2010

# Modeling and simulation of nanoparticle aggregation in colloidal systems

Sergiy Markutsya
*Iowa State University*

**Modeling and simulation of nanoparticle aggregation in colloidal systems**

by

Sergiy Markutsya

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Mechanical Engineering

Program of Study Committee:
Shankar Subramaniam, Major Professor
Rodney O. Fox,
Monica H. Lamm,
Pranav Shrotriya
Sriram Sundararajan
R. Dennis Vigil

Iowa State University

Ames, Iowa

2010

## DEDICATION

I would like to dedicate this dissertation to my wife Svitlana and to my daughter Alina without whose support I would not have been able to complete this work. I would also like to thank my family for their loving guidance and support during the writing of this work.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere thanks and deep appreciation to those who helped me with various aspects of conducting research and the writing of this dissertation. First and foremost, I would like to thank Dr. Shankar Subramaniam for his constant guidance, encouragement, patience and support throughout this research and the writing of this dissertation. His scientific knowledge, technical skills and enlightened views on the research process has left a deep impression on me. How fortunate I am to have been a graduate student in his laboratory.

I would also like to thank my POS committee members for their efforts and contributions to this work: Dr. Rodney Fox, Dr. Monica Lamm, Dr. Pranav Shrotriya, Dr. Sriram Sundararajan, and Dr. Dennis Vigil.

My special appreciation extends to all my lab-mates: Ragul Garg, Madhu Pai, Xu Ying, Vidyapati, Tenneti Sudheer, Ravi Kolakaluri, Karthikeyan Devendran, Mohammad Mehrabadi, Bo Sun, and Christopher Schmitz. Thank you for your friendship and essential support. You made my life in the lab comfortable and enjoyable.

My thanks to the Department of Mechanical Engineering and its professional and administrative staff, especially Amy Carver, Jannet Huqqard, Sherrie Nystrom, Janelle Miranda, Hap Steed, and Nate Jensen for their flawless and cheerful assistance during the past years.

Finally, I would like to extend love and thanks to my family, especially to my wife Svitlana for her love, patience and enduring support during my graduate studies at ISU. I am grateful for my mother and my mother-in-law for their help with caring for our daughter Alina, whose entry into our lives gave us extra motivation to pursue our dreams.

# CHAPTER 1. INTRODUCTION

Nanoparticles are widely used as building blocks in nanotechnology research and they offer the promise of creating new materials and new applications in the nanoscale range. Moreover, properties of such materials differs from bulk material properties [1]. These novel properties are observable only at the nano-scale dimensions have already found their first commercial applications [2]. For example, latex nanoparticles are used for a variety of biological applications [3].

Two high-rate synthesis methods are commonly used in the industry: aerosol reactors in a gaseous environment and colloidal reactors in a fluid environment [1, 4]. In both methods the synthesis of the particles occurs in turbulent reactors due to the reaction of chemical precursors and the formation of nuclei, which rapidly grow due to surface addition and/or aggregation. Such a synthesis process subjects nanoparticle aggregates to a spatially homogeneous, time-varying shear flow and is characterized by the variety of time–scales and length–scales from size of single particle to the size of particles aggregate. The next generation of applications will require improvement in the quality of the monodispersity, purity, and uniform surface chemistry of nanoparticles [4]. Since the aggregation of the sheared colloidal nanoparticles is the important part of this process, a better understanding of the sheared aggregation phenomenon will help to improve synthesis methods.

Aggregating systems that are studied in the literature can be classified by: (a) their compositions: for example solute particles in solvent (latex particles in solvent), polymer chains in colloidal systems, etc; (b) concentration of substance; (c) presence (or absence) of external force: gravitational force, shear flow, etc. In aggregating systems a rich variety of phenomena is observed. Competition between the physical mechanisms of interparticle attraction, intensity

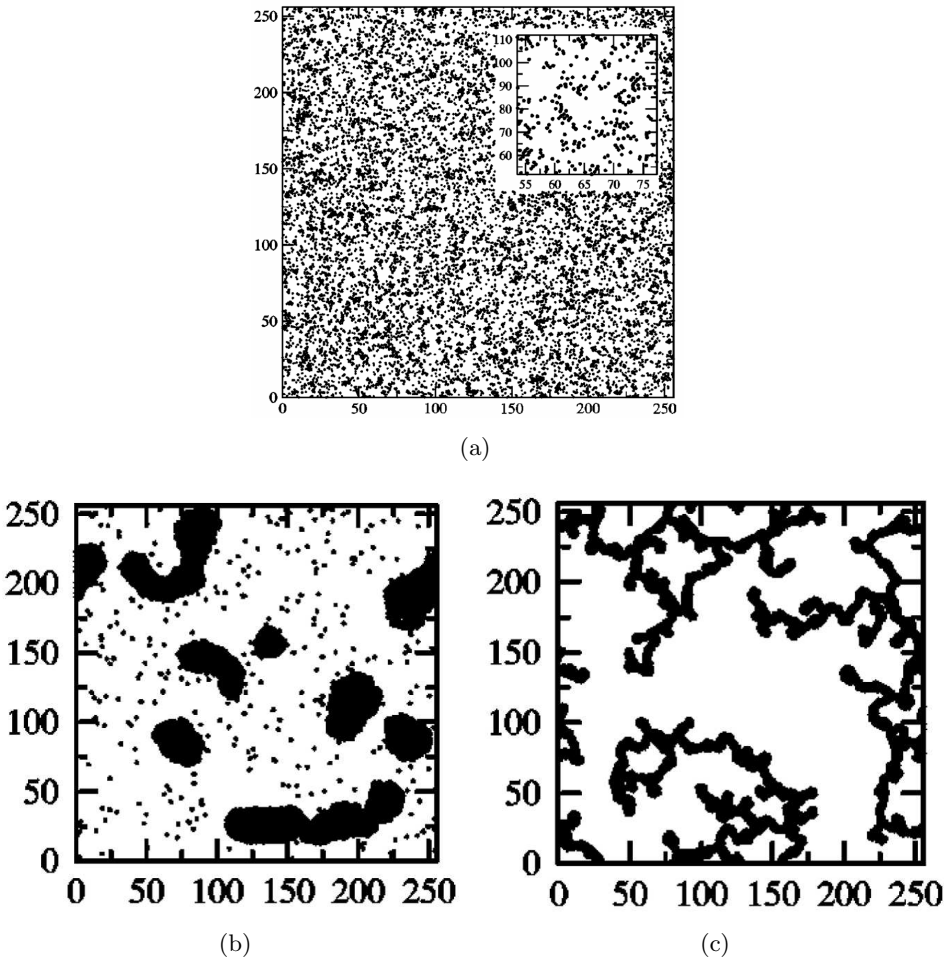



Figure 1.1: Figures taken from Cerda *et al.* [5] for different aggregating regimes: (a) system at equilibrium with dimensionless potential well–depth $\hat{U} = U/k_BT = 3.125$; (b) non–equilibrium aggregation with dimensionless potential well–depth $\hat{U} = U/k_BT = 4.0$; (c) gelation with dimensionless potential well–depth $\hat{U} = U/k_BT = 7.0$.

of external shear flow, and thermal energy determines whether a system will evolve reversibly or irreversibly as shown, for example, on Figure 1.1. These different regimes are observed by varying the strength of interparticle interaction related to thermal energy ($\hat{U} = U/k_BT$). When interparticle interaction is weak relative to thermal energy and is lesser than some critical value then the formation of only small aggregates is observed (Figure 1.1a). As interparticle interaction become stronger, larger aggregates are formed (Figure 1.1b,c).

Application of external shear force to the nanoparticle aggregation system is important because in large–scale reactors, the flow is turbulent and aggregating nanoparticles will be subjected to time varying shear flow at the Kolmogorov scale. In such a system the wide range of time and length scales are present. The wide range of time scales is introduced by the presence of short-time Brownian motion and the long-time hydrodynamic behavior of solvent. The wide range of length scales occurs due to the size separation of clusters of colloidal nanoparticles and solvent molecules. However, because of the scale separation between nanoparticle clusters and Kolmogorov scale, for a first approximation the flow can be treated as locally uniform time-varying shear. Once formed, aggregates do not break apart, and to introduce breakage some external forces such as shear flow must be introduced into the system. When the thermal energy and kinetic energy associated with external shear flow is able to overcome the interparticle interactions then a reversible change is expected. If interparticle attraction dominates, then an irreversible change is expected. In sheared colloidal systems, aggregation may occur due to particle–cluster (monomer addition) and cluster–cluster aggregation. The breakage and restructuring of these clusters is promoted by shear flow, and all these processes are related to irreversible changes. Recent experiments observe restructuring of clusters in the presence of external shear force [6]. However, there is no complete explanation of such behavior, which is a good reason to use computational approach for such a case.

Before discussing the characterization of aggregation outcomes, it is useful to clarify some terminology specific to aggregation. Colloidal aggregation is sometimes classified as reversible or irreversible depending on the system's characteristics (Figure 1.1a and Figure 1.1b,c correspondingly). However, the thermodynamic definition of a reversible process and reversible

aggregation phenomenon are different. Thermodynamically a reversible change is one that is performed quasi-statically such that the system remains infinitesimally close to thermodynamic equilibrium. Such changes can always be reversed and the system brought back to its original thermodynamic state without causing any changes in the thermodynamic state of the universe [7]. But when we are talking about a reversible aggregation process we mean a system at non-equilibrium steady state (NESS). In aggregating systems, once aggregation starts it continues irreversibly. Reversibility (due to aggregate breakage) may occur only through shear flow or increase in thermal energy. Thus, we can conclude that aggregating systems cannot be in thermodynamic equilibrium, instead, aggregating systems are in non–equilibrium steady states (NESS). An example of aggregating irreversible system is represented on Figure 1.1b from simulations performed by Cerda *et al.* [5] for 2D systems where large and dense aggregates are formed together with the presence of single particles. After an irreversible change the system cannot be brought back to its original thermodynamic state without causing a change in the thermodynamic state of the universe [7]. Based on this definition we can conclude that an irreversible colloidal nanoparticle aggregation leads to such non–equilibrium steady state as gelation which is a first order phase transition (as in a first order phase transition a system either absorbs or releases a fixed amount of energy). An aggregate structure that corresponds to the gelation stage is represented on Figure 1.1c from Cerda *et al.* [5] work. In this case large aggregates with a ramified structure are formed that occupy all the system's volume. At this stage no single particles are observed.

There also have been efforts to classify this phase–change behavior of aggregating systems on a phase diagram. Anderson and Lekkerkerker [8] described all these regimes with the phase diagram for the colloid–polymer systems. In these systems polymer is added to colloidal systems to produce an attraction between the particles. By varying the relative size of polymer and the colloid; the polymer concentration and colloid volume fraction the range of particle–particle interaction can be tuned and a variety of phase diagrams can be realized. Anderson and Lekkerkerker [8] reported that the aggregating outcome depends on the initial conditions and slight change in one condition may significantly change the outcome. They conclude that it

is difficult to reliably predict the transition mechanisms of colloid and colloid–polymer systems.

The processes of aggregate formation and aggregate breakage have been investigated from an experimental and computational perspective [9, 10, 11, 12]. Researchers agree that colloidal particle aggregating phenomena is very complex and multiscale problem where aggregation outcome depends significantly on the initial conditions. *However*, a single unified aggregation map that would determine different aggregating outcomes based on the initial parameters of the sheared aggregating system is not available. Such an aggregation map would be very useful when designing efficient turbulent reactors used for synthesis of the particles with good size control of product.

In principle, such an aggregation regime map could be generated based on a purely theoretical description of aggregation; or using experimental approach; or using a computation approach. Colloidal particles aggregation phenomena is not completely described yet therefore pure theoretical approach for describing aggregating phenomena is not appropriate. Experimental approach allows to measure aggregation in real systems. However, it is not feasible to control and measure all the parameters that determine aggregation phenomena. Therefore, in this dissertation a computational approach is adopted to develop a fundamental understanding of colloidal aggregation.

The focus area of this dissertation with respect to the work of other researchers is shown in Figure 1.2. The system complexity axis on this map represents model approaches used to study aggregation processes beginning from the simplest model LJ systems to more detailed and complicated systems such as protein molecules. The solute–concentration axis represents the range of solute densities, while the shear axis represents increase in shear flow intensity in the system. Dark–gray areas represent work of other researchers, such as Hobbie [13] who had performed experimental studies of depletion–driven phase separation for dilute polystyrene spheres. Aggregation processes under shear flow for dilute latex nanoparticles were studied by Chakrabarti, Sorensen, *et al.* [14]. Aggregation in systems with dense polymeric spherical nanoparticles are performed by Lekkerkerker [9] as well as Shepherd [10] for systems with and without shear. On this map the focus of the present work is represented with light–gray

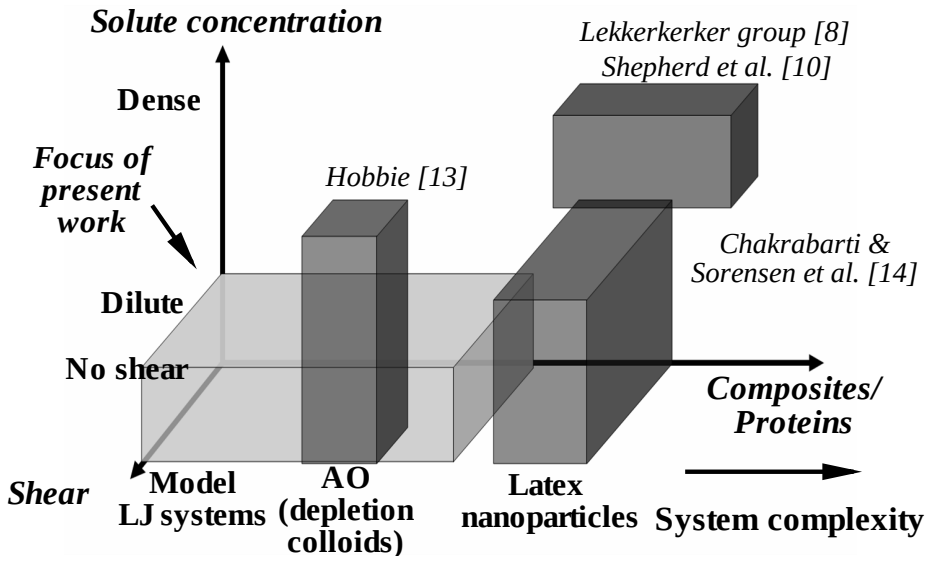



Figure 1.2: The focus area of this dissertation relative to other work on colloidal aggregating system under shear.

area which represents dilute systems of solute particles in solvent, with and without shear flow, for LJ model systems and depletion colloids. Since this work is a computational study, corresponding model systems are used instead of physical ones due to feasibility limitations of numerical methods. Therefore, the focus of this dissertation is on simple LJ model systems and depletion colloids. Dilute systems are chosen to compare our aggregating results with experimental results obtained by Mokhtari *et al.* [14] for latex nanoparticles. And shear flow is applied to aggregating systems for the reasons given before.

A fundamental understanding of changes in aggregate structure due to presence of the external shear flow is required to correctly describe aggregation growth and breakage processes. Therefore, an efficient numerical model that would accurately predict aggregation phenomenon in colloidal nanoparticles systems must be chosen.

Currently, the following simulation approaches for aggregation are commonly used [11], [15]-[18]

1. Molecular dynamics (MD), which is a microscale method (described in Chapter 2).

2. Mesoscale methods, such as Langevin dynamics (LD) and Brownian dynamics (BD)

(described in Chapter 2), stochastic rotational dynamics (SRD), and dissipative particle dynamics (DPD).

3. Monte Carlo methods, such as lattice Monte Carlo (LMC) method and off-lattice Monte Carlo (OLMC) method.

However, requirement for a significant aggregation statistics leads to consideration of large simulation systems that can significantly increase computational costs and decrease simulation efficiency. Ideally the model which is chosen to predict sheared colloidal nanoparticle aggregation should accurately describe physico-chemical interactions of relatively large physical systems, and at the same time, simulate at a low computational cost. In reality this is hard to achieve. In many cases if the model is very accurate it is usually not efficient and cannot be used to simulate a physical problem. On the other hand, more efficient models usually are not very accurate in terms of representing the physics, thereby limiting their applicability. Thus, a computational model which is chosen to predict a sheared aggregation of colloidal nanoparticles should maintain the balance between the level of accuracy and computational efficiency.

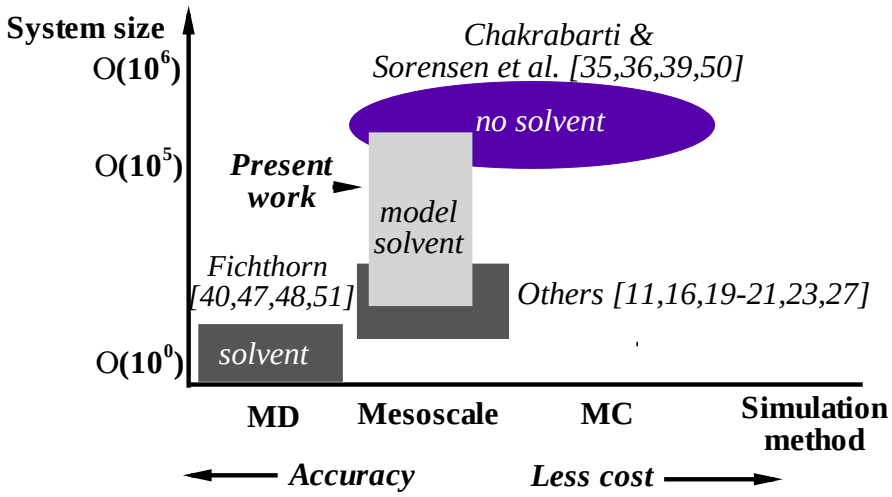



Figure 1.3: Dependence of the maximum system size versus simulation method.

To study sheared aggregation we focus on coarse-grained (or mesoscale) simulation methods such as a Langevin dynamics. These methods are computationally efficient when compared with microscale methods such as molecular dynamics, and they have the ability to accurately represent the aggregate structure when compared with the Monte Carlo methods. The hierarchy of these methods is represented on Figure 1.3. From this Figure we can observe decrease in the maximum system size (represents the number of solute particles in a system) for models with more detailed solvent representation such as MD. And increase in the system size when solvent effect is removed (Monte Carlo methods). The mesoscale methods still include solvent effect through solvent modeling that allows to decrease computational cost and increase system size to get good aggregation statistics.

As we show in Chapter 2, using MD to simulate aggregation phenomenon for realistic systems is too expensive. At the same time the off-lattice Monte Carlo (OLMC) simulation has limitations in simulating aggregate restructuring, because it is not capable of representing restructuring of the cluster after the cluster is formed. On the other hand, the mesoscale methods such as BD, LD, DPD, and SRD have the promise of low cost and accurate representation of aggregation structure on today's computers, but it applicability for simulating non-equilibrium systems should be established. The development of LD and BD methods for solving sheared aggregation problems requires consideration of the following points:

1. The current LD and BD models are not adequate for aggregation.

2. Numerical accuracy and approach not well established in context of aggregation. This leads to Chapter 2.

3. Model accuracy of LD and BD is not satisfactory for aggregating systems when compared with established MD approach. This motivates the need for improved BD model with potential mean force (PMF) that accounts for solvent interaction in non-equilibrium aggregating systems which leads to Chapter 3.

4. The minimum set of characteristics and metrics required for complete description of sheared aggregation phenomenon is not established and the correspondent aggregation

map is not defined. This leads to development of new characteristics and metrics fully described in Chapter 4.

Also accurate modeling of physico-chemical interactions is required. This can be achieved by developing the coarse–grained particle interaction potentials derived from quantum mechanics calculations, that are suitable for large scale nanoparticle aggregation simulations. In this case, the atomic models for surface molecules of polystyrene nanoparticles can be developed to calculate surface-molecule, surface-surface, and molecule-molecule interaction forces. These results can be validated by atomic force microscopy (AFM) measurements of polystyrene-polystyrene nanoparticles, and used in simulation of nanoparticles aggregation as a physical potential.

## 1.1 Simulation Approaches

In this subsection the various approaches used to simulate nanoparticle aggregation are briefly reviewed and their advantages and disadvantages are considered.

### 1.1.1 Molecular Dynamics

Molecular dynamics (MD) simulation is an established technique that can simulate colloidal nanoparticle aggregation [15]. In the MD approach, solute and solvent particles interact through a modeled, intermolecular potential, and the positions and velocities of these particles evolve in time according to Newton's equations of motion. In most MD simulations, the intermolecular potential energy is taken to be the sum of isolated pair interactions, which is called the *pairwise additivity* assumption. The main difficulty with such an approach is that it cannot be used to model aggregation of a realistic system of colloidal nanoparticles. The requirements of large size separation between nanoparticles and solvent molecules ($d_{NP} \sim 40$ nm, $d_{solv} \sim 0.3$ nm and $d_{NP}/d_{solv} \sim 100$ at solvent molecules volume fraction $\lambda_{solv} \sim 0.45$ and very low nanoparticle volume fraction $\lambda_{solute} \sim 0.005$), and the large number of nanoparticles that are modeled to have good statistics of aggregated clusters lead to an enormous number of solvent molecules in the system (on the order of $10^{10}$). Moreover, calculation of intermolecular

forces between solvent molecules in MD would require resolving time scales on the order of $\hat{\tau}_F = 0.125$. However, the time scale of evolution of cluster statistics is much larger and is on the order of $\hat{\tau}_{cl} \sim 40,000.00$. Therefore, simulating any colloidal system even far from realistic physical parameters is a challenging and sometimes even impossible task. Alternative approaches are needed to resolve this problem. One alternative is to use Monte Carlo approaches for nanoparticle aggregation simulation.

### 1.1.2 Monte Carlo Simulation Algorithms

Based on the off-lattice Monte Carlo (OLMC) simulation, several methods are frequently used to model nanoparticle aggregation. These models include diffusion-limited aggregation (DLA), diffusion-limited cluster aggregation (DLCA), ballistic-limited aggregation (BLA), ballistic-limited cluster aggregation (BLCA), reaction-limited aggregation (RLA), and reaction-limited cluster aggregation (RLCA).

In DLA models, particles diffuse through a random-walk from distant points and finally arrive and stick to the surface of the growing aggregate [16, 19]. In the DLCA model, the particles and clusters move in random-walk trajectories, which represent the Brownian motion of the particles and clusters in a dense fluid [20, 21]. According to this model, particles and aggregates are moved randomly, and when the distance between centers of two particles approach "cluster distance" $r_{cl}$ (the maximum distance between two neighbor particles which belong to the same cluster) they irreversibly link. After this linking if the distance between any pair of particles in two different clusters appears to be less than $r_{cl}$, two clusters move apart along their approach path until the separation is equal to the cluster distance. Thus, stickiness probability $p_{stick}$ for these two approaches is unity. DLCA is the more appropriate model when simulating colloidal aggregation because in reality, aggregates grow not only due to cluster-monomer interaction but also due to the cluster-cluster interaction. Both DLA and DLCA models allow simulating the aggregation of systems with more than a million nanoparticles, which gives good statistics of aggregates. However, these approaches can only be applied if the interparticle interactions are smaller than $k_B T_{ref}$, where $k_B$ is the Boltzmann constant and

$T_{ref}$ is the reference temperature. If the interactions are large compared with $k_B T_{ref}$ for some length scale $l$, the structure of the resulting aggregates will be changed over this length scale, and fractal dimensionality may be changed if the interactions are sufficiently strong [22].

In BLA models, particles and clusters are added to each other through linear paths. Each path is chosen randomly from all possible paths that could result in a collision between the particle and cluster [23]. Similarly, the BLCA approach models cluster-cluster collisions in addition to the particle-cluster collision used in BLA [24]. BLA and BLCA approaches were developed in the 1960s. These models were used because at that time it was not sufficiently efficient to carry out simulations with random walk trajectories which are implemented in DLA and DLCA models. The use of BLA and BLCA approaches for colloidal nanoparticle aggregation modeling is limited due to the assumption of linear paths between collisions, which can lead to incorrect cluster size distributions.

In RLA and RLCA models, particles and clusters (just particles in the RLA case) follow random walk trajectories, but they do not form a new cluster each time they come into contact [25]. Instead, they continue their random walk paths and many collisions are usually required before a pair of clusters will join. This behavior is dictated by the presence of a repulsive barrier in the particle-particle pair potential. Only when this barrier is overcome will the short-range attractive force finally hold two clusters together. This process is identical to DLA and DLCA models with a small sticking probability $p_{stick}$. A disadvantage of this method is the very large amount of computer time required if the sticking probability is small.

All the models represented above (DLA, DLCA, BLA, BLCA, RLA, RLCA) help to characterize complexity of aggregate structures by extracting important characteristics, such as the size distribution of aggregating clusters, dimensionality of the cluster structure, and local concentration of the particles in the system in very efficient way. However, there are several features which do not allow to obtain correct results from simulations of the aggregation of colloidal nanoparticles:

(i) All these models imply an irreversible linking of particle-cluster and cluster-cluster when new clusters are formed. This prohibits rearrangement of nanoparticles within a cluster

as well as dissociation of one cluster into two or more. This will bring to incorrect results in aggregation structure, especially when nanoparticle aggregation is simulated with shear force. In this case, depending on the magnitude of the shear force the aggregates structure can be rearranged which significantly changes fractal dimension $D_f$ of aggregates.

(ii) The structure of clusters obtained with these models depends on model parameter such as cluster distance $r_{cl}$, which makes it impossible to implement any of these models for the dynamic simulation of the colloidal nanoparticle aggregation.

(iii) With these models it is impossible to use interaction potentials which can be derived from quantum mechanics calculations for a realistic physical system.

To overcome these problems and keep simulation efficiency it is also proposed to use off-lattice Dynamic Monte Carlo simulation (DMC) for 1-d case [26]. However, to extent this approach to 2-d and 3-d cases will require a number of enhancements, the most significant of which are the inclusion of rotational motion, intra-cluster relaxation and reactions and improved calculation of reaction rate constants.

### 1.1.3 Mesoscale Methods

Another alternative to MD simulations are mesoscale models. There is a wide variety of mesoscale simulation techniques to model the dynamics of colloidal suspensions. Among these techniques the Langevin dynamics (LD) and Brownian dynamics (BD) techniques are the most frequently used to simulate diffusion problems. On the other hand such techniques as dissipative particle dynamics (DPD), and stochastic rotational dynamics (SRD) are most advanced one.

*Dissipative Particle Dynamics*

Dissipative particle dynamics [17, 27] is an alternative to standard MD techniques and includes hydrodynamic and Brownian fluctuations. In DPD, fluid molecules themselves are not represented, but instead, groups of molecules called dissipative particles are considered. These dissipative particles are simulated to obtain a flow field. These dissipative particles

interact with each other dissipatively, exchange momentum, and move randomly like Brownian particles. The main advantage of DPD is that in this method the multibody hydrodynamic interactions among colloidal particles are automatically reproduced through the interactions with dissipative particles. However, even though the number of dissipative particles in DPD is less than in MD, this method is still computationally expensive, because the dissipative particles still interact through the pairwise potentials and the number of colloid nanoparticles $N_p$ is much smaller than the number of dissipative particles $N_{dp}$.

*Stochastic Rotational Dynamics*

In the SRD approach [11], all the space in SRD is partitioned using a rectangular grid. Nanoparticles and solvent particles move in continuous space according to Newton's laws of motion, excluding solvent-solvent interaction. This excluded interaction is modeled by collision events at discrete times called collision time steps. At these collision events, solvent particles inside each cell exchange momentum by rotating their velocity vector relative to the center of mass velocity of the cell around a randomly chosen axis. This method is more efficient than MD because there is no direct computation of solvent-solvent interactions; but instead this solvent-solvent collision is simulated at every collision step which is much greater than the computational time step. This method successfully models aggregation of colloidal nanoparticles [28].

In the LD/BD approaches, [18] it is assumed that the collisions of colloidal nanoparticles with solvent molecules induce their random displacement. As a result, the positions and velocities of the colloidal nanoparticles change accordingly. In LD and BD the local momentum is not conserved; however, it satisfies the average momentum conservation (ensemble average). This approach is very attractive for simulation of colloidal nanoparticle aggregation due to its simplicity and efficiency [29]. In the present work Brownian dynamics method will be described and analyzed in detail with respect to the aggregation of colloidal nanoparticles.

## 1.2 Research Objectives

A need to develop a numerical method for sheared aggregating system that would be able to predict aggregation outcome accurately and efficiently was identified earlier. To infer physics and get complete understanding of aggregating phenomenon as well as characterize aggregating structure the following questions should be answered:

1. How does a colloidal system aggregate in the absence of shear?
2. What are the characteristics that control aggregation phenomenon? What are the metrics that characterize aggregate structures?

Careful description of aggregation phenomenon for non–sheared systems allows to create the "reference systems" before introducing the shear flow. The shear flow when introduced into aggregating system not only introduce additional time and length scales but also may cause aggregate restructure and change in aggregates spatial configuration. The need for an aggregation map and a better understanding of the physical mechanisms in sheared aggregating systems promotes the following questions:

1. What is the aggregation regime map for non–sheared and sheared aggregating systems that would predict aggregate outcome for different aggregating regimes?
2. What are the sources for aggregate restructuring when shear flow is applied? When shear flow is applied how does the aggregate structure change?
3. What are the characteristics and metrics that control and describe aggregation outcome for sheared aggregating systems? Are these different than those used for non–sheared aggregating systems?

When using the proposed Langevin dynamics (LD) method, which is one of the mesoscale methods, for modeling aggregating phenomenon some of the questions with regard to LD applicability to solve this kind of problems should be answered:

1. Is it appropriate to use a Lennard–Jones potential as the mean–field potential in LD model for reproducing aggregation phenomenon? Which coarse–graining method should

we use to extract the mean–field potential in aggregating system which is not in thermal equilibrium?

2. Is LD model a valid approach for representing aggregation phenomenon? Is this approach accurate when reproducing aggregate structure? How to verify LD method with respect to the aggregation phenomenon under shear flow?

Answering to all these questions allows complete description of the aggregation phenomenon of colloidal particles simulated with mesoscale LD model.

In summary, the principal objectives of this study are to:

1. Gain a better fundamental understanding of aggregation in sheared and non–sheared model systems.

2. Delineate aggregation outcomes for sheared and non–sheared cases in regime maps that identify the key dimensionless parameters that determine aggregation.

3. Characterize and understand the role of aggregating structure in sheared and non–sheared systems using mesoscale simulation method.

4. Establish the requirements for a numerically accurate mesoscale simulation method.

5. Develop a coarse–graining procedure to infer mesoscale interaction potentials for atomic interaction for aggregating systems.

## 1.3 Original Contributions of this Dissertation

1. Developed semi–analytical coarse–graining approach to infer mesoscale interaction potentials in aggregating systems to calculate the relative acceleration between solute particles in a solvent bath and to reproduce the pair correlation function and cluster size distribution in the RLA and DLA regimes. Developed computational algorithms and code to implement this approach.

2. Established guidelines for numerical accuracy of simulations of aggregation problems using mesoscale approaches such as LD and BD. Characterized numerical convergence and accuracy of LD and BD in simple test problems.

3. Gained a fundamental understanding of aggregation in non–sheared systems by explaining the dependence of aggregate structure on interaction potential.

4. Characterized aggregation outcomes of LD and BD simulations of non–sheared systems on a $\hat{D} - \hat{\varepsilon}$ plane, and showed that dimensionless well–depth of the interaction potential controls the aggregation outcome, while the dimensionless diffusion only controls the rate at which this outcome is reached.

5. Explained the scaling of aggregate size with dimensionless shear rate through a simple mechanistic model.

6. Performed a budget analysis of the energy balance equation for LD simulations of sheared and non–sheared aggregating systems. In the non–sheared case at steady state the principal balance is between dissipation and random (Wiener) terms, while the force–velocity correlation term is negligible. In sheared systems, at steady state, the force–velocity correlation is large compared to the non–sheared case and contributes significantly to the principal balance with dissipation and random terms. The unsteady evolution of energy in velocity fluctuation shows that the production (position–velocity correlation) of velocity fluctuation arising from position–velocity correlation provides a mechanism to transfer the energy from mean energy to the velocity fluctuating energy. The initial rapid increase in velocity fluctuation results in high dissipation and transfer of this energy into force–velocity autocorrelation term that is a signature of aggregate restructuring.

7. Characterized aggregate anisotropy and restructuring in sheared systems. Aggregate anisotropy is characterized by anisotropy coefficient and restructuring by LPED (local volumetric potential energy density).

8. Identified the relevant dimensionless parameter (a dimensionless force ratio ) that char-

acterizes aggregation outcomes of LD simulations of sheared systems in the $Pe-\hat{\varepsilon}$ plane. Aggregate size and structure are characterized by specifying the maximum radius of gyration, fractal dimension, and LPED that allows the delineation of three distinct aggregation regimes on the $Pe-\hat{\varepsilon}$ map: (a) non–aggregating regime, (b) aggregating regime with less dense local structure, (c) aggregating regime with compact local structure.

## 1.4 Outline of the Dissertation

The dissertation is structured as following: convergence of mesoscale BD/LD method and its simulation accuracy is outlined in Chapter 2. Also, in this chapter characteristics and metrics that control and characterize aggregate structure for non–sheared aggregating systems are determined and used to create regime aggregating map for non–sheared systems. Evidence of inapplicability of LJ potential in BD/LD models for reproducing aggregation phenomenon in reaction limited regime is presented in Chapter 3. Importance of the relative acceleration approach for developing a new method to derive the mean–field potential is described in the same chapter. A new improved BD/LD model is proposed that allow to improve aggregating statistics for non–sheared systems from diffusion limited to reaction limited regimes. Chapter 4 deals with sheared aggregating systems in the context of complete description of the effect of shear on aggregate structure/outcome. A new method for the kinetic energy analysis is proposed that allows to determine the source for the aggregate restructuring when shear is applied. Effect of shear on global and local structure of aggregates and on the maximum size of aggregates is outlined in the same chapter. The aggregating map is introduced that allows to predict the aggregating outcome/structure based on the initial characteristics. Chapter 5 presents the conclusions of this work and some ideas on future work.

# CHAPTER 2. ON BROWNIAN DYNAMICS SIMULATION OF AGGREGATION

A paper is published in *Industrial and Engineering Chemistry Research*

Sergiy Markutsya, Shankar Subramaniam, R. Dennis Vigil, Rodney O. Fox

## 2.1 Abstract

Accurate simulation and control of nanoparticle aggregation in chemical reactors requires that population balance equations be solved by using realistic expressions for aggregation and breakage rate kernels. Obtaining such expressions requires that atomistic simulation approaches that can account for microscopic details of particle collisions be used. In principle, molecular dynamics simulations can provide the needed microscopic information, but because of the separation in length scales between the aggregates and solvent molecules, such simulations are too costly. Brownian dynamics simulations provide an alternative to the molecular dynamics approach for simulation of particle aggregation, but there has been no systematic attempt to validate the Brownian dynamics method for this class of problems. In this work we attempt to develop a better understanding of Brownian dynamics simulations of aggregation by (1) developing convergence criteria, (2) determining criteria for aggregation to occur in BD simulations using dimensionless variables, and (3) directly comparing BD and MD simulation predictions for a model aggregation problem.

## 2.2 Introduction

In recent years there has been an explosion of interest in the synthesis of nanoparticles because they serve as building blocks for materials with novel mechanical, optical, electric,

magnetic, thermal, chemical, and biological properties.[1] Consequently, the ability to predict and control nanoparticle aggregation in reactors used to synthesize these particles is of prime importance.[2] The traditional approach to modeling colloidal particle aggregation at the reactor scale is to employ mean-field rate equations, also known as population balance equations (PBEs). For example, the case of irreversible aggregation in a well-stirred batch reactor can be represented by the much-studied discrete PBE

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k} K_{ij}c_i c_j - c_k \sum_{i=1}^{\infty} K_{ki}c_i, \tag{2.1}$$

where $c_k$ is the concentration of particles with mass $k$ and $K_{ij}$ is a symmetric matrix of rate constants describing the aggregation of particles with masses $i$ and $j$. This PBE can also be formulated in continuous form, and it has been elaborated to include mechanisms such as nucleation, growth, breakage, and feed and removal. More recently, with the introduction of the direct quadrature method of moments,[3] multivariate forms of the PBE have received increased attention, corresponding to an increased interest in predicting and controlling not only the particle size distribution but also particle morphology.

In order to solve (either analytically or numerically) equations of the type (2.1), the functional form of the aggregation kernel, which depends upon particle transport mechanisms and microscopic details of the particle collision events, must be specified. To this end, the aggregation kernel is often decomposed into the product of a collision efficiency, $0 < \alpha_{ij} \leq 1$, and a collision frequency function, $\beta_{ij}$, such that $K_{ij} = \alpha_{ij}\beta_{ij}$. Approximate expressions for the collision frequency function, $\beta_{ij}$, have been derived for certain limiting cases, such as when the motion of the aggregates can be considered to be Brownian[4] (particle sizes smaller than the characteristic shear gradients) or for the instance in which particles are large relative to shear gradients but smaller than the Kolmogorov micro-scale.[5] The derivation of these expressions, however, requires the invocation of a number of ad-hoc assumptions, such as the neglect of long range particle-particle interactions and the assumption that all aggregates are spherical. Although the latter assumption can be relaxed so that particles have an arbitrary fractal dimension, $d_f$, it is still necessary to invoke assumptions concerning the mobility (both translational and rotational) of fractal aggregates. Derivation of an analytical expression for

the collision efficiency $\alpha_{ij}$ is even more problematic, since the probability that an aggregation event occurs upon collision of particles of sizes $i$ and $j$ can in general be expected to depend upon many microscopic details including the strength of particle-particle forces, and the morphology, angle of approach, and momenta of the colliding particles. Of course $\alpha_{ij}$ is averaged over these microscopic collision variables so that it depends explicitly only upon measurable bulk properties and on the particle size variables, $i$ and $j$, but in order to perform the required averaging over the microscopic collision variables, an atomistic simulation approach must be used that can generate the relevant particle configuration ensembles.

Atomistic simulation methods such as molecular dynamics (MD) can in principle provide the detailed information concerning collision, aggregation, and breakage events that is needed to derive realistic expressions for aggregation (and breakage) rate kernels, because they explicitly represent all molecules in the system (both solute and solvent) and compute the motion of these molecules using classical Newtonian mechanics.[6] However, in order to carry out such simulations, information is required concerning the interaction forces between all of the constituent molecules. Usually, these forces are assumed to be pairwise additive so that it is only necessary to define force laws between each type of molecule (e.g. solute-solute, solute-solvent, and solvent-solvent). Typically these forces are obtained by differentiating presumed intermolecular potential energy functions (such as the well-known Lennard-Jones potential) fitted to experimental data. These presumed potential energy functions mimic the competition between near-range repulsions arising from the overlap of electronic shells and long-range attractive Van der Waals forces. Hence, interaction potentials typically display a potential energy minima at intermediate distances that arises from the balance of the longer-range attractive forces and short range repulsive forces. More recently, there have been efforts to avoid the use of presumed interaction potentials by instead using coarse-graining procedures to compute these interaction potentials using information obtained from quantum mechanical calculations.[7, 8]

Even when accurate pairwise interaction potentials are available, however, other problems with using the MD approach for simulation of aggregation remain. In particular, the sep-

aration in scales between the sizes of the solvent molecules (typically $10^{-10} - 10^{-9}$ m) and nanoparticle aggregates (usually $10^{-8} - 10^{-7}$ m) requires that an enormous number of solvent molecules be simulated, especially for dilute systems. For example, consider Figure 2.1, which shows the CPU time required for each simulation time step as a function of the total number of molecules (solute and solvent) simulated using the MD simulation software LAMMPS[9]. Results for two sets of MD simulations are shown, each carried out under identical conditions

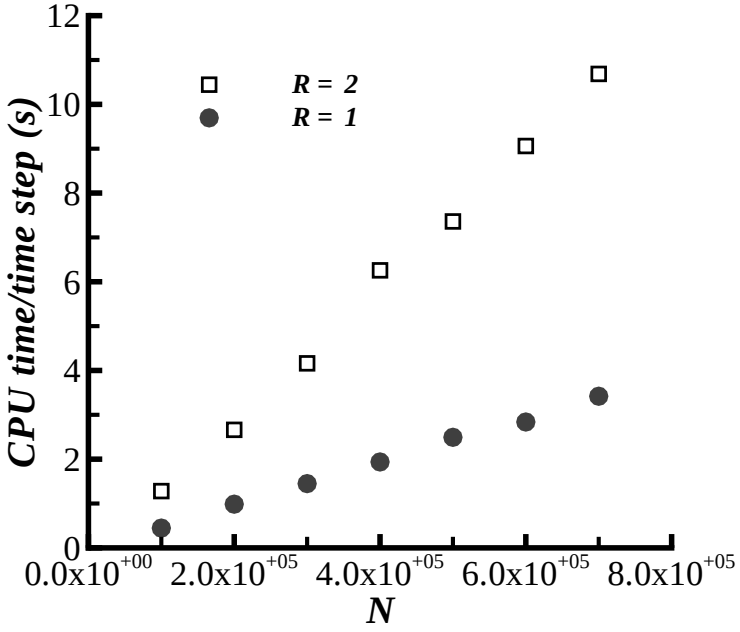



Figure 2.1: Dependence of MD simulation CPU time on the number of Lennard-Jones particles, $N$, and the solute/solvent diameter ratio, $R$ for non-aggregating particles. All other simulation parameters are identical in the two sets of simulations. Solute volume fraction was chosen to be 0.01.

except for the solute/solvent diameter ratio used (equal solute and solvent sizes in one case, solute diameter twice that of the solvent in the other case). It is readily apparent that the CPU time scales approximately linearly with the number of molecules, but that the CPU time grows more rapidly with increasing solute/solvent size ratio, since as size ratio increases the number of solvent molecules involved in solute-solvent interaction increase. In view of the fact that

realistic simulations would require solute/solvent size ratios on the order of at least 10-100, it is evident that MD simulation of aggregation, even when using nano-scale primary particles, is computationally demanding. Furthermore, the dynamic range of the largest aggregates to the primary nanoparticles can itself be two to three orders of magnitude in light-scattering experiments.[10] Hence, to obtain a meaningful statistical distribution of aggregates, it is clear that very large systems will need to be simulated. All these factors contribute to the conclusion that MD simulation of aggregation with existing simulation packages and hardware is computationally prohibitive.

In order to circumvent the computational limitations that result from the large number of solvent molecules required in MD simulations of nanoparticle aggregation, the Brownian dynamics (BD) approach can be used. In this method, the solute-solvent interactions are incorporated into Langevin equations for solute particles, and therefore there is no need to track solvent molecules explicitly. For example, in an isotropic system if particles are sufficiently small so that they are unaffected by fluid shear, the $i$-th solute particle position, $\mathbf{r}_i$, and velocity, $\mathbf{v}_i$, can be described by

$$d\mathbf{r}_i = \mathbf{v}_i dt, \tag{2.2}$$

and

$$d\mathbf{v}_i = -\gamma \mathbf{v}_i dt + \frac{1}{m_i}\mathbf{F}(\{\mathbf{r}_i\})dt + \sqrt{2\gamma}\sigma_{v_\infty} d\mathbf{W}_i. \tag{2.3}$$

In the above equations, $m_i$ is the mass of particle $i$, $\gamma$ is the frictional coefficient, $\mathbf{F}(\{\mathbf{r}_i\})$ is the net force exerted on the $i$-th particle due to its interactions with all other particles, $\sigma^2_{v_\infty}$ is the equilibrium velocity variance ($= k_B T_\infty / m_i$), and $d\mathbf{W}_i$ is a Wiener process increment. For cases in which the relaxation time for the particle velocities $1/\gamma$ is short compared with the relaxation time for particle position (which includes most cases of practical interest for particles suspended in liquids), Equations 2.2 and 2.3 can be integrated so that only the following position Langevin (PL) equation must be evolved:[11]

$$d\mathbf{r}_i = \frac{\mathbf{F}(\{\mathbf{r}_i\})}{m_i \gamma} dt + \sigma_{v_\infty}\sqrt{\frac{2}{\gamma}} d\mathbf{W}_i. \tag{2.4}$$

The advantage of using BD simulations rather than MD simulations in terms of computational cost is evident in Table 2.1, which compares results for MD simulations (using LAMMPS)

and BD simulations (of the position equation 2.2 and the velocity Langevin equation 2.3 implemented in an in-house code) of identical systems with an order of magnitude separation in length scales between the solute particles and solvent molecules. In particular, a comparison

Table 2.1: Comparison of MD and position-velocity Langevin BD simulation time for 31 non-aggregating solute particles. All solvent and solute particle interactions were modeled using Lennard-Jones potentials with well depth $\varepsilon$ and particle radius $\sigma$. The time increment in both types of simulations was fixed at $5 \times 10^{-15}$ seconds.

| | Molecular Dynamics | | Brownian Dynamics | |
|---|---|---|---|---|
| | Solvent | Solute | Solvent | Solute |
| $\sigma$ (m) | $2.85 \times 10^{-10}$ | $4.0 \times 10^{-9}$ | N/A | $4.0 \times 10^{-9}$ |
| $m$ (kg) | $1.33 \times 10^{-26}$ | $3.686 \times 10^{-23}$ | N/A | $3.686 \times 10^{-23}$ |
| $\varepsilon$ (kg-m$^2$/s$^2$) | $1.073 \times 10^{-21}$ | $1.646 \times 10^{-20}$ | N/A | $1.646 \times 10^{-20}$ |
| N | 146,840 | 31 | N/A | 31 |
| time steps/CPU sec (1 processor) | 0.06 | | 100 | |
| Time for $10^7$ steps (20 processors) | 115 days | | 2 hrs | |

of the number of simulation time steps executed per second of CPU time demonstrates that there is more than three orders of magnitude speedup in the BD simulations as a result of the fact that individual solvent molecules are not simulated, and positions and velocities are calculated only for solute particles. This speedup is a necessity for simulating aggregation in colloidal systems, where the number of solute particles and the aggregate sizes are relatively large.

Although several investigators have employed the BD approach to simulate particle aggregation, [13, 14, 15, 16, 17, 18] we are not aware of any systematic effort to establish the legitimacy and accuracy of this approach with respect to aggregation. Furthermore, it has not been demonstrated that BD simulations of aggregation duplicate the predictions produced by corresponding MD simulations, nor is it understood in general how to establish correspondence between the two types of simulations. In order to address these issues, the following questions must be answered: (1) What are the minimal requirements for numerical convergence of BD simulations of aggregation? (2) Under what conditions is particle aggregation significant

in BD simulations? (3) How accurate are BD simulations of aggregation? and (4) How can model coefficients for BD simulations of aggregation be obtained from MD simulations or other methods? The remainder of this paper describes efforts to begin to address questions (1) - (3). Specifically, in section 2.3 we estimate requirements for convergence of BD simulations of aggregation by considering a simple model problem with a known analytical solution and by computing the deterministic and statistical contributions to the error. In section 2.4 we carry out a dimensional analysis in order to delineate regions in parameter space where significant aggregation occurs in BD simulations. The regions in parameter space where the PVL to PL reduction are admissible are also identified. In section 2.5 we consider a model problem for directly comparing predictions of MD and BD simulations of aggregation, and in the last section we conclude the paper by suggesting an approach for addressing question (4) above.

## 2.3 Convergence of Brownian Dynamics Simulations

The ultimate goal of performing BD or MD simulations is to extract statistics. For aggregating systems, these statistics are usually the cluster size distribution, or its moments. While the numerical convergence requirements of MD and BD simulations of equilibrium non-aggregating systems are reasonably well understood, the same is not true for aggregating systems. In order to gain an understanding of convergence criteria for BD simulations of aggregation phenomena, information is needed concerning how the error associated with evolving Eqs. 2.2 and 2.3 with finite integration step size ($\Delta t$) affects estimates of aggregation statistics.[1] It is also necessary to determine how these estimates are affected by the averaging procedure (for example by the use of multiple independent simulations or time averaging). Following the standard approach,[21] the error in any estimate can be decomposed into a deterministic and statistical part. The deterministic error is due to the finite integration step size, and it arises from the numerical approximations involved in integrating Eqs. 2.2 and 2.3. In contrast, the statistical error depends on the number of samples. It is important to note that in aggregating systems the number of samples is not the number of particles $N$, but is the number of independent realizations of the $N$-particle system.

Calculation of the deterministic and statistical components of the error associated with using BD simulations requires that a test problem with a known analytical solution be chosen. Any such test problem will by necessity be relatively simple, and we propose one such problem here that bears similarity to the classical Kramer's problem.[19] Presently, we consider the one-dimensional motion of a single particle immersed in a fluid in the absence of fluid shear and under the influence of the ramp-well potential depicted in Figure 2.2 and defined by:

$$U(x) = \begin{cases} \infty, & 0 < x < \sigma \\ -\varepsilon \dfrac{x - x_a}{\sigma - x_a}, & \sigma \leq x \leq x_a \\ 0, & x_a \leq x \leq L \\ \infty, & x > L \end{cases} \tag{2.5}$$

The systematic force in Eq. 2.3 can be found by differentiating the above expression so

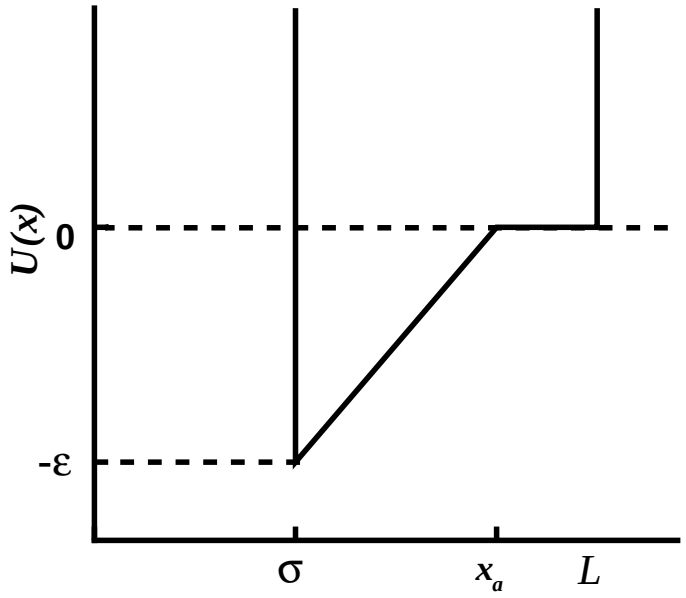



Figure 2.2: Illustration of the ramp-well potential.

that $F_x = -dU/dx$. Hence, the essential feature of the ramp well potential is that it produces a constant force of attraction, in contrast to the more commonly-used square well, which is

everywhere zero except at the boundaries of the well where the force is singular.

Klyatskin [20] has derived a separable analytical solution for the position-velocity probability density function of a multiparticle system evolving by the position-velocity Langevin equations. Here, we calculate the solution for the single particle position-velocity probability density function, $p(x,v)$, which can be decomposed into the product $p(x,v) = p_x p_v$ in the ramp-well test problem. The position probability distribution is given by

$$p_x = C_x \exp\Big(-\frac{U(x)}{k_B T_\infty}\Big). \tag{2.6}$$

The velocity probability $p_v$ is given by the Maxwell distribution function

$$p_v = \sqrt{\frac{m}{2\pi k_B T_\infty}} \exp\Big(-\frac{mv^2}{2k_B T_\infty}\Big). \tag{2.7}$$

The constant $C_x$ can be found by normalization, and it can subsequently be shown that the probability that the particle lies in the interval $\sigma < x < x_a$ is given by

$$p_a = \frac{1 - e^{-\hat{\varepsilon}}}{1 - e^{-\hat{\varepsilon}}\Big[1 + \dfrac{\hat{\varepsilon}(L - x_a)}{(\sigma - x_a)}\Big]}, \tag{2.8}$$

where $\hat{\varepsilon} = \varepsilon / k_B T_\infty$. The probability $p_a$ can in some sense be considered to be a "trapping" probability corresponding to the system being in an aggregated state. In comparing BD simulation predictions with the analytical solution given by Eq. 2.8, we have chosen the system parameters $(\sigma, \varepsilon, x_a, m, T)$ to satisfy two cases with $p_a = 0.74$, and $p_a = 0.90$. Brownian dynamics simulations were then carried out using one-dimensional versions of the position-velocity Langevin equations 2.2 and 2.3. In the simulations, initial particle positions were chosen randomly using a uniform distribution in the interval $\sigma < x < L$, and the initial velocity was chosen to be a Gaussian corresponding to $T_\infty$.

As was discussed above, the total error associated with the BD simulations arises from at least two sources. A deterministic error, $D_p$ is incurred due to the fact that a finite time step $\Delta t$ must be used to integrate Eqs. 2.2 and 2.3. Furthermore, as a consequence of the fact that only a finite number of samples $M$ can be computed, a statistical error $S_p$ is also incurred. Hence, the total error is given by $e = D_p + S_p$. In ergodic statistically stationary systems the

statistical error can be reduced either by averaging over longer times in a single simulation or by carrying out multiple independent simulations. However, in an aggregation-dominated system that produces a single volume spanning cluster, the system can become trapped in a gelled state and therefore may not sample the accessible states with the proper frequency. Hence, in order to develop convergence criteria that are applicable to general problems involving aggregation including those that produce gelled states, we carry out $M$ multiple independent simulations executed using fixed time step sizes, $\Delta t$. Each independent simulation was carried out for $2 \times 10^7$ time steps. For each independent simulation using time step $\Delta t$, the estimate for the probability $\{p\}_{\Delta t}$ that the particle resided in the interval $\sigma < x < x_a$ was computed. The ensemble average for $M$ such simulations using time step $\Delta t$ is denoted $\{p\}_{\Delta t,M}$. Therefore, the total error $e$ can be decomposed as follows:

$$\begin{aligned} e &= \{p\}_{\Delta t,M} - p_a \\ &= \{p\}_{\Delta t,M} - \{p\}_{\Delta t,\infty} + \{p\}_{\Delta t,\infty} - p_a \\ &= S_p + D_p \end{aligned} \tag{2.9}$$

In the above expression $\{p\}_{\Delta t,\infty}$ is the expected value of the trapping probability for an infinite number of independent simulations carried out using an integration time step $\Delta t$. In practice this quantity must be approximated by carrying out a finite but large number of simulations. We approximated $\{p\}_{\Delta t,\infty}$ by choosing $M = 1 \times 10^7$.

The deterministic error $D_p = \{p\}_{\Delta t,\infty} - p_a$ will depend upon the nature of the numerical integration scheme used,[21] and for example using a first order in time method one expects that $D_p \sim \Delta t$. We have verified this prediction and we find that $D_p$ can be kept below 0.06% for $\sigma_{v_\infty} \Delta t/\sigma \leq 0.004$, where $\sigma$ is the particle radius. Assuming that the errors for individual simulations are normally distributed, it can be expected that the statistical error $S_p$ obeys

$$S_p = \{p\}_{\Delta t,M} - \{p\}_{\Delta t,\infty} \sim \frac{1}{\sqrt{M}}. \tag{2.10}$$

Figure 2.3 demonstrates that this prediction is indeed fulfilled. Also it shows that even for high $p_a = 0.90$ the statistical error magnitude remains similar to that for $p_a = 0.74$. Therefore the statistical error of 30% for a single simulation requires that at least 100 simulations be

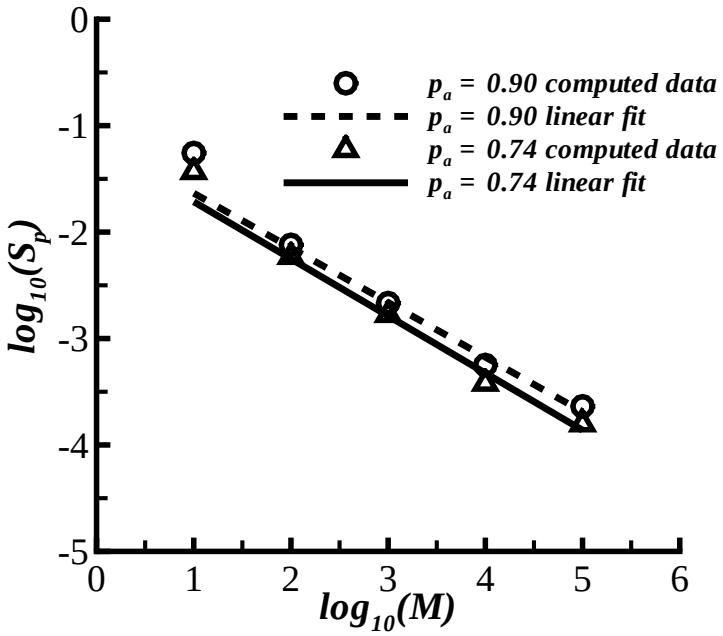



Figure 2.3: Dependence of statistical error, $S_p$ on number of independent simulations, $M$ for a time step $\sigma_{v_\infty}\Delta t/\sigma = 0.002$. The slope of the linear fit is -0.52 for $p_a = 0.9$, and is -0.54 for $p_a = 0.74$.

performed in order to reduce the statistical error so that it is comparable with the deterministic error, $D_p$. Moreover, because these results were obtained for a simple one-dimensional simulation with only a single particle, the number of simulations required to converge the predictions of BD simulations of the aggregation of a large number of particles may in many cases be prohibitive or may require the development of other methods for more rapidly reducing the statistical error. Furthermore, the rate of convergence will depend on the aggregation statistic that one seeks to extract from the BD simulations, with higher moments of the cluster size distribution converging more slowly. This analysis also demonstrates that calculations from a single BD simulation of an aggregation process are likely not converged statistically.

## 2.4 Aggregation Regime

In principle, given sufficient computational power and memory, converged BD simulations of aggregation can be executed if pairwise particle interaction potentials are known. However, it is not necessarily the case that significant aggregation will always occur in these systems, depending upon several system parameters. In this section we develop two important non-dimensional parameters and use them to characterize clustering outcomes and thereby delineate a criterion for aggregation to occur in BD simulations.

Table 2.2 lists several relevant characteristic scales for BD simulations of interacting particles in the absence of fluid shear (see Appendix A for details), and selecting from among these we identify the dimensionless reduced potential well depth and diffusivity,

$$\begin{aligned} \hat{\varepsilon} &= \frac{\varepsilon}{k_B T_\infty} \\ \hat{D}_\infty &= \frac{D_\infty}{\sigma}\sqrt{\frac{m}{k_B T_\infty}}. \end{aligned} \tag{2.11}$$

The particle volume fraction is also an important dimensionless parameter that is likely to influence clustering outcomes, but we will consider only cases of low particle loading ($< 1\%$ by volume) so that variations in this parameter can be neglected. The product $\hat{\varepsilon}\hat{D}_\infty$ can be interpreted as the ratio of the frictional and systematic force time scales.[1] Therefore if $\hat{\varepsilon}\hat{D}_\infty << 1$ (as is the case for nanoparticles suspended in liquids) there is sufficient separation in time scales such that the BD simulations can be carried out using a position-only Langevin scheme obtained by integrating Eq.2.3.[11]

In order to quantify the clustering of particles, we calculate the extent of aggregation, $0 \leq \xi < 1$, defined as

$$\xi = 1 - \frac{M_0(t)}{M_0(0)}. \tag{2.12}$$

where $M_0$ is the zeroth moment or total concentration of clusters. Hence $\xi$ is an aggregation progress variable that approaches unity as the system mass accumulates in a single cluster. Three-dimensional Brownian dynamics simulations were carried out using our in-house

[1] Although there are three terms in the BD velocity evolution equation 2.3, the coefficient of the noise term is related to that of the frictional term by the fluctuation-dissipation theorem. Therefore, there are only two independent timescales in that equation.

Table 2.2: Characteristic length, time, and velocity scales in BD simulations. The parameter $\varepsilon$ represents the intermolecular potential energy minimum, or well depth.

| Parameter | Dimension | Description |
|---|---|---|
| $\sigma$ | Length | Particle Size |
| $r_c$ | Length | Interaction Potential Cutoff Distance |
| $1/\gamma$ | Time | Velocity Relaxation Time |
| $\tau_F = \frac{\sigma}{\varepsilon}\sqrt{mk_BT_\infty}$ | Time | Force Time Scale |
| $\sqrt{\frac{k_BT_\infty}{m}}$ | Velocity | Velocity Standard Deviation |

BD code to evolve 10000 primary particles with random non-overlapping initial positions. Particle-particle interactions were modeled by Lennard-Jones potentials, and simulations were continued until the clustering index $\xi$ approached steady state. Other simulation details are provided in Table 2.3. Simulations were carried out for several fixed values of $\hat{\varepsilon}\hat{D_\infty}$, and the results are shown in Figure 2.4. It is evident that the extent of aggregation depends most sensitively on the value of the reduced interaction potential well depth, $\hat{\varepsilon}$, and in fact $\hat{\varepsilon} >\approx 2$ is a necessary condition for significant aggregation to occur. Hence, for sufficiently small values of $\hat{\varepsilon}$ corresponding to high temperatures or shallow interaction potential well depths, colliding particles have low probability of sticking because thermal fluctuations are large enough for the particles to overcome the potential energy barrier that otherwise would keep them together. Therefore $\hat{\varepsilon}$ controls how "sticky" the particles are and it must play a major role in determining the collision efficiency function, $\alpha_{ij}$. In contrast, $\xi$ is relatively insensitive to the value of the reduced diffusivity. This latter observation is consistent with the fact that the Gibbs stationary solution of the Fokker-Planck equation corresponding to Eqs. 2.2 and 2.3 yields a Boltzmann distribution of particle coordinates independent of diffusivity.

Table 2.3: Simulation parameters used to produce Figure 2.4. Particle interactions were modeled using Lennard-Jones potentials and simulations were carried out using the LAMMPS[9] software package.

| Parameter | Description | Value |
|---|---|---|
| $N$ | Number of Particles | 10000 |
| $f_v$ | Particle Volume Fraction | 0.005 |
| $\sigma$ | Particle Diameter | $3.4 \times 10^{-10}$ m. |
| $T$ | Temperature | 121 K |
| $\frac{\sigma\gamma}{\sigma_{v_\infty}}$ | Dimensionless Friction Coefficient | 1.31 |

## 2.5 Simulation Accuracy

Although the Brownian dynamics method has been used by many investigators to simulate aggregation processes, little consideration has been given to the accuracy of such simulations even if statistically converged results can be obtained. Here we use the word "accuracy" in reference to how well the BD simulation predictions of aggregation reproduce those obtained from corresponding MD simulations, since the BD technique is essentially a reduction of the MD method. Because this reduction is obtained by eliminating the explicit representation of solvent molecules and replacing solvent-solute interactions with a mathematical model consisting of a stochastic fluctuating force and a deterministic frictional term, any discrepancies between predictions of the two methods are likely due to breakdowns in the assumptions and approximations implicit in these terms.

The accuracy of BD simulations for dilute non-aggregating systems has previously been considered by Giro *et al*[22]. These investigators considered the situation in which the solute particles are identical to the solvent molecules, and they showed that the BD simulations closely reproduce the equilibrium solute-solute radial distribution function, $g(r)$. However, they also

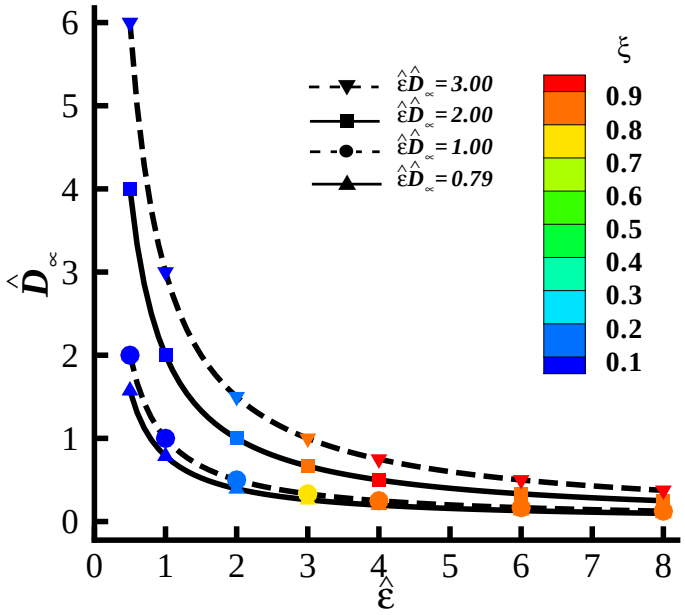



Figure 2.4: Clustering index (see color legend) as a function of reduced interaction potential well depth, $\hat{\varepsilon}$ and reduced diffusivity, $\hat{D}_\infty$. Each curve represents constant $\hat{\varepsilon}\hat{D}_\infty$. The region bounded by $\hat{\varepsilon}\hat{D}_\infty \ll 1$ represents the regime of validity of the position and velocity Langevin to PL reduction.

found that the BD-computed solute diffusivities are larger than those predicted by the MD method, and they attributed this discrepancy to the fact that the the frictional coefficient $\gamma$ in the Langevin Equation 2.3 is assumed to be constant, whereas a more realistic description (particularly for liquids) requires that the frictional coefficient be replaced by a time-dependent memory function. The fact that the BD method can accurately compute the equilibrium solute-solute radial distribution function and yet incur noticeable error in the calculation of diffusivity is perhaps to be expected for reasons mentioned in the previous section - namely that the stationary solution of the Fokker-Planck equation is independent of the diffusivity. Hence, one expects that in general, BD predictions of system dynamics will not match the predictions of corresponding MD simulations, but that equilibrium quantities can be well-predicted by BD simulations. If follows, therefore, that BD simulations of the early stages of an aggregation

process far from equilibrium may differ substantially from corresponding MD calculations. In spite of this observation (and the fact that MD simulation of systems with large aggregate-solvent size scale separation is not feasible), we have endeavored to perform MD simulations with sufficiently large numbers of solute particles such that a particle size distribution can be computed (at least during the early stages of aggregation) with the aim of directly comparing these MD predictions with corresponding BD calculations.

We seek a computationally tractable model system of aggregation appropriate for comparison of BD and MD methods, within the limitations discussed above. Consequently, we follow Giro's example and carry out simulations using equal-diameter Lennard-Jones solute and solvent particles. However, in contrast with the work of Giro, the solute-solute interaction potential well depth, $\varepsilon$, was chosen such that solute aggregation was favored (as was discussed in the previous section). Additionally, the ratio of the mass of a single solute primary particle to a solvent molecule, $m_{solute}/m_{solvent} = 50$, was chosen to be relatively large to ensure that the solute particles had lower mobility than the solvent molecules, despite the fact that they have equal size. All MD simulations were carried out using LAMMPS on an IBM eServer Blue Gene which consists of 1024 dual-core PPC440 CPUs running at 700Mhz, with 512MB of RAM per node. Each run on the Blue Gene took up to 5 hours on 1024 CPUs, and other simulation details are listed in Table 2.4. In the case of BD simulations, the position-velocity equations were used because the position-only reduction is not applicable for this set of parameters.

In order to determine the accuracy of the BD simulations for aggregating systems, in Fig. 2.5 we compare the extent of aggregation $\xi$ (as defined in 2.12) with that obtained from MD simulations for the system described in Table 2.4. It is clear that on the basis of the dimensionless time used to compare the two methods, the BD calculation predicts significantly more aggregation than does the MD simulation. The large disparity between the two curves suggests that the proper time scaling relation between the BD and MD is not given by $\hat{t} = \sigma t/\sigma_{v_\infty}$, although it is unclear what the correct relation should be. Hence, in order to provide a better basis of comparison for the two methods, we shall compare the predicted cluster size distributions at the same extent of aggregation, $\xi$.

Table 2.4: Simulation parameters used to produce Figure 2.5. Particle interactions were modeled using Lennard-Jones potentials. MD simulations were carried out using the LAMMPS[9] software package.

| Parameter | Description | Value (MD) | Value (BD) |
|---|---|---|---|
| $N_{solvent}$ | Number of Solvent Particles | 809,787 | N/A |
| $N_{solute}$ | Number of Solute Particles | 10,000 | 10,000 |
| $\varepsilon/k_BT_\infty$ | Reduced Well Depth | 8 | 8 |
| $f_{v,solvent}$ | Solvent Volume Fraction | 0.44 | N/A |
| $f_{v,solute}$ | Solute Volume Fraction | 0.005 | 0.005 |
| $\sigma$ | Particle Diameter | $3.4 \times 10^{-10}$ m. | $3.4 \times 10^{-10}$ m. |
| $D_\infty/\sigma\sigma_{v_\infty}$ | Dimensionless Diffusion Coefficient | N/A | 0.787 |
| $\sigma t_{stop}/\sigma_{v_\infty}$ | Dimensionless Simulation Time | 329.8 | 329.8 |

Figure 2.6 shows a direct comparison of the cluster size distributions computed using corresponding MD and BD simulations at $\xi = 0.89$. Although the two cluster size distributions appear to compare favorably in general, we have omitted from this plot the data for monomer frequency (which is quite large for the MD case) in order to depict in detail the comparisons for clusters. In fact the agreement between the MD and BD simulations is very poor for the monomer frequency (471 monomers in the MD simulation and only 97 in the BD simulation), and hence the BD simulations predict a larger number average cluster size (10.2 particles/cluster versus 8.5 particles/cluster for MD). If the monomers are de-emphasized by computing the mass-average cluster size (ratio of the second to first moment of the cluster size distribution), the mean particle size is 20.2 particles/cluster for the MD simulations and 15.7 for BD simulations. The larger mass-average particle size in the MD simulations (despite the fact that the MD simulations produce a much larger population of monomers) is a reflection

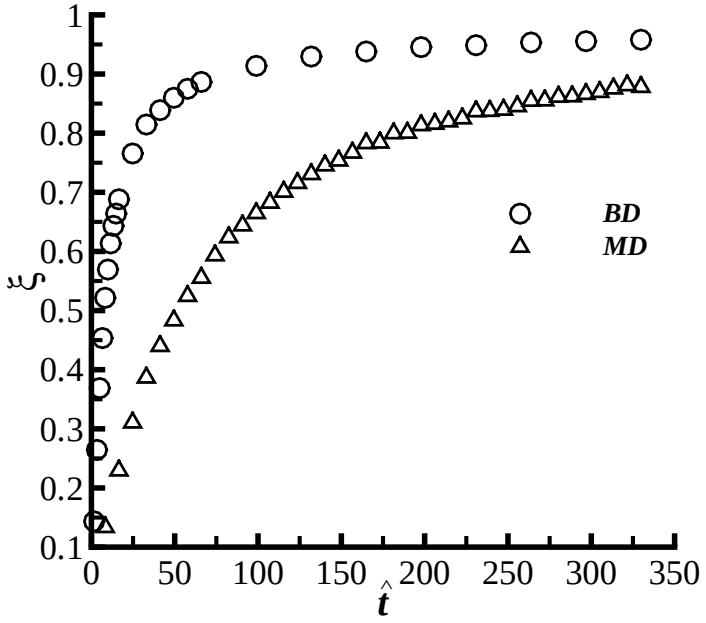



Figure 2.5: Extent of aggregation as a function of dimensionless time, $\hat{t} = \sigma t/\sigma_{v_\infty}$, for BD and MD simulations described in Table 2.4.

of the fact that the tail of the cluster size distribution (at large size) for the MD case decays more slowly than in the BD case.

An alternative method for comparing the cluster size distributions computed using the MD and BD simulation methods is to employ a dynamic scaling relation. In particular, it has been observed for a very wide range of aggregation processes that cluster size distributions can be collapsed by employing the following scaling ansatz:[23]

$$N_k = s^{-2}(t)\phi(k/s(t)), \tag{2.13}$$

where $N_k$ is the concentration of clusters containing $k$ monomers, $s(t)$ is the mass-averaged particle size, and $\phi$ is a scaling function. If Eq.2.13 is valid, then a plot of $s^2 N_k$ vs. $k/s$ should collapse the cluster size distributions for all sufficiently large values of $t$ such that the self-preserving regime has been reached. Figures 2.7 and 2.8 show such plots for the MD and BD cases, respectively. Despite the relatively large statistical error associated with

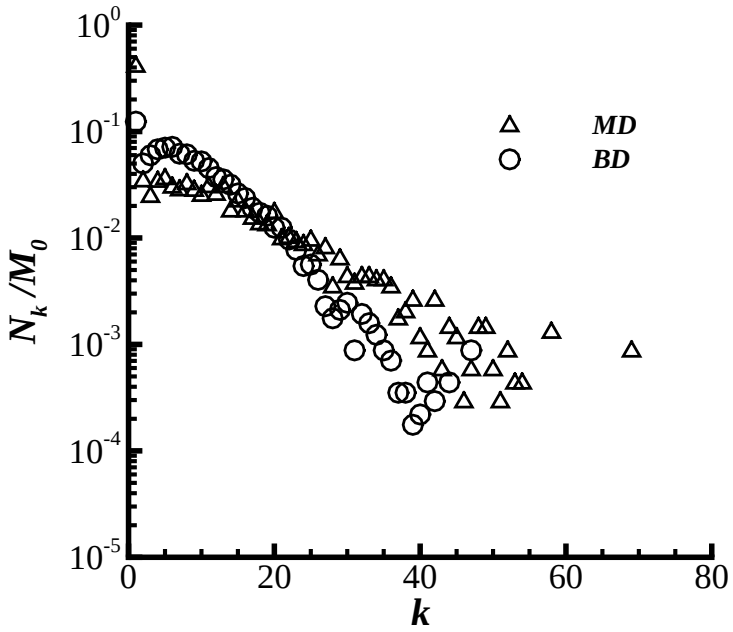



Figure 2.6: Comparison of cluster size distributions obtained from MD and BD simulations of aggregation carried out under conditions specified in Table 2.4 and at the same extent of aggregation $\xi = 0.89$.

only carrying out a small number of independent simulations, in both cases the cluster size distributions do appear to fall on universal curves when plotted using Eq.2.13. However, comparison of Figures 2.7 and 2.8 demonstrates that the shape of the scaling functions are clearly different for the MD and BD cases. Consequently, it can be concluded that the BD simulations produce different cluster size distributions than the MD simulations, independent of any difficulties in comparing them due to lack of information concerning the proper time scaling to be used. In particular, we see that the MD simulations generate cluster size distributions that decay monotonically in size, whereas the BD simulations produce cluster size distributions that exhibit a maxima in $N_k$.

The morphology of the aggregates generated by MD and BD simulations can be compared by computing the volume fractal dimension, $d_f$, as illustrated in Figures 2.9 and 2.10 for $\xi = 0.89$. Both types of simulations produce clusters with $d_f \approx 2.5$, which is a relatively

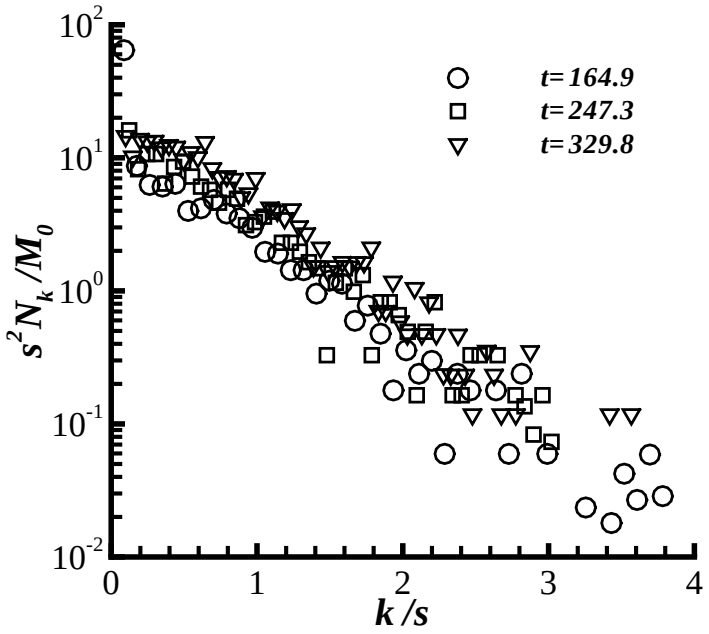



Figure 2.7: Scaled cluster size distributions for MD simulations.

large value indicating that the clusters are quite compact. Indeed, this fractal dimension is comparable to the value produced in processes with diffusion-limited growth by monomer addition[24]. Hence one could infer that the collisions between small clusters and large clusters are more important than are the collisions between two large clusters in both the MD and BD simulations, even at large extents of aggregation.

## 2.6 Summary and Discussion

In the introduction we posed the question as to how BD model coefficients can be obtained from MD simulations or other methods. Although the Giro *et al.* study[22] showed that in non-aggregating dilute systems the potential of mean force for BD could be inferred by curve-fitting the equilibrium pair correlation function $W(r) = -k_B T_\infty \ln g_\infty(r)$, this approach is not feasible in aggregating systems. One reason is because the pair correlation function is itself evolving as the system aggregates. It is possible that matching the pair correlation function

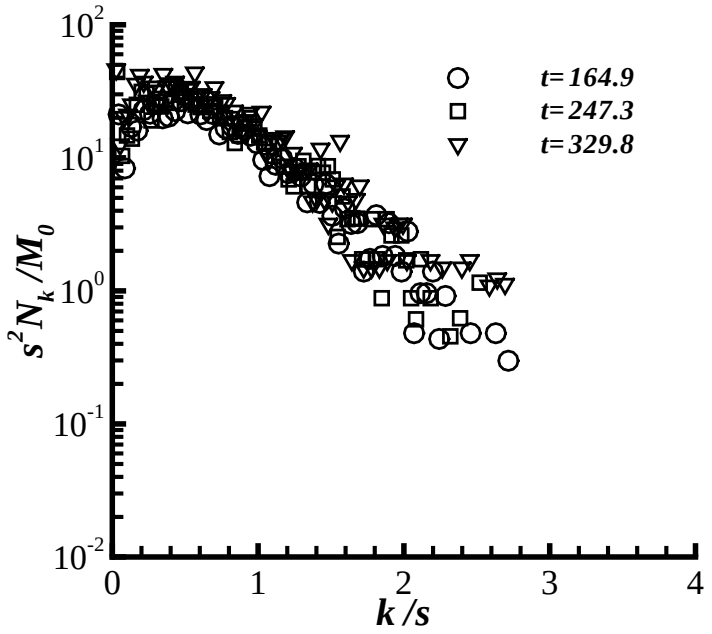



Figure 2.8: Scaled cluster size distributions for BD simulations.

from a BD simulation to the corresponding MD simulation of an aggregating system is a necessary condition for an accurate simulation. Subramaniam and Pai [25] outline an approach for deriving the evolution equation of the pair correlation function in MD simulations that reveals the importance of the relative velocity and relative acceleration between particle pairs, conditional on their separation distance. It is possible that matching the conditional relative acceleration statistics from MD to BD through the potential for mean force specification can guarantee the matching of the pair-correlation function.

However, it is important to note that the cluster size distribution that determines important aggregation statistics contains more information regarding connectivity of the monomers in clusters that is not available in the pair-correlation function. The requirement for matching moments of the cluster size distribution, and their relation to the pair correlation function, can provide a rational specification for model coefficients in the BD equations.

Clearly the progress of aggregation as characterized by $\xi$ is another important quantity that

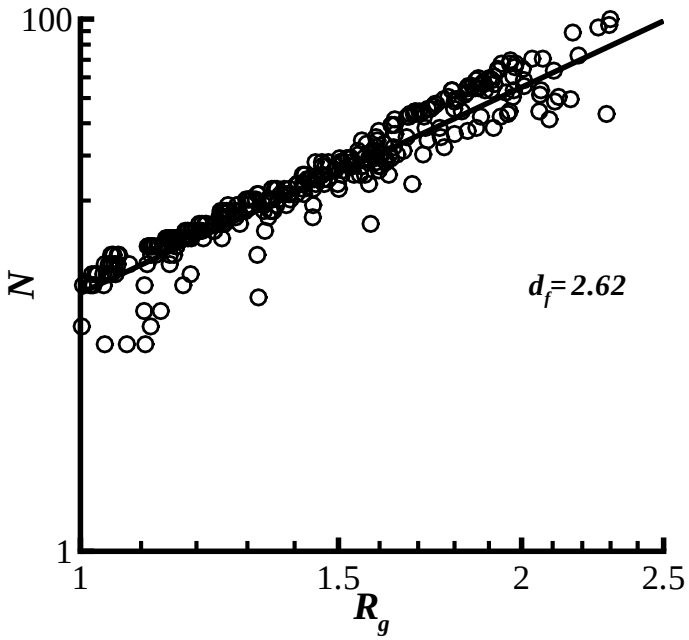



Figure 2.9: Number of monomers as a function of radius of gyration, $R_g$ for MD simulations described in Table 2.4. The slope of the linear fit is the volume fractal dimension, $d_f$.

BD simulations should capture accurately. Although this may seem to be closely related to the accuracy of BD in predicting the diffusivity of monomers, the issues involved are somewhat more complex. As noted earlier, even the notion of computing a diffusivity from mean-squared displacements of the monomers is questionable in aggregating systems, and therefore it is unclear whether the trends in predicted diffusivity from dilute non-aggregating BD simulations can be used to infer the physics of aggregating systems. Secondly, it seems more likely that the mean relative velocity between particle pairs conditional on their separation (or the implied second-order diffusivity [25]) determines aggregation, rather than the single particle diffusivity.

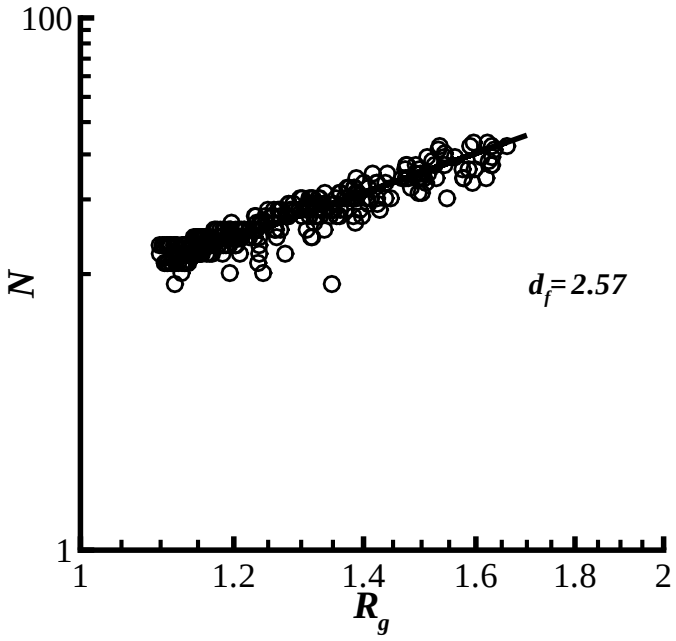



Figure 2.10: Number of monomers as a function of radius of gyration, $R_g$ for BD simulations described in Table 2.4. The slope of the linear fit is the volume fractal dimension, $d_f$.

## 2.7 Acknowledgement

The authors would like to thank Monica Lamm for helpful discussions and for pointing out the work of Giro and coworkers.

## Bibliography

[1] Roco, M. C. Nanoparticles and nanotechnology research. *J. Nanoparticle Res.* **1999**, *1*, 1.

[2] Friedlander, S. K. Nanoparticles and their structures: the next generation. *J. Nanoparticle Res.* **1999**, *1*, 159.

[3] Marchisio, D. L. and Fox, R. O. Solution of population balance equations using the direct quadrature method of moments. *J. Aerosol Sci.* **2005**, *36*, 43-73.

[4] von Smoluchowski, M. Versuch einer mathematischen theorie der koagulation-skinetik kolloider losungen. *Phys. Chem.* **1917**, *92*, 129.

[5] Saffman, P. G. and Turner, J. S. On the collision of drops in turbulent clouds. *J. Fluid. Mech.* **1956**, *1*, 16-30.

[6] Frenkel, D. and Smit, B. *Understanding Molecular Simulation*; Academic Press: San Diego, 2002.

[7] Pranami, G.; Slipchenko, L.; Lamm, M. H.; and Gordon, M. S. A coarse-grained forcefield for polystyrene derived from the effective fragment potential. *in preparation.*

[8] Adamovic; Li, H.; Lamm, M. H.; and Gordon, M. S. Modeling styrene-styrene interactions. *J. Phys. Chem. A* **2006**, *110*, 519525.

[9] Plimpton, S. J. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.* **1995**, *117*, 1-19. Software download site: http://lammps.sandia.gov/index.html.

[10] Sorensen, C. M. Light scattering by fractal aggregates: A review. *Aerosol Sci. Tech.* **2001**, *35*, 648-687.

[11] Ermak, D. L. and McCammon, J. A. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* **1978**, *69*, 1352-1360.

[12] Moskala, A. and Payatakes, A. C. Estimation of the diffusion coefficient of aerosol particle aggregates using Brownian simulation in the continuum regime. *Aerosol Sci.* **2006**, *37*, 1081-1101.

[13] Bos, M. T. A. and van Opheusden, J. H. J. Brownian dynamics simulation of gelation and aging in interacting colloidal systems. *Phys. Rev. E* **1996**, *53*, 5044-5050.

[14] Lodge, J. F. M. and Heyes, D. M. Brownian dynamics simulations of Lennard-Jones gas/liquid phase separation and its relevance to gel formation. *J. Chem. Soc. Faraday Trans.* **1997**, *93*, 437-448.

[15] Mellema, M.; van Opheusden, J. H. J.; and van Vliet, T. Relating colloidal particle interactions to gel structure using Brownian dynamics simulations and the Fuchs stability ratio. *J. Chem. Phys.* **1999**, *111*, 6129-6135.

[16] Puertas, A. M.; Fernandez-Barbero, A.; and de las Nieves, F. J. Structure factor scaling in colloidal charge heteroaggregation. *Eur. Phys. J. E* **2005**, *18*, 335-341.

[17] Cerda, J. J.; Sintes, T.; Sorensen, C. M.; and Chakrabarti, A. Structure factor scaling in colloidal phase separation. *Phys. Rev. E* **2004**, *70*, 051405.

[18] Videcoq, A.; Han, M.; Abelard, P.; Pagnoux, C.; Rossignol, F.; and Ferrando, R. Influence of the potential range on the aggregation of colloidal particles. *Physica A* **2007**, *374*, 507-516.

[19] Gardiner, C. W. *Handbook of Stochastic Methods*; Springer: Berlin, 2004.

[20] Klyatskin, V. I. *Stochastic Equations Through the Eye of the Physicist*; Elsevier: Amsterdam, 2005.

[21] Kloeden, P. E. and Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer-Verag: Berlin, 1992.

[22] Giro, A.; Guardia, E.; and Padro, J. A. Langevin and molecular dynamics simulations of Lennard-Jones liquids. *Mol. Phys.* **1985**, *55*, 1063-1074.

[23] Vicsek, T. and Family, F. Dynamic scaling for aggregation of clusters. *Phys. Rev. Lett.* **1984**, *52*, 1669-1672.

[24] Witten, T. A. and Sander, L. M. Diffusion-limited aggregation: a kinetic critical phenomenon. *Phys. Rev. Lett.* **1981**, *47*, 1400-1403.

[25] Subramaniam, S. and Pai, M. G. Transport of fluctuations in clustering particle systems. *Phys. Rev. Lett.* (in review).

# CHAPTER 3. COARSE-GRAINING APPROACH TO INFER MESOSCALE INTERACTION POTENTIALS FROM ATOMISTIC INTERACTIONS FOR AGGREGATING SYSTEMS

A paper is submitted in *Physical Review E*

Sergiy Markutsya, Shankar Subramaniam

## 3.1 ABSTRACT

A coarse-graining (CG) approach is developed to infer mesoscale interaction potentials in aggregating systems, resulting in an improved potential of mean force for Langevin dynamics (LD) and Brownian dynamics (BD) simulations. Starting from the evolution equation for the solute pair correlation function, this semi–analytical CG approach identifies accurate modeling of the relative acceleration between solute particles in a solvent bath as a reliable route to predicting the time–evolving structural properties of non–equilibrium aggregating systems. Noting that the solute–solvent pair correlation function attains a steady state rapidly compared to characteristic aggregation time scales, this CG approach derives the effective relative acceleration between a pair solute particles in the presence of this steady solute–solvent pair correlation by formally integrating the solvent force on each solute particle. This results in an improved potential of mean force that explicitly depends on the solute–solute and solute–solvent pair potentials with the capability of representing both solvophilic and solvophobic interactions that give rise to solvation forces. This approach overcomes the difficulty in specifying the LD/BD potential of mean force in aggregating systems where the solute pair correlation function evolves in time, and the Kirkwood formula $U(r) = -k_B T \ln g(r)$ that is applicable in equilibrium diffusion problems cannot be used. LD simulations are compared with molecu-

lar dynamics (MD) simulations for a model colloidal system interacting with Lennard-Jones pair potentials to develop and validate the improved potential of mean force. LD simulations using the improved potential of mean force predict a solute pair correlation function that is in excellent match with MD in all aggregation regimes, whereas using the unmodified MD solute-solute pair potential in LD results in a poor match in the reaction–limited aggregation regime. The improved potential also dramatically improves the predicted extent of aggregation and evolution of cluster size distributions that exhibit the same self–similar scaling found in MD. This technique of coarse–graining MD potentials to obtain an improved potential of mean force can be applied in a general multiscale framework for non–equilibrium systems where the evolution of aggregate structure is important.

## 3.2 INTRODUCTION

In high-rate methods for nanoparticle synthesis, nuclei that are formed by turbulent mixing of chemical precursors grow rapidly by surface addition and aggregation. A fundamental understanding of aggregation in colloidal systems is needed to optimize these methods such that they yield nanoparticle aggregates of desired size. While a complete characterization of nanoparticle aggregation will depend on the specific chemical precursors and solvent, several general characteristics of the phenomenon have been studied in a simplified generic system of latex nanoparticles [1, 2] in water destabilized by the addition of $MgCl_2$ salt. Experiments show the emergence of different aggregate structures depending on the extent of aggregation, and the duration and intensity of applied shear [1, 2, 3]. The structure of aggregates in turn affects the aggregation rate, resulting in a coupled nonlinear phenomenon. Therefore, a simulation method used to predict aggregation in colloidal systems must accurately describe the structure of aggregates.

A variety of simulation approaches have been employed to study aggregation, ranging from population balance approaches at the macroscale to molecular dynamics at the microscale (MD) [4]. Each level of description represents a trade-off between the fidelity with which the physico-chemical interactions are represented and the associated computational cost. Monte

Carlo methods are frequently used to simulate aggregation because their computational cost scales favorably for large systems. They are classified on the basis of the physico-chemical regimes of aggregation that they are intended to describe. In non-sheared systems, aggregation outcomes emerge from a dynamic balance between interparticle attractive forces arising from a solute interaction potential and random thermal motions that can disrupt this attraction. If on average the attractive forces dominate then aggregates or clusters of solute particles form. Aggregation regimes can be defined in terms of reaction-diffusion terminology based on whether the rate-limiting step arises from diffusion or reaction. Thus, in the diffusion–limited aggregation (DLA) regime, once the aggregating particles have diffused close enough towards each other they always stick together and form an aggregate due to strong attractive forces between them. In the reaction–limited aggregation (RLA) regime, particles do not always aggregate every time they come into contact because the sticking probability is less than 1. At a microscopic level, this can be due to the presence of a repulsive barrier in the solute interaction potential. The most commonly used Monte-Carlo methods to simulate aggregation in colloidal systems are diffusion-limited cluster aggregation (DLCA), ballistic-limited cluster aggregation (BLCA), and reaction-limited cluster aggregation (RLCA) [5, 6, 7, 8, 9, 10, 11]. However, because of their simple sticking probability rules for cluster formation, the MC-based approaches are not able to account for rearrangement of aggregates within a cluster due to thermal motion, or restructuring of the aggregates under shear.

Molecular dynamics simulations yield unparalleled physical insight into aggregation in colloidal systems and enable prediction of colloidal structure for different colloidal systems such as systems of solvophilic or solvophobic solute particles. However, because all the interparticle forces (solute-solute, solute-solvent, and solvent-solvent) are represented in MD, the large proportion of solvent molecules in a dilute system makes the MD system very large. Also, disparity in solute and solvent particle size (e.g., latex nanoparticles $\sim 10 - 100 \times 10^{-9}$m to water molecules $\sim 10^{-10}$m) slows down the MD calculations significantly because the relatively large nanoparticles have many solvent neighbors whose interactions must be accounted for [12]. The need for a wide dynamic range of aggregate sizes to reliably extract aggregate statistics

also demands MD simulation of a very large number of particles.

The need for reliable statistical characterization of aggregate size and structure in aggregation problems motivates us to examine coarse-graining techniques for molecular dynamics. Mesoscale methods such as Brownian dynamics (BD) [13, 14], dissipative particle dynamics (DPD) [15, 16], and stochastic rotational dynamics (SRD) [17] are established approaches to coarse-graining MD. However, the predictive capability of these coarse-graining approaches in aggregation problems needs to be validated. While the classical problem of the diffusion of large inertial solute particles in a bath of solvent molecules is what motivated the BD approach, its capability to accurately predict aggregate structure in colloidal systems is not established. Brownian or Langevin dynamics is a simulation approach that is formally obtained by using a projection operator technique on the MD equations, which effectively replaces the solute-solvent interactions by frictional and random forces [13, 14]. DPD and SRD models account for solute-solvent interactions more faithfully, with resultant increase in their computational expense relative to LD. While all these mesoscale methods are orders of magnitude faster than MD, the cost of SRD and DPD depends on the solute/solvent volume fraction and size ratios. The cost of LD is independent of the solute/solvent size ratio, but can depend on the solute volume fraction if hydrodynamic interactions are included. Among these mesoscale models we have chosen the LD approach for simulating aggregation because it is the most computationally efficient, while still being capable of representing aggregate structure. However, because the LD coarse-graining approach was not originally developed for aggregation problems it needs further testing and development.

The LD model requires the specification of a potential of mean force to account for the effect of solvent molecules on solute interparticle interaction. Typically the LD model is used in equilibrium systems where the Kirkwood formula [18]

$$U_{AA}^{LD}(r) = -k_B T \ln g_{AA}^{\mathrm{eqm}}(r) \tag{3.1}$$

is used to specify the potential of mean force. However, for non-equilibrium aggregating systems the choice of appropriate interaction potential between solute particles in the presence of solvent molecules is not well established. The interaction of a pair of nanoparticle aggregates

in the presence of solvent molecules has been studied using MD simulations [19, 20] and experiment [21]. The simulations [19, 20] show that the presence of solvent molecules changes the interaction between an aggregate pair. The nature of this change in the interaction depends on the geometry and size of the aggregates, and also very strongly on the solute-solvent interaction behavior (solvophilic solute particles vs. solvophobic particles). These works [19, 20, 21] point to the importance of the solute-solvent interaction, and motivate its introduction into the LD potential of mean force. However, the results of these studies cannot be directly applied to the LD potential of mean force because they do not take into account the statistical distribution of solute and solvent particles. What is needed is a statistical approach to connect the microscale solute-solvent interactions to the mesoscale solute-solute interactions, which is critical to establish sound coarse-graining procedures for aggregation problems.

There are several coarse–graining methods that are commonly used to derive an effective potential such as energy–based coarse–graining, Boltzmann Inversion (BI), Iterative Boltzmann Inversion (IBI), Inverse Monte Carlo (IMC), and Force Matching (FM) approaches [22, 23]. Useful categorizations of CG methods are based on the target (structure or thermodynamics) and method of linking (forces, effective interactive potentials) [22], or whether the method is iterative or non–iterative [23]. In energy–based coarse–graining approach [24, 25] the CG potential is developed to fit free energies in the system. This method is useful for processes such as lipid membrane association. However, energy–based coarse–graining does not guarantee reproduction of the atomistic structure of the system, which is essential for aggregation problems. BI, IBI, and IMC approaches [26, 27, 28] are structure–based methods that reproduce a pre–defined target *equilibrium* structure described by a set of radial distribution functions obtained from full molecular simulations of the reference system. These approaches are not suitable for predicting the time–evolving structure of an aggregating system. The FM approach [29] is used to reproduce thermodynamic properties of systems at *equilibrium*. In this method the difference between the instantaneous CG forces and the forces from full molecular simulation are minimized using a least–squares fitting procedure. However, this method does not guarantee an exact reproduction of local structural properties such as the pair distribution function,

nor is it applicable to time–evolving non–equilibrium systems. In summary, these existing CG approaches do not address the problem of time–evolving structure in non–equilibrium systems, which characterizes aggregation.

In this work we propose such a new method for deriving the potential of mean force for the LD model that accurately captures aggregation statistics in non-equilibrium systems. The CG method developed in our work is a non–iterative, semi–analytical, force–matching method that exploits rapid relaxation of the solute–solvent pair correlation function to derive an analytical expression for the effective potential that is used to predict time–evolving structure (including pair correlation function) in a non–equilibrium system. In the context of accurately computing aggregation statistics, we identify the need for LD to accurately reproduce the solute pair correlation function. This in turn leads us to derive the evolution of the second-order density (unnormalized pair correlation function) corresponding to the MD and LD dynamical equations. This transport equation for second-order density is used as a route to improve the LD model based on the relative acceleration concept. This theoretical basis provides the necessary connection between microscale and mesoscale interactions for coarse-graining of aggregating systems.

We propose an improved LD model where we modify the direct interaction between solute particles based on their MD pair potential by adding a potential that accounts for the presence of solvent molecules. We computed key aggregation statistics from MD and compared with the standard and improved LD models. The results show significant improvement in LD prediction of solute particle pair correlation function, dynamic scaling of the cluster size distribution, and extent of aggregation with the improved LD model.

The rest of this paper is organized as follows. First, we describe the simulation methods used in this work. We assess the performance of the standard LD method in DLA and RLA regimes by comparing with benchmark MD simulations. This comparison reveals the need for an improved LD model in the RLA regime. The analysis of solute relative acceleration from the MD simulations allows us to directly quantify the importance of solute-solvent interaction in the RLA regime. The transport equation for the unnormalized pair correlation function is

used as a route to build an improved LD model. The algorithm for computing the improved LD potential of mean force that takes into account influence of solvent molecules is then described. Then the aggregation statistics calculated from the improved LD method are analyzed and compared with the benchmark MD simulations, revealing a marked improvement in the LD model predictions. Generalization of the approach to other mesoscale methods, as well as the assumptions and limitations underlying the improved LD model are then discussed. The principal conclusions are then summarized. Appendix C gives the details of the pair correlation calculation for binary mixtures. A detailed derivation of the transport equation for the two–particle density leading to a phase–space expression of the relative acceleration between two particles is given in Appendix D. Appendix B describes a method for analytical calculation of the relative acceleration, and its comparison with results from numerical simulation for simple test problems.

## 3.3 METHODS

We simulate aggregation of solute particles immersed in liquid solvent using molecular dynamics and Langevin dynamics models, in order to assess the accuracy of LD and propose improvements. Direct MD calculation of physical systems (e.g., 20 nm latex nanoparticles in water [30]) is computationally prohibitive even with state–of–the–art numerical implementations because of the large size separation between solute and solvent particles [12, 31]. A model system with solute particles (denoted particles of type $A$) and solvent particles (type $B$) of the same size ($\sigma = 0.34$ nm) is chosen so that MD computations can be performed in reasonable time. This enables a direct comparison of LD simulations with MD results for the same system. We account for the lower mobility of the solute particles relative to the solvent molecules by assigning them higher mass $m_A = 2000$ amu, as compared to $m_B = 40$ amu for the solvent particles. Although our model system does not represent the size-separated case of nanoparticles aggregating in water, it is useful because it gives insight into a system that would otherwise be impossible to simulate. In fact, there are colloidal systems with hydrocarbon solvents such as $n$-decane where the solvent molecule size approaches that of nanoparticle clusters [32, 33].

Furthermore, using molecular dynamics simulations of size-separated systems with few solute particles, Qin and Fichthorn [19, 20] show that the solvation force for solvophilic particles when compared with the van der Waals force is important even when the solute particles are 9 times larger than the solvent molecules. Results from our model system show that we do indeed capture these essential solute-solvent interactions that manifest as solvation forces.

### 3.3.1 Molecular Dynamics

In MD the solute-solute, solute-solvent and solvent-solvent interactions are described by a truncated Lennard-Jones (12-6) potential,

$$U_{\alpha\beta}^{MD}(r_{ij}) = U_{\alpha\beta}^{LJ}(r_{ij}) = \begin{cases} 4\varepsilon_{\alpha\beta}\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^{6}\right], & r_{ij} \leq r_{\text{cut}} \\ \\ 0, & r_{ij} > r_{\text{cut}}, \end{cases} \tag{3.2}$$

where $\sigma$ is the particle diameter, $r_{ij}$ is the separation between centers of particles $i$ and $j$, $r_{\text{cut}}$ is the cutoff distance chosen to be $2.5\sigma$, and $\varepsilon_{\alpha\beta}$ is the potential well-depth between $\alpha$ and $\beta$ particles ($\alpha, \beta = A, B$). The dimensionless potential well depth is defined as

$$\hat{\varepsilon}_{\alpha\beta} = \frac{\varepsilon_{\alpha\beta}}{k_B T_{\text{ref}}}, \tag{3.3}$$

where $T_{\text{ref}}$ is the reference temperature chosen to be 121 K in these simulations.

Two cases of aggregation are simulated: one in the DLA regime and one in the RLA regime. These regimes are identified by generating a stability map from MD simulations to delineate different aggregation regimes in the $\hat{\varepsilon}_{AA} - \hat{D}_{\infty}$ parameter space. A similar stability map was generated for Langevin dynamics [12] to characterize and efficiently probe the parameter space for aggregating systems. This stability map reveals whether a system initialized at a given point in the parameter space will lead to large clusters or not, and it also gives an estimate of how long the formation of aggregates will take. In this map the extent of aggregation $\xi$ (which is defined as $\xi = 1 - M_0(t)/M_0(0)$, where $M_0(t)$ is the zeroth moment or total concentration of clusters at time $t$) is used as a metric to determine the extent of aggregation in the dimensionless self-diffusion coefficient $\hat{D}_{\infty}$ and dimensionless well-depth $\hat{\varepsilon}_{AA}$ space. When

the extent of aggregation approaches unity we identify these regions as belonging to the DLA regime. If the extent of aggregation is closer to zero we identify these regions as belonging to the RLA regime. The dimensionless well-depth corresponding to these regimes is listed in Table 3.1. In both these systems the dimensionless number density ($\hat{n} = n\,\sigma^3$) is 0.01 for

Table 3.1: Parameters used in MD simulations.

| Aggregation regime | $\hat{\varepsilon}_{AA}$ | $\hat{\varepsilon}_{AB}$ | $\hat{\varepsilon}_{BB}$ | $N_A$ | $N_B$ |
|---|---|---|---|---|---|
| DLA | 8.0 | 2.83 | 1.0 | 10,000 | 813,218 |
| RLA | 4.0 | 2.0 | 1.0 | 10,000 | 813,218 |

the solute particles, while it is 0.85 for solvent molecules. All simulations are performed in a cubic domain with periodic boundary conditions. Each side of the cube is $98.53\,\sigma$ in particle units, resulting in a total number of solute particles $N_A = 10,000$, and total number of solvent molecules $N_B = 813,218$.

The MD simulations are performed using the LAMMPS [34] software package. The MD simulations correspond to the NVT ensemble, which is appropriate for comparison with the constant–temperature LD simulations. The initial spatial configuration of the particles is specified to ensure no overlaps between particles. This is accomplished by spatially distributing the solute particles according to a hard–core Matérn point process [35], and by placing the solvent molecules at FCC lattice sites. All particles are assigned a Maxwellian velocity distribution corresponding to the reference temperature. From this configuration the system is allowed to evolve and equilibrate to the initial condition for the aggregation simulations by allowing particles to interact with a dimensionless well–depth $\hat{\varepsilon}_{\alpha\beta} = 1$.

### 3.3.2 Langevin Dynamics

Langevin dynamics is an approach to efficiently simulate the evolution of solute particles by modeling the solute-solvent interactions in terms of frictional and random terms, and through modification of the solute pair potential in the presence of solvent molecules. For this model system, the relative magnitude of the timescales corresponding to the frictional and pairwise

interaction force terms requires evolution of the full position and velocity Langevin equation set for accurate LD simulations [13, 14, 12]. The LD equations for evolution of the position $\mathbf{r}_i$ and velocity $\mathbf{v}_i$ of the $i$-th solute particle are [13, 14, 12]

$$d\mathbf{r}_i = \mathbf{v}_i dt, \tag{3.4}$$

$$d\mathbf{v}_i = -\gamma \mathbf{v}_i dt + \frac{1}{m_i}\mathbf{F}_i dt + \sqrt{2\gamma}\sigma_{v_\infty} d\mathbf{W}_i, \quad i = 1, \ldots, N, \tag{3.5}$$

where $m_i$ is the mass of particle $i$, $\gamma = k_B T_{\text{ref}}/m_i D_\infty$ is the friction coefficient, $D_\infty$ is the self-diffusion coefficient of the solute particles in solvent at infinite dilution, $\mathbf{F}_i = -\nabla_{\mathbf{r}} U_{AA}^{LD}$ where $U_{AA}^{LD}$ is the potential of mean force between solute particles in the presence of solvent, $\sigma_{v_\infty}^2 = k_B T_{\text{ref}}/m_i$ is the stationary velocity variance, $d\mathbf{W}_i$ is a Wiener process increment, and $N$ is the total number of solute particles.

As noted earlier, the specification of the appropriate potential for mean force in aggregating systems is not straightforward because Eq. 3.1 cannot be used when the solute pair correlation function $g_{AA}$ is evolving in time through states not in equilibrium. It is useful to decompose $U_{AA}^{LD}$ as follows

$$U_{AA}^{LD} = U_{AA}^{MD} + \tilde{U}_2, \tag{3.6}$$

where $U_{AA}^{MD}$ is the potential corresponding to solvent–explicit MD simulations and $\tilde{U}_2$ is a correction to account for the modeled solute–solvent interactions. The simplest choice is to take $U_{AA}^{LD} = U_{AA}^{MD}$, which corresponds to the solute pair potential *un*modified by the presence of solvent molecules. This is identical to the LJ solute pair-potential used in the MD calculations (Eq. 3.2).

The self-diffusion coefficient at infinite dilution is a required input parameter for the LD model. The self-diffusion coefficient is extracted from MD simulation of the corresponding system, but with only 125 solute particles initially located at minimum $19\,\sigma$ from each other. This setup satisfies an infinite dilution condition for each individual particle and allows to obtain 125 multiple independent simulations in one run to improve statistics. These MD calculations are done in two stages. In the first stage the system is allowed to equilibrate for approximately $t = 0.89\sigma^2/D_\infty$ where $\sigma^2/D_\infty$ is the diffusion timescale. Then, over the next

$t = 0.89\sigma^2/D_\infty$ diffusion timescales the mean squared displacement of the solute particles is computed, and the self-diffusion coefficient at infinite dilution is obtained from the expression

$$\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle = 6D_\infty t, \tag{3.7}$$

where $t$ is time for which the system has evolved in the second stage. LD simulations are performed for the same DLA and RLA cases described in Table 3.1 using an in-house code for solving Eqs. 3.4-3.5 [12].

### 3.3.3 Aggregation Statistics

Aggregates resulting from the MD and LD simulations are characterized by calculating the pair correlation function $g(r)$, and cluster size distribution (CSD). The pair correlation function $g(r)$, and its Fourier transform the structure factor, are useful in characterizing aggregate structure. The structure factor can also be inferred from light scattering experiments, and the fractal dimension $d_f$ of the aggregates can be extracted from the structure factor [3] (see Appendix E for details). The expression for the pair correlation function in a binary system is

$$g_{\alpha\beta}(r) = \frac{\langle N_{\alpha\beta}(r, \Delta r) \rangle}{N_\alpha n_\beta V(r, \Delta r)}, \tag{3.8}$$

where $\langle N_{\alpha\beta}(r, \Delta r) \rangle$ is the average number of $\alpha$–$\beta$ pairs with a $\beta$ particle in a shell $(r, \Delta r)$ separated by $r$ from an $\alpha$ particle, $N_\alpha$ is the total number of $\alpha$ particles in the system, $n_\beta V(r, \Delta r)$ is the expected number of $\beta$ particles in the shell $(r, \Delta r)$ with $n_\beta$ denoting the number density of $\beta$ particles, and $V(r, \Delta r)$ being the volume of the spherical shell (details are provided in Appendix C). We use the sample mean over all $\alpha$ particles to estimate the ensemble average $\langle N_{\alpha\beta}(r, \Delta r) \rangle$, leading to the following estimate for the pair correlation from particle data

$$g_{\alpha\beta}(r) \approx \frac{1}{N_\alpha n_\beta V(r, \Delta r)} \left( \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} N_{\alpha\beta}^{(i)}(r, \Delta r) \right).$$

The cluster size distribution (CSD) $N_k(k)$ gives a statistical description of the relative occurrence of the number of clusters $N_k$, each composed of $k$ monomers. The CSD is constructed by first initializing all particles as individual clusters and then recursively checking the another criterion to merge clusters. A particle is defined as belonging to a cluster if its center is within

$r_{cl} = 1.4\,\sigma$ of at least one particle already in the cluster. If a particle simultaneously belongs to two clusters this criterion is used to merge clusters.

## 3.4 PERFORMANCE OF LD MODEL

Here we compare the pair correlation function $g(r)$ and cluster size distribution obtained from Langevin dynamics with the MD results for the DLA and RLA systems described in Table 3.1. Simulations are evolved to a nondimensional time $\hat{t} = tD_\infty/\sigma^2 = 86.5$, where $\sigma^2/D_\infty$ is the diffusion timescale. The time step is chosen based on a previous study in which the resolution requirements for accurate numerical simulation of aggregating systems using LD and MD were established [12].

In the DLA regime the pair correlation function for solute particles $g_{AA}(\hat{r})$ and cluster size distribution predicted by LD compare well with MD (Figs. 3.1a and 3.2a). However, in the RLA regime the LD predicts a significantly different pair correlation function (Fig. 3.1b) and cluster size distribution (Fig. 3.2b), as compared to MD.

The difference in $g_{AA}(\hat{r})$ calculated from MD simulations for different aggregation regimes can be explained by considering the different values of sticking probability for the solute particles. In the DLA regime the sticking probability for solute particles is higher than in the RLA regime. This decrease in sticking probability in the RLA regime occurs in part due to the lower well-depth $\hat{\varepsilon}_{AA} = 4.0$ for pairwise interaction of the solute particles. In addition, in RLA the solute-solvent interaction is relatively strong and therefore solvent particles ($B$-type particles) attach to the solute particles ($A$-type particles). Therefore, solvent particles can block solute particles from aggregating because they are of comparable size and the solute-solute interaction is relatively weak. In contrast, in DLA regime a strong solute pair potential $\hat{\varepsilon}_{AA} = 8.0$ is able to overcome the blockage effects of the $B$-type particles and large aggregates are formed. Therefore, the interaction between $A$ and $B$-type particles represented by $\hat{\varepsilon}_{AB}$ plays an important role in aggregation.

Based on these results we conclude that using the unmodified LJ potential for LD is adequate in the DLA regime for non size-separated model systems. However, in the RLA regime

the unmodified LJ potential is unable to capture the structure and size distribution of aggregates. This motivates the development of an improved potential of mean force for LD that is capable of accurately describe aggregate structure in both DLA and RLA regimes.

## 3.5 RELATIVE ACCELERATION

Since the use of unmodified MD potential in LD does not always result in good prediction of the solute pair correlation $g_{AA}(r)$, we investigate the evolution of $g_{AA}(r)$ to gain an insight into developing a better LD potential of mean force. The position-velocity pair correlation function $g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)$ is related to the two-particle density $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)$ by

$$\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t) = n^2 g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t), \tag{3.9}$$

where $n$ is the number density of particles. This leads us to consider the evolution of the two-particle density (or unnormalized pair correlation function) corresponding to the MD and LD dynamical equations. The two-particle density $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)$ is defined as [36]

$$\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t) = \langle f_1' f_2' \rangle, \tag{3.10}$$

where $f_k' = \sum_{i=1}^{N} \delta(\mathbf{v}_k - \mathbf{V}^{(i)}) \delta(\mathbf{x}_k - \mathbf{X}^{(i)})$, $\{\mathbf{X}^{(i)}, \mathbf{V}^{(i)}, i = 1, ...N\}$ are the position and velocity of the $N$ particles in the ensemble, the product $\langle f_1' f_2' \rangle$ is formed over distinct pairs ($j \neq i$) over all realizations of the multiparticle system. If the $i^{th}$ particle evolves according to $d\mathbf{X}^{(i)}/dt = \mathbf{V}^{(i)}$, and $d\mathbf{V}^{(i)}/dt = \mathbf{A}^{(i)}$, then the evolution of $\rho^{(2)}$ can be derived from Eq. 3.10. Thus, after differentiation of Eq. 3.10 and with additional assumptions of statistical homogeneity in both position space as well as velocity space the following equation for evolution of $\rho^{(2)}$ is obtained [36],

$$\frac{\partial \rho^{(2)}(\mathbf{r}, \mathbf{w}, t)}{\partial t} + \nabla_{\mathbf{r}} \cdot \left( \mathbf{w} \rho^{(2)} \right) + \nabla_{\mathbf{w}} \cdot \left( \langle \Delta \mathbf{A} | \mathbf{r}, \mathbf{w}, t \rangle \rho^{(2)} \right) = 0, \tag{3.11}$$

where $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$ is the pair-relative separation, $\mathbf{w} = \mathbf{v}_2 - \mathbf{v}_1$ is the pair-relative velocity, and $\langle \Delta \mathbf{A} | \mathbf{r}, \mathbf{w}, t \rangle = \langle \mathbf{A}^{(2)} | \mathbf{r}, \mathbf{w}, t \rangle - \langle \mathbf{A}^{(1)} | \mathbf{r}, \mathbf{w}, t \rangle$ is the average relative acceleration. The two-particle density evolves by a transport equation that contains two terms: one is a transport term in relative pair-separation space $\mathbf{r}$ that contains the pair-relative velocity $\mathbf{w}$, and the other

is a transport term in pair relative velocity space $\mathbf{w}$ that contains the conditional expectation of pair-relative acceleration $\langle \Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle$.

If the dynamical equations that govern solute particle evolution corresponding to the LD model (Eqs. 3.4,3.5) are used in deriving Eq. 3.11, then the evolution of $\rho^{(2)}_{AA}$ implied by LD is obtained. In the case of MD, by using the appropriate dynamical equations for particles (solute particles and solvent molecules), Eq. 3.11 can be written for $\rho^{(2)}_{AA}$, $\rho^{(2)}_{BB}$, and $\rho^{(2)}_{AB}$. Of course, only $\rho^{(2)}_{AA}$ from the MD can be compared with LD.

Comparing the evolution of $\rho^{(2)}_{AA}$ for MD and LD, it is clear that if the average relative acceleration term $\langle \Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle$ is accurately modeled, then the $\rho^{(2)}_{AA}$ evolution will be identical. The average relative acceleration $\langle \Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle$ is conditioned on pair-relative separation $\mathbf{r}$ and pair-relative velocity $\mathbf{w}$, and is difficult to estimate accurately from simulations because of high statistical error arising from few samples. It is useful to consider a class of mesoscale models (to which LD belongs) that decompose this conditional relative acceleration as

$$\langle \Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle = \langle \Delta\mathbf{A}|\mathbf{w},t\rangle + \langle \Delta\mathbf{A}|\mathbf{r},t\rangle. \tag{3.12}$$

In LD, the term $\langle \Delta\mathbf{A}|\mathbf{w},t\rangle$ is modeled as $-\gamma\mathbf{w}$, and we provisionally accept this model as adequate. We focus on improved modeling of the remaining term $\langle \Delta\mathbf{A}|\mathbf{r},t\rangle$, the average pair-relative acceleration conditioned on pair separation.

In both MD and LD, it is useful to decompose the average relative acceleration $\langle \Delta\mathbf{A}|\mathbf{r},t\rangle$ into a direct contribution to the relative acceleration $\langle \Delta\mathbf{A}^D|\mathbf{r},t\rangle$ and an indirect part:

$$\langle \Delta\mathbf{A}|\mathbf{r},t\rangle = \langle \Delta\mathbf{A}^D|\mathbf{r},t\rangle + \langle \Delta\mathbf{A}^I|\mathbf{r},t\rangle. \tag{3.13}$$

In MD, the direct contribution to the relative acceleration between two solute particles arises from $\Delta\mathbf{A}^D = \mathbf{F}^{MD}_{ij}/m_i - \mathbf{F}^{MD}_{ji}/m_j = \mathbf{F}^{MD}_{ij}\,(1/m_i + 1/m_j)$, where $m_i$ and $m_j$ are the masses of $i^{th}$ and $j^{th}$ particles respectively. The indirect contribution to relative acceleration arises from the interaction of particles (1) and (2) with ($a$) other solute particles and ($b$) solvent molecules, as shown in Figure 3.3. In LD, the direct contribution to relative acceleration between two solute particles arises from $\Delta\mathbf{A}^D = \mathbf{F}_{ij}\,(1/m_i + 1/m_j)$, where $\mathbf{F}_{ij}$ is the mean force between two solute particles in the presence of solvent. The indirect contribution to relative acceleration

in LD arises from the modeled interaction of particles (1) and (2) with other solute particles (the presence of solvent is implicit).

Given a model for $U^{LD}(r)$, the potential of mean force in LD, we can compute the modeled indirect contribution to relative acceleration and compare this with the corresponding result from MD where the interactions with solvent are explicitly represented. Prior to carrying out this comparison, we construct a simple test to verify the accuracy of our calculation of the indirect relative acceleration. Details of the test, which verify our computation of the relative acceleration with approximate analytic solutions, are given in Appendix B.

Figure 3.4 shows the indirect average relative acceleration of a pair of $A$-type particles to other $A$-type particles for both MD and LD simulations. We see that the indirect average relative acceleration calculated from MD and LD simulations match closely for systems in the DLA regime (Figure 3.4a), whereas they are significantly different for systems in the RLA regime (Figure 3.4b). Since in these simulations $U^{LD} = U^{MD}$ we attribute the mismatch between the MD and LD results in the RLA regime to the mismatch in the pair correlation function (Figure 3.1b).

In order to investigate the influence of solvent particles on the indirect relative acceleration between a pair of $A$-$A$ type particles, we calculate the indirect relative acceleration between $A$-$A$ type particles due to the effect of only $B$-type particles from MD simulations (Figure 3.5). Here we see that in the DLA regime the magnitude of relative acceleration due to $B$-type particles (Figure 3.5a) is much smaller than the effect of $A$-type particles only (Figure 3.4a). However, in the RLA regime, the relative acceleration due to $A$-type particles (Figure 3.4b, MD results) and $B$-type particles (Figure 3.5b) are of comparable magnitude. Clearly in the RLA regime, this effect of solvent in the relative acceleration that is unaccounted for in the LD model is responsible for the mismatch of the pair correlation of solute particles in Figure 3.1b. These results lead us to conclude that although the effect of solvent is not significant in the DLA regime, it must be accounted for in the RLA regime.

The analytical expression for the indirect relative acceleration

$$\langle \Delta \mathbf{A}^I | \mathbf{r} \rangle = \int \mathbf{A}^{I:(2)}(\mathbf{r}'', \mathbf{r}) g(\mathbf{r}'') d\mathbf{r}'' - \int \mathbf{A}^{I:(1)}(\mathbf{r}', \mathbf{r}) g(\mathbf{r}') d\mathbf{r}'$$

derived in Appendix B, which is valid for dilute systems where an approximate expression for conditional three–particle statistics holds, reveals that the indirect average relative acceleration obtained from MD and LD simulations can be matched only when both the interaction potential and the pair correlation function are correctly described by the LD simulations. We now describe an improved LD potential that takes into account the presence of solvent molecules.

## 3.6 IMPROVED BD MODEL

Since accurate simulation of aggregation using LD requires the incorporation of solvent effects in the potential of mean force, an approach to generate an improved LD potential needs to be developed. The theoretical framework described by Likos [37] is used as a basis to build such an improved potential. Likos' theoretical framework for a two-component solute ($A$) - solvent ($B$) system defines an effective Hamiltonian $\mathcal{H}_{\text{eff}}$ that depends solely on positions and velocities of the solute particles, and which can be formally related to the full two-component Hamiltonian $\mathcal{H}$ of the system [37]. This full two-component Hamiltonian (in our case defined by the MD) is

$$\mathcal{H} = \mathcal{H}_{AA} + \mathcal{H}_{BB} + \mathcal{H}_{AB}, \tag{3.14}$$

where each $\mathcal{H}_{\alpha\beta}$ contains the interactions between $\alpha$-type and $\beta$-type particles only. The effective Hamiltonian $\mathcal{H}_{\text{eff}}$ is related to the full Hamiltonian $\mathcal{H}$ by the following expression

$$\exp(-\beta\mathcal{H}_{\text{eff}}) = \text{Tr}_B[\exp(-\beta\mathcal{H})], \tag{3.15}$$

where $\beta = 1/k_B T_{\text{ref}}$, and $\text{Tr}_B$ is the trace over solvent molecules, representing the multiple integral over the positions and momenta of all degrees of freedom of the particles $B$, which are solvent molecules in our case. The final expression for the effective Hamiltonian after invoking the pair-potential approximation is [37]

$$\mathcal{H}_{\text{eff}} = \sum_{i=1}^{N_1} \frac{\mathbf{p}_i^2}{2M} + \sum_{i=1}^{N_1} \sum_{j=i+1}^{N_1} U_{\text{eff}}(|\mathbf{r}_i - \mathbf{r}_j|; n_A, n_B, T_{\text{ref}}) + F_0, \tag{3.16}$$

which can be thought of as the sum of an effective Hamiltonian of the $N_1$ solute particles conditional on the solvent macrostate specified by the solvent number density $n_B$ and the reference

temperature $T_{\text{ref}}$. The volume terms $F_0$ do not depend on the particle coordinates [37]. In the above expression $M$, $\mathbf{p}_i$, $\mathbf{r}_i$ are the mass, momentum and position of $i^{th}$ solute particle. The effective interaction potential $U_{\text{eff}}$ is given by [37]

$$U_{\text{eff}}(r = |\mathbf{r}_i - \mathbf{r}_j|; n_A, n_B, T_{\text{ref}}) = U_{AA}(r) + \tilde{U}_2(r; n_A, n_B, T_{\text{ref}}), \tag{3.17}$$

where $U_{AA}$ represents the direct interaction between solute particles, and $\tilde{U}_2$ is their interaction mediated by the solvent molecules (also called the *solvation potential*), $n_A$ and $n_B$ are number densities of solute particles and solvent molecules respectively, and $T_{\text{ref}}$ is the system reference temperature.

In the context of an improved LD model Eq. 3.17 tells us that the $U_{AA}$ potential represents the direct interaction between solute particles, with the $\tilde{U}_2$ potential accounting for the interaction between solute particles mediated by solvent molecules. The direct interaction potential $U_{AA}$ is simply the MD interaction potential between solute particles $U_{AA}^{MD}$. The solvation potential $\tilde{U}_2$ has been successfully extracted by different researchers [19, 20, 38]. In all these works the solvation potential $\tilde{U}_2$ (or solvation force $f_s(r)$) has been found for a liquid between two fixed parallel surfaces at separation $r$ [38], or between two fixed spheres at separation $r$ [19, 20]. However, to the best of our knowledge, a model for $\tilde{U}_2$ in dynamic aggregating systems with a *statistical* description of solute particles and solvent has not been proposed yet.

In the present work we propose the following algorithm for calculation of $\tilde{U}_2$ based on the hypothesis that the pair correlation function for the solvent molecules distributed around solute particle $g_{AB}(r', t)$ ($r'$ is the solute-solvent separation) reaches a steady state value $g_{AB}^{ss}(r')$ over a time scale that is much smaller than characteristic aggregation time scales. Indeed this hypothesis is supported by MD simulations as shown in Figure 3.6. This steady configuration of solvent molecules relative to solute particles induces a potential $U_{AB}(r')$ on each solute particle, which we model using the Kirkwood formula

$$U_{AB}(r') = -k_B T_{\text{ref}} \ln g_{AB}^{ss}(r'). \tag{3.18}$$

We now calculate the relative force between a pair of solute particles (marked (1) and (2) in Fig. 3.13) induced by this potential $U_{AB}$ based on the relative acceleration idea introduced

in the previous section. As shown in Appendix B, the average indirect relative acceleration between the two solute particles (1) and (2) in Fig. 3.13 separated by $\mathbf{r}$ that is induced by the solute–solvent potential $U_{AB}$ is given by

$$\langle \Delta \mathbf{A}^I_{AA}|\mathbf{r}\rangle = \int \mathbf{A}^{I:(2)}_{AB}(\mathbf{r}'',\mathbf{r}) h_{AB}(\mathbf{r}''|\mathbf{r}) d\mathbf{r}'' - \int \mathbf{A}^{I:(1)}_{AB}(\mathbf{r}',\mathbf{r}) h_{AB}(\mathbf{r}'|\mathbf{r}) d\mathbf{r}'. \tag{3.19}$$

This equation gives the average indirect relative acceleration conditional on solute pair separation $\mathbf{r}$, in terms of conditional three–particle statistics $h_{AB}(\mathbf{r}''|\mathbf{r})$ and $h_{AB}(\mathbf{r}'|\mathbf{r})$ that are defined in Appendix B, by integrating out the effect of solvent as shown in Fig. 3.13. In general the conditional three-particle statistics of systems are unknown, but assuming that in a dilute system $h_{AB}(\mathbf{r}'|\mathbf{r})$ and $h_{AB}(\mathbf{r}'|\mathbf{r})$ are well approximated by

$$h_{AB}(\mathbf{r}'|\mathbf{r}) \cong g_{AB}(\mathbf{r}') \tag{3.20}$$

$$h_{AB}(\mathbf{r}''|\mathbf{r}) \cong g_{AB}(\mathbf{r}''), \tag{3.21}$$

results in the following expression for the indirect relative acceleration that is completely specified by the solute–solvent potential $U_{AB}$ and solute–solvent pair correlation $g_{AB}$:

$$\langle \Delta \mathbf{A}^I_{AA}|\mathbf{r}\rangle = \int \mathbf{A}^{I:(2)}_{AB}(\mathbf{r}'',\mathbf{r}) g_{AB}(\mathbf{r}'') d\mathbf{r}'' - \int \mathbf{A}^{I:(1)}_{AB}(\mathbf{r}',\mathbf{r}) g_{AB}(\mathbf{r}') d\mathbf{r}', \tag{3.22}$$

where $\mathbf{A}^{I:(k)}_{AB}(\mathbf{r}',\mathbf{r}), k = 1,2$ is computed as

$$\mathbf{A}^{I:(1)}_{AB}(\mathbf{r}',\mathbf{r}) = \frac{\mathbf{F}^{(1)}_{AB}(\mathbf{r}')}{m} = -\frac{1}{m_A}\nabla U_{AB}(r') \tag{3.23}$$

$$\mathbf{A}^{I:(2)}_{AB}(\mathbf{r}'',\mathbf{r}) = \frac{\mathbf{F}^{(2)}_{AB}(\mathbf{r}'')}{m} = -\frac{1}{m_A}\nabla U_{AB}(r'') \tag{3.24}$$

where $m_A$ is the mass of solute particles.

Analytical expressions have been derived for the relative acceleration in the simple 1-D and 2-D test problems shown in Appendix B. In 1-D the integrals in Eq. 3.22 are calculated as follows. The centers of the pair of solute particles 1 and 2 (each of size $\sigma_A$) are located such that they are separated by a distance $r$. Then the "probe" solvent particle $p$ of size $\sigma_B$ is inserted at all allowable locations along the line (for 1-D case) that satisfy the conditions

$$|r'| = |r_1 - r_p| \geq 1/2(\sigma_A + \sigma_B) \tag{3.25}$$

$$|r''| = |r_2 - r_p| \geq 1/2(\sigma_A + \sigma_B) \tag{3.26}$$

to exclude overlap of solute and solvent particles. For each location of the "probe" particle $p$, the force $\mathbf{F}_{AB}^{I:(1)}(r',r)$ exerted by the probe particle on solute particle 1, and the force $\mathbf{F}_{AB}^{I:(2)}(r'',r)$ exerted by the probe particle on solute particle 2 is computed for each set of values $r$, $r'$, and $r''$ with $U = U_{AB}$. These forces are weighted by the pair correlation function $g_{AB}$ and integrated according to Eq. 3.22, to obtain the average indirect relative acceleration. By repeating the procedure for every value of the solute pair separation $r$ over a desired range, the average indirect relative acceleration is calculated as a function of solute pair separation. While analytical expressions can be derived for the relative acceleration in the simple 1-D and 2-D test problems shown in Appendix B, the 3-D integrals in Eq. 3.22 needed to calculate the relative acceleration in the improved LD model have to be performed numerically.

The relative force between two solute particles induced by the $U_{AB}$ potential is denoted

$$m_A\langle\Delta\mathbf{A}_{AA}^I|\mathbf{r}\rangle = -\nabla_{\mathbf{r}}\tilde{v}_2(r;n_A,n_B,T_{\text{ref}})$$

The final step is to add the force resulting from this semi-analytical calculation to the direct $A$-$A$ interaction and complete the LD specification of mean force as

$$\mathbf{F} = -\nabla_{\mathbf{r}}U_{AA,\text{eff}}^{LD} = -\nabla_{\mathbf{r}}U_{AA} - \nabla_{\mathbf{r}}\tilde{U}_2 = -\nabla_{\mathbf{r}}U_{AA} - C_2\nabla_{\mathbf{r}}\tilde{v}_2. \tag{3.27}$$

In the above expression we have modeled the solvation potential $\tilde{U}_2(r;n_A,n_B,T_{\text{ref}})$ by introducing a model coefficient $C_2$ to relate it to the potential $\tilde{v}_2$, such that

$$\tilde{U}_2(r;n_A,n_B,T_{\text{ref}}) = C_2\tilde{v}_2(r;n_A,n_B,T_{\text{ref}}), \tag{3.28}$$

where $\tilde{v}_2(r;n_A,n_B,T_{\text{ref}})$ is the mean potential between $A$-$A$ particles induced by the solute-solvent potential $U_{AB}(r')$. For the case where the LJ potential is used in the MD, the potential $U_{AA} = U_{AA}^{MD} = U_{AA}^{LJ}$.

The model coefficient $C_2$ is specified to match the indirect average relative acceleration for a single solute particle pair. The indirect average relative acceleration between all pairs of solute particles at a separation $r$ due to the presence of solvent molecules is calculated from an MD simulation that explicitly accounts for the solute-solvent and solvent-solvent interactions. The modeled indirect average relative acceleration for the solute particle pair is given by $-C_2\nabla_{\mathbf{r}}\tilde{v}_2$.

The coefficient $C_2$ is determined by a least-squares fit of the model to the indirect average relative acceleration results obtained from MD.

In Figure 3.7 we compare the improved LD potential $U_{\text{eff}}$ with the Lennard-Jones potential $U_{AA}^{LJ}$ and the modeled solvation potential $\tilde{U}_2$ for two cases: one with $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$ typical of the RLA regime, and another with $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$ typical of the DLA regime. In the DLA regime (Figure 3.7a) the modification of the LD potential due to solvation effects is not significant and the improved LD potential is very close to the Lennard-Jones potential. On the other hand, in the RLA regime (Figure 3.7b), the modification due to solvation effects is significant, and the improved LD potential and the Lennard-Jones potential are significantly different. Two interesting features of the improved LD potential in the RLA regime are worth noting in Figure 3.7b. The first is that the primary minimum of the improved LD potential (at $\hat{r} = 1.1$) has a well depth that is less than that of the LJ potential. The second is the appearance of a secondary minimum (at $\hat{r} = 2.1$). These two minima are separated by a potential barrier. The formation of this secondary minimum is a characteristic feature of the RLA regime where aggregation occurs only if the solute particle has sufficient energy to overcome a potential barrier. Since the improved LD potential is practically identical to the unmodified MD potential in the DLA regime, we do not expect the improved LD potential to change the structure of aggregates significantly. In the DLA regime the unmodified MD potential is capable of capturing aggregate structure, and so we expect that the improved LD model will not deviate significantly from these good results. Based on the potentials shown in Figure 8, we expect that in the RLA regime the improved LD potential will significantly change the predicted aggregate structure. In this section we have derived an improved LD potential for aggregating systems that takes into account the presence of solvent molecules.

## 3.7 RESULTS WITH IMPROVED BD MODEL

We repeated the LD simulations reported in previous sections with the improved LD interparticle force (Eq. 3.27) to ascertain if the improved LD potential gives a better prediction of aggregate structure and other aggregation statistics. In Figure 3.8a we compare the pair

correlation function predicted by the improved LD model with that from MD simulations and find reasonable agreement in the DLA regime. There is no significant difference in the improved LD model's predictions from those of the original LD model in the DLA regime where $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$ (compare Figure 3.8a with Figure 3.1a). Therefore, the improved LD model is as good as the original LD model in the DLA regime. However, in the RLA regime with $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$ there is a significant improvement in the pair correlation function predicted by the improved LD model (Figure 3.8b) when compared with the original LD result (Figure 3.1b). The improved LD model's representation of solvent effects on the potential of mean force result in an excellent match with the MD pair correlation function (Figure 3.8b), whereas the original LD with the unmodified MD potential (Lennard-Jones) resulted in a significant difference in the predicted pair correlation function (Figure 3.1b). Since aggregation is a time-dependent phenomenon, we also check if the improved LD model's prediction of $g_{AA}(\hat{r}, t)$ is accurate at different time instants. The absolute relative error in $g_{AA}(\hat{r}, t)$ between model and MD is integrated over all $r$ and averaged over three different time instants to quantify temporal accuracy. The maximum error in both RLA and DLA regimes is 26% for the improved LD model. For comparison, the same error for the original LD model is 600%.

We also compare the normalized cluster size distribution predicted by the improved LD model with that of MD simulations. The cluster size distribution (CSD) is formed by calculating the number of clusters $N_k$ that include $k$ monomers from the steady solute particles positions. Clusters are defined using the Stillinger criterion that is based purely on instantaneous physical proximity of the solute particles. When calculating $N_k$ we assume that two neighboring solute particles belong to the same cluster if they are separated by a distance less than $1.4\sigma$. The cluster sizes $N_k$ are normalized by the zeroth moment of the cluster size distribution $M_0(t)$, which characterizes the total number of clusters formed at time $t$. The CSD predicted by the improved LD model agrees very well with the MD result in both DLA and RLA regimes (see Figures 3.9a and 3.9b). Recall that the original LD model did not predict the CSD well in the RLA regime (cf. Figure 3.2b). The improvement in prediction of the CSD with the improved LD model is a significant result because the CSD contains topological

information that is not there in the pair correlation function.

Since in aggregation the CSD evolves in time, it is worthwhile to examine scaled cluster size distributions computed using the MD and LD (with unmodified LJ and improved potential) using a dynamic scaling relation. In particular, it has been observed for a very wide range of aggregation processes that cluster size distributions can be collapsed by employing the following scaling ansatz [40]:

$$N_k = s^{-2}\phi\left(\frac{k}{s(t)}\right), \tag{3.29}$$

where $N_k$ is the concentration of clusters containing $k$ monomers, $s(t)$ is the mass-averaged particle size, and $\phi$ is a scaling function. If Eq. 3.29 is valid, then a plot of $s^2 N_k$ vs $k/s$ should collapse the cluster size distributions for all sufficiently large values of $t$ in the self-preserving regime. Figures 3.10a and 3.10b show such plots for the RLA regime case ($\hat{\varepsilon}_{AA} = 4.0$) using MD, and LD with the unmodified LJ potential. While the MD shows a universal scaling for the cluster size according to Eq. 3.29, the same is not evident in the LD with unmodified potential. Furthermore, the shapes of the function $\phi$ are very different. Figure 3.10c shows the same function $\phi$ for LD with the improved potential, and we see a marked improvement in both the collapse and the comparison with the MD result. This is a remarkable justification for the validity of the relative acceleration concept as a basis for generating the improved LD potential. For the DLA regime case ($\hat{\varepsilon}_{AA} = 8.0$), Figures 3.10d and 3.10e show the same scaled cluster size distrubution plots using MD, and LD with the unmodified LJ potential similar to results previously reported in [12]. In the DLA regime the scaled cluster size distributions do appear to fall on universal curves for both cases when plotted using Eq. 3.29, but the shape of the scaling functions are clearly different. In the DLA regime the improved LD potential has a smaller effect on the scaled cluster size distributions as seen in Figure 8(f). While the maximum size of aggregates compares better with MD using the improved LD potential, the peak is more enhanced showing a slight departure from the MD result. Nevertheless as Figures 3.10c and 3.10f show, the improved LD potential gives a far better agreement for the scaled cluster size distribution with MD than the unmodified LD potential.

To address the issue of proper time evolution of aggregation at the coarse–grained level we

have included plots of the extent of aggregation $0 \leq \xi < 1$, defined as

$$\xi = 1 - \frac{M_0(0)}{M_0(t)} \tag{3.30}$$

where $M_0$ is the zeroth moment or total concentration of clusters. Hence $\xi$ is an aggregation progress variable that approaches unity as the system mass accumulates in a single cluster. We plot $\xi$ obtained from MD, unmodified LD (LJ), and improved LD simulations for both DLA and RLA regimes. In the RLA regime, LD with the unmodified LJ potential (Fig. 3.11a predicts a very rapid increase in the extent of aggregation when compared with MD. However, the improved LD potential results in a dramatic improvement in the match of predicted extent of aggregation $\xi$ with MD (Fig. 3.11b). This establishes that the coarse-graining approach is able to represent mesoscale time evolution accurately. In the DLA regime, even LD with the unmodified LJ potential (Fig. 3.11c) is reasonably close to the MD result and we ascribe the difference to the simple model for the frictional term. As a result, even LD with the improved potential shows the same difference with MD in the extent of aggregation (Fig. 3.11d).

We now examine the indirect average relative acceleration between solute particle pairs due to other solute particles, since this was the quantity that motivated the model development. The match between the improved LD model and the MD result in both RLA and DLA regimes (Figures 3.12a and 3.12b) is much better than with the original LD model. In the RLA regime the magnitude of the relative acceleration and its variation with separation $r$ are much improved in comparison with that obtained from the original LD model (Figure 3.4). This result establishes the validity of our modeling approach: namely, that calibrating the effect of solvent on the average relative acceleration of solute particles, in conjunction with our semi-analytical integration procedure that accounts for the statistical configuration of solute-solvent and solute-solute pairs, is a successful route to mesoscale coarse-graining of MD potentials in aggregating systems.

In summary, the comparison of aggregation statistics—the pair correlation function ($g_{AA}(\hat{r})$), the cluster size distribution (CSD), and the indirect average relative acceleration ($\langle \Delta \mathbf{A}^I | \mathbf{r} \rangle$)—predicted by the improved LD model with MD indicates that the improved LD potential developed in this work is successful in accurately modeling aggregation in a model system.

## 3.8 DISCUSSION

In this section the broader implications of this new approach to modeling the potential of mean force between solute particles in aggregating systems are now discussed. The CG approach we have developed in this work is substantially different from existing approaches in many ways. While existing CG approaches apply to structural properties such as RDF of *equilibrium* systems our approach addresses the problem of *non–equilibrium* time evolution of structure. Our approach of deriving the evolution equation for the second–order density and identifying the conditional relative acceleration gives a sound physical basis for coarse–graining because it relates the effective potential to the *dynamics* of the system (in contrast to adjustment of equilibrium structure using model coefficients). Our method results in an analytical form of the correction potential without fitting parameters that is based on physical reasoning, and we can interpret the improved potential in terms of ratio of solute-solute to solute-solvent interaction potential well depth. The model coefficient in our approach only determines the relative strength of the correction but does not alter the shape of the correction potential. In our approach the numerical integration of the probe particle over physical space to determine the correction term to potential interaction needs to be performed only once (i.e., it is not an iterative method).

This new modeling approach establishes a connection between microscale simulation methods (such as MD) and mesoscale simulation methods (such as LD) based on the pair correlation function and relative acceleration. This approach can be generalized to other mesoscale approaches (e.g., DPD), and to other aggregation statistics of interest. The improved LD model can be applied to any MD potential (LJ is only chosen here as a standard example), and it is not limited to isotropic potentials. The model can be used for both solvophilic and solvophobic solute particles and is not intrinsically limited by the number of solute types in the system. However, each of these generalizations will modify the specific form of the improved LD potential of mean force and extension of the approach described here.

We now critically examine the improved model in terms of its level of complexity, the underlying assumptions and associated limitations. The model requires as input

1. the solute-solute MD potential $U_{AA}^{MD}$,

2. a form of the steady state solute-solvent pair correlation function $g_{AB}^{ss}$that is used to compute $U_{AB}$, and

3. the relative acceleration of a single pair of solutes from MD in order to determine the coefficient $C_2$.

This model is more complicated than existing LD models, but the need for these inputs is justified on the basis that the complex physico-chemical interactions in aggregation require these minimal inputs to be predictive. It is well established that aggregation behavior depends on whether the solute-solvent interactions are solvophobic or solvophilic [19, 39]. The key input that is needed to model this effect is the steady state solute-solvent pair correlation function $g_{AB}^{ss}$. The coefficient $C_2$ is a scale factor that is needed to quantitatively match the relative acceleration, and it arises from assumptions that are needed to derive this improved model. Without it we anticipate the model will qualitatively predict the correct trends, but it would not quantitatively match the MD data. The need for these inputs from two MD calculations is a less desirable feature of this improved model, even though these computationally inexpensive MD are for simple systems involving a single solute particle and pair of solutes. Later we discuss some approaches to remedy this aspect by developing analytical expressions for the steady state solute-solvent pair correlation function $g_{AB}^{ss}$and $C_2$. First we review the assumptions underlying this modeling approach.

The model is based on the assumption that the solute-solvent pair correlation attains a steady state on time scales that are much shorter than aggregation timescales. In other words, the arrangement of the solvent molecules relative to the solute does not significantly change even when aggregation of solute particles occurs. Indeed this hypothesis is supported by MD simulations as shown in Figure 3.6. It is conceivable that the interparticle interactions for some systems may violate this assumption, and in those cases this model would be inapplicable. The Kirkwood formula (cf. Eq. 3.18) that is used to infer the solute-solvent potential from the steady state solute-solvent pair correlation is strictly valid only for dilute systems

at equilibrium. For higher solvent molecule density, correction terms to the $g^{ss}_{AB}(r')$ need to be introduced [41]. However, these would significantly complicate the proposed model. For simplicity these are neglected. Another assumption is invoked when the mean force between a pair of solute particles that is induced by the solute-solvent interaction potential $U_{AB}(r')$ is computed. Specifically, three-particle densities such as $h(\mathbf{r}'|\mathbf{r})$ and $h(\mathbf{r}''|\mathbf{r})$ are approximated by two-particle densities $g^{ss}_{AB}(r')$ and $g^{ss}_{AB}(r'')$, respectively (see Appendix B for details). This assumption is strictly valid only for dilute systems. Since aggregation is a two–particle problem at low solute densities, this approach is reasonable and going to higher-order closures is probably not warranted. The model coefficient $C_2$ is a scale factor that accounts for the quantitative errors introduced by these assumptions.

We now examine approaches to analytically specify the model inputs without resorting to MD simulations for each specific system. One approach towards an analytical specification of $g^{ss}_{AB}$ is to propose a weighted form that reduces to appropriate limiting cases, such as: (a) aggregating system with solvophilic particles with $\varepsilon_{AA} > \varepsilon_{AB} > \varepsilon_{BB}$; (b) aggregating system with solvophobic particles with $\varepsilon_{AA} > \varepsilon_{BB} > \varepsilon_{AB}$; (c) non-aggregating system with solvophilic particles with $\varepsilon_{AA} \approx \varepsilon_{AB}$, and $\varepsilon_{AB} > \varepsilon_{BB}$; (d) non-aggregating system with solvophobic particles with $\varepsilon_{AA} \approx \varepsilon_{AB}$, and $\varepsilon_{AB} < \varepsilon_{BB}$. It is worthwhile to examine if the $C_2$ could also be specified as an analytical function of the physical parameters of the problem for each of the limiting cases described above. The modeling problem is to determine how this coefficient depends on the pair–potential well depths $\varepsilon_{AA}$ and $\varepsilon_{AB}$ (or $\varepsilon_{AB}/\varepsilon_{AA}$), the solute number density $n_A$ and the solvent number density $n_B$. It is unclear if the dependence on the potential well–depths is necessary because the effect of the potentials is already accounted for in the $\tilde{U}_2$ term. It is probably also important to characterize the dependence of $C_2$ on the size ratio of solute to solvent. If an analytical form of $C_2$ can be developed for the limiting cases (a)–(d), then the coefficient $C_2$ for any problem could also be modeled as a weighted average. Finally, the limit $C_2 \to 0$ should be imposed when $\varepsilon_{AB} \ll \varepsilon_{AA}$, which is the case when solute-solvent interaction is negligible and the solvation correction should tend to zero.

Present work does not include the effect of hydrodynamic interactions (HI) since HI effect is

not significant for equisize systems considered is this work. However, even if HI forces were to be accounted for in a BD (Langevin dynamics) model, these would appear in the frictional term (first term on the right hand side of Eq 3.5), that does not negate the contribution of present work which concerns the improved specification of potential of mean force (second term on the right hand side of Eq 3.5). Our estimates indicate that the correction force arising from our improved BD potential that accounts for solvent interactions are important in aggregation, even if HI forces are included in size–separated systems. This point is elaborated more in Appendix G.

## 3.9 CONCLUSIONS

We simulated aggregation in a model system using both molecular dynamics and Langevin dynamics in order to determine the accuracy of the LD model using the MD as a benchmark. The structure of the aggregates is inferred from the solute–particle pair–correlation function. It is found that using the unmodified MD solute-solute pair potential in LD results in accurate prediction of the aggregate structure in the diffusion–limited regime, but not in the reaction–limited regime. This finding motivates the development of an improved LD model for the potential of mean force between solute particles. The transport equation for the solute particle pair–correlation function (or second–order density) informs us that improved modeling of the relative acceleration between a pair of solute particles in solvent is necessary for accurate prediction of the aggregate structure. We propose an improved model for the potential of mean force in LD by decomposing the relative acceleration between a pair of solute particles into solute-solute and solute-solvent interactions. We exploit the fact that the solute-solvent pair-correlation rapidly reaches a steady state (relative to aggregation timescales), and we approximate the solute-solvent potential using the standard Kirkwood formula. The effect of the solute-solvent interaction on solute relative acceleration is semi-analytically computed by integrating the effect of a test solvent molecule on a pair of solute particles. Incorporating the effect of the solvent in this manner leads us to an improved specification of the potential of mean force between solute particles in the LD model. The improved

LD model dramatically improves results for the aggregate structure in both reaction–limited and diffusion–limited regimes. Moreover, the proposed model allows microscale interactions to be related to mesoscale interactions, thereby addressing a critical need in multiscale simulation. This improved LD model also gives better prediction of the cluster size distribution in both regimes. The model has the capability of representing the effect of solvent on aggregation in both solvophilic and solvophobic systems.

## 3.10 ACKNOWLEDGMENTS

This work has been partially supported by the National Science Foundation through Grant No. CTS-0403864. The authors would like to thank Monica Lamm, Dennis Vigil and Rodney Fox for useful discussions. Computer time on the IBM Blue GENE/L at Iowa State University is gratefully acknowledged.

## Bibliography

[1] Lattuada, M., Wu, H., Sandfuhler, P., Sefcik, J., and Morbidelli, M. Modelling of aggregation kinetics of colloidal systems and its validation by light scattering measurements. *Chem. Eng. Science*, 59:1783–1798, 2004.

[2] Selomulya, C., Bushell, G., Amal, R., and Waite, T. D. Aggregation mechanisms of latex of different particle sizes in a controlled shear environment. *Langmuir*, 18:1974–1984, 2002.

[3] Fry, D., Chakrabarti, A., Kim, W., and Sorensen, C. M. Structural crossover in dense irreversibly aggregating particulate systems. *Phys. Rev. E*, 69:061401, 2004.

[4] Ulberg, D. E., Churaev, N. V., Ilyin, V. V., and Malashenko, G. L. Molecular-dynamics simulation of the aggregation of colloidal particles. *Colloids & Surfaces A Physicoch. & Engin. Aspects*, 80:93–102, 1993.

[5] Witten, T. A., and Sander, L. M. Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.*, 47:1400–1403, 1981.

[6] Meakin, P., and Vicsek, T. Internal structure of diffusion-limited aggregates. *Phys. Rev. A*, 32(1):685–688,1985

[7] Meakin, P. Formation of fractal clusters and networks by irreversible diffusion-limited aggregation. *Phys. Rev. Lett.*, 51:1119–1122, 1983.

[8] Kolb, M., Botet, R., and Jullien, R. Scaling of kinetically growing clusters. *Phys. Rev. Lett.*, 51:1123–1126, 1983.

[9] Vold, M. J. Computer simulation of floc formation in a colloidal suspension. *J. Colloid & Sci.*, 18:684–695, 1963.

[10] Sutherland, D. N. A theoretical model of floc structure. *J. Colloid & Inter. Sci.*, 25:373–380, 1967.

[11] Jullien, R., and Kolb, M. Hierarchical model for chemically limited cluster-cluster aggregates. *Journal of Physics*, A17:L639–L643, 1984.

[12] Markutsya, S., Subramaniam, S., Vigil, R. D., and Fox, R. O. On Brownian dynamics simulation of nanoparticle aggregation. *Ind. Eng. Chem. Res.*, 47:3338–3345, 2008.

[13] Ermak, D. L. Computer-simulation of charged-particles in solution .1. Technique and equilibrium properties. *J. Chem. Phys.*, 62:4189–4196, 1975.

[14] Turq, P., Lantelme, F., and Friedman H. L. Brownian dynamics - its application to ionic solutions. *J. Chem. Phys.*, 66:3039, 1977.

[15] Hoogerbrugge, P. J., and Koelman, J. M. V. A. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.*, 19:155–160, 1992.

[16] Español, P., and Warren, P. B. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.*, 30:191–196, 1995.

[17] Malevanets, A., Kapral, R. Mesoscopic model for solvent dynamics. *J. Chem. Phys.*, 110:8605–8613, 1999.

[18] Giro, A., and Guardia, E. Langevin and molecular dynamics simulations of Lennard-Jones liquids. *Mol. Phys.*, 55(5):1063–1074, 1985.

[19] Qin, Y., and Fichthorn, K. A. Molecular-dynamics simulation of forces between nanoparticles in a Lennard-Jones liquid. *J. Chem. Phys.*, 119 (18):9745–9784, 2003.

[20] Qin, Y., and Fichthorn, K. A. Molecular-dynamics simulation of the forces between colloidal nanoparticles in n-decane solvent. *J. Chem. Phys.*, 127 (18):144911, 2007.

[21] Velegol, D., Holtzer, G. L., Radovic-Moreno, A. F., and Cuppet J. D. Force measurements between sub-100 nm colloidal particles. *Langmuir*, 123:1275–1280, 2007.

[22] Peter, C., and Kremer, K. Multiscale simulation of soft matter systems - from the atomistic to the coarse–grained level and back. *Soft Matter*, 5:4357–4366, 2009.

[23] Ruhle V., Junghans, C., Lukyanov, A., Kremer, K., and Andrienko, D. Versatile object–oriented toolkit for coarse–graining applications. *J. Chem. Theory Comput.* 5:3211–3223, 2009.

[24] Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., and de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B*, 111:7812–7824, 2007.

[25] Monticelli, L., Kandasamy, S. K., Periole, X., Larson, G. R., Tieleman, D. P., and Marrink, S. J. The MARTINI coarse–grained force field: extension to proteins. *J. Chem. Theory Comput.*, 4:819–834, 2008.

[26] Lyubartsev, A. P., and Laaksonen, A. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Phys. Rev. E*, 52(4):3730–3737, 1995.

[27] Muller-Plathe, F. Coarse–graining in polymer simulation: from the atomistic to the mesoscopic scale and back. *Chem. Phys. Chem* , 3:754–769, 2002.

[28] Peter, C., Delle Site, L., and Kremer, K. Classical simulations from the atomistic to the mesoscale and back: coarse graining an azobenzene liquid crystal. *Soft Matter* 4:859–869, 2008.

[29] Izvekov, S., and Voth G. A. Multiscale coarse graining of liquid–state systems. *J. Chem. Phys.* 123:134105, 2005.

[30] Mokhtari, T., Chakrabarti, A., Sorensen, C. M., Cheng, C., and Vigil, D. The effect of shear on colloidal aggregation and gelation studied using small-angle light scattering. *J. Coll. Interf. Science*, 327:216–223, 2008.

[31] Veld, P. J., Plimpton, S. J., and Grest, G. S. Accurate and efficient methods for modeling colloidal mixtures in an explicit solvent using molecular dynamics. *Comp. Phys. Comm*, 179:320–329, 2008.

[32] Qin, Y., and Fichthorn, K. A. Solvophobic solvation at large and intermediate length scales: Size, shape, and solvent effects. *Phys. Rev. E*, 74:020401(R), 2006.

[33] Fichthorn, K. A., and Qin, Y. Molecular dynamics simulation of the forces between colloidal nanoparticles in Lennard-Jones and n-decane solvent. *Granular Matter* 10:105–111, 2008.

[34] Plimpton, S. J. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.*, 117:1–19, 1995.

[35] Stoyan, D., and Stoyan, H. *Fractals, random shapes, and point fields : methods of geometrical statistics*; John Wiley & Sons, 1994.

[36] Pai, M. G. and Subramaniam, S. Second-order transport due to fluctuations in clustering particle systems. *60th Annual Meeting of the Division of Fluid Dynamics, The American Physical Society*, Salt Lake City, UT, November 2007.

[37] Likos, C.N., Effective interactions in soft condensed matter physics. *Physics Reports*, 348:267–439, 2001.

[38] Snook, I.K., and Megen, W., Calculation of solvation forces between solid particles immersed in a simple liquid. *J. Chem. Soc. Faraday Trans. II*, 77:181–190, 1981.

[39] Sheng, Y-J., Wang, T-Y., Chen, W. M., and Tsao, H-K., A-B diblock copolymer micelles: effects of soluble-block length and component compatibility. *J. Phys. Chem. B*, 111:10938–10945, 2007.

[40] Vicsek, T. and Family, F. Dynamic scaling for aggregation of clusters. *Phys. Rev. Lett.*, 52:1669-1672, 1984.

[41] McQuarrie, D.A. *Statistical mechanics*; Harper Collins Publishers, 1976.

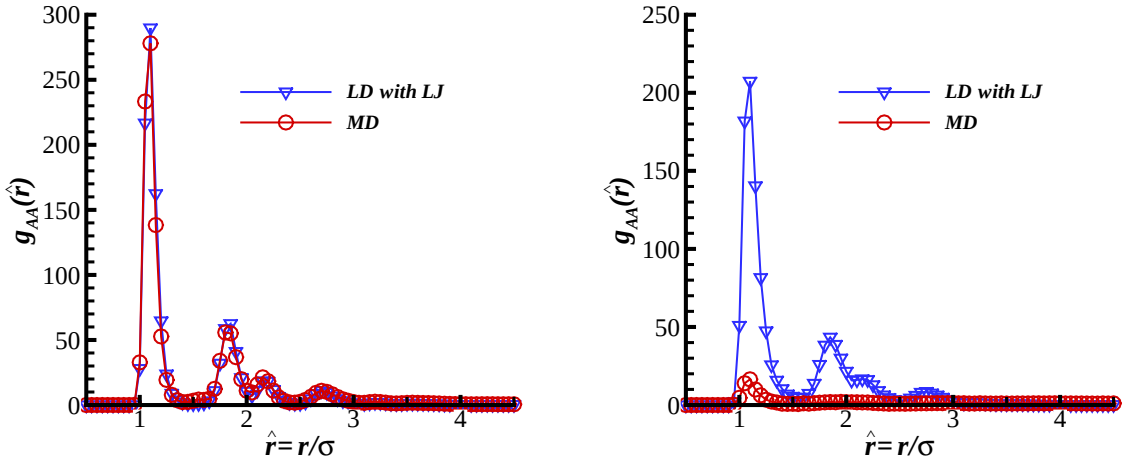



Figure 3.1: Comparison of $g_{AA}(\hat{r})$ predicted by the LD model with MD simulation data at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$.

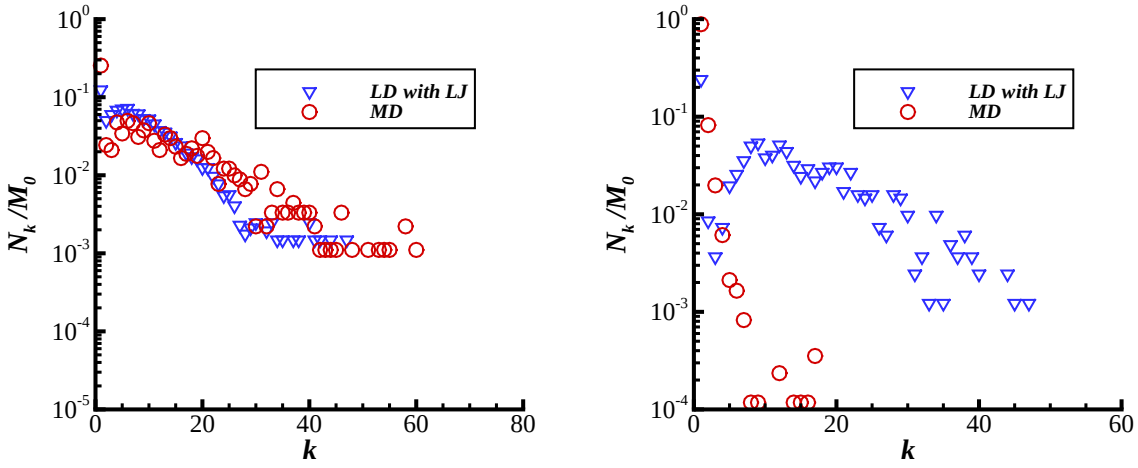



Figure 3.2: Comparison of the cluster size distribution predicted by the LD model with corresponding MD simulation data at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$

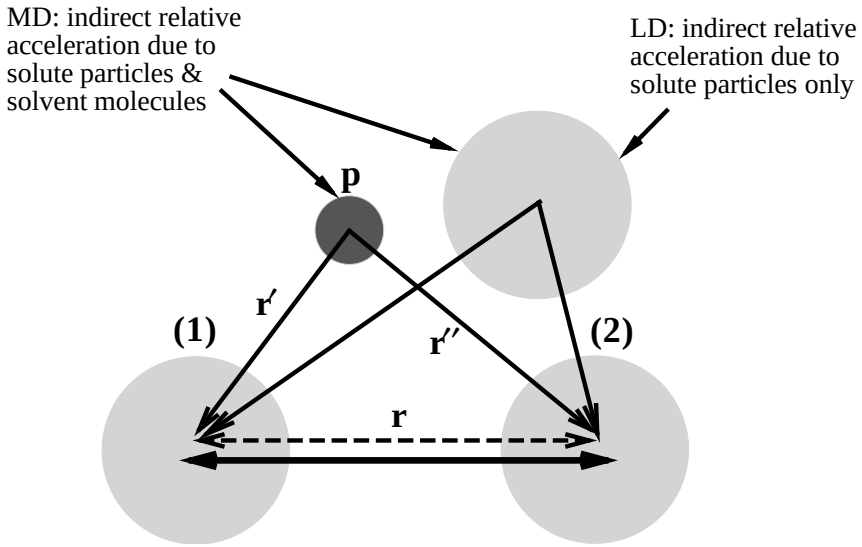



Figure 3.3: Direct relative acceleration between solute particles (1) and (2) (solid arrow between particles (1) and (2)) occurs due to their direct interaction. Indirect relative acceleration between solute particles (1) and (2) (dashed arrow) occurs (a) in MD due to interaction of particles (1) and (2) with probe solute particle and probe solvent molecule $p$; (b) in LD due to interaction of particles (1) and (2) with probe solute particle only.

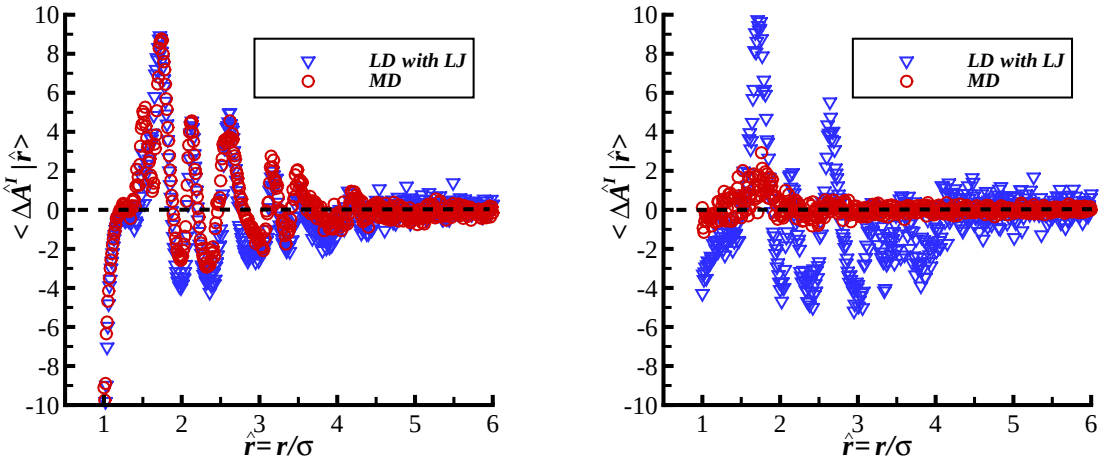



Figure 3.4: Indirect average relative acceleration between $A$-$A$ pairs resulting solely from other $A$-type particle interactions. MD simulations (10,000 $A$-type solute particles and 813,218 $B$-type solvent particles) compared with LD model predictions (10,000 $A$-type solute particles) at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$. Indirect average relative acceleration is scaled as $\langle \Delta \hat{\mathbf{A}}^I | \hat{\mathbf{r}} \rangle = \langle \Delta \mathbf{A}^I | \mathbf{r} \rangle \, \sigma_A m_A / \varepsilon_{AA}$.

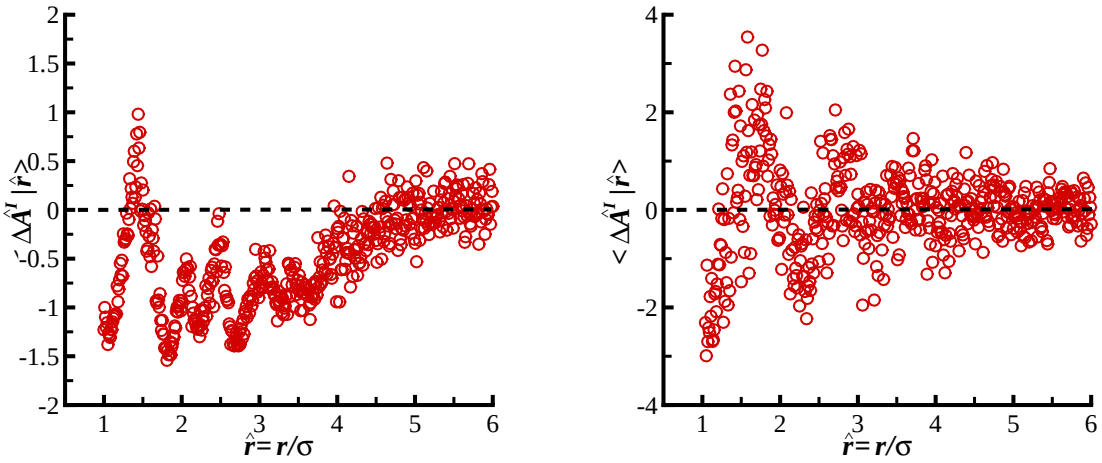



Figure 3.5: Indirect average relative acceleration between $A$-$A$ pairs resulting solely from $B$ particles. MD simulation of 10,000 $A$-type solute particles and 813,218 $B$-type solvent particles at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$.

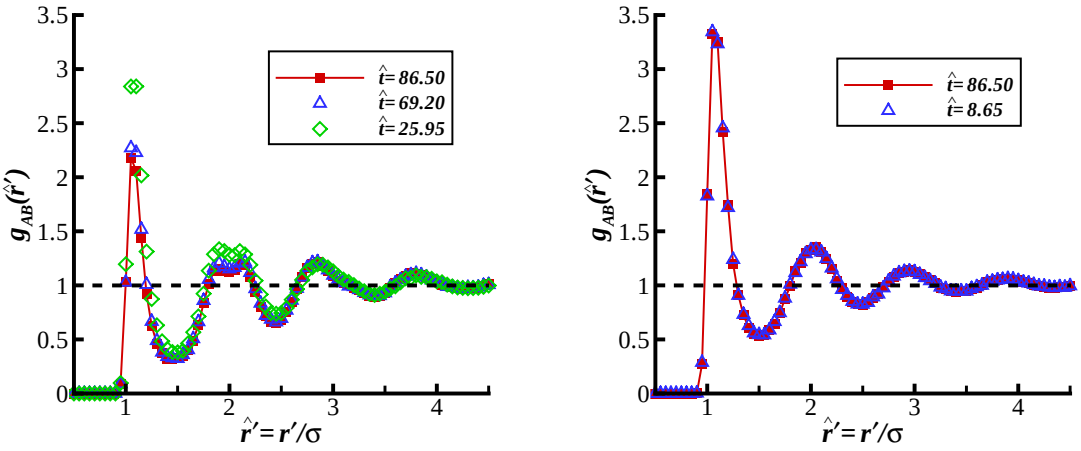



Figure 3.6: Evolution of the pair correlation function $g_{AB}(\hat{r}', t)$ from MD simulation: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$. Scaled time $\hat{t} = t\, D_\infty/\sigma^2$.

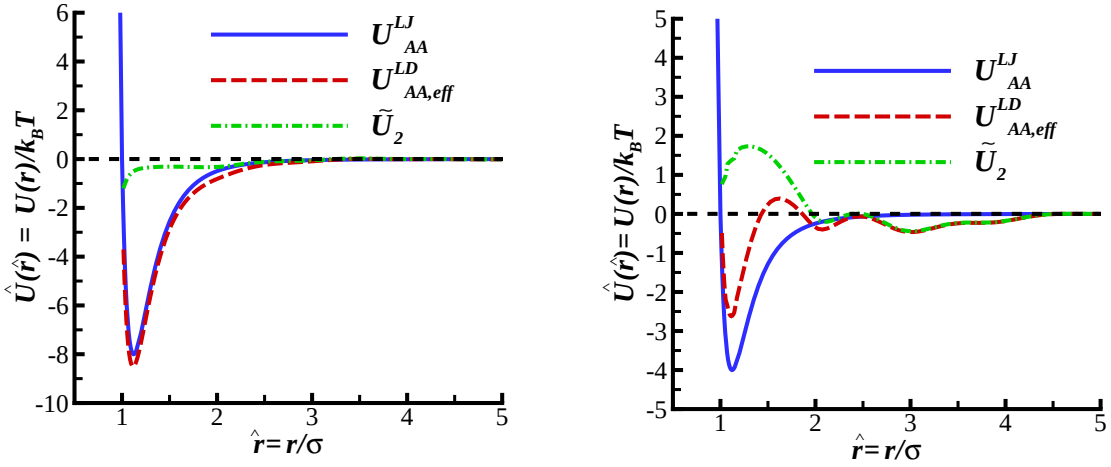



Figure 3.7: Comparison of the improved LD potential $U^{LD}_{AA,\text{eff}}$ and the modeled solvation potential $\tilde{U}_2$ with the Lennard-Jones potential $U^{LJ}_{AA}$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$, $C_2 = 0.51$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$, $C_2$=3.15.

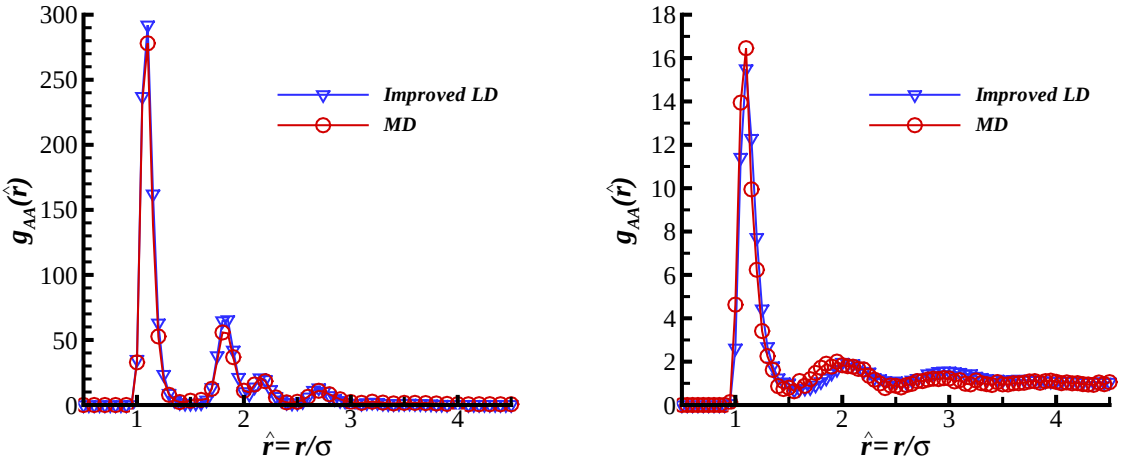



Figure 3.8: Comparison of $g_{AA}(\hat{r})$ predicted by improved LD model with corresponding MD result at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$.

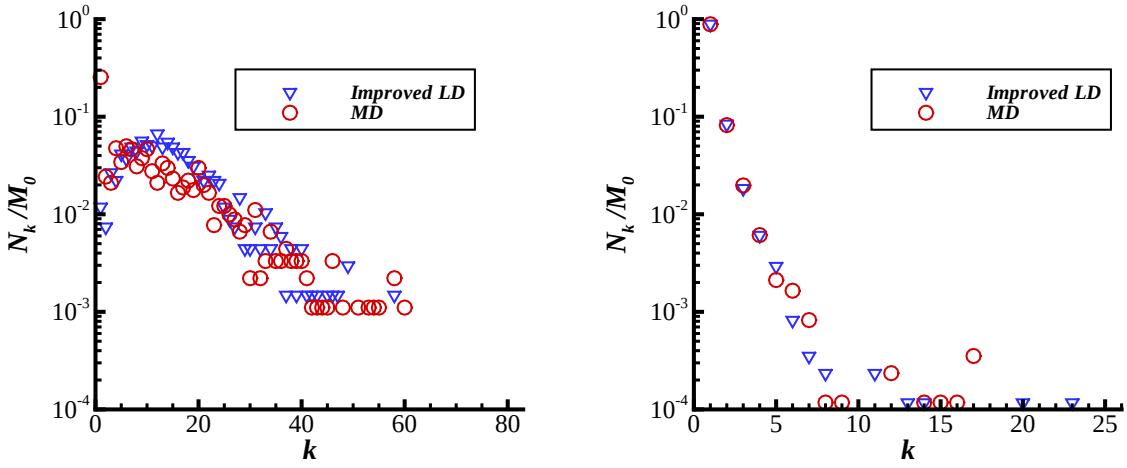
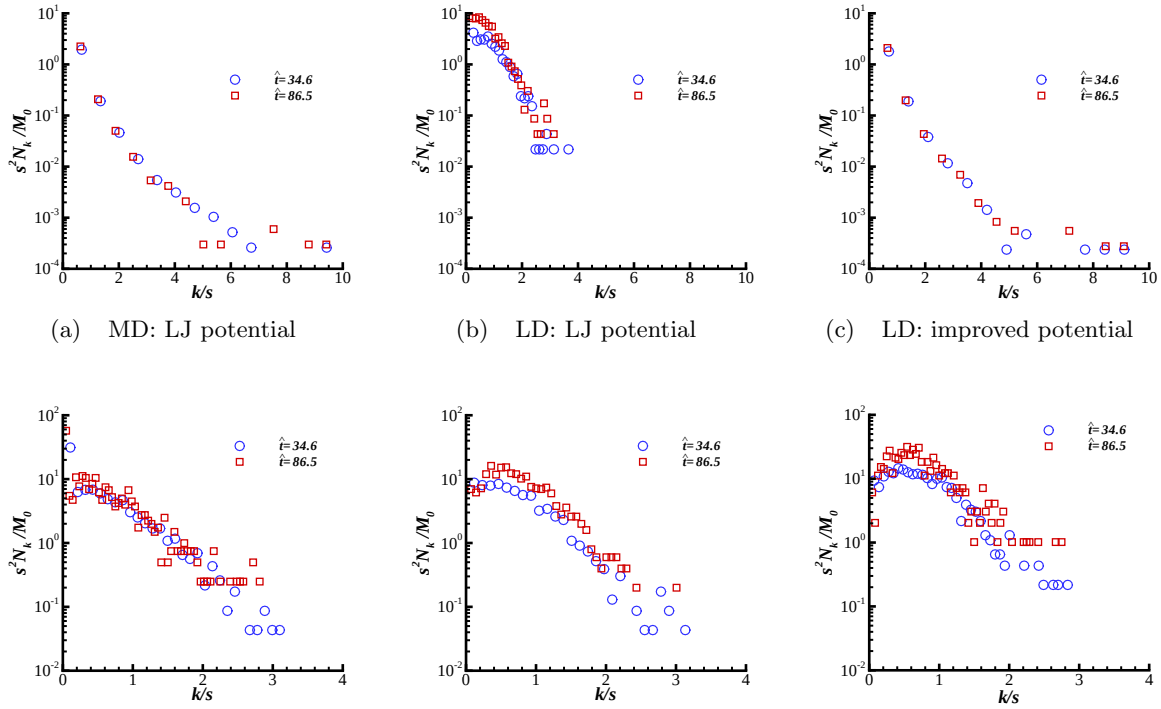


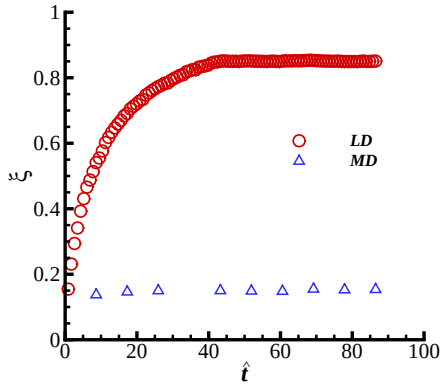

Figure 3.9: Comparison of the cluster size distribution predicted by improved LD model with corresponding MD result at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$.



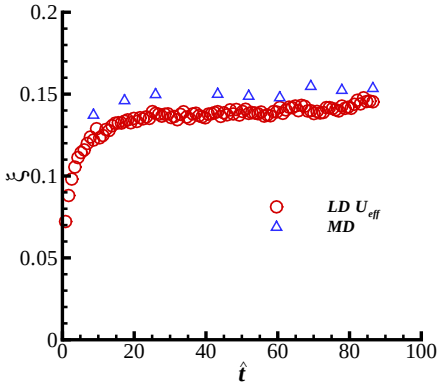

(a) MD: LJ potential (b) LD: LJ potential (c) LD: improved potential

(d) MD: LJ potential (e) LD: LJ potential (f) LD: improved potential

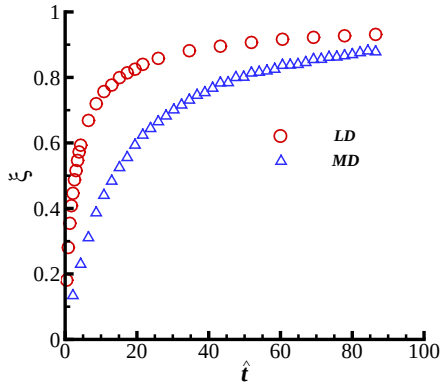
Figure 3.10: Normalized cluster size distributions in the RLA regime $\hat{\varepsilon}_{AA}/\hat{\varepsilon}_{BB} = 4.0$ [top panel (a),(b) and (c)], and the DLA regime $\hat{\varepsilon}_{AA}/\hat{\varepsilon}_{BB} = 8.0$ [bottom panel (d), (e) and (f)] at different times $\hat{t} = tD_\infty/\sigma^2$ for MD simulations with LJ potential [left column (a) & (d)]; LD simulations with LJ potential [middle column (b) & (e)]; LD with improved potential [right column (c) & (f)].



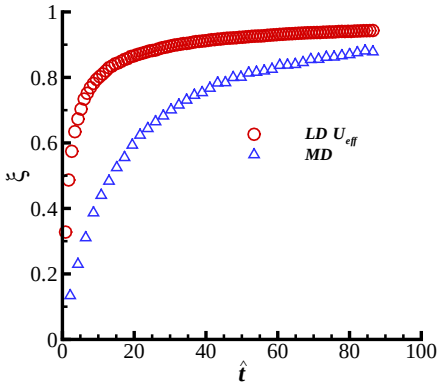

(a) MD and LD (LJ potential): RLA




(b) MD and LD (improved potential): RLA




(c) MD and LD (LJ potential): DLA




(d) MD and LD (improved potential): DLA

Figure 3.11: Comparison between MD and LD (unmodified LJ and improved potential) of the extent of aggregation $\xi$ for RLA regime: $\hat{\varepsilon}_{AA}/\hat{\varepsilon}_{BB} = 4.0$ [top panel: (a) and (b)], and DLA regime: $\hat{\varepsilon}_{AA}/\hat{\varepsilon}_{BB} = 8.0$ [bottom panel: (c) and (d)]. LD simulations with unmodified LJ potential are compared with MD in the left column [(a) & (c)], while LD with the improved potential is compared with MD in the right column [(b) & (d)].

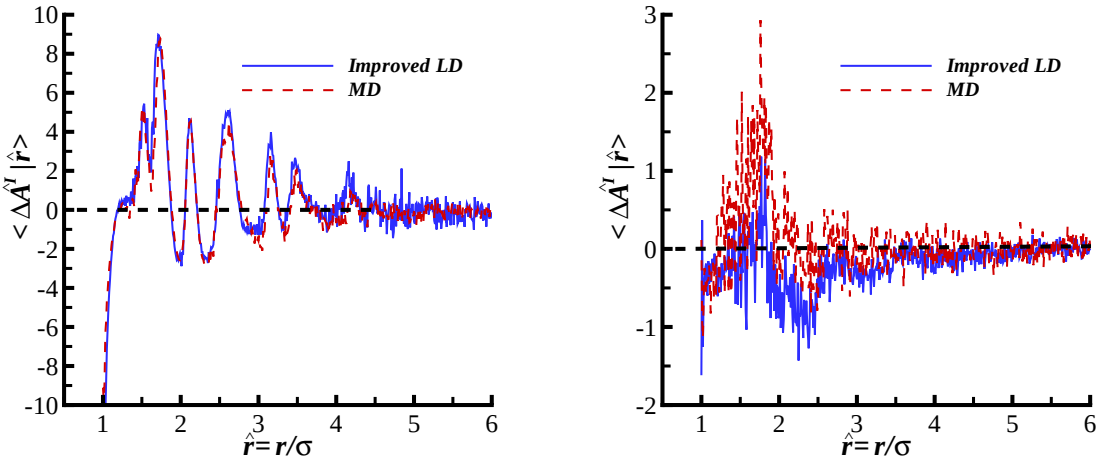



Figure 3.12: Indirect average relative acceleration between $A$-$A$ pairs resulting solely from other $A$ particle interactions. Improved LD model (10,000 $A$-type particles) compared with MD simulation (10,000 $A$-type particles and 813,218 $B$-type particles) at time $\hat{t} = 86.5$: a) DLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 8.0$; b) RLA regime: $\varepsilon_{AA}/\varepsilon_{BB} = 4.0$.

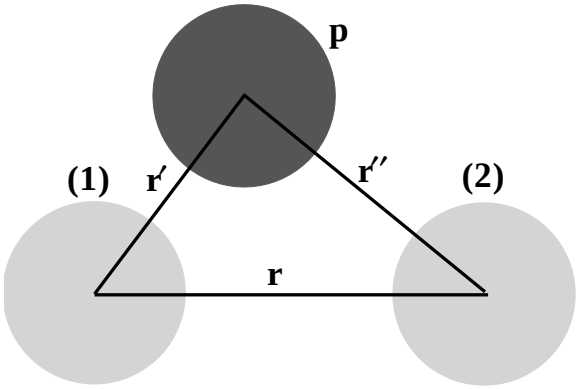



Figure 3.13: Schematic showing a "probe" solvent molecule $p$ which can occupy any point in 3-d space except volumes of solute particles 1 and 2, thus defining the domain of integration for the relative acceleration calculation.

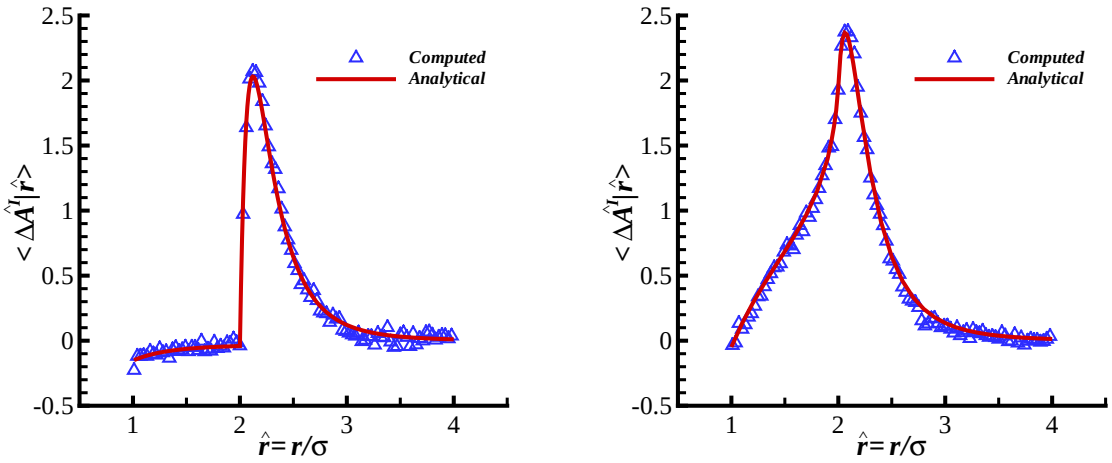



Figure 3.14: Comparison of computed indirect average relative acceleration with the analytical result at $\hat{n} = 0.1$: a) 1-d case, computations with 150,000 particles averaged over 3,000 multiple independent trials, b) 2-d case, computations with 823,000 particles averaged over 240 multiple independent trials.

# CHAPTER 4. EFFECT OF SHEAR ON COLLOIDAL AGGREGATION OF A MODEL SYSTEM USING LANGEVIN DYNAMICS SIMULATION

This chapter is a manuscript in preparation titled "Effect of Shear on Solloidal Aggregation of a Model System Using Langevin Dynamics Simulation" co–authored with S. Subramaniam.

## 4.1 Introduction

Aggregation of colloidal nanoparticles is a non-equilibrium multiscale problem that is characterized by a wide range of length and time scales ranging from those associated with a monomer to superaggregates comprising tens of thousands of monomers. Usually colloidal aggregation in real physical systems occurs in the presence of external forces due to gravity [1, 2] or shear flow [3, 4, 5, 6, 7, 8, 9, 10, 11]. In this work we focus on the effects of shear. Shear flow affects the size and structure of aggregates, and the rate at which they are formed. A rich variety of phenomena are observed depending on the shear rate, the initiation time [11] (i.e., whether shear is applied after some aggregation has taken place, or right from the onset of aggregation) and the duration of time over which the system is subjected to shear.

These phenomena have been investigated experimentally by studying the influence of externally applied shear on the aggregation of latex nanoparticles [8, 11]. If shear is applied after aggregates have already formed it can change aggregate structure. In experiments it has been found that moderate shear flow (characterized by the Pèclet number based on the monomer diameter $\mathrm{Pe}_\sigma = 1 \div 5$) results in a more compact aggregate structure than that found in non-sheared systems [3, 8, 11]. In these experiments, changes in aggregate structure are inferred from the fractal dimension of the aggregates $d_f$ that is obtained from light scattering analysis. However, 2D LD simulations of Cerda *et al.* [12] report that shear weakens the effective

interaction between particles by stretching the separation between primary particles in aggregates, thereby forming less compact structures. These changes in structure are characterized by computing a mean interaction energy attributed to the physical bonds that hold the cluster together [12]. At this point it is not clear why in some cases shear results in the formation of more compact aggregates with higher fractal dimension $d_f$, while in others it produces less compact structures,.

Computational approaches are well suited to answering such questions because one can easily vary parameters such as dimensionless interaction well depth and the Pèclet number to investigate their effect on aggregation outcomes. In earlier work we showed that although molecular dynamics (MD) simulations of aggregation in dilute systems with full solvent interactions are still too computationally expensive, mesoscale methods such as Langevin dynamics (LD) with modeled solvent interactions scale favorably to larger systems while retaining the capability of representing structure in aggregating systems [16]. A coarse–graining procedure we recently developed to specify the potential of mean force in LD for aggregating systems yields time–evolving structure in non–equilibrium aggregating systems that matches very well with MD simulations in both diffusion–limited and reaction–limited regimes [13]. With this improved potential of mean force, we have confidence that LD simulations of aggregation reliably predict important aggregation statistics such as the extent of aggregation, time–evolving solute pair correlation function, and dynamically scaled cluster size distribution when compared with MD simulations of smaller model systems [13]. The essential features of this improved LD model are described in Section 4.2. In this work we use this improved LD model to study sheared aggregating systems.

The effect of shear on aggregate structure is not easy to characterize because shear affects aggregates differently on smaller length scales (on the order of ten monomer diameters) as compared to larger length scales (on the order of hundred monomer diameters). The reason for this probably lies in the mechanism underlying the formation of aggregates at different scales. As shown by Sorensen and co–workers [11], the small–scale structure of aggregates arises from aggregation by monomer addition, while the large–scale structure corresponds to

cluster-cluster aggregation. Here again mesoscale computational approaches are well suited to addressing the influence of shear on aggregate structure at different scales because we can easily control the relative magnitude of scales (length, time, velocity and force) by appropriately choosing simulation parameters .

We present a scale analysis of the sheared aggregation problem in Section 4.3 where we identify length, time and velocity scales corresponding to micro, meso and macroscales. Starting from a set of physical parameters $\mathscr{P}$ that characterize the sheared aggregating system, we use dimensional analysis to identify a (nonunique) set of dimensionless parameters $\Pi$. We show that a subset of these dimensionless parameters that corresponds to length scale and time scale ratios is useful in demarcating scale–separated and scale–overlap regimes. Lekkerkerker et al. [14] have proposed a phase diagram for aggregating systems. Also Chakrabarti et al. [15] classified aggregation regimes into (a) system at equilibrium, (b) non–equilibrium aggregation, and (c) gelation. Inspired by these works we seek to identify the most important dimensionless parameters from the set $\Pi$ that will enable us to construct a regime map for sheared aggregation. We propose to use LD simulations to identify the metrics that characterize aggregation outcomes so that we can distinguish different aggregation outcomes on this regime map. Identifying the relevant metrics requires a fundamental understanding of how shear affects aggregate structure.

Shear affects aggregation by providing additional kinetic energy to particles and aggregates, thereby increasing the rate of aggregation [11, 12]. This motivates an energy analysis of the LD simulation data for systems aggregating under shear to link aggregate restructuring to energy transfer. The energy balance in sheared aggregating systems is analyzed in Section 4.4. We characterize the changes in aggregate structure due to shear in Section 4.3, and describe how the redistribution of energy corresponds to structural changes in the aggregates. Using the insights gained from the energy budget analysis in Section 4.4, we propose new metrics in Section 4.5.3 to characterize the compactness of anisotropic aggregates that are formed under shear.

Aggregate breakage is observed [6, 9] when sufficiently high shear is imposed. Thus shear

limits the maximum size of aggregates [3, 6, 7, 8, 9], with the maximum attainable size decreasing with increasing shear rate. This effect is qualitatively understood but a quantitative prediction of the maximum size of aggregates has not yet been proposed to the best of our knowledge. In Section 4.6 we propose a simple mechanistic model for shear–induced aggregate breakage by computing the relative magnitude of shear force to interparticle attraction force. This simple model gives good predictions for the maximum size of aggregates in sheared systems. In Section 4.7 we combine the dimensional analysis of Section 4.3 with metrics obtained from LD simulations in Section 4.5.3 to propose a regime map for sheared aggregation that demarcates different aggregation regimes and enables us to classify sheared aggregating systems based on their initial physical parameters.

## 4.2 Improved Langevin Dynamics Simulation

Langevin dynamics (LD) is a mesoscale simulation method that can capture the effect of shear on aggregate restructuring [13]. The Langevin dynamics model is used to simulate aggregation of solute particles immersed in liquid solvent that is subjected to uniform (spatially homogeneous), steady shear flow. Langevin dynamics allows us to simulate the evolution of larger systems of solute particles than MD [16] because in LD the solute-solvent interactions are modeled through frictional and random terms, and through modification of the solute pair interaction potential in the presence of solvent molecules [13]. For a model system we consider, the relative magnitude of time scales corresponding to the frictional and pairwise interaction force terms requires evolution of both position and velocity Langevin equations for accurate LD simulations [16, 17, 18]. The LD equations for evolution of the position $r_\alpha^{(i)}$ and velocity $v_\alpha^{(i)}$ of the $i$-th solute particle in a sheared solvent flow are

$$dr_\alpha^{(i)} = v_\alpha^{(i)} dt, \qquad \alpha = 1, 2, 3, \tag{4.1}$$

$$dv_\alpha^{(i)} = -\gamma v_\alpha^{(i)} dt + \gamma u_\alpha^{(i)} dt + \frac{1}{m^{(i)}} F_\alpha(r^{(i)}) dt + \sqrt{2\gamma}\sigma_{v_\infty} dW_\alpha^{(i)}, \quad i = 1, \ldots, N, \tag{4.2}$$

where $\alpha$ represents Cartesian coordinates, $m^{(i)}$ is the mass of $i$-th particle, $u_\alpha^{(i)} = G r_\beta^{(i)} \delta_{1\alpha}\delta_{2\beta}$, $(\alpha, \beta = 1, 2, 3)$ is the mean fluid velocity (of solvent molecules) due to imposed shear flow,

$G$ is the uniform shear rate, $\gamma = k_B T_{\text{ref}}/m^{(i)} D_\infty$ is the friction coefficient, $D_\infty$ is the self-diffusion coefficient of the solute particles in solvent at infinite dilution, $\sigma^2_{v_\infty} = k_B T_{\text{ref}}/m^{(i)}$ is the stationary velocity variance, $dW^{(i)}_\alpha$ is a Wiener process increment, $N$ is the total number of solute particles, and $F(\mathbf{r})_\alpha = -\nabla_{\mathbf{r}} U^{LD}$ where $U^{LD}$ is the effective LD potential of mean force between solute particles in the presence of solvent. This effective LD potential is calculated according to a coarse–graining method developed by Markutsya and Subramaniam [13] as

$$U^{LD}(r_{ij}) = U^{LJ}(r_{ij}) + \tilde{U}_2(r_{ij}) = \begin{cases} 4\varepsilon\left[\left(\dfrac{\sigma}{r_{ij}}\right)^{12} - \left(\dfrac{\sigma}{r_{ij}}\right)^{6}\right] + \tilde{U}_2(r_{ij}), & r_{ij} \le r_{\text{cut}} \\ \\ 0 & r_{ij} > r_{\text{cut}}, \end{cases} \tag{4.3}$$

where $\sigma$ is the particle diameter, $r_{ij}$ is the scalar separation or distance between centers of particles $i$ and $j$, $r_{\text{cut}}$ is the cutoff distance chosen to be $2.5\,\sigma$, $\tilde{U}_2$ is the correction term that takes into account effect of solvent molecules on interparticle interaction potential [13], and $\varepsilon$ is the potential well-depth between particles. The self-diffusion coefficient at infinite dilution $D_\infty$ and the correction to the interparticle potential $\tilde{U}_2(r_{ij})$ are required input parameters for the improved LD model. The self-diffusion coefficient at infinite dilution $D_\infty$ is extracted directly from MD simulation of the corresponding system, and the correction to the interparticle potential $\tilde{U}_2(r_{ij})$ is obtained using the semi–analytical coarse–graining procedure described previously [13].

The LD simulations are performed using the LAMMPS [19] software package. The initial spatial configuration of the particles is specified to ensure non–overlapping particles. This is accomplished by spatially distributing the solute particles according to a hard–core Matérn point process [20]. The solute particles are assigned a Maxwellian velocity distribution corresponding to their reference temperature $T_{\text{ref}}$. From this configuration the system is allowed to equilibrate to the initial condition for the aggregation simulations by allowing particles to interact with a Lennard–Jones potential with dimensionless well–depth $\hat{\varepsilon} = \varepsilon/k_B T_{\text{ref}} = 1$, where $k_B$ is the Boltzmann constant, and $T_{\text{ref}}$ is the reference temperature.

## 4.3 Analysis of Colloidal Aggregation Under Shear

Even though we do not represent the solvent molecules explicitly in LD simulations, the aggregation of colloidal particles in the presence of shear is studied. Still introduces a wide range of length scales from monomer size (tens of nanometers) to size of large aggregates (order of micrometers). There is also a wide range of time scales ranging from Brownian motion of particles (order of femtoseconds) and the fluid time scale associated with shear (order of milliseconds).

A system of colloidal particles aggregating in the presence of shear can be characterized by a set of physical parameters $\mathscr{P}$ that includes characteristic scales of length, time, velocity, and interparticle force. Aggregation introduces clusters with characteristic length scale $R_g$ which is meso scale and shear flow introduces macro scales as represented in Table 4.1. A dimensional analysis based on the Buckingham $\Pi$ theorem allows us to reduce these physical parameters to a non–unique set of dimensionless parameters $\Pi$. The dimensionless parametes that represent length and time scale rations of macro (or meso) to meso (or micro) scales are useful in characterizing scale–separated or scale–overlap regimes. However, it is not known *a priori* which set of dimensionless parameters is most useful for characterizing aggregation outcomes. While metrics such as the radius of gyration $R_g$, fractal dimension $d_f$, extent of aggregation $\xi$ are used for non–sheared systems, the appropriate metrics for sheared aggregating systems is not yet established.

Table 4.1: Micro, meso, amd macro scales. Where $\tau_v$ is the characteristic time, $R_g$ is the aggregate radius of gyration, $L$ is the box length, and $G$ is the shear rate.

| | Micro | Meso | Macro |
|---|---|---|---|
| Length | $\sigma$ | $R_g$ | $L$ |
| Time | $\sigma/\sigma_{v_\infty}$ | $\tau_v$ | $G^{-1}$ |
| Velocity | $\sigma_{v_\infty}$ | $R_g/\tau_v$ | $LG$ |

Although the procedure of determining physical and dimensionless parameters for the sheared aggregating systems is well known, an aggregation regime map that would uniqly

identify aggregation outcome for such systems does not exist. In this paper we describe an attempt to create an aggregation regime map that would identify the difference in aggregating outcomes for the systems characterized by different input values of dimensionless characteristics.

### 4.3.1 Scale–separated and Scale–overlap Regimes

For sheared aggregating systems, the dimensional physical parameters are scaled to corresponding non-dimensional counterparts in the manner described earlier for non-sheared aggregating systems [16] (see Appendix A for details). However, because shear is added to the aggregating system, an additional dimensionless parameter (Pèclet number Pe) related to the shear rate $G$ is introduced and defined as

$$\mathrm{Pe}_\sigma = \frac{1}{4}\frac{G\sigma^2}{D_\infty}, \tag{4.4}$$

where $\sigma$ is the particle diameter, and $D_\infty$ is the self–diffusion coefficient of particle. Superscript $(\sigma)$ identifies that Pèclet number is calculated for a monomer particle of size $\sigma$.

When shear flow is introduced into aggregation systems it creates an additional time scale associated with the shear rate $G$ and additional length scale $R_g^{max}$ . This additional length scale arises because the shear flow limits the maximum size of aggregates. It should be clarified that aggregating system may be characterized by different sets of parameters: set of the *microscale* parameters such as velocity autocorrelation time scale for a single particle $\tau_v^{(1)}$, and particle size $\sigma$. Or set of *mesoscale* parameters such as velocity autocorrelation time scale for a cluster containing $k$ monomers $\tau_v^{(k)}$, and a radius of gyration of aggregate $R_g$ as described in Table 4.1. In this paper the aggregation outcome is characterized by a set of microscale parameters. Since in a sheared aggregating system two time scales such as shear time scale (represented by $1/G$) and diffusion time scale (represented by the velocity autocorrelation time scale $\tau_v^{(1)}$) the ratio of these time scales $G\tau_v^{(1)}$ corresponds to the Deborah number $De$ and can be proposed as a characteristics of aggregation outcome. This ratio is defined based on the diffusion time scale

for a single particle as

$$G\tau_v^{(1)} = 4\mathrm{Pe}_\sigma \frac{D_\infty}{\sigma^2}\tau_v^{(1)} = 4\mathrm{Pe}\sigma\frac{\tau_v^{(1)}}{t_{ref}}, \tag{4.5}$$

where $t_{ref} = \sigma^2/D_\infty$ is the reference time. By analogy, aggregation outcome can be monitored by the ratio of microcale length scale to mesoscale length scale as

$$\frac{\sqrt{D_\infty \tau_v^{(1)}}}{R_g^{max}}, \tag{4.6}$$

where $R_g^{max}$ represents the maximum size of aggregates which is a mesoscale parameter, and $\sqrt{D_\infty \tau_v^{(1)}}$ represents the microscale length scale and corresponds to the distance a single particle will travel during time $\tau_v^{(1)}$.

From these time and length scale ratios, the scale–separated regime can be determined as

$$G\tau_v^{(1)} \ll 1; \;\; \frac{\sqrt{D_\infty \tau_v^{(1)}}}{R_g^{max}} \ll 1, \tag{4.7}$$

and scale–overlap regime is defined as

$$G\tau_v^{(1)} \geq 1; \;\; \frac{\sqrt{D_\infty \tau_v^{(1)}}}{R_g^{max}} \geq 1. \tag{4.8}$$

Based on this analysis an aggregating map can be represented as shown in Figure 4.1. From this figure, different values for the time ratio $G\tau_v^{(1)}$ should correspond to scale-separated or scale-overlap regimes. A proposed aggregation map is constructed by analogy to the aggregation map reported for non–sheared aggregating systems [16] with an additional axis for the Pèclet number to represent shear. For the non–sheared case it can be concluded that: a) aggregation in a system may occur only if the potential well–depth $\hat{\varepsilon}$ will be greater than some critical value of the well–depth $\hat{\varepsilon}_{cr}$; b) the diffusion coefficient does not control an aggregation outcome, but only controls the speed of an aggregation process.

Based on these conclusions, the final aggregation map can be constructed by dropping the diffusion axis for simplicity and systems with $\hat{\varepsilon} > \hat{\varepsilon}_{cr}$ are chosen for the simulations. From a practical design consideration and a fundamental scientific perspective, it is of interest to relate aggregation outcomes to the physical parameters of the problem. However, it is very

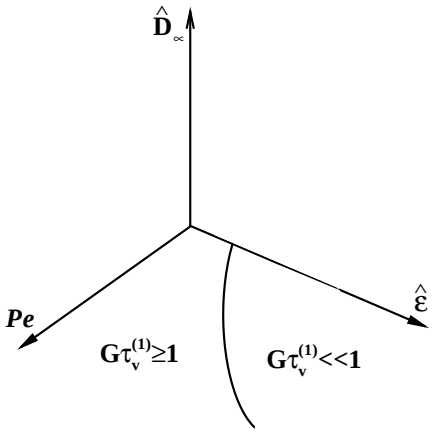



Figure 4.1: Space of dimensionless parameters in which we scale to characterize aggregation outcomes for $G\tau_v^{(1)}$ as a function of the dimensionless potential well depth $\hat{\varepsilon}$, dimensionless diffusion coefficient $\hat{D}_\infty$, and Péclet number Pe.

difficult to simulate realistic physical systems because of their range in length and time scales. Furthemore, a typical dilute colloidal system with volume fraction $f_v = 0.4\%$ in a domain volume of 1cm$^3$ contains up to $10^{15}$ monomer particles . Therefore, for simulation purposes a model systems are considered as reported in Table 4.2. Initially all of the systems are simulated with no shear for dimensionless time $\hat{t} = tD_\infty/\sigma^2 = 3244$ until good statistics of aggregate

Table 4.2: Parameters used in LD simulations to produce Figure 4.2.

| $\hat{\varepsilon}$ | Pe | $N$ |
|---|---|---|
| 8.0 | 2.1 | 300,000 |
| 8.0 | 8.0 | 300,000 |
| 50.0 | 2.1 | 300,000 |
| 50.0 | 8.0 | 300,000 |

structure is obtained. Then the uniform shear flow is applied for $\hat{t} = tD_\infty/\sigma^2 = 113$. Such a simulation strategy allows us to distinguish changes in the aggregate structure that may occur when the uniform shear flow is applied.

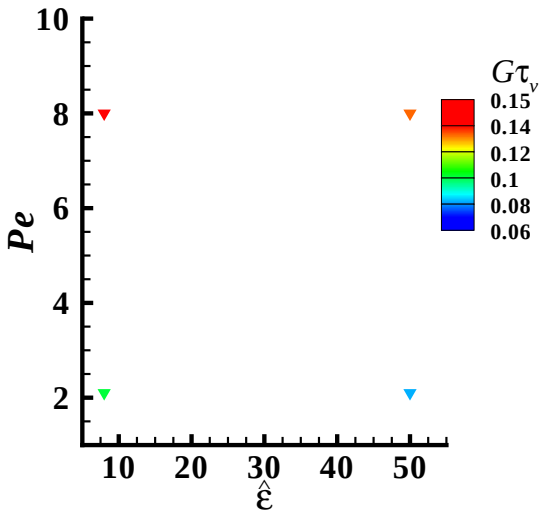



Figure 4.2: $G\tau_v^{(1)}$ as a function of the dimensionless potential well depth $\hat{\varepsilon}$ and Péclet number Pe for LD simulations.

For all LD simulations of systems described in Table 4.2, a velocity autocorrelation time scale for a single particle $\tau_v^{(1)}$ is computed and the regime map of $G\tau_v^{(1)}$ as a function of the dimensionless potential well depth $\hat{\varepsilon}$ and Péclet number Pe is constructed as shown on Figure 4.2.

From the regime map represented on Figure 4.2 it is possible to conclude that for all simulated systems reported in Table 4.2, the scale–separated regime is observed and it is not possible to attain scale–overlap regime. Moreover, application of shear flow does not change $G\tau_v^{(1)}$ parameter significantly since increase in the shear rate $G$ causes a decrease in the velocity autocorrelation time $\tau_v^{(1)}$, and this phenomena requires more detailed explanation.

The possible explanation of this phenomena is based on the fact that as intensity of the shear flow (Pe) increases, the maximum size of aggregates $R_g^{max}$ decreases [3, 6, 7, 8, 9]. This decrease in the size of aggregates occur because particles are leaving the aggregate surface as shear is applied. As result, those particles that leave aggregate lose their "memory" faster than those left in aggregates, and overall, the average velocity autocorrelation time calculated for a single particle $\tau_v^{(1)}$ decreases.

The proposed regime map (Figure 4.2) represents similar values of time ratios for different aggregating systems. Two conclusions may be drawn from this fact: different aggregating systems chosen in Table 4.2 give a similar aggregation outcome, or the additional dimensionless characteristics should be introduced to represent and classify aggregation outcome. In order to determine which conclusion is correct, additional metrics $\mathscr{M}$, that characterize aggregation outcomes should be introduced for the systems described in Table 4.2.

## 4.4 Energy Balance in Sheared Aggregating System

Before analyzing aggregation outcome due to effect of shear, it is important to analyze the redistribution of energy in the system when shear flow is applied. By applying shear into the aggregating system, the kinetic energy of the system increase due to an increase in the kinetic energy in mean velocity $\langle E_{mean} \rangle = m/2\langle \mathbf{u} \cdot \mathbf{u} \rangle$, and due to an increase in the kinetic energy in fluctuating particle velocity $\langle E_{fluct} \rangle = m/2\langle \mathbf{v}' \cdot \mathbf{v}' \rangle$, where $\mathbf{u}$ is the mean velocity and $\mathbf{v}'$ is the fluctuating particle velocity. Evolution of these kinetic energies is shown on Figure 4.3 for a system with $\hat{\varepsilon} = 8$ and Pe $= 2.1$, where $\hat{t} = 0$ corresponds to the time when shear flow is applied to the aggregating system. As can be seen from this figure, the kinetic energy in the mean velocity rises to some steady value in a very short time $\hat{t} \approx 2/\gamma = 0.52$ in $\sigma^2/D_\infty$ units. This time–scale is associated with the time needed for particles to attain the velocity of flow. Then, after reaching a steady value the kinetic energy in mean velocity remains constant. The kinetic energy in fluctuating particle velociti increases by picking up energy from the kinetic energy in mean velocity, however, after particles attain the velocity of flow the kinetic energy in fluctuating particle velocity dropped down to some steady value. Due to such significant changes in kinetic energy of fluctuating particle velocity we focus on the kinetic energy in fluctuating particle velocity $\langle E_{fluct} \rangle = m/2\langle \mathbf{v}' \cdot \mathbf{v}' \rangle$. To identify this energy redistribution, the evolution of average fluctuation particle velocity correlation function is considered

$$\frac{\partial \langle v'_\alpha v'_\beta \rangle}{\partial t} = \lim_{\Delta t \to 0} \frac{\Delta \langle v'_\alpha v'_\beta \rangle}{\Delta t} = \lim_{\Delta t \to 0} \frac{\langle v'_\alpha(t+\Delta t) v'_\beta(t+\Delta t) - v'_\alpha(t) v'_\beta(t) \rangle}{\Delta t} \tag{4.9}$$
$$= \lim_{\Delta t \to 0} \frac{\langle v'_\alpha(t) \Delta v'_\beta(t) \rangle + \langle v'_\beta(t) \Delta v'_\alpha(t) \rangle + \langle \Delta v'_\alpha(t) \Delta v'_\beta(t) \rangle}{\Delta t}, \quad \alpha, \beta = 1, 2, 3,$$

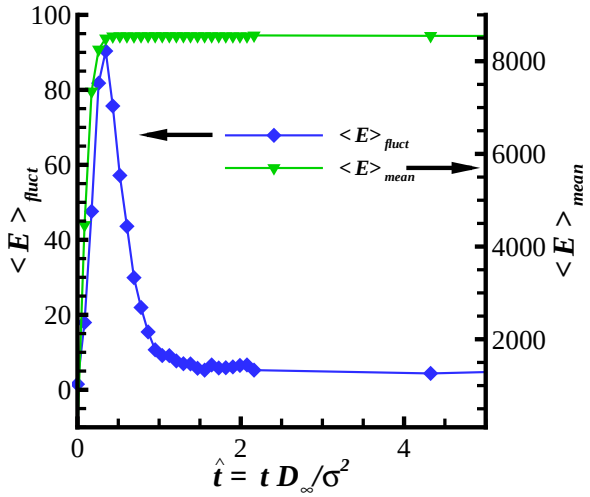



Figure 4.3: Evolution of kinetic energy in mean velocity $\langle E_{mean} \rangle$ and kinetic energy in fluctuating particle velocity $\langle E_{fluct} \rangle$ in $k_B T_{ref}$ units for system with $\hat{\varepsilon} = 8$ and Pe = 2.1.

where $v'_\alpha = v_\alpha - \langle u_\alpha^{(f)} \rangle \delta_{1\alpha}$ is the fluctuation velocity, $v_\alpha$ is the instantaneous velocity, and $\langle u_\alpha^{(f)} \rangle$ is the mean fluid velocity, $\alpha$ and $\beta$ are Cartesian components. By substituting the expression for the velocity increment from Eq. 4.2 the following equation is obtained

$$\frac{\partial \langle v'_\alpha v'_\beta \rangle}{\partial t} = -2\gamma \langle v'_\alpha(t) v'_\beta(t) \rangle + \gamma G \left[ \langle v'_\alpha(t) r'_k(t) \rangle \delta_{1\beta} \delta_{2k} + \langle v'_\beta(t) r'_l(t) \rangle \delta_{1\alpha} \delta_{2l} \right] \quad (4.10)$$
$$+ \frac{1}{m} \left[ \langle v'_\alpha(t) F'_\beta(t) \rangle + \langle v'_\beta(t) F'_\alpha(t) \rangle \right] + 2\gamma \sigma^2_{v_\infty} \delta_{\alpha\beta},$$

or in symbolic terms

$$\frac{\partial \langle v'_\alpha v'_\beta \rangle}{\partial t} = -2\gamma C_{\alpha\beta} + P_{\alpha\beta} + S_{\alpha\beta} + R_{\alpha\beta}. \quad (4.11)$$

In this equation the first term in RHS is the dissipation term with

$$C_{\alpha\beta} = \langle v'_\alpha(t) v'_\beta(t) \rangle, \quad (4.12)$$

second term is production term due to imposed shear with

$$P_{\alpha\beta} = \gamma G \left[ \langle v'_\alpha(t) r'_k(t) \rangle \delta_{1\beta} \delta_{2k} + \langle v'_\beta(t) r'_l(t) \rangle \delta_{1\alpha} \delta_{2l} \right], \quad (4.13)$$

third term is the $vF$ term with

$$S_{\alpha\beta} = \frac{1}{m} \left[ \langle v'_\alpha(t) F'_\beta(t) \rangle + \langle v'_\beta(t) F'_\alpha(t) \rangle \right], \quad (4.14)$$

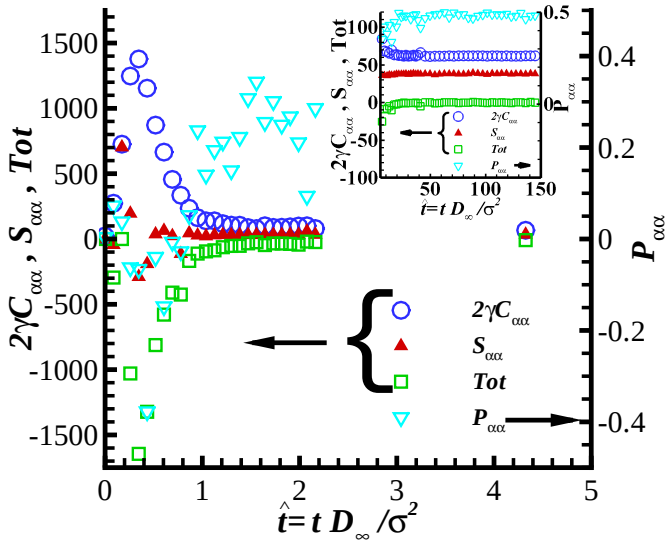



Figure 4.4: Evolution of the trace of each component in Eq. 4.11 in $\sigma^3_{v_\infty}/\sigma$ units for system with $\hat{\varepsilon} = 8$ and Pe $= 2.1$. Inset represents the same values at longer time when system reaches a steady–state.

and the last one is the random source term with

$$R_{\alpha\beta} = 2\gamma\sigma^2_{v_\infty}\delta_{\alpha\beta}. \tag{4.15}$$

Evolution of the trace of each of these components is calculated for a system with $\hat{\varepsilon} = 8$ and Pe $= 2.1$ as shown in Figure 4.4. The evolution of random term $R_{\alpha\alpha}$ is not shown since it remains unchanged. When shear is applied to the aggregating system then particles attain the velocity of flow in a very short time $\sim 2/\gamma = 0.52$ in $\sigma^2/D_\infty$ units (Figure 4.3) that causes a rapid and significant increase in dissipation term (fluctuating energy) observed in Figure 4.4. In fact, fluctuating energy is pumping up the energy from the kinetic energy of the mean velocity as shown on Figure 4.3. This redistribution of kinetic energy of the mean velocity into the fluctuating energy is the driving force for the restructuring of aggregates since kinetic energy of fluctuating particles in aggregates becomes so high that it overcomes interparticle interaction energy. Then with time, particle energy dissipates to some steady value as shown on the insight of Figure 4.3. The balance of all the components in Eq. 4.11 is calculated at $\hat{t} = 140$ in $\sigma^2/D_\infty$ units and reported in Table 4.3 for the two representative systems where the last column

Table 4.3: Energy budget for non–sheared and representative sheared aggregation systems.

| Pe | $\hat{\varepsilon}$ | $-2\gamma C_{\alpha\alpha}$ | $P_{\alpha\alpha}$ | $S_{\alpha\alpha}$ | $R_{\alpha\alpha}$ | Sum,% |
|---|---|---|---|---|---|---|
| 0 | 8 | -22.843 | 0 | 0.063 | 22.876 | 0.43 |
| 0 | 50 | -30.475 | 0 | 0.292 | 30.501 | 1.04 |
| 2.1 | 8 | -61.873 | 0.481 | 38.277 | 22.876 | 0.39 |
| 2.1 | 50 | -155.64 | 0.640 | 125.664 | 30.501 | 0.75 |

represents the left hand side of Equation 4.11, reported in percents of dissipation term $2\gamma C_{\alpha\alpha}$ and it is expected to be zero. This balance analysis satisfies the energy balance of canonical systems and validates the computational code for the numerical simulation of aggregates in the presence of shear flow.

Several important conclusions can be drawn from this energy analysis for canonical systems at steady state. First, when aggregation is simulated with no shear (Pe = 0), production term $P_{\alpha\alpha}$ has zero value, and $S_{\alpha\alpha}$ term is very small (represents velocity–force correlation) as can be seen from Table 4.3. However, when shear is applied, production term slightly increased, when $S_{\alpha\alpha}$ term increased significantly that caused the rise of dissipation term $C_{\alpha\alpha}$. Also, from Table 4.3 it can be observed that at the same shear rate (or the same Pe) the $S_{\alpha\alpha}$ term increases as potential well–depth $\hat{\varepsilon}$ increases. Since $S_{\alpha\alpha}$ can be considered as a work done on the system per unit time and unit mass, it is observed that for stronger interparticle interaction between pairs of particles, the work done on the aggregating system due to shear flow is larger when comparing with systems with weaker interparticle interactions even if the same shear flow with the same Pèclet number is applied.

From the energy analysis for the fluctuating particle velocities, it is possible to conclude that for non-sheared aggregating systems, dissipation and random terms are the major terms for the energy exchange process. When shear flow is applied, the production term $P_{\alpha\alpha}$ is non zero but remains very small when the work done on the aggregating system increases significantly ($S_{\alpha\alpha}$ term), and is larger for systems with larger potential well–depth $\hat{\varepsilon}$. By tracking the time evolution of each component in the energy balance equation (Figure 4.4), it

is observed that as shear flow is applied to the aggregating system the dissipation term $2\gamma C_{\alpha\alpha}$ rises rapidly and significantly and then dissipates to some steady value (Figure 4.4). This effect promotes restructuring in aggregate structure since kinetic energy in fluctuating particle velocity becomes much stronger than interparticle force even when it is observed for a short initial time. Thus, analysis of energy balance allows us to conclude that the application of shear flow into aggregating systems leads to the restructuring of already formed aggregates.

Restructuring of aggregates predicted by energy analysis should be appropriately captured and analyzed. To capture those changes the results for sheared aggregating systems need to be compared with the reference aggregating systems with no shear. Then, some new metrics may be introduced to characterize and analyze the restructuring effect. In the next section a comparison of non-sheared aggregating systems and sheared aggregating systems is performed and the need for new metrics are discussed.

## 4.5 Effect of Shear on Aggregation Structure

In this section we analyze the aggregating results obtained with LD model for systems described in Table 4.2 with and without shear. All LD simulations performed according to the following simulation strategy: first, aggregation simulation with no shear is performed. Then, the LD simulation with applied uniform shear flow continues for the same systems and the aggregates structure changes due to the shear flow are quantified through the analysis of characteristics parameters. The particle-particle interactions for these aggregating systems are described by an effective potential introduced in Eq. 4.3.

### 4.5.1 Aggregation Without Shear

As a first step two aggregating systems described in Table 4.2 with volume fraction $f_v = 0.0169$ are simulated with no shear. The dimensionless potential well–depths used in present study (Table 4.2) satisfy the aggregation regime as reported in previous work [16]. This choice of potential well–depths significantly reduce the effect of solvation potential $\tilde{U}_2$ in an effective potential in Eq. 4.3. All the LD simulations are evolved to a dimensionless time

$\hat{t} = tD_\infty/\sigma^2 = 3244$, where $\sigma^2/D_\infty$ is the diffusion time scale, since the continuation of simulation has no significant influence onto the aggregation statistics since at this time systems consist of mostly large significantly separated aggregates. Thus, further aggregation growth is possible only due to cluster–cluster aggregation that requires a significantly long simulation time.

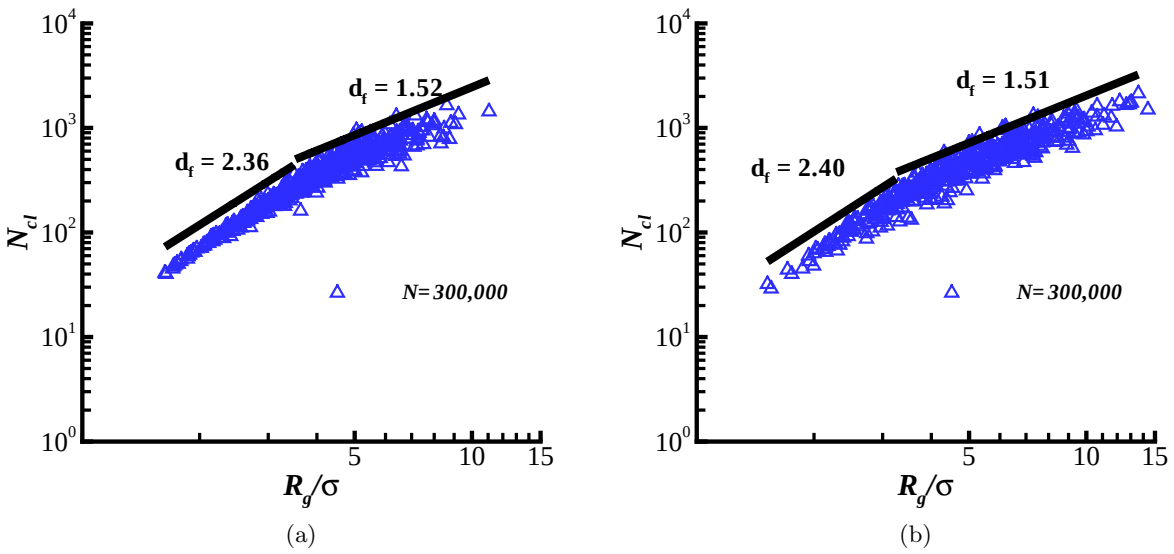
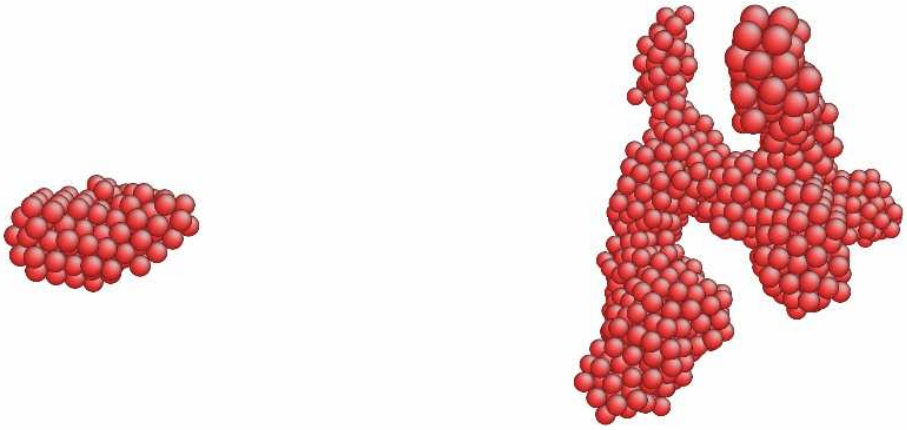



Figure 4.5: The fractal dimension $D_f$ from the LD with effective potential $U_{eff}^{LD}$ at time $\hat{t} = 3244$: a) simulations are done with $\hat{\varepsilon} = 8.0$; b) simulations are done with $\hat{\varepsilon} = 50.0$.

The first step in the aggregate analysis is to determine aggregates in systems. Clusters are determined based on the minimum neighbor distance criteria. With this method two particles are assigned to the same cluster if the distance between their centers is less than some distance $r_{cl}$. This distance is typically chosen between $1.1\sigma < r_{cl} < 1.5\sigma$. In the present study $r_{cl} = 1.4\sigma$ which is chosen based on the radial distribution function $g(r)$ where $r = 1.4\sigma$ corresponds to the first minimum.

The fractal dimension parameter $D_f$ is the first most commonly used metric to describe the structure of aggregates. The maximum value of fractal dimension depends on the dimensionality of space $D$ as $D_f^{max} = D$. Thus, for 3D solid material the fractal dimension is $D_f = 3$.



(a) (b)

Figure 4.6: Snapshots for two typical aggregates for $\hat{\varepsilon} = 50.0$ at time $\hat{t} = 3244$: a) aggregate containing 150 monomers with the radius of gyration $R_g = 2.8\ \sigma$; b) aggregate containing 966 monomers with the radius of gyration $R_g = 7.1\ \sigma$.

Typically, when fractal structure is formed in $D$-space, its fractal dimension is lesser than the space dimension. Thus, in $3D$ system a flat paper sheet has fractal dimension 2, stretched linear polymer chain will have fractal dimension 1.

The fractal dimension $D_f$ is related to the number of monomers in a cluster $k$ and to the radius of gyration of the cluster $R_g$ by the fundamental relation,

$$k = k_0 \left( \frac{R_g}{a} \right)^{D_f}, \tag{4.16}$$

where $a = \sigma/2$ is the particle radius, $\sigma$ is the particle diameter, and $k_0$ is the constant. For isotropic system the radius of gyration for each cluster can be calculated as

$$R_g = \left( \sum_{i=1}^{k} \left( \mathbf{r}^{cm} - \mathbf{r}^{(i)} \right)^2 \right)^{1/2}, \tag{4.17}$$

where $\mathbf{r}^{cm}$ is the coordinate of the center of mass of cluster, and $\mathbf{r}^{(i)}$ is the coordinate of $i$-th particle in the cluster.

According to Eq. 4.16 fractal dimension for aggregates can be determined by plotting the number of monomers in each cluster $k$ versus the cluster radius of gyration $R_g$ by using

logarithmic scale. By extracting a slope from this logarithmic plot a fractal dimension value is obtained as shown on Figure 4.5. From this figure two regions with larger and smaller fractal dimensions separated by inflection point $R_g^{in} \approx 4\,\sigma$ are observed for all systems with different potential well–depths (Figure 4.5). All the aggregates with $R_g < R_g^{in}$ are characterized with high fractal dimension $D_f \approx 2.4$, when for the larger aggregates with $R_g > R_g^{in}$ the fractal dimension is much smaller ($D_f \approx 1.5$). Such a difference in fractal dimension for different sizes of aggregates outline the difference in their structures. This difference in structures at different length scales is attributed to the fact that the Lennard-Jones potential used in present study is a long–range potential. Therefore, with very deep well–depth (large potential well–depth $\hat{\varepsilon}$) "*fat fractals*" are formed that are characterized by compact local structure and ramified structure at largerl ength scale. In this case a fractal dimension $d_f$ can be reliably extracted only for very large aggregates containing more than $100,000$ monomer particles. To satisfy this requirement very large systems with large number of particles should be simulated for a very long computational time that may be not achieved. If short–range potential would be used then much thinner aggregates would form on the local length scale and cluster structure would be similar at different length scales.

To observe the structural difference two typical structures of aggregates are plotted for two regions with different fractal dimensions $D_f$ for $\hat{\varepsilon} = 50.0$ as shown on Figure 4.6. For smaller aggregates with $R_g < R_g^{in}$ (Figure 4.6a) clusters with a very compact monomer packing are observed. This suggests that the driving mechanism for the cluster formation is the monomer addition. On the other hand, larger aggregates with $R_g > R_g^{in}$ (Figure 4.6b) are characterized by ramified structure which is formed due to aggregation of smaller clusters, thus the cluster–cluster aggregation mechanism is suggested. Therefore, an inflection point $R_g^{in}$ serves as the critical size of aggregate when aggregation mechanism due to the particle addition changes to cluster–cluster aggregation.

### 4.5.2 Shear–induced Aggregation Mode

After characterizing structures of aggregates formed without shear flow, the aggregation simulation is continued with application of moderate shear flow (Pe = 2.1) to the systems. Such a shear–induced aggregations are evolved to a dimensionless time $\hat{t} = 113$ in $\sigma^2/D_\infty$ units. For these simulations a fractal dimension parameter $D_f$ is extracted in a similar manner as it is described above as shown on Figure 4.7. However, shear flow introduces spatial anisotropy into aggregates, therefore Eq. 4.17 can not be used for the radius of gyration calculation any more. Instead, inertia ellipsoid with axes length $d_\alpha = a, b$, and $c$ for $\alpha = 1, 2$ and 3 for 3D system is generated for each cluster according to the following procedure. The shape of any aggregate containing $k$ solute particles can be described by its moment of inertia tensor $\mathbf{I}$ with components

$$I_{\alpha\beta} = \frac{1}{k}\sum_{n=1}^{k}\left(r_\alpha^{(n)} - r_\alpha^{cm}\right)\left(r_\beta^{(n)} - r_\beta^{cm}\right), \qquad \alpha, \beta = 1, 2, 3, \tag{4.18}$$

where $r_\alpha^{(n)}$ is the $\alpha$-th component of position of $n$-th solute particle in the cluster, and $r_\alpha^{cm}$ is the $\alpha$-th component of the cluster center of mass. Then the singular–value decomposition of inertia tensor $\mathbf{I}$ is performed as

$$\mathbf{I} = \mathbf{USV}^T, \tag{4.19}$$

where $\mathbf{U}$ is the unitary matrix, $\mathbf{V}^T$ is the conjugate transpose of the unitary matrix $\mathbf{V}$, and $\mathbf{S}$ is the diagonal matrix of principal axes of inertia ellipsoid with components

$$S_{\alpha\alpha} = d_\alpha^2, \qquad \alpha = 1, 2, 3. \tag{4.20}$$

Then the squares of principal radii of gyration $R_\alpha^2$, for $\alpha = 1, 2, 3$, are calculated as

$$R_\alpha^2 = \frac{1}{3}\left(d_\beta^2 + d_\gamma^2\right), \qquad \alpha \neq \beta \neq \gamma, \tag{4.21}$$

and the radius of gyration is calculated as [21]

$$R_g = \sqrt{\frac{1}{2}\left(R_1^2 + R_2^2 + R_3^2\right)}. \tag{4.22}$$

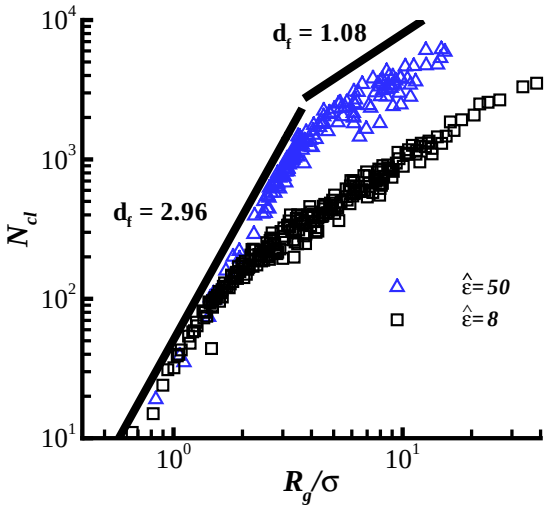



Figure 4.7: The fractal dimension $D_f$ from the LD with effective potential $U_{eff}^{LD}$ under shear flow with Pe = 2.1 at time $\hat{t} = 113$.

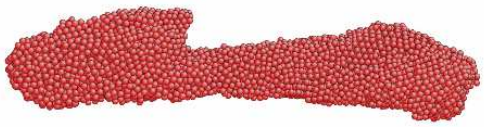


Figure 4.8: Snapshots for typical aggregate for $\hat{\varepsilon} = 50.0$ at time $\hat{t} = 113$ for aggregate containing 7144 monomers with the radius of gyration $R_g = 15\ \sigma$.

Therefore, Eq. 4.22 should be used when calculating the radius of gyration $R_g$ for aggregates in the sheared aggregating systems.

The fractal dimension parameter $D_f$ has two distinct values with inflection point $R_g^{in}$ (Figure 4.7) similarly to previously observed behavior (Figure 4.5). However, application of the shear flow induces significant changes into aggregates structure (Figure 4.8) when comparing with non–shear case (Figure 4.6b). Application of the shear flow increases the fractal dimension for the local length scale from $D_f \approx 2.4$ to $D_f \approx 3.0$ which corresponds to the formation of more compact structure in smaller aggregates. On the other hand, the fractal dimension for the larger length scale decreases from $D_f \approx 1.5$ (Figure 4.5) to $D_f \approx 1.1$ (Figure 4.7) that can be explained by formation of cigar–like shapes of aggregates instead of globular–like shapes that increase the radius of gyration $R_g$ (Figure 4.8). Thus, application of moderate shear flow into aggregation system causes the compacting effect at the local length scale and spatial rearrangement of aggregates by forming cigar–like shapes at the global length scale. Moreover, in systems with uniform shear flow, the inflection point value $R_g^{in}$ is found to be different for different interparticle potential well–depths $\hat{\varepsilon}$ (Figure 4.7), that results in different local

length scales of aggregates for different $\hat{\varepsilon}$. This effect may be explained by the fact that the maximum size of aggregates in the direction perpendicular to the shear flow is controlled by the strength of shear flow (controlled by Pèclet number) and by the strength of the interparticle force (controlled by the dimensionless potential well–depth $\hat{\varepsilon}$). Therefore, the relation between these two forces will determine the maximum size of aggregates in the direction perpendicular to the shear flow $R_g^{max}$.

It is possible to conclude that as shear flow is applied to the aggregating system its structure on local and global length scale significantly changes. Those changes are observed through changes in such metric as fractal dimension $D_f$ and visually when ramified structure of aggregates changes to cigar–like structure on a global length scale. However, even fractal dimension $D_f$ allows to determine aggregate structure changes due to shear it can not determine the changes in local structure of a single aggregate since $D_f$ is extracted only from the set of different sizes of aggregates. At this point the global structure of aggregates is characterized with fractal dimension $D_f$. Now, we would like to introduce metric that characterize the local structure of individual aggregates.

### 4.5.3 Characterization of Local Structure

Some studies predict that the shear flow may compact the local structure of the aggregates and so called "compactness effect" can be observed [12]. On the other hand if shear flow is strong the aggregates with less dense local structure are formed. To investigate the isotropy of aggregates and their interparticle local distribution the dimensionless local volumetric potential energy density (LPED) $\hat{U}/\hat{V}_{cl}$ is calculated for each aggregate as

$$\frac{\hat{U}}{\hat{V}_{cl}} = \frac{U}{V_{cl}}\frac{\sigma^3}{\varepsilon} = \left| \frac{1}{kV_{cl}} \sum_{i=1}^{k} \sum_{j>i}^{k} \frac{\sigma^3}{\varepsilon} U^{LJ}(r_{ij}) \right|, \tag{4.23}$$

where $\sigma$ is the solute particle diameter, $\hat{\varepsilon}$ is the potential well–depth, $k$ is the number of monomer particles in cluster (does not include particles that lay on the surface of cluster) to exclude size effect, $V_{cl} = kV_m$ is the volume of all the bulk particles in a cluster, $V_m$ is the volume of a single particle, $U^{LJ}(r_{ij})$ is the Lennard–Jones interaction potential between

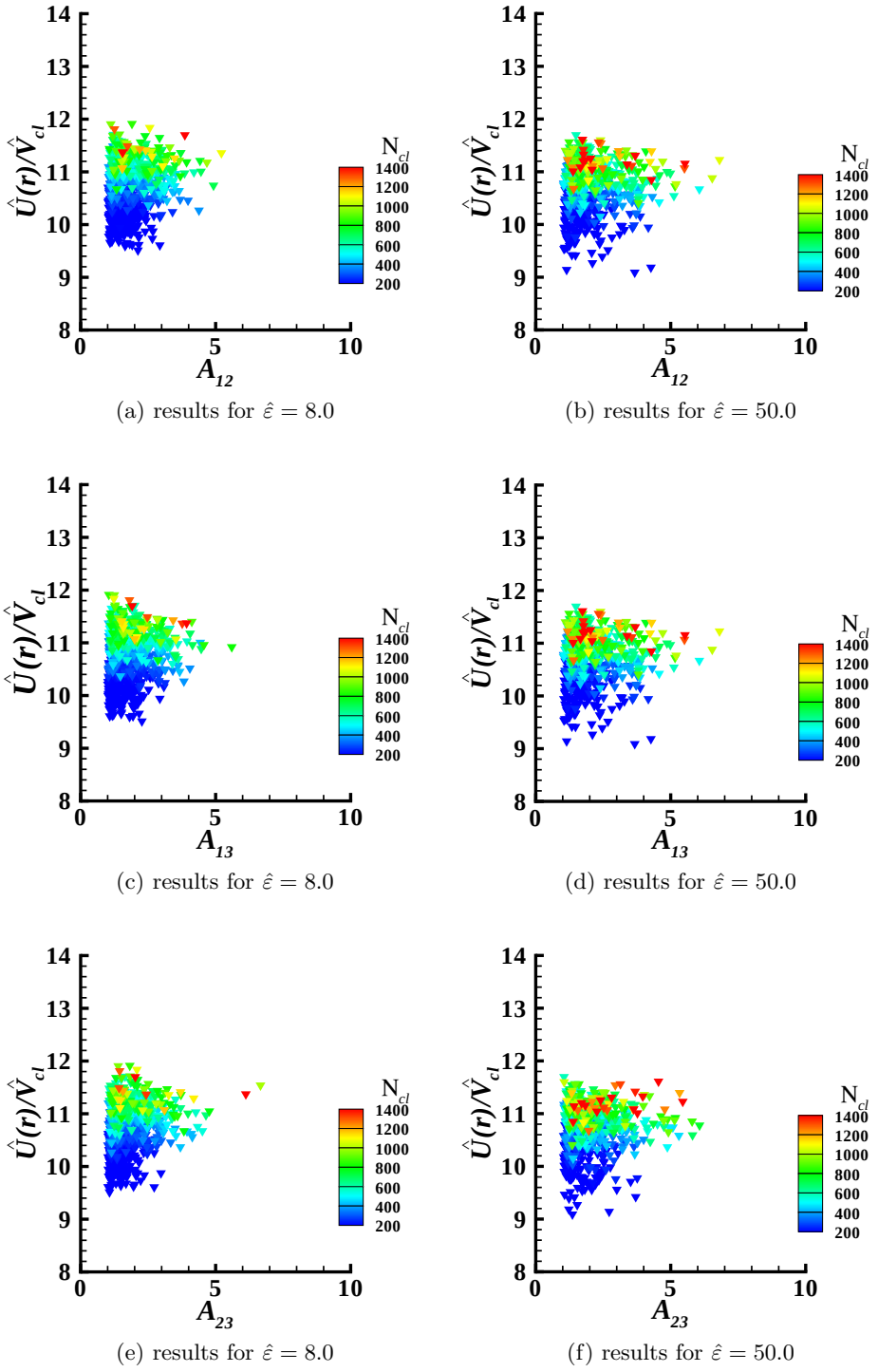
particles $i$ and $j$ defined in Eq. 4.3, and $r_{ij}$ is the separation distance between solute particles $i$ and $j$. The dimensionless LPED determines the potential energy in the system per unit volume and it describes compactness of the aggregate structure. Compactness of aggregate structures can be characterized by the average number of nearest neighbors in the aggregate. Thus, for highly packed structure, when the number of nearest neighbors is higher the LPED is higher than for a porous structure. However, the LPED is very sensitive to the presence of structural anisotropy in aggregate structure (since anisotropy can limit the number of neighbors in some direction). Therefore, the dimensionless LPED parameter $\hat{U}/\hat{V}_{cl}$ should be represented together with anisotropy parameter $A_{\alpha\beta}$.

The anisotropy parameter is calculated after determining the equivalent inertia ellipsoid with principle axes for each cluster according with Eq. 4.18-Eq. 4.20, and is defined as

$$A_{\alpha\beta} = \frac{d_\alpha}{d_\beta}, \tag{4.24}$$

where the principle axes are rearranged as $d_1 > d_2 > d_3$. With such specification the anisotropy factor value $A_{\alpha\beta}$ is always greater than one while $\alpha < \beta$. This specification of an anisotropy is different to those proposed in literature [21] where the mean shape anisotropy for the ensemble of cluster $\langle A_{\alpha\beta} \rangle$ is calculated. However, the mean shape anisotropy parameter does not fit the needs of present work where anisotropy for every cluster in the system should be defined.

For isotropic aggregating systems the LPED as a function of the anisotropy factor $A_{\alpha\beta}$ is calculated first as shown on Figure 4.9. On these plots each point corresponds to a single cluster and only clusters with $k > 100$ are shown. For both isotropic aggregating system with $\hat{\varepsilon} = 8$ and $\hat{\varepsilon} = 50$, a similar pattern is observed for all anisotropy factors $A_{\alpha\beta}$. This result supports that observed systems do not have preferential direction and validates statistically isotropic assumption for non-sheared aggregating systems. However, even for isotropic systems the maximum magnitude of anisotropy factors is as large as 5. Also, we can observe that the smaller clusters with $100 < k < 400$ are characterized by lower local volumetric potential energy density (Figure 4.9), which is caused by poor statistics for small aggregates when calculating the potential energy due to LJ interaction potential, since the number of particles at aggregate surface is relatively large to the total number of particles in aggregate. Thus, by excluding the




(a) results for $\hat{\varepsilon} = 8.0$

(b) results for $\hat{\varepsilon} = 50.0$

(c) results for $\hat{\varepsilon} = 8.0$

(d) results for $\hat{\varepsilon} = 50.0$

(e) results for $\hat{\varepsilon} = 8.0$

(f) results for $\hat{\varepsilon} = 50.0$

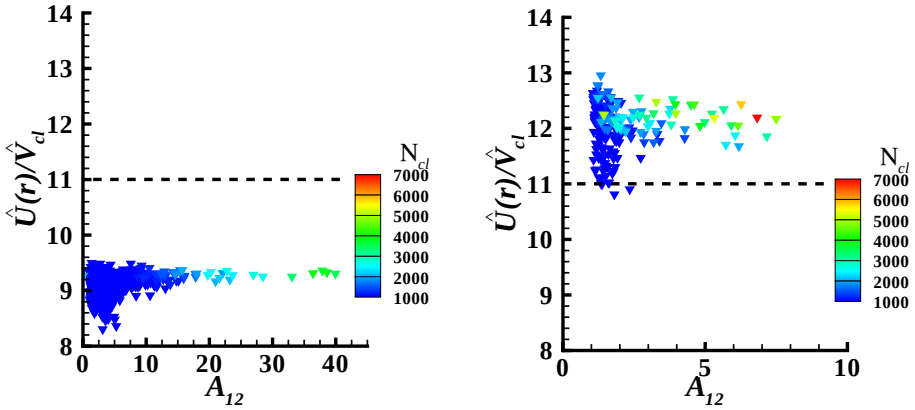
Figure 4.9: For aggregation without shear the dimensionless local volumetric potential energy density $\hat{U}(r)/\hat{V}_{cl}$ in $\varepsilon/\sigma^3$ units as a function of the anisotropy $A_{ij}$, where $i, j = 1, \ldots, 3$ for $\hat{\varepsilon} = 8.0$ and $\hat{\varepsilon} = 50.0$. Color legend represents the number of monomers in each cluster.

region with smaller aggregates the variation in the local volumetric potential energy density is within 10%.

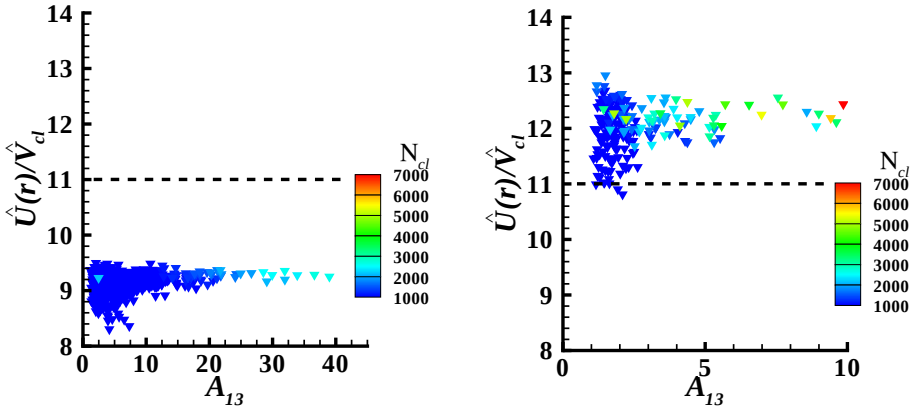
Based on these results we conclude that the formation of clusters with branch–like structure through the coalescence of smaller aggregates explains the decrease of the fractal dimension $D_f$. Moreover, the local structure (at small length scale $< R_g^{in}$) of the large aggregates is very different to their global structure (at large length scale) as seen on Figure 4.6b. From anisotropy analysis the anisotropy factors $A_{\alpha\beta}$ are found to be identical with no preferential direction that supports isotropy assumption. For non-sheared systems the average value for the local volumetric potential energy density $\langle \hat{U}/\hat{V}_{cl} \rangle = 10.8$ is similar for systems with the interaction potential well–depth $\hat{\varepsilon} = 8$ and $\hat{\varepsilon} = 50$.

By applying shear flow with Pe = 2.1 to aggregating systems with different potential well–depths $\hat{\varepsilon}$ the figure of dimensionless LPED as a function of the anisotropy factors $A_{\alpha\beta}$ is generated as shown on Figure 4.10. On this figure only clusters with $k > 100$. By comparing these data for systems with a uniform shear flow with non-shear case (Figure 4.9) several important features can be distinguished. A significant asymmetry in anisotropy factors $A_{12}$ and $A_{13}$ is observed for all systems with different potential well–depths (Figure 4.10a,b,c,d) which is explained by application of the shear flow along $x$ direction (which corresponds to $\alpha = 1$ in $A_{\alpha\beta}$). At the same time the magnitude for the anisotropy factor $A_{23}$ is in the same range as for non-sheared systems (Figure 4.9e,f and Figure 4.10e,f). Application of the moderate shear flow to the aggregate system significantly changes the magnitude of anisotropy factors along shear flow direction ($A_{12}$ and $A_{13}$) from 5 for non-sheared systems to up to 40 for $\hat{\varepsilon} = 8$ and up to 10 for $\hat{\varepsilon} = 50$. It is possible to conclude that shear flow introduces anisotropy into the aggregating system by changing a globular aggregate structure to cigar–like structure.

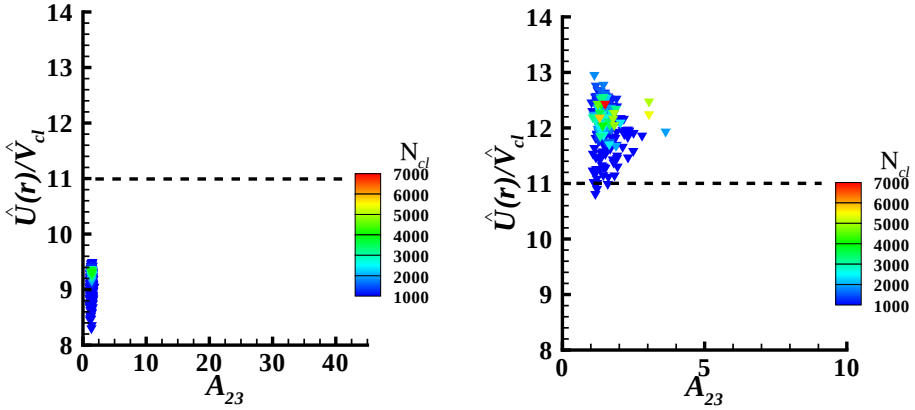
Moreover, this "stretching" of aggregates due to shear flow changes their local structure which is perfectly detected by the dimensionless LPED. For systems with weaker interparticle interaction ($\hat{\varepsilon}$=8) the shear force with Pe = 2.1 is strong enough to significantly separate particle pairs inside of aggregate and as result the average number of nearest neighbors decreases (aggregates become less compact). As result the dimensionless LPED decreases to 9.2 from




(a) results for $\hat{\varepsilon} = 8.0$ (b) results for $\hat{\varepsilon} = 50.0$




(c) results for $\hat{\varepsilon} = 8.0$ (d) results for $\hat{\varepsilon} = 50.0$




(e) results for $\hat{\varepsilon} = 8.0$ (f) results for $\hat{\varepsilon} = 50.0$

Figure 4.10: Local volumetric potential energy density $\hat{U}(r)/\hat{V}_{cl}$ in $\varepsilon/\sigma^3$ units as a function of the anisotropy $A_{ij}$, where $i, j = 1, \ldots, 3$ for sheared abbregating systems with Pe = 2.1. The color legend represents the number of monomers in each cluster. Dashed line represents average LPED value for non-sheared case.

10.8 for non-sheared systems as can be seen on Figure 4.10a,c,e. When interparticle interaction is stronger ($\hat{\varepsilon} = 50$), the same shear flow with Pe = 2.1 is not able to stretch pair of particles inside of aggregate to the same extent as in the case of $\hat{\varepsilon} = 8$. However, some displacements of particles due to shear force occur that initiates their local rearrangement. This local rearrangement allows particles to occupy more energy stable locations than they could have due to aggregation process with no shear. As result of these local rearrangements more compact structures are formed that is supported by increasing value of the average dimensional LPED to 12.1.

From these results it is possible to conclude that the presenting of shear flow in aggregation systems may change local structure of aggregates in several ways. When the force due to shear flow between pair of particles is significant by comparing with the force due to interparticle potential interaction the aggregates with cigar–like structures are formed. These cigars are very long along the shear flow direction with less compact local structure (when comparing with non-sheared aggregates). In the cases where force due to shear flow is relatively small in comparison with force due to interparticle potential interaction we still observe aggregates with cigar–like structures. However, in this case "cigars" are shorter. Moreover, shear flow reduces stresses in aggregates due to local rearrangement of particles and as result more compact local structures of aggregates are formed. With this analysis we can extract the length to width ratio for aggregates, however, to complete characterization of the aggregating structure the method for prediction of the maximum size of aggregates $R_g^{max}$ in the direction perpendicular to the shear flow needs to be introduced.

## 4.6 Prediction of Maximum Size of Aggregates Under Shear

Competition between interparticle attraction and shear is presented in sheared aggregation systems. On one side present attraction forces between particles due to interparticle interaction (represented by $\hat{\varepsilon}$ in the present study) are driving forces for the aggregation process. On the other side shear flow introduces the forces between particles and its intensity corresponds to Péclet number Pe. The relative magnitude of these forces will determine the maximum size of

aggregates in the direction perpendicular to the shear flow.

In this section the algorithm to determine maximum size of aggregates $R_g^{max}$ is proposed based on the idea of a relative acceleration between pair of particles calculations reported in previous work [13]. First, let us consider an aggregate and choose a pair of particles $i$ and $j$ in this aggregate with condition that $i$-th particle is located at the center of mass of the aggregate. These particles are separated by distance $\mathbf{r}^{(i),(j)} = \mathbf{r}^{(i)} - \mathbf{r}^{(j)}$ and have relative velocity $\mathbf{w} = \mathbf{v}^{(i)} - \mathbf{v}^{(j)}$. The relative acceleration between these particles $\Delta\mathbf{A}^{(i),(j)}$ then can be calculated as a sum of the relative acceleration between particles due to interparticle potential interaction $\Delta\mathbf{A}_{pot}^{(i),(j)}$ and the relative acceleration between particles due to shear flow $\Delta\mathbf{A}_{sh}^{(i),(j)}$ as

$$(\Delta\mathbf{A}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}, \mathbf{W} = \mathbf{w}) = (\Delta\mathbf{A}_{pot}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}) + (\Delta\mathbf{A}_{sh}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}, \mathbf{W} = \mathbf{w}) \quad (4.25)$$

To improve statistics the calculations for all the $j^{th}$ particles in the cluster and for all the clusters is performed and the average relative acceleration is obtained

$$\langle\Delta\mathbf{A}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}, \mathbf{W} = \mathbf{w}\rangle = \langle\Delta\mathbf{A}_{pot}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}\rangle + \langle\Delta\mathbf{A}_{sh}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}, \mathbf{W} = \mathbf{w}\rangle \quad (4.26)$$

Negative value for the mean relative acceleration corresponds to the particles acceleration toward each other when positive value corresponds to the particle acceleration away from each other.

The relative acceleration between a pair of particles due to potential interaction is calculated by computing the force on each particle due to all other particles in the aggregate and by taking the difference of these forces

$$\langle\Delta\mathbf{A}_{pot}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}\rangle = \frac{1}{m^{(i),(j)}}\langle\Delta\mathbf{F}_{pot}^{(i),(j)}|\mathbf{R} = \mathbf{r}^{(i),(j)}\rangle = \quad (4.27)$$
$$\frac{1}{M}\sum_{n}^{M}\left[\sum_{i\neq j,j=1}^{k}\frac{1}{m^{(i),(j)}}\left[-\nabla_{\mathbf{r}^{(i),(j)}}U^{LD}(|\mathbf{r}^{(i),(j)}|)\right] - \sum_{j\neq i,i=1}^{k}\frac{1}{m^{(i),(j)}}\left[-\nabla_{\mathbf{r}^{(i),(j)}}U^{LD}(|\mathbf{r}^{(i),(j)}|)\right]\right],$$

where $m^{(i),(j)} = (m^{(i)} + m^{(j)})/2$ is the mean mass of particles $i$ and $j$, $k$ is the number of monomer particles in the cluster, $M$ is the number of clusters in the system at time $t$, and $U^{LD}$ is the interparticle potential calculated from Eq. 4.3.

The relative acceleration between a pair of particles due to shear and relative velocity between those particles is calculated based on the following observations. A pair of particles is characterized by the relative fluctuation velocity $w'_\alpha$ that is calculated as

$$w'_\alpha = w_\alpha - \Delta\langle u_\alpha\rangle^{(i),(j)}\delta_{1\alpha}, \quad \alpha = 1,2,3 \tag{4.28}$$

where $w_\alpha$ is the relative instantaneous velocity, and $\Delta\langle u_\alpha\rangle^{(i),(j)}\delta_{1\alpha}$ is the relative mean velocity due to shear flow. When shear flow is applied to the colloidal system the particles in the system attain a velocity of shear flow in a very short time ( $\hat{t} = O(1/\hat{\gamma})$, where $\hat{\gamma}$ is the dimensionless friction coefficient defined in Eq. 4.2). Therefore, a final relative acceleration is calculated as the difference of the relative acceleration due to instantaneous velocities and the relative acceleration due to fluctuation velocity as

$$\langle \Delta A^{(i),(j)}_{sh\ \alpha} | R_\alpha = r^{(i),(j)}_\alpha, W_\alpha = w_\alpha\rangle = \frac{1}{M}\sum_n^M \gamma\left[GR_\beta\delta_{2\beta} - w'_\alpha\right]\delta_{1\alpha}, \tag{4.29}$$

where $GR_\beta\delta_{2\beta}$ is the relative instantaneous velocity, and $M$ is the number of clusters in the system at time $t$. By substituting calculated results from Eq. 4.27 and Eq. 4.29 into Eq. 4.26 the distribution of the average relative acceleration between pair of particles separated by $\mathbf{r}^{(i),(j)}$ is created as shown on Figure 4.11. On these plots, the average relative acceleration is calculated along the line connecting a center of particles pair, such that $\langle \Delta\mathbf{A}_r\rangle = \langle \Delta A\rangle \mathbf{r}/|\mathbf{r}|$. From these figures the maximum size of aggregates in the direction perpendicular to the direction of the shear flow can be found as the separation after which the average relative acceleration is always greater than zero that can be represented mathematically as

$$R_g^{max}\{\mathbf{r} : \Delta\mathbf{A}_r > 0 \ \forall\ r > r^*\}. \tag{4.30}$$

This definition is valid even for the cases when some oscillations of the average relative acceleration are observes (Figure 4.11b) since they are related to the local aggregate structure. This method has been applied to different aggregating systems with shear flow, and results of analysis are represented in Table 4.4. In this table $R^{max}_{g\ num}$ is the maximum size of aggregates in $\sigma$ units calculated from simulation snapshot at time $t$, $R^{max}_{g\ an}$ is the maximum size of aggregates in $\sigma$ units calculated by using proposed algorithm. We note there is a relatively small difference

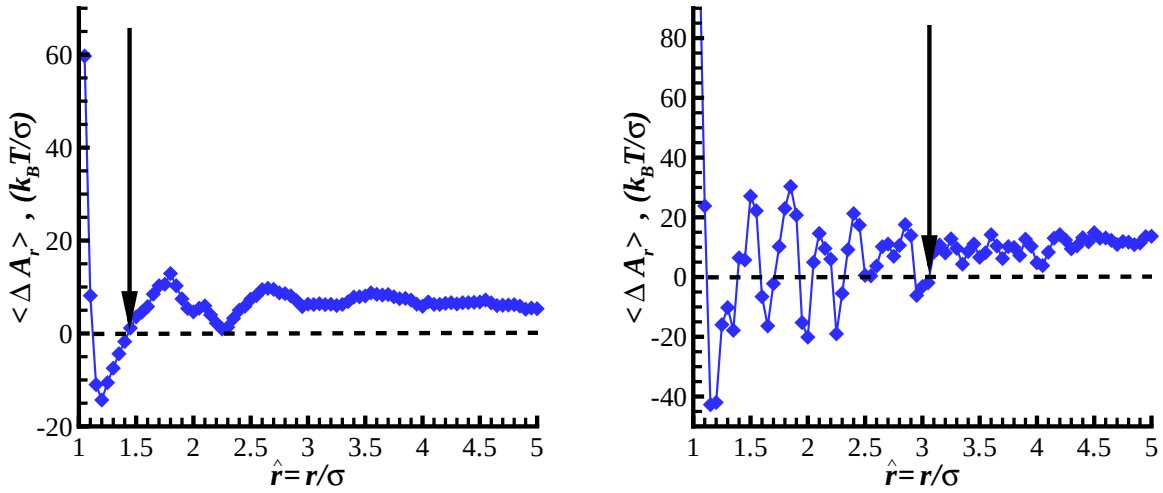



Figure 4.11: Sum of the relative accelerations due to potential interaction and the relative acceleration due to shear at time $\hat{t} = 113$: a) system with $\hat{\varepsilon} = 8.0$; b) system with $\hat{\varepsilon} = 50.0$. Arrays show prediction of the maximum size of aggregates $R_g^{max}$.

between calculated and predicted data (less than 10% for most of the systems). This allows us conclude that proposed algorithm can be used for a reasonable prediction of aggregates sizes in the direction perpendicular to the shear flow.

Table 4.4: Maximum size of aggregates calculation.

| Pe | $\hat{\varepsilon}$ | $R_{g\ num}^{max}$ | $R_{g\ an}^{max}$ | % difference |
|---|---|---|---|---|
| 0.5 | 8 | 2.88 | 3.09 | 6.8 |
| 1.0 | 8 | 1.94 | 2.33 | 16.7 |
| 2.1 | 8 | 1.46 | 1.43 | 2.1 |
| 4.0 | 8 | 1.37 | 1.44 | 4.9 |
| 2.1 | 50 | 2.87 | 3.06 | 6.2 |

These results support the reasonable scientific prediction that when intensity of shear flow increases (while interparticle interaction potential remains the same) the size of aggregates decreases, and for a critical shear flow aggregation may not occur at all. By increasing the

potential well depth (while shear flow intensity remain unchanged) aggregates of larger sizes are formed.

With the proposed method the influence of the shear flow onto aggregate structure can be characterized quantitatively by using the relative acceleration approach and based on the competition of the forces due to potential interaction and shear flow.

The proposed method for analytical prediction of the maximum size of aggregates $R_g^{max}$ in shear flow is based on the consideration of only two forces between pair of particles: attraction force due to interparticle interaction, and repulsion force due to shear. The fact that consideration of only these two forces gives good prediction for $R_g^{max}$ suggests that the concurrency of these two forces is controlling the aggregation outcome. However, this point will be discussed with more details in the next section.

After analyzing LPED and anisotropy factor the sheared aggregating systems are completely described through the set of dimensionless parameters and through the introduced metrics such as fractal dimension $D_f$, radius of gyration $R_g$, local volumetric potential energy density LPED, and anisotropy parameter $A_{\alpha\alpha}$. With all of these characteristics in hand it would be beneficial to construct an aggregation map with clear identification of aggregation outcomes based on input parameters such as Pèclet number and dimensionless potential well–depth $\hat{\varepsilon}$. In the next section development of such an regime map for aggregation is discussed.

## 4.7 A Regime Map for Aggregation Under Sheared

As described in Section 4.6 by introducing shear flow into aggregating system we are not only introducing the new time scale $1/G$, but also an additional shear force that occurs between a pair of particles. When shear flow is applied to the aggregating system particles quickly attain flow velocity ($\sim 1/\gamma$). Because of the presence of the velocity gradient in a shear flow a pair of particles at different locations along the velocity gradient will attain different velocities. This difference in velocity for a pair of particles is a source of an additional force due to shear flow that tends to tear apart particles. Now, aggregation outcome does not solely depend on the magnitude of pair–interaction force that arises from interaction potential, but instead it

depends on relation of magnitudes of the pair–interaction force and the shear force. Based on reasoning proposed above it is useful to introduce such characteristics as the relative force $f_{i,sh}$ that is determined as

$$f_{pot,sh} = \frac{|F_{pot}(r_{min})|^{max}}{F_{sh}(r_{min})}, \tag{4.31}$$

where $|F_{pot}(r_{min})|^{max}$ is the maximum absolute interparticle force between pair of particles separated by $r_{min}$, calculated as

$$|F_{pot}(r_{min})|^{max} = \left| 24\frac{\varepsilon}{\sigma}\left[2\left(\frac{\sigma}{r_{min}}\right)^{13} - \left(\frac{\sigma}{r_{min}}\right)^{7}\right]\right|, \tag{4.32}$$

where $r_{min}$ is the pair separation that satisfies the maximum attraction force between pair of particles. $F_{sh}(r_{min})$ is the force due to shear between a pair of particles separated by $r_{min}$, calculated as

$$F_{sh}(r_{min}) = m\gamma G r_{min} = m\gamma 4 Pe D_{\infty}\frac{r_{min}}{\sigma^2}. \tag{4.33}$$

By substituting Eq. 4.32 and Eq. 4.33 into Eq. 4.31 and representing values in dimensionless form we will get

$$f_{pot,sh} = \frac{6\hat{\varepsilon}}{Pe}\left|\left[2\left(\frac{\sigma}{r_{min}}\right)^{14} - \left(\frac{\sigma}{r_{min}}\right)^{8}\right]\right|, \tag{4.34}$$

where dimensionless units are computed as $\hat{\varepsilon} = \varepsilon/k_B T_{ref}$, $\hat{\gamma} = \gamma\sigma/\sigma_{v_\infty}$, $\hat{D}_\infty = D_\infty/\sigma\sigma_{v_\infty}$, $\sigma^2_{v_\infty} = k_B T/m$, and by taking into account the fact that $\hat{\gamma}\hat{D}_\infty = 1$.

By analyzing Eq. 4.31 and Eq. 4.34 it is possible to conclude that when $f_{pot,sh} < 1$ the shear force is stronger than an iterparticle interaction force and in such a system the breakage of aggregates to monomers is expected. When $f_{pot,sh} > 1$ formation of aggregates in a system is expected (this can be achieved by increasing dimensionless potential well depth $\hat{\varepsilon}$ or decreasing shear flow intensity Pe as can be seen from Eq. 4.34). It should be noted that the relative force is defined based on particle–particle interaction force which may not be an appropriate one when cluster–cluster aggregation occur, however this should not disturb the aggregation map but only may change the magnitude of the relative force $f_{pot,sh}$.

By performing improved LD numerical simulations of the aggregation in the presence of shear flow of different Pèclet numbers Pe and with different potential well–depth $\hat{\varepsilon}$ such metric

as the relative force $f_{pot,sh}$ is calculated and an aggregation map is constructed as shown on Figure 4.12.

Table 4.5: ($d_f$, $R_g^{max}$, LPED) Values as a function of Pèclet number Pe and potential well–depth $\hat{\varepsilon}$.

| $\hat{\varepsilon}$ | Pe | | | | | |
|---|---|---|---|---|---|---|
| | 2.1 | | | 8.0 | | |
| 8 | 1.08 | 1.5 | 9.2 | 0 | 0.5 | NA |
| 50 | 1.08 | 2.9 | 12.1 | 1.13 | 2.0 | 10.0 |

This map represents the relative force $f_{pot,sh}$ as a function of the dimensionless interparticle potential well–depth $\hat{\varepsilon}$, and Pèclet number Pe. On this map a color legend represents different values for the relative force $f_{pot,sh}$. Based on the value of the relative force $f_{pot,sh}$ three different regions are identified: (a) no aggregation; (b) aggregaion with less dense local structure (when compare with non–sheared system); (c) aggregation with compact local structure. The dashed line represents the border between regions where aggregates do not form $f_{pot,sh} < 1$ and the region where formation of aggregates is observed $f_{pot,sh} > 1$. Dotted line separates regions when aggregates with less dense and compact local structure are formed. To validate the choice for these regions such metrics as the fractal dimension $d_f$, the maximum size of aggregates in the direction perpendicular to the shear flow $R_g^{max}$, and the local volumetric potential energy density LPED are calculated for systems at different regions and results are represented in Table 4.5. Thus, for aggregating non dense region two systems with significantly different initial conditions with ($\hat{\varepsilon}$;Pe)=(8;2.1) and (50;8.0) are selected. Since these two systems are in the same region on aggregating map their metrics should be of similar values even their initial conditions are significantly different which is perfectly satisfied as shown in Table 4.5 and on Figure 4.12. An as it was predicted the LPED value for both systems is lesser than for non–sheared aggregating system (LPED=10.8). System which initial conditions predict formation of more compact aggregates with ($\hat{\varepsilon}$;Pe)=(50;2.1) is characterized with LPED=12.1, and system where no aggregation is predicted is fully supported by values of proposed metrics. Thus, a proposed aggregation map (represented on Figure 4.12) is able to predict different

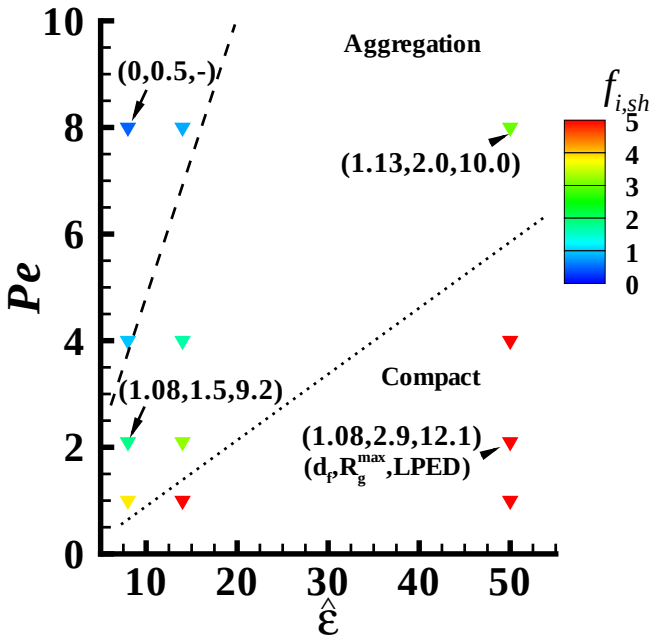



Figure 4.12: Relative force$f_{pot,sh}$ as a function of the dimensionless interparticle potential well–depth $\hat{\varepsilon}$, and Pèclet number Pe. The dashed line represents the boundary between non–aggregating and aggregating systems, and dotted line identifies region when a compactness of the local structure is observed. Values in brackets for selected systems represent $D_f$, $R_g^{max}$, and LPED correspondently.

outcomes for different aggregating regimes for sheared aggregating system that are confirmed with the metrics calculated for some of the systems (Table 4.5).

Based on analysis of metrics for sheared aggregating systems it is possible to conclude that such characteristic as a relative force $f_{pot,sh}$ is an adequate characteristics that is very useful for predicting an aggregation outcome.

## 4.8 Discussion

In this section we would like to summarize all of the processes that are taking place when shear flow is applied to the aggregating systems. At the beginning, when shear flow is applied to the aggregating system all the particles in aggregates attain the flow velocity in a very short time that is order of dissipation time scale $\sim 1/\gamma$. During this period the magnitude of the fluctuating particle velocity increases significantly because the kinetic energy of mean velocity transfers to the kinetic energy of fluctuating particle velocity. Then fluctuating kinetic energy dissipates to some smaller steady value. This rapid rise in fluctuating kinetic energy causes aggregates restructure that causes different structure outcomes depending on the system's parameters. We have determined that aggregates structure outcome depends not only on the shear intensity Pe but also on the magnitude of the interparticle force represented by the potential well–depth $\hat{\varepsilon}$. Therefore the sheared aggregation outcome depends on the ratio of the interparticle force to the shear force, represented by $f_{pot,sh}$. When shear intensity is dominant ($f_{pot,sh} < 1$) then energy that is transferred from kinetic energy of mean velocity into fluctuating kinetic energy overcomes the interparticle interaction energy and aggregates break down to the monomer–size aggregates, thus aggregate breakage is observed. When shear intensity is such that shear force is of the same order of magnitude as the interparticle force and $f_{pot,sh} > 1$ then aggregates which size is larger than the maximum allowable size of aggregate $R_g^{max}$ will break to smaller ones that will satisfy $R_g < R_g^{max}$. At the same time the local structure of these aggregates is getting less compact since the shear is not able to break aggregates to the monomer–size it still large enough to increase the average neighbor distance $R_{nn}$. And finally, when shear intensity is weak relative to potential interaction ($f_{pot,sh} \gg 1$) then all the

aggregates with sizes $R_g > R_g^{max}$ will break to smaller clusters to satisfy $R_g < R_g^{max}$ and the local structure of these aggregates is getting more compact. More compact structure is formed because shear reduces internal stress within aggregates and particles in aggregates are able to occupy more energetically preferable positions.

Our results determine that application of shear flow also changes the global structure of aggregates by formation of cigar–like aggregates for $f_{pot,sh} < 1$. In this case the reduction in the fractal dimension $d_f$ (which characterize the global structure of aggregate) is observed. Our findings are consistent with experimental results for the moderate shear flow [11]. Although, these authors are also reporting that the fractal dimension of aggregates increases when high intensity shear flow is applied. However, this phenomena is observed for systems where very large aggregates are able to form with sizes $R_g > 50\sigma$. We could not observe this phenomena with our model systems because we could not achieve this size of aggregates in a reasonable computation time and suitable number of particles for 3D system. On the other hand 2D simulations of the sheared aggregating system may be useful that will allow to reach the size when compactness effect for the global structure may be observed.

## 4.9 Conclusions

We have performed numerical simulation of sheared colloidal particles aggregation in model systems by using Langevin dynamics model with improved interparticle interaction potential. For these systems the set of dimensionless parameters that is able to distinguish scale–separated and scale–overlap regimes was determined. However, this set of dimensionless parameters is not the best for representation of aggregation outcomes and other metrics are used for this purpose. To understand the aggregates restructuring process due to shear flow we propose the method of the energy evolution analysis as such that allows to capture redistribution of the flow energy into the fluctuating energy which is the source of the aggregates restructuring. With this energy analysis we are able to gain a fundamental understanding of restructure and/or breakage processes that caused by imposed shear flow. The effect of shear flow onto the local and global structure of aggregates is studied with the local volumetric potential energy density

LPED and the maximum radius of gyration $R_g^{max}$ correspondently. It is observed that shear flow dramatically change the structure of aggregates on both local and global length scales. On the local length scale shear flow cases the formation of more or less compact structures depending on the shear flow intensity characterized by Pèclet number Pe. On the global length scale the size of aggregates in the direction perpendicular to the shear flow is limited by $R_g^{max}$ and its value depends on the Pe as well. We proposed a new method for $R_g^{max}$ prediction that yields results consistent with those obtained from the direct size distribution calculations. With full analysis of the sheared aggregating systems we propose the aggregating map based on new metric $f_{pot,sh}$ which is the ratio of interparticle force to the shear force. This map allows to determine different aggregating outcomes based on the initial parameters of the sheared aggregating systems such as interparticle force and shear flow rate. This map can be used when planning new aggregating experiments or when comparing outcomes from several different aggregating systems.

## 4.10 Acknowledgments

This work has been partially supported by the National Science Foundation through Grant No. CTS-0403864.

## Bibliography

[1] Gonzalez, A. E. Colloidal aggregation with sedimentation: computer simulations. *Phys. Rev. Lett*, 86(7):1243–1246, 2001.

[2] Manley, S., Skotheim, J. M., Mahadevan, L., and Weitz, D. A. Gravitational collapse of colloidal gels. *Phys. Rev. Lett*, 94:218302, 2005.

[3] Selomulya, C., Bushell, G., Amal, R., and Waite, T. D. Aggregation mechanisms of latex of different particle sizes in a controlled shear environment. *Langmuir*, 18(6):1974–1984, 2002.

[4] Harada, S., Tanaka, R., Nogami, H., Sawada, M., and Asakura, K. Structural change in non-fractal particle cluster under fluid stress. *Colloids & Surfaces A*, 302:396–402, 2007.

[5] Wang, L., Marchisio, D. L., Vigil, R. D., and Fox, R. O. CFD simulation of aggregation and breakage process in laminar Taylor-Couette flow. *J. Colloid & Interf. Sci.*, 282:380–396, 2005.

[6] Soos, M., Moussa, A. S., Ehrl, L., Sefcik, J., Wu, H., and Morbidelli M. Effect of shear rate on aggregate size and morphology investigated under turbulent conditions in stirred tank. *J. Colloid & Interf. Sci.*, 319:577–589, 2008.

[7] Kikuchi, Y., Yamada, H., Kinimori, H., Tsukada, T., Hozawa, M., Yokoyama, C., and Kubo, M. Aggregation behavior of latex particles in shear flow confined between two parallel plates. *Langmuir*, 21(8):3273–3278, 2005.

[8] Serra, T., and Casamitjana, X. Structure of the aggregates during the process of aggregation and breakup under a shear flow. *J. Colloid & Interf. Sci.*, 206:505–511, 1998.

[9] Serra, T., Colomer, J., and Casamitjana, X. Aggregation and breakup of particles in a shear flow. *J. Colloid & Interf. Sci.*, 187:466–473, 1997.

[10] Satoh, A., Chantrell, R. W., and Coverdale, G. N. Brownian dynamics simulations of ferromagnetic colloidal dispersions in a simple shear flow. *J. Colloid & Interf. Sci.*, 209:44–59, 1999.

[11] Mokhtari, T., Chakrabarti, A., Sorensen, C. M., Cheng, C., and Vigil, D. The effect of shear on colloidal aggregation and gelation studied using small-angle light scattering. *J. Coll. Interf. Science*, 327:216–223, 2008.

[12] Cerdà, J. J., Sintes, T., Holm, C., Sorensen, C. M., and Charkabarti, A., Shear effect on crystal nucleation in colloidal suspensions. *Phys. Rev. E*, 78:031403, 2008.

[13] Markutsya, S., Subramaniam, S. Coarse–graining approach to infer mesoscale potentials from atomistic interactions for aggregating systems. *Submitted to Phys. Rev. E*

[14] Anderson, V. J., and Lekkerkerker, H. N. W. Insights into phase transition kinetics from colloid science. *Nature*, 416:811–815, 2002.

[15] Cerdà, J. J., Sintes, T., Sorensen, C. M., and Charkabarti, A., Kinetics of phase transformations in depletion-driven colloids. *Phys. Rev. E*, 70:011405, 2004.

[16] Markutsya, S., Subramaniam, S., Vigil, R. D., and Fox, R. O. On Brownian dynamics simulation of nanoparticle aggregation. *Ind. Eng. Chem. Res.*, 47:3338–3345, 2008.

[17] Ermak, D. L. Computer-simulation of charged-particles in solution .1. Technique and equilibrium properties. *J. Chem. Phys.*, 62:4189–4196, 1975.

[18] Turq, P., Lantelme, F., and friedman H. L. Brownian dynamics - its application to ionic solutions. *J. Chem. Phys.*, 66:3039, 1977.

[19] Plimpton, S. J. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.*, 117:1–19, 1995.

[20] Stoyan, D., and Stoyan, H. *Fractals, random shapes, and point fields : methods of geometrical statistics*; John Wiley&Sons, 1994.

[21] Fry, D., Mohammad, A., Chakrabarti, A., and Sorensen, C. M. Cluster shape anisotropy in irreversibly aggregating particulate systems. *Langmuir*, 20:7871–7879, 2004.

# CHAPTER 5. CONCLUSIONS AND FUTURE WORK

## 5.1 Summary and conclusions

A mesoscale simulation method using Langevin dynamics (LD) method with an improved potential of mean force that is capable of accurately capture aggregation simulation from diffusion limited to reaction limited regimes is successfully used to accurately predict aggregation outcomes for colloidal systems with and without shear. Based on these simulations, maps of aggregating regime in sheared and non–sheared systems have been developed for Lennard–Jones (LJ) dilute systems. The principal findings of this study are:

1. A novel coarse–graining approach of deriving an effective potential for LD simulations is proposed. This effective potential takes into account both direct interaction between solute particles (Lennard-Jones potential) and also interaction between solute particles initiated by presence of solvent molecules. In aggregating systems if unmodified LJ potential is used as the potential of mean force, then the aggregation structure predicted by LD does not match with that observed in MD for equisize systems in the reaction limited regime. Applying of this effective potential into LD model significantly improve aggregate structure. As result a good matching is obtained for MD and improved LD simulations from diffusion limited to reaction limited regimes.

2. Application of the principle of statistical/stochastic equivalence is proposed to match aggregation statistics obtained from MD and LD simulations. In this method the evolution of the second-order density for MD model is derived. The average relative acceleration between nanoparticle pairs is identified as an important link between MD and mesoscale models such as LD in both DLA and RLA regimes.

3. Aggregation outcomes are represented on regime maps for sheared and non–sheared systems. These maps identify the relevant dimensionless parameters that determine aggregation for dilute system. For non–sheared case these are potential well–depth $\hat{\varepsilon}$ that controls aggregation outcome, and dimensionless diffusion $\hat{D}_{\infty}$ that only controls the rate at which this outcome is reached. For sheared case these are potential well–depth $\hat{\varepsilon}$ and dimensionless shear flow $Pe$. The aggregation outcome is characterized with the extent of aggregation $\xi$ for non–sheared systems and by the ratio of interparticle force to the force due to shear $f_{i,sh}$ for sheared case.

4. The extent of aggregation $\xi$, cluster size distribution (CSD), radial distribution function $g(r)$, the average relative acceleration between pair of particles, the local volumetric potential energy density (LPED), and the maximum size of aggregates in the presence of shear flow $R_g^{max}$ are important statistics that are useful in analyzing and classifying the structure of aggregates for sheared and non–sheared aggregating systems.

## 5.2 Secondary findings

1. Molecular dynamics approach is not feasible to simulate an aggregation of colloidal nanoparticles for systems of physical sizes using existing computational algorithms and resources due to presence of very large number of solvent particles in such systems.

2. The condition that allows for simplification of LD position and velocity–Langevin equations to BD position-Langevin namely, that relaxation timescales be much smaller than configuration relaxation time scales is met by the physical system considered in the present work when shear is not applied, but is not justified for the model system. However, when shear flow is applied this reduction to position–Langevin may be not allowed even for physical systems, since shear flow introduces an additional time scale into the aggregating system.

3. To satisfy accuracy requirements when simulating aggregation processes, a significant number of multiple independent simulations (MIS) is needed. The reason for this is that

the aggregation process is evolving in time and it is impossible to improve statistics by performing time averaging as can be done for equilibrium systems. Therefore, accurate simulation of the aggregation process is computationally demanding.

4. Light scattering analysis (LSA) is a powerful tool for analyzing the aggregate structure from experiment. It allows us to extract such aggregation characteristics as the fractal dimension $D_f$ for a range of aggregate length scales. The same analysis can be applied to simulation data, thereby enabling a direct comparison of experiments and computer simulation. The application of this tool is limited by requirements of a large dynamic range for precise measurements of $D_f$. This leads to requirements of large cluster sizes and small nanoparticles volume fraction (in order to shift the ideal gel point radius of gyration). Thus, meaningful comparison of LSA applied to computer system requires a significant speed up of LD simulations.

In summary, new method of deriving an effective potential for LD mesoscale simulations is proposed. With this improved potential, both aggregate structure (described by the radial distribution function $g(r)$ and cluster size distribution) and force field (relative acceleration) are accurately captured. In addition, detailed characterization of aggregation outcome is performed for colloidal systems with and without shear.

## 5.3 Future work

Recently, a simulation of protein molecules in solvent is of a great interest for the biological applications. Typical proteins are the linear polymers build from series of amino acids characterized by different physical and chemical properties. Simulation of such complicated structures in solvent bath is a very hard problem. Applicability of microscale methods (such as MD) or its modification such as accelerated MD for solving this problem is very limited, due to very large number of solvent molecules presented in the system. From this perspective mesoscale methods (such as BD/LD) with an effective potential that would account for the effect of solvent molecules could be more appropriate to solve this kind of problems.

Before applying mesoscale methods to solve for protein problems it is important to develop a general approach for aggregating phenomenon. For this we need to answer the following questions: (a) How aggregating outcome depends on the solute–solvent interaction (systems with solvophobic/solvophilic particles)? (b) When introducing solute/solvent size separation how aggregating outcome would change? (c) How aggregating map will change for the different interparticle potentials? With mesoscale BD/LD method we would be able to answer to all these questions: questions (a) and (b) can be answered through deriving correspondent effective potential using methodology described in this dissertation; to get answer for question (c) different interparticle potentials can be easily implemented into the BD/LD approach.

In current work the hydrodynamic interactions (HI) are neglected because particle size is very small. However, when size–separated systems are considered the HI effect might be significant and effect of HI should be included in LD model. Implementation of effective algorithm for HI into mesoscale LD method is one of the important step that needs to be done for accurate representation of large particle aggregating phenomenon in liquid bath.

Some additional thoughts about sheared aggregating process are needed such as when shear flow is introduced to the aggregating system we observe formation of long cigar–like aggregates that are aligned along the shear flow direction. However, other studies observe rotation of aggregates when shear flow is applied. This point must be clarified by increasing simulation box size and increasing observation time. It is possible that tumbling of aggregates can not occur when aggregates length is comparable with the simulation box size. And longer simulation time would allow to observe aggregating process far from initial unsteady point.

The effect of aggregates structure on their diffusion coefficient needs some further investigation. It is possible to predict the diffusion coefficient for spherical aggregates of different sizes, however the diffusion coefficient for the fractal aggregates with ramified structure is unknown. Further study should be done to understand dependence of the diffusion coefficient on the fractal dimension of aggregates $d_f$ and on the aggregate radius of gyration $R_g$.

The numerical accuracy of mesoscale approaches may be improved by implementing more accurate numerical integration of stochastic differential equations (SDE). Currently explicit

Euler method is used, however, implementation of higher stochastic integration schemes such as Milstein scheme, implicit Euler and Milstein schemes, balanced methods etc. would be beneficial.

# APPENDIX A. BUCKINGHAM PI ANALYSIS

Buckingham Pi Theory relies on the identification of variables involved in a process. Several steps shoul be procedeed to perform a nondimensional analysis.

1. List all the variables that govern the process.

   In our case we have: potential well depth $\varepsilon$, nanoparticle diameter $\sigma$, nanoparticle position $\mathbf{r}$, nanoparticle velocity $\mathbf{v}$, nanoparticle mass $m$, nanoparticle diffusion coefficient $D_\infty$, $k_B T_{ref}$ where $k_B$ is Boltzmann constant, reference temperature $T_{ref}$, time $t$, and particle number density $\lambda$. Thus all together there are $n = 9$ variables.

2. Between all variables in the system mark a few of them as "Repeating Variables". This step is most difficult in a dimensional analysis.

   In our case these are: $\sigma$, $m$, $k_B T_{ref}$ and $k = 3$

3. Define how many non-dimensional numbers is in system. In this case it is $n - k = 6$. Our problem has four non-dimensional numbers: $\Pi_1$, $\Pi_2$, $\Pi_3$, $\Pi_4$, $\Pi_5$, and $\Pi_6$.

4. Define the non-dimensional numbers by grouping the variables into $n-k$ groups. So each group has all the repeating variables and one non-repeating variable. For our problem we have:

$$\Pi_1 = \Pi_1(\sigma, m, k_B T_{ref}, \varepsilon) \quad (A.1)$$

$$\Pi_2 = \Pi_2(\sigma, m, k_B T_{ref}, D_\infty) \quad (A.2)$$

$$\Pi_3 = \Pi_3(\sigma, m, k_B T_{ref}, t) \quad (A.3)$$

$$\Pi_4 = \Pi_4(\sigma, m, k_B T_{ref}, \lambda) \quad (A.4)$$

$$\Pi_5 = \Pi_5(\sigma, m, k_B T_{ref}, \mathbf{r}) \quad (A.5)$$

$$\Pi_6 = \Pi_6(\sigma, m, k_B T_{ref}, \mathbf{v}) \quad (A.6)$$

Let

$$\Pi_1 = \sigma^a m^b (k_B T_{ref})^c \varepsilon \quad (A.7)$$

$$\Pi_2 = \sigma^d m^e (k_B T_{ref})^f D_\infty \quad (A.8)$$

$$\Pi_3 = \sigma^g m^h (k_B T_{ref})^i t \quad (A.9)$$

$$\Pi_4 = \sigma^j m^k (k_B T_{ref})^l \lambda \quad (A.10)$$

$$\Pi_5 = \sigma^m m^n (k_B T_{ref})^o \mathbf{r} \quad (A.11)$$

$$\Pi_6 = \sigma^p m^q (k_B T_{ref})^r \mathbf{v} \quad (A.12)$$

5. Express each variable in terms of its dimensions.

| Variable | $\varepsilon$ | $\sigma$ | $\mathbf{r}$ | $\mathbf{v}$ | m | $D_\infty$ | $k_B T_{ref}$ | t | $\lambda$ |
|---|---|---|---|---|---|---|---|---|---|
| Dimension | $ML^2/T^2$ | L | L | L/T | M | $L^2/T$ | $ML^2/T^2$ | T | $1/L^3$ |

Substituting these dimensions into $\Pi_{1-6}$ at the previous step:

$$\Pi_1 = (L)^a(M)^b(ML^2/T^2)^c(ML^2/T^2) \quad \text{(A.13)}$$

$$\Pi_2 = (L)^d(M)^e(ML^2/T^2)^f(L^2/T) \quad \text{(A.14)}$$

$$\Pi_3 = (L)^g(M)^h(ML^2/T^2)^i(T) \quad \text{(A.15)}$$

$$\Pi_4 = (L)^j(M)^k(ML^2/T^2)^l(1/L^3) \quad \text{(A.16)}$$

$$\Pi_5 = (L)^m(M)^n(ML^2/T^2)^o(L) \quad \text{(A.17)}$$

$$\Pi_6 = (L)^p(M)^q(ML^2/T^2)^r(L/T) \quad \text{(A.18)}$$

or

$$a + 2c + 2 = 0;\ b + c + 1 = 0;\ -2c - 2 = 0$$

$$d + 2f + 2 = 0;\ e + f = 0;\ -2f - 1 = 0$$

$$g + 2i = 0;\ h + i = 0;\ -2i + 1 = 0$$

$$j + 2l - 3 = 0;\ k + l = 0;\ -2l = 0$$

$$m + 2o + 1 = 0;\ n + o = 0;\ -2o = 0$$

$$p + 2r + 1 = 0;\ q + r = 0;\ -2r - 1 = 0$$

Solving these equations yields,

$$a = 0;\ b = 0;\ c = -1$$

$$d = -1;\ e = 1/2;\ f = -1/2$$

$$g = -1;\ h = -1/2;\ i = 1/2$$

$$j = 3;\ k = 0;\ l = 0$$

$$m = -1;\ n = 0;\ o = 0$$

$$p = 0;\ q = 0.5;\ r = -0.5$$

Non-dimensional numbers become:

$$\hat{\varepsilon} = \frac{\varepsilon}{k_B T_{ref}} \tag{A.19}$$

$$\hat{D}_\infty = \frac{D_\infty}{\sigma}\sqrt{\frac{m}{k_B T_{ref}}} \tag{A.20}$$

$$\hat{t} = \frac{t}{\sigma}\sqrt{\frac{k_B T_{ref}}{m}} \tag{A.21}$$

$$\hat{\lambda} = \lambda\sigma^3 \tag{A.22}$$

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{\sigma} \tag{A.23}$$

$$\hat{\mathbf{v}} = \mathbf{v}\sqrt{\frac{m}{k_B T_{ref}}} \tag{A.24}$$

By introducing the velocity variance $\sigma_{v_\infty}^2 = k_B T_{ref}/m$ these equations reduce to

$$\hat{\varepsilon} = \frac{\varepsilon}{k_B T_{ref}} \tag{A.25}$$

$$\hat{D}_\infty = \frac{D_\infty}{\sigma\sigma_{v_\infty}} \tag{A.26}$$

$$\hat{t} = t\frac{\sigma}{\sigma_{v_\infty}} \tag{A.27}$$

$$\hat{\lambda} = \lambda\sigma^3 \tag{A.28}$$

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{\sigma} \tag{A.29}$$

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\sigma_{v_\infty}} \tag{A.30}$$

It is clear that there are only three parameters that characterise the system (except position, velocity, and time), such as scaled particle number density $\hat{\lambda}$ (a particle volume fraction $\alpha$ can be used instead), a scaled nanoparticle diffusion coefficient $\hat{D}_\infty$, and a scaled potential well depth $\hat{\varepsilon}$.

# APPENDIX B. RELATIVE ACCELERATION CALCULATION

The average relative acceleration between two particles conditional on their separation $\mathbf{r}$ is decomposed into direct and indirect contributions (cf. Eq. 3.13). Assuming pairwise interactions, an expression for the average indirect relative acceleration is derived in terms of three-particle statistics. An approximation for dilute systems is used to derive a closed–form expression for the indirect relative acceleration conditional on pair separation as a function of two-particle statistics (pair correlation function) and the pair potential. Simple 1-D and 2-D test systems are constructed to compare and verify numerical simulation results with this closed–form analytical expression. The results show an excellent match between the numerical simulation and the analytical expression, thereby verifying our numerical simulation. The tests give insight into the nature of the indirect relative acceleration, and also draw attention to special numerical accuracy requirements for calculating the same.

We first consider a system of identical particles experiencing pairwise additive interactions through an isotropic pair potential $U(r)$ . The direct relative acceleration between particles (1) and (2) (see Figure 3.3), each with mass $m$ and separated by $r = |\mathbf{r}|$, is simply

$$\langle \Delta \mathbf{A}^D | \mathbf{r} \rangle = \frac{2\mathbf{F}_{(2)(1)}}{m} = -\frac{2}{m} \nabla U(r).$$

Here we drop the time dependence in the relative acceleration expressions for simplicity. We seek to derive a similar expression for the indirect relative acceleration between particles (1) and (2) conditional on their separation $r$. Toward this end, we first write out the *unconditional* average indirect relative acceleration between particles (1) and (2) in terms of the *conditional* average indirect relative acceleration as

$$\langle \Delta \mathbf{A}^I \rangle = \frac{1}{n^2} \int \langle \Delta \mathbf{A}^I | \mathbf{r} \rangle \rho^{(2)}(\mathbf{r}) d\mathbf{r} = \int \langle \Delta \mathbf{A}^I | \mathbf{r} \rangle g(\mathbf{r}) d\mathbf{r}. \tag{B.1}$$

The same *unconditional* average indirect relative acceleration can be written in terms of the acceleration induced by a probe particle $p$ located at $\mathbf{x}_p$ on each particle (1) and (2) as

$$\langle \Delta \mathbf{A}^I \rangle = \frac{1}{n^3} \int \mathbf{A}^{I:(2)}(\mathbf{r}'', \mathbf{r}) \rho^{(3)}(\mathbf{r}'', \mathbf{r}) d\mathbf{r}'' d\mathbf{r} - \frac{1}{n^3} \int \mathbf{A}^{I:(1)}(\mathbf{r}', \mathbf{r}) \rho^{(3)}(\mathbf{r}', \mathbf{r}) d\mathbf{r}' d\mathbf{r} \tag{B.2}$$

where $\mathbf{r}'' = \mathbf{x}_2 - \mathbf{x}_p$ and $\mathbf{r}' = \mathbf{x}_1 - \mathbf{x}_p$ as shown in Fig. 3.3. While this expression is exact for pairwise interacting systems, it requires knowledge of the three-particle density term $\rho^{(3)}(\mathbf{r}', \mathbf{r})$, which is unknown in general. Also, in order to extract the *conditional* average indirect relative acceleration from Eq. B.2, we need to define appropriate conditional third–order statistics.

The three-particle density $\rho^{(3)}(\mathbf{r}', \mathbf{r})$ can normalized to a three-particle correlation $h(\mathbf{r}', \mathbf{r})$ as $\rho^{(3)}(\mathbf{r}', \mathbf{r}) = n^3 h(\mathbf{r}', \mathbf{r})$, analogous to the normalization of the two-particle density to obtain the pair correlation $\rho^{(2)}(\mathbf{r}) = n^2 g(\mathbf{r})$. The three-particle correlation function $h(\mathbf{r}', \mathbf{r})$ can be expressed in terms of conditional three-particle statistics $h(\mathbf{r}'|\mathbf{r})$ and the pair correlation function as

$$h(\mathbf{r}', \mathbf{r}) = h(\mathbf{r}'|\mathbf{r}) g(\mathbf{r}). \tag{B.3}$$

Substituting all these expressions into Eq. B.2, the following expression for the average indirect relative acceleration conditional on pair separation results:

$$\langle \Delta \mathbf{A}^I | \mathbf{r} \rangle = \int \mathbf{A}^{I:(2)}(\mathbf{r}'', \mathbf{r}) h(\mathbf{r}''|\mathbf{r}) d\mathbf{r}'' - \int \mathbf{A}^{I:(1)}(\mathbf{r}', \mathbf{r}) h(\mathbf{r}'|\mathbf{r}) d\mathbf{r}'. \tag{B.4}$$

This equation gives the average indirect relative acceleration conditional on pair separation in terms of conditional three–particle statistics.

In Eq. B.4 for the average indirect relative acceleration conditional on pair separation $\mathbf{r}$, the integrals are taken over all possible separation distances $\mathbf{r}'$ between particle (1) and the probe particle $p$, and $\mathbf{r}''$ between particle (2) and probe particle $p$, as shown in Fig. 3.13. The limits of integration for $\mathbf{r}''$ and $\mathbf{r}'$ in Eq. B.4 need to respect the geometrical constraint $\mathbf{r}' + \mathbf{r} = \mathbf{r}''$, and to avoid overlaps of the probe particle with (1) and (2) we must satisfy $|\mathbf{r}''| > (\sigma_2 + \sigma_p)/2$, and $|\mathbf{r}'| > (\sigma_1 + \sigma_p)/2$ (see Fig. 3.13). We see from Eq. B.4 that for pairwise interactions (assumed in MD and BD) the presence of the third particle only affects the limits of integration. In general, the conditional three-particle statistics of systems are unknown.

In order to verify our numerical calculation of relative acceleration, we consider dilute systems where the conditional three particle statistics can be approximated such that the conditional average indirect relative acceleration can be analytically calculated from the pair correlation function and pair potential. Assuming a dilute system where $h(\mathbf{r}'|\mathbf{r})$ and $h(\mathbf{r}'|\mathbf{r})$ are approximated by

$$h(\mathbf{r}'|\mathbf{r}) \cong g(\mathbf{r}'), \tag{B.5}$$

$$h(\mathbf{r}''|\mathbf{r}) \cong g(\mathbf{r}''), \tag{B.6}$$

Eq. B.4 simplifies to

$$\langle \Delta \mathbf{A}^I|\mathbf{r}\rangle = \int \mathbf{A}^{I:(2)}(\mathbf{r}'',\mathbf{r})g(\mathbf{r}'')d\mathbf{r}'' - \int \mathbf{A}^{I:(1)}(\mathbf{r}',\mathbf{r})g(\mathbf{r}')d\mathbf{r}', \tag{B.7}$$

where $\mathbf{A}^{I:(k)}(\mathbf{r}',\mathbf{r})$ can be computed as

$$\mathbf{A}^{(k)}(\mathbf{r}',\mathbf{r}) = \frac{\mathbf{F}^{(k)}(\mathbf{r}')}{m} = -\frac{1}{m}\nabla U(r'), \tag{B.8}$$

where $k$ is 1 or 2, and $m$ is the mass of particles. Now Eq. B.7 for the conditional average indirect relative acceleration can be calculated analytically for a specified pair correlation and pair potential.

A dilute system with dimensionless number density $\hat{n} = n\sigma^3 = 0.1$ is considered with spherical particles whose centers are distributed according to a Matérn hard-core point process [30], which has an analytic form for the pair correlation function $g(r)$. Since the analytical calculations of the average relative acceleration in 3-D (dimension of the space in which the sphere centers are distributed) is very challenging, we consider simpler 1-D and 2-D systems. In the 1-D system sphere centers are distributed on a line, and in the 2-D system the sphere centers are distributed in a plane. Particles undergo pairwise interactions governed by the Lennard-Jones potential (Eq. 3.2) with potential well depth $\hat{\varepsilon} = 1.0$. The solid lines in both panels of Figure 3.14 show the analytical result for the conditional average indirect relative acceleration. Positive values of relative acceleration indicate that the other particles induce an effective repulsive force between the pair, while negative values indicate effective attraction. The effect of dimensionality is seen by comparing panels (a) and (b) of Figure 3.14 in the range $1 < \hat{r} < 2$.

The slightly attractive behavior in the 1-D case for $1 < \hat{r} < 2$ is a consequence of the restricted geometric arrangements that are possible when all three particle centers are distributed on a line (the probe particle cannot be inserted between particles (1) and (2) below the minimum separation of $\hat{r} = 2$). On the other hand, if particles are distributed in a plane (or in 3-D) this restriction is absent and the effective force is always repulsive.

For the computations a system of 150,000 identical particles in the 1-D case and 823,000 identical particles in the 2-D case were generated according to the Matérn process with hardcore distance $h = \sigma$, and reduced number density $\hat{n} = 0.1$. For each of the $N(N-1)/2$ solute particle pairs separated by $r$, the indirect relative acceleration for this pair due to all other $N-2$ particles is calculated (the direct interaction between particles in the pair is excluded). Subsequently this data is binned in separation space $r$, and the conditional average indirect relative acceleration $\langle \Delta \mathbf{A}^I | \mathbf{r} \rangle$ is computed. Multiple independent simulations corresponding to different particle configurations are performed to reduce statistical error. The results of these computations are shown in Figure 3.14, where $\langle \Delta \hat{\mathbf{A}}^I | \hat{\mathbf{r}} \rangle = \langle \Delta \mathbf{A}^I | \mathbf{r} \rangle \sigma m / \varepsilon$ and $\hat{r} = r/\sigma$. The range of interaction $\hat{r}$ for these computations is larger than the cutoff distance typically used for the potential calculations as shown in Figure 3.14. Excellent agreement is found between $\langle \Delta \mathbf{A}^I | \mathbf{r} \rangle$ calculated analytically using Eq. B.7 and computed data for 1-D and 2-D cases. These results verify our computation of the indirect average relative acceleration.

In aggregation problems the relative acceleration needs to be computed for a mixture of at least of two types of particles. To verify the computations for a mixture of particles we randomly tagged a specified fraction (29.5% in our test) of particles as type $A$ particles and the rest as type $B$ particles. The potential well depth for $A$-$A$ interactions was varied from that of $B$-$B$ interactions such that $\varepsilon_{AA}/\varepsilon_{BB} = 8$. We then computed the indirect average relative acceleration between $A$ type particles due to all other particles. We obtained a good match between our computed results and the analytical expression but these results are not shown here for brevity. However, we do observe wider spread in the computational data than in the previous case which is attributed to a significant reduction of the number of pairs that are involved in the relative acceleration computation.

These test runs for the indirect average relative acceleration verify our computations for the pure solvent system of identical particles as well as for the mixture of different types of particles. Good agreement between computations and the analytical expression is observed for both 1-D and 2-D cases. When the relative acceleration is extracted from a mixture, additional independent simulations are needed to compensate for fewer samples.

# APPENDIX C. DERIVATION OF PAIR CORRELATION EXPRESSION FOR A BINARY MIXTURE

Here we derive the expression for the pair–correlation function $g_{\alpha\beta}(r), \alpha, \beta \in (A, B)$ in a binary system that contains two types of particles $A$ and $B$ whose centers are distributed as statistically homogeneous and isotropic point fields. The *second factorial moment measure* of a point field (see Stoyan and Stoyan [30]) is generalized to a binary system with two particle types as

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_1 \times \mathscr{V}_2) = \langle N_\alpha(\mathscr{V}_1)\ [N_\beta(\mathscr{V}_2) - 1] \rangle\,, \tag{C.1}$$

where $\mathscr{V}_1$ and $\mathscr{V}_2$ are sets in physical space, $N_\alpha(\mathscr{V}_1)$ is the number of $\alpha$ particles in region $\mathscr{V}_1$, and $N_\beta(\mathscr{V}_2)$ is the number of $\beta$ particles in region $\mathscr{V}_2$. The second factorial moment measure $\mu^{(2)}_{\alpha\beta}(\mathscr{V}_1 \times \mathscr{V}_2)$ has a density $\rho^{(2)}_{\alpha\beta}(\mathbf{x}_1, \mathbf{x}_2)$ such that it can be written as an integral

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_1 \times \mathscr{V}_2) = \int_{\mathscr{V}_1} \int_{\mathscr{V}_2} \rho^{(2)}_{\alpha\beta}(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \tag{C.2}$$

This *second-order product density* $\rho^{(2)}_{\alpha\beta}(\mathbf{x}_1, \mathbf{x}_2)$ is the unnormalized pair correlation function.

For a statistically homogeneous point field the second-order product density $\rho^{(2)}_{\alpha\beta}(\mathbf{x}_1, \mathbf{x}_2)$ depends only on the pair separation $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$. It is then convenient to transform $\mathscr{V}_1 \times \mathscr{V}_2$ to $\mathscr{V}_R \times \mathscr{V}_r$ in $(\mathbf{R}, \mathbf{r})$ space with $\mathbf{R} = (\mathbf{x}_1 + \mathbf{x}_2)/2$ and $\rho^{(2)}_{\alpha\beta}(\mathbf{R}, \mathbf{r})J = \rho^{(2)}_{\alpha\beta}(\mathbf{x}_1, \mathbf{x}_2)$, where the Jacobian of the transformation $J = |\partial(\mathbf{x}_1, \mathbf{x}_2)/\partial(\mathbf{R}, \mathbf{r})|$ is unity, leading to

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_1 \times \mathscr{V}_2) = \mu^{(2)}_{\alpha\beta}(\mathscr{V}_R \times \mathscr{V}_r) = \int_{\mathscr{V}_R} \int_{\mathscr{V}_r} \rho^{(2)}_{\alpha\beta}(\mathbf{R}, \mathbf{r}) d\mathbf{R} d\mathbf{r}. \tag{C.3}$$

For homogeneous and isotropic point fields, the second-order product density $\rho^{(2)}_{\alpha\beta}$ depends only on the scalar separation distance $r = |\mathbf{r}|$, and can be written as

$$\rho^{(2)}_{\alpha\beta}(r) = n_\alpha n_\beta g_{\alpha\beta}(r), \tag{C.4}$$

where $n_\alpha$ and $n_\beta$ are the number densities of the $\alpha$-type and $\beta$-type particles, respectively. Substituting this expression into Eq. C.3, we obtain

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_R \times \mathscr{V}_r) = \int_{\mathscr{V}_R} \int_r n_\alpha n_\beta g_{\alpha\beta}(r) 4\pi r^2 d\mathbf{R}\, dr, \tag{C.5}$$

where the integral over $\mathscr{V}_r$ has been simplified using a spherical volume element $4\pi r^2 dr$. Noting that

$$\langle N_\alpha(\mathscr{V}_R)\rangle = \int_{\mathscr{V}_R} n_\alpha d\mathbf{R}, \tag{C.6}$$

and considering the case where $\mathscr{V}_r$ is a spherical shell with volume $V(r,\Delta r) = 4\pi r^2 \Delta r$ we obtain

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_R \times \mathscr{V}_r) = \langle N_\alpha(\mathscr{V}_R)\rangle\, n_\beta g_{\alpha\beta}(r) 4\pi r^2 \Delta r, \tag{C.7}$$

provided $\Delta r$ is smaller than the scale of variation of $g_{\alpha\beta}(r)$.

Noting that the equivalent expression for $\langle N_\alpha(\mathscr{V}_1)\,[N_\beta(\mathscr{V}_2) - 1]\rangle$ in Eq. C.1 is

$$\mu^{(2)}_{\alpha\beta}(\mathscr{V}_R \times \mathscr{V}_r) = \langle N_\alpha(\mathscr{V}_R)\,[N_\beta(\mathscr{V}_r) - 1]\rangle\,,$$

leads to the following estimate for the pair correlation from particle data

$$g_{\alpha\beta}(r) = \frac{\langle N_{\alpha\beta}(r,\Delta r)\rangle}{\langle N_\alpha\rangle\, n_\beta V(r,\Delta r)} \tag{C.8}$$

where $\langle N_{\alpha\beta}(r,\Delta r)\rangle$ is the average number of $\alpha$–$\beta$ pairs with a $\beta$ particle in a shell $(r,\Delta r)$ separated by $r$ from an $\alpha$ particle. For the NVT ensemble considered in these simulations the total number of $\alpha$ and $\beta$ particles is a constant, so it is appropriate to replace $\langle N_\alpha\rangle$ in Eq. C.8 by the total number of $\alpha$ particles $N_\alpha$ in the domain, leading to

$$g_{\alpha\beta}(r) = \frac{\langle N_{\alpha\beta}(r,\Delta r)\rangle}{N_\alpha n_\beta V(r,\Delta r)} \tag{C.9}$$

In the computations, we estimate $\langle N_{\alpha\beta}(r,\Delta r)\rangle$ as follows. For the $i$th $\alpha$-particle we compute the number of $\beta$-particles $N^{(i)}_{\alpha\beta}(r,\Delta r)$ whose centers $\mathbf{r}^{(j)}$ relative to the $\alpha$ particle are located at a distance $|\mathbf{r}^{(j)}| \in (r,\Delta r)$. The average number of such pairs $\langle N(r,\Delta r)\rangle$ is estimated by averaging over all the $\alpha$ particles

$$\frac{1}{N_\alpha}\sum_{i=1}^{N_\alpha} N^{(i)}_{\alpha\beta}(r,\Delta r)$$

Substituting this estimate into Eq. C.9 leads to

$$g_{\alpha\beta}(r) \approx \frac{1}{N_\alpha n_\beta V(r,\Delta r)} \left( \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} N_{\alpha\beta}^{(i)}(r,\Delta r) \right) \tag{C.10}$$

The remaining factor $n_\beta V(r,\Delta r)$ in the denominator is simply the expected number of $\beta$ particles in the shell $(r,\Delta r)$.

# APPENDIX D. DERIVATION OF THE TRANSPORT EQUATION FOR THE TWO–PARTICLE DENSITY $\rho^{(2)}$

In the Klimontovich approach [31, 32], the ensemble of particles is characterized by a fine–grained density function $f_1'$ that is defined in a six-dimensional position-velocity space $[\mathbf{x}, \mathbf{v}]$ as

$$f_1'(\mathbf{x}, \mathbf{v}, t) \equiv \sum_{i=1}^{N} f_1'^{(i)} = \sum_{i=1}^{N} \delta(\mathbf{x} - \mathbf{X}^{(i)}(t))\delta(\mathbf{v} - \mathbf{V}^{(i)}(t)), \tag{D.1}$$

where the shortened notation

$$f_1'^{(i)} = \delta(\mathbf{x} - \mathbf{X}^{(i)}(t))\delta(\mathbf{v} - \mathbf{V}^{(i)}(t))$$

is used to represent the delta function associated with the $i$th particle. The number of particles in any region $\mathscr{B}$ in $[\mathbf{x}, \mathbf{v}]$ space can be obtained by integrating the fine–grained density $f_1'$ as follows:

$$N(\mathscr{B}) = \int_{\mathscr{B}} f_1' d\mathbf{x}\, d\mathbf{v}. \tag{D.2}$$

The ensemble average of the Klimontovich fine–grained density function $f_1'$ is the one–particle density function $f$, which is written as

$$f(\mathbf{x}, \mathbf{v}, t) = \langle f_1' \rangle = \left\langle \sum_{i=1}^{N} f_1'^{(i)} \right\rangle = \left\langle \delta(\mathbf{x} - \mathbf{X}^{(i)}(t))\delta(\mathbf{v} - \mathbf{V}^{(i)}(t)) \right\rangle. \tag{D.3}$$

Integrating the one–particle density over velocity space results in the number density $n(\mathbf{x}, t)$ that forms the basis for the continuum hydrodynamic description

$$n(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t)\, d\mathbf{v}, \tag{D.4}$$

which in turn can be integrated over physical space to obtain the expected number of particles:

$$\langle N \rangle = \int n(\mathbf{x}, t)\, d\mathbf{x}. \tag{D.5}$$

In order to characterize structural properties such as the pair correlation function, we need to consider the two–particle density. The one–point fine–grained density in the Klimontovich approach can be extended to its two–particle counterpart as follows [33]:

$$f_1' f_2' = \sum_{i=1}^{N} f_1'^{(i)} \sum_{\substack{j=1 \\ j\neq i}}^{N} f_2'^{(j)} = \sum_{i=1}^{N} \delta(\mathbf{x}_1 - \mathbf{X}^{(i)}(t))\delta(\mathbf{v}_1 - \mathbf{V}^{(i)}(t)) \sum_{\substack{j=1 \\ j\neq i}}^{N} \delta(\mathbf{x}_2 - \mathbf{X}^{(j)}(t))\delta(\mathbf{v}_2 - \mathbf{V}^{(j)}(t)) \tag{D.6}$$

where $[\mathbf{x}_k, \mathbf{v}_k, k = 1, 2]$ are the Eulerian coordinates of the position–velocity phase space for the particle pair. (The summation over distinct pairs $j \neq i$ is necessary for the definition of the two–particle density, whose integral is the second factorial measure. If all pairs are included, an atomic contribution arises in the second moment measure that does not have a density [30, 34].) The ensemble average of the two–particle fine–grained density function $f_1' f_2'$ is the two–particle density $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)$, which is defined as

$$\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t) \equiv \langle f_1' f_2' \rangle. \tag{D.7}$$

Integrating the two–particle density over the velocity spaces results in the unnormalized pair–correlation function

$$\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, t) = \int \rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)\, d\mathbf{v}_1 d\mathbf{v}_2, \tag{D.8}$$

which in turn can be integrated over a region $\mathscr{B}$ in physical space to obtain the second factorial moment measure:

$$\langle N(\mathscr{B})\ [N(\mathscr{B}) - 1] \rangle = \int \rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, t) d\mathbf{x}_1 d\mathbf{x}_2. \tag{D.9}$$

Substituting Eq. D.6 into Eq. D.7, and differentiating Eq. D.7 with respect to time results in the evolution equation for the two–particle density $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)$:

$$\frac{\partial \rho^{(2)}}{\partial t} =$$

$$\left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \Big( -\mathbf{V}^{(i)}(t)\frac{\partial}{\partial \mathbf{x}_1}\left[\delta(\mathbf{x}_1-\mathbf{X}^{(i)}(t))\delta(\mathbf{v}_1-\mathbf{V}^{(i)}(t))\delta(\mathbf{x}_2-\mathbf{X}^{(j)})\delta(\mathbf{v}_2-\mathbf{V}^{(j)}(t))\right] \right.$$
$$-\mathbf{V}^{(j)}\frac{\partial}{\partial \mathbf{x}_2}\left[\delta(\mathbf{x}_1-\mathbf{X}^{(i)}(t))\delta(\mathbf{v}_1-\mathbf{V}^{(i)}(t))\delta(\mathbf{x}_2-\mathbf{X}^{(j)}(t))\delta(\mathbf{v}_2-\mathbf{V}^{(j)}(t))\right]$$
$$-\mathbf{A}^{(i)}\frac{\partial}{\partial \mathbf{v}_1}\left[\delta(\mathbf{x}_1-\mathbf{X}^{(i)}(t))\delta(\mathbf{v}_1-\mathbf{V}^{(i)}(t))\delta(\mathbf{x}_2-\mathbf{X}^{(j)}(t))\delta(\mathbf{v}_2-\mathbf{V}^{(j)}(t))\right]$$
$$\left. -\mathbf{A}^{(j)}\frac{\partial}{\partial \mathbf{v}_2}\left[\delta(\mathbf{x}_1-\mathbf{X}^{(i)}(t))\delta(\mathbf{v}_1-\mathbf{V}^{(i)}(t))\delta(\mathbf{x}_2-\mathbf{X}^{(j)}(t))\delta(\mathbf{v}_2-\mathbf{V}^{(j)}(t))\right] \Big) \right\rangle, \quad \text{(D.10)}$$

where $\mathbf{V}^{(i)} = \partial \mathbf{X}^{(i)}/\partial t$ represents the velocity of the $i$th particle, and $\partial \mathbf{V}^{(i)}/\partial t = \mathbf{A}^{(i)} = \mathbf{F}^{(i)}/m$ represents the acceleration of the $i$th particle. Here we have used the chain rule and the following identity:

$$\frac{\partial}{\partial a} f(a-b) = -\frac{\partial}{\partial b} f(a-b). \quad \text{(D.11)}$$

Now substituting the relation

$$a \cdot \delta(a-b) = b \cdot \delta(a-b), \quad \text{(D.12)}$$

in Eq. D.10 leads to

$$\frac{\partial \rho^{(2)}}{\partial t} = -\left\langle \frac{\partial}{\partial \mathbf{x}_1}\left[\mathbf{v}_1 f_1' f_2'\right]\right\rangle - \left\langle \frac{\partial}{\partial \mathbf{x}_2}\left[\mathbf{v}_2 f_1' f_2'\right]\right\rangle$$
$$-\frac{\partial}{\partial \mathbf{v}_1}\left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{A}^{(i)} f_1'^{(i)} f_2'^{(j)} \right\rangle - \frac{\partial}{\partial \mathbf{v}_2}\left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{A}^{(j)} f_1'^{(i)} f_2'^{(j)} \right\rangle. \quad \text{(D.13)}$$

We now define the following functions in phase spase:

$$\langle \mathbf{A}^{(1)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t\rangle \equiv \frac{1}{\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)} \left[ \left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{A}^{(i)} f_1'^{(i)}(\mathbf{x}_1, \mathbf{v}_1, t) f_2'^{(j)}(\mathbf{x}_2, \mathbf{v}_2, t) \right\rangle \right]$$

$$\langle \mathbf{A}^{(2)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t\rangle \equiv \frac{1}{\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t)} \left[ \left\langle \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathbf{A}^{(j)} f_1'^{(i)}(\mathbf{x}_1, \mathbf{v}_1, t) f_2'^{(j)}(\mathbf{x}_2, \mathbf{v}_2, t) \right\rangle \right]$$

if $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t) > 0$, and substituting these definitions into Eq. D.13 results in the following evolution equation for $\rho^{(2)}$:

$$\frac{\partial \rho^{(2)}}{\partial t} = -\frac{\partial}{\partial \mathbf{x}_1}\left(\mathbf{v}_1 \rho^{(2)}\right) - \frac{\partial}{\partial \mathbf{x}_2}\left(\mathbf{v}_2 \rho^{(2)}\right) - \frac{\partial}{\partial \mathbf{v}_1}\left(\langle \mathbf{A}^{(1)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t \rangle \rho^{(2)}\right) - \frac{\partial}{\partial \mathbf{v}_2}\left(\langle \mathbf{A}^{(2)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t \rangle \rho^{(2)}\right). \tag{D.14}$$

Introducing the pair relative separation $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$ and the pair relative velocity $\mathbf{w} = \mathbf{v}_2 - \mathbf{v}_1$, and assuming statistical homogeneity in physical space and velocity space, leads to the following form for the evolution of the two–particle density

$$\frac{\partial \rho^{(2)}}{\partial t} + \frac{\partial}{\partial \mathbf{r}}\left[\mathbf{w} \rho^{(2)}\right] + \frac{\partial}{\partial \mathbf{w}}\left[\langle \Delta \mathbf{A}^{(2)(1)} | \mathbf{r}, \mathbf{w}; t \rangle \rho^{(2)}\right] = 0, \tag{D.15}$$

where

$$\langle \Delta \mathbf{A}^{(2)(1)} | \mathbf{r}, \mathbf{w}; t \rangle = \langle \mathbf{A}^{(2)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t \rangle - \langle \mathbf{A}^{(1)} | \mathbf{x}_1, \mathbf{x}_2, \mathbf{v}_1, \mathbf{v}_2, t \rangle$$

is the average relative acceleration between particles 1 and 2. The angle brackets represent averaging over all three–particle (and higher multiparticle) statistics.

# APPENDIX E. LIGHT SCATTERING ANALYSIS

Light scattering analysis (LS) is proposed as a direct method for investigation of aggregate structure [35, 36, 37, 38]. This method allows measurement of $D_f$ even for a single cluster. In addition, the light scattering technique gives cluster structure information for the wide range of scales: from monomer size to the geometric size of a cluster. This feature provides for the discovery of any possible structure changes at different length scales, especially for the relatively large clusters. This gives the most complete description of the aggregate structure.

For the system of $N$ nanoparticles, the intensity of elastically scattered light can be represented as

$$I(\mathbf{q}) = NF(\mathbf{q})S(\mathbf{q}) \tag{E.1}$$

where $\mathbf{q}$ is the scattering wave vector, which is defined as

$$q = |\mathbf{q}| = \frac{4\pi}{\lambda_l} sin(\theta/2) \tag{E.2}$$

where $\theta$ is the scattering angle, $\lambda_l$ is the wavelength of light, $S(\mathbf{q})$ is the static structure factor, where

$$S(\mathbf{q}) = \frac{1}{N} \sum_{k}^{N} \sum_{l}^{N} exp[i\mathbf{q} \cdot (\mathbf{r}_k - \mathbf{r}_l)] \tag{E.3}$$

Due to a spherical shape of nanoparticles with uniform density, Eq. E.3 can be reduced to

$$S(\mathbf{q}) = \frac{1}{N} \left| \sum_{k}^{N} exp[i\mathbf{q} \cdot \mathbf{r}_k] \right|^2 = \frac{1}{N} \left| \sum_{k}^{N} [cos(\mathbf{q} \cdot \mathbf{r}_k) + i\ sin(\mathbf{q} \cdot \mathbf{r}_k))] \right|^2 \tag{E.4}$$

and $F(\mathbf{q}) \rightarrow F(q)$ is the form factor for a sphere,

$$F(q) = \left[ 3 \frac{sin(qa) - qa\ cos(qa)}{(qa)^3} \right]^2 \tag{E.5}$$

Because all modeled systems are isotropic in cluster position and orientation it is valid to perform a spherical averaging by selecting over 200 different $\mathbf{q}$ values of constant magnitude, so $S(\mathbf{q}) \rightarrow S(q)$. This was done by creating the set of angles $(\theta, \phi)$ according to a uniform differential solid angle $d\Omega$, and $q$ is calculated from:

$$\mathbf{q} = |\mathbf{q}|(sin(\theta)cos(\phi)\mathbf{e}_1 + sin(\theta)sin(\phi)\mathbf{e}_2 + cos(\theta)\mathbf{e}_3) \tag{E.6}$$

where $\mathbf{e}_i$ is the $i$-th Cartesian unit vector.

In case of a self-similar fractal aggregate with a fractal dimension $D_f$, $I(q)$ has the following three regimes: the first regime is for small values of $q$ (the so-called Rayleigh regime), where $I(q) = N$ the number of monomers per cluster. The second regime is for intermediate values of $q$, where $I(q) \sim q^{-D_f}$. The third regime is for very large $q$ when Porod's law can be applied, so $I(q) \sim q^{-4}$ [39].

To validate the in-house LS code, a system of 100,000 nanoparticles where distributed in a 3-d square lattice with period $d_l = 2$. Positions of the light scattering peaks in the crystal lattice must be distributed according to Bragg's law

$$2d_l \sin(\theta/2) = n\lambda_l \tag{E.7}$$

where $\lambda_l$ is the wavelength of an incident light, $n$ is the integer corresponding to the order of intensity peak, and $\theta$ is the angle between incident light and the scattering panels. The value for the wavelength is chosen arbitrarily to be $\lambda_l = 0.251$.

Table E.1: Validation of LS code

| n | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\theta$ (Bragg's law) | 7.201 | 14.431 | 21.719 | 29.097 |
| $\theta$ (LS code) | 7.197 | 14.435 | 21.719 | 29.094 |

For such a system, an LS analysis is performed and correspondent positions of peaks are computed and compared with analytically expected ones(Table E.1). The excellent match of computational data with analytical values verifies the LS code and allows us to use it for systems with unknown structures.

In BD (as well as in MD) the conservative force is not calculated for all the possible pair separations between nanoparticles. Instead, such parameter as a cut-off distance $r_c$ is introduced. In this way, force is calculated for the particles separated by a cut-off distance or smaller. All other separations do not influence the force calculations. This approach significantly decreases simulation time without interfering with computational accuracy. Such an approach is applicable for a monotonically decaying interaction potential which is very close to zero value at $r > r_c$. Therefore, it would be interesting to see how LS results depend on the $r_c$ value. For this purpose BD simulations of the aggregation of colloid nanoparticles which interact through the Lennard-Jones potential were performed for $r_c = 2.5\sigma = 5a$ and $r_c = 1.5\sigma = 3a$, where $a$ is the nanopartile radius. All the parameters for these simulations are described in Table E.2. Figure $E.1$ represents LS analysis of the largest clusters obtained from these BD simulations. The largest clusters consist of $N = 8,900$ and $N = 8,717$ nanoparticles for $r_c = 3a$ and $r_c = 5a$ correspondingly. The positions of cut-off distances are marked by a dashed line and labeled correspondently. According to the previous work [36], in the interval $a/R_{g,G} < qa < 2$ structure factor $S(q)$ should be $\sim q^{-D_f}$, where $D_f = 1.78$ is the fractal dimension, and $R_{g,G}$ is the ideal

Table E.2: Simulation parameters used to produce Figure E.1. Particle interactions are modeled using Lennard-Jones potentials.

| Parameter | Description | Value |
|---|---|---|
| $N_p$ | Number of Nanoparticles | 108,882 |
| $\varepsilon/k_BT_\infty$ | Reduced Well Depth | 4 |
| $f_{v,nanop}$ | Solute Volume Fraction | 0.035 |
| $\sigma$ | Particle Diameter | $3.4 \times 10^{-10}$ m. |
| $D_\infty/\sigma\sigma_{v_\infty}$ | Dimensionless Diffusion Coefficient | 0.524 |
| $t_{stop}D_\infty/\sigma^2$ | Dimensionless Simulation Time | 8.65 |
| $\sigma\Delta t/\sigma_{v_\infty}$ | Dimensionless Computational Time Step | 0.0025 |

gel point radius of gyration, and its defined as

$$R_{g,G} = a\left[k_0^{-1}\left(\frac{D_f+2}{D_f}\right)^{d/2}\alpha\right]^{1/(D_f-d)} \tag{E.8}$$

where $k_0 \simeq 1.3$, $d$ is the space dimension, and $\alpha$ is the nanoparticle volume fraction. However, in this case $D_f \simeq 3$ for the range $(a/r_c < qa < 2)$ for both cut-off distances. Such behavior can be explained that by implementing Lennard-Jones potential into BD, the structure of the aggregate cannot be described correctly for the scale of a potential's range. At the same time, by setting up the cut-off distance $r_c \to \infty$ the LS results should be similar to those obtained for $r_c = 5a$, because for LJ potential as $r > 5a$, the interaction potential $U \to 0$. Thus, the only part of the LS curve with $qa < a/r_c$ can be taken into account for further analysis. In the LS plot there are two ranges for $qa$ that can be clearly defined. In the range $(R_{g,G} < qa < a/r_c)$, results are similar with the well-known results for DCLA where $D_f \sim 1.8$, when for $(qa < R_{g,G})$ the results are close to the DLA, where $D_f \sim 2.5$ [39]. However, it is hard to make a final decision about fractal dimension values because of a very short dynamic range where $D_f$ can

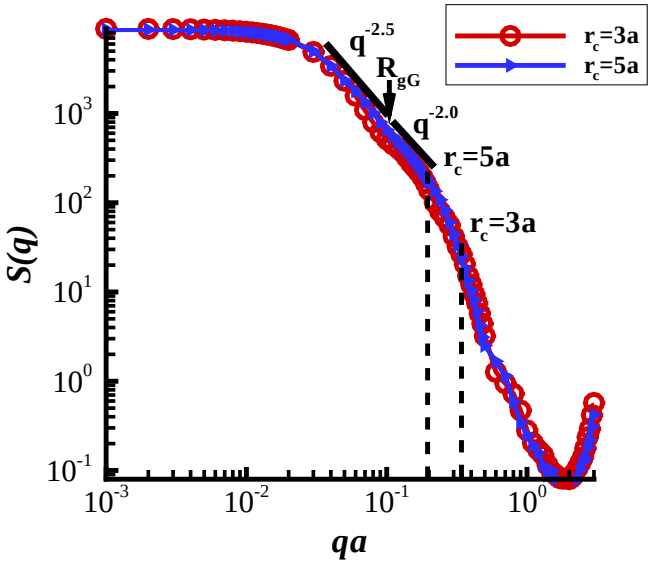



Figure E.1: Structure of the largest cluster for 3-D BD simulation with implemented LJ potential for $r_c = 5.0a$ and $r_c = 3.0a$.

be measured. An increase of dynamic range can be reached in a different ways. The first is to decrease the cut-off distance $r_c$. In this case, the dynamic range will increase, however the simulation accuracy reduces significantly due to a loss of accuracy during force calculations (even LS results look similar for both values of $r_c$ in Figure E.1). Therefore, the first approach cannot be accepted. The second way is to shift the ideal gel point radius of gyration to the left. This can be done by decreasing the nanoparticle volume fraction. However, the aggregation process takes longer for systems with a lower volume fraction. Moreover, the number of nanoparticles in the system should be increased to maintain a high possibility of large cluster formation. For example, by reducing the nanoparticle volume fraction to $\alpha = 0.015$, and increasing the number of nanoparticles in a system to $N = 500,000$, the computational time for BD simulation would be 50-70 hours on a single processor in order to obtain the maximum cluster containing up to $50,000$ nanoparticles. For even larger systems, a single processor BD simulation cannot be done in a reasonable computational time and some additional steps to speed up BD calculations are required. Thus, to speed up BD simulations the BD code must be parallelized.

# APPENDIX F. EVOLUTION OF THE SECOND-ORDER DENSITY FOR MD MODEL

Differentiating Eq. 3.10 with respect to time results in

$$\begin{aligned}
&\frac{\partial \rho^{(2)}(x_1, x_2, v_1, v_2, t)}{\partial t} = \\
&\langle -\sum_{i=1}^{N}\sum_{j=1}^{N} V_k^{(i)} \frac{\partial}{\partial x_{1k}} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)}) \\
&-\sum_{i=1}^{N}\sum_{j=1}^{N} A_k^{(i)} \frac{\partial}{\partial v_{1k}} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)}) \\
&-\sum_{i=1}^{N}\sum_{j=1}^{N} V_k^{(j)} \frac{\partial}{\partial x_{2k}} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)}) \\
&-\sum_{i=1}^{N}\sum_{j=1}^{N} A_k^{(j)} \frac{\partial}{\partial v_{2k}} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)})\rangle
\end{aligned} \tag{F.1}$$

where $V_k^{(i)}$ and $V_k^{(j)}$ represent the velocity of $i$-th and $j$-th nanoparticles correspondently along Cartesian coordinate $k$, and $A_k^{(i)}$ and $A_k^{(j)}$ represent the acceleration of $i$-th and $j$-th nanoparticles correspondently along Cartesian coordinate $k$. By substituting the relation

$$a \cdot \delta(a - b) = b \cdot \delta(a - b), \tag{F.2}$$

an expression is changing to

$$\begin{aligned}
&\frac{\partial \rho^{(2)}(x_1, x_2, v_1, v_2, t)}{\partial t} = \\
&\langle -\frac{\partial}{\partial x_{1k}}(v_{1k}\langle f_1' f_2'\rangle) - \frac{\partial}{\partial v_{1k}} \sum_{i=1}^{N}\sum_{j=1}^{N} (A_k^{(i)} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)})) \\
&-\frac{\partial}{\partial x_{2k}}(v_{2k}\langle f_1' f_2'\rangle) - \frac{\partial}{\partial v_{2k}} \sum_{i=1}^{N}\sum_{j=1}^{N} (A_k^{(j)} \delta(\mathbf{v}_1 - \mathbf{V}^{(i)})\delta(\mathbf{x}_1 - \mathbf{X}^{(i)})\delta(\mathbf{v}_2 - \mathbf{V}^{(j)})\delta(\mathbf{x}_2 - \mathbf{X}^{(j)}))\rangle
\end{aligned} \tag{F.3}$$

By defining the following function in phase space:

$$\langle A_k^{(i)}|x_1,x_2,v_1,v_2,t\rangle\rho^{(2)}(x_1,x_2,v_1,v_2,t) = \tag{F.4}$$
$$\sum_{i=1}^{N}\sum_{j=1}^{N}(A_k^{(i)}\delta(\mathbf{v}_1-\mathbf{V}^{(i)})\delta(\mathbf{x}_1-\mathbf{X}^{(i)})\delta(\mathbf{v}_2-\mathbf{V}^{(j)})\delta(\mathbf{x}_2-\mathbf{X}^{(j)}))$$

An expression for the second-order density is

$$\frac{\partial\rho^{(2)}(x_1,x_2,v_1,v_2,t)}{\partial t} = -\frac{\partial}{\partial x_{1k}}\left(v_{1k}\rho^{(2)}(x_1,x_2,v_1,v_2,t)\right) \tag{F.5}$$
$$-\frac{\partial}{\partial x_{2k}}\left(v_{2k}\rho^{(2)}(x_1,x_2,v_1,v_2,t)\right) - \frac{\partial}{\partial v_{1k}}\left(\langle A_k^{(i)}|x_1,x_2,v_1,v_2,t\rangle\rho^{(2)}(x_1,x_2,v_1,v_2,t)\right)$$
$$-\frac{\partial}{\partial v_{2k}}\left(\langle A_k^{(j)}|x_1,x_2,v_1,v_2,t\rangle\rho^{(2)}(x_1,x_2,v_1,v_2,t)\right)$$

Now, by changing variables in the following way $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$ and $\mathbf{w} = \mathbf{v}_2 - \mathbf{v}_1$, the final expression can be written in the vector form as

$$\frac{\partial\rho^{(2)}(\mathbf{r},\mathbf{w},t)}{\partial t} + \frac{\partial}{\partial\mathbf{r}}\left(\mathbf{w}\rho^{(2)}(\mathbf{r},\mathbf{w},t)\right) \tag{F.6}$$
$$-\frac{\partial}{\partial\mathbf{w}}\left(\langle\mathbf{A}^{(i)}|\mathbf{r},\mathbf{w},t\rangle\rho^{(2)}(\mathbf{r},\mathbf{w},t)\right) + \frac{\partial}{\partial\mathbf{w}}\left(\langle\mathbf{A}^{(j)}|\mathbf{r},\mathbf{w},t\rangle\rho^{(2)}(\mathbf{r},\mathbf{w},t)\right) = 0$$

And finally,

$$\frac{\partial\rho^{(2)}(\mathbf{r},\mathbf{w},t)}{\partial t} + \nabla_{\mathbf{r}}\cdot\left(\mathbf{w}\rho^{(2)}(\mathbf{r},\mathbf{w},t)\right) + \nabla_{\mathbf{w}}\cdot\left(\langle\Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle\rho^{(2)}(\mathbf{r},\mathbf{w},t)\right) = 0 \tag{F.7}$$

where $\mathbf{r}$ represents the pair separation vector, $\mathbf{w}$ represents the relative velocity vector, and $\langle\Delta\mathbf{A}|\mathbf{r},\mathbf{w},t\rangle$ represents the expected relative acceleration $\Delta\mathbf{A} = \mathbf{A}^{(j)} - \mathbf{A}^{(i)}$ conditional on relative velocity and pair separation.

In Molecular Dynamics case, an acceleration, experienced by $i$-th particle is

$$\mathbf{A}_{MD}^{(i)} = \frac{d\mathbf{V}^{(i)}}{dt} = \frac{\mathbf{F}_{MD}^{(i)}}{m} \tag{F.8}$$

where $\mathbf{F}_{MD}^{(i)}$ is the force that $i$-th particle experienced due to an interaction with all other particles.

Thus, in MD case an expression for the evolution of the second order density is

$$\frac{\partial\rho_{MD}^{(2)}(\mathbf{r},\mathbf{w},t)}{\partial t} + \nabla_{\mathbf{r}}\cdot\left(\mathbf{w}\rho_{MD}^{(2)}(\mathbf{r},\mathbf{w},t)\right) + \nabla_{\mathbf{w}}\cdot\left(\langle\Delta\mathbf{A}_{MD}|\mathbf{r},\mathbf{w},t\rangle\rho_{MD}^{(2)}(\mathbf{r},\mathbf{w},t)\right) = 0 \tag{F.9}$$

where $\Delta\mathbf{A}_{MD} = \mathbf{A}_{MD}^{(j)} - \mathbf{A}_{MD}^{(i)}$.

# APPENDIX G. HYDRODYNAMIC EFFECT

Here we provide estimates of the HI forces relative to the correction force from improved potential. In order to determine the importance of hydrodynamic interaction (HI) forces on aggregation, it is useful to split the HI into short-range (lubrication) HI and long range HI. Based on rough scaling arguments it appears that it is only meaningful to compare the short-ranged HI effect on aggregation to the effect of modifying the BD potential of mean force to account for solvation forces. While the long-range HI will affect the aggregate structure on large scales, the small-scale aggregate structure will be strongly influenced by solvation forces that are modeled by the improved potential of mean force in BD that we propose.

## G.1 Short-range lubrication forces

We estimate the lubrication force correction to the BD equations and compare this to the correction force arising from our improved potential that accounts for solvent effects. We show that for the equi-sized case the correction force from our improved potential is larger than, or comparable to, the short-ranged lubrication force (see Fig. G.1) (details are given below). For size-separated systems, such a direct comparison is not possible because the lubrication force applies to a pair of nanoparticle aggregates, each composed of several monomers, whereas the correction force from our improved potential applies to each monomer in these aggregates. The net effect of the correction to monomer interactions from our improved potential on the relative acceleration between a pair of stationary nanoparticle aggregates is known to be comparable to the van der Waals force from the work of Fichthorn's group [40]. Precise estimates of the scaling (with solute/solvent size ratio) of relative magnitude and range (in dimensionless separation) of the solvation force to lubrication force between nanoparticle aggregates requires

larger MD calculations than are currently available in the literature.

The lubrication force between two equi-sized solute particles is computed as [41, 42, 43]

$$\mathbf{F}_L = 6\pi\mu a_{np}\mathbf{W}\left[\frac{1}{8\varepsilon} - \frac{9}{40}\ln(2\varepsilon) + 0.99\right]. \tag{G.1}$$

In this equation $\mu$ is the dynamic viscosity of the solvent, $a_{np}$ is the solute particle radius, $\mathbf{W}$ is the relative velocity between the solute particles, $\varepsilon = h/\sigma_{np}$ is a dimensionless interparticle separation where $h$ is the minimum distance between the clear surfaces of the two solute particles, and $\sigma_{np} = 2a_{np}$ is the diameter of the solute particle. This equation is valid for $\varepsilon < 1$, but it breaks down when the separation distance $h$ becomes comparable to surface asperities, as seen from the singularity in the lubrication force at zero $\varepsilon$ [44]. The unbounded growth of the repulsive lubrication force with decreasing separation is unphysical because it predicts an infinite initial relative momentum (or attraction force) to overcome the lubrication force in order for particles to collide. In reality particles do collide with finite initial relative momentum (or attraction force) and this reflects the breakdown of the lubrication theory due to non-continuum effects [45]. Therefore, theoretical corrections truncate the lubrication force [46] for $\varepsilon < \varepsilon_{crit}$ and predict a constant value of $\mathbf{F}_L$ for $\varepsilon < \varepsilon_{crit}$. The truncated lubrication force $\mathbf{F}_{TL}$ is calculated as

$$\mathbf{F}_{TL} = 2.2\mu\mathbf{W}\frac{a_{np}^2}{\Delta}, \tag{G.2}$$

where $\Delta$ is the characteristic size of solute particle asperities, which we estimate as $0.1a_{np}$ for calculations.

In the BD model with no HI that we use in our work, the hydrodynamic force between two solute particles depends only on the relative velocity $\mathbf{W}$ (corresponding to the last term inside the square brackets on the right hand side of the expression for the lubrication force $\mathbf{F}_L$ in Eq. G.1) and does not depend on the relative separation $\mathbf{r}$ (accounted for by the first two terms inside square brackets on the right hand side of Eq. G.1). Including short-range HI into our BD model would correct the force between two solute particles by the contribution to the

lubrication force that accounts for particle separation only $\mathbf{F}^h_{TL}$, which is

$$\mathbf{F}^h_{TL} = \begin{cases} 3\pi\mu\sigma_{np}\mathbf{W}\left[\dfrac{1}{8\varepsilon} - \dfrac{9}{40}\ln(2\varepsilon)\right], & \varepsilon > \varepsilon_{crit}; \\ 2.2\mu\mathbf{W}\dfrac{a^2_{np}}{\Delta} - 3\pi\mu\sigma_{np}\mathbf{W}, & \varepsilon < \varepsilon_{crit}, \end{cases} \tag{G.3}$$

or in the dimensionless form

$$\hat{\mathbf{F}}^h_{TL} = \mathbf{F}^h_{TL}\frac{\sigma_{np}}{k_B T_{ref}} = \begin{cases} 3\pi\mu\sigma^2_{np}\dfrac{\mathbf{W}}{k_B T_{ref}}\left[\dfrac{1}{8\varepsilon} - \dfrac{9}{40}\ln(2\varepsilon)\right], & \varepsilon > \varepsilon_{crit}; \\ 2.2\mu\sigma_{np}\dfrac{\mathbf{W}}{k_B T_{ref}}\dfrac{a^2_{np}}{\Delta} - 3\pi\mu\sigma^2_{np}\dfrac{\mathbf{W}}{k_B T_{ref}}, & \varepsilon < \varepsilon_{crit}, \end{cases} \tag{G.4}$$

where $k_B$ is the Boltzmann constant, and $T_{ref}$ is the reference temperature.

Now we compare the correction $\mathbf{F}^h_{TL}$ to solute interparticle force in our BD model arising from short-range HI using the lubrication theory with the correction to force $\mathbf{F}_2$ arising from the improved BD potential described in the original MS. Figure G.1 shows the comparison

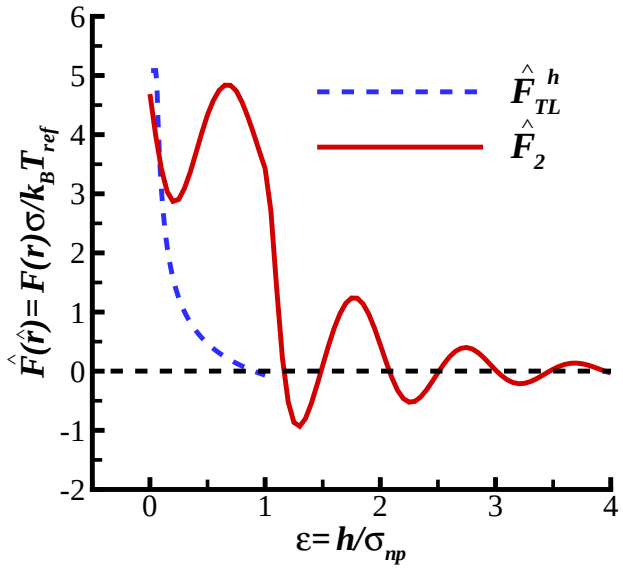



Figure G.1: Truncated dimensionless lubrication force $\hat{\mathbf{F}}^h_{TL}$ and correction force $\hat{\mathbf{F}}_2$ as function of $\varepsilon$ for equisize particles.

of dimensionless lubrication force and dimensionless correction force for equi-size particles. For equi-size particles the magnitude of the dimensionless correction force $\hat{\mathbf{F}}_2$ from improved potential is larger than the magnitude of the dimensionless lubrication force $\hat{\mathbf{F}}^h_{TL}$ for $\varepsilon > 0.2$. Therefore, short-range HI do not overwhelm the potential correction we propose for equi-sized

particles. Colloidal systems of nanoparticles in hydrocarbon solvents such as $n$-decane have solvent molecules comparable in size to solute particles [47, 48]. For size-separated systems, one could form rough estimates based on the scaling of lubrication and solvation forces. However, the estimates for the lubrication force correction depend on assumptions regarding the scale of surface asperities. Estimates for scaling of the solvation force must rely on only two size ratios that are available from the work of Fichthorn and co-workers. Therefore, more simulation data on the scaling of both these forces in size-separated systems are needed for a definitive conclusion.

## G.2 Long-range hydrodynamic interactions

The effect of large range HI on the structure in colloidal systems is studied by Heyes [49] where he compared characteristics of systems simulated by BD with and without many-body hydrodynamics. In his work he shows that the structure of colloidal system that is characterized by pair correlation function $g(r)$ remains statistically the same for systems with and without long-range HI (see Fig. G.2). Although the long term self-diffusion coefficient $D_L$ changes when HI is introduced, the pair correlation and structure are unaffected. Although it is possible that in aggregating systems long-range HI might change the structure of large-scale aggregates, these results indicate that they will not affect the structure of small-scale aggregates significantly.

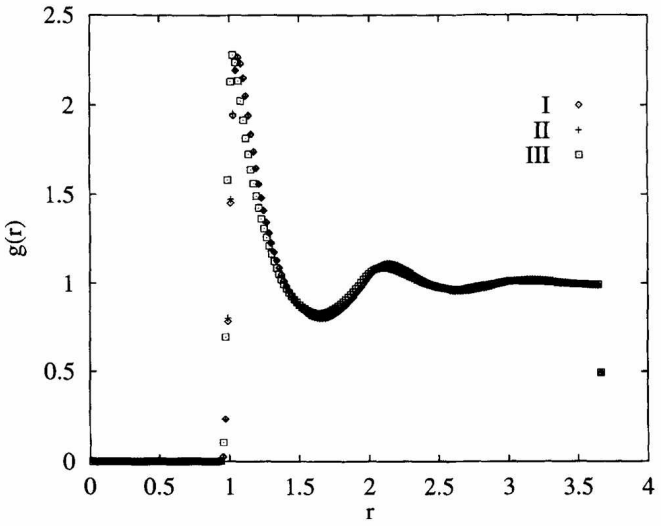



Figure G.2: The radial distribution function for the state point $N = 256$ and $\phi = 0.3403$ using the three equations of motion/algorithms: I, no many-body hydrodynamics; II, with many-body hydrodynamics; and III, many-body hydrodynamics with an incomplete algorithm. Figure is taken from Heyes work [49].

# BIBLIOGRAPHY

[1] Roco, M. C. Nanoparticles and nanotechnology research. *J. Nanop. Rec.*, 1:1–6, 1999.

[2] Salata, O. V. Application of nanoparticles in biology and medicine. *J. Nanobiotechnology*, 2:3, 2004.

[3] Hussain, N., and Florence, A. T. Utilizing bacterial mechanisms of epithelial cell entry: invasin-induced oral uptake of latex nanoparticles. *Pharm Res.*, 15:153–156, 1998.

[4] Friedlander, S. K. Nanoparticles and their structures: the next generation. *J. Nanoparticle Rec.*, 1:159–160, 1999.

[5] Cerdà, J. J., Sintes, T., Sorensen, C. M., and Charkabarti, A., Kinetics of phase transformations in depletion-driven colloids. *Phys. Rev. E*, 70:011405, 2004.

[6] Sorensen, C. M. Light scattering by fractal aggregates: A review. *Aerosol Sci. Tech.*, 35:648–687, 2001.

[7] Reichl, L.E. *A modern course in statistical physics*; University of Texas Press, 1980.

[8] Anderson, V. J., and Lekkerkerker, H. N. W. Insights into phase transition kinetics from colloid science. *Nature*, 416:811–815, 2002.

[9] de Hoog, E. H. A., Kegel, W. K., van Blaaderen, A., and Lekkerkerker, H. N. W. Direct observation of crystallization and aggregation in a phase-separating colloid-polymer suspension. *Phys. Rev. E*, 64:021407, 2001.

[10] Shepherd, R. F., Conrad, J. C., Rhodes, S. K., Link, D. R., Marquez, M., Weitz D. A., and Lewis, J. A. Microfluidic assembly of homogeneous and janus colloid–filled hydrogel granules *Langmuir* 22:8618–8622, 2006.

[11] Malevanets, A., Kapral, R. Mesoscopic model for solvent dynamics. *J. Chem. Phys.*, 110:8605–8613, 1999.

[12] Lodge, J. F. M., and Heyes, D. M. Brownian dynamics simulations of Lennard-Jones gas/liquid phase separation and its relevance to gel formation. *J. Chem. Soc. Faraday Trans.*, 93:437–448, 1997.

[13] Hobbie, E. K. Depletion-driven phase separation and reversible aggregation in confined colloidal mixtures. *Langmuir*, 15:8807–8812, 1999.

[14] Mokhtari, T., Chakrabarti, A., Sorensen, C. M., Cheng, C., and Vigil, D. The effect of shear on colloidal aggregation and gelation studied using small-angle light scattering. *J. Coll. Interf. Science*, 327:216–223, 2008.

[15] Ulberg, D. E., Churaev, N. V., Ilyin, V. V., and Malashenko, G. L. Molecular-dynamics simulation of the aggregation of colloidal particles. *Colloids & Surfaces A Physicoch. & Engin. Aspects*, 80:93–102, 1993.

[16] Witten, T. A., and Sander, L. M. Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.*, 47:1400–1403, 1981.

[17] Hoogerbrugge, P. J., and Koelman, J. M. V. A. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.*, 19:155–160, 1992.

[18] Ermak, D. L. Computer-simulation of charged-particles in solution .1. Technique and equilibrium properties. *J. Chem. Phys.*, 62:4189–4196, 1975.

[19] Meakin, P., and Vicsek, T. Internal structure of diffusion-limited aggregates. *Phys. Rev. A*, 32(1):685–688,1985

[20] Meakin, P. Formation of fractal clusters and networks by irreversible diffusion-limited aggregation. *Phys. Rev. Lett.*, 51:1119–1122, 1983.

[21] Kolb, M., Botet, R., and Jullien, R. Scaling of kinetically growing clusters. *Phys. Rev. Lett.*, 51:1123–1126, 1983.

[22] Meakin, P. A historical introduction to computer models for fractal aggregates. *J. Sol-Gel Sci. & Tech.*, 15:97–117, 1999.

[23] Vold, M. J. Computer simulation of floc formation in a colloidal suspension. *J. Colloid & Sci.*, 18:684–695, 1963.

[24] Sutherland, D. N. A theoretical model of floc structure. *J. Colloid & Inter. Sci.*, 25:373–380, 1967.

[25] Jullien, R., and Kolb, M. Hierarchical model for chemically limited cluster-cluster aggregates. *Journal of Physics*, A17:L639–L643, 1984.

[26] Salazar, R., and Gelb, L. D. Off-lattice Dynamic Monte Carlo simulations of aggregation in one dimension. *Physica A*, 356:190–195, 2005.

[27] Español, P., and Warren, P. B. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.*, 30:191–196, 1995.

[28] Lee, S. H., and Kapral, R. Cluster structure and dynamics in a mesoscopic solvent. *Physica A*, 298:56–68, 2001.

[29] Sullivan, F., and Mountain, R. D. An efficient algorithm for the Brownian dynamics simulation of aggregation. *Comp. Phys. Comm.*, 42(1):43–49, 1986.

[30] Stoyan, D., and Stoyan, H. *Fractals, random shapes, and point fields : methods of geometrical statistics*; John Wiley&Sons, 1994.

[31] Klimontovich, Yu. L. *Statistical physics*; Harwood Academic Publishers GmbH, Switzerland, 1986.

[32] Nicholson, D. R. *Introduction to plasma theory*; Krieger, Malabar, FL, 1992.

[33] Krall, N. A., and Trivelpiece, A. W. *Principles of plasma physics*; McGraw-Hill, New York, 1973.

[34] Stoyan, D., Kendall, W. S., and Mecke, J. *Stochastic geometry and its applications*; John Wiley and Sons, Chichester, New York, Brisbane, Toronto and Singapore, 1987.

[35] Oh, C., and Sorensen C. M. The effect of overlap between monomers on the determination of fractal cluster morphology. *J. Colloid & Interf. Sci.*, 193:17–25, 1997.

[36] Fry, D., Chakrabarti, A., Kim, W., and Sorensen C. M. Structural crossover in dense irreversibly aggregating particulate systems. *Phys. Rev. E*, 69:061401, 2004.

[37] Lattuada, M., Wu, H., Sandkuhler, P., Sefcik, J., and Morbidelli, M. Modelling of aggregation kinetics of colloidal systems and its validation by light scattering measurements. *Chem. Engin. Science*, 59:1783–1798, 2004.

[38] Fox, R. O. (private communication), 2006.

[39] Oh, C., and Sorensen C. M. Structure factor of diffusion-limited aggregation clusters: Local structure and non-self-similarity. *Phys. Rev. E*, 57(1):784–790, 1998.

[40] Qin, Y., and Fichthorn K. A. Molecular-dynamics simulation of forces between nanoparticles in a Lennard-Jones liquid. *J. Chem. Phys.*, 119 (18):9745–9784, 2003.

[41] Cooley, M. D. A., and O'Neill, M. E. On the slow motion generated in a viscous fluid by the approach of a sphere to a plane wall or stationary sphere. *Mathematika* 16:37–49 1969.

[42] Hansford, R. E. On converging solid spheres in a highly viscous fluid. *Mathematika* 17: 250–254, 1970.

[43] Kim, S., and Karrila, S. *Microhydrodynamics, Principles and Selected Applications*; Dover Pub., Inc., 2005.

[44] Koch, D. L. Kinetic-theory for a monodisperse gas-solid suspension. *Phys. Fluids A-Fluid Dynam.* 2 (10):1711–1723, 1990.

[45] Sundararajakumar, R. R., and Koch, D. L. Non-continuum lubrication flows between particles colliding in a gas. *J. Fluid Mech.* 313:283–308, 1996.

[46] Jenkins, J. T., and Koenders, M. A. Hydrodynamic interaction of rough spheres. *Granular Matter* 7 (1):13–18, 2005.

[47] Qin, Y., and Fichthorn, K. A. Solvophobic solvation at large and intermediate length scales: Size, shape, and solvent effects. *Phys. Rev. E*, 74:020401(R), 2006.

[48] Fichthorn, K. A., and Qin, Y. Molecular dynamics simulation of the forces between colloidal nanoparticles in Lennard-Jones and n-decane solvent. *Granular Matter* 10:105–111, 2008.

[49] Heyes, D. M. Brownian dynamics simulations of self and collective diffusion of near hard sphere colloidal liquids: inclusion of many-body hydrodynamics. *Molecular Physics* 87 (2):287–297, 1996.

[50] Charkabarti, A., Fry, D., and Sorensen, C. M. Molecular dynamics simulation of the transition from dispersed to solid phase. *Phys. Rev. E*, 69:031408, 2004.

[51] Qin, Y., and Fichthorn K. A. Molecular-dynamics simulation of the forces between colloidal nanoparticles in n-decane solvent. *J. Chem. Phys.*, 127 (18):144911, 2007.