Structural Analysis and Optimisation of Jammed Granular Particles

Esma Kurban

A thesis submitted to the University of London for the degree of
Doctor of Philosophy

School of Mathematical Sciences
Queen Mary University of London
United Kingdom

Queen Mary
University of London

January 2023

Declaration

I, Esma Kurban, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third parties' copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university. The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: Esma Kurban   Date: January 2023

Abstract

Granular matter consisting of maximally dense random particle packings has been studied in science and industry. However, understanding the relationship between particle shape and random packing density has been challenging, so the question of "what shape packs with the highest density?" remains unanswered. While experiments and theory are restricted to investigating a few common shapes, numerical simulations allow studying the properties of packings of a diverse range of shapes. This thesis utilises simulations and machine learning methods to investigate particle geometry's role in packing behaviour by exploring a high-dimensional shape space.

The first two parts of the thesis generate disordered packings of monodisperse and binary mixtures of frictionless dimers in three dimensions by a gravitational pouring protocol in LAMMPS and analyse their structural properties by packing density, contact statistics, and several order metrics. The results show that monodisperse dimers exhibit a non-monotonic behaviour in the packing density with the aspect ratio α and undergo significant structural rearrangements up to the characteristic peak at $\alpha_{\max} \approx 1.4 - 1.5$. This unique density maximum also exists in the binary case, irrespective of the variation in shape or mixture composition, accompanied by similar microscopic rearrangements. The outcomings indicate that the packing density of binary mixtures is independent of the segregation state by holding an ideal mixing law.

The final section of the thesis applies a Random forest regressor to a dataset of the packing density represented as a function of particle shape obtained by overlapping spheres to predict novel dense packing shapes. The regression model is applied to the data represented in the lower-dimensional space by Principal component analysis (PCA) and Kernel PCA separately. The findings illustrate that the regressor predicts dense packing shapes, and their novelty depends on the dimensionality reduction method. While it optimises the trimer that already exists in the dataset with the highest density for PCA, Kernel one predicts a distinct shape.

Acknowledgement

Thanks to QMUL for giving me a nice place to do my research. I want to start with what PhD has truly meant for me. I think it is the most fascinating yet, at the same time, tough journey one can experience in life. I have enjoyed a lot every moment of the process of creating new ideas and doing actual research, yet suffered from the anxieties it caused. During my PhD, I have developed both personally and professionally. There are many people whom I should thank a lot for guiding, supporting and inspiring me during this journey.

There is no doubt that my supervisor Adrian needs to be thanked first of all and the most. His knowledge and guidance have greatly inspired me and made it possible to do the research you will read in this thesis. So Adrian, thanks a lot for your time with our discussions and your patience with my questions. Thanks, particularly for giving me the opportunity and guidance to develop my ideas to be an independent researcher. I also thank Prof. Hernan Makse and Dr. Kuang Liu for very helpful discussions, and Dr. Leah K. Roth and Prof. Heinrich M. Jaeger for sharing their dataset, which I used in Chapter 5 of this thesis.

I want to thank also my talented colleagues in the School of Mathematics, with whom I had inspiring scientific discussions and shared our cultures during lunch or coffee breaks. It is impossible not to mention our happy days in France House; many thanks for being a part of those memories: Marco, Jin, Ye, Tiago, Asma, Francesco, Gianmichele, Marica and all other PhD friends. Marco, thanks particularly for

all your support. Jin and Ye, many thanks for being great flatmates, our lovely Europe trips and our Chinese hotpot day. Thanks, Tiago and Enrico, for sharing your passion for research, which greatly motivated me.

I have met great friends during my time in London. Thanks, Naz, Safa, Rana, and Katherina, for accompanying me with museum visits, art events and enjoying this city.

I can not thank my flatmates Su and Halime enough for their great support and for being my family in this city. I can imagine how difficult the pandemic days would be without you. Thanks, Halime, for encouraging me to be less anxious in life. A special thanks go to Su, whom I admired greatly for being such a brilliant researcher and a great friend. Even though our research disciplines are different, philosophical discussions with her and même love of science greatly inspired me. Many thanks for being a big part of my personal development during the PhD. Thanks, my dear friends Asiye, Begüm, and Seda, for always being there for me.

I owe almost everything I am to my family. Their endless support and love helped me to stay strong despite all ups and downs of the PhD. I especially want to thank my father, who encourages me the most to contribute to science.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

What are Granular Materials?

Granular materials are composed of solid, macroscopic particles that are large enough such that their thermal motion fluctuations are negligible. They are ubiquitous in everyday experience, for example, sand, coffee, rice and bearing balls. They can be observed in various sizes, from sand grains to icebergs or asteroids. The fact that they have characteristics reminiscent of solids, liquids, or gases makes them unique in industrial applications, i.e. they are easily manipulated and processed to create products such as in the pharmaceutical industry, and agriculture [1]. They also play an important role in geological processes, e.g. landslides and erosion, so it is essential to understand the properties of granular matter, which has been a challenge to statistical physics as an extreme example of a system far from equilibrium.

Jamming transition

Granular materials exhibit jamming phase transition, which enables them to change from fluid-like plasticity to solid-like rigidity [2]. Pouring sugar into a cup is a simple example of a fluid-to-solid transition caused by a density increase. In order to explore different configurations of granular materials, external driving forces are needed since thermal fluctuations are negligible, and there is no kinetic energy due to the dissipative nature of the interactions between grains. For example, tapping sugar particles in the cup decreases the free volume per particle, increasing the packing density ϕ, the fraction of total space covered by the particles, until reaching a critical state (jammed) with ϕ_{j}. In granular matter, a jammed state is achieved when the particles are arrested in a static configuration where they only touch their neighbours and withstand a sufficiently small applied stress [3]. This phase transition, in an athermal ensemble first proposed by S.F. Edwards and co-workers, analogously to the phase transition in thermal hard spheres [4]. Unlike thermal systems, jamming is governed by two parameters: the available free volume per particle and the applied shear stress. If the former is large enough to allow the particles to move or the latter is increased, the system yields and dilates where jamming and rigidity are lost. This thesis studies the structural properties of jammed granular particles that are mechanically stable, i.e. there must be force and torque balance on every particle in the system. Several parameters, such as particle shape, size distribution, friction, and elasticity, determine the macroscopic quantities like packing fraction, pressure and contact number. Besides

granular matter, jammed particle packings have also been investigated to study soft materials such as colloidal suspensions, compressed emulsions, cells, DNA, and protein packing [5].

Packing Problem

Jammed particle packings have a long history dating back to finding the densest packing of equal-sized objects conjectured by Kepler [6, 7]. Kepler conjectured that the face-centred-cubic (FCC) crystal arrangement of spheres in three dimensions gives the highest packing density, which is $\phi_{\text{fcc}} = \pi/(3\sqrt{2})) \approx 0.7405$ [6]. This arrangement was proven as the densest Bravais lattice packing by Gauss [8]. However, a full mathematical proof of Kepler's conjecture has remained an unsolved problem for almost four centuries until Hales announced his proof of Kepler's conjecture [9]. While investigating the structure of a liquid, Bernal suggested that it can be considered a heap of particles, hence started experimental studies of disordered hard-sphere packings in the 1960s [10, 11]. The packings were obtained by pouring equally sized spherical particles into a large container, vertically vibrating the system until achieving maximum densification. They extrapolated the measured packing fractions to eliminate finite-size effects due to boundary conditions. These experiments yielded a range of densities for steel balls with a maximum of $\phi_{\text{rcp}} \approx 0.64$, where the packing of ϕ_{rcp} is referred to as the "random close packing (RCP)" in the literature [12, 13]. This density coincides with the jamming phase transition point of the sphere packings, so-called point J, at which crystalline structures start to form [14]—further increasing the density results in ordered

packings. Many studies have observed several power-law scalings to point J, such as for packing fraction, contact number, and pressure; see [5] for more details. The determination of the order in sphere packings has also been well-studied. Different parameters have been proposed to measure the order, such as the bond-orientational order parameter [15] or Minkowski tensors [16, 17]. The densities for disordered systems are generally found as strongly dependent on the packing protocol and friction, falling in the range $\phi_j \approx 0.55 - 0.64$ [5]. Disordered packings of frictional hard spheres were first investigated by Bernal and Scott in [10, 13], where the lower bound ($\phi_j \approx 0.55$) of the achieved packing densities is referred to as the "random loose packing (RLP)". Besides the packing density, empirical studies were also interested in the contact number z to investigate the mechanical stability of the packing. Determination of the contact number is, in general, challenging for experimental work. Using X-ray tomography to resolve coordinates of a high number of spheres, Aste *et al.* observed that the contact number exhibits a monotonic increase from 4 to 7 over the range $\phi_j \approx 0.55 - 0.64$ [18–20]. Several numerical algorithms have been developed to simulate disordered sphere packings and study their properties. One of the geometrical algorithms, the rate-dependent densification algorithm, starts with generating points randomly in a cubic lattice with periodic boundary conditions, where each point is the centre of an inner and an outer sphere, and the inner one is the actual sphere. The algorithm iteratively eliminates overlappings by slowly reducing the outer diameter until the inner and outer diameters coincide [21]. In the Lubachevsky-Stillinger algorithm, the spheres are grown from randomly

generated points at a fixed rate, and the elastic contacts occur during the evolution of the spheres [22, 23]. The dynamic algorithms imitate experiments by simulating spheres exposed to forces and underlying contacts. For example, Silbert *et al.* generated packings by pouring spheres under gravity into a rectangular box with periodic boundary conditions, where molecular dynamics (MD) was used for the simulations [24]. The numerical simulations are useful for investigating the packing properties; however, the final states of the packings or the resulting densities have been observed to be strongly protocol-dependent. For example, the rate-dependent densification algorithm achieved $\phi_j = 0.642 - 0.649$, a "drop and roll" procedure found $\phi_j \approx 0.6$, see [25] for more details. By pointing out this result, Torquato *et al.* argued that the concept of "RCP" is not well defined since one should sacrifice randomness to increase the packing density [26]. Hence, they replaced "RCP" with a concept of "maximally random jammed (MRJ) packing", based on minimising a bond-orientational order metric among all jammed states. In terms of theoretical work of jamming, there have been successful studies, for example, based on replica theory [27, 28], mean-field theory [29, 30]. Overall, there has been much progress in understanding the characteristics of jammed sphere packings [5].

Packings of Non-Spherical Particles

Beyond spherical particles, there has been increased attention to investigating disordered packings of non-spherical particles that, in fact, better represent materials in nature and are used in industrial applications [2, 46]. In particular, it appeals to scientists and engineers to understand the effect of the particle shape on the packing

Table 1-A: Overview of maximal packing fractions $\phi_{\max}$ obtained for random packings of a selection of regular shapes. Note that the packings are generated with different protocols. Tetrahedron attains a maximum packing fraction in disordered arrangement, which exceeds that for spheres in the FCC crystal ($\phi_{\text{fcc}} = 0.7405$). General ellipsoids and lens-shaped particles reach maximal packing fractions close to this value.

Particle shape	$\phi_{\max}$ simulation	$\phi_{\max}$ experiment	$\phi_{\max}$ theory
Sphere	0.645 [31]	0.64 [10]	0.634 [29], 0.68 [27]
M&M candy		0.665 [32]	
Dimer	0.703 [33]		0.707 [34]
Oblate ellipsoid	0.709 [32, 35]		
Prolate ellipsoid	0.716 [32]		
Spherocylinder	0.725 [36, 37]		0.731 [34]
Lens-shaped particle			0.736 [34]
Tetrahedron	0.7858 [38]	0.76 [39]	
Cube		0.67 [40]	
Octahedron	0.697 [41]	0.64 [40]	
Dodecahedron	0.716 [41]	0.63 [40]	
Icosahedron	0.707 [41]	0.59 [40]	
General ellipsoid	0.735 [32, 42]	0.74 [43]	
Superball	0.674 [44]		
Trimer	0.729 [45]		

density, thereby identifying the shape that randomly packs with the maximum density to design new materials. Additional degrees of freedom that arise due to asphericities hinder the development of a general theory for the packing problem. Therefore, empirical investigations for the dense disordered packings of non-spherical shapes have been performed on a case-by-case basis, mostly focusing on shapes like ellipsoids [32, 35, 42, 43, 47–49], spherocylinders [36, 37, 50–57], dimers [33, 58], and polyhedral or polygonal shapes [39, 40, 59–67]. Recently, particles constructed by assembling or bending a specific particle shape, such as trimers [45, 68] (composed of three overlapping spheres) and curved spherocylinders [69–71] have attracted much interest. The results have shown that non-spherical shapes can generally achieve

denser maximal packing densities than spheres; see Table 1-A. Ulam's conjecture in the context of ordered packings, recently also formulated for disordered packings [41], states that the sphere indeed packs the worst among all convex shapes [72]. The densest disordered packing for tetrahedra with $\phi_j = 0.7858$ has been found in [38]. General ellipsoids and lens-shaped particles can reach maximal packing fractions of $\phi_j \approx 0.74$. Rotationally symmetric elongated shapes such as ellipsoids, spherocylinders, and dimers exhibit a non-monotonic behaviour in the packing density with the aspect ratio α by possessing a peak at $\alpha \approx 1.4 - 1.5$. Beyond the peak, the packing density decreases as the aspect ratio increases. The strong dependence on packing protocol in the empirical work causes significant variations in the resulting peak locations and magnitudes even for the same shape. For example, the results reported for spherocylinders show a large variance [36, 50–54, 56]. Baule *et al.* have made considerable progress in analytically estimating packing densities of axisymmetric particles such as dimers, lens-shaped particles, and spherocylinders by using the mean-field Edwards approach, where all microscopic jammed states at a fixed volume of granular systems are assumed as equiprobable and macroscopic quantities can be calculated by taking averages over them [34]. The predicted packing densities are an upper bound of the empirical findings [34].

Previous work has reported that packings of convex elongated shapes possess different mechanical and vibrational properties than sphere packings near jamming transition [73]. For example, unlike frictionless hard-sphere packings, which are generally observed as isostatic, frictionless ellipsoids generate hypostatic packings

over a wide range of aspect ratios [32, 74–78]. Isostaticity means that the total number of contacts in the system is equal to the number of degrees of freedom, i.e., the packings are mechanically stable. Hypostatic packings, on the other hand, possess fewer contacts than required for mechanical stability. One would expect that those packings are not mechanically stable; however, previous work on ellipsoidal packings has shown that they actually do not possess zero-frequency modes in the jammed phase $\phi > \phi_{\mathrm{j}}$ [76–78]. Despite some floppy modes (they are not exactly zero-frequency modes) being induced, the repulsive interactions between ellipsoids can stabilise them. Note that, like ellipsoidal shape, spherocylinders also generate hypostatic packings [79, 80]. Compared to convex elongated shapes, structural properties of jammed packings of dimers have been less investigated. Unlike the former, dimer packings have been observed as isostatic [58, 73, 77]. Although the packing fraction ϕ_{j} is characterised similarly by the aspect ratio for both convex elongated shapes and dimers, some other macroscopic properties, such as the contact number, strongly depend on the microscale geometric features of the individual particles [77]. In Chapter 3 of this thesis, I explore the structural properties of disordered packings of monodisperse frictionless symmetric dimers in three dimensions generated by simulations in LAMMPS, a molecular dynamics simulation platform [81]. For the simulations, I use soft particles and a pouring packing protocol, i.e. exposing the particles to a gravitational force which causes them to pack into jammed disordered phases. In granular material simulations, a soft particle refers to a grain that can be deformed or overlapped by some

amount, usually much less than the particle radii. Focusing on several macroscopic parameters, such as packing fraction, contact number, and order metrics, reveal significant structural features of the dimer packings on both microscopic and macroscopic levels.

Most studies of non-spherical particle packings mainly focused on the monophasic case. However, particle aggregates in nature generally have polydispersity in both shape and size. Previous studies of binary and polydisperse packings of spherical particles show that the packing density increases due to varying the size distribution of the particles [82–92]. Non-spherical particle mixtures, on the other hand, have been less investigated and restricted to a few common shapes. A universal density maximum at a unique aspect ratio has been observed in jammed packings of binary spherocylinder mixtures, irrespectively of the mixture composition [53, 55, 56, 93–96]. Moreover, the mixtures satisfy a remarkable empirical ideal mixing law, which states that the total packing volume is independent of the segregation state, i.e. the two subsystems of the single species are in some sense decoupled [56, 93, 97, 98]. In Chapter 4 of this thesis, I investigate whether these findings are also valid in disordered packings of binary dimer mixtures. I also show the microscopic structural properties that accompany these observations.

The shape is an infinitely variable parameter that makes a rigorous systematic exploration of the densest random packings infeasible. There has been recently increased interest in making this search possible by using advanced developments

in computational methods. For example, Jaeger and collaborators used artificial evolutionary algorithms to map the possible space of the constitutive particle shapes obtained by glueing or overlapping spherical particles of varying radii [2, 45, 68, 99]. The evolutionary algorithm starts with a given shape and then mutates the composite particles until achieving maximal packing density. They have found maximal packing densities of $\phi_{\max} \approx 0.73$ for trimer-shaped particles. Machine learning algorithms can be used to explore the high-dimensional shape space in the packing problem. In Chapter 5 of this thesis, I examine the feasibility of using machine learning approaches to predict novel dense packing shapes. I present a framework that applies a Random forest regressor to a dataset from Ref. [45] of the packing density ϕ_{j} as a function of particle shape, 5-sphere, obtained by overlapping five spheres to identify new optimal shapes to maximise ϕ_{j}.

1.1 Outline of the thesis

Chapter 2 provides information about the simulation platform, LAMMPS and the gravitational pouring protocol used to generate disordered packings of monodisperse symmetric dimers, binary mixtures of dimers, and monodisperse 5-sphere particles. The simulation parameters for all packings are given.

Chapter 3 is devoted to investigating the structural properties of the disordered monodisperse dimer packings in three dimensions. I first obtain packing density, Voronoi volume statistics, and contact and coordination numbers in the analysis.

Then, I introduce five types of contact configuration and study how their fractions vary with the aspect ratio. I also calculate orientational order parameter, bond-orientational order parameters, and radial distribution function. Finally, I determine the polar and azimuthal angles of the bond vectors between neighbouring particles. The results show dimers exhibit a non-monotonic behaviour in the packing density with the aspect ratio α and undergo significant structural rearrangements as α increases up to α_{max}. The metrics remain largely unchanged for larger aspect ratios.

Chapter 4 presents the packing properties of dimer-sphere and dimer-dimer mixtures. I first measure the packing density as a function of both shape and mixture composition. I then determine both partial and total contact and coordination numbers. Finally, I calculate fractions of distinct contact configurations for various mixture compositions. The findings show that a unique density maximum exists in the binary mixtures, irrespectively of the variation in shape or mixture composition of the two species. This density maximum is accompanied by similar microscopic rearrangements of contacts as in monodisperse dimer packings. Moreover, the binary mixtures hold an ideal mixing law, i.e., the packing density is independent of the segregation state.

Chapter 5 is devoted to applying machine learning algorithms to a dataset from Ref. [45] of the packing density as a function of particle shape obtained by overlapping five spheres to predict novel dense packing shapes. I first introduce the dataset and reduce the high input dimension by Principal component analysis

(PCA) and Kernel PCA, respectively. Then, I apply a Random forest regressor to the reduced data by the two methods separately to predict dense packing shapes. Finally, I generate LAMMPS simulations for disordered packings of those shapes and compare the simulation values for packing density and the predicted ones. The findings illustrate that the regressor predicts dense packing shapes, and their novelty depends on the dimensionality reduction method.

Chapter 6 presents the conclusions and a brief discussion of the directions for possible future work.

Chapter 2

Simulation Method

In this thesis, I generate simulations for disordered packings of monodisperse symmetric dimers, binary mixtures of dimer particles, and monodisperse 5-sphere particles in three dimensions for the following three sections, respectively. Both dimers and 5-sphere particles are frictionless and obtained by overlapping spheres. For dimers, the constituent spheres have the same diameter d and mass m, and they are defined by the aspect ratio α, which is the length ratio over the width, see Fig. 2.1. One obtains a 5-sphere particle by placing the constituent spheres in sequence [45]. The spheres are allowed to overlap and also vary in diameter. The 5-sphere particle shape is defined by fixing the diameter and coordinate of the first sphere and by the relative coordinates and diameters of the other four spheres, see Fig. 2.2. Representing a particle shape by overlapping spheres is an efficient way to detect contacts between particles in simulations. Moreover, this model can be
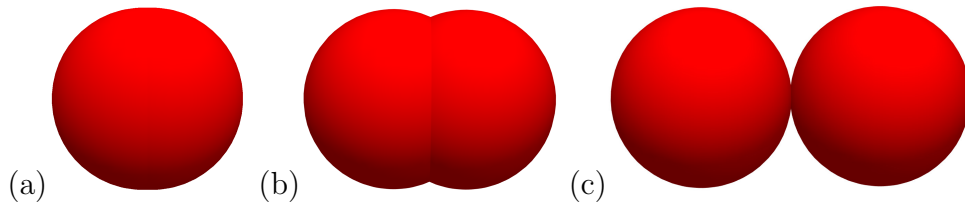
Figure 2.1: Dimer shape defined by the aspect ratio: (a) $\alpha = 1.05$, (b) $\alpha = 1.4$, (c) $\alpha = 2$.
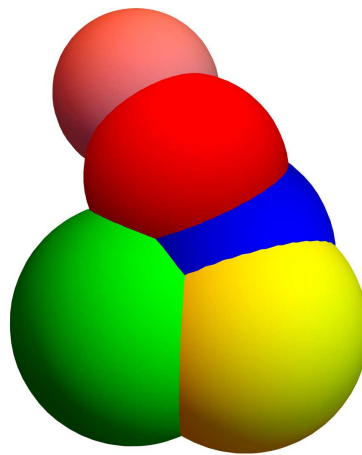
Figure 2.2: 5-sphere particle shape obtained by overlapping five spheres of varying radii. The shape is defined by fixing the diameter and position of the reference sphere (coloured in green) and by the relative coordinates and diameters of the other four spheres under overlapping constraints [45].

easily implemented in molecular dynamics (MD) simulations of jammed particle packings. In MD simulations, the atoms or molecules are allowed to interact for a fixed period of time, and their trajectories are determined by numerically solving Newton's equations of motion [100].

Numerical simulations for disordered packings of non-spherical particles have employed different algorithms. For example, the Lubachevsky-Stillinger algorithm has been applied to ellipsoids in [32], spherocylinder packings have been obtained

with a mechanical contraction algorithm in [50, 54, 101], with Monte Carlo method in [51], and with a digital packing algorithm in [52]. Faure *et al.* presented a dynamic algorithm to simulate the sedimentation of dimers in a parallelepipedic container [33]. Other studies of dimer packings have used a compression/decompression method followed by energy minimisation [58, 73, 77]. Here, I use a gravitational pouring protocol in the molecular dynamics platform LAMMPS [81, 100] to generate disordered packings. In this protocol, N particles are put at random positions above the packing at a certain height and let settle under gravity into a three-dimensional box. The lateral ($\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$-plane) boundary conditions are chosen to be periodic, and the box of side length $\approx 20d$ is bounded in the $\hat{\mathbf{z}}$-direction by a rough surface at the bottom (implemented by the "fix wall/gran hertz/history" command) and an open top. The rough surface is constructed by taking a slice from a random packing of particles of the same diameter d as those in the bulk. During a simulation run, a gravitational force acts on the particles in the $\hat{\mathbf{z}}$-direction. The pouring protocol makes use of LAMMPS' "fix pour" command, which repeatedly inserts particles into the simulation box every few timesteps within a specified insertion region $30 - 40d$ above the bottom and releases them until N particles have been added overall. In the insertion region, particles are added with random positions and orientations and without any overlap. Particles are only inserted again after the previously inserted particles have fallen out of the insertion region under the gravitational force. This simulation model has been used to generate disordered packings of monodisperse spheres in [24], where interactions between particles are
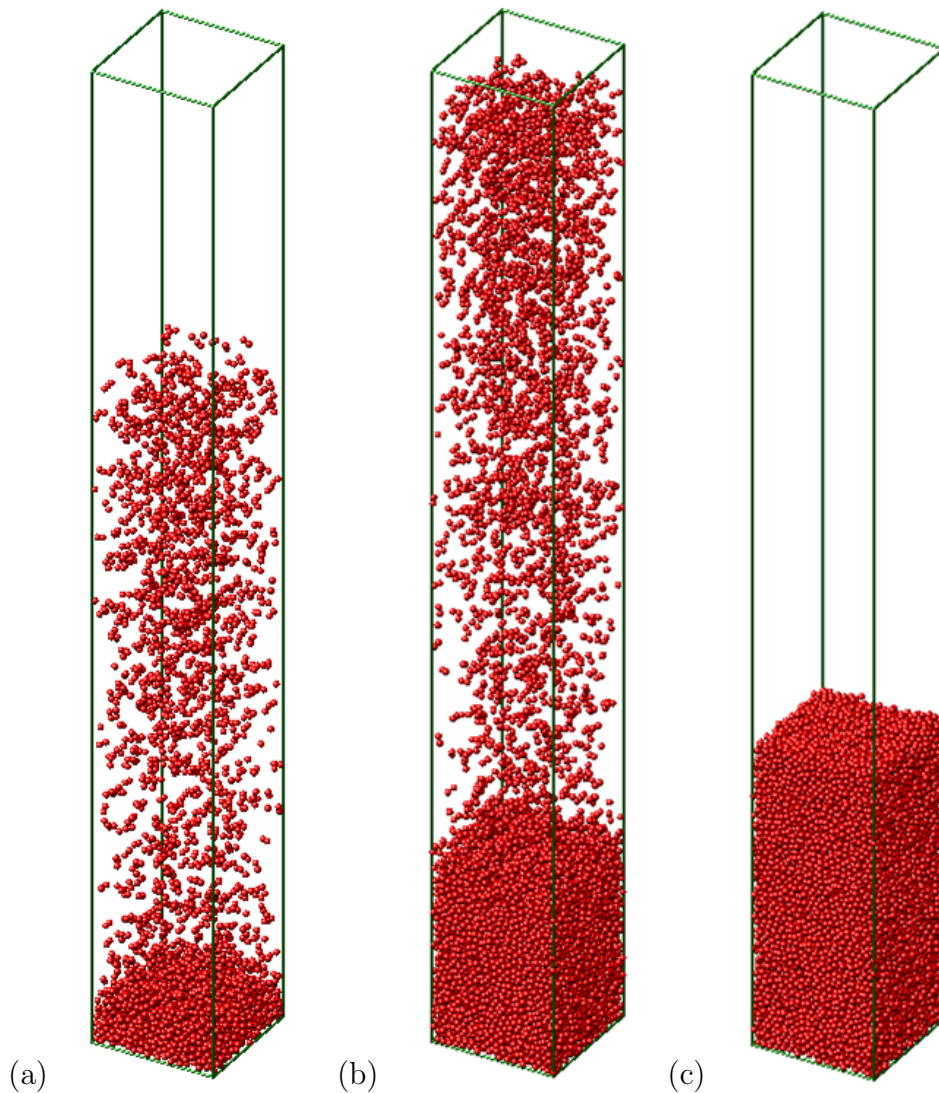
Figure 2.3: Three configurations for LAMMPS simulations of disordered packings of 12,000 monodisperse dimers of $\alpha = 2$ (final configuration on the right).

modelled as spring-dashpots (with Hookean and Hertzian models). Roth *et al.* also used this protocol with the Hertzian contact model in LAMMPS to simulate disordered packings of monodisperse 5-sphere particles [45].

Here, since I model dimers and 5-sphere particles as overlapping spheres, the interaction between two composite particles can be determined from the pairwise

interaction of spheres, which is assumed to follow the spring-dashpot model in LAMMPS as in the studies of sphere packings of [24, 102]. LAMMPS can implement two models for calculating the contact forces between the spheres; Hookean and Hertzian. A Hookean model is suitable for symmetric dimers since it does not take into account relative particle sizes. This model is also convenient for dissipating residual kinetic energy and hence reaching a static state quickly [24]. In the Hookean model, when two spheres i and j having positions $\mathbf{r}_i$ and $\mathbf{r}_j$, respectively, are in contact, they experience a relative normal compression with overlap $\delta = d - r_{ij}$, where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and $r_{ij} = |\mathbf{r}_{ij}|$. The resulting force on particle i is $\mathbf{F}_{ij} = \mathbf{F}_{ij}^{\mathrm{n}} + \mathbf{F}_{ij}^{\mathrm{t}}$, where $\mathbf{F}_{ij}^{\mathrm{n,t}}$ are the normal and tangential contact forces, respectively, given as [102]:

$$\mathbf{F}_{ij}^{\mathrm{n}} = K_{\mathrm{n}}\delta\,\mathbf{n}_{ij} - \frac{m}{2}\gamma_{\mathrm{n}}\mathbf{v}_{\mathrm{n}} \qquad \mathbf{F}_{ij}^{\mathrm{t}} = -K_{\mathrm{t}}\Delta\mathbf{s}_t - \frac{m}{2}\gamma_{\mathrm{t}}\mathbf{v}_{\mathrm{t}}. \qquad (2.1)$$

Here, $\mathbf{n}_{ij} = \mathbf{r}_{ij}/r_{ij}$, $\mathbf{v}_{\mathrm{n,t}}$ are the normal and the tangential components of the relative velocity of the spheres i and j, and $K_{\mathrm{n,t}}$ and $\gamma_{\mathrm{n,t}}$ are the elastic and viscoelastic constants, respectively. The quantity $\Delta\mathbf{s}_{\mathrm{t}}$ denotes the elastic tangential displacement between the spheres [102]. The corresponding contact force on particle j is then given by Newton's third law, i.e. $\mathbf{F}_{ji} = -\mathbf{F}_{ij}$. The total force $\mathbf{F}_i^{\mathrm{tot}}$ and torque τ_i^{tot} on sphere i in a gravitational field $\mathbf{g} = -g\,\hat{\mathbf{z}}$ are then given as:

$$\mathbf{F}_i^{\mathrm{tot}} = m\,\mathbf{g} + \sum_{i\neq j}\mathbf{F}_{ij}^{\mathrm{n}} + \sum_{i\neq j}\mathbf{F}_{ij}^{\mathrm{t}}, \qquad (2.2)$$

$$\tau_i^{\text{tot}} = -\frac{1}{2} \sum_{i \neq j} \mathbf{r}_{ij} \times \mathbf{F}_{ij}^{\text{t}}, \tag{2.3}$$

where the sum runs over all j spheres in contact with sphere i. LAMMPS treats a composite particle (e.g. dimer or 5-sphere particle) defined by fixed distances between its constituent spheres as an independent rigid body (implemented by the "fix rigid/nve/small" command). The total force and torque on each particle rigid body are computed as the sum of the forces and torques on its constituent spheres in every time step. The coordinates, velocities, and orientations of the constituent spheres are then updated so that the particle moves and rotates as a single entity.

Throughout the investigation the basic units are set as $d = 1$, $m = \pi/6$ (the density of the particle is set to 1, so the mass is equal to the particle volume), and $g = 1$. Distances, times, velocities, forces and elastic constants are then measured in units of d, $\sqrt{d/g}$, $\sqrt{gd}$, mg, mg/d, respectively. For both monodisperse dimer packings and binary mixtures, I use $K_{\text{n}} = 2 \times 10^5 mg/d$ and set $\gamma_{\text{t}} = 0$. In the monodisperse case, I also simulate harder dimers with a normal spring constant $K_{\text{n}} = 2 \times 10^6 mg/d$ and softer ones with $K_{\text{n}} = 2 \times 10^4 mg/d$ to examine the effect of particle hardness on the contact number of the dimers at small aspect ratios. The total number of particles, the aspect ratio α range, material parameter values, and time step used in the simulations of monodisperse dimer packings and binary mixtures are given in Tables 2-A and 2-B, respectively. The choice of most of the material parameter values follows the discussion in [24].

Simulations are run until the system reaches a static equilibrium when the

Table 2-A: The number of particles, the aspect ratio α range, material parameter values, and time step Δt used in the simulations of monodisperse dimer packings. $\alpha = 1$ corresponds to sphere packings.

N	α	K_{n} (mg/d)	$K_{\mathrm{t}}/K_{\mathrm{n}}$	γ_{n} $\sqrt{g/d}$	Δt ($\sqrt{d/g}$)
15,000	1, 1.0005-1.1	2×10^4	2/7	15	0.003
12,000-15,000	1, 1.0005-2	2×10^5	2/7	50	0.001
15,000	1, 1.0005-1.1	2×10^6	2/7	150	0.0003

Table 2-B: The number of particles N, the aspect ratio α range, material parameter values, and time step Δt used in the simulations of dimer-sphere and dimer-dimer mixtures. $\alpha = 1$ corresponds to sphere packings.

mixtures	N	α	K_{n} (mg/d)	$K_{\mathrm{t}}/K_{\mathrm{n}}$	γ_{n} $\sqrt{g/d}$	Δt ($\sqrt{d/g}$)
dimer-sphere	12,000-18,000	1, 1.05-2	2×10^5	2/7	50	0.001
dimer-dimer	12,000	1.05-2	2×10^5	2/7	50	0.001

kinetic energy per particle is less than $10^{-8} mgd$ for small K_{n} and up to three orders of magnitude less for large K_{n}. For example, when $K_{\mathrm{n}} = 2 \times 10^5 mg/d$, the simulation takes $3 - 8 \times 10^6 \Delta t$ to reach equilibrium, which depends on the chosen aspect ratio and also on the random initial configurations when particles are poured into the container. I show the configurations at three different time steps for the simulations of $12,000$ monodisperse dimers of $\alpha = 2$ in Fig. 2.3.

The Hookean model is unsuitable for the 5-sphere particle packings, which require determining the interactions between two spheres of unequal radii (a_i, a_j). A Hertzian model, on the other hand, takes into account the relative size effect with radii-dependent prefactors. For the Hertzian model, the overlapping distance is found as $\delta = a_i + a_j - r_{ij}$, and the normal and tangential forces in Eq. 2.1 are

Table 2-C: The number of particles N, material parameter values, and time step Δt used in the simulations of 5-sphere particle packings.

N	K_{n} (mg/d^2)	$K_{\mathrm{t}}/K_{\mathrm{n}}$	γ_{n} $(\frac{\sqrt{g/d}}{d})$	Δt $(\sqrt{d/g})$
6,000	2×10^5	2/7	50	0.001

then replaced by

$$\mathbf{F}_{ij}^{\mathrm{n}} = \sqrt{\delta}\sqrt{\frac{a_i a_j}{a_i + a_j}}\left(K_{\mathrm{n}}\delta\,\mathbf{n}_{ij} - \frac{m_{\mathrm{eff}}}{2}\gamma_{\mathrm{n}}\mathbf{v}_{\mathrm{n}}\right) \qquad \mathbf{F}_{ij}^{\mathrm{t}} = \sqrt{\delta}\sqrt{\frac{a_i a_j}{a_i + a_j}}\left(-K_{\mathrm{t}}\Delta\mathbf{s}_t - \frac{m_{\mathrm{eff}}}{2}\gamma_{\mathrm{t}}\mathbf{v}_{\mathrm{t}}\right),$$

$$(2.4)$$

where m_{eff} is the effective mass of the two spheres of mass m_i and m_j and calculated as $m_{\mathrm{eff}} = \frac{m_i m_j}{m_i + m_j}$. While the normal push-back force for two particles is a linear function of the overlapping distance in the Hookean model, it becomes proportional to the area of overlap of the two particles in the Hertzian one. Thus the elastic constants K_{n}, K_{t} have units of force per area (mg/d^2) and the viscoelastic ones γ_{n}, γ_{t} are in units of $\frac{\sqrt{g/d}}{d}$. For a 5-sphere particle, d and m refer to the diameter and the mass of the largest constituent sphere, respectively and are set as $d = 1$ and $m = \pi/6$. Note that the density of each constituent sphere is set to 1, so the mass is equal to the corresponding sphere volume. The number of particles, material parameter values, and time step for the simulations of 5-sphere particle packings are given in Table 2-C. These values are chosen to be in line with the packings in the dataset from Ref. [45], where the pouring protocol in LAMMPS has also been used to generate them.

All simulations are run on Queen Mary's Apocrita HPC computers [103].

Chapter 3

Structural Analysis of Disordered Dimer Packings

3.1 Introduction

Jammed disordered particle packings have been used as a model to understand the structures of liquid crystals, glasses, self-assembly of nanoparticles, biological systems, and granular media [25]. While there has been considerable recent progress in the understanding of jammed sphere packings [28], the effect of particle shape on the properties of jammed packings has been much less explored [46]. It is well-known that spheres achieve maximal random packing densities of $\phi_j \approx 0.64$ for a wide range of packing protocols. However, denser packings can also be achieved for specific protocols; see the discussion in [5]. Recent empirical studies have shown that many

non-spherical shapes can generally pack denser than spheres. For example, many polyhedra [38, 41, 61, 63, 67], ellipsoids [32, 43, 48], spherocylinders [36, 50–54, 56], and dimers [33, 58], as well as irregular shapes such as those composed of a number of overlapping spheres [45, 68] achieve packing densities $\phi_j \geq 0.7$, with the densest disordered packings so far found for tetrahedra at $\phi_j \approx 0.78$ [38]. Rotationally symmetric elongated shapes such as ellipsoids, spherocylinders, and dimers exhibit a similar dependence of the packing density on the aspect ratio α. The packing density increases as soon as the aspect ratio deviates from the sphere value ($\alpha = 1$) until reaching a peak at $\alpha \approx 1.4 - 1.5$, and beyond that, it decreases, following, e.g., an approximate scaling behaviour $\phi_j \sim 1/\alpha$ for spherocylinders [104]. However, the mechanisms that underlie this non-monotonic behaviour in the packing density remain unknown.

Although the packing fraction ϕ_j is characterised similarly by the aspect ratio for both convex and non-convex elongated shapes, the contact number variation with the aspect ratio strongly depends on the microscale geometric features of the individual particles [77]. Ellipsoids have been observed to generate hypostatic packings for small aspect ratios [32, 74–78], and such underconstrained packings have also been found for spherocylinders [79, 80]. On the other hand, dimer packings are isostatic for all aspect ratios [58, 73, 77]. Schreck *et al.* have found key differences between the mechanical and vibrational properties of the packings of ellipses and dimers, e.g. the isostatic dimer packings do not possess quartic vibrational modes as found for hypostatic ellipse packings [77]. Shiraishi *et al.*

have also observed similar results for two-dimensional dimer packings in [73] and for three-dimensional dimer packings in [58]. It is crucial to study the packing properties of dimers to investigate the effect of non-convexity on mechanical and structural properties of disordered packings since a dimer represents the simplest case for non-convex shapes and can be defined by a single parameter, the aspect ratio, which is the ratio of the length over the width.

This chapter presents a structural analysis of dimer packings simulated with the MD platform LAMMPS using a gravitational pouring protocol. My goal is to identify structural features that characterise the peak in the packing density by focusing on details of the contact statistics as well as positional and orientational order metrics. In this context, it is important to emphasise the role of the packing protocol. Torquato *et al.* have shown that the notion of "random close packings" is not well-defined due to the interplay between the packing density and the degree of order in sphere packings that arises by tuning the protocol parameters [26]. Jammed packings of non-spherical particles also exhibit strong protocol dependence even for the same shape. For example, the results reported for ϕ_j of spherocylinders exhibit a significant variance [36, 50–54, 56]. Here, I use a specific pouring protocol for all the packings and study the effect of shape variation on the packing properties.

Previous studies of disordered packings of elongated particles obtained inconsistent results regarding ordering effects, which might be due to different protocols and boundary conditions used. For example, simulations of prolate ellipsoids by pouring

into a container under gravity found considerable nematic order, whereby the ellipsoids' symmetry axes (the semi-major axes) tend to lie within the plane normal to the gravity direction [47–49]. Buchalter *et al.* have explained this ordering effect as a result of the particles' tendency to minimise the gravitational potential energy [47]. On the other hand, simulations that compress or inflate the non-spherical particles from an initial random state, such as the Lubachevsky-Stillinger algorithm (applied to ellipsoids [32, 74]) or a mechanical contraction algorithm (applied to spherocylinders [54, 57, 101]) do not find any significant order as is also observed with other geometric simulation methods [36, 37]. While 3D experiments of ellipsoids [43] and elongated colloids [105] did not observe any signatures of order, experiments of asymmetric dumbbells in 2D showed strong orientational correlations between neighbours due to mutual restrictions on positions [106]. To my knowledge, the order characteristics of dimers in 3D have not been investigated.

This chapter is structured as follows: In Sec. 3.2, I present the analysis results; the packing fraction, Voronoi volume statistics, contact and coordination numbers, contact configurations, and orientational/positional order metrics. Finally, in Sec. 3.3, I summarize the findings.

3.2   Structural Analysis

3.2.1   Packing Fraction

Packing fraction is one of the most common macroscopic variables studied for the properties of granular assemblies. I calculate the packing fraction of the dimer packings for various aspect ratios. The packing density is determined for the bulk region shown in Fig. 3.1. The particles within $5 - 8d$ from the container floor are excluded from the bulk region since they can be highly crystallised. The thickness of this crystallised region depends on many factors, such as the box dimension and the pouring height. Excluding the particles within $5 - 8d$ provides results that are largely unaffected by the crystallisation. The particles within $5d$ from the upper-most particles have also been excluded from the bulk because their Voronoi volumes can not be decided accurately due to deficiencies in their neighbourhood.
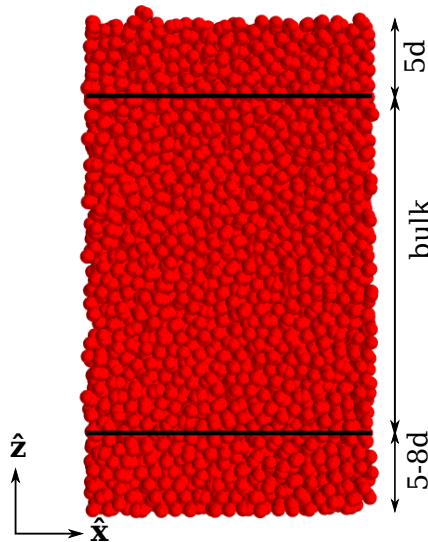
Figure 3.1: The bulk region shown in the $\hat{\mathbf{x}}$-$\hat{\mathbf{z}}$-plane.

In order to determine the packing density in the bulk region, I calculate the Voronoi volume of each dimer in the bulk. The Voronoi volume of a particle contains all points closer to this particle than any other one in the system; see Appendix A.1 for more details regarding the Voronoi tessellation and volume. A built-in package (Voro++) in LAMMPS provides the computation of the Voronoi volume of the individual spheres in the packing using a conventional Voronoi tessellation. The Voronoi volume W_i of a dimer is then found by summing the Voronoi volumes of its two constituent spheres. The bulk volume V_b occupied by N_b dimers in the bulk is calculated as $V_b = \sum_{i=1}^{N_b} W_i$. I then obtain the packing fraction as $\phi_j = N_b V_\alpha / V_b$, where V_α is the volume of a dimer with aspect ratio α. The volume of a dimer V_α is found by subtracting the overlap volume from the sum of its constituent sphere volumes. The overlap volume contains two equal spherical caps whose volume can be calculated exactly; see Appendix B.1. Note that a dimer is considered to be part of the bulk region only if the centres of both constituent spheres are within the bulk. All average quantities discussed in the following are calculated for dimers in the bulk only.

The packing fraction ϕ_j of the dimers is plotted as a function of the aspect ratio α and is shown in Fig. 3.2. As can be seen from Fig. 3.2, the packing fraction ϕ_j has a non-monotonic relationship with α, i.e., it increases as α increases until reaching a peak at $\phi_j = 0.707$ for $\alpha = \alpha_{\max} = 1.4$, beyond that it decreases. These results are in agreement with previous studies [33, 58] and also show reasonably good agreement with results from a mean-field calculation [34], shown in Fig. 3.2.
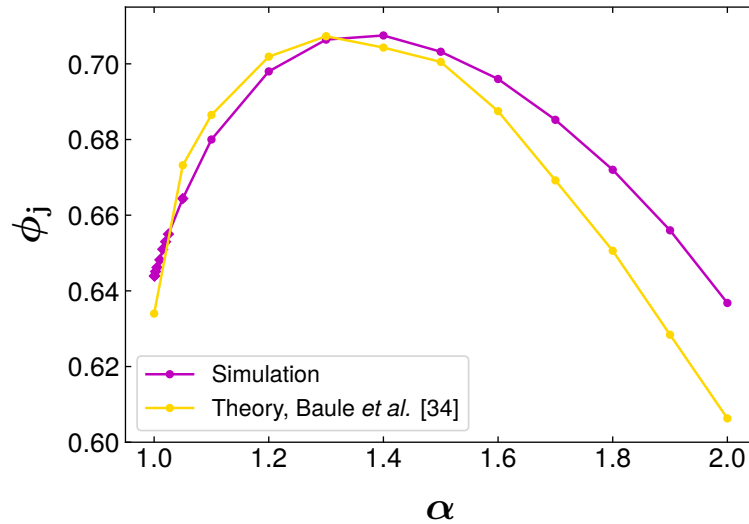
Figure 3.2: The packing fraction ϕ_j as a function of the dimer aspect ratio α. Simulation values of ϕ_j are shown averaged over 10 independent simulation runs for $\alpha \geq 1.1$ (dots), and for a single run for $\alpha < 1.1$ (diamonds). The standard deviation for each dot is approximately ± 0.0004.

In the mean-field calculations, the non-linear behaviour of the packing density with the aspect ratio has been explained based on the excluded volume. The density increases upon deviation of the aspect ratio from 1 due to the additional orientational degrees of freedom, whereby a particle can be oriented to efficiently reduce the space left by its neighbours and the change in the ratio of the excluded volume over the particle volume from the spherical value is slight. For larger aspect ratios, the density decreases since the excluded volume effects increase whereas the average number of contacts remains constant. The dependence of the average contact number on the aspect ratio is deduced by evaluating the probability of finding stable degenerate configurations, which is then used to predict the density $\phi(\alpha)$. Systematic deviations between the simulations and the mean-field theory are in particular visible in the behaviour for larger aspect ratios $\alpha > 1.5$, which

are likely due to the strong mean-field assumptions. In fact, the mean-field theory relies on a reduction of higher-order positional correlations to pair correlations and also neglects orientational correlations between particles. The latter become more significant for particles of larger aspect ratios, see Sec. 3.2.4.1. Nevertheless, the mean-field theory has successfully explained the non-linear behaviour of packing densities of rotationally symmetric elongated particles. I attempt to extend the mean-field approach to asymmetric dimers (obtained by overlapping two spheres of varying diameter). However, this is unavailable due to a lack of analytical expression for the Voronoi boundary between two particles, which is necessary to define two central quantities in the framework: the Voronoi excluded volume and surface. The details of the mean-field theory and the Voronoi boundary construction for asymmetric dimers are provided in Appendix A.2.

3.2.2   Voronoi Volume Statistics

The volume function plays the key role in the description of granular matter by the Edwards thermodynamics [4]. Voronoi tessellation is a convenient way of determining the system volume, where Voronoi cell volume distributions characterize the structural properties of granular assemblies. Theoretical modelling of such distributions, however, remains challenging. Aste *et al.* proposed an analytical model for sphere packings that provides a well-predicted distribution of Voronoi cell volumes [7, 107]. They found a so-called k-Gamma distribution, with probability density function $f(\rho) \propto \rho^{k-1} \exp(-k\rho)$, where the quantity ρ is the rescaled volume

$$\rho = \frac{W - W_{\min}}{\overline{W} - W_{\min}}. \tag{3.1}$$

This model significantly depends on $W_{\min}$, the minimum attainable cell volume. The average volume from the distribution coincides with $\overline{W}$, and the variance is

$$\sigma_w^2 = \frac{(\overline{W} - W_{\min})^2}{k}, \tag{3.2}$$

so the value of k parameter can be empirically calculated as $k = (\overline{W} - W_{\min})^2 / \sigma_w^2$ [7, 107], which is associated with a granular temperature, known as the compactivity. For sphere packings, the parameter k varies between 11 and 15, which agrees with the neighbour number contributes to the Voronoi cell of any given sphere. Schaller *et al.* extended the k-Gamma model for disordered packings of uniaxial ellipsoidal particles [108]. They first identified the densest local structures for each aspect ratio by numerically minimizing the Set Voronoi cell volume around a central particle and then determined $W_{\min}(\alpha)$ to validate k-Gamma model predictions. For the investigated aspect ratio regime, they observed that the model satisfactorily predicts the full distribution of the Voronoi volumes for ellipsoid packings, where the best-fit value for k varies between 12 and 17.

(a)

(b)
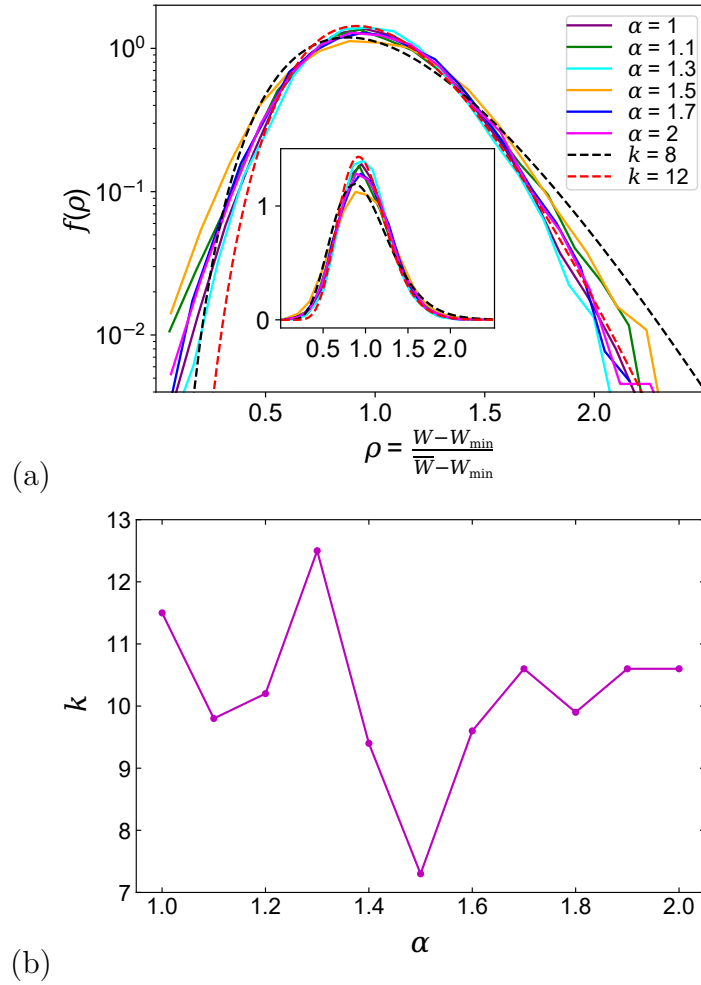
Figure 3.3: Distributions of the rescaled Voronoi cell volumes in disordered sphere ($\alpha = 1$) and dimer packings. (a) The simulation data for various α is compared with two k-Gamma distributions ($k = 8$ and $k = 12$). The inset illustrates the same data on a linear scale. (b) The best-fit k values are presented as a function of α.

Here, I investigate the Voronoi volume distributions by the k-Gamma model for dimer packings with different aspect ratios and sphere packing. Since the densest local structures of dimers are unknown, $W_{\min}(\alpha)$ is determined from the simulations as the minimum Voronoi cell volume among the dimers in the bulk. As can be seen from Fig. 3.3(a), for both spheres and dimers, the model curves satisfactorily reproduce the distributions of the Voronoi volumes. However, the

data exhibits deviations from the k-Gamma curve at small and large volumes. A similar behaviour at small volumes of ellipsoids has also been observed due to some specific packing motifs [108]. Although they found the best-fit value for k generally increases with elongation of the ellipsoids, such a variation of k with α has not been detected for the dimers, where it takes values between 7 and 13, see Fig. 3.3(b). I suspect this difference arises due to the discrepancy in determining $W_{\min}(\alpha)$.

3.2.3 Contact and coordination numbers

The contact number z, defined as the average number of contacts per particle, plays an essential role in the mechanical stability or rigidity of jammed particle packings [109]. Isostatic conjecture estimates a minimum value of z below which the system loses rigidity [110]. To satisfy the isostatic condition, the total number of interparticle contacts should equal the total number of degrees of freedom in packings of frictionless particles. This property implies that $z = 2d_{\mathrm{f}}$, where d_{f} denotes the degrees of freedom of a particle ($z = 6$ for spheres, $z = 10$ for rotationally symmetric shapes such as spheroids, $z = 12$ for fully asymmetric shapes). Packings with z smaller or larger than the isostatic value are referred to as hypostatic and hyperstatic, respectively. Previous studies have shown that the isostaticity in general holds for sphere packings [5]. For rotationally symmetric shapes, the isostatic conjecture would predict a discontinuous jump in z from 6 to 10 as soon as the shape deviates from the sphere. However, such a discontinuity has not been observed for packings of convex elongated shapes. They instead

have been generally found hypostatic with a smooth increase from the spherical isostatic z value [32, 50, 54, 74]. Nonetheless, it has been shown that the convex curvature of some contacts allows these hypostatic packings to be mechanically stable [74, 80, 111]. Unlike convex elongated shapes, dimer packings are observed as isostatic [58, 73, 77].

Here, I determine z for the dimer packings. I refer z to the average number of contact points of a dimer. I also present the coordination number cn as the average number of neighbours of a dimer, whereby a neighbour is defined as another dimer with which at least one contact point is shared. While $z = cn$ for smooth convex shapes like spheres, ellipsoids, and spherocylinders, $z \geq cn$ for concave shapes like dimers, since two particles can share more than one contact point. In general, two dimers A and B share a contact point if the separation vector of two spheres i and j, with sphere i in dimer A and sphere j in dimer B, satisfies $r_{ij} \leq d$, which can be detected with high numerical precision. Two dimers can thus share up to four different contact points. Due to the soft interaction potential, the contact "point" is strictly a small overlap region, which creates some complications at small dimer aspect ratios. I introduce these problems in the following.

The behaviour of cn as a function of α and the associated distributions of cn for a set of aspect ratios are shown in Fig. 3.4(a). A smooth increase of $cn(\alpha)$ is observed for $\alpha > 1$ with a maximum at $cn = 8.34$ for $\alpha = 1.5$ followed by a slight decay. The qualitative behaviour is in line with the results of [58], where
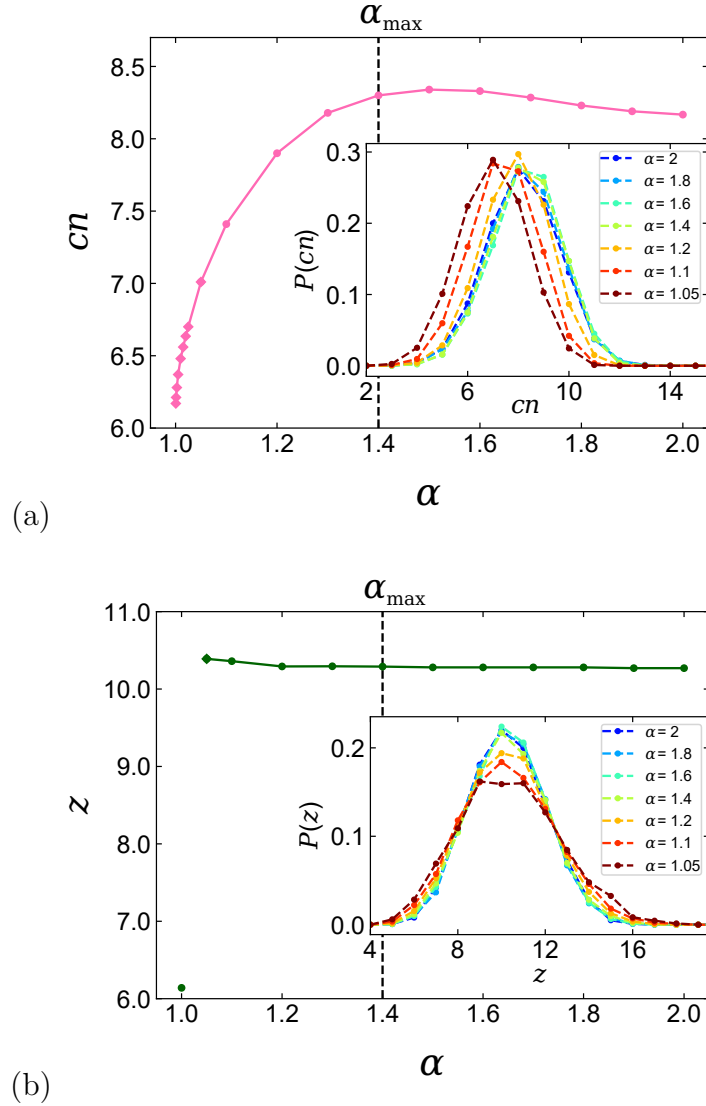
Figure 3.4: (a) The coordination number *cn vs.* α and distributions $P(cn)$ for various aspect ratios (inset). (b) The contact number *z vs.* α and distributions $P(z)$ (inset). The values of *cn* and *z* are shown averaged over 10 independent simulation runs for $\alpha \geq 1.1$ and $\alpha = 1$ (dots), and for a single run for $1 < \alpha < 1.1$ (diamonds).

dimer packings were generated using an energy minimization protocol, although the values of *cn* here are consistently larger over the range of aspect ratios. The distributions $P(cn)$ are approximately symmetric and Gaussian (Fig. 3.4(a,inset)).

On the other hand, the contact number *z* does not exhibit such a smooth

increase, see Fig. 3.4(b). I first establish the baseline for sphere packings at $\alpha = 1$ with the simulation protocol and find that $z = 6.14$ for spheres. This value is slightly above the isostatic value of $z = 6$. This difference is due to the gravitational packing protocol and the interaction potential with non-zero softness. The studies of sphere packings [24, 33] using a similar protocol also found comparable values for z. The smallest aspect ratio found for dimer packings with a reliable contact number is $\alpha = 1.05$, for which $z = 10.39$. For larger aspect ratios, z decreases slightly, but then remains unchanged at $z = 10.28$ for $\alpha > 1.2$. The difference with the isostatic value $z = 2d_{\mathrm{f}} = 10$ is approximately of the same magnitude as the difference for spheres using the packing protocol. By comparison, the studies in [58, 73, 77] find that dimers are almost exactly isostatic, which is thus in line with my findings. The observation of a constant z for all aspect ratios of dimers is an important difference with the behaviour of convex elongated shapes such as spheroids and spherocylinders, which are hypostatic ($z < 2d_f$) at small aspect ratios and show a smooth increase upon shape deformation from the sphere like the coordination number *cn* here.

I observe that for very small aspect ratios $\alpha \in (1, 1.05)$ the calculation of z is unreliable, since the particle model used here leads to incorrect contact detections: the overlap regions due to the particle softness can extend far enough into the dimer as to create a contact with an interior sphere as illustrated in Fig. 3.5.

Shiraishi *et al.* also observed such problematic contact configurations for dimers
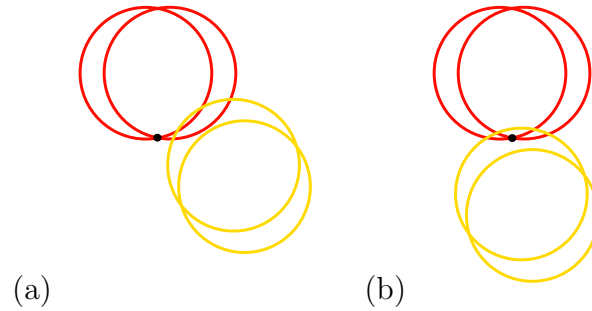
Figure 3.5: Illustrations of "double" and "cusp" contacts shown in 2D as discussed in [73]. (a) Double contact: the yellow sphere is embedded into the red dimer so deeply that it contacts both red spheres. (b) Cusp contact: the yellow sphere contacts both red spheres by covering the cusp point (black point) of the red dimer.

in [58, 73], where they are separated into "double" and "cusp" contacts, see Fig. 3.5. They investigated the contact number of dimer packings using a compression protocol with soft particle interactions for various packing densities ϕ. For large enough values of the excess packing density $\Delta\phi = \phi - \phi_j$, where ϕ_j denotes the packing density at jamming onset, "double" and "cusp" contacts were observed. In their analysis, these contacts could thus be avoided by setting an upper limit for $\Delta\phi$ at each aspect ratio studied and they observed that this upper limit approaches zero as $\alpha \to 1$. Here, I find that the occurrence of these configurations depends on the stiffness value K_n of the particles as shown in Fig. 3.6, where it can be seen that the threshold aspect ratio, at which double and cusp contacts occur, is shifted to smaller aspect ratios for larger K_n. For any value of K_n, double and cusp contacts will occur at sufficiently small aspect ratios and thus the contact number very close to the sphere shape can not be reliably established. For $K_n = 2 \times 10^5$, double and cusp contacts do not occur for $\alpha \geq 1.05$, which is thus the lower limit of α used in the contact number analysis.
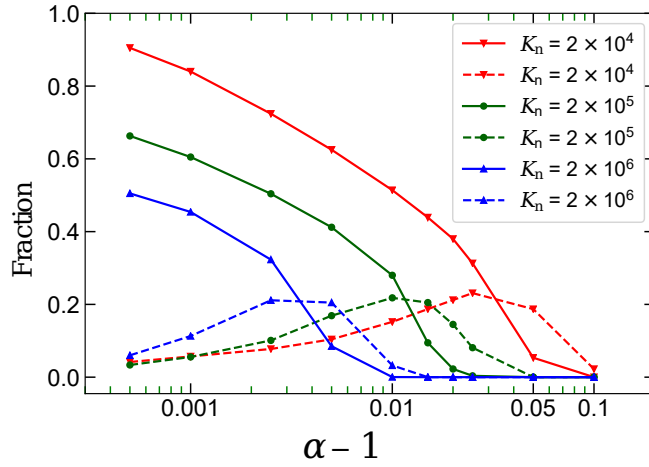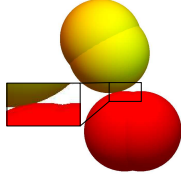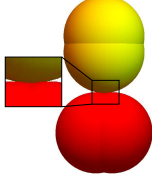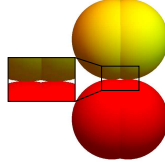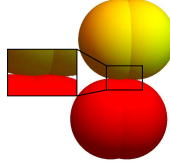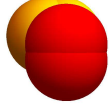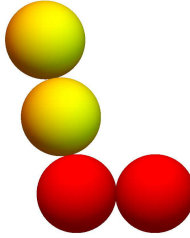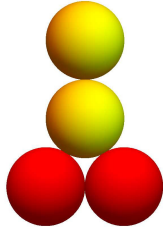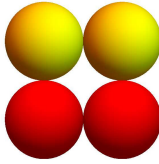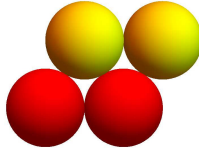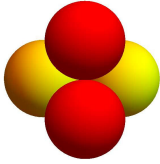
Figure 3.6: The fraction of double (solid lines) and cusp contacts (dashed lines) in the dimer packings for small α and three normal spring constants K_{n}.

In order to better understand the packing microstructure, I define five distinct contact configurations according to the number of contact points that are shared by two neighbouring dimers, as illustrated in Table 3-A. Excluding the regime $\alpha \in [1, 1.05)$, I display how the fraction of each configuration type changes as a function of α in Fig. 3.7. Even though the average number of contacts z is approximately constant over this range of α, the underlying contact configurations change significantly with α. Most notably, the two most common contact configurations, Type 1 and Type 2, increase and decrease, respectively, as α increases up to around α_{max} and remain approximately unchanged for $\alpha > \alpha_{\mathrm{max}}$. The remaining contact configurations confirm this trend, showing the strongest variations in the regime $\alpha < \alpha_{\mathrm{max}}$. Overall, contact configurations, in which spheres of neighbouring dimers only have one contact point (Type 1 and Type 3) increase, while those with multiple contact points (Types 2,4,5) decrease as the packing becomes denser up

Table 3-A: Five distinct contact configurations of two dimers. We show illustrations for aspect ratios $\alpha = 1.2$ and $\alpha = 2$. The total number of contact points for each type is: one (Type 1), two (Type 2,3), three (Type 4), four (Type 5).

α	Type 1	Type 2	Type 3	Type 4	Type 5
1.2					
2					

to the packing density peak at α_{max}. This trend is somewhat counter-intuitive, since the Type 2,4,5 configurations correspond to more optimal local arrangements between two dimers, which locally reduce the packing density. Similar results for the fractions of these five configuration types have been found for packings of shapes composed of four overlapping spheres [112].

I re-assign the problematic double and cusp contacts to infer the microscopic properties of the small aspect ratio regime. For example, a double contact as in Fig. 3.5(a), which creates two overlaps of sphere pairs and is thus counted as two contact points, could be counted as only one, effectively ignoring the incorrect overlap with the interior sphere. This can be done likewise for other contact configurations, which require a careful consideration of the relative position and orientation of the overlapping dimer pair, see the full discussion in Appendix B.4.
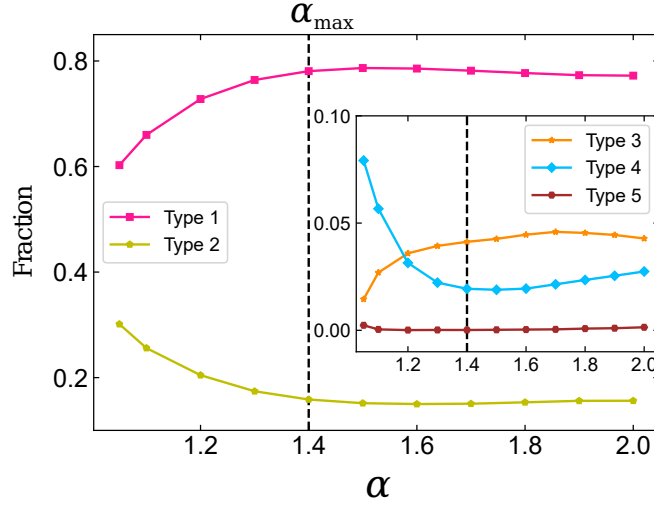
Figure 3.7: The fractions of the five contact configuration types of Table 3-A for packings of dimers with different α. Each data point is shown averaged over 10 independent simulation runs.

Re-assigning contacts in this way leads to a rapid but smooth decrease of z to the corresponding value of spheres $z \approx 6$ as $\alpha \to 1$ (Fig. B5), but also exhibits seemingly unphysical behaviour, such as sharp peaks in the fractions of the Type 1–5 contact configurations around $\alpha \approx 1.05$, i.e., at the aspect ratio where double and cusp contacts start to occur (Fig. B6).

3.2.4   Order metrics

I employ several order metrics to measure global and local ordering in the dimer packings at various aspect ratios. The nematic orientational order parameter and the orientational pair correlation function are used to evaluate orientational ordering. Translational ordering is investigated with bond orientational order parameters, the radial distribution function and bond angle distributions. All calculations are made for the particles within the bulk volume so as to discard the crystallized region

observed at the bottom of the container.

3.2.4.1 Metrics for orientational order

The nematic orientational order parameter S has traditionally been applied to identify different ordered phases of liquid crystals by characterising the average molecular orientation [113]. S is defined as:

$$S = \langle P_2(\cos\beta) \rangle \approx \frac{1}{N_{\mathrm{b}}} \sum_{i}^{N_{\mathrm{b}}} P_2(\cos\beta_i) \tag{3.3}$$

where $P_2(x) = \frac{1}{2}\left(3x^2 - 1\right)$ is the second Legendre polynomial and β_i the angle between the orientation of dimer i and the so-called director, which specifies the average orientation of the particles. The dimer orientation is described by the unit vector $\mathbf{u}^{(i)}$ measured along the dimer's long axis, see Fig. 3.8.

I apply this parameter to the dimer packings to quantify the global orientational order. When all $\mathbf{u}^{(i)}$ are randomly oriented, $S = 0$, while if all $\mathbf{u}^{(i)}$ are oriented in a plane normal to the director, $S = -0.5$, which corresponds to a perfect oblate phase. When all $\mathbf{u}^{(i)}$ are aligned with the director, the packings have perfect nematic order with $S = 1$.

In order to determine the director and S, I first evaluate the tensor Ω defined as:

$$\Omega_{kl} = \frac{1}{N_{\mathrm{b}}} \sum_{i}^{N_{\mathrm{b}}} \left(\frac{3}{2} u_k^{(i)} u_l^{(i)} - \frac{1}{2}\delta_{kl} \right) \tag{3.4}$$
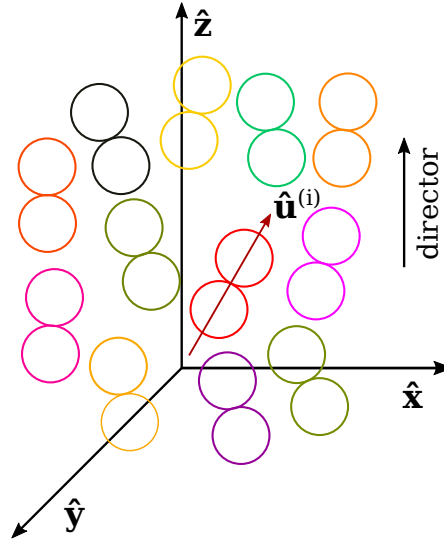
Figure 3.8: Orientation of dimer particles in the packings, $\mathbf{u}^{(i)}$ is unit vector of dimer i along its long axis describes its orientation. The director is found aligned with $\hat{\mathbf{z}}$-axis for all aspect ratios.

Denoting by $\lambda_{\max}$ the eigenvalue of Ω with the largest absolute value, I identify the director as the eigenvector corresponding to $\lambda_{\max}$. For all aspect ratios, I find that the director is aligned with the $\hat{\mathbf{z}}$-axis (gravity direction). S is then obtained directly as:

$$S = \lambda_{\max}. \tag{3.5}$$

I also determine the orientational pair correlation function S_2 in order to quantify local ordered structures at a radial distance r from a reference particle. S_2 is calculated as:

$$S_2(r) = \langle P_2(\cos \beta_{ij}) \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \rangle \approx \frac{\sum\limits_{i=1}^{N_b} \sum\limits_{j \in n_i(r)} P_2(\cos \beta_{ij}(r))}{\sum\limits_{i=1}^{N_b} |n_i(r)|} \tag{3.6}$$
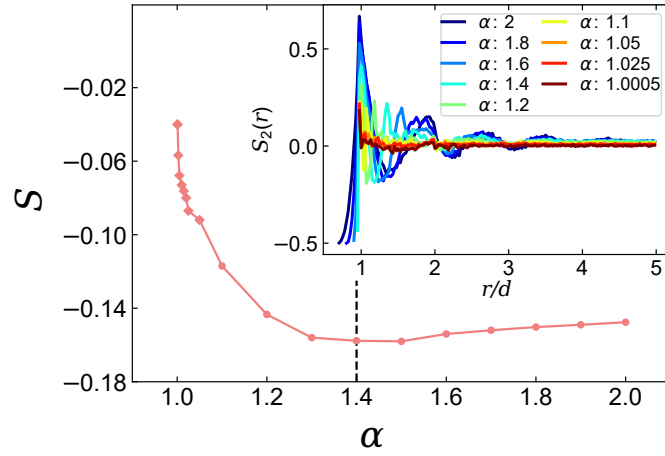
Figure 3.9: The nematic orientational order parameter S *vs.* the aspect ratio α. Values of S are shown averaged over 10 independent simulation runs for $\alpha \geq 1.1$ (dots), and for a single run for $\alpha < 1.1$ (diamonds). Inset: the orientational pair correlation function S_2 *vs.* r/d for various aspect ratios.

where $\cos\beta_{ij} = \mathbf{u}^{(i)} \cdot \mathbf{u}^{(j)}$ and $n_i(r)$ denotes the set of particles in a spherical shell

of width $\Delta(r) = 0.025d$ at a distance r from the centre of dimer i in the bulk. The

expression $|n_i(r)|$ refers to the size (cardinality) of the set $n_i(r)$. I should note that

the spherical shell considered in S_2 can extend into the boundary region beyond

the bulk and thus include particles in partially crystallized regions, although the

effect on the average should be small. In general, due to the non-periodic boundary

conditions in the $\hat{\mathbf{z}}$-direction the packings are not rotationally invariant and thus

the restriction to a radial coordinate is only an approximation.

I present the dependence of S and $S_2(r)$ on the aspect ratio α in Fig. 3.9. S

changes rapidly as α increases from the sphere value, reaching its minimum at

around α_{max} at which the densest packing is achieved and remaining approximately

constant for $\alpha > \alpha_{\mathrm{max}}$, in line with the behaviour of cn and the different contact

types. Interestingly, the behaviour of $S(\alpha)$ as $\alpha \to 1$ appears almost singular, but

the range of values is not sufficient to identify a clear power-law. The minimum of S at ≈ -0.16 indicates slight oblate ordering, where the dimers' long axes are oriented close to the horizontal plane normal to the direction of gravity. This ordering is thus in agreement with that observed in simulation studies of prolate ellipsoids using also pouring under gravity, where the packing density is found as a non-monotonic function of the aspect ratio [47–49]. In order to compare the magnitude of the orientational ordering with these studies, I also calculate the order parameter χ used in [47–49], which is defined in Eq. (B.3). Note that in [47], χ smoothly increases with the aspect ratio up to $\alpha \approx 1.5$ that achieves the densest packings, and it decreases for larger aspect ratios, whereas in [48, 49], it monotonically increases upon further elongation over the observed range of aspect ratios. I find a maximum of $\chi \approx 0.32$ for $\alpha_{\max} = 1.4$ at which the maximum in the packing density occurs. By comparison, in [47] the maximum is $\chi \approx 0.4$ for $\alpha \approx 1.5$, while [48] and [49] find $\chi \approx 0.25$ and $\chi \approx 0.5$, respectively, for $\alpha \approx 1.5$.

The plot of S_2 in Fig. 3.9(inset) demonstrates how orientational correlations become more long-range for larger aspect ratios. For small α, correlations decay rapidly within the first coordination shell, while for large α oscillations in S_2 are visible over the whole range of r/d, which is here limited by $r/d = 5$, i.e., the width of the boundary region on top of the bulk region that restricts the maximum radius of the spherical shell used in Eq. (3.6).

3.2.4.2 Bond orientational order parameters

The bond-orientational order metrics q_l and Q_l introduced by Steinhardt *et al.* [15] have most commonly been used to quantify translational order in disordered packings of spherical particles [26, 114–121]. While Q_l is widely accepted as a well-defined parameter to measure global ordering in a packing, it has been suggested that the local order parameter q_l needs more caution to reliably identify local crystalline structures in these systems [16, 122]. It was assumed that higher values of q_6 are associated with higher degrees of order [123] and averages $\langle q_6 \rangle$ have been used to quantify the overall degree of order for disordered sphere packings [118]. However, it has been found that some local configurations of disordered sphere packings that are clearly non-crystalline have exhibited the same values of q_6 as hcp or fcc crystals [16]. Therefore, in this analysis, I use recently introduced local order parameters defined by Eslami *et al.* [124] to improve the accuracy of determining local translational order in the dimer packings.

Steinhardt *et al.* [15] associated with every bond joining a particle and its neighbours a set of spherical harmonics:

$$q_{lm}(i) = \frac{1}{|NN(i)|} \sum_{j \in NN(i)} Y_{lm}(\theta_{ij}, \phi_{ij}) \tag{3.7}$$

where the Y_{lm} are spherical harmonics and θ_{ij}, ϕ_{ij} denote the polar and azimuthal angles which define the orientation of the vector (bond) pointing from the reference particle i to another particle j, see Fig. 3.10. $NN(i)$ contains the set of neighbour
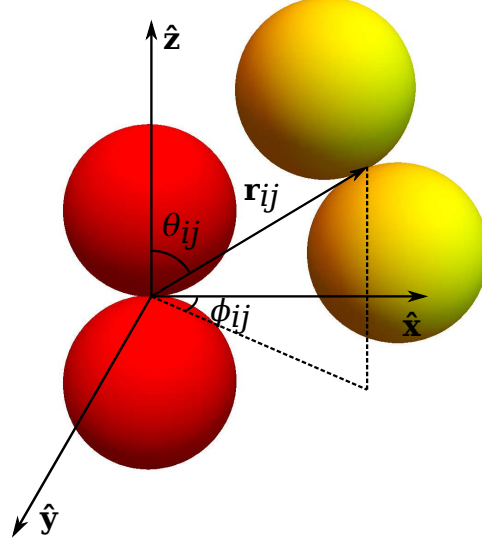
Figure 3.10: Parametrization of the separation vector (bond vector) $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ connecting the reference particle i (red) with j (yellow). The definitions of the polar and azimuthal angles, θ_{ij} and ϕ_{ij}, respectively, are indicated.

indices for particle i, which are defined as those particles j that have at least one contact with i.

The local orientational order parameter $q_l(i)$ of particle i is then defined as the following rotational invariant combination of q_{lm}:

$$q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2}. \tag{3.8}$$

Moreover, the global orientational order parameter Q_l is defined as

$$Q_l = \left(\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |Q_{lm}|^2 \right)^{1/2}, \tag{3.9}$$

where

$$Q_{lm} = \frac{1}{N_{\mathrm{b}}} \sum_{i=1}^{N_{\mathrm{b}}} q_{lm}(i). \tag{3.10}$$

Previous investigations of disordered sphere packings have shown that the choice of neighbourhood definition in Eq. 3.7 can significantly influence the numerical values and the qualitative trend of $q_l(i)$ [122]. Several different criteria were used for $NN(i)$, for example all neighbours within $1.2d$ in [15], Voronoi neighbours in [114, 123]. Recently, Eslami *et al.* introduced the local order parameters $\bar{\bar{q}}_l(i)$ to improve the determination of liquid and different crystallized phases [124]. Note that the liquid phase is related to structural disorder only. Starting from the q_{lm} of Eq. (3.7), where the neighbours in the first coordination shell are counted, I first determine

$$\tilde{q}_l(i) = \frac{1}{|NN(i)|} \sum_{j \in NN(i)} \sum_{m=-l}^{l} \hat{q}_{lm}(i)\hat{q}_{lm}^*(j), \tag{3.11}$$

where $\hat{q}_{lm}^*(j)$ is the complex conjugate of $\hat{q}_{lm}(j)$ and $\hat{q}_{lm}(i)$ is defined as follows:

$$\hat{q}_{lm}(i) = \frac{q_{lm}(i)}{\left(\sum_{m=-l}^{l} |q_{lm}(i)|^2 \right)^{1/2}} \tag{3.12}$$

Then the order parameters $\bar{\bar{q}}_l(i)$ are obtained by averaging over the first coordination shell of particle i:

$$\bar{\bar{q}}_l(i) = \frac{1}{1 + |NN(i)|} \left[\tilde{q}_l(i) + \sum_{j \in NN(i)} \tilde{q}_l(j) \right]. \tag{3.13}$$

The advantage of $\bar{\bar{q}}_l(i)$ over q_l is that they can distinguish the liquid phase and different crystalline phases in a more accurate way [124]. They indicate in fact the correlation between the order in the first and the second coordination shell of a reference particle [124]. It has been observed that $\bar{\bar{q}}_6(i)$ is large ≈ 1 for crystalline phases, while $\bar{\bar{q}}_6(i)$ assumes values close to zero for disordered (liquid) phases, which thus allows to easily discriminate between such phases. On the other hand, the values of $\bar{\bar{q}}_4(i)$ are sensitive to the crystal type, so $\bar{\bar{q}}_4(i)$ is able to distinguish bcc, fcc, and hcp crystals.

The pairs $(\bar{\bar{q}}_4, \bar{\bar{q}}_6)$ for each dimer in the bulk region of the packing are displayed in Fig. 3.11 for various aspect ratios. By comparing these results to empirical data for liquid, bcc, hcp, and fcc phases of Lennard-Jones particles from [124], I observe that the distributions at large aspect ratios ($\alpha > 1.4$) are quite clearly in a liquid phase where $-0.05 < \bar{\bar{q}}_4 < 0.3$ and $0 < \bar{\bar{q}}_6 < 0.4$. As the aspect ratio decreases, the region occupied by $\bar{\bar{q}}_4$ and $\bar{\bar{q}}_6$ expands and approaches the region occupied by the bcc/hcp crystal phases indicating the presence of a large proportion of dimers exhibiting some local translational order intermediate between a liquid and bcc/hcp crystalline order. The distributions of $\bar{\bar{q}}_4$ and $\bar{\bar{q}}_6$ with respect to the particle positions for a few aspect ratios are displayed in Fig. 3.12 and Fig. 3.13, respectively. For both $\bar{\bar{q}}_4$ and $\bar{\bar{q}}_6$, the local clusters with high values are homogeneously distributed along the bulk region in the packing.

To quantify global translational ordering, I calculate the averages $\langle \bar{\bar{q}}_4 \rangle$, $\langle \bar{\bar{q}}_6 \rangle$ and
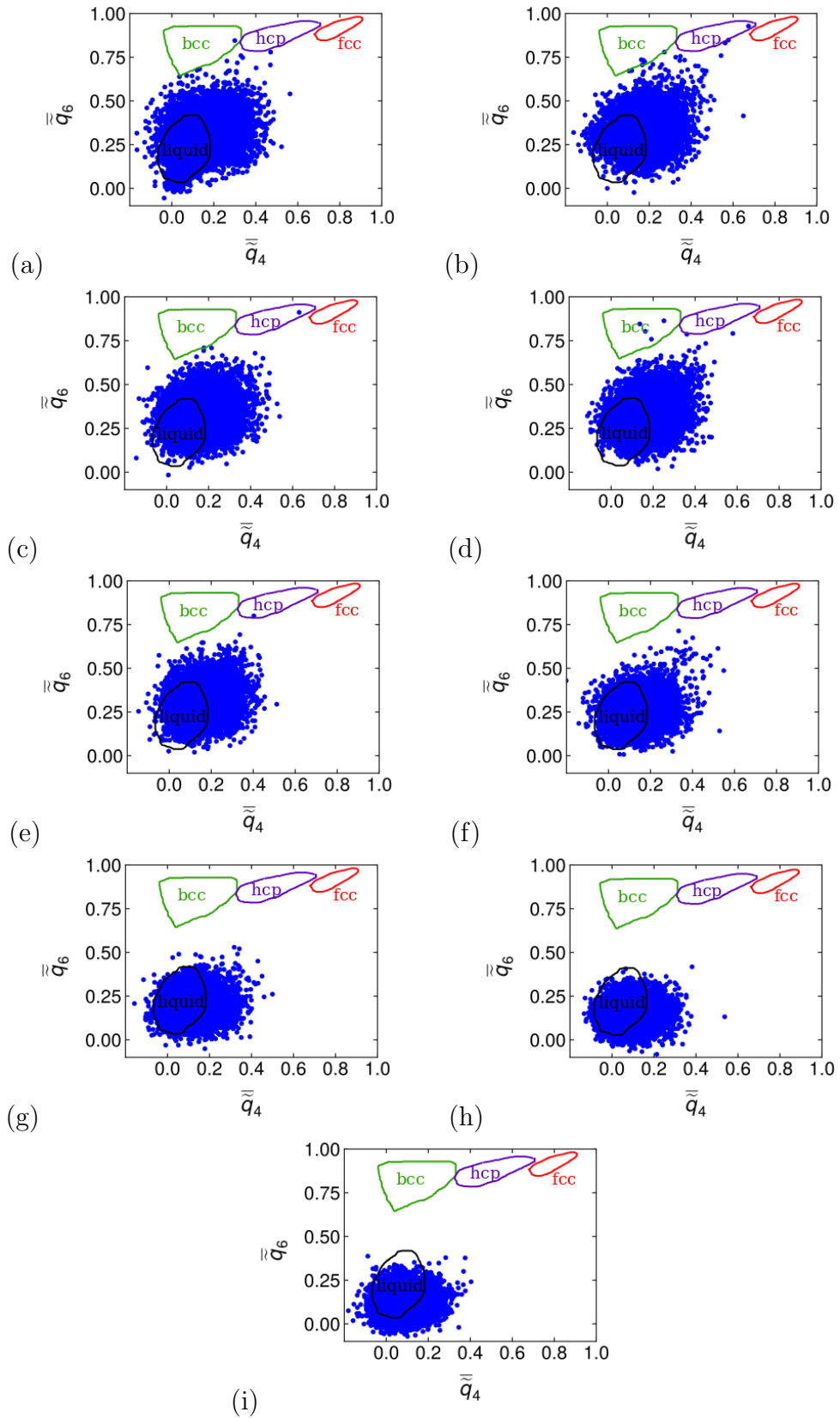
Figure 3.11: The local order parameters $\overline{\overline{q}}_4$ and $\overline{\overline{q}}_6$ for the packings of (a) $\alpha = 1.0005$, (b) $\alpha = 1.025$, (c) $\alpha = 1.05$, (d) $\alpha = 1.1$, (e) $\alpha = 1.2$, (f) $\alpha = 1.4$, (g) $\alpha = 1.6$, (h) $\alpha = 1.8$, (i) $\alpha = 2$. Every data point corresponds to a dimer in the bulk region of the packing. The sketched regions for bcc, hcp, fcc, and liquid phases of Lennard-Jones particles are taken from [124].
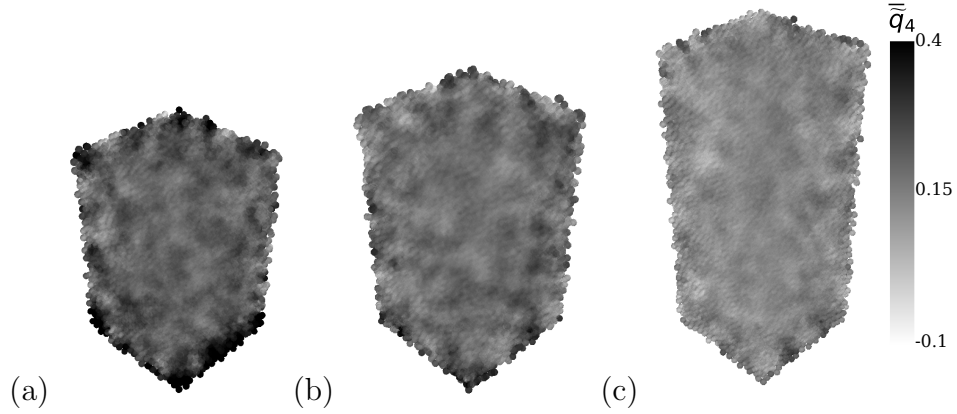
Figure 3.12: The local order parameter $\bar{\bar{\tilde{q}}}_4$ distributions along the packing for (a) $\alpha = 1.05$, (b) $\alpha = 1.4$, (c) $\alpha = 2$. The excluded particles from the bulk are also shown here, exhibiting high crystallization at small aspect ratios.
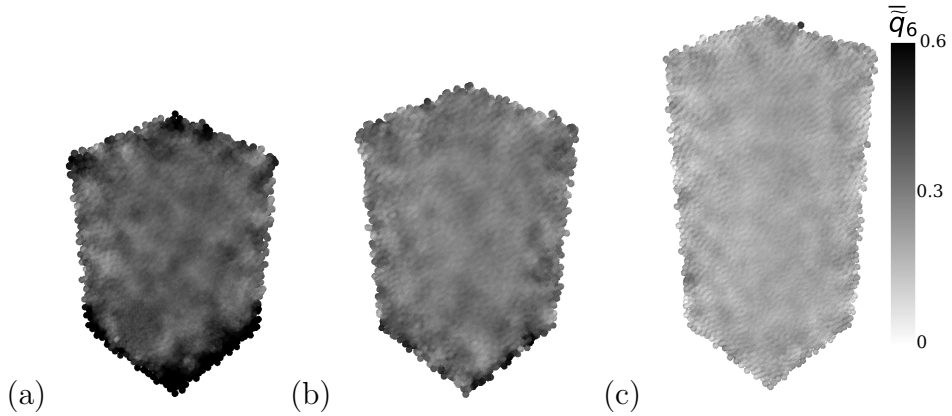
Figure 3.13: The local order parameter $\bar{\bar{\tilde{q}}}_6$ distributions along the packing for (a) $\alpha = 1.05$, (b) $\alpha = 1.4$, (c) $\alpha = 2$. The excluded particles from the bulk are also shown here, exhibiting high crystallization at small aspect ratios.

compare their values with the global order parameters Q_4, Q_6 for different aspect ratios, see Fig. 3.14. While Q_4 is close to zero for all aspect ratios, there is a slight increase in Q_6 for $\alpha < 1.4$ implying some global ordering at small aspect ratios. In line with the observations in Fig. 3.11, both $\langle \bar{\bar{\tilde{q}}}_4 \rangle$ and $\langle \bar{\bar{\tilde{q}}}_6 \rangle$ are non-zero and monotonically decreasing as α increases, whereby $\langle \bar{\bar{\tilde{q}}}_6 \rangle$ varies over a larger range than $\langle \bar{\bar{\tilde{q}}}_4 \rangle$. Both averages are considerably larger than the corresponding averages of a fluid phase, which were determined as $\langle \bar{\bar{\tilde{q}}}_4 \rangle \approx 0.06$ and $\langle \bar{\bar{\tilde{q}}}_6 \rangle \approx 0.2$. Overall,
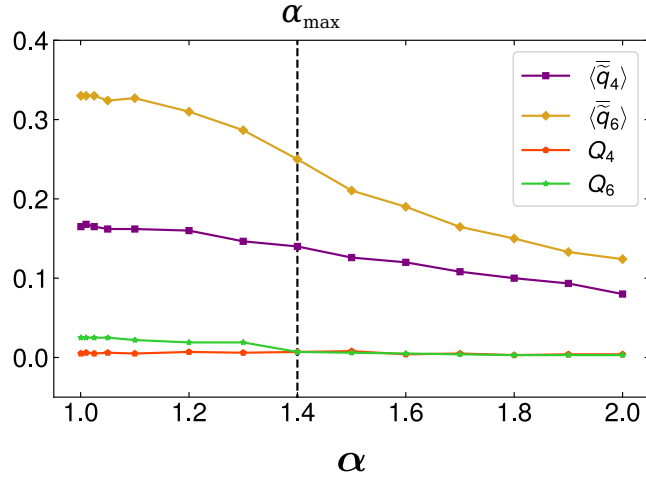
Figure 3.14: The global bond orientational order parameters Q_4, Q_6 and the averages $\langle \overline{\widetilde{q}}_4 \rangle$ and $\langle \overline{\widetilde{q}}_6 \rangle$ vs. α. By comparison, $\langle \overline{\widetilde{q}}_4 \rangle \approx 0.06$ and $\langle \overline{\widetilde{q}}_6 \rangle \approx 0.2$ for the liquid phase of Lennard-Jones particles [124].

the results indicate that at large aspect ratios the packing is more translationally disordered than at small aspect ratios.

3.2.4.3   Radial distribution function

I calculate the radial distribution function $g(r)$ to further examine the translational correlations between the dimers. The radial distribution function of the bulk dimers is determined as

$$g(r) = \frac{\sum\limits_{i=1}^{N_{\mathrm{b}}} |n_i(r)|}{N_{\mathrm{b}} \iota V_{\mathrm{shell}}(r)}, \tag{3.14}$$

where $n_i(r)$ denotes the set of particles in a spherical shell of width $\Delta(r) = 0.025d$ at a distance r from the centre of dimer i in the bulk, ι is the particle number density, and $V_{\mathrm{shell}}(r)$ is the volume of the shell. As discussed for the orientational

correlation function $S_2(r)$, Eq. (3.6), the restriction to a radial coordinate is only an approximation due to the fact the packings are not rotationally invariant. As before the spherical shell can extend into the boundary region beyond the bulk. I plot $g(r)$ as a function of r/d for various aspect ratios in Fig. 3.15. As can be seen, for small aspect ratios, $g(r)$ exhibits the characteristic shape of sphere packings with a main peak at $r/d = 1$ and a split second-peak at $r/d \approx 1.7$ and $r/d \approx 2$ [14, 19, 24, 36, 50, 101, 117, 125]. For larger aspect ratios, these sharp peaks broaden and reduce in height. These results are consistent with the variation of bond orientational correlations with the aspect ratio discussed above, where elongation in the dimers results in a reduction of translational correlations.
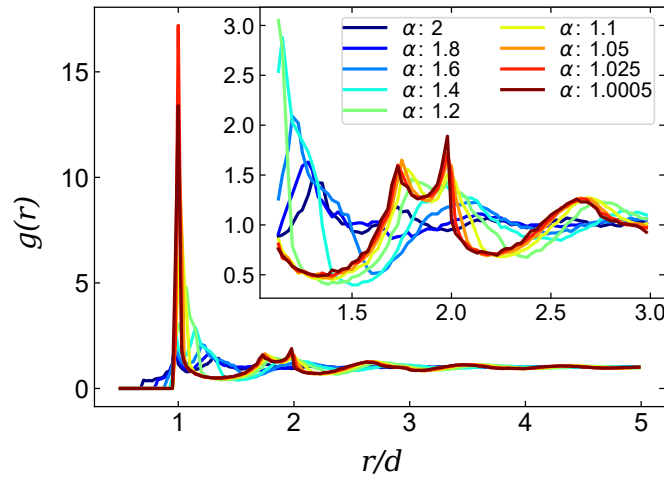
Figure 3.15: The radial distribution function $g(r)$ of the dimer packings, Eq. (3.14), for different α. Inset: enlargement of the regime $r/d \in [1.125, 3]$.

3.2.4.4 Bond angle distribution

Here, I measure the probability for a dimer to have a contact at a particular direction relative to its long axis. For each dimer pair i, j, I determine the polar

angle θ_{ij} and the azimuthal angle ϕ_{ij} of the bond vector $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ in the reference

frame of particle i, see Fig. 3.10.
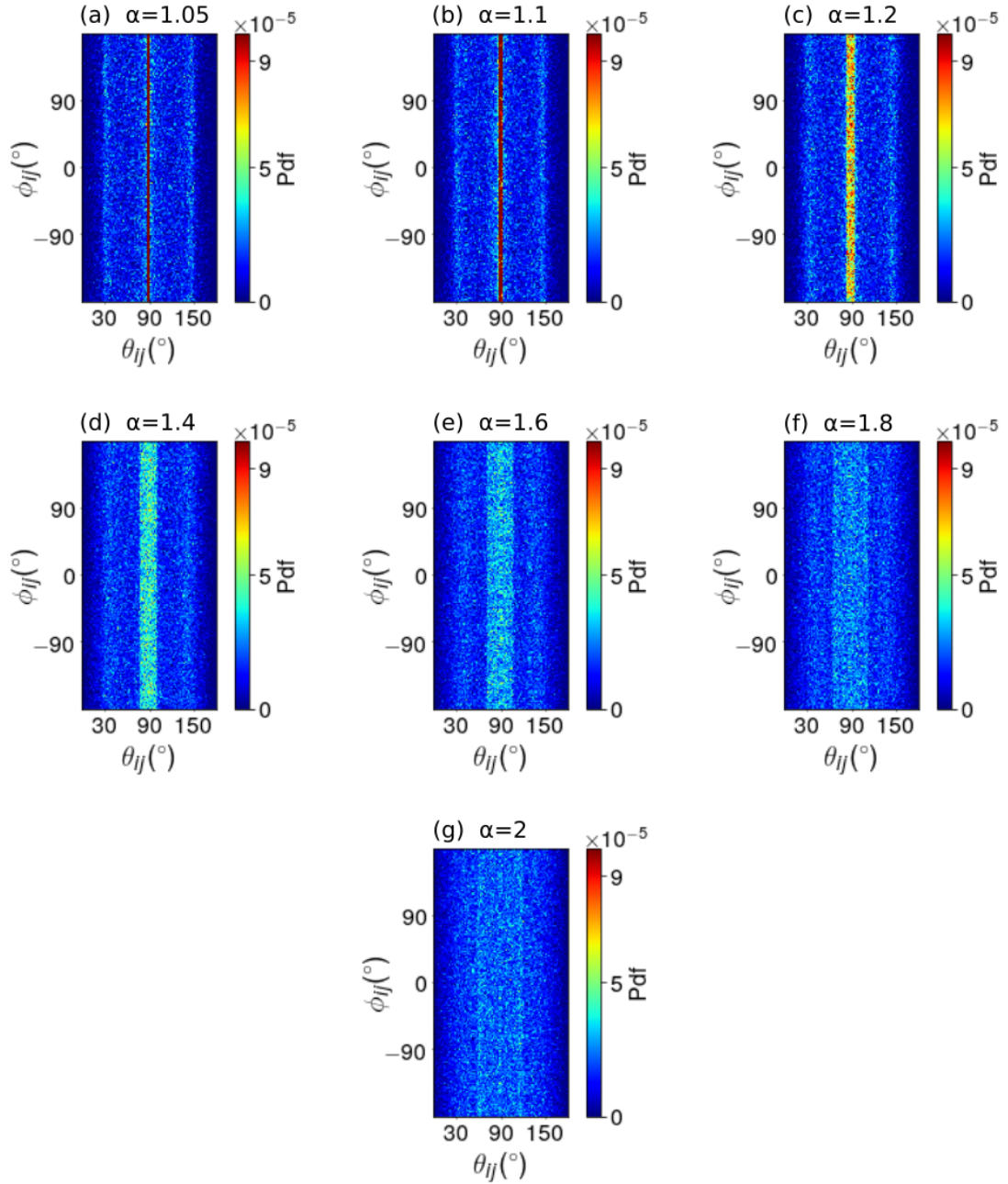
Figure 3.16: PDFs of the polar and azimuthal angles θ_{ij}, ϕ_{ij} of the bond vectors $\mathbf{r}_{ij}$ for all neighbour pairs i, j and different aspect ratios.

The probability density functions (PDFs) of θ_{ij} and ϕ_{ij} are shown for various

aspect ratios in Fig. 3.16. It can be clearly seen from Fig. 3.16 that at small aspect

ratios dimers have primarily contacts at $\theta_{ij} = 90°$. As the aspect ratio increases,

the band around 90° widens and finally disappears at $\alpha = 2$. For small aspect

ratios, there are also symmetric secondary peaks visible at $\theta_{ij} = 30°$ and $\theta_{ij} = 150°$,

with all contacts occurring within the range $\theta_{ij} \in [30°, 150°]$ up to $\alpha \approx 1.4$.
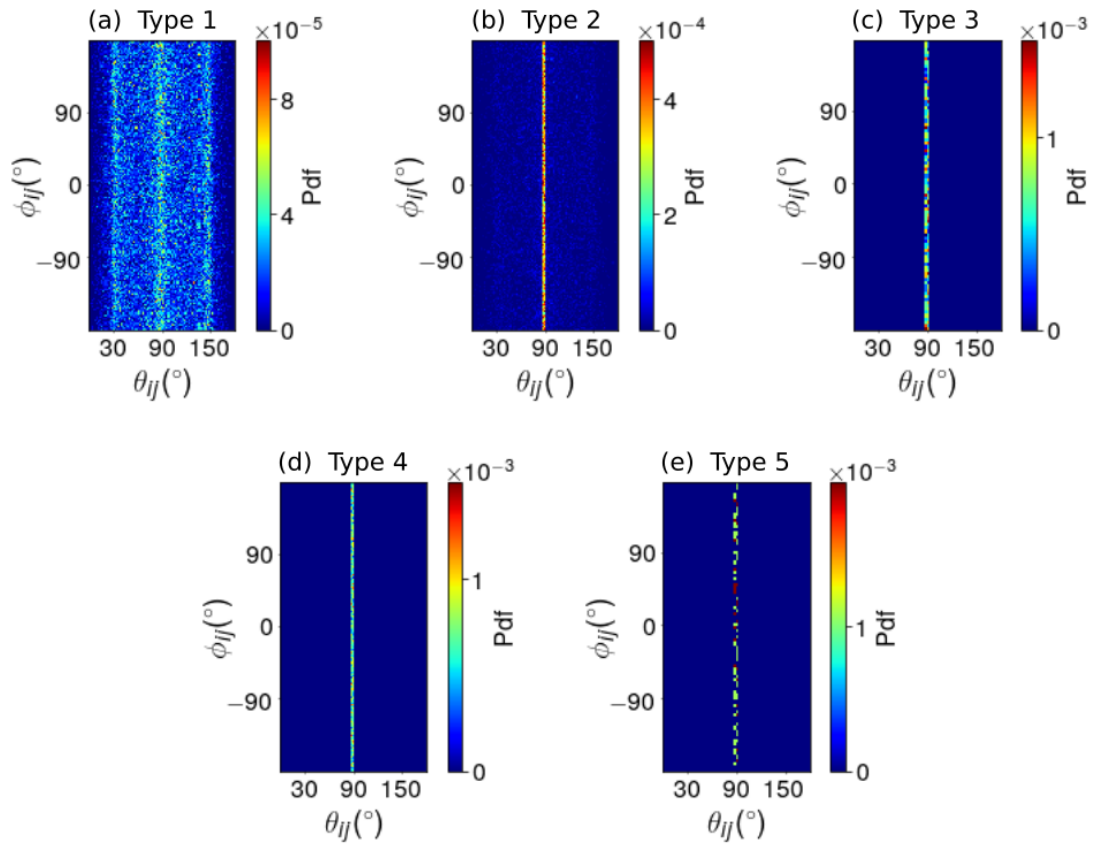
Figure 3.17: PDFs of the polar and azimuthal angles θ_{ij}, ϕ_{ij} of the bond vectors $\mathbf{r}_{ij}$ for all neighbour pairs i, j with a specific contact type. Aspect ratio: $\alpha = 1.05$.

I further refine the PDFs of θ_{ij}, ϕ_{ij} according to the contact configuration

type between neighbouring dimers to get a better insight into the origin of these

structures, see Figs. 3.17–3.19. For aspect ratio $\alpha = 1.05$ (Fig. 3.17), it can be

seen that for Type 2—5 only configurations with $\theta_{ij} \approx 90°$ are possible due to the

geometric constraint of these configuration types. The structure observed in the overall bond diagram at very small aspect ratios (Fig. 3.16a,b) is thus primarily due to Type 1 configurations and the peak at $\theta_{ij} \approx 90°$.
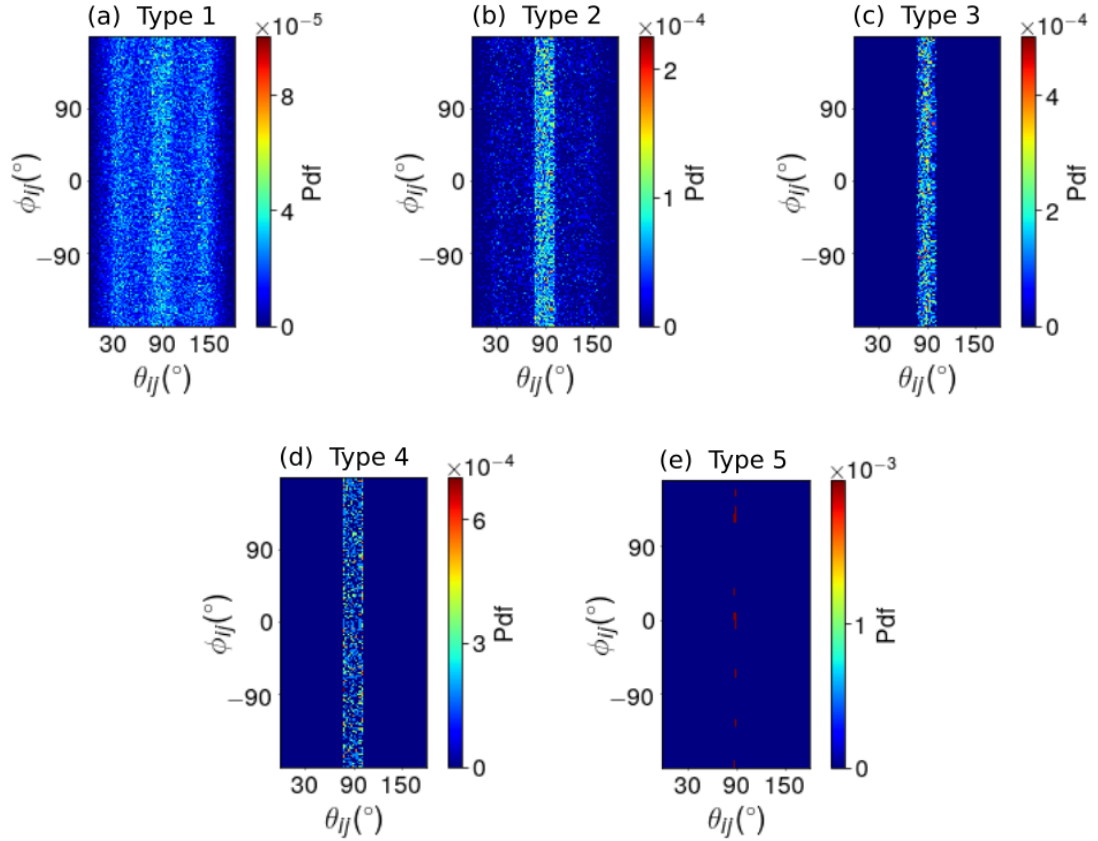
Figure 3.18: PDFs of the polar and azimuthal angles θ_{ij}, ϕ_{ij} of the bond vectors $\mathbf{r}_{ij}$ for all neighbour pairs i, j with a specific contact type. Aspect ratio: $\alpha = \alpha_{\max} = 1.4$.

For larger aspect ratios $\alpha = \alpha_{\max} = 1.4$ and $\alpha = 2$, the bands for Type 2—4 widen due to the increase in possible relative orientations that still satisfy the contact constraint (see Figs. 3.18,3.19). This excludes Type 5 configurations which are available only in a narrow width of possible polar angles by definition. As expected, Type 1 configurations with only a single contact point between neighbours, which thus least constrains the relative orientations, exhibit a wide band of possible

polar angles at all aspect ratios, see Figs. 3.17(a),3.18(a),3.19(a). Interestingly, this band still exhibits some structure, with a main peak at $\theta_{ij} = 90°$ and symmetric secondary peaks at $\theta_{ij} = 30°$ and $\theta_{ij} = 150°$ for both $\alpha = 1.05$ and $\alpha = 1.4$, which disappear for $\alpha = 2$.
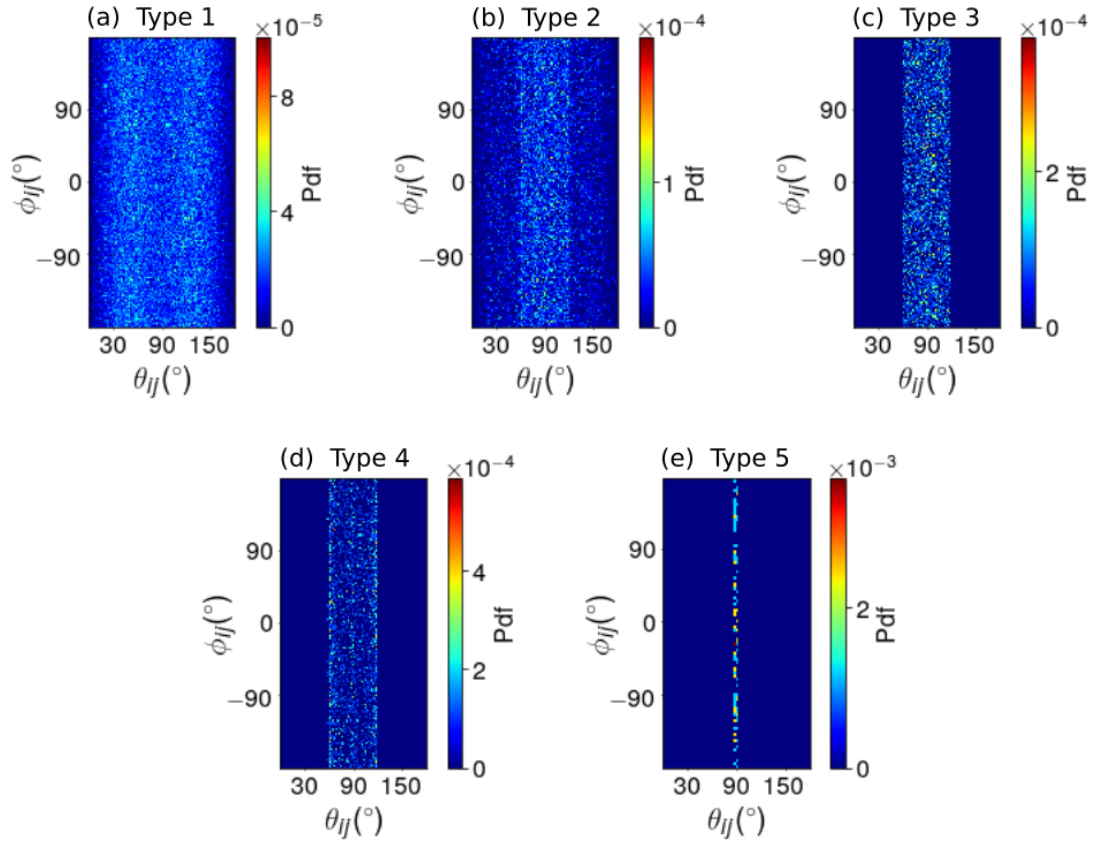
Figure 3.19: PDFs of the polar and azimuthal angles θ_{ij}, ϕ_{ij} of the bond vectors $\mathbf{r}_{ij}$ for all neighbour pairs i, j with a specific contact type. Aspect ratio: $\alpha = 2$.

3.3 Conclusions

In this chapter, I have identified the structural features that accompany the formation of the peak in the packing density of elongated non-spherical particles. The results have shown that (i) the coordination number cn; (ii) the fractions of

Type 1–4 contact configurations; and (iii) the nematic order parameter S undergo rapid changes upon deforming spheres into dimers with aspect ratios up to $\alpha \approx \alpha_{\max}$, while further elongation of the dimers leaves these metrics largely unchanged. This observation highlights that the peak in the packing density of Fig. 3.2 arises due to microscopic re-arrangements up to $\alpha \approx \alpha_{\max}$ and subsequent excluded volume effects: the contact configurations remain statistically unchanged for $\alpha > \alpha_{\max}$, but since the particles are longer the packing can sustain more empty space while being mechanically stable, in line with the phenomenological description of spherocylinder packings using the random contact equation, which predicts a decay $\phi_{\mathrm{j}} \sim 1/\alpha$ [104].

I have also studied the distributions of Voronoi volume for the dimer packings and shown that they can be scaled using a shifted k-Gamma probability distribution akin to sphere packings. Unlike other metrics measured in this study, the shape parameter k has not exhibited a general trend in the aspect ratio variation. In this study, I have also shown that the mean-field theory can not be extended to asymmetric dimers since it is unavailable to provide an analytically tractable expression for the VB between two particles.

An important result of this chapter is that using the particle model (overlapping spheres) to generate dimers has some problematic contact configurations. Although sphere-to-sphere interactions are easy to detect, this model does not allow to resolve the contact configurations at very small aspect ratios when interactions are not truly hard. Therefore, this difficulty in the simulations has hindered examining the

analytical predictions from effective medium theory on the contact number scaling for very small shape deformations [126]. The problematic double and cusp contacts should generally occur for shapes composed of overlapping (soft) spheres as used, e.g., in the optimisation studies of [45, 68], which also might prevent a detailed analysis of the contact properties of such simulated packings.

The investigation of ordering effects in the dimer packings highlights the competition between orientational and translational correlations between particles due to the growth in the aspect ratio. While the translational correlations are more significant for small aspect ratios, the elongation induces the dimers to have more orientationally ordered local structures (with slight global oblate ordering) but less translational order akin to those in a liquid. Dimers at large aspect ratios thus exhibit structures that resemble a liquid crystal in terms of these metrics. I should emphasize that the structural features identified here might be specific to the gravitational packing protocol used and might not occur in dimer packings obtained with other packing methods, such as energy minimization from a random initial configuration [58]. Nevertheless, due to the simplicity of the protocol, which is also relevant in many real-world scenarios, I expect these results would be significant in understanding the packing density and structural properties of granular matter composed of non-spherical particles.

Chapter 4

Disordered Packings of Binary Mixtures of Dimer Particles

4.1 Introduction

Jammed particle packings have been studied to understand the structures of amorphous materials such as powders, reinforcing fibres, granular matter and glasses [127]. Most studies mainly focused on monodisperse packings of spherical and non-spherical particles for which a plethora of experimental and theoretical results are available [5, 25, 28, 46]. Elongated non-spherical particles such as ellipsoids [32, 43, 47, 48], spherocylinders [36, 50–54, 56], and dimers [33, 34, 58] exhibit a non-monotonic variation of the packing density upon deviation from the spherical shape, with a maximum at specific aspect ratios, which has also been

observed for dimers in Chapter 3. Although these shapes are more complicated than spherical particles due to the additional rotational degrees of freedom, they do not fully represent natural materials yet since polydispersity in both shape and size is inevitable for particle aggregates in nature.

In fact, numerous studies of binary and polydisperse packings of spherical particles demonstrate that changing the size distribution of the particles improves the packing density [82–92]. Investigations of non-spherical particle mixtures, on the other hand, are less common and restricted in a few common shapes. Studies of jammed packings of spherocylinder–sphere mixtures with the same diameter report a density maximum occurring at the same aspect ratio of the spherocylinders as in monodisperse packings, regardless of the relative volume fraction of the two species [55, 56, 93]. The density maximum is also present under different conditions, such as when spherocylinders and spheres have equal volume or different diameters [53, 94, 95]. Binary mixtures of two species of spherocylinders with the same diameter but different aspect ratios exhibit likewise a density maximum at a unique aspect ratio of one species, when the shape of the second component is kept fixed [53, 95, 96]. This unique aspect ratio is irrespective of the fixed aspect ratio of the second component and agrees with the value at which the packing density of the monodisperse spherocylinder packing is maximal.

Non-spherical particle mixtures also satisfy a remarkable empirical ideal mixing law, which states that the inverse packing density is a linear superposition of the

inverse packing densities of pure (single species) phases weighted by their relative volume fraction [56, 93, 97, 98]. For example, the total packing density ϕ_j of spherocylinder–sphere mixtures is then given by

$$\phi_j^{-1} = X_1\phi_1^{-1} + (1 - X_1)\phi_2^{-1}, \tag{4.1}$$

where X_1 denotes the relative volume fraction of spherocylinders with a monodisperse packing density of ϕ_1 and likewise ϕ_2 is the packing density of monodisperse spheres with volume fraction $1 - X_1$. For binary mixtures the ideal mixing law establishes a linear dependence of ϕ_j^{-1} on the relative volume fraction of one species. Clearly, since the relative volume fraction is fixed by the setup, the packing density of the mixture is then fully determined by the packing densities of the pure phases and thus independent of the segregation state. This independence implies further that particle orientations are completely uncorrelated showing an interesting similarity with a plastic crystal [56].

All these results provide a new route to optimise the packing densities of granular materials by mixing non-spherical particles, which are relevant in industrial applications. For example, the prediction of the spherocylinder aspect ratio that maximises the packing density of a binary mixture has been used in metallurgy, transportation, agricultural, and chemical industries [37]. Furthermore, the ideality observed in binary mixtures of spherocylinders is relevant to designing materials that are free of the "Brazil-nut effect", which occupy a constant volume irrespectively

of particle dispersion [56]. This effect occurs as a result of component segregation, where the larger particles are found at the top of a granular material containing a mixture of variously sized particles upon shaking, vibrating or shearing. For example, one commonly observes that the larger particles reach the top when opening a box of breakfast cereal, which is undesirable. Understanding the effect of shape and size on mixing and segregation is of great interest to food manufacturing and pharmaceuticals [128]. Therefore, it is a fundamental question whether the universal density maxima and the ideality in mixing exist also for binary mixtures of elongated non-convex particles. In this chapter, I focus on investigating the properties of disordered packings of both dimer-sphere and dimer-dimer mixtures generated by using the gravitational pouring protocol in LAMMPS. I take into account the combined effect of shape and size by using the same diameter for both components in the packings. In addition to the packing density, I also aim to investigate the contact statistics of the mixtures. In Chapter 3, I have shown that the emergence of the density maximum is associated with microstructural re-arrangements that are manifest in the contact statistics of local configurations. Here, I examine whether such re-arrangements are also relevant in the behaviour of the packing density of binary dimer mixtures.

This chapter is organized as follows. In Section 4.2 and Section 4.3, I present results on my analysis of the packing density, contact and coordination numbers, and contact configurations for dimer-sphere and dimer-dimer mixtures, respectively. Finally, I conclude in Section 4.4 with a discussion of the findings.

4.2 Dimer-Sphere Mixtures

4.2.1 Packing Density

I calculate the total packing density of the dimer-sphere mixtures for various relative dimer volume fractions, X_d which is defined as:

$$X_\mathrm{d} = \frac{N_\mathrm{d}V_\mathrm{d}}{N_\mathrm{d}V_\mathrm{d} + N_\mathrm{s}V_\mathrm{s}} = \frac{\phi_\mathrm{d}}{\phi_\mathrm{j}} \tag{4.2}$$

where ϕ_i is the volume fraction of one of the two components $i \equiv \mathrm{d(dimers)}$ and $i \equiv \mathrm{s(spheres)}$, $\phi_\mathrm{j} = \phi_\mathrm{d} + \phi_\mathrm{s}$ is the total packing density, N_i and V_i are the number of particles and the volume of one particle of component i, respectively. An example for disordered solid of dimer-sphere mixtures is displayed in Fig. 4.1.
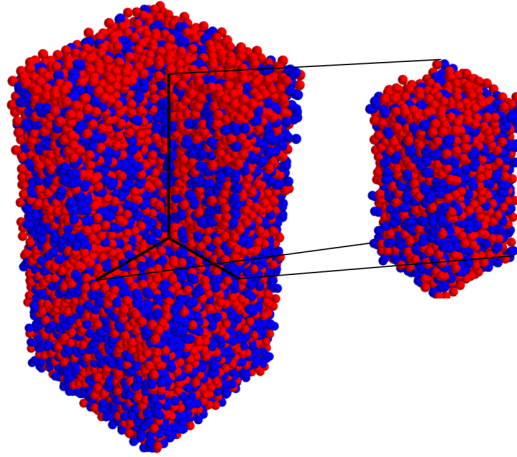
Figure 4.1: A disordered solid of mixtures of dimers (coloured in blue) with $\alpha = 1.4$ and spheres (coloured in red) for the relative dimer volume fraction $X_\mathrm{d} = 0.5$.

For the density calculation of binary mixtures, I use the same definition as in Sec. 3.2 for a bulk region, i.e., excluding the particles within $5 - 8d$ from the

container floor and the ones within $5d$ from the upper-most particles, see Fig. 4.2. Then, I measure the Voronoi volume of each particle in the bulk. As in Sec. 3.2, the Voronoi volume W_l of a dimer l is found by summing the Voronoi volumes of its two constituent spheres, where the Voronoi volume W_j of a sphere j is provided by LAMMPS. The total bulk volume occupied by N_s spheres and N_d dimers in the bulk is calculated as $V_b = \sum_{j=1}^{N_s} W_j + \sum_{l=1}^{N_d} W_l$. I obtain the total packing density as $\phi_j = (N_s V_s + N_d V_d)/V_b$. All average quantities discussed in the following are calculated for the bulk particles only. For both dimer–sphere and dimer–dimer mixtures, I run ten independent simulations and average all data points in the following plots over them.
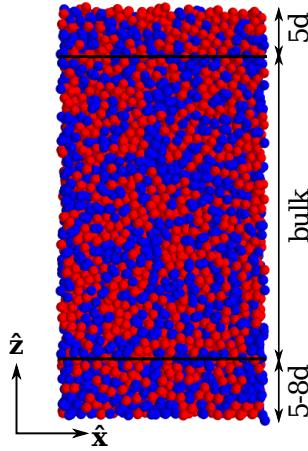
Figure 4.2: The bulk region shown in the $\hat{\mathbf{x}}$-$\hat{\mathbf{z}}$-plane.

I plot the total packing density of the mixture as a function of the dimer aspect ratio α in Fig. 4.3(a) for different relative dimer volume fractions. There are two limiting cases: pure sphere and pure dimer packings are obtained when $X_d = 0$ and $X_d = 1$, respectively. It can be seen from Fig. 4.3(a) that ϕ_j exhibits a

non-monotonic behaviour with the dimer aspect ratio by yielding a maximum at $\alpha \approx 1.4$, as in the case of monodisperse dimer packings [33, 34, 58], independently of the relative volume fraction. Moreover, the packing density monotonically grows upon the increase in the relative amount of dimers up to the absolute maximum ($\phi_{j} = 0.707$) that has been achieved in the study of monodisperse dimer packings in Chapter 3. The appearance of a unique maximum is also in agreement with previous studies of packings of spherocylinder-sphere mixtures with the same diameter condition for various compositions [55, 56, 93] and have been explained by the competition between local caging (a short spherocylinder can be oriented to minimize the space left by its contacting neighbours) and excluded volume effects. Other studies simulated the packings of spherocylinder–sphere mixtures only for one fixed spherocylinder volume fraction and with different diameters, they still found a maximum in the packing density at one unique rod aspect ratio [53, 94, 95].

In order to investigate the validity of the ideal mixing law Eq. (4.1) in the dimer-sphere mixtures, I plot the inverse packing density as a function of the dimer volume fraction X_{d} for several aspect ratios. As shown in Fig. 4.3(b), the different curves are indeed well described by a linear relationship; see Appendix C.1 for the statistical analysis of the consistency of the mixing law with the data. This linearity suggests that the packing density of the mixture is independent of the segregation state, i.e., a completely mixed packing has the same volume as one consisting of two separate phases, each composed of only dimers or spheres, respectively, as discussed in detail in [56].
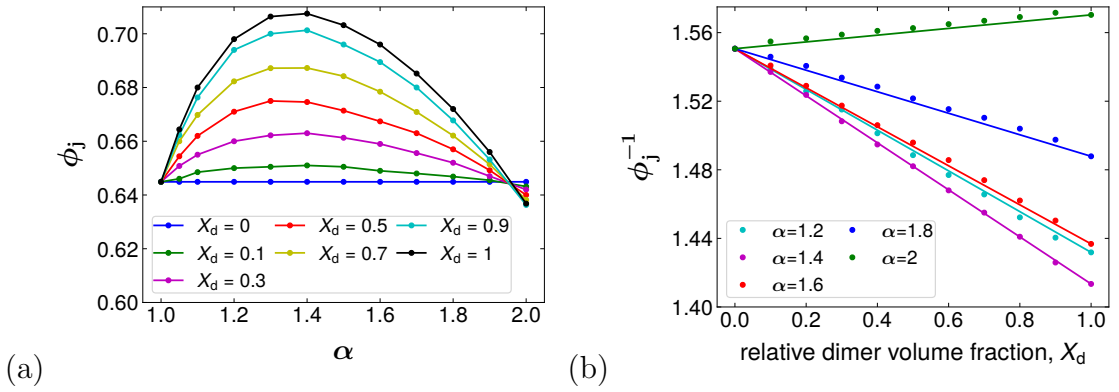
Figure 4.3: (a) The packing density ϕ_j of binary mixtures of dimer and sphere as a function of the dimer aspect ratio α. Each data point is shown averaged over 10 independent simulation runs, with a standard deviation approximately ± 0.0004, (b) The inverse packing density ϕ_j^{-1} of the mixtures as a function of the relative dimer volume fraction X_d for several dimer aspect ratios. Solid lines are obtained by the ideal mixing law Eq. 4.1 without any fit parameter: the lines are fully specified by the packing densities of pure monodisperse sphere and dimer packings (when $X_d = 0$ and $X_d = 1$, respectively).

4.2.2 Contact and coordination numbers

I measure the contact number z and the coordination number cn for the mixtures of dimers and spheres. Considering different pair interactions, I define four types of contact numbers: dimer-to-dimer z_{dd}, dimer-to-sphere z_{ds}, sphere-to-sphere z_{ss}, and sphere-to-dimer z_{sd}, which denote the average number of contact points of the former component with the latter one. The contact number of dimer particles z_d and the contact number of spheres z_s are found as, respectively:

$$z_d = z_{dd} + z_{ds} \tag{4.3}$$

$$z_s = z_{sd} + z_{ss}. \tag{4.4}$$

The overall contact number z is then defined as:

$$z = \frac{z_\mathrm{d} N_\mathrm{d} + z_\mathrm{s} N_\mathrm{s}}{N_\mathrm{d} + N_\mathrm{s}}. \tag{4.5}$$

Similarly, there are four types of coordination number: dimer-to-dimer cn_dd, dimer-to-sphere cn_ds, sphere-to-sphere cn_ss, and sphere-to-dimer cn_sd, which denote the average number of neighbours of the former component with the latter one. The coordination number of dimer particles cn_d and the coordination number of spheres cn_s are found as, respectively:

$$cn_\mathrm{d} = cn_\mathrm{dd} + cn_\mathrm{ds} \tag{4.6}$$

$$cn_\mathrm{s} = cn_\mathrm{sd} + cn_\mathrm{ss}. \tag{4.7}$$

The overall coordination number cn is then defined as:

$$cn = \frac{cn_\mathrm{d} N_\mathrm{d} + cn_\mathrm{s} N_\mathrm{s}}{N_\mathrm{d} + N_\mathrm{s}}. \tag{4.8}$$

In Fig. 4.4(a), I show the overall contact number z as a function of the relative dimer volume fraction X_d. As can be seen, for all aspect ratios, z monotonically increases from the contact number of monodisperse sphere packings, 6.14 to that of pure dimer ones, 10.28 as X_d grows. The overall coordination number cn of the mixtures is also strongly dependent on the relative dimer volume fraction,
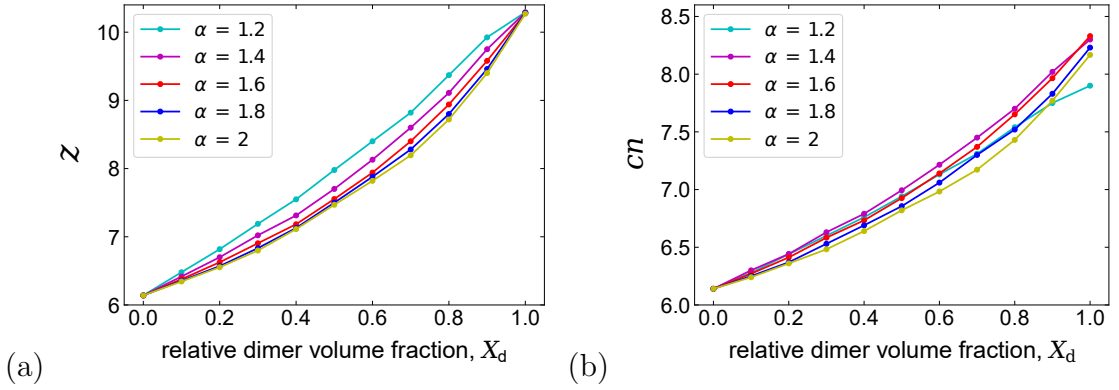
Figure 4.4: (a) The overall contact number z, (b) The overall coordination number cn of dimer–sphere mixtures vs the relative dimer volume fraction X_d for various dimer aspect ratios.

as displayed in Fig. 4.4(b). The variation of cn with the aspect ratio becomes

noticeable and approaches that for monodisperse dimer packings as X_d increases.
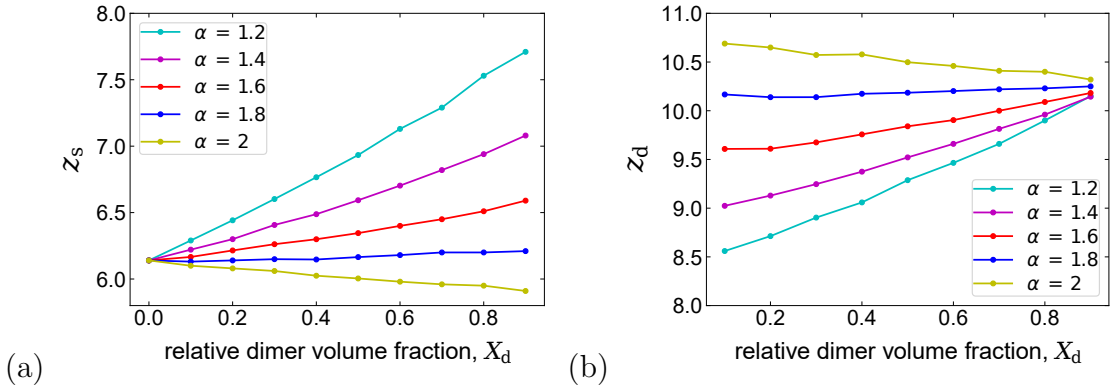
Figure 4.5: (a) The contact number of spheres z_s (b) The contact number of dimers z_d of binary mixtures vs the relative dimer volume fraction X_d for several dimer aspect ratios.

The contact numbers of spheres and dimers, z_s and z_d are calculated separately

and shown in Fig. 4.5 as a function of X_d. Both z_s and z_d vary approximately

linearly with X_d exhibiting a positive slope for $\alpha \leq 1.8$ and a negative slope for

$\alpha = 2$. When $\alpha = 2$, $z_s < 6$ for large dimer volume fractions, indicating that the

spheres have fewer contacts than required for mechanical stability in this regime.
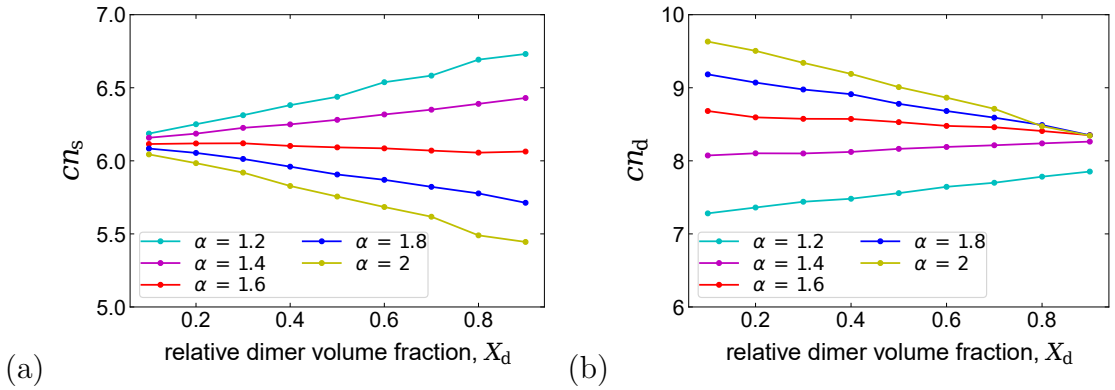
Figure 4.6: (a) The coordination number of spheres cn_s, (b) The coordination number of dimers cn_d of binary mixtures vs the relative dimer volume fraction X_d for several dimer aspect ratios.

Similarly, while dimers of $\alpha = 2$ have more contacts than in the monodisperse case ($z_d > 10.28$), shorter dimers lose some of their contacts in the dimer–sphere mixtures. This decrease in z_d is more dramatic for small aspect ratios in low dimer concentration regimes. Overall, it can be seen that the isostatic condition is generally violated for the individual components in the disordered packings of dimer–sphere mixtures. While one species is hyperstatic, the other one is always hypostatic. I also calculate the coordination numbers of spheres and dimers (cn_s, cn_d) separately and display them as a function of X_d in Fig. 4.6. The dependence on X_d is also approximately linear with similar trends as for the contact numbers, but the change in slope from positive to negative occurs now already for $\alpha > 1.4$. The validity of a linear relationship of z_s, z_d, cn_s, and cn_d as a function of X_d highlights that also the component-wise contact and neighbour numbers satisfy a simple superposition principle for mixing and are fully specified by the corresponding numbers of monodisperse packings.
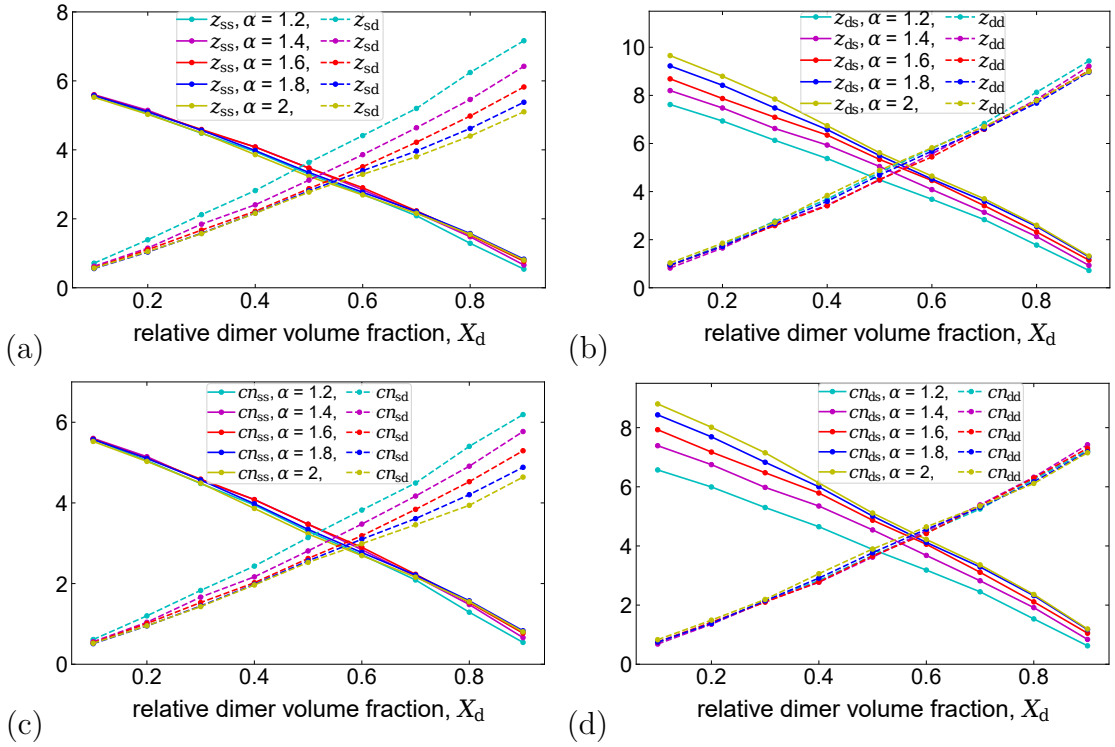
Figure 4.7: The partial contact numbers for binary mixtures (a) sphere-to-sphere z_{ss} and sphere-to-dimer z_{sd}, (b) dimer-to-dimer z_{dd} and dimer-to-sphere z_{ds} vs the relative dimer volume fraction X_d, The partial coordination numbers (c) sphere-to-sphere cn_{ss} and sphere-to-dimer cn_{sd}, (d) dimer-to-sphere cn_{ds} and dimer-to-dimer cn_{dd} vs the relative dimer volume fraction X_d for various dimer aspect ratios.
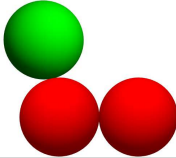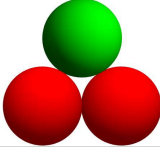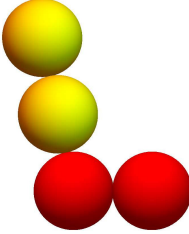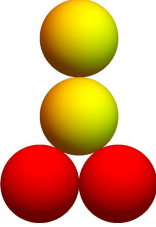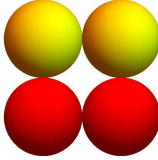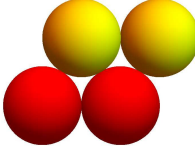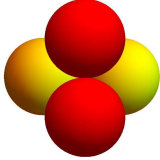
The partial contact and coordination numbers calculated for each combination of pair components are plotted with X_d in Fig. 4.7. Interestingly, both the contact and coordination numbers measured for the interaction between the same components exhibit a slight dependence on the aspect ratio for all volume fractions. On the other hand, the contacts between different species strongly depend on the dimer aspect ratio. The growth rate of z_{sd} and cn_{sd} with X_d is considerably higher for small aspect ratios, while the decline in z_{ds} and cn_{ds} is more significant for larger ones. These results imply that the variations of the contact and coordination numbers of spheres and dimers with the aspect ratio for the same X_d, displayed

in Fig. 4.5 and Fig. 4.6, are primarily due to the interactions between different components.

4.2.3   Contact configurations

In order to better understand changes in the microstructure of the packing due to shape variations and mixing, I investigate five distinct contact configurations, defined in Chapter 3 for monodisperse dimer packings according to the number of contact points shared by two neighbouring particles, see Table 4-A. I count the number of contact types occurring in the mixture per component pair and calculate the fractions of these contacts as shown in Fig. 4.8. Here, the Type $1_{\mathrm{sd,ds}}$ fraction refers to the fraction of Type 1 contacts among all contacts between dimers and spheres, i.e., the fractions of Type $1_{\mathrm{sd,ds}}$ and Type $2_{\mathrm{sd,ds}}$ add up to 1. Likewise, the fractions of the five different types of dimer–dimer contacts add up to 1. Fig. 4.8 shows that these fractions change initially upon increasing the dimer aspect ratio α up to the region at which the density peak occurs ($\alpha \approx 1.4$), but remain approximately unchanged for $\alpha > 1.4$. Moreover, it can be seen that the fraction of Type 1 contacts increases, while that of Type 2 contacts decreases as the mixture packs more dense for increasing α, which is somewhat counterintuitive, since Type 2 configurations are locally more compact, see Table 4-A. These observations are analogous to the monodisperse case, which further supports the qualitative picture that the peak in the packing density arises due to the interplay of structural rearrangements for small α and subsequent excluded volume effects with unchanged

Table 4-A: Two distinct contact configurations of sphere and dimer (sd,ds) and five distinct contact configurations of two neighbouring dimers (dd). Illustrations are shown for aspect ratio $\alpha = 2$. The total number of contact points for each type is: one (Type 1), two (Type 2,3), three (Type 4), four (Type 5).

Pair	Type 1	Type 2	Type 3	Type 4	Type 5
sd,ds					
dd					

structure.

Surprisingly, all the different fractions in Fig. 4.8 show almost no variations with a change of X_{d}. Indeed the fractions for the dimer–dimer contact configurations are almost identical to the monodisperse case. I believe that this is not a simple consequence of the normalization of these fractions. In fact, it could have been expected that the fraction of Type 1 dimer–dimer configurations is different when there are a lot of dimers available as contacts than if there are few, but this is not the case.
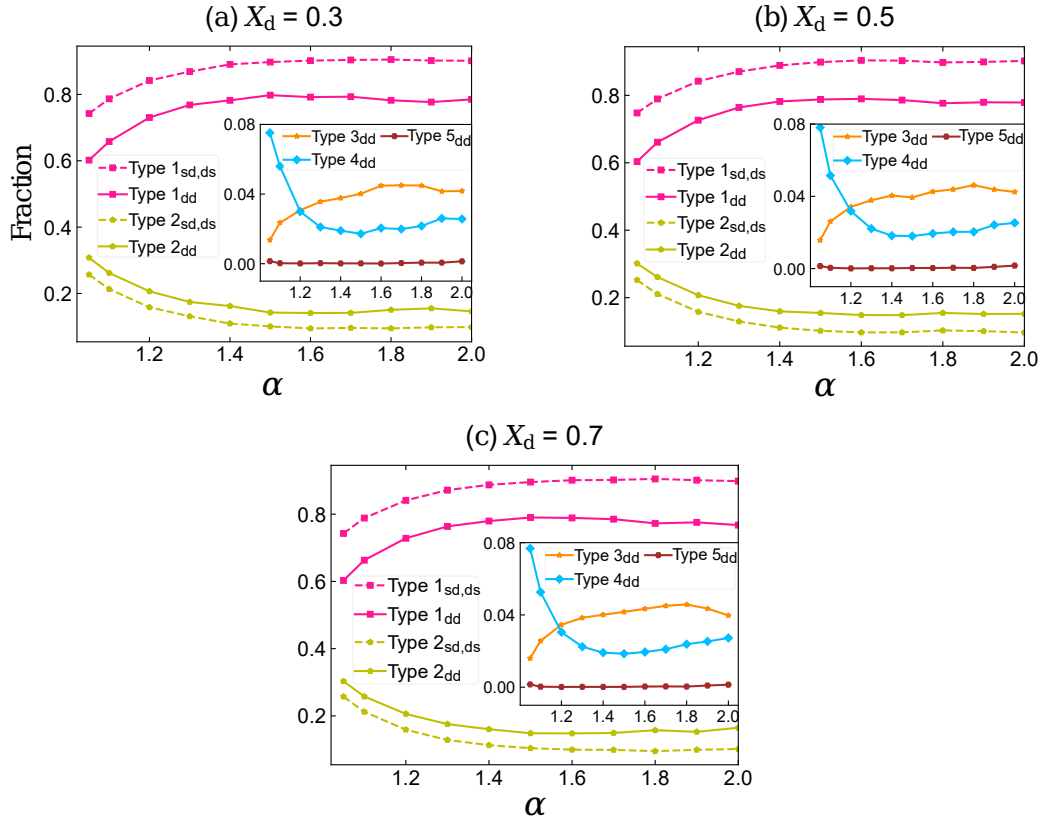
Figure 4.8: The fractions of the contact configuration types of Table 4-A between the components of binary mixtures, contact pairs are indicated by lower indices: sphere-to-dimer, dimer-to-sphere and dimer-to-dimer $(\mathrm{sd}, \mathrm{ds}, \mathrm{dd})$ for different dimer volume fractions (a) $X_\mathrm{d} = 0.3$, (b) $X_\mathrm{d} = 0.5$, (c) $X_\mathrm{d} = 0.7$.

4.3 Dimer–Dimer Mixtures

4.3.1 Packing Density

I measure the bulk packing density ϕ_j of the binary dimer mixtures for various mixture compositions. The two species of dimers (dimer 1 and dimer 2) have different aspect ratios, α_1 and α_2, respectively. The relative volume fraction of dimer 1, X_1 is defined as:

$$X_1 = \frac{N_1 V_1}{N_1 V_1 + N_2 V_2},$$

(4.9)
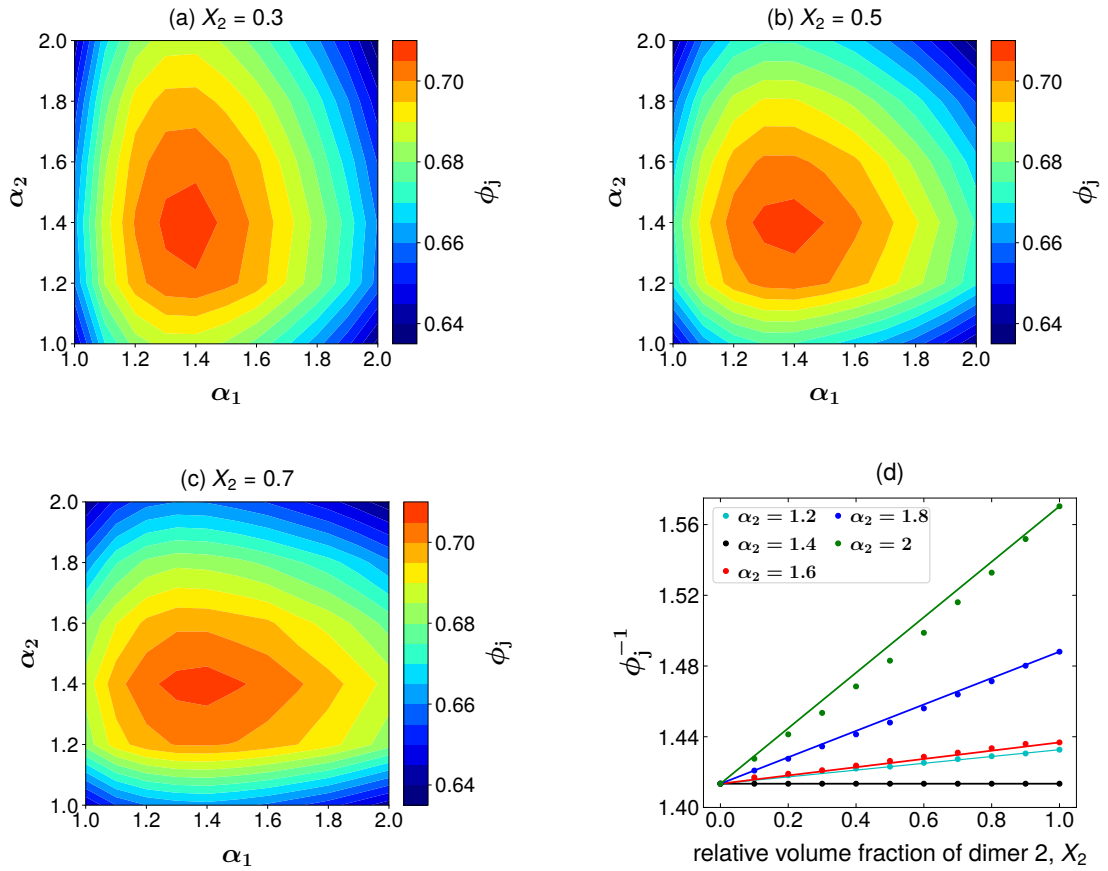
Figure 4.9: The mixing packing density map of binary dimer mixtures for different aspect ratios of the first and the second component, α_1 and α_2, respectively is shown for three volume fractions of the first component (a) $X_2 = 0.3$, (b) $X_2 = 0.5$, (c) $X_2 = 0.7$. Note that figures (a) and (c) contain the same information and each data point is averaged over 10 independent simulation runs, with a standard deviation approximately ± 0.0004. (d) The inverse packing density ϕ_j^{-1} of mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with various aspect ratios α_2 as a function of X_2. Solid lines are obtained by the ideal mixing law Eq. 4.1 without any fit parameter: the lines are fully specified by the packing densities of pure monodisperse dimer 1 and dimer 2 packings (when $X_2 = 0$ and $X_2 = 1$, respectively).

where N_i and V_i (i=1,2) are the number of particles and the volume of the i-th component, respectively. I follow the same steps as in Sec. 4.2 for the packing density calculation. The mixture packing density variation with α_1 and α_2 is shown in a heat map for three different X_2 values in Fig. 4.9(a–c). For all fixed aspect ratios of one component (e.g. α_1), ϕ_j exhibits a non-monotonic relationship with

the aspect ratio of the second species α_2: it increases up to a peak at $\alpha_2 \approx 1.4$ and subsequently decays, whereby the peak always occurs at $\alpha_2 \approx 1.4$ (the aspect ratio at which the maximum packing density of the monodisperse dimer packing has been found in Chapter 3). This behaviour is irrespective of the relative volume fractions of the two components. As in the case of dimer–sphere mixtures, the packing density of binary dimer mixtures never exceeds the maximum density observed for monodisperse dimer packings at $\phi_j = 0.707$ in Chapter 3. These results agree with the findings of disordered packings of binary spherocylinders [53, 95, 96]. They observed a universal density maxima at one unique spherocylinder aspect ratio (1.35-1.5) of one species regardless of the shape of the second component in the system. I also plot the inverse packing density of the mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with various aspect ratios α_2 as a function of X_2 and compare them with the results from the ideal mixing law Eq. (4.1) in Fig. 4.9(d). As can be seen, the curves exhibit a linear relationship with the volume fraction when $\alpha_2 < 2$. However, the inverse density curve is slightly concave-upward, exhibiting systematic deviations from the ideal mixing law for $\alpha_2 = 2$. A statistical analysis confirming the validity of the mixing law for $\alpha_2 < 2$ can be found in Appendix C.1. For $\alpha_2 = 2$, a perfect fit of the data can be obtained using a fourth-order polynomial instead of a linear curve, see Fig. C1(b).

4.3.2 Contact and coordination numbers

For the contact and coordination number analysis, I focus on the mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with various aspect ratios, α_2. I use the same definitions for partial contact and coordination numbers of the two components as in Sec. 4.2. I also measure the overall contact and coordination numbers of the mixtures. For all calculations, I consider only bulk particles.
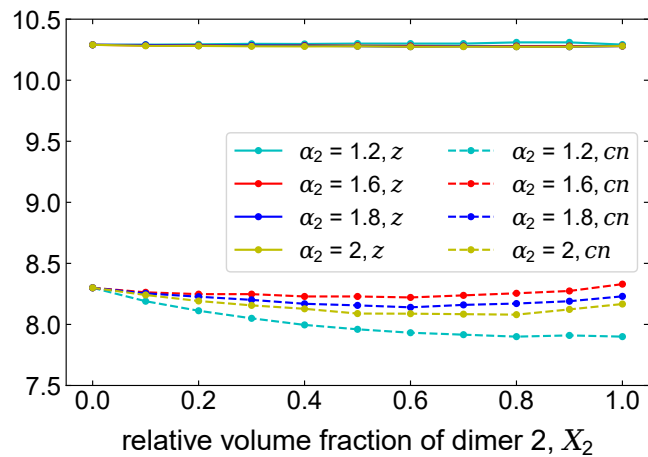
Figure 4.10: The overall contact number z and the overall coordination number cn of the mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with different aspect ratios α_2 vs the relative volume fraction of the second component X_2.
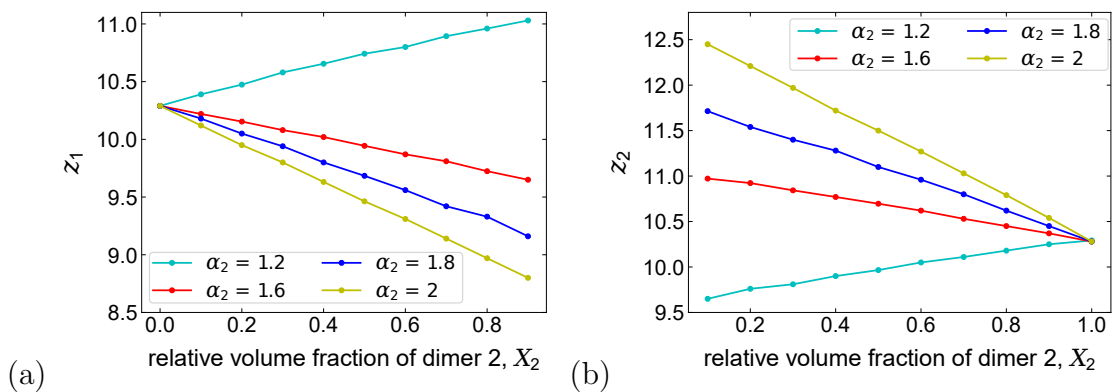
Figure 4.11: The contact number of each dimer species vs X_2. Different α_2 are shown and $\alpha_1 = 1.4$. (a) z_1, (b) z_2.
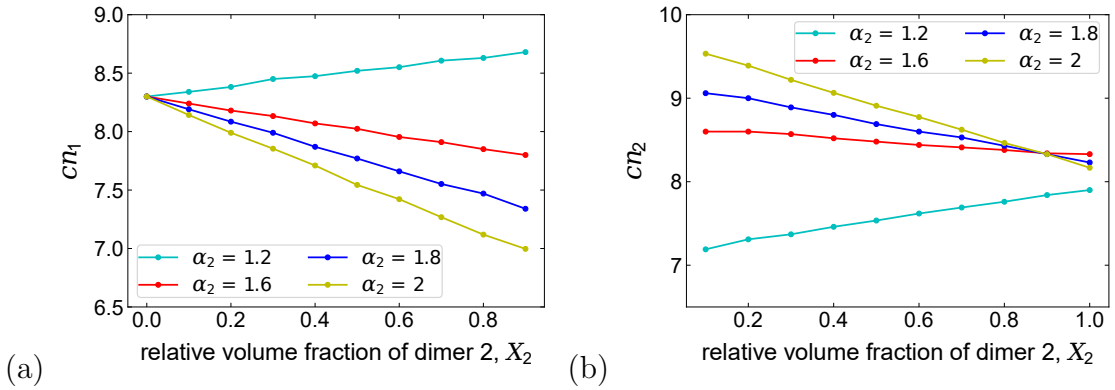
Figure 4.12: The coordination number of each dimer species vs X_2. Different α_2 are shown and $\alpha_1 = 1.4$. (a) cn_1, (b) cn_2.

The overall contact number z and the overall coordination number cn are shown as a function of the relative volume fraction of the second species in Fig. 4.10. As can be seen, the coordination number cn is sensitive to both changes in α_2 and X_2, while z is essentially constant with $z = 10.3$, the same value of monodisperse dimer packings as found in Chapter 3. Therefore, disordered packings of binary dimer mixtures also satisfy the isostatic condition regardless of the relative volume fraction of the components. In this context, I should mention that previous studies of bidisperse and polydisperse sphere packings also found a constant mean contact number, irrespectively of the particle size distribution and the relative amount of different components [129–133]. Although the size effect is the only parameter in the sphere case whereas I consider the joint effect of shape and size in the packings when α_2 is varied, a constant mean contact number seems a generic result of mixing two components having the same isostatic value for their disordered monodisperse packings.

I calculate the contact and coordination numbers of the two components (z_1,

z_2, cn_1 and cn_2) separately and display them as a function of X_2 in Fig. 4.11 and Fig. 4.12, respectively. All these numbers satisfy an approximate linear relationship as in the case of dimer–sphere mixtures. Fig. 4.11 shows that the contact number of the longer dimer species is always hyperstatic with a higher coordination number than in the monodisperse case ($X_2 = 0$), while the shorter dimer species is always hypostatic. Both contact and coordination numbers exhibit a crossover from positive to negative slopes for $\alpha > 1.2$. Comparing with the dimer–sphere case (Fig. 4.5b), it can be seen that dimers mixed with spheres need one to two fewer contacts at each aspect ratio than dimers mixed with another dimer species. On the other hand, the coordination numbers of the dimers are very similar (compare Fig. 4.6b and Fig. 4.12b) indicating that the extra contacts arise from multiple contacts between neighbouring dimer pairs. The coordination number cn_2 exhibits an intersection of the $\alpha_2 > 1.4$ curves for $X_2 = 0.9$, whose origin is not clear.

For binary sphere packings, individual contact numbers of large and small particles have been observed as a combined result of two factors: geometrical and statistical [132]. The contacts of large and small spheres increase as the volume fraction of the former component decreases, governed by geometrical and statistical effects, respectively. To investigate those effects on binary dimer mixtures, I display the variations of the partial contact and coordination numbers with the relative volume fraction of components in Fig. 4.13. Although the number of contacts between the longer dimer species decreases statistically as their volume fraction decreases, their contact number still rises due to the interaction with the shorter ones
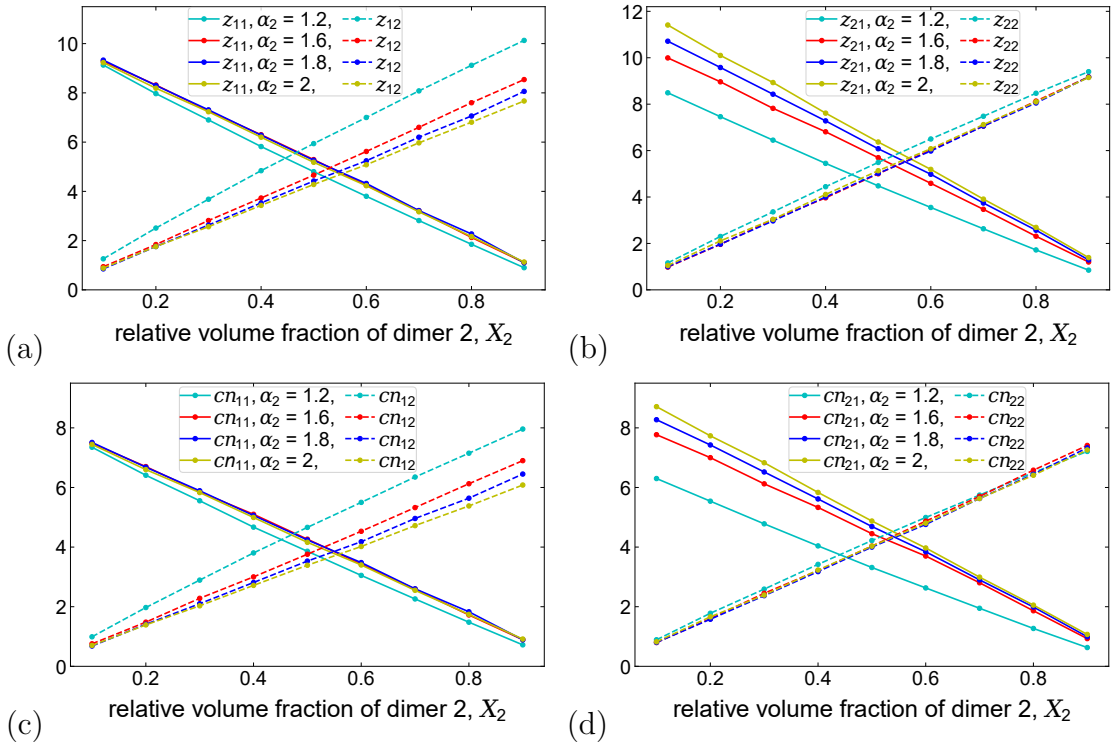
Figure 4.13:   The partial contact numbers in the mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with various α_2 (a) z_{11}, dimer 1-to-dimer 1 and z_{12}, dimer 1-to-dimer 2, (b) z_{21}, dimer 2-to-dimer 1, and z_{22}, dimer 2-to-dimer 2 vs X_2. The four types of coordination number measured in the mixtures (c) cn_{11}, dimer 1-to-dimer 1 and cn_{12}, dimer 1-to-dimer 2, (d) cn_{21}, dimer 2-to-dimer 1 and cn_{22}, dimer 2-to-dimer 2 vs X_2.

(geometrical effect). On the other hand, the statistical effect is more significant for the increase in the contact number of shorter dimer species as the volume fraction of the longer dimers decreases. Both partial contact and coordination numbers measured for the interaction between the same components exhibit an approximately linear relationship with the mixture composition, which is not the case for binary sphere packings [132]. As in the case of dimer–sphere mixtures, the dependence of the contact and coordination numbers of the two dimer species on the aspect ratio (α_2) for the same relative volume fraction (X_2) is primarily due to

the interaction between different species.
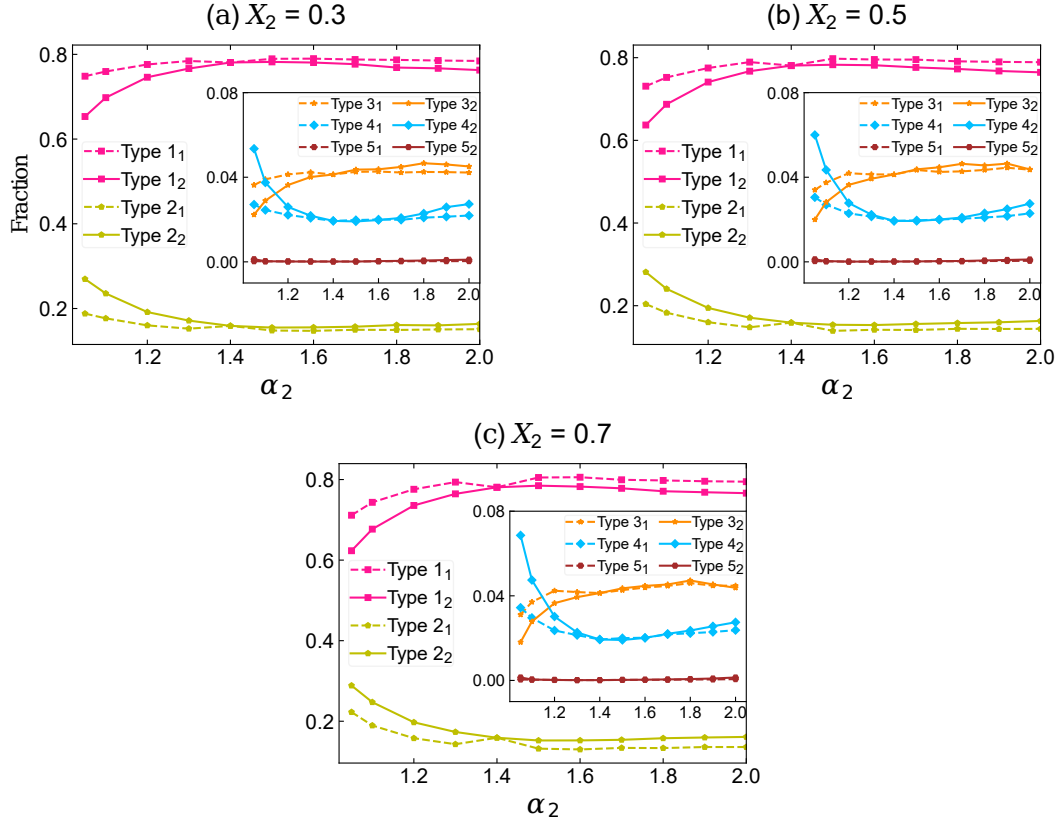
Figure 4.14: The fractions of the five contact configuration types of Table 4-A as a function of α_2 for each component of the binary dimer mixtures with $\alpha_1 = 1.4$: the notation Type 1_2 refers to the fraction of Type 1 contacts on dimer 2, etc. (a) $X_2 = 0.3$, (b) $X_2 = 0.5$, (c) $X_2 = 0.7$.

4.3.3 Contact configurations

I also determine the fractions of the five contact configuration types of Table 4-A for each component (dimer 1 and 2) and display them as a function of the aspect ratio of the second component α_2 for three different volume fractions in Fig. 4.14. Comparing with the dimer–sphere case, the fractions of the dimer 2 contact configurations are very close to those of the dimers in dimer–sphere

mixtures and thus also to those of monodisperse dimer packings. However, a slight dependence on X_2 can be observed, particularly for small α_2. The Type 1/Type 2 fractions of the dimer 1 contact configurations are considerably smaller/larger than those of spheres, but always larger/smaller than those of dimer 2. The curves become slightly flattened when X_2 decreases, i.e., there are in particular more Type 1 and fewer Type 2 configurations for small α_2, which is expected for the dimer 1 fractions, since they need to approach constants when $X_2 \to 0$, but why those of dimer 2 change likewise is unclear. Overall, as in the dimer–sphere and monodisperse dimer case, it can be seen that the fractions show the most considerable variation in the regime $\alpha_2 < 1.4$ and remain approximately unchanged for $\alpha_2 \geq 1.4$.

4.4   Conclusions

In this study, I have shown that the packing densities of dimer–sphere and dimer–dimer mixtures exhibit a non-monotonic variation in the packing density as the aspect ratio of one species changes while keeping the shape of the second component fixed, confirming previous results of spherocylinder–sphere and ellipsoid–sphere mixtures. The findings also confirm the validity of the ideal mixing law Eq. (4.1) for both types of mixtures, highlighting the independence of the packing density on the segregation state. Somewhat surprising is the observation that the packing density of dimer–sphere and dimer–dimer mixtures is always below the maximum packing density of monodisperse dimers (with $\alpha = 1.4$), while bidisperse spheres of

different diameters, e.g., pack denser than monodisperse spheres. However, this behaviour follows from the ideal mixing law since the packing densities of the pure sphere or dimer phases are below the maximum and thus the total packing density as well, see Eq. (4.1). Another manifestation of this ideal mixing property is evident in the linear behaviour of the component-wise contact and coordination numbers, which is not observed for the corresponding total contact and coordination numbers. The analysis of the contact configurations confirms the qualitative picture that the peak in the packing density arises due to the competition of locally optimal rearrangements and excluded volume effects, which is here manifest in the significant variation of the configuration statistics as the dimer is elongated until the maximal packing density is achieved at $\alpha = 1.4$.

In future work, it would be interesting to understand further what kind of observables exhibit similar ideal mixing properties in the mixture. The large parameter space complicates any systematic analysis of mixtures of non-spherical particles. In this study, using the same diameter for spheres that constitute the dimers leads to considering the combined effect of shape and size on the properties of mixtures. Disentangling the effect of shape and size variation as attempted, e.g. in [96] could shed further valuable insight.

Chapter 5

Optimisation of Packing Density by Machine Learning Algorithms

5.1 Introduction

Finding optimum arrangements of particles of a given shape has been a long-standing problem in science since the time of Kepler [6, 134]. It has great importance also in industrial applications of granular materials. For example, there has recently been increased attention to designing new materials by exploiting the complex structures that result from assemblies of particles [2, 46, 61]. The fact that shape is an infinitely variable parameter makes a rigorous systematic exploration of the densest random packings infeasible. Jaeger and collaborators approached this problem by employing a more general shape representation and artificial evolutionary algorithms

to optimise packing densities [2, 45, 68, 99]. They have found maximal packing densities of $\phi_{\mathrm{max}} \approx 0.73$ for trimer-shaped particles.

The difficulties in searching for dense packing shapes arising from the shape space's high-dimensionality can be circumvented by using machine learning methods. Recent developments in algorithms and computational power have made it available to produce large datasets. Machine learning methods have been then applied to large datasets of especially high-dimensional input data for several tasks such as classification, regression, and dimensionality reduction. There has been recently increased attention to employing machine learning algorithms in materials science, especially for tasks where experiments have limitations [135]. For example, experiments require high resources, equipment, and human intuition to discover new materials. Recent developments in computational methods, such as Monte Carlo simulations and molecular dynamics, led researchers to explore design space more efficiently. The combination of experiments and simulations has produced enormous data, making it available to use machine learning algorithms to predict new materials [135], which cuts the time and cost of designing. In the context of granular matter, they have been employed to identify flow defects [136], to reveal correlations between particle size distributions and the mechanical properties [137], and to characterise permeability [138].

In this study, I investigate the feasibility of using machine learning approaches to predict novel particle shapes that achieve high densities when randomly packed. I

use a dataset from Ref. [45] including 5800 distinct particle shapes (each composed of five overlapping spheres) and their corresponding packing densities, i.e. $(\mathbf{X}_i, \phi_i)$ pairs where $\mathbf{X}$ represents a shape vector, as training data. I present a framework that applies a Random forest regressor to the dataset and finds an optimal shape to maximise ϕ_j. In order to facilitate optimising the regression function that yields predictions of new dense packing shapes, the high input dimension of the dataset is first reduced by two different dimensionality reduction methods, Principal component analysis (PCA) and Kernel PCA. Then, the regression model is applied separately to the data represented in the two-dimensional space by PCA and Kernel PCA. Discretizing the plane of the two principal components is used to find new maxima. The validity of predicted shapes is tested with simulations generated by a gravitational pouring protocol in LAMMPS. I also explore the results without the dimensionality reduction, i.e., by fitting a regression function to the dataset in the original high-dimensional input space, and using a gradient-free method, Constrained Optimisation By Linear Approximation (COBYLA) to optimise the regression function.

This chapter is organised as follows. I describe the dataset in detail in Section 5.2. In Section 5.3 and Section 5.4, I explain the dimensionality reduction methods: PCA and Kernel PCA, and the Random Forest regression method, respectively. In Section 5.5, I explain the predictive framework, which is a combination of dimensionality reduction, regression, and optimisation, and present the results. I describe the model without the dimensionality reduction step and show the shape

it predicts in Section 5.6. Finally, I summarise the findings and discuss the future work in Section 5.7.

5.2 Dataset

In this study, I present a framework to predict novel dense packing shapes by applying a Random forest regression model to the dataset from Ref. [45]. Roth *et al.* used an evolutionary algorithm, Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [139], that performs mutations to given shapes until maximal packing densities are achieved. In this algorithm, a population of 40 different particle shapes (each composed of n overlapping spheres) undergoing evolution are represented as a multivariate Gaussian distribution. Representing a compound particle is based on so called blueprint rules that use ordered lists of bearings to specify where to place each sphere [68, 99]. It starts with placing a sphere at the origin. Each bearing is taken from the list and used to draw a ray from the centre of the previous sphere placed along the direction it specifies. Then, another sphere is slid along this ray from infinity towards the origin and placed at the point where it is in contact with the previously built shape. The constituent spheres are placed with this process in sequence. In [68, 99], since the spheres are only allowed to touch, three parameters describe a constituent sphere: two angles that define a bearing relative to the previous sphere placed and a diameter. Roth *et al.* extended these rules by adding a fourth parameter that specifies the degree of overlapping between a sphere and the previous sphere placed. The boundaries on the CMA-ES

search were imposed by fixing a range of all the variables. At each generation in the evolution, 40 disordered packings, one for each particle shape, were simulated by a pouring protocol in LAMMPS in parallel. The packing densities are calculated and ranked to produce the next generation of particle shapes, mutated to explore shape space for potentially higher packing densities. The optimiser converges on an asymptotic value for the packing density over a large number of generations. For particle shapes consisting of $n = 5$ overlapping spheres, the asymptotic solution of packing density has been obtained over 145 generations, where trimer-shaped particles have been observed to achieve the highest packing fraction of $\phi_{\max} \approx 0.73$ [45], see Fig. 5.1. The dataset I use for the predictive framework consists of the shape
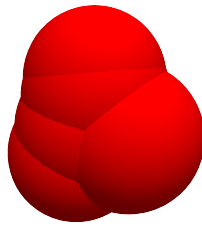
Figure 5.1: The trimer shape that achieved the highest packing fraction of $\phi_{\max} = 0.7367$ in the 5-sphere dataset from Ref. [45].

information of those 5800 (40 shapes that undergo evolution over 145 generations) distinct particle shapes and their corresponding packing densities, i.e. $(\mathbf{X}_i, \phi_i)$ pairs. Note that a 5-sphere particle shape in the dataset is represented in a vector $\mathbf{X}$ that contains the diameters and the spatial coordinates of its constituent spheres, i.e., $\mathbf{X} = (x_1, y_1, z_1, d_1, x_2, y_2, z_2, d_2, ..., x_5, y_5, z_5, d_5)$, where d_a is the diameter of constituent sphere a, so there are in total 20 dimensions. It is possible to characterise a 5-sphere particle shape by only 13 dimensions since there are seven redundant

variables. Fixing the coordinate origin to the centre of the first sphere and measuring the diameters of the other four spheres relative to that of the first sphere reduces four degrees of freedom. There are also reductions due to global rotations of the shape, i.e. the second sphere is aligned along one axis, which reduces two degrees of freedom, and the third sphere of all shapes is oriented in the same plane, which reduces one more degree of freedom, see Fig. 5.2. Hence, it can be then expressed as

$$\hat{\mathbf{X}} = (x_2 - x_1, d_1 - d_2, x_3 - x_1, y_3 - y_1, d_1 - d_3, x_4 - x_1, y_4 - y_1, z_4 - z_1, d_1 - d_4, x_5 - x_1, y_5 - y_1, z_5 - z_1, d_1 - d_5),$$

where the diameter of the first sphere, d_1 is fixed and corresponds to the largest diameter.
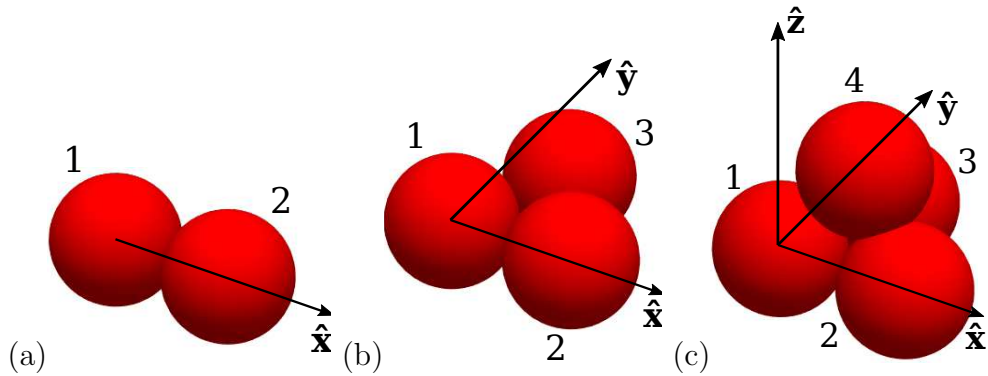
Figure 5.2: Reduction of three degrees of freedom in the description of a 5-sphere particle shape due to global rotations, (a) the second sphere (2) is aligned along $\hat{\mathbf{x}}$ axis, (b) the third sphere (3) is oriented in the $\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$-plane, (c) all the coordinates of sphere 4 (x_4, y_4, z_4) are needed, which is also valid for sphere 5.

5.3   Dimensionality Reduction Methods

Applying machine learning algorithms to high-dimensional datasets has been challenging because the amount of training data needed for a reliable analysis grows exponentially as the number of features increases. Moreover, data analysis tools are

best illustrated in low-dimensional spaces, so interpreting the high-dimensional data by those tools is very difficult. A common approach to overcome the curse of dimensionality is to reduce the data to a low-dimensional space before applying machine learning algorithms. The reduced data should represent sufficient information for a reliable model. There are two types of dimensionality reduction: feature selection and feature extraction. While the former uses a subset of the original features, the latter transforms the original ones via a function (linear or non-linear) to produce a new feature set. Principal Component Analysis (PCA) and Kernel Principal Component Analysis (Kernel PCA) are standard feature extraction methods, which are explained in the following.

5.3.1 Principal Component Analysis

Principal component analysis (PCA) is the most commonly used linear method for dimensionality reduction and feature extraction [140, 141]. PCA transforms the data to a new coordinate system in which the new components are orthogonal to each other and capture as much variance as possible [142]. PCA procedure is explained in the following steps:

1. Given a dataset $\{\mathbf{X}_i\}$, where $i = 1, 2, ..., H$, H is the number of samples and each $\mathbf{X}_i$ is a D-dimensional vector, to project the data onto an f-dimensional subspace, where $f < D$, firstly the centred covariance matrix $\mathbf{C}$ of the dataset

is determined as:

$$\mathbf{C} = \frac{1}{H} \sum_{i=1}^{H} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^{\mathsf{T}}, \tag{5.1}$$

where $\bar{\mathbf{X}}$ is the mean vector, which is D-dimensional vector where each value represents the sample mean of a feature column in the dataset as:

$$\bar{\mathbf{X}} = \frac{1}{H} \sum_{i=1}^{H} \mathbf{X}_i. \tag{5.2}$$

2. Then, the eigenvectors and eigenvalues of the covariance matrix $\mathbf{C}$ are calculated from

$$\mathbf{C}\mathbf{U} = \lambda \mathbf{U}. \tag{5.3}$$

The magnitude of an eigenvalue λ_e corresponding to an eigenvector $\mathbf{U}^e$ equals the amount of variance in the direction of $\mathbf{U}^e$, where $e = 1, 2, ..., D$. The eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_D\}$ are sorted in decreasing magnitude, and the top f eigenvectors that contain most of the information are chosen to reduce the dimension.

3. A $D \times f$ dimensional transformation matrix $\mathbf{G}$ is formed from the top f eigenvectors. The projected dataset $\{\mathbf{X}'_i\}$ on the eigenvector basis are called *principal components* and $\mathbf{X}'_i$ is obtained as:

$$\mathbf{X}'_i = \mathbf{X}_i \mathbf{G}. \tag{5.4}$$

4. Note that $\mathbf{X}_i$ can be approximated from the reduced data as:

$$\widetilde{\mathbf{X}}_i = \mathbf{X}_i' \mathbf{G}^\mathsf{T}.$$ (5.5)

PCA can only capture linear structures in the data. For a more complicated dataset, using PCA might not be suitable. Kernel PCA, on the other hand, allows one to extract nonlinear structures [143]. Kernel PCA first maps the data into some feature space F via a (usually nonlinear) function ψ and then performs linear PCA on the mapped data. As the feature space, F might be very high dimensional, Kernel PCA employs kernel methods instead of carrying out the mapping ψ explicitly [144].

5.3.2 Kernel Principal Component Analysis

5.3.2.1 Feature Extraction

Kernel Principal Component Analysis starts with projecting each data point $\mathbf{X}_i$ in the original D-dimensional feature space to a point $\psi(\mathbf{X}_i)$ in M-dimensional feature space F, where usually $M \gg D$. The projected new features are assumed to have zero mean:

$$\frac{1}{H} \sum_{i=1}^{H} \psi(\mathbf{X}_i) = \mathbf{0}.$$ (5.6)

The covariance matrix of the projected features, $\mathbf{C}_\mathrm{F}$ is $M \times M$ and determined as:

$$\mathbf{C}_\mathrm{F} = \frac{1}{H} \sum_{i=1}^{H} \psi(\mathbf{X}_i)\psi(\mathbf{X}_i)^\mathsf{T}.$$ (5.7)

Then, its eigenvalues and eigenvectors are given by:

$$\mathbf{C}_{\mathrm{F}}\mathbf{V} = \lambda\mathbf{V}, \tag{5.8}$$

where the eigenvectors $\mathbf{V}$ can be expressed with coefficients $\{h_i\}$, where $i = 1, 2, ..., H$ as:

$$\mathbf{V} = \frac{1}{H}\sum_{i=1}^{H} h_i\psi(\mathbf{X}_i). \tag{5.9}$$

By substituting $\mathbf{C}_{\mathrm{F}}$ and $\mathbf{V}$ from Eq. 5.7 and Eq. 5.9, respectively in Eq. 5.8, one obtains the following equation

$$\frac{1}{H}\sum_{i=1}^{H}\psi(\mathbf{X}_i)\psi(\mathbf{X}_i)^{\mathsf{T}}\sum_{j=1}^{H} h_j\psi(\mathbf{X}_j) = \lambda\sum_{i=1}^{H} h_i\psi(\mathbf{X}_i). \tag{5.10}$$

The "kernel trick" is used by defining a function K that calculates the dot product in F as:

$$K(\mathbf{X}_i, \mathbf{X}_j) = \psi(\mathbf{X}_i)^{\mathsf{T}}\psi(\mathbf{X}_j). \tag{5.11}$$

Multiplying both sides of Eq. 5.10 with $\psi(\mathbf{x}_l)^{\mathsf{T}}$ and using the kernel function results in the following equation

$$\frac{1}{H}\sum_{i=1}^{H} K(\mathbf{X}_l, \mathbf{X}_i)\sum_{j=1}^{H} h_j K(\mathbf{X}_i, \mathbf{X}_j) = \lambda\sum_{i=1}^{H} h_i K(\mathbf{X}_l, \mathbf{X}_i). \tag{5.12}$$

It can be expressed with matrix notation as:

$$\mathbf{K}^2\mathbf{h} = H\lambda\mathbf{Kh}, \tag{5.13}$$

where

$$\mathbf{K}_{ij} = K(\mathbf{X}_i, \mathbf{X}_j) \tag{5.14}$$

and $\mathbf{h}$ denotes the column vector with entries $h_1, ..., h_H$. As $\mathbf{K}$ is symmetric, it has a set of eigenvectors which spans the whole space, so all solutions $\mathbf{h}$ of Eq. 5.13 can be found from

$$\mathbf{Kh} = H\lambda\mathbf{h}. \tag{5.15}$$

Note that $\mathbf{K}$ is positive semi-definite, so its eigenvalues are non-negative, and give the solutions $H\lambda$ of Eq. 5.13. To diagonalize $\mathbf{K}$, let $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_H$ denote the eigenvalues, and $\mathbf{h}^1, ..., \mathbf{h}^H$ the corresponding complete set of eigenvectors, with λ_p being the first nonzero eigenvalue. Since the vectors $\mathbf{V}$ in F are needed to be normalized, $\mathbf{h}^1, ..., \mathbf{h}^H$ are also normalized, i.e.

$$\mathbf{V}^e \cdot \mathbf{V}^e = 1 = \sum_{i,j=1}^{H} h_i^e h_j^e (\psi(\mathbf{X}_i)^\mathsf{T}\psi(\mathbf{X}_j)) \tag{5.16}$$

for all $e = p, ..., H$. To extract principal components, one needs to compute projections on the eigenvectors $\mathbf{V}^e$ in F ($e = p, ..., H$). For a test point $\mathbf{X}$, with an

image $\psi(\mathbf{X})$ in F,

$$PC^e(\mathbf{X}) = \mathbf{V}^e \cdot \psi(\mathbf{X}) = \sum_{i=1}^{H} h_i^e K(\mathbf{X}_i, \mathbf{X}) \tag{5.17}$$

are called its kernel principal components. To reduce the dimension of the original dataset $\{\mathbf{X}_i\}$ to a low-dimensional space $(f < D)$, the top f principal components corresponding to the highest eigenvalues are chosen.

If the projected features $\psi(\mathbf{X}_i)$ do not have zero mean, the kernel matrix $\mathbf{K}$ can be centred by using the following equation:

$$\widetilde{\mathbf{K}} = \mathbf{K} - \mathbf{1}_H \mathbf{K} - \mathbf{K} \mathbf{1}_H + \mathbf{1}_H \mathbf{K} \mathbf{1}_H, \tag{5.18}$$

where $\mathbf{1}_H$ is the $H \times H$ matrix with all elements equal to $1/H$. The advantage of kernel methods is that there is no need to compute $\psi(\mathbf{X}_i)$ explicitly. The kernel matrix can be directly obtained from the training data set $\{\mathbf{X}_i\}$. The standard steps of Kernel PCA dimensionality reduction can be summarized as:

1. Construct the kernel matrix $\mathbf{K}$ from the dataset by using a kernel function, Eq. 5.14.

2. Compute the centralized kernel matrix $\widetilde{\mathbf{K}}$ from Eq. 5.18.

3. Use Eq. 5.15 to find $\mathbf{h}$ (substitute $\mathbf{K}$ with $\widetilde{\mathbf{K}}$).

4. Calculate the kernel principal components by using Eq. 5.17 and choose the

top f principal components to reduce the input dimension of the data.

The results of Kernel PCA strongly depend on the choice of the kernel function. The Radial Basis Function (RBF) or Gaussian kernel, one of the most widely used kernels, is defined as:

$$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\eta \|\mathbf{x} - \mathbf{y}\|^2\right). \tag{5.19}$$

It measures the similarity between two vectors based on their Euclidean distance. The parameter η plays a significant role in this similarity, for example the kernel is very sensitive to noise in the data for small η values.

5.3.2.2 Reconstruction of Pre-Images

Producing principal components $PC^e(\mathbf{X})$ using Kernel PCA is generally used for feature extraction and data visualization. To reconstruct the $\psi(\mathbf{X})$ of a vector $\mathbf{X}$ from its projections $PC^e(\mathbf{X})$ onto the first f principal components in F, a projection operator $\mathcal{T}_f$ has been defined in [145] as:

$$\mathcal{T}_f \psi(\mathbf{X}) = \sum_{e=1}^{f} PC^e(\mathbf{X}) \mathbf{V}^e. \tag{5.20}$$

If f is large enough, then $\mathcal{T}_f \psi(\mathbf{X}) \approx \psi(\mathbf{X})$. Most of the time, however, one is interested in reconstruction in the original input space I rather than in F. Since it is difficult to obtain the exact pre-image $\mathbf{X}$, one aims to find an approximation $\widetilde{\mathbf{X}}$ such that

$$\psi(\widetilde{\mathbf{X}}) \approx \mathcal{T}_f \psi(\mathbf{X}). \tag{5.21}$$

A good approximation can be achieved by minimizing

$$\epsilon(\widetilde{\mathbf{X}}) = \|\psi(\widetilde{\mathbf{X}}) - \mathcal{T}_f \psi(\mathbf{X})\|^2. \tag{5.22}$$

The techniques proposed to solve this nonlinear optimization problem often employ gradient descent or nonlinear iteration methods [145, 146]. However, they have some drawbacks, such as being computationally inefficient and having numerical issues. To find $\widetilde{\mathbf{X}}$, Bakir *et al.* developed a technique that learns a pre-image map from the lower-dimensional space to the original input space, i.e. $\Xi^j \colon \mathbb{R}^f \to I$, where $j = 1, 2, ..., D$, by solving the learning problem:

$$\Xi^j = \arg\min_{\Xi^j} \sum_{i=1}^{H} \mathcal{L}(\mathbf{X}_i, \Xi(\mathcal{T}_f \psi(\mathbf{X}_i))) + \Lambda \Delta(\Xi), \tag{5.23}$$

where $\mathcal{L}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$, Δ is a regularizer and $\Lambda \geq 0$ [146]. Considering the learning problem (5.23) as a regression problem for the H points with a kernel K yield a pre-image mapping $\Xi^j(\mathcal{T}_f \psi(\mathbf{X})) = \sum_{i=1}^{H} o_i^j K(\mathcal{T}_f \psi(\mathbf{X}), \mathcal{T}_f \psi(\mathbf{X}_i))$, then coefficients $\{o_i^j\}$ can be calculated by using kernel ridge regression, see [146] for more details. To find pre-image $\acute{\mathbf{X}}$ of a new data on the first f principal components $PC^e(\acute{\mathbf{X}})$, where $e = 1, 2, ..., f$, the learned pre-image mapping from the training dataset can be used, i.e. $\Xi^j(\mathcal{T}_f \psi(\acute{\mathbf{X}})) = \sum_{i=1}^{H} o_i^j K(\mathcal{T}_f \psi(\acute{\mathbf{X}}), \mathcal{T}_f \psi(\mathbf{X}_i))$, where $\mathcal{T}_f \psi(\acute{\mathbf{X}}) = \sum_{e=1}^{f} PC^e(\acute{\mathbf{X}}) \mathbf{V}^e$.

5.4   Random Forests

Random Forests is one of the most commonly used methods in machine learning for classification and regression tasks [147]. A random forest is an ensemble of de-correlated decision trees. A decision tree has a flowchart-like structure that learns simple decision rules from the input data features to predict the value of a target variable. The model is called a classification or regression tree when the outcome takes a discrete set of values or continuous ones, respectively. A tree's decision nodes (features) have branches representing the values for the feature tested, and leaf nodes represent a decision on class labels or numerical values. As the tree goes deeper, decision rules become more complex, and the model gains higher accuracy. However, the deeper trees usually result in overfitting their training datasets, i.e. have low bias but high variance. Breiman *et al.* introduced a bootstrap aggregating or bagging technique, which is an ensemble of trees to reduce the variance [148]. It repeatedly trains a classification or regression tree on a random sample with replacement of the training set. Then, predictions for a new, unseen sample are made by averaging the predictions from all the individual trees on the sample for the regression and allowing the trees to vote for the most popular label for the classification. Ho *et al.* improved the bootstrapping method by letting trees grow in a random subset of the features [149, 150] and used the "Random Forest" term for the first time in [149]. Other attempts have been shown to improve the bagging method, such as by using additional randomization [151]. Breiman *et al.* provided a substantial definition of Random Forests in [152]. The trees are

decorrelated by a learning process through random selection of input features in Random Forests. The nonlinear and non-overfitting nature gives power to Random Forest by exhibiting good performance in machine learning tasks. One drawback of Random Forests for regression tasks is the lack of extrapolation since it can only predict the values between the minimum and maximum of the training dataset. The algorithm for a Random Forest regression model is summarized as follows:

Algorithm 1 Random Forest for Regression

---

1. For $b = 1, ..., B$
   - Generate a random sample $\mathbf{Z}^*$ with replacement of the training dataset.
   - A Random Forest tree T_b to the sample $\mathbf{Z}^*$ is grown by recursively repeating the following steps on the non-leaf branches until the maximum depth of the tree is reached.
     - Choose l_b features at random from the t features.
     - Pick the best feature among the l_b for splitting.
     - Divide the dataset into subsets based on values of the selected feature.
   - Calculate the average of the subset for the leaf node, i.e. when there is no need for further splitting.
2. Prediction for a new point x' is made: $\mu(x') = \frac{1}{B} \sum\limits_{b=1}^{B} T_b(x')$

5.5 Predictive Framework

In this section, I introduce a framework including dimensionality reduction, regression and optimisation of the regression function to predict novel dense packing shapes. The dimensionality reduction step is employed to facilitate the optimisation step, which is challenging in a high-dimensional shape representation due to the difficulty of converging to a global maximum and implementing constraints to obtain physical shapes only; see Section 5.6 for further information. Moreover, representing

the data in lower dimensional space makes data visualisation and interpretation easier. The steps in the framework that uses the 5-sphere dataset from Ref.[45] $(\mathbf{X}_1, \phi_1), ..., (\mathbf{X}_{5800}, \phi_{5800})$ based on the 20-dimensional shape representation $\mathbf{X}$ as a training data can summarised as:

1. Applying dimensionality reduction on the dataset $\{\mathbf{X}_i\}$, where $i = 1, 2, ..., 5800$ to obtain two-dimensional input data, the principal components, $\{\mathbf{PC}_i\}$ for the regression.

2. Fitting a regression function to the training data $(\mathbf{PC}_1, \phi_1), ..., (\mathbf{PC}_{5800}, \phi_{5800})$.

3. Optimising the packing density that the regressor predicts.

4. Applying the inverse dimensionality reduction to identify the shapes corresponding to the maximal densities.

5. Testing the optimal shapes with the same simulation algorithm as the training data.

5.5.1 Implementation of the framework in Python

PCA and Kernel PCA

I use standard PCA and Kernel PCA classes implemented by the Scikit-learn [153] package in Python for dimensionality reduction of the 5-sphere dataset $\{\mathbf{X}_i\}$, where $i = 1, 2, ..., 5800$, 5800 is the number of samples and each $\mathbf{X}_i$ is a 20-dimensional vector as described in Section 5.2. Before running PCA and Kernel

Table 5-A: Performance of Kernel PCA with the RBF kernel by varying the η parameter.

	η					
	0.002	0.02	0.2	2	20	200
Average MSE	2.14	1.77	2.1	2.25	2.39	2.4

PCA, the dataset is standardised to prevent any bias for some variables. For Kernel PCA, I use the RBF kernel and determine the parameter η with GridSearchCV, which is a hyperparameter optimisation technique implemented in Python. The computer compares the performance for different values for a hyperparameter with k-fold cross validation and a scorer. For the parameter η, I define a scorer that calculates the mean squared error (MSE) of the reconstruction of the original input data from the reduced one and uses 3-fold cross-validation. The average MSE of three folds for each η value is displayed in Table 5-A. The minimum mean MSE is considered the best score, so $\eta = 0.02$ is chosen for the Kernel PCA model. I compare the principal components, $\{\mathbf{PC}_i\}$, where $i = 1, 2, ..., 5800$ and each $\mathbf{PC}_i$ is a two-dimensional vector, extracted from the two methods in the following.

Random Forest Regressor

The dataset $(\mathbf{PC}_1, \phi_1), ..., (\mathbf{PC}_{5800}, \phi_{5800})$ based on the two-dimensional shape representation $\mathbf{PC}$ is trained on a Random Forest regressor in Scikit-learn. Note that the regression is applied separately to the reduced data by PCA and Kernel PCA. I use 100 decision trees for training, determined with GridSearchCV by using a 3-fold cross-validation and the coefficient of determination as a scorer. I select mean-squared error as its objective or function, which needs to be minimized and

is equal to variance reduction as a feature selection criteria. Since there are only two features, the maximum number of features (l_b) when searching for the best split is chosen as 2. The trees go deeper until all leaves contain two samples.

Optimisation

In order to find the maximal densities, I first discretise the plane of the principal components (PC^1,PC^2) and then allow the Random forest regressor to predict the packing densities for the regular grid points on the plane. Discretising is done for different grid spacings to observe the predicted density as a function of the grid size. For each grid size, the grid point corresponding to the maximum is transformed back to the original high dimensional space (20-dimensional shape representation) to visualize the shape. For PCA, the mapping from the reduced data to the original input space is achieved by the transformation matrix in Eq. 5.5. Scikit-learn uses a learned pre-image mapping as described in Section 5.3.2.2 for the inverse transformation in Kernel PCA. Finally, the validation of the predicted shapes is tested by generating simulations for disordered packings with the pouring protocol in LAMMPS.

5.5.2   Principal Components

In PCA, explained variance for each eigenvector is the ratio of related eigenvalue and sum of eigenvalues of all eigenvectors, i.e. $\frac{\lambda_e}{\lambda_1+\lambda_2+...+\lambda_D}$, where $e = 1, 2, ..., D$. The explained variance ratio then represents the variance captured by a particular eigenvector. I display the individual and cumulative explained variance with the

number of reduced dimensions (principal components) for the 5-sphere dataset in
Fig. 5.3(a). As can be seen from Fig. 5.3(a), PCA captures all the information
already with 16 dimensions, so it predicts that four variables are redundant. How-
ever, it does not detect global rotations, which reduce three more variables. The
first two principal components contain 44 % of the information. The explained
variance ratio can also be calculated for Kernel PCA, although the interpretation of
it is not as straightforward as for PCA. As can be shown in Fig. 5.3(b), there are in
total 5800 (the number of samples in the dataset) principal components for Kernel
PCA, and the first two principal components capture 35 % of the information in
feature space.
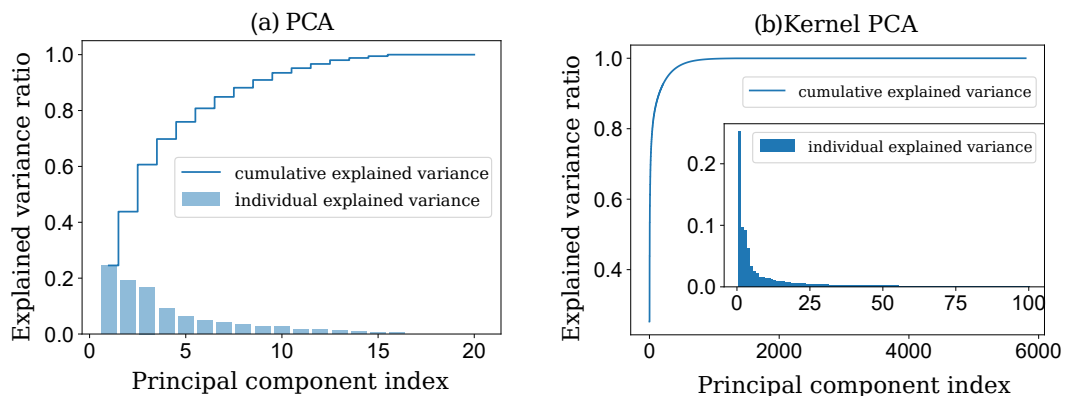
Figure 5.3: Explained variance ratio *vs.* the number of reduced dimensions for the
5-sphere dataset from Ref. [45], individual and cumulative explained variance for
(a) PCA, (b) Kernel PCA.

The first two principal components (PC^1, PC^2) of 5800 samples obtained by
PCA are shown in Fig. 5.4(a). Furthermore, the packing density corresponding to
the first principal component for each sample is displayed in Fig. 5.4(b). As can be
seen, a classification between high and low-density regions is unclear for the PCA

results. On the other hand, Kernel PCA provides a good separation of samples
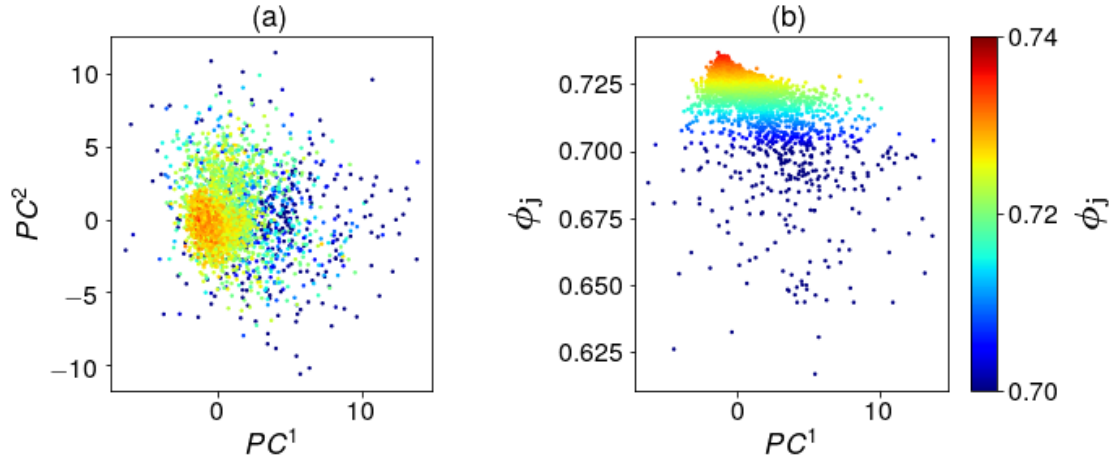
Figure 5.4: (a) The first two principal components (PC^1,PC^2) of 5800 samples obtained by PCA, (b) The corresponding packing density ϕ_j *vs.* the first principal component PC^1.
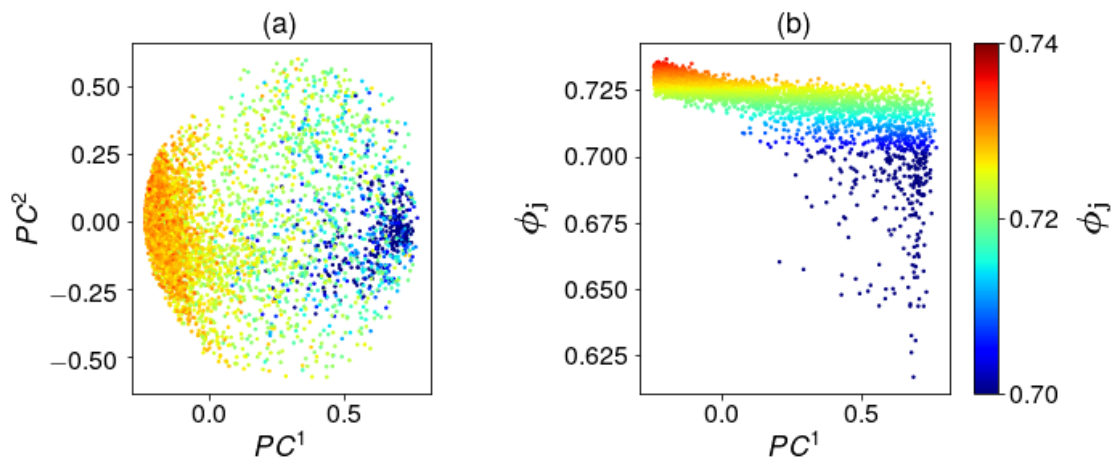
Figure 5.5: (a) The first two principal components (PC^1,PC^2) of 5800 samples obtained by Kernel PCA, (b) The corresponding packing density ϕ_j *vs.* the first principal component PC^1. Two classes with high and low densities, coloured in red and blue, respectively, are observed.

with high (red coloured) and low (blue coloured) packing densities, see Fig. 5.5(a),

so the pairs of principal components for the two groups can be easily distinguished.

Moreover, compared to PCA, Kernel PCA produces results that can better fit a
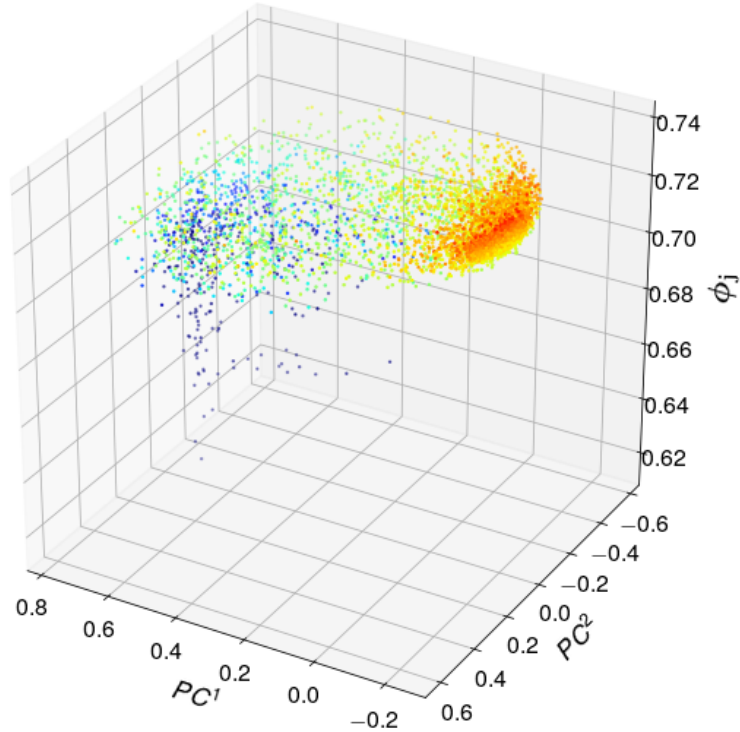
Figure 5.6: The corresponding packing densities ϕ_j *vs.* the first two principal components (PC^1,PC^2) of 5800 samples obtained by Kernel PCA. The two classes with high and low densities, coloured in red and blue, respectively, can be clearly seen.

regression model, as can be seen in Fig. 5.4(b) and Fig. 5.5(b). The two classes with high and low densities can be clearly seen in Fig. 5.6. In PCA, the principal components are linear combinations of all the variables. For Kernel PCA, on the other hand, they are obscure since the non-linear mapping function is not known.

5.5.3   Predicted Shapes

I run a Random forest regressor on the reduced dataset $\{\mathbf{PC}_i\}$, where $\{i = 1, 2..., 5800\}$ and each $\mathbf{PC}_i$ is a two-dimensional vector with the two principal com-
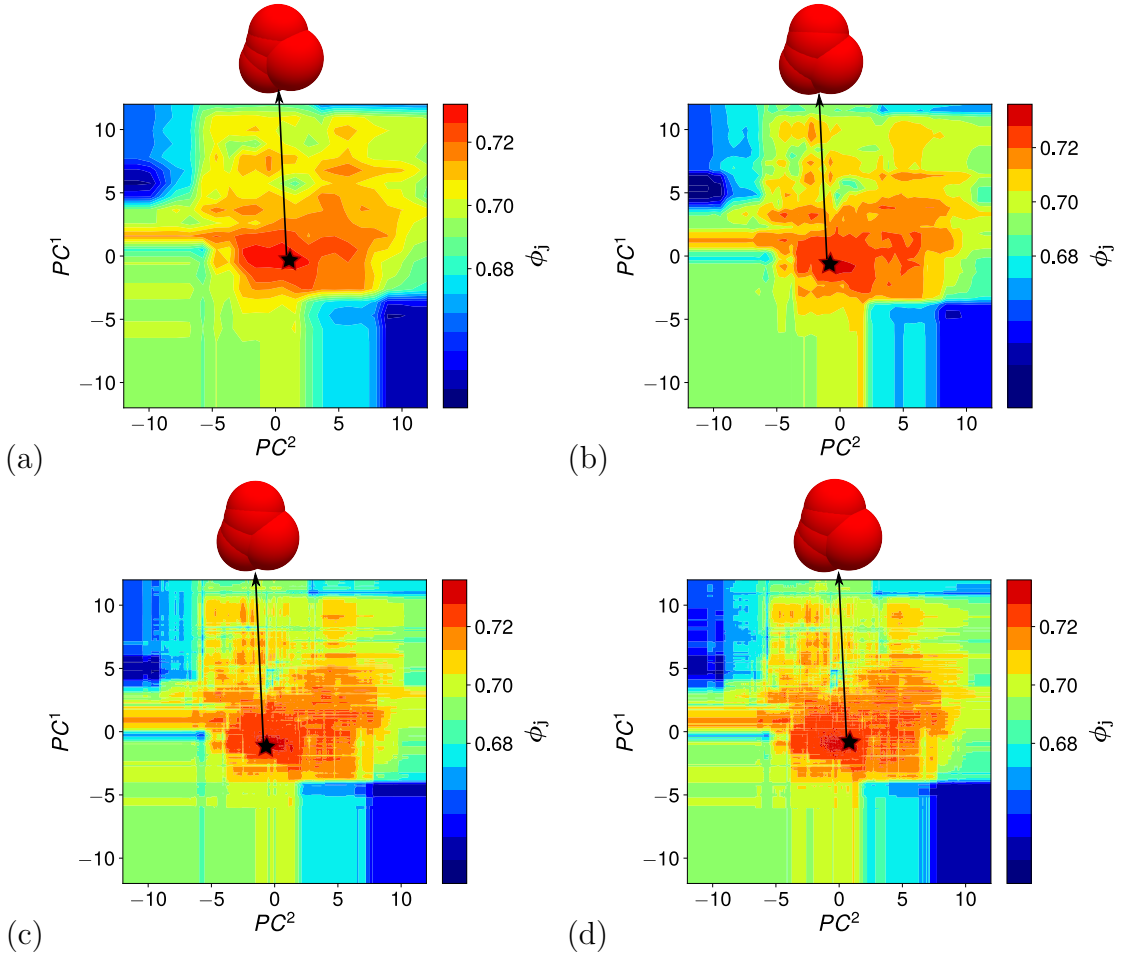
Figure 5.7: Predicted packing densities ϕ_j for regular grid points on the plane of the principal components found by PCA, where $PC^1 \in [-12, 12]$, $PC^2 \in [-12, 12]$, (a) Grid size=1, (b) Grid size=0.5 (c) Grid size=0.05, (d) Grid size=0.001. The star indicates the point corresponding to the densest packing shape predicted for each grid size.

ponents, PC^1 and PC^2. The model is applied separately to the two reduced datasets by PCA and Kernel PCA. The regressor is allowed to predict packing densities for regular grid points on the plane of the principal components, where $PC^1 \in [-12, 12]$, $PC^2 \in [-12, 12]$ for PCA and $PC^1 \in [-0.8, 0.8]$, $PC^2 \in [-0.8, 0.8]$ for Kernel PCA. To see the dependence of the predicted maximal density on the grid spacing, I use different step lengths, see Fig. 5.7 and Fig. 5.8. For each grid spacing, I find the
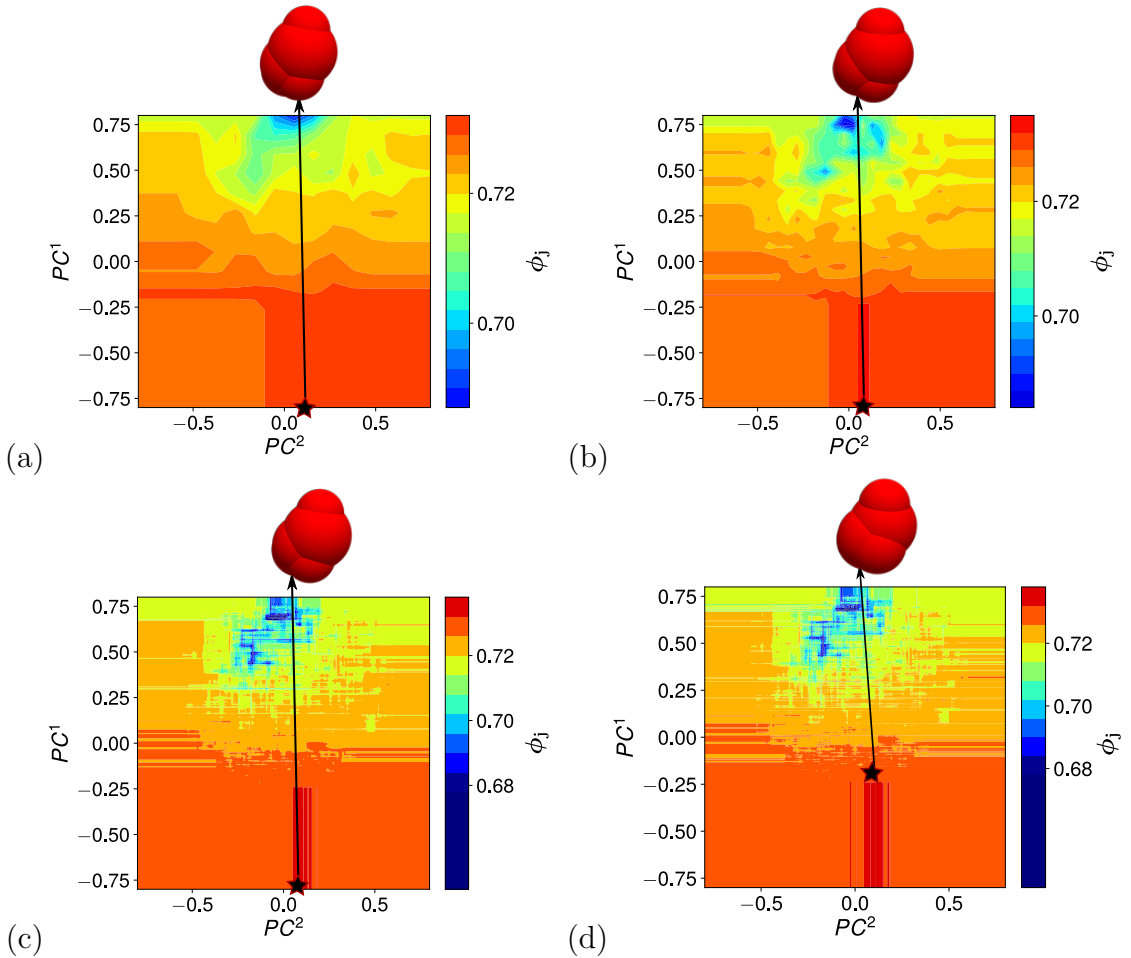
Figure 5.8: Predicted packing densities ϕ_j for regular grid points on the plane of the principal components found by Kernel PCA, where $PC^1 \in [-0.8, 0.8]$, $PC^2 \in [-0.8, 0.8]$, (a) Grid size=0.1, (b) Grid size=0.05 (c) Grid size=0.005, (d) Grid size=0.0001. The star indicates the point corresponding to the densest packing shape predicted for each grid size.

grid point corresponding to the maximal density and identify the particle shape representing that point in the original high-dimensional space. To construct the particle shape from the two-dimensional space, while PCA uses a transformation matrix in Eq. 5.5, Kernel PCA employs a learned pre-image mapping by solving the problem Eq. 5.23. Since the data is standardised at the beginning, the pre-images obtained after transformation are needed to scale back to the coordinates and

diameters. The predicted densities for different grid spacings and the shapes that correspond to the maximal densities are shown in Fig. 5.7 and Fig. 5.8 for the regression models applied to the reduced data by PCA and Kernel PCA, respectively. For both models, the grid point corresponding to the predicted maximal density approaches the point ((-1.366, 0.442) for the PCA data, (-0.196, 0.1365) for the Kernel PCA data) with $\phi_{\max} = 0.7367$ as the grid size decreases. For the model using the PCA data, the regressor predicts the trimer shape, the densest packing shape found in the dataset, with slight variations in the grid spacing, see Fig. 5.7. On the other hand, novel shapes are predicted for the model using the Kernel PCA data, see Fig. 5.8.

5.5.4 Validation of the Predicted Shapes

To test the validation of the predicted shapes, I generate simulations for disordered packings of those shapes in LAMMPS. For the simulations, I use the same packing-generation protocol in Chapter 3 for disordered dimer packings. Note that Roth *et al.* also used the same protocol to generate simulations for the 5-sphere particle packings in the dataset. I calibrate simulation parameters by ensuring to achieve the same packing density for the trimer shape, the densest packing shape in the dataset. Then, I use the same material parameters in the simulations to generate packings for the predicted shapes. The number of particles and material parameters can be found in Chapter 2. As in Section 3.2, I determine a bulk region by excluding the particles within $5 - 8d$ from the container floor and $5d$ from the upper-most

particles, where d is the diameter of the largest sphere of a 5-sphere particle, to calculate the packing density. The Voronoi volume W_i of a 5-sphere particle i is found by summing the Voronoi volumes of its five constituent spheres. I determine the bulk volume V_b occupied by N_b 5-sphere particles in the bulk by summing their individual volumes, i.e. $V_b = \sum_{i=1}^{N_b} W_i$. I then obtain the packing density as $\phi_j = N_b V_{fs}/V_b$, where V_{fs} is the volume of a 5-sphere particle, calculated by using the Monte Carlo method; see Appendix D.1. Note that the Voro++ package in LAMMPS uses the radical tessellation, also called the power diagram, for Voronoi volume calculations when spheres in the packing have different diameters [81]. The bisecting plane in the conventional Voronoi tessellation is replaced by the radical one, which collects all the points having the same power for the two spheres [154]. The power of a point with respect to a sphere is the square of the length of a line segment from that point to its tangent point with the sphere. A point is then counted inside the associated Voronoi volume of the sphere, which minimises its power. This tessellation causes an error in the packing density calculations due to incorrect Voronoi volumes of the particles that cross the boundaries of the bulk region; however, this error should be small since the number of bulk particles is high. The correct Voronoi cell of a given particle collects all points closer to that particle's surface than to the surface of any other in the packings; see Appendix A.1 for the VB construction between two spheres of unequal radii.

Plotting the densities as a function of the grid size, I compare the maximal packing densities predicted by the Random forest regressor and the simulation
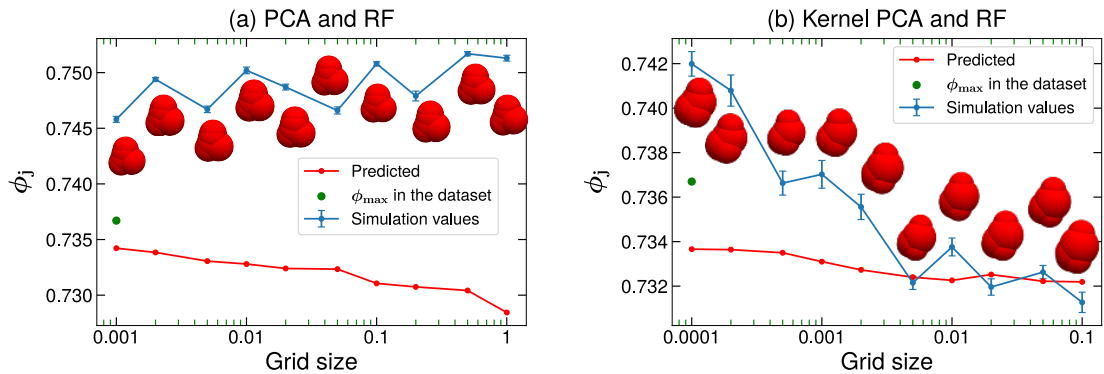
Figure 5.9: The predicted packing densities and the simulation results for the optimal shapes obtained by the Random Forest (RF) regressor applied to the reduced data *vs.* Grid size. The simulation values are averaged over 5 realizations. The results are shown for the data reduced by (a) PCA, (b) Kernel PCA. The green data point indicates the maximal packing density ($\phi_{\max} = 0.7367$) in the dataset achieved for the trimer shape. While the regressor improves the density by varying the trimer shape slightly for the PCA data, novel dense packing shapes are predicted for the Kernel PCA one.

results for the predicted shapes in Fig. 5.9. As can be seen from Fig. 5.9(a) and Fig. 5.9(b), the dependence of the predicted densities on the grid size is similar for the two models applied to the data reduced by PCA and Kernel PCA, respectively. For both reduced data, the regressor predictions increase and approach the maximum in the dataset ($\phi_{\max} = 0.7367$) as the grid size decreases. On the other hand, the predicted shapes and their packing densities calculated from the simulations are different for the two regression models. The simulation values follow a similar trend to the predictions with the grid size for the data obtained by Kernel PCA, whereas they exhibit slight improvements as the grid size increases for the PCA data. While the regressor predicts the trimer shape as the optimal one for the PCA data, with slight variations in the grid size, novel shapes are found for the Kernel one, see Figs. 5.9(a,b). Moreover, the changes in the predicted shape

with the grid size are more dramatic for the regression model applied to the Kernel PCA data. Overall, the results highlight that the predicted shapes for both models achieve higher packing densities in the simulations than the predicted values and the maximum in the dataset ($\phi_{\max} = 0.7367$). For example, the trimer shapes can reach maximal densities as high as $\phi_j \approx 0.75$, and the predicted shape for the smallest grid size by the regressor for the Kernel PCA data packs the densest with $\phi_j \approx 0.742$. The predicted densities always stay below the simulation values because the Random forest regressor can not extrapolate, i.e. it can not predict values greater than the maximum in the training dataset.

5.6 Regression and Optimisation in the High-Dimensional Shape Space

I explore the results without the dimensionality reduction step, i.e., applying a Random regressor to the training data in the high-dimensional shape space, $(\hat{\mathbf{X}}_1, \phi_1), ..., (\hat{\mathbf{X}}_{5800}, \phi_{5800})$, and optimising the regression function. Note that the optimisation of the regression function requires a gradient-free method since the Random forest regression is non-parametric. Moreover, since the optimisation is performed in the high-dimensional shape space, the method should also allow implementing constraints to obtain physical shapes only. When considering the space of shapes composed of five overlapping spheres, valid ones are constrained to the range of values for which two spheres placed consecutively overlap, imposing

non-linear constraints; see the blueprint rules to obtain a 5-sphere shape explained in Section 5.2. Like the dataset, the centre of the first sphere is placed in the origin, and its diameter is fixed. The diameters of the other four constituent spheres are constrained to range from zero to the diameter of the first one. Hence, to reduce the number of constraints for the optimisation, I use 13-dimensional shape representation $\hat{\mathbf{X}}$, rather than the 20-dimensional vector $\mathbf{X}$ since it contains the same information. Note that the 13-dimensional shape representation $\hat{\mathbf{X}}$ is obtained by hand.

The dataset $(\hat{\mathbf{X}}_1, \phi_1), ..., (\hat{\mathbf{X}}_{5800}, \phi_{5800})$ based on 13-dimensional shape representation, $\hat{\mathbf{X}} = (x_2 - x_1, d_1 - d_2, x_3 - x_1, y_3 - y_1, d_1 - d_3, x_4 - x_1, y_4 - y_1, z_4 - z_1, d_1 - d_4, x_5 - x_1, y_5 - y_1, z_5 - z_1, d_1 - d_5)$ is trained on a Random Forest regressor in Scikit-learn. I use 100 decision trees for training and set the maximum number of features $l_b = 13$, determined with GridSearchCV by using a 3-fold cross-validation and the coefficient of determination as a scorer. I select mean-squared error as its objective or function. The trees go deeper until all leaves contain two samples. The optimisation is also performed in Python, by using the Constrained Optimization BY Linear Approximation (COBYLA) algorithm built-in SciPy [155] to find a maximal density. This method is gradient-free and also allows the implementation of constraints required to obtain valid shapes, but is not guaranteed to converge to a global maximum. The resulting densest packing shape that corresponds to the maximum is displayed in Fig. 5.10. The density is measured from the LAMMPS simulations for this shape as $\phi_j = 0.741$ with an error bar of ≈ 0.0004, higher than
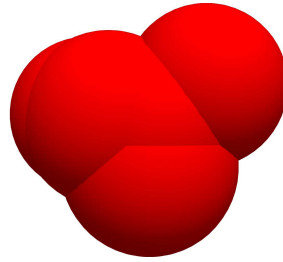
Figure 5.10: The optimal shape obtained by the regression and optimisation performed in the high-dimensional shape space. The Constrained Optimization BY Linear Approximation (COBYLA) algorithm built-in SciPy is used for the optimisation. The density of this shape is determined from the simulations as $\phi_j = 0.741$, with an error bar of ≈ 0.0004.

the predicted density value 0.729.

5.7 Conclusions

In this chapter, I have examined the applicability of machine learning algorithms

to predict novel dense packing shapes. I have developed a predictive framework

by applying a Random Forest regressor to the dataset of 5-sphere particle packing

densities. The input dimension is reduced before running the regressor to facilitate

the optimisation of the regression function and to better visualise the results. Two

different dimensionality reduction methods have been employed here, PCA and

Kernel PCA. Reducing the data to two dimensions, Kernel PCA has shown better

performance in classifying the data into high and low packing densities. The

regressor has been applied to the two reduced data separately, resulting in different

predicted shapes. The packing densities are estimated from simulations for the

disordered packings of those shapes to test their validity. The results show that the

Random Forest regressor predicts dense packing shapes, although the predictions

of the regressor are not highly accurate. Interestingly, the novelty of predicted shapes has been observed as strongly dependent on the dimensionality reduction methods. While the regressor applied to the PCA data slightly changes the trimer shape, the densest packing shape in the dataset, to maximise the packing density ϕ_j, it predicts novel shapes for the Kernel PCA data. This result is because PCA and Kernel PCA use different ways to reconstruct the pre-images from the reduced data.

I have also performed the regression and optimisation in the high-dimensional shape space. The optimal shape found by this method packs more efficiently than the densest packing shape in the dataset; however, its packing density does not reach as high as the densities of the shapes obtained by the framework consisting of the dimensionality reduction step. This result might imply that the dimensionality reduction aids in searching for a global maximum. To further understand this, observing the prediction variations with the number of reduced dimensions might be tempting. I have found empirically that the optimisation in the reduced space does not require any additional constraint, e.g. the overlap constraint, and always leads to overlapping particle shapes.

In this study, I have used only one regression method, which has some drawbacks, i.e. it can not predict values greater than the maximum in the training dataset. To find more distinct shapes, applying different regression methods and comparing the predictions would be tempting. The training dataset used here is itself the

output of an optimisation algorithm, so the shapes in the dataset are not uniformly distributed in the shape space and are already clustered in the region of high values of the packing densities. Nevertheless, the machine learning approach performed here has considerably improved the packing densities found in [45]. To further verify and investigate the dense packing shapes found in this study, it would be interesting to employ other simulation methods using the gravitational packing protocol and also experiments to produce disordered packings of these shapes. Overall, using machine learning algorithms in conjunction with simulations can help explore dense packing shape space and also shed light on understanding the relationship between particle shape and packing density.

Chapter 6

Conclusions and future work

This thesis investigates the shape effect on the structural properties of disordered particle packings based on a particle representation model of overlapping spheres. I have generated disordered packings of frictionless symmetric dimers simulated by the gravitational pouring protocol in LAMMPS and studied their properties at the microscopic and macroscopic levels. I also have examined whether machine learning algorithms can be applied to identify novel dense packing shapes. The key contributions of this thesis are summarised as follows:

1. The characteristic peak of the packing density at $\alpha_{\mathrm{max}} \approx 1.4 - 1.5$ observed previously for rotationally symmetric elongated shapes by numerical, experimental and theoretical work has been validated for dimer packings. Investigating the contact number and several order metrics has revealed that the formation of this peak is accompanied by significant microscopic rearrange-

ments between neighbouring dimers, increasing nematic order and decreasing local translational order. The results imply that the fall in the packing density for larger aspect ratios ($\alpha > \alpha_{\mathrm{max}}$) is due to the increased excluded volume effects.

2. Investigations of binary mixtures of dimers are in line with the findings for binary mixtures of spherocylinders. They possess a universal peak in the packing density at a unique aspect ratio, irrespectively of the variation in shape or mixture composition. They also satisfy a remarkable empirical ideal mixing law, which states that the total packing volume is independent of the segregation state. The microscopic rearrangements observed up to α_{max} for the mixtures of dimers highlight the qualitative explanation of the density maxima that manifest competition between local caging and excluded volume effects.

3. Two dimensionality reduction methods, PCA and Kernel PCA, and a Random Forest regressor have been applied to the 5-sphere particle packing dataset to identify new shapes of maximal packing densities. The results have shown that the regressor has found optimal shapes, and the novelty of predicted shapes depends on the dimensionality reduction method. The simulation values are higher than the predicted densities by the regressor.

Future Work

In this thesis, I have used a specific packing protocol for the simulations, so the results on ordering in the dimer packings might be due to gravity. Therefore, it is necessary to investigate ordering for packings generated with other simulation methods or experiments. In the context of order, it should be noted that the results strongly depend on the metrics. The order metrics measured here, such as bond orientational order parameters and radial distribution function, have been chiefly used for quantifying order in spherical particle packings. Hence, devising incisive order metrics for packings of non-spherical particles is essential.

In this study, I have investigated only the effect of the particle shape on the dense packings, so I have not considered the effect of friction on the dimer packings. It is also crucial to study disordered packings of frictional dimers in the future. Previous work on sphere packings showed that introducing friction results in lower densities, and it reduces the contact number required for mechanical stability, i.e., $z = d_{\mathrm{f}} + 1$. These results also hold for packings of elongated shapes where a frictional contact can constrain both the translational and rotational degrees of freedom of a particle at the contact point. Delaney *et al.* extended the concept of the random loose packing limit ("RLP") defined for spheres to disordered packings of ellipsoidal particles simulated by settling the particles into a viscous liquid [48]. Their findings highlighted that this settling technique causes particles to have a high orientational ordering, and introducing the friction causes a decrease in this

ordering, but a considerable amount still remains. Hence, they discussed that the applicability of sedimentation or pouring techniques is problematic in the definition of "RLP" of non-spherical particles. Investigation of the effect of the friction on the disordered packings of dimer particles with metrics measured in this thesis can shed light on understanding the microscopic structures of the loose packings of elongated shapes, and hence the concept of "RLP". I expect the microscopic rearrangements between neighbouring particles observed at small aspect ratios would vanish due to the reduction of the contact number to satisfy mechanical stability, and the slight oblate ordering would decrease in the presence of friction. The dependence of the packing density on the aspect ratio would not be significant for small aspect ratio regime. However, the packing density still would decrease for large aspect ratios due to the excluded volume effects. For binary mixtures, investigating individual shape and size effects would be interesting.

In Chapter 5, most particles in the dataset used for regression have been found in the same class (trimer-shaped) of high packing densities, which are already optimised. When using PCA to reconstruct the shape from the reduced data, the Random Forest regressor has identified dense packing shapes near the trimer-shaped class. Hence, the results are dependent on the methods as well as the quality of the training dataset. It would be interesting to produce a more extensive dataset of distinct particle shapes with a wide range of packing densities, which can then be used to explore shape space by applying different regression methods.

Appendix A

Voronoi Boundary Construction for Asymmetric Dimers

The complexity that arises due to particle geometry in granular matter leads to a lack of a unified theoretical framework to predict packing densities. A mean-field theory has been developed to estimate packing densities of both spheres and anisotropic particles with rotational symmetry in [29, 30, 34]. The results for dimers, spherocylinders, and lens-shaped particles have shown a good agreement with the highest densities obtained by empirical studies [34]. The authors have discussed that this framework can be suitable for other shapes such as tetrahedra, trimers and irregular polyhedra. Here, I investigate whether the mean-field approach can be extended to more general shapes, focusing on asymmetric dimers.

The theory relies on a statistical mechanics framework proposed by S.F. Edwards

to treat granular materials [5]. He postulated that all microscopic jammed states at a fixed total volume of granular systems are equally probable, and macroscopic observables can be calculated by taking averages over them [4]. Moreover, the volume is the crucial macroscopic quantity in granular matter, analogous to the energy in thermal systems [5]. A volume function is then necessary to define the total volume as a function of the positions and orientations of particles in the granular system [3]. Voronoi tessellation is a convenient way to partition the total volume V of into non-overlapping cells associated with each particle in the system [156, 157].

A.1   The Voronoi Volume

The Voronoi volume of a particle i, W_i contains all points closer to this particle than any other one. The packing density of N monodisperse particles can be calculated as $\phi = NV_0/V$, where V_0 is the particle volume and $V = \sum_{i=1}^{N} W_i$. The determination of W_i requires finding the Voronoi boundary between two particles i and j, which is defined as the hypersurface consisting of all points that are equidistant to the particles' surfaces [5, 34, 158, 159]. Fixing the coordinate system at the centre of mass of particle i and assuming its orientation is also fixed, the VB can be parametrized in terms of a given direction $\hat{\mathbf{c}}$ [29, 30, 34]. A point on the VB between particle i and j along this direction is found at $s\hat{\mathbf{c}}$, where s depends on their relative position $\mathbf{r}_{ij}$ and orientation $\hat{\mathbf{t}}_{ij}$: $s = s(\mathbf{r}_{ij}, \hat{\mathbf{t}}_{ij}; \hat{\mathbf{c}})$. The value of s is measured as half the minimal distance to the surfaces of both particles along the direction $\hat{\mathbf{c}}$. For
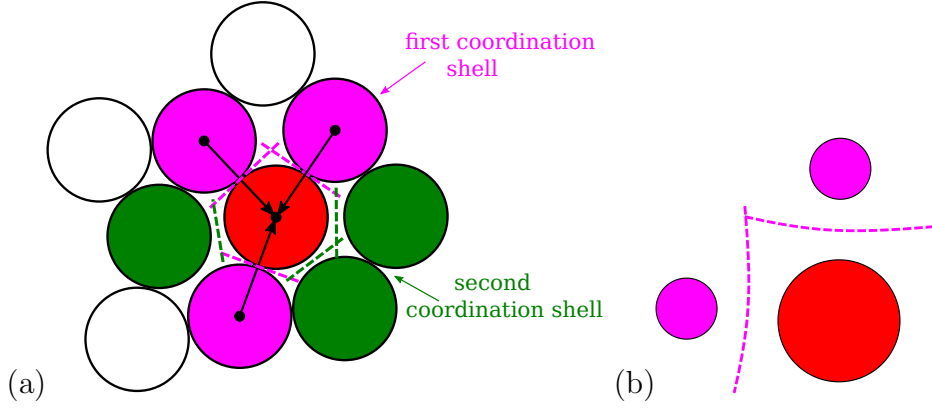
Figure A1: (a) Voronoi tessellation in a packing of monodisperse spheres (illustrated in 2D for simplicity). The Voronoi Boundary (VB) is the flat plane perpendicular to the vector that separates the spheres' centres at half the distance. The VBs between the reference particle (red) and the particles in the first and second coordination shell are indicated with dashed lines (pink and green, respectively), (b) The VB becomes curved when the spheres have different radii.

equal radii spheres, the VB can be obtained between two points at the centres of the spheres, so s then follows the equation $(s\hat{\mathbf{c}})^2 = (s\hat{\mathbf{c}} - \mathbf{r}_{ij})^2$ and can be calculated as:

$$s = \frac{r_{ij}}{2\hat{\mathbf{c}} \cdot \hat{\mathbf{r}}_{ij}}, \tag{A.1}$$

where $r_{ij} = |\mathbf{r}_{ij}|$. Hence, the VB between two spheres of equal radii is the flat plane perpendicular to the vector $\mathbf{r}_{ij}$ that separates the spheres' centres at half the distance, see Fig. A1(a). It becomes a curved surface when the spheres have different radii a_i and a_j as shown in Fig. A1(b), and s can be calculated from the equation $s - a_i = \sqrt{(s\hat{\mathbf{c}} - \mathbf{r}_{ij})^2} - a_j$ as in [160]:

$$s = \frac{1}{2} \frac{r_{ij}^2 - (a_i - a_j)^2}{\hat{\mathbf{c}} \cdot \mathbf{r}_{ij} - (a_i - a_j)}. \tag{A.2}$$
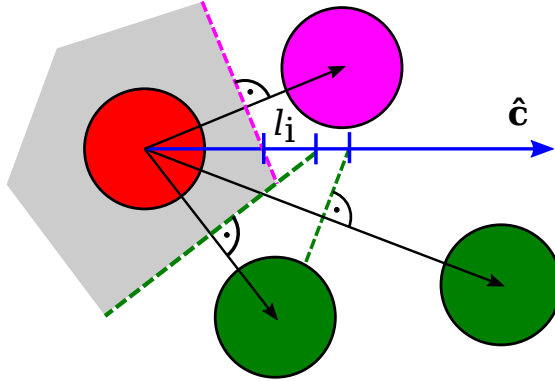
Figure A2: The volume of a Voronoi cell (coloured grey) associated with the reference particle (red) contains all points closer to the surface of that particle than to the surface of any other one. The minimum among all VBs in a given direction $\hat{\mathbf{c}}$ gives the cell boundary $l_i(\hat{\mathbf{c}})$. The contribution of the pink particle to the VB is smaller than that of the green particles, so it defines the boundary in that direction.

A.2   The Mean-field Theory

Song *et al.* obtained the exact mathematical formula for the Voronoi volume, W_i in three dimensions by the orientational integral [29]

$$W_i = \frac{1}{3} \oint \mathrm{d}\hat{\mathbf{c}}\, l_i(\hat{\mathbf{c}})^3, \tag{A.3}$$

$$l_i(\hat{\mathbf{c}}) = \min_{j:s>0} s(\mathbf{r}_{ij}, \hat{\mathbf{t}}_{ij}, \hat{\mathbf{c}}), \tag{A.4}$$

where $l_i(\hat{\mathbf{c}})$, the boundary of the Voronoi cell, is found by a global minimization of the VB in the direction $\hat{\mathbf{c}}$ over all other $N-1$ particles in the packing (see Fig. A2). Since this process requires precise knowledge of the microscopic configurations of all particles, the determination of W_i remains challenging. The mean-field theory avoids this difficulty by presenting a coarse-grained Voronoi volume $\overline{W}$, which is

the ensemble average of W_i over all particles in the packing as [34]:

$$\overline{W} = \left\langle \frac{1}{3} \oint \mathrm{d}\hat{\mathbf{c}}\, l_i(\hat{\mathbf{c}})^3 \right\rangle_i = \frac{1}{3} \oint \mathrm{d}\hat{\mathbf{c}}\, \langle l_i(\hat{\mathbf{c}})^3 \rangle_i = \frac{1}{3} \oint \mathrm{d}\hat{\mathbf{c}} \int_{c^*(\hat{\mathbf{c}})}^{\infty} \mathrm{d}c\, c^3 p(\mathbf{c}), \qquad (A.5)$$

where $p(\mathbf{c})$ represents the probability density to find the VB at c in the direction $\hat{\mathbf{c}}$, and $c^*(\hat{\mathbf{c}})$, the minimum value of the VB along $\hat{\mathbf{c}}$, corresponds to the particle's hard core boundary in the direction $\hat{\mathbf{c}}$. For example, for spheres of radius R_s, $c^*(\hat{\mathbf{c}}) = R_\mathrm{s}$. Substituting the cumulative distribution function $P(\mathbf{c})$, where $p(\mathbf{c}) = -\frac{\mathrm{d}}{\mathrm{d}c}P(\mathbf{c})$, in Eq. A.5 and applying integration by parts results in:

$$\overline{W}(z) = \int \mathrm{d}\mathbf{c}\, P(\mathbf{c}, z), \qquad (A.6)$$

where $P(\mathbf{c}, z)$ depends on the average contact number z, and can be interpreted as the probability to find $N-1$ particles outside a volume Υ centred at $\mathbf{c}$, i.e. they do not contribute a shorter VB [34]. Taking into account both contact and bulk particle contributions to $P(\mathbf{c}, z)$ leads to the following self-consistent equation:

$$\overline{W}(z) = \int \mathrm{d}\mathbf{c}\, \exp\left\{ -\frac{V^*(\mathbf{c})}{\overline{W} - V_0} - \sigma(z)S^*(\mathbf{c}), \right\} \qquad (A.7)$$

where $V^*(\mathbf{c})$ and $S^*(\mathbf{c})$ are referred to as the Voronoi excluded volume and surface by extending Onsager's hard-core excluded volume concept to disordered particle packings, see Fig. A3. The surface density $\sigma(z)$ for a packing with an average contact number z measures the available surface for contacts, and it is found
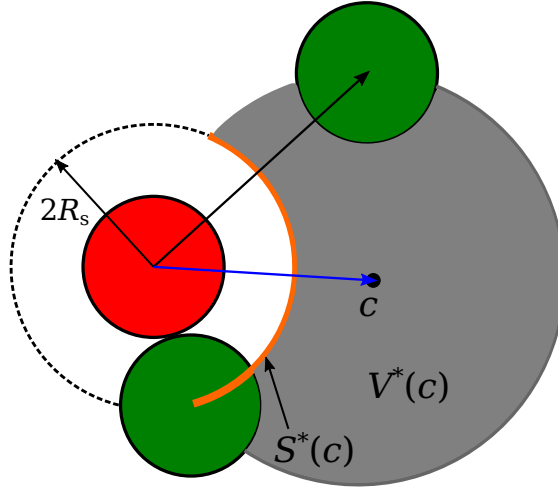
Figure A3: The VB of the reference sphere (red) in a given direction at the value c is related to the volume $V^*(\mathbf{c})$ (grey volume) and the surface $S^*(\mathbf{c})$ (orange line) excluded by Υ for bulk and contacting particles, respectively. $V^*(\mathbf{c})$ and $S^*(\mathbf{c})$ take into account the overlap between Υ and the hard-core excluded volume and boundary, respectively [5].

from the simulations of the local contact configurations. Determining $\overline{W}(z)$ from Eq. A.7 simplifies the packing density calculation as a function of z to $\phi = V_0/\overline{W}(z)$ [29, 30, 34]. Since both $V^*(\mathbf{c})$ and $S^*(\mathbf{c})$ follow an exclusion condition on the Voronoi boundary (VB) between two particles, it is essential to find the analytical expression for the VB between two particles to solve Eq. A.7 [158]. However, it is challenging to obtain that for non-spherical particles [157].

Baule *et al.* provided an analytical expression for the VB by decomposing rotationally symmetric non-spherical particles into overlapping and intersecting spheres of equal-radii [34]. In this approach, the VB between two particles consists of segments constructed by considering the interaction between a constituent sphere on each of the two objects. In order to determine the cell boundary in a given direction, one needs to identify the correct pair of spheres. The method for the
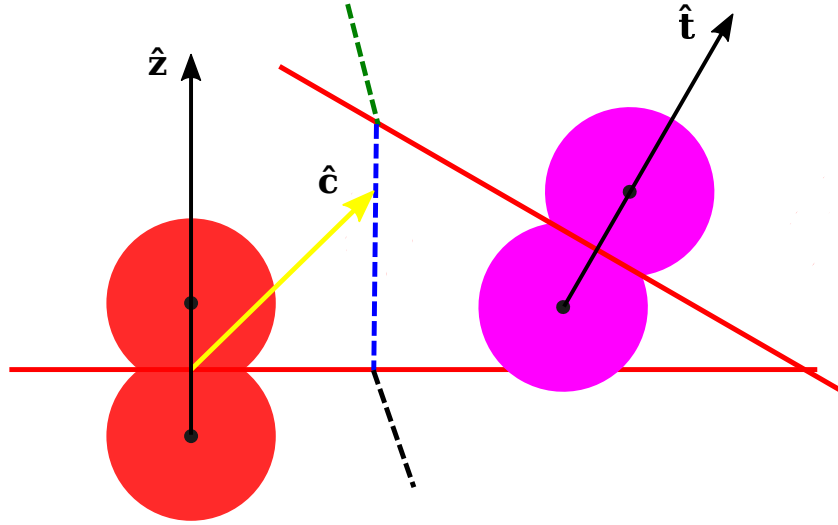
Figure A4: The VB between two symmetric dimers of a given relative position and orientation consists of the VBs (dashed lines) between particular spheres on each of the two particles. There are four sphere interactions, and the separation lines (indicated in red) determine the correct pair to construct the VB in a given direction $\hat{c}$. For example, the green part in the VB is due to the interaction between the two upper spheres [34].

identification depends on the density of overlapping or intersecting spheres defining the particle shape. For dimers, for instance, there are four different VBs between the points at the centre of each constituent sphere, and the correct one in a given direction depends on the relative orientation of the dimers, see Fig. A4. A spherocylinder consists of many spheres, so the VB is then determined by the interaction between four points and two lines.

Here, I apply this approach to asymmetric dimers obtained by overlapping two spheres of unequal radii. Three parameters define an asymmetric dimer: the radius of the large sphere (1) a_1, the radius of the small sphere (2) a_2, and the distance between the centres of the two spheres, b, see Fig. A5. The centre of the coordinate system is set to the centre of sphere 1 of the reference dimer i, and
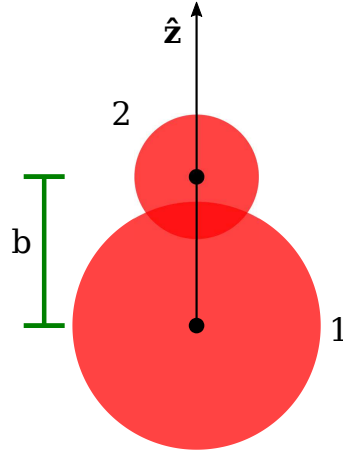
Figure A5: An asymmetric dimer is obtained by overlapping two spheres of unequal radii (a_1 and a_2), with a separation of the two sphere centres b.

its orientation is fixed along the direction $\hat{\mathbf{z}}$ as shown in Fig. A5. The VB is then parametrized in terms of a given direction $\hat{\mathbf{c}}$ from sphere 1. Four different sphere interactions generate the VB between two asymmetric dimers i and j, so there are four candidate VBs for the direction $\hat{\mathbf{c}}$. The four values for the VB are calculated by separation vectors between the points at the sphere centres for each pair, which are given as:

$$\mathbf{r}_{11}, \qquad \mathbf{r}_{12} = \mathbf{r}_{11} + b\hat{\mathbf{t}}, \qquad \mathbf{r}_{21} = \mathbf{r}_{11} - b\hat{\mathbf{z}}, \qquad \mathbf{r}_{22} = \mathbf{r}_{11} - b\hat{\mathbf{z}} + b\hat{\mathbf{t}}, \qquad (A.8)$$

where the subscript 11 denotes sphere 1 of dimer i and sphere 1 of dimer j (indicated blue in Fig A6), and $\hat{\mathbf{t}}$ is the relative orientation of dimer j. The VB $s\hat{\mathbf{c}}_{11} = s_{11}\hat{\mathbf{c}}$ is due to the interaction between sphere 1 of dimer i and sphere 1 of dimer j. Since the two spheres have the same radius, the value of s_{11} follows Eq. A.1 and can be

found as:

$$s_{11} = \frac{r_{11}}{2\hat{\mathbf{c}} \cdot \hat{\mathbf{r}}_{11}}. \tag{A.9}$$

Taking into account the different radii a_1 and a_2, s_{12} can be calculated directly from Eq. A.2:

$$s_{12} = \frac{1}{2} \frac{r_{12}^2 - (a_1 - a_2)^2}{\hat{\mathbf{c}} \cdot \mathbf{r}_{12} - (a_1 - a_2)}. \tag{A.10}$$

The equation $(s_{22}\hat{\mathbf{c}} - b\hat{\mathbf{z}})^2 = ((s_{22}\hat{\mathbf{c}} - b\hat{\mathbf{z}}) - \mathbf{r}_{22})^2$ gives the value of s_{22} as:

$$s_{22} = \frac{1}{2} \frac{r_{22}}{\hat{\mathbf{c}} \cdot \hat{\mathbf{r}}_{22}} + \frac{b\hat{\mathbf{z}} \cdot \hat{\mathbf{r}}_{22}}{\hat{\mathbf{c}} \cdot \hat{\mathbf{r}}_{22}}. \tag{A.11}$$

Although obtaining the values of s_{11}, s_{12}, and s_{22} is straightforward, s_{21} can not be determined since it requires a solution for the equation $\sqrt{(s_{21}\hat{\mathbf{c}} - b\hat{\mathbf{z}})^2} - a_2 = \sqrt{(s_{21}\hat{\mathbf{c}} - b\hat{\mathbf{z}} - \mathbf{r}_{21})^2} - a_1$. For symmetric dimers, the separation lines can identify the correct interaction for the VB by tessellating space into four areas, see Fig. A4. However, such an identification is not possible for the asymmetric case. Although the VB can be constructed intuitively, as seen in Fig. A6, an analytical expression for that is not available due to s_{21} can not be determined in the parametrisation, and the separation curves do not identify the correct pair. Therefore, the mean-field theory can not be extended to asymmetric dimers. In terms of numerical aspect, Schaller *et al.* have developed an algorithm to compute the Set Voronoi diagram for more general shapes, which is based on the discretisation of the particles' bounding surfaces [159].
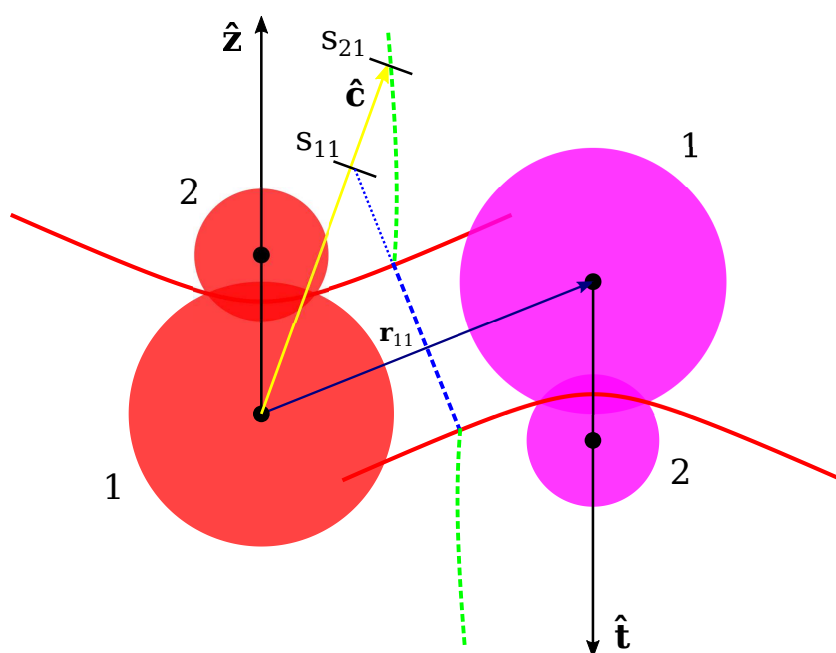
Figure A6: The VB between asymmetric dimers consists of both flat (blue) and curved (green) surfaces depending on the interacting pair of spheres. The separation lines (indicated red) to identify the correct pair for the VB in the direction $\hat{\mathbf{c}}$ become curved for the asymmetric ones. There is no analytical solution to determine the VB.

Appendix B

Appendix of Chapter 3

B.1 Calculation of the dimer volume

The overlap volume of the two constituent spheres of a dimer contains two equal spherical caps which lie above/below the plane through the cusp points at the dimer's centre, see Fig. B1. The volume of a spherical cap V_{cap} of height h is found as:

$$V_{\text{cap}} = \frac{1}{3}\pi h^2 (3R_{\text{s}} - h) \tag{B.1}$$

where R_{s} is the sphere radius. The dimer volume V_α is then calculated by subtracting the overlap volume from the sum of its constituent sphere volumes $V_{\text{sphere}} = \frac{4}{3}\pi R_{\text{s}}^3$ as:

$$V_\alpha = 2V_{\text{sphere}} - 2V_{\text{cap}} \tag{B.2}$$
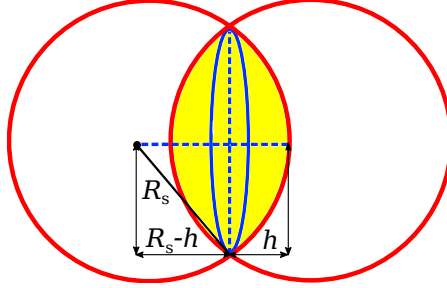
Figure B1: The overlap volume of a dimer contains two equal spherical caps of height h (coloured in yellow).

B.2 Algorithm for the identification of double and cusp contacts

Double and cusp contacts are identified by checking if there is any overlap between the circle enclosing the cusp on the dimer surface and a contacting sphere of its neighbouring dimer, see Fig. B2(a). This circle with centre $\mathbf{c}_c$, radius R_c and unit normal $\mathbf{w}$ and a sphere with centre $\mathbf{c}_s$, radius R_s are shown in Fig. B2(b). The next steps are followed for the identification:

1. The distance $d_{cs} = |\mathbf{w} \cdot (\mathbf{c}_c - \mathbf{c}_s)|$ between the plane of the circle and the sphere's centre is calculated to check if the plane cuts the sphere or not. If $d_{cs} > R_s$ then there is no intersection, so the plane passes above/below the sphere entirely.

2. If there is an intersection, i.e., $d_{cs} < R_s$, it will be between the original circle and a new one formed where this plane meets the sphere, with centre $\mathbf{c}_p = \mathbf{c}_s + d_{cs}\mathbf{w}$.

3. If $d_{cs} = R_s$ then this is the sole point of intersection with the plane, otherwise a new circle with radius R_p occurs as displayed in Fig. B2(c), where $R_p = \sqrt{R_s^2 - d_{cs}^2}$. Then, the problem has been reduced to a circle-circle interaction.

4. If $|\mathbf{c}_p - \mathbf{c}_c| < R_c + R_p$, then there is overlap between the circle and the sphere, so the contact is identified as a cusp contact. If there is no overlap, then the contact is either a double contact or a Type 2 configuration.

5. To distinguish a double and a Type 2 configuration, two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ from the contacting sphere's centre to the centres of the constituting spheres of the reference dimer are determined as illustrated in Fig. B3. The projections of these two vectors onto the unit normal $\mathbf{w}$ of the circle enclosing cusp are determined and the directions of these projections are checked. If both of them have the same direction, the contact is identified as a double contact, otherwise it is regarded as a Type 2 configuration.
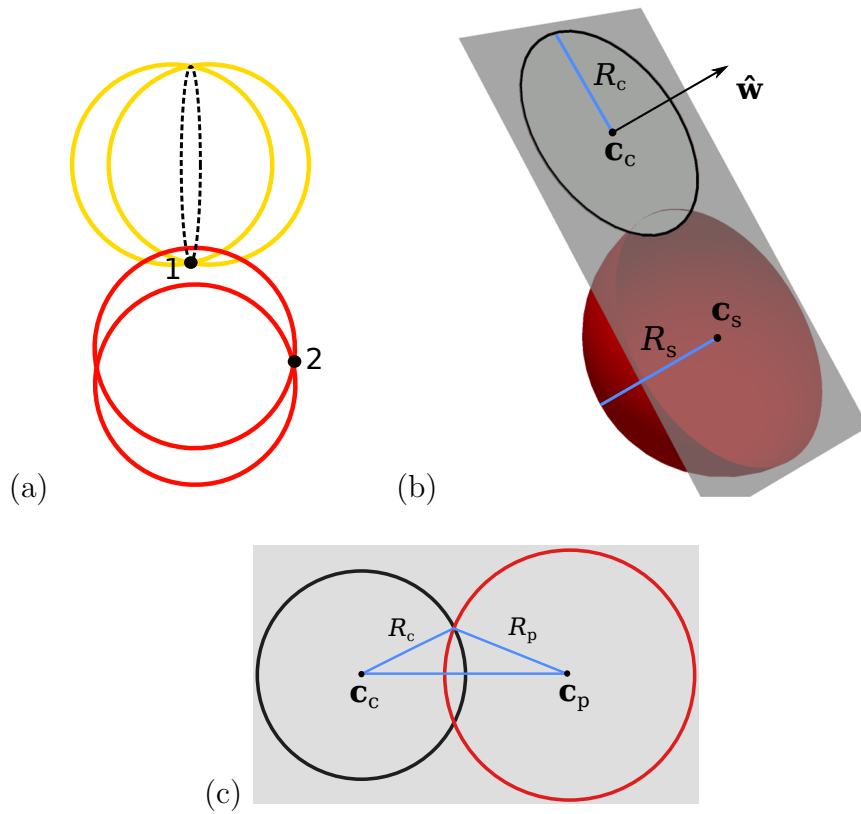
Figure B2: Detecting double and cusp contacts. (a) First, it is checked if there is any overlap between the black circle (dashed) enclosing cusp located on the yellow dimer's surface and the contacting sphere of the red dimer. If there is an overlap between the circle and the sphere, it is identified as a cusp contact. (b) 3D Visualization of the circle and sphere interaction, it is determined if the plane of the circle cuts the sphere or not. (c) If the plane of the circle cuts the sphere, it forms a new circle (red) and then it is checked if there is overlap between the original circle and the new red circle.
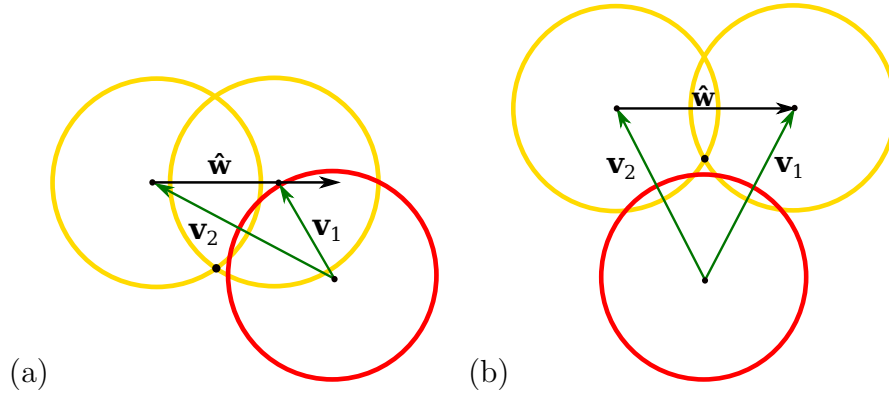
Figure B3:   Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ from the contacting red sphere's centre to the centres of the constituting spheres of the yellow dimer are determined. The projections of these two vectors onto the unit normal $\mathbf{w}$ of the circle enclosing cusp are determined and the directions of these projections are checked. (a) If both of them have the same direction, it is identified as a double contact (b) otherwise it is regarded as Type 2 configuration.

B.3   The order parameter χ

In [47] the following order parameter has been introduced to measure the orientational order of prolate ellipsoids

$$\chi = \frac{3}{2} \left\{ \frac{1}{N_b} \sum_i^{N_b} \cos 2 \left(\beta_i - \frac{\pi}{2} \right) - \frac{1}{3} \right\}, \tag{B.3}$$

where β_i is the angle between the semi-major (long) axis of particle i and the $\hat{\mathbf{z}}$-axis (gravity direction). The parameter χ of Eq. (B.3) takes values in the interval $[-2, 1]$: when all particles are randomly oriented, $\chi = 0$, while if all particles' long axes are oriented in the horizontal plane normal to the gravity direction $\chi = 1$. When the long axes of particles are oriented along the gravity direction we have $\chi = -2$. A plot of χ as a function of α for our dimer packing data is shown in Fig. B4.
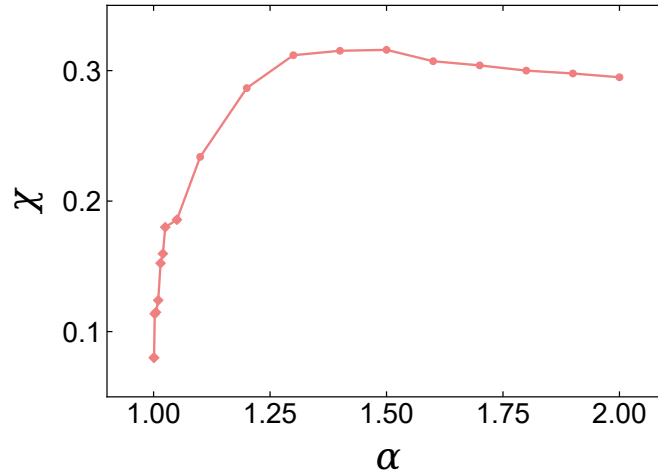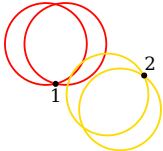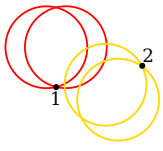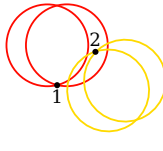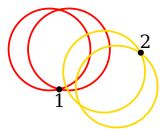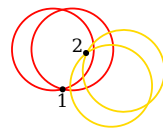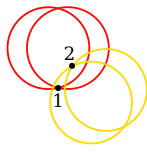
Figure B4: The orientational order parameter χ *vs.* α. Values of χ are shown averaged over 10 independent simulation runs for $\alpha \geq 1.1$ (dots), and for a single run for $\alpha < 1.1$ (diamonds).

B.4 Mapping between different contact configuration types

I introduce a heuristic method to re-assign configurations with double and cusp contacts to one of the Type 1, 2, and 4 configurations. The precise mapping depends on the number and the location of double and cusp contacts as summarized in Table B-I. In general, double contacts are mapped to one contact point and cusp contacts to two. For Type 3 configurations, no double or cusp contacts have been found. For Type 5 configurations, two cusp contacts do occur, which leave the configuration as Type 5 after the mapping.

With this mapping, I count a smaller number of contact points and thus the average number of contacts z decreases. In fact, a rapid but smooth decrease of

Table B-I: Two-dimensional illustrations of configurations with double and cusp contacts. These configurations are re-assigned to Type 1, 2, and 4 as indicated in the table.

Configuration type	Re-assigned configuration type		
	Type 1	Type 2	Type 4
Type 2	A double contact is counted as one contact point: two contact points are reduced to one.		
Type 4	Two overlapping double contacts are counted as one contact point: three contact points are reduced to one.	One double and one cusp contact (cusp 2 overlaps with the red sphere) are counted as two contact points: three contact points are reduced to two.	
Type 5	Two overlapping double contacts are counted as one contact point: four contact points are reduced to one.	Two distinct double contacts are counted as two contact points: four contact points are reduced to two.	One double contact (cusp 1 is not covered by one of the yellow spheres) and one cusp contact (cusp 1 overlaps with the other yellow sphere) are counted as three contact points: four contact points are reduced to three.

z is observed as $\alpha \to 1$, whereby z approaches the corresponding value of spheres (Fig. B5). Resolving the contact counting by Type 1–5 configurations, as expected, the fraction of Type 1 configurations now increases for $\alpha < 1.05$, while the fractions of Type 2,4,5 configurations decreases in the same regime (Fig. B6). In fact, the adjusted counting of contact points leads to sharp peaks at $\alpha \approx 1.05$, i.e., at the aspect ratio at which double and cusp contacts start to occur, that appear unphysical.
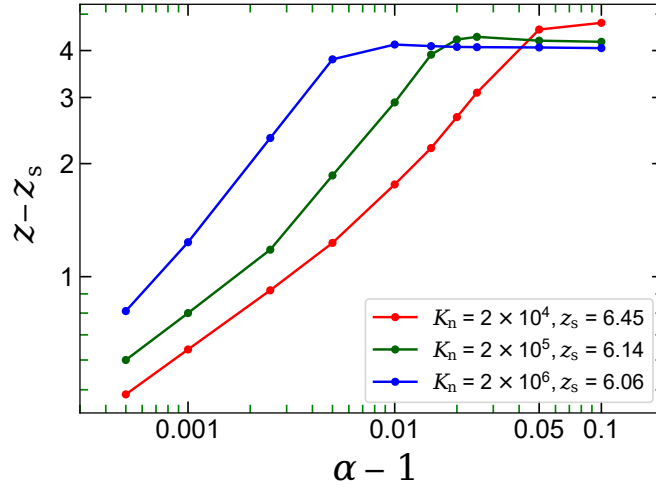
Figure B5: A double-logarithmic plot of $z - z_s$ *vs.* $\alpha - 1$ for three different normal spring constants K_n. I define z_s as the contact number of the corresponding sphere packing, which approaches the isostatic value $z_s = 6$ as the particle hardness increases.
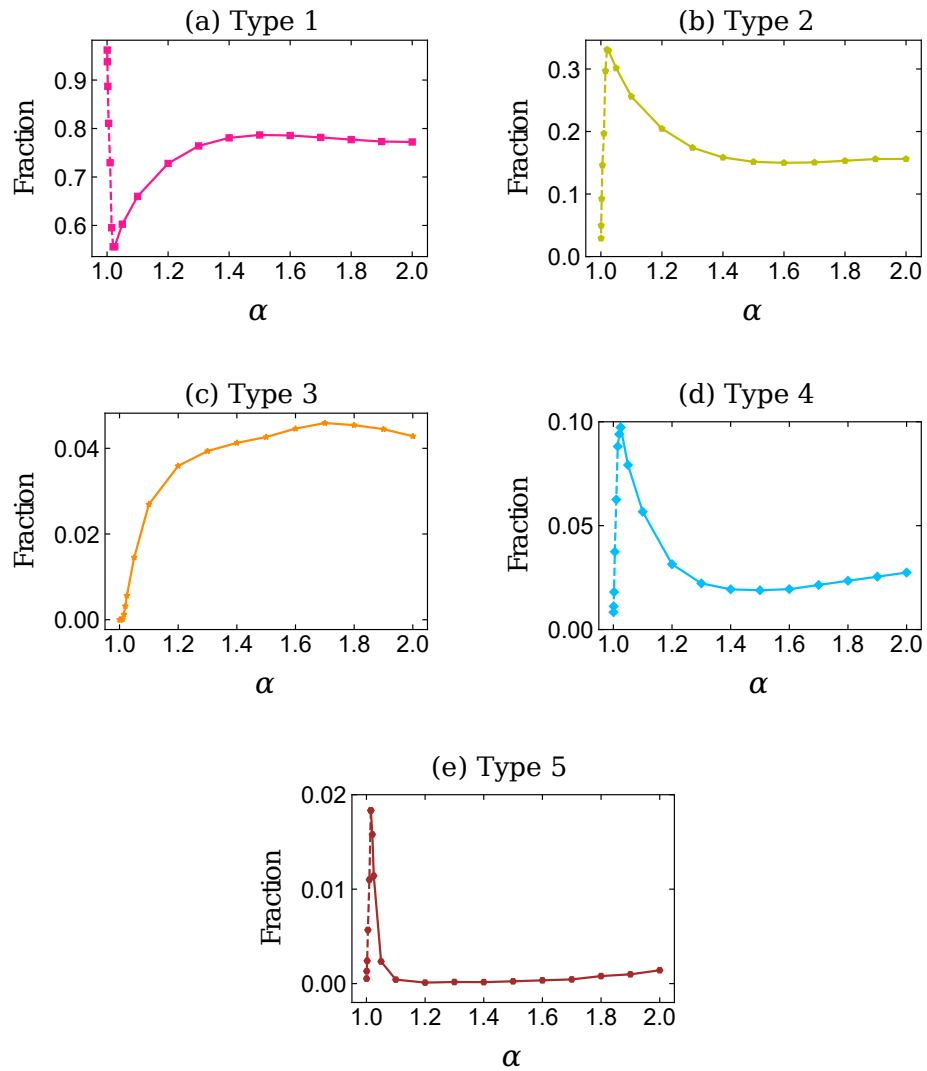
Figure B6: The fractions of the contact configurations of Type 1–5 *vs.* the aspect ratio α. For $\alpha \geq 1.05$ the data shown is the same as in Fig. 3.7, but in the regime $\alpha < 1.05$ (dashed lines) the contact counting has been adjusted by re-assigning configurations with double and cusp contacts to Type 1, 2, and 4 configurations as summarized in Table B-I.

Appendix C

Appendix of Chapter 4

C.1 Statistical analysis of the consistency of the ideal mixing law with the binary dimer mixtures data

To evaluate whether the agreement of the mixing law Eq. 4.1 with the data of dimer–sphere and dimer–dimer mixtures is statistically significant, I have fitted a linear line to the data by ordinary least squares regression and then checked whether the mixing law falls within the 99% confidence interval around the best-fit slope. I have shown the results for dimer–sphere and dimer–dimer mixtures in Figs. C1(a,b), respectively. As can be seen, the ideal mixing law is consistent with a linear fit since it falls within the 99% confidence interval for all aspect ratios in dimer–sphere

mixtures and for $\alpha_2 < 2$ in dimer–dimer mixtures. For the mixtures with $\alpha_2 = 2$,
since there are clearly systematic deviations from the linear behaviour, which is
visible in Fig. 4.9(d), the linear regression fit is not valid anymore. Therefore, I
have fitted a polynomial model, which shows that contributions up to degree 4 are
statistically significant. This polynomial fit almost perfectly agrees with the data
as shown in Figs. C1(b), thus highlighting the onset of nonlinear behaviour for a
sufficiently large difference between the volumes of the two components.
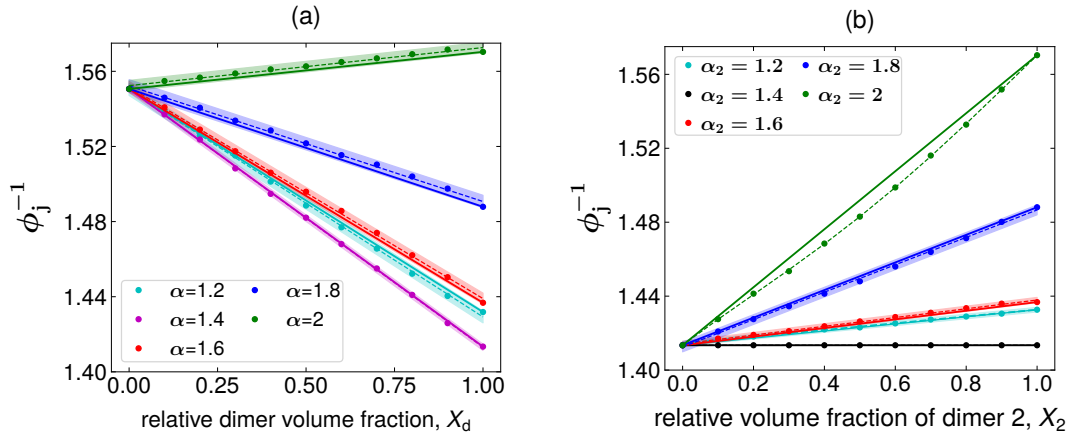
Figure C1: Statistical analysis of the mixing law's consistency with the binary
mixtures' data. (a) The inverse packing density ϕ_j^{-1} of the dimer–sphere mixtures
as a function of the relative dimer volume fraction X_d for several dimer aspect
ratios (same data as shown in Fig. 4.3(b)). Solid lines are obtained from Eq. (4.1),
and dashed lines are fitted to the data points with linear regression. Each shaded
region shows the 99% confidence interval for the regression slope. (b) The inverse
packing density ϕ_j^{-1} of mixtures of dimer 1 with $\alpha_1 = 1.4$ and dimer 2 with various
aspect ratios α_2 as a function of X_2 (same data as shown in Fig. 4.9(d)). Solid
lines are obtained from Eq. (4.1), and dashed lines are fitted to the data points
with linear regression for $\alpha_2 < 2$ (including the 99% confidence interval) and with
a fourth-order polynomial for $\alpha_2 = 2$.

Appendix D

Appendix of Chapter 5

D.1   Calculation of the 5-sphere particle volume

In order to calculate the volume V_{fs} of a 5-sphere particle, I use the Monte Carlo integration. The particle is placed into a box by ensuring that all constituent spheres are inside it, and a large number of points N_{box} is randomly generated in the box. The distance d_{ps} between a point and the centre of an individual sphere is calculated to determine how many points are in the molecule. If $d_{\mathrm{ps}} \leq a_i$ is satisfied for any sphere, where $i = 1, ..., 5$ and a_i is the radius of sphere i, then the point is counted in the molecule. Finally, the volume box V_{box} is multiplied by the ratio of the number of points inside the molecule, N_{in} to the total number of points to determine the volume of the molecule, i.e. $V_{\mathrm{fs}} = V_{\mathrm{box}} N_{\mathrm{in}} / N_{\mathrm{box}}$. To determine N_{box}, I set a criterion that keeps generating points until the variance is less than 10^{-6}.

Appendix E

Presentations at conferences

During my PhD, I have presented my work on *"Structural analysis of disordered dimer packings"* (Chapter 3) at:

- SEPnet Student-Led Soft Matter Conference, online, United Kingdom, 2020.

  Title: "Structural analysis of randomly packed dimer particles".

Appendix F

Author's publications

The following research works have been completed in collaboration with Dr. Adrian Baule.

- **Kurban, E.** and Baule, A. Structural analysis of disordered dimer packings. *Soft Matter* 17, 8877-8890 (2021). This work has been represented in Chapter 3 of this thesis.

- **Kurban, E.** and Baule, A. Disordered packings of binary mixtures of dimer particles. This paper is submitted to "*Journal of Physics Communications*" and is under review. This work has been represented in Chapter 4 of this thesis.

References

[1] Heinrich M. Jaeger, Sidney R. Nagel, and Robert P. Behringer. Granular solids, liquids, and gases. *Reviews of Modern Physics*, 68:1259–1273, 1996.

[2] Heinrich M. Jaeger. Celebrating soft matter's 10th anniversary: toward jamming by design. *Soft Matter*, 11:12–27, 2015.

[3] Hernán A. Makse, Jasna Brujić, and Sam F. Edwards. Statistical mechanics of jammed matter. In *The Physics of Granular Media*. Wiley-VCH, 2005.

[4] Sam F. Edwards and R. B. S. Oakeshott. Theory of powders. *Physica A: Statistical Mechanics and its Applications*, 157:1080–1090, 1989.

[5] Adrian Baule, Flaviano Morone, Hans J. Herrmann, and Hernán A. Makse. Edwards statistical mechanics for jammed granular matter. *Reviews of Modern Physics*, 90:015006, 2018.

[6] Johannes Kepler. Strena seu de nive sexangula (The six-cornered snow ake). 1611. URL: `http://www.thelatinlibrary.com/kepler/strena.html`.

[7] Tomaso Aste and Tiziana Di Matteo. Emergence of gamma distributions in granular materials and packing models. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 77:021309, 2008.

[8] Carl F. Gauss. Besprechung des buchs von l. a. seeber: Untersuchungen über die eigenschaften der positiven ternären quadratischen formen. *Göttingsche Gelehrte Anzeigen See also J. reine angew. Math.*, 20:312–320, 1840.

[9] Thomas C. Hales. A proof of the kepler conjecture. *Annals of Mathematics*, 162:1065–1185, 2005.

[10] John Desmond Bernal and J. Mason. Packing of spheres: Co-ordination of randomly packed spheres. *Nature*, 188:910–911, 1960.

[11] John Desmond Bernal. The bakerian lecture, 1962 the structure of liquids. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 280:299 – 322, 1964.

[12] G. David Scott. Packing of spheres: Packing of equal spheres. *Nature*, 188:908–909, 1960.

[13] G. David Scott and D. Marc Kilgour. The density of random close packing of spheres. *Journal of Physics D: Applied Physics*, 2:863–866, 1969.

[14] Corey S. O'Hern, Leonardo E. Silbert, Andrea J. Liu, and Sidney R. Nagel. Jamming at zero temperature and zero applied stress: the epitome of disorder. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 68:011306, 2003.

[15] Paul Steinhardt, David Nelson, and Marco Ronchetti. Bond-orientational order in liquids and glasses. *Physical Review B: Condensed Matter and Materials Physics*, 28:784–805, 1983.

[16] Sebastian Kapfer, Walter Mickel, Klaus Mecke, and Gerd E. Schröder-Turk. Jammed spheres: Minkowski tensors reveal onset of local crystallinity. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 85:030301, 2012.

[17] Gerd E. Schröder-Turk, Walter Mickel, Matthias Schröter, Geoff P. Delaney, Mohammad Saadatfar, Timothy John Senden, Klaus Mecke, and Tomaso

Aste. Disordered spherical bead packs are anisotropic. *Europhysics Letters*, 90:34001, 2010.

[18] Tomaso Aste, Mohammad Saadatfar, Arthur Sakellariou, and Timothy John Senden. Investigating the geometrical structure of disordered sphere packings. *Physica A: Statistical Mechanics and its Applications*, 339:16–23, 2004.

[19] Tomaso Aste. Variations around disordered close packing. *Journal of Physics: Condensed Matter*, 17:S2361, 2005.

[20] Tomaso Aste, Mohammad Saadatfar, and Timothy John Senden. Local and global relations between the number of contacts and density in monodisperse sphere packs. *Journal of Statistical Mechanics: Theory and Experiment*, 2006:P07010 – P07010, 2006.

[21] W. Steven Jodrey and Elmer M. Tory. Computer simulation of close random packing of equal spheres. *Phyical Review A*, 32:2347–2351, 1985.

[22] Boris D. Lubachevsky and Frank H. Stillinger. Geometric properties of random disk packings. *Journal of Statistical Physics*, 60:561–583, 1990.

[23] Boris D. Lubachevsky, Frank H. Stillinger, and Elliot N. Pinson. Disks vs. spheres: Contrasting properties of random packings. *Journal of Statistical Physics*, 64:501–524, 1991.

[24] Leonardo Silbert, Deniz Ertaş, Gary S. Grest, Thomas Halsey, and Dov Levine. Geometry of frictionless and frictional sphere packings. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 65:031304, 2002.

[25] Salvatore Torquato and Frank H. Stillinger. Jammed hard-particle packings: From kepler to bernal and beyond. *Reviews of Modern Physics*, 82:2633–2672,

2010.

[26] Salvatore Torquato, Thomas M. Truskett, and Pablo G. Debenedetti. Is random close packing of spheres well defined? *Physical Review Letters*, 84:2064–2067, 2000.

[27] Giorgio Parisi and Francesco Zamponi. Mean-field theory of hard sphere glasses and jamming. *Reviews of Modern Physics*, 82:789–845, 2010.

[28] Patrick Charbonneau, Jorge Kurchan, Giorgio Parisi, Pierfrancesco Urbani, and Francesco Zamponi. Glass and jamming transitions: From exact results to finite-dimensional descriptions. *Annual Review of Condensed Matter Physics*, 8(1):265–288, 2017.

[29] Chaoming Song, Ping Wang, and Hernán A. Makse. A phase diagram for jammed matter. *Nature*, 453:629–632, 2008.

[30] Chaoming Song, Ping Wang, Yuliang Jin, and Hernán A. Makse. Jamming i: A volume function for jammed matter. *Physica A: Statistical Mechanics and its Applications*, 389:4497–4509, 2010.

[31] Monica L. Skoge, Aleksandar Donev, Frank H. Stillinger, and Salvatore Torquato. Packing hyperspheres in high-dimensional euclidean spaces. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 74:041127, 2006.

[32] Aleksandar Donev, Ibrahim Cisse, David Sachs, Evan Variano, Frank H. Stillinger, Robert Connelly, Salvatore Torquato, and Paul M. Chaikin. Improving the density of jammed disordered packings using ellipsoids. *Science (New*

York, N.Y.), 303:990–993, 2004.

[33] Sylvain Faure, Aline Lefebvre-Lepot, and Benoît Semin. Dynamic numerical investigation of random packing for spherical and nonconvex particles. In *ESAIM: Proceedings*, volume 28, pages 13–32, 2009.

[34] Adrian Baule, Romain Mari, Lin Bo, Louis Portal, and Hernán Makse. Mean-field theory of random close packings of axisymmetric particles. *Nature communications*, 4:2194, 2013.

[35] Chengjie Xia, Kuan Zhu, Yixin Cao, Haohua Sun, Binquan Kou, and Yujie Wang. X-ray tomography study of the random packing structure of ellipsoids. *Soft Matter*, 10:990–996, 2014.

[36] Jian Zhao, Shuixiang Li, Ruiping Zou, and Aibing Yu. Dense random packings of spherocylinders. *Soft Matter*, 8:1003–1009, 2012.

[37] Lingyi Meng, Yang Jiao, and Shuixiang Li. Maximally dense random packings of spherocylinders. *Powder Technology*, 292:176–185, 2016.

[38] Amir Haji-Akbari, Michael Engel, Aaron S. Keys, Xiaoyu Zheng, Rolfe G. Petschek, Peter Palffy-Muhoray, and Sharon C. Glotzer. Disordered, quasicrystalline and crystalline phases of densely packed tetrahedra. *Nature*, 462:773, 2009.

[39] Alexander Jaoshvili, Andria Esakia, Massimo Porrati, and Paul M. Chaikin. Experiments on the random packing of tetrahedral dice. *Physical Review Letters*, 104:185501, 2010.

[40] Jessica Baker and Arshad Kudrolli. Maximum and minimum stable random packings of platonic solids. *Physical Review E: Statistical, Nonlinear, and*

Soft Matter Physics, 82:061304, 2010.

[41] Yang Jiao and Salvatore Torquato. Maximally random jammed packings of platonic solids: Hyperuniform long-range correlations and isostaticity. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 84:041309, 2011.

[42] Paul M. Chaikin, Aleksandar Donev, Weining Man, Frank H. Stillinger, and Salvatore Torquato. Some observations on the random packing of hard ellipsoids. *Industrial & Engineering Chemistry Research*, 45:6960–6965, 2006.

[43] Weining Man, Aleksandar Donev, Frank H. Stillinger, Matthew Sullivan, William Russel, David Heeger, Souheil Inati, Salvatore Torquato, and Paul M. Chaikin. Experiments on random packings of ellipsoids. *Physical Review Letters*, 94:198001, 2005.

[44] Yang Jiao, Frank H. Stillinger, and Salvatore Torquato. Distinctive features arising in maximally random jammed packings of superballs. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 81:041304, 2010.

[45] Leah K. Roth and Heinrich M. Jaeger. Optimizing packing fraction in granular media composed of overlapping spheres. *Soft Matter*, 12:1107–1115, 2016.

[46] Adrian Baule and Hernán A. Makse. Fundamental challenges in packing problems: from spherical to non-spherical particles. *Soft Matter*, 10:4423–4429, 2014.

[47] B. J. Buchalter and R. Mark Bradley. Orientational order in amorphous

packings of ellipsoids. *Europhysics Letters*, 26(3):159–164, 1994.

[48] Gary Delaney, James E. Hilton, and Paul Cleary. Defining random loose packing for nonspherical grains. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 83:051305, 2011.

[49] Jieqing Gan and Aibing Yu. Dem study on the packing density and randomness for packing of ellipsoids. *Powder Technology*, 361:424–434, 2020.

[50] Stephen R. Williams and Albert P. Philipse. Random packings of spheres and spherocylinders simulated by mechanical contraction. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 67:051301, 2003.

[51] Charlles R. A. Abreu, Frederico W. Tavares, and Marcelo Castier. Influence of particle shape on the packing and on the segregation of spherocylinders via monte carlo simulations. *Powder Technology*, 134:167–180, 2003.

[52] Xiaodong Jia, Mingle Gan, Richard A. Williams, and D. Rhodes. Validation of a digital packing algorithm in predicting powder packing densities. *Powder Technology*, 174:10, 2007.

[53] Monika Bargiel. Geometrical properties of simulated packings of spherocylinders. In *Computational Science–ICCS2008*, volume 5102, pages 126–135, 2008.

[54] Alan Wouterse, Stefan Luding, and Albert P. Philipse. On contact numbers in random rod packings. *Granular Matter*, 11:169–177, 2009.

[55] Peng Lu, Shuixiang Li, J. Zhao, and L. Meng. A computational investigation on random packings of sphere-spherocylinder mixtures. *Science China Physics, Mechanics and Astronomy*, 53:2284–2292, 2010.

[56] Andriy V. Kyrylyuk, Marie Anne van de Haar, Laura Rossi, Alan Wouterse, and Albert P. Philipse. Isochoric ideality in jammed random packings of non-spherical granular matter. *Soft Matter*, 7:1671–1674, 2011.

[57] Claudia Ferreiro-Córdova and Jeroen S. van Duijneveldt. Random packing of hard spherocylinders. *Journal of Chemical and Engineering Data*, 59:3055–3060, 2014.

[58] Kumpei Shiraishi, Hideyuki Mizuno, and Atsushi Ikeda. Mechanical and vibrational properties of three-dimensional dimer packings near the jamming transition. *Journal of the Physical Society of Japan*, 89:074603, 2020.

[59] Salvatore Torquato and Yang Jiao. Dense packings of the platonic and archimedean solids. *Nature*, 460:876–879, 2009.

[60] Joost de Graaf, René van Roij, and Marjolein Dijkstra. Dense regular packings of irregular nonconvex particles. *Physical Review Letters*, 107:155501, 2011.

[61] Pablo Damasceno, Michael Engel, and Sharon Glotzer. Predictive self-assembly of polyhedra into complex structures. *Science (New York, N.Y.)*, 337:453–7, 2012.

[62] Salvatore Torquato and Yang Jiao. Organizing principles for dense packings of nonspherical hard particles: not all shapes are created equal. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 86:011102, 2012.

[63] Robert Shepherd, Jacinta Conrad, Tapan Sabuwala, Gustavo Gioia, and Jennifer A. Lewis. Structural evolution of cuboidal granular media. *Soft Matter*, 8:4795–4801, 2012.

[64] Max Neudecker, Stephan Ulrich, Stephan Herminghaus, and Matthias Schröter.

Jammed frictional tetrahedra are hyperstatic. *Physical Review Letters*, 111:028001, 2013.

[65] Anjan Prasad Gantapara, Joost de Graaf, René van Roij, and Marjolein Dijkstra. Phase diagram and structural diversity of a family of truncated cubes: degenerate close-packed structures and vacancy-rich states. *Physical Review Letters*, 111:015501, 2013.

[66] Elizabeth R. Chen, Daphne Klotsa, Michael Engel, Pablo F. Damasceno, and Sharon C. Glotzer. Complexity in surfaces of densest packings for families of polyhedra. *Physical Review X*, 4:011024, 2014.

[67] Lufeng Liu, Zhuoran Li, Yang Jiao, and Shuixiang Li. Maximally dense random packings of cubes and cuboids via a novel inverse packing method. *Soft Matter*, 13:748–757, 12 2017.

[68] Marc Z. Miskin and Heinrich M. Jaeger. Evolving design rules for the inverse granular packing problem. *Soft Matter*, 10:3708–3715, 2014.

[69] Lingyi Meng, Shuixiang Li, Peng Lu, Teng Li, and Weiwei Jin. Bending and elongation effects on the random packing of curved spherocylinders. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 86:061309, 2012.

[70] Lingyi Meng, Shuixiang Li, and Xiaohu Yao. Maximally dense random packings of intersecting spherocylinders with central symmetry. *Powder Technology*, 314:49–58, 2017.

[71] Lingyi Meng, Chao Wang, and Xiaohu Yao. Non-convex shape effects on the dense random packing properties of assembled rods. *Physica A: Statistical*

Mechanics and its Applications, 490:212–221, 2018.

[72] Martin Gardner. *The Colossal Book of Mathematics: Classic Puzzles, Paradoxes, and Problems.* Norton, 2001.

[73] Kumpei Shiraishi, Hideyuki Mizuno, and Atsushi Ikeda. Vibrational properties of two-dimensional dimer packings near the jamming transition. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 100(1):012606, 2019.

[74] Aleksandar Donev, Robert Connelly, Frank H. Stillinger, and Salvatore Torquato. Underconstrained jammed packings of nonspherical hard particles: Ellipses and ellipsoids. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 75:051304, 2007.

[75] Zorana Zeravcic, Ning Xu, Andrea J. Liu, Sidney R. Nagel, and Wim van Saarloos. Excitations of ellipsoid packings near jamming. *Europhysics Letters*, 87:26001, 2009.

[76] Mitch Mailman, Carl F. Schreck, Corey S. O'Hern, and Bulbul Chakraborty. Jamming in systems composed of frictionless ellipse-shaped particles. *Physical Review Letters*, 102:255501, 2009.

[77] Carl F. Schreck, Ning Xu, and Corey S. O'Hern. A comparison of jamming behavior in systems composed of dimer- and ellipse-shaped particles. *Soft Matter*, 6:2960–2969, 2010.

[78] Carl F. Schreck, Mitch Mailman, Bulbul Chakraborty, and Corey S. O'Hern. Constraints and vibrations in static packings of ellipsoidal particles. *Physical*

Review E: Statistical, Nonlinear, and Soft Matter Physics, 85:061305, 2012.

[79] Theodore A. Marschall and Stephen Teitel. Compression-driven jamming of athermal frictionless spherocylinders in two dimensions. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 97:012905, 2018.

[80] Kyle VanderWerf, Weiwei Jin, Mark D. Shattuck, and Corey S. O'Hern. Hypostatic jammed packings of frictionless nonspherical particles. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 97:012909, 2018.

[81] *LAMMPS*. URL: `http://lammps.sandia.gov`.

[82] A. E. R. Westman and H. R. Hugill. The packing of particles. *Journal of the American Ceramic Society*, 13(10):767–779, 1930.

[83] R. K. McGeary. Mechanical packing of spherical particles. *Journal of the American Ceramic Society*, 44:513–522, 1961.

[84] Stephen Yerazunis, Joseph W. Bartlett, and Alfred H. Nissan. Packing of binary mixtures of spheres and irregular particles. *Nature*, 195:33–35, 1962.

[85] Stephen Yerazunis, Stephen W. Cornell, and Barry Wintner. Dense random packing of binary mixtures of spheres. *Nature*, 207:835–837, 1965.

[86] Dongling He, Nduka N. Ekere, and Lele Cai. Computer simulation of random packing of unequal particles. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 60:7098–7104, 1999.

[87] Erik Santiso and Erich Müller. Dense packing of binary and polydisperse hard spheres. *Molecular Physics*, 100:2461–2469, 08 2002.

[88] Anuraag R. Kansal, Salvatore Torquato, and Frank H. Stillinger. Computer generation of dense polydisperse sphere packings. *The Journal of Chemical*

Physics, 117:8212–8218, 2002.

[89] Alexander Bezrukov, Monika Bargiel, and Dietrich Stoyan. Statistical analysis of simulated random packings of spheres. *Particle and Particle Systems Characterization*, 19:111–118, 2002.

[90] Kai de Lange Kristiansen, Alan Wouterse, and Albert P. Philipse. Simulation of random packing of binary sphere mixtures by mechanical contraction. *Physica A: Statistical Mechanics and its Applications*, 358:249–262, 2005.

[91] Adam B. Hopkins, Yang Jiao, Frank H. Stillinger, and Salvatore Torquato. Phase diagram and structural diversity of the densest binary sphere packings. *Physical Review Letters*, 107:125501, 2011.

[92] Lingyi Meng, Peng Lu, and Shuixiang Li. Packing properties of binary mixtures in disordered sphere systems. *Particuology*, 16:155–166, 2014.

[93] Av Andriy Kyrylyuk, Alan Wouterse, and Albert P. Philipse. Random packings of rod-sphere mixtures simulated by mechanical contraction. In *AIP Conference Proceedings*, volume 1145, page 211, 2009.

[94] Av Andriy Kyrylyuk, Alan Wouterse, and Albert P. Philipse. Percolation and jamming in random heterogeneous materials with competing length scales. *Colloid and Polymer Science*, 137:29–33, 2010.

[95] Av Andriy Kyrylyuk and Albert P. Philipse. Effect of particle shape on the random packing density of amorphous solids. *Physica Status Solidi A*, 208:2299–2302, 2011.

[96] Lingyi Meng, Peng Lu, Shuixiang Li, J. Zhao, and Teng Li. Shape and size effects on the packing density of binary spherocylinders. *Powder Technology*,

228:284–294, 2012.

[97] Aibing Yu, Ruiping Zou, and Nicholas Standish. Modifying the linear packing model for predicting the porosity of nonspherical particle mixtures. *Industrial Engineering Chemistry Research*, 35:3730–3741, 1996.

[98] Ye Yuan, Lufeng Liu, Yuzhou Zhuang, Weiwei Jin, and Shuixiang Li. Coupling effects of particle size and shape on improving the density of disordered polydisperse packings. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 2018.

[99] Marc Z. Miskin and Heinrich M. Jaeger. Adapting granular materials through artificial evolution. *Nature materials*, 12 4:326–31, 2013.

[100] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.

[101] Alan Wouterse, Stephen R. Williams, and Albert Philipse. Effect of particle shape on the density and microstructure of random packings. *Journal of Physics: Condensed Matter*, 19:406215, 2007.

[102] Leonardo Silbert, Deniz Ertaş, Gary S. Grest, Thomas Halsey, Dov Levine, and Steven Plimpton. Granular flow down an inclined plane: Bagnold scaling and rheology. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 64:051302, 2001.

[103] Thomas King, Simon Butcher, and Lukasz Zalewski. *Apocrita - High Performance Computing Cluster for Queen Mary University of London*, 2017. URL: https://doi.org/10.5281/zenodo.438045.

[104] Albert P. Philipse. The random contact equation and its implications for

(colloidal) rods in packings, suspensions, and anisotropic powders. *Langmuir*, 12:1127–1133, 1996.

[105] Stefano Sacanna, Laura Rossi, Alan Wouterse, and Albert P. Philipse. Observation of a shape-dependent density maximum in random packings and glasses of colloidal silica ellipsoids. *Journal of Physics: Condensed Matter*, 19:376108, 2007.

[106] Youngkyu Han and Mahn Kim. Effect of bond and asymmetry of 2d-dumbbells on their structure in high concentration regime. *Soft Matter*, 8:9015–9021, 2012.

[107] Tomaso Aste, Tiziana di Matteo, Mohammad Saadatfar, Timothy John Senden, Matthias Schroter, and Harry L. Swinney. An invariant distribution in static granular media. *Europhysics Letters*, 79:24003, 2007.

[108] Fabian Schaller, Robert Weigel, and Sebastian Kapfer. Cuddling ellipsoids: Densest local structures of uniaxial ellipsoids. *Physical Review X*, 6:041032, 2016.

[109] Martin van Hecke. Jamming of soft particles: geometry, mechanics, scaling and isostaticity. *Journal of Physics: Condensed Matter*, 22:033101, 2010.

[110] Shlomo Alexander. Amorphous solids: Their structure, lattice dynamics and elasticity. *Physics Reports-review Section of Physics Letters*, 296:65–236, 1998.

[111] Jean-Noël Roux. Geometric origin of mechanical properties of granular materials. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*,

61:6802–36, 2000.

[112] Emilien Azéma, Farhang Radjaï, Baptiste Saint-Cyr, Jean-Yves Delenne, and Philippe Sornay. Rheology of three-dimensional packings of aggregates: Microstructure and effects of nonconvexity. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 87:052205, 2013.

[113] Pierre Gilles de Gennes and Jacques Prost. *The Physics of Liquid Crystal*. Clarendon Press, second edition, 1993.

[114] Patrick Richard, Luc Oger, Jean-Paul Troadec, and A. Gervois. Geometrical characterization of hard-sphere systems. *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 60:4551–8, 1999.

[115] Patrick Richard, A. Gervois, Luc Oger, and Jean-Paul Troadec. Order and disorder in hard-sphere packings. *Europhysics Letters*, 48:415–420, 1999.

[116] Thomas M. Truskett, Salvatore Torquato, and Pablo G. Debenedetti. Towards a quantification of disorder in materials: distinguishing equilibrium and glassy sphere packings. *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 62:993–1001, 2000.

[117] Tomaso Aste, Mohammad Saadatfar, and Tim Senden. Geometrical structure of disordered sphere packings. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 71:061302, 2005.

[118] Klaus Lochmann, Alexey V. Anikeenko, Antje Elsner, Nikolai N. Medvedev, and Dietrich Stoyan. Statistical verification of crystallization in hard sphere

packings under densification. *The European Physical Journal B: Condensed Matter and Complex Systems*, 53:67–76, 2006.

[119] Alan Wouterse and Albert P. Philipse. Geometrical cluster ensemble analysis of random sphere packings. *The Journal of Chemical Physics*, 125:194709, 2006.

[120] Yuliang Jin and Hernán Makse. A first-order phase transition at the random close packing of hard spheres. *Physica A: Statistical Mechanics and its Applications*, 389:5362–5379, 2010.

[121] Wen-Sheng Xu, Zhao-Yan Sun, and Li-Jia An. Dense packing in the monodisperse hard-sphere system: A numerical study. *The European Physical Journal E: Soft Matter and Biological Physics*, 31:377–82, 04 2010.

[122] Walter Mickel, Sebastian Kapfer, Gerd E. Schröder-Turk, and Klaus Mecke. Shortcomings of the bond orientational order parameters for the analysis of disordered particulate matter. *The Journal of Chemical Physics*, 138:044501, 2013.

[123] Anuraag Kansal, Salvatore Torquato, and Frank H. Stillinger. Diversity of order and densities in jammed hard-particle packings. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 66:041109, 2002.

[124] Hossein Eslami, Parvin Sedaghat, and Florian Müller-Plathe. Local bond order parameters for accurate determination of crystal structures in two and three dimensions. *Physical Chemistry Chemical Physics*, 20:27059–27068, 2018.

[125] Aleksandar Donev, Salvatore Torquato, and Frank H. Stillinger. Pair correlation function characteristics of nearly jammed disordered and ordered hard-sphere packings. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 71:011105, 2005.

[126] Carolina Brito, Harukuni Ikeda, Pierfrancesco Urbani, Matthieu Wyart, and Francesco Zamponi. Universality of jamming of nonspherical particles. *Proceedings of the National Academy of Sciences*, 115(46):11736–11741, 2018.

[127] Salvatore Torquato. *Random Heterogeneous Materials*. Springer, 2002.

[128] Parmesh Gajjar, Chris G. Johnson, James Andrew Carr, Kevin Chrispeels, J.M.N.T. Gray, and Philip J. Withers. Size segregation of irregular granular materials captured by time-resolved 3d imaging. *Scientific Reports*, 11, 2021.

[129] Martin J. Powell. Computer-simulated random packing of spheres. *Powder Technology*, 25(1):45–52, 1980.

[130] Michitaka Suzuki and Toshio Oshima. Co-ordination number of a multi-component randomly packed bed of spheres with size distribution. *Powder Technology*, 44(3):213–218, 1985.

[131] Luc Oger, Jean Paul Troadec, Daniel E. Bideau, John Allan Dodds, and Martin J. Powell. Properties of disordered sphere packings i. geometric structure: Statistical model, numerical simulations and experimental results. *Powder Technology*, 46:121–131, 1986.

[132] David J. Pinson, Ruiping Zou, Aibing Yu, Paul Zulli, and M.J. McCarthy. Coordination number of binary mixtures of spheres. *Journal of Physics D:*

Applied Physics, 31:457–462, 1998.

[133] Ruiping Zou, Xiangyang Bian, David J. Pinson, Ruzhen Yang, Aibing Yu, and Paul Zulli. Coordination number of ternary mixtures of spheres. *Particle and Particle Systems Characterization*, 20:335–341, 2003.

[134] Denis Weaire and Tomaso Aste. *The pursuit of perfect packing.* CRC Press, 2008.

[135] Jonathan Schmidt, Mário R. G. Marques, Silvana Botti, and Miguel A. L. Marques. Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials*, 5:1–36, 2019.

[136] Ekin Dogus Cubuk, Samuel S. Schoenholz, Jennifer M. Rieser, Brad D. Malone, Joerg Rottler, Douglas J. Durian, Efthimios Kaxiras, and Andrea J. Liu. Identifying structural flow defects in disordered solids using machine-learning methods. *Physical Review Letters*, 114:108001, 2015.

[137] Ignacio González Tejada and Pablo Antolin. Use of machine learning for unravelling hidden correlations between particle size distributions and the mechanical behavior of granular materials. *Acta Geotechnica*, 17:1443 – 1461, 2022.

[138] Joost van der Linden, Guillermo Narsilio, and Antoinette Tordesillas. Machine learning framework for analysis of transport through complex networks in porous, granular media: A focus on permeability. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 94:022904, 2016.

[139] Nikolaus Hansen, Sibylle D. Müller, and Petros Koumoutsakos. Reducing

the time complexity of the derandomized evolution strategy with covariance matrix adaptation (cma-es). *Evolutionary Computation*, 11:1–18, 2003.

[140] Karl Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

[141] Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:498–520, 1933.

[142] Ian Jolliffe. *Principal component analysis*. Springer Verlag, 2002.

[143] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In *Advances in Kernel Methods-Support Vector Learning*, pages 327–352. MIT Press, 1999.

[144] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5):1299–1319, 1998.

[145] Sebastian Mika, Bernhard Schölkopf, Alex Smola, Klaus-Robert Müller, Matthias Scholz, and Gunnar Rätsch. Kernel pca and de-noising in feature spaces. In *Advances in Neural Information Processing Systems 11*, pages 536–542. MIT Press, 1999.

[146] Gökhan Bakir, Jason Weston, and Bernhard Schölkopf. Learning to find pre-images. In *Advances in Neural Information Processing Systems*, pages 449–456. MIT Press, 2004.

[147] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer New York Inc., New York, USA, 2008.

[148] Leo Breiman. Bagging predictors. *Machine Learning*, 24:123–140, 1996.

[149] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282, 1995.

[150] Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:832–844, 1998.

[151] Thomas Dietterich. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning*, 40:139–157, 2000.

[152] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.

[153] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Betrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Edouard Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[154] Franz Aurenhammer. Power diagrams: Properties, algorithms and applications. *SIAM Journal on Computing*, 16:78–96, 1987.

[155] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones,

Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[156] Franz Aurenhammer. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23:345–405, 1991.

[157] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, Sung Nok Chiu, and David Kendall. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams.* Wiley- Blackwell, second edition, 2000.

[158] Louis Portal, Maximilien Danisch, Adrian Baule, Romain Mari, and Hernán A. Makse. Calculation of the voronoi boundary for lens-shaped particles and spherocylinders. *Journal of Statistical Mechanics: Theory and Experiment*, 2013:11009, 2013.

[159] Fabian M. Schaller, Sebastian C. Kapfer, Myfanwy E. Evans, Matthias Hoffmann, Tomaso Aste, Mohammad Saadatfar, Klaus Mecke, Gary W. Delaney, and Gerd E. Schröder-Turk. Set voronoi diagrams of 3d assemblies of aspherical particles. *Philosophical Magazine*, 93:3993 – 4017, 2013.

[160] Maximilien Danisch, Yuliang Jin, and Hernán A. Makse. Model of random packings of different size balls. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 81:051303, 2010.