

# Paths in a Network of Polydisperse Spherical Droplets

Johannes Josef Schneider<sup>1,2</sup>, Alessia Faggian<sup>3</sup>, Silvia Holler<sup>3</sup>, Federica Casiraghi<sup>3</sup>, Jin Li<sup>2</sup>, Lorena Cebolla Sanahuja<sup>3</sup>, Hans-Georg Matuttis<sup>4</sup>, Martin Michael Hanczyc<sup>3,5</sup>, David Anthony Barrow<sup>2</sup>, Mathias Sebastian Weyland<sup>1</sup>, Dandolo Flumini<sup>1</sup>, Peter Eggenberger Hotz<sup>1</sup>, Patrik Eschle<sup>1</sup>, Aitor Patiño Diaz<sup>3</sup>, and Rudolf Marcel Fuchsli<sup>1,6</sup>

<sup>1</sup>Institute of Applied Mathematics and Physics, School of Engineering, Zurich University of Applied Sciences, Technikumstr. 9, 8401 Winterthur, Switzerland

<sup>2</sup>School of Engineering, Cardiff University, Queen's Buildings, 14-17 The Parade, Cardiff CF24 3AA, Wales, United Kingdom

<sup>3</sup>Laboratory for Artificial Biology, Department of Cellular, Computational and Integrative Biology, University of Trento, Via Sommarive, 9 - 38123 Povo, Italy

<sup>4</sup>Department of Mechanical Engineering and Intelligent Systems, The University of Electrocommunications, Chofu Chofugaoka 1-5-1, Tokyo 182-8585, Japan

<sup>5</sup>Chemical and Biological Engineering, University of New Mexico, MSC01 1120, Albuquerque, NM, 87131-0001, USA

<sup>6</sup>European Centre for Living Technology, S.Marco 2940, 30124 Venice, Italy  
scnj@zhaw.ch, johannesjosefschneider@googlemail.com

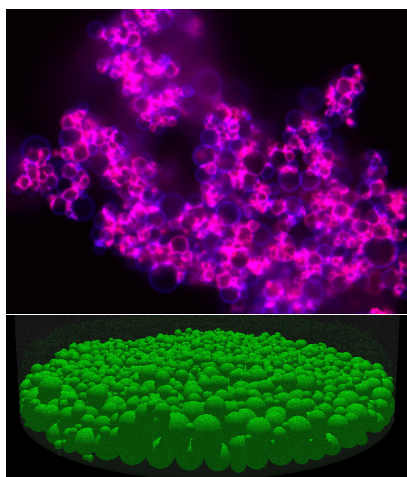


Figure 1: Top: Snapshot of an experimentally found cluster of oil droplets in water recorded with a confocal microscope. Bottom: Final configuration of 2,000 multidisperse particles at the bottom of a cylinder in a computer simulation.

## Abstract

We simulate the movement and agglomeration of oil droplets in water under constraints, like confinement, using a simplified stochastic-hydrodynamic model. In the analysis of the network created by the droplets in the agglomeration, we focus on the paths between pairs of droplets and compare the computational results for various system sizes.

## Introduction

Spatial arrangements of hard spheres are widely studied in physics, as these systems serve as simple models for granu-

lar matter, colloidal systems, and molecular crystals (Reiss et al., 1996; Russel et al., 1989; Mitarai and Nakanishi, 2003; Metcalfe et al., 1995; Zallen, 1998). Mostly, monodisperse and bidisperse systems are considered, i.e., all spheres exhibit the same radius value or one of two different values, but sometimes also multidisperse systems with radii chosen from a finite set of prespecified radius values have been investigated (Müller et al., 2009; Schneider et al., 2009). Also packings of non-spherical items for which the determination of overlaps is more difficult, such as ellipsoids (Pfleiderer and Schilling, 2007; Matuttis and Chen, 2014) and spherocylinders, have been studied. It has, e.g., been found that a random packing of ellipsoids with a specific aspect ratio (M & M candies) is denser than a random packing of spheres (Donev et al., 2004). Furthermore, arrangements of particles with long-range interactions, in confinement, and under constraints, such as shear forces and repulsive walls, (Ricci et al., 2007; Ochoa et al., 2006) have been investigated.

In our collaboration, we intend to develop a probabilistic compiler (Flumini et al., 2020; Weyland et al., 2020) to aid the three-dimensional agglomeration of droplets filled with various chemicals in a specific way in order to e.g. allow the creation of desired macromolecules via a successive reaction scheme (Schneider et al., 2020a,b). Neighboring droplets can form connections, either by forming bilayers (Li and Barrow, 2017) or by getting glued to each other by matching pairs of single-stranded DNA (Hadorn et al., 2012). Chemicals contained within the droplets can move to neighboring droplets either directly, as hydrophobic compounds can be exchanged between adjacent oil droplets at the contact face, or, if the oil droplets are contained in a hull comprised of amphiphilic molecules like phospholipids, through pores within these bilayers. Thus, a complex bi-

layer network is created, with the droplets being the nodes of this graph and the existing connections being the edges between the corresponding droplets. In such bilayer networks, a controlled successive reaction scheme can be effectuated to produce the intended macromolecules. Within the scope of this paper, we present computational results for basic simulations of a simplified agglomeration process of a polydisperse system of droplets, mimicking experiments. Here we want to focus on the question which influence the density of the droplets has on the agglomeration process of the particles and on some specific properties of the networks created, which are of crucial importance for the gradual reaction scheme intended. In order to focus on these questions and to exclude effects from other experimental properties, we simulate the droplets as hard spheres and ignore details of the surface structure of the particles, attractive forces as well as adhesion effects. As the extension of the bilayers is very small and as due to their small radii (Aprin et al., 2015), the droplets keep their spherical shape during the experiments, as shown in Fig. 1, this simplified approach is justified.

Networks can be in general analyzed

- either on an elementary level, i.e., by considering properties of the single nodes within a network, like the degree of a node (the number of nodes a node is attached to via edges),
- or by considering groups of nodes, e.g., by finding cliques of nodes which are fully connected with each other within such a group (Marino and Kirkpatrick, 2018),
- or by taking the overall network into account, e.g., by determining whether the network is dominated by a large cluster and even percolating (Stauffer and Aharony, 1994; Stauffer, 1986; Naftaly et al., 1991),
- or by taking a local-global attitude, e.g., by investigating the role some specific nodes play for the overall network (Schneider and Kirkpatrick, 2005).

Often random networks are considered in which nodes are connected with randomly selected edges (Bollobás, 2011). But in real-world networks, like genetic networks or the World Wide Web, the degrees of the nodes often follow a scale-free power-law distribution (Barabási and Albert, 1999) due to the tendency that a newly added node preferentially attaches itself to nodes with higher degree. In contrast, Gaussian distributions are found for degree numbers in random networks. We are mainly interested in the time evolution of the network formed by the connections between the various droplets. In this paper, we focus on the paths between pairs of droplets, i.e., on the question whether paths between them exist and, if yes, how large the geodesic distances between them are. In some applications of artificial chemistry we intend to perform on such a network

of droplets, the maximum geodesic distance determines the maximum number of steps in the gradual chemical reaction scheme.

## Simulation Details

We consider polydisperse systems of oil droplets in water, which are modeled as hard spheres. The radii of the particles are randomly chosen from a uniform distribution in the range of  $10 - 50 \mu\text{m}$ . Initially, they are randomly placed in a cylinder of height 4 mm and radius 1 mm without overlaps among particles or between particles and walls. The particles are initialized with zero velocity. As we are interested in the effect of particle density on the agglomeration process and on the properties of the evolving networks, we perform 100 simulations each for various numbers  $N$  of particles, with  $N = 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2500, 3000, 3500, 4000, 4500, \text{ and } 5000$ . The presented results are averages of these 100 simulation runs.

The simulation is divided in time steps of  $\delta t = 10^{-5} \text{s}$ . In each time step, the particles are subjected to various forces: All particles are filled with oil of the density  $\rho = 1.23 \text{kg/l}$ , so that they sink in water due to gravity reduced by the buoyant force. Furthermore, the three spatial components of the velocity vectors  $\vec{v}_i$  are subjected to random velocity changes. Each velocity component  $v_i^{x,y,z}$  is changed independently by a uniform random variable chosen from the interval  $[-0.05|v_i^{x,y,z}|, 0.05|v_i^{x,y,z}|]$ . The particles are also subjected to the Stokes friction force

$$\vec{F}_{i,S} = -6\pi\eta r_i \vec{v}_i, \quad (1)$$

with the radius  $r_i$  of the particle, the current velocity  $\vec{v}_i$ , and the viscosity  $\eta = 0.891 \text{mPas}$  of water at  $25^\circ\text{C}$ . The concept of added mass is used (Stokes, 1851). This virtual mass is the inertia added to the mass of the particle, because an acceleration or deceleration of a body in water must move or deflect some volume of surrounding fluid when it moves through it. For a spherical particle with radius  $r_i$  far away from other boundaries, the added mass is given by  $\frac{2}{3}\pi r_i^3 \rho_{\text{fluid}}$ , i.e., it is half of the mass of the fluid displaced by the particle.

After the new velocities of the particles are determined in this way, their positions are updated according to

$$\vec{x}_i(t) = \vec{x}_i(t - \delta t) + \vec{v}_i(t) \times \delta t. \quad (2)$$

Due to the stochastic nature of the random velocity changes, only this Euler scheme is suitable for the determination of velocities and positions of the resulting stochastic differential equation of motion (Kloeden and Platen, 2013).

After the determination of these new positions, some conditions are enforced: First, it is checked whether a particle

collides or even overlaps with a boundary of the cylinder at its new position and whether the overlap would increase if the velocity vector of the particle remains unchanged. In this case, the collision normal is determined and the velocity vector of the particle is updated according to the standard collision rules with an elasticity factor of 0.9. If there is an overlap, the position of the particle is updated in order to resolve the overlap. Analogously, then checks for collisions and overlaps between pairs of particles are performed and their velocity vectors and positions are updated accordingly. The overlaps have indeed to be eliminated as otherwise they partially remain and can even increase over time, especially in the regime of slow velocities.

In total,  $10^7$  time steps of duration  $\delta t = 10^{-5}$ s are performed during a simulation run. A simulation run thus covers in total a time span of  $\mathcal{T} = 100$ s which is sufficient for finishing the agglomeration process within the dimensions of the cylindrical container, as the smallest particles have a radius of at least  $r_{\text{smallest}} = 10\mu\text{m}$  in our simulations. Shorter  $\delta t$  and longer  $\mathcal{T}$  would be needed for smaller  $r_{\text{smallest}}$ , as  $\delta t$  scales with  $r_{\text{smallest}}^2$  and  $\mathcal{T}$  scales with  $r_{\text{smallest}}^{-2}$  due to the Stokes friction force, because of which the sink velocity is smaller for smaller spherical particles. Thus, reducing  $r_{\text{smallest}}$  by e.g. a factor of 10 would result in an increase of the number of necessary time steps and thus of the computing time by a factor of  $10^4$ . The computing time for one simulation also depends on the number  $N$  of particles in a quadratic way, i.e.,  $\mathcal{T} \sim N^2$ . For example, it took roughly 12 hours on a standard laptop for  $N = 2000$  particles. The bottom part of Fig. 1 exhibits a final configuration of a simulation containing 2000 particles.

At the end of each 1000th time step, a configuration is recorded for the network analysis, such that we get a set of  $10^4$  stored configurations and thus of networks from each simulation run, which are equally spaced in time with a time delay of  $\Delta t = 10^{-2}$ s between each pair of successive configurations.

## Network Analysis

### General Remarks

As a first step in the network analysis, we need to create a network from a spatial configuration of droplets. For answering the question whether an edge exists between a pair  $(i, j)$  of droplets, we need to determine the distance  $D(i, j)$  between their midpoints. Let  $(x_i, y_i, z_i)$  be the triple of midpoint coordinates of droplet  $i$ , then the Euclidean spatial distance between two droplets  $i$  and  $j$  is given by

$$D(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}. \quad (3)$$

Two droplets  $i$  and  $j$  overlap if  $D(i, j)$  is smaller than the sum of their respective radii  $r_i$  and  $r_j$ . Assuming that an edge between a pair of droplets exists if there is a not yet completely resolved overlap between them or if they exactly

touch each other or if they get very close to each other, we can define an  $N \times N$  adjacency matrix  $\eta$  with

$$\eta(i, j) = \begin{cases} 1 & \text{if } D(i, j) - (r_i + r_j) < 0.1\mu\text{m} \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

This symmetric matrix  $\eta$  contains all information about the network formed by the droplets. For example, the number  $e$  of edges is given as

$$e = \sum_{i < j} \eta(i, j). \quad (5)$$

As  $\eta$  turns out to be extremely sparse, we create neighborhood lists  $\mathcal{N}(i)$  for each node  $i$ , containing all nodes to which node  $i$  is connected by an edge:

$$\mathcal{N}(i) = \{j | \eta(i, j) = 1\} \quad (6)$$

The number of elements of  $\mathcal{N}(i)$  is called the degree  $k(i)$  of node  $i$ . Using these neighborhood lists, the computation time for the determination of some network properties can be significantly reduced compared to the use of the adjacency matrix alone.

We performed 100 simulations each for various numbers  $N$  of particles and recorded  $10^4$  configurations from each simulation. The results shown below are ensemble averages over these 100 simulation runs or values derived from these averages. Thus, each curve in the left pictures of Figs. 2, 3, 4, 5, 6, and 8 is comprised of  $10^4$  data points, which are ensemble averages of 100 simulation runs, so that  $10^6$  networks had to be evaluated for each curve.

### Geodesic Distances

As mentioned above, we intend to perform a gradual reaction scheme on this network of droplets, i.e., neighboring droplets should exchange chemicals either directly at the contact face or through pores in the bilayers, the resulting intermediary reaction products serve as educts for a next step with the next droplet connected. Thus, we are interested in the question whether such reaction paths exist and how long they are, i.e., how many steps we can perform gradually. Thus, in this paper, we focus on the investigation of distances between nodes. Hereby we are not interested in the Euclidean distance  $D(i, j)$  between the midpoints of the droplets, as given in Eq. (3), or the minimum distance of points within a pair of droplets, as given by  $D(i, j) - (r_i + r_j)$ , but in the minimum number of edges one has to cross when trying to get from node  $i$  to node  $j$  on a path comprised by the edges of the network. For example, this distance  $d(i, j)$  between pairs  $(i, j)$  of nodes, which is also called geodesic distance (González and Cascone, 2014), equals 1 if  $\eta(i, j) = 1$  and  $d(i, j) = 2$  if  $\eta(i, j) = 0$  but  $\eta(i, n) = \eta(n, j) = 1$  for at least one node  $n \neq i, j$ . If no path from some node  $i$  to another node  $j$  can be found, one sets  $d(i, j) = \infty$ . As  $\eta$  is very sparse, we use the Dijkstra

algorithm to determine these geodesic distances (Dijkstra, 1959). Please note that the matrix  $d$  is symmetric, as we do not apply preferential directions to our edges.

## Computational Results

### Growth of the Network

**Mean degrees** Before we investigate the existence of paths and the lengths of geodesic distances, we have a look at some observables describing the time evolution of the size of the network. We start off our investigation by considering the average degree  $\langle k \rangle$  of the particles, which is usually seen as one of the most important observables when performing a local network analysis. But it is also related to the overall number  $e$  of edges by the equation

$$N \times \langle k \rangle = 2e. \quad (7)$$

Figure 2 shows for some selected numbers  $N$  of droplets that the average  $\langle k \rangle$  of the degrees sigmoidally increases in time. As the inset reveals,  $\langle k \rangle$  indeed increases in a double-sigmoidal way. There is a first sigmoidal increase at short time scales, in which first small groups, especially pairs of droplets, are formed with some small probability. However, we are mainly interested in the subsequent second sigmoidal increase, as this main increase is the one in which the network formation takes place. This second sigmoidal increase is the more pronounced the larger the number  $N$  of droplets is. We denote the final value of the degree of node  $i$  as  $k_f(i)$  and its average over all nodes as  $\langle k_f \rangle$ . We have a closer look at the increase of the final mean values  $\langle k_f \rangle$  of the degrees with increasing  $N$  in the right picture in Fig. 2. While  $\langle k_f \rangle$  can be nicely fitted linearly for very small  $N$  and quadratically for small  $N$ , we find criticalities for larger  $N$ , such that we get the overall behavior

$$\langle k_f \rangle = \begin{cases} c_1 N & \text{for } N \leq 40 \\ c_1 N + c_2 N^2 & \text{for } N \leq 80 \\ c_3 \frac{N^\alpha}{(N_{\text{crit},1} - N)^\gamma} & \text{for } 40 \leq N \leq 900 \\ c_4 \tanh\left(c_5 (N - N_{\text{crit},2})^\beta\right) & \text{for } 850 \leq N \end{cases} \quad (8)$$

with the prefactors  $c_1, \dots, c_5$ , the critical exponents  $\alpha, \beta$ , and  $\gamma$ , and the critical numbers  $N_{\text{crit},1}$  and  $N_{\text{crit},2}$  of particles. Various fits of the functions in Eq. 8 with similar fitting qualities result in  $\alpha = 1 \dots 1.1$ ,  $\gamma = 0.4 \dots 0.5$ ,  $\beta = 0.1 \dots 0.18$ ,  $N_{\text{crit},1} = 940 \dots 1000$ , and  $N_{\text{crit},2} = 780 \dots 845$ . Please note that the prefactor  $c_4 = 6 \dots 7$  provides an estimate for  $\langle k_f \rangle$  in the limit  $N \rightarrow \infty$ . While the values for prefactors and critical numbers of particles will change if altering the simulation parameters, the theory of critical phenomena decrees that the critical exponents  $\alpha, \beta$ , and  $\gamma$  and the form of the functions stay identical.

**Cluster Numbers and Cluster Sizes** Next, we would like to have a look at a prominent example of global network

analysis, studying cluster numbers and cluster sizes. A cluster is defined as a subset of nodes in which a path exists between any pair of nodes in this subset. Please note that we count clusters consisting of one node only also as clusters. Figure 3 shows the sigmoidal decrease of the number  $N_c$  of clusters, normalized by the number  $N$  of droplets in order to better compare the results for various  $N$ . Each curve starts at a value of 1, as each droplet forms a cluster of its own at the beginning. Generally, we find sigmoidal decreases of  $N_c$  in time. With increasing  $N$ , the increase first becomes steeper and then less steep again. The final values  $N_{c,f}$  for the number of clusters, which are shown in the right picture of Fig. 3, exhibit a very interesting behavior: the data points can be well fitted to a parabolic function for small  $N$ .  $N_{c,f}$  first increases with increasing  $N$  till  $N = 450$  to a value of  $\sim 261.41$  and decreases afterwards till  $N = 1400$  to a value of  $\sim 17.56$ . This symmetry can be easily explained: For small  $N$ , the bottom of the cylinder is gradually filled with a two-dimensional agglomeration of droplets, forming clusters, with increasing  $N$ . This behavior is mirrored in the bottom area free of droplets, which is gradually split in a first increasing number of clusters of connected free areas, but then more and more of these free area clusters vanish. For larger values of  $N$ , the droplets need to be stacked in a three-dimensional way and the number of droplet clusters fluctuates between 12.02 and 25.68.

Another quantity which needs to be considered is the size of the largest cluster Stauffer and Aharony (1994); Stauffer (1986); Naftaly et al. (1991), in order to find out whether one large cluster is dominating the overall system and whether even the whole network is percolating. For this purpose, we measure the size  $C_{\text{max}}$  of the largest cluster, i.e., the number of nodes contained in the largest cluster. The left picture in Fig. 4 shows that  $C_{\text{max}}$  normalized by  $N$  increases sigmoidally to almost 1 for  $N \geq 1000$ . The final values  $C_{\text{max},f}$  are shown in the right picture of Fig. 4. We find that these final values increase almost linearly with  $N$  for  $N \geq 1000$ , they are only slightly smaller than  $N$ . Almost all nodes (aside from 12-27) are part of the largest cluster. The situation is different for smaller  $N$ , as can be nicely seen in the inset. Here we find that the number of particles not being part of the largest cluster first increases, then peaks at  $N = 550$  with a value of 468, and afterwards decreases again. This structure is reminiscent of the parabolic shape in the right picture of Fig. 3, but here the form of the peak is less symmetric.

### Existence of Paths

In order to better understand the results for geodesic distances, we first need to know how many paths in the network exist. We count the number  $N_p$  of existing paths between pairs of nodes, i.e., the number of geodesic distances which have a finite value. Please note that while there is a geometric restriction for the maximum value of a degree of

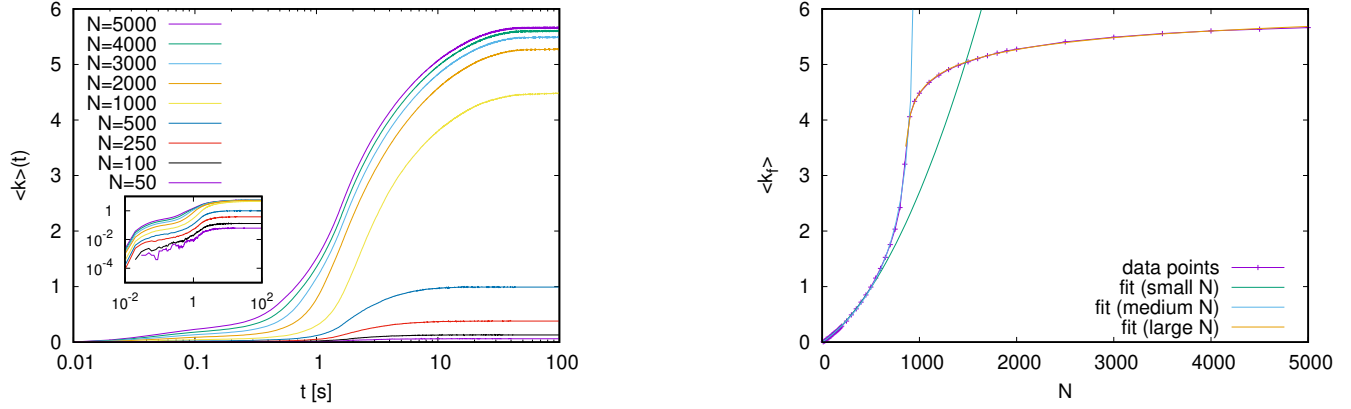


Figure 2: Left: Time evolution of the average value  $\langle k \rangle$  of degrees for various numbers  $N$  of particles. In the inset, the curves are redrawn in a double-logarithmic plot. Right: Average  $\langle k_f \rangle$  of final degree values vs. number  $N$  of particles: The displayed fit function for small  $N$  is given by  $f(N) = 1.2 \times 10^{-3} \times N + 1.5 \times 10^{-6} \times N^2$ , the fit function for intermediate  $N$  with  $40 \leq N \leq 900$  by  $f(N) = 2.222848282 \times 10^{-2} \times (982.95 - N)^{-0.4597383} \times N^{1.064929}$ , and the fit function for large  $N$  with  $850 \leq N \leq 5000$  by  $f(N) = 6.4592928303905470 \times \tanh(0.41066892597996169 \times (N - 843.03747786400800)^{0.14512724870611132})$ .

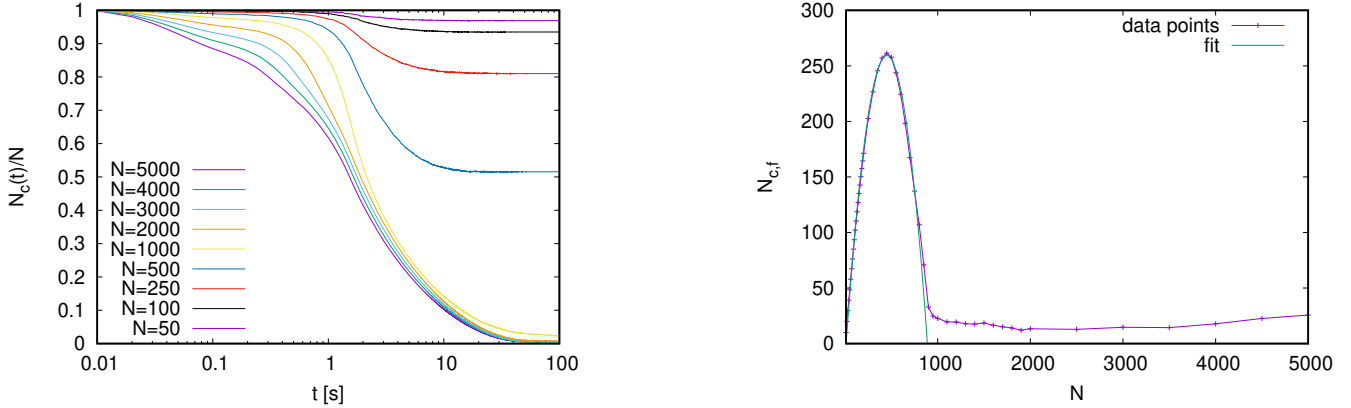


Figure 3: Left: Time evolution of the number  $N_c$  of clusters, normalized by  $N$ , for various numbers  $N$  of particles. Right: Final values  $N_{c,f}$  of the number of clusters for various numbers  $N$  of particles. The parabolic shape for small  $N$  is fitted with  $f(N) = 260 - 0.00135 \times (N - 450)^2$ .

a node, which is also called the kissing number (Schneider et al., 2022), there is no such restriction for  $N_p$ .  $N_p$  can take values up to  $N \times (N - 1)/2$ , for which all possible paths between pairs of nodes exist. In Fig. 5, we have a look at the fraction

$$f_p = \frac{2N_p}{N \times (N - 1)} \quad (9)$$

of existing paths. This fraction is restricted to the range from  $f_p = 0$ , for which no edge exists in the network, to  $f_p = 1$ , for which the network is connected, i.e., for which a path exists from any node to every other node. The inset in the left picture of Fig. 5 with the double-logarithmic plot reveals that  $f_p$  exhibits a double-sigmoidal increase over time. We are interested in the second sigmoidal increase. The right

picture in Fig. 5 shows the final values of  $f_p$  at the end of the simulation runs for various  $N$ . Here we find again a double-sigmoidal increase. Even for the largest values of  $N$  used, the fraction of existing paths does not reach a value of 1 but only of  $\sim 0.99$ . Thus, there are either singular droplets or very small groups of droplets which are not connected with the remaining system. These droplets could lie at the bottom of the cylinder without touching other droplets.

### Geodesic Distances

In the next step, when intending to measure a mean geodesic distance, we have to define how exactly we intend to measure it. The main problem here is how to take those distances between pairs of nodes into account for which no path exists. We defined above that  $d(i, j) = \infty$  in this case, but this defi-

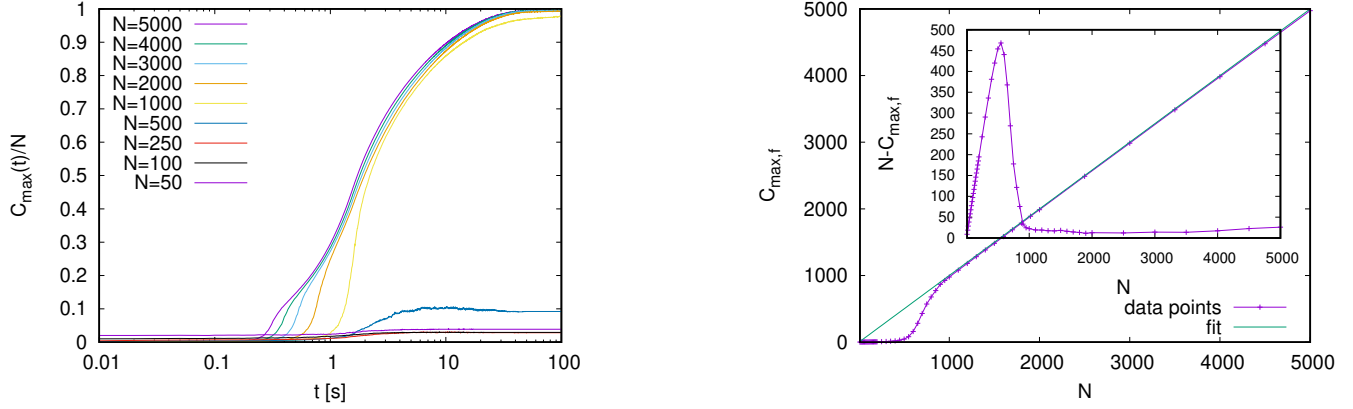


Figure 4: Left: Time evolution of the size  $C_{\max}$  of the largest cluster, normalized by  $N$ , for various numbers  $N$  of particles. Right: Final values of the size  $C_{\max,f}$  of the largest cluster for various numbers  $N$  of particles. The fit is given by  $f(N) = N$ . The inset displays the difference of the numbers of particles and the final size of the largest cluster, i.e., the number of particles not included in the largest cluster.

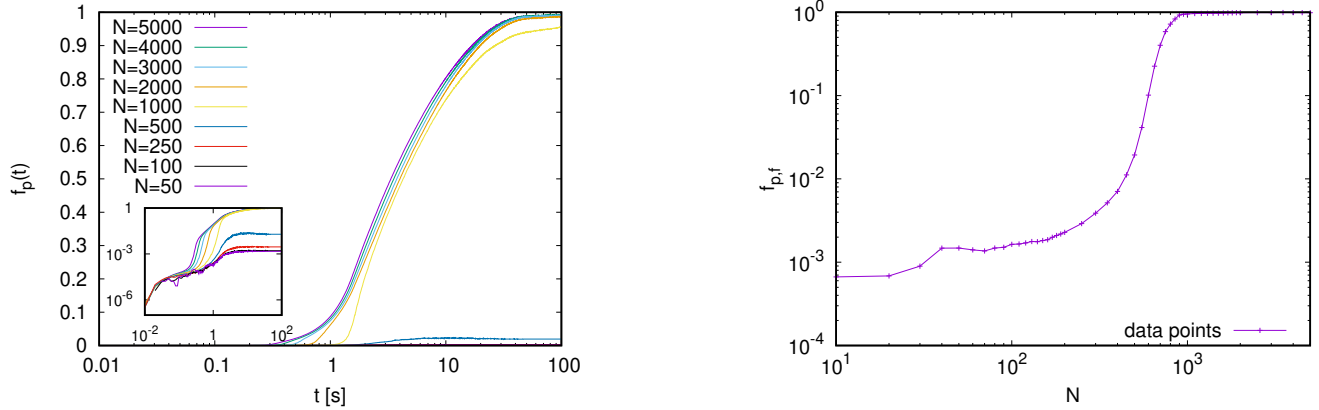


Figure 5: Left: Time evolution of the fraction  $f_p$  of existing paths between pairs of nodes for various numbers  $N$  of particles. In the inset, the curves are redrawn in a double-logarithmic plot. Right: Final values  $f_{p,f}$  for the fraction of existing paths for various numbers  $N$  of particles.

nition is not suitable for taking averages. If no path between a pair  $(i, j)$  of nodes exist, some authors set  $d(i, j)$  to an arbitrarily chosen value  $v \geq N$ , as for an existing path, a distance can only take a value in the range  $1 \leq d(i, j) \leq N - 1$ . As the results then depend strongly on the value  $v$ , we make another choice and set

$$\langle d \rangle = \begin{cases} \frac{1}{N_p} \sum_{\substack{i < j \\ \exists \text{ path } i \rightarrow j}} d(i, j) & \text{if } e > 0 \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Therefore, if there is at least one edge and thus at least one path in the network, then we take the average of the geodesic distances of the existing paths only. Otherwise, we indicate with a value of 0 that no path exists. But if there is exactly one edge and thus one path in the network, then  $\langle d \rangle = 1$ . At the beginning, one will thus see a transition from 0 to

1, as the probability for the existence of a first edge increases. With an increasing number of edges, also longer distances can occur, such that the mean value increases. Figure 6 shows the computational results for  $\langle d \rangle$ . As already mentioned above, we generally plot the ensemble average over 100 simulations. The inset nicely shows the transition whether some small network exists for small  $N$ . And also otherwise the graphic exhibits some interesting properties for  $\langle d \rangle$  and its time evolution, the most important of them being the intermediary maximum occurring for larger numbers of droplets.

If we have a look at the final values  $\langle d_f \rangle$  of the mean distances for various  $N$ , we find that it first increases to some intermediate maximum, decreases again and then slightly increases afterwards with increasing  $N$ , as shown in the right picture of Fig. 6. The maximum lies in a range of  $N$ , in

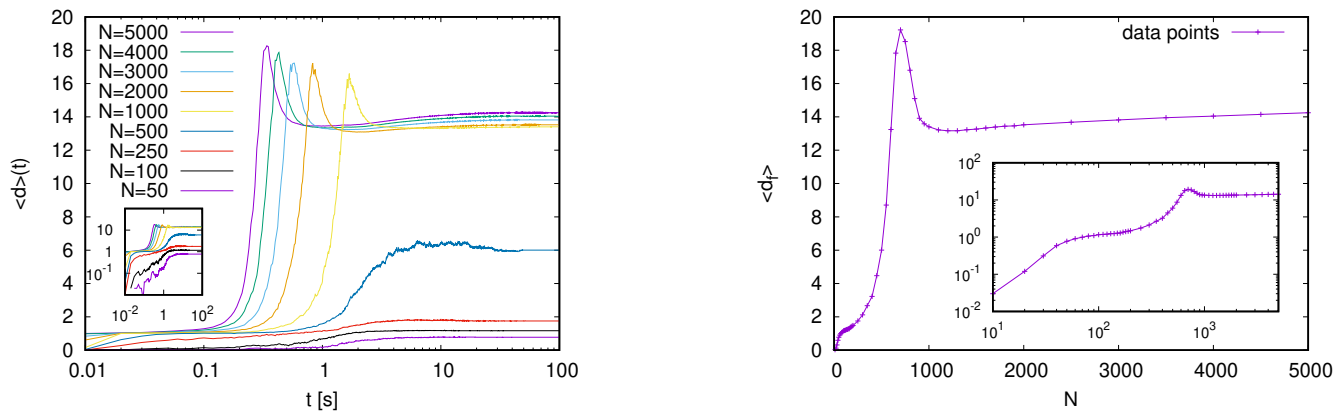


Figure 6: Left: Time evolution of the mean value  $\langle d \rangle$  of the geodesic distances for various numbers  $N$  of particles. In the inset, the curves are redrawn in a double-logarithmic plot. Right: Final values  $\langle d_f \rangle$  for the mean value of geodesic distances for various numbers  $N$  of particles. The inset, in which the curve is redrawn in a double-logarithmic plot, reveals a linear increase of  $\langle d_f \rangle$  for small  $N$ .

which the transition from a quasi two-dimensional network at the bottom of the cylinder to a three-dimensional network starts to take place. (This maximum is missing in the left picture of Fig. 6, as no curves for  $500 < N < 1000$  are shown there.) Obviously, shorter paths can at first be found by adding further droplets in the third dimension. But as the particles have to be stacked in the third dimension even further with further increasing  $N$ , the mean distance starts to increase again.

Similarly, the intermediate maxima in the curves for the largest values of  $N$  can be explained. In Fig. 7, we have a closer look at these intermediate sharp maxima which occur only for  $N \geq 800$ . (Wide maxima which are higher than the final values also occur at slightly smaller values of  $N$ .) The larger  $N$ , the earlier the maximum occurs at a time  $t_m$ :

$$t_m \propto \frac{1}{N} \quad (11)$$

For the largest values of  $N$  considered here, the height of the intermediate maximum increases with increasing  $N$ .

In the next step, we would like to consider the maximum distance of the network. Here again we first have to properly define how we intend to measure this maximum distance. In accordance to the definition of the mean distance in Eq. (10), we define the maximum distance as

$$d_{\max} = \begin{cases} \max\{d(i, j) | \exists \text{ path } i \rightarrow j\} & \text{if } e > 0 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Please note that when it comes to considering maximum distances, two types of maximum distances are often discriminated (Boitmanis et al., 2006; Erdős et al., 1989): Other authors first start off with determining a maximum distance  $d_{i, \max}$  for each node defined as the maximum of distances for all paths starting at node  $i$  and ending at some other node.

Then the diameter  $\mathcal{D}$  of a network is defined as the maximum of these maxima,

$$\mathcal{D} = \max_i d_{i, \max}, \quad (13)$$

and the so-called radius  $\mathcal{R}$  of a network as the minimum of these maxima,

$$\mathcal{R} = \min_i d_{i, \max}. \quad (14)$$

In this sense, our maximum distance  $d_{\max}$  corresponds to the diameter of the network.  $d_{\max}$  is of special interest for us, as we need to know the maximum number of steps possible for the gradual reaction scheme which shall be governed by the compiler we intend to develop.  $d_{\max}$  corresponds to or provides a good estimate for this maximum number of steps. Contrarily, we are not interested in the radius of our networks, as we have already seen that isolated droplets or small isolated groups of droplets can lie at the bottom of the cylinder. But these isolated parts, which would define  $\mathcal{R}$ , play no further role in our considerations.

Figure 8 shows the time evolution of the maximum distance  $d_{\max}$ . At first glance, one sees the similarity to the time evolution of the mean distance  $\langle d \rangle$  in Fig. 6: Again we get a double-sigmoidal increase: at first, the maximum value increases to 1, such that we now know that only pairs of droplets are formed in this stage, with the probability for pair formation increasing with time and with increasing  $N$ . The subsequent main sigmoidal increase can be analyzed as for  $\langle d \rangle$ . Again we have a look at the final values for various  $N$ , which are shown in the right picture of Fig. 8. The final values first sharply increase till  $N = 700$ , for which an ensemble average of the final maximal distances of 50.85 is found. Then it decreases again till  $N = 1300$ . For large  $N$ , we find a slight increase of the final values of  $d_{\max}$ .

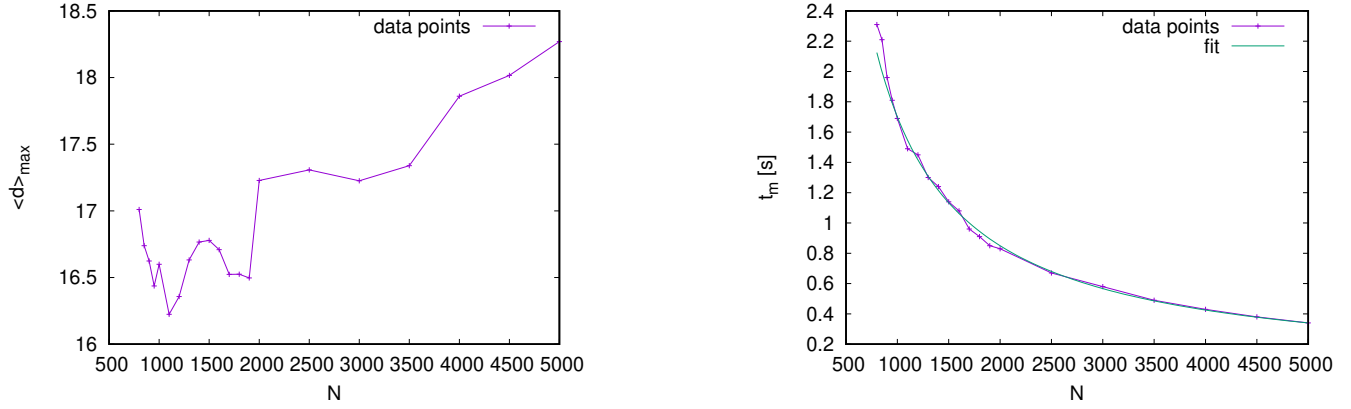


Figure 7: Left: Heights of intermediate maxima of  $\langle d \rangle$  for  $N \geq 800$ . Right: Times at which the intermediate maxima occur for  $N \geq 800$ . The fit function is given by  $1.7 \times 10^3 \text{s} \times N^{-1}$ .

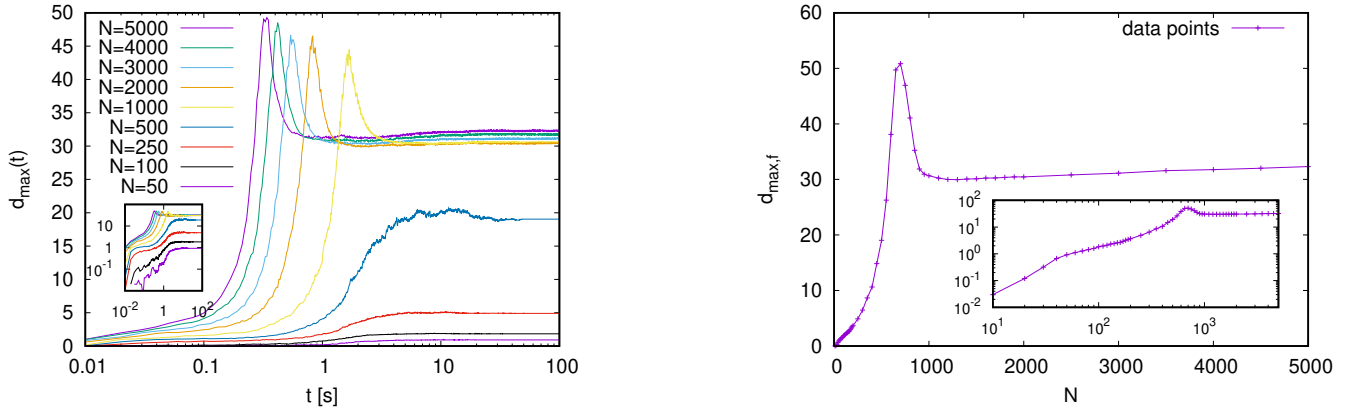


Figure 8: Left: Time evolution of the maximum  $d_{\max}$  of the geodesic distances for various numbers  $N$  of particles. In the inset, the curves are redrawn in a double-logarithmic plot. Right: Final values  $d_{\max,f}$  for the maximum value of geodesic distances for various numbers  $N$  of particles. The inset, in which the curve is redrawn in a double-logarithmic plot, reveals a linear increase of  $d_{\max,f}$  for small  $N$ .

We also find intermediate maxima again for  $N \geq 800$ . The time  $t_{\max,m}$  at which  $d_{\max}$  exhibits its intermediate maximum is also rather the same as the time the intermediate maxima occur for  $\langle d \rangle$  and we find again the power law

$$t_{\max,m} \propto \frac{1}{N}, \quad (15)$$

as shown in Fig. 9. We also find that the heights of these intermediate maxima increase with increasing  $N$  for large  $N$ . Besides these sharp maxima which we get for  $N \geq 800$ , we also find wide intermediate maxima with fluctuating heights for slightly smaller values of  $N$ .

## Conclusion and Outlook

In this paper, we presented results of simulations for the agglomeration of droplets. As we are interested in the effects

of varying numbers of particles on the agglomeration process and on the resulting droplet networks, we study a very simplified system, in which the droplets are represented as hard spheres, subjected to gravity reduced by buoyancy, as well as Stokes friction, added mass effect, random velocity changes, and almost-elastic impacts. Connections between these particles are virtually formed if they (almost) touch or overlap. The particles gradually agglomerate at the bottom of the cylindrical container. The analysis of this agglomeration process shows that the results for the time evolution and the final outcome strongly depend on the number of particles. In particular, we find two transition regimes: at small numbers  $N$  of particles, we find an over time gradually increasing number of droplets lying finally at the bottom of the cylinder where they either stay isolated or gradually form pairs and then some slightly larger groups with other parti-



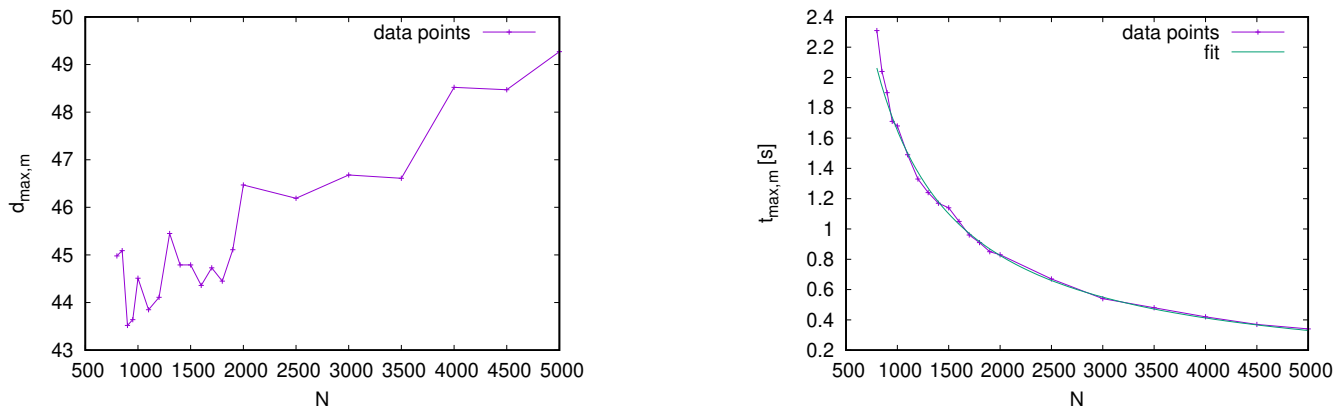


Figure 9: Left: Heights of intermediate maxima of  $d_{\max}$  for  $N \geq 800$ . Right: Times at which the intermediate maxima occur for  $N \geq 800$ . The fit function is given by  $1.65 \times 10^3 \text{s} \times N^{-1}$ .

cles reaching the bottom as well. When increasing  $N$  even further, a network of droplets is created at a bottom layer of the cylinder. During this regime, we find an increase of the mean and maximum geodesic distances, followed by a decrease for larger system sizes. For very large numbers of particles, the time evolution of the observables at first more or less reflects the final outcome for small and then intermediate numbers of particles, showing e.g. intermediate maxima for the geodesic distances (the earlier, the larger the number of particles is), before displaying the properties of truly three-dimensional networks of particles. As only large values of the maximum distance allow very extended gradual reaction schemes governed by our compiler, we suspect that it is misleading to believe that the structures optimum for our purposes need to be three-dimensional. Instead, when using a unary system of droplets, we should consider working with the largest cluster in a quasi two-dimensional layer at the bottom of the cylinder, which exhibits a fractal dimension (Falconer, 2003) significantly smaller than 2.

We intend to continue our investigations by measuring clustering coefficients, fractal dimensions, the locations of droplets with differing radii, and the importance some droplets might have for the overall network. Furthermore, we plan to extend our investigations first to binary systems, in which two particle types  $A$  and  $B$  are present and connections can only be forged between pairs of  $A - B$  but not  $A - A$  or  $B - B$  and then to ternary systems, in which there are three particle types  $A$ ,  $B$ , and  $C$  with connections between adjacent pairs of  $A - B$  particles but in which the additional  $C$ -particles are unable to form any connections. Hereby we want to study the breakdown of the size of the largest cluster with increasing density of  $C$ -particles and find out whether there is a regime as well in which the maximum geodesic distance between droplets increases to suitable values. Furthermore, we want to add gluing forces between particles to find out how they change the results.

## Acknowledgements

This work has been partially financially supported by the European Horizon 2020 project *ACDC – Artificial Cells with Distributed Cores to Decipher Protein Function* under project number 824060.

## References

- Aprin, L., Heymes, F., Laureta, P., Slangena, P., and Le Floch, S. (2015). Experimental characterization of the influence of dispersant addition on rising oil droplets in water column. *Chemical Engineering Transactions*, 43:2287–2292.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Boitmanis, K., Freivalds, K., Lediņš, P., and Opmanis, R. (2006). Fast and simple approximation of the diameter and radius of a graph. In Álvarez, C. and Serna, M., editors, *Experimental Algorithms*, pages 98–108, Berlin, Heidelberg. Springer.
- Bollobás, B. (2011). *Random Graphs, 2nd ed.*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271.
- Donev, A., Cisse, I., Sachs, D., Variano, E. A., Stillinger, F. H., Connelly, R., Torquato, S., and Chaikin, P. M. (2004). Improving the density of jammed disordered packings using ellipsoids. *Science*, 303(5660):990–993.
- Erdős, P., Pach, J., Pollack, R., and Tuza, Z. (1989). Radius, diameter, and minimum degree. *Journal of Combinatorial Theory, Series B*, 47:73–79.
- Falconer, K. (2003). *Fractal Geometry: Mathematical Foundations and Applications, 2nd edition*. John Wiley & Sons, Chichester, West Sussex.
- Flumini, D., Weyland, M. S., Schneider, J. J., Fellermann, H., and Füchsli, R. M. (2020). Towards programmable chemistries.

- In Cicirelli, F., Guerrieri, A., Pizzuti, C., Socievole, A., Spezzano, G., and Vinci, A., editors, *Artificial Life and Evolutionary Computation, Wivace 2019, Communications in Computer and Information Science*, volume 1200, pages 145–157.
- González, J. A. and Cascone, M. H. (2014). Geodesic distribution in graph theory: Kullback-Leibler-Symmetric. *Revista de Matemática: Teoría y Aplicaciones*, 21:249–260.
- Hadorn, M., Boenzli, E., Sørensen, K. T., Fellermann, H., Eggenberger Hotz, P., and Hanczyc, M. M. (2012). Specific and reversible dna-directed self-assembly of oil-in-water emulsion droplets. *Proceedings of the National Academy of Sciences*, 109(50):20320–20325.
- Kloeden, P. and Platen, E. (2013). *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer, Berlin, Heidelberg.
- Li, J. and Barrow, D. A. (2017). A new droplet-forming fluidic junction for the generation of highly compartmentalised capsules. *Lab on a chip*, 17:2873–2881.
- Marino, R. and Kirkpatrick, S. (2018). Revisiting the challenges of maxclique. *CoRR*, abs/1807.09091.
- Matuttis, H.-G. and Chen, J. (2014). *Understanding the Discrete Element Method: Simulation of Non-Spherical Particles for Granular and Multi-Body Systems*. John Wiley & Sons, Singapore Pte. Ltd.
- Metcalfe, G., Shinbrot, T., McCarthy, J. J., and Ottino, J. M. (1995). Avalanche mixing of granular solids. *Nature*, 374:39–41.
- Mitarai, N. and Nakanishi, H. (2003). Hard-sphere limit of soft-sphere model for granular materials: Stiffness dependence of steady granular flow. *Phys. Rev. E*, 67:021301.
- Müller, A., Schneider, J. J., and Schömer, E. (2009). Packing a multidisperse system of hard disks in a circular environment. *Phys. Rev. E*, 79:021102.
- Naftaly, U., Schwartz, M., Aharony, A., and Stauffer, D. (1991). The granular fracture model for rock fragmentation. *Journal of Physics A: Mathematical and General*, 24(19):L1175–L1184.
- Ochoa, J. G. D., Binder, K., and Paul, W. (2006). Molecular dynamics simulations of the embedding of a nano-particle into a polymer film. *J. Phys.: Condens. Matter*, 18(10):2777–2787.
- Pfleiderer, P. and Schilling, T. (2007). Simple monoclinic crystal phase in suspensions of hard ellipsoids. *Phys. Rev. E*, 75:020402.
- Reiss, H., Ellerby, H. M., and Manzanares, J. A. (1996). Ornstein-zernike-like equations in statistical geometry: Stable and metastable systems. *The Journal of Physical Chemistry*, 100(14):5970–5981.
- Ricci, A., Nielaba, P., Sengupta, S., and Binder, K. (2007). Ordering of two-dimensional crystals confined in strips of finite width. *Phys. Rev. E*, 75:011405.
- Russel, W. B., Saville, D. A., and Schowalter, W. R. (1989). *Colloidal Dispersions*. Cambridge Monographs on Mechanics. Cambridge University Press.
- Schneider, J. J., Barrow, D. A., Li, J., Weyland, M. S., Flumini, D., Eggenberger Hotz, P., and Fuchslin, R. M. (2022). Geometric restrictions to the agglomeration of spherical particles. In *Artificial Life and Evolutionary Computation, Wivace 2021, Communications in Computer and Information Science*. accepted for publication.
- Schneider, J. J. and Kirkpatrick, S. (2005). Selfish versus unselfish optimization of network creation. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(08):P08007.
- Schneider, J. J., Müller, A., and Schömer, E. (2009). Ultrametricity property of energy landscapes of multidisperse packing problems. *Phys. Rev. E*, 79:031122.
- Schneider, J. J., Weyland, M. S., Flumini, D., and Fuchslin, R. M. (2020a). Investigating three-dimensional arrangements of droplets. In Cicirelli, F., Guerrieri, A., Pizzuti, C., Socievole, A., Spezzano, G., and Vinci, A., editors, *Artificial Life and Evolutionary Computation, Wivace 2019, Communications in Computer and Information Science*, volume 1200, pages 171–184.
- Schneider, J. J., Weyland, M. S., Flumini, D., Matuttis, H.-G., Morgenstern, I., and Fuchslin, R. M. (2020b). Studying and simulating the three-dimensional arrangement of droplets. In Cicirelli, F., Guerrieri, A., Pizzuti, C., Socievole, A., Spezzano, G., and Vinci, A., editors, *Artificial Life and Evolutionary Computation, Wivace 2019, Communications in Computer and Information Science*, volume 1200, pages 158–170.
- Stauffer, D. (1986). Percolation and cluster size distribution. In Stanley, H. E. and Ostrowsky, N., editors, *On Growth and Form: Fractal and Non-Fractal Patterns in Physics*, pages 79–100. Springer Netherlands, Dordrecht.
- Stauffer, D. and Aharony, A. (1994). *Introduction to Percolation Theory*, 2nd ed. Taylor & Francis, London.
- Stokes, G. G. (1851). On the effect of the internal friction of fluids on the motion of pendulums. *Transactions of the Cambridge Philosophical Society*, 9:8–106.
- Weyland, M. S., Flumini, D., Schneider, J. J., and Fuchslin, R. M. (2020). A compiler framework to derive microfluidic platforms for manufacturing hierarchical, compartmentalized structures that maximize yield of chemical reactions. *Artificial Life Conference Proceedings*, (32):602–604.
- Zallen, R. (1998). *The Physics of Amorphous Solids*. A Wiley-Interscience publication. Wiley.