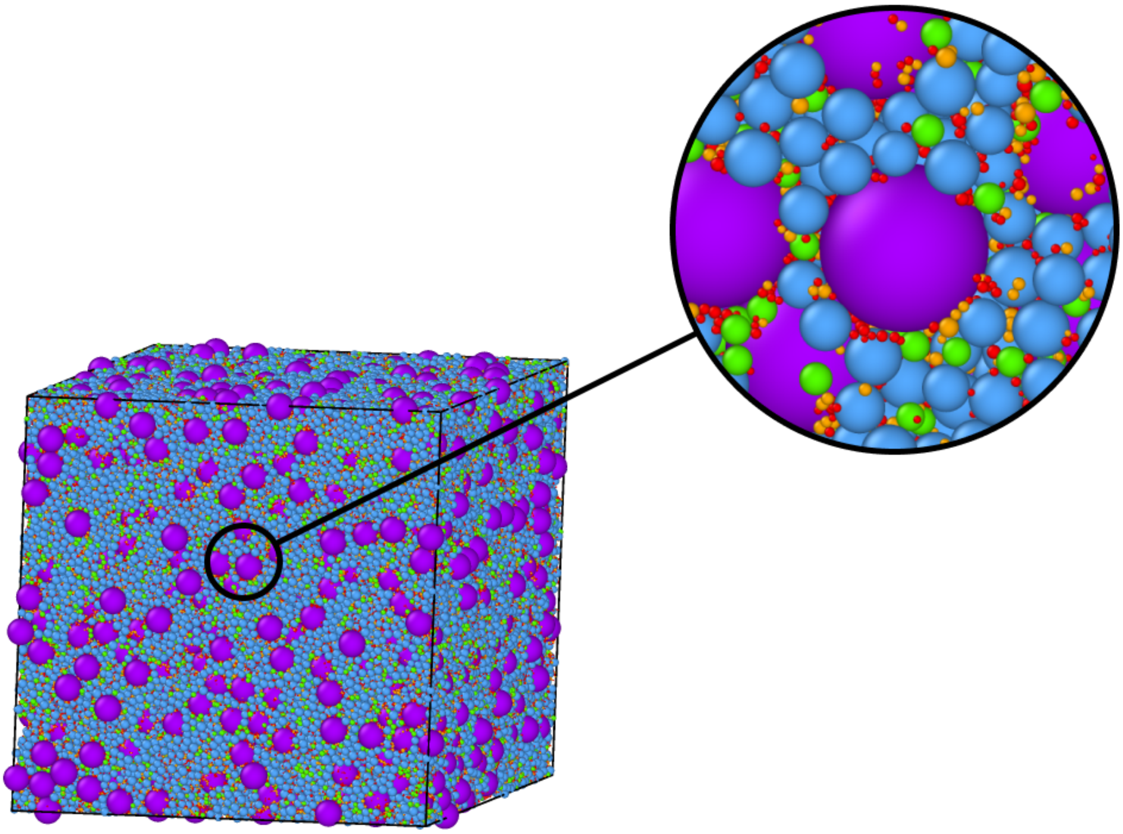


Bulk Properties of Highly Polydisperse Particle Assemblies

Master's thesis in Engineering Mathematics and Computational Science

IVAN FLENSBURG
FILIP WESTBERG

DEPARTMENT OF MATHEMATICAL SCIENCES

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2023
www.chalmers.se

Master's thesis 2023

Bulk Properties of Highly Polydisperse Particle Assemblies

FILIP WESTBERG
IVAN FLENSBURG

Department of Mathematical Sciences
Chalmers University of Technology
Gothenburg, Sweden 2023

Bulk Properties of Highly Polydisperse Particle Assemblies
FILIP WESTBERG
IVAN FLENSBURG

Supervisor: Luis Martin de Juan & Astrid Boje, Oral Product Development, Pharmaceutical Technology & Development, AstraZeneca, Gothenburg, Sweden
Examiner: Alexey Geynts, Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden

Master's Thesis 2023
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A multi-component polydisperse particle packing simulated with DEM.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2023

Bulk Properties of Highly Polydisperse Particle Assemblies
FILIP WESTBERG
IVAN FLENSBURG
Department of Mathematical Sciences
Chalmers University of Technology

Abstract

Pharmaceutical oral drug products often come in the form of highly polydisperse multi-component granular powders. Drug development is limited by experimental testing, making in silico methods a valuable alternative approach. To this end, a framework was developed using Discrete Element Method which simulates arbitrary assemblies at or below jamming density, of spherical frictionless particles.

Analytic formulations for the component concentration homogeneity are compared against simulation results, where that of Stange proved to be more appropriate than Hilden. The packing fraction variability influences the content homogeneity. This variability was then compared to that of a packing model, and show some proportionality, though requiring further work to model one through the other. A simulation scale should have been formulated through a maximum packing fraction variability and content inhomogeneity, as its lack was visible due to its effect on the content homogeneity.

Keywords: Discrete Element Method, Content Uniformity, highly polydisperse

Acknowledgements

We'd like to thank our supervisors Luis, Astrid, our examiner Alexey and Sohan for sharing their considerable knowledge and guidance along the way.

This project has been funded by AstraZeneca.

Filip Westberg, Ivan Flensburg, Gothenburg, July 2023

List of Acronyms

Below is the list of acronyms that have been used throughout this thesis listed in alphabetical order:

COR	Coefficient of Restitution
COV	Coefficient of Variation
DEM	Discrete Element Method
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator
LMPM	Linear Mixing Packing Model
PSD	Particle Size Distribution
UDU	Unit-Dose-Uniformity
VDW	Van der Waal
API	Active Pharmaceutical Ingredient
NPH	The NPH ensemble
RCP	Random Close Packing
MC	Master Curve

Contents

List of Figures

List of Tables

1

Introduction

1.1 Background

Within the pharmaceutical industry, the process of developing a new product requires thorough testing and validation. Oral products, such as capsules and tablet commonly consist of polydisperse granular material within the µm-scale, which can be divided into the active pharmaceutical ingredients (API) and excipients. Excipients are generally used for the sake of stability, dose accuracy and release among other purpouses, and often compose most of the weight of a pharmaceutical product [1]. The amount of API in each particular capsule or tablet is strictly controlled, with limits on dose variability [2]. The problem of assuring this unit-dose-uniformity (UDU) is magnified for low-dose products, where individual grains of API may compose non-trivial fractions of the total API in a dose. The properties of the particle size distributions (PSD) and component densities may then be essential in characterizing the UDU of a pharmaceutical product, and can be used to estimate it [3].

In silico capability to simulate and predict properties of pharmaceutical products can accelerate product development and reduce the need for experimental testing, which may ultimately reduce costs and yield more sustainable products. The Discrete element method (DEM) [4] is a widely used framework to simulate an assembly of particles in which the dynamics of individual particles and their interactions with neighboring particles are given by Newton's equations of motion. Simulation entails defining initial conditions and force laws between interacting particles, construction of neighbor lists keeping track of particle-particle interactions and subsequently updating particle motion and position using time integration. The computational intensity of updating neighbor lists over the entire simulation run using the most common frameworks has been an issue for polydisperse packings, giving rise to the development of a new method [5]. This was recently implemented in the widely used open-source software for DEM, LAMMPS (Large-scale Atomic/Molecular Massively Parallel simulator) [6], enabling the study of arbitrarily polydisperse packings in a feasible timeframe [7].

1.2 Aim

This thesis aims to create a framework which enables the study of bulk product characteristics of arbitrary polydisperse multi-component particle packings. This is achieved by using the DEM framework through the software LAMMPS to simulate particle packings. The generated packings are then used to study bulk characteristics. In the context of pharmaceutical drug development these are:

- Coordination number between components
- Coordination number distributions
- Measurement of homogeneity of final packing
- Estimation of packing fraction variation and UDU and comparison with simulation results.

1.3 Assumptions

In order to achieve results within the designated timespan for this work some delimitations were made, simplifying the work process. Although many granular materials of interest have irregular shapes, all particles are represented as spheres due its simple shape. Furthermore, every particle component is characterized by a density, mass fraciton and PSD which is limited to a Log-Normal distribution given the provided expected diameter and variance.

Particle motion does not include twisting or rolling, and material specific static and dynamic friction are not considered. Any type of friction would require validation data and is thus disregarded. Moreover, particles are assumed to always maintain its shape. This includes particles retaining their shape after a deformation during a particle-particle collision.

2

Theory

2.1 Granular material

Granular material is any material composed of many individial solid particles, irrespective of particle size [8], though in the context of granular materials within pharmaceutical products, particles are often defined in microns (µm). A collection of particles with such sizes are also called granular powders, following the size classification in [8]. A material consisting of granular powder often comes in widely varying sizes, and thus necessitates that the material is described in terms of a PSD rather than a mean particle size.

At sizes $< 100\,\mu\text{m}$ particles are found to resist gravitational forces and adhere to other particles due to interparticle forces such as Van der Waals forces (VDW) [9], making adhesion and cohesion an important characteristic of granular powder. While at a scale where VDW forces have a considerable effect, granular powders are unaffected by temperature due to its insignificant energy scale in relation to potential energy by gravity. Furthermore particle-particle interactions are highly dissipative due to static friction and inelastic collisions [10].

Typical of granular material is the broad distribution of contact forces, and that often a subset of particles are under stress [11]. Applied load is largely transmitted by a subset of particles, with forces several times larger than others occuring even under isotropic compression [12].

2.1.1 Particle size distribution

Granular material in pharmaceutical products are often described using Log-normal ($\mathcal{LN}$) distributions. For multi-component packings, each components' particle size distribution is represented by a $\mathcal{LN}$-distribution defined here as

$$\mathcal{LN}(x;\mu,\sigma) = \frac{1}{x\sqrt{2\pi}\sigma}\exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right),$$

with distribution parameters μ and σ. Given a granular material with expected diameter μ_X and variance σ_X^2, the distribution parameters are given by

$$\mu = \ln\left(\frac{\mu_X^2}{\sqrt{\mu_X^2+\sigma_X^2}}\right), \tag{2.1}$$

$$\sigma^2 = \ln\left(1+\frac{\sigma_X^2}{\mu_X^2}\right). \tag{2.2}$$

2.1.2 Coordination number

An important characteristic of granular materials is its coordination number, which is the number of contacts for a given particle. In the case of multi-component packings it is of high interest to identify the contact distribution between components in order to draw conclusions about how they interact with each other, where an example could be the contact distribution of the API in a pharmaceutical product. Following the definitions in [13], a packing with N particles and M unique contacts the classical definition of coordination numbers, also denoted *global* coordination number is given by

$$Z = \frac{M}{N}. \tag{2.3}$$

This also includes particles which have too few contacts to contribute to the structural stability of the packing. Such particles are denoted *rattlers*. For frictionless spheres, rattlers can be identified by the 1-stable condition [14], which states that each sphere has to have at least 4 contacts such that the contacting spheres center coordinates can be separated by a 2D-plane passing through the center coordinate of the reference particle. For the latter condition there exists linear programming solutions [15], although in this report this is disregarded for the purposes of simplification. Instead rattlers are determined only based on the number of contacts, independent of the geometrical location of contacting particles. Thus a corrected coordination number, also denoted *non-rattler* coordination number in [16] is of interest, defined as

$$Z_{\mathrm{nr}} = \frac{M_{Z\geq 4}}{N_{Z\geq 4}}, \tag{2.4}$$

where $N_{Z\geq 4}$ is total number of particles with at least 4 contacts, and $M_{Z\geq 4}$ denotes total number of contacts of particles with at least 4 contacts.

2.1.3 Jamming

For granular materials, the point at which its behaviour transitions from a fluid to a solid, with increasing packing fraction is generally called jamming, often accompanied by a rapid increase of the coordination number [17]. The system of particles is able to sustain non-zero applied stress beyond the jamming point [11]. The lowest packing fraction at which this occurs is denoted by ϕ_J [18]. In this setting, the packing fraction ϕ_J is shown to correspond to the random close packing, ϕ_{RCP}, defined as the maximum volume fraction of solid spheres when packed randomly. A common example of this is $\phi_{RCP} \approx 0.64$ for monosized spheres [18].

A jammed state is one of mechanical equilibrium, where the net force on particles are zero. One key parameter characterising the state of a particle system is the coordination number of particles contributing to the structural stability, i.e. non-rattlers. Given the global coordination number Z, the number of unknown contact forces equal $\frac{NZ}{2}$. Then for force balance equations to hold in a 3D system $\frac{NZ}{2} \geq 3N$ where for *isostatic* packings $Z = 6$ [11].

2.2 DEM modeling

Originally formulated by Cundall & Strack [4], DEM is a framework used to simulate the dynamics of numerous bodies of arbitrary shape and deformability. Either the body is deformable or rigid, also referred to as soft and hard formulations. Hard-sphere formulation involves instantaneous and thus only pairwise collisions, resulting in fewer force evaluations than the soft-sphere formulation. It does however lose applicability in denser systems where multiple and long-lasting contacts are expected between particles [19]. The soft-sphere model is more applicable in the case of granular materials due to their elasticity and deformability, and is thus what is used in this report. The formulation entails deformable particles, but the deformation is reversible and original shapes are retained once the particle loses its contact. Particle-particle interactions occur through their shared contact area and the magnitude of contact force is given by force-displacement laws which is a function of overlap, relative velocity and material properties [19].

2.2.1 Contact model

The contact model determines how particles interact with each other, i.e. force vectors acting on particles. The dynamics of non-rotating frictionless particles are described by Newton's second law of motion which states that for a particle with constant mass m_i, velocity $\vec{v}_i$ and center of mass coordinate $\vec{x}_i$ its acceleration at timestep t is given by

$$m_i \frac{d\vec{v}_i(t)}{dt} = m_i \frac{d^2\vec{x}_i(t)}{dt^2} = \sum_{j \in \mathcal{L}_i(t)} \vec{F}_{ij}^{n}, \tag{2.5}$$

where $\vec{F}_{ij}^{n}$ denotes normal forces summed over all particles j present within the neighbor list $\mathcal{L}_i(t)$ at time t for particle i.

In this report two different contact models are considered, namely that of Hertz [20, 21, 22] and Johnson-Kendall-Roberts (JKR) [23]. The Hertzian contact model explains the force-displacement relationship of elastic spheres such that for a pair of contacting spheres i, j the normal force is given by

$$\vec{F}_{ij}^{n} = \vec{F}_{ij}^{ne,hertz} + \vec{F}_{ij}^{nd},$$

where $\vec{F}_{ij}^{ne,hertz}$ is the normal elastic component relating to the shared contact area, and $\vec{F}_{ij}^{nd}$ is the normal viscous damping component causing dissipative collisions. However, the Hertzian contact model does not capture the adhesive property of granular powder, which requires a more complex model. The theoretical contact model JKR adresses this issue by extending on the Hertizan model using an adhesive force so that the normal force is given by

$$\vec{F}_{ij}^{n} = \vec{F}_{ij}^{ne,jkr} + \vec{F}_{ij}^{ne,adh} + \vec{F}_{ij}^{nd},$$

where the normal elastic component $\vec{F}_{ij}^{ne,jkr}$ relates to the shared contact area in a similar manner to the Hertizan model, $\vec{F}_{ij}^{ne,adh}$ is the normal adhesive force component, and $\vec{F}_{ij}^{nd}$ is the damping component. The following sections aims to describe the framework of the Hertz and JKR models in further detail.

2.2.1.1 Hertz contact model

The contact mechanics for particles are described using *effective* properties of their mass m, radius R, Poisson ratio v and Young's modulus E. For two contacting particles i, j, these effective quantities are

$$\begin{aligned}\frac{1}{m} &= \frac{1}{m_i} + \frac{1}{m_j},\\ \frac{1}{R} &= \frac{1}{R_i} + \frac{1}{R_j},\\ \frac{1}{E} &= \frac{1-v_i^2}{E_i} + \frac{1-v_j^2}{E_j}\end{aligned}$$

Hertzian contact theory predicts a semi-ellipsoidal normal pressure distribution $p(r)$ over the contact area with radius a

$$p(r) = p_0\left[1 - \frac{r^2}{a^2}\right]^{1/2}.$$

with p_0 as unknown - termed the maximum Hertz pressure. The normal displacement of sphere i over the contact area is then given by

$$u_i(r) = \frac{\pi p_0(1-v_i^2)}{4aE_i}(2a^2 - r^2), \tag{2.6}$$

so that Eq. 2.6 for spheres i and j satisfies the boundary condition

$$u_i(r) + u_j(r) = \delta - \frac{r^2}{R}, \tag{2.7}$$

where δ is the summed overlap distance of both spheres. Substituting for the effective properties in Eq. 2.7 yields the boundary condition

$$\frac{4p_0}{4aE}(2a^2 - r^2) = \delta - \frac{r^2}{R^2}.$$

With $r = 0$, the overlap is $\delta = \frac{\pi p_0 a}{2E}$. Subtituting this into the above equation, the contact radius becomes

$$\frac{4p_0}{4aE}(2a^2 - r^2) = \frac{\pi p_0 a}{2E} - \frac{r^2}{R^2} \ : \ a = \frac{\pi p_0 R}{2E}.$$

The contact radius can be expressed in terms of the overlap such that

$$\begin{aligned}a &= \frac{\delta R}{a}\\ a^2 &= R\delta.\end{aligned}$$

The normal contact force, i.e. normal elastic component is then calculated from integrating the normal pressure over the contact area

$$F_{ne} = 2\pi\int_0^a p(r)rdr = \frac{2}{3}p_0\pi a^2.$$

Substituting this into the expression for a, then

$$a^3 = \frac{3F_{ne}R}{4E},$$

and finally reordering the equation yields the expression for the normal elastic component

$$F_{ne} = \frac{4Ea^3}{3R} = \frac{4E\sqrt{R\delta^3}}{3}.$$

The normal damping constant using the Tsuji model [24] is given by

$$\eta_n = \alpha\sqrt{\frac{4mEa}{3}}. \tag{2.8}$$

Marshall [25] showed that the normal coefficient of restitution (COR) e, was only dependent on the damping parameter α. For a given COR e, the damping parameter could then be adequately approximated by a sixth degree polynomial.

$$\alpha(e) = 1.2728 - 4.2783e + 11.087e^2 - 22.348e^3 + 27.467e^4 - 18.022e^5 + 4.8218e^6. \tag{2.9}$$

The duration of impact, without damping, was derived by [26] so that given two colliding particles i, j the collision time is

$$\Delta t_c = 2.865\left[\frac{m^2}{RE^2|\vec{v}_{n,rel}|}\right]^{1/5}. \tag{2.10}$$

2.2.1.2 Johnson-Kendall-Roberts contact model

The JKR contact model for DEM incorporates the effect of adhesion, by including an attractive normal force, dependent on a specified surface energy γ. The contact normal force, consisting of the normal elastic component and adhesive force is expressed as

$$F_n = \frac{4Ea^3}{3R} - 2\pi a^2\sqrt{\frac{4\gamma E}{\pi a}},$$

where γ is the surface energy of the interaction and a is the radius of the contact zone, implicitly defined by

$$\delta_n = \frac{a^2}{R} - 2\sqrt{\frac{\pi\gamma a}{E}}.$$

Note that for $\gamma = 0$, the contact radius reduces to the Hertzian case $a^2 = \delta_n R$. The surface energy of particle interactions is in turn calculated through geometric mixing

$$\gamma = \sqrt{\gamma_i\gamma_j}$$

The critical overlap $\delta_c = \frac{a_0^2}{2(6)^{1/3}R}$ is the particle overlap at the point of no applied external force $F_n = F_c$. The parameter a_0 is the area of contact under no external load, given by

$$a_0 = \left(\frac{9\pi\gamma R^2}{E}\right)^{1/3}.$$

The relations between the contact force F_n and the contact area a, as well as between the overlap (relative approach) δ_n and critical overlap δ_c are [25]

$$\frac{F_n}{F_c} = 4\left[\left(\frac{a}{a_0}\right)^3 - \left(\frac{a}{a_0}\right)^{3/2}\right] \tag{2.11}$$

$$\frac{\delta_n}{\delta_c} = 6^{1/3}\left[2\left(\frac{a}{a_0}\right)^2 - \frac{4}{3}\left(\frac{a}{a_0}\right)^{1/2}\right] \tag{2.12}$$

where $F_n \geq -F_c$ and $F_c = 3\pi\gamma R$ is the *pulloff force*, the maximum tensile force during contact, the maximum tensile force between the particles [27]. Fig. 2.1 below shows the normalized force F_n/F_c and contact area a/a_0 versus the normalized overlap distance δ_n/δ_c.

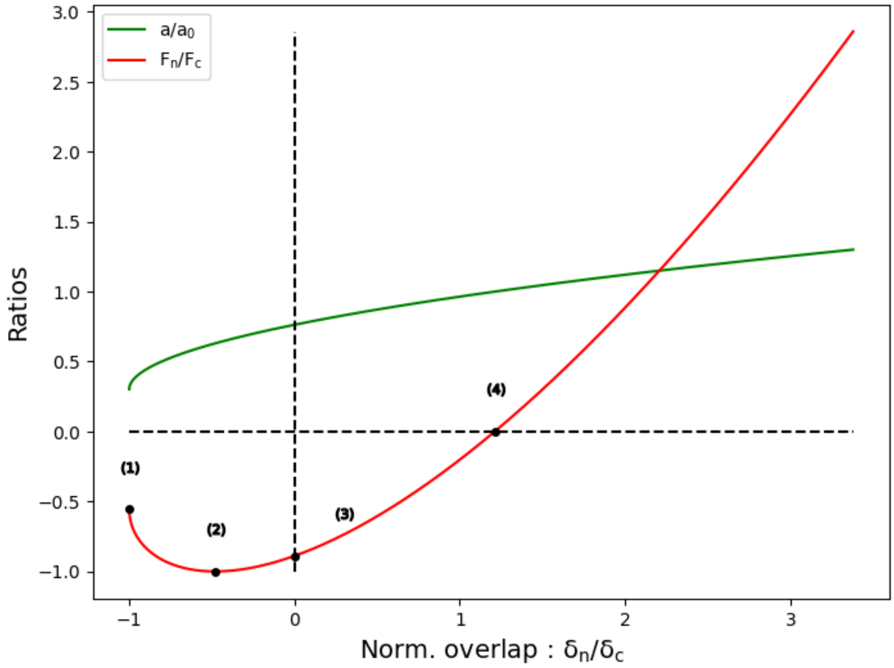


Figure 2.1: Ratios of F_n/F_c and a/a_0 plotted against the normalized overlap δ_n/δ_c. The marked points correspond to: (1) the tear-off distance at particle separation, (2) the maximum tensile force $F_n = -F_c$, (3) initial contact $\delta_n = 0$ and (4) equilibrium point $F_n = 0$.

If elastic dissipation is disregarded, all kinetic energy has been regained at the point of zero overlap (3). The hysteretic tensile force between (3) and (1) then induces a first-contact energy loss W, calculated as

$$W = 7.09\left(\frac{w^5 R^4}{E^2}\right)^{1/3},$$

where $w = 2\gamma$, [28, 29]. A lower bound for the sticking velocity given this energy dissipation is then given by (see Appendix A.1 for details)

$$v_s = 1.483\left[\frac{w^5}{r_{max}^5 E^2}\right]^{1/6}.$$

The surface energy is specified through constraining the maximum relative overlap at equilibrium $\tilde{\delta}_c \leq \epsilon$ (see Appendix A.1 for details) so that

$$\gamma = \frac{E}{2}\left[\frac{\epsilon}{A}\right]^{3/2}.$$

2.2.2 Contact detection

In order to calculate the particle-particle interactions in a computationally feasible time, it is imperative that the construction of contact lists is done in an efficient manner. A commonly method is to divide the 3D domain containing the particles into same-sized cubic cells with side-length s, such that the cell is large enough to inscribe the largest particle. All positions of the particles $\vec{x}_i = (x_i, y_i, z_i)^\top$ are then mapped to the cell corresponding to the integer coordinate

$$\vec{c}_i = \Big(\lfloor x_i/s \rfloor, \lfloor y_i/s \rfloor, \lfloor z_i/s \rfloor\Big)^\top. \tag{2.13}$$

Neighbors can then be identified by including particles within the same cell as the reference particle, and its neighboring cells. By implementing Verlet style neighbor lists [30], a subset of all neighboring particles construct the neighbor list if their distance to the reference particle is smaller than the neighbor cutoff distance

$$R_n = R_f + R_s, \tag{2.14}$$

where R_f is the force cutoff distance and R_s is the skin distance displayed in Fig. 2.2. The skin allows same neighbor list to be used over multiple timesteps, given that particles have not moved too far. A large skin yields larger neighbor lists and thus higher memory usage, but a lower reneighboring frequency. Usually the reneighboring occurs every few timesteps or when a particle within a neighbor list has moved a certain distance.

Fig. 2.2 shows an example of how neighbor lists are constructed for a single-component monodisperse packing. The force cutoff distance is set to each particles' radius. The cell size s is set by the diameter of the largest present particle plus the skin distance. Neighboring particles are collected from the grey region. The neighbor list is narrowed down by only collecting particles with intersecting skins marked as dark-green dashed lines.

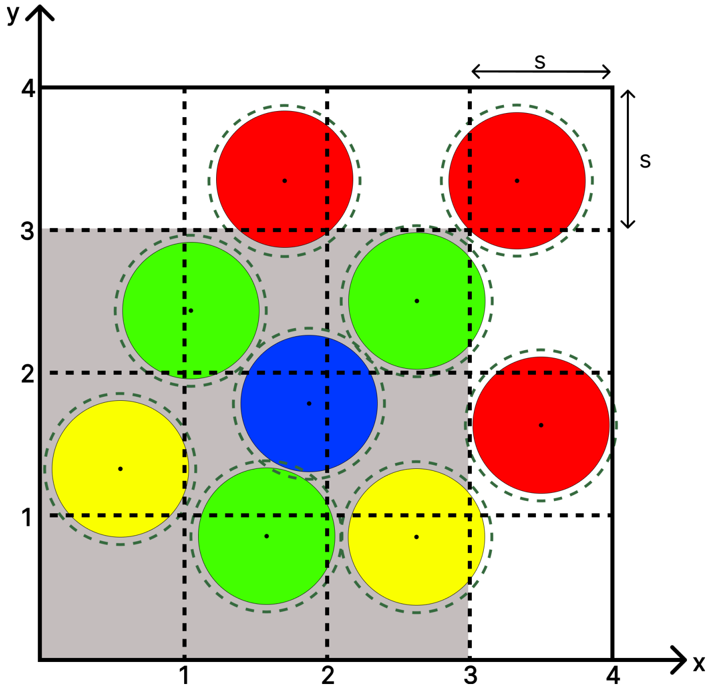


Figure 2.2: Neighbor list construction of a monodisperse packing in 2D. Dashed black lines represent cell boundaries with width $s = 1$. Neighbor check region is represented by grey coloring. Dashed dark-green lines represent skin distance. Reference particle is blue and is located in cell $(1, 1)$. Particles not included in check region are colored red. Particles included in check region but not in neighbor lists are colored yellow. Particles included in neighbor lists are colored green.

While above mentioned framework is efficient for monodisperse systems, polydisperse systems may result in an unnecessarily large number of neighbor checks due to the largest particle dictating the size of each cell, i.e. more particles will be contained within each cell. The issue can be alleviated by implementing a hierarchical grid [5], which works with the same principle as shown in Fig. 2.2, but L grids are stacked over the same domain with each level $h \in [1, L]$ having varying cell sizes s_h. Integer representation of a particle at level h is similar to Eq. 2.13 so that

$$\vec{c}_i = \Big(\lfloor x_i / s_h \rfloor, \lfloor y_i / s_h \rfloor, \lfloor z_i / s_h \rfloor, h \Big)^\top. \tag{2.15}$$

For particle i, its level of insertion $h(i)$, i.e. at which level h in the hierarchy the particle will be located, is determined depending on the particle radius r_i so that

$$h(i) = \left\{ \min_{1 \leq \mathrm{h} \leq \mathrm{L}} \mathrm{h} : \mathrm{s_h} \geq 2\mathrm{r_i} \right\}.$$

Contact detection is performed in two steps. First, the cell of particle i and its neighbors at level h are searched in the same manner as previously described, from which neighbor lists are constructed using Eq. 2.14. Second, a cross-level search is performed where neighbor lists are constructed based on particles at lower levels of insertion than the rerefence particle i, i.e. levels $1 \leq h < h(i)$. The cross-level search region $\mathcal{R}_i$ for particle i at level $h < h(i)$ is defined as

$$\mathcal{R}_i = \Big\{(x, y, z) \in \mathbb{R}^3 \;\Big|\; |x - x_i| \le \alpha,\ |y - y_i| \le \alpha,\ |z - z_i| \le \alpha\Big\},$$

where $\alpha = r_i + 0.5 s_h$. The coordinates are subsequently mapped to cells for level h using Eq. 2.15. Thus, same-level particles are searched for as described previously, lower-level particles are searched using cross-search, and higher-level particles are not searched at all, avoiding unnecessary double searches of particle pairs.

Fig. 2.3 shows an example of how multi-level neighbor lists are constructed for a bidisperse packing. As previously, force cutoff is set to particle radius. Cell sizes s_1 and s_2 are set by the diameter of the largest present particle plus the skin distance for each level $h \in [1, 2]$. Neighboring particles are collected from the shaded area. Neighbor list are constructed using the Verlet method.

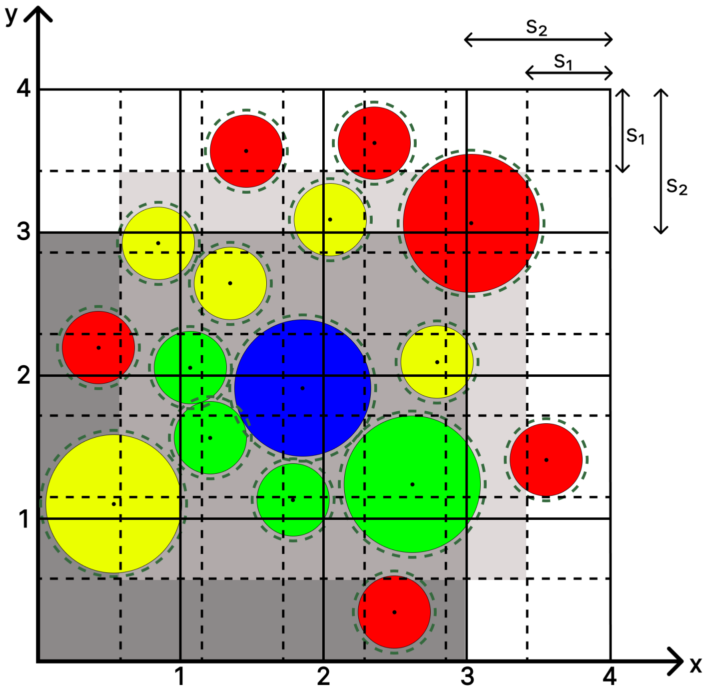


Figure 2.3: Neighbor list construction of a bidisperse packing in 2D using multigrid. Black lines are cell boundaries for level 2 with width $s_2 = 1$ and dashed lines are cell boundaries for level 1 with width $s_1 = 4/7$. Dark grey is neighbor check region for same level particles. Light grey is cross-level check region $\mathcal{R}$, where boundaries are rounded upwards to nearest cell boundary of level 1. Dashed dark-green lines represent skin distance. Reference particle is blue located in cell (1, 1, 2). Particles not included in check regions for its belonging level are colored red. Particles included in check region but not in neighbor lists are colored yellow. Particles included in neighbor lists are colored green.

In the bidisperse case it is intuitive to use a hierarchy with two levels, having cell size s_h based on the diameter of each particle type. However, in a general polydisperse case it is not as trivial determining the number of levels and the cell size for each level. More information regarding selection of these parameters and how this is implemented can be found in Appendix B.2. Furthermore a simple benchmark was conducted and is presented in Appendix C.3 between the two previously described methods, showing that the multilevel method is superior for increasing polydispersities.

2.2.3 Time integration

Given that contact lists and forces can be calculated on each particle, its position and velocity is updated for the next timestep using explicit time integration. The timestep size Δt should be short enough such that any particle-particle collisions occurs over multiple timesteps to ensure a high resolution. To this end, a critical timestep $\Delta t_{crit} \geq \Delta t$ needs to be determined, which is dependent on the system defined contact model.

There are many integration methods available which comes with its own benefits and drawbacks with respect to numerical accuracy, computational intensity through number of force evaluations and extra variables. A common integration method with good numerical stability at reasonable computational cost is the multi-step method Velocity-Verlet [30]. By rearranging Eq. 2.5 to express the change in velocity of particle i at timestep t yields

$$\vec{a}_i(t) = \frac{d\vec{v}_i(t)}{dt} = \frac{d^2\vec{x}_i(t)}{dt^2} = \frac{1}{m_i} \sum_{j \in \mathcal{L}_i(t)} \vec{F}^n_{ij}.$$

Then the position for the next timestep is given by

$$\vec{x}_i(t + \Delta t) = \vec{x}_i(t) + \vec{v}_i(t)\Delta t + \frac{1}{2}\vec{a}_i(t)\Delta t^2.$$

Given that the normal force is a function of overlap which is kept track of through the neighbor list $\mathcal{L}_i$ the updated acceleration is given by

$$\vec{a}_i(t + \Delta t) = \frac{1}{m_i} \sum_{j \in \mathcal{L}_i(t+\Delta t)} \vec{F}^n_{ij}.$$

Finally the updated velocity is

$$\vec{v}_i(t + \Delta t) = \vec{v}_i(t) + \frac{1}{2}\Big(\vec{a}_i(t) + \vec{a}_i(t + \Delta t)\Big)\Delta t$$

2.3 LAMMPS

LAMMPS is a classical, parallell molecular dynamics software written in C++, released under the GNU general public licence (GPL) [6]. The software is parallelisable by dividing the simulation box into processor subdomains with distributed memory (using MPI). The simulation box can be orthogonal or triclinic, as in this report. Each processor stores necessary information about its owned atoms (particles), and for atoms in other processor subdomains within a cutoff, so called *ghost atoms*. Each processor has a Verlet neighbor list as described in Sec. 2.2.2, indicating particles that are within a certain distance of each other. How often the neighbor list is rebuilt has significant influence on computation times, which can be varied through the skin distance. The default neighbor cutoff distance used is formulated in Eq. 2.14, allowing for the neighbor list to be reused over several timesteps. A *pair style* is the particular force model used, corresponding to the contact models defined in Sec. 2.2.1.1 and 2.2.1.2. Velocities and positions of particles are derived in upcoming timesteps by performing numerical integration using Velocity-Verlet, as described in Sec. 2.2.3.

LAMMPS calculations are executed through input script text files, which supports programming-like commands as described in the documentation. These commands can be categorized by, but not limited to

- *atom*: Per-atom properties (Atom represents particle in this report).
- *pair*: Pairwise potentials and properties between atoms.
- *fix*: Operations that are applied during timestepping.
- *compute*: Simple mathematical calculations.
- *dump*: Output of simulation handling.

2.3.1 Triclinic and periodic cell

The triclinic cell in LAMMPS is a parallelepiped with its origin in $(x_{\mathrm{lo}}, y_{\mathrm{lo}}, z_{\mathrm{lo}})$ and spanned by the edge column vectors $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$, defining the matrix

$$\mathbf{H} = \left[\vec{a}, \vec{b}, \vec{c}\right] = \begin{bmatrix} \Delta x & xy & xz \\ 0 & \Delta y & yz \\ 0 & 0 & \Delta z \end{bmatrix}, \tag{2.16}$$

where $\Delta x = x_{\mathrm{hi}} - x_{\mathrm{lo}}, \Delta y = y_{\mathrm{hi}} - y_{\mathrm{lo}}, \Delta z = z_{\mathrm{hi}} - z_{\mathrm{lo}}$ and xy, xz, yz are the tilt factors. Fig. 2.4 shows a visual representation of the triclinic cell defined by its edge vectors $\vec{a}, \vec{b}$ and $\vec{c}$ and including internal cell angles α, β, γ.

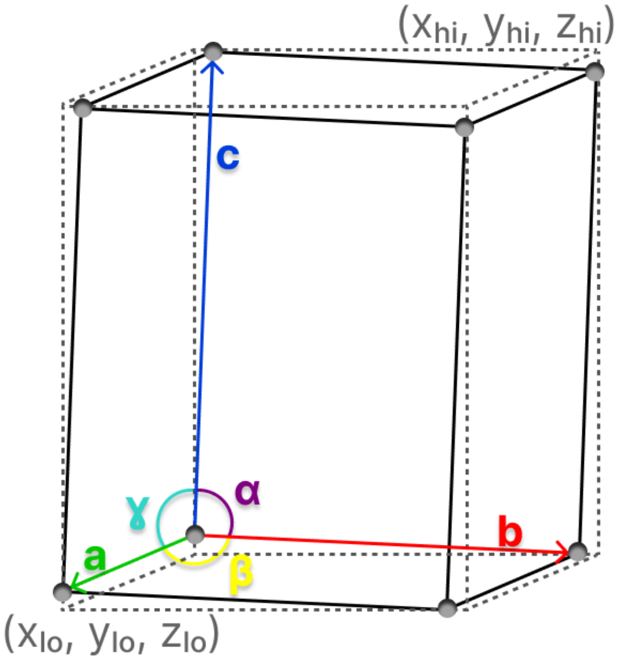


Figure 2.4: Triclinic cell in 3D defined by its edge vectors $\vec{a}, \vec{b}$ and $\vec{c}$ and its origin $(x_{\text{lo}}, y_{\text{lo}}, z_{\text{lo}})$ and the diagonally opposing side $(x_{\text{hi}}, y_{\text{hi}}, z_{\text{hi}})$. Dashed lines represents an orthogonal cell which encloses the triclinic cell.

The volume of a triclinic cell is given by

$$V_{\text{cell}} = \|\vec{a}\|\|\vec{b}\|\|\vec{c}\|\left(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos^2\alpha\cos^2\beta\cos^2\gamma\right)^{1/2}, \quad (2.17)$$

where cell angles are related to the tilt factors by

$$\begin{aligned}\cos\alpha &= \frac{xy \cdot xz + \Delta y \cdot yz}{\|\vec{b}\|\|\vec{c}\|},\\ \cos\beta &= \frac{xz}{\|\vec{c}\|},\\ \cos\gamma &= \frac{xy}{\|\vec{b}\|}.\end{aligned}$$

For a particle with coordinate $\vec{x}$ it can be represented as its unwrapped coordinate in a neighboring cell identified by the vector

$$\vec{n} = (n_1, n_2, n_3)^{\top} \in \mathcal{N}^3 \backslash \{(0,0,0)\},$$

where $\mathcal{N} = \{-1, 0, 1\}$, yielding $3^3 - 1$ unique vectors $\vec{n}$, one for each neighboring cell. The $\vec{n}$ denotes the offset from the origin, which cell image the pair potential is from. Then the unwrapped coordinate is expressed by the translation

$$\vec{x}_{uw} = \vec{x} + \mathbf{H}\vec{n}. \quad (2.18)$$

In the case of periodic cells, meaning particles interact and can be moved across the boundaries to the opposing side of the cell, this is done using Eq. 2.18. An example of how this may look can be found in Fig. 3.2.

2.3.2 Pressure calculation

The internal pressure for the set of particles $\{\vec{x}_i \in \mathbb{R}^d\}_{i=1}^{N}$ inside a volume V with temperature T is calculated in LAMMPS as

$$P_{\text{int}} = \frac{N k_b T}{V} + \frac{1}{dV} \sum_{i=1}^{N'} \vec{F}_i \cdot \vec{x}_i \tag{2.19}$$

where $\vec{F}_i$ is the force acting on particle i, k_b is the Boltzmann constant and the temperature is given by

$$T = \frac{\sum_{i=1}^{N} m_i \vec{v}_i^2}{3 k_b N},$$

where m_i and $\vec{v}_i$ is the mass and velocity of particle i. As LAMMPS divides the pressure calculation over the processor subdomains, they calculate their individual contributions, including over ghost atoms, which is then summed across processors, such that $N' \neq N$ *in general*. The particle potentials are two-body, using this and disregarding contribution of periodic images, the internal pressure can be expressed as

$$P_{int} = \frac{N k_b T}{V} + \frac{1}{3V} \sum_{i=1}^{N} \sum_{j>i}^{N} \vec{F}_{ij} \cdot \vec{r}_{ij}$$

where for zero particle overlaps in the initial state is

$$P_{int} = \frac{N k_b T}{V}$$

2.3.3 Isobaric-isoenthalpic ensemble

The isobaric-isoenthalpic ensemble (NPH) implemented in LAMMPS [6] propagates the motion of the particles in the cell, while keeping the internal pressure, specified by Eq. 2.19 nearly constant, through changing the volume and geometry of the simulation cell - acting as a barostat. The independent variables of the system are the number of particles N, the pressure P and the enthalpy H. The particle equations of motion are those specified in Shinoda et al. [31]. As the contact model dissipates energy, the kinetic energy of the system gradually decreases, such that a larger proportion of the internal pressure stems from the virial term calculated by LAMMPS, this being the second term in Eq. 2.19.

Through parameters P_{damp} and T_{target}, the responsiveness of the simulation cell to pressure fluctuations can be specified through the cell inertia

$$W = (N+1) k_b T_{\text{target}} P_{\text{damp}}^2. \tag{2.20}$$

If no thermostat is initialized the target temperatue T_{target} is set to the system temperature when the barostat is initialized, as this temperature may poorly reflect the temperature at a steady state, this can be manually set. The inertia being too low means that the simulation box tilts too quickly, and the simulation stops.

2.4 Voronoi tessellation

In order to study the local properties of the packing the end state of the simulation is tessellated by a Laguerre-Voronoi tesselation. The Voronoi tesselation of a set of points $D := \{\vec{x}_i\}_{i=1}^N$ is defined by associating a cell C_i with its corresponding nuclei $\vec{x}_i$. The cell C_i is defined as the set of point closer to the nuclei $\vec{x}_i$ than any other, i.e.

$$C_i = C(\vec{x}_i, D) = \{\vec{y} \in R^k : \|\vec{y} - \vec{x}_i\| \leq \|\vec{y} - \vec{x}_j\| \quad \forall \vec{x}_j \in D \setminus \vec{x}_i\}.$$

As this tessellation does not take into account the effect of polydisperse particles a weighting is introduced by introducing the distance $d(y, (x, w)) = \|y - x\|^2 - w$ such that cells are calculated as

$$C((x_i, r_i), D) = \{y \in R^k : d(y, (x_i, r_i^2)) \leq d(y, (x_j, r_j^2)) \quad \forall x_j \in D \setminus x_i\},$$

where r_i is the radius of the particle with center x_i. The Laguerre tesselation cells are convex polytopes filling the whole space, in this case R^k.

The tessellation of the periodic cell is done by applying the Voro++ [32] functionality available in LAMMPS. Voro++ is designed specifically to be of use in material physics by calculating the tessellation on specified particles, allowing for complex boundary conditions. It allows calculation of cell volumes, face areas, and finding the particles neighboring a specific particle. A neighbor in this context is a particle sharing a face with the particle in question. The tessellation then yields quantities allowing for the study of local correlations of for example orientation and size.

2.5 Packing fraction models

Models for the packing fraction attempt to estimate the packing fraction ϕ given certain specified input parameters, some being the particle size distribution, particle shapes and component volume fractions among others. This thesis uses the model of Yu and Standish [33], developed using data from binary mixtures. This model was modified by Frings et al. [34], ultimately writing a Python implementation [35] used in this work.

2.5.1 Structural Universality

By binning particles and calculating the specific volume $e_i = \frac{1}{\phi_i}$ and normalized area of each bin, Yuan et al. [36] found that a so called *master curve* of specific volume versus relative normalized area appears for a variety of packings of different particle size distributions. Given this master curve $f : (D_i) \to e_i$, with D_i being individual particle diameters and e_i, the specific volume of the Voronoi cell containing it, the bulk packing fraction can be estimated by

$$\phi = \frac{V_p}{V} = \frac{\sum_{i=1}^N D_i^3}{\sum_{j=1}^N f(D_j) D_j^3}$$

The specific volume is studied by constructing a radical Voronoi tesselation (see Sec. 2.4) of each simulation, binning sorted particles into k bins $\mathbb{B}_i$ of equal number of particles and calculating the bin specific, mean normalized area or volume $A_i^{(m)}$ and mean specific volume e_i, defined by

$$\begin{aligned} A^{(m)} &= \frac{\sum_{j=1}^{N} D_j^m}{N}, \quad m = 2, 3 \\ A_i^{(m)} &= \frac{\sum_{j \in \mathbb{B}_i} D_j^m}{|\mathbb{B}_i| A^{(m)}}, \\ e_i &= \frac{\sum_{j \in \mathbb{B}_i} V_{cell,j}}{|\mathbb{B}_i| \sum_{j \in \mathbb{B}_i} V_j} \end{aligned}$$

where $V_{cell,j}$ is the radical Voronoi cell volume of particle j, $|\mathbb{B}_i|$ is the number of particles in bin i and $m = 2, 3$. In the case of [36], they specify $m = 2$, and study the specific volume as a function of the normalized area.

2.5.2 Linear Mixing Packing Model

A packing of spherical particles with a continuous PSD, of N components is considered. The solid volume of particles belonging to class i is S_i, with its fractional solid volume of particles X_i then being

$$X_i = \frac{S_i}{\sum_{j=1}^{N} S_j}, \quad \sum_{i=1}^{N} X_i = 1, \quad X_i \geq 0 \quad \forall i = 1, \ldots, N$$

The space of mixtures with N components can then be considered geometrically as a $N-1$-dimensional simplex [37], with vertices representing single component mixtures. Some models for the mixture specific volume $e = \frac{1}{\phi}$, incorporate the effect of component interactions by including products of class solid volume fractions X_i.

Yu and Standish [33] developed a geometric model for the packing fraction of multi-component mixtures, which disregards interaction between three or more components. For binary packings, they argue that there exists a *critical ratio of entrance* r_c, where for size ratios smaller than this, controlling components then form the skeleton of the packing. Small particles are then added by the *filling mechanism*, where small particles percolate through the skeleton of the large particles, which can be seen in Fig. 2.5 (a). From geometric considerations they derive that $r_c \approx 0.154$. [33].

Above the critical size ratio, particles cannot percolate through the pores of the large particle skeleton. In this case the packing is formed by the *occupation mechanism*, where neither partial specific volume is undisturbed, see Fig. 2.5(c). The controlling component is dependent on both component size ratios, and solid volume fraction. Consider as a small number of fines are added to a packing of coarse particles, for small values of X_{fine} the fines fit into the pores, but at a certain point, the pore space runs out and the fines are incorporated into the packing skeleton, the packing is then being formed by occupation. For multi-component mixtures the controlling component can consist of components of different sizes. The packing fraction is then estimated through interaction functions between particle sizes [33] and identification of the controlling components.

The controlling components $n, n+1, \ldots, m$ for a given class are then identified by

$$n = \begin{cases} 1, & d_i/r_c > d_1 \\ j, & d_{j-1} \geq d_i/r_c > d_j, \quad j = 2, \ldots i \end{cases} \tag{2.21}$$

$$m = \begin{cases} j, & d_j \geq d_i r_c > d_{j+1}, \quad j = i, \ldots N-1 \\ N, & d_N > d_i r_c \end{cases} \tag{2.22}$$

$$\tag{2.23}$$

where d_j are component diameters. This is a modification of the original Yu and Standish model by Frings et al. [34]. The size ratio between classes is defined as

$$r_{ij} = \begin{cases} d_i/d_j, & d_i \leq d_j \\ d_j/d_i, & d_i > d_j \end{cases}$$

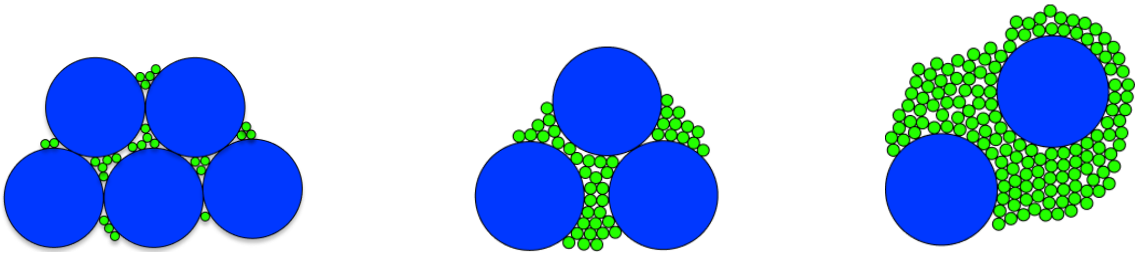

(a) Filling (b) Interacting (c) Occupating

Figure 2.5: LMPM mechanisms

The observation that the total specific volume e of a binary mixture could be well described by the maximum of two linear functions of the solid volume fractions motivated the idea that the relation between e and X_i for multi-component mixtures was also linear [34], and that the true specific volume was the maximum of these segments

$$e = \max(e_i^*), \quad e_i^* = e_i - S_{ij} X_j$$

where $e_i = \frac{1}{1-\eta_{0,i}}$, with $\eta_{0,i}$ being the unmixed porosity for component i and S_{ij} is an experimentally fitted slope. The Yu and Standish model is a generalization of this, where for a mixture of N components, with controlling components $n, \ldots, m$, they express the partial specific volume as

$$e_i^* = \sum_{j=1}^{n-1} (e_j - (e_j - 1) b(r_{ij})) X_j e_i^{mix} \sum_{j=n}^{m} X_j + \sum_{j=m+1}^{N} e_j (1 - a(r_{ij})) X_j$$

where $e_i = \frac{1}{1-\eta_{0,i}}$, with $\eta_{0,i}$ being the component porosity. This is set to some value η_0 for all components, which in this case should lie close to 0.36, the porosity at close random packing, while for different real life materials this porosity may be drastically different [33]. The first and third terms are then formed by filling and occupation mechanisms (see Fig. 2.5 (a), (c)), with interaction functions $a(\cdot), b(\cdot)$, while the second term is an interaction term of the controlling mixture defined by $[n, m]$, with the specific volume of the mixing component e_i^{mix} as unknown. This is formed by the interaction mechanism (Fig. 2.5 (b)), and can be estimated by

$$\begin{aligned} e_i^{mix} &= \sum_{i=n}^{m} e_i X_i + \sum_{j=n}^{m-1} \sum_{k=j+1}^{m} X_j X_k \left[c(r_{jk}, n_0^{mix}) \frac{1}{Y^2} + d(r_{jk}, n_0^{mix}) \frac{(X_j - X_k)}{Y^3} \right] \\ Y &= \sum_{i=n}^{m} X_i . \end{aligned}$$

where

$$v_0^{mix} = \frac{\sum_{j=m}^{n} X_j e_j}{\sum_{k=m}^{n} X_k} = \frac{1}{1 - n_0^{mix}}$$

The interaction terms $c(\cdot), d(\cdot)$ quantify the effect of mixing of different size ratios, and are always negative, decreasing the specific volume v_0^{mix} (increasing packing fraction) through joint interaction [34]. The form of all interaction functions can be seen in Appendix D. The components n, m associated with the chosen $e_i = \max_j(e_j^*)$ then specify the controlling mixture [34].

For multimodal packings, the bin diameters and solid volume fraction can simply be calculated from the particle types. The initial porosity is specified as $n_{0,i} \equiv n_0$ for all bins, then being the only fitting parameter.

2.6 Content uniformity - Theoretical models

Low dose drug products such as tablets and capsules, require strict control of the uniformity of the active pharmaceutical ingredient (API). The largest particles may embody an unacceptably large percentage of the total dosage, and can ultimately make sufficient dosage control impossible. Particle size control is then essential. Tying together UDU with PSD descriptors is then necesssary [3]. An indicator of dose uniformity is the Coefficient of Variation (COV) C_v, defined as the standard deviation over the mean for some quantity of interest, $C_v = \sigma/\mu$, in this case the dose. The United States Pharmacopeia put limits on the C_v such that $C_v \leq 6.25\%$. While there are factors other than particle size that influence C_v, these are not considered here, though the maximum allowed C_v should be considerably smaller than the limit 6.25% to account for these additional errors [2].

There are many mixing indexes that have been developed over the years. These are ways of quantifying the degree to which a mixture is sufficiently random - with a completely random mixture commonly meaning that the probability of finding a particle of a specific component is the same over the entire mixture [38]. Estimating the sample variability using statistics means that comparation with physical samples is possible. To evaluate their quality of mixing, a "ground truth" of best possible mixing can then be supplied. Both of the following estimates assume particles to be spherical.

A necessity for the study of content homogeneity is to introduce a *scale of scrutiny*, which is the scale at which the homogeneity of a particle mixture is important given its use. For pharmaceutical purposes, a single tablet may form this scale [39].

2.6.1 Stange

Stange derived [40] an expression for the homogeneity of multi-component mixtures based on their mass fractions $\tau_{m,i}$ and PSD properties, homogeneity being quantified by the C_v of the component mass fraction

$$\left(\frac{\sigma_i}{\tau_{m,i}}\right)^2 = \frac{1}{m}\left[\left(\frac{1-\tau_{m,i}}{\tau_{m,i}}\right)^2 \tau_{m,i}\bar{w}_i\left[1+\frac{s^2_{m(i)}}{\bar{w}_i}\right] + \sum_{j\neq i}\tau_{m,j}\bar{w}_j\left(1+\frac{s^2_{m(j)}}{\bar{w}_j}\right)\right] \tag{2.24}$$

here $\bar{w}_i$ is the mean mass of component i, $s_{m(i)}$ is the component mass standard deviation and m is the total sample mass (all components). Assuming that the components are of a single size, such that $s_{m(i)} \equiv 0$ and $\bar{w}_i \equiv w$ then the expression becomes

$$\sigma_i^2 = \frac{w}{m}\tau_{m,i}(1-\tau_{m,i})$$

with equal densities this mixing variance is equivalent to that of Hoggs' [41]

$$\sigma_i^2 = \left(\frac{d}{D}\right)^3 \frac{P(1-P)}{\phi},$$

where d is the particle diameter, D the diameter of the sphere containing the requisite particle volume, including porous volume, P the particle number fraction and $\phi = 1-\eta$ is the packing fraction. The number of particles is given by $N_s = \frac{1}{\phi}\left(\frac{d}{D}\right)^3$ which makes it consistent with Lacey as well [41]. An expression based on volume fractions $\tau_{v,j}$, not including porosity was derived from this by Mort [41]:

$$\begin{aligned}\left(\frac{\sigma_{v(i)}}{\tau_{v,i}}\right)^2 = \frac{1}{V\phi}\Bigg[\left(\frac{1-\tau_{v,i}}{\tau_{v,i}}\right)^2 \tau_{v,i}\left(\bar{v}_i+\frac{s^2_{v(i)}}{\bar{v}_i}\right) + \\ \sum_{j\neq i}^{n}\tau_{v,j}\left(\bar{v}_j+\frac{s^2_{v(j)}}{\bar{v}_k}\right)\Bigg],\end{aligned} \tag{2.25}$$

where V is the volume of the particles V_p, including porosity. The Stange model includes the effect of PSD width through the mass standard deviation $s_{m(i)}$. For multimodal distributions this becomes inappropriate [41], though in this case, component distributions are monomodal. The variance of the concentration $\sigma^2_{c(i)}$ is derived from that of the volume fraction as $\sigma^2_{c(i)} = (\phi\sigma_{v(i)})^2$, assuming ϕ to be constant. Then the concentration standard deviation $\sigma_{c(i)}$ and COV C_v is expressed as

$$\sigma_{c(i)} = \tau_{v,i}\sqrt{\frac{\phi}{V}\left[\left(\frac{1-\tau_{v,i}}{\tau_{v,i}}\right)^2 \tau_{v,i}\left(\bar{v}_i+\frac{s^2_{v(i)}}{\bar{v}_i}\right) + \sum_{j\neq i}^{n}\tau_{v,j}\left(\bar{v}_j+\frac{s^2_{v(j)}}{\bar{v}_k}\right)\right]}, \tag{2.26}$$

$$C_v = \frac{\sigma_{c(i)}}{C}. \tag{2.27}$$

where is the component bulk concentration. This analytical estimate is based on the inclusion of whole particles, meaning that for volumes that are small compared to the particles, $\sigma_{c(i)}$ could be overestimated [41].

2.6.2 Hilden

Hilden et al. [3] developed a low-dose drug load model for the dose C_v, incorporating the sample descriptor $D[6,3]$, reported by a particle size analyzer. In general the moment ratio $D[a,b]$ is expressed as

$$D[a,b] = \left(\frac{\sum_{i=1}^{N} d_i^a n_i}{\sum_{j=1}^{N} d_j^b n_j}\right)^{\left(\frac{1}{a-b}\right)} = \left(\frac{\sum_{i=1}^{N} d_i^a}{\sum_{j=1}^{N} d_j^b}\right)^{\left(\frac{1}{a-b}\right)} , \quad a,b \in \mathcal{Z}_+$$

where d_i is the diameter of bin i, while n_i is its frequency. As the particles in this case are not binned $n_i \equiv 1$. Under the assumption that the API is randomly distributed, the COV C_v of drug product is then expressed as

$$C_v = \sqrt{\frac{\pi\rho}{6G} D[6,3]^3}, \tag{2.28}$$

with G being the target dose of drug (mass), ρ the component density and $D[6,3]$ is over the API particle diameters. The Hilden model assumes a low concentration of API, as particle overlap is not considered and neither is the PSD of other components. For higher concentrations Hilden et al. argue that the API concentration becomes more uniform, meaning the model provides a conservative, higher estimate in these cases [3]. The expression above Eq. 2.28 is not the one used for the concentratioin, instead the standard deviation of the drug mass σ_m is considered, expressed by Hilden as

$$\sigma_m = \frac{\pi\rho}{6}\sqrt{\sum_{i=1}^{N} d_i^6}.$$

When subsampling the simulation cell, the entire component mass is not included, instead the sum $\sum_{i=1}^{N} d_i^6$ is adjusted by the fraction of total volume sampled $\frac{V}{V_{cell}}$, such that

$$\sigma_m = \frac{\pi\rho}{6}\sqrt{\frac{V_s}{V_{cell}}\sum_{i=1}^{N} d_i^6}.$$

Since the concentration $C = \frac{m}{\rho V}$, the standard deviation of the concentration and C_v of the same is then

$$\sigma_c = \frac{\pi}{6}\sqrt{\frac{1}{V_s V_{cell}}\sum_{i=1}^{N} d_i^6}, \tag{2.29}$$

$$C_v^{\text{Hilden}} = \frac{\sigma_c}{C}, \tag{2.30}$$

Then the estimates of Stange and Hilden both have the same proportional dependence on sampling volume V, i.e $\sigma_c \propto V^{-1/2}$.

3

Method

3.1 Simulation

Given a K-component packing, with component-wise parameters $\{\mu_k, \sigma_k, \tau_k, \rho_k\}_{k=1}^K$, where τ_k, ρ_k is mass fraction and true density of component k, N_k particle diameters are drawn for component k by

$$d_1, d_2, \ldots,\ d_{N_k} \sim \mathcal{LN}(\mu_k, \sigma_k).$$

In the case that μ_k, σ_k represent the expected diameter and standard deviation, the distribution parameters used for $\mathcal{LN}$ are given using Eq. 2.1, 2.2. Samples are drawn in a manner such that mass fractions τ_k deviate within a set tolerance parameter ϵ (see Appendix B.3). The total sample size over all components $N = \sum_{k=1}^K N_k$ is bounded from below by the minimum number of particles required to satisfy the mass fraction tolerance ϵ. After a total of N particles are drawn the diameters

$$\{d_i\}_{i=1}^N = \bigcup_{k=1}^K \{d_1, d_2, \ldots,\ d_{N_k}\},$$

were normalized by the smallest diameter present so that $\min\{d_i\}_{i=1}^N = 1$. Units used are such that $k_B = 1$ and the minimum diameters forms the length scale.

The initial simulation box is a cube with sidelengths $\|\vec{a}\|$, $\|\vec{b}\|$ and $\|\vec{c}\|$ are given by

$$\|\vec{a}\| = \|\vec{b}\| = \|\vec{c}\| = \left(\frac{\pi}{6\phi_0}\sum_{i=1}^N d_i^3\right)^{1/3},$$

with the initial packing fraction $\phi_0 = 0.05$. The initial particle coordinates $\{\vec{x}_i\}_{i=1}^N$ were drawn in a random manner within the initial simulation box such that particles were not in contact with each other, using a method further described in Appendix B.1. The hierarchy levels used for contact detection within LAMMPS were determined based on Appendix B.2. The skin distance was naively set to $R_s = 1$ for the entire simulation run, i.e. the diameter of the smallest particle. Finally box and particle data were written to a file in a LAMMPS-readable format, so that it could be executed using a LAMMPS-script. The script can be separated into two parts; a compression phase and an adhesion phase.

The compression phase uses the Hertzian contact model (Sec. 2.2.1.1), with $E = 5$ as the Young's modulus set so that maximum relative overlap does not exceed 1% (see Appendix A.3, Fig. A.3). The damping was specified through a

COR $e = 0.1$ that was determined based on simulation testing (see Appendix A.3, Fig. A.6).

Increasing the packing fraction was achieved through the NPH-ensemble which compresses the triclinic and periodic cell. The ensemble parameters were chosen as $P_{\text{damp}} = 10^4$, $T_{\text{target}} = 10^{-5}$ to set the relaxation rate (Eq. 2.20), and $P_{\text{ext}} = 10^{-6}$ to set the applied external isotropic pressure. The system was initialized with no overlap and at $T = 0$, such that $P_{\text{int}} = 0$.

At the initial stage of the simulation the timestep size Δt_0 was set to a fraction $T_0 = 40$ of the Hertzian collision time (Eq. 2.10) for the smallest particles so that

$$\Delta t_0 = \frac{2.865}{T_0}\left[\frac{m^2}{RE^2 v_{max}}\right]^{1/5} = \frac{2.865}{T_0}\left[\frac{\pi^2}{36E^2 v_{max}}\right]^{1/5},$$

where the smallest necessary timestep occured for the smallest particles, such that $\frac{m^2}{R}|_{r=1/2} = \frac{\pi^2}{36}$. As the particles were initially at rest, the maximum velocity v_{max} was conservatively set to unity, to avoid simulation failure. After an initial compression period the timestep was switched to

$$\Delta t = \frac{2.865}{T}\left[\frac{\pi^2}{36E^2 v_{max}}\right]^{1/5}, \tag{3.1}$$

where $T = 10$ was selected based on testing simulation stability on example packings (see Appendix A.3, Fig. A.6). Then every 25.000:th timestep the timestep size Δt was updated as v_{max} changed to increase simulation speed.

The exit conditions for the compression phase were formulated based on intensive parameters, that are insensitive to polydispersity. The first was the Cundall parameter C [42] defined as the sum total net force across the packing, divided by the sum total force

$$C = \frac{\sum_p |f_p|}{\sum_c |f_c|} = \frac{\sum_{i=1}^{N} \left|\sum_{j\setminus i} \vec{f}_{ij}\right|}{\sum_{i=1}^{N} \sum_{j\setminus i} |\vec{f}_{ij}|}. \tag{3.2}$$

This then gives quantitative insight into the ratio of net particle force and the sum total contact force across the packing, in a sense quantifying the stability of the packing. The second monitored parameter is based on the contribution of kinetic energy to the pressure

$$PR = \frac{\frac{NT}{V}}{P_{\text{int}}} = \frac{NT}{P_{\text{int}} V}, \tag{3.3}$$

where the pressure was calculated by LAMMPS as described in Sec. 2.3.2. The parameters C and PR required long simulation times to stabilize, while the change in packing fraction ϕ drops off quicker. the The adhesion phase was initialized, when exit tolerances for the compression phase were satisfied, $C \leq 10^{-2}, PR \leq 10^{-3}$. The adhesion phase was used to ensure that small particles, nonphysically suspended in pores adhere to the packing. The introduction of the adhesion contact model should not disrupt the structure of contacting particles too much, while allowing zero-contact particles to stick. The exit criteria should be stringent enough, and the surface energy small enough that this holds. The surface energy (see Appendix A.1) was specified as

$$\gamma = \frac{E}{2}\left[\frac{10^{-2}}{3.094}\right]^{3/2}.$$

This was selected based on the same condition as for the Hertz contact model, i.e. that the maximum relative overlap does not exceed 1% (see Appendix A.3, Fig. A.4). When introducing the JKR contact model, the group of zero-contact particles $\mathcal{J}$ are identified and their velocities $\vec{v}_j$, $j \in J$ scaled by

$$\vec{v}_j := \frac{\|\vec{v}_n\|}{\sqrt{3}} \frac{\vec{v}_j}{v_{max}},$$

where $\vec{v}_n = 10\vec{v}_s$, with $\vec{v}_s$ the sticking velocity defined by Eq. A.1. The velocities should be bounded by $10\vec{v}_s$, hence the factor $\frac{1}{\sqrt{3}}$. The factor of 10, was determined through the stick and rebound tests in Appendix A.2 for COR $e = 0.1$. The timestep is kept as that specified in Eq. 3.1 over the entire adhesion phase and is not updated as in the Hertzian model, due to it resulting in simulation instabilities. For the adhesive phase no exit condition was implemented, something which is further discussed in Sec. 5.1.

Throughout the simulation run, particle-wise, pair-wise and scalar properties were written to data files. These included particle and component identifiers, coordinates, pair-wise contacts and distances, volume density, pressure ratio and Cundall parameter. Python scripts were written to handle the data and create statistics of interest. Visualisation of the packing at individual timesteps was achieved using the open-source software OVITO Basic (Open Visualisation Tool) [43].

3.2 Concentric shell sampling method

The component-wise homogeneity for a given sampling scale in a simulation was calculated through an approach similar to [41]. First a radial mixing scale L was defined, then a set of random points $\{\vec{p}_m\}_{m=1}^{M}$ were selected in the cell at random. Concentric spheres S_j of thickness $j\Delta L$ where $\Delta L \ll L$ centered on $\vec{p}_m$ were constructed, i.e

$$S_{j,m} = \{\vec{y} : d(\vec{p}_m, \vec{x}) \leq j\Delta L\}, \tag{3.4}$$

where $d(\cdot, \cdot)$ denotes the euclidean distance. For every sphere radius $R_j := j\Delta L$, the total volume of particles contained within the concentric sphere was calculated. For a particle with coordinate $\vec{x}$, radius r and distance to center of concentric sphere $d := d(\vec{p}_m, \vec{x})$ intersecting the j:th shell the intersecting volume was calculated by

$$V_{\text{intersection}} = \frac{\pi\left(R_j + r - d\right)^2\left(d^2 + 2dr - 3r^2 + 2dR_j + 6rR_j - 3R_j^2\right)}{12d}.$$

Below follows a schematic image displaying the concentric shell method applied over one point $M = 1$, a three-component monosized packing in 2D, where ΔL represents inner shell thickness, and $10\Delta L$ represents shell thickness at $j = 10$.

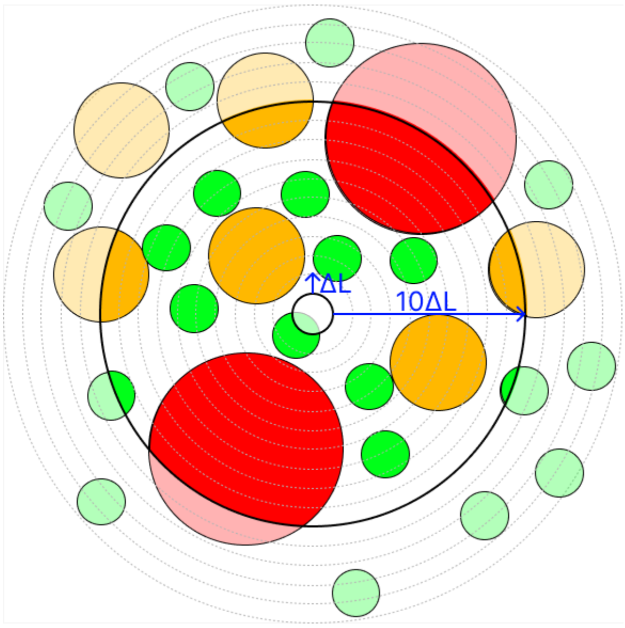


Figure 3.1: Concentric shells applied over three-component monosized packing in 2D. Each color assignment (red, orange, green) represents particle belonging to the same component. Outer black line represents shell thickness for $j = 10$, i.e. $S_{10,1}$, where particle volume inscribed within it is contributing to the total particle concentration. Particle volume within shaded regions is not contributing to the concentration of $S_{10,1}$, but will be for larger j.

When selecting M random points within the cell, the radial mixing scale may extend beyond the cell boundaries, resulting in underestimating component concentrations. Given the fact that particles were placed within a periodic cell, the coordinates of all particles $\{\vec{x}_i\}_{i=1}^{n}, \vec{x}_i \in \mathbb{R}^3$ were unwrapped in all $3^3 - 1$ directions as described in Sec. 2.3.1, yielding $\{\vec{x}_u\}_{u=1}^{27n}$, with corresponding index notation for the set of particle diameters $\{d_u\}_{u=1}^{27n}$. This allowed for selecting any amount of random points such that the concentric shell intersects the cell boundary without any local density fluctuations when crossing it, as displayed in Fig. 3.2. The schematic figure contains a monosized particle packing with diameters d, where a point has been selected next to the rightmost boundary of the 2D triclinic cell defined by the its edge vectors $\vec{a}$ and $\vec{b}$. The particles intersecting cell boundaries, depicted as light-green were unwrapped to their opposing side. The radial mixing scale L for any 3D cell was limited by the upper bound

$$L \leq \frac{1}{2}\min(\|\vec{a}\|, \|\vec{b}\|, \|\vec{c}\|) \tag{3.5}$$

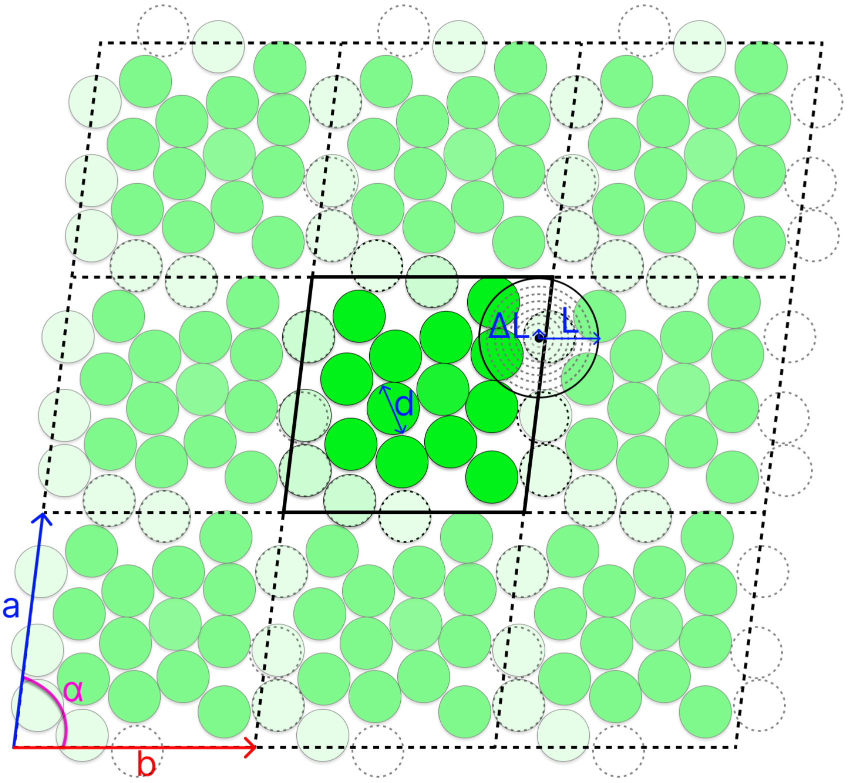


Figure 3.2: 2D illustration of cell unwrapping in all $3^2 - 1$ directions from the reference cell in the middle with highlighted particles. Lighter green representing particles partially over the cell boundaries.

Given a multi-component packing consisting of components $k = 1, \dots, K$, particle volume concentrations were given by

$$C_{k,j,m} = \frac{6V_{k,j,m}}{\pi(j\Delta L)^3}, \tag{3.6}$$

where $V_{k,j,m}$ denotes total volume of particles of component k inside the sphere of radius $j\Delta L$ centered at point $\vec{p}_m$. The per-component mean and standard deviation with respect to shell thickness over each concentric shell centered around its random point was calculated so that

$$\mu_{k,j} = \frac{1}{M}\sum_{m=1}^{M} C_{k,j,m}, \tag{3.7}$$

$$\sigma_{k,j} = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(C_{k,j,m} - \mu_{k,j})^2}. \tag{3.8}$$

These are then the sample mean and standard deviation, conditioned on a sampling scale specified by the shell thickness. From this, the COV for a given shell thickness, i.e. mixing scale, and number of random points was given by

$$CV_{k,j} = \frac{\sigma_{k,j}}{\mu_{k,j}}. \tag{3.9}$$

Then, the concentration COV was compared with theoretical models described in Sec. 2.6.

4

Results

4.1 Simulation

An example packing was defined based on nominal pharmaceutical drug product composition (see Table E.1). The packing contains 605.000 particles with a polydispersity $\lambda = 53$. The maximum mass fraction error was $\epsilon < 10^{-3}$, but note that no condition was set on maximum packing fraction variance by using a fitted packing fraction model. Fig. 4.1 shows the component-wise probability densities of sampled particles for given input parameters.

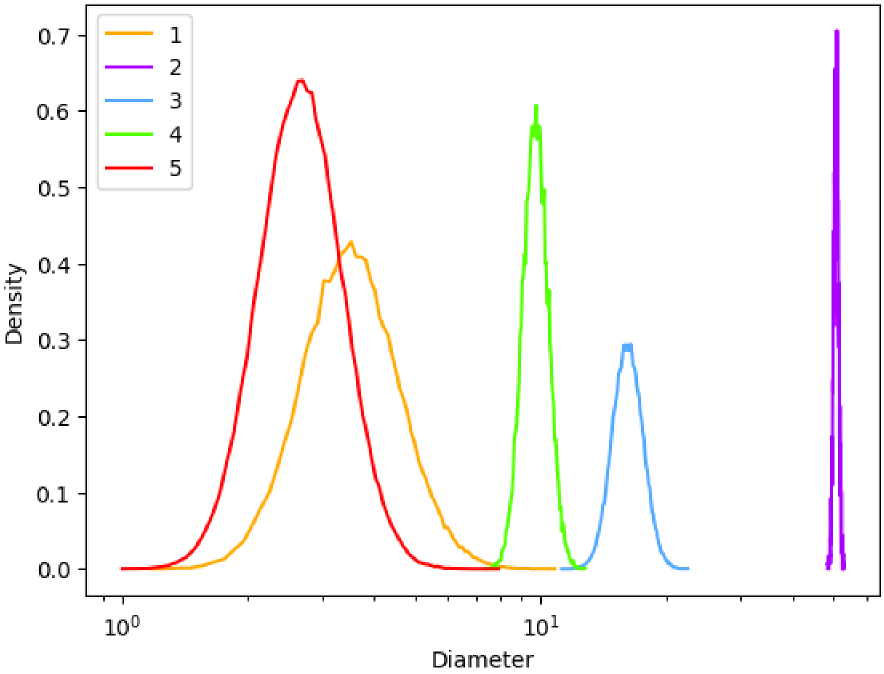


Figure 4.1: Component-wise sampled diameter distributions for example packing in E.1. (1, orange): API, (2, purple): Plastic filler, (3, blue): Elastic filler, (4, green): Disintegrant and (5, red): Lubricant.

4.1.1 Compression phase

Fig. 4.2 shows the development of the volume density ϕ, global coordination number Z (Eq. 2.3) and non-rattler coordination number Z_{nr} (Eq. 2.4). The density increases rapidly over the first ≈ 250.000 timesteps as a consequence of the NPH-ensemble applied on the simulation cell, to then slow down due to reaching a jamming state. Z_{nr} increases from 0 to around 12 at first timestep, due to the many collisions occuring for the small subset of large particles. For higher volume density small particles increasingly gain contacts, stabilising Z_{nr} at a value of 6, corresponding with the isostatic condition of frictionless spheres (see Sec. 2.1.3). At the same time Z does not increase much during the compression phase, indicating a large share of the particles are rattlers which move freely inside the pores of the larger particles.

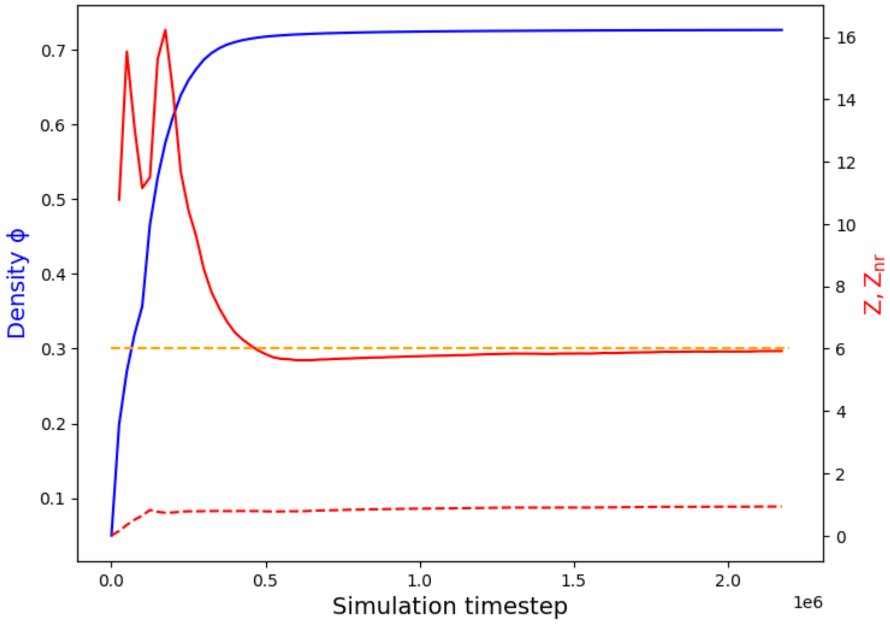


Figure 4.2: Development of ϕ, Z and Z_{nr} during compression phase over a linear timescale. Red line is Z_{nr}, dashed red line is Z and dashed orange line corresponds to the isostatic condition of $Z = 6$. Blue line is ϕ.

Fig. 4.3 shows the development of the exit conditions used for the compression phase, monitoring the Cundall parameter C (Eq. 3.2) and Pressure ratio PR (Eq. 3.3). Both share the downward trend, but develop in different orders of magnitude. Even after ϕ, Z_{nr} and Z stops developing after ≈ 500.000 timesteps as shown in Fig. 4.2 PR and C keeps decreasing, indicating that the particles has not yet reached a stable state in all aspects. PR reaches below its tolerance after 480.000 timesteps, and C after 2.150.000, at which point the compression phase is terminated.

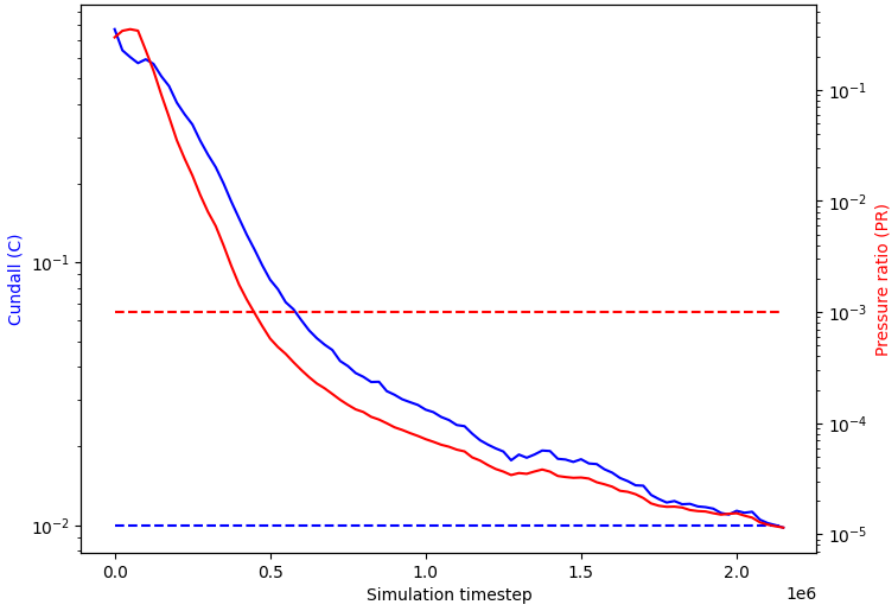


Figure 4.3: Development of exit conditions during compression phase. Red dashed line corresponds to maximum tolerance of the pressure ratio, and blue dashed line corresponds to maximum tolerance of the Cundall parameter.

4.1.2 Adhesion phase

The following results shows the adhesion phase where the first timestep is same as the last timestep of the compression phase. Fig. 4.4 shows the development of contacts for small and large particles during adhesion phase, where small particles are those within the 95:th percentile and large are those above it, when ordering by particle diameter. Small particles rapidly gain new contacts with both other small particles and large particles, as a consequence of the adhesive property due to switching contact model to JKR. Large particles do not develop new contacts with other large particles, indicating that the structure of large particles in the packing is stable when switching contact model.

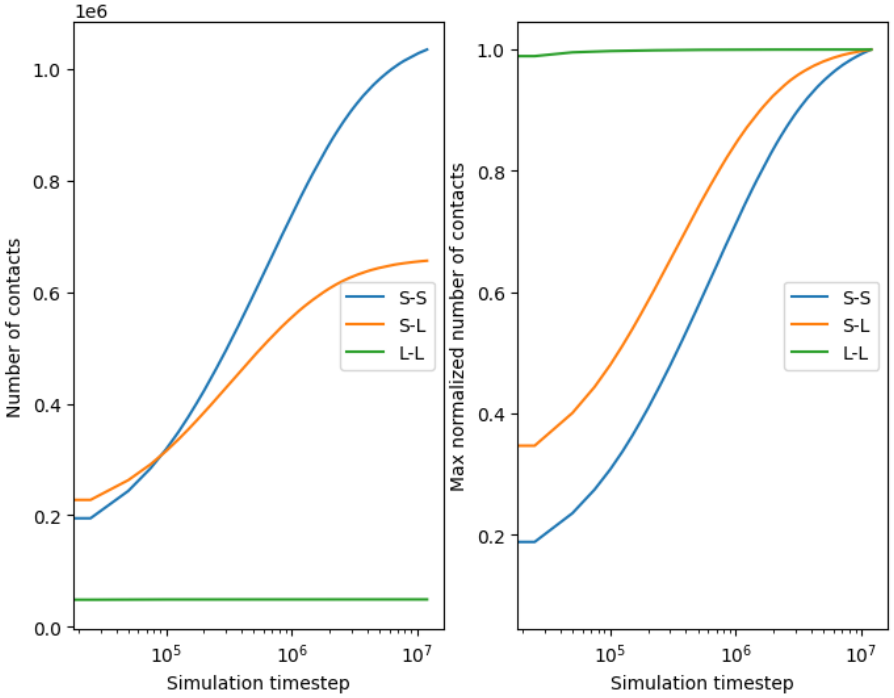


Figure 4.4: Development of contacts for small and large particles during adhesion phase. $S-S$ denotes contacts between small particles, $S-L$ contacts between small and large particles, $L-L$ contacts between large and large particles. Small particles are those within the 95:th percentile and large are those above it, when ordering by particle diameter. Left: Total number of contacts for each kind of contact pair, over a logarithmic timescale. Right: Contacts for each kind of contact pair, each normalized by the maximum number of contacts reached during adhesion phase, over a logarithmic timescale.

Fig. 4.5 shows the number of particles divided into groups based on coordination number and normalized by total number of particles. In the beginning $> 80\%$ of the particles have zero contacts $N_{Z=0}$, which has declined to $\approx 0\%$ after 1.000.000 timesteps. Particles in this group move to the group of particles containing one contact $N_{Z=1}$, reaching a maxima after $\approx$ 200.000 timesteps, after which particles find another contact resulting in $N_{Z=2}$ reaching a maxima at $\approx$ 500.000 timesteps, to then finally find a third contact $N_{Z=3}$. This group is stable for the last 5.000.000 timesteps. Non-rattler $N_{Z\geq4}$ and the inversely proportional rattler group $N_{Z<4}$ indicate that the packing still develops contacts even at the end of the adhesive phase.

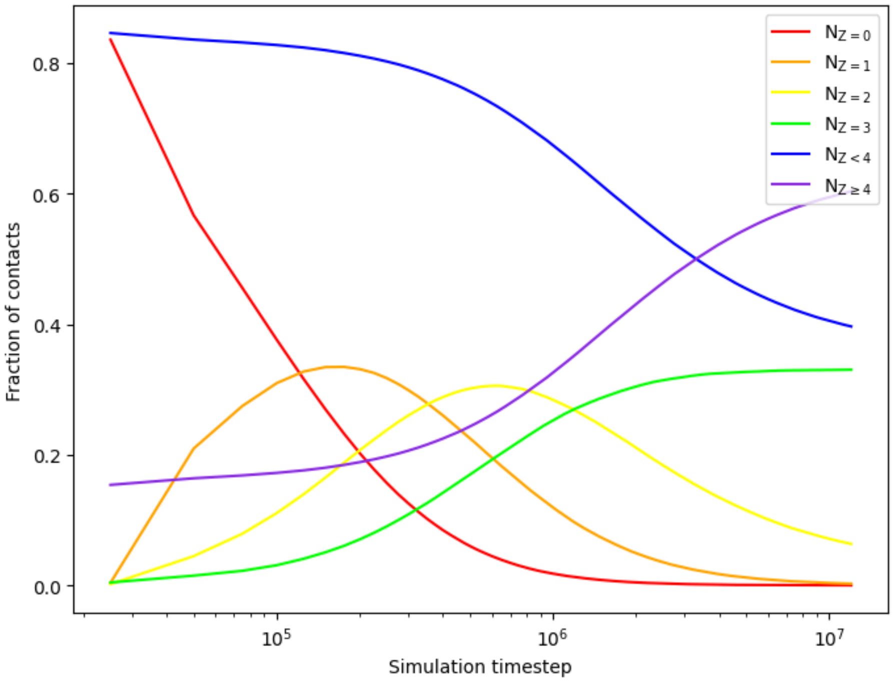


Figure 4.5: Share of total number of particles grouped by coordination number during adhesion phase with logarithmic timescale.

Fig. 4.6 shows the development in global coordination number Z during adhesion phase, where ΔZ is the absolute difference of Z with previous timestep. The change in number of non-rattler particles $\Delta N_{Z\geq4}$ normalized by the total number of particles is also included. Z increases at a declining rate, as shown by ΔZ. At the end of the adhesion phase the rate of change in coordination number is at its smallest.

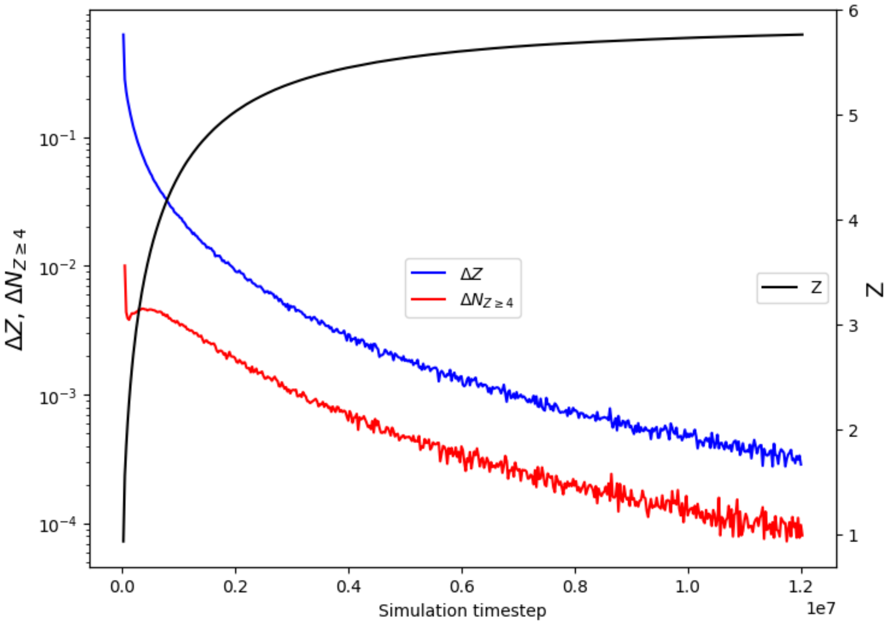
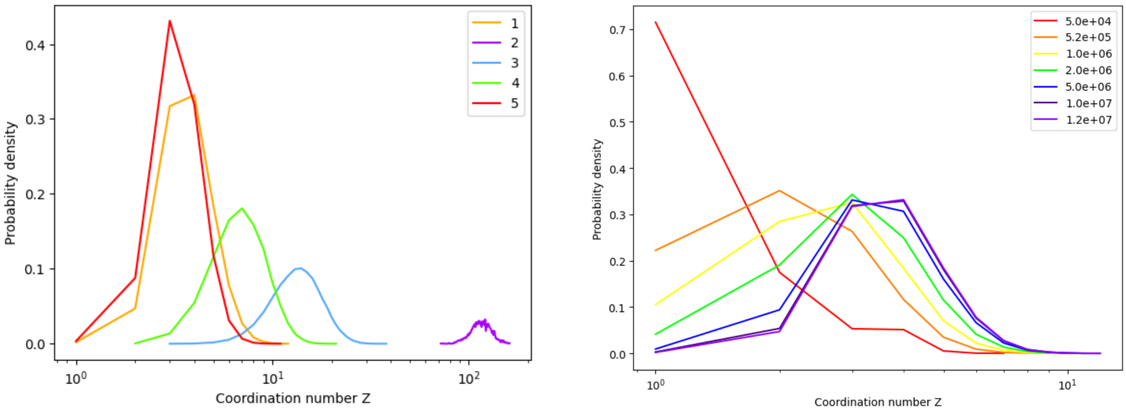


Figure 4.6: Development of global coordination number and change in fraction of non-rattler particles during adhesion phase.

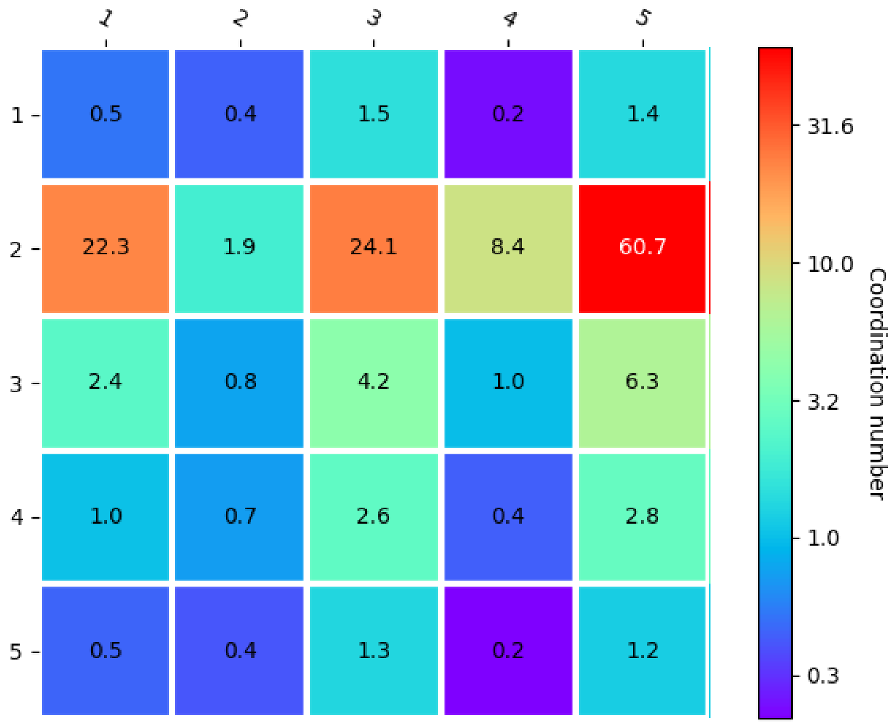
Fig. 4.7(a) shows the component-wise coordination number distribution for discrete number of contacts $Z = 1, 2, 3, \ldots$, after the adhesion phase is finished. All components have the appearance of some left skewed probability density function, where its mode is ordered in the same manner as the component-wise mean particle size as given by Table E.1. For increasing coordination number the component-wise mode has a lower density due to the component assuming a larger set of discrete coordination numbers. Fig. 4.7(b) displays the change in coordination number of the API during the adhesion phase. Initially the mode is at coordination number 1, which increases for timestamps further into the adhesion phase. The distribution for the last three timestamps are similar, indicating that the API component has stopped developing contacts.



(a) All components after adhesion phase. **(b)** API during adhesion phase.

Figure 4.7: Probability density as a function of discrete coordination numbers: a) component-wise after adhesion phase is finished, with color-scheme corresponding to that in Fig. 4.9. b) For component 1 at 7 different timestamps of the adhesion phase.

Fig. 4.8 shows the pairwise coordination number between components. The matrix should be read row-wise, i.e. for row 2, column 1 the interpretation is: The coordination number for particles belonging to component 2 and having contacts with component 1 is 22.3 (number of contacts between component 2 and 1 divided by number of particles in component 2). The matrix shows that component 2 has the highest coordination number with all other components, which is due to it consisting of the largest particles. It shares on average 60.7 contacts per particle with component 5, which contains the smallest particles in the packing. At the same time component 5 with the smallest particles has the lowest coordination number with all other components.



	1	2	3	4	5
1	0.5	0.4	1.5	0.2	1.4
2	22.3	1.9	24.1	8.4	60.7
3	2.4	0.8	4.2	1.0	6.3
4	1.0	0.7	2.6	0.4	2.8
5	0.5	0.4	1.3	0.2	1.2

Figure 4.8: Pairwise coordination number between components with indexing in the same order as component ID:s are arranged in table E.1. Color is graded on a logarithmic scale to improve interpretability.

.

Below follows a visual representation of the packing after adhesion phase, with color coding corresponding to that in Fig. 4.7 and 4.1

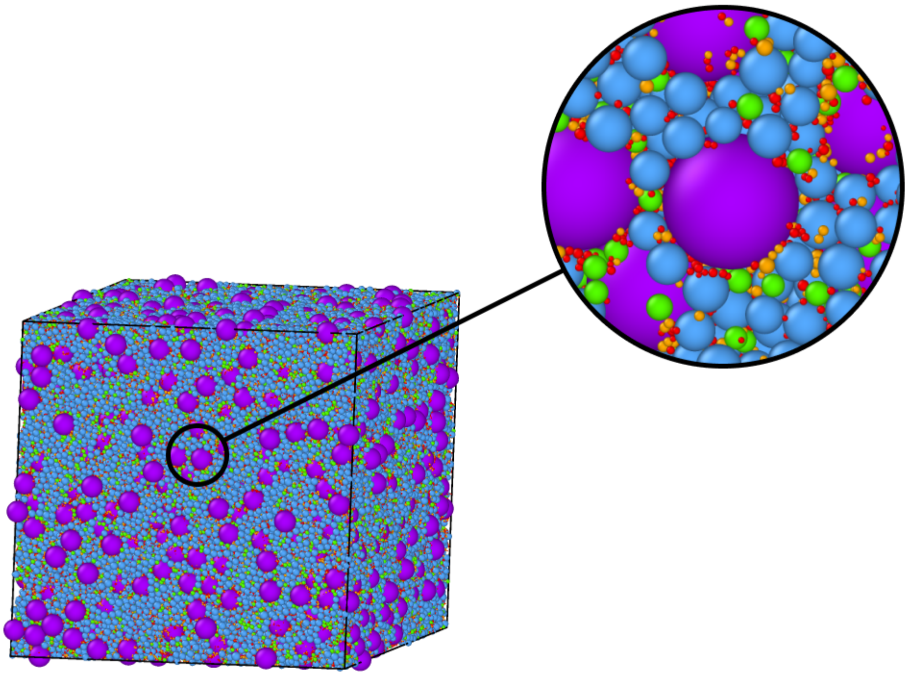


Figure 4.9: Packing of 605.000 particles given by Table E.1, after adhesion phase is finished. (1, orange): API, (2, purple): Plastic filler, (3, blue): Elastic filler, (4, green): Disintegrant and (5, red): Lubricant.

4.2 Packing models fitting and performance

Fig. 4.10 shows the relationship between the specific volume e and normalized area and volume, $A_i^{(m)}, m = 2, 3$. Comparing the normalization procedure, normalization by volume $m = 3$, yields a closer grouping for the first three packings (Appendix E, Table E.2, E.3, E.4, E.5) of the specific volume, but also for its sample standard deviation. It can be seen that for $m = 3$ that the specific volume curve is in no way universal, having a packing that deviates entirely (4), while for large $A_i^{(3)}$ the grouping also disappears. The specific volume curve of the packings 1,2,3 collapse as a function of $A_i^{(3)}$ for $A_i^{(3)} < 10$. This grouping can also be seen from the standard deviation of the specific volumes in Fig 4.11 for $m = 3$.

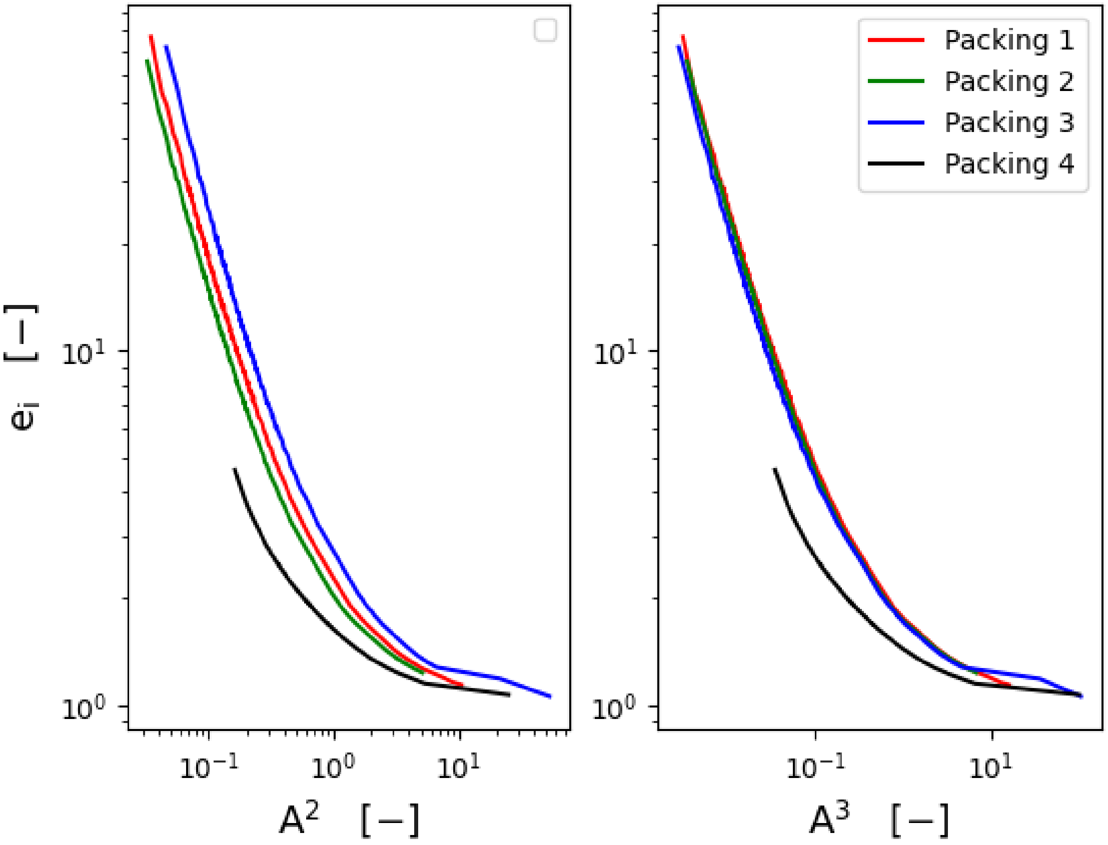


Figure 4.10: left: Bin normalized area $A_i^{(2)}$ Right: Normalized volume $A_i^{(3)}$ versus bin specific volume e_i for four multicomponent packings specified in Appendix E, using 200 bins. The curve grouping is noticeably closer for $m = 3$ than for $m = 2$, while packing 4 clearly breaks the trend.

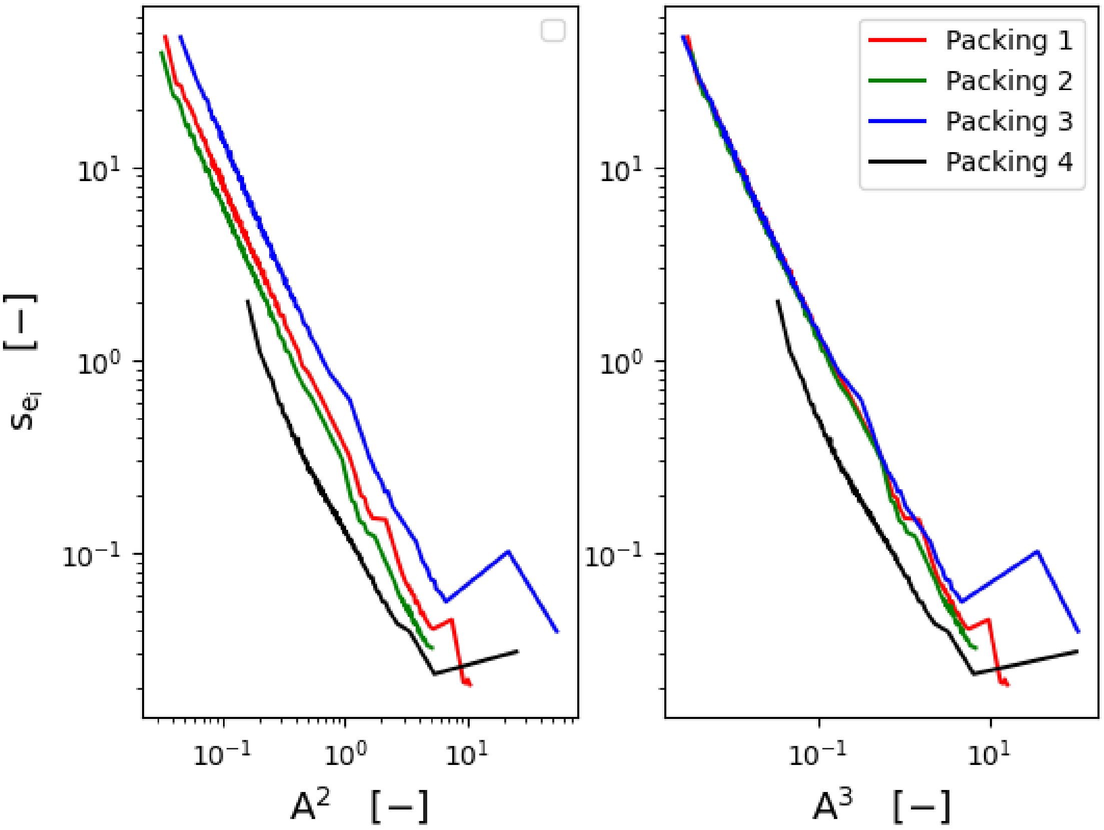


Figure 4.11: Left: Normalized area $A_i^{(2)}$ Right: Normalized volume $A_i^{(3)}$, versus standard deviation of the specific volume s_{e_i} for four multimodal packings specified in Appendix E, using 200 bins.

This section shows the results of fitting the LMPM model as well as comparing it against simulation results. Component diameters are set to the average component diameters. First a porosity η_0 minimizing the error between simulation and LMPM packing fractions are calculated. This is then used for all classes $\eta_{i,0} \equiv \eta_0$. The error minimizing porosity $\eta_{0,min} = 0.361$ from Fig. 4.12 yields the initial packing fraction $\phi_0 = 0.639$, close to the RCP packing fraction of monosized spheres [33]. This is reassuring, as this is the numerical interpretation of the parameter.

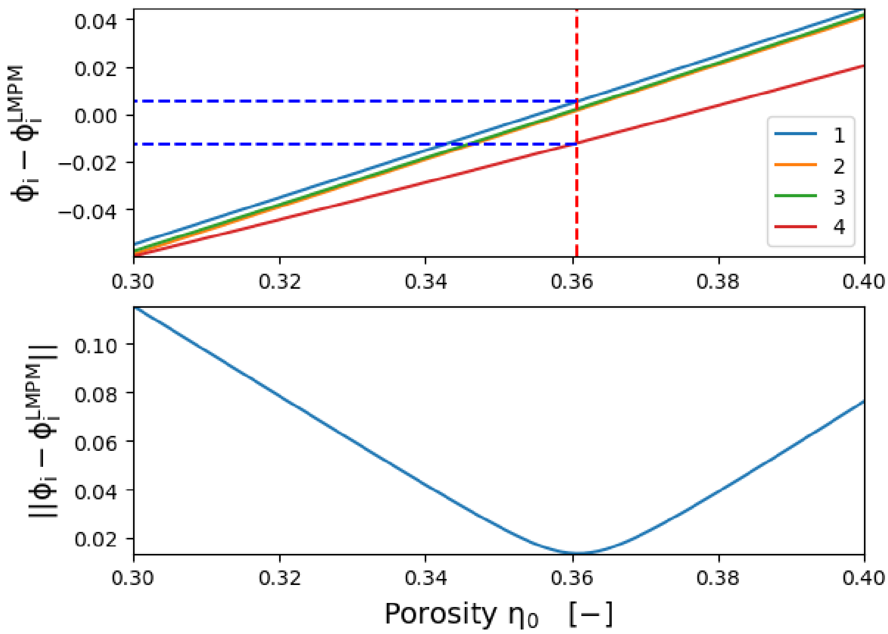


Figure 4.12: Comparison of packing fraction of simulations ϕ with estimations ϕ_{LMPM} using the LMP model. The lower plot shows the error $\phi - \phi^{\mathrm{LMPM}}$ vs. porosity, while the upper plot shows the difference $\phi - \phi_{\mathrm{LMPM}}$ for different packings. The red dashed line shows the error minimizing porosity $\eta_{0,\mathrm{min}} \approx 0.361$ and the blue shows the difference at $\eta_{0,\mathrm{min}}$. The maximum difference between simulation and the LMPM at $\eta_{0,\mathrm{min}}$ is 1.22%.

Bimodal packings, with modes of equal numbers of particles, with means $d_i, d_j \in \{5, 15\}$ and standard deviations $s_i = s_j \in \{1, \sqrt{2}\}$, were created. The total number of particles N was varied from 500 to 10000, with 40 simulations for each particle count N. Based on these the standard deviation of the packing fraction ϕ was calculated. The simulations were run until the expected change in ϕ over $8 \cdot 10^6$ timesteps were less than 1%. Based on these simulations the standard deviations $\sigma_\phi^{\mathrm{sim}} |\ N$ were then calculated. The packing fraction estimated by the LMPM is deterministic but as the component means d_i and solid volume fractions Y_i are stochastic, so is its prediction. The standard deviations of its predictions $\sigma_\phi^{\mathrm{LMPM}} |\ N$ are then compared against those of simulations.

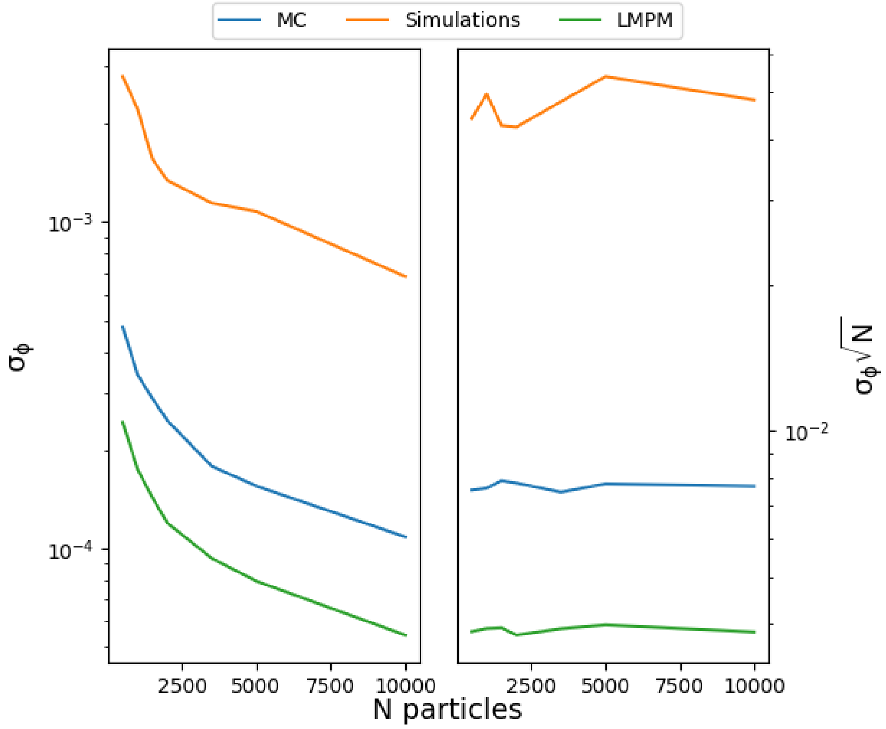


Figure 4.13: For means $[5, 15]$ and standard deviations $[1, 1]$. Left: Comparison of the standard deviation of the packing fraction for simulations, using the LMPM σ_ϕ^{LMPM} and by using the grouped specific volume curve, σ_ϕ^{MC} in Fig. 4.10 to calculate the packing fraction. The LMPM and MC standard deviations were based on 1000 particle resamplings. Right: The same standard deviations multiplied by $\sqrt{N}$. That $\sigma_\phi^{\text{LMPM}}$ is proportional to $N^{-1/2}$ is clearer than that σ_ϕ^{sim} is so.

Assuming that the standard deviation of the packing fraction is proportional to the LMPM standard deviation, a prefactor d^*, is calculated to minimize the error $||\sigma_\phi^{\text{sim}} - \sigma_\phi^{\text{LMPM}}||$ over all particle counts. Then $d^*\sigma_\phi^{\text{LMPM}}$ could be used as an approximation of σ_ϕ^{sim}, as shown in Fig. 4.14. The d^* is unique to each realization of $\sigma_\phi^{\text{LMPM}}|N$ and so has some distribution. By resampling and recalculating d^* repeatedly this distribution can be studied.

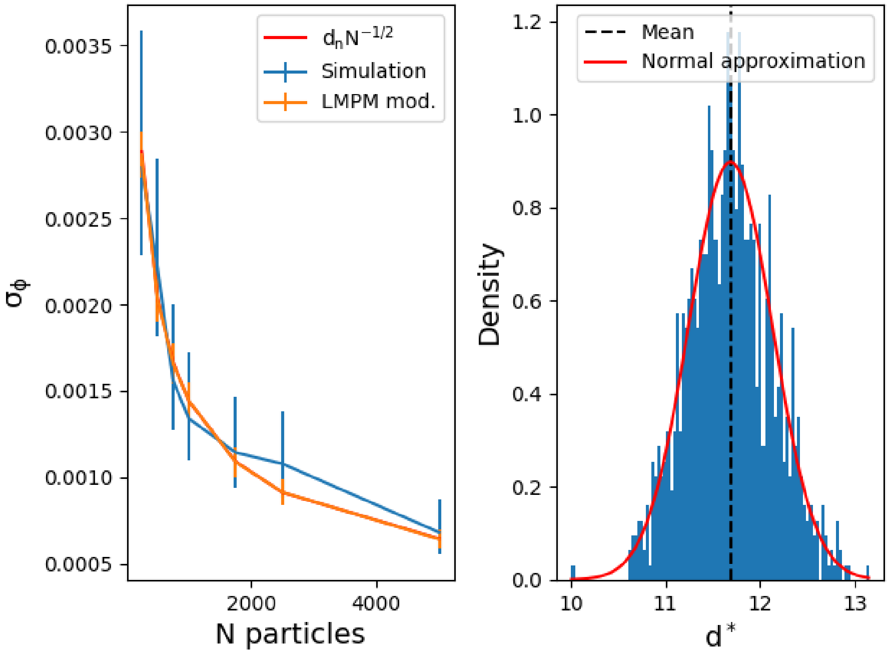


Figure 4.14: The standard deviation of the LMPM packing fraction is multiplied by a factor d^*, that minimizes the error between $\sigma_\phi^{\text{sim}}|\ N$ and $\sigma_\phi^{\text{LMPM}}|\ N$. Left : The mean and standard deviation of $d^*\sigma_\phi^{\text{LMPM}}|\ N$ over 1000 resamplings compared with $d^*\sigma_\phi^{\text{sim}}|\ N$. Also included is $\frac{d_n}{\sqrt{N}}$, where $d_n = \arg\min_d ||\sigma_\phi^{\text{sim}} - \frac{d}{\sqrt{N}}||$. Right: Empirical distribution of d^* and its normal approximation $\mathcal{N}(d^*; 11.69, 0.44)$, $C_v \approx 3.8\%$

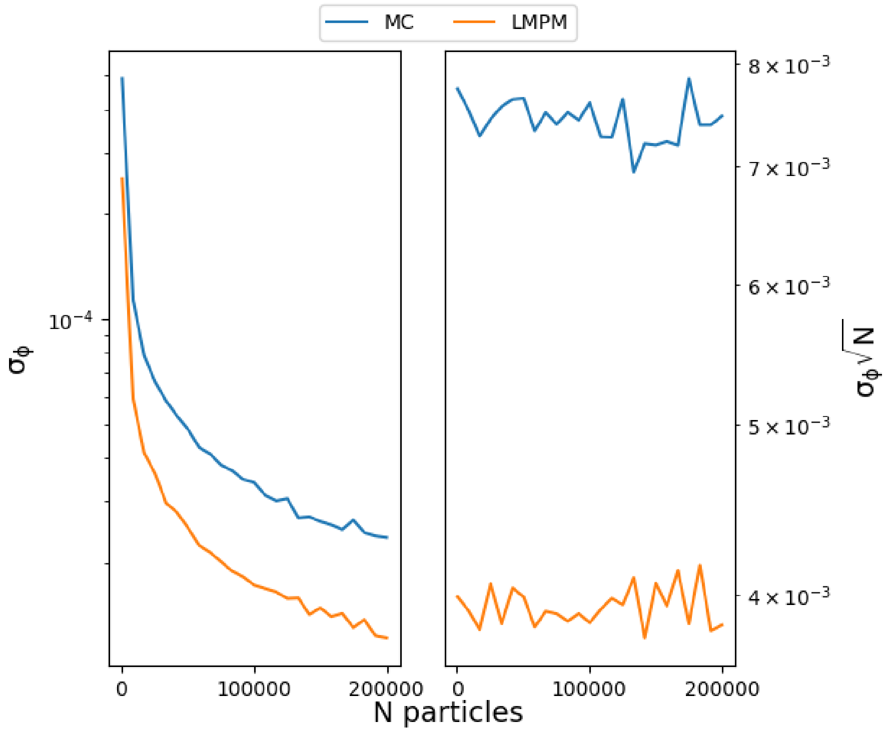


Figure 4.15: When increasing the particle count beyond the number simulated, the standard deviation σ_ϕ of the LMPM and using the master curve can be seen to be roughly proportional to $N^{-1/2}$

Based on Figs. 4.14, C.10, C.11 and C.12 distributions of d^* then seems well approximated by Gaussian distributions, with $C_v \in [3.6, 3.9]\%$. for the different PSDs. The means of d^* lie in the range $[6.96, 18.45]$, with $C_v \approx 42\%$. The variability in d^* then seem dominated by the chosen PSD. The range of d^*, means that it cannot be assumed constant, but neither does it span orders of magnitude. From Fig. 4.15 it can be seen that $s_\phi^{\mathrm{LMPM}} \propto N^{-1/2}$ roughly. Based on the four simulation series shown in Figs. 4.14, C.10, C.11, C.12 it can be hypothesized that $\sigma_\phi^{\mathrm{sim}}$ is roughly proportional to $\sigma_\phi^{\mathrm{LMPM}}$, such that $d^*\sigma_\phi^{\mathrm{LMPM}} \approx \sigma_\phi^{\mathrm{sim}}$.

4.3 Content uniformity

The concentric shell algorithm was applied across four different multi-component packings, where 1000 random coordinates were sampled within the cell. Concentration profiles and the COV for each component was calculated radially and compared to the Hilden, Eq. 2.29 and Stange, Eq. 2.26 analytic estimates. The shell thickness was specified as 0.1, a tenth of the smallest particle diameter. The calculated C_v of the concentration for different packings and components can be seen in Figs. 4.16, 4.17.

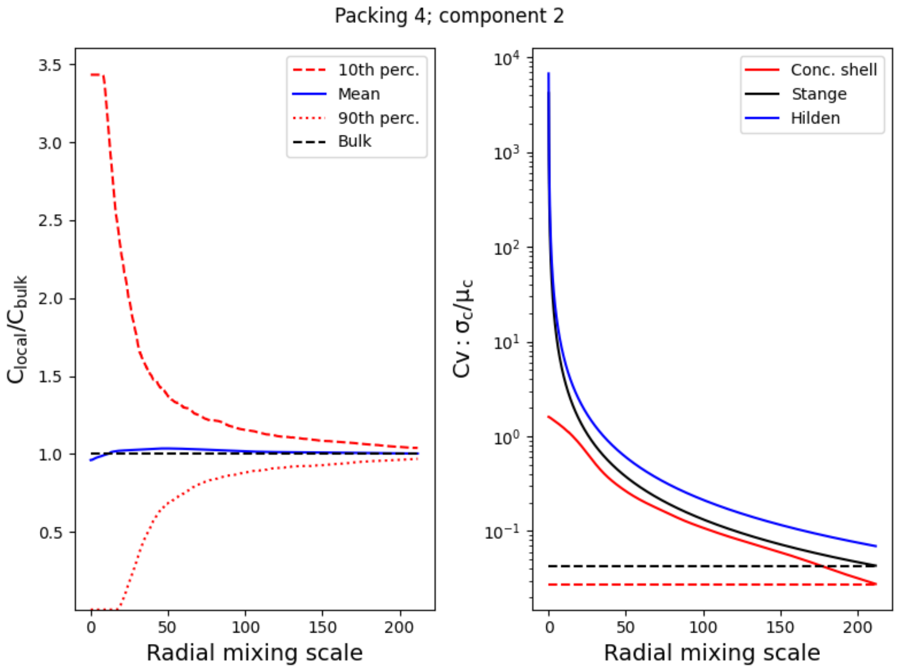
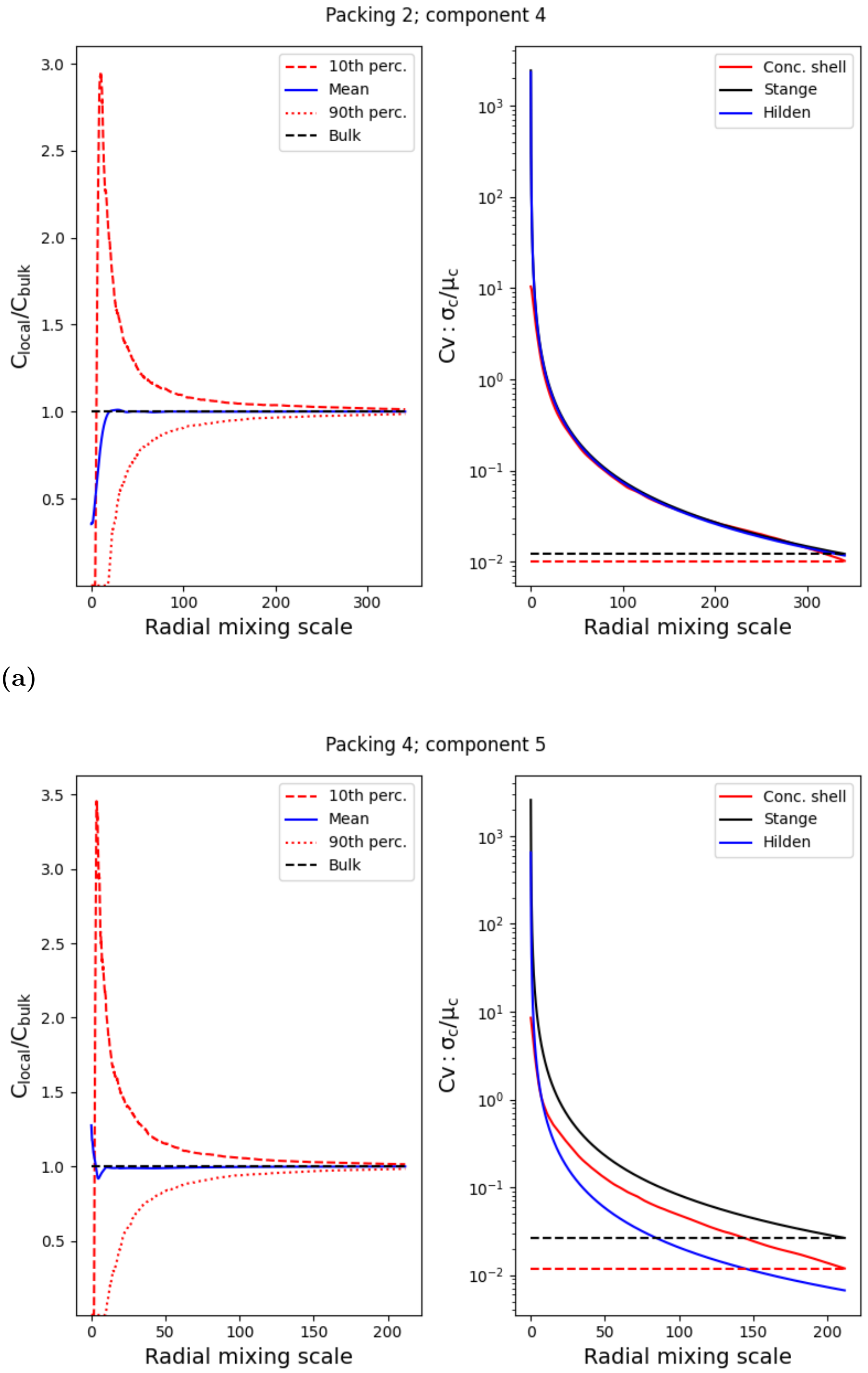


Figure 4.16: On the left is the bulk mean (blue) and percentiles (red) of the local component concentration normalized by the bulk concentration. On the right is the COV of the concentration calculated using the concentric shell method $C_c^{\text{Conc.}}(r)$ (red), $C_c^{\text{Stange}}(r)$ from 2.26 (black) and $C_c^{\text{Hilden}}(r)$ from Eq. 2.29 (blue). The dashed lines on the right show the values of $C_c^{\text{Conc.}}(r)$ and $C_c^{\text{Stange}}(r)$ for $r = L$, the maximum radial mixing scale. The size ratio $\lambda \approx 50$. It can be seen in this case that $C_c^{\text{Hilden}} > C_c^{\text{Stange}}$ contrary to Fig. 4.17. A decay of $C_c^{\text{Conc.}}(r)$ faster than $C_c^{\text{Stange}}(r)$ can be seen for r close to L. In this case Hilden overestimates Stange.



(a)

(b)

Figure 4.17: The figure content is similar to Fig. 4.16. (a) : A packing of size ratio $\lambda \approx 20$ where $\frac{C_c^{\text{Stange}}(L)}{C_c^{\text{Conc.}}(L)} = 1.193$ with good agreement for a smaller radial mixing scale. Hilden and Stange are approximately equal. (b) : Here $\lambda \approx 50$ and $\frac{C_c^{\text{Stange}}(L)}{C_c^{\text{Conc.}}(L)} = 2.226$. Hilden here underestimates Stange.

As the coordinates are randomly sampled, for components of low volume fraction sampled coordinates then tend to lie in voids between particles. The 10th percentile of the concentration then tends to be close to zero for small radial mixing scales, as shown above in the left side of Fig. 4.17.

The results for the remaining components of the two packings are shown in C.1. For all packings studied, the Stange estimate overestimates $C_c^{\text{Conc.}}$ for low radial mixing scales, but for some components $C_c^{\text{Conc.}}$ is slightly larger than C_c^{Stange} for large radial mixing scales. The Hilden estimate is inconsistent when comparing with that calculated and Stange, sometimes underestimating (Fig. 4.17), sometimes overestimating (Fig, 4.16) the concentration. The inconsistency of Hilden with regards to Stange can be explained analytically, see Sec. 4.3.1.

For components with good agreement between C_c^{Stange} and $C_c^{\text{Conc.}}$, the radial mixing scale is large enough that $C_c^{\text{Stange}} \approx 1\%$, these are the packings with a smaller size ratio. While for components with larger disagreement $C_c^{\text{Stange}} \geq 1\%$ up to 4%, a larger radial mixing scale is then needed to achieve as low a C_c^{Stange}.

The disagreement of C_c^{Stange} and $C_c^{\text{Conc.}}$ is consistently such that the magnitude of the disagreement $\frac{C_c^{\text{Stange}}}{C_c^{\text{Conc.}}} \leq 2.5$, is within an order of magnitude. For some components $C_{c(i)}^{\text{Shell}}$ falls off at the largest radial mixing scales.

4.3.1 Hilden and Stange discrepancy

The Hilden and Stange predictions show discrepancies when compared which can be explained analytically. Consider setting the sample volume to that of the cell, $V = V_{cell}$, with an n-component mixture with mean particle masses w_i and variances s_i^2. The estimate of Stange Eq. 2.26 can be rewritten by noting that

$$\frac{w_i^2 + s_i^2}{w_i} = \frac{\pi}{6}\frac{E_{f_i}[d_i^3]^2 + [E_{f_i}[d_i^6] - E_{f_i}[d_i^3]^2]}{E_{f_i}[d_i^3]} = \frac{\pi}{6}\frac{E_{f_i}[d_i^6]}{E_{f_i}[d_i^3]} = \frac{\pi}{6}D_i[6,3]^3.$$

Defining the component mass fractions τ_i, their density ρ_i and the component i dose G_i, the original expression for the C_v of the drug dose by Hilden et al. [3], can be compared to that of Stange Eq. 2.24, after expressing Stange in terms of $D_i[6,3]$

$$C_{m(i)}^{\text{Hilden}} = \sqrt{\frac{\pi}{6G}\frac{1}{\tau_i}\rho_i D_i[6,3]^3}$$

$$C_{m(i)}^{\text{Stange}} = \sqrt{\frac{\pi}{6G}\left[\frac{\tau_{j\setminus i}^2}{\tau_i}\rho_i D_i[6,3]^3 + \sum_{j\neq i}^{n}\tau_j\rho_j D_j[6,3]^3\right]}$$

where these are the COV of mass fractions. Rough equivalence between them then requires that

$$\rho_i D_i[6,3] \approx \tau_{j\setminus i}^2\rho_i D_i[6,3]^3 + \tau_i\sum_{j\neq i}^{n}\tau_j\rho_j D_j[6,3]^3.$$

The difference between the predictions of Hilden and Stange can then be explained by the expression Eq. 4.3.1 above. The Stange expression is then a generalization of the Hilden one, including the effect of both non-API volume fractions through $\tau_{j\setminus i}$, and their particle size distribution through $D_j[6,3]^3$, $j \neq i$. For high volume fractions $\tau_i \not\approx 0$ then $\tau_{j\setminus i} \not\approx 1$. And as $\tau_i \not\approx 0$, the effect of non-API polydispersity (the sum on the right, Eq. 4.3.1) increases the discrepancy from Hilden. The Hilden estimate then need not be conservative for increasing volume fractions τ_i.

Then $C_{m(i)}^{\text{Stange}}$ can be expressed as a function of $C_{m(i)}^{\text{Hilden}}$ like

$$\left(C_{m(i)}^{\text{Stange}}\right)^2 = \left(\tau_{j\setminus i} C_{m(i)}^{\text{Hilden}}\right)^2 + \sum_{i \neq j}^{n} \left(\tau_j C_{m(j)}^{\text{Hilden}}\right)^2$$

where the total sample mass is $G = \frac{G_i}{\tau_i}$. The square of the homogeneity measure of Stange for a component can then be written as a weighted sum of the squares of Hilden for all components. A condition for the estimation of Stange to be conservative with regards to Hilden is then that

$$C_{m(i)}^{\text{Hilden}} \geq \sqrt{\frac{\sum_{i \neq j}^{n} \left(\tau_j C_{m(j)}^{\text{Hilden}}\right)^2}{1 - \tau_{j\setminus i}^2}}$$

5

Discussion

5.1 Simulation

The compression phase achieves a steady development of the density, as shown in Fig 4.2 while ensuring particle overlaps are maintained low at end of phase (Appendix A.3, Fig. A.3). As the packing develops a force-bearing network it is able to withstand the external pressure applied by the NPH-ensemble causing the density to stop developing. Additionally the isostatic condition is satisfied, indicating that a jammed state has been achieved. In this regard the exit conditions in Fig. 4.3 work well. Switching contact model and removing the external pressure did not disrupt the structure of contacting particles too much as shown by Fig. 4.4.

The adhesion phase had no set exit condition, as it was unsure what the final packing structure should look like. Given particles are frictionless it is theoretically possible for a particle which has cohered to another particle to move along its surface indefinitely. This does not happen since the sliding particles come into contact and cohere to other neighboring particles until they are geometrically constrained, a claim supported by Fig. 4.5. With a high share of particles initially having 0 contacts, they quickly find particles to cohere to, followed by sliding along the particle surface and cohering to other neighboring particles, increasing their number of contacts. Even at the end of the phase the figure suggests particles develop new contacts, where $N_{Z=3}$ remains constant due to it gaining about the same number of particles from $N_{Z=2}$ group as it loses to the $N_{Z\geq 4}$ group. Even though the change in global coordination number ΔZ is very low at the end of phase, as displayed in Fig. 4.6 the packing is still developing. Thus setting an exit condition based on ΔZ requires that the threshold is low enough to guarantee particles do not develop further contacts. Another option could be to use the change in number of non-rattler particles $N_{Z\geq 4}$ normalized by total number of particles. This relates well to the 1-stable condition defined in Sec. 2.1.2 and thus makes more sense in the context of measuring how many particles that are geometrically constrained by its neighbors. A third option which could be viable is monitoring the velocity of particles. When a particle has come to a halt it is assumed to be geometrically constrained for example. However all of these proposals are only viable for frictionless system. In a system with friction and no external pressure the particles would *probably* go to a halt sooner and stop developing contacts even though they are not geometrically constrained by their neighbors.

The component-wise coordination number distributions after adhesion phase in Fig. 4.7(a) suggest some left-skewed probability distribution. The previously mentioned behavior of particles reaching a state where they are geometrically constrained by their neighbors and thus not developing new contacts is also supported by Fig 4.7b. In a state with friction the adhesion phase would terminate earlier and the component-wise coordination number distributions would *probably* assume a form similar to that shown in Fig 4.7(b) at earlier timestamps.

5.2 Content uniformity

Both the Hilden and Stange predictions deviate from the calculated COV for some packing components, but the Stange prediction is more reliable. The discrepancy between the Hilden and Stange estimates is studied in Sec. 4.3.1, and is explainable by the dilute API assumption of Hilden [3], and through only considering one component. The Hilden estimate is based purely on the API PSD, and so it is natural that the concentric shell algorithm deviates from this for high volume fractions τ_i, and when the contribution of the other components to the content uniformity are large.

For several components it can be seen that the COV falls off at the highest radial mixing scales. Different to how the concentric shell method was used by Mort [41] where a unique structure was generated for each sampling sphere, here the concentric shells subsample a packing and are allowed to share domains. The radial mixing scale is large compared to the size of the cell, the largest sampling shell includes $40 - 50\%$ of the total cell volume, meaning that for large mixing scales the different sample volumes share large parts of the packing structure, on average. The concentration standard deviation then artificially decreases.

For small sampling scales $r \approx 10$, the concentric shell method yields considerably smaller COV than either Hilden or Stange. The cause of this was deduced by Mort [41] to be that the concentric shell algorithm constructs intersected sample volumes, therefore representing the small scale mixing of components better, as both Hilden and Stange is based on the inclusion of whole particles. The effect of slicing particles decrease as the radial mixing scale increases. Keeping in mind then that Eq. 2.26 is for an ideal mixture, then that $C_c^{\text{Conc.}} \geq C_c^{\text{Stange}}$ for large enough mixing scales, makes sense. The effect of intersecting sampling volumes could then explain why $C_c^{\text{Conc.}} \leq C_c^{\text{Stange}}$ for an even larger radial mixing scale.

Packings of size ratio $\lambda \approx 20$ (packings 1,2), yielded better agreement between the results of the concentric shell algorithm and that of Stange than when $\lambda \approx 50$, but even then they are within an order of magnitude of each other. A hypothesis for the cause of this discrepancy is then that the radial mixing scale L is too small compared to the size of the largest particles for packings 3,4.

5.3 Packing fraction variability

The variability of the packing fraction was investigated in Sec. 4.2, by fitting the LMPM model to simulation data and comparing the packing fraction standard deviations of it and simulation. It was seen that the correcting factors d^* roughly follow Gaussian distributions with non-constant means and standard deviations, though the COVs were grouped $C_v \in [3.6, 3.9]\%$. The wide range of the mean of d^*, for packings that are quite similar, means that it cannot be assumed constant. Without a structured approach to select d^*, i.e modeling its distribution $f(d^*)$ given packing PSDs, a conservative approach could be to specify it as the maximum of those calculated $d = \max\{d^*\}$ or the maximum of the means.

Rigorously testing whether s_ϕ^{sim} is roughly proportional to $\sigma_\phi^{\text{LMPM}}$, requires further simulations series (for increasing N), as deviation from this can be seen for larger N for one series, Fig. C.10. It should be noted that simulation results are dependent on compression history, the initial distribution of particles and the formulation of the exit conditions, and cannot merely be described by the particle sample, which fully determine the LMPM prediction.

The large difference between $\sigma_\phi^{\text{LMPM}}$ and σ_ϕ^{sim} can perhaps be accounted for by the use of the component means as bin diameters. Instead using bins $\mathcal{B}_i := \{d_j \in [D_{i-1}, D_i]\}$ defined by some spacing $D_1, \dots, D_k$ could decrease this error, as the LMPM would be more responsive to individual particles. That $\sigma_\phi^{\text{LMPM}} \propto N^{-1/2}$ may be due to the convergence in distribution of the sample means, though this ignores the effect of the solid volume fractions X_i.

5.4 Granocentric model

A way to holistically study the packing fraction variability, content uniformity as well as microstructural relationships like the master curve could be through the local geometric, stochastic model of Clusel et al. [44, 45], for polydisperse hard-sphere packings. The model is able to predict local density fluctuations, coordinate number and neighbor distributions after fitting to experimental data. The packing structure is modeled by filling the available solid angle by stochastically sampled neighboring particles, from the particle size distribution $P(r)$, with solid angles Ω_i such that the solid angle threshold Ω^* is reached. One of the underlying assumptions of the model is that $p(r|r_c) = P(r)$, but for higher polydispersities the size distribution of neighbors for a given particle i is particle dependent, and not necessarily that of the bulk [46]. Particle randomly determined to be non-contacting are then move an average distance δ^* from the central particle. For a given realization, the model then predicts the number of neighbors n, contacts z and the total filled solid angle Ω^*. Calculating the packing fraction is done constructing a cell similar to those of the radical Voronoi tesselation.

Further developments of the model study the effect of adhesion [47], and include the ability to calculate the number of rattlers and the radial distribution function $g(r)$ [48]. Testing this approach on larger polydispersities and for multi-modal PSDs could then perhaps give insight into the microstructure of these packings.

The specific volume curve shows how binned particles in general packs, but distributional information that is obscured when using mean properties (e and $A^{(m)}$)), should be looked at. Using a granocentric model the specific volume curve in 2.5.1 could be studied, through the specific volume distribution for a given particle size $P(e|r)$, where $e = \frac{V_p}{V_{cell}}$. Its microstructural cause could then perhaps be explained. Similar to the way the variability of the packing fraction can be studied using the LMP model, this can also be done using a granocentric model, where the local variability of the packing fraction is available.

6

Conclusion

The aim of the thesis was to develop a framework in order to study bulk product characteristics. A DEM simulation framework was developed in order to create highly polydisperse multi-component packings of adhesive, frictionless spheres. From this, bulk characteristics were collected and presented. However, the validity of these results are dependent on homogeneity measurements ensuring that the simulation scale is large enough to be representative. An attempt was made to relate packing fraction variation to the simulation scale using packing fraction models fitted to simulation results but this was not achieved satisfactorily. There is limited evidence showing that the packing fraction standard deviation of simulations roughly follow $N^{-1/2}$, and can be approximated through an upscaled standard deviation of the LMPM. This should be tested with further simulations, of increasing N to conclusively prove. Theoretical underpinnings for or against this should be sought as well. It should be tested if a further discretization of the PSDs for the LMPM yield closer standard deviation fits, without using the prefactor d^*. Though a more fruitful area of research might be to study the local uniformity of the packing fraction for larger simulations (through e.g concentric shell analysis and use of the LMPM), as this more closely approximates the batching of small doses.

The applicability of the UDU estimation of Hilden, in comparison with that of Stange is conditional on both component volume fractions as well as the particulars of component PSDs. It was shown to be a simplification of the Stange multi-component homogeneity index, in the limit of small component volume fraction and disregarding non-API polydispersity, with Stange being expressable in terms of Hilden. Homogeneity measurements using the concentric shell method requires a sufficiently large radial mixing scale to converge with Stange. But the issue of highly intersecting and correlating sampling volumes counteract this.

Bibliography

[1] G. Pifferi and P. Restani, “The safety of pharmaceutical excipients,” *Il Farmaco*, vol. 58, no. 8, pp. 541–550, 2003.

[2] J. Hilden, C. L. Burcham, S. D. Stamatis, J. Miesle, and C. A. Coutant, “Guidance on drug substance particle size controls for solid oral dose forms,” *Particles and Nanoparticles in Pharmaceutical Products: Design, Manufacturing, Behavior and Performance*, pp. 277–302, 2018.

[3] J. Hilden, M. Schrad, J. Kuehne-Willmore, and J. Sloan, “A first-principles model for prediction of product dose uniformity based on drug substance particle size distribution,” *Journal of pharmaceutical sciences*, vol. 101, no. 7, pp. 2364–2371, 2012.

[4] P. Cundall and O. Strack, “A discrete numerical model for granular assemblies,” *Geotechnique*, vol. 29, pp. 47–65, Mar. 1979.

[5] V. Ogarko and S. Luding, “A fast multilevel algorithm for contact detection of arbitrarily polydisperse objects,” *Computer physics communications*, vol. 183, no. 4, pp. 931–936, 2012.

[6] T. et al., “LAMMPS-a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales,” *Computer Physics Communications*, vol. 271, p. 108171, 2022.

[7] T. Shire, K. J. Hanley, and K. Stratford, “DEM simulations of polydisperse media: efficient contact detection applied to investigate the quasi-static limit,” *Computational Particle Mechanics*, vol. 8, no. 4, pp. 653–663, 2021.

[8] R. M. Nedderman, *Statics and Kinematics of Granular Materials.* Cambridge University Press, 1992.

[9] J. Seville, C. Willett, and P. Knight, “Interparticle forces in fluidisation: a review,” pp. 261–268, Dec. 2000. NEPTIS-7 Meeting on Emereging Frontiers in Fluidization, Dec, 1998. OTSU, Japan ; Conference date: 01-12-2000.

[10] H. M. Jaeger, S. R. Nagel, and R. P. Behringer, “Granular solids, liquids, and gases,” *Reviews of Modern Physics*, vol. 68, pp. 1259–1273, Oct 1996.

[11] R. P. Behringer and B. Chakraborty, “The physics of jamming for granular materials: a review,” *Reports on Progress in Physics*, vol. 82, no. 1, p. 012601, 2018.

[12] C. Thornton, “Granular dynamics, contact mechanics and particle system simulations,” *A DEM study. Particle Technology Series*, vol. 24, 2015.

[13] F. Göncü, O. Durán, and S. Luding, “Constitutive relations for the isotropic deformation of frictionless packings of polydisperse spheres,” *Comptes Rendus Mécanique*, vol. 338, no. 10-11, pp. 570–586, 2010.

[14] A. Bezdek, K. Bezdek, and R. Connelly, “Finite and uniform stability of sphere packings,” *Discrete and Computational Geometry*, vol. 20, pp. 111–130, 07 1998.

[15] D. et al., “A linear programming algorithm to test for jamming in hard-sphere packings,” *Journal of Computational Physics*, vol. 197, no. 1, pp. 139–166, 2004.

[16] J. M. Monti, J. T. Clemmer, I. Srivastava, L. E. Silbert, G. S. Grest, and J. B. Lechman, “Large-scale frictionless jamming with power-law particle size distributions,” *Physical Review E*, vol. 106, no. 3, p. 034901, 2022.

[17] F. Göncü, O. Durán, and S. Luding, “Jamming in frictionless packings of spheres: determination of the critical volume fraction,” in *AIP Conference Proceedings*, vol. 1145, pp. 531–534, American Institute of Physics, 2009.

[18] C. S. O’Hern, L. E. Silbert, A. J. Liu, and S. R. Nagel, “Jamming at zero temperature and zero applied stress: The epitome of disorder,” *Physical review E*, vol. 68, p. 011306, Jul 2003.

[19] N. et al., *Coupled CFD-DEM Modeling: Formulation, Implementation and Application to Multiphase Flows.* 10 2016.

[20] H. Hertz, “Ueber die berührung fester elastischer körper.,” vol. 1882, no. 92, pp. 156–171, 1882.

[21] R. D. Mindlin, “Compliance of Elastic Bodies in Contact,” *Journal of Applied Mechanics*, vol. 16, pp. 259–268, 04 2021.

[22] R. D. Mindlin and H. Deresiewicz, “Elastic Spheres in Contact Under Varying Oblique Forces,” *Journal of Applied Mechanics*, vol. 20, pp. 327–344, Sept. 1953.

[23] K. Johnson, K. Kendall, and A. Roberts, “Surface energy and contact of elastic solids,” *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 324, pp. 301–313, 09 1971.

[24] Y. Tsuji, T. Tanaka, and T. Ishida, “Lagrangian numerical simulation of plug flow of cohesionless particles in a horizonal pipe,” *Powder Technology*, vol. 71, pp. 239–250, 09 1992.

[25] J. Marshall, “Discrete-element modeling of particulate aerosol flows,” *Journal of Computational Physics*, vol. 228, no. 5, pp. 1541–1561, 2009.

[26] M. Murgai, “Application of the hertz theory of impact to explosion phenomena,” *The Journal of Chemical Physics*, vol. 22, no. 10, pp. 1687–1689, 1954.

[27] A. Chokshi, A. Tielens, and D. Hollenbach, “Dust coagulation,” *Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 407, no. 2, p. 806-819.*, vol. 407, pp. 806–819, 1993.

[28] C. Thornton and Z. Ning, “A theoretical model for the stick/bounce behaviour of adhesive, elastic-plastic spheres,” *Powder technology*, vol. 99, no. 2, pp. 154–162, 1998.

[29] S. Chen, S. Li, and M. Yang, “Sticking/rebound criterion for collisions of small adhesive particles: Effects of impact parameter and particle size,” *Powder Technology*, vol. 274, pp. 431–440, 2015.

[30] L. Verlet, “Computer experiments on classical fluids. I. thermodynamical properties of lennard-jones molecules,” *Physical review*, vol. 159, pp. 98–103, Jul 1967.

[31] W. Shinoda, M. Shiga, and M. Mikami, “Rapid estimation of elastic constants by molecular dynamics simulation under constant stress,” *Physical Review B*, vol. 69, no. 13, p. 134103, 2004.

[32] C. Rycroft, “Voro++: A three-dimensional voronoi cell library in c++,” tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2009.

[33] A.-B. Yu and N. Standish, “Estimation of the porosity of particle mixtures by a linear-mixture packing model,” *Industrial & engineering chemistry research*, vol. 30, no. 6, pp. 1372–1385, 1991.

[34] R. M. Frings, M. G. Kleinhans, and S. Vollmer, “Discriminating between pore-filling load and bed-structure load: a new porosity-based method, exemplified for the river rhine,” *Sedimentology*, vol. 55, no. 6, pp. 1571–1593, 2008.

[35] R. et al., “On the use of packing models for the prediction of fluvial sediment porosity,” *Earth Surface Dynamics Discussions*, pp. 1–29, 2023.

[36] Y. Yuan, W. Deng, and S. Li, “Structural universality in disordered packings with size and shape polydispersity,” *Soft Matter*, vol. 16, no. 18, pp. 4528–4539, 2020.

[37] J. A. Cornell, “Experiments with mixtures: a review,” *Technometrics*, vol. 15, no. 3, pp. 437–455, 1973.

[38] L. Fan, S. Chen, and C. Watson, “Annual review solids mixing,” *Industrial & Engineering Chemistry*, vol. 62, no. 7, pp. 53–69, 1970.

[39] G. et al., “Multi-scale homogeneity analysis of co-milled powders: Development of a reverse approach to assess quality of mixtures,” *Powder Technology*, vol. 400, p. 117263, 2022.

[40] K. Stange, “Die mischgüte einer zufallsmischung aus drei und mehr komponenten,” *Chemie Ingenieur Technik*, vol. 35, no. 8, pp. 580–582, 1963.

[41] P. R. Mort and R. E. Riman, “Determination of homogeneity scale in ordered and partially ordered mixtures,” *Powder technology*, vol. 82, no. 1, pp. 93–104, 1995.

[42] W. F. Oquendo-Patiño and N. Estrada, “Densest arrangement of frictionless polydisperse sphere packings with a power-law grain size distribution,” *Granular Matter*, vol. 22, pp. 1–8, 2020.

[43] A. Stukowski, “Visualization and analysis of atomistic simulation data with OVITO-the Open Visualization Tool,” *Modelling and simulation in Materials Science and Engineering*, vol. 18, JAN 2010.

[44] M. Clusel, E. I. Corwin, A. O. Siemens, and J. Brujić, “A ‘granocentric’model for random packing of jammed emulsions,” *Nature*, vol. 460, no. 7255, pp. 611–615, 2009.

[45] N. et al., “A statistical mechanics framework captures the packing of monodisperse particles,” *Soft Matter*, vol. 7, no. 24, pp. 11518–11525, 2011.

[46] C. B. O’Donovan, E. I. Corwin, and M. E. Möbius, “Mean-field granocentric approach in 2d & 3d polydisperse, frictionless packings,” *Philosophical Magazine*, vol. 93, no. 31-33, pp. 4030–4056, 2013.

[47] W. Liu, S. Chen, C.-Y. Wu, and S. Li, “Unified size-density and size-topology relations in random packings of dry adhesive polydisperse spheres,” *Physical Review E*, vol. 99, no. 2, p. 022901, 2019.

[48] Z. et al., “Structure of marginally jammed polydisperse packings of frictionless spheres,” *Physical Review E*, vol. 91, no. 3, p. 032302, 2015.

[49] S. Chen, W. Liu, and S. Li, “A fast adhesive discrete element method for random packings of fine particles,” *Chemical Engineering Science*, vol. 193, pp. 336–345, 2019.

[50] V. Ogarko and S. Luding, “A study on the influence of the particle packing fraction on the performance of a multilevel contact detection algorithm,” in *Particles II: proceedings of the II International Conference on Particle-Based Methods: fundamentals and applications*, pp. 364–370, CIMNE, 2011.

[51] D. Krijgsman, V. Ogarko, and S. Luding, “Optimal parameters for a hierarchical grid data structure for contact detection in arbitrarily polydisperse particle systems,” *Computational particle mechanics*, vol. 1, no. 3, pp. 357–372, 2014.

[52] M. Capece, K. R. Silva, D. Sunkara, J. Strong, and P. Gao, “On the relationship of inter-particle cohesiveness and bulk powder behavior: Flowability of pharmaceutical powders,” *International journal of pharmaceutics*, vol. 511, no. 1, pp. 178–189, 2016.

A

Parameter studies

A.1 Surface energy

The first contact energy loss of the JKR contact model W means that the initial v_i and rebound velocity v_r are related by

$$\frac{mv_i^2}{2} - \frac{mv_r^2}{2} = W \;:\; v_i^2 - v_r^2 = \frac{2W}{m}.$$

With a rebound velocity precisely zero, $v_r = 0$, then the initial velocity is called the *sticking velocity* v_s, such that a sticking criterion can be determined [12]

$$v_s = \sqrt{\frac{2W}{m}} = \left(\frac{14.18}{m}\right)^{1/2} \left(\frac{w^5 R^4}{E^2}\right)^{1/6}$$

The effective mass is expressed as

$$m := \frac{m_i m_j}{m_i + m_j} = \frac{4\pi}{3} \frac{\rho_i \rho_j r_i^3 r_j^3}{\rho_i r_i^3 + \rho_j r_j^3} = \frac{4\pi}{3} \frac{r_i^3 r_j^3}{r_i^3 + r_j^3} = \frac{4\pi}{3} f(r_i, r_j)$$

since $\rho_i \equiv 1$ for all components. The sticking velocity is then

$$v_s = 1.84 \left[\frac{w^5 R^4}{f(r_i, r_j)^3 E^2}\right]^{1/6}$$

By introducing the size ratio $\frac{r_j}{r_i} = \eta$, where $\eta \in \left[\frac{1}{2r_i}, \frac{\lambda}{2r_i}\right]$ the fraction of f and R can be expanded into

$$\frac{R^4}{f(r_i, r_j)^3} = \frac{(r_i^3 + r_j^3)^3}{(r_i r_j)^5 (r_i + r_j)^4} = \frac{(1+\eta^3)^3}{(r_i \eta)^5 (1+\eta)^4} = \frac{g(\eta)}{r_i^5}$$

A minima for $g(\cdot)$ is located at $\eta_m = \frac{1+\sqrt{6}}{2}$, where $g(\eta_m) \approx 0.274$. Then the sticking velocity of a particle of size r_i can be bounded from below by

$$v_s \geq 1.483 \left[\frac{w^5}{r_i^5 E^2}\right]^{1/6} \geq 1.483 \left[\frac{w^5}{r_{max}^5 E^2}\right]^{1/6},$$

with $r_{\max} = \lambda/2$. In the case of a particle of radius r impacting a wall

$$v_s = 1.84 \left[\frac{\left(\frac{w}{r}\right)^5}{E^2}\right]^{1/6}.$$

When specifying the JKR pull-off force through the surface energy, the particle overlap at equilibrium δ_e is considered

$$\delta_e = \left(\frac{4}{3}\right)^{2/3} \frac{a_0^2}{2(6)^{1/3}R} = \left(\frac{4}{3}\right)^{2/3} \frac{\left[\frac{9\pi\gamma R}{E}\right]^{2/3}}{2(6)^{1/3}R} = A\left[\frac{\gamma}{E}\right]^{2/3} R^{1/3}$$

$$A = \frac{(12\pi)^{2/3}}{2(6)^{1/3}} \approx 3.094$$

Then the relative overlap of particle i involved in an interaction is

$$\tilde{\delta}_e = \frac{\delta_c}{r_i} = A\left[\frac{\gamma}{E}\right]^{2/3} \frac{\left[\frac{r_i r_j}{r_i + r_j}\right]^{1/3}}{r_i} = A\left[\frac{\gamma}{E}\right]^{2/3} \left[\frac{r_j}{r_i^2(r_i + r_j)}\right]^{1/3}$$

For the radii $r_i, r_j \in [1/2, \lambda]$

$$\frac{r_j}{r_i^2(r_i + r_j)} \leq \frac{4r_j}{\frac{1}{8} + \frac{r_j}{4}} \leq \frac{4\lambda}{1+\lambda} \leq 4$$

Then

$$\tilde{\delta}_c \leq 4^{1/3} A\left[\frac{\gamma}{E}\right]^{2/3} = \epsilon,$$

where ϵ is the maximum allowed relative overlap at equilibrium. Given the effective Young's modulus E, then for particle-particle and particle-wall interactions the surface energy is specified such that

$$\gamma = \frac{E}{2}\left[\frac{\epsilon}{A}\right]^{3/2}. \tag{A.1}$$

A.2 Stick and rebound

When introducing viscous damping into the contact model, the sticking velocity increases from the merely adhesive case v_s [49]. The adhesive JKR model is introduced to make small particles adhere, the stick and rebound behaviour with regards to velocity is then investigated for small radii, for various surface energies γ. A granular wall is introduced to the simulation, which is modeled in LAMMPS by a sphere of infinite mass and radius, such that

$$R = r_i$$
$$m = m_i$$

The surface energy is specified by $\gamma = \frac{E}{2}\left[\frac{\tilde{\delta}}{A}\right]^{3/2}$ where $\tilde{\delta} \in [10^{-2}, 10^{-4}]$ is the allowed relative overlap between the particle and a granular wall. Collision velocities lie in the interval $[0.5v_s^*, 20v_s^*]$, where

$$v_s^* = 1.483\left[\frac{\left(\frac{w}{R}\right)^5}{E^2}\right]^{1/6}.$$

Through Eq. 2.9 the damping parameter α is determined from the CoR e^*. The actual CoR after a particle-wall interaction is here denoted by e, while the visco-elastic sticking velocity is denoted by v_s.

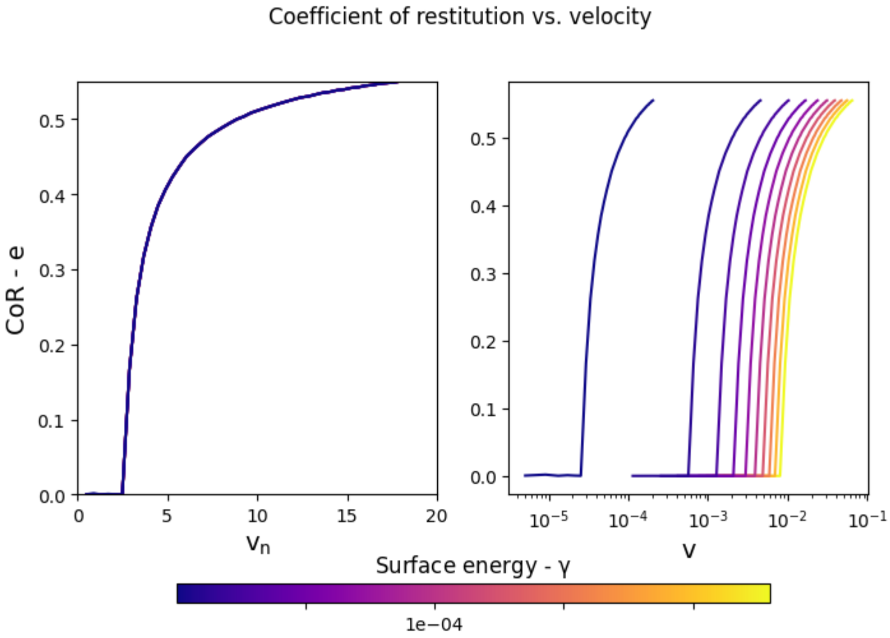


Figure A.1: Left: The is the CoR e for a particle-wall collision versus the dimensionless velocity $v_n = v/v_s^*$, it can be seen that e collapses as a function of v_n. Right: e versus v, for a selection of surface energies g. The magnitude of damping α was specified such that $e^* = 0.5$ in the non-adhesive case.

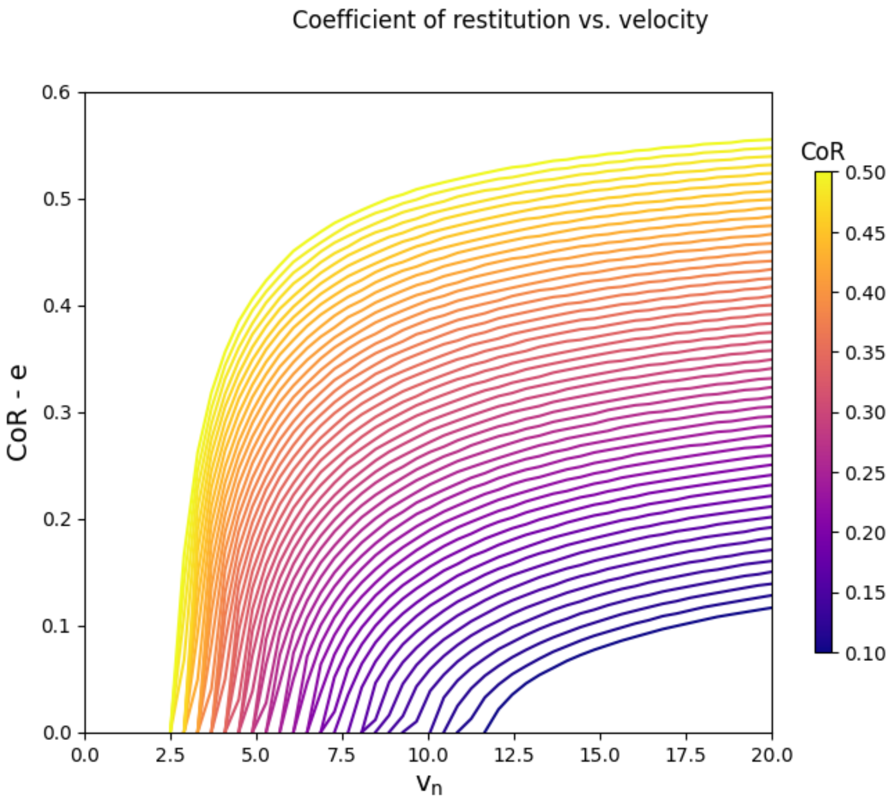


Figure A.2: Shown here is the CoR e for a particle-wall collision versus the dimensionless velocity $v_n = v/v_s^*$. The colorbar shows the e^* specified through Eq. 2.9. For a high enough collision velocity $v - s$. For decreasing e^* the sticking velocity v_s increases as expected.

Chen et al. [49] found that the sticking behaviour of adhesive particles using the JKR model with viscoelastic damping $F \propto \sqrt{a}$ could be described using the damping magnitude α and a parameter A^*, where

$$A^* = \left(\frac{E}{v_0^2}\right)^{-1/3} \left(\frac{\gamma}{v_0^2 r}\right)^{5/6} \ : \ \frac{1}{A^*} \propto v_0 r^{5/6}$$

With α fixed, they found the CoR to increase with increasing $1/A^*$ (and therefore radius) outside of the sticking region $\{(\alpha, A^*) : e = 0\}$. The sticking velocity used in this paper is then determined from that of the smallest particles.

A.3 Simulation parameters

The Youngs' modulus $E = 5$ was specified such that the maximum relative overlap between particles $\tilde{\delta} \leq 10^{-2}$, after compression with an applied pressure $P = 10^{-6}$. The surface energy was then calculated from E.

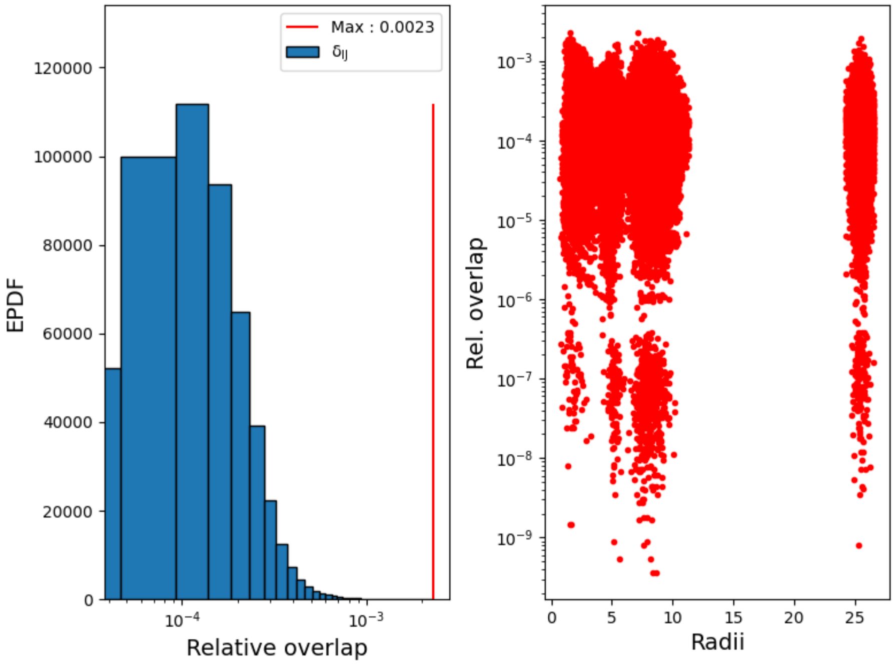


Figure A.3: Left : Histogram of the relative overlaps $\tilde{\delta}$ (blue), with the maximum shown at the red line. Right: A scatter plot of relative overlap vs. particle radii. The maximum $\tilde{\delta} \leq 10^{-2}$ as sought through the Youngs' modulus E.

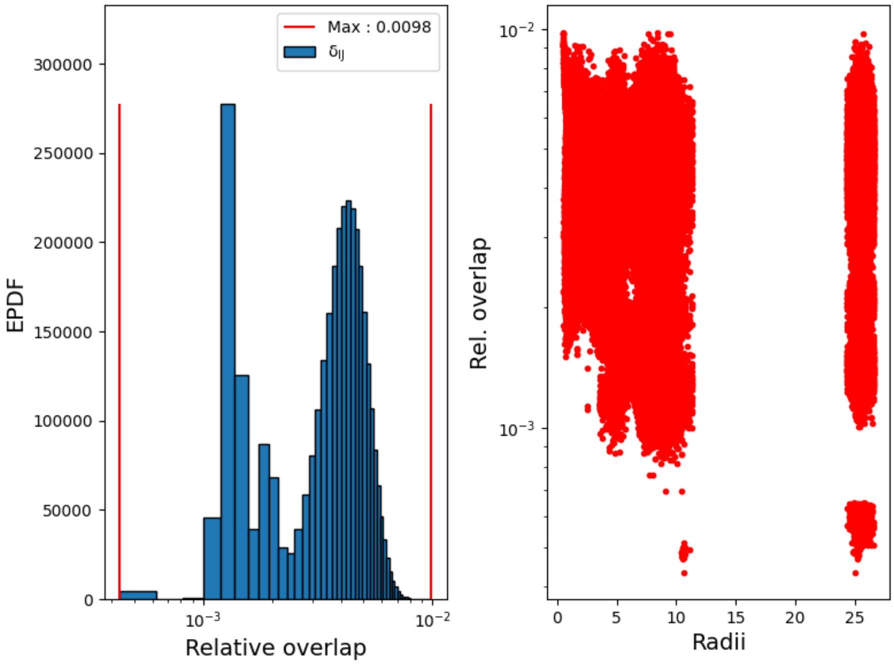


Figure A.4: Relative overlaps after applying the JKR contact model. Left: Histogram of the relative overlaps $\tilde{\delta}$ (blue), with the maximum shown by the red line. Right: Scatter plot of relative overlap vs. particle radii. The maximum $\tilde{\delta} = 0.0098 \approx 10^{-2}$ agrees well with the surface energy criteria.

The timestep was specified as a fraction of the Hertzian collision time Eq. 2.10. The results of varying the timestep fraction is shown below, in Fig. A.6.

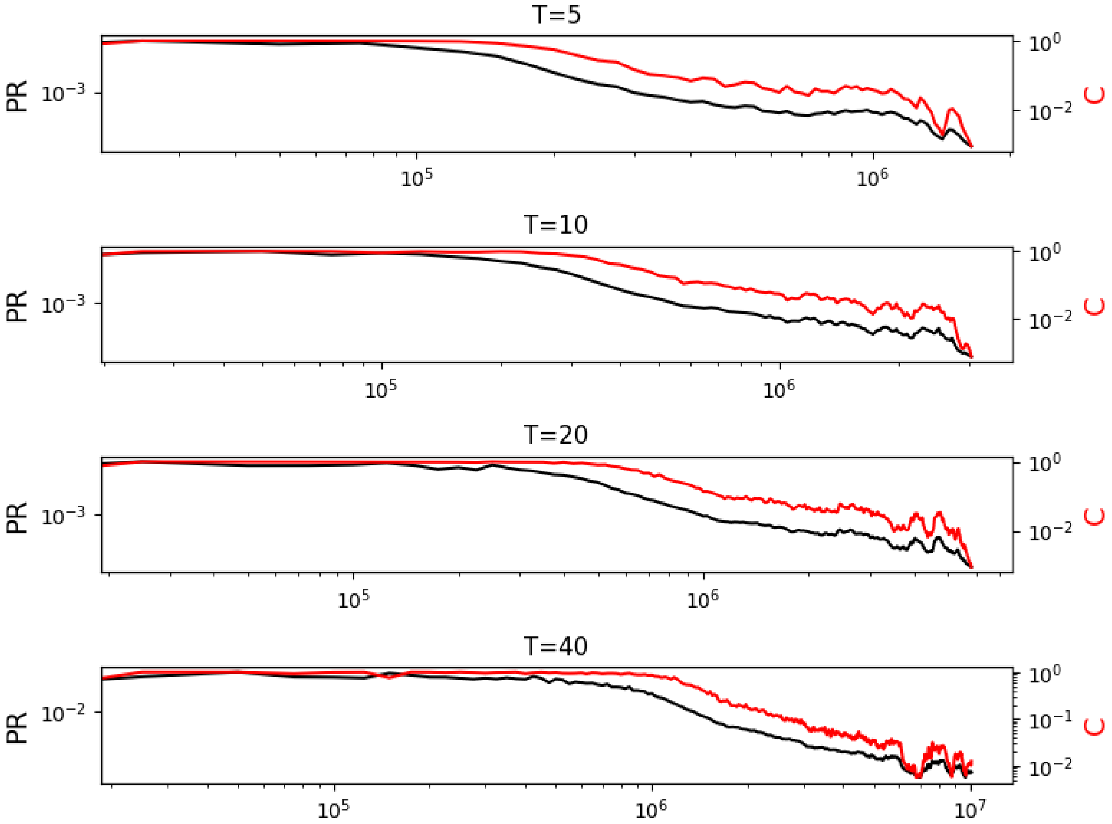


Figure A.5: The figure shows the evolution of the Cundall parameter C and the kinetic pressure ratio PR as the simulation progresses for 10.000 particles with diameters $d \sim \mathcal{LN}(1,1), d \in [1,60]$. The x-axis shows the discrete timesteps at which data is recorded, separated by 25.000 timesteps. T specifies what fraction of the Hertzian timestep is used. The simulations were run until $C < 10^{-3}$ and $PR < 10^{-5}$ or for 10^{7} timesteps.

Using $T = 10$ as the fractional timestep, the effect of varying the CoR e was done on the same packing.

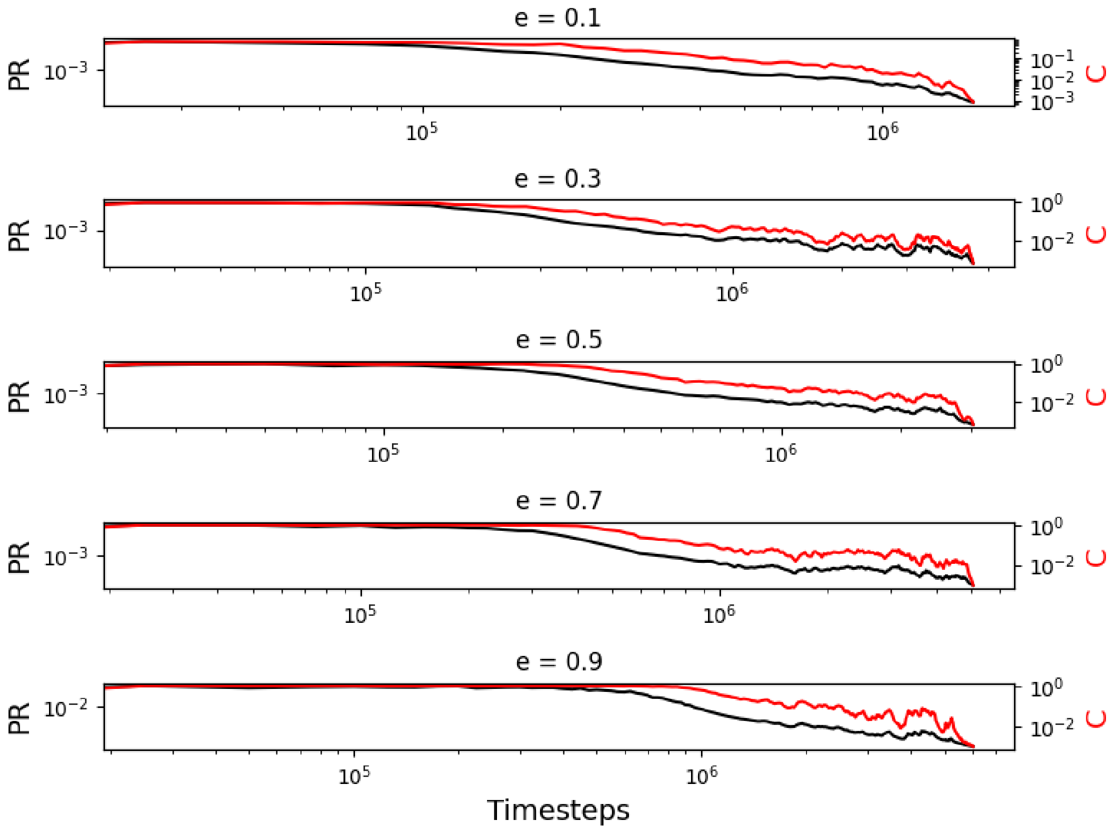


Figure A.6: The figure shows the evolution of the Cundall parameter C and the kinetic pressure ratio PR as the simulation progresses for 10.000 particles with diameters $d \sim \mathcal{LN}(1,1), d \in [1,60]$. The x-axis shows the discrete timesteps at which data is recorded, separated by 25.000 timesteps. By e is meant the CoR determining the damping force.

B

Algorithms

B.1 Non-overlapping initial state

The initial coordinates of the sampled particles are specified by a python implementation of an algorithm influenced by the algorithm of [50]. Given h collection intervals, the cubic box is divided into h-levels. For each collection interval the x, y, z-coordinates are segmented into 500 lists, where particles are binned when placed. By dividing the separate coordinate axes into intervals, instead of creating a grid, the number of lists can be kept manageable, but instead necessitates intersection checks between possible collisions. A collection interval j is specified by the radius of its largest particle R_j. For a particle i with radius r_i belonging to the j:th collection interval, collision checks are run against particles in cells collection interval k in cell level k, within a distance $r_i + R_k,\ k \leq j$. Particles are appended if no collisions are detected.

B.2 Collection interval determination

The *multi* neighbor style of LAMMPS requires that the maximum diameter (so called *collection intervals*) of all levels is supplied. The interpretation of these is equivalent to the hierarchies and their corresponding cell sizes explained in 2.2.2. Based on suggestions in [5, 51], the number of levels and collection intervals is chosen such that the average number of particles per cell for a given level is roughly equal, i.e. given L levels

$$m_l \approx m, \quad m_l = \frac{N_l}{N_l^c}, \quad l = 1, \dots, L \tag{B.1}$$

where N_l is the number of particles assigned to level l, and N_l^c is the number of cells of level l. The diameters are assumed to be sorted. Concretely, the problem then is to choose $L-1$ indeces $\{j_i\}_{i=1}^{L-1} = \mathcal{J}_m \subset [1, N]$ such that Eq. B.1.

As the size of the cells of a particular level is based on the diameter of the largest particle, the number of cells of level l is given by

$$N_l^c = \left(\frac{D_{j_l}}{S}\right)^3, \tag{B.2}$$

where S is the side length of the volume the particles are placed in, excepting level L where $j_{i_L} = N$. Assuming there is only one level, $m_0 = N\left(\frac{D_N}{S}\right)^3$, increasing the number of levels, means that $\sum_{l=1}^{L} m_l \leq m_0$. Given a particle sample of size N with sorted diameters $\{D_i\}_{i=1}^N$, m_0 is calculated, after which a stepsize $\Delta m = m_0/10000$ is introduced. For a given $m_k = k\delta m$, the level assignment is done by incrementally assigning particle i in the sorted list to level l, until

$$\frac{N_l(i)}{N_l^c(i)} > m_k$$

upon which $l := l + 1$. For a given m, the number of levels and their size is then determined. As m_k is incrementally increased the number of levels decrease, such that $\mathcal{J}_{m_k}$ can be reused as the starting point for m_{k+1}, to speed up this process. Many m yield the same level l, the smallest m is then chosen such that the number of particles at this level is the maximum possible. This then supplies a way to get $m(\mathrm{L}, \mathrm{PSD})$, the number of particles per cell, given a number of levels and particular PSD. In [5] they find that the computation time for contact detection T_{CD} scales as

$$T_{CD} \sim NL(m + K)$$

where $K = 0.3$ is an overhead parameter, that they find to be insensitive to the packing fraction ϕ [50]. Based on this the number of levels and collection intervals for a particular sample of a PSD, is then the level at which T_{CD} is minimal.

B.3 Mass fraction balancing

Given desired mass fractions and current mass fractions τ_k for each component $\tilde{\tau}_k$, $k = 1, .., K$, so that

$$\tau_k = \rho_k V_k, \tag{B.3}$$

where ρ_k, V_k is the density and total volume for particles of component k. Particles are sampled from the component which has the highest relative discrepancy from the target mass fraction as identified by

$$k := \arg\max_k \frac{\tilde{\tau}_k - \tau_k}{\tilde{\tau}_k}.$$

One particle diameter is drawn so that $d \sim \mathcal{LN}(\mu_k, \sigma_k)$, where μ_k, σ_k are the distribution parameters. The particle volume is added to the volume of the component so that

$$V_k := V_k + \frac{\pi d^3}{6}$$

Subsequently the current mass fractions and the component with highest relative discrepancy are re-evaluated using Eq. B.3, and the above scheme is reiterated until it satisfies the condition

$$\max_k \frac{\tilde{\tau}_k - \tau_k}{\tilde{\tau}_k} < \epsilon,$$

where ϵ is some relative error tolerance.

C

Miscellaneous results

C.1 Content uniformity

Here the results of the concentric shell algorithm for two different packings are presented, using a radial shell thickness of 0.1, and 1000 randomly sampled points. The following figures show the comparison between the analytical estimates of the COV and that calculated using concentric shells.

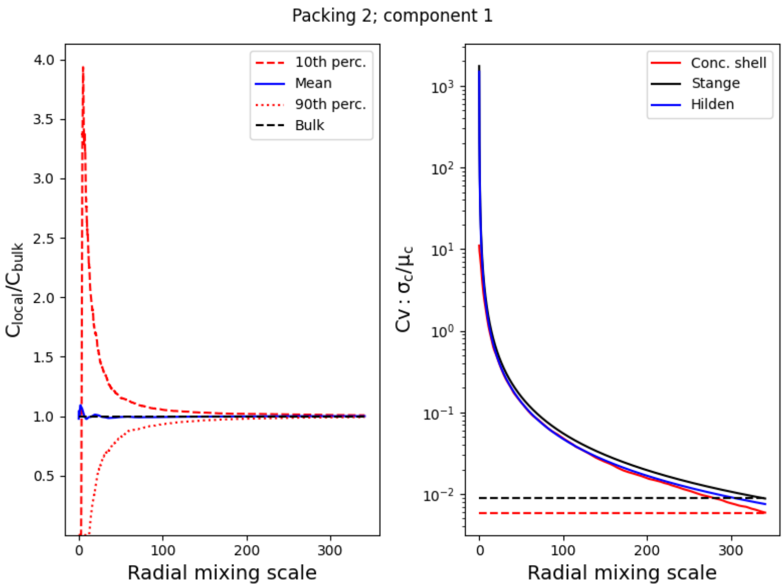


Figure C.1: Left: Bulk (black), mean (blue) and percentiles (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maximum radial mixing scale for Stange and the concentric shell method.

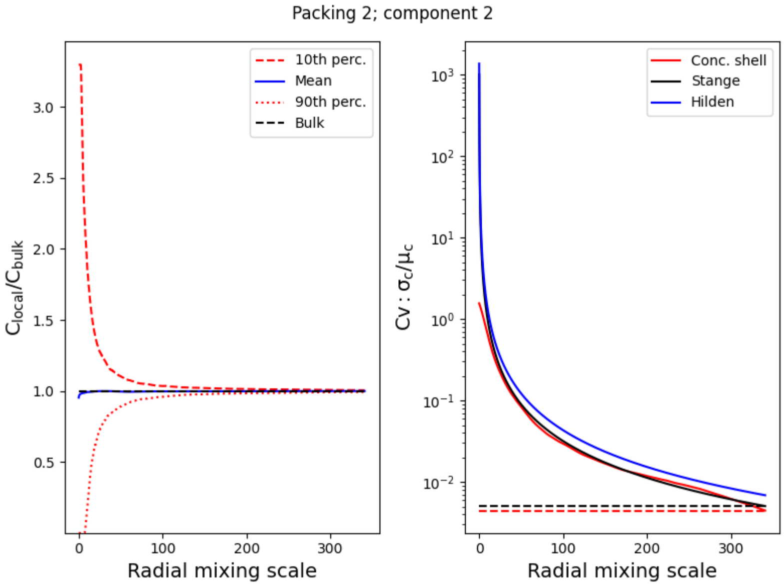
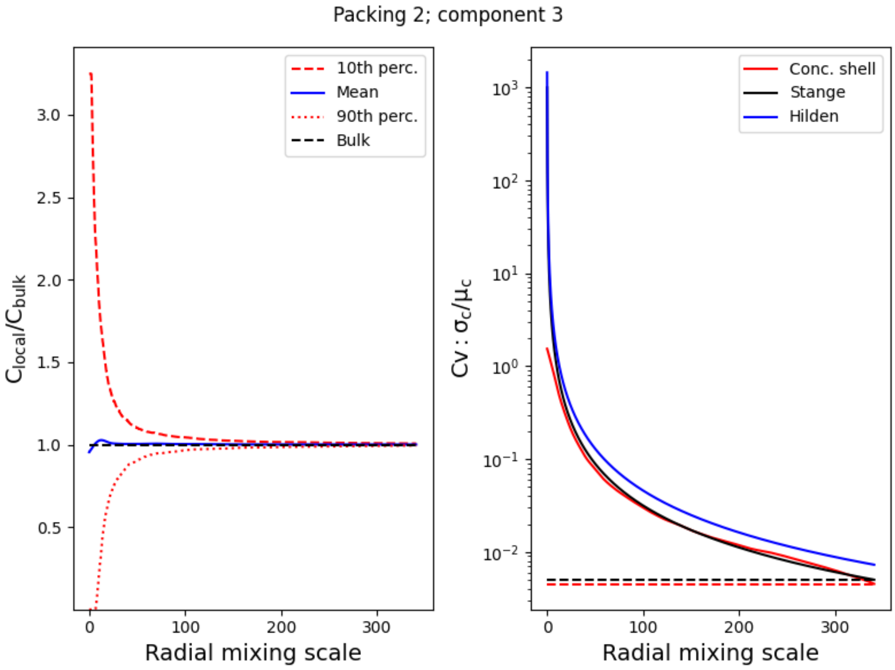


Figure C.2: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.



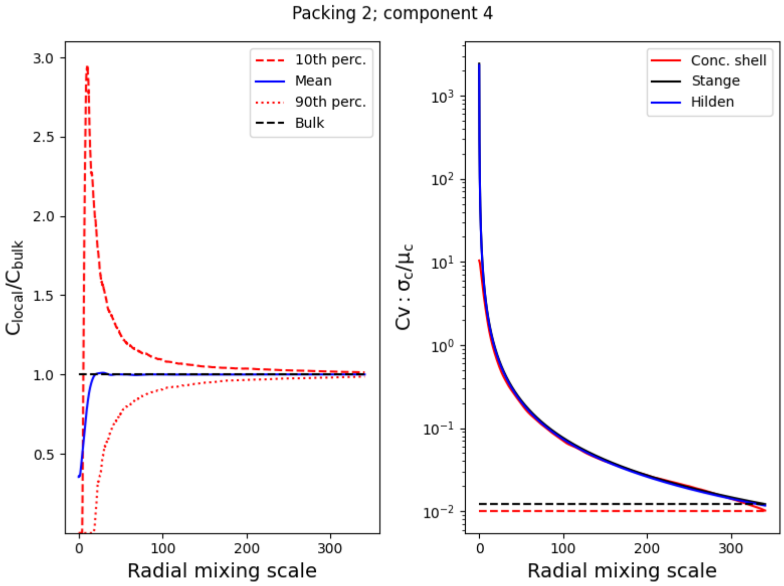


Figure C.3: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

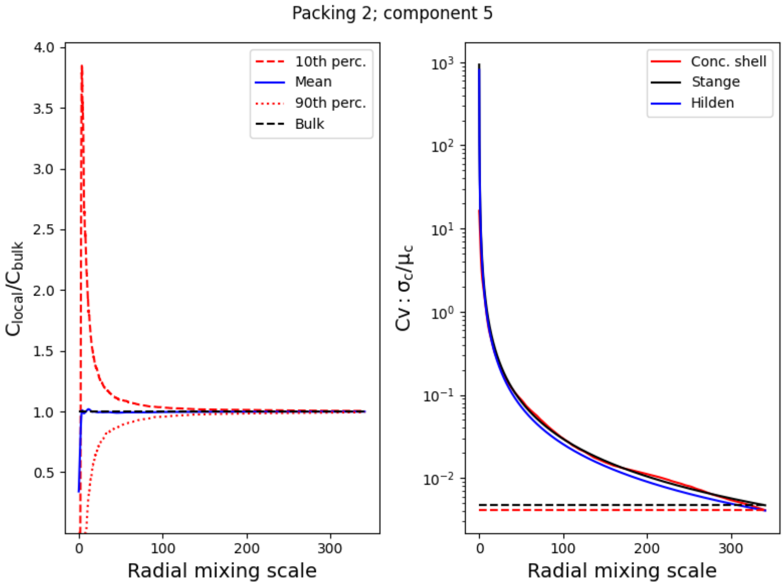


Figure C.4: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

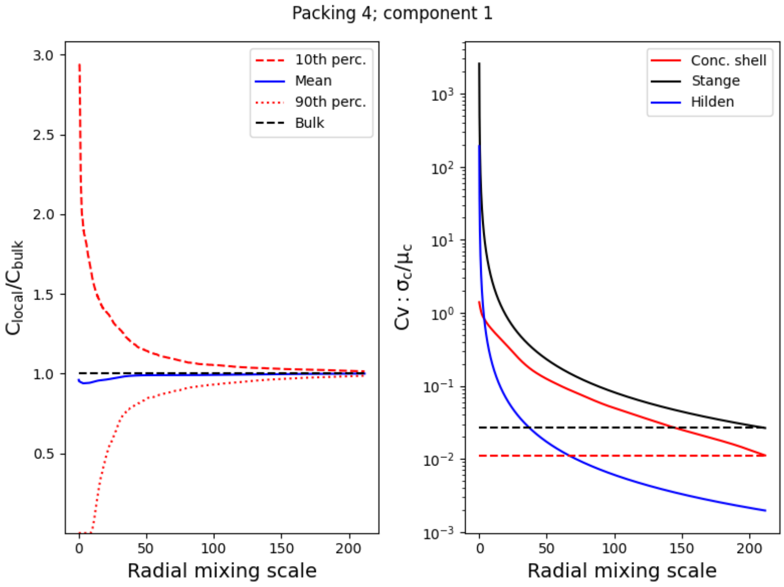


Figure C.5: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

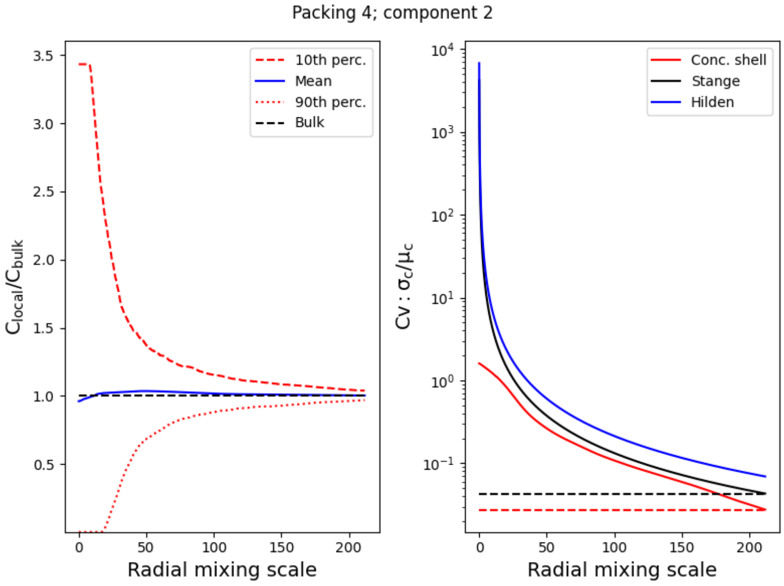


Figure C.6: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

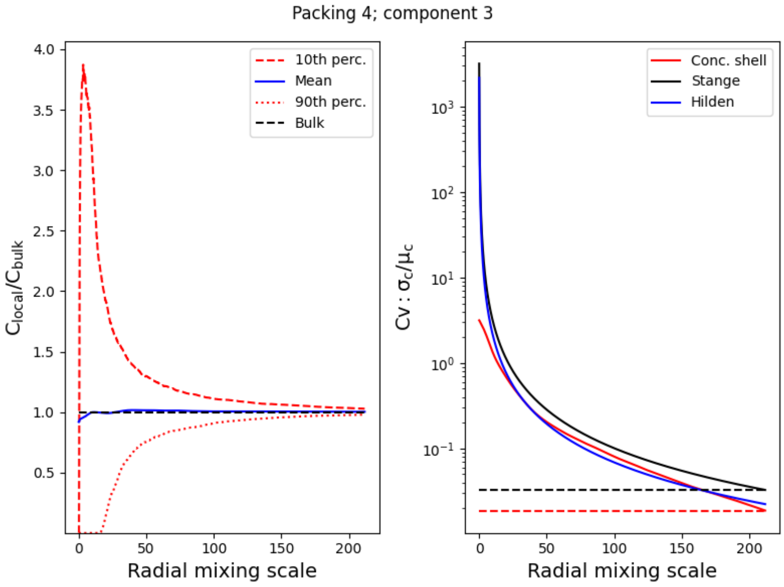


Figure C.7: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

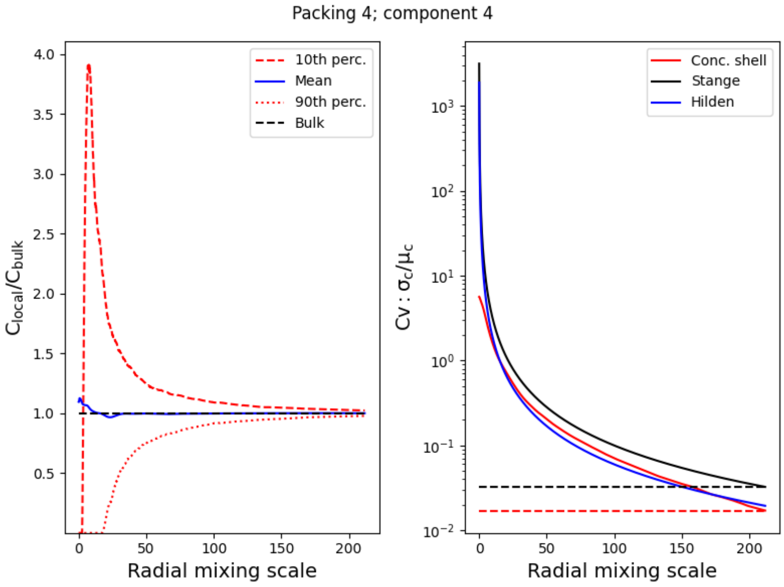


Figure C.8: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

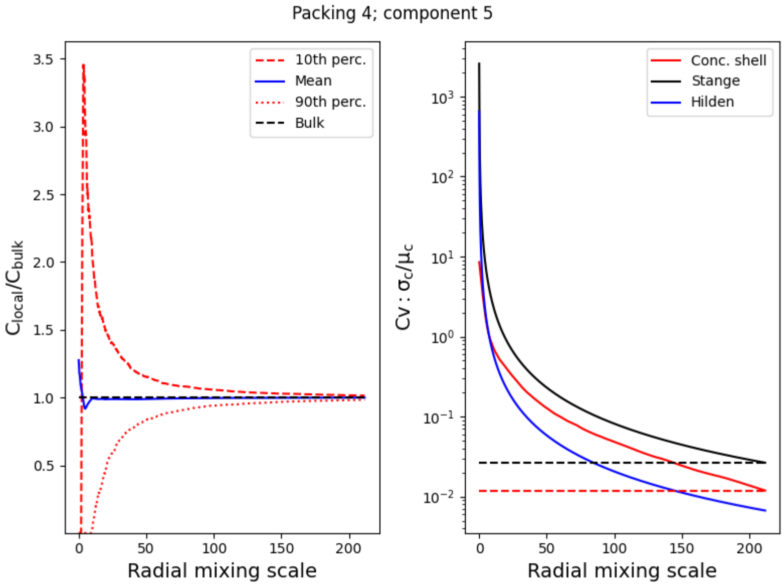


Figure C.9: Left: Bulk (black), mean (blue) and percentiles (10th, 90th) (red) of the component concentrations. Right: The COV calculated using the concentric shell algorithm (red), the Stange estimate, Eq. 2.26 and that of Hilden, Eq. 2.29. The dashed lines show the COV at the maxium radial mixing scale for Stange and the concentric shell method.

C.2 LMPM comparison

The standard deviation of the packing fraction is here compared between simulations s_ϕ^{sim} and the LMP model. The LMPM standard deviation is fit to the simulation data using the parameter d^*, detailed in Sec. 4.2

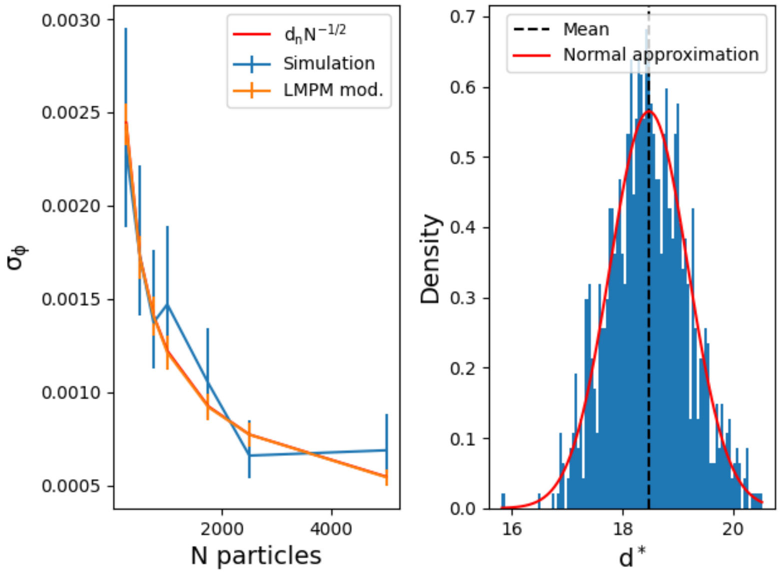


Figure C.10: A bidisperse lognormal mixture with mean diameters $\mu = [5, 10]$ and variance $\sigma^2 = [1, 1]$. Left : Comparison between s_ϕ^{Sim}, including Chi-square confidence interval, and the mean with standard deviation of $d^* s_\phi^{\text{LMPM}}$, for 1000 realizations, as well as a fit of $\frac{d_n}{\sqrt{N}}$. Right: Density histogram of d^* over 1000 realizations, with a normal approximation overlayed $\mathcal{N}(18.480.71)$, $C_v \approx 3.8\%$.

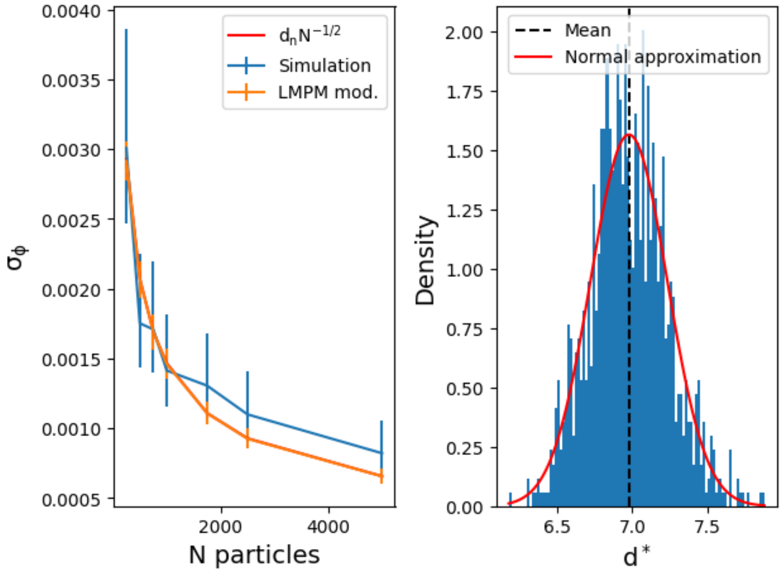


Figure C.11: A bidisperse lognormal mixture with mean diameters $\mu = [5, 15]$ and variance $\sigma^2 = [2, 2]$. Left : Comparison between s_ϕ^{Sim}, including Chi-square confidence interval, and the mean with standard deviation of $d^* s_\phi^{\text{LMPM}}$, for 1000 realizations, as well as a fit of $\frac{d_n}{\sqrt{N}}$. Right: Density histogram of d^* over 1000 realizations, with a normal approximation overlayed $\mathcal{N}(6.980.25)$, $C_v \approx 3.7\%$.

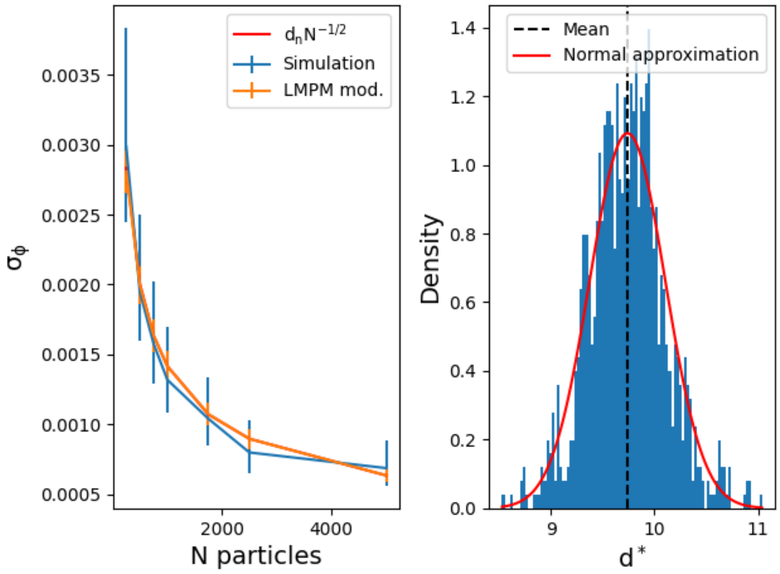


Figure C.12: A bidisperse lognormal mixture with mean diameters $\mu = [5, 10]$ and variance $\sigma^2 = [2, 2]$. Left : Comparison between s_ϕ^{Sim}, including Chi-square confidence interval, and the mean with standard deviation of $d^* s_\phi^{\mathrm{LMPM}}$, for 1000 realizations. Right: Density histogram of d^* over 1000 realizations, with a normal approximation overlayed $\mathcal{N}(9.74, 0.37)$, $C_v \approx 3.8\%$.

C.3 Neighbor list benchmark

Fig. C.13 shows runtimes normalized by shortest time between the standard bin method as shown in Fig. 2.2 and the multilevel method as shown in Fig. 2.3. The figure suggests that relative gains increase for higher polydispersities λ.

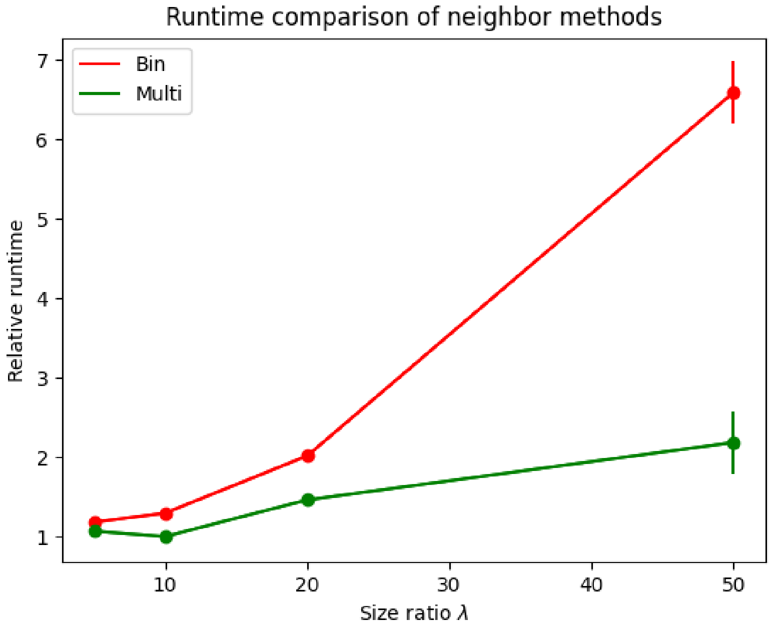


Figure C.13: Runtimes for 5 different packings containing 10000 particles with diameters $d \sim \mathcal{LN}(1,1),\ d \in [1, \lambda]$ at each size ratio $\lambda \in [5, 10, 20, 50]$ normalized by shortest runtime, with error bars representing one standard error.

D

LMPM interaction functions

The filling and occupation interaction functions of the LMPM model are specified by

$$\begin{aligned} a(r_{ij}) &= (1-r_{ij})^{A_1} + A_2 r_{ij}(1-r_{ij})^{A_3} \\ b(r_{ij}) &= (1-r_{ij})^{B_1} + B_2 r_{ij}(1-r_{ij})^{B_3}, \end{aligned}$$

where the constants are

$$\begin{aligned} A_1 &= 3.3, \quad A_2 = 2.8 \quad A_3 = 2.7 \\ B_1 &= 2.0, \quad B_2 = 0.4, \quad B_3 = 3.7. \end{aligned}$$

The interaction functions $c(r_{ij}, n_0)$ and $d(r_{ij}, n_0)$ are defined as

$$c(r_{ij}, n_0) = \begin{cases} 10^{C_1 n_0} C_2 (C_3 + C_4 r_{ij} + C_5 r_{ij}^2 + C_6 r_{ij}^3 + C_7 r_{ij}^4), \quad r_{ij} \leq 0.741 \\ 0, \quad r_{ij} > 0.741 \end{cases},$$

with the constants

$$\begin{aligned} &C_1 = 10.288,\ C_2 = -1.4566,\ C_3 = -1.0002,\ C_4 = 0.1126 \\ &C_5 = 5.8455,\ C_6 = -7.9488,\ C_7 = 3.1222 \end{aligned}$$

and

$$d(r_{ij}, n_0) = \begin{cases} 0, \quad r_{ij} > 0.741 \\ (D_1 + D_2 n_0 + D_3 n_0^2 + D_4 n_0^3 + D_5 n_0^4)(D_6 + D_7 r_{ij} + D_8 r_{ij}^2 + D_9 r_{ij}^3 + D_{10} r_{ij}^4) \end{cases},$$

with the constants

$$\begin{aligned} &D_1 = -1.3092,\ D_2 = 15.039,\ D_3 = -37.453,\ D_4 = 40.869,\ D_5 = -17.110 \\ &D_6 = -1.0029,\ D_7 = 0.3589,\ D_8 = 10.970,\ D_9 = 22.197,\ D_{10} = 12.434 \end{aligned}$$

E

Packing tables

Nominal pharmaceutical drug product compositions drawn from [52]:

Table E.1: Packing 0: Multi-component PSD parameters, mass fractions and densities for packing used in simulation results in Sec. 3.1.

Type	Component ID	Density (kg/m^3)	Mass fraction (%)	Mean (µm)	Standard deviation
API	1	1291	1	16	4.22
Plastic filler	2	1562	47	214	2.82
Elastic filler	3	1539	47	68	5.65
Disintegrant	4	1585	4	41	2.83
Lubricant	5	1092	1	12	2.74

Table E.2: Packing 1: Multi-component PSD parameters, mass fractions and densities

Type	Component ID	Density (kg/m^3)	Mass fraction (%)	Mean (µm)	Standard deviation
API	1	1291	1	16	4.22
Plastic filler	2	1563	47	106	3.15
Elastic filler	3	1539	47	68	5.65
Disintegrant	4	1585	4	41	2.83
Lubricant	5	1092	1	12	2.74

Table E.3: Packing 2: Multi-component PSD parameters, mass fractions and densities

Type	Component ID	Density (kg/m^3)	Mass fraction (%)	Mean (µm)	Standard deviation
API	1	1291	1	16	4.22
Plastic filler	2	1562	47	66	3.81
Elastic filler	3	1539	47	68	5.65
Disintegrant	4	1585	4	41	2.83
Lubricant	5	1092	1	12	2.74

Table E.4: Packing 3: Multi-component PSD parameters, mass fractions and densities.

Type	Component ID	Density (kg/m^3)	Mass fraction (%)	Mean (µm)	Standard deviation
API	1	1291	1	16	4.22
Plastic filler	2	1562	70.5	214	2.82
Elastic filler	3	1539	23.5	68	5.65
Disintegrant	4	1585	4	41	2.83
Lubricant	5	1092	1	12	2.74

Table E.5: Packing 4: Multi-component PSD parameters, mass fractions and densities.

Type	Component ID	Density (kg/m^3)	Mass fraction (%)	Mean (µm)	Standard deviation
API	1	1291	40	16	4.22
Plastic filler	2	1562	41.25	214	2.82
Elastic filler	3	1539	13.7	68	5.65
Disintegrant	4	1585	4	41	2.83
Lubricant	5	1092	1	12	2.74

Table E.6: Number of particles and their polydispersity in packings 0,1,2,3 and 4 with inputs from Table E.1, E.2, E.3, E.4 and E.5.

Comp. ID / Packing ID	1	2	3	4	5	Total count	Poly-dispersity
0	125362	2500	77108	29291	370101	604362	53
1	61565	10000	37758	14326	180407	304056	29.2
2	63875	43044	39412	15000	188594	349925	23.7
3	100968	3000	30925	23371	295403	453667	52.7
4	1140852	500	5108	6646	84309	1237415	49

DEPARTMENT OF SOME SUBJECT OR TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY