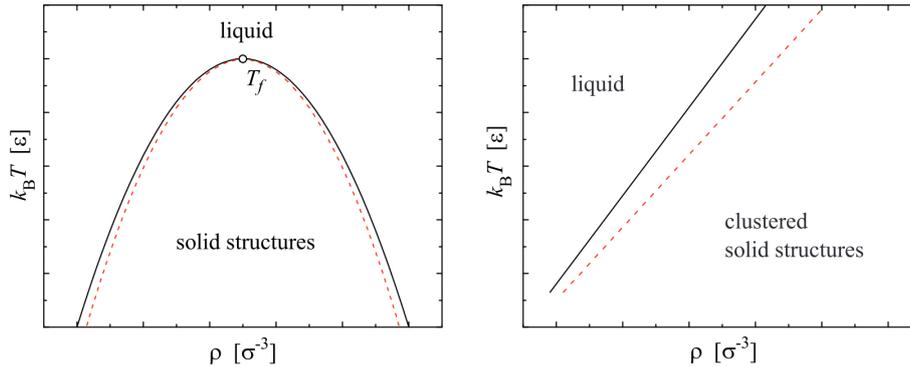


Figure 2.2.: Schematic representation of the phase diagram in the  $(\rho, T)$ -plane of systems where particles interact *via*  $Q^+$  (left) and  $Q^\pm$  (right) potentials. The lines denote the phase boundary of the solid (dashed red) and the fluid (solid black) phase. Between these lines is the coexistence region.  $T_f$  denotes the freezing temperature below which re-entrant melting takes place. Taken from [43].

determines whether the cluster crystal phase is the most stable phase for high densities or not.

### $Q^+$ potentials

Potentials whose Fourier transform monotonically decays to zero belong to the  $Q^+$  class. Upon compression at constant temperature these systems will crystallise into a single occupancy crystal and upon further compression the crystal will melt into a fluid (see left panel of figure 2.2). This is called *re-entrant melting*. Re-entrant melting was first observed by Stillinger in the 1970s [8] in simulations of the Gaussian core model (GCM) and confirmed by Lang *et al.* by means of integral theory. The potential of the GCM is given by:

$$\phi(r) = \epsilon e^{-(r/\sigma)^2}. \quad (2.3)$$

Other examples of soft matter systems for which re-entrant melting has been observed in simulations are star polymers [16], microgels [7] and diblock copolymers [17]. Re-entrant melting was also observed in a series of small angle X-ray scattering (SAXS) experiments performed with copolymer micelles [18–20]

### $Q^\pm$ potentials

Potentials whose Fourier transform has an oscillatory behaviour in  $k$ -space, and therefore can have some negative components, belong to the  $Q^\pm$  class. Upon compression, these systems always crystallise at sufficiently high densities and/or sufficiently low temperatures into a crystal of multiple occupancy (see right panel of figure 2.2). So far clustering behaviour has been observed in coarse grained simulations of systems where particles interact *via* the

generalised exponential model (GEM- $n$ ) [21–23, 25, 26]:

$$\phi_n(r) = \epsilon \exp [-(r/\sigma)^n]. \quad (2.4)$$

These potentials have a negative component in the Fourier transform (and therefore form cluster crystals at sufficiently high densities or low temperatures) if  $n > 2$ . GEM- $n$  systems will be further discussed in section 2.4. Only very recently the formation of cluster crystals has been observed in monomer resolved simulations of amphiphilic dendrimers [28]. An experimental verification of the clustering behaviour is still missing.

## 2.3. Cluster crystals

### Stability of the cluster crystal phase

Mladek *et al.* computed in reference [37] the free energy of a system of ultrasoft particles in the cluster crystal phase in the following way: the free energy of the system is expressed in the MFA as the sum of ideal and the excess contribution:

$$F[\rho] = F_{\text{id}}[\rho] + F_{\text{ex}}[\rho] = k_B T \int d^3r \rho(\mathbf{r}) [\ln [\rho(\mathbf{r})\Lambda^3] - 1] + \frac{1}{2} \iint d^3r d^3r' \phi(|\mathbf{r} - \mathbf{r}'|) \rho(\mathbf{r}) \rho(\mathbf{r}') \quad (2.5)$$

where  $\Lambda$  is the de Broglie wavelength. The one-particle density  $\rho(r)$  in the crystalline phase can be expressed as a sum over Gaussians, centred at the Bravais lattice sites  $\{\mathbf{R}\}$  of the ordered phase. However, one has to take into account that now the number of lattice sites  $N_c$  is not equal to the number of particles  $N$  due to the possible multiple occupancy; thus  $\rho(r)$  has to be normalised by the cluster occupation  $n_{\text{occ}} = N/N_c$ :

$$\rho(\mathbf{r}) = n_{\text{occ}} \left(\frac{\alpha}{\pi}\right)^{3/2} \sum_{\{\mathbf{R}\}} e^{-\alpha(\mathbf{r}-\mathbf{R})^2} = \sum_{\{\mathbf{R}\}} \rho_l(\mathbf{r} - \mathbf{R}) \quad (2.6)$$

$$\rho_l(\mathbf{r}) = n_{\text{occ}} \left(\frac{\alpha}{\pi}\right)^{3/2} e^{\alpha r^2}, \quad (2.7)$$

where  $\alpha$  is the localization parameter. The free energy per particle  $F[\rho]/N = f(n_{\text{occ}}, \alpha)$  is now a function of the cluster occupation and  $\alpha$ . The free energy can be divided into the ideal, the inter- and the intra-cluster contribution (for more details see [37]):

$$f(n_{\text{occ}}, \alpha) = f_{\text{id}}(n_{\text{occ}}, \alpha) + f_{\text{inter}}(n_{\text{occ}}, \alpha) + f_{\text{intra}}(n_{\text{occ}}, \alpha) \quad (2.8)$$

$$f_{\text{id}}(n_{\text{occ}}, \alpha) = k_B T [\log n_{\text{occ}} + 3/2 \log(\alpha\sigma^2/\pi) - 5/2 + 3 \log(\Lambda/\sigma)] \quad (2.9)$$

$$f_{\text{inter}}(n_{\text{occ}}, \alpha) = n_{\text{occ}} \sqrt{\frac{\alpha}{8\pi}} \sum_{\mathbf{R} \neq 0} \int_0^\infty dr \frac{r}{R} [e^{-\alpha(r-R)^2/2} - e^{-\alpha(r+R)^2/2}] \phi(r) \quad (2.10)$$

$$f_{\text{intra}}(n_{\text{occ}}, \alpha) = (n_{\text{occ}} - 1) \sqrt{\frac{\alpha^3}{2\pi}} \int_0^\infty dr r^2 e^{-\alpha r^2/2} \phi(r) \quad (2.11)$$

A detailed study of these three contributions provides an insight into why clustering can occur. In figure 2.3 the authors of [37] plotted the different contributions and their sum for a particular realisation of a system where particles interact *via* the GEM-4 potential at  $k_B T = 1.0$  and  $\rho\sigma^3 = 9.0$  in an fcc lattice as functions of  $n_{\text{occ}}$  for a fixed  $\alpha$ -value. We see that the inter-cluster contribution (dashed line) decreases with  $n_{\text{occ}}$ : if the cluster

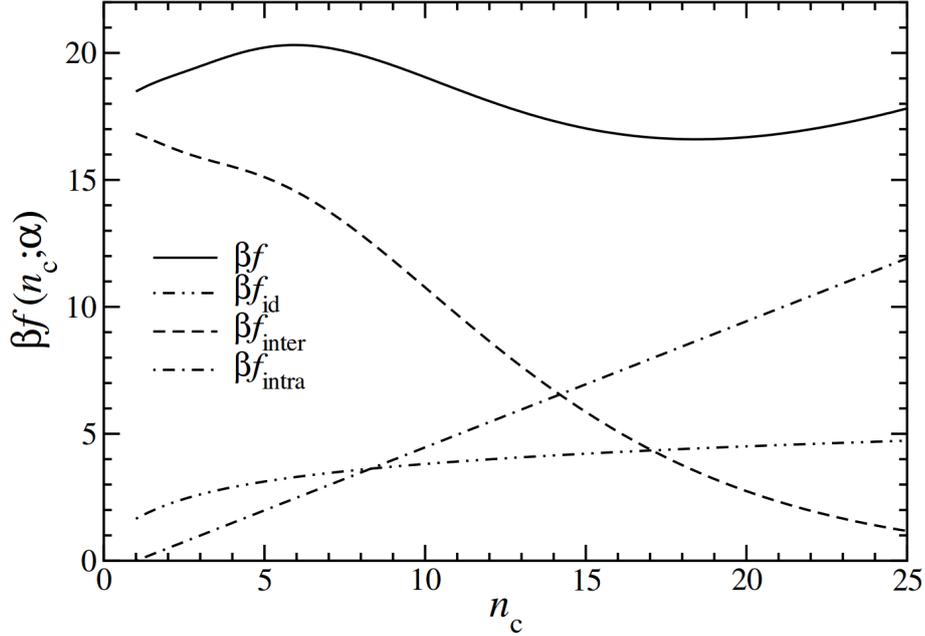


Figure 2.3.: Free energy per particle in a GEM-4 system at  $k_B T = 1.0$  and  $\rho\sigma^3 = 9.0$  as a function of the cluster occupation: the dashed line presents the inter-cluster contribution, the dash-dotted line presents the intra-cluster contribution, the dash-dot-dot line presents the ideal contribution and the solid line the sum over all of them. Taken from [37]

occupation increases the number of clusters decreases; if the number of clusters in a fixed volume decreases, these can then be further away separated and therefore the repulsion they exert on each other decreases. On the other hand, the intra-cluster contribution (dash-dotted line) increases with the cluster occupation: as clusters are more highly occupied the particles feel a stronger repulsion from the other particles in the cluster. Finally the term  $\log n_{occ}$  in the ideal contribution expresses the entropy loss due to particle aggregation. One can see that in total three competing terms lead to a global minimum of the free energy at a finite  $n_{occ}$ , which represents the equilibrium cluster occupation for this particular state point.

### Freezing properties

Likos *et al.* studied in [23] the freezing properties of cluster crystals. By transforming the expression for  $\rho(r)$  in equation (2.6) into Fourier space we obtain:

$$\rho_{\mathbf{K}} = \frac{1}{v_c} \int_{\mathcal{C}} d^3 r e^{i\mathbf{K}\mathbf{r}} \rho(\mathbf{r}) = \frac{1}{v_c} \int_{\mathbb{R}^3} d^3 r e^{i\mathbf{K}\mathbf{r}} \rho_l(\mathbf{r}) = \frac{n_{occ}}{v_c} e^{-K^2/(4\alpha)} = \rho e^{-K^2/(4\alpha)} \quad (2.12)$$

where  $\int_{\mathcal{C}}$  means the integral over the elementary unit cell and  $v_c = V/N_c$  and the second integral extends over the entire space. If now one introduces equation (2.12) into equation (2.5),

the free energy per particle can be expressed as a function of  $n_{\text{occ}}$ ,  $\alpha$ ,  $T$  and  $\rho$ :

$$\begin{aligned} \frac{F_{\text{id}} + F_{\text{ex}}}{N} &= \tilde{f}(n_{\text{occ}}, \alpha^*; T^*, \rho^*) = & (2.13) \\ &= T^* \left[ \ln n_{\text{occ}} + \frac{3}{2} \ln \left( \frac{\alpha^*}{\pi} \right) - \frac{5}{2} + 3 \ln \left( \frac{\Lambda}{\sigma} \right) \right] + \frac{\rho^*}{2} \sum_{\mathbf{Y}} \tilde{\phi}(Y) e^{-Y^2/(2\alpha^*)} \end{aligned}$$

where the asterisks denote reduced units, namely  $\rho^* = \rho\sigma^3$ ,  $T^* = k_B T/\epsilon$ ,  $\alpha^* = \alpha\sigma^2$  and  $\mathbf{Y} = \mathbf{K}\sigma$ . The free energy of the crystal  $f_{\text{sol}}(T^*, \rho^*)$  is obtained by minimizing equation (2.14) with respect to  $n_{\text{occ}}$  and  $\alpha^*$

$$f_{\text{sol}}(T^*, \rho^*) = \min_{[n_{\text{occ}}, \alpha^*]} \tilde{f}(n_{\text{occ}}, \alpha^*; T^*, \rho^*) \quad (2.14)$$

As a first approximation one can ignore the sum over reciprocal lattice vectors (RLVs) in equation (2.14) above all RLV beyond the first shell, whose length is  $Y_1 = K_1\sigma$ . This is justified by the exponentially decaying factor  $e^{-Y^2/(2\alpha^*)}$  and the fact that  $\tilde{\phi}(Y)$  quickly decays to zero. If we now express the density in terms of the lattice constant  $a$

$$\rho = \frac{zn_{\text{occ}}}{a^3} \quad (2.15)$$

where  $z$  is lattice dependent integer coefficient of the order unity ( $z = 4$  for a fcc lattice,  $z = 2$  for a bcc lattice, etc ...) and introduce the quantity  $\alpha a^2 = \gamma^{-1}$  we can perform a change of variables which simplifies the mathematics. We also drop the constant term  $3 \ln \left( \frac{\Lambda}{\sigma} \right)$  which also appears in the free energy of the fluid and does not affect any phase boundaries. Equation (2.14) then reads as:

$$\begin{aligned} \tilde{f}(n_{\text{occ}}, \alpha^*; T^*, \rho^*) &= T^* \left[ \ln \rho^* - \ln z - \frac{3}{2} [\ln(\pi\gamma) - 1] - 1 \right] + & (2.16) \\ &+ \frac{\rho^*}{2} \tilde{\phi}(0) + \frac{\rho^* \xi_1}{2} \tilde{\phi}(Y_1(n_{\text{occ}})) e^{Y_1(n_{\text{occ}})^2/(2\alpha^*)} \end{aligned}$$

where  $\xi_1$  is the coordination number of the reciprocal lattice. Now the dependency on  $n_{\text{occ}}$  is hidden in the length of the first RLV: the length of the first shell of RLVs in a Bravais lattice of a lattice constant  $a$  scales as  $K_1 = \zeta/a$ , where  $\zeta$  is again a lattice-dependent integer. Together with equation (2.15) this implies that

$$Y_1(n_{\text{occ}}) = \zeta \left( \frac{\rho^*}{zn_{\text{occ}}} \right)^{1/3}. \quad (2.17)$$

We introduce the parameter  $\gamma$  in the exponent in equation (2.16) *via* the equality  $Y_1^2/2\alpha^* = \gamma\zeta^2/2$ , and therefore the dependency on  $n_{\text{occ}}$  remains only in  $\tilde{\phi}(Y_1(n_{\text{occ}}))$

The minimization of  $\tilde{f}$  in equation (2.16) with respect to  $n_{\text{occ}}$  determines the value for the

lattice spacing:

$$\frac{\partial \tilde{f}}{\partial n_{\text{occ}}} = 0 \Rightarrow \tilde{\phi}'(Y_1)Y_1^4 = 0 \quad (2.18)$$

One solution for the equation  $Y_1 = 0$  is discarded because it leads to the non-physical solution of having  $n_{\text{occ}} = \infty$ . The other solution is  $Y_1 = y^*$  being  $y^*$  the value at which the Fourier transform of the potential attains its negative minimum. This condition determines the spacing of the lattice, because  $K_1 = Y_1/\sigma = \zeta/a$ :

$$a = \zeta/(y^* \sigma), \quad (2.19)$$

which only depends on the shape of the potential *via*  $y^*$  and not on the density  $\rho$  or temperature  $T$ .

We can introduce  $\phi(Y_1) = \phi(y^*) < 0$  in equation (2.16). Furthermore we can see that the term  $T^*[\ln \rho^* - 1]$  and  $\rho^* \tilde{\phi}(0)/2$  correspond to the ideal and excess free energy  $f_{\text{fl}}$  of a uniform fluid of density  $\rho^*$  [23], and introduce the difference  $\Delta f \equiv f - f_{\text{fl}}$ :

$$\Delta f(n_{\text{occ}}(y^*), \gamma, T^*, \rho^*) = -\frac{3T^*}{2} \left[ \ln(\gamma\pi) + 1 + \frac{2 \ln z}{3} \right] + \frac{\xi_1 \rho^*}{2} \tilde{\phi}(y^*) e^{-\gamma \zeta^2/2}. \quad (2.20)$$

The value of  $\gamma$  is restricted to small values so that the Gaussians centred on the different lattice sites do not overlap: a very generous upper limit would be  $\gamma \leq 0.005$ . In the regime of these values the first term in equation (2.20) is positive. This term accounts for the entropic cost of localization that crystals pay. This cost has to be compensated by the second term, which is only possible if  $\phi(y^*) < 0$ . The most favourable crystal structure will be that one which maximised  $\xi_1$ , the coordination number of the reciprocal lattice. This is the BCC lattice, whose reciprocal lattice is an fcc. This result is a consequence of the sum over the RLV only being performed in the first shell. Inclusion of higher order shells, can under suitable thermodynamic conditions, stabilise the fcc in favour of the BCC.

By minimizing equation (2.16) with respect to  $\gamma$  and imposing  $\Delta f = 0$  one obtains the freezing line.

$$\frac{\partial \tilde{f}}{\partial \gamma} \Rightarrow \frac{3T^*}{2\gamma} + \frac{\xi_1 \zeta^2 \rho^*}{4} \tilde{\phi}(y^*) e^{-\gamma \zeta^2/2} \quad (2.21)$$

$$\Delta f = 0 \Rightarrow \frac{\xi_1 \rho^*}{2} \tilde{\phi}(y^*) = \frac{3T^*}{2} \left[ \ln(\gamma\pi) + 1 + \frac{2 \ln z}{3} \right] \quad (2.22)$$

Introducing  $z = 2$  and  $\zeta = 2\sqrt{2}\pi$  for the BCC lattice we obtain:

$$\frac{T_f^*}{\rho_f^*} \cong 1.393 |\tilde{\phi}(y^*)|. \quad (2.23)$$

Equation (2.23) is a major result since it allows to estimate the freezing line (and therefore the

phase diagram) of a system of ultrasoft particles just by computing the Fourier transform of the interaction potential and finding its minimum. Results on the phase diagram of a GEM-4 cluster-forming system shown in references [21, 23] show a very good agreement between equation (2.23), results obtained by a full minimization of the density functional and results obtained by Monte Carlo (MC) simulations.

### 2.3.1. Bulk modulus

The response to compression of a cluster crystal shows very interesting features due to the multiple occupancy of the lattice. This response can be measured by the bulk modulus; *i.e.* the second derivative of the free energy with respect to the volume at constant temperature and number of particles:

$$B = V \left( \frac{\partial^2 F}{\partial V^2} \right)_{N,T} = -V \left( \frac{\partial P}{\partial V} \right)_{N,T} \quad (2.24)$$

which account for the response of the system to a change in volume. A general variation of the free energy of a crystal can be written as [22]:

$$dF = -SdT - PdV + \mu dN + \mu_c dN_{\text{cells}} \quad (2.25)$$

where  $S$  is the entropy,  $P$  the pressure,  $\mu$  the chemical potential and  $\mu_c$  the cell chemical potential conjugate to the number of unit cells  $N_{\text{cells}}$  in the system. At constant  $T$  and  $N$ ,  $N_{\text{cells}}$  takes the optimum value such that  $\mu_c = 0$ . In single occupancy crystals  $N_{\text{cells}}$  is considered to be  $N_{\text{cells}} = N/z$ , where  $z$  is the previously introduced lattice dependent parameter of order unity. The corresponding term is then usually neglected because the number of vacancies is so small that the effect of keeping the term is negligible. In multiple occupancy crystals the scenario is completely different. If the  $N_{\text{cells}}$  term is kept the bulk modulus has two contributions:

$$B = -V \left( \frac{\partial P}{\partial V} \right)_{N,T,N_{\text{cells}}} - V \left( \frac{\partial P}{\partial N_c} \right)_{N,T,V} \left( \frac{\partial N_{\text{cells}}}{\partial V} \right)_{N,T} = B_{\text{vir}} - B_{\text{del}}. \quad (2.26)$$

The first term,  $B_{\text{vir}}$ , corresponds to the intuitive reaction of a crystal to compression: when the volume is reduced the particles (or clusters of particles) move closer. The second term corresponds to the deletion of certain lattice positions. Mladek *et al.* computed in [22] the magnitude of these two contributions for a GEM-4 cluster crystal and found that the correction introduced by  $B_{\text{del}}$  results in a reduction of the original value ( $B_{\text{vir}}$ ) of over 45 %.

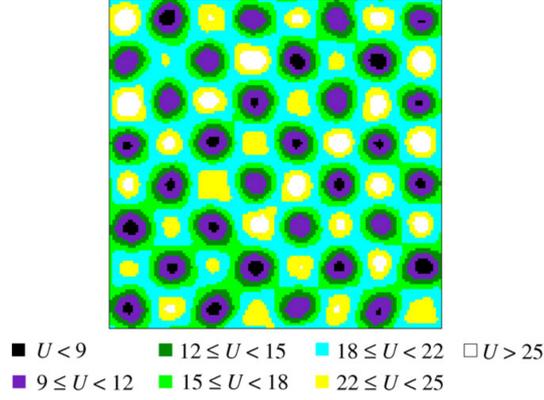


Figure 2.4.: Projection of the potential energy  $U$  of a GEM-8 model (measured in units of  $\epsilon$ ) on a [001] lattice plane, for a configuration of the fcc crystal at  $\rho^* = 5.0$  and  $T^* = 0.667$ . Darker colours represent lower energies (see legend) and are centred around lattice sites. Taken from [44].

### 2.3.2. Dynamics in cluster crystals

The optimum (ideal) value of the cluster occupation of a cluster crystal is  $n_{\text{occ}} = N/N_c$ , which in general is a non-integer value. In a realistic system, whose clusters are of course populated by an integer number of particles, a considerable polydispersity in the lattice sites population can be expected. Thus the system is full of *defects*, in the sense that the occupation of the clusters is not ideal. Mechanical stability must be maintained on average, by means of dynamical events that guarantee the equality of the time-average  $n_{\text{occ}}$  for all sites [44]. Moreno *et al.* performed in reference [24] a thorough study of the dynamic features of a cluster crystal of particles interacting *via* the GEM-8 potential. Figure 2.4 shows the energy landscape that the authors calculated for the 2-dimensional crystal. Through an activated mechanism particles which have a sufficiently high energy are able to escape from their host cluster and migrate between lattice sites maintaining the lattice structure unaffected. This *activated hopping* mechanism results in finite values for the diffusivity of the GEM particles [24]. Moreno *et al.* reported that for a representative state point of a solid fcc GEM-8 crystal more than 80% of the particles in the system have moved to a different cluster after a sufficiently long observation time.

In a later contribution [45] Strauss *et al.* calculated the distribution of the sizes of *kicking* and *catching* clusters, *i.e.* the relative occupation of the clusters which lose a particle and the clusters which receive it with respect to the average cluster occupation. It was observed that in the majority, kicking clusters were over-occupied and catching clusters were under-occupied. However, the overlap between these two distributions was considerable. Particle hopping is therefore not just a mechanism to relax unbalanced cluster occupancy but there are also substantial occupancy fluctuations in which particles start hopping from their cluster at random, without being induced by any external influence.

## 2.4. Generalised exponential model

The generalised exponential model (equation (2.4)) has been so far the system most commonly used to study cluster crystals on the level of coarse grained particles [21–23, 25, 26]. The parameter  $n$  controls the steepness of the potential and takes values from  $n = 2$  (the Gaussian core model, GCM) to  $n = \infty$  (the penetrable sphere model, PSM). It can be shown

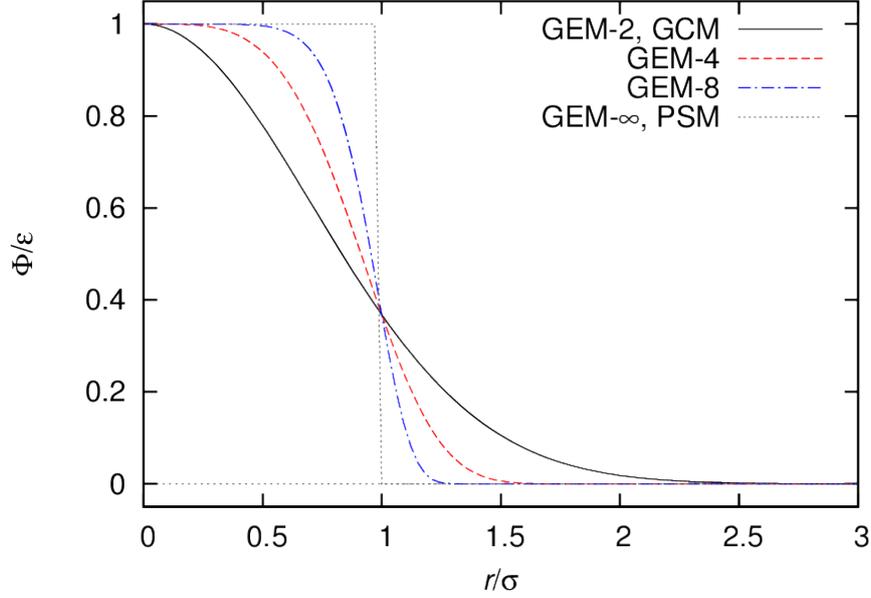


Figure 2.5.: Four members of the GEM- $n$  family with  $n = 2$  (GCM), 4, 8 and  $\infty$  (PSM).

that GEM- $n$  potentials show clustering when  $n > 2$ : we start from the inverse Fourier transform of the potential:

$$\phi(r) = \frac{1}{2\pi^2} \int_0^\infty \frac{\sin(kr)}{kr} \tilde{\phi}(k) k^2 dk. \quad (2.27)$$

The second derivative of the potential at  $r = 0$  is given by:

$$\frac{d^2\phi(r=0)}{dr^2} = -\frac{1}{6\pi^2} \int_0^\infty \tilde{\phi}(k) k^4 dk. \quad (2.28)$$

If the left hand side of equation (2.28) is larger than or equal to zero, then  $\tilde{\phi}(k)$  must necessarily have negatives parts and  $\phi(r)$  is a  $Q^\pm$  potential. The second derivative of a GEM- $n$  potential reads as follows:

$$\frac{d^2\phi(r)}{dr^2} = \frac{\epsilon n}{\sigma^2} \exp[-(r/\sigma)^2] \left[ n \left(\frac{r}{\sigma}\right)^{2(n-1)} - (n-1) \left(\frac{r}{\sigma}\right)^{n-2} \right] \quad (2.29)$$

One can see that  $\frac{d^2\phi(r=0)}{dr^2} = 0$  for  $n > 2$ .

Clustering is a consequence of the *steepness* of the potential. When clusters are formed, if the inter-cluster separation is large enough, particles sitting in the cluster only experience overlaps with the particles in their own cluster. The steeper the potential, the smaller this inter-cluster separation needs to be, and in turn, the smaller the occupation of the clusters (because if the required separation between the clusters is smaller, more clusters can be formed). Therefore clustering is advantageous when the potential is steep, the inter-cluster separations is small, and the cost of sparing a lot of partial overlaps (which would occur in a non-cluster configuration) is just a full overlap with a small number of cluster members. If the potential is *shallow* (as in the  $n = 2$ ) the inter-cluster separation has to be large (in figure 2.5 at  $r = 1.5\sigma$  GEM-4, 8 and  $\infty$  have vanished, while GEM-2 is still  $\sim 0.1\epsilon$ ). The cluster occupation  $n_{\text{occ}}$  has to be higher, and the big amount of full overlaps in the cluster configuration does not compensate the partial overlaps in the non-cluster configuration.

## 2.5. Dendrimers

Dendrimers are polymeric macromolecules with a tree-like structure (see figure 2.6) which are usually embedded in a microscopic solvent. Our interest in dendrimers lies in the fact that they are under certain conditions polymeric macromolecules which have an effective COM interaction with a negative Fourier component [46]. In 2011 Lenz *et al.* produced stable cluster fluid [27] and crystal [28] dendrimer configurations. Dendrimers can therefore be considered as a promising candidate macromolecule for an experimental realisation of a cluster crystal.

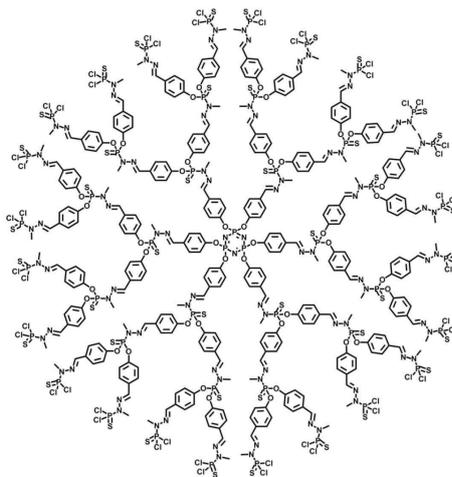


Figure 2.6.: Structure of a fourth-generation dendrimer. From <http://www.nanodic.com/PicturePreview.php?ch=nanomaterial&term=Dendrimer&g=6&type=JPG> (active 29.07.2014).

### 2.5.1. History of dendrimers

Dendrimers were synthesised for the first time in 1978 by Vögtle and co-workers [47]. Interest in these molecules grew rapidly due to the long lasting discussion on which conformation dendrimers would take: a *dense-shell* conformation, where end-groups reside in the periphery of the macromolecule leaving a hollow core, or a *dense-core* conformation where the end-groups fold back into the core. After some progress in the theory and experiments which made it possible to map chemical structures onto simulation data [48] it was finally shown that both conformations could exist: dendrimers with flexible bonds would show a dense-core conformation (theory and simulations [48–50] and experimental corroboration [51, 52]) and dendrimers with stiff bonds would show a dense-shell conformation (simulations [53] and experiments [54]). Furthermore it was shown that changing the pH or the salt concentration of the solvent could bring a conformation change from dense-core to dense-shell and vice versa (simulations [39] and experiments [55, 56]) making dendrimers the perfect candidate for the carriage of small molecular substances and boosting once again interest in them.

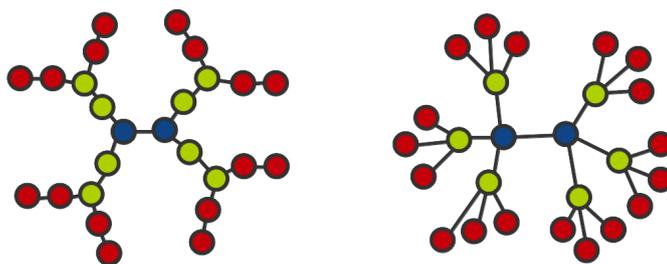


Figure 2.7.: Schematic representation of two dendrimers. Monomers are coloured according to the generation they belong to (blue is 0, green is 1, red is 2). Left: an  $f=3$ ,  $p=1$ ,  $g=2$ ; right:  $f=4$ ,  $p=0$ ,  $g=2$ .

Dendrimers have nowadays many technological and medical applications: they are used for gene transduction and drug delivery, as bio-sensors or magnetic resonance image contrast agents or even for the prevention of infection with HIV and other sexually transmitted diseases [57–64]. The synthesis of dendrimers is however rather time consuming and expensive and therefore their commercial availability is up to now limited to only a few substances [65].

### 2.5.2. Modelling dendrimers

The architecture of a dendrimer is characterised by three quantities: the functionality  $f$ , the generation  $g$  and the spacer length  $p$ . We start from two central monomers, onto which  $(f - 1)$  branches of  $p$  monomers are grafted. The starting two monomers belong to the  $0^{\text{th}}$  generation, the  $(f - 1)$  attached branches belong to the first generation. Then  $(f - 1)$  branches would be attached to each of the first generation monomers, forming the second generation monomers, and so on until  $g$  generations have been grafted. In figure 2.7 examples of a  $f=3$ ,  $p=1$ ,  $g=2$  and a  $f=4$ ,  $p=0$ ,  $g=2$  dendrimer are given.

The behaviour of the dendrimer is determined by the architecture and the chemical nature of the monomers (interactions with the solvent, size, softness, etc... ). In computer simulations the architecture of the polymer is taken into account by the bonding interactions and the chemical nature of the monomers *via* the excluded volume interactions.

*Athermal dendrimers* were the first kind of dendrimers to be simulated by Götze *et al.* in reference [66]. The bead-spring model was used, where excluded volume was modelled either by a hard sphere (in Monte Carlo simulations) or a cut and shifted (purely repulsive) Lennard Jones (LJ) potential (in Molecular Dynamics simulations). The *athermal* terminology stems from the  $T$ -independent nature of the hard sphere potential. Although a cut and shifted LJ potential is not strictly  $T$ -independent, the parameters chosen in reference [66] make it close enough to a hard sphere potential, so that we can still say that the dendrimers are athermal. Bonds were represented with threads or the finite extensible non-linear elastic (FENE) model

(see below). Such a model leads to effective Gaussian interactions [67], which are not able to show clustering.

In order to obtain effective interactions with an oscillatory Fourier transform a distinction between the monomers had to be introduced [46]: monomers in the outer shell (“S” monomers) are assumed to be solvophilic, the fact that the monomers tend to maximise the contact with the solvent (which is not simulated explicitly) results in an effective repulsion between each other. The monomers in the core (“C” monomers) are assumed to be solvophobic, their repulsion with the solvent results in a net effective attraction between C monomers. These kinds of dendrimers are called *amphiphilic dendrimers*. Instead of one potential for the excluded volume interaction between the monomers we have to consider three different ones:  $\phi_{SS}(r)$  and  $\phi_{SC}(r)$  (purely repulsive) and  $\phi_{CC}(r)$  (attractive).

The FENE potential is used for the bonded interaction [39, 48]:

$$\beta\phi^{\text{FENE}}(r) = -\kappa R^2 \ln \left[ 1 - \left( \frac{r-l}{R} \right)^2 \right], \quad (2.30)$$

$l$  determines the equilibrium bond-length,  $R$  determines how much the bond is allowed to stretch from the equilibrium bond-length  $l$ , and  $\kappa$  determines the elasticity of the bond. In the left panel of figure 2.8 we plot  $\phi^{\text{FENE}}(r)$  for three different combinations of the parameters  $(\kappa, l, R)$ . The three curves correspond to bonds with an equilibrium length of  $l = 3$ . The red curve corresponds to  $\kappa = 40$  and  $R = 0.6$ , the green curve to  $\kappa = 60$  and  $R = 0.6$  and the blue curve to  $\kappa = 40$  and  $R = 0.3$ .

The Morse potential for the excluded volume is given by:

$$\beta\phi_{\nu\mu}^{\text{Morse}}(r) = \epsilon_{\nu\mu} \left[ \left( e^{-\alpha_{\nu\mu}(r-\sigma_{\nu\mu})} - 1 \right)^2 - 1 \right], \quad (2.31)$$

with  $\nu, \mu = S, C$ . The parameter  $\epsilon$  determines the strength of the interaction,  $\sigma$  determines the radius of the monomer and  $\alpha$  tunes the attraction, *i.e.*  $\alpha$  accounts for the interaction with solvent. In the right panel of figure 2.8 we plot  $\phi^{\text{Morse}}(r)$  for four different combinations of the parameters  $(\epsilon, \sigma, \alpha)$ . The minimum value of the potential is placed at  $r = \sigma$ . The red ( $\epsilon = 0.7, \alpha = 6.0, \sigma = 1.0$ ) and the green ( $\epsilon = 0.7, \alpha = 6.0, \sigma = 1.5$ ) curve have a strong attractive well at  $r = \sigma$ : they represent a solvophobic interaction. By reducing the value of  $\epsilon$  (red to green curve) the attraction between the monomers is reduced (the depth of the well is reduced) and the softness of the monomers is increased (the divergence of the potential in the origin is weaker, therefore the monomers can come closer). Finally, by reducing the value of  $\epsilon$  and increasing the value of  $\alpha$  (red to pink curve) the attractive well completely disappears and therefore, we change the interaction from solvophobic to solvophilic.

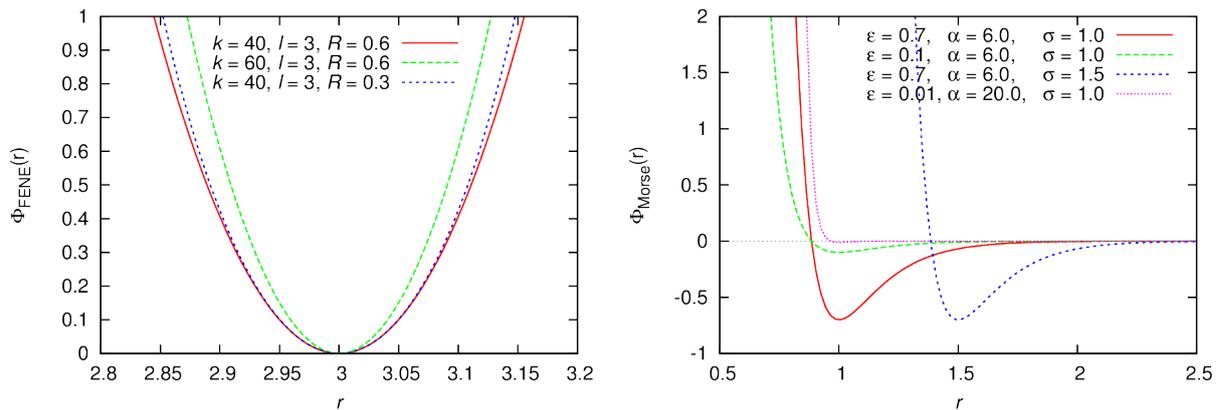


Figure 2.8.: Left: FENE potential for different sets of  $(\kappa, l, R)$  parameters as shown in the legend. Right: Morse potential for different sets of  $(\epsilon, \alpha, \sigma)$  parameters as shown in the legend.

### 2.5.3. Cluster forming amphiphilic dendrimers

Mladek *et al.* simulated [46] a set of 7 amphiphilic dendrimers, called D1 to D7. The architecture of the D dendrimers is given by  $g = 2$ ,  $p = 0$ , and  $f = 3$ . Monomers in the zero and first generation were assumed to be solvophobic and monomers in the second generation were assumed to be solvophilic. The parameters in the Morse and FENE potential were varied for the seven models within the limits of what can be realised experimentally [39]. The monomer density profile is defined as:

$$\rho(r)_{C/S} = \left\langle \sum_i \delta(\mathbf{r}_{\text{COM}} - r_i^{C/S}) \right\rangle \quad (2.32)$$

where  $\mathbf{r}_{\text{COM}}$  is the position of the centre of mass of the dendrimer,  $\mathbf{r}_i$  is the position of the  $i$ -th monomer and the angular brackets denote an average over equilibrium configurations. A study of  $\rho(r)_{C/S}$  showed that in contrast to athermal dendrimers, shell and core monomers were more segregated, end groups did not fold back and therefore the core is less populated, favouring clustering. Effective interactions were calculated for the seven models using the Umbrella sampling method [68] and the Widom particle insertion method [46]. All the seven dendrimer models did indeed show an oscillatory Fourier transforms.

Lenz *et al.* used [27] D2 dendrimers in monomer resolved simulations, the parameters of the interaction potentials are specified in table 2.1. The simulations were performed at densities where the fluid phase is expected. At low density the system formed an homogeneous fluid and on increasing density it spontaneously transformed into a fluid of clusters of dendrimers.

In a later contribution [28] Lenz *et al.* used a modified dendrimer model, the so-called  $r$ -family to go deep into the solid cluster crystal phase. The  $r$  dendrimer model makes a distinction between the central bond ( $g = 0$ ) and the rest of the core-core bonds. Lenz *et*

Morse	$\epsilon_{\nu\mu}$	$\alpha_{\nu\mu}$	$\sigma_{\nu\mu}$	FENE	$\kappa$	$l$	$R$
CC	0.714	6.4	1	CC	40	1.875	0.375
CS	0.014	19.2	1.25	CS	20	3.75	0.75
SS	0.014	19.2	1.5				

Table 2.1.: Morse and FENE parameters of the  $D2$  model.

*al.* used the so-called  $r6$  model (see table 2.2) which has larger shell monomers, a considerably larger equilibrium central bond length and a longer range of attraction between the core monomers. These features made the minimum of the Fourier transform of the effective potential deeper. The transition to the cluster crystal phase occurs at lower densities (see equation (2.23)), facilitating the computational effort to simulate this phase. As a first step Lenz *et al.* simulated fluid systems at different densities. At sufficiently high densities small clusters (with occupations below 3) disappeared and the system spontaneously developed local crystalline order. This was confirmed by the radial distribution function of the COMs of the dendrimers:

$$g(r) = \frac{1}{N} \left\langle \sum_{i=1}^N \sum_{j=1}^N \delta(\mathbf{r} - \mathbf{r}_j + \mathbf{r}_i) \right\rangle \quad (2.33)$$

which measures the probability of finding two particles separated by a distance  $r$ . The radial distribution function showed a peak centred at a value close to the predicted lattice spacing  $a$  (see equation (2.19)) which was clearly separated from the second peak. Lenz *et al.* prepared several starting configurations where the clusters of dendrimers were positioned at the lattice positions of perfect fcc and bcc crystals with different cluster occupations and lattice spacings (close to the predicted one). Unstable crystals melted quickly, while mechanically stable systems remained in the crystalline clustered arrangement throughout the whole simulation. Figure 2.9 shows a snapshot of a stable crystal. In panel (a) the monomers are coloured according to the generation number: red for  $g = 2$ , blue for  $g = 1$  and black  $g = 0$ . In panel (b) the positions of the COMs are shown in green for this particular snapshot; a sphere surrounding them which contains the latter to a probability of 99% accounts for the oscillations of the COMs around the lattice site. These spheres do not overlap, confirming the prediction that the clusters are well localised at their lattice sites.

Hopping events were observed for the low cluster occupation configurations ( $n_c = 4$ ). The main discrepancy with the effective pair potential picture arose at high densities, where an unlimited growth of the cluster occupancy with density is predicted in the CG picture. Instead, in the MR simulations, big clusters split into smaller clusters, the distance between the clusters decreased and entered the length regime of the bonds between the dendrimer's monomers. Dendrimers stretched and were shared between different clusters forming a percolating network.

Morse	$\epsilon_{\nu\mu}$	$\alpha_{\nu\mu}$	$\sigma_{\nu\mu}$	FENE	$\kappa$	$l$	$R$
CC	0.714	1.8	1	CC ( $g = 0$ )	60	3.1875	0.6375
CS	0.01785	6.0	1.75	CC ( $g \neq 0$ )	60	1.875	0.375
SS	0.01785	6.0	2.5	CS	30	3.5625	0.7125

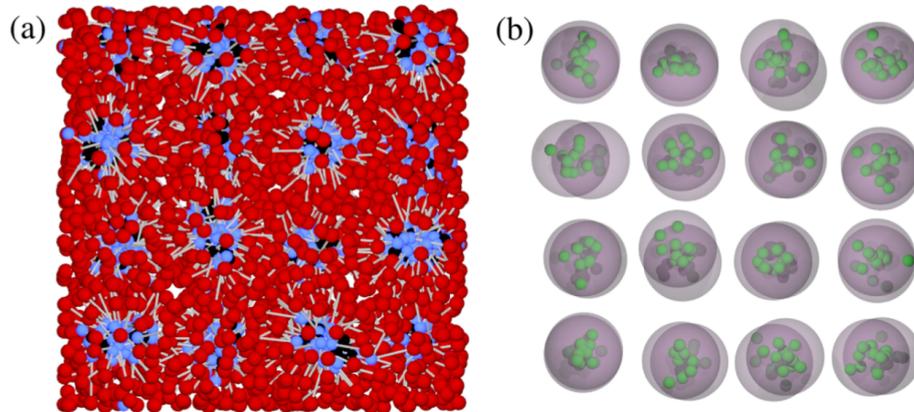
Table 2.2.: Morse and FENE parameters of the  $r6$  model.

Figure 2.9.: Figure taken from [28] where a snapshot of a stable cluster crystal is taken. Panel (a) shows the monomers coloured by red for  $g = 2$ , blue for  $g = 1$  and black for  $g = 0$ . Panel (b) shows the positions of the COMs (green) and an outer sphere which contains the position of the COM to a probability of 99%.



# 3. Methods

## 3.1. Simulation techniques

Ever since computers became widely spread in the past century, computer simulations have become an indispensable tool in sciences. Bridging between experiments and theory they complement both of them: on the one hand, in a simulation which tries to reproduce experimentally observable properties of a system one has access to microscopic data: positions, velocities and accelerations of individual particles. The interplay between different interactions can be disentangled and the origin of certain behaviours can be understood. In a computer simulation one can also perform “experiments” at extreme conditions, such as high pressure or very low temperatures which are otherwise very difficult or even impossible to access in a lab. On the other hand, with the help of computers one can numerically solve theoretical problems which do not have an analytical solution. However, computer simulations do suffer from limitations related to the system size and the length of simulations runs: in simulations one can handle typically ensembles of up to  $10^6$  particles, highly parallelised advanced techniques running on modern computer clusters can reach up to  $10^9$  particles; however, even at that level one is far away from systems of macroscopic size ( $N_A \sim 10^{23}$ ); on the other hand within a reasonable amount of time one can simulate a molecular system over time scales in the order of milliseconds, so that events taking place on larger time scales will only be accessible by sacrificing details of the system to build up a simpler model which can explore a longer time window in the same computing time. Thus, in practical applications a compromise has to be found between the level of accuracy of the microscopic details and the available time span, this situation is not always satisfactory.

### 3.1.1. Molecular dynamics

Molecular dynamics (MD) is a computer simulation technique which numerically integrates Newton’s second law of motion for an ensemble of particles interacting *via* a given potential  $U(\mathbf{r}^N)$ , where  $\mathbf{r}^N$  is an  $N$ -dimensional vector and every component is the position one of the  $N$  particles in the system. In this thesis we will only consider classical systems. The simulation begins with an initialization of the system, where positions ( $\mathbf{r}$ ) and velocities ( $\mathbf{v}$ ) are assigned to all the particles and an equilibration phase. After the system is equilibrated one can measure static and dynamic properties of the system. The term “equilibrated” has a different meaning for different ensembles. In this work we use the canonical ( $NVT$ ) ensemble, where the number of particles  $N$ , the volume  $V$  and the temperature  $T$  are kept constant. For such simulations the system is considered to be equilibrated once the internal

potential energy  $U(\mathbf{r}^N)$  fluctuates around a constant (equilibrium) value and there is no drift, *i.e.*  $U(\mathbf{r}^N)$  has reached a minimum. During an MD simulation Newton's second law of motion is integrated:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{f}_i}{m_i} = -\frac{1}{m_i}\nabla_i U(\mathbf{r}^N) \quad i = 1, \dots, N, \quad (3.1)$$

where  $\mathbf{r}_i$  denotes the position of particle  $i$ ,  $\mathbf{f}_i$  is the force acting on it and  $m_i$  is its mass. Several different numerical integration algorithms have been proposed in literature [68]. In this work we chose the velocity-Verlet algorithm [69], which even though it is fairly simple, has a very satisfactory performance. The velocity-Verlet algorithm is time reversible, area preserving in phase-space and has a good performance in the energy conservation. Positions are written as Taylor-expansions of the time-step  $\Delta t$ :

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{\mathbf{f}_i(t)}{2m_i}(\Delta t)^2 + \dots \quad (3.2)$$

If we truncate this expression at  $(\Delta t)^2$  we obtain the Euler algorithm. This algorithm is neither time reversible nor area preserving in phase space and suffers from a catastrophic energy drift [68]. To overcome this problem, the velocity-Verlet algorithm computes the velocities as functions of the forces at different time steps:

$$\mathbf{v}_i(t + \Delta t) \cong \mathbf{v}_i(t) + \frac{\mathbf{f}_i(t + \Delta t) + \mathbf{f}_i(t)}{2m_i}\Delta t. \quad (3.3)$$

Thus integration with respect to time is performed in the following way:

- (1) Update of the positions

$$\mathbf{r}_i(t + \Delta t) \cong \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{\mathbf{f}_i(t)}{2m_i}(\Delta t)^2. \quad (3.4)$$

- (2) First update of the velocity

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_i(t) + \frac{\mathbf{f}_i(t)}{2m_i}\Delta t. \quad (3.5)$$

- (3) Force calculation

$$\mathbf{f}_i(t + \Delta t) = \mathbf{f}_i(\mathbf{r}_i(t + \Delta t)) = -\nabla_i U(\mathbf{r}^N(t + \Delta t)). \quad (3.6)$$

- (4) Second update of the velocity

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t + \Delta t/2) + \frac{\mathbf{f}_i(t + \Delta t)}{2m_i}\Delta t. \quad (3.7)$$

This algorithm conserves the total energy,  $E$ , and therefore the temperature  $T$  is not constant.

For simulations in the canonical ( $NVT$ ) ensemble, where the temperature  $T$  instead of the energy is conserved, a thermostat is required which maintains a constant temperature. In this work we use the Andersen [70] and the Nosé-Hoover thermostat [71–74]. The Andersen thermostat mimics a temperature bath in the simulation by randomly giving particles a velocity drawn from the appropriate Boltzmann distribution for the given temperature  $T$ , corresponding to stochastic collisions between the particles in the system and in the bath. The implementation of this thermostat is very simple but the dynamics of the system become artificial. When reliable measurements of dynamic quantities are required one has to use a more elaborate thermostat. The Nosé-Hoover thermostat defines a deterministic integration method which keeps the temperature around a desired value. This is achieved by using an extended Lagrangian  $\mathcal{L}_{\text{Nosé}}$  and making thus the heat bath to an integral part of the system by adding an artificial variable  $s$  associated with an artificial mass  $Q$ :

$$\mathcal{L}_{\text{Nosé}} = \sum_{i=1}^N \frac{m_i}{2} s^2 \mathbf{v}_i^2 - U(\mathbf{r}^N) + \frac{Q}{2} \left( \frac{ds}{dt} \right)^2 - \frac{L}{\beta} \ln s. \quad (3.8)$$

The extended system generates a microcanonical ensemble of  $6N+2$  degrees of freedom and it can be shown [68] that if the parameter  $L$  takes the value  $L = 3N+1$ , averages of quantities in this extended ensemble reduce to the canonical average.

### 3.1.2. Multi particle collision dynamics

In simulations of systems of mesoscopic particles embedded in a microscopic solvent, one encounters the difficulty of treating very different time- and length-scales. The solvent particles are several orders of magnitude smaller than the solute ones and therefore their motion is several orders of magnitude faster. The first idea that could come to our minds is to simply include a set of solvent particles in the simulation with a proper potential describing the interactions between the solute and the solvent particles, and between the solvent particles among themselves. Since the motion of the solvent particles is much faster, a time step that is orders of magnitude smaller than the one used in a simulation of only solute particles would be needed. In addition the number of solvent particles will also be orders of magnitude larger than the number of solute particles, so that the computational time needed for the force calculation, which scales like  $\mathcal{O}(N^2)$  being  $N$  the total number of particles, will dramatically increase. Being the time step smaller and the computational time needed for the force calculation bigger, the time window that can be simulated within a reasonable amount of computing time will be ridiculously small.

A conventional solution to this problem is to simulate the solvent as a continuum background to the system: the properties of the solvent are included in the inter-particle interaction

potential of the solute particles: for example solvophobic solute particles will feel a net attraction to each other as a consequence of the fact that they are repelled by the solvent. However, such an approach does not take into account hydrodynamic interactions (HI). To visualise their effect we consider two solute particles separated by a large distance: one of the particles is moving. This movement will push the solvent particles aside, inducing *via* the solvent a force on the other solute particles, which otherwise would not have felt the movement of the first particle. These hydrodynamic interactions are very important for a proper description of the dynamic behaviour of colloidal systems.

Malevants and Kapral suggested in 1999 [34] a simulation method which reproduces the dynamic behaviour of the system but can be simulated over long time- and length-scales within a reasonable computing time: the so-called multi particle collision dynamics technique (MPCD). The MPCD method includes  $N_s$  solvent particles of mass  $m_s$  to the simulation of the  $N$  solute particles of mass  $m$ . The technique combines conventional MD steps with MPCD steps: after every  $\tau$  MD steps a MPCD step is performed.  $\tau$  controls the coupling between the solvent and the solute. In the MD steps the solvent particles are ignored and only positions, velocities and forces of the solute particles are updated, as described in subsection 3.1.1. MPCD steps consist of two parts: the streaming step and the collision step. In the streaming step the solvent particles propagate ballistic ally:

$$\mathbf{r}'_i(t + \Delta t) = \mathbf{r}'_i(t) + \mathbf{v}'_i(t)\Delta t, \quad (3.9)$$

where  $\mathbf{r}'_i$  and  $\mathbf{v}'_i$  are the position and velocity of the  $i$ th solvent particle (the apostrophe means that  $i$  is a solvent particle). In the collision step the volume is divided into  $N_{\text{cells}}^{\text{MPCD}}$  of length  $c$  and the solvent and solute particles are sorted into these cells. In each cell a rotation of the relative velocities of the particles with respect to the velocity  $\mathbf{u}_j$  of the centre of mass of the cell is performed ( $j$  is the index of the cell):

$$\mathbf{v}'_i(t + \Delta t) = \mathbf{u}_j(t) + \Omega(\alpha) [\mathbf{v}'_i(t) - \mathbf{u}_j(t)]. \quad (3.10)$$

Here  $\Omega(\alpha)$  is a norm conserving matrix performing a rotation of the relative velocities around a certain vector  $\hat{\zeta}$  of angle  $\alpha$ :

$$\Omega(\alpha) = \begin{bmatrix} \cos \alpha + \zeta_x^2(1 - \cos \alpha) & \zeta_x \zeta_y(1 - \cos \alpha) - \zeta_z \sin \alpha & \zeta_x \zeta_z(1 - \cos \alpha) + \zeta_y \sin \alpha \\ \zeta_y \zeta_x(1 - \cos \alpha) + \zeta_z \sin \alpha & \cos \alpha + \zeta_y^2(1 - \cos \alpha) & \zeta_y \zeta_z(1 - \cos \alpha) - \zeta_x \sin \alpha \\ \zeta_z \zeta_x(1 - \cos \alpha) - \zeta_y \sin \alpha & \zeta_z \zeta_y(1 - \cos \alpha) + \zeta_x \sin \alpha & \cos \alpha + \zeta_z^2(1 - \cos \alpha) \end{bmatrix}; \quad (3.11)$$

as a consequence the total energy and momentum of the solute particles are conserved:

$$\sum_{i \in \text{cell}_j} \mathbf{v}'_i = \sum_{i \in \text{cell}_j} (\mathbf{u}_j + \Omega(\alpha) [\mathbf{v}'_i - \mathbf{u}_j]) \quad (3.12)$$

$$\sum_{i \in \text{cell}_j} |\mathbf{v}'_i|^2 = \sum_{i \in \text{cell}_j} |\mathbf{u}_j + \Omega(\alpha) [\mathbf{v}'_i - \mathbf{u}_j]|^2. \quad (3.13)$$

In addition, there is a momentum transfer between solute and solvent particles during the collision step. Since the solvent particles are not included in the force calculation and the MPCD step is only performed every  $\tau$  steps (which is typically of the order of 50) the computational effort for including the solvent in the simulation is very low.

The components of the vector  $\hat{\zeta}$  are chosen randomly for every cell in each time step by picking two numbers  $x_1$  and  $x_2$  from a random distribution between 0 and 1 and expressing the components of  $\hat{\zeta}$  as:

$$\begin{aligned} \zeta_x &= 2x_1 \sqrt{1 - x_1^2 - x_2^2} \\ \zeta_y &= 2x_2 \sqrt{1 - x_1^2 - x_2^2} \\ \zeta_z &= 1 - 2(x_1^2 + x_2^2). \end{aligned} \quad (3.14)$$

The angle  $\alpha$  is kept constant during the whole simulation. It can be chosen from an analytic expression (see [34]) which expresses the kinetic and collisional viscosity  $\nu_{\text{kin}}$  and  $\nu_{\text{col}}$  as a function of  $\alpha$  to reproduce the desired value of these two quantities:

$$\nu_{\text{kin}} = \frac{k_B T \Delta t \rho}{\rho a^3} \left( \frac{5\rho}{(4 - 2 \cos \alpha - 2 \cos(2\alpha))(\rho - 1)} - \frac{1}{2} \right) \quad (3.15)$$

$$\nu_{\text{col}} = \frac{1 - \cos \alpha}{18 \rho a \Delta t} (\rho - 1), \quad (3.16)$$

where  $\rho = N/V$  is the density.

The mean free path of the solvent particles is given by  $\lambda = \Delta t \sqrt{k_B T}$ . It has been shown [75] that Galilean invariance is violated if  $\lambda < c/2$ . To avoid this non-physical behaviour all the cells are shifted by a random vector whose components take values from  $[-c/2, c/2]$  before every collision step (see figure 3.1). In this way, even if the particles do not move far enough to abandon the cell they are in they will not always collide with the same set of particles.

To keep the temperature of the system constant velocity rescaling shows to be good enough. During the equilibration steps the velocities of the solute and the solvent particles are rescaled by a factor  $\lambda$  dependent on the current (curr.) kinetic energy so that the system has the

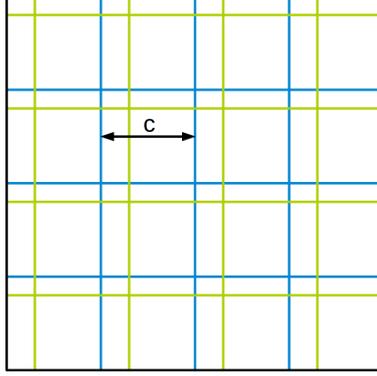


Figure 3.1.: Division of the simulation box into  $N_{\text{cells}}^{\text{MPCD}}$  of size  $c$ . To preserve Galilean invariance the cells are shifted (from blue to green, for example) by a random vector whose components take values from the interval  $[-c/2, c/2]$  before every collision step

desired (des.)  $T$  value:

$$K^{\text{curr.}} = \frac{1}{2} \sum_i m_i \mathbf{v}_i \quad (3.17)$$

$$K^{\text{des.}} = \frac{d}{2} k_B T \quad (3.18)$$

$$\lambda \equiv \frac{K^{\text{curr.}}}{d/2k_B T} \quad (3.19)$$

$$\mathbf{v}_i^{\text{scaled}} = \frac{\mathbf{v}_i}{\lambda} \quad (3.20)$$

After the equilibration steps no thermostat is needed and the temperature will fluctuate around the desired value  $T$ .

## 3.2. Computation of effective interactions: Multi-scale coarse graining method

In this work we have used the multi-scale coarse graining method (MSCG) proposed by Izvekov and Voth [76] to compute the *effective* interactions of polymeric macromolecules *via* a suitable coarse graining approach: in the end, the macromolecules are reduced to effective particles placed in their centre of mass (COM). This method is based on a force matching algorithm. When applying this method one has to run a simulation of an ensemble of monomer resolved molecules. The positions of individual monomers  $\mathbf{r}^n$  and the forces acting on them  $\mathbf{f}^n(\mathbf{r}^n)$  are stored during the simulation. Each set of data  $\{\mathbf{r}^n, \mathbf{f}^n(\mathbf{r}^n)\}$  will be referred to as a configuration. These configurations are then mapped to the positions and acceleration of the COMs of the macromolecules. Finally a variational principle which minimises a functional (defined later on in the text) is used to compute the functional form of the effective force field. The main advantage of this method compared to other existing methods (see [12, 13]) is that the effective interactions between two macromolecules can be extracted from configurations taken from systems at a finite density. The other methods compute effective pair interactions based on simulations of two isolated molecules (i.e. in the zero-density limit) and therefore do not include many-body effects. Their neglect can for example lead to an underestimation of the repulsion that two effective particles exert on each other when they approach, a repulsion which stems from the excluded volume interaction of the monomers; this effect will of course be much weaker in the zero-density limit than at finite density.

### 3.2.1. General formulation of the MSCG method

#### Low and high resolution model

We will refer to the high resolution model as the *atomic model*, where every basic monomer unit of a macromolecule will be referred to as an *atom* and the low resolution model as *coarse grained model* where every basic unit will be referred to as a *coarse grained (CG) site*. Capital letters will be used for quantities related to the coarse grained model and lower case letters will stand for quantities related to the atomic model.

We consider a system of  $N$  macromolecules. Each macromolecule is composed of  $n_m$  monomers,  $n = Nn_m$  is the total number of monomers. The interactions between the monomers are described by excluded volume and bonding potentials. We aim to coarse grain each macromolecule to an effective particle placed in its COM and compute the effective interaction between these coarse grained sites.

We start from an atomic model for the system consisting of  $n$  atoms, where Cartesian coordinates  $\mathbf{r}^n = \{\mathbf{r}_1, \dots, \mathbf{r}_n\}$  and  $\mathbf{p}^n = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$  are used to specify positions and momenta

of all atoms of masses  $m_i$ ,  $i = 1, \dots, n$ . The atomistic Hamiltonian is given by

$$h(\mathbf{r}^n, \mathbf{p}^n) = \sum_{i=1}^n \frac{1}{2m_i} \mathbf{p}_i^2 + u(\mathbf{r}^n). \quad (3.21)$$

The equilibrium probability density of the system in the canonical ensemble is given by

$$\begin{aligned} p_{rp}(\mathbf{r}^n, \mathbf{p}^n) &= p_r(\mathbf{r}^n) p_p(\mathbf{p}^n) \\ p_r(\mathbf{r}^n) &\propto \exp\left(-\frac{u(\mathbf{r}^n)}{k_B T}\right) \\ p_p(\mathbf{p}^n) &\propto \exp\left(-\frac{1}{k_B T} \sum_{i=1}^n \frac{1}{2m_i} \mathbf{p}_i^2\right) \end{aligned} \quad (3.22)$$

In order to define our low resolution (CG) model we need to define the linear mapping operators  $\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n)$  and  $\mathbf{M}_{\mathbf{P}}^N(\mathbf{p}^n)$  which take the atomistic coordinates  $\mathbf{r}^n$  and  $\mathbf{p}^n$  as input parameters and produce the Cartesian coordinates  $\{\mathbf{M}_{\mathbf{R}_1}(\mathbf{r}^n), \dots, \mathbf{M}_{\mathbf{R}_N}(\mathbf{r}^n)\}$  and  $\{\mathbf{M}_{\mathbf{P}_1}(\mathbf{p}^n), \dots, \mathbf{M}_{\mathbf{P}_N}(\mathbf{p}^n)\}$  of the  $N$  coarse grained sites, which is a point of the phase space of the CG coordinates  $\mathbf{R}^N = \{\mathbf{R}_1, \dots, \mathbf{R}_N\}$  and  $\mathbf{P}^N = \{\mathbf{P}_1, \dots, \mathbf{P}_N\}$ :

$$\begin{aligned} \mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) &= \{\mathbf{M}_{\mathbf{R}_1}(\mathbf{r}^n), \dots, \mathbf{M}_{\mathbf{R}_N}(\mathbf{r}^n)\}, \\ \mathbf{M}_{\mathbf{P}}^N(\mathbf{p}^n) &= \{\mathbf{M}_{\mathbf{P}_1}(\mathbf{p}^n), \dots, \mathbf{M}_{\mathbf{P}_N}(\mathbf{p}^n)\}, \end{aligned} \quad (3.23)$$

$$\begin{aligned} \mathbf{M}_{\mathbf{R}_I}(\mathbf{r}^n) &= \sum_{i=1}^n c_{Ii} \mathbf{r}_i, \\ \mathbf{M}_{\mathbf{P}_I}(\mathbf{p}^n) &= M_I \sum_{i=1}^n c_{Ii} \frac{\mathbf{p}_i}{m_i}, \end{aligned} \quad (3.24)$$

for  $I=1, \dots, N$ .  $M_I$  is the mass of the effective particle.

There are several conditions that the mapping between both models must fulfil, which will be derived in what follows. The mapping operator should be translationally symmetric in the sense that if all atoms are translated by a vector  $\mathbf{T}$  all the CG sites should be translated by the same vector  $\mathbf{T}$ . In order to impose translational symmetry to the mapping operator, the coefficients  $c_{Ii}$  for  $I = 1, \dots, N$  and  $i = 1, \dots, n$  must satisfy the condition

$$\sum_{i=1}^n c_{Ii} = 1 \quad (3.25)$$

It is convenient at this point to define for every CG site  $I$ , the set of *involved* atoms  $\mathcal{I}_I = \{i | c_{Ii} \neq 0\}$  and *specific* atoms  $\mathcal{S}_I = \{i | c_{Ii} \neq 0 \text{ and } c_{Ji} = 0 \text{ for all } J \neq I\}$ , which will be used in what follows.

The Hamiltonian of the coarse grained system is given by

$$H(\mathbf{R}^N, \mathbf{P}^N) = \sum_{I=1}^N \frac{1}{2M_I} \mathbf{P}_I^2 + U(\mathbf{R}^N) \quad (3.26)$$

where  $U(\mathbf{R}^N)$  is yet unknown and has to be computed. The equilibrium probability density in the canonical ensemble for the CG model is given by

$$\begin{aligned} P_{RP}(\mathbf{R}^N, \mathbf{P}^N) &= P_R(\mathbf{R}^N) P_P(\mathbf{P}^N) \\ P_R(\mathbf{R}^N) &\propto \exp\left(-\frac{U(\mathbf{R}^N)}{k_B T}\right) \\ P_P(\mathbf{P}^N) &\propto \exp\left(-\frac{1}{k_B T} \sum_{I=1}^N \frac{1}{2M_I} \mathbf{P}_I^2\right). \end{aligned} \quad (3.27)$$

### Consistency Condition

The coarse grained model has to be consistent with the underlying atomistic model in phase space in the sense that if two simulations are performed in the same thermodynamic conditions in the CG and the MR representation, the distribution of the positions and velocities of the CG sites in the CG representation and the COMs of the macromolecules in the MR representation have to be equal. To this end it must be satisfied that the equilibrium probability density of the CG model in terms of coordinates and momenta given by equation (3.27) is equal to that implied by the equilibrium probability density of the atomistic model in terms of the respective coordinates and momenta given in equation (3.22) in combination with the mapping operator, equation (3.24).

This implies

$$\begin{aligned} P_R(\mathbf{R}^N) = p_R(\mathbf{R}^N) &\equiv \int d\mathbf{r}^n p_r(\mathbf{r}^n) \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N) \\ P_P(\mathbf{P}^N) = p_P(\mathbf{P}^N) &\equiv \int d\mathbf{p}^n p_p(\mathbf{p}^n) \delta(\mathbf{M}_{\mathbf{P}}^N(\mathbf{p}^n) - \mathbf{P}^N) \end{aligned} \quad (3.28)$$

where

$$\begin{aligned} \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N) &\equiv \prod_{I=1}^N \delta(\mathbf{M}_{\mathbf{R}_I}(\mathbf{r}^n) - \mathbf{R}_I) \\ \delta(\mathbf{M}_{\mathbf{P}}^N(\mathbf{p}^n) - \mathbf{P}^N) &\equiv \prod_{I=1}^N \delta(\mathbf{M}_{\mathbf{P}_I}(\mathbf{p}^n) - \mathbf{P}_I). \end{aligned} \quad (3.29)$$

This implies that:

$$\exp\left(-\frac{1}{k_B T} U(\mathbf{R}^N)\right) \propto \int d\mathbf{r}^n \exp\left(-\frac{1}{k_B T} u(\mathbf{r}^n)\right) \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N) \quad (3.30)$$

$$\exp\left(-\frac{1}{k_B T} \sum_{I=1}^N \frac{1}{2M_I} \mathbf{P}_I^2\right) \propto \int d\mathbf{r}^n \exp\left(-\frac{1}{k_B T} \sum_{i=1}^n \frac{1}{2m_i} \mathbf{p}_i^2\right) \delta(\mathbf{M}_{\mathbf{P}}^N(\mathbf{p}^n) - \mathbf{P}^N) \quad (3.31)$$

Equation (3.30) fixes the consistency condition in configuration space and determines thus the coarse grained potential  $U(\mathbf{R}^N)$

$$U(\mathbf{R}^N) = -k_B T \ln z(\mathbf{R}^N) + (\text{const.}) \quad (3.32)$$

where

$$z(\mathbf{R}^N) \equiv \int d\mathbf{r}^n \exp(-u(\mathbf{r}^n)/k_B T) \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N). \quad (3.33)$$

The gradient of this potential with respect to  $\mathbf{R}_I$  determines the CG force field acting on the CG site with index  $I$ :

$$\begin{aligned} \mathbf{F}_I(\mathbf{R}^N) &= \\ &= -\frac{\partial U(\mathbf{R}^N)}{\partial \mathbf{R}_I} = \frac{k_B T}{z(\mathbf{R}^N)} \frac{\partial z(\mathbf{R}^N)}{\partial \mathbf{R}_I} = \\ &= -\frac{k_B T}{z(\mathbf{R}^N)} \int d\mathbf{r}^n \exp(-u(\mathbf{r}^n)/k_B T) \left[ \prod_{J \neq I} \delta(\mathbf{M}_{\mathbf{R}_J}(\mathbf{r}^n) - \mathbf{R}_J) \right] \frac{\partial}{\partial \mathbf{R}_I} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \mathbf{r}_i - \mathbf{R}_I\right) \end{aligned} \quad (3.34)$$

where the sum over runs over atoms  $i$  belonging to the set of involved atoms of CG  $I$   $\mathcal{I}_I$  and the relation  $\delta(\mathbf{M}_{\mathbf{R}}^N - \mathbf{R}^N) = \prod_{K=1}^N \delta(\mathbf{M}_{\mathbf{R}_K} - \mathbf{R}_K)$  and  $\delta'[f] = -\delta[f']$  have been used. If the partial derivatives with respect to  $\mathbf{R}_I$  in equation (3.34) are expressed as a linear combination of partial derivatives with respect to atomistic coordinates,  $\partial/\partial \mathbf{r}_i$ , an integration by parts can be performed. Thus equation (3.34) can be expressed in terms of the atomistic forces  $\partial u(\mathbf{r}^n)/\partial \mathbf{r}_i$ . The identity

$$\frac{\partial}{\partial \mathbf{R}_I} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \mathbf{r}_i - \mathbf{R}_I\right) = -\frac{1}{c_{Ik}} \frac{\partial}{\partial \mathbf{r}_k} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \mathbf{r}_i - \mathbf{R}_I\right) \quad (3.35)$$

holds for every  $k \in \mathcal{I}_I$ . Although this identity may be used in equation (3.34), if the atom  $k$  is involved in other CG sites  $J \neq I$ , *i.e.*  $c_{Jk} \neq 0$ , the integration by parts will be complicated by the dependence of the remaining  $(N-1)$  mapping operators  $\mathbf{M}_{\mathbf{R}_J}(\mathbf{r}^n)$  on  $\mathbf{r}_k$ . This is not a problem when the mapping operators are computing the position of the COM, but the solution to such a situation will be derived for completeness. To avoid this, one can define a set of coefficients  $d_{Ii}$  such that  $d_{Ij} \neq 0$  only if  $i \in \mathcal{S}_I$  and such that  $\sum_{j \in \mathcal{S}_I} d_{Ij} = 1$  for all  $I$ . One can then reformulate equation (3.35) as

$$\frac{\partial}{\partial \mathbf{R}_I} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \mathbf{r}_i - \mathbf{R}_I\right) = -\sum_{j \in \mathcal{S}_I} \frac{d_{Ij}}{c_{Ij}} \frac{\partial}{\partial \mathbf{r}_j} \delta\left(\sum_{i \in \mathcal{I}_I} c_{Ii} \mathbf{r}_i - \mathbf{R}_I\right) \quad (3.36)$$

One must assume that for every CG site there is at least one atom which is specific to the site in order to obtain such an equation for every  $I$ . Now the integration by parts can be performed:

$$\mathcal{F}_I(\mathbf{r}^n) \equiv \sum_{j \in \mathcal{S}_I} \frac{d_{Ij}}{c_{Ij}} \mathbf{f}_j(\mathbf{r}^n) \quad (3.37)$$

$$\mathbf{F}_I(\mathbf{R}^N) = \langle \mathcal{F}_I(\mathbf{r}^n) \rangle_{\mathbf{R}^N} = \frac{\int d\mathbf{r}^n \exp(-u(\mathbf{r}^n)/k_B T) \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N) \mathcal{F}_I(\mathbf{r}^n)}{\int d\mathbf{r}^n \exp(-u(\mathbf{r}^n)/k_B T) \delta(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n) - \mathbf{R}^N)}. \quad (3.38)$$

Equation (3.38) provides an expression for the force on the coarse grained site  $I$  for a given CG configuration  $\mathbf{R}^N$  in terms of an equilibrium average of  $\mathcal{F}_I(\mathbf{r}^n)$ , taken over atomic configurations  $\mathbf{r}^n$  which are mapped onto the given CG configuration.  $\mathcal{F}_I(\mathbf{r}^n)$  itself is a linear combination of the atomic forces acting on the atoms which are specific to site  $I$ . Equation (3.38) provides the basis for a variational principle and an applicable algorithm for calculating the many-body potential of mean force (PMF) from simulations of the atomic system.

In order to guarantee consistency not only in configuration space but also in phase space additional conditions are required. Equation (3.31) represents the consistency condition in momentum space. In order to be able to express the right hand side of the equation as a product of separate factors (as the left hand side), we must require that no atom  $i$  is involved in more than one CG site. Therefore, we define the mass of CG site as:

$$M_I = \left( \sum_{i \in \mathcal{I}_I} \frac{c_{Ii}^2}{m_i} \right)^{-1} \quad (3.39)$$

## Variational Principle

We consider the vector space of CG force fields: every element of the vector space is a set of real continuous functions,  $\mathbf{G}(\mathbf{R}^N) = \{\mathbf{G}_1(\mathbf{R}^N), \dots, \mathbf{G}_N(\mathbf{R}^N)\}$ , of the CG configurations  $\mathbf{R}^N$ . We shall refer to these sets of functions as a single function. This vector space contains the atomistically consistent CG force field  $\mathbf{F}(\mathbf{R}^N)$  determined by equation (3.38). The functional

$$\chi^2[\mathbf{G}] = \frac{1}{3N} \left\langle \sum_{I=1}^N |\mathcal{F}_I(\mathbf{r}^n) - \mathbf{G}_I(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n))|^2 \right\rangle \quad (3.40)$$

is defined, where the angular brackets denote an canonical ensemble average within the atomistic model and  $\mathcal{F}_I(\mathbf{r}^n)$  is the atomistic force associated with site  $I$  (see equation (3.37)). The global minimum of such functional is achieved when  $\mathbf{G}_I(\mathbf{R}^N) = \mathbf{F}_I(\mathbf{R}^N)$  for  $I =$

1, ..., N. We define the norm in this vector space as:

$$\|\mathbf{G}\| = \left\langle \sum_{I=1}^N |\mathbf{G}_I(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}^n))|^2 \right\rangle^{1/2} = \left( \int d\mathbf{R}^N \sum_{I=1}^N |\mathbf{G}_I(\mathbf{R}^N)|^2 p_R(\mathbf{R}^N) \right)^{1/2} \quad (3.41)$$

With equation (3.41) we can define a distance between two elements  $\mathbf{G}$  and  $\mathbf{G}'$  of the vector space as  $\|\mathbf{G} - \mathbf{G}'\|$ .

The dimension of the vector space is infinite, we may therefore define a set of  $N_D$  linearly independent vectors  $\mathcal{G}_D(\mathbf{R}^N) = \{\mathcal{G}_{1:D}(\mathbf{R}^N), \dots, \mathcal{G}_{N:D}(\mathbf{R}^N)\}$  for  $D=1, \dots, N_D$  of the vector space (of CG force fields), which form a finite but incomplete basis and consider the subspace spanned by these. Each component of a vector in the given subspace is represented as

$$\mathbf{G}_I(\mathbf{R}^N : \phi) = \sum_{D=1}^{N_D} \phi_D \mathcal{G}_{I:D}(\mathbf{R}^N). \quad (3.42)$$

It can be shown that a minimum of  $\chi^2[\mathbf{G}]$  in this subspace exists and that it is *unique* [77]. This minimum is the member of the subspace which minimises the distance (as defined in equation (3.41)) to the CG force field  $\mathcal{F}(\mathbf{r}^n)$

In the limit that the set of  $N_D$  basis functions is sufficiently complete and provided that the simulation data are sufficiently accurate to evaluate ensemble averages with sufficiently small statistical error an accurate numerical approximation of the many-body force field  $\mathbf{F}$  can be obtained using this variational principle.

In practice (i) the basis set will not be complete and (ii) the data will be affected by some statistical error. The accuracy of the obtained CG force field will therefore rely on the choice of the basis set and the appropriateness of the simulated data.

Once we have the  $N_D$  basis functions and a set  $n_t$  configurations, we can express equation (3.40) as:

$$\chi^2[\mathbf{G}] = \frac{1}{3Nn_t} \sum_{t=1}^{n_t} \sum_{I=1}^N |\mathcal{F}_I(\mathbf{f}_t^n(\mathbf{r}_t^n)) - \sum_{D=1}^{N_D} \phi_D \mathcal{G}_{I:D}(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}_t^n))|^2 \quad (3.43)$$

It is convenient at this point to adopt a more compact matrix notation:  $\underline{\mathbf{f}}$  denotes a column vector of dimension  $3n_tN$  where the three components of every  $\mathcal{F}_I(\mathbf{r}_t^n) = \sum_{i \in \mathcal{I}_I} \mathbf{f}_i(\mathbf{r}_t^n)$  vector are ordered one after another for every CG site (one after another as well) in every sampled configuration  $\mathbf{r}_t^n$ .  $\underline{\phi}$  is a vector of dimension  $N_D$ .  $\underline{\mathcal{G}}_D$  is a column vector of dimension  $3n_tN$  where the three components of every  $\mathcal{G}_{I:D}(\mathbf{M}_{\mathbf{R}}^N(\mathbf{r}_t^n))$  vector are ordered one after another for every CG site (one after another as well) in every sampled configuration  $\mathbf{r}_t^n$ .  $\underline{\mathcal{G}}$  is a matrix of  $N_D$  columns and  $3n_tN$  rows, where every columns is a  $\underline{\mathcal{G}}_D$  vector for  $D=1, \dots, N_D$ . With

such a notation equation (3.43) transforms in:

$$\chi^2[\mathbf{G}] = |\mathbf{f} - \underline{\phi}\underline{\mathcal{G}}| \quad (3.44)$$

and the problem of finding the minimum to  $\chi^2[\mathbf{G}]$  turns into finding the  $\underline{\phi}$  vector which satisfies:

$$\underline{\mathcal{G}}\underline{\phi} = \underline{f} \quad (3.45)$$

Solving equation (3.44) will not be trivial, as in general  $\underline{\mathcal{G}}$  will be big ( $3n_t N \times N_D$ ) and sparse. In this work we use the LSQR algorithm [78, 79] of Paige and Saunders.

### 3.2.2. Effective pair interactions

#### Zero-density limit

In this thesis we are interested in computing the effective interactions of macromolecular polymeric colloids (polymers) when coarse-grained to effective particles placed in its centre of mass. As a first approach, one assumes that the system can be described *via* pairwise additive forces. Defining a cut-off  $R_{max}$  for the radial component of the interaction and dividing the distance from 0 to  $R_{max}$  into  $N_D$  intervals, one can use a set of  $N_D$  basis functions  $f_D(R)$  and express the force acting on a particular CG site with index  $I$  as:

$$\mathbf{F}_I(\mathbf{R}^N) = \sum_{J \neq I} \sum_{D=1}^{N_D} \phi_D f_D(R_{IJ}) \hat{\mathbf{R}}_{IJ} \quad (3.46)$$

with  $R_{IJ} = |\mathbf{R}_J - \mathbf{R}_I|$  and  $\hat{\mathbf{R}}_{IJ}$  the unitary vector pointing in the direction  $\mathbf{R}_J - \mathbf{R}_I$ , so that

$$\mathcal{G}_{I:D}(\mathbf{R}^N) = \sum_{J \neq I} f_D(R_{IJ}) \hat{\mathbf{R}}_{IJ}. \quad (3.47)$$

In this work we use for  $f_D(R)$  linear splines, which are defined as:

$$A_D(R) \equiv \frac{R_{D+1} - R}{R_{D+1} - R_D} \quad (3.48)$$

$$B_D(R) \equiv \frac{R - R_D}{R_{D+1} - R_D} \quad (3.49)$$

$$f_D(R) \equiv \begin{cases} B_{D-1}(R) & \text{if } R_{D-1} < R \leq R_D \\ A_D(R) & \text{if } R_D < R \leq R_{D+1} \\ 0 & \text{otherwise.} \end{cases} \quad (3.50)$$

With the set of  $n_t$  saved configurations ( $\{\mathbf{r}^n, \mathbf{f}^n(\mathbf{r}^n)\}$ ) one can then:

- (i) compute the positions of the centres of mass  $M_{\mathbf{R}}^N(\mathbf{r}^n)$  as a function of the atomic positions *via* the mapping operator in equation (3.23),
- (ii) compute the vector  $\mathbf{f}$  by mapping the atomic forces to the force exerted on the COM of the molecule making use of equation (3.37),
- (iii) use the computed vector  $M_{\mathbf{R}}^N(\mathbf{r}^n)$  and equation (3.46) to compute  $\underline{\mathcal{G}}$
- (iv) solve equation (3.45) and obtain a solution for  $\underline{\phi}$ .

For the set of  $n_t$  saved configurations, a solution will be found for all the coefficients  $\phi_D$ . This means, that the set of configurations must be big enough so that all distances in the  $N_D$  intervals can be sampled.

In a conventional MD simulation at zero-density the two polymers will keep during most of the simulation the equilibrium distance; thus other distances will not be sampled sufficiently often. In such cases, we need to apply some tricks to the MD simulation. We propose the utilization of two springs between the COM of the polymers ( $\mathbf{M}_{\mathbf{R}_1}(\mathbf{r}^n), \mathbf{M}_{\mathbf{R}_2}(\mathbf{r}^n)$ ) and the geometric centre of these two points,  $\mathbf{R}_{\text{geo}} = (\mathbf{M}_{\mathbf{R}_1}(\mathbf{r}^n) + \mathbf{M}_{\mathbf{R}_2}(\mathbf{r}^n))/2$ . The springs exert an extra force on each atom:

$$\mathbf{F}_{\text{spring}}^{1-2}(\mathbf{r}^n) = \alpha (|\mathbf{R}_{\text{geo}} - \mathbf{M}_{\mathbf{R}_{1-2}}(\mathbf{r}^n)| - r_l) \frac{\mathbf{R}_{\text{geo}} - \mathbf{M}_{\mathbf{R}_{1-2}}(\mathbf{r}^n)}{|\mathbf{R}_{\text{geo}} - \mathbf{M}_{\mathbf{R}_{1-2}}(\mathbf{r}^n)|}, \quad (3.51)$$

where  $\alpha$  is the strength of the spring, and  $r_l$  is the equilibrium length of the spring which shall be varied during the simulation in order to sample all the relevant configurations.  $\mathbf{F}_{\text{spring}}^1(\mathbf{r}^n)$  acts on the monomers in  $\mathcal{I}_1$  and  $\mathbf{F}_{\text{spring}}^2(\mathbf{r}^n)$  acts on the monomers in  $\mathcal{I}_2$ . Along a simulation  $r_l$  has been varied from  $r_l = R_{\text{max}}/2$  down to  $r_l = 0$ . The force exerted by the spring can easily be subtracted from the atomic forces, having therefore no effect on the final effective force.

### Finite densities

To compute the effective pair interaction at finite densities we have performed  $NVT$  simulations for a sufficiently large number of polymers  $N$ . Again we need to have enough samples of all the distances between the COMs, therefore we have to simulate our system at densities at which the radial distribution function does not vanish in the  $R$ -range of interest. This, unfortunately, rules out the possibility of computing effective potentials in the solid phase.

# 4. Diffusion in cluster crystals in the explicit presence of a solvent

## 4.1. Introduction

In this part of the thesis we focus on the dynamical behaviour of particles in a GEM cluster crystals. Several features of this hopping processes were investigated in detail by means of molecular dynamics (MD), Monte Carlo (MC) and Brownian dynamics (BD) simulations [24, 25, 30]: it was found that particles propagate in a cluster crystal between lattice sites through an activated hopping mechanism [24, 30] and the question was raised and discussed whether the diffusive motion of the particles corresponds to anomalous diffusion [25], a feature that is typical for Lévy flights [80–83].

However, all the preceding studies did not accurately incorporate the effects of the solvent and the ensuing hydrodynamic interactions (HI). Of course the influence of the HI can be considered to be of less relevance in investigations dedicated to the *static* properties, such as the phase diagram. However, as we proceed to *dynamic* properties, such as diffusion, the solvent often plays a significant role, and therefore should not (or even must not) be neglected in a faithful description of the system. The relevance of the solvent for the dynamics has been shown, for instance, for charged colloidal suspension [32] by means of Schmoluchowski dynamics and for short polymer chains [33] by means of multi-particle collision dynamics simulations (MPCD). We will demonstrate in this contribution that this is also the case for our system at hand.

We *explicitly* take into account the solvent by performing MPCD simulations (see subsection 3.1.2). We investigate the influence of the solvent by varying the number of MD steps performed between every MPCD step. We have considered two set-ups: (i) In section 4.2 we have investigated, similar as in [25], the hopping processes of a pure system of cluster-forming GEM-4 particles. (ii) In section 4.3 we have added non-cluster-forming colloidal (Gaussian) particles to a cluster crystal, similar as in [30], where the clustering particles interact *via* a GEM-8 potential.

## 4.2. Hopping processes in a pure system

We simulate a system of  $N = 3367$  GEM-4 particles.  $\epsilon$  and  $\sigma$  in equation (2.4) will be used as the corresponding energy and length units. We will take  $m_s$ , the mass of the solvent

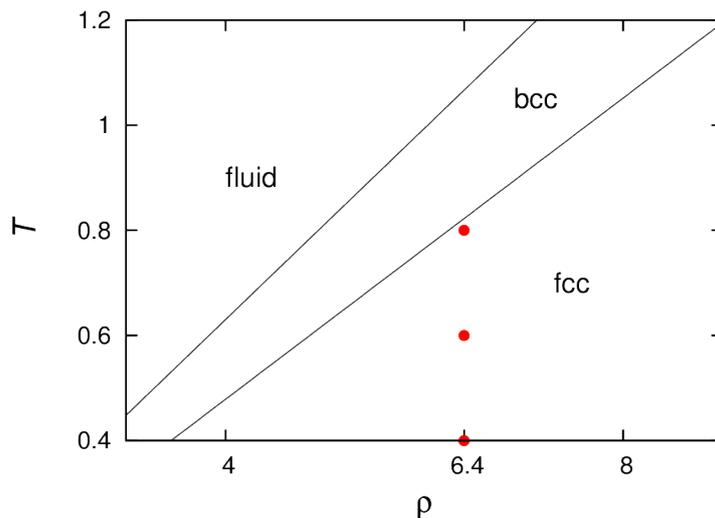


Figure 4.1.: Schematic phase diagram of a system of GEM-4 particles in the  $(T, \rho)$ -plane, indicating the range of stability of the fluid phase and of the bcc and fcc cluster phases respectively. Filled circles specify the systems investigated in this contribution. Taken from [84].

particles as the mass unit and the rest of the quantities will be expressed in reduced units, as specified in appendix A. The particles are arranged in a fcc lattice with a cubic lattice constant  $a = 2.018$ , calculated with equation (2.19). The temperature of the systems is varied between  $T = 0.4, 0.6$  and  $0.8$  in units of  $k_B/\epsilon$ . In figure 4.1 the phase diagram of a GEM-4 system is shown. For convenience we have marked those state points where simulations have been carried out; they all correspond to states where the system forms a cluster crystal. It is clear that they all correspond to cluster crystals.

In what follows we specify the values of the parameters used in the MPCD simulation, defined in subsection 3.1.2: the mass of the solvent particles,  $m_s$ , is used as the mass unit, therefore  $m_s = 1$ . The mass of the GEM particles was then set  $t$  used as the mass unit, therefore  $m_s = 1$ . The angle of the rotation in the collision step is  $\alpha = 130^\circ$ , the cell size is  $c = \sigma$  and every cell is initially filled with 30 solvent particles. The time step for the MD algorithm is  $\Delta t_{MD} = 0.002$ . An MPCD step was performed every  $\tau$  MD steps so that  $\Delta t_{MPCD} = \tau \Delta t_{MD}$ ; for  $\tau$  we assumed the following values:  $\tau = 50, 100, 250, 500$  and  $\infty$  (the latter one corresponds to a pure MD simulation) to study the influence of the solvent. Each simulation was extended over a total time of  $t_{total} = 10000t^*$ , positions and velocities of the colloidal particles were recorded along the simulation.

### 4.2.1. Mean-squared displacement

We first analysed the mean-squared displacement (MSD):

$$\delta r_i^2(t) = \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle \quad \text{for } i = 1, \dots, N. \quad (4.1)$$

From its time dependence, one can draw conclusions whether the particles of the system propagate (predominantly) ballistically, *i.e.*  $\delta r^2(t) \sim t^2$ , or if they show diffusive behaviour (be it either sub-diffusive, *i.e.*  $\delta r^2(t) \sim t^{<1}$ , or normal, *i.e.*  $\delta r^2(t) \sim t^1$ ).

Figure 4.2 shows  $\delta r^2(t)$ , *i.e.* the average of  $\delta r_i^2(t)$  taken over all solute particles as a function of time  $t$ , for the three different temperatures investigated  $T = 0.4, 0.6$  and  $0.8$ . For each temperature different  $\tau$  values have been considered. According to the evolution of  $\delta r^2(t)$ , four different time regimes can be identified:

- (i) the (trivial) ballistic regime at small  $t$ -values;
- (ii) a region where  $\delta r^2(t)$  shows oscillations;
- (iii) a plateau-like region;
- (iv) and finally the long-time regime, corresponding to normal diffusion, *i.e.*  $\delta r^2(t) \sim 6Dt$ , where  $D$  is the diffusion coefficient.

Depending on the combination of the  $T$ - and  $\tau$ -values, some of the regimes are covered by or merged with other regimes.

At high temperatures, we observe for vanishing and weak HI (*i.e.* for  $\tau \gtrsim 250$ ) characteristic high-frequency oscillations due to single particle vibrational modes of the particles within their clusters. With increasing influence of the solvent, these oscillations are gradually suppressed, leading to a plateau that extends approximately over one decade in time. This behaviour is due to the fact that for small  $\tau$ -values the particles collide more frequently with the surrounding solvent particles and thus quickly lose their memory about their original flight direction, *i.e.* effects of inertia become completely negligible.

For larger  $t$ -values, the thermal energy (which is now of the order of magnitude of the energy barrier between two clusters) allows solute particles to spontaneously jump from their original cluster to a neighbouring lattice site, leading to a normal diffusion at large time-scales. The onset of this regime varies by one order of magnitude over the  $\tau$ -range considered in this contribution. However, the presence of the solvent does not have an effect on the diffusion constant.

As the temperature is decreased, the high-frequency modes show the same  $\tau$ -dependence as those observed for  $T = 0.8$ ; they are slightly more pronounced and extend over a significantly larger time-regime. The decreasing temperature has a distinct influence on the long-time behaviour of the mean-squared displacement: with decreasing temperature the plateau region continuously extends over a larger time span. At the lowest temperature considered ( $T = 0.4$ )

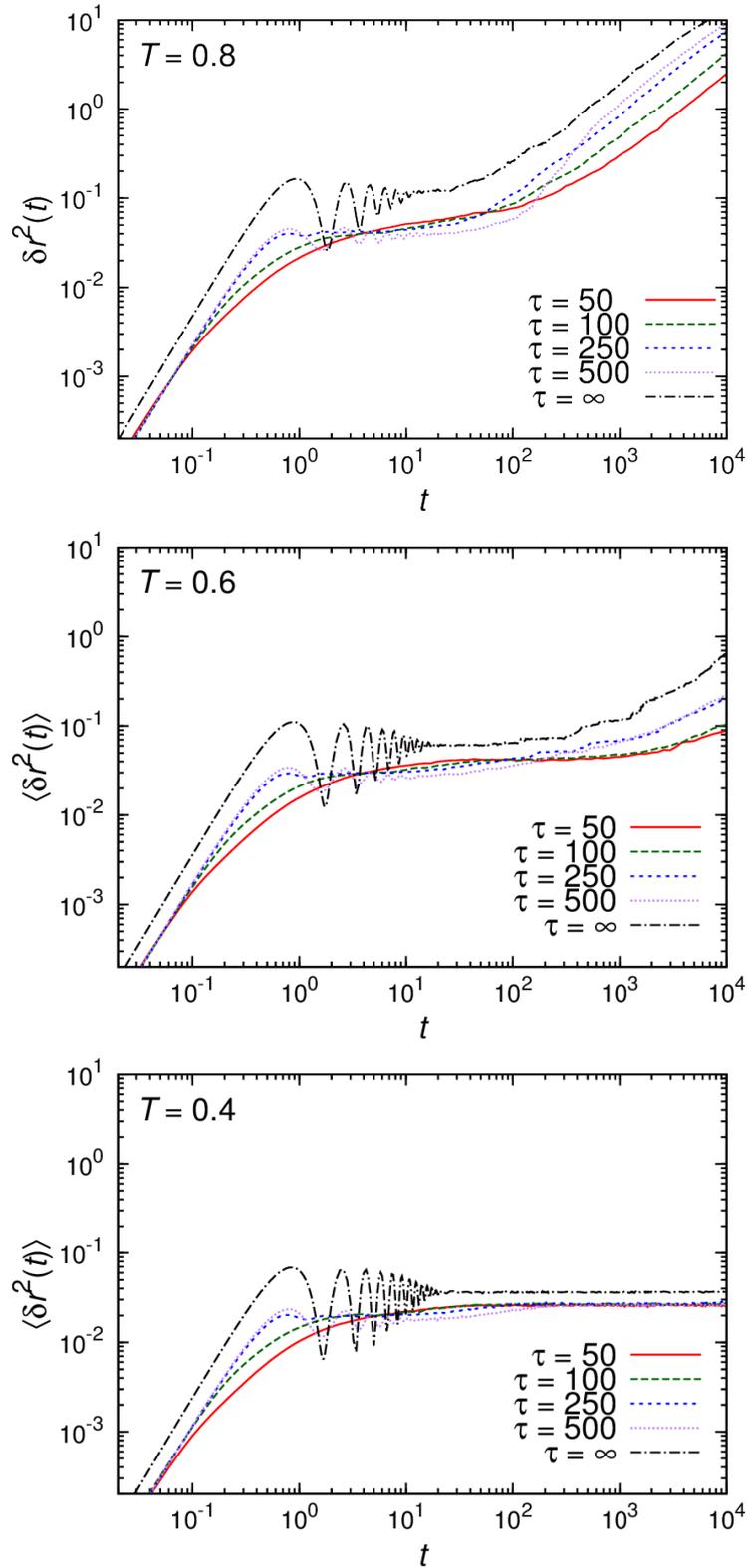


Figure 4.2.: Mean-squared displacement,  $\delta r^2(t)$ , as a function of time  $t$  for three different temperatures as labeled for a system of cluster-forming GEM-4 particles in the presence of a solvent. Results were obtained *via* MD (black) and MPCD simulations with  $\tau = 50$  (red), 100 (green), 250 (blue) and 500 (purple). Taken from [84]

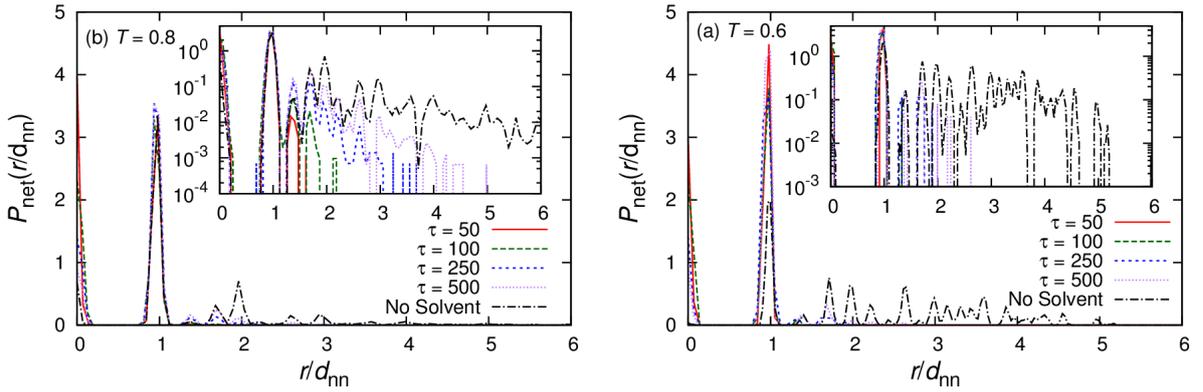


Figure 4.3.: Net jump length distribution  $P_{\text{net}}(r/d_{\text{nn}})$  as a function of  $r/d_{\text{nn}}$  for  $T = 0.6$  and  $0.8$  for different  $\tau$  values for a system of cluster-forming GEM-4 particles in the presence of a solvent on a semi-logarithmic scale. Results were obtained MD (black) and MPCD simulations with  $\tau = 50$  (red),  $100$  (green),  $250$  (blue) and  $500$  (purple). Taken from [84]

the curves of the mean-squared displacement remain flat (i.e. the diffusion constant vanishes) over the observed time range. This is a consequence of the fact that the very small thermal agitations essentially suppress any particle exchange between lattice site.

### 4.2.2. Jump length distribution

In order to obtain a better understanding of the diffusive behaviour of the solute particles we have analysed the individual hopping processes of particles between clusters in more detail. We have used the same algorithm to identify clusters in the system as the one outlined in the appendix of [25]. We have recorded the trajectories of all particles along the simulations runs in terms of the cluster that they belong to. We consider that a hopping event starts when a particle abandons its host cluster, i.e. jumps to a different one. A hopping event finishes when a particle stays in a cluster for a time longer than  $t^*$ . We have chosen  $t^* = 2$ , which is of the order of a few vibrational modes, as can be seen in figure 4.2. We have verified that taking  $t^* = 3$  or  $4$  has no quantitative influence on the results. Three different types of hopping events have been identified:

- (a) particles hops directly from the original cluster to the final clusters;
- (b) the particle visits some intermediate clusters before it reaches the final cluster;
- (c) or the particle visits some intermediate clusters and returns to the original one.

The net jump length of a hopping event is defined as  $r_{\text{net}} = |\mathbf{r}_{\text{final}} - \mathbf{r}_{\text{initial}}|$ , where  $\mathbf{r}_{\text{final/initial}}$  is the position of the centre of mass of the final/initial cluster.

In figure 4.3 we have plotted the net jump length distribution  $P_{\text{net}}(r/d_{\text{nn}})$  as a function of  $r/d_{\text{nn}}$  where  $d_{\text{nn}}$  is the distance to the nearest neighbour for the different  $\tau$  values at  $T = 0.6$  and  $0.8$ . At  $T = 0.4$  no jump events were identified. This correlates nicely to the results in figure 4.2 where the curves of the MSD are flat for  $T = 0.4$ , i.e. no diffusion takes place at all. In contrast, at high temperatures ( $T = 0.6$  and  $0.8$ ) the peaks in figure 4.3 and the linear increase of the MSD in figure 4.2 provide a consistent picture that particle hopping is the dominant mass transport mechanism in cluster crystals. From the data in figure 4.3 it can be seen that (i) at higher temperatures the jump length increases and (ii) that the surrounding solvent plays a crucial role in the jump events of the GEM particles: for small  $\tau$  values  $P_{\text{net}}(r/d_{\text{nn}})$  shows distinct peaks at short distances and rapidly goes to zero with growing distance. In contrast, pure MD simulations display sizeable peaks even for large distances.

These conclusions become more evident from the data compiled in table 4.1, where the length of the largest observed jump for every run is displayed. The net jump length of the hopping

$\tau$	$r_{\text{max}}/d_{\text{nn}}, T = 0.6$	$r_{\text{max}}/d_{\text{nn}}, T = 0.8$
50	1	1.4
100	1.3	2.2
250	1.9	3.6
500	2.6	5
$\infty$	6	28

Table 4.1.: Largest net jump length for the various simulations at  $T = 0.6$  and  $0.8$ .

events can be up to a factor of 20 larger if the solvent is ignored.

The peaks of  $P_{\text{net}}(r/d_{\text{nn}})$  at  $r = 0$  correspond to jump of class (c), i.e. jumps where the particle visits an intermediate cluster and returns to the original one. Integration over the corresponding peaks in figure 4.3 leads to the data visualised in figure 4.4, where the total number of jumps are divided into jumps terminated in the initial clusters, in the nearest neighbour cluster or in any of the remaining clusters. At  $T = 0.8$  and for strong solute-solvent coupling ( $\tau = 50$ ) nearly all jumps are restricted to the cluster of origin (43%) and the nearest neighbour cluster (56%). In contrast, in the complete absence of the solvent ( $\tau = \infty$ ) only very few jumps (7%) terminate in the clusters of origin, while a sizeable number of jumps extend beyond the nearest neighbour distance (36%). At  $T = 0.6$  we observe the same behaviour. Thus a strong coupling between the solute and the solvent particles seems to introduce an additional energetic barrier which hinders long net jump lengths.

Finally from the data presented in figure 4.3 we can definitely exclude the occurrence of Lévy flights in cluster crystals in the presence of an explicit solvent: as the data displayed in figure 5 of [25]  $P_{\text{net}}(r)$  shows in the absence of a solvent (i.e. in MD simulations) a power law decay, i.e.  $\sim 1/r^{1+\alpha}$ , with  $\alpha \sim 2.2$  [25], which represents an  $\alpha$ -value which is slightly larger than the upper limit required for Lévy flights [80–83]. Increasing the influence of the solvent

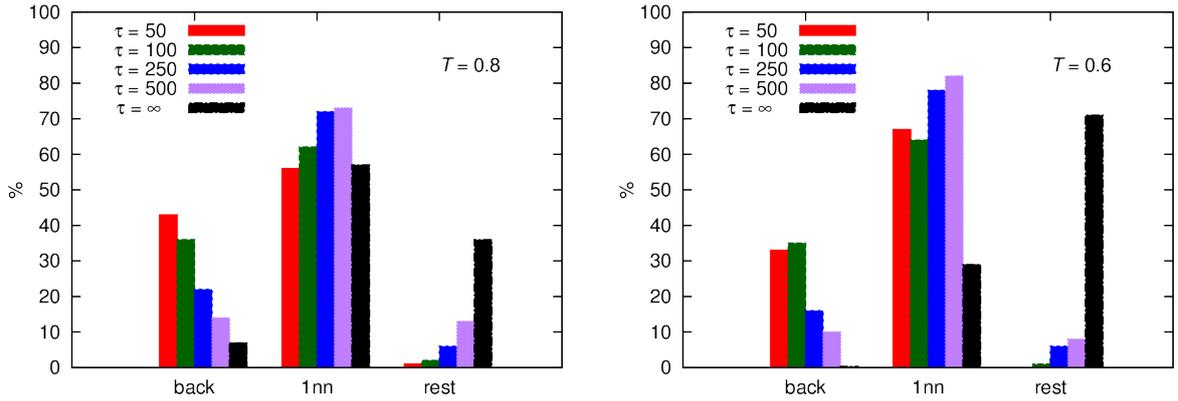


Figure 4.4.: Integrated values of the net jump length distribution,  $P_{\text{net}}(r/d_{\text{nn}})$  in figure 4.3, integrated over distances corresponding to the initial clusters (“back”), the nearest neighbour clusters (“1nn”) and the remaining clusters (“rest”) in per cent for different  $\tau$  values. Taken from [84].

by decreasing  $\tau$  we observe that the long-distance behaviour of  $P_{\text{net}}(r)$  is characterised by increasingly larger  $\alpha$ -values.

### 4.2.3. Distribution of the angle enclosed by successive jumps

To further characterise the nature of the long range jumps in the cluster crystal we have extracted from our simulation data the correlation between the net jump length  $r_{\text{net}}$  and the angle  $\theta$  enclosed by two successive segments of the trajectory of a jump event (schematic representation in the right panel of figure 4.5). The left panel of figure 4.5 shows the possible consecutive jumps in the fcc unit cell. We consider all jumps of type (b) and (c), as classified earlier in the text, and consider all the angles enclosed by the different segments of the jump.

In figures 4.6 and 4.7 we have depicted the the encountered angles  $\theta$  as functions of the net jump length (empty grey circles), the distribution of the net jump length  $P_{\text{net}}(r/d_{\text{nn}})$  as in figure 4.3 (solid line) and the average over all the encountered angles  $\theta$  at a given  $r_{\text{net}}$  value,  $\langle \theta \rangle$  as a function of the net jump length  $r_{\text{net}}$  (filled coloured circle). Figure 4.6 shows the results for  $T = 0.8$  and figure 4.7 for  $T = 0.6$  for all the different  $\tau$  values considered. The  $\tau$ -dependence of the results is the same for both temperatures: in the explicit presence of a solvent, the majority of jumps are rather short ranged. The preferred  $\theta$  values are  $\theta = 60^\circ$  and  $90^\circ$ , as the particles only jump back and forth to their second nearest neighbours, but do not move along straight trajectories through the crystal. This picture completely changes when the HI are neglected: especially for longer jump lengths,  $\theta$  values of  $120^\circ$  and  $180^\circ$  dominate, corresponding to slightly deflected or perfectly straight pathways trough the crystal. Thus, the incorporation of HI significantly modifies the diffusion behaviour both on a quantitative but also qualitative level: while in the MD simulations the diffusion is dominated by long ranged and strongly correlated ballistic flights, the explicit presence of a solvent reduces the

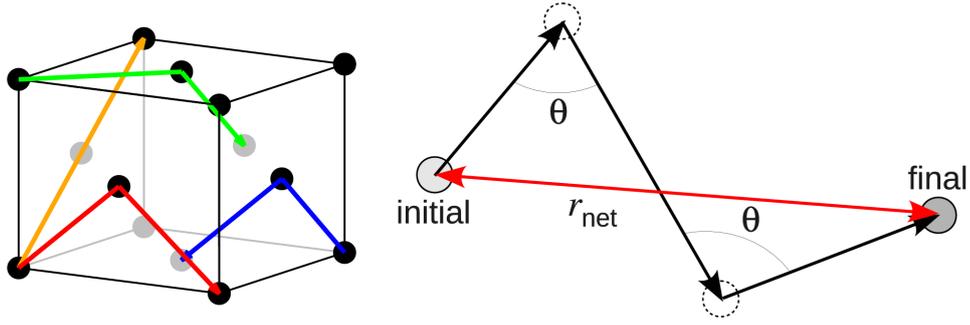


Figure 4.5.: Left panel: figure 5 of [25], showing the possible consecutive jumps of a particle to nearest neighbours in the fcc unit cell. The full black circles indicate the fcc lattice sites. Blue arrow:  $60^\circ$ ,  $r = d_{\text{nn}}$ . Red arrow:  $90^\circ$ ,  $r = \sqrt{2}d_{\text{nn}}$ . Green arrow:  $120^\circ$ ,  $r = \sqrt{3}d_{\text{nn}}$ . Orange arrow:  $180^\circ$ ,  $r = 2d_{\text{nn}}$ . Right panel: Schematic sketch of a jump event, visualizing the net jump length,  $r_{\text{net}}$  and the related angles  $\theta_1$  and  $\theta_2$ . Taken from [84]

responsible mechanism for mass transport to random jumps to nearby lattice sites.

Figure 4.8 shows the entire trajectory of two selected jumps at  $T = 0.8$ , the left panel corresponds to a simulation with  $\tau = \infty$  and right one to  $\tau = 50$ . The entire trajectory can be seen in the following videos:

- $\tau = \infty$  → [https://www.youtube.com/watch?v=zix85\\_6ropM](https://www.youtube.com/watch?v=zix85_6ropM)
- $\tau = 50$  → <https://www.youtube.com/watch?v=EG30bVmfYd8>.

The videos show a short clip of the simulation. The GEM-4 particles are showed as semi-transparent green spheres. A particle which will perform a hopping event within this time window is selected and drawn as a solid blue sphere. For a better visualization of the trajectory the particle leaves a blue wake behind it as it moves.

The reader can see a nice wrap-up of the results: in the video for  $\tau = \infty$  it can be observed how the particles oscillate around the centre of mass of the cluster they belong to, corresponding to the oscillations observed in figure 4.2 for small time values, whereas in the video for  $\tau = 50$  these oscillations are lost due to the collisions with the solvent which are appreciable from the bumpy trajectories of the particles. The particle in the video for  $\tau = \infty$  performs a long jump following a straight trajectory which is interrupted when the particle passes close by the centre of mass of a cluster by which it is attracted. The first times that this happens the particle enters an orbit around the cluster and continues its trajectory through the crystal. As the particle slowly loses its kinetic energy (that enables it to abandon its host cluster) it temporarily oscillates around one cluster performing now smaller orbits for a time shorter than  $t^*$  (and therefore the jump does not terminate here). The particle is still able to escape this cluster. Finally the particle has lost too much kinetic energy and is eventually trapped in the terminal cluster. The particle in the video for  $\tau = 50$  performs a completely different trajectory: once the particle leaves its cluster, it is not able

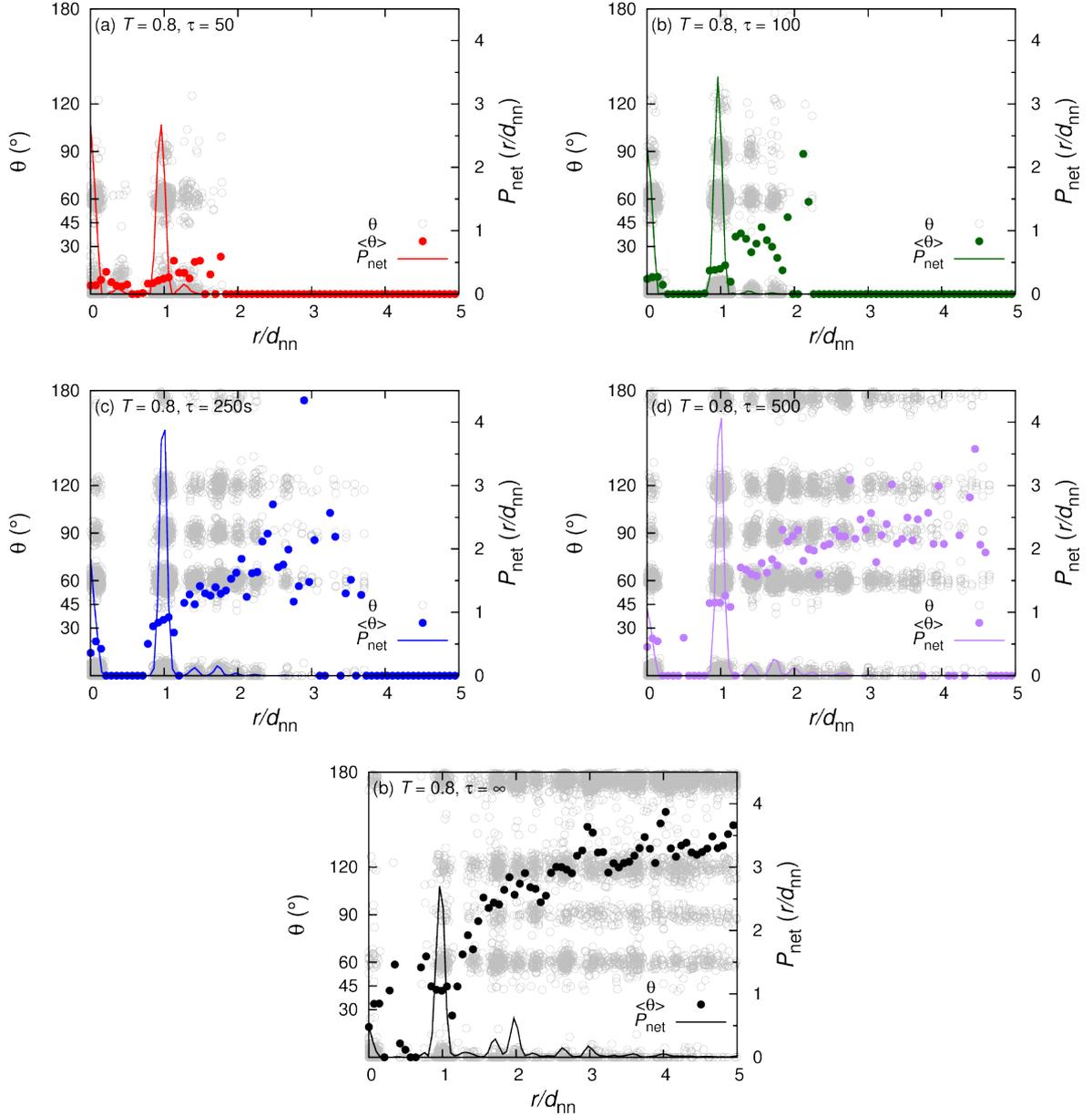


Figure 4.6.: Correlation between the angle  $\theta$  and the net jump length  $r_{\text{net}}$  (see figure 4.3) of a jump event as a function of  $r/d_{\text{nn}}$  at  $T = 0.8$  for different  $\tau$  values for a system of cluster-forming GEM-4 particles in the presence of a solvent. The solid lines correspond to the net jump length distribution  $P_{\text{net}}(r/d_{\text{nn}})$ . Open grey circles represent  $\theta$  values identified in the trajectory of a jump event that extends over a net jump length of  $r/d_{\text{nn}}$  and filled circles correspond to the average of all  $\theta$ -values at a fixed jump length, denoted by  $\langle \theta \rangle$ . Taken from [84]

#### 4. Diffusion in cluster crystals in the explicit presence of a solvent

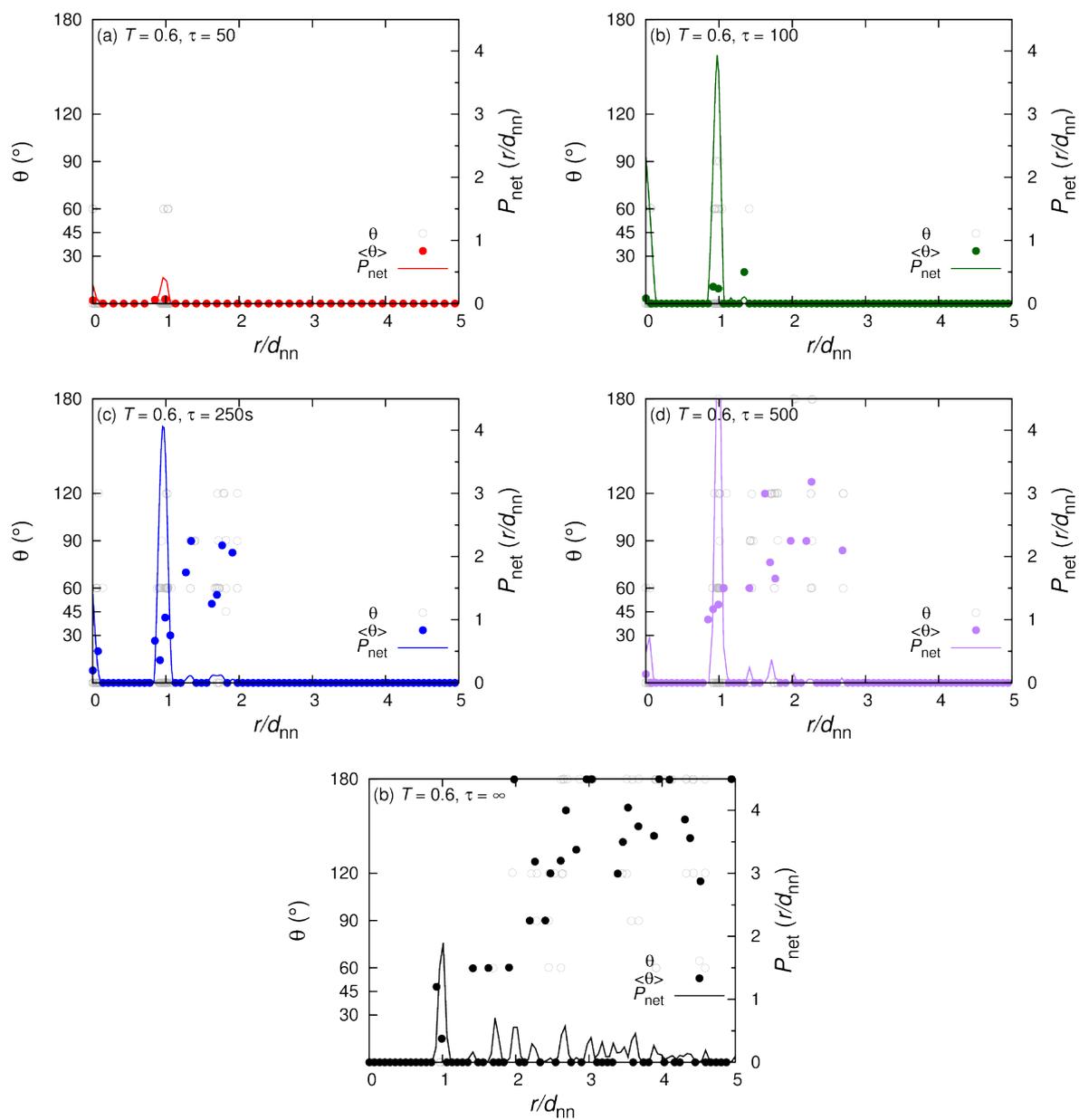


Figure 4.7.: Same as figure 4.6 but for  $T = 0.6$ .

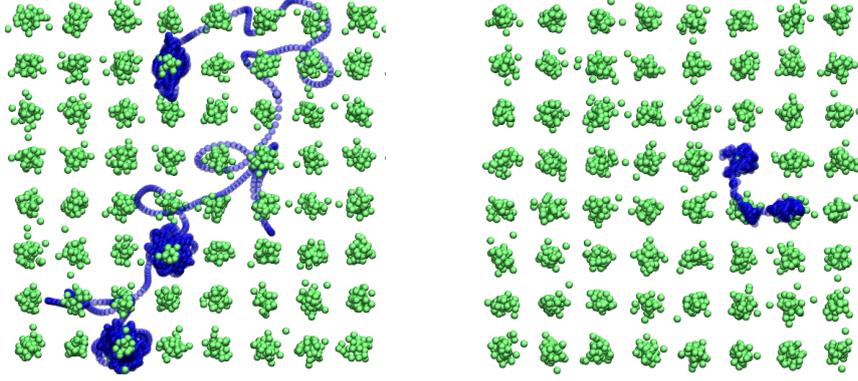


Figure 4.8.: Selected jump trajectories for  $T = 0.8$  observed over the time window in which the jump takes place. The panels correspond to a simulation with  $\tau = \infty$  (left panel) and  $\tau = 50$  (right panel). The GEM-4 particles are showed as semi-transparent green spheres. A particle which will perform a hopping event during the time window of observation is selected and marked as a solid blue sphere. For a better tracking of its trajectory the particle leaves a blue wake behind it as it moves.

to move as fast as its counterpart in the video for  $\tau = \infty$  due to its collisions with the solvent. It quickly loses the kinetic energy and is trapped in a nearest neighbour cluster.

#### 4.2.4. Self-part of the van Hove correlation function

Finally, we have studied the influence of the properties of the solvent on time-dependent correlation functions through the self-part of the van Hove correlation function, defined as:

$$G_s(r, t) = \left\langle \frac{1}{N} \sum_{i=1}^N \delta[r - |\mathbf{r}_i(t) - \mathbf{r}_i(0)|] \right\rangle. \quad (4.2)$$

This function correlates the position of a tagged particle (with index  $i$ ) at a time  $t = 0$  with its position at some time  $t > 0$ . It thus provides information about the spatio-temporal correlations of a particle [85, 86].

The results for  $G_s(r, t = 6000)$ , evaluated at  $T = 0.8$ , are depicted in figure 4.9. For all  $\tau$ -values considered in this contribution we observe cascades of peaks (with two pronounced peaks at  $r = 0$  and  $d_{\text{nn}}$ ). In contrast, the exponential decay at large distances depends in a highly sensitive way on  $\tau$ : smaller  $\tau$ -values (and therefore smaller mean free paths) lead to a faster decay, meaning that the hopping processes are limited to shorter jump lengths. For completeness we mention that for none of the  $\tau$ -values and time ranges investigated in this contribution a Gaussian shape of the self-part of the van Hove correlation function has been observed [85, 86]: however, we expect to observe this behaviour at considerably larger  $t$ -values.

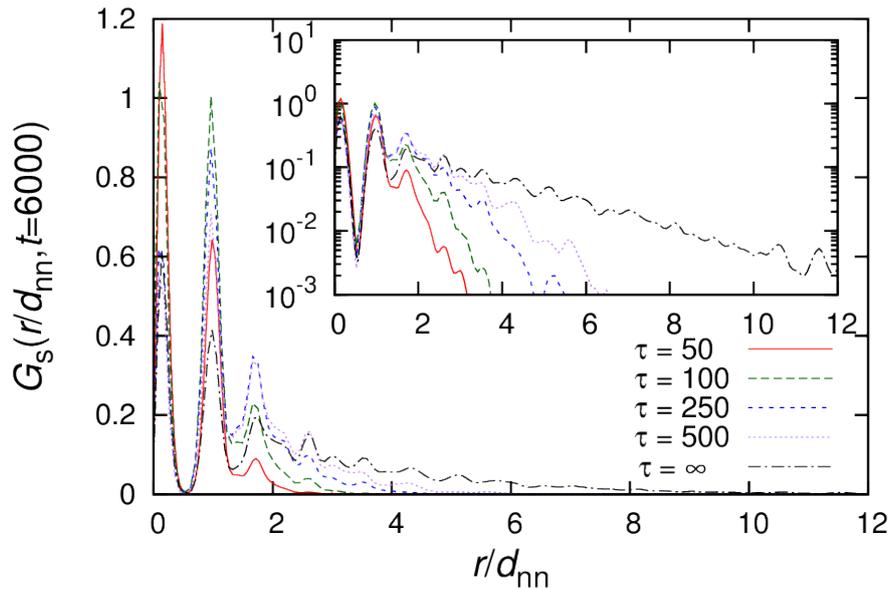


Figure 4.9.: Main panel: self-part of the van Hove correlation function,  $G_s(r, t)$ , at  $t = 6000$  as a function of  $r/d_{nn}$  at  $T = 0.8$  for a system of cluster-forming GEM-4 particles. Results were obtained *via* MD (black) and MPCD simulations with  $\tau = 50$  (red), 100 (green), 250 (blue) and 500 (purple). The inset shows the same data on a semi-logarithmic scale. Taken from [84]

### 4.3. Hopping processes in a binary system

Now that the migration of cluster-forming particles through cluster crystals in the explicit presence of a solvent is better understood, we consider a more complex scenario in what follows: we add ultrasoft, *non-cluster-forming* particles to the set-up considered in section 4.2. In the following, we will denote these particles as type-A particles, while the cluster-forming particles will be specified by the label B (introducing their concentration,  $c_B$ ). We have chosen the index  $n = 8$  for the GEM potential describing the interaction between the type-B particles. The choice for this  $n$ -value is motivated by the investigations presented in [30] and is justified by the fact that the change from  $n = 4$  (section 4.2) to  $n = 8$  (section 4.3) will affect the properties of the cluster crystal only on a *quantitative* but not on a *qualitative* level. The additional mesoscopic particles interact *via* a GEM-2 (Gaussian) potential, guaranteeing that these particles definitely will not form stable clusters [11]. For the cross interaction between the two types of ultrasoft particles we have assumed – similar as in [30] – a GEM-4 potential, *i.e.* an interaction that also supports cluster formation. For the type dependent  $\sigma$ -parameter entering in equation (2.4) we chose  $\sigma_{AA} = 0.3\sigma$ ,  $\sigma_{AB} = 0.6\sigma$ , and  $\sigma_{BB} = \sigma = 1$ . We investigate the system both in an MD simulation ( $\tau = \infty$ ) and an MPCD simulation ( $\tau = 50$ ).

We consider again three different temperatures,  $T = 0.8, 0.6$  and  $0.4$ , and four different

values for the concentrations of the type- $B$  particles,  $c_B = 0.65, 0.8, 0.95$  and  $1.0$ . For every concentration we choose a suitable box size such that a fcc lattice with four unit cells in each spatial direction can fit in it. The spacing of the lattice is obtained *via* the theoretical derivations in [87]. As argued in this contribution on the basis of a simple mean field approximation [87, 88], the corresponding states are selected to be located on the so-called  $\lambda$ -line, *i.e.* where the system is prone to crystallise. The total number of particles is adjusted such that we achieve a total density  $\rho = 6.4$ , as in section 4.2. The parameters of the system are specified in table 4.2

$c_B$	$a$	$L$	$\langle N_c \rangle$
0.65	1.930	7.718	7.57
0.80	1.896	7.584	8.73
0.95	1.864	7.456	9.76
1.00	1.857	7.428	10.25

Table 4.2.: System parameters of all investigated binary systems: concentration of the cluster-forming type-B particles,  $c_B$ , lattice constant of the cluster crystal,  $a$  (in units of  $\sigma$ ), simulation box size  $L$  (in units of  $\sigma$ ), and average cluster size  $\langle N_c \rangle$ .

### 4.3.1. Mean-squared displacement

We start again our discussion with the mean-squared displacement presented in figure 4.10. We present our results for type-A particles,  $\delta r_A^2(t)$ , in the left column and for type-B particles,  $\delta r_B^2(t)$  in the right column. data for the three different temperatures are shown in rows, as labelled, for all the different values of the concentrations  $c_B$  considered. MPCD results are shown with solid lines and MD with dot-dashed ones. The results for all three temperatures are on a qualitative level comparable. With decreasing temperature (and the ensuing reduced thermal agitations) the hopping processes become less frequent. Reading from figure 4.10 we can see that we will not find any hopping events at  $T = 0.4$  (both for MD and MPCD at all concentrations) and at  $T = 0.6$  these will only occur for  $c_B = 0.65$  in the MD simulation. We therefore only concentrate in what follows on the results at  $T = 0.8$ .

A first, qualitative analysis of the mean-squared displacement of the type-A (GEM-2) particles,  $\delta r_A^2(t)$  shown in the left panels of figure 4.10, reveals the expected and rather rapid cross-over from the short-time ballistic motion to the long-time diffusive behaviour; the fact that – irrespective of the concentration  $c_B$  and the simulation scheme – all  $\delta r_A^2(t)$ -curves are characterised in their long-time limit by the same slope provides evidence that the diffusion constant has in all cases the same value. Thus the cluster-crystal has essentially no impact on the diffusion constant of the mobile particles. Further, in none of the curves an intermediate oscillatory region is observed, indicating that the A-particles behave as a fluid confined in a crystalline matrix of clustered B-particles. Throughout, a higher concentration

in B-particles delays the onset of the diffusive behaviour: obviously due to the increase in the number of type-B particles, the interstitial, non-cluster-forming type-A particles experience through the GEM-4 cross interaction stronger bonds to the clusters which delay the diffusion and thus slow down the dynamics. Considering now the results of each type of the simulation schemes separately, we observe distinctive differences: in the MD simulations we see a rather direct and slightly concentration-dependent transition from the ballistic to the diffusive regime. In the MPCD simulations on the other hand, the ballistic behaviour terminates at  $t \sim 10^{-1}$ , followed by an intermediate regime that extends roughly over three time-decades and is characterised by an exponent of  $t$  of approximately 1.25 to 1.3; eventually at  $t \sim 10^2$  the diffusive behaviour sets in [see also the discussion of  $G_s^A(r, t)$  below]. The fact that in the explicit presence of the solvent the onset of the diffusive behaviour is delayed by two orders of magnitude with respect to the MD simulation data emphasises the eminent role of the solvent.

For the type-B (GEM-8) particles we observe a similar time dependence of  $\delta r_B^2(t)$  as the one observed for the pure GEM-4 case (see section 4.2 and figure 4.2): the MD-data for  $\delta r_B^2(t)$  show – after the ballistic regime – pronounced oscillations for all concentrations investigated, reflecting single particle vibrational modes of the particles in the clusters; eventually – and strongly dependent on  $c_B$  – the linear diffusive behaviour sets in. The fact that the onset of the diffusive regime shifts with increasing concentration of type-B particles to larger times is related to the fact that cluster-crystals with higher  $c_B$ -values are characterised by larger occupancy numbers (see table 4.2); these larger clusters exert stronger binding forces on the mobile cluster-forming type-B particles and thus delay their diffusive behaviour. In the MPCD simulations, these oscillations are essentially suppressed due to the presence of the solvent. Instead we find a broad time-regime where  $\delta r_B^2(t)$  is essentially constant. Only towards the end of the investigated time range (*i.e.*  $t \sim 10^4$ ), a  $c_B$ -dependent onset of the linear behaviour can be observed: the smaller  $c_B$ , the earlier the diffusive behaviour sets in (see also the related discussion in the previous paragraph).

### 4.3.2. Self-part of the van Hove correlation function

We conclude this section by evaluating the self-part of the van Hove correlation functions,  $G_s^A(r, t)$  and  $G_s^B(r, t)$ , for different values of concentration  $c_B$  and different values of time  $t$ ; throughout,  $T$  was set to 0.8. The respective results are plotted in figure 4.11. By comparing the data presented in panels a and b, we can trace how  $G_s^A(r, t)$  develops at relatively short times: for very small  $t$ -values (*i.e.*,  $t = 80$ ), a strong correlation between the non-clustering type-A particles and the solvent particles can be observed, reflected in the MPCD data in well-defined peaks for  $r \sim d_{nn}$ , which emerge for all concentrations; at  $t = 800$ , these peaks are smeared out over larger distances, but are still sizeable. In contrast, the MD simulation results for  $G_s^A(r, t = 80)$  show no particular structure, indicating the rapid loss of spatio-temporal correlations of the type-A particles and at  $t = 800$  this function has essentially vanished. This distinct difference emphasises the eminent role of the solvent in supporting

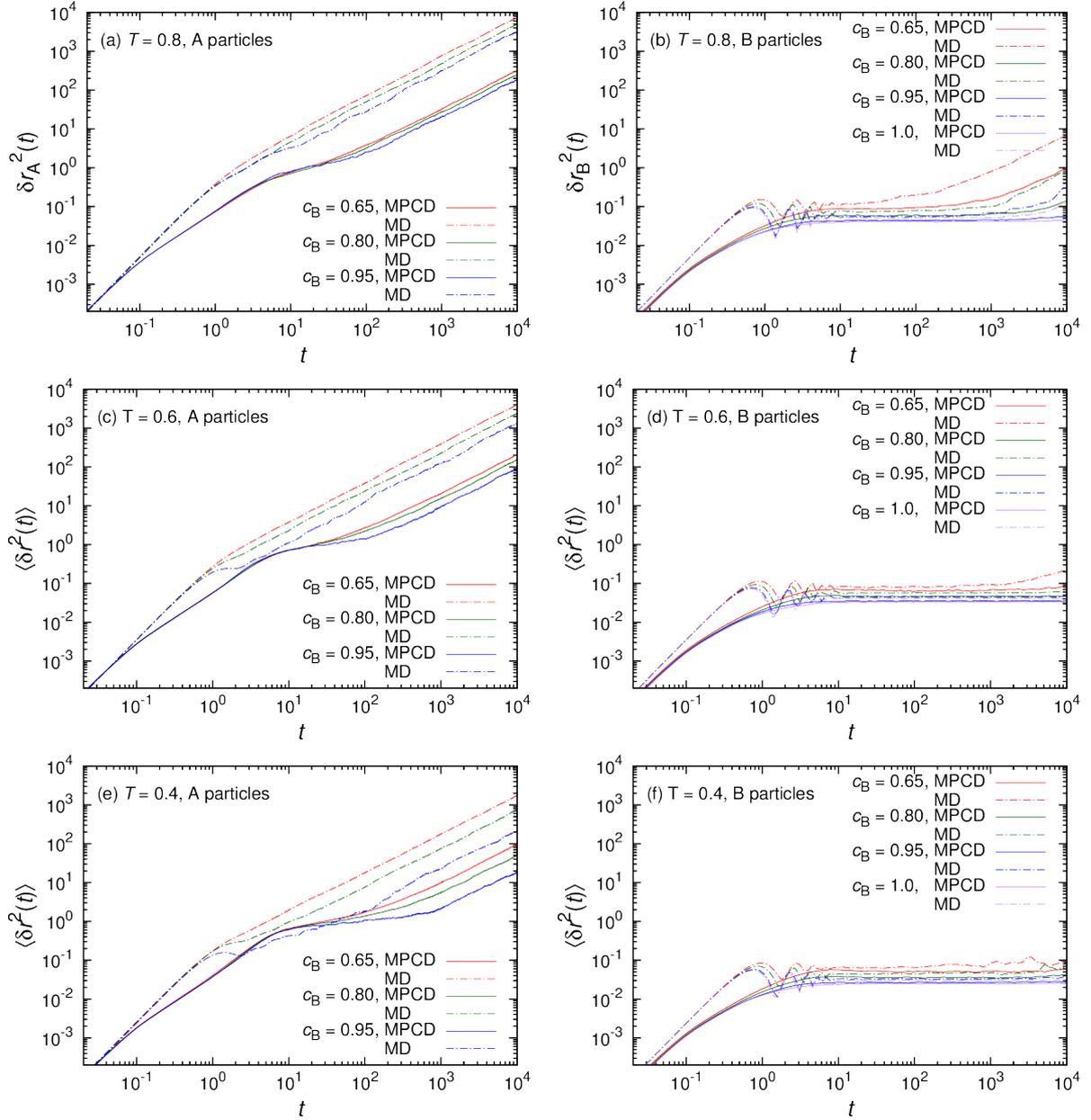


Figure 4.10.: Mean-squared displacement,  $\delta r^2(t)$ , as a function of time  $t$ , left panel for the type-A particles and right one for type-B particles, for three different temperatures as labeled for a system of cluster-forming GEM-8 particles in the presence of non-cluster-forming GEM-2 particles and a solvent. Dot-dashed lines correspond to MD simulations results, while solid lines display MPCD data. Different colours correspond to different values of  $c_B$  (as-labelled). Taken from [84]

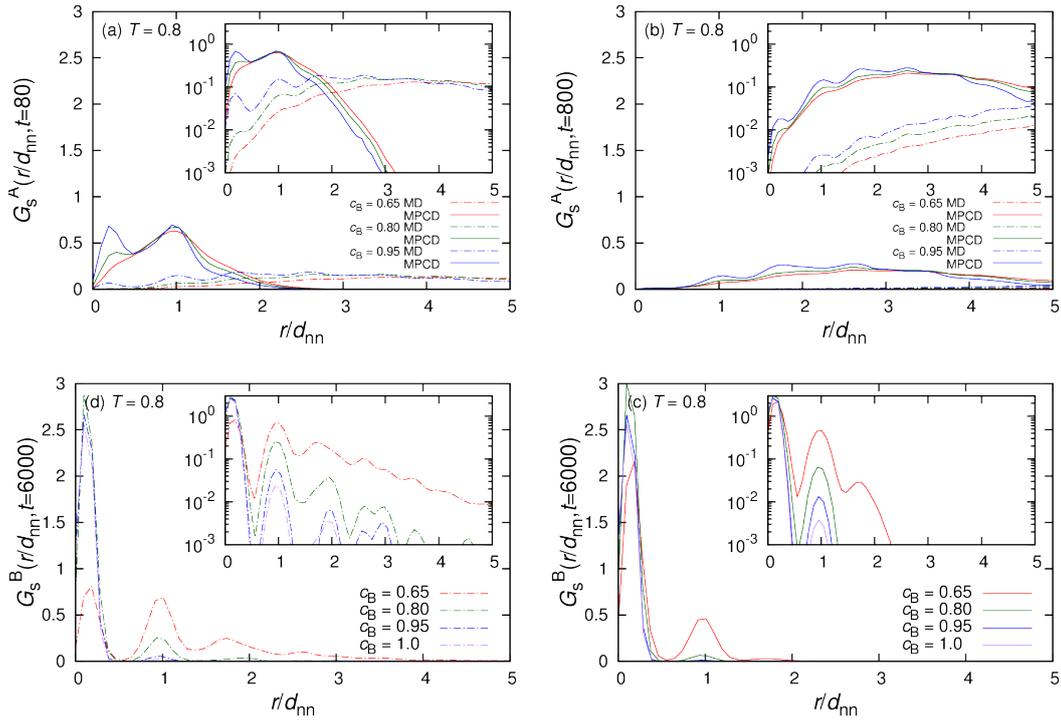


Figure 4.11.: Self-parts of the van Hove correlation functions,  $G_s^A(r, t)$  and  $G_s^B(r, t)$ , as functions of distance. The upper panels show the results for the type-A particles at  $T = 0.8$  at  $t = 80$  (left) and  $t = 800$  (right) for concentrations of the type-B particles  $c_B = 0.65, 0.8$  and  $0.95$  as labelled. The lower panels show the results for the type-B particles at  $T = 0.8$  at  $t = 6000$  observed *via* MD (left) and MPCD (right) simulations for concentrations of the type-B particles  $c_B = 0.65, 0.8, 0.95$  and  $1.0$  as labelled. Taken from [84]

spatio-temporal correlations between the mobile, non-cluster-forming type-A particles over relatively large time-spans. This phenomenon is also reflected by the intermediate region observed in  $\delta r_A^2(t)$  for  $10^{-1} \lesssim t \lesssim 10^2$  discussed above. Concluding we note that the loss of spatio-temporal correlations is fully confirmed by the Gaussian shape of  $G_s^A(r, t)$  [85, 86]; the diffusion constants, that we obtain from the respective fits are in satisfactory agreement with the respective values extracted from the mean-squared displacements.

Finally,  $G_s^B(r, t)$ , is displayed in panels c and d of figure 4.11 for  $t = 6000$ . Within the MPCD simulation scheme, this function shows for all values of  $c_B$  pronounced peaks at  $r \sim 0$  and  $r \sim d_{nn}$  and then vanishes rapidly for larger distances, providing evidence that the solvent significantly slows down the mobility of the cluster-forming particles. In contrast, in the absence of the solvent well-defined peaks up to  $r \sim 3d_{nn}$  are visible in the corresponding data of the MD simulations. In both cases, a strong  $c_B$ -dependence in the heights of the peaks is visible.

# 5. Compression experiments on a two dimensional GEM-4 system

## 5.1. Introduction

So far, the static properties of ultrasoft cluster-forming systems have been studied extensively, providing evidence that they show quite a few unexpected and intriguing features as compared to their hard matter counterparts [21, 22]. In contrast, only very little is known about the response of these systems to external fields such as pressure or mechanical shear [31]. This chapter is dedicated to the study of the response of an ultrasoft nanocrystal of particles interacting *via* the GEM-4 potential to an external hydrostatic pressure. Our investigations are motivated by the fact that previous studies on the equilibrium properties have provided evidence that cluster crystals exhibit a highly unusual response to compression, due to the inherent softness of the constituent particles: increasing the density in cluster crystals does not induce a decrease in the lattice constant (as one would intuitively expect), but leads instead to a linear increase in the average occupation number per lattice site [21, 23, 37]. As mentioned in section 2.3, DFT calculations within the mean field approximation predict for cluster crystals a density independent lattice constant. This conjecture was confirmed by Mladek *et al.* [21] *via* independent *NVT* simulations performed at different density values. In these simulations the authors found a linear dependence of the cluster occupation as a function of density (see figure 4 of [21]).

The volume of the cluster crystal is given by

$$V = N_{\text{cells}} a^3, \quad (5.1)$$

where  $a$  is the spacing of the lattice, as defined in section 2.3 and  $N_{\text{cells}}$  is the number of unit cells in the crystal. At the same time,  $N_{\text{cells}}$  can be expressed as

$$N_{\text{cells}} = \frac{N}{z \langle n_{\text{occ}} \rangle}, \quad (5.2)$$

where  $z$  is the lattice dependent constant value, defined in section 2.3 and  $\langle n_{\text{occ}} \rangle$  is the average cluster occupation in the crystal. If we insert equations (5.1) and (5.2) in the definition of the density  $\rho$  we end up with

$$\rho = \frac{N}{V} = \frac{z n_{\text{occ}}}{a^3}. \quad (5.3)$$

Therefore, a linear dependence of  $n_{\text{occ}}$  on  $\rho$  confirms that the lattice spacing  $a$  is  $\rho$ -independent.

Thus, it is clear that after a compression, and once equilibrium has been recovered, some lattice positions must have disappeared and particles which were populating the clusters in those lattice positions must have migrated to other surviving clusters. This redistribution of particles is realised *via* two key processes, namely *particle hopping* and *cluster merging*. Particle hopping is a characteristic transport mechanism in cluster-forming systems [24, 25, 84] where particle jumps extending over multiple lattice site are commonplace (both in- and out-of-equilibrium). In contrast, merging of clusters essentially never occurs under equilibrium conditions due to the high energy barriers that separate adjacent clusters in the crystal. Under these conditions merging of two clusters would be an energetically highly unfavourable process. This novel response to compression is unique for ultrasoft clustering systems, setting them apart from conventional hard matter crystals, which react to pressure exclusively *via* a reduction of the lattice constant.

We have exposed an ultrasoft nanocrystal of a limited number of particles interacting *via* a GEM-4 potential (equation (2.4)) to compression. For computational reasons and in an effort to enhance the transparency of the processes we have restricted ourselves to a two-dimensional nanodrop where the clusters form a regular, hexagonal lattice. The motivation to study a finite system instead of a bulk crystal lies on our initial hypothesis that merging events will occur close to the surface that establishes the contact of the nanocrystal with the pressure bath. In an effort to study these effects we have realised *NPT* conditions by bringing the system into contact with an external combined pressure- and temperature-bath, formed by ideal gas particles [36, 89, 90]. The target pressure ( $P_t$ ) and target temperature ( $T_t$ ) of the combined baros- and thermostat are controlled *via* the number of ideal gas particles and their velocities, respectively. Pressure and temperature are transferred from this reservoir to the nanocrystal through a repulsive cross-interactions. The trajectories of all particles involved are calculated in standard molecular dynamics simulations. The pressure or temperature is increased repeatedly in steps of  $\Delta P$  or  $\Delta T$  at fixed  $T_t$  or  $P_t$ , thereby temporarily inducing non-equilibrium conditions under which both hopping and merging processes can occur.

In section 5.2 we give a detailed description method used to simulate the pressure and temperature bath. In subsection 5.2.1 we elaborate on the protocol used to compress and expand or heat and cool the nanocrystal. Section 5.3 presents the results of this chapter grouped in three subsections: subsection 5.3.1 elaborates on the effect of using different compression rates in a compression experiment. We evaluate how different  $\Delta P$  values affect the final structural conformation of the nanocrystal and we examine under which circumstances the nanodrop is able to recover equilibrium. In subsection 5.3.2 we analyse the results of several compression experiments performed at various temperatures with different compression rates. We study the path followed by the system to recover equilibrium:

- which clusters will disappear?
- What will happen with the particles belonging to those clusters?

- Does the distribution of the kinetic energy in the system play a role in the response of the nanocrystal to compression? Do over-heated clusters react in a different way than the under-heated ones?
- Is the response of the system affected by the temperature and the compression rate at which the experiment is performed?
- What role do hopping mechanisms play?
- Where do the merging events occur preferentially? In the core of the nanocrystal or in the surface in contact with the pressure bath?

Finally in subsection 5.3.3 we will discuss how we can extract the equation of state of the system from a combination of these compression experiments with subsequent annealing runs.

## 5.2. Method

We follow the method for applying hydrostatic pressure in a simulation of a nanocrystal presented by Grünwald *et al.* in [36, 89, 90], using an ideal gas as the pressure medium and thermostat. We surround a system of particles interacting *via* the GEM-4 potential by an ensemble of ideal –non-interacting– gas particles. The cross interaction between the ideal gas and the GEM-4 particles is given by:

$$\Phi(r) = \epsilon_b \left( \frac{\sigma_b}{r} \right)^{12}, \quad (5.4)$$

where  $\sigma_b$  is the interaction range and  $\epsilon_b$  the interaction strength. We set  $\epsilon_b = 1$  and cut and shift the potential so that it vanishes at  $r_c = 2\sigma_b$ . The choice of  $\sigma_b$  is guided by two considerations: on the one hand a large  $\sigma_b$  will produce a large number of interaction partners, which will worsen the efficiency of the algorithm due to longer force calculations. On the other hand, if  $\sigma_b$  is too small the gas particles will be able to penetrate in the GEM-4 system, and this unphysical effect has to be avoided. The size of  $\sigma_b$  is therefore determined by the crystal structure. In our simulations we found  $\sigma_b = 1$  to be an optimum value.

As there is no interaction between the ideal gas particles, the ideal gas is described by the statistics of the particles enclosed in finite volume around the GEM-4 system. The surface of this volume is given by the overlapping spheres of radius  $r_c$  centred in the particles of the nanocrystal surface. For computational reasons this minimum interaction volume is replaced by a more generous volume, built up by cubic cells (of length  $r_c$ ) and delimited by orthogonal surfaces. All the cells which could contain ideal gas particles separated by a distance equal or smaller than  $r_c$  from at least one particle of the GEM-4 system are filled with ideal gas particles. Figure 5.1 shows a snapshot of the two-dimensional adaptation of this scheme. The GEM-4 particles are shown in red and the ideal gas particles in grey. Thin grey lines delimit the cells.

The pressure and the temperature of an atmosphere of ideal gas particles surrounding a nanocrystal are determined by the velocity distribution and the density distribution of the ideal particles. The main idea of this method for applying hydrostatic pressure is that pressure and temperature in the ideal gas, and consequently in the nanocrystal, are controlled by producing the right flow of particles into the volume and removing all particles that leave the volume occupied by the ideal gas. Grünwald *et al.* performed a detailed derivation of the statistics of an ideal gas in reference [89] which determines that the number of ideal gas particles  $N_b$  that flow through a rectangular surface of area  $A$  (in 3D) or a line of length  $A$  (in 2D) in a time interval  $\Delta t_{\text{flow}}$  is given again by a Poisson distribution  $P_{\text{Poisson}}(N_b)$ :

$$P_{\text{Poisson}}(N_b) = \frac{\langle N_b \rangle^{N_b}}{N_b!} e^{-\langle N_b \rangle}, \quad (5.5)$$

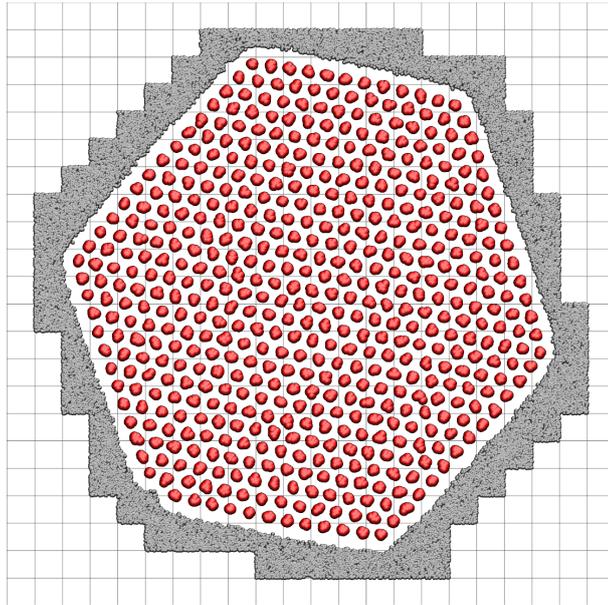


Figure 5.1.: The minimal volume of cells around a GEM-4 crystal (red); the ideal gas particles (grey) are injected or removed at the outer surface of the ideal gas. Taken from [91].

where the mean is given by:

$$\langle N_b \rangle = \Delta t_{\text{flow}} A P \left( \frac{1}{2\pi m_b k_B T} \right)^{1/2} \quad (5.6)$$

where  $m_b$  is the mass of the ideal gas particles. The components of the velocities of the entering particles follow different distributions in the parallel and the perpendicular directions to the surface (in 3D) or line (in 2D)  $A$  under consideration. The parallel components of the velocity follow a Maxwell-Boltzmann distribution given by:

$$P_{\text{M-B}}(v_{\parallel}) = \sqrt{\left[ \frac{m}{2\pi k_B T} \right]^d} \exp \left( -\frac{m v_{\parallel}^2}{2k_B T} \right), \quad (5.7)$$

where  $d$  is the number of dimensions in which the simulation is being carried out. The probability distribution of the perpendicular components  $p(v_{\perp})$  is given by:

$$P(v_{\perp}) = \frac{m_b}{k_B T} v_{\perp} \exp \left( -\frac{m_b v_{\perp}^2}{2k_B T} \right), \quad (5.8)$$

since only those particles outside the minimal volume with a sufficiently high velocity will reach the surface.

The equations of motion of the GEM-4 and ideal gas particles are integrated in a standard MD simulation, using the velocity-Verlet integration scheme with a time-increment of  $\delta t =$

0.002 (see subsection 3.1.1). As the ideal gas bath acts as a thermostat for the GEM-4 system, no additional thermostat is used in the simulation. The standard MD algorithm is adapted to include the peculiarities of the ideal gas bath.

The main steps of the simulation algorithm are the following:

1. propagate the positions of the GEM-4 particles for one time step (first step of the velocity-Verlet algorithm).
2. Check if the minimal volume has changed due to the propagation of positions of the particles in the crystal. Remove cells and gas particles which are no longer needed.
3. Propagate the positions of the gas particles for one time step (first step of the velocity-Verlet algorithm). Remove all gas particles that have left the minimal volume.
4. Place  $N_{\text{in}}$  particles on the surface of the minimal volume with positions drawn from a uniform distribution.  $N_{\text{in}}$  is a number drawn from a Poisson distribution  $P_{\text{Poisson}}(N_{\text{in}})$  (equation (5.5)) with an average given by equation (5.6). Assign to these particles the appropriate parallel and perpendicular velocities with respect to the entering surface. We introduce in equations (5.6) and (5.8)  $T_t$  and  $P_t$ , the target pressure and temperature that we impose to the GEM-4 system.
5. If necessary add new cells and fill them with  $N_{\text{app}}$  gas particles.  $N_{\text{app}}$  is a number drawn from a Poisson distribution with an average given by  $\langle N_{\text{app}} \rangle = \frac{P}{k_B T} v_{\text{cell}}$ , where  $v_{\text{cell}}$  is the volume of the cells, namely  $r_c^d$ . The positions of the particles are uniformly distributed over the cell volume and their velocity distribution is Maxwell-Boltzmann (equation (5.7)).
6. Compute forces and propagate all velocities of the GEM-4 and gas particles (second step of the velocity-Verlet algorithm).

Configurations are saved every 1000 time steps. At this point, relevant quantities of the system as the temperature  $T_m$  in the GEM-4 system, the pressure  $P_m$  or the density  $\rho$  are measured. We define the measured temperature in the GEM-4 system  $T_m$  *via* the kinetic energy of the GEM-4 particles:

$$T_m k_B = \frac{2}{d} m \sum_{i=1}^N \frac{\mathbf{v}_i^2}{2} \quad (5.9)$$

where  $N$  is the number of GEM-4 particles and  $m$  is their mass. The measured pressure  $P_m$  is defined *via* the virial:

$$P_m = \rho_t k_B T_t + \frac{1}{3V} \left\langle \sum_{i=1}^{(N+N_b)-1} \sum_{j=i+1}^{(N+N_b)} \mathbf{f}(\mathbf{r}_{ij}) \mathbf{r}_{ij} \right\rangle, \quad (5.10)$$

where  $\mathbf{f}(\mathbf{r}_{ij})$  is the force that particle  $j$ , separated by  $\mathbf{r}_{ij}$  from particle  $i$ , exerts on the latter

and  $\rho_t$  is the total density defined as:

$$\rho_t = \frac{N + N_b}{n_{\text{cell}}v_c}, \quad (5.11)$$

being  $n_{\text{cell}}$  the number of active cells (see figure 5.1). The density of the system of GEM-4 particles  $\rho$  is defined as:

$$\rho = \frac{N}{n_{\text{cell}}v_c - \frac{N_b k_B T_b}{P_m}}, \quad (5.12)$$

where  $T_b$  is the temperature measured from the kinetic energy of the ideal gas bath particles. We define the quantities  $\delta T$  and  $\delta P$  as:

$$\delta T = \frac{|T_m - T_t|}{T_t}, \quad (5.13)$$

$$\delta P = \frac{|P_m - P_t|}{P_t}, \quad (5.14)$$

which measure the deviations of the temperature and pressure measured in the GEM-4 particles ( $T_m$  and  $P_m$ ) are from the target temperature and pressure ( $T_t$  and  $P_t$ ).

### 5.2.1. Compression/expansion or heating/cooling protocol

By appropriately varying  $T_t$  or  $P_t$  in step number 4 of the simulation algorithm we can heat, cool, compress or expand the GEM-4 system. These procedures are characterised by three quantities:  $\Delta P$  or  $\Delta T$ ,  $t_{\text{wait}}$  and the tolerance range: after a pressure or temperature change ( $\Delta P$  or  $\Delta T$ ) has been applied, the GEM-4 system is perturbed and we have to let the system relax to *equilibrium*. We consider that the system is back in *equilibrium* when the quantities  $\delta T$  and  $\delta P$  are within a defined tolerance range. In our simulations we defined our tolerance range to be 0.02 (2%), *i.e.* the system is in equilibrium if  $\delta T \leq 0.02$  and  $\delta P \leq 0.02$ . Once the system is in *equilibrium* and before we apply the next change  $\Delta P$  or  $\Delta T$ , we wait during a time  $t_{\text{wait}}$  during which we measure relevant quantities of the system. In the results presented in this thesis we use  $t_{\text{wait}} = 500$ . We define the time between two pressure or temperature steps are applied as  $t_P$  or  $t_T$ , where the sub-index indicates that the target pressure or temperature is kept constant during this time.  $t_P$  or  $t_T$  are the sum of the time needed to recover *equilibrium* and  $t_{\text{wait}}$ . For a given  $\Delta P$  or  $\Delta T$  the value of  $t_P$  or  $t_T$  varies along the simulation. Additional and independent runs under the same conditions provide evidence that the values of  $t_P$  or  $t_T$  can differ for a specific pressure or temperature value; nevertheless, these variations in the equilibration protocol do not affect the relevant data about the structure of the GEM-4 system that we have extracted from these runs and that we present in the following sections. This is demonstrated for the density  $\rho$  and the pressure  $P_m$  in figure 5.2 where we compare  $P_m(t)$  and  $\rho(P_m)$  for two independent runs: although the

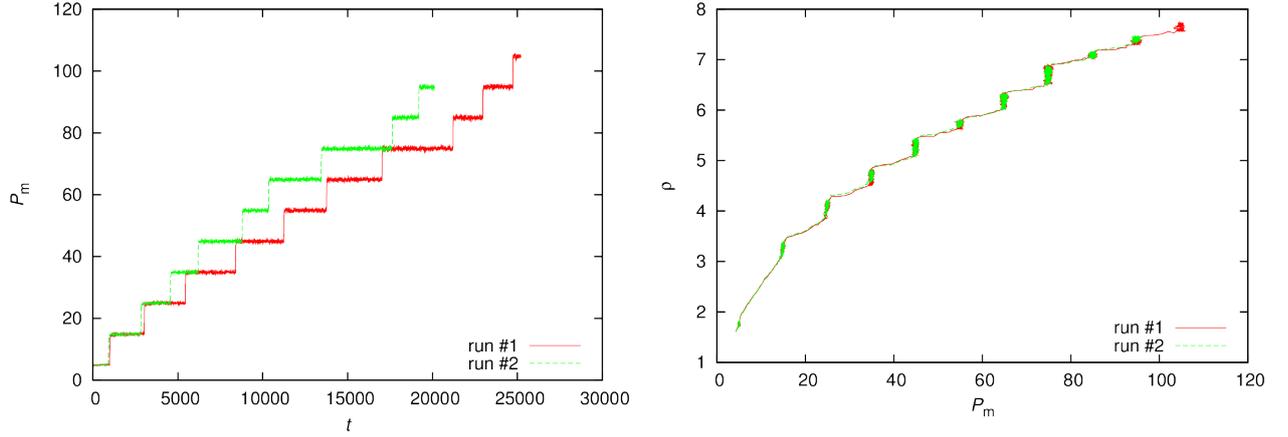


Figure 5.2.: The measured pressure  $P_m$  as a function of time (left) and the density as a function of  $P_m$  (right) for two independent runs at  $T_t = 0.4$  and  $\Delta P = 10$ .

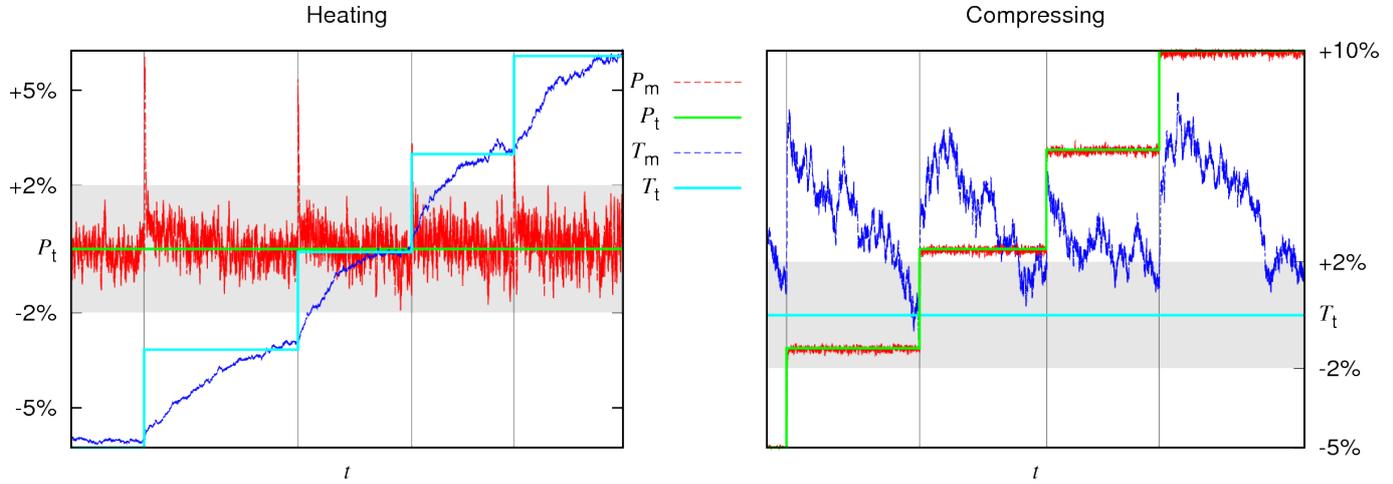


Figure 5.3.: Typical behaviour of the measured pressure  $P_m$  (red) and temperature  $T_m$  (blue) during a heating (left) and a compressing (right) process. The target temperature  $T_t$  and pressure  $P_t$  are marked in cyan and green, respectively.

values of  $t_P$  are very different, indicated by the difference in the length of the plateaus in the left panel of figure 5.2, data for the density as a function of the pressure show a perfect agreement between the two runs.

In figure 5.3 we show the typical behaviour of the GEM-4 system during a heating (left panel) and a compression run (right panel). We can observe that after every increment in temperature by  $\Delta T$  (left panel) the measured pressure increases and respectively in the right panel, after every increment in pressure by  $\Delta P$  the measured temperature also increases. There is however a fundamental difference between the behaviour of  $P_m$  and  $T_m$ :  $P_m$  instantly adapts to  $P_t$  both during a compression (when a  $\Delta P$  is applied) and in a heating (after the initial perturbation induced by  $\Delta T$ ). In contrast,  $T_m$  approaches  $T_t$  in an much slower way both during a heating (once  $\Delta T$  has been applied) and during a compression (after the initial perturbation induced by  $\Delta P$ ).

## 5.3. Results

In the following we present the results obtained in a simulation of a two dimensional GEM-4 system surrounded by an ideal gas. Reduced units, as specified in appendix A, will be used. If not stated otherwise, the ensemble of GEM-4 particles is of the size  $N = 6144$ .

In subsection 5.3.2 we compress the GEM-4 system at constant temperature. We explore the temperatures  $T_t = 0.3, 0.4, 0.5, 0.6$  and  $0.7$ , and investigate how the GEM-4 system reacts to compression. In subsection 5.3.3 we concentrate on  $T = 0.4$  and we use the ideal gas bath to extract the equation of state  $\rho_{\text{equ}}(P_m, T_t = 0.4)$ . To this end we combine compressions at constant temperature with alternating annealing processes at constant pressure, where the temperature is raised until the system melts completely and then cooled down again to  $T_t = 0.4$ .

### 5.3.1. About the compression rate

Before we focus on the analysis of the reaction of a GEM-4 system to compression we explore the influence of  $\Delta P$  on our results. We start our discussion by analysing its effect on the measured temperature. In figure 5.4 we show how the  $T_m$  evolves as a pressure increment is applied to a GEM-4 system at  $T_t = 0.4$ ,  $P_t = 55$ . We compare the effect of four different pressure increments:  $\Delta P = 1, 2, 10$  and  $25$ . The curves are shown until a waiting time  $t_{\text{wait}} = 500$  has been completed. The inset shows the average value of  $t_P$ ,  $\langle t_P \rangle$ . We see how with increasing  $\Delta P$  the subsequent jump in  $T_t$  increases. Consequently,  $t_P$  increases.

In figure 5.5 we elaborate more on  $\langle t_P \rangle$  by plotting the values of  $\langle t_P \rangle$  as a function of temperature for different  $\Delta P$  values. We see that although  $\langle t_P \rangle$  increases with  $\Delta P$ , the proportionality constant is less than one. Therefore,  $\Delta P$  must be adjusted to the purpose of the simulation: if we are interested in measuring equilibrium quantities, as for example the equation of state, we should rather take a small  $\Delta P$  value so that  $T_m$  deviates as little as possible from the  $T_t$ ; if we are interested in reaching high pressure values we should use a big  $\Delta P$  value in order to reduce the simulation runtime; once the desired pressure value has been attained, we wait for equilibrium to be recovered.

As for the influence of  $T_t$  on  $\langle t_P \rangle$ , for a fixed  $\Delta P$  value,  $\langle t_P \rangle$  decreases with increasing  $T_t$ . We attribute this behaviour to the role of thermal fluctuations: with increasing  $T_t$ , they become more pronounced and enable rearrangements of the particles in the system, which are necessary to achieve the equilibrium configuration.

Naturally, the question arises whether the structural response of the GEM-4 system to a compression is affected by the choice  $\Delta P$  *i.e.* whether the configuration of the GEM-4 particles at a certain  $P_t$  value is affected by the chosen  $\Delta P, t_{\text{wait}}$  and tolerance parameter to achieve this pressure.

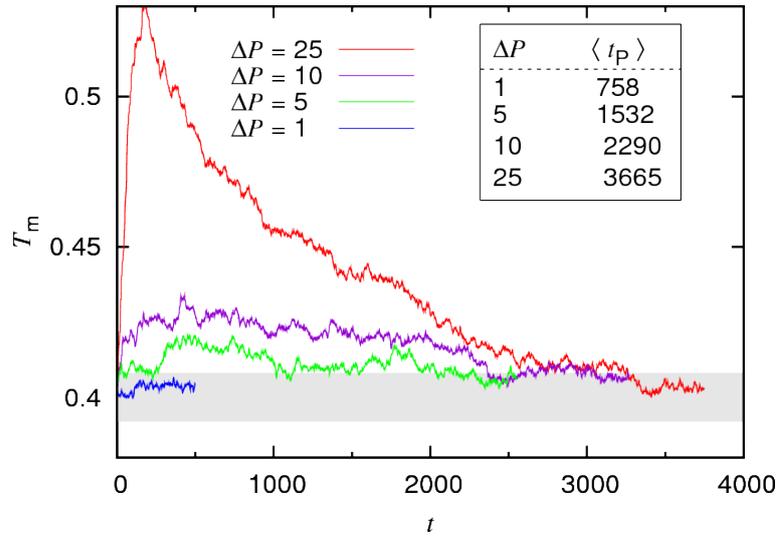


Figure 5.4.: Evolution of the  $T_t$  in a GEM-4 at  $T_t = 0.4$ ,  $P_t = 55$  after a pressure increment of  $\Delta P = 1, 5, 10$  and  $25$  is applied (colours as labeled). The waiting is  $t_{\text{wait}} = 500$ . The curves are shown during the respective time  $t_P$ . The inset shows the value of  $\langle t_P \rangle$  for the different values of  $\Delta P$ .

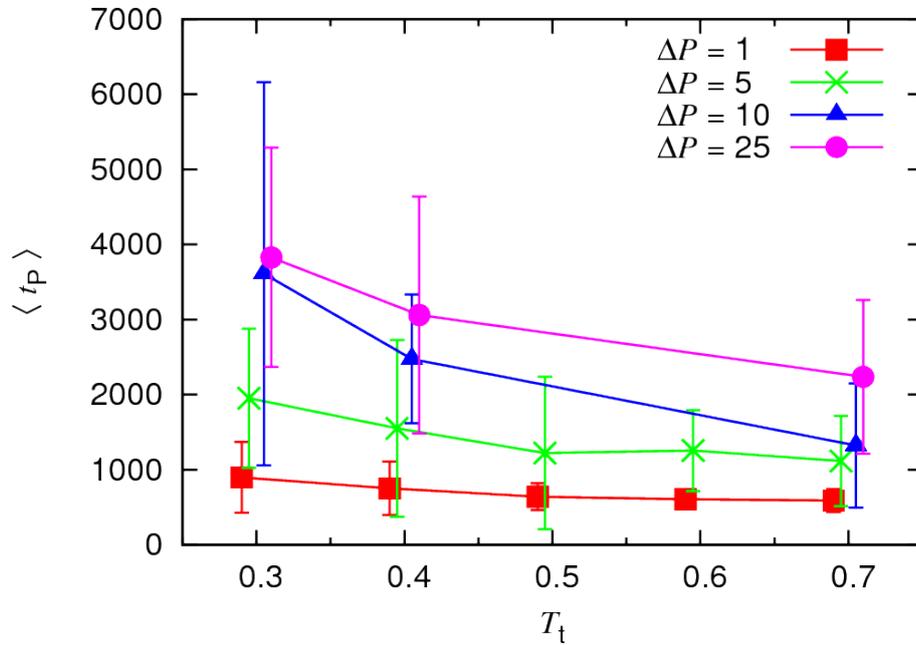


Figure 5.5.:  $\langle t_P \rangle$  for different compression simulations with different  $T_t$  and  $\Delta P$  values (as labeled).

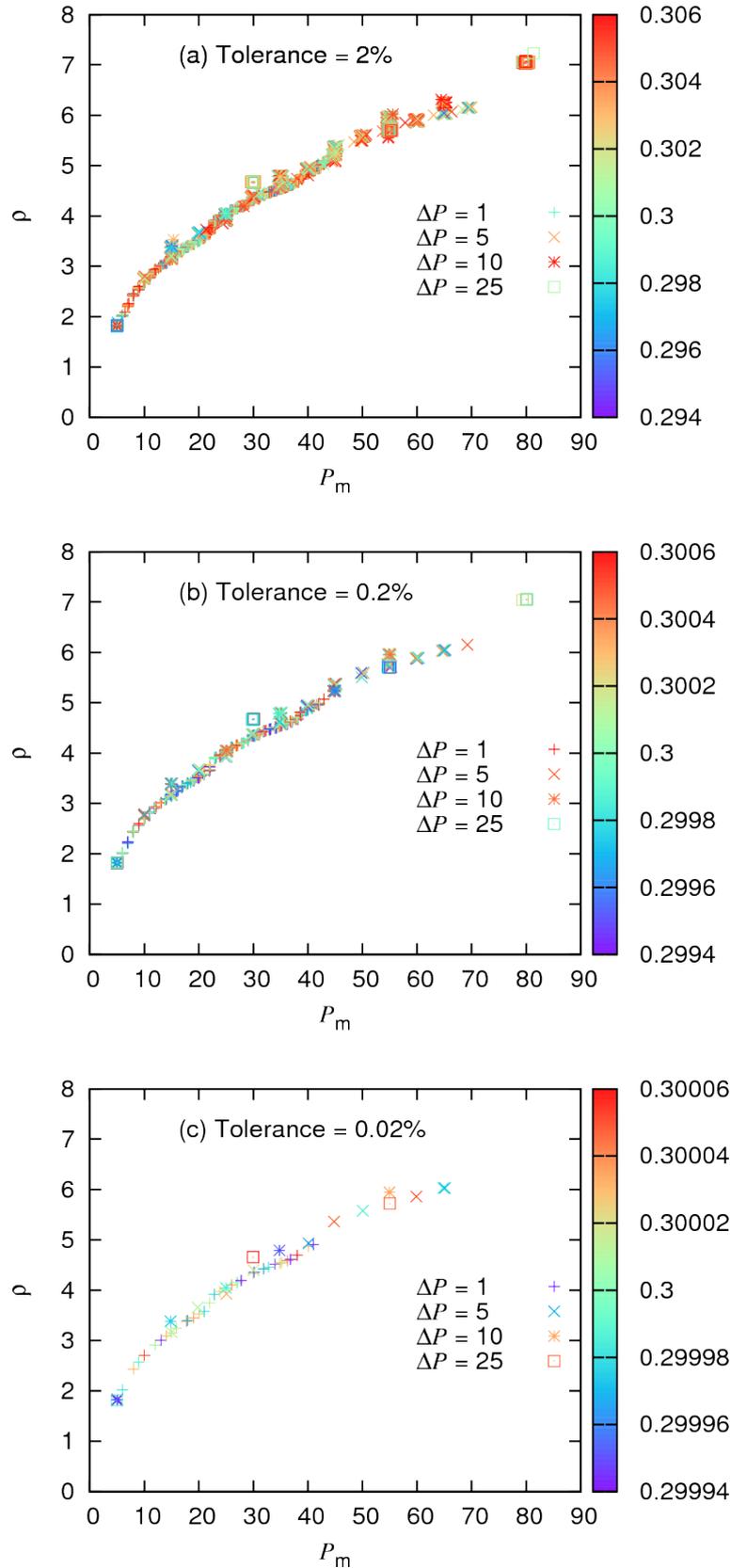


Figure 5.6.: Measurements of the density  $\rho$  as a function of the pressure  $P_m$  during a compression run in a GEM-4 system of cluster-forming particles at  $T_t = 0.3$  for  $\Delta P = 1, 5, 10$  and  $25$  as specified in the legend. The data points are coloured according to the instantaneous temperature  $T_t$  of the GEM-4 system. Panels (a), (b) and (c) show those points where  $\delta T = |T_m - T_t|/T_t$  is smaller than 0.02 (2%), 0.002 (0.2%) and 0.0002 (0.02%) respectively.

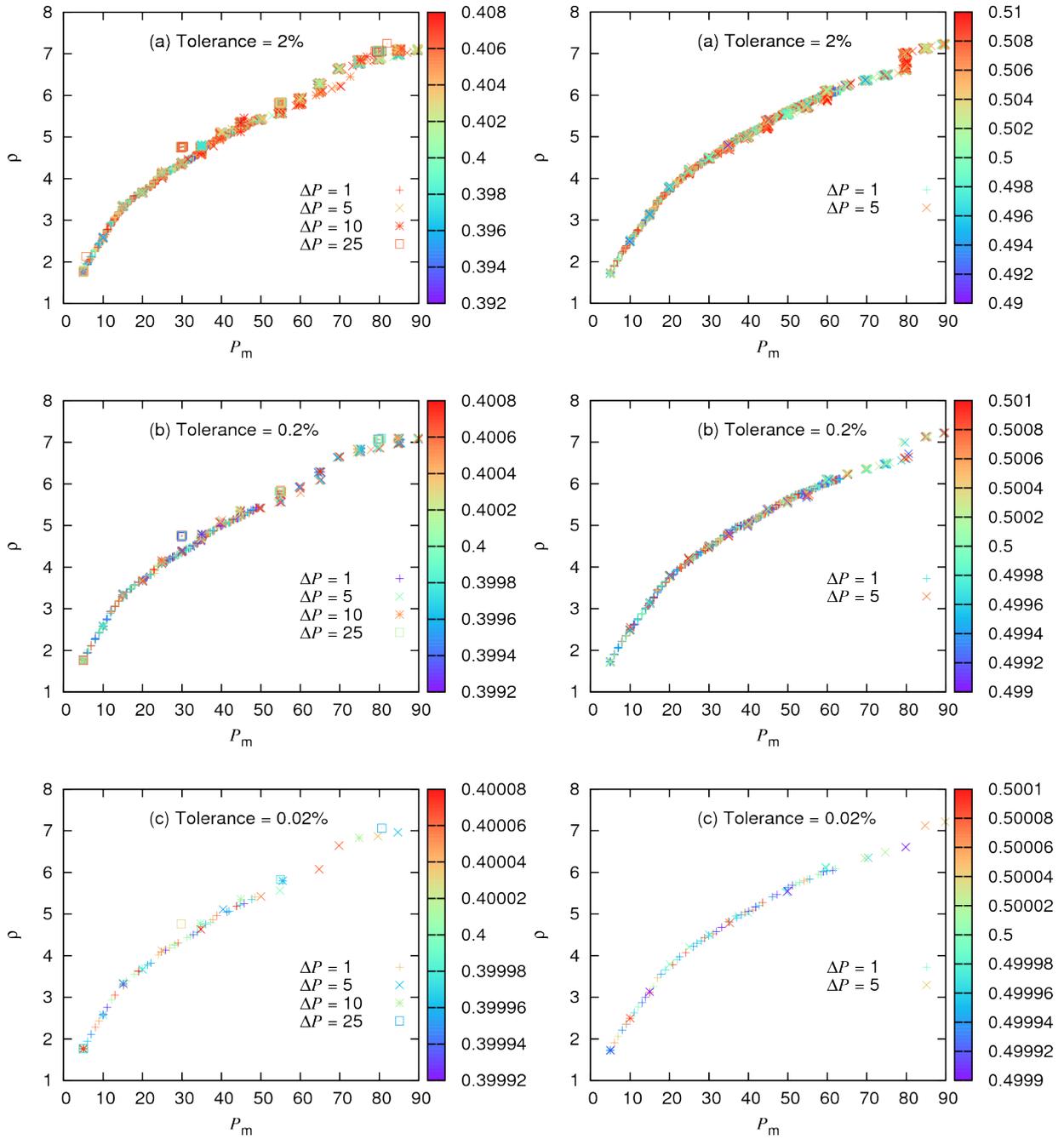


Figure 5.7.: Same as figure 5.6 but for  $T_t = 0.4$  (left) and  $T_t = 0.5$  (right).

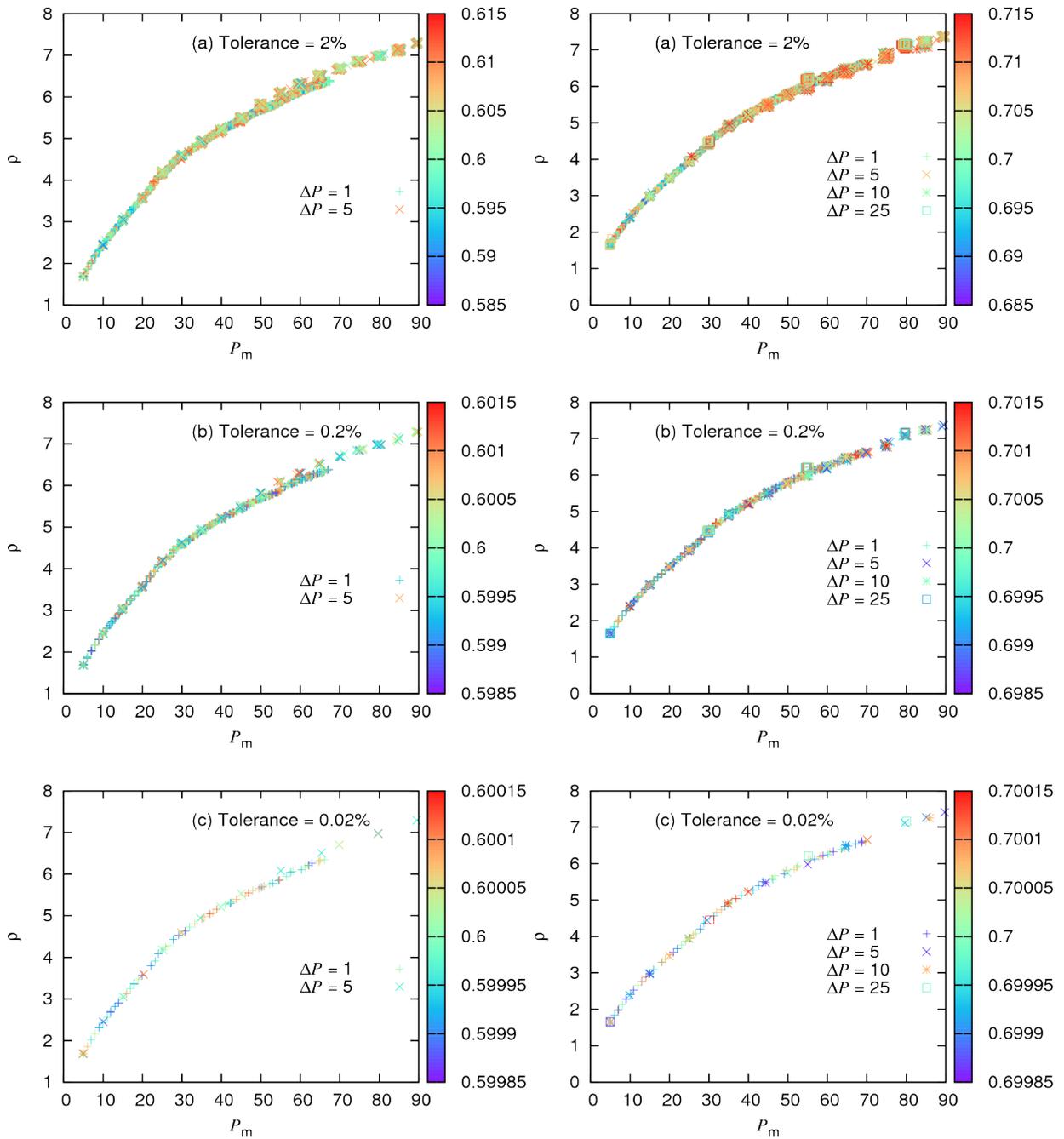


Figure 5.8.: Same as figure 5.6 but for  $T_t = 0.6$  (left) and  $T_t = 0.7$  (right).

To examine this question we present in figures 5.6 to 5.8 a superposition of the measured  $\rho(P_m)$  curves during a compression run for  $T_t = 0.3, 0.4, 0.5, 0.6$  and  $0.7$ , respectively, with different values for  $\Delta P$ , as specified in the legend.

In panels (a) of the three figures we show the data for  $\rho(P_m)$  where  $\delta T \leq 0.02$  (2%); the symbols are coded according to  $T_m$ . For each  $T_t$  we observe that although all the  $\rho(P_m)$  data essentially follow the same curve, there is a considerable scattering of the results. We observe that the scattering is in general stronger for larger  $\delta T$ -values. We test the effect of the tolerance parameter by imposing a more restrictive criterion on the data shown in the next two panels for each  $T_t$ : panels (b) show only the  $\rho(P_m)$  data where  $\delta T \leq 0.002$  (0.2%) and panels (c) present those results where  $\delta T \leq 0.0002$  (0.02%). The sets of data belong, however, to the same simulation, which was run with a tolerance parameter of 0.02 (2%) (and not to new simulations with a smaller one). In panels (b) and (c) of figures 5.6 to 5.8 we see how data that are located further away from the average curve are eliminated as the tolerance is decreased. It is specially clear in figure 5.8 where we see that for  $T_t = 0.6$  in the pressure regime  $50 \lesssim P_t \lesssim 70$  the  $\rho(P_m)$  curve for  $\Delta P = 5$  shows  $\rho$ -values clearly larger than those in the  $\Delta P = 1$  curve. As the tolerance is reduced, data belonging to the ( $\Delta P = 5$ )-curve are eliminated, meaning that the states in the ( $\Delta P = 1$ )-curve are closer to equilibrium results. We notice that with increasing temperature there is less scattering of the data. Again, as  $T_t$  increases, it is easier for the system to stay closer to equilibrium thanks to enhanced thermal fluctuations.

$T_t$	$\Delta P$	$\delta T < 0.002$	$\delta T < 0.0002$
0.3	1	19.4%	1.96%
	5	6.1%	0.61%
	10	2.5%	0.25%
	25	1.7%	0.16%
0.4	1	9.9%	0.98%
	5	7.6%	0.71%
	10	6.7%	0.67%
	25	9.1%	0.91%
0.5	1	32.5%	3.22%
	5	11.4%	1.18%
0.6	1	35.0%	3.48%
	5	10.6%	1.06%
0.7	1	37.3%	3.67%
	5	12.3%	1.18%
	10	7.8%	0.81%
	25	3.3%	0.29%

Table 5.1.: Percentage of  $\rho(P_m)$  measurements at different  $T_t$  values with  $\delta T \leq 0.002$  and  $\delta T \leq 0.0002$  with respect to the number of measurements where  $\delta T \leq 0.02$  for different  $\Delta P$  values.

We quantify our findings by summarizing in table 5.1 the percentages of data which “survive”

the more restrictive criteria in panels (b) and (c) with respect to the number of data shown in panel (a). We observe that indeed at higher  $T_t$  values these percentages are bigger. We also observe that at a certain temperature  $T_t$ , the system stays closer to equilibrium for smaller  $\Delta P$  values. The percentage of “surviving points” for  $\Delta P = 25$  is in all of the cases very low, meaning that this compression rate is too fast for our waiting time ( $t_{\text{wait}}$ ).

### 5.3.2. Compression experiments

We start our compression experiments for the GEM-4 system with a random configuration of the particles in a square box with periodic boundary conditions. After an equilibration phase we remove the periodic boundary conditions and surround the system with the ideal gas particles at an initial pressure of  $P_t = 5$  and the corresponding desired target temperature  $T_t$  (0.3 to 0.7). The system is at this point in a liquid state. Then we start the compression protocol described in subsection 5.2.1. Along the simulation we save the positions and velocities of the GEM-4 particles for the subsequent cluster analysis (described in appendix B). By visual inspection we select a moment in the simulation where we consider the system to be in the ordered phase; then we can start to identify *via* the cluster analysis hopping and merging events. We keep track of the number of hopping particles ( $N_h$ ) and the number of particles involved in a merging event ( $N_m$ ). We perform the simulation during three days on a supercomputer using 12 parallel cores. Only in the case of  $\Delta P = 1$  we extend the simulation over six days so that we are able to achieve at the end of the experiments pressure values that are comparable to those obtained for other  $\Delta P$  values.

#### Merging and hopping activity along a compression run

In figure 5.9 we present in a multi-panel figure the measured temperature,  $T_m$ , pressure,  $P_m$ , and the number of particles involved in a hopping event,  $N_h$ , and in a merging event,  $N_m$ , along the entire simulation for different  $T_t$  and  $\Delta P$  values: rows are ordered by  $T_t$  and columns by  $\Delta P$ , as labeled. The grey vertical lines delimit the time intervals during which  $P_t$  is kept constant. At the beginning the system is in the fluid phase and no information about  $N_h$  and  $N_m$  is available.

We start by analysing the behaviour of  $T_m$  during the corresponding runs and its consequences on the progress of the compression. For all temperatures and all  $\Delta P$ -values we observe that: (i) variations in the measured temperature  $T_m$  correlate with variations in the hopping and merging activities, (ii) after every increase in pressure by  $\Delta P$  a sudden jump in  $T_m$  is observed and (iii) as anticipated in subsection 5.2.1 the time-intervals during which the target pressure is kept constant differ in their length.

We next turn our attention to the merging and hopping activity. Once the system is in the ordered phase, as we start our analysis: we observe an initial phase where hopping processes dominate over merging events ( $N_h > N_m$ ); we anticipate that this intermediate time-regime

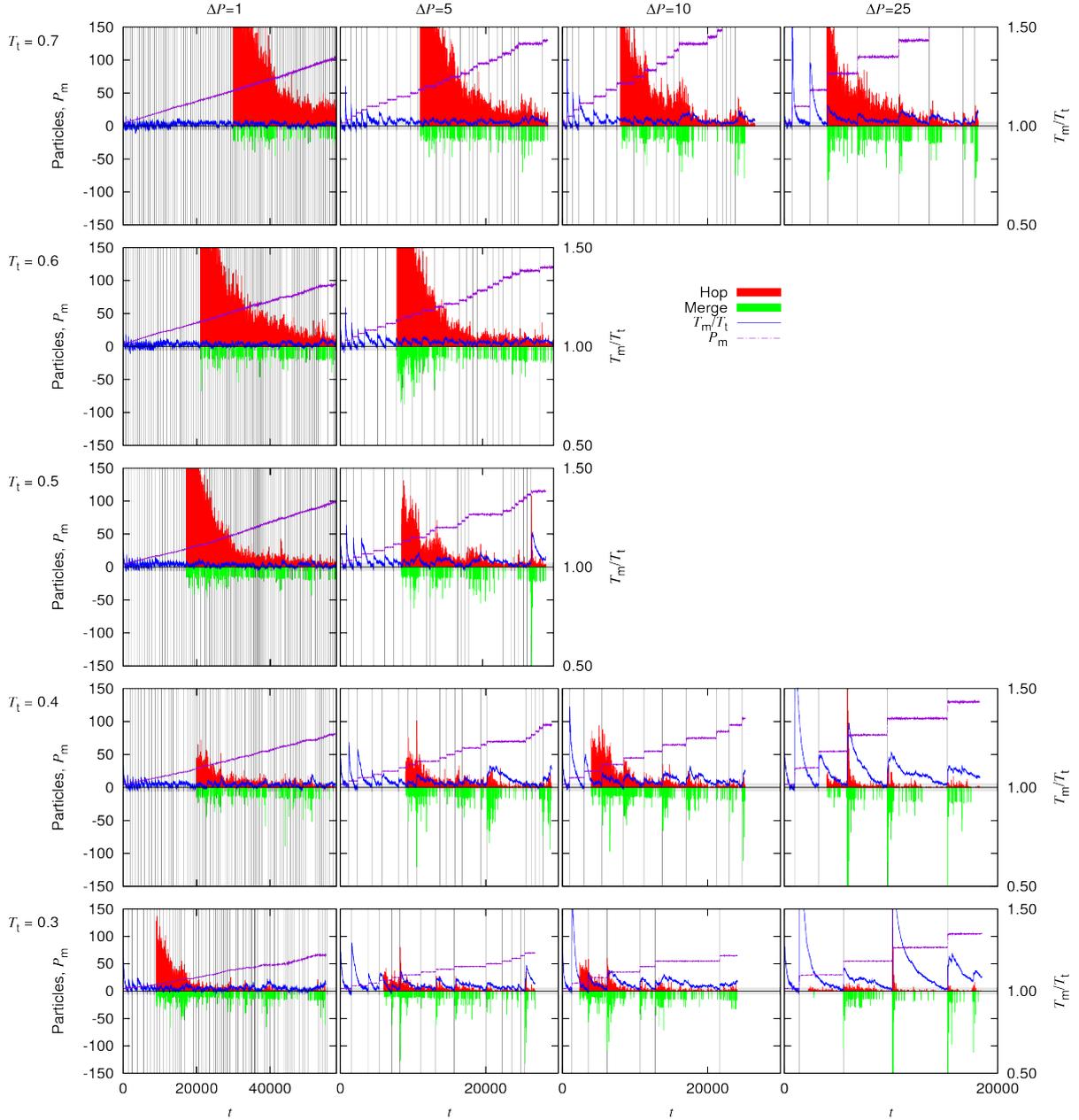


Figure 5.9.:  $T_m/T_t$ ,  $P_m$ ,  $N_h$  and  $N_m$  (as labeled) as functions of (simulation) time during a compression a run at  $T_t = 0.7$  (first row), 0.6 (second row), 0.5 (third row), 0.4 (fourth row) and 0.3 (last row) using  $\Delta P = 1$  (first column), 5 (second column), 10 (third column) and 25 (fourth column).

corresponds to a low order phase measured by the order parameter  $\Psi_6$  (equation (5.17)) which will be discussed later on. Then, as the system attains a higher order configuration,  $N_h$  stabilises around a constant value. On the contrary, for  $N_m$  we do not distinguish a different behaviour along the compression experiment, with the exception of the data for  $T_t = 0.6$  with  $\Delta P = 5$  where we do see an initially higher value of  $N_m$ .

We observe that at low target temperatures the pressure increment  $\Delta P$  has a strong influence on when the hopping and merging activities occur. For small  $\Delta P$  values these events are observed over the entire run, while for large  $\Delta P$  values hopping and merging occur immediately after the pressure step has been increased. As the target temperature increases, this distinction in the behaviour at low and high  $\Delta P$  values disappears and the hopping and merging activity extends all over the run.

As for the impact of the target temperature on the merging and hopping activities, we observe that while this quantity has little impact on  $N_m$ , the target temperature  $T_t$  has a very strong influence on  $N_h$ : at low  $T_t$ , the value of  $N_m$  is considerably larger than  $N_h$  while at high  $T_t$  both quantities are of comparable size. As we measure the cluster merging activity we are purely recording the reaction of the GEM-4 system to a compression stem (is driven out of equilibrium): merging of clusters essentially never occurs under equilibrium conditions due to the high energy barriers that separate adjacent clusters. On the other hand, particle hopping is a characteristic transport mechanism in cluster-forming systems, which also takes place in equilibrium. In chapter 4 that the hopping activity increases with temperature. Since we do not have data to decouple the amount of hopping activity which would be present if the system was in equilibrium from the hopping activity which is a response to the compression, we cannot draw a conclusion whether this increase in  $N_h$  with  $T_t$  means that the reaction of a cluster-forming system to a compression (in terms of the ratio between  $N_h$  and  $N_m$ ) is different at high  $T_t$  values.

### Analysis of the clusters involved in a merging event

To better elucidate the nature of merging events we try to identify which clusters are prone to merge by comparing the probability distributions of the occupation number, kinetic energy, spatial distribution and number of nearest neighbours of the clusters involved in a merging event compared to the rest of the clusters in the system:

- Cluster occupation:

In figure 5.10 we provide a more detailed study of the complex interplay between hopping and merging processes of a cluster-forming system. In the left panel of figure 5.10 we plot the probability distribution of the size of the clusters involved in a merging event normalised by the average cluster size measured in the system at the moment where the merging event occurs. For each  $T_t$  we merge the data obtained from different  $\Delta P$  values. In the right panel of figure 5.10 we show the same data but now we separate it by  $\Delta P$  values, and for each  $\Delta P$  value we merge the data obtained from

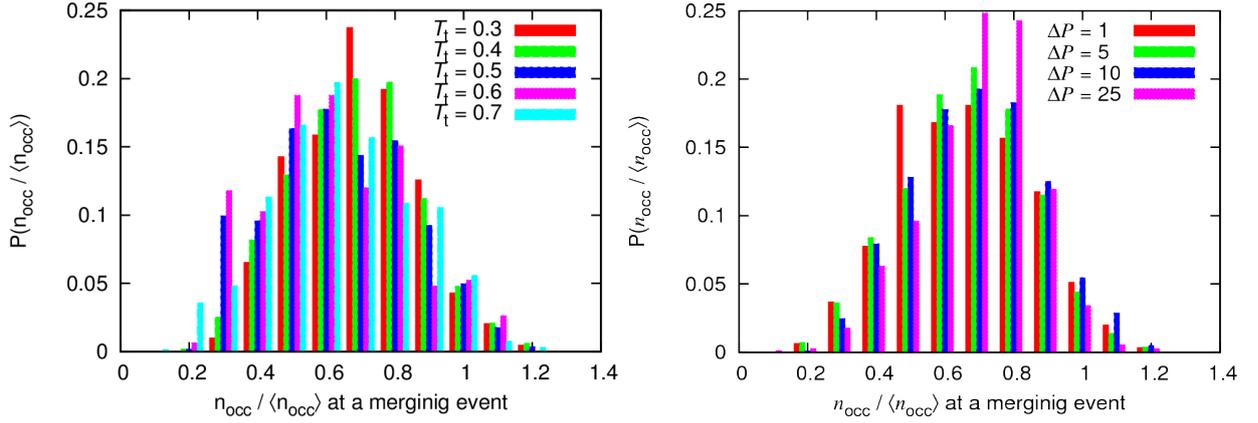


Figure 5.10.: Probability distribution of the size of the clusters involved in a merging event normalised by the instantaneous global average cluster occupation. (Left) Results are shown for simulations at different  $T_t$  (Data for different  $\Delta P$  values are merged in the same plot). Taken from [91]. (Right) Results are shown for simulations at different  $\Delta P$  (Data for different  $T_t$  values are merged in the same plot).

different  $T_t$ . We find that, essentially irrespective of  $T_t$  and  $\Delta P$ , the distributions have their maxima at  $n_{\text{occ}}/\langle n_{\text{occ}} \rangle \sim 0.7$ , *i.e.* clusters that are involved in merging processes are smaller in size as compared to other clusters in the system.

We can thus make the following conclusion: ultrasoft, cluster-forming systems respond to a compression by:

- 1) activated hopping processes which lead to
- 2) a spatially heterogeneous cluster size distribution;
- 3) under-occupied clusters feel a strong repulsion from larger neighbouring clusters;
- 4) and they are forced to approach each other and eventually merge.

Figure 5.10 demonstrates the important role of hopping mechanisms in the response of a cluster crystal to compression: merging events mainly occur when small clusters are pushed together by their larger neighbours. This can only occur if there is an heterogeneous cluster size distribution in the system, which is achieved through the hopping activity.

- Kinetic energy:

We also address the question whether the distribution of the kinetic energy among the particles in the system has an effect on which clusters are merging. In figure 5.11 we plot for every  $T_t$ - and  $\Delta P$ -value combination, the probability distribution of the normalised temperature of the clusters,  $T_{\text{cluster}}/T_m$ , of all the clusters in the system ( $P_{\text{system}}(T_{\text{cluster}}/T_m)$ ) and compare it to the distribution of the normalised temperature of the clusters involved in merging events ( $P_{\text{merging}}(T_{\text{cluster}}/T_m)$ ). The temperature of a

cluster is defined as:

$$T_{\text{cluster}} = \frac{1}{2}mv_{\text{COM}}^2. \quad (5.15)$$

For all  $T_t$ - and  $\Delta P$ -values we observe that the maxima of both distributions (whole system and only merging clusters) roughly coincide. For small  $T_t$ - and/or  $\Delta P$ -values both distributions coincide over the whole range of  $T_{\text{cluster}}/T_m$ . As  $T_t$  or  $\Delta P$  increase,  $P_{\text{system}}(T_{\text{cluster}}/T_m)$  attains for small ( $T_{\text{cluster}}/T_m < 0.5$ ) and large ( $T_{\text{cluster}}/T_m > 1.5$ ) values of the normalised temperature larger values than  $P_{\text{merging}}(T_{\text{cluster}}/T_m)$ : the normalised temperatures of the clusters involved in a merging event concentrate around  $T_{\text{cluster}}/T_m = 1$ .

- Spatial distribution of the merging events:

We next analyse the spatial distribution of the merging clusters in the system by calculating the probability distribution of their relative radial coordinate  $d_{\text{rel}}$ . The relative radial coordinate measures the distance between the COM of a cluster and the COM of the whole system. It can take values between 0 and 1, so that  $d_{\text{rel}} = 0$  means that the cluster is positioned in the centre of mass of the whole GEM-4 system and  $d_{\text{rel}} = 1$  means that the cluster is positioned on the surface in contact with the pressure bath.

In figure 5.12 we plot the normalised probability distribution of  $d_{\text{rel}}$ , *i.e.*  $P(d_{\text{rel}})/d_{\text{rel}}$ , of the clusters involved in a merging event for different  $T_t$  values. Again, for each  $T_t$  we merge the data from different  $\Delta P$  values. We do not observe any trend in the data, meaning that the merging events are homogeneously distributed in space. Therefore, our original hypothesis that the frequency of merging events is higher close to the surface cannot be confirmed.

- Number of nearest neighbours:

Finally, in figure 5.13 we plot the number of neighbouring clusters within a radius of  $r_c = 1.5$ , of the clusters involved in a merging event before (left panel) and after (right panel) the event. The choice of  $r_c$  is justified by the data presented in figure 5.16 (see below) in view of the fact that we find that the average distance to the nearest neighbour cluster lies in the range [1.25,1.45]. We see from figure 5.13 that the clusters which are merging are in general a defect of the hexagonal lattice with 5 nearest neighbours.

In figure 5.14 we present for a time window between  $t = 4000$  and  $t = 13000$  the superposition of the  $xy$ -positions at which merging events occur during a compression run at  $T_t = 0.7$  and  $\Delta P = 10$ . The points are coloured according to the time at which the events occur. We observe cascades of consecutive merging events which occur preferably close to each other in space. The system is using merging events to delete disclinations in the crystal.

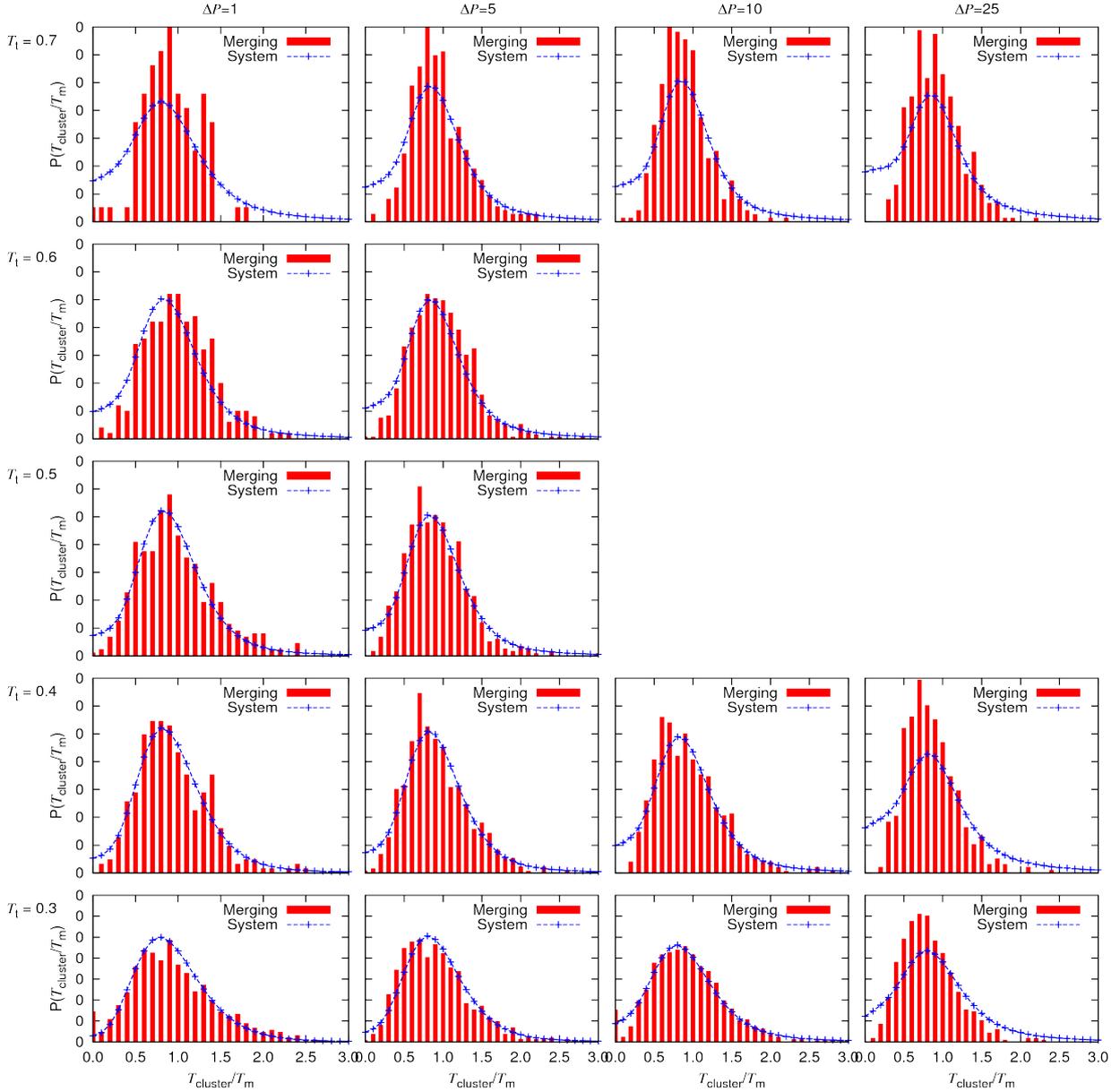


Figure 5.11.: Probability distribution of the normalised temperature of the clusters ( $T_{\text{cluster}}/T_m$ ) measured in the states where  $\delta T$  and  $\delta P < 0.002$  (*equilibrium* states) during the compression with different  $\Delta P$ - and  $T_i$ -values, as labeled and probability distribution of the normalised temperature of the clusters involved in merging events.

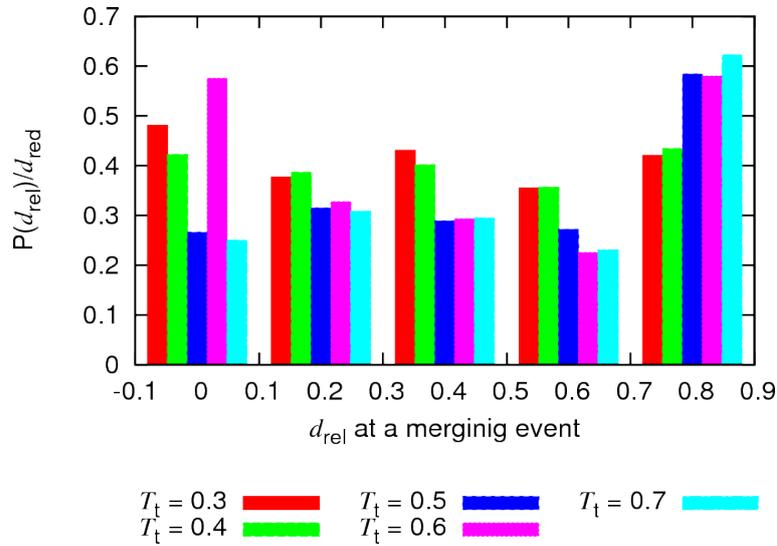


Figure 5.12.: Probability distribution of  $d_{\text{rel}}$  of the clusters involved in a merging event divided by  $d_{\text{red}}$ . Results are shown for simulations at different target temperatures  $T_t$  (Data from different  $\Delta P$  values are merged in the same plot).

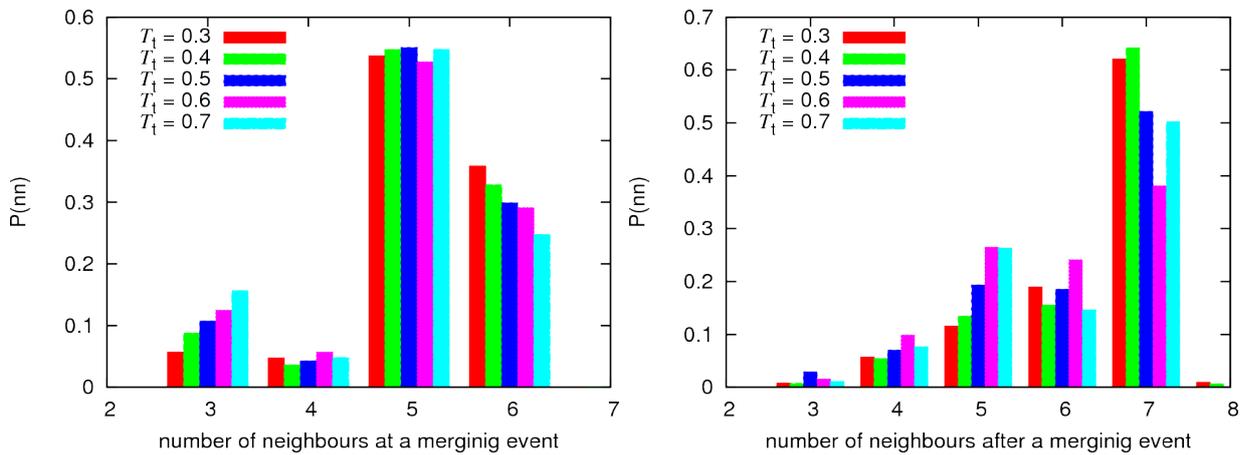


Figure 5.13.: Probability distribution of the number of nearest neighbours within a radius  $r_c = 1.5$  of the clusters before (left) and after (right) a merging event for different target temperatures as labeled. Data for different  $\Delta P$  are merged in the same plot.

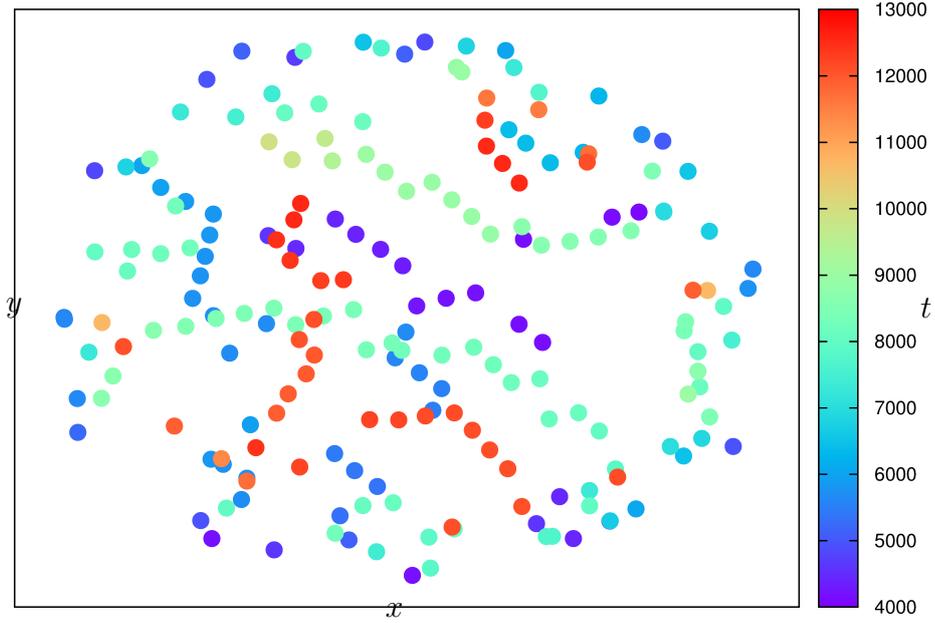


Figure 5.14.: Positions where merging events occur coloured according to the time within the simulation at which they occur. The data is taken from a clip between  $t = 4000$  and  $t = 13000$  of a compression run at  $T_t = 0.7$  with  $\Delta P = 10$ .

### Cluster occupation and distance to the nearest neighbour along a compression run

We next analyse in figure 5.16 for the different values of  $T_t$  and  $\Delta P$  how the density  $\rho$ , the distance to the nearest neighbour cluster  $d_{nn}$ , and the average cluster occupation  $\langle n_{occ} \rangle$  evolve during a compression run. The distance to the nearest neighbour cluster is defined as the position of the highest peak of the radial distribution function  $g(r)$  of the positions of the centres of mass of the clusters:

$$g(r) = \frac{1}{N_c} \sum_{i=1}^{N_c} \sum_{j=i+1}^{N_c} \delta(\mathbf{r} - \mathbf{r}_j^{cm} + \mathbf{r}_i^{cm}) \quad (5.16)$$

where  $N_c$  is the number of clusters, and  $\mathbf{r}_i^{cm}$  is the position of the centre of mass of the  $i$ -th cluster. To help the reader to visualise how  $d_{nn}$  is defined, we plot in figure 5.15 the radial distribution function for a configuration in the fluid phase, when crystallization is starting and in the crystallised state. The vertical arrows indicate for each configuration where  $d_{nn}$  is located in each case.

In figure 5.16 we observe that  $d_{nn}$  assumes in the initial fluid phase values of  $d_{nn} \simeq 0.5$  (like the blue curve in figure 5.15). As compression proceeds, an additional peak in  $g(r)$  appears at  $r \simeq 1.4$ . There is a narrow transition region observed, where the value of  $d_{nn}$  changes between these two values (0.5 and 1.4). As the density increases, the peak at  $r \simeq 1.4$  prevails (like in the green curve in figure 5.15) and finally, one indicator that the system forms an

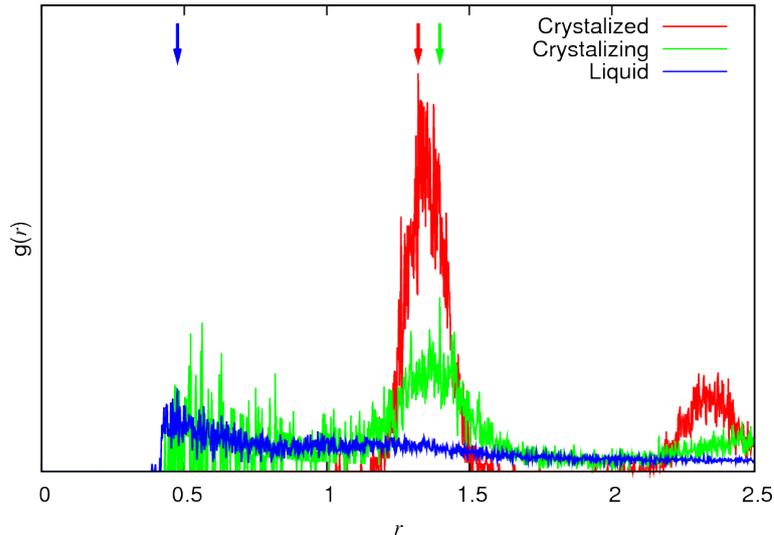


Figure 5.15.: Radial distribution function of the centres of mass of the clusters in a fluid (blue), a crystallizing (green) and a crystallised (red) configuration. The vertical arrows mark with the corresponding colour the positions of the highest peaks, i.e. the value of  $d_{\text{nn}}$  for each case.

ordered phase is the disappearance of the first peak at  $r \simeq 0.5$  (like in the red curve in figure 5.15).

As the density in a cluster crystal increases as a reaction to the increase in pressure, the system can react *via* two mechanisms:

- either a shrinkage of the lattice spacing, reflected in a decrease of  $d_{\text{nn}}$ ;
- or the deletion of some clusters, reflected in the increase of  $n_{\text{occ}}$ .

Figure 5.16 provides evidence about the rate of the two mechanisms. We observe two distinct behaviours for  $\rho$ : at low a compression rates ( $\Delta P = 1$ )  $\rho$  increases in a monotonous way, while for large compression rates ( $\Delta P = 10$  and  $25$ )  $\rho$  increases in two steps: a sudden initial increase immediately after the pressure step is applied and a second smoother one spanning over a longer time window. For  $\Delta P = 5$  the behaviour of  $\rho$  lies between these two cases: for  $T_t = 0.3$  the two steps are observed at the beginning of the compression while for the higher values of  $T_t$  the increase of  $\rho$  is rather monotonous. In the cases where  $\rho$  increases in two steps we observe that the initial decrease in volume correlates with a sudden decrease in  $d_{\text{nn}}$  accompanied by a sudden increase in  $n_{\text{occ}}$ . In the second step where  $\rho$  increases,  $d_{\text{nn}}$  slowly increases and the volume reduction is achieved by the steady increase of  $n_{\text{occ}}$ . In the case where  $\rho$  increases monotonically, it does it in the same way as in this second step.

Therefore, we can conclude that the value of  $\Delta P$  has a strong influence on how  $\rho$  increases: for small  $\Delta P$ -values there is no sudden shrinkage of the lattice constant immediately after the pressure is increased. As  $\Delta P$  increases this sudden shrinkage of  $d_{\text{nn}}$  becomes more and more pronounced.

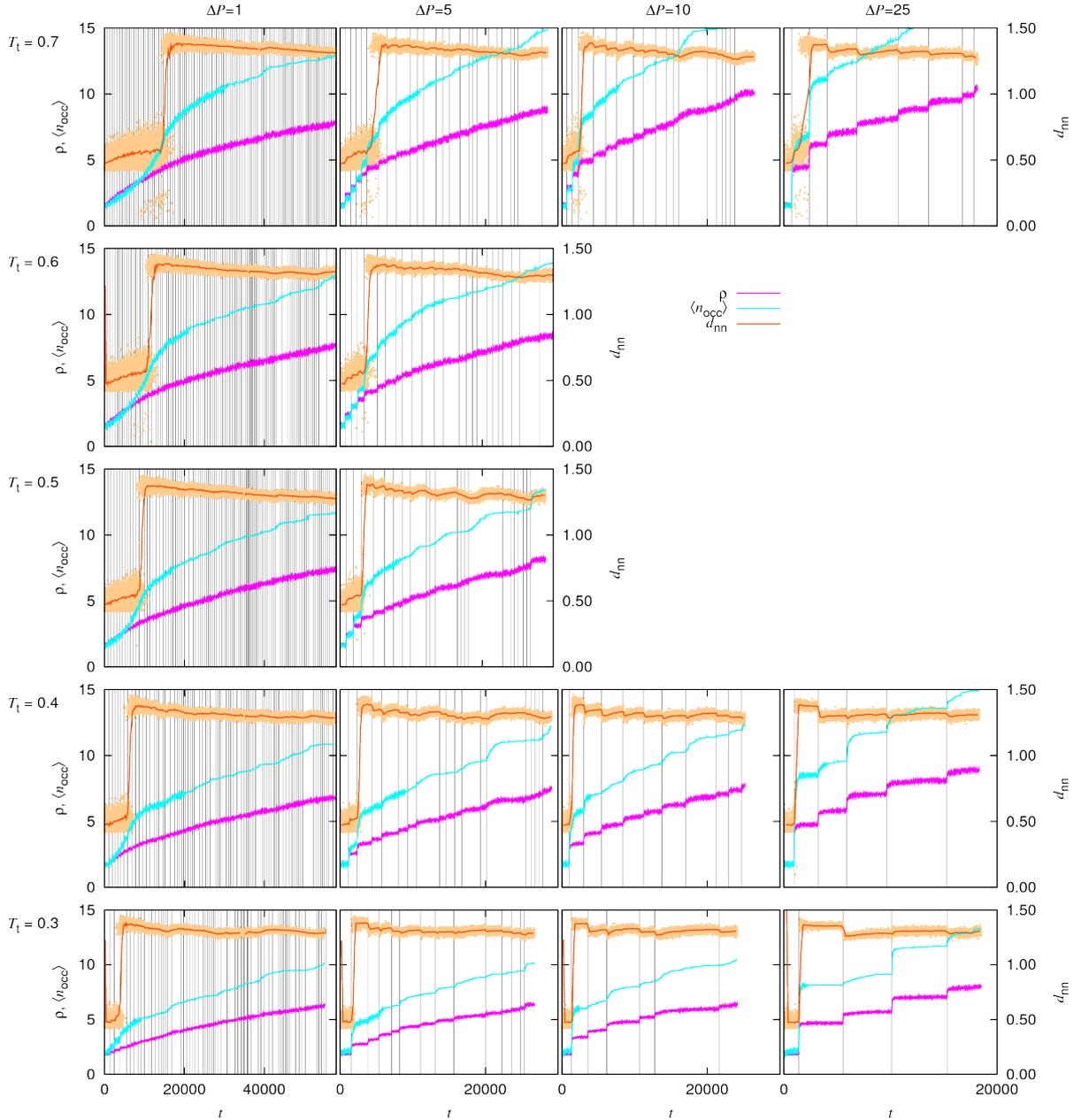


Figure 5.16.: Density  $\rho$  (pink), average cluster occupation  $\langle n_{occ} \rangle$  (cyan) and distance to the nearest neighbour (orange) as a function of time. The dark-orange curve presents a running average over 100 configurations (light-orange) corresponding to  $10^5$  time steps. The panels are organised by  $T_t$  in rows and by  $\Delta P$  in columns, as labeled.

For all the  $T_t$ - and  $\Delta P$ -values we observe, as a general trend and omitting the sudden decreases in  $d_{nn}$  sometimes present, that the spacing of the lattice shows an initial monotonous decay with time and levels off to a value of  $d_{nn} \sim 1.28$ . Density functional theory calculations based in a mean field approximation (section 2.3) predict a lattice spacing of  $a = 1.42$ , *i.e.* a value considerably larger than our result of  $d_{nn} = 1.28$ . The results obtained *via* annealing processes to be presented in subsection 5.3.3 will help to enlighten the discrepancy between the theoretically calculated value of the lattice spacing and the measured one.

### Cluster occupation as a function of density

We plot in figure 5.17 the average cluster occupation at equilibrium (*i.e.* when the values of  $\delta T$  and  $\delta P$  lie within the range of tolerance) as functions of the density  $\rho$ . We see that even though we observed a small decrease in  $d_{nn}$  with increasing pressure we can still observe the linear dependence of  $\langle n_{occ} \rangle_{equ}$  with the density, as predicted in section 2.3.

### Order parameter along a compression run

Finally, we pose the question whether the sudden increase of the density, which we observe at high  $\Delta P$ -values, will favour the formation of defects in the crystal, *i.e.* whether the number of defects in the cluster crystal compressed at a rate of  $\Delta P = 25$  is bigger as in a crystal compressed by  $\Delta P = 1$ . To answer this question, we plot in figure 5.18 the (cluster) order parameter  $\Psi_6$ , which quantifies the hexagonal order of the clusters in the system, as a function of density for the different target temperatures and different compression rates. Following reference [92], we used:

$$\Psi_6 = \frac{1}{N} \sum_{j=1}^N \frac{1}{n_j} \sum_{k=1}^{n_j} \exp(i6\Theta_{jk}), \quad (5.17)$$

where  $n_j$  denotes the number of nearest neighbours of cluster  $j$ , and  $\Theta_{jk}$  the angle between between the selected cluster and its nearest neighbours. By definition,  $\Psi_6$  attains a value of 1 for a perfectly hexagonal arrangement, while we found values  $\Psi_6$  ranging between 0.3 and 0.4 in the disordered liquid state. For compression rates where we have sudden decreases of  $d_{nn}$ , as in  $\Delta P = 25$ , we see that every time that a pressure increase is applied the order parameter decreases to later recover the original value. In the contrary, for  $\Delta P = 1$ , where  $d_{nn}$  changes rather smoothly upon pressure increase, we see that also  $\Psi_6$  has a continuous evolution. We observe that for all the compression rates, except for  $\Delta P = 25$ , the order parameter attains at equilibrium a value of  $\Psi_6 \simeq 0.95$ . With these results at hand and the results presented in subsection 5.3.1 (where we saw that for  $\Delta P = 25$  the GEM-4 system spent the great majority of the run time out-of-equilibrium) we can conclude that a compression at a rate of  $\Delta P = 25$  in combination with the chosen tolerance parameter and  $t_{wait}$  is too fast for the system at hand, since it does not have enough time after the pressure is increased to

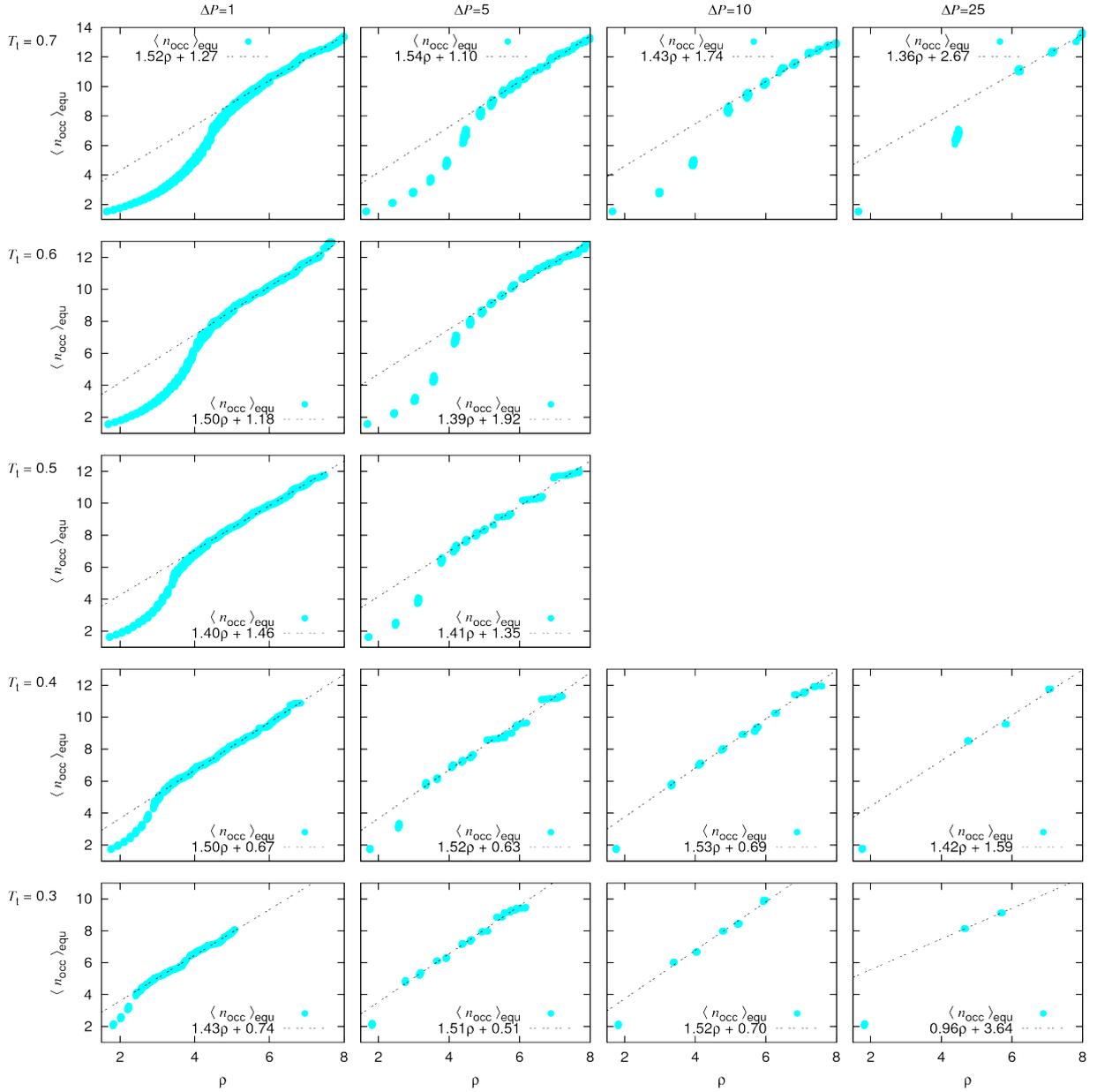


Figure 5.17.: Average cluster occupation number  $\langle n_{\text{occ}} \rangle$  in equilibrium (when  $\delta T$  and  $\delta P$  are less than the tolerance parameter) as a function of density  $\rho$ . The black dashed line is a linear fit to the data. The panels are organised by  $T_t$  and  $\Delta P$ , as labeled.

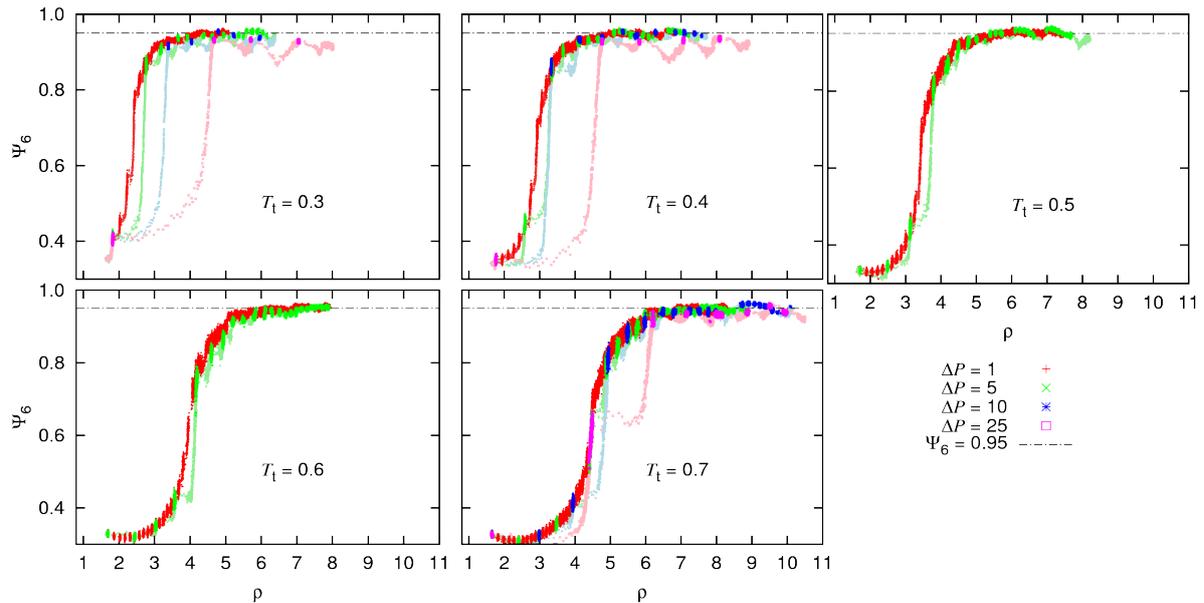


Figure 5.18.: Order parameter  $\Psi_6$  as a function of density during a compression run at different  $T_t$  (in different panels) and  $\Delta P$ , as labeled. Symbols in faded colour represent measurements out-of-equilibrium states while symbols with solid colour represent data when  $\delta T$  and  $\delta P$  lie within the range of tolerance. The dash-dotted black line marks the value  $\Psi_6 = 0.95$ .

realise the necessary cluster rearrangements in order to correct the defects in the crystal and is therefore not able to recover equilibrium before the next pressure increment is applied.

We have also checked if the fact that the order parameter does not increase above  $\Psi_6 \simeq 0.95$  is due to the size of the ensemble. In figure 5.19 we show the order parameter as a function of density measured during the compression of a system with  $N = 6144$  and  $N = 20736$  particles. We see that in both cases  $\Psi_6$  saturates at the same value:  $\Psi_6 \simeq 0.95$ , ruling out the origin of this behaviour in size effects. We anticipate that the formation of defects in the crystal is inherent to the method and that an annealing process, where temperature is raised at constant pressure and then decreased to the original value, helps to attain a higher order in the system. We will deepen in this issue in subsection 5.3.3.

### Miscellaneous

To conclude this section we provide links to some videos of a compression experiment performed at  $T_t = 0.7$  using  $\Delta P = 10$ . In the first video we show the full compression run at a high speed; in the second and third ones we select a particular time window of the simulation and play the movie in slow motion so that details can be better appreciated. The third video corresponds to the time window presented in figure 5.14.

- First video → <https://www.youtube.com/watch?v=bjN-A108E8A>

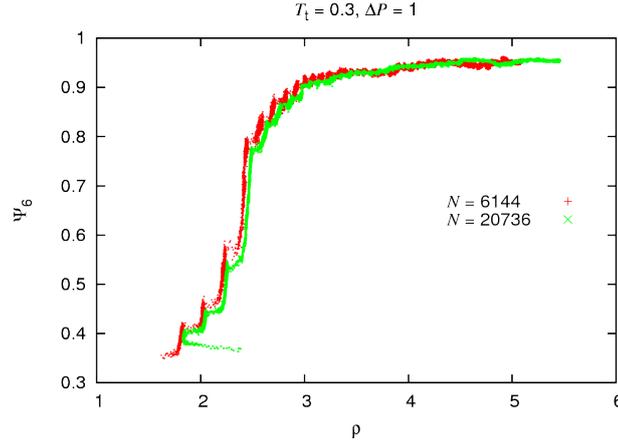


Figure 5.19.: Order parameter  $\Psi_6$  as a function of density during a compression run at  $T_i = 0.3$  using  $\Delta P = 1$  in a system of  $N = 6144$  particles (red) and  $N = 20736$  (green).

- Second video → <https://www.youtube.com/watch?v=CinfhJjMbvq>
- Third video → <https://www.youtube.com/watch?v=TaPnsfHzuVo>

The videos show four panels organised as follows: in the top left panel we show in the title running labels indicating the time ( $t$ ), the measured pressure (here labeled as  $P$  instead of  $P_m$ ), the measured temperature (here labeled as  $T$  instead of  $T_m$ ), the density ( $\rho$ ) and the average cluster occupation (here labeled as  $n_{occ}$  instead of  $\langle n_{occ} \rangle$ ). In the plot we show the density  $\rho$  (green) and the measured temperature  $T$  (red) as a function of the measured pressure  $P$  during the whole run. The grey horizontal bar represents the tolerance interval of temperature where  $\delta T < 0.002$ . Two big black crosses indicate the position in the  $\rho$ - and  $P$ -curve where we are. In the top right panel we plot the positions of the particles in grey and the centre of mass of the clusters with open circles, coloured according to the  $\Psi_6$  order parameter value for each cluster. The coloured bar in the right hand side shows the legend and the black arrow head points in the legend the average  $\Psi_6$  value in the system at that moment. In the bottom left panel we do a similar plot as in the top right one, but the centres of mass of the clusters are now coloured according to the cluster occupation. Again, the black arrow head points to the average  $n_{occ}$  value in the system at that moment. Finally, in the bottom right panel we plot the particles as filled circles coloured according to their kinetic energy and the centres of mass of the clusters as opened circles coloured according to  $T_{cluster}$ , as defined in equation (5.15).

In the bottom left panel merging events are easy to spot because there is a sudden increase in the cluster occupation and therefore a strong change in the colour of the cluster. Since clusters which have merged remain for a while over-occupied, the cascades of subsequent merging events introduced previously on the text are easy to identify.

In the top right panel merging events are also easy to spot because right before the event occurs, the clusters which are going to merge have a low  $\Psi_6$  value. One can see how as

merging events occur, the order in the whole system increases.

Finally, in the bottom right panel we see how the kinetic energy is homogeneously distributed in the system and it does not play a significant role in merging events. One could expect the surface of the crystal in contact with the ideal gas to be slightly hotter and the centre of the system to be slightly cooler, but this is not the case.

### 5.3.3. Equation of state

In what follows we concentrate on the results obtained at  $T_t = 0.4$  using a compression rate of  $\Delta P = 1$  and analyse whether the obtained  $\rho(P_m)$  curve corresponds to the equation of state  $\rho_{\text{equ}}(P, T = 0.4)$ . We choose the smallest value of  $\Delta P$ , namely  $\Delta P = 1$ , firstly because in this way we obtain a more detailed  $\rho(P_m)$  curve, since we are sampling more (equilibrium)  $\rho(P_m)$  points in the  $P_m$ -axis, and secondly because we have seen in subsection 5.3.1 that the smaller the  $\Delta P$ -value the closer  $\delta T$  will stay to zero, *i.e.* the closer the system will remain to equilibrium. We test if the  $\rho(P_m) = \rho_{\text{equ}}(P, T = 0.4)$  by selecting particle configurations of different state points along the  $\rho(P_m)$  curve and start simulations where the pressure is decreased, *i.e.*  $\Delta P = -1$ ; this corresponds to an expansion run. If the system reproduces the same  $\rho(P_m)$  curve, *i.e.* if  $\rho_{\text{comp}}(P) = \rho_{\text{exp}}(P)$ , we can argue that the GEM-4 system is in thermodynamic equilibrium during the whole run and thus  $\rho(P_m)$  corresponds, at the given  $T$ -value to the equation of state.

However, this is not the case, as it can be observed in figure 5.20. In the main panel we plot  $\rho(P_m)$  during a compression run in red and during three expansion runs in blue. The big black circles mark the initial state points of the expansion runs. We identify at  $P_m \sim 15$  the transition between the fluid and the ordered state. In the fluid state  $\rho_{\text{comp}}(P) = \rho_{\text{exp}}(P)$ , while in the ordered phase the expansion curves first follow parallel trajectories in the  $\rho, P$  plane (at high  $P_m$ ), which do not coincide with the compression trajectory. Then a second regime is observed in the blue curves at low  $P_m$ -values, where  $\rho$  as a function of  $P_m$  becomes steeper.

In order to test if the fact that  $\rho_{\text{comp}}(P) \neq \rho_{\text{exp}}(P)$  is a hysteresis effect, we take two particle configurations along two of the expansion runs and compress them. Results on  $\rho(P_m)$  during these new compression runs are shown in dark red the inset of figure 5.20, where data of the initial compression and expansion runs is shown in light red or blue. The initial state points of the new compression runs are marked with a circle. The inset corresponds to the  $(\rho, P_m)$  regime specified by a grey rectangle in the main panel. From a hysteresis effect we would expect the data of the new compression runs to follow trajectories which are parallel to the initial compression run. In the contrary, the dark red curve follows the trajectory of the light blue curve and when the point where the light blue and light red curve meet is reached, *i.e.* the  $\rho(P_m)$ -value at which the expansion run was launched, the slope of the dark red curve changes and it follows the light red one.

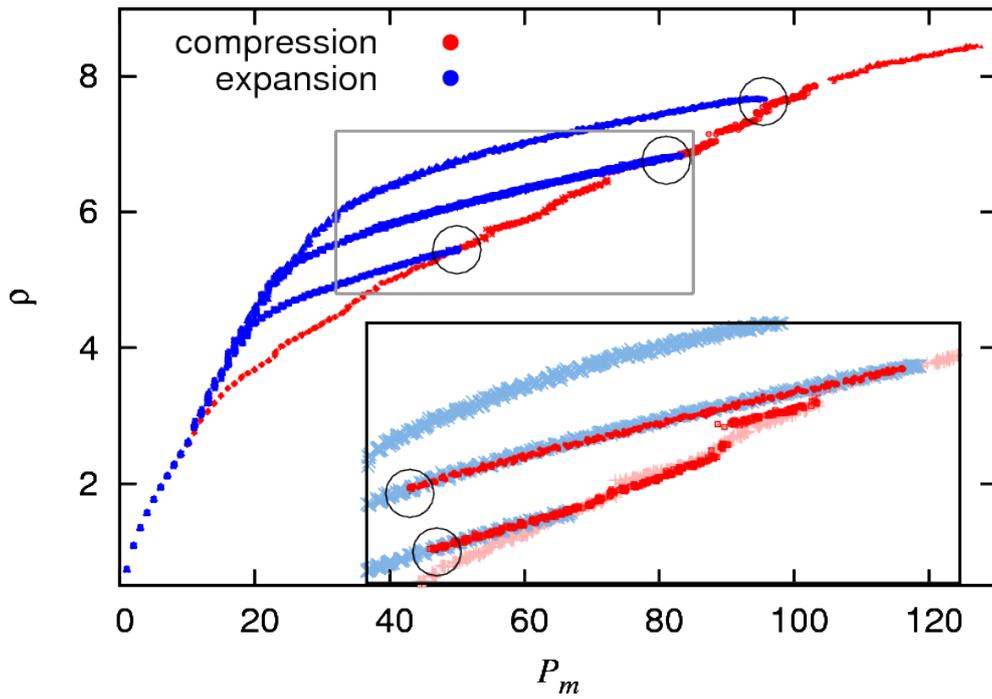


Figure 5.20.: Density  $\rho$  as a function of the pressure  $P_m$  during compression (red) and expansion (blue) runs in a GEM-4 system of cluster-forming particles at  $T_t = 0.4$ , where  $\Delta P = \pm 1$ . The inset shows the  $(\rho, P)$ -regime marked with a grey rectangle in the main panel. The initial compression run is shown in light-red and the expansions runs are shown in light-blue. New compression runs are shown in red. The big black circles mark the starting state points of the expansion runs (main panel) and the new compression runs (inset).

### Cluster occupation *vs* lattice constant

To better elucidate what is happening in the GEM-4 system along the different branches in the  $(\rho, P_m)$ -plane we take a closer look at the two different mechanisms through which a cluster crystal can accommodate a change in volume: (i) the shrinkage/growth of the lattice constant, reflected in a shrinkage/growth of the position of the main peak of the radial distribution function  $d_{nn}$  and (ii) the deletion/creation of new lattice positions, reflected in a increase/decrease of the average cluster occupation  $\langle n_{occ} \rangle$  (obtained *via* a cluster analysis as described in appendix B).

We first analyse the evolution of  $\langle n_{occ} \rangle$ . In figure 5.21 we show in panel (a) the different branches of  $\rho(P_m)$  in the  $(\rho, P_m)$ -plane obtained in the compression and expansion runs coloured by the respective value of the pressure. In panel (b) we show  $\langle n_{occ} \rangle$  as a function of  $\rho$  along the compression and expansion runs, and the points are again coloured according to  $P_m$ . In both panels we have marked with a circle the  $\rho$ - and the  $\langle n_{occ} \rangle$ -values at the moment when the expansion runs were launched. We can classify the different branches into three categories according to the behaviour of  $\langle n_{occ} \rangle$ :

1. along the initial compression branch, marked by a dashed line,  $\langle n_{occ} \rangle$  increases linearly with the density. We shall refer to this branch as the *compression branch*, as the data corresponds only to data from compression runs.
2. Along the final part of the expansion runs, marked by a dash-dotted line,  $\langle n_{occ} \rangle$  is again a linear function of the density. We shall refer to this branch as the *expansion branch*, since the data corresponds only to data of expansion runs.
3. Along the initial parts of the expansion runs, indicated with solid lines,  $\langle n_{occ} \rangle$  is constant (there are no cluster merging or splitting events). We shall refer to these branches as the *intermediate branches* (lower, middle and upper), since the data correspond either to compression or to expansion runs (inset of figure 5.20) and they connect the *compression* and *expansion branches*.

We can conclude that the two different slopes observed in the  $\rho(P_m)$  curves during expansion runs correspond to an initial volume expansion at constant average cluster occupation and a second volume expansion *via* a reduction of  $\langle n_{occ} \rangle$ , *i.e.* *via* a production of new lattice positions.

We present the information about the lattice spacing by showing in figure 5.22 the same plot as in panel (b) of figure 5.21 but now the data is coloured according to the  $d_{nn}$ -value. In the low density range ( $\rho \lesssim 3$ ) the distance to the nearest neighbour is very short because the system is in a fluid configuration. The abrupt change in  $d_{nn}$  at  $\rho \sim 3$  is an artefact of the cluster analysis: at this point we are in the ordered phase and we add additional steps to the analysis (appendix B, step 4 is responsible for the jump in  $d_{nn}$ ). We observe different values of  $d_{nn}$  in the different branches:

1. In the *compression branch*, for  $3 \lesssim \rho \lesssim 5$ ,  $d_{nn}$  decreases until it reaches a value,

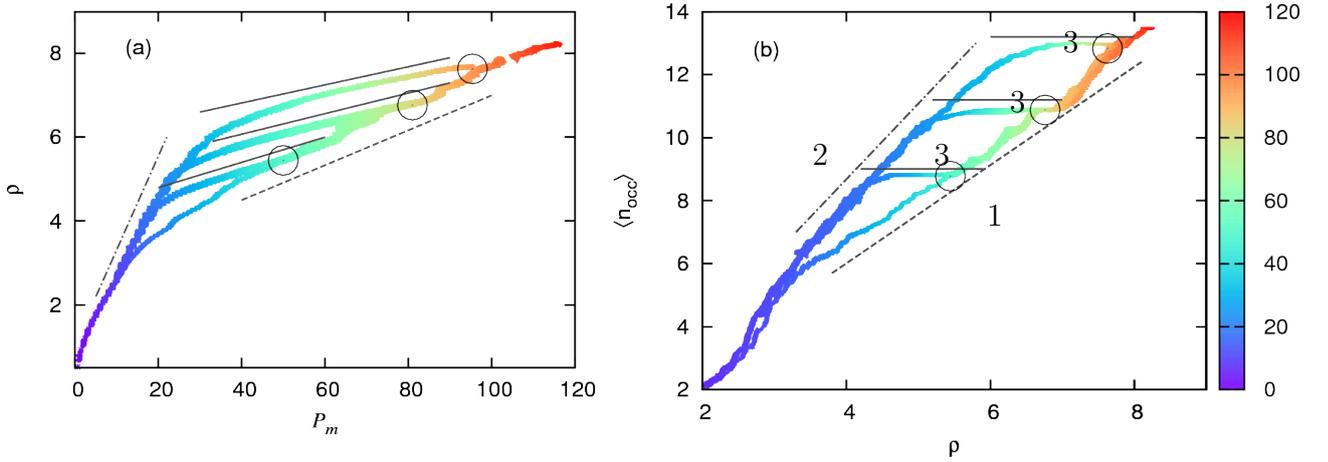


Figure 5.21.: Panel (a): density as a function of pressure in a GEM-4 system at  $T_t = 0.4$  during compression and expansion runs where  $\Delta P = \pm 1$ . Panel (b): average cluster occupation as a function of the density for the same runs. The data are coloured in both panels according to the respective  $P_m$ -value. The dotted line indicates the compression branch (1). The solid lines indicates the intermediate branches (lower, middle and upper) (3). The dash-dot line indicates the expansion branch (2).

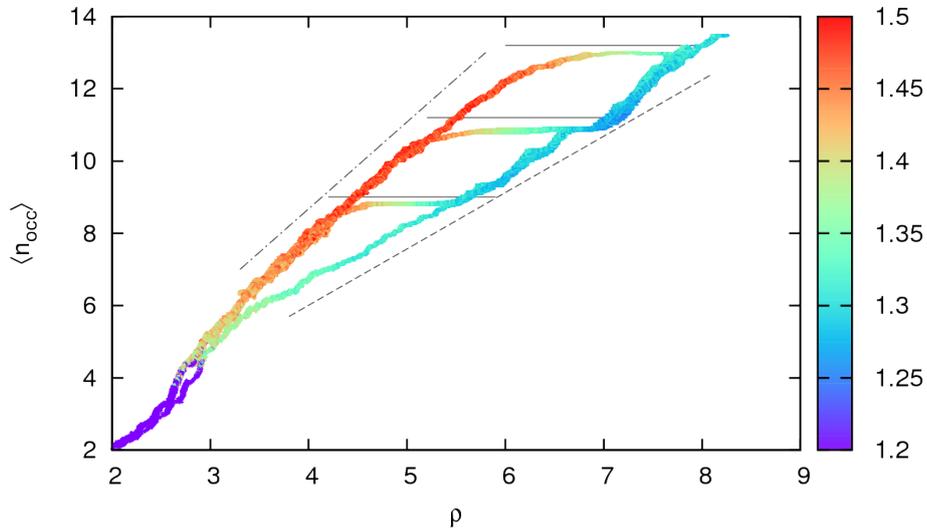


Figure 5.22.: Average cluster occupation as a function of the density in a GEM-4 system at  $T_t = 0.4$  during compression and expansion runs where  $\Delta P = \pm 1$ . The data are coloured according to the respective  $d_{mn}$  value. The dotted line demarcates the compression branch. The solid lines demarcate the intermediate branches (lower, middle and upper). The dash-dot line demarcates the expansion branch.

$d_{\text{nn}} \sim 1.28$ , which remains essentially constant for the rest of the branch. This value corresponds to the minimum saturated value that we observed in figure 5.16 during the compression experiments in subsection 5.3.2.

2. In the *expansion branch* we also observe a constant  $d_{\text{nn}}$  value,  $d_{\text{nn}} \sim 1.48$ , at high densities and a small variation again in the same  $\rho$ -regime ( $3 \lesssim \rho \lesssim 5$ ).

These two values are the maximum and minimum  $d_{\text{nn}}$  values observed.

3. In the intermediate branches  $d_{\text{nn}}$  varies between  $d_{\text{nn}}^{\text{max}}$  and  $d_{\text{nn}}^{\text{min}}$ .

We conclude that if an ordered system of GEM-4 particles is compressed or expanded, variations in  $d_{\text{nn}}$  and  $\langle n_{\text{occ}} \rangle$  are completely decoupled: its immediate reaction is to reduce or enhance the spacing of its lattice,  $a$ , within a range  $[d_{\text{nn}}^{\text{max}}, d_{\text{nn}}^{\text{min}}]$  while keeping  $\langle n_{\text{occ}} \rangle$ , *i.e.* the system will move along the *intermediate branches*. If the system which has reached a state where  $a = d_{\text{nn}}^{\text{min}}$  is further compressed, the inter-cluster interaction energy will be high enough to activate the cluster merging processes: the system will achieve the volume reduction maintaining  $a = d_{\text{nn}}^{\text{min}}$  by deleting lattice positions, *i.e.*  $\langle n_{\text{occ}} \rangle$  increases. On the other hand, if the system which has reached a state where  $a = d_{\text{nn}}^{\text{max}}$  is further expanded, the inter-cluster repulsion will not be strong enough to stabilize the clusters of overlapping particles: thus these aggregates will start to split, maintaining  $a = d_{\text{nn}}^{\text{max}}$  the volume will increase by creating new lattice positions, *i.e.*  $\langle n_{\text{occ}} \rangle$  decreases.

### Annealing processes

The question which density value corresponds to an equilibrium state at a given pressure, *i.e.* what is the equation-of-state  $\rho_{\text{equ}}(P_{\text{m}})$ , remains still unanswered. Based on the DFT calculations within the MFA [23], we can assure that for each  $P_{\text{m}}$  there is one unique value for  $\rho$ ,  $a$  and  $\langle n_{\text{occ}} \rangle$  which corresponds to the equilibrium state. We assume that in our experiments we are observing different values  $\rho$ ,  $a$  and  $\langle n_{\text{occ}} \rangle$  along the different branches because the system is trapped in metastable states separated from the equilibrium state by a large energetic barrier. We thus launch several annealing processes for selected  $P_{\text{m}}$ -values starting from particle configurations at that pressure selected from different state points in the compression, expansion and intermediate branches.

During the annealing processes we maintain the target pressure constant and increase the target temperature until the GEM-4 system loses the ordered structure. Then we cool the system down to the initial temperature  $T_{\text{t}} = 0.4$ . We hope that the energy introduced into the system by the heating processes allows the GEM-4 particles to overcome the energetic barrier between the initial and the equilibrium state and expect all the annealing runs starting from different  $\rho$ -values (in the different branches) at the same  $P_{\text{m}}$ -value, to end up at the same density, *i.e.*  $\rho_{\text{equ}}(P_{\text{m}})$ .

In figure 5.23 we show the trajectory followed in the  $(\rho, T_{\text{m}})$ -plane by three annealing processes launched from different densities at  $P_{\text{t}} = 70$ . The values of  $\rho$  in the initial configura-

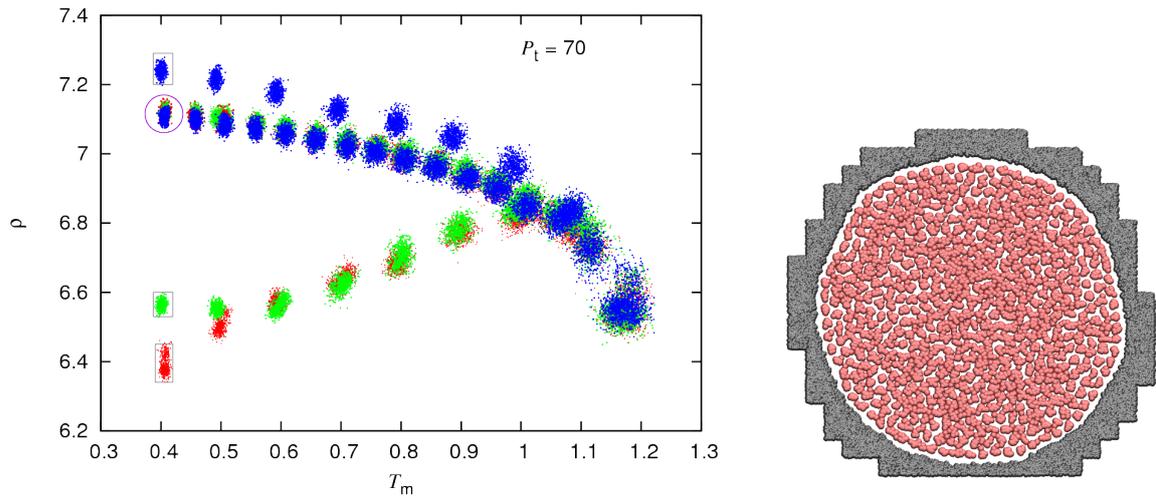


Figure 5.23.: Left:  $\rho$  vs  $T$  along annealing processes of a system of GEM-4 cluster-forming particles launched at  $P_t = 70$  from the compression (red), middle- (green) and upper-intermediate (blue) branches. The starting points of a process are marked with a rectangle. As the run arrives at  $T_m = 1.2$ ,  $\Delta T$  is changed from  $\Delta T = 0.1$  to  $\Delta T = -0.05$ . The final points are marked with a violet circle. Right: A snapshot of the system at the moment where  $\Delta T$  is switched.

tions are marked with a grey rectangle, the corresponding particle configurations were taken from the compression branch (red), the middle intermediate branch (green) and the upper intermediate branch (blue). During the heating process, the green and the red trajectories merge at  $T_m = 0.6$  and follow for the rest of the annealing run the same  $\rho(T_m)$  curve. The blue trajectory meets the red and the green at  $T_m = 1.1$ , we observe, looking at the particle configurations, that this is the point where the system starts to melt. As the annealing processes reach a temperature of  $T_m = 1.2$ , the temperature increment  $\Delta T$  is changed from  $\Delta T = 0.1$  to  $\Delta T = -0.05$ . A snapshot of the particle configuration taken at that moment is shown in the right hand side: it can be seen that the GEM-4 system has lost the ordered structure. In the subsequent process the three cooling runs follow essentially the same trajectory; eventually they reach a final common point  $\rho_{\text{equ}}(P_m = 70)$ , marked with a violet circle. Figure 5.23 demonstrates that the annealing processes lead indeed to an equilibrium state.

The annealing process was repeated for  $P_m = 30, 40, 53, 60$  and  $80$ . The reader can imagine that the corresponding curves in the  $(\rho, T_m)$ -plane are similar to those shown in figure 5.23, of course the starting densities are different, leading, however, for each  $P_m$ -value to a common  $\rho_{\text{equ}}(P_m)$ . Those results obtained for different  $P_m$ -values are summarized in figure 5.24, where the original  $\rho(P_m)$  curves from the compression (light-red) and the expansion (light-blue) are shown (see figure 5.20). Additionally, the results from the annealing processes are shown: the respective starting  $\rho(P_m)$ -values (marked in figure 5.23 by the grey rectangles) are plotted in light-grey and the final  $\rho_{\text{equ}}(P_m)$ -values for all the annealing runs (highlighted

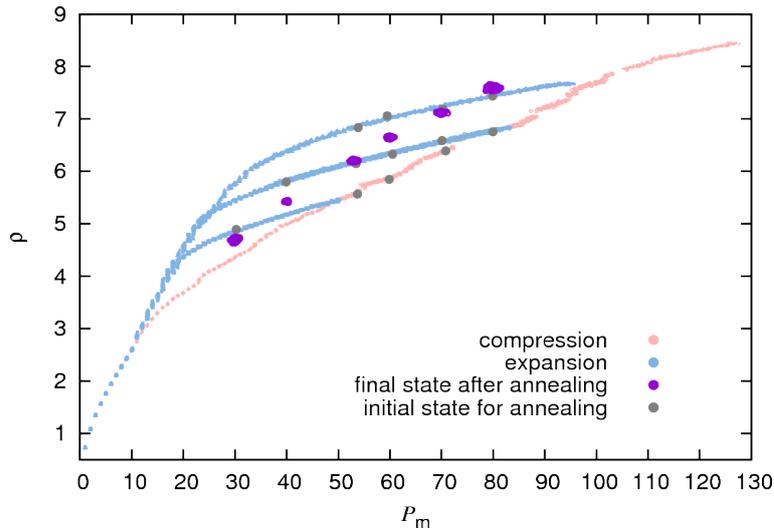


Figure 5.24.:  $\rho(P_m)$  curves for the compression (light-red) and expansion (light-blue) processes of a system of GEM-4 cluster-forming particles at  $T_t = 0.4$ . The grey circles denote the starting points of the annealing processes; along those runs the pressure of the system is kept constant, while the temperature is first raised until the system has melted and is afterwards decreased to the original target value  $T_t = 0.4$ . The violet points denote the final density measured after annealing in the different runs.

in figure 5.23 by a violet circle) are shown in violet. We see how through the respective annealing processes we have succeeded in finding, for every  $P_m$ -value a final  $\rho$ -value common to all the runs starting from different state points. We assume the curve followed by the purple data to be the equation-of-state  $\rho_{\text{equ}}(P_m)$ .

Finally, we have a look at the details of the structure that the system forms along the  $\rho_{\text{equ}}(P_m)$  curve. We characterise the arrangement of the clusters *via* the average distance to the nearest neighbour cluster,  $d_{\text{nn}}$ , and via the hexagonal order parameters,  $\Psi_6$  defined in equation (5.17) (both refer to the COM of the clusters). In figure 5.25 we show the correlation between  $d_{\text{nn}}$  and  $\Psi_6$  for all the state points investigated; data from the compression (red) and from the expansion (blue) runs. These results are compared with the correlation values obtained after the annealing processes (violet) and the value of the lattice spacing predicted by DFT calculations in the MFA. On the one hand we observe that the order parameter has slightly increased after the annealing processes, where it attains its maximum value: obviously annealing is able to heal out lattice defects in the cluster crystal. On the other hand we see that the obtained value for  $d_{\text{nn}} \sim 1.37$  is slightly below the one predicted by DFT ( $d_{\text{nn}} = 1.42$ ). We attribute this difference to the fact that we are investigating a nanodrop and not a bulk system.

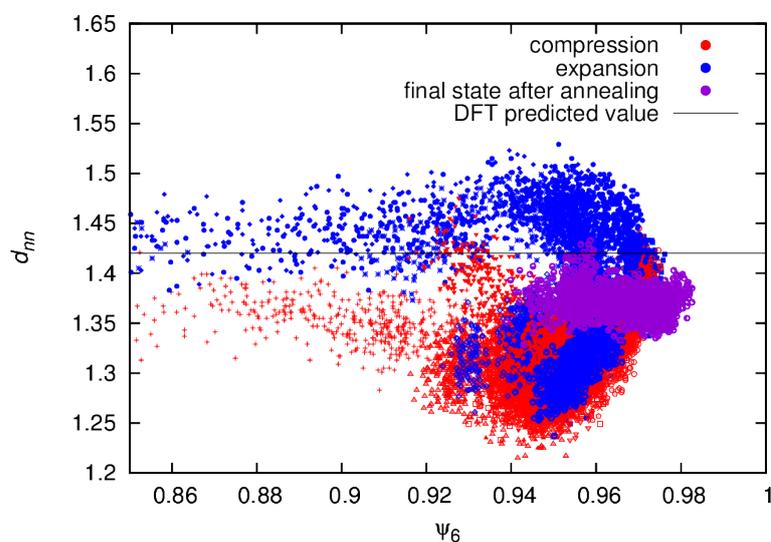


Figure 5.25.: Correlation plot of  $d_{nn}$  and  $\Psi_6$  (see text) for the compression (red) and the expansion (blue) processes of a system of GEM-4 cluster-forming particles at  $T_t = 0.4$ . The violet points denote the data obtained at the end of the different annealing runs. The DFT predicted value within the mean field approximation for a 2D hexagonal lattice of GEM-4 particles is shown with a solid line.

# 6. Amphiphilic, cluster-forming polymer chains

## 6.1. Introduction

In this chapter of the thesis we study the behaviour of different polymeric structures with the intention of designing an alternative macromolecule to dendrimers which can form a cluster crystal at certain  $\rho, T$  conditions and aims to be easy to synthesise in a lab. Motivated by the results presented by Meijer and co-workers [29], we explore the possibility of assembling a cluster crystal with polymeric linear chains decorated by amphiphilic side groups. In reference [29], Meijer *et al.* synthesised a series of water soluble random copolymers

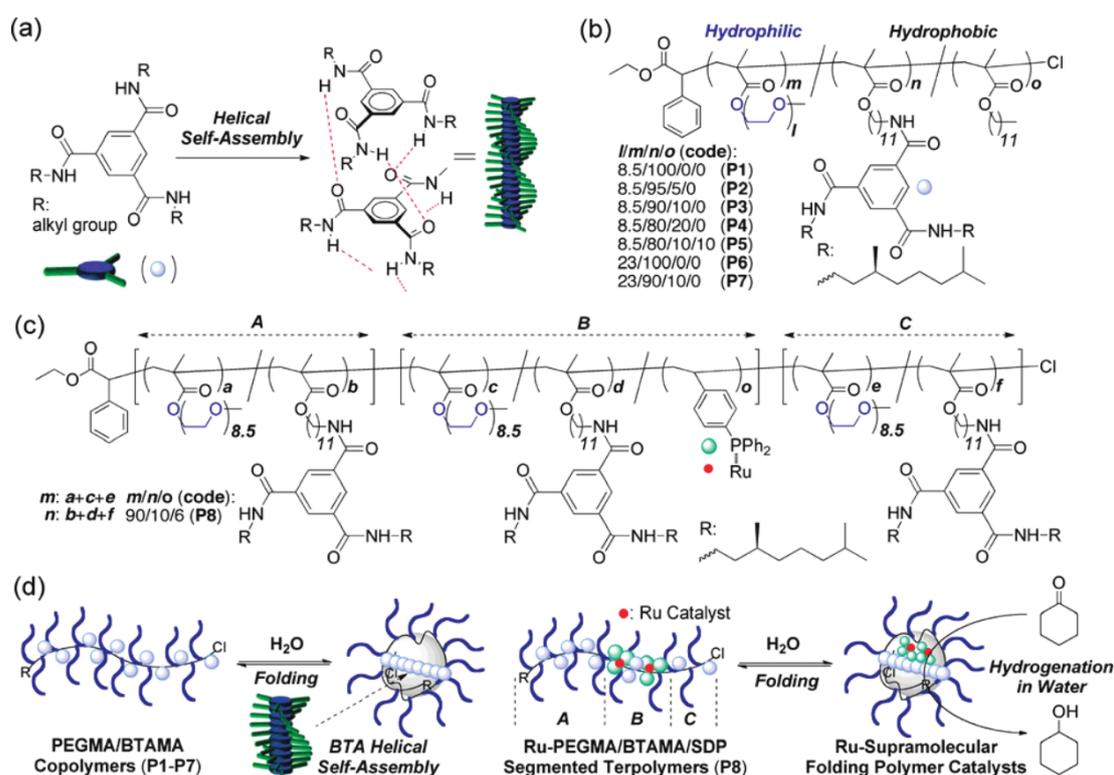


Figure 6.1.: (a) Helical self-assembly of chiral BTAs. (b) PEGMA/BTAMA copolymers (P1-7) with  $l$ =average number of glycol units,  $m$ =average number of PEGMA units,  $n$ =average number of BTAMA units and  $o$ =average number of dodecyl methacrylate units. (c) Ru-PEGMA/BTAMA/SDP segmented terpolymer. (d) Supramolecular single-chain folding of polymers in water affording a compartmentalised catalyst. Taken from [29].

(PEGMA/BTAMA copolymers P1-7) built up with blocks of hydrophilic poly(ethylene glycol) (PEG) and hydrophobic chiral benzene-1,3,5-tricarboxamide (BTA) substituents, where the different P1-7 copolymers contained different PEG and BTA concentrations (see panel (b) of figure 6.1). Their motivation to synthesise these polymers was to arrive at a polymeric chain carrying spatially separated catalytic units. The structural characterization of PEGMA/BTAMA copolymers by various scattering techniques strongly supported the view that intramolecular folding of the pendant BTAs results in collapsed single-chain nanoparticles [29].

The amphiphility of the PEGMA/BTAMA chains and their ability to fold into a hollow structure are very promising for the formation of cluster crystals. Therefore we made an *in silico* design of an amphiphilic polymer chain; we varied several parameters in the model, as the length of the chain, the nature of monomers (solvophobic, neutral or solvophilic), the bond lengths and the monomer size and softness, which appear as parameters in the monomeric potentials; and we calculated the zero-density effective interaction to test *via* the criterion presented in subsection 2.2.1 if these chains will freeze into a cluster crystal. Once we find a model whose Fourier transform of the effective interaction has negative components, we estimate the spacing of the lattice (following the guidelines in section 2.3) and simulate a cluster crystal. We start our simulation by positioning the polymers in a ordered crystal structure and check if the crystal remains stable or if on the contrary the crystal melts.

The rest of the chapter is organised as follows: in section 6.2 we present the model system, in section 6.3 we show the effective interactions of the different variations of the model calculated at zero-density and the estimated freezing line and finally in section 6.4 we present the results of monomer resolved and coarse grained simulations of a selected model in the fluid and crystal phase.

## 6.2. Model

We design two kinds of amphiphilic chains (AC): in the first one (left panel of figure 6.2) we construct the backbone of the polymer with neutral monomers (C) and we attach solvophobic (A) and solvophilic (B) side groups. We denote these chains as “ABC chains”. In the second kind of AC (right panel of figure 6.2) we dispense with the neutral monomers and construct the backbone with solvophobic (A) monomers and attach solvophilic (B) side groups to it. We denote these chains as “AB chains”.

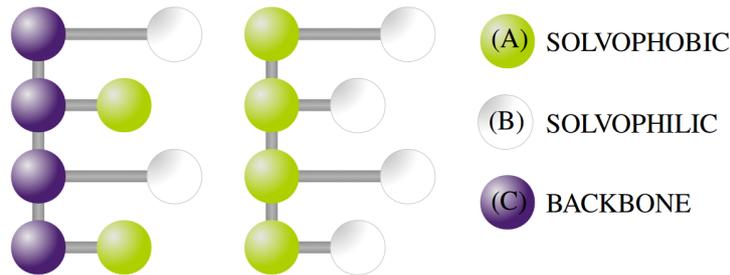


Figure 6.2.: Left: Sketch of an ABC chain. The backbone is formed by neutral (C) monomers, attached to it are solvophobic (A) and solvophilic (B) side groups. Right: Sketch of an AB chain. The backbone is formed by solvophobic (A) monomers and attached to it are solvophilic (B) side groups.

We will use  $\epsilon_{\text{unit}} = \epsilon_{AA}$  and  $\sigma_{\text{unit}} = 0.5\sigma_{BB}$  of the Morse interaction in our first model AC-1 as the energy and length units, respectively (see table 6.1). All the monomers in the chain have the same mass,  $m$ , which we will use as the mass unit. The rest of the quantities in the text are expressed in reduced units as described in appendix A.

The number of A, B or C monomers in the chain are designated by  $N_A$ ,  $N_B$  and  $N_C$ , respectively. The total length of the chain  $L$  is defined as  $L = N_A + N_B + N_C$ , and is varied from  $L = 100$  to  $L = 20$ . In the case of the ABC chains, the ratio between  $N_A$  and  $N_B$  is varied for a given  $L$  and  $N_C$  is always  $N_C = N_A + N_B$ . In the case of AB chains  $N_A = N_B$  always. All the monomers in the chains interact with each other *via* the Morse potential (equation (2.31)) and bonded monomers interact *via* the FENE potential (equation (2.30)). In tables 6.1 and 6.2 we list the parameters of the Morse and FENE interaction of AC-1 to AC-14, where AC-1 to AC-7 are ABC chains and AC-8 to AC-14 are AB chains.

Morse	AC-1	AC-2	AC-3	AC-4	AC-5	AC-6	AC-7
$\epsilon_{AA}$	1.0	1.15	0.7	1.0			
$\epsilon_{BB}$	0.05						
$\epsilon_{CC}$	0.05						
$\epsilon_{AB}$	0.05						
$\epsilon_{BC}$	0.05						
$\epsilon_{CA}$	0.05						
$\sigma_{AA}$	0.8			0.9	1.0	0.8	
$\sigma_{BB}$	2.0						
$\sigma_{CC}$	1.7						
$\sigma_{AB}$	2.0						
$\sigma_{BC}$	1.7						
$\sigma_{CA}$	1.7						
$\alpha_{AA}$	2.5			2.2	1.75	2.5	
$\alpha_{BB}$	2.5						
$\alpha_{CC}$	2.5						
$\alpha_{AB}$	2.5						
$\alpha_{BC}$	2.5						
$\alpha_{CA}$	2.5						
FENE	AC-1	AC-2	AC-3	AC-4	AC-5	AC-6	AC-7
$k_{AC}$	40						
$k_{AB}$	40						
$k_{CC}$	40						
$l_{AC}$	4.0					4.5	3.0
$l_{AB}$	4.0					4.5	3.0
$l_{CC}$	1.5	1.6	1.5				
$R_{AC}$	0.75						
$R_{AB}$	0.75						
$R_{CC}$	0.75						

Table 6.1.: Parameters of the Morse and FENE interactions between the monomers in the ABC chains (see text). Blank positions in the table mean that the parameter takes the same value as in the in the column to the left. We mark in blue those parameters which differ from the parameters of AC-1.

Morse	AC-8	AC-9	AC-10	AC-11	AC-12	AC-13	AC-14
$\epsilon_{AA}$	0.714	0.5	0.714	0.6	0.7		0.714
$\epsilon_{BB}$	0.01785	0.015				0.01785	
$\epsilon_{AB}$	0.01785	0.015				0.01785	
$\sigma_{AA}$	1.0						
$\sigma_{BB}$	2.5		2.0	2.5			
$\sigma_{AB}$	1.75		1.5	1.75			
$\alpha_{AA}$	1.8				2.5	2.0	1.9
$\alpha_{BB}$	6.0						
$\alpha_{AB}$	6.0						
FENE	AC-8	AC-9	AC-10	AC-11	AC-12	AC-13	AC-14
$k_{AA}$	60						
$k_{AB}$	30						
$l_{AA}$	3.1875						
$l_{AB}$	3.5625						
$R_{AA}$	0.6375						
$R_{AB}$	0.7125						

Table 6.2.: Parameters of the Morse and FENE interactions between the monomers in the AB chains (see text). Blank positions in the table mean that the parameter takes the same value as in the in the column to the left. We mark in blue those parameters which differ from the parameters of AC-8.

## 6.3. Effective interactions

### 6.3.1. ABC chains

Chain length  $L = 100$

We start our investigations with the ABC chain AC-1 of length  $L = 100$  with  $N_A/N_B = 1$ , for which we measure a radius of gyration  $R_g = 3.64$ . We present the effective interaction calculated in the zero-density limit together with its Fourier transform and some snapshots of two interplaying ABC chains in figure 6.3.

We immediately see that the chosen model is very unrealistic: the solvophobic monomers attract each other very strongly, making the attraction of the effective interaction at a separation  $r = 0$  of the centres of mass stronger than  $-5k_B T$ . Furthermore, based on the Fourier transform of the effective interaction we can predict *via* equation (2.23) that the system will freeze at  $T = 1.0$  into a cluster crystal at the extremely low density of  $\rho_{fr} = 0.007$ , where  $\rho_{fr}$  is defined in units of the radius of gyration.

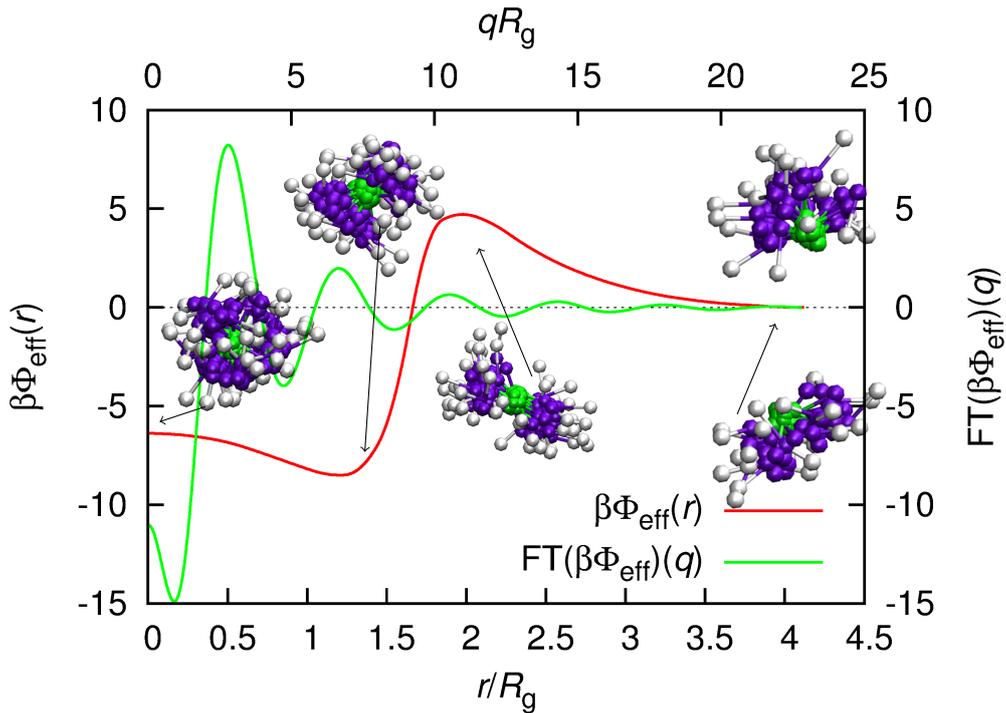


Figure 6.3.: Effective interaction  $\Phi_{\text{eff}}(r)$  (red, bottom  $x$ -axis) of an ABC chain AC-1 of length  $L = 100$  with  $N_A/N_B = 1$ . The Fourier transform of the effective interaction is shown in green, (top  $x$ -axis). Some snapshots of two interacting AC-1 chains separated by different distances are shown.

We thus try to create more realistic models by changing the attractiveness between the monomers (controlled by  $\epsilon_{\nu\mu}$ ,  $\nu, \mu = A, B, C$ ), the bond lengths ( $l_{\nu\mu}$ ), the size of the monomers

	$N_A/N_B$	$R_g$
AC-2	1/2	4.76
AC-2	1/3	6.22
AC-3	1	5.08
AC-4	1/2	4.73
AC-5	1/2	4.04

Table 6.3.: Radius of gyration,  $R_g$ , of different ABC chain models of length  $L = 100$ .

( $\sigma_{\nu\mu}$ ) and their softness ( $\alpha_{\nu\mu}$ ): compared to AC-1 (i) in model AC-2 the attraction between the solvophobic monomers is 15% stronger and the backbone bonds are slightly longer (we use  $N_A/N_B = 1/2$  and  $1/3$ ), (ii) in model AC-3 the attraction between the solvophobic monomers is 30% weaker (we maintain  $N_A/N_B = 1$ ), (iii) in model AC-4 the solvophobic monomers are 12.5% bigger and 12% softer (we use  $N_A/N_B = 1/2$ ) and (iv) in model AC-5 the solvophobic monomers are 20% bigger and 30% softer (we use  $N_A/N_B = 1/2$ ). The effective interactions are shown in figure 6.4 and the corresponding radii of gyration are shown in table 6.3.

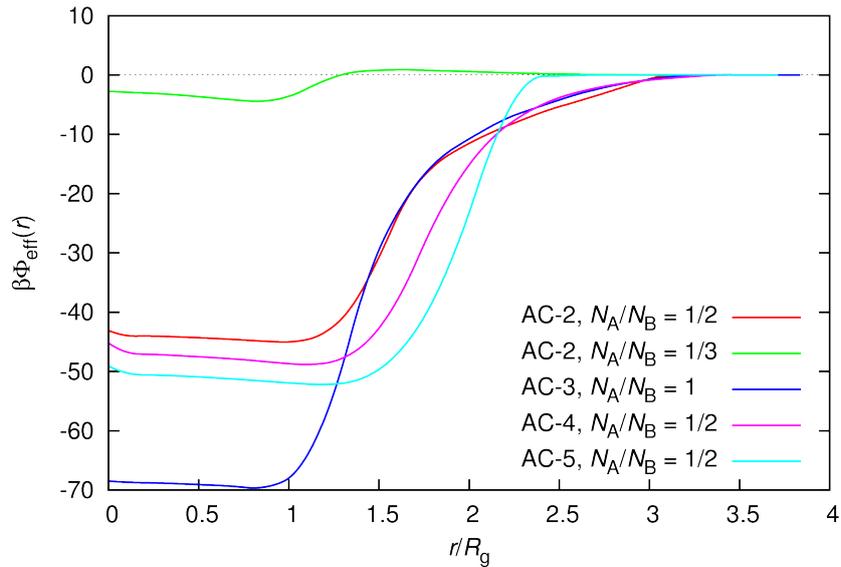


Figure 6.4.: Effective interaction  $\Phi_{\text{eff}}(r)$  of ABC chains of length  $L = 100$  AC-2 with  $N_A/N_B = 1/2$  (red), AC-2 with  $N_A/N_B = 1/3$  (green), AC-3 with  $N_A/N_B = 1$  (blue), AC-4  $N_A/N_B = 1/2$  (pink) and AC-5 with  $N_A/N_B = 1/2$  (cyan). The parameters of the models are listed in table 6.1.

We see that except for case where  $N_A/N_B = 1/3$  (AC-2) the effective interactions are even more unrealistic than in our initial model. Even though the effective interaction of the AC-2 chain with  $N_A/N_B = 1/3$  has become, compared to AC-1 with  $N_A/N_B = 1$ , less attractive at  $r = 0$  ( $\phi_{\text{eff}}(r = 0) \simeq -2k_B T$ ), this value is still too large and the freezing density  $\rho_{\text{fr}} = 0.04$  at  $T = 1.0$  is still very low for the model to be realistic.

From the results presented in figure 6.4 we see that the only change in the model which leads to a better effective interaction, in comparison to AC-1 (not so strong attraction at  $\Phi_{\text{eff}}(r=0)$  and higher freezing density), is the reduction of  $N_A/N_B$ . However, if we look at the left-most snapshot presented in figure 6.5, where the folding process of two AC-2 chains as they approach each other is shown, we see that the volume occupied by the non-attractive monomers (B and C) of the molecules is in this case very big. A further reduction of  $N_A/N_B$  will strongly hinder the formation of clusters, as the first two molecules forming a cluster will easily overlap but then it will become more and more difficult for other molecules to join the cluster as the shell is occupied by repulsive monomers. Therefore, we decide to instead reduce the length  $L$  of the ABC chains. Like this, we could have a similar configuration as the one in the left-most panel of figure 6.5, with the same number of monomers, but a cluster of higher occupation.

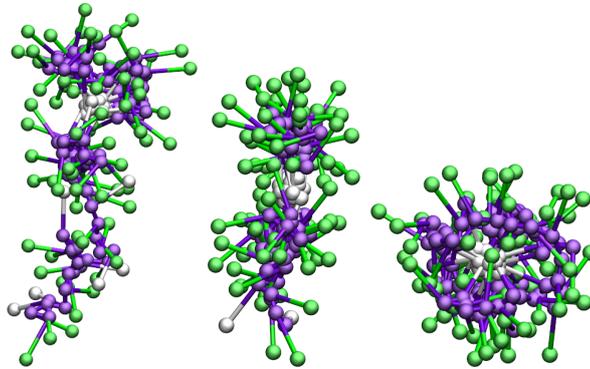


Figure 6.5.: Three snapshots of the folding process of two ABC chains AC-2 of length  $L = 100$  and  $N_A/N_B = 1/3$  as they approach each other.

### Chain length $L = 60, 40$ and $30$

We reduce  $L$  to 60, 40 and 30 and consider AC-1 and AC-6 models for different  $N_A/N_B$  values. In comparison to AC-1, AC-6 has smaller solvophobic (A) monomers and longer bonds. In the top panel of figure 6.6 we present the effective interactions of ABC chains with  $N_A/N_B = 1$  of length 60 and 40. We see that the effective interactions are still too attractive at short distances ( $\Phi_{\text{eff}}(r=0) > -2k_B T$ ). In the middle panel of figure 6.6 we reduce  $N_A/N_B$  to 1/2 and we see that the situation improves, especially for  $L = 30$  where we start to get GEM-like (purely repulsive) effective interactions. The radius of gyration and estimated freezing density of each model are summarised in table 6.4 (for the case of  $N_A/N_B = 1$  we do not compute the freezing density as we can already discard the model by looking at  $\Phi_{\text{eff}}(r=0)$ ). For the chains of length  $L = 60$  we obtain at  $T = 1.0$  a freezing density of  $\rho_{\text{fr}} = 0.2$  for AC-1 and  $\rho_{\text{fr}} = 1.3$  for AC-6. For  $L = 30$  we obtain  $\rho_{\text{fr}} = 0.91$  for AC-1 and  $\rho_{\text{fr}} = 0.89$  for AC-6. Finally, in the lower panel of figure 6.6 we reduce  $N_A/N_B$  to 1/3 and we get for  $L = 30$  effective interactions which give a good fit to a GEM potential with an exponent slightly below 3.

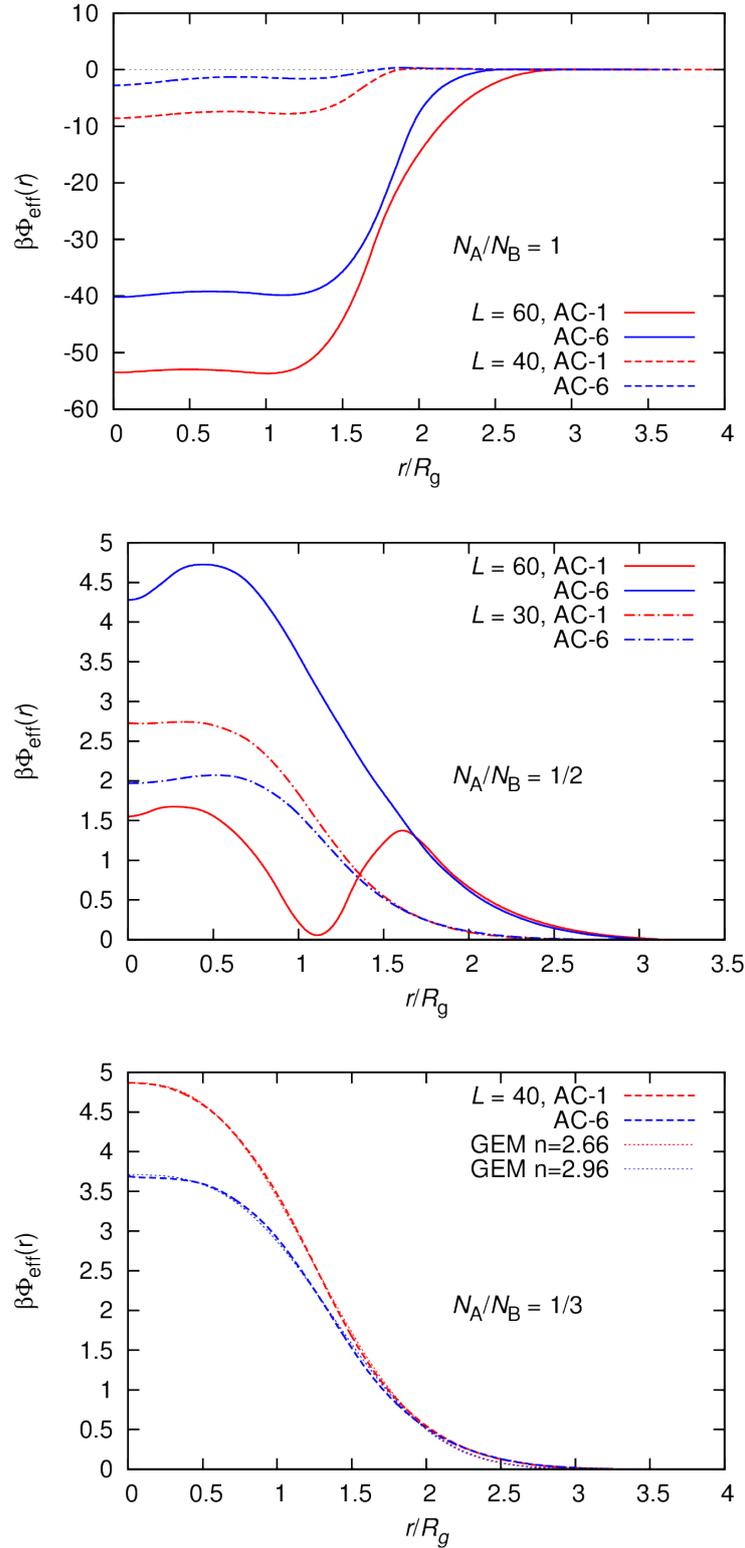


Figure 6.6.: Effective interaction  $\Phi_{\text{eff}}(r)$  of ABC chains AC-1 (red) and AC-6 (blue) of length  $L = 60$  (solid line), 40 (dashed line) and 30 (dot-dashed line). The top panel shows results for  $N_A/N_B = 1$ , the middle panel for  $N_A/N_B = 1/2$  and the lower panel for  $N_A/N_B = 1/3$ , where also a fit to a GEM-like curve is shown (dotted line). The parameters of the models are listed in table 6.1.

	L	$N_A/N_B$	$R_g$	$\rho_{fr}$ at $T = 1.0$
AC-1	60	1	4.29	-
AC-6	60	1	4.43	-
AC-1	40	1	3.82	-
AC-6	40	1	4.06	-
AC-1	60	1/2	4.64	0.2
AC-6	60	1/2	4.85	1.3
AC-1	30	1/2	3.76	0.91
AC-6	30	1/2	3.98	0.89
AC-1	40	1/3	4.15	1.21
AC-6	40	1/3	4.32	0.99

Table 6.4.: Radius of gyration,  $R_g$  and estimated freezing density,  $\rho_{fr}$ , at  $T = 1.0$  (expressed in units of  $R_g$ ) of different ABC chain models.

	$R_g$	$\rho_{fr}$ at $T = 1.0$
AC-1	3.38	0.94
AC-3	3.39	0.99
AC-6	3.98	1.31
AC-7	2.90	1.19

Table 6.5.: Radius of gyration,  $R_g$ , and estimated freezing density,  $\rho_{fr}$ , at  $T = 1.0$  (expressed in units of  $R_g$ ) of different ABC chain models of length  $L = 20$  and  $N_A/N_B = 1$ .

We are moving in the right direction towards a realistic design of cluster-forming chains; by decreasing the length of the chains and the ratio between A and B monomers, the effective interactions have become positive for all  $r$  values, assuming GEM-like shapes and obtaining freezing densities around  $\simeq 1$ . We decide to pursue in this direction and to further reduce  $L$  to 20 and target for similar results (GEM-like potentials) but maintaining  $N_A/N_B = 1$ .

### Chain length $L = 20$ and $N_A/N_B = 1$

We present in figure 6.7 the effective interactions of ABC chains AC-1, AC-3, AC-6 and AC-7 of length  $L = 20$  and  $N_A/N_B = 1$ . All the effective potentials are positive for all  $r$ -values and have a local minimum at small  $r$ -values which will favour clustering. All the potentials show Fourier transforms with negative components. The radius of gyration and estimated freezing density for each model are given in table 6.5. Although all these systems are very promising towards the design of a realistic model, we would still prefer to have a the freezing density of  $\simeq 0.3 - 0.4$  so that the simulation of a bulk cluster crystal is computationally less expensive.

For this reason we decide to simulate AB chains, hoping that by getting rid of the neutral backbone the freezing densities will be considerably lower.

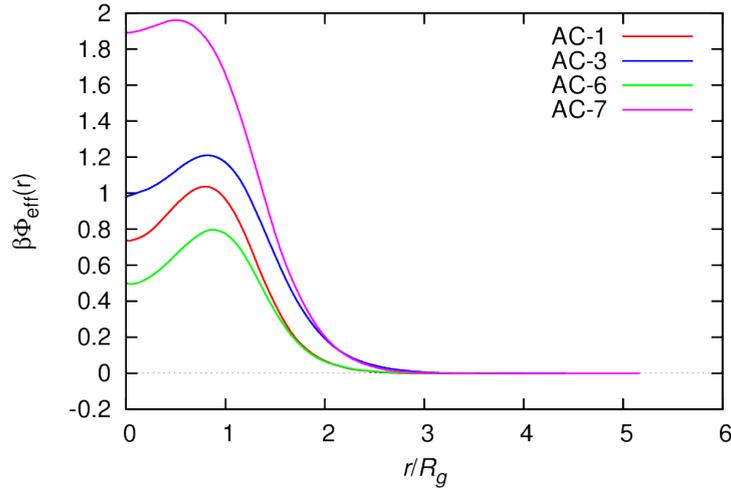


Figure 6.7.: Effective interaction  $\Phi_{\text{eff}}(r)$  of ABC chains AC-1 (red), AC-3 (blue), AC-6 (green) and AC-7 (pink) of length  $L = 20$  and  $N_A/N_B = 1$ . The parameters of the models are listed in table 6.1.

### 6.3.2. AB chains

We calculate the effective interactions of AB chains (right panel of figure 6.2) of length  $L = 20$ . In the AB  $N_A/N_B = 1$ . We define chains AC-8 to AC-14, specified in table 6.2. The effective interactions are shown in figure 6.8 and the radii of gyration and freezing densities at  $T = 1.0$  are given in table 6.6.

It is surprising to see how in AC-10 a decrease of just a 20% of the size of monomers A with respect to AC-8, brings about such a dramatic change in the effective interaction. This demonstrates how sensitive effective interactions are to small details of the monomer interactions. Discarding AC-10, due to the strong ( $< -5k_B T$ ) attraction at short distances, the most promising effective interaction is that of AC-8. The freezing density is rather low ( $\rho_{\text{fr}} = 0.33$ ) and the attraction at the origin is smaller than  $k_B T$ .

	$R_g$	$\rho_{\text{fr}}$ at $T = 1.0$
AC-8	4.73	0.33
AC-9	4.98	1.56
AC-10	4.40	-
AC-11	4.88	0.73
AC-12	5.03	2.05
AC-13	4.87	0.74
AC-14	4.79	0.45

Table 6.6.: Radius of gyration,  $R_g$ , and estimated freezing density,  $\rho_{\text{fr}}$ , at  $T = 1.0$  (expressed in units of  $R_g$ ) of different AB chain models of length  $L = 20$  and  $N_A/N_B = 1$ .

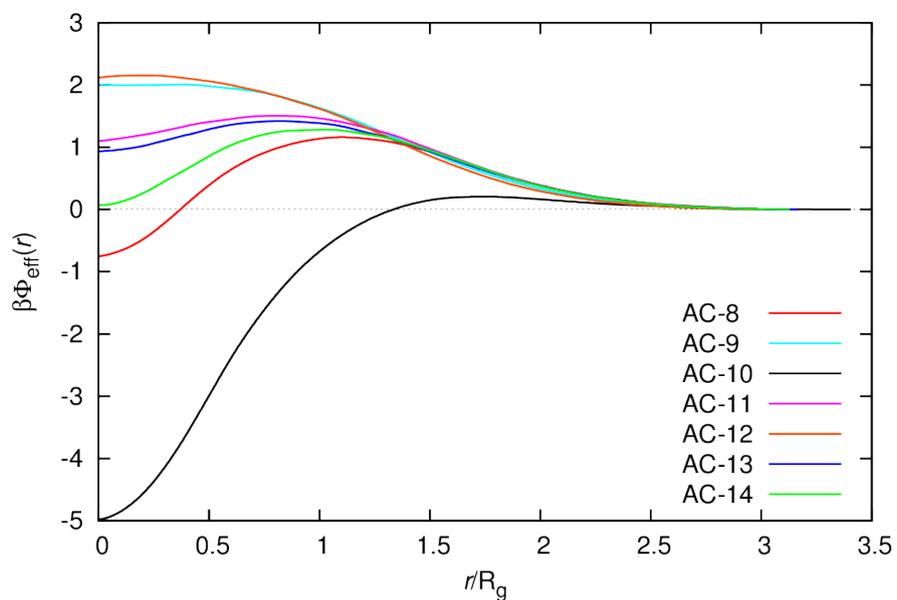


Figure 6.8.: Effective interaction  $\Phi_{\text{eff}}(r)$  of AB chains AC-6 (red), AC-9(cyan), AC-10 (black), AC-11 (pink), AC-12 (orange), AC-13 (blue) and AC-14 (green) of length  $L = 20$ . The parameters of the models are listed in table 6.2.

## 6.4. Monomer resolved and coarse grained simulations of AB chains

From all the designed models of amphiphilic polymer chains we choose our most promising one: the AB chains AC-8 of length  $L = 20$ . We estimate the freezing line by using equation (2.23) and select four state points in the fluid phase, and three in the crystal phase. The phase diagram with the selected state points is shown in figure 6.9.

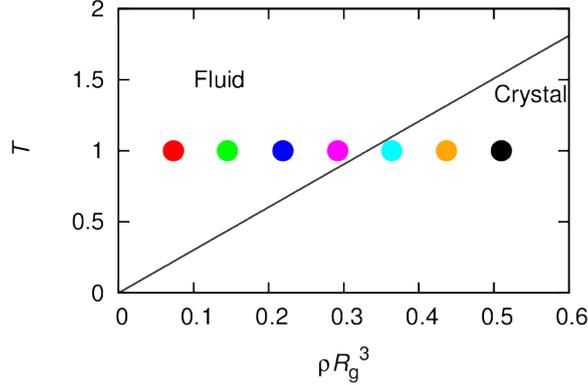


Figure 6.9.: Phase diagram  $T, \rho$  of a system of AB chains AC-8 of length  $L = 20$ . The density is expressed in units of  $R_g$ . The freezing line is estimated by DFT calculations within the MFA (section 2.3). The filled circles correspond to the state points where simulations have been carried out.

### 6.4.1. Simulations in the fluid state

We perform monomer resolved and coarse grained simulations of a system of AC-8 AB chains in a fluid configuration over a time  $t = 1500$  using a time step of  $\delta t = 0.002$  in both cases (MR and CG). In the coarse grained representation we use the effective potential calculated at zero-density, presented in figure 6.8. The number of particles and volume of the simulation box are given in table 6.7.

In the coarse grained simulation we initialise our system by giving the particles random positions in the box and random velocities drawn from a Maxwell-Boltzmann distribution. In the monomer resolved simulation we have to take care because the Morse potential diverges at a monomer separation  $r = 0$ , therefore we have to avoid generating initial configurations where the monomers are initialised too close to each other. This would produce very large initial accelerations on the monomers, which would then move very far in the first time step of the MD integration algorithm and the bonds would break. Therefore, we first initialise each chain one by one isolated from the rest of the system: we place the monomers following a random walk of steps of the length of the corresponding equilibrium bond length. Every

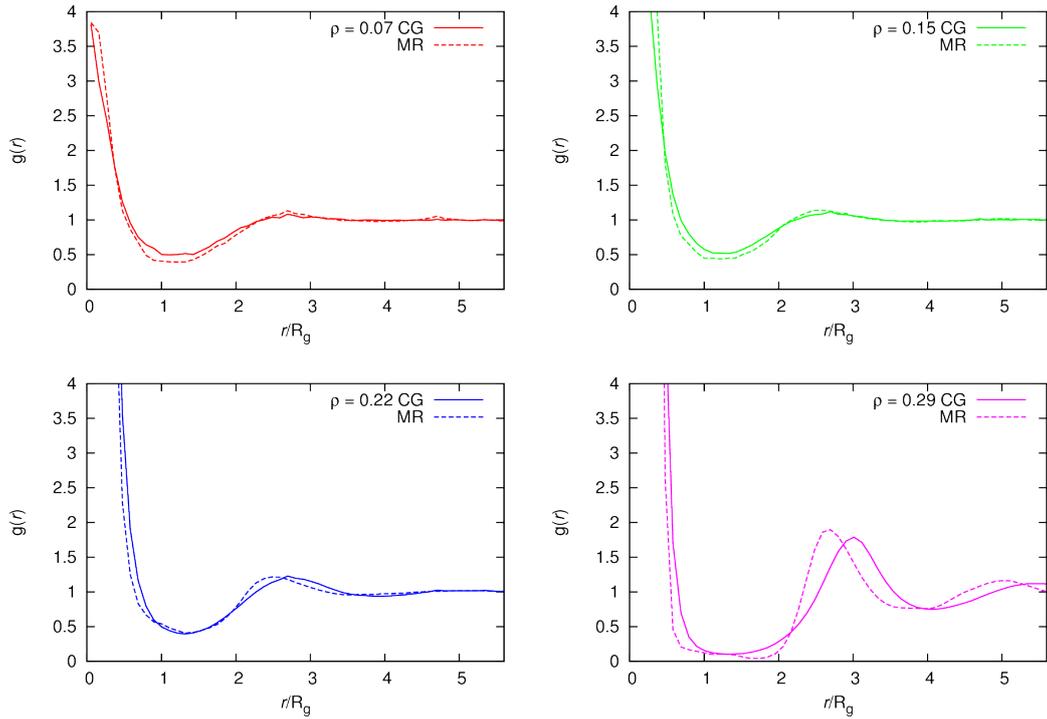


Figure 6.10.: Radial distribution function of AB chains AC-8 of length  $L = 20$  in the coarse grained (solid line) and in the monomer resolved (dashed line) representation at a density  $\rho = 0.07$  (up left),  $\rho = 0.15$  (up right),  $\rho = 0.22$  (down left) and  $\rho = 0.29$  (down right), where the density is expressed in units of the radius of gyration of an isolated chain.

time a new position is drawn for the random walk, we first check if it is overlapping with any of the already existing monomers of the chain. We consider that two monomers are overlapping if the Morse force acting on them is greater than  $10.0\epsilon_{\text{unit}}/\sigma_{\text{unit}}$ . If this is the case we discard this position and draw a new one. Once an isolated chain has been created in this way, a random point in the box is chosen for its centre of mass. We check if by placing the chain at this point an overlap with any monomer in the other already existing chains occurs. If this is not the case we accept this position and proceed with the next chain. If on the contrary there is an overlap, we first randomly rotate the chain, and re-check. If after 5000 random rotations we still didn't find a valid orientation of the chain, we choose a new position for the COM and repeat the process. If after 2000 new positions for the COM in the box we still have an overlap, we delete all the existing chains in the box and start from zero again. The velocities of the monomers are also randomly drawn from a Maxwell-Boltzmann distribution.

In figure 6.10 we present the measured radial distribution functions obtained in simulations of fluid states at different density values. At low densities ( $\rho < 0.2$ ) we obtain a very good agreement between the MR and CG  $g(r)$ . As the density increases many-body interactions become more relevant and since the effective potential that we are using in the CG repre-

sentation does not take these effects into account we start to observe disagreements between the CG and the MR  $g(r)$ . We see cluster formation both for MR and CG simulations at all densities, indicated by the peak of  $g(r)$  at  $r = 0$ . In table 6.7 we show the average cluster occupation and the maximum cluster size found in the system, both of them in the MR simulation, which were identified by performing the first step in the cluster analysis presented in appendix B with  $r_C = 3.0\sigma_{\text{unit}} = 0.65R_g$ . We see that even at the lowest density some quite large clusters are forming.

$\rho$	$V$	$N$	$\langle n_{\text{occ}} \rangle$ (MR)	$n_{\text{occ}}^{\text{max}}$ (MR)
0.07	$54^3$	108	1.25	5
0.15	$54^3$	216	2.25	6
0.22	$54^3$	324	3.48	7
0.29	$54^3$	432	3.63	8

Table 6.7.: Density (in units of  $R_g$  of an isolated chain, see table 6.6), volume and number of chains used in the MR and CG simulations of the fluid phase. We also show the average and maximum cluster size found in the MR simulations.

### 6.4.2. Simulations in the crystal state

We now turn to the simulation of fcc cluster crystals of AC-8 AB chains. Based on the DFT calculations within the MFA, once the system has crossed the freezing line, it will freeze in a bcc crystal. Then, upon increasing density, there will be a phase transition to a fcc lattice [23]. We can predict, based on the behaviour of systems with a similar interaction potentials [21, 28], that the range of stability of the bcc phase will be very narrow in the phase diagram, and therefore we simulate for all densities an fcc lattice. Applying equation (2.19) to our system in hand, we obtain  $a = 18.0\sigma_{\text{unit}} = 3.8R_g$ . We construct an fcc lattice with three cubic unit cells in each direction, *i.e.* with 108 lattice sites. We vary the occupation number of the clusters in order to obtain different densities, as specified in table 6.8. The system is simulated for a time  $t = 1500$  using  $\delta t = 0.002$  both in MR and CG representations.

For the coarse grained simulation we initialise the CG particle on the points of the lattice and assign them random velocity from a Maxwell-Boltzmann distribution. In the monomer

$\rho$	$V$	$N$	$\langle n_{\text{occ}} \rangle$
0.36	$54^3$	540	5
0.44	$54^3$	648	6
0.51	$54^3$	756	7

Table 6.8.: Density (in units of  $R_g$  of an isolated chain, see table 6.6), volume and number of chains used in the MR and CG the simulations in the cluster crystal phase. We also show the average cluster occupation.

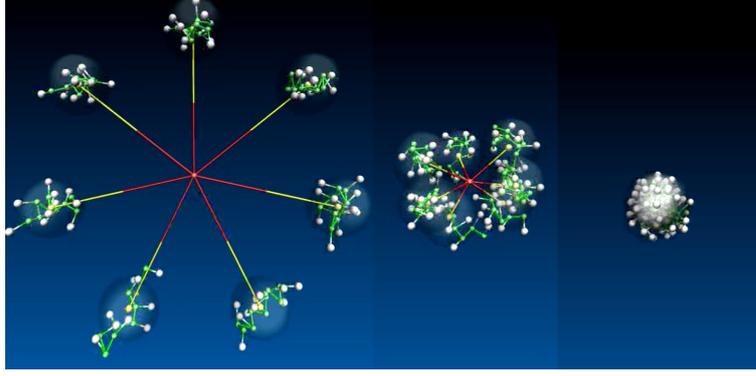


Figure 6.11.: Three snapshots of the cluster-forming process. (Left) Initialisation of the process:  $n_{\text{occ}}$  dendrimers are initialised equally spaced in a circumference of radius long enough so that there are no overlaps.  $n_{\text{occ}}$  springs are attached to the COM of each dendrimer and the COM of the whole system. (Centre) The rest length of the springs is slowly decreases in order to bring the dendrimers closer to a cluster configuration. (Right) Final cluster configuration.

resolved simulation we first create  $n_{\text{occ}}$  isolated chains, as described earlier in the text (subsection 6.4.1). We then position the AB chains equally separated on a circumference of radius large enough so that no overlap between the monomers occurs (see right panel of figure 6.11). We attach  $n_{\text{occ}}$  springs to each of the COMs of the AB chains ( $\mathbf{R}_{\text{COM}}^I$  for  $I = 1, \dots, n_{\text{occ}}$ ) and to the centre of the circumference ( $\mathbf{r}_0$ ). These springs exert a force on the COMs, *i.e.* on all the monomers of the AB chains, given by

$$\mathbf{F}_{\text{spring}}^I(\mathbf{R}_{\text{COM}}^I) = \alpha (|\mathbf{R}_{\text{COM}}^I - \mathbf{r}_0| - r_l) \frac{\mathbf{R}_{\text{COM}}^I - \mathbf{r}_0}{|\mathbf{R}_{\text{COM}}^I - \mathbf{r}_0|}, \quad (6.1)$$

where  $\alpha$  is the strength of the spring and  $r_l$  its rest length. We use  $\alpha = 1.0\epsilon_{\text{unit}}/\sigma_{\text{unit}}^2$  and slowly decrease  $r_l$  from the radius of this circumference to  $r_l = 0$ . A video of how this process is realised is shown in <https://www.youtube.com/watch?v=miWwT7erbv0>, and three snapshots of taken from the video are shown in figure 6.11.

Repeating this process, we generate 24 clusters. We define the size of the cluster as the distance from the centre of mass of the whole cluster to the furthest monomer. We pick the four smallest clusters, position them in an fcc cubic unit cell of size  $a$  and discard the rest of the clusters. We check for overlaps, if these occur we perform a random rotation of each of the four clusters and re-check again. If after 1500 rotations we still have overlaps, we re-start the process by generating 24 new clusters. If on the contrary, we don't get an overlap, we repeat this unit cell three times in each direction and generate the fcc crystal.

In figure 6.12 we present the radial distribution functions obtained in all the simulations of the cluster crystals. We observe that in the coarse grained simulation for  $\rho = 0.36$  and  $\rho = 0.44$  the crystal is melting while for  $\rho = 0.51$  it remains stable. We attribute this behaviour to the fact that a bcc lattice might be more stable for  $\rho = 0.36$  and  $\rho = 0.44$ .

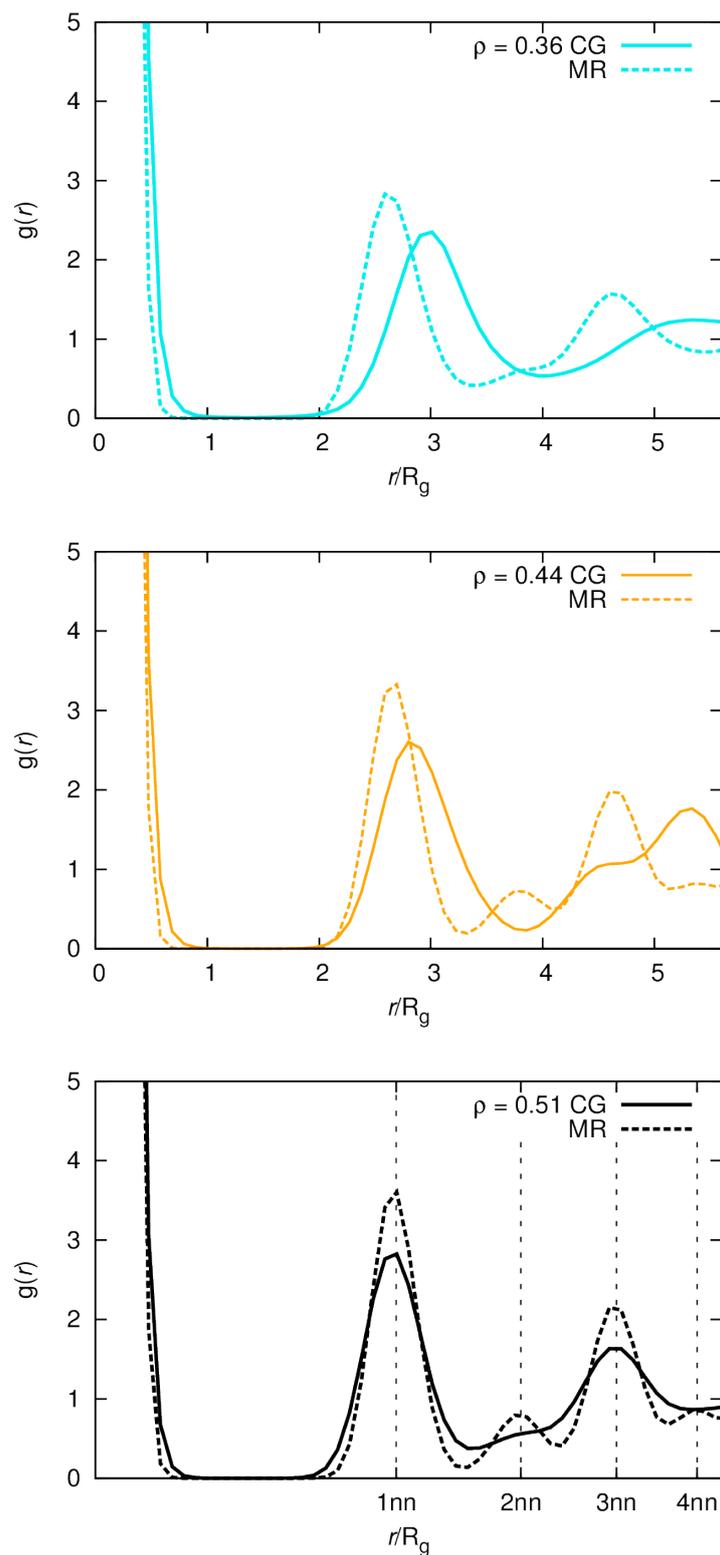


Figure 6.12.: Radial distribution function of AB chains AC-8 of length  $L = 20$  in the CG (solid line) and MR (dashed line) representation at densities  $\rho = 0.36$  (up),  $\rho = 0.44$  (middle) and  $\rho = 0.51$  (bottom) in units of  $R_g$  of an isolated chain (see table 6.6). The systems are initialised in an fcc lattice with  $a = 18.0\sigma_{\text{unit}}$ . The vertical dashed lines in the lowest panel show the positions of the nearest neighbours of an fcc lattice with the corresponding  $a$ -value.

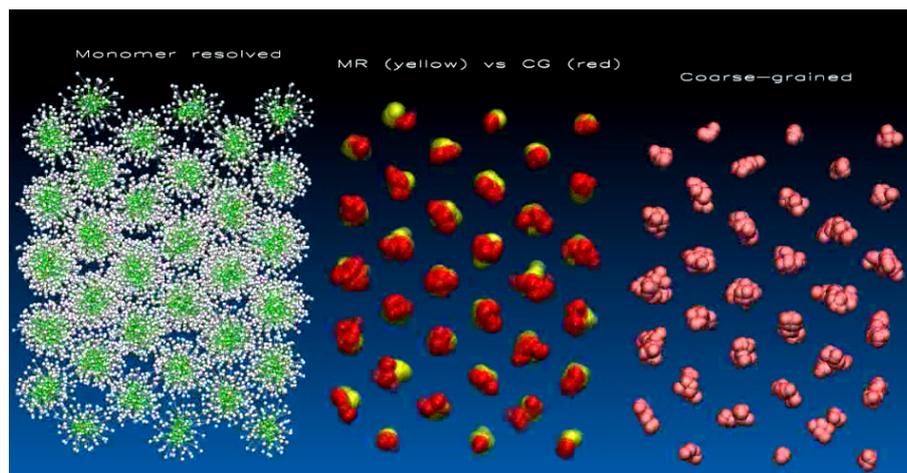


Figure 6.13.: Snapshot of the video presented in <https://www.youtube.com/watch?v=xGoynyf0gBU>. See the text for a description.

In the monomer resolved simulations, although for  $\rho = 0.36$  and  $0.44$  the peaks of  $g(r)$  are slightly smeared, for all  $\rho$ -values, the positions of the maxima coincide with the positions of the maxima in a fcc lattice, *i.e.* the peaks are not moving from the initial configuration: **the fcc lattice is stable in the MR level**.

In <https://www.youtube.com/watch?v=xGoynyf0gBU> we present a video with the data from the MR and CG simulation at  $\rho = 0.51$ , figure 6.13 shows a snapshot of this video. The frame is divided into three panels: the left one shows the monomer resolved simulation with the solvophobic monomers (A) in green and the solvophilic (B) ones in white. The right panel shows the data from the CG simulation, where the CG particles are shown in solid red. In the middle panel MR and CG data are merged: CG particles are shown in transparent red and the COMs of the AB chains in the MR simulation are shown in transparent yellow. The first and most important conclusion that one can draw from this video is that we have successfully designed a polymeric structure which forms **stable** cluster crystals. Secondly, we can observe how the behaviour of the CG particles and that of the MR chains, represented by their COMs, differs: CG particles are more mobile, the clusters are more spread and some particles follow orbits around the cluster very big in comparison to what the MR COMs do. This is again due to the fact that in the CG simulation we are using the effective potential calculated in the zero-density, which does not account for many-body effects.

With these results we terminate our investigations on amphiphilic chains. The model which we have designed is stable in an fcc cluster crystal configuration at a density sufficiently low so that a MR simulation is affordable. We have observed spontaneous cluster formation in the fluid phase, where starting from a random configuration clusters of up to 8 chains are formed, and we have seen that a system initialised in an fcc lattice with  $a = 18.0\sigma_{\text{unit}}$  remains stable during a long simulation time. A study of spontaneous crystallization is still missing as well as free energy calculations of the optimum lattice spacing.

# 7. Effective interactions of polymeric macromolecules computed at finite densities

## 7.1. Introduction

In this chapter of the thesis we focus on the computation of effective interactions of polymeric macromolecules. Effective interactions are commonly computed for two reasons: (i) To build a coarse grained model of a system of polymers whose simulations will spare a lot of computing time in comparison to a monomer resolved simulation and (ii) to gain insight into the interaction between the macromolecules and predict the phases that a bulk system will form. To date the most commonly used methods calculate the effective pair interactions by averaging out the internal degrees of freedom of two isolated and interacting macromolecules [12, 13], *i.e.* in the zero-density limit. However, it is to expect that many-body effects, which are not contained in the zero-density effective pair interaction, will make the behaviour of the macromolecule in a bulk system be different as in the super-dilute regime. Consequently, correlation functions from CG simulations using the effective potential calculated at  $\rho = 0$  will not reproduce the correlation functions of the MR simulation and the predictions based on this approximation, as for example the prediction of clustering behaviour based on the Fourier transform of the effective potential, might turn out to be erroneous as they are not fulfilled in the MR simulations. An example of a system where many-body effects play a crucial role is flexible polymer rings [14]: effective interactions calculated at  $\rho = 0$  have negative Fourier components, *i.e.* predict clustering at high densities which, however, does not occur. Due to many-body effects, related to the deformation of the rings, the behaviour of the MR system deviates from the behaviour of the CG system.

To circumvent the problems of zero-density effective pair potentials, one can go one step further in the computation of the effective interaction, and try to calculate it at finite densities [93, 94]. Such an approach implies of course going one (or several) steps further in the complexity of the procedure: to start with, one needs to perform a MR simulation of the system at the given *finite* density. An ensemble of at least  $N \sim 10^2$  will be needed to produce the proper statistics, therefore the computational effort is orders of magnitude higher compared to a zero-density simulation where only  $N = 2$  macromolecules are simulated. Bolhuis *et al.* proposed in [93] a method to calculate the effective interaction potentials at finite densities. The authors performed monomer resolved simulations of linear polymer chains and computed the radial distribution function,  $g(r)$ , of the centre of mass of the

polymers. They computed the effective interaction by inverting the  $g(r)$  using integral equation techniques from the theory of simple fluids. This method is designed to reproduce the pair correlation functions; however it involves integral equation techniques which are approximate methods, due to the use of a closure relation.

Here we suggest to use an alternative method, the so-called multi-scale coarse grained method developed by Voth *et al.* [15, 94] explained in section 3.2. This method is based on a force matching algorithm, which matches the parameters of a functional form of the coarse-grained force to reproduce the COM acceleration generated by the monomeric forces. It allows for the computation of effective interactions at finite densities, and offers an absolute flexibility in the functional form of the potential: the interaction can be expressed as a sum of two-body and higher order terms, which can be functions of the inter-particle distance ( $r$ ), the angle enclosed by three particles, dihedrals, etc ...

In our calculations we build the functional form as an  $r$ -dependent two-body term, and explore its density dependence by calculating the effective interaction with data obtained from MD simulations at various  $\rho$ -values. The only requirement of this set-up is that we need a MR simulation where the range of distances between two polymers spanning from  $r = 0$  to  $r = R_{\max}$  is visited,  $R_{\max}$  being the distance beyond which one can assume that two polymers do not interact, *i.e.* we need a state which produces a  $g(r)$  which does not vanish in the range  $0 \leq r \leq R_{\max}$ . Therefore, the effective interaction can only be computed in fluid states.

This chapter is organised as follows: in section 7.2 we present the effective interactions computed at finite densities of non-clustering polymers in subsection 7.2.1 and of clustering amphiphilic dendrimers in subsection 7.2.2. In section 7.3 we present the effective interactions of an isolated dendrimer with a cluster of dendrimers to gain insight into why CG simulations with finite density effective pair potentials do not always reproduce pair correlations functions of MR simulations if the system forms clusters.

## 7.2. Effective interactions at finite densities

In order to compute the effective interactions at finite density in the different systems listed in the following subsections, we first perform MD simulations of an ensemble of  $N \sim 10^2$  to  $10^3$  polymers using the LAMMPS simulation package [95]. We perform  $NVT$  simulations with the Nose-Hoover thermostat (subsection 3.1.1). We will specify in each system the values of the parameters in the MD simulations, such as the time step  $\Delta t$  (see equation (3.2)) and the artificial mass of the thermostat  $Q$  (see equation (3.8)). We will use for  $Q$  the notation of LAMMPS which works with the parameter  $T_{\text{damp}}$ , defined as:  $Q = L\beta T_{\text{damp}}^2$ , where  $L$  is the parameter appearing in equation (3.8), which takes the value  $L = 3N + 1$ .

To create the initial configuration we used the same procedure as described in subsection 6.4.1, where isolated polymers are generated and then fitted one by one into the simulation box, by trying different orientations and positions in the box, until no overlap with the other polymers is found. In some of the cases we directly fitted the polymers into a box of the desired size, while in other cases we first put them in a very large box, generating thereby a very dilute configuration without problems due to overlaps of the monomers. We then used the *fix deform* LAMMPS command, which starts with an initial box size and rescales at every time step the positions of all the particles to arrive finally to the target box size. We will specify in each system which method we used.

We first equilibrate the system, *i.e.* we wait until the total energy is constant, and then we start to produce the necessary data for the coarse graining procedure: we save a sufficient amount of configurations, where the positions and the forces acting on the centres of mass (COMs) of the polymers are recorded, each of them taken every 500 to 5000 time steps. We consider that we have saved a sufficient amount of configurations when, by adding more configurations in the coarse graining procedure, the resulting effective potential remains unchanged. To compute the two-body effective interactions we use the Multi-Scale Coarse Graining procedure described in section 3.2. The parameters of the MSCG such as  $R_{\text{max}}$ ,  $N_{\text{D}}$  (see equation (3.42)) and the number of configurations  $n_t$  will be specified in each system.

### 7.2.1. Non-clustering systems

We start our study with non-clustering system, *i.e.* linear polymer chains and athermal dendrimers.

#### Linear polymer chains

Linear polymer chains are made up of  $N_{\text{m}}$  connected monomers of equal mass. We model the excluded volume interaction between the monomers with a truncated and shifted Lennard

$T$	$\rho$	$N$	$L_{\text{box}}$	$R_g$	$\rho^*$	$\Delta t$	$T_{\text{damp}}$	$N_D$	$R_{\text{max}}$
0.2	0.005	1000	58.48	2.84	0.115	0.0001	0.45	10	20
0.2	0.075	1000	51.09	2.79	0.163	0.0001	0.45	15	20
0.2	0.01	1000	46.41	2.74	0.206	0.00005	0.09	10	20
0.2	0.05	1000	27.14	1.96	0.377	0.001	0.1	10	20
0.1	0.08	1000	23.20	1.68	0.380	0.001	0.09	10	23

Table 7.1.: System parameters (temperature  $T$ , density  $\rho$ , number of polymers  $N$ , box size  $L_{\text{box}}$ , radius of gyration  $R_g$  and density  $\rho^*$  in units of  $R_g$ ), parameters of the MD simulation (time step  $\Delta t$  and damping factor  $T_{\text{damp}}$ ) and parameters of the MSCG (dimension of the basis  $N_D$ , basis choice and cut-off of the effective interaction  $R_{\text{max}}$ ) of the different simulations performed with linear chains of length  $N_m = 20$ .

Jones potential:

$$\Phi_{\text{LJ}}(r) = \epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right]$$

$$\Phi_{\text{LJcut+shift}}(r) = \begin{cases} \Phi_{\text{LJ}}(r) - \Phi_{\text{LJ}}(r_c) & \text{if } r < r_c = 1.2\sigma \\ 0 & \text{otherwise,} \end{cases} \quad (7.1)$$

which is a purely repulsive potential and the bonding interaction with a FENE potential (see equation (2.30)). We use  $\epsilon$  and  $\sigma$  in equation (7.1) as the energy and length units and the mass of the monomers as the mass unit. Other quantities are expressed in terms of those units, except  $\rho^*$ , the density expressed in units of the radius of gyration  $R_g$ .

In our simulations we used chains of a length of  $N_m = 20$ . The parameters of the FENE potential are given by  $l = 0$ ,  $R = 4.0$  and  $\kappa = 2.5$ .

The different state points that we study are listed in table 7.1, where the temperature, density, number of polymers, box size, the measured radius of gyration and the density expressed in units of the radius of gyration are shown. We also show the parameters of the MD simulation,  $\Delta t$  and  $T_{\text{damp}}$ , and the parameters of the MSCG:  $N_D$ , basis choice and  $R_{\text{max}}$ .

To produce the initial configuration for the lowest densities ( $\rho = 0.005, 0.0075$  and  $0.01$ ) we directly fitted the isolated polymer chains into a box of the corresponding size for that density. For the highest densities ( $\rho = 0.05$  and  $\rho = 0.08$ ) we took a configuration with  $\rho = 0.01$  and we reduced the box size to the desired length within  $5 \times 10^3$  time steps.

The initial configurations generated by the procedure described in subsection 6.4.1 were very far from equilibrium and therefore we were forced to use a very small time step for the MD simulation (see table 7.1). Those configurations generated using the *fix deform* command were much closer to equilibrium, allowing us to use larger time steps. We saved the positions and forces acting on the COMs of the polymers every 500 time steps and used  $n_t = 10000$  configurations for the coarse-graining procedure.

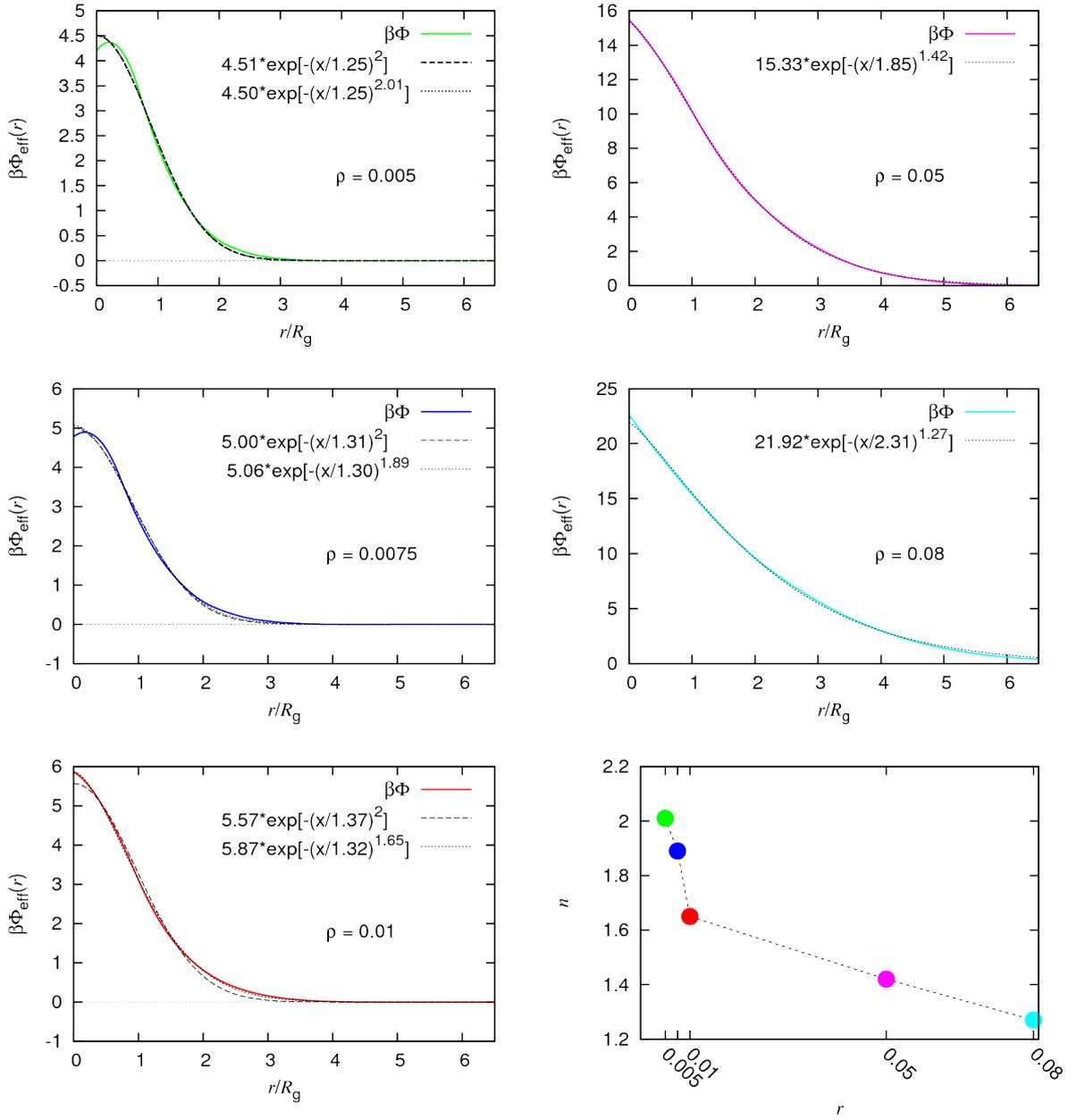


Figure 7.1.: Effective interactions  $\Phi_{\text{eff}}(r)$  as functions of  $r$  of linear chains of length  $N_m = 20$  calculated at different densities as labelled. The dashed lines show fits to a GEM-2 function and the dotted lines show fits to a GEM- $n$  function, where the index  $n$  is allowed to vary. In the lower right panel we show the dependence of  $n$  on  $\rho$ .

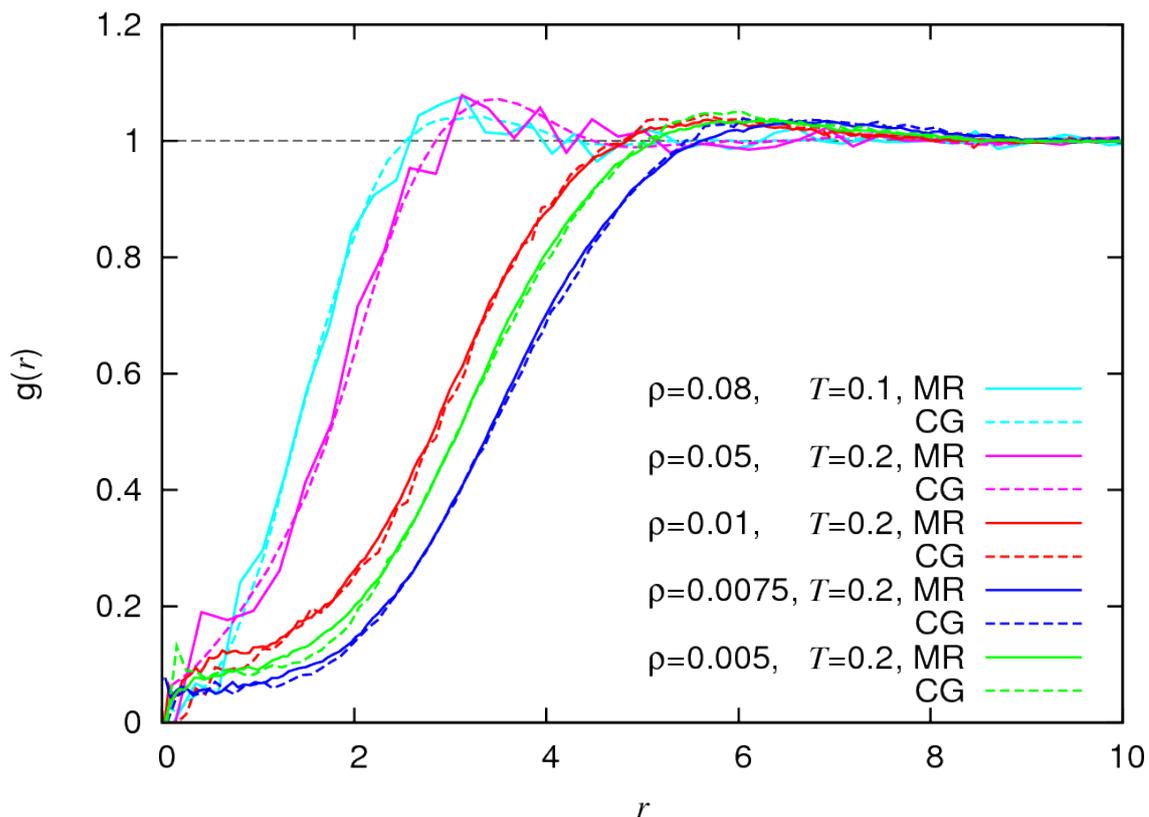


Figure 7.2.: Comparison of the radial distribution function  $g(r)$  as a function of  $r$  of the COMs of linear chains in a MR simulation and of the CG particles in a simulation using the effective potentials presented in figure 7.1.  $T$  and  $\rho$  as labelled.

The resulting effective potentials,  $\Phi_{\text{eff}}(r)$ , are shown in figure 7.1 for the different state points. Additionally, two fits of  $\beta\Phi_{\text{eff}}(r)$  are shown: the first one to a GEM-2 potential, motivated by the fact that the effective potential calculated at zero-density can be very well approximated by a Gaussian function [93], and the second one to a GEM- $n$  potential, where the index  $n$  is allowed to vary. We show this index  $n$  as a function of  $\rho$  in the lower right panel of figure 7.1. For densities up to  $\rho = 0.01$  we obtain effective potentials that are essentially a GEM-2 curve ( $n = 2$  is the threshold beyond which cluster formation is predicted). As the density increases,  $n$  decreases and the potential becomes steeper at the origin. This tendency is in good agreement with the results presented by Bolhuis *et al.* [93].

We compare the radial distribution function,  $g(r)$ , computed during the MD simulation ( $g(r)$  of the COMs of the polymers) and in the CG simulation (with the corresponding effective potential), both of them performed at the same  $\rho$ - and  $T$ -values. Results are shown in figure 7.2.

We observe that, apart from the noise due to the lack of good statistics, the agreement between the CG and MR curves is perfect.

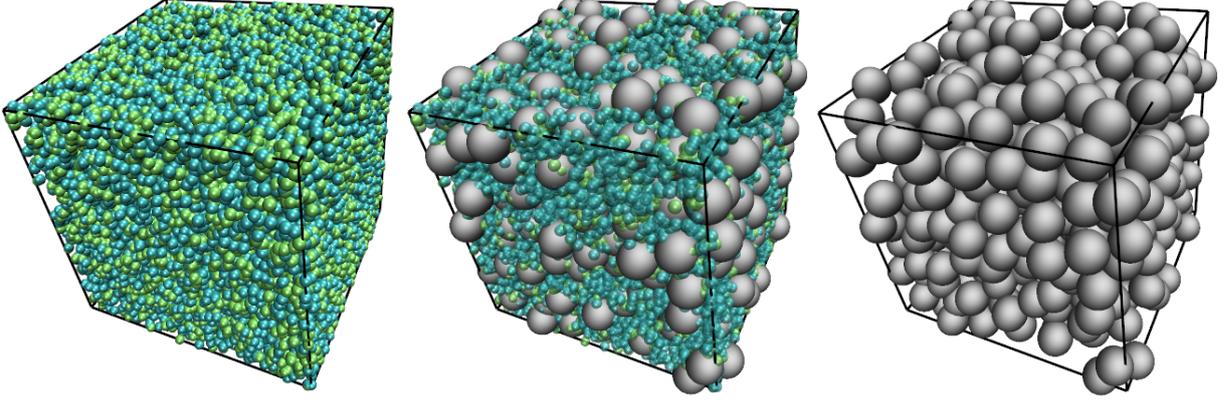


Figure 7.3.: Snapshot of a simulation of 500 athermal dendrimers at a density  $\rho = 2\rho_{\text{overlap}}$ . In the left panel we plot the monomers of size  $\sigma = 1$ , in blue the shell monomers and in green the core ones. In the right panel we plot spheres of size  $R_g = 3.41$  positioned on the COMs of the dendrimers. In the middle panel we combine both plots.

### Athermal dendrimers

The second system that we consider in our study are athermal dendrimers. We use  $G = 4$  dendrimers from reference [96], modelled with a truncated and shifted LJ potential (equation (7.1)) for the excluded volume interaction and a FENE potential (equation (2.30)) for the bonds (with  $l = 0$ ,  $R = 10$ ,  $\kappa = 0.5$ ). In reference [96] Götze *et al.* calculated the effective interactions at zero-density of athermal dendrimers of variable flexibility (tuned by  $\kappa$ ) and generation number  $G$ . The authors found a generic Gaussian shape of the effective potential of athermal dendrimers, irrespective of the flexibility and generation number. In a later contribution [97] the same authors compared the pair correlation functions calculated in a MR simulation and in a CG simulation using the zero-density effective potential for  $G = 4$  athermal dendrimers and found that there was a very good agreement up to the overlap density  $\rho_{\text{overlap}}$ , defined as:

$$\frac{4\pi}{3}\rho_{\text{overlap}}(1.5R_g)^3 = 1. \quad (7.2)$$

In our work we calculate the effective interaction at  $\rho = 2\rho_{\text{overlap}}$  and temperature  $T = 1.0$ , using a box size of  $L_{\text{box}} = 51.8$  and  $N = 500$  (we measure  $R_g = 3.41$ ). We generate the initial configuration in a box of  $L_{\text{box}} = 300$ ; using LAMMPS we reduce the box size until the desired value. We show a snapshot of the system in figure 7.3, where the reader can see that the system is very dense. Within a LAMMPS MD simulation we generate equilibrium trajectories using a time step  $\Delta t = 0.0005$  and a damping factor  $T_{\text{damp}} = 0.09$ . We save the positions and forces acting on the COMs every 1000 time steps. In the left panel of figure 7.4 we show the obtained effective interaction  $\Phi_{\text{eff}}(r)$ , computed with  $n_t = 12000$  configurations,  $N_D = 8$  and  $R_{\text{max}} = 20$ , and a fit to a GEM- $n$  function. At this density we find a value

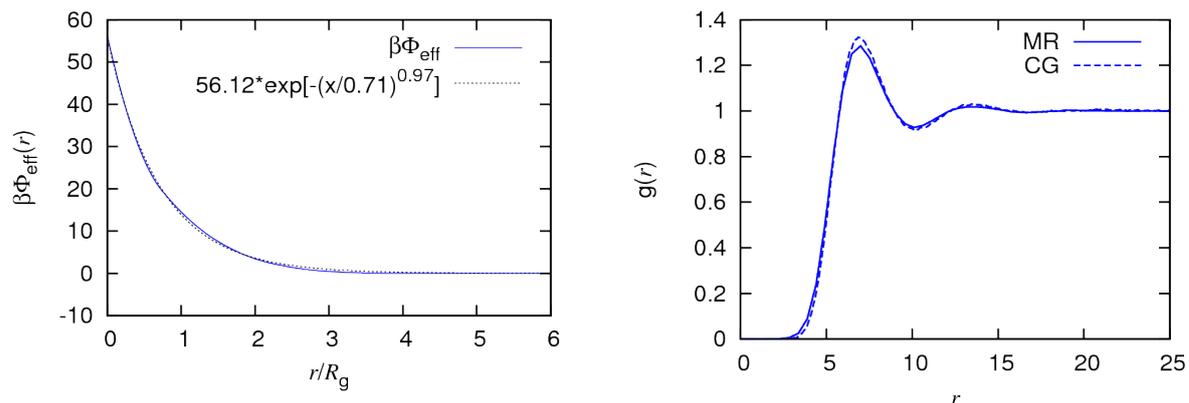


Figure 7.4.: Left: Effective interaction  $\Phi_{\text{eff}}(r)$  of athermal dendrimers of  $G = 4$  calculated at  $\rho = 2\rho_{\text{overlap}}$  and  $T = 1.0$ . Right: Comparison of  $g(r)$  measured in the MR simulation and in a CG simulation using the effective potential in the left panel, at the same  $\rho$  and  $T$  conditions.

for the exponent of  $n = 0.97$ , and therefore, since at zero density  $n = 2$  [96], we observe the same tendency as in the linear chains: as the density increases the potential becomes steeper at the origin and  $n$  attains values substantially smaller than  $n = 2$ . In the right panel of figure 7.4 we compare the  $g(r)$  measured in the MR simulation and in a CG simulation using the computed effective potential. The agreement is very good; thus we can conclude that, at least for this system, the range of densities for which a CG description produces reliable results can extend beyond the overlap density if effective potentials calculated at finite densities are used.

### 7.2.2. Clustering systems

Proven the success of effective interactions calculated at finite density in systems which do not show clustering, we next consider clustering systems, namely  $D2$  [38] and  $r6$  [28] amphiphilic dendrimers (see subsection 2.5.3).

#### D2 dendrimers

$D2$  amphiphilic dendrimers were first introduced by Mladek *et al.* [38] in 2008 as an improvement over athermal dendrimers towards clustering: by introducing the solvophilic shell and solvophobic core monomers the effective interaction between the COMs of the dendrimers attains a local minimum at  $r = 0$  which leads to oscillations in the Fourier transform of the effective potential. In 2010 Lenz *et al.* [27] performed extensive monomer resolved simulations of an ensemble of second generation  $D2$  dendrimers in the fluid state. In their investigations the authors found a qualitative agreement between the pair correlation functions in the MR

$N$	$L_{\text{box}}$	$\rho$	$R_g$	$\rho^*$
500	36	0.011	3.295	0.38
2000	51.5	0.015	3.303	0.52
2000	48.2	0.018	3.316	0.65
2000	44.9	0.022	3.334	0.82
2000	41.6	0.028	3.359	1.05

Table 7.2.: Number of dendrimers  $N$ , box size  $L_{\text{box}}$ , density  $\rho$ , radius of gyration  $R_g$  and the density  $\rho^*$  in units of  $R_g$  of the considered systems of second generation D2 dendrimers.

and CG simulations (using the zero-density effective potential) at low and intermediate densities: the zero-density effective potential predicted the formation of stable, finite aggregates of dendrimers in the fluid phase which were indeed found in their MR simulations. However, the discrepancies between the MR and the CG simulations became bigger and bigger, as the predicted freezing density  $\rho_{\text{freeze}}^0$  was approached. At higher  $\rho$ -values the CG system froze into a cluster crystal while the MR system remained in a fluid configuration.

In our work we consider a D2 system at five different densities and  $T = 1.0$ , where the units of energy and length are given by  $1.4\epsilon_{\text{CC}}$  and  $\sigma_{\text{CC}}$ , *i.e.* the parameters in the Morse interaction of two core monomers (table 2.1). The mass of the monomers is used as mass unit and the rest of the quantities are expressed in reduced units. The number of dendrimers  $N$ , the box size  $L_{\text{box}}$ , the density  $\rho$ , the measured radius of gyration  $R_g$  and the density in units of  $R_g$  are shown in table 7.2. We explore a density range comparable to the one presented in reference [27]. We perform MR simulations with LAMMPS using a time step  $\Delta t = 0.0005$  and a damping factor  $T_{\text{damp}} = 0.09$ . The initial configurations are produced by compressing with the *fix deform* command a very dilute configuration ( $\rho = 2.5 \times 10^{-4}$ ), which in turn was produced by the procedure described in subsection 6.4.1. We save positions and forces acting on the COMs of the polymers every 1000 steps and produce  $n_t = 25000$  configurations for the coarse graining procedure. We compute the effective pair interaction by using the MSCG with a basis of dimension  $N_D = 13$  and  $R_{\text{max}} = 20.0$ .

The resulting effective interactions are shown in figure 7.5. We observe that there is a strong dependency of  $\Phi_{\text{eff}}(r)$  on the density: with increasing  $\rho$  the potential becomes longer ranged and the local attraction at short distances (characterised by the local minimum at  $r = 0$ ) disappears.

In table 7.3 we present for each different density the freezing density  $\rho_{\text{freeze}}$  at  $T = 1.0$  (estimated *via* equation (2.23)) computed with the effective potentials presented in figure 7.5. We observe that for the lowest densities ( $\rho^* = 0.38$  and  $0.52$ ),  $\rho_{\text{freeze}}$  decreases with respect to  $\rho_{\text{freeze}}^0$ . As  $\rho^*$  increases above 0.6,  $\rho_{\text{freeze}}$  takes values close to  $\rho^*$ . In figure 7.6 we show the cluster size distribution obtained in the MR simulations computed by following only step one of the cluster analysis algorithm detailed in appendix B using  $r_c = 3.0$ , correspondingly  $\sim 0.9R_g$ . This cluster analysis is very simplified in comparison to the one presented in

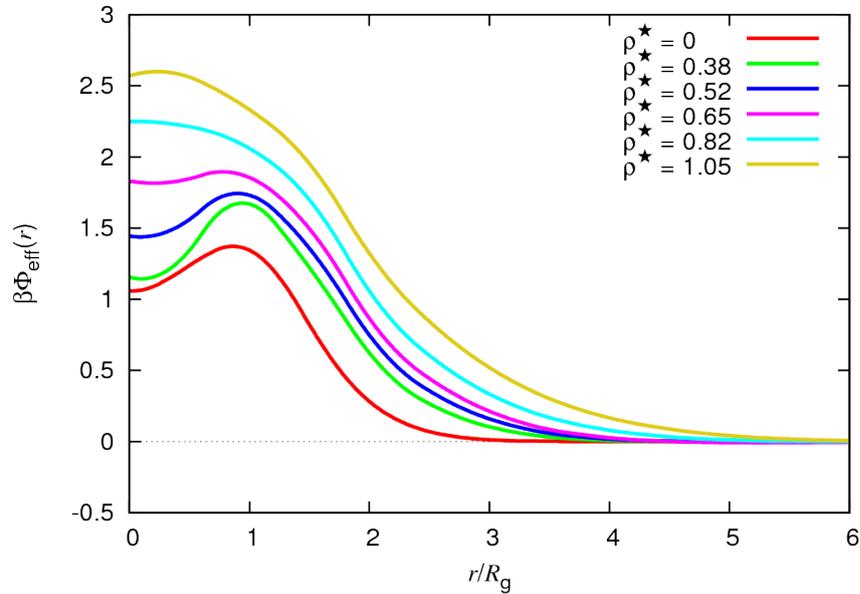


Figure 7.5.: Effective interactions  $\beta\Phi_{\text{eff}}(r)$  of a D2 second generation dendrimer system calculated at different densities, as labeled.

$\rho^*$	$\rho_{\text{freeze}}$
0	0.78
0.38	0.57
0.52	0.60
0.65	0.67
0.82	0.81
1.05	1.07

Table 7.3.: Predicted freezing density  $\rho_{\text{freeze}}$  at  $T = 1.0$  of a system of D2 second generation dendrimers using the effective interactions computed at the corresponding  $\rho^*$  in a system of second generation D2 dendrimers

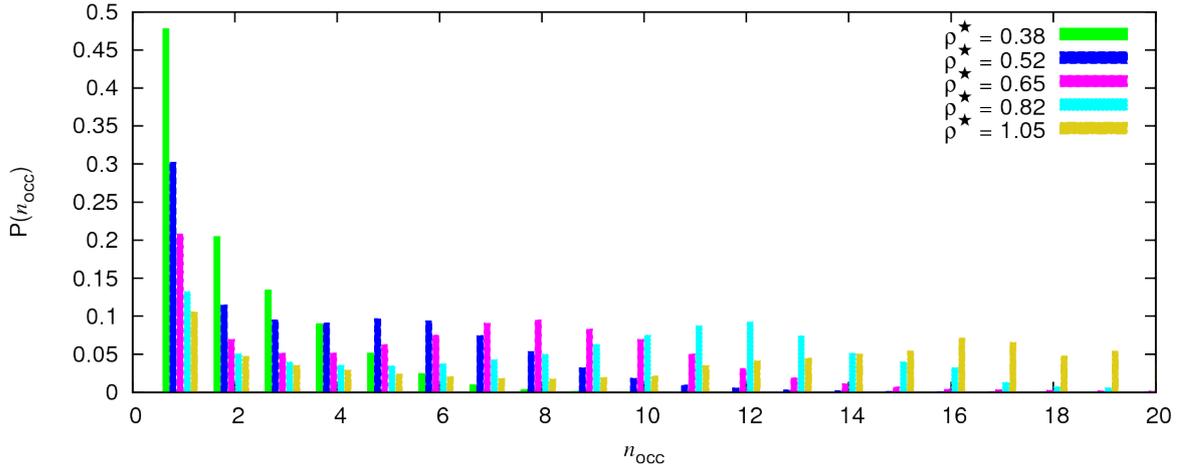


Figure 7.6.: Cluster size distribution  $P(n_{\text{occ}})$  as a function of the occupation number  $n_{\text{occ}}$  in the different MR simulations of a fluid state at different  $\rho^*$  values of a  $G = 2$  D2 dendrimer system.

reference [27], since we do not check if the dendrimers in small clusters ( $n_{\text{occ}} \leq 2$ ) are located within almost a linear chain of dendrimers between different clusters (and should then be considered as isolated dendrimers). Our simplified analysis might therefore underestimate the number of isolated dendrimers; however the ensuing results can give us some preliminary insight into the system. We observe that for  $\rho^* = 0.38$  approximately  $\sim 47\%$  of the dendrimers are isolated and that  $P(n_{\text{occ}})$  monotonically decreases with  $n_{\text{occ}}$ . For  $\rho^* = 0.52$  the number of isolated dendrimers decreases down to 30% and we find a local maximum of  $P(n_{\text{occ}})$  at  $n_{\text{occ}} = 6$ . As  $\rho^*$  further increases the number of isolated dendrimers goes down while the local maximum in  $P(n_{\text{occ}})$  continues to grow and shifts towards higher values of  $n_{\text{occ}}$ . These results in combination with the data compiled in table 7.3 let us conclude that the system might be trapped in a metastable equilibrium when  $\rho^* \gtrsim 0.65$  and that a crystal might have a lower free energy than the fluid.

In figure 7.7 we compare  $g(r)$  computed both in the MR simulation (RDF-MR) and in CG simulations, either using the zero-density potential (RDF-CG<sub>0</sub>) or with the corresponding effective interaction calculated at the same density (RDF-CG <sub>$\rho^*$</sub> ). In all cases we compare  $g(r)$  obtained in simulations carried out at the same  $\rho$  and  $T$  values. On the one hand, we observe that RDF-CG<sub>0</sub>, as already anticipated by Lenz *et al.* [27], shows a poor qualitative agreement with RDF-MR at low densities ( $\rho < 0.7$ ) and resoundingly fails as  $\rho_{\text{freeze}}^0$  is approached. On the other hand, there is only a good agreement between RDF-CG <sub>$\rho^*$</sub>  and RDF-MR at very low densities ( $\rho < 0.6$ ). For higher densities even though RDF-CG <sub>$\rho^*$</sub>  is closer to RDF-MR than RDF-CG<sub>0</sub>, this agreement is very far away from what we found in the non-clustering systems.

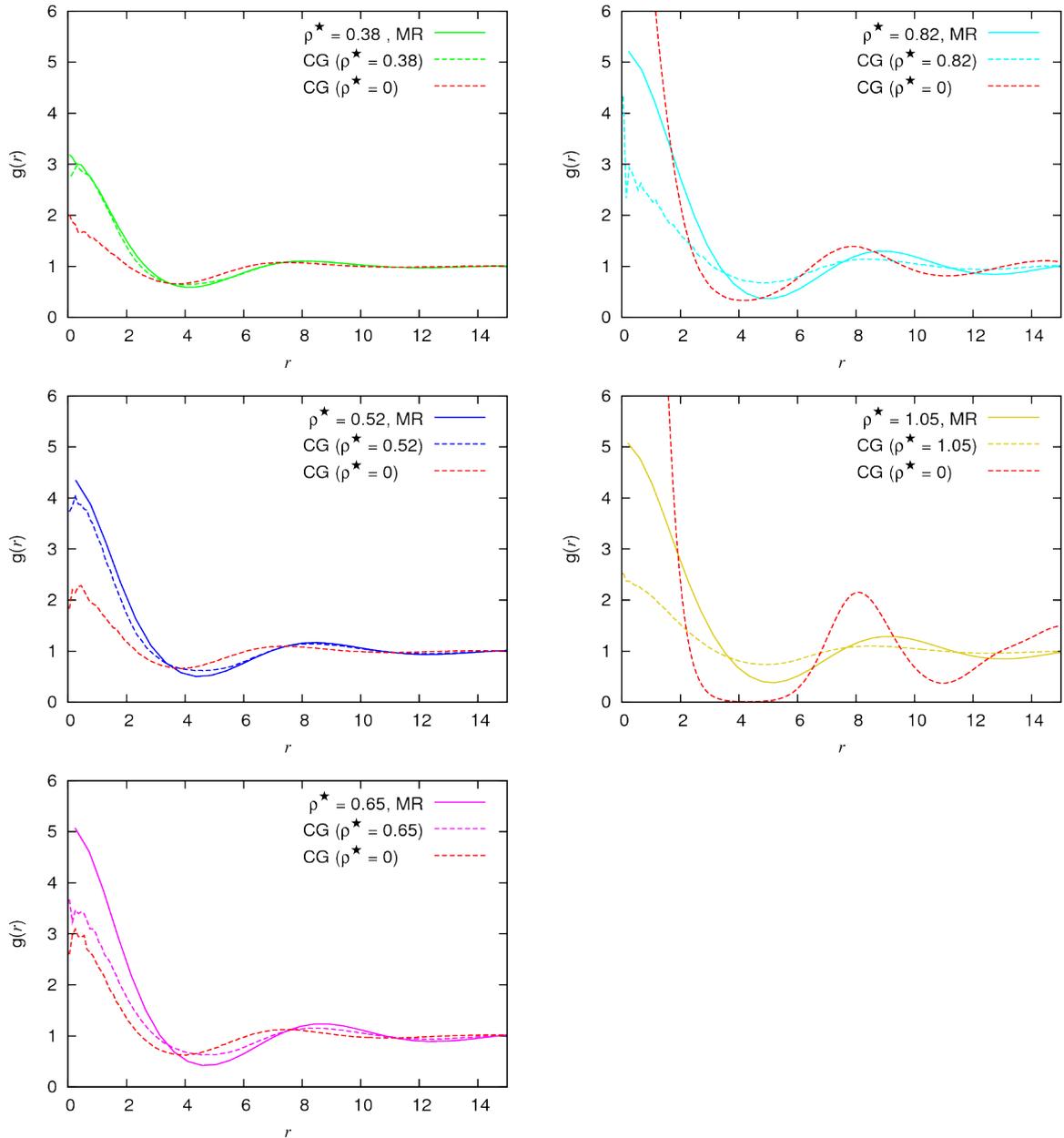


Figure 7.7.: Comparison of  $g(r)$  computed in a MR simulation (solid line) and a CG simulation (dashed line) of  $G = 2$  D2 dendrimers. The density at which the effective potential was computed is marked in brackets. All the panels correspond to simulations of a fluid state at different  $\rho^*$ , as labeled.

## r6 dendrimers

The next system which we study are second generation r6 amphiphilic dendrimers. This dendrimer model was introduced by Lenz *et al.* in reference [28], where they proved the existence and stability of cluster crystals in a MR level. The model was designed to make a MR simulation of a cluster crystal computationally less expensive: while for a D2 dendrimer system the predicted freezing density at  $T = 1.0$  is  $\rho_{\text{freeze}}^0 = 0.78$ , for a r6 system  $\rho_{\text{freeze}}^0 = 0.141$ .

$N$	$L_{\text{box}}$	$\rho$	$R_g$	$\rho^*$
2000	138.27	$0.7 \times 10^{-3}$	3.49	0.033
1500	100.00	$1.5 \times 10^{-3}$	3.47	0.065
1280	87.36	$1.9 \times 10^{-3}$	3.46	0.084
1520	87.36	$2.3 \times 10^{-3}$	3.45	0.099
1760	87.36	$2.6 \times 10^{-3}$	3.50	0.115

Table 7.4.: Number of dendrimers  $N$ , box size  $L_{\text{box}}$ , density  $\rho$ , radius of gyration  $R_g$  and the density  $\rho^*$  in units of  $R_g$  of the considered systems of second generation r6 dendrimers.

We consider six different systems, whose densities cover the range from  $0 \leq \rho^* \lesssim \rho_{\text{freeze}}^0$  at  $T = 1.0$ , where the units of energy and length are given by  $1.4\epsilon_{\text{CC}}$  and  $\sigma_{\text{CC}}$ , *i.e.* the parameters in the Morse interaction of two core monomers (table 2.2). The mass of the monomers is used as mass unit and the rest of the quantities are expressed in reduced units. Characteristic quantities of the systems are shown in table 7.4. The initial configurations for the simulations carried out at  $\rho^* = 0.033$  and  $0.065$  were generated by the procedure described in subsection 6.4.1. For the other densities, the initial configuration was produced from a cluster-fluid configuration at  $\rho^* = 0.131$  provided by D. A. Lenz [28]: for lower  $\rho^*$  values a sufficient amount of randomly chosen dendrimers were deleted. This “brutal” procedure to generate the starting configurations did not provide very good results, since in some cases up to  $10^8$  time steps were needed to recover equilibrium. We therefore recommend for future simulations to rather use the *fix deform* command and to compress a very dilute configuration.

MD trajectories were produced using LAMMPS with a time step of  $\delta t = 0.0005$ , positions and forces acting on the COMs of the dendrimers were recorded every 5000 time steps. The effective potentials for the different  $\rho^*$  values are shown in figure 7.8. We used  $R_{\text{max}} = 20$  and  $N_{\text{D}} = 20$ . In strong contrast to what we found for D2 dendrimers, we now observe a very mild dependency of  $\Phi_{\text{eff}}(r)$  on  $\rho^*$ .

In table 7.5 we present  $\rho_{\text{freeze}}$  computed from the Fourier transform of  $\Phi_{\text{eff}}(r)$  at different  $\rho^*$  values. We find again, as for the D2 case, that  $\rho_{\text{freeze}}$  values obtained at finite densities are lower than  $\rho_{\text{freeze}}^0$ . For  $\rho^* = 0.099$ ,  $\rho_{\text{freeze}}$  is very close to  $\rho^*$  and for  $\rho^* = 0.115$  it is even lower than  $\rho^*$ . In these two cases we might again have metastable equilibria; the crystal might

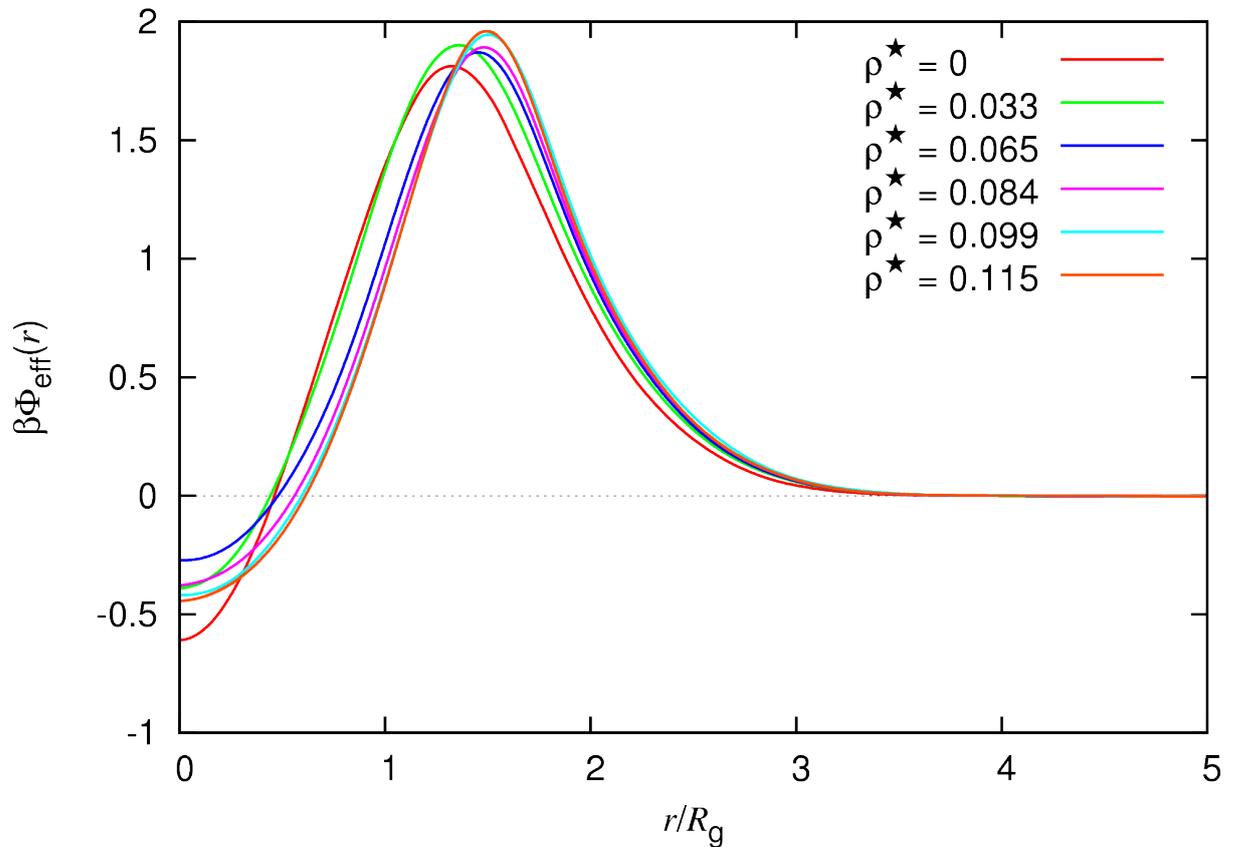


Figure 7.8.: Effective interactions  $\beta\Phi(r)$  of a r6 second generation dendrimers calculated at different densities, as labeled.

$\rho^*$	$\rho_{\text{freeze}}$
0	0.141
0.033	0.113
0.065	0.111
0.084	0.104
0.099	0.097
0.115	0.098

Table 7.5.: Predicted freezing density  $\rho_{\text{freeze}}$  at  $T = 1.0$  using the effective interaction computed at the corresponding  $\rho^*$ -value in a system of second generation r6 dendrimers.

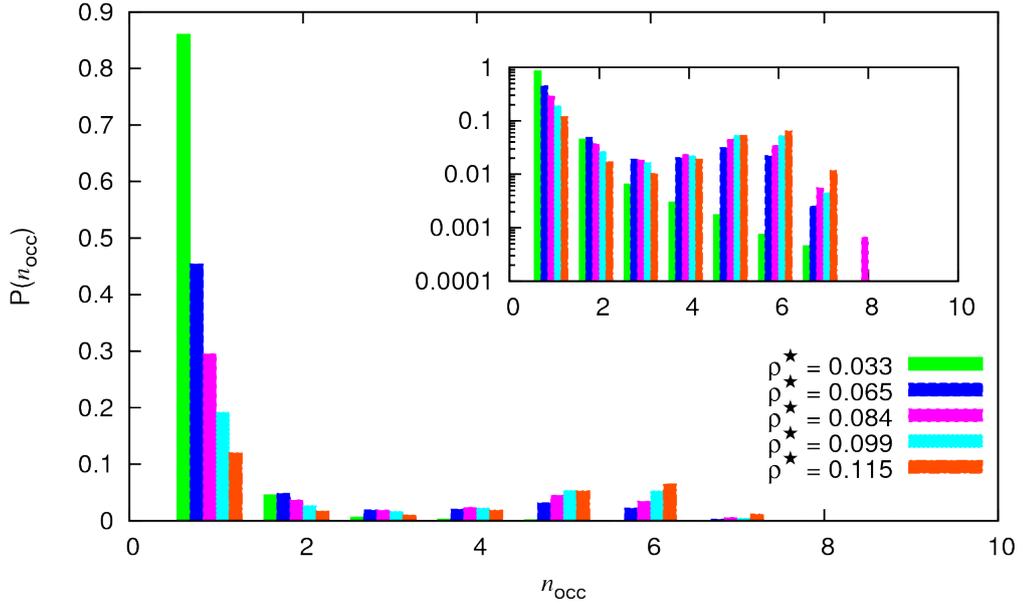


Figure 7.9.: Cluster size distribution  $P(n_{\text{occ}})$  as a function of the occupation number  $n_{\text{occ}}$  in the different MR simulations of a fluid state at different  $\rho^*$  values of a  $G = 2$  r6 dendrimer system.

have a lower free energy as the fluid state.

We plot in figure 7.9 the cluster size distribution in the MR resolved simulations at different  $\rho^*$ -values, using step 1 of the cluster analysis described in appendix B with  $r_C = 3.0$ , which is of the order of  $\sim 0.85R_g$ . For this systems we do not find such big clusters as for the D2 systems (see figure 7.6), the largest  $n_{\text{occ}}$ -value observed is 6. We find again that except for the lowest density, there is a local maxima at  $n_{\text{occ}} > 1$ , which becomes more pronounced and shifts towards larger  $n_{\text{occ}}$  values as  $\rho^*$  increases.

Finally, in figure 7.10 we compare the RDF computed in the MR simulation (RDF-MR) and in CG simulations, either using the zero-density potential (RDF-CG<sub>0</sub>) or the corresponding effective interaction calculated at the same density (RDF-CG <sub>$\rho^*$</sub> ). This time we find that not even at very low densities RDF-CG <sub>$\rho^*$</sub>  is able to reproduce RDF-MR. Again, as observed for high  $\rho^*$ -values of D2 dendrimers, RDF-CG <sub>$\rho^*$</sub>  is closer to RDF-MR than RDF-CG<sub>0</sub>, but there is no quantitative agreement between RDF-CG <sub>$\rho^*$</sub>  and RDF-MR.

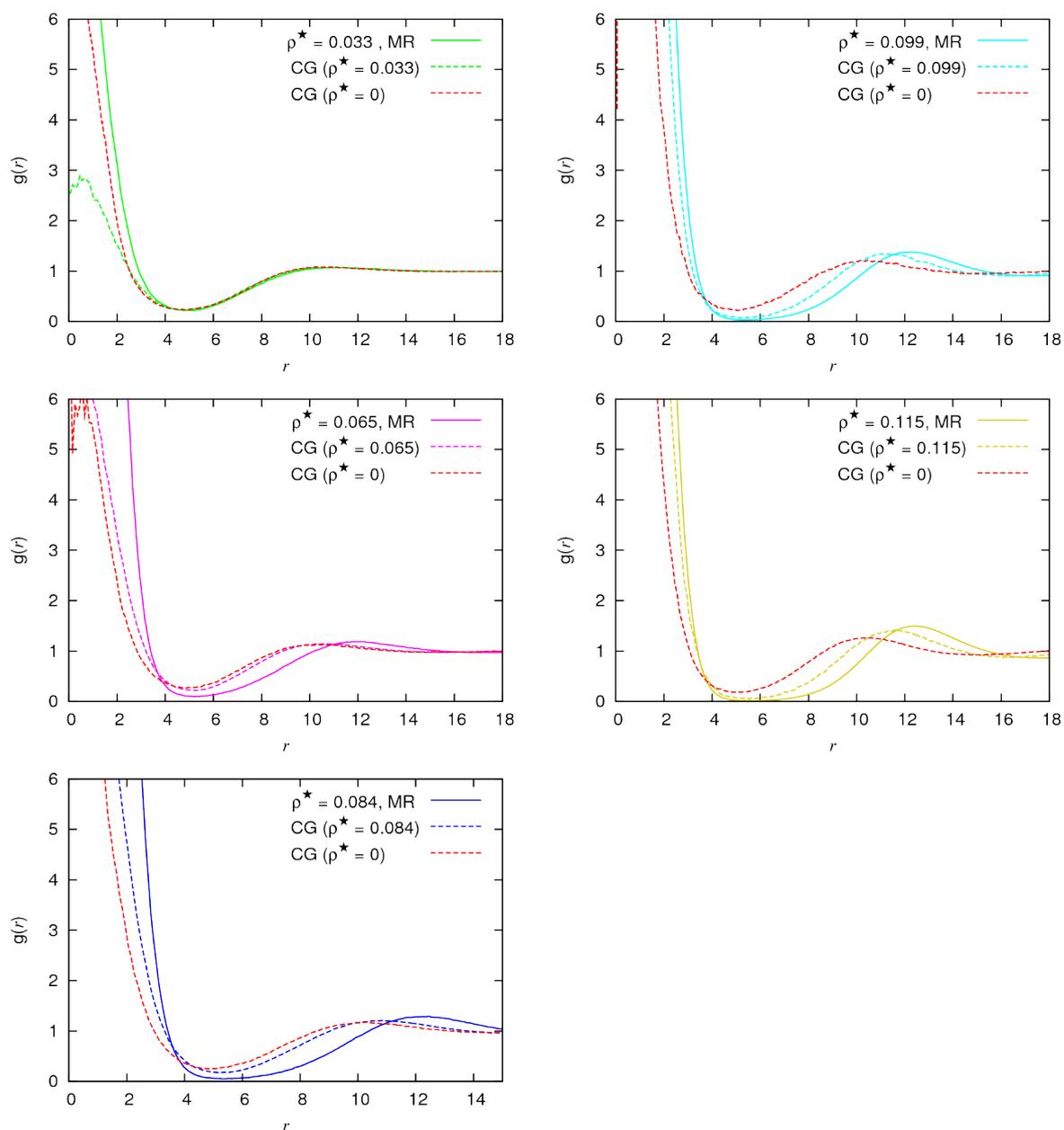


Figure 7.10.: Comparison of  $g(r)$  computed in a MR simulation (solid line) and a CG simulation (dashed line) of  $G = 2$  r6 dendrimers. The density at which the effective potential was computed is marked in brackets. All the panels correspond to simulations of a fluid state at different  $\rho^*$ , as labeled.

### 7.3. Effective interaction of a single dendrimer with a cluster

Analysing the results at hand, we can definitely conclude that for non-clustering systems effective interactions computed at finite densities represent a significant improvement compared to effective potentials computed in the zero-density limit. However, this is not the case for clustering systems, where we have only observed a good agreement between pair correlation functions in MR and CG $_{\rho^*}$  simulations in a very narrow range of low densities for D2 systems and no agreement at all for r6 ones. We speculate that the fundamental problem behind these disagreements is that a cluster fluid has a strongly varying heterogeneous density distribution. By applying the MSCG method, as we have done in section 7.2, we are averaging the internal degrees of freedom (DOF) of an isolated dendrimer with those of a dendrimer inside a clusters of considerably big size (figures 7.6 and 7.9). The fact that the average of the DOF of a dendrimer strongly depends on the local density, *i.e.* the size of the cluster which it is populating, rather than on the global density of the system would serve as an explanation of the encountered failure of finite density effective potentials.

To test this hypothesis, we compute in this section the interaction between a single dendrimer and a cluster of dendrimers with  $n_{\text{occ}} = 2$  to 6. We perform the calculations for  $G = 2$  r6 dendrimers, as in subsection 7.2.2. We compute the effective potentials at zero-density, *i.e.* by considering an isolated pair (formed by the single dendrimer and the cluster of dendrimers). To properly sample the range of distances between the dendrimer and the cluster of dendrimers, we use two springs each of them attached to the centre of mass of one of the entities ( $\mathbf{R}_{\text{isolated}}$  or  $\mathbf{R}_{\text{cluster}}$ ) and the geometrical centre of the whole system,  $\mathbf{R}_{\text{geo}}$ , defined as:

$$\mathbf{R}_{\text{geo}} = \frac{\mathbf{R}_{\text{isolated}} + \mathbf{R}_{\text{cluster}}}{2}. \quad (7.3)$$

The COM of the cluster is defined as:

$$\mathbf{R}_{\text{cluster}} = \frac{1}{n_{\text{occ}}} \sum_{I=1}^{I=n_{\text{occ}}} \mathbf{R}_I, \quad (7.4)$$

where  $\mathbf{R}_I$  is the COM of the  $I$ -th dendrimer in the cluster. The spring exerts a force on the COM of the dendrimer/cluster of dendrimers given by:

$$\mathbf{F}_{\text{isolated/cluster}}(\mathbf{R}_{\text{isolated/cluster}}) = \alpha_t \left( |\mathbf{R}_{\text{geo}} - \mathbf{R}_{\text{isolated/cluster}}| - r_L^t \right) \frac{\mathbf{R}_{\text{geo}} - \mathbf{R}_{\text{isolated/cluster}}}{|\mathbf{R}_{\text{geo}} - \mathbf{R}_{\text{isolated/cluster}}|}, \quad (7.5)$$

where  $\alpha_t$  is the strength of the spring and  $r_L^t$  is the rest length. The force that the springs exerts on the COM of the dendrimer/cluster of dendrimers is subtracted from the accelerations of the monomers, so that the springs do not have any effect on the coarse grained effective

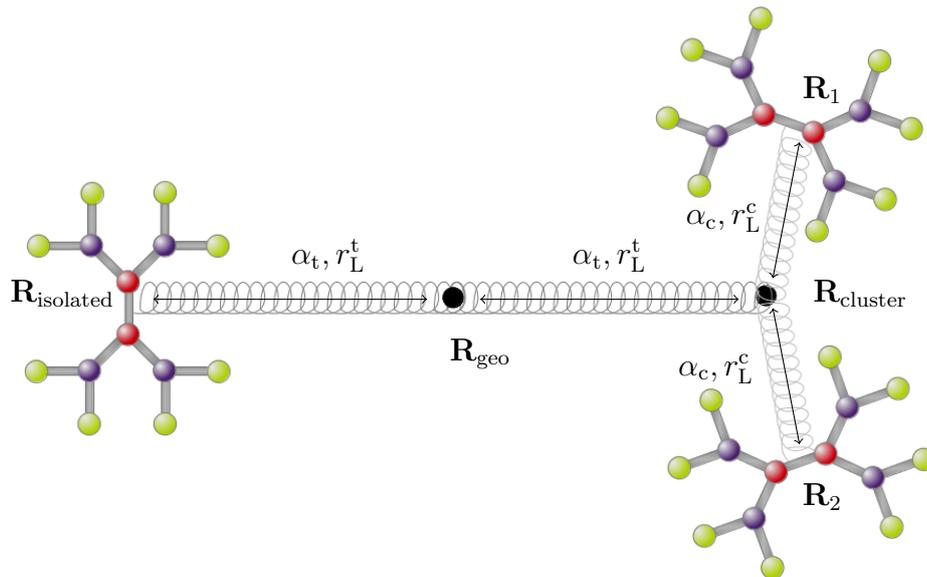


Figure 7.11.: Set up for the computation of the effective interaction between a dendrimer and a cluster of dendrimers of  $n_{occ} = 2$ . Two springs (dark-grey) are attached to the COM of the cluster and the isolated dendrimer, their rest length is being reduced from  $R_{max}/2$  to 0 along the simulation to sample all the distances between  $\mathbf{R}_{isolated}$  and  $\mathbf{R}_{cluster}$ . Additionally two springs (light-grey) are attached to dendrimers in the cluster to keep the cluster together. The rest length of these springs are adjusted to reproduce the value of a cluster of dendrimers of  $n_{occ} = 2$  observed in the bulk system.

potential. The rest lengths of these two springs,  $r_L^t$ , are reduced from  $R_{max}/2$  to zero, where  $R_{max}$  is the distance cut-off of the effective potential, beyond which it is considered that the isolated dendrimer and the cluster do not interact any more. We use  $R_{max} = 20$ . The strength of this spring is set to  $\alpha_t = 5.0$ , the influence of this parameter will be discussed later. Additionally, there are  $n_{occ}$  springs acting on the cluster between  $\mathbf{R}_{cluster}$  and  $\mathbf{R}_I$  to keep the cluster together (see figure 7.11 for a schematic representation of this set up for  $n_{occ}=2$ ), which exert a force on the COMs of the dendrimers in the cluster given by:

$$\mathbf{F}_I(\mathbf{R}_I) = \alpha_c (|\mathbf{R}_{cluster} - \mathbf{R}_I| - r_L^c) \frac{\mathbf{R}_{cluster} - \mathbf{R}_I}{|\mathbf{R}_{cluster} - \mathbf{R}_I|}. \quad (7.6)$$

The forces of these springs are also subtracted from the monomeric accelerations. The rest lengths of these springs,  $r_L^c$ , are kept constant during the whole simulation, and this values are chosen to reproduce the observed size of the clusters in bulk systems. In figure 7.12 we show the radius of gyration of the clusters  $R_g^{cluster}$  as a function of the cluster size  $n_{occ}$  measured in the bulk simulations at different density values in subsection 7.2.2; here the

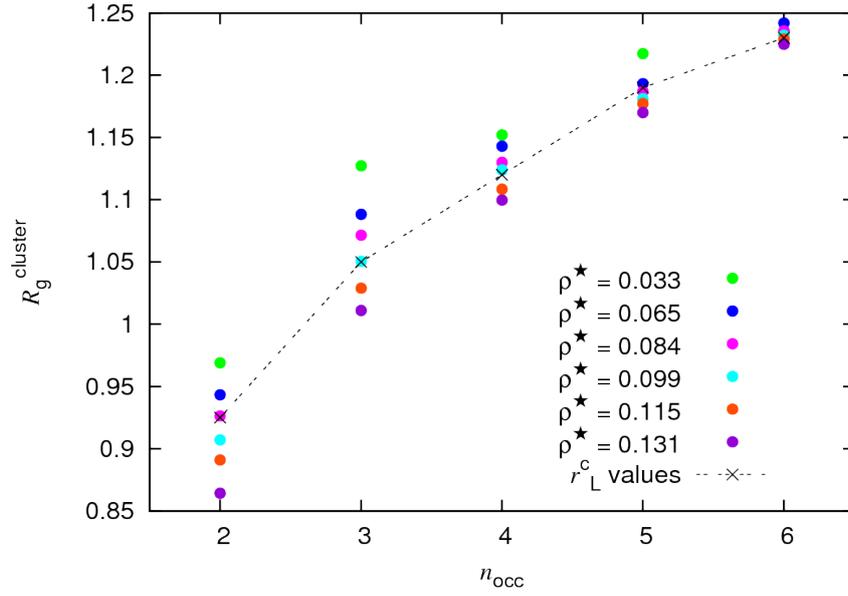


Figure 7.12.:  $R_g^{\text{cluster}}$  as a function of the cluster occupation  $n_{\text{occ}}$  measured in the MR simulations of r6 second generation dendrimers carried out at different density  $\rho^*$  values, as labeled, presented in subsection 7.2.2. We mark with black crosses the rest length value that we choose for each cluster size.

radius of gyration of a cluster is defined as:

$$R_g^{\text{cluster}} = \frac{1}{n_{\text{occ}}} \sum_{I=1}^{I=n_{\text{occ}}} (\mathbf{R}_{\text{cluster}} - \mathbf{R}_I)^2. \quad (7.7)$$

We also show the chosen values for  $r_L^c$ . We observe that for a given cluster size, the value of  $R_g^{\text{cluster}}$  varies very much with the density. We choose for  $r_L$  an intermediate value within the observed range of  $R_g^{\text{cluster}}$ . To test the effect of  $r_L^c$  on the effective potential, we show in figure 7.13 the effective interaction obtained by using different choices  $r_L^c$ , and consequently having different  $R_g^{\text{cluster}}$ . It can be seen that the value of  $R_g^{\text{cluster}}$  has a very little impact on the effective potential. Therefore, the arbitrariness in the choice of  $r_L^c$  does not present a problem.

For the strength of the springs in the cluster we chose  $\alpha_c = 2.0$ . The following considerations took us to the choice of the values of  $\alpha_t$  and  $\alpha_c$ :

- although the value of  $\alpha_{c/t}$  does not have a direct influence on the effective potential, since the force exerted by the spring on the COMs of the dendrimers is subtracted from the monomer accelerations, we want to keep these values small to avoid a possible loss of precision when the spring force is subtracted.
- We need a value of  $\alpha_t$  which is large enough so that  $\mathbf{R}_{\text{isolated}}$  and  $\mathbf{R}_{\text{cluster}}$  stay at the desired distance of  $\mathbf{R}_{\text{geo}}$ , dictated by  $r_L^t$  (see figure 7.11).

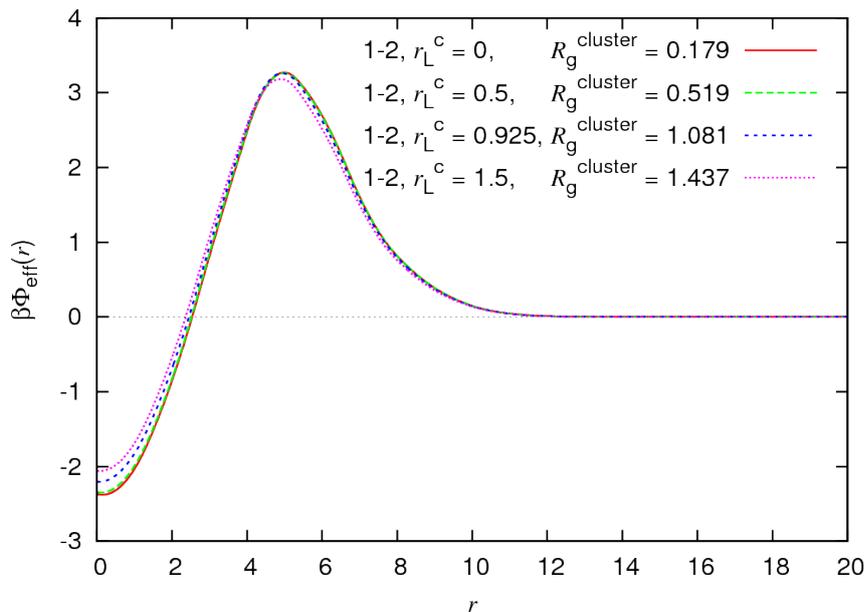


Figure 7.13.: Effective interaction  $\Phi_{\text{eff}}(r)$  between an isolated dendrimer and a cluster of dendrimers of size of  $n_{\text{occ}} = 2$  using different values of  $r_L^c$ ,  $\alpha_c = 2.0$  and  $\alpha_t = 5.0$ . The measured values of  $R_g^{\text{cluster}}$  are also shown in the legend.

- We need a value of  $\alpha_c$  large enough so that  $R_g^{\text{cluster}}$  reproduces the value dictated by  $r_L^c$ . In figure 7.14 we plot  $R_g^{\text{cluster}}$  as a function of the distance between an isolated dendrimer and cluster of dendrimers of size  $n_{\text{occ}} = 2$ . We see that the fluctuations of  $R_g^{\text{cluster}}$  around  $r_L^c$  increase as  $\alpha_c$  is decreased.
- Finally, we observe in figure 7.14 that  $R_g^{\text{cluster}}$  decreases for small  $r$ -values. This could be due to the fact that  $\alpha_t > \alpha_c$ , since for small  $r$ -values  $r_L^t$  tends to zero. To check if this is the case, we perform a new calculation with  $\alpha_t < \alpha_c$ , namely  $\alpha_t = 1.0$  and  $\alpha_c = 2.0$ . The effective interactions for both  $(\alpha_t, \alpha_c)$  combinations and the values of  $R_g^{\text{cluster}}$  as a function of  $|\mathbf{R}_{\text{isolated}} - \mathbf{R}_{\text{cluster}}|$  are shown in figure 7.15. We observe that (i)  $R_g^{\text{cluster}}$  decreases at small  $r$ -values less if  $\alpha_c/\alpha_t > 1$  than if  $\alpha_c/\alpha_t < 1$ , but it still decreases, *i.e.* this effect is not only a consequence of the choice of the  $a$ -values, but also due to the attraction of the dendrimers at short distances, and (ii) the reduction of  $R_g^{\text{cluster}}$  does not have a strong influence on the effective potential.

Once we are sure about effect of the used the parameters ( $\alpha_t = 5.0$ ,  $\alpha_c = 2.0$  and  $r_L^c$  as can be read from figure 7.12) on the final results, we proceed to the discussion of the obtained effective potentials. We perform these calculations at two levels of description, the MR level, where every monomer of the dendrimers is explicitly simulated, and the CG level, where the dendrimers are approximated by CG particles interacting *via* the zero-density effective potential. The resulting potentials are shown in figure 7.16 for  $n_{\text{occ}} = 2$  to 6. In figure 7.17 we plot the effective potentials, normalised by the  $n_{\text{occ}}$  in the MR picture (up) and in the CG one (below). The effective interactions were calculated in all the cases with  $R_{\text{max}} = 20$

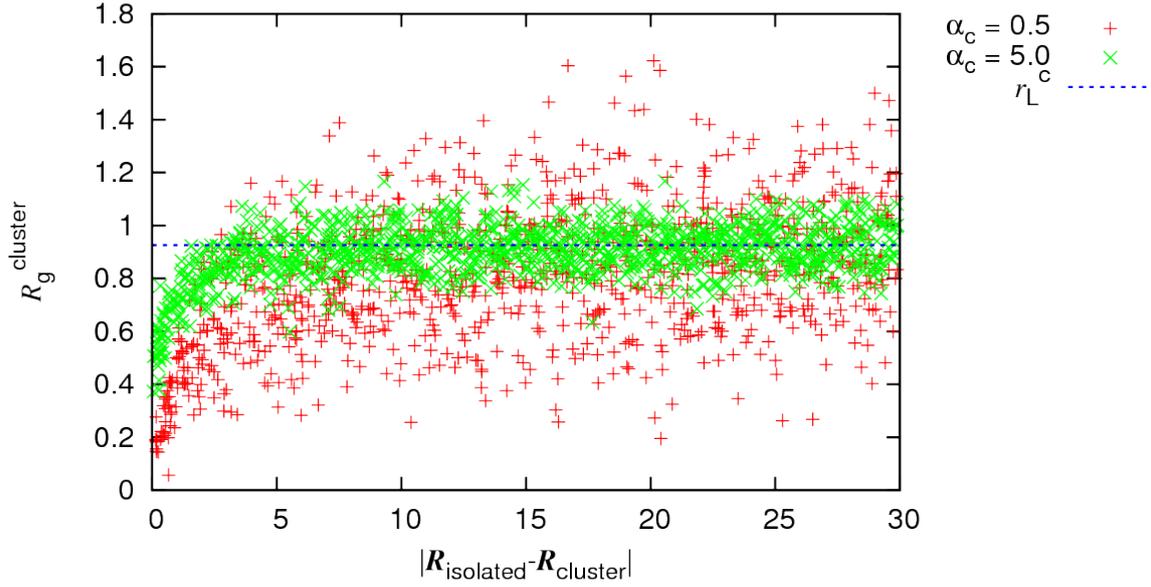


Figure 7.14.:  $R_g^{\text{cluster}}$  as a function of the distance between an isolated dendrimer and cluster of dendrimers of size  $n_{\text{occ}} = 2$ . In both the simulations  $r_L^c$  was 0.925 while  $\alpha_c$  was varied, as labeled.

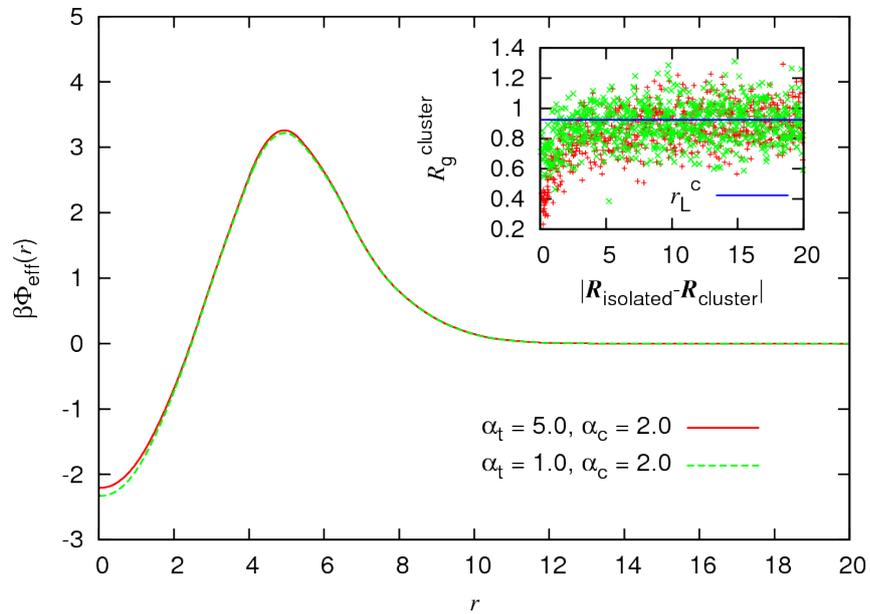


Figure 7.15.: Effective interaction  $\Phi_{\text{eff}}(r)$  between an isolated dendrimer and a cluster of  $n_{\text{occ}} = 2$ , where  $r_L^c = 0.925$  and  $\alpha_c = 2.0$  in both cases. The red curve was calculated with  $\alpha_t = 5.0$  and the green one with  $\alpha_t = 1.0$ . In the inset  $R_g^{\text{cluster}}$  as a function of  $|\mathbf{R}_{\text{isolated}} - \mathbf{R}_{\text{cluster}}|$  is shown.

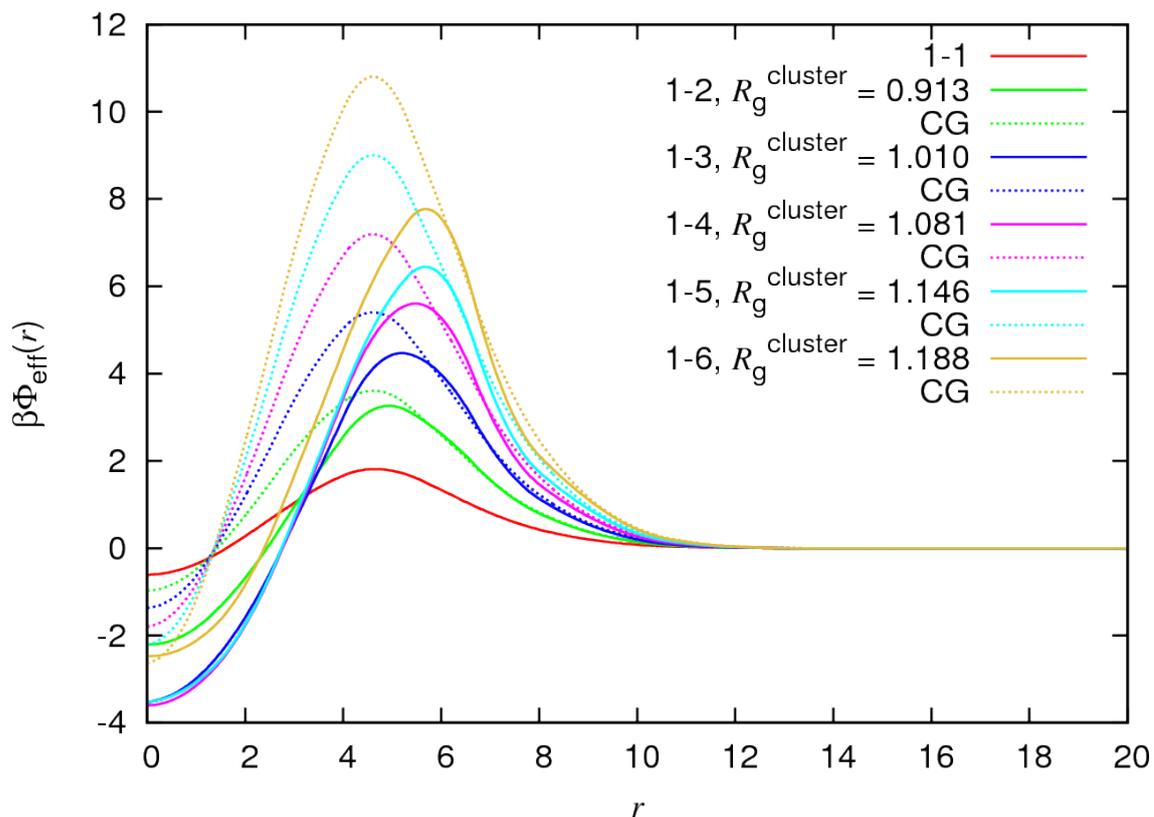


Figure 7.16.: Effective interactions between an isolated dendrimer and a cluster of  $n_{occ} = 2$  to 6 calculated at the MR (solid line) and CG (dotted line), as labeled.

and  $N_D = 15$ . As expected, we observe significant differences in the CG and MR effective potentials. When the CG effective potentials are normalised by  $n_{occ}$ , all the curves collapse to a common curve: the zero-density effective potential used in the CG simulation. On the contrary, the normalised MR effective potentials differ very much. The maximum of the MR effective interaction shifts to larger values of  $r$  as  $n_{occ}$  increases. This is probably due to the fact that the total radius of gyration of the cluster, calculated with the monomer positions (and not just with the COMs of the dendrimers) increases with  $n_{occ}$ . Unfortunately we do not have measurements of the total radius of gyration (because we only saved the positions of the COMs of the dendrimers), but we speculate that if the  $x$ -axis was normalised by these value, the maxima of the effective potentials would coincide in the same point. It is remarkable how for a given  $n_{occ}$  value the MR effective potential has a smaller maximum, meaning that cluster formation will be more favoured in the MR picture as in the CG one.

This observations help us understand why in subsection 7.2.2 we observed that pair correlation functions in the MR simulations are not always reproduced in CG simulations using effective potentials calculated at finite densities. The behaviour of a dendrimer depends very much on the size of the cluster it belongs to, and in clustering systems in a fluid configuration we always have a large polydispersity in the size of the clusters.

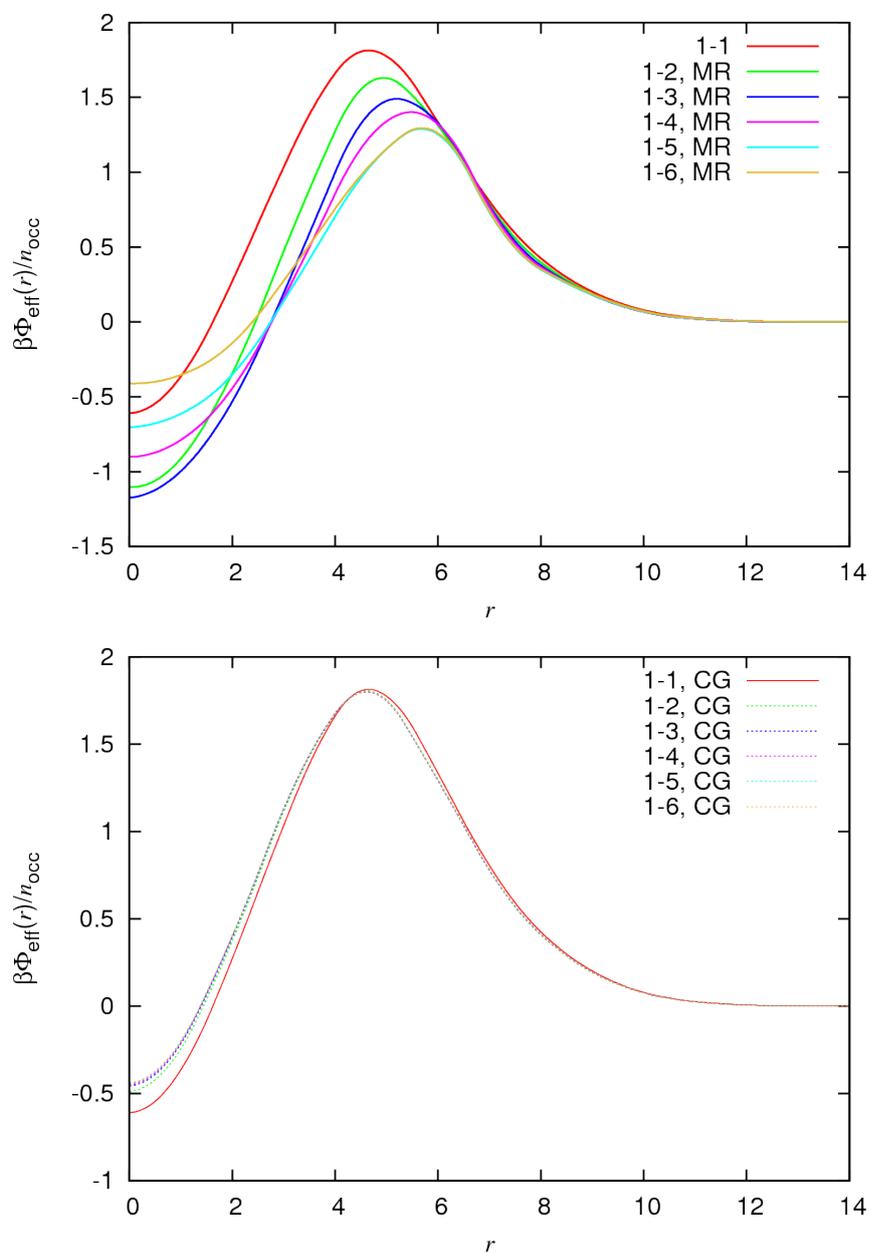


Figure 7.17.: Effective interactions between an isolated dendrimer and a cluster of  $n_{\text{occ}} = 2$  to 6 normalised by  $n_{\text{occ}}$  in the MR (left panel) and CG (right panel) level.



## 8. Conclusions & Outlook

In this work, we have studied several properties of cluster-forming ultrasoft systems of polymeric macromolecules at different description levels: on the monomer resolved (MR) level, where each monomer of the polymers is considered explicitly, and on the coarse-grained (CG) level, where the internal degrees of freedom of the molecules are averaged out and each polymer is represented by an effective CG particle placed in its centre of mass.

Via CG simulations of the generalised exponential model (GEM) we have studied the influence of hydrodynamic interactions due to the presence of the solvent in diffusion and hopping processes in a pure cluster crystal of GEM-4 particles and in a binary mixture of cluster-forming GEM-8 particles and non-cluster-forming Gaussian GEM-2 particles. We have used the multi-particle collision dynamics method [34, 35], which explicitly includes the solvent and takes hydrodynamic interactions into account as faithfully as possible. By tuning a suitable parameter of the simulation, the coupling between the solute and the solvent can be controlled. We have found that the presence of the solvent has a dramatic effect on the dynamics of a cluster crystal: upon increasing the coupling parameter the number of particles involved in a hopping event as well as the length of their trajectory during a jump dramatically decrease and the inner cluster vibrations are suppressed: the solvent acts as a damping buffer. By studying the angles enclosed by successive jumps in a hopping event, we have seen that the nature of the trajectories is strongly influenced by the presence of the solvent: at vanishing or low coupling between the solvent and the solute hopping particles tend to perform straight or slightly deflected trajectories, while at high coupling we have mainly observed small angles ( $30^\circ$  to  $90^\circ$ ). We also were able to answer the yet open question of whether Lévy flights [80, 83] take place in cluster crystals: through our analysis we excluded their occurrence.

We have also studied at the CG level the response of a cluster crystal to external compression. We have exposed a system of cluster-forming purely repulsive GEM-4 particles to a pressure bath composed of ideal gas particles. By controlling the number of ideal gas particles and the distribution of their velocities we were able to tune the pressure and temperature inside the GEM-4 system. We have observed that in a simple compression experiment at constant temperature the spacing of the lattice remains, irrespective of the temperature and the compression rate, almost constant, though at a value lower than the equilibrium value, and the cluster occupancy increases linearly with density. The GEM-4 system achieves this behaviour by first increasing the hopping activity and producing an heterogeneous cluster size distribution and then merging smaller clusters (of the size of  $\sim 70\%$  of the average cluster occupancy in the system) which are pushed together by their neighbouring bigger clusters. We have learnt that in order to recover the equilibrium configuration of a GEM-4 cluster crystal one has to combine compression runs at constant temperature with annealing processes at

constant pressure, where the temperature in the GEM-4 system is increased until the system begins to melt and then lowered to the original value. In this way, the system receives an energy kick which allows it to make the necessary cluster rearrangements to recover equilibrium, which at constant temperature represent a (too high) energetic barrier separating the metastable from the real equilibrium. Finally, we have learnt *via* a combination of compression, annealing and expansion experiments that a cluster crystal remains (meta)stable within a range of lattice spacing of 10% around the equilibrium value. When the system is compressed or expanded it will first shrink or expand within this range, and only once an extreme value is adopted will cluster merging or cluster splitting events take place. In the future, it would be interesting to study the response to compression of a cluster crystal in the explicit presence of a solvent. Our study has shown that the hopping activity is largely suppressed when hydrodynamic interactions are taken into account, but we also have seen that the increase in the hopping activity is necessary to generate the heterogeneous cluster size distribution in the system which leads to merging of the smaller clusters. Will merging events still take place if the solvent is present? Will the lattice constant remain unchanged and the cluster occupation linearly increase with density?

At the MR level of description we have worked on the *in silico* design of an amphiphilic linear chain composed by a solvophobic backbone decorated by solvophilic side groups. We have come up with a combination of the parameters in the model which leads to an effective interaction potential whose Fourier transform takes negative values, and therefore predicts the formation of cluster crystals at high densities. We have performed MR simulations of the amphiphilic chains in their cluster crystal configuration and we have proven the stability of the system.

Finally, we have studied the appropriateness of effective interactions computed at finite density for the description of bulk phases of ultrasoft polymeric macromolecules. We have used the so-called multi-scale coarse graining method [15] developed by Izvekov and Voth, which allows for the computation of effective interactions at finite densities *via* a force matching method, and to our knowledge, has not been used up to date in the field of ultrasoft matter. We have seen that in non-clustering systems effective potentials computed at finite densities represent a significant improvement with respect to effective potentials computed at zero density: if the former are used in CG simulations (instead of the latter) the range of densities in which pair correlation functions from MR simulations can be perfectly reproduced is widely increased. However, we have learnt that the cluster fluid phase cannot be properly described by finite density effective potentials, because such a system is a mixture of very different local densities. To understand this issue, we have computed the effective interaction of an isolated dendrimer and a cluster of dendrimers of sizes ranging from 2 to 6. We have performed these calculations at two description levels: on the MR level, where each monomer is considered explicitly; and on the CG level, where each dendrimer is represented as a CG particle interacting *via* the zero-density effective potential. We have observed that while the effective potential between a dendrimer and a cluster computed in CG simulations is additive (by rescaling the potential by the cluster occupation we obtain in each case the

---

zero-density potential), this is not the case for the effective potentials obtained in the MR simulations. This proves that the average over the internal degrees of freedom of a dendrimer differs very much depending on the size of the cluster it belongs to, and therefore averaging in the same way the internal degrees of freedom of dendrimers within very different local densities is a very weak approximation for the description of the system. We suggest to study if an effective interaction including a two-body and a three-body term would lead to a more reliable description clustering systems. The multi-scale coarse graining method offers an outstanding flexibility for the functional form of the effective interaction and allows for the calculation of effective three-body interactions. These calculations were in our plan, but unfortunately due to lack of time, we could not carry them out.



# Appendices



## A. Reduced units

The parameters  $\epsilon$  and  $\sigma$ , as defined in equation (2.1) are used as the respective energy and length units. The mass of the particle (in each system we specify which particle),  $m$ , will be used as the mass unit. The rest of the quantities are expressed as reduced dimensionless quantities:

- the number-density:  $\rho^* = \rho\sigma^2$ ,
- the temperature:  $T^* = k_B T/\epsilon$ ,
- the pressure:  $P^* = P\epsilon/(k_B\sigma^2)$ ,
- the time unit:  $t^* = \sigma\sqrt{m/\epsilon}$
- and the spring strength:  $\alpha^* = \epsilon/\sigma^2$ .

For simplicity the stars will be dropped.



## B. Cluster Analysis

The cluster analysis is carried out in a five-step procedure: in **step one** of our cluster analysis of a given particle configuration we have used the scheme detailed in the Appendix of reference [25]. Then we calculated the individual ( $n_{\text{occ}}$ ) and the average ( $\langle n_{\text{occ}} \rangle$ ) occupancy of each cluster as well as its spatial extent, quantified by its radius of gyration,  $r_g$ .

The first class of clusters that requires special treatment in **step two** are oversized clusters, *i.e.* all clusters with an occupation number larger than  $N_{\text{occ}}^{\text{max}} = 1.6\langle n_{\text{occ}} \rangle$ . Such clusters can appear either because (a) two under-occupied clusters have merged or (b) because a hopping particle is located at this very instant between two clusters and the cluster analysis could not identify them as separate aggregates (“fused clusters”). A suitable quantity for distinguishing these two cases is the radius of gyration  $r_g$ , and we set  $r_g^{\text{max}} = 1.4\sigma$  to identify fused clusters. This test is carried out *via* the following iterative algorithm: we identify for each particle of such a cluster its average number of neighbours; all particles that have less neighbours than this average value are put aside for the moment and the cluster analysis is repeated with the remaining particles, whereby the original cluster meanwhile might have split up into two (or more) (sub-)clusters. If the radii of gyration of the resulting clusters are still larger than  $1.4\sigma$ , the value of  $r_c$  is reduced by 10% and steps 1 and 2 are repeated until the  $r_g$ -values of the emerging clusters are smaller than  $1.4\sigma$ . The isolated particles that have been previously put aside during the algorithm are then reintegrated and assigned to the respective nearest cluster.

In **step three** we focus on small clusters ( $n_{\text{occ}} < 3$ ) which are most probably migrating particles. They are assigned to the respective nearest cluster, but are not included in the evaluation of its centre-of-mass.

Finally, we focus in **step four** on clusters that are separated by a distance that is smaller than  $0.65\sigma$ , which we consider as merged clusters. This particular value for the threshold distance represents a reasonable choice in view of the fact that the lattice constant of our cluster crystal assumes values close to  $1.37\sigma$ .

Finally we identify and analyse in **step five** merging and cluster separation events. If in a compression step two clusters have merged from one configuration  $I$  to the subsequent one ( $I + 1$ ), one of the cluster identities is maintained for the further analysis while the other one is put idle. If, in contrast, in an expansion step, a cluster has separated into two (sub-)clusters from one configuration to the following one, the subcluster that is closer to the original aggregate keeps the identity of the latter, while for the other subcluster a new identity is created. All other aggregates of configuration ( $I + 1$ ) will carry on the cluster identity of those aggregates that are closest in configuration  $I$  to their actual position. If the association of a particle to a cluster changes from one configuration to the subsequent

one, this particle is identified as a hopping particle; relevant information about the particle and the cluster it belonged to is stored. Through this criterion also particles involved in merging processes are identified as hopping particles; using now the information obtained in the preceding cluster-tracking analysis we can finally separate hopping particles from those that are involved in merging processes in a final step.

# Bibliography

- [1] C. N. Likos, *Physics Reports* **348**, 267 (2001).
- [2] C. Kittel, *Introduction to Solid State Physics* (John Wiley & Sons, Inc, 2005).
- [3] V. Parsegian, *Van der Waals forces: a handbook for biologists, chemists, engineers, and physicists* (Cambridge University Press, 2006).
- [4] A. A. Louis, P. G. Bolhuis, J. P. Hansen, and E. J. Meijer, *Phys. Rev. Lett.* **85**, 2522 (2000).
- [5] C. N. Likos, M. Schmidt, H. Löwen, M. Ballauf, and D. Pötschke, *Macromolecules* **34**, 2914 (2001).
- [6] M. Bohn and D. W. Heermann, *J. Chem. Phys.* **132**, 044904 (2010).
- [7] D. Gottwald, C. N. Likos, G. Kahl, and H. Löwen, *Phys. Rev. Lett.* **92**, 068301 (2004).
- [8] F. H. Stillinger, *J. Chem. Phys.* **39**, 1976 (65).
- [9] C. Marquest and T. A. Witten, *J. Phys.* **50**, 1267 (1989).
- [10] C. N. Likos, M. Watzlawek, and H. Löwen, *Phys. Rev. E* **58**, 3135 (1998).
- [11] C. N. Likos, A. Lang, M. Watzlawek, and H. Löwen, *Phys. Rev. E* **63**, 031206 (2001).
- [12] G. M. Torrie and J. P. Valleau, *J. Comp. Phys.* **23**, 187 (1977).
- [13] B. Widom, *J. Chem. Phys.* **39**, 2802 (1963).
- [14] A. Narros, A. J. Moreno, and C. N. Likos, *Soft Matter* **6**, 2435 (2010).
- [15] S. Izvekov and G. A. Voth, *J. Phys. Chem. B* **109**, 7 (2004).
- [16] M. Watzlawek, C. N. Likos, and H. Löwen, *Phys. Rev. Lett.* **82**, 5289 (1999).
- [17] M. W. Matsen and M. Schick, *Phys. Rev. Lett.* **72**, 2660 (1994).
- [18] G. A. M. ell and P. A. Gast, *Phys. Rev. E* **54**, 5447 (1996).
- [19] G. A. McConnell and P. A. Gast, *Macromolecules* **30**, 435 (1997).
- [20] P. A. Gast, *Curr. Opinion Colloid Interface Sci.* **2**, 258 (1997).
- [21] B. M. Mladek, D. Gottwald, G. Kahl, M. Neumann, and C. N. Likos, *Phys. Rev. Lett.* **96**, 045701 (2006).
- [22] B. M. Mladek, P. Charbonneau, and D. Frenkel, *Phys. Rev. Lett* **99**, 235702 (2007).

- [23] C. N. Likos, B. M. Mladek, D. Gottwald, and G. Kahl, *J. Chem. Phys.* **126**, 224502 (2007).
- [24] A. J. Moreno and C. N. Likos, *Phys. Rev. Lett.* **99**, 107801 (2007).
- [25] D. Coslovich, L. Strauss, and G. Kahl, *Soft Matter* **7**, 2127 (2011).
- [26] N. Wilding and P. Sollich, *EPL* **101**, 1 (2013).
- [27] D. A. Lenz, B. M. Mladek, C. N. Likos, G. Kahl, and R. Blaak, *J. Phys. Chem. B* **115**, 7218 (2010).
- [28] D. A. Lenz, R. Blaak, C. N. Likos, and B. M. Mladek, *Phys. Rev. Lett.* **109**, 228301 (2012).
- [29] T. Terashima, T. Mes, T. F. A. De Greef, M. A. J. Gillisen, P. Besenius, A. R. A. Palmans, and E. W. Meijer, *J. Am. Chem. Soc.* **133**, 4247 (2011).
- [30] M. Camargo, A. J. Moreno, and C. N. Likos, *JSTAT* p. P10015 (2010).
- [31] A. Nikoubashman, G. Kahl, and C. N. Likos, *Phys. Rev. Lett* **107**, 068302 (2011).
- [32] G. Nägele and P. Baur, *Physica A* **245**, 297 (1997).
- [33] M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper, *EPL* **68**, 106 (2004).
- [34] A. Malevanets and R. Kapral, *J. Chem. Phys.* **110**, 8605 (1999).
- [35] G. Gompper, T. Ihle, D. M. Kroll, and R. G. Winkler, *Adv. Polym. Sci.* **221**, 1 (2009).
- [36] M. Grünwald, E. Rabani, and C. Dellago, *Phys. Rev. Lett.* **96**, 255701 (2006).
- [37] B. M. Mladek, D. Gottwald, G. Kahl, M. Neumann, and C. N. Likos, *J. Phys. Chem. B* **111**, 12799 (2007).
- [38] B. M. Mladek, G. Kahl, and C. N. Likos, *Phys. Rev. Lett* **100**, 028301 (2008).
- [39] P. Welch and M. Muthukumar, *Macromolecules* **31**, 5892 (1998).
- [40] A. A. Louis, *J. Phys.: Condens. Matter* **14**, 9187 (2002).
- [41] J. L. Barrat and J. P. Hansen, *Basic Concepts for Simple and Complex Liquids* (Cambridge University Press, 2003).
- [42] J. G. Kirkwood, *Phase Transitions in Solids* (Wiley, New York, 1951).
- [43] A. Nikoubashman, Ph.D. thesis, Technische Universität Wien (2012).
- [44] C. N. Likos, B. M. Mladek, A. J. Moreno, D. Gottwald, and G. Kahl, *Comput. Phys. Commun.* **179**, 71 (2008).
- [45] L. Strauss, Master's thesis, Technische Universität Wien (2009).
- [46] B. M. Mladek, Ph.D. thesis, Technische Universität Wien (2007).

- [47] E. Buhleier, W. Wehner, and F. Vögtle, *Synthesis-Stuttgart* **2**, 155 (1978).
- [48] M. Ballauf and C. N. Likos, *Angew. Chem. Int. Ed.* **43**, 2998 (2004).
- [49] R. L. Lescanec and M. Muthukumar, *Macromolecules* **23**, 2280 (1990).
- [50] T. C. Zook and G. T. Pickett, *Phys. Rev. Lett.* **90**, 015502 (2003).
- [51] B. J. Bauer, B. Hammouda, R. Briber, and D. A. Tomalia, *Polym. Mater. Sci. Eng.* **67**, 340 (1997).
- [52] T. J. Prosa, B. J. Bauer, E. J. Amis, D. A. Tomalia, and R. Scherrenberg, *J. Polym. Sci., Part B: Polym. Phys.* **35**, 2913 (1997).
- [53] C. N. Likos and M. Ballauf, *Functional Molecular Nanostructures of the Series Topics in Current Chemistry*, vol. Chap. Equilibrium Structure of Dendrimers - Results and Open Questions (Springer-Verlag, Berlin, 2005).
- [54] S. Rosenfeldt, N. Dingenouts, D. Poetschke, M. Ballauf, A. J. Beeresheim, K. Muellen, and P. Linder, *Angew. Chem. Int. Ed.* **116**, 111 (2004).
- [55] J. K. Young, G. R. Baker, G. R. Newkome, K. F. Morris, and C. S. Johnson, *Macromolecules* **27**, 3464 (1994).
- [56] R. Briber, B. Bauer, B. Hammouda, and D. Tomalia, *Polym. Mater. Sci. Eng.* **67**, 430 (1992).
- [57] B. Helms and E. W. Meijer, *Science* **313**, 929 (2006).
- [58] L. J. Twyman, A. E. Beezer, R. Esfand, M. J. Hardy, and J. C. Mitchell, *Tetrahedrom Lett.* **40**, 1743 (1999).
- [59] N. Malik, R. Wiwattanapatapee, R. Klopsch, K. Lorenz, H. Frey, J. W. Weener, E. W. Meijer, W. Paulus, and R. Duncan, *J. Control. Release* **65**, 133 (2000).
- [60] A. E. Beezer, A. S. H. King, I. K. Martin, J. C. Mitchel, L. J. Twyman, and C. F. Wain, *Tetrahedrom* **59**, 3873 (2003).
- [61] F. Aulenta, W. Hayes, and S. Reannard, *Eur. Polym. J.* **39**, 1741 (2003).
- [62] N. Nishiyama, A. Iriyama, W. D. Jang, K. Miyata, K. Itaka, Y. Inoue, H. Takahashi, Y. Yanagi, Y. Tamaki, H. Koyama, et al., *Nat. Mater.* **4**, 934 (2005).
- [63] C. C. Lee, J. A. MacKay, J. M. J. Frechet, and F. C. Szoka, *Nat. Biotechnol.* **23**, 1517 (2005).
- [64] T. D. McCharthy, P. Karellas, S. A. Henderson, M. Giannis, D. F. O'Keefe, G. Heery, J. R. A. Paull, B. R. Matthews, and G. Holan, *Mol. Pharmaceutics* **2**, 2005 (312).
- [65] P. Antoni, D. Nystroem, C. J. Hawker, A. Hult, and M. Malkoch, *Chem. Commun.* **22**, 2249 (2007).

- [66] I. O. Göetze, H. M. Harreis, and C. N. Likos, *J. Chem. Phys.* **120**, 7761 (2004).
- [67] C. N. Likos, S. Rosenfeldt, N. Dingenouts, M. Ballauff, P. Lindner, N. Werner, and F. Vögtle, *J. Chem. Phys.* **117**, 1869 (2002).
- [68] D. Frenkel and B. Smit, *Understanding Molecular Simulations* (Academic Press, 2002).
- [69] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76**, 637 (1982).
- [70] H. C. Andersen, *J. Chem. Phys.* **72**, 2384 (1980).
- [71] S. Nosé, *J. Chem. Phys.* **81**, 511 (1984).
- [72] S. Nosé, *Mol. Phys.* **52**, 255 (1984).
- [73] W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- [74] W. G. Hoover, *Phys. Rev. A* **34**, 2499 (1986).
- [75] T. Ihle and D. M. Kroll, *Phys. Rev. E* **63(2)**, 020201 (2001).
- [76] S. Izvekov and G. A. Voth, *J. Phys. Chem. B* **109** (2005).
- [77] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, and H. C. Andersen, *J. Chem. Phys.* **128** (2008).
- [78] C. C. Paige and M. A. Saunders, *ACM Transactions on Mathematical Software* **8**, 1 (1982).
- [79] C. C. Paige and M. A. Saunders, *ACM Transactions on Mathematical Software* **8**, 2 (1982).
- [80] M. F. Schlesinger, G. M. Zaslavsky, and J. Klafter, *Nature* **363**, 31 (1993).
- [81] W. Paul and J. Baschnagel, *Stochastic Processes. From Physics to Finance* (Springer, 1999).
- [82] R. Klages, G. Radons, and I. M. Sokolov, *Anomalous Transport* (Wiley-VCH (Weinheim), 2008).
- [83] G. M. Zaslavsky, *Phys. Rep.* **371**, 461 (2002).
- [84] M. Montes-Saralegui, A. Nikoubashman, and G. Kahl, *J. Phys: Condens. Matter* **25**, 195101 (2013).
- [85] J. P. Boon and S. Yip, *Molecular Hydrodynamics* (Dover Publications, 1992).
- [86] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, 2006).
- [87] S. D. Overduin and C. N. Likos, *EPL* **85**, 26003 (2009).
- [88] S. D. Overduin and C. N. Likos, *J. Chem. Phys.* **131**, 034902 (2009).

- [89] M. Grünwald and C. Dellago, *Molecular Physics* **104**, 3709 (2006).
- [90] M. Grünwald, C. Dellago, and P. L. Geissler, *J. Chem. Phys.* **127**, 1547187 (2007).
- [91] M. Montes-Saralegui, A. Nikoubashman, and G. Kahl, *J. Chem. Phys.* **141**, 124908 (2014).
- [92] K. J. Strandburg, *Rev. Mod. Phys.* **60**, 161 (1988).
- [93] P. G. Bolhuis, A. A. Louis, J. P. Hansen, and E. J. Meijer, *J. Chem. Phys.* **114**, 4296 (2001).
- [94] S. Izvekov, P. W. Chung, and B. M. Rice, *J. Chem. Phys.* **133**, 064109 (2010).
- [95] S. Plimpton, *J. Comp. Phys.* **117**, 1 (1995).
- [96] I. O. Götze, H. M. Harreis, and C. N. Likos, *J. Chem. Phys.* **120**, 16 (2004).
- [97] I. O. Götze and C. N. Likos, *J. Phys.: Condens. Matter* **17**, S1777 (2005).

The work presented here resulted in the following original publications:

- M.Montes-Saralegui, A. Nikoubashman and G. Kahl  
*Hopping and diffusion of ultrasoft particles in cluster crystals in the explicit presence of a solvent,*  
J. Phys.: Condens. Matter, **25** 195101 (2013)
- M.Montes-Saralegui, A. Nikoubashman and G. Kahl  
*Merging and hopping processes in systems of ultrasoft, cluster forming particles under compression,*  
J. Chem. Phys., **141** 124908 (2014)
- M.Montes-Saralegui, and G. Kahl  
*Derivation of the equation of state of a cluster forming system via Molecular Dynamic simulations,*  
in preparation
- M.Montes-Saralegui, A. Nikoubashman and G. Kahl  
*Effective interactions of ultrasoft polymeric macromolecules computed at finite densities,*  
in preparation
- M.Montes-Saralegui, A. Nikoubashman and G. Kahl  
*Cluster forming amphiphilic polymer chains,*  
in preparation

This work has been supported by the Marie Curie ITN-COMPLOIDS (Grant Agreement No.234810), by the FWF under Project No. P19890-N16 and by the Vienna Graduate School on Computational Materials Science.

Computer time on the Vienna Scientific Cluster (VSC-1) is gratefully acknowledged.

## Acknowledgements

- Thanks to **Gerhard Kahl**, for giving me the opportunity of doing this PhD. Thanks for being such a human person and making possible this nice working atmosphere. Thank you also for your patience and for all what you have taught me.
- Thanks to **Arash Nikoubashman** for motivating me to do this PhD, for taking care of me as an old brother, for being a mentor and always being ready to help with whatever would pop up. Thanks also for the speakers (which very much helped me going through this thesis), they are awesome!
- Thanks to my work colleagues: **Dieter, Emanuela, Jan, Arash**, the lunch king **Günther, Giannis, Ulf, Andy, Martina**, what?-it's-true!-**Moritz, Ismene, Cecilia, Clara** and last but not least, but-**Silvano**. Thanks for the nice lunch breaks and for always being ready to help with any doubt or problem. It has been a pleasure working with you!
- Thanks to **Christoph Dellago** and **Michael Grünwald** for their readiness to help implementing their method.
- Gracias al conjunto de **profesores de la facultad de física** de la universidad de Zaragoza por la excelente formación que me han dado, con especial mención a **Fernando Bartolomé**, sin quien todo esto no hubiera sido posible.
- Gracias a **Alberto**, porque tampoco ibas a faltar en esta y porque sé que eres el único que se va a leer esta tesis de pe a pa.
- Danke an der **Familie Doppelbauer**, weil ihr meine Familie gewesen seid, als meine so weit weg war.
- Gracias a **mi extensa familia** incluyendo tíos, primos, cuñados..., por vuestro apoyo y esfuerzo en comprender mi trabajo. Gracias también a las pequeñas de la familia, cuyas fotos y videos han amenizado las largas jornadas de trabajo. Finalmente, una especial mención a mi tío **Agustín Falcón**, muchas gracias por tu ineterés y tu crítica siempre constructiva.
- Gracias a mi hermana **Lucía** por siempre creer en mi, por tu cariño, tu apoyo y por poner luz (y sentido común) a los momentos oscuros. Gracias por marcarme el camino a seguir: siempre hacer lo contrario a lo que hayas hecho tu antes, y sobre todo gracias por no haber estudiado física, porque una tesis en arbitraje internacional me hubiera matado.
- Gracias a **mis padres** por habermelo dado todo en la vida. Gracias por vuestro apoyo incondicional y vuestra guia.
- Gracias a **Günther** por tu cariño, por tu apoyo, por tu ayuda, por tu paciencia, por tus cuidados... por todo!

# Curriculum Vitae

## 1) Personal data

Name	Marta Montes Saralegui
Date of birth	22. February 1988
Place of birth	San Sebastián, Spain
Citizenship	Spain

## 2) Education

since 11/2011	<b>PhD studies</b> <ul style="list-style-type: none"> <li>• <i>Technische Universität Wien</i> (Vienna, Austria)</li> </ul>
09/2006 - 06/2011	<b>Licenciatura en Física</b> <ul style="list-style-type: none"> <li>• <i>Universidad de Zaragoza</i> (Zaragoza, Spain)</li> <li>• 9<sup>th</sup> and 10<sup>th</sup> semesters studied at the <i>Technische Universität Wien</i> within the Erasmus exchange program (Ibercaja excellence scholarship)</li> </ul>
09/2004 - 06/2006	High school scientific-technological diploma <ul style="list-style-type: none"> <li>• <i>St. Patrick's English School</i> (San Sebastián, Spain)</li> </ul>

## 3) Further achievements

07/2010 - 08/2010	<b>Introduction to the investigation</b> scholarship by the CSIC <ul style="list-style-type: none"> <li>• “<i>Simulation of Light Absorption and Emission Spectra in Systems with Biological Interest</i>”</li> <li>• Under supervision of Prof. Angel Rubio, <i>Centro de Física de Materiales</i> (San Sebastián, Spain)</li> </ul>
-------------------	---