The physics of collective cell motion

structure, jamming and alignment

Valk, Marijke

Award date:
2018

Graduation thesis

THE PHYSICS OF COLLECTIVE CELL MOTION: STRUCTURE, JAMMING AND ALIGNMENT

Marijke Valk
Theory of Polymers and Soft matter,
Department of Applied Physics,
Eindhoven University of Technology
July 18, 2018

Supervised by:
Dr. L.M.C. Janssen and Prof. Dr. C. Storm, Eindhoven University of Technology
In collaboration with: Dr. J.P. Butler, Harvard School of Public Health

HARVARD
SCHOOL OF PUBLIC HEALTH

Abstract

Collective motion in confluent cell layers plays an important role in biological processes such as embryonic development, wound healing, cancer and asthma. In these processes the cellular collective can undergo a jamming transition from a fluid-like to a solid-like state reminiscent of a glass transition in molecular fluids or a jamming transition in granular systems. Using both simulated data from the self-propelled Voronoi model and experimental data obtained from MDCK-cells, we study this jamming phenomenon, the accompanying structural and dynamical changes and the analogy between cells and glasses via correlation functions and mode coupling theory (MCT) obtained from glassy physics. These static and dynamic correlation functions show a growth in structural order and a dynamical slowing down upon approaching the jamming transition and are therefore a powerful tool to characterize this jamming transition. Based on structural information, standard MCT predicts the dynamics in the unjammed regime correctly and this shows that dynamical changes are governed by structural changes in confluent cell layers and supports the notion that collective migration exhibits glass-like dynamics. For the analysis of collective cellular motility, MCT provides us with a new analysis framework to determine dynamics based on only one snapshot instead of expensive measurements during a long period. In addition, this MCT analysis in confluent cell layers extends the applicability of mode coupling theory to a new field of science, namely biophysics. We also investigate jamming and alignment in an extension of the self-propelled Voronoi model that explicitly includes neighbor-alignment as an organizing principle in collective motion. We show that increasing the radius that determines the neighbors to which a cell aligns its velocity to increases the total alignment in the system and can cause the cellular collective to flow like a flock, in agreement with experimental results. Furthermore, we show that the switch on of this neighbor-alignment can fluidize a jammed layer and can re-solidify this layer upon further increasing the radius of the neighbor-alignment. The combination of jamming and alignment results in a rich phase diagram with four different states: a stationary liquid, a stationary solid, a liquid flock and a solid flock.

Contents

1 Introduction

Cells are the building blocks of the human body and of all other living organisms and their motion plays an important role in many biological processes. Nevertheless, the fundamental rules governing this motion are still not well understood. Single cells already exhibit very complex motion [1, 2], but in many biological situations cells are coordinating their movements together as a collective. This resulting collective cellular motion is even more complicated and has been listed as one of the ten major unsolved mysteries in biology [3]. Recently, observations in different physiological and pathophysiological processes have revealed that cells can stay quiescent and fixed around their position in some situations while these same cells can mobilize and migrate cooperatively over great distances in other situations [4–12]. So just as coffee beans in a chute or grain in a silo can flow in some circumstances and jam in others, a cellular collective can unjam and flow like a fluid or instead jam and rigidify like a solid [13–19].

In physiological processes like embryonic development and wound healing, differences in behavior between the sticky cells in the solid state and the collective motion in the fluid state play an important role. Starting from a single cell, cell division and differentiation combined with the collective migration of the cells through the whole body make it possible for organs and the human body as a whole to develop [20, 21]. In this unjammed collective, each individual cell is guided by chemical and mechanical stimuli from its neighbors. In a jammed state on the other hand, each individual cell performs the physiological functions that its type dictates. Epithelial cells, for example, line the inner and outer surface of all organs and cavities in the human body and in a solid state serve as a barrier to protect, separate, sense, absorb and secrete [14]. Another example of the function of collective motion of the cells is shown in Figure 1 where epithelial cells move cooperatively towards a wound to remove harmful bacteria and to close the layer [5, 22–26]. As soon as opposite cells meet, the layer starts to jam again. The fact that the speed of a single cell does not determine the wound healing rate, shows that the cooperation between cells play an important role in this process [25].

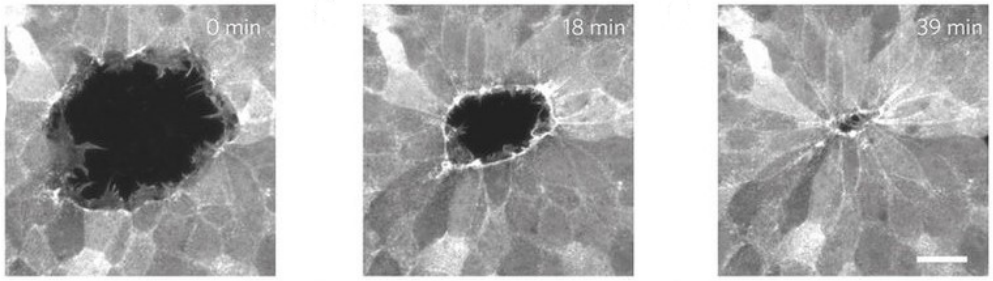

Figure 1: Cells working together in wound healing. The black area in the left panel is a wound and cells are collectively moving towards this wound (center panel) to close it (right panel) [25].

In pathophysiological processes like cancer and asthma, unjamming can be a malfunction or failure of human tissue cells. Cancer [15, 16, 27–30] is characterized by uncontrolled cell growth, invasion into surrounding tissues and metastases to other parts of the body. Important in cancer metastasis is the epithelial-mesenchymal transition (EMT), which is a process in which epithelial cells gain migratory and invasive properties and become mesenchymal stem cells that can differentiate into a variety of cell types. But recently, another gateway to cellular migration was proposed, namely an unjamming transition in which cells retain their phenotype. It is yet unknown how these two processes, EMT and unjamming, are intertwined, but the idea that influencing the unjamming transition might be a key to cancer treatments is growing. More established is the role of unjamming in airway epithelium cells and their malfunctioning by asthmatic patients [10]. The jammed phase of these epithelial cells represents a mature and quiescent layer and a good working barrier function, while the unjammed phase is associated with an immature and harmed layer. To illustrate the difference between those two states, Figure 2 shows the cell shape of human bronchial epithelial cells (HBEC) in a jammed and immobile state (left panel) and in an unjammed and rapidly migrating state (right panel). While breathing, the airway epithelium is subject to repeated mechanical perturbations and injuries due to exposure to, for example, pollutions and viruses. These disruptions can cause the cellular collective to unjam to repair the epithelial layer. During this repair, the layer explores different configurations and then resettles into a recovered and remodelled solid-like state. In this final state, the layer is jammed again. But in asthmatic patients this jamming turns out to be delayed and the cellular collective stays in a fluid-like state for a longer time period, resulting in deterioration of the airway function and unwanted structural remodeling of the airway wall.

Collective motion of cell tissues, as described in the examples above, is regulated by many complex physical and molecular factors which are strongly effected by each other [31]. This causes cells to exhibit a much greater

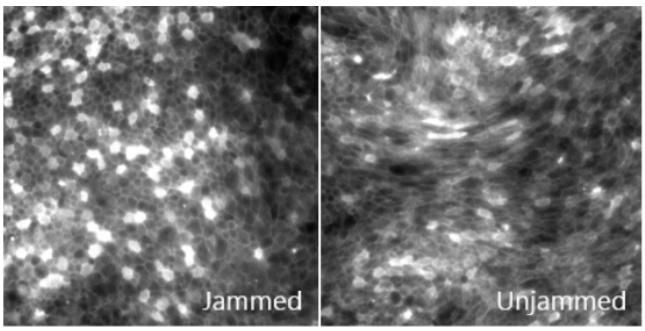

Figure 2: Human bronchial epithelial cells (HBEC) in a jammed and immobile state (left panel) and in an unjammed and rapidly migrating state (right panel) [14]. Although the movements of the cells are not visible in these pictures, there is a clear difference in the cell shapes in the jammed and unjammed state.

range of behavior than possibly be found in non-biological materials [32]. Cells can grow, deform, divide and die. Signaling molecules can alter a cell's mechanical properties, complex feedback mechanisms can cause them to change their behavior and long-range chemical communication can influence their movements. Furthermore, cells are highly heterogeneous and can have different phenotypes making them communicate differently. It is therefore very complicated to take into account all these factors in describing biological processes. Therefore, a unifying framework might be highly valuable for understanding the influence of underlying factors and make the interactions between them better understood [14]. It has been shown that the transition between solid-like and fluid-like biological tissues is accompanied by large-scale collectivity, a slowing down of the dynamics and a dynamic heterogeneity exhibited by multicellular packs and swirls (see Figure 3), even in situations where no change is observed in the properties of the constituent cells and the structure of the collective [6, 9, 10, 13, 14, 29, 32, 33]. Strikingly, these characteristics are also found in molecular fluids and particulate systems close to, respectively, the glass or jamming transition. The suggestion that this fluid to solid transition in biological processes is similar to a glass or jamming transition is called the jamming hypothesis and has the potential to be a unifying framework that links diverse biological processes and gives insights into the laws governing collective cellular dynamics [4, 13, 14, 34].

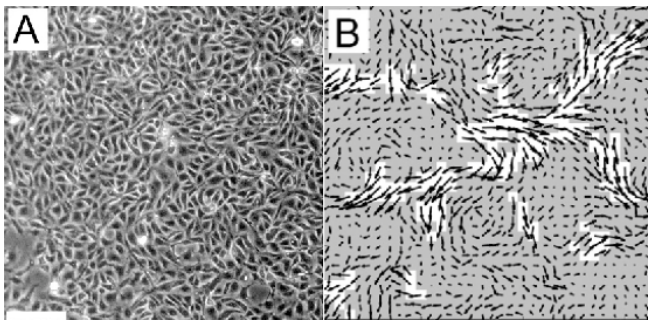

Figure 3: A snapshot of Madine-Darby canine kidney cells (MDCK) close to a jamming transition. The velocity field (right panel) shows swirls of cells with approximately the same velocity magnitude and direction, which indicate spatially heterogeneity [4].

So what happens at this glass or jamming transition in non-living materials? In the well-known fluid to solid transition, a spontaneous ordering of the particles or molecules takes place from an amorphous, disordered fluid state to a well-ordered, crystallized solid state. But in a jamming or glass transition, this spontaneous structural ordering is lacking and disorder exists in both liquid- and solid-like phases [F1]. Figure 4 shows an example of the structural differences between this crystallized (left panel) and the so-called glassy state (right panel). In molecular fluids, this glassy state is reached by rapidly cooling a material allowing no time for reordering to the crystalline ground state. Thermal fluctuations are now insufficient to drive local structural arrangement and the system will stay away from a thermodynamic equilibrium. Consequently, the resulting packing geometries remain disordered. Instead of decreasing the temperature, in particulate systems the transition from a liquid-like to a solid-like state can be reached by increasing the density. Despite the differences in origins, the phenomenological characteristics of both transitions are the same. Both the jamming and the glass transition are characterized by a slowing down of the motions of the individual components due to caging of the immediate neighbors as shown in Figure 5. In the left panel, particles are in a liquid or unjammed state and can move around freely without being hindered by their neighbors, but the particles in the right panel are hindered escaping from the confinement of their neighbors. This causes the particles to be caged and dynamically arrested.

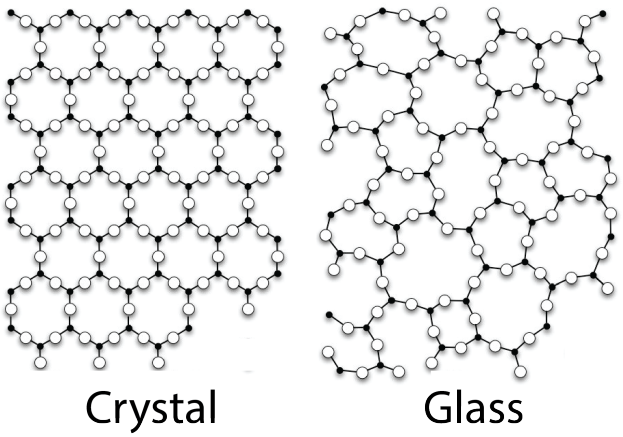

Figure 4: Two possibilities for solidification of a simple fluid: a crystalline state (left panel) and a glassy state (right panel) (source: https://steemit.com/science/@timsaid/myth-or-fact-11-glass-is-a-liquid, date of access: 07-06-2018).

These particles exhibit a solid-like character on short times and a liquid-like on long time scales, because it will take a long time before the particles escape the confinement of their direct neighbors and obtain new neighbors. Another feature of the glass and jamming transition is the growth of swirls or clusters that move collectively in the same direction at similar speed (see the right panel of Figure 3). Because of these similarities between the glass and jamming transition in molecular fluids and particulate systems, they are now seen as part of a bigger jamming picture in which there are many possible ways to unjam a material [35, 36]. In Figure 6 a jamming phase diagram is drawn showing three of these possibilities: increasing the temperature, decreasing the density or increasing the (shear) stress [37].

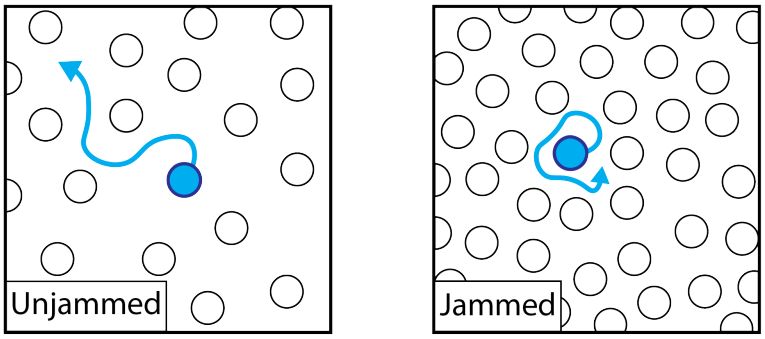

Figure 5: Differences between an unjammed, fluid-like state in which particles can move around freely (left panel) and a jammed, solid-like state in which the particles are caged by their neighbors (right panel).

Living materials might show other pathways to unjam and Figure 7 shows a proposal for a jamming phase diagram for cellular layers. According to this diagram, unjamming can be obtained by decreasing the cell density, decreasing the adhesion between cells or increasing the ability of a cell to move autonomously, which is called motility or activity of a cell [34]. Other examples that can influence the jamming or unjamming transition are cellular or substrate stiffness, stretch or shear loading or cellular volume. At different positions in Figure 7 a sketch of the corresponding cellular layer is given, showing characteristics reminiscent to molecular fluids or particulate systems near a glass or jamming transition. In the unjammed state, cells can move freely and independently of their neighbors. Going towards a jammed situation, these neighbors serve as a cage with some specific strength that can be expressed in the form of an energy barrier. A cell has to overcome this barrier to escape, to see new neighbors and to make cellular rearrangements possible. When the caging becomes stronger, the energy barriers become higher and it becomes more difficult for a cell to escape from its neighbors. Near the jamming transition, cells can only move when they cooperate in a collective fashion and dynamic heterogeneity arises in the tissue. Finally, when the barrier becomes so big that cells can only overcome this barrier rarely, the cell remains trapped in its cage and the layer is said to be jammed. Concluding, the phase diagram in Figure 7 illustrates the importance of understanding the jamming phenomenon in biological processes because a small change in the value of a parameter can have a large influence on the collective behavior of cells [14].

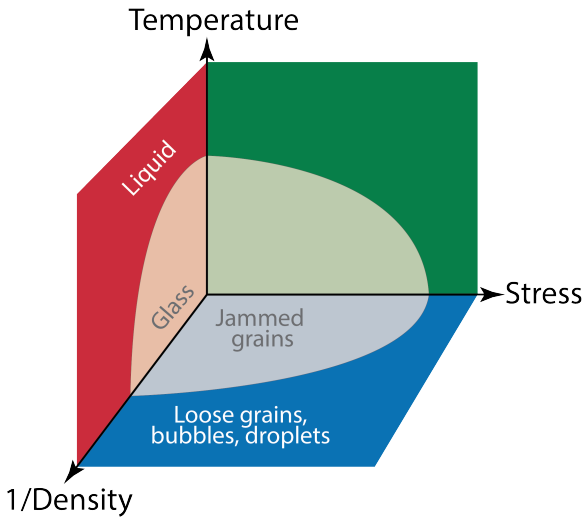

Figure 6: Proposal for a jamming phase diagram that unifies the glass and jamming transition [35].

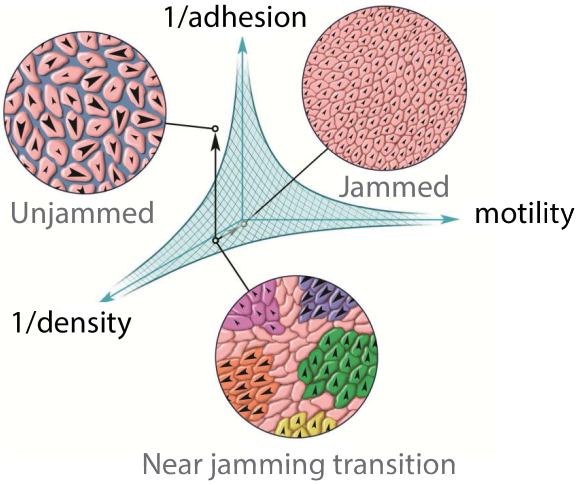

Figure 7: Proposal for a jamming phase diagram for cell layers, depending on adhesion, density and motility. The insets show typical structures and velocity vectors for a jammed state, an unjammed state and a state near the jamming transition [34].

Hence, it is very likely that there is some kind of analogy between collective cell behavior and that of simple fluids or particulate systems near a glass or jamming transition, and this might provide the opportunity to solve many challenging questions in biology such as the following. What are the forces driving collective cellular motion [38]? What is the purpose of jamming and unjamming in biological processes [39]? What is the relation of cell jamming to the dominant mechanisms of collective cellular migration or is cell jamming itself a dominant mechanism [14]? Is unjamming the cause of migratory events or is it only a result [34]? In embryonic development, which signals causes the cells to unjam and migrate? In wound-healing, what controls the unjamming in the cellular layer [13]? In cancer, what is the link between EMT and the unjamming transition [14]? If EMT is not the principal role in cancer as is suggested, might unjamming then plays this role [40]? Can the use of mechanical properties give rise to better treatments than the current therapies that often only consider the biochemical aspects of the disease [34]? In asthma, can we accelerate the process of jamming or reverse the unjamming as a remedy [10]?

However, before we can start thinking about answering these burning and interesting questions, a better understanding of the cellular jamming phenomenon and its analogy to simple fluids is necessary. In this thesis, we aim to get a deeper insight into this collective cellular behavior and we work towards establishing the analogy with glassy fluids or jammed particles further. We do this by applying analysis methods and a predictive theory borrowed from glassy physics to living epithelial cells and by extending the current available cell models with a model that explicitly incorporates a mechanism for collective behavior explicitly. Our research goals are summarized by the three research questions in the next subsection. Furthermore, in Subsection 1.2 the outline of this thesis is explained.

1.1 Research questions

1. What do static and dynamic correlation functions obtained from glassy physics teach us about the jamming phenomenon in epithelial cell layers in experiments and simulations?
The principle of jamming and its characteristics are not well defined yet in cells and, therefore, broader and more fundamental knowledge is needed to establish this phenomenon and to make jamming a general term in biological processes. In glassy physics, static and dynamic correlation functions are frequently used to analyze structure and changes in this structure over time, but these functions are rarely used in studying cell tissues. By treating cell layers in the same way as glass-forming liquids, we aim to extend this knowledge in terms of structure and dynamics and possibly push the analogy between cells and glasses further. To start with, this will provide us with a new lens to look at a successful model for confluent cell layers and already existing experimental data. Just as in molecular fluids or granular particles, we expect a relation between structure and dynamics in confluent cell layers. This would call for applying a specific theory of glasses and jammed materials, called mode coupling theory (MCT).

2. By applying mode coupling theory on epithelial cells, can we broaden the applicability of MCT, extend the analogy between glasses and cells, and, ultimately, define a new analysis framework for epithelial cells?
Mode coupling theory is a promising theory from glassy physics that predicts dynamics based on the structure of a system. The similarities between the jamming transition in biological processes and the glass transition motivates us to apply MCT on simulated and experimental data of epithelial cells. The motivation for this MCT approach is two fold as it can give insights into the analogy between glasses and cells, but also can help to validate MCT in general. Because MCT is not an exact theory, it is only successful in some cases and fails in other cases. Validating MCT predictions for confluent cell layers can substantiate this theory and contribute to its applicability. Furthermore, the intended applicability of MCT to epithelial cells has several benefits to the research of these cell layers. Until now, only simulation models to explain experimental results exist, but a theory is lacking. Such a theory can give additional insights in the underlying physical principles and has the power to make predictions that can be tested in simulations and experiments. Finally, if mode coupling turns out to be applicable to confluent cell layers, it can provide us with an alternative way to define jamming in which only one snapshot is needed instead of many snapshots over long time period.

3. How does a model for epithelial cells behave in terms of alignment and jamming when a principle regulating collective motion is explicitly included?
The existing models for confluent cell layers do not incorporate an explicit rule that regulates the motion of a cell depending on the movements of its neighbors. Nevertheless, experiments show situations in which this collaboration might be a dominant mechanism in collective cell motion. Therefore, we propose a model that includes this mechanism explicitly and will look at how different parameters affect alignment and jamming. The natural next step is testing mode coupling theory for this new model, but that will be left for future research.

1.2 Outline

The outline of this thesis is as follows. In the next section we go more deeply into the physics of glasses and the glass transition for simple fluids. We start with defining the properties of the glass transition in Subsection 2.1 and in Subsection 2.2, the static and dynamic correlation functions are derived. Subsection 2.3 focuses on mode coupling theory and a sketch of the derivation is given, as well as a description of its applicability. In Section 3, we return to biophysics and focus on models for confluent cell layers. Subsection 3.1 lists important properties of these cell layers that have to be taken into account when designing a model. Different model classes exist for these confluent cell layers and these are described in Subsection 3.2, together with their pros and cons for our research. We chose to use two different models, and their details as well as their connection with the literature are described in Subsection 3.3 and Subsection 3.4. In Subsection 3.5 we summarize our computational details.

The first model we investigate is the Voronoi rotational diffusion-model (VRD) and the results are presented in Section 4. Subsection 4.1 describes the general working of this model, Subsections 4.2 and 4.3 focus on respectively the structure and the dynamics using correlation functions and Subsection 4.4 contains a mode coupling theory analysis. In Section 5 experimental data is analyzed in the same way as the data from the VRD-model. In Subsection 5.1 a first exploration of the data set is given, Subsections 5.2 and 5.3 discuss respectively the structure and dynamics using correlation functions and mode coupling theory is applied in Subsection 5.4. In Section 6 we investigate a second model, called the Voronoi Vicsek-model (VV). A first exploration of the model is given in Subsection 6.1 and Subsection 6.2 contains a broad parameter analysis. In Subsection 6.3 the results of the VV-model regarding alignment are compared to experimental results and in Subsection 6.4 jamming in this model is investigated. In Section 7 we look at future research possibilities, with in Subsection 7.1 different possibilities to extend the incorporation of the motility in our research, in Subsection 7.2 a review of other research possibilities and in 7.3 some concluding remarks about what the field to which this research belongs might bring in the far future. Finally, our conclusions are summarized in Section 8.

2 Theory of glasses and the glass transition

In Section 1 is explained that confluent cell layers can exhibit the same features as molecular fluids near a glass transition or granular particles near a jamming transition, but the physics of such a transition in biological systems is not well established yet. In this thesis we examine this transition through a glassy lens and therefore we explain in this section the physics that we need to do so [41, 42]. We mostly refer to molecular liquids and glasses in this section, but we assume that within the unifying jamming framework [35] all the principles can be extended to granular materials and the jamming transition. Nevertheless, we will call the molecules 'particles' for clarity reasons. In Subsection 2.1 the properties of glasses as described in Section 1 are explained in further detail. Static and dynamic correlation functions are a widely used analysis method in glassy physics and they are defined in Subsection 2.2. These correlation functions are necessary to derive mode coupling theory (MCT), the only first principle theory to describe the dynamics of glasses. A sketch of the derivation of MCT, its applicability and its successes and failures are explained in respectively Subsection 2.3.1, 2.3.2 and 2.3.3. We conclude in Subsection 2.4 with a description of how we can transfer the physics of glasses to the physics of confluent cell layers.

2.1 Properties of glasses

In Section 1 is described that a material can reach a glassy state via supercooling, in which the formation of an ordered array and a corresponding crystalline state is suppressed (left panel of Figure 4). In Figure 8 the difference between glass and crystal forming is sketched. At the melting point T_m, a liquid normally freezes in a crystalline state resulting in an abrupt decrease in volume and enthalpy (crystal line in Figure 8). But when the liquid is cooled down fast enough, this first-order phase transition is avoided as each particle becomes trapped in a local cage preventing the system to reach a completely ordered state (two glass lines in Figure 8). At some temperature T_g the liquid becomes so viscous that it does not flow on any practical time scale anymore and from that moment on it is called a glass. Hence, the glass transition is not a clear phase transition between two equilibrium states, but it is an out-of-equilibrium phenomenon. The location of the glass transition is not a material property either, as it depends on the cooling rate (glass 1 and 2 in Figure 8) and the glass temperature T_g is empirically determined. One characteristic of the glass transition is the excessive growth of the viscosity near the transition as is shown for several materials in Figure 9. The temperature for which the viscosity reaches 10^{12} Pa (or 10^{13} poise in Figure 9) is often used as definition for the glass temperature. Other possible measures are the diffusion coefficient obtained from the mean squared displacement and the relaxation time obtained from a dynamic correlation function, which are respectively discussed later in this subsection and in Subsection 2.2.

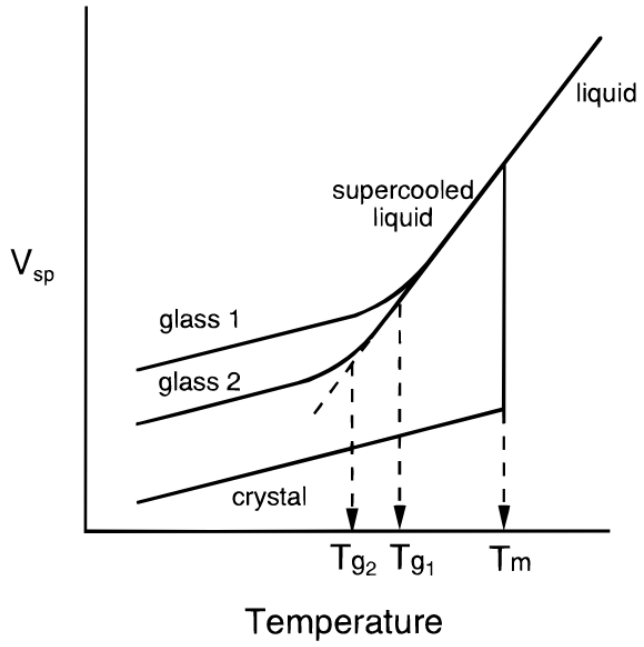

Figure 8: A material can reach a glassy or crystalline state depending on the cooling rate [43].

Figure 9 shows that many different materials exhibit glassy behavior in which the viscosity increases drastically, but also shows the broad range of temperatures at which this happens. This behavior can be unified using an Angell plot, where all temperatures are scaled to the glass temperature (see Figure 10). Materials that show linear growth on the Angell plot are called strong glass formers and materials showing a concave line are fragile glass formers. These fragile glass formers show super Arrhenius behavior, meaning that the energy barriers for structural rearrangements grow nonlinearly when the temperature decreases.

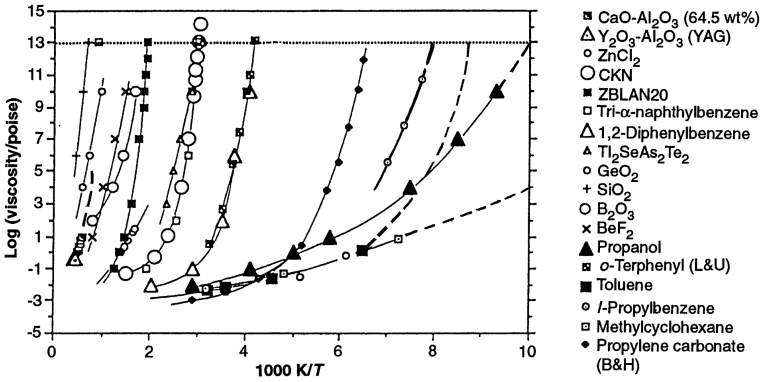

Figure 9: Growth of viscosity for several materials upon supercooling [44].

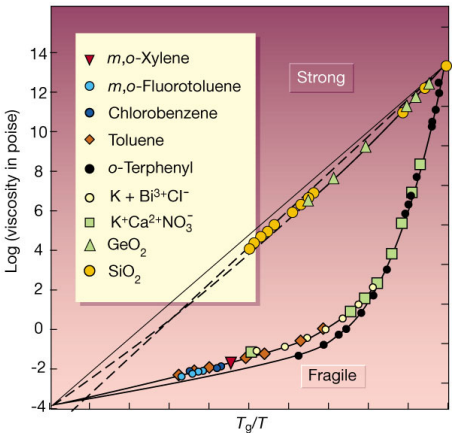

Figure 10: Angell plot of the viscosity for several materials showing the difference between strong and glass formers [43].

Besides the viscosity, this huge dynamical slowing down near the glass transition can also be probed via the mean squared displacement (MSD) given by

$$\mathrm{MSD}(t) = \frac{1}{N}\sum_{i=1}^{N} |\boldsymbol{r}_i(t) - \boldsymbol{r}_i(0)|^2, \tag{1}$$

with $\boldsymbol{r}_i(t)$ the location of particle i at time t. The diffusion coefficient D is the long-time limit of the slope of the MSD, given by

$$D = \lim_{t\to\infty} \frac{\mathrm{MSD}}{2dt}, \tag{2}$$

with d the dimensionality of the system. Three phases can be distinguished in a typical graph of the MSD as displayed in Figure 11. At short times particles behave ballistic because they are not aware of other particles around them and the slope of the MSD on a log-log plot is two. Once a particle senses it neighbors and undergoes collisions, the dynamics slows down and a plateau arises in the MSD. When waiting long enough, a particle can escape from the cage of its neighbor and the MSD grows again. The behavior in this regime is called diffusive when the slope of the MSD on a log-log plot is one or subdiffusive when this slope is smaller than one. However, when the caging of the particles is too strong such that hopping events are not possible, the diffusive regime disappears and the MSD stays at a plateau value. In this case, the system behaves non-ergodically. Because it is impossible to check the dynamical behavior of a system at infinitely long times, the material is said to be a glass when the diffusion coefficient is lower than a threshold value after some measurable time.

Another feature of glass formation is the presence of dynamic heterogeneity (see the example in Figure 3). A consequence of this dynamic heterogeneity is the breakdown of the Stokes-Einstein relation which relates the diffusion D to the viscosity η, for example via

$$D = \frac{k_B T}{6\pi\eta r} \tag{3}$$

for spherical particles, with k_B the Boltzmann constant, T the temperature and r the radius of a particle. Diffusion and viscosity are two ways to measure dynamics in a system and a breakdown between this relation implies that the two measures do not give similar results anymore. This breakdown can happen because there are different time scales at play around the glass transition.

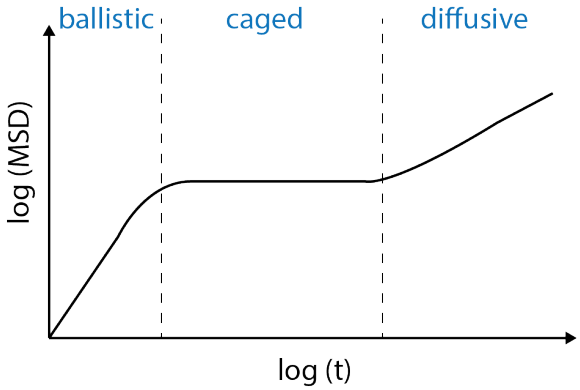

Figure 11: Typical mean squared displacement of a glass.

Barely anything happens microscopically at the glass transition and the structure of a liquid looks similar upon supercooling and yet, the glass transition is accompanied with huge changes in the dynamics and macroscopical quantities such as the viscosity and the MSD. Understanding the origin of the glass transition is a difficult question and has been listed as one of the 125 unanswered questions in science in 2015 [45]. Different theories have been proposed to describe the behavior of materials near the glass transition of which some are more successful and promising than others. The only first principle theory, meaning that it starts from established physics without using empirical models or fitting parameters, is mode coupling theory. Because the version of the mode coupling theory that we will apply has a static correlation function as input and a dynamic correlation as output, we need to define these correlation functions first. Because these correlation functions are also of interest on its own, we explain them in detail in the next subsection.

2.2 Static and dynamic correlation functions

Starting from the basis definition of a correlation function, we derive four different correlation functions based on the density, which are either in real space or Fourier space and either static or dynamic. These four functions are called the radial distribution function, the Van Hove distribution Function, the static structure factor and the intermediate scattering function. In Appendix B we give the full derivations of these functions based on Refs. [46–48] together with own interpretations and we summarize the main equations here. We start with a formal description of the correlation functions in Subsection 2.2.1 and then explain how these functions can be measured using simulated or experimental data in Subsection 2.2.2.

2.2.1 Correlation functions in theory

In a system with N particles we define the dynamic variable $A(\boldsymbol{r}, \boldsymbol{R}(t))$, with Fourier transform $\hat{A}_{\boldsymbol{k}}(t)$, dependent on position $\boldsymbol{r}$ and the particles' coordinates given by $\boldsymbol{R}(t) \equiv \{\boldsymbol{R}_1(t), \boldsymbol{R}_2(t), ..., \boldsymbol{R}_N(t)\}$, where $\boldsymbol{R}_i(t)$ is the position of particle i at time t. The fluctuation $\delta A(\boldsymbol{r}, \boldsymbol{R}(t))$ of this variable is given by

$$\delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \equiv A\left(\boldsymbol{r}, \boldsymbol{R}(t)\right) - \langle A(\boldsymbol{r}, \boldsymbol{R}(t))\rangle\,, \tag{4}$$

and the Fourier transform $\delta\hat{A}_{\boldsymbol{k}}(t)$ is

$$\delta\hat{A}_{\boldsymbol{k}}(t) = \int \mathrm{d}\boldsymbol{r}\delta A(\boldsymbol{r}, \boldsymbol{R}(t))e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) - \left\langle \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t))\right\rangle. \tag{5}$$

$\langle ...\rangle$ is the ensemble average given by

$$\langle ...\rangle \equiv \int \mathrm{d}\boldsymbol{\Gamma} f_{\mathrm{eq}}(\boldsymbol{\Gamma})..., \tag{6}$$

where $\boldsymbol{\Gamma}$ is a $4N$-dimensional phase space coordinate $\boldsymbol{\Gamma} \equiv \{\boldsymbol{R}_1, ..., \boldsymbol{R}_N, \boldsymbol{P}_1, ..., \boldsymbol{P}_N\} \equiv \{\boldsymbol{R}, \boldsymbol{P}\}$ in a two- dimensional system with $\boldsymbol{P}_i$ the momentum of particle i and $f_{\mathrm{eq}}(\boldsymbol{\Gamma})$ the equilibrium probability distribution given in a canonical ensemble by

$$f_{\text{eq}}(\mathbf{\Gamma}) \equiv \frac{e^{-\beta H(\mathbf{\Gamma})}}{\int \mathrm{d}\mathbf{\Gamma}' e^{-\beta H(\mathbf{\Gamma}')}}, \tag{7}$$

with $\beta = 1/(k_B T)$ and $H(\mathbf{\Gamma})$ the Hamilton function. The correlation function $C(\Delta \boldsymbol{r}, \tau)$ of $A(\boldsymbol{r}, \boldsymbol{R}(t))$ in real space for a system that is translation and time-invariant is given by

$$C(\Delta \boldsymbol{r}, \tau) = \langle \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \delta A(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \rangle \tag{8}$$

and the correlation function $C(\boldsymbol{k}, \tau)$ of $\hat{A}_{\boldsymbol{k}}(t)$ in Fourier space by

$$C(\boldsymbol{k}, \tau) = \frac{1}{V} \left\langle \delta A^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) \delta A_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \right\rangle, \tag{9}$$

with V the two dimensional volume of the system. This function correlates the variable A at two points Δr apart and at two times τ apart. We choose the variable $A(\boldsymbol{r}, \boldsymbol{R}(t))$ to be the local density function $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ defined by (with δ the Dirac delta-function):

$$\rho(\boldsymbol{r}, \boldsymbol{R}(t)) \equiv \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)), \tag{10}$$

and in Fourier space by

$$\hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \int \mathrm{d}\boldsymbol{r} \rho(\boldsymbol{r}, \boldsymbol{R}(t)) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \sum_{i=1}^{N} e^{i\boldsymbol{k}\cdot\boldsymbol{R}_i(t)}. \tag{11}$$

Equations 4 -11 are the ingredients we need to determine the four correlation functions we are interested in, namely the Van Hove distribution function, the radial distribution function, the intermediate scattering function and the static structure factor.

The Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$, also called the dynamic density correlaton function, is defined as the correlation function (Equation 8) of the local density function (Equation 10) in real space divided by the average density ρ_0 for normalization purposes, which results in

$$G(\Delta \boldsymbol{r}, \tau) = \frac{1}{\rho_0} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j=1}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t+\tau)) \right\rangle - \rho_0. \tag{12}$$

The radial distribution function $g(\Delta \boldsymbol{r})$ is the static variant of the Van Hove distribution function divided by ρ_0 for normalization purposes and with the trivial contribution of an ideal gas subtracted:

$$g(\Delta \boldsymbol{r}) = \frac{1}{\rho_0^2} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j \neq i}^{N} \delta(\boldsymbol{r} + + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t) \right\rangle. \tag{13}$$

The intermediate scattering function $F(\Delta \boldsymbol{r}, \tau)$, also called the dynamic correlation function in Fourier space, is defined as the correlation function (Equation 9) for the local density function (Equation 11) in Fourier space divided by ρ_0 for normalization purposes:

$$F(\boldsymbol{k}, \tau) \equiv \frac{1}{\rho_0} \frac{1}{V} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t))} \right\rangle. \tag{14}$$

The static structure factor $S(\boldsymbol{k})$ is the static variant of the intermediate scattering function and is given by

$$S(\boldsymbol{k}) = \frac{1}{\rho_0} \frac{1}{V} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))} \right\rangle. \tag{15}$$

2.2.2 Correlation functions in measurements

To calculate the four correlation functions using measured data obtained from simulations and experiments, we need to define a specific implementation of the ensemble average $\langle ... \rangle$ of Equation 6. In theory, $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ is known for the complete phase space $\boldsymbol{R}$ at time t and for every position $\boldsymbol{r}$. While in measurements the value of $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ is only known at points $\boldsymbol{r}$ where and at times t when measured. Therefore we average over N^{ss} different snapshots m in time and over the complete volume of the snapshot to calculate the ensemble average:

$$\langle \quad \rangle \equiv \left\langle \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \right\rangle_t \equiv \frac{1}{N^{ss}} \sum_{m=1}^{N^{ss}} \frac{1}{V} \int \mathrm{d}\boldsymbol{r}, \tag{16}$$

where $\langle ... \rangle_t = \frac{1}{N^{ss}} \sum_{m=1}^{N^{ss}}$. Applying this implementation of the bracket function to the radial distribution function $g(\Delta\boldsymbol{r})$ of Equation 13 gives

$$g(\Delta\boldsymbol{r}) = \frac{1}{\rho_0 N} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(\Delta\boldsymbol{r} - (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))) \right\rangle_t . \tag{17}$$

Besides $g(\Delta\boldsymbol{r})$ being a correlation function of local densities, Equation 17 suggests another explanation. $g(\Delta\boldsymbol{r})$ is namely an average of N density functions $\rho(\boldsymbol{r})$, one for each particle i as origin, and divided by the result of this average for an ideal gas, namely ρ_0. Thus, the radial distribution function measures how many particles there are on average at vector $\Delta\boldsymbol{r}$ from a particle with respect to this number in an ideal gas and is therefore proportional to the probability $p(\Delta\boldsymbol{r})$ of finding a particle at position $\Delta\boldsymbol{r}$ given that there is a particle at the origin.

When the system is isotropic we can change from the vector $\Delta\boldsymbol{r}$ to its magnitude $|\Delta\boldsymbol{r}| = \Delta r$:

$$g(\Delta r) = \frac{1}{\rho_0} \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(\Delta r - R_{ij}(t)) \right\rangle_t , \tag{18}$$

with $R_{ij}(t) = |\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t)|$. The average number of particles $\mathrm{d}n$ in a shell of thickness $\mathrm{d}r$ at distance r from a particle is given by $\mathrm{d}n(\Delta r) = \rho_0 g(\Delta r) \mathrm{d}^2(\Delta r)$.

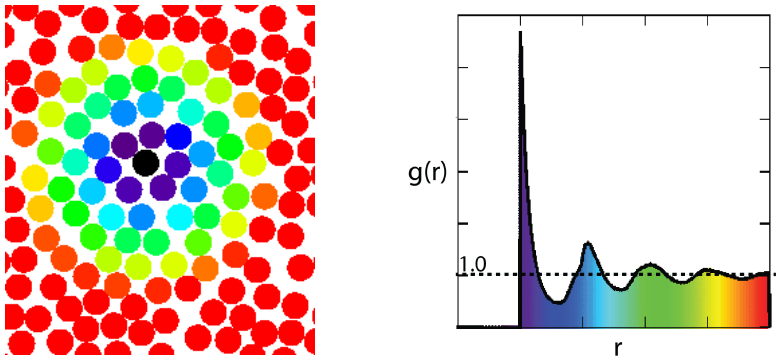

Figure 12: A typical structure of a fluid (left panel) with its corresponding radial distribution function (right panel), where the colors display the contribution of each particle to the radial distribution function (source: http://www.physics.emory.edu/faculty/weeks//idl/gofr.html, Acces date: 07-05-2018).

An example of a radial distribution of a liquid is given in Figure 12, with in the left panel a typical structure and in the right panel the corresponding radial distribution function. $g(\Delta r) = 0$ means it is impossible to find a particle at distance Δr from another particle. In this example particles have a finite size and cannot overlap with each other, making $g(\Delta r) = 0$ for small Δr. $g(\Delta r) = 1$ means that the probability of finding a particle at Δr uncorrelated to whether there is a particle at the origin and equal to the probability of finding two particles at distance Δr in an ideal gas. For the liquid in this example particles are uncorrelated on the long-range as $g(\Delta r) = 1$ for large Δr. The radii in the left panel of Figure 12 are given in reduced units of

the molecular diameter, and the peak at one is caused by the first circle of particles that touch the center particle in the right panel of Figure 12. The second peak is caused by the second circle around the target particle, etcetera. Although the particles have a finite size, the radial distribution function is calculated for the center of masses. Note furthermore that the radial distribution function is not normalized in the sense that $\lim_{\Delta \boldsymbol{r} \to \infty} \int \mathrm{d}(\Delta \boldsymbol{r}) g(\Delta \boldsymbol{r}) = \infty$.

Applying the bracket implementation of Equation 16 to the Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$ in Equation 12 gives for an isotropic system

$$G(\Delta r, \tau) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\Delta r - |\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t)|) \right\rangle_t . \tag{19}$$

The Van Hove distribution function is a dynamic variant of the radial distribution function and is proportional to the probability of finding a particle at position $\Delta \boldsymbol{R}$ at time $t+\tau$ given that there is a particle at the origin at time t. To draw a graph of $G(\Delta r, \tau))$, Δr is usually chosen to be $\overline{\Delta r} = \arg\max(g(\Delta r))$ although other values of Δr are also possible. The graph is then plotted as a function of τ, usually with a log scale on the x - axis. Two examples of the Van Hove distribution function are shown in Figure 13, with faster dynamics in the left panel than in the right panel, which results in a faster decay of $G(\overline{\Delta r}, \tau)$. Normally $\lim_{\tau \to \infty} G(\overline{\Delta r}, \tau) = 0$ and the ergodicity is broken when this is not the case.

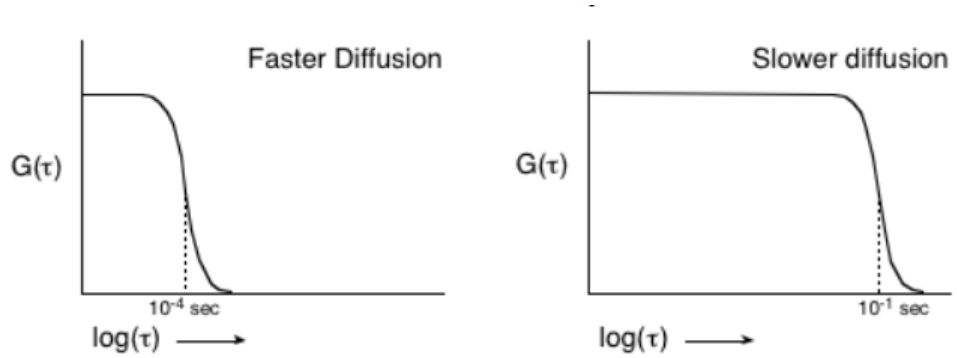

Figure 13: Two examples of the Van Hove distribution function $G(\overline{\Delta r}, \tau)$ with faster dynamics in the left panel than in the right panel.

Applying the bracket implementation of Equation 16 to the static structure factor $S(\boldsymbol{k})$ in Equation 15 gives

$$S(\boldsymbol{k}) = \frac{1}{N} \left\langle \left(\sum_{i=1}^{N} \cos \boldsymbol{k} \cdot \boldsymbol{R}_i(t) \right)^2 + \left(\sum_{i=1}^{N} \sin \boldsymbol{k} \cdot \boldsymbol{R}_i(t) \right)^2 \right\rangle_t , \tag{20}$$

or in the isotropic case

$$S(k) = \frac{1}{N} \left\langle \left(\sum_{i=1}^{N} \cos \boldsymbol{k} \cdot \boldsymbol{R}_i(t) \right)^2 + \left(\sum_{i=1}^{N} \sin \boldsymbol{k} \cdot \boldsymbol{R}_i(t) \right)^2 \right\rangle_t . \tag{21}$$

$S(\boldsymbol{k})$ and $S(k)$ measure how much the structure of the system contains a periodicity matching the wave vector $\boldsymbol{k}$. Another way to calculate the static structure factor is via the Fourier transform of the radial distribution function, which is explained in Appendix B.2.3.

An example of a graph of the static structure factor for a liquid is given in Figure 14. $S(k) > 1$ means that the structural correlations over the inverse length scale k are more present than in an ideal gas and $S(k) < 1$ means that these correlations are less present. In a densely packed system with circularly symmetric hard particles the first peak is around $k = \frac{2\pi}{\sigma}$ (but not exactly) with σ the diameter of the particles because this reflects the inverse length scale at which particles touch each other. The higher order peaks are contributions from waves with smaller wave lengths that also match the periodicity of the system.

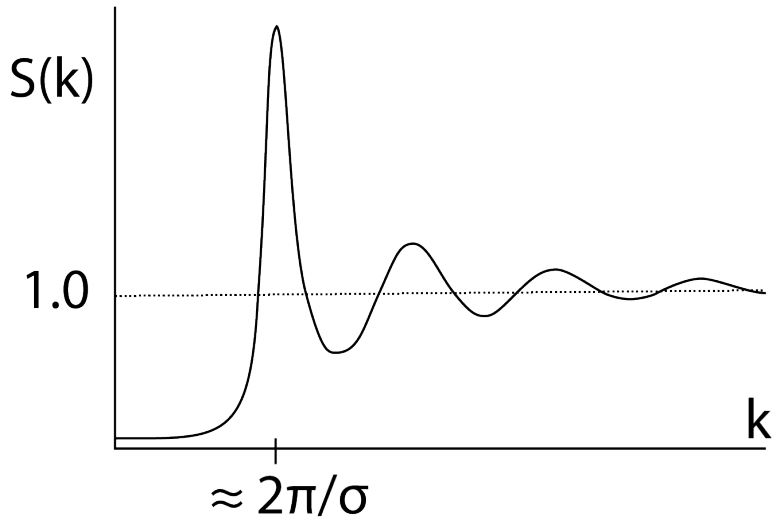

Figure 14: A typical static structure factor of a liquid [49].

Applying the bracket implementation of Equation 16 to the intermediate scattering function $F(\boldsymbol{k}, \tau)$ in Equation 14 gives in an isotropic system

$$F(k,\tau) = \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right) + \left(\sum_{i=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right)\right\rangle_t . \tag{22}$$

The intermediate scattering function measures how much of the structure of a system is still present after some time τ. Another way to calculate the intermediate scattering function is via the Fourier transform of the Van Hove correlation function, which is explained in Appendix B.2.4.

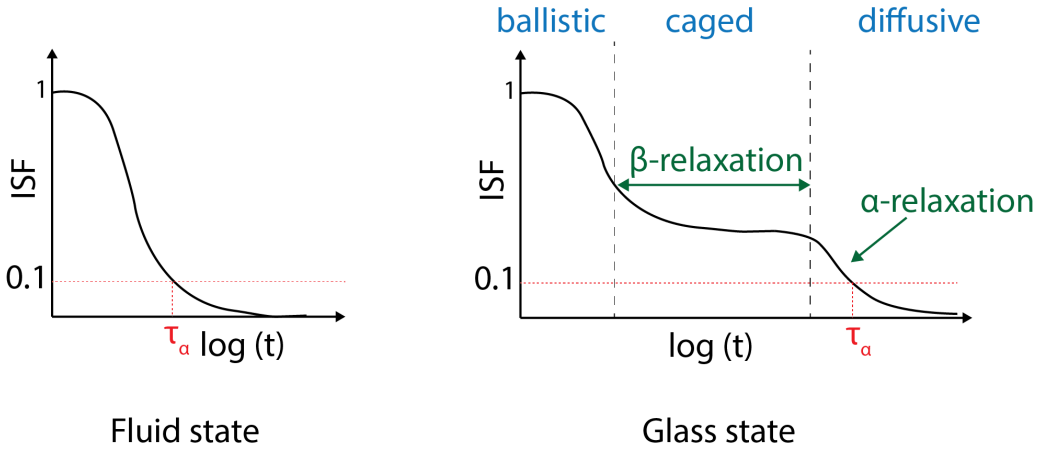

Figure 15: The intermediate scattering function for a fluid state (left) and a glassy state (right), where the β-relaxation and α-relaxation are defined.

Two examples of an intermediate scattering function are given in Figure 15. Normally $F(k,\tau)$ is drawn with k the k-value of the first peak of $S(k)$ (thus for $\overline{k} = \arg\max(S(k))$), with a logarithmic scale for the time τ and divided by $F(k,0)$ such that $F(k,\tau)$ is normalized for $t = 0$. Usually $\lim_{\tau\to\infty} F(\overline{k},\tau) = 0$ and the ergodicity is broken when this is not the case. The intermediate scattering function in the left panel of Figure 15 decreases faster than in the right panel and is typical for a fluid. In the right panel a plateau arises, called the β-relaxation regime. After the plateau, the intermediate scattering function decays further to zero and this is called the α-relaxation regime. This plateau in Figure 15 is the same plateau as we already saw in the mean squared displacement (Figure 11) and is typical for a glass. Thus, the intermediate scattering function can be used to characterize a glassy state and a frequently used criterion is the relaxation time τ_α. This relaxation time is the time it takes until the normalized $F(k,t)$ reaches 0.1 or some other small value. A system is said to be in

a glassy state when this relaxation time is longer than can be measured in simulations or experiments. The value of the intermediate scattering function at the end of the measurement function is called the non-ergodicity parameter. The decay of the intermediate scattering function in the early β-regime, the late β-regime and the α-regime can be fitted to different scaling laws [50].

The intermediate scattering function can be split up in a self $F_s(k,\tau)$ and a distinct part $F_d(k,\tau)$:

$$F(k,\tau) = F_s(k,\tau) + F_d(k,\tau), \tag{23}$$

where the self part is the correlation of the position of a particle at time t with the position of the same particle at time $t+\tau$ and the distinct part is the correlation of the position of a particle at time t with the positions of all other particles at time $t+\tau$. Thus, $F_s(k,\tau)$ is the contribution to Equation 22 of $i=j$ and $F_d(k,\tau)$ the contribution of $i \neq j$. For example, $F_s(k,\tau)$ is defined as

$$F_s(k,\tau) = \frac{1}{N}\left\langle \sum_{i=1}^{N} \left(\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t+\tau))\right) + \sum_{i=1}^{N} \left(\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t+\tau))\right)\right\rangle_t \tag{24}$$

$F_s(k,\tau)$ is also called the incoherent intermediate scattering function and the full version $F(k,\tau)$ the coherent intermediate scattering function.

Summarizing, all correlations functions have the same origin and can be deduced from the Van Hove correlation function, which is a dynamic correlation in real space. Removing the time dependence of the Van Hove correlation function results in the radial distribution function, which is a static correlation in real space. Taking the Fourier transform of the Van Hove correlation function results in the intermediate scattering function, which is a dynamic correlation function in Fourier space. Finally, removing the time dependence in the intermediate scattering function results in the static structure factor, which is a static correlation in Fourier space.

In the remainder of this thesis we switch from Δr to r and from τ to t in the arguments of the correlation functions. Furthermore, we abbreviate the radial distribution function, static structure factor and coherent and incoherent intermediate scattering function respectively with RDF, SSF and coherent and incoherent ISF. The algorithms we have used for solving these correlation functions are explained in Appendix E.2.

Mode coupling theory uses the static correlation function as input and predicts the coherent and incoherent intermediate scattering function. Having derived these correlation functions in this section, we are ready to discuss mode coupling theory in the next subsection.

2.3 Mode coupling theory

Mode coupling theory (MCT) of the glass transition was first derived in 1984 by Leutheusser [51] and Bengtzelius *et al.* [52]. Starting from a microscopic description, mode coupling theory ends up with a complete description of the dynamics of a supercooled liquid and is therefore a powerful tool for the interpretation an prediction of the behavior of glass forming liquids. The original derivation of mode coupling theory was done for glass-forming fluids via the Mori-Zwanzig projection operator. Later, mode coupling theory was also derived for Brownian dynamics via a Smoluchowski operator and the MCT-equation turned out to be similar [53]. We give the derivation for glass-forming fluids in a two dimensional system based on Reichman and Charbonneau [49] in Appendix C.1 and summarize the main equations and approximations in Subsection 2.3.1. In Subsection 2.3.2 we describe how the MCT-equation changes for a Brownian system as well as some other possible modification to the MCT-equations. In Subsection 2.3.3 we explain some of the successes and failures of mode coupling theory.

2.3.1 Derivation of standard MCT-equations

The Newtonian dynamics of any set $\boldsymbol{A}(\boldsymbol{r}_1(t),...,\boldsymbol{r}_N(t),\boldsymbol{p}_1(t),...,\boldsymbol{p}_N(t))$ of phase space variables, in short written as $\boldsymbol{A}(t)$, dependent on the positions $\boldsymbol{r}_i(t)$ and momenta $\boldsymbol{p}_i(t)$ of all N particles i at time t in a classical system is given by

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = \{\mathcal{H},\boldsymbol{A}\} = i\mathcal{L}\boldsymbol{A}(t), \tag{25}$$

where $\mathcal{H}$ is a classical Hamiltonian, $\{...\}$ are Poisson brackets and $\mathcal{L}$ is the Liouvillian. The Liouville equation is based on the principle that a phase-space distribution function is constant along the trajectories of the system and $\mathcal{L}$ is given by

$$i\mathcal{L} \equiv -\sum \frac{\partial}{\partial \boldsymbol{r}_i}\frac{\partial \boldsymbol{r}_i}{\partial t} + \frac{\partial}{\partial \boldsymbol{p}_i}\frac{\partial \boldsymbol{p}_i}{\partial t}. \tag{26}$$

The first step in the derivation of mode coupling theory is to apply the Mori-Zwanzig formalism [54], which is a widely used formalism to rewrite an equation of motion based on a separation of time scales. The reason we apply this principle for supercooled liquids is that they exhibit fluctuations on a microscopic time scale while other dynamical processes such as the relaxation occur on a time scale that can be many magnitudes of orders larger. We are only interested in these so-called slow variables and we will assume that fast variables influence these slow variables only via a fluctuation force. But before applying this assumption, we first split the slow variables from the fast variables with the Mori-Zwanzig principle, which results in still an exact equation of motion. To do so, we define a projection operator $\mathcal{P}$ that projects a function f along the direction of the slow variable $\boldsymbol{A}$,

$$\mathcal{P}f \equiv \frac{(\boldsymbol{A}, f)}{(\boldsymbol{A}, \boldsymbol{A})}\boldsymbol{A}, \tag{27}$$

where $(..., ...)$ denotes the ensemble average of a dyadic product. Furthermore, we have adapted the notation $\boldsymbol{A}(\tau) \to \boldsymbol{A}$ and $\boldsymbol{A}(\tau + t) \to \boldsymbol{A}(\tau)$ to distinguish between $\boldsymbol{A}$ at an arbitrary time τ and at a time t later. We rewrite Equation 25 to $\boldsymbol{A}(t) = e^{i\mathcal{L}t}\boldsymbol{A}$ and divide this equation via the projection operator of Equation 27 into parts parallel and orthogonal to $\boldsymbol{A}$. After some calculations this results in

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = i\boldsymbol{\Omega}\cdot\boldsymbol{A}(t) - \int_0^t \mathrm{d}\tau \boldsymbol{M}(\tau)\cdot\boldsymbol{A}(t-\tau) + \boldsymbol{f}(t), \tag{28}$$

where $i\boldsymbol{\Omega} = (\boldsymbol{A}, i\mathcal{L}\boldsymbol{A})\cdot(\boldsymbol{A}, \boldsymbol{A})^{-1}$ is a so-called frequency matrix and $\boldsymbol{f}(t) = e^{i(1-\mathcal{P})\mathcal{L}t}i(1-\mathcal{P})\mathcal{L}\boldsymbol{A}$ is the fluctuation force orthogonal to $\boldsymbol{A}$. Furthermore, $\boldsymbol{M}(t) = (\boldsymbol{f}, \boldsymbol{f}(t))\cdot(\boldsymbol{A}, \boldsymbol{A})^{-1}$ is called the memory function. Equation 28 is called a generalized Langevin equation. The correlation matrix of $\boldsymbol{A}(t)$ is given by $\boldsymbol{C}(t) = \langle\delta\boldsymbol{A}^*\delta\boldsymbol{A}(t)\rangle$ (see Subsection 2.2) and using Equation 28 gives for the equation of motion for $\boldsymbol{C}(t)$ in Fourier space

$$\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t} = i\boldsymbol{\Omega}\cdot\boldsymbol{C}(t) - \int_0^t \mathrm{d}\tau \boldsymbol{M}(\tau)\cdot\boldsymbol{C}(t-\tau), \tag{29}$$

where the last term of Equation 28 containing fluctuation noise disappears because $\boldsymbol{f}(t)$ is orthogonal to $\boldsymbol{A}$. Note that Equation 29 is still exact as there are no assumptions made yet. However, the memory kernel $\boldsymbol{M}(t)$ cannot be determined exactly and therefore has to be approximated.

To obtain the equation of motion for the intermediate scattering function $F(\boldsymbol{k}, t)$, we use for the specific implementation of $\delta\boldsymbol{A}$ the density fluctuations in Fourier space and its derivative, the longitudinal density current. This implementation is chosen because $F(\boldsymbol{k}, t)$ depends on these density fluctuations as shown by Equation 22. Thus,

$$\delta\boldsymbol{A}(t) = \begin{bmatrix} \sum_i e^{\boldsymbol{k}\cdot\boldsymbol{r}_i} - (2\pi)^2\rho\delta(\boldsymbol{k}) \\ \frac{1}{m}\sum_i(\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \end{bmatrix} \equiv \begin{bmatrix} \delta\rho_{\boldsymbol{k}} \\ j^L_{\boldsymbol{k}} \end{bmatrix}, \tag{30}$$

where m is the mass. Working out Equation 29 with the specific implementation of $\boldsymbol{A}(t)$ given by Equation 30 we obtain for the exact equation of motion of the intermediate scattering function

$$\frac{\mathrm{d}^2F(\boldsymbol{k}, t)}{\mathrm{d}t^2} + \frac{k^2k_BT}{mS(\boldsymbol{k})}F(\boldsymbol{k}, t) + \int_0^t \mathrm{d}\tau M(\boldsymbol{k}, t-\tau)\frac{\mathrm{d}F(\boldsymbol{k}, \tau)}{\mathrm{d}\tau}, \tag{31}$$

with k_B the Boltzmann constant, T the temperature and $M(\boldsymbol{k}, t)$ a memory term given by $\frac{m\langle R_{-\boldsymbol{k}}, R_{\boldsymbol{k}}(t)\rangle}{Nk_BT}$ with $R_{\boldsymbol{k}}(t) = \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} - i\frac{kk_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}}$. This memory term $M(\boldsymbol{k}, t)$ cannot be solved exactly and we therefore rewrite this memory kernel using four approximations. We split $M(\boldsymbol{k}, t)$ into a contribution $M^0(\boldsymbol{k}, t)$ at short times and a contribution at long times, which we call $M^{\mathrm{MCT}}(\boldsymbol{k}, t)$. The first approximation to rewrite the short term contribution $M^0(\boldsymbol{k}, t)$ using a δ-function at the origin: $M^0(\boldsymbol{k}, t) \approx A\delta(t)$, with A a constant. The reason for this

approximation is that we are only interested in the long-term dynamics. From now on we focus on the so-called mode coupling memory kernel $M^{\mathrm{MCT}}(\boldsymbol{k},t)$ and it can be derived that this depends to leading order on a pair of density functions (see Appendix C.1). Therefore, the second approximation is to project the random force $R_{\boldsymbol{k}}$ onto these pair densities, which results in the following dependence relation for the memory kernel

$$M(\boldsymbol{k},t) \sim \sum_{\boldsymbol{q},\boldsymbol{q}'} |V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2 \left\langle \delta\rho_{-\boldsymbol{q}'}(0)\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}(0)\delta\rho_{-\boldsymbol{q}}(t)\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\right\rangle , \tag{32}$$

with $\boldsymbol{q}$ a wave vector and V a so-called vertex that we will define later. This equation depends on an ensemble average of the product of four densities, which is unknown. Therefore, the third approximation is the factorization of this four-point density correlation function into a product of two-point densities:

$$\left\langle \delta\rho_{-\boldsymbol{q}'}(0)\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}(0)\delta\rho_{-\boldsymbol{q}}(t)\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\right\rangle \approx 2\left\langle \delta\rho_{-\boldsymbol{q}}(0)\delta\rho_{\boldsymbol{q}}(t)\right\rangle \left\langle \delta\rho_{\boldsymbol{q}-\boldsymbol{k}}(0)\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}(t)\right\rangle , \tag{33}$$

in which we recognize in the last step the definitions of $F(\boldsymbol{k},t)$. This third approximation is uncontrolled, which means that there is no a priori reason why this approximation might be applied. The last approximation that is needed to calculate the memory kernel is rewriting $\langle\delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}\delta\rho_{\boldsymbol{k}}\rangle$ by applying the convolution approximation to [55]

$$\langle\delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}\delta\rho_{\boldsymbol{k}}\rangle \approx NS(\boldsymbol{k})S(\boldsymbol{q})S(\boldsymbol{q}-\boldsymbol{k}). \tag{34}$$

Finally, this leads to the mode coupling theory equation

$$\frac{\mathrm{d}^2F(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{k^2k_BT}{mS(\boldsymbol{k})}F(\boldsymbol{k},t) + \int_0^t \mathrm{d}\tau\left(M^0(\boldsymbol{k},t-\tau) + M^{\mathrm{MCT}}(\boldsymbol{k},t-\tau)\right)\frac{\mathrm{d}F(\boldsymbol{k},\tau)}{\mathrm{d}\tau}, \tag{35}$$

with the MCT-part of the memory kernel, $M^{\mathrm{MCT}}(\boldsymbol{k},t)$, given by

$$M^{\mathrm{MCT}}(\boldsymbol{k},t) = \frac{\rho k_BT}{8\pi^2m}\int \mathrm{d}\boldsymbol{q}|\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2F(\boldsymbol{q},t)F(\boldsymbol{q}-\boldsymbol{k},t), \tag{36}$$

with ρ the number density and the vertices $\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}$ given by

$$\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}} = (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(\boldsymbol{q}) + \hat{\boldsymbol{k}}\cdot(\boldsymbol{k}-\boldsymbol{q}))c(\boldsymbol{k}-\boldsymbol{q}), \tag{37}$$

with the so-called direct correlation function defined as

$$c(\boldsymbol{k}) = \frac{1}{\rho}\left(1 - \frac{1}{S(\boldsymbol{k})}\right). \tag{38}$$

The mode coupling theory of Equation 35 is a self-consistent second order integro-differential equation for the intermediate scattering function and consists of an infinite amount of coupled equations, one for every possible $\boldsymbol{k}$. Each equation is coupled to all other equations via the vertices and this gives rise to the name mode coupling theory. The MCT-equation has the form of the equation of motion of a damped harmonic oscillator, which is given by $\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \omega^2x + 2\xi\omega\frac{\mathrm{d}x}{\mathrm{d}t} = 0$, with x the position, ω the frequency of the corresponding undamped oscillator and ξ the damping coefficient. This damping coefficient appears in the MCT-equation via the memory kernel and this memory kernel can therefore be seen as a time dependent damping, which takes into account the history of the system.

The effect of the memory kernel is a nonlinear feedback mechanism that accounts for caging of the particles as can be understand as follows. The dominant contributions in the MCT-equation come from the first peak in $S(\boldsymbol{k})$. When $S(\boldsymbol{k})$ is stronger peaked, the vertices given in Equation 37 are bigger. This results in an increase in $M(\boldsymbol{k},t)$, which increases the damping and slows down the decay of $F(\boldsymbol{k},t)$. This influences the memory kernel again because it depends on $F(\boldsymbol{k},t)$ and this again influences $M(\boldsymbol{k},t)$, etcetera. This non-linear feedback mechanism leads to strong dependence of the relaxation time on the static structure factor and can even cause non-ergodicity breaking.

A simple interpretation of mode coupling theory is that it takes into account how much a particle is caged via the static structure factor, determines which inverse length scales dominate the dynamics and takes these length scales into account via the coupling of these modes. When a particle is highly caged, the dynamics only depends on the nearest neighbors and is hardly influenced by other particles far away. The system is near a glass transition. However, when a particle is barely caged the dynamics is also influenced by particles far away and the dynamics also depend on these long length scales.

A last important note about MCT is that it does not depend on the interaction potential. The specific form of the potential is not taken into account in the derivation of MCT in Appendix C.1 and logically does not appear in the MCT-equation given by Equation 35. The effect of the potential is, simply speaking, hidden in the static structure factor. This makes it possible to apply mode coupling theory on many different systems. Nevertheless, due to the approximations, and especially the third approximation in which we factorize a four-point density correlation function into a product of two-point density correlation functions, the applicability of mode coupling theory is never guaranteed.

2.3.2 Different versions of MCT

In the previous subsection we derived a mode coupling theory for glass-forming, Newtonian fluids, but it is possible to extend the theory to Brownian particles. This is done by Szamel and Löwen [56], based on the memory function described by B. Cichocki and Hess [57] in which the equation of motion is governed by the Smoluchowksi $\mathcal{S}$ rather than the Liouvillian operator of Equation 25. This derivation starts with the Smoluchowski equation for the probability density $P(\mathbf{\Gamma}, t)$ (with $\mathbf{\Gamma} \equiv \{\boldsymbol{r}_1, ..., \boldsymbol{r}_N, \boldsymbol{p}_1, ..., \boldsymbol{p}_N\}$ the phase space configuration) for interacting Brownian particles,

$$\frac{\partial P(\mathbf{\Gamma}, t)}{\partial t} = \mathcal{S} P(\mathbf{\Gamma}, t), \tag{39}$$

where $\mathcal{S}$ is the Smoluchowski operator of which we omit the definition here. Using this Smoluchowski operator, the intermediate scattering function can be written as

$$F(\boldsymbol{k}, t) = \frac{1}{N} \left\langle \hat{\rho}_{-\boldsymbol{k}} e^{\mathcal{S}t} \hat{\rho}_{\boldsymbol{k}} \right\rangle. \tag{40}$$

From this equation the MCT-equation for Brownian particles is derived applying the same four approximations as used in Subsection 2.3.1 [56]. Furthermore, we assume an isotropic system such that we can switch from the vector $\boldsymbol{k}$ to its magnitude k. This leads to (with D_0 the self- diffusivity constant)

$$\frac{\mathrm{d}^2 F(k,t)}{\mathrm{d}t^2} + \frac{k^2 D_0}{S(k)} F(k,t) + \int_0^t \mathrm{d}\tau \left(M^0(k, t-\tau) + M^{\mathrm{MCT}}(k, t-\tau) \right) \frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau}, \tag{41}$$

with the MCT-part of the memory kernel, $M^{\mathrm{MCT}}(k,t)$, given by

$$M^{\mathrm{MCT}}(k,t) = \frac{\rho D_0}{8\pi^2} \int \mathrm{d}\boldsymbol{k} |\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2 F(q,t) F(|\boldsymbol{q}-\boldsymbol{k}|, t), \tag{42}$$

with the vertices $\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}$ given by

$$\tilde{V}_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}} = (\hat{\boldsymbol{k}} \cdot \boldsymbol{q}) c(q) + \hat{\boldsymbol{k}} \cdot (\boldsymbol{k}-\boldsymbol{q})) c(|\boldsymbol{k}-\boldsymbol{q}|), \tag{43}$$

and the direct correlation function $c(k)$ given by

$$c(k) = \frac{1}{\rho} \left(1 - \frac{1}{S(k)} \right). \tag{44}$$

Thus besides the adaptation to an istropic system, the only change to the MCT-equation of Subsection 2.3.1 is

$$\frac{k_B T}{m} \to D_0. \tag{45}$$

The MCT-equation given by Equation 41 predicts the behavior of the coherent intermediate scattering function. A similar equation exists for the incoherent intermediate scattering function and in Appendix C.2 is explained what changes in the derivation of the MCT-equation [58]. The resulting MCT-equation for the incoherent ISF is

$$\frac{\mathrm{d}^2 F_s(k,t)}{\mathrm{d}t^2} + k^2 D_0 F_s(k,t) + \int_0^t \mathrm{d}\tau \left(M^0(k,t-\tau) + M_s^{\mathrm{MCT}}(k,t-\tau) \right) \frac{\mathrm{d}F_s(k,\tau)}{\mathrm{d}\tau}, \tag{46}$$

with

$$M_s^{\mathrm{MCT}}(k,t) = \frac{\rho D_0}{8\pi^2} \int \mathrm{d}\boldsymbol{q} |\tilde{V}^s_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2 F(q,t) F_s(|\boldsymbol{q}-\boldsymbol{k}|,t), \tag{47}$$

with

$$\tilde{V}^s_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}} = (\hat{\boldsymbol{k}} \cdot \boldsymbol{q}) c(q), \tag{48}$$

and the direct correlation function $c(q)$ unaltered. The MCT-equation for the incoherent ISF depends on the coherent ISF which makes it necessary to calculate the latter first via MCT.

Equations 41 and 46 define the two versions of the MCT-equation that we use in this thesis and the algorithm to solve them is given in Appendix E.3. Besides these two versions, many modifications to these equations exists. First of all, the full MCT-equation given by Equations 41 and 46 is very complicated due to the coupling of the equations for different $\boldsymbol{k}$ and can only be solved numerically. Therefore, a simplified version is often studied to determine the characteristics of the MCT-equation. This simplified version neglects the $\boldsymbol{k}$-dependence and is called a schematic MCT-equation [59]. Furthermore, MCT-equations are derived for multidisperse mixtures [60, 61] and for specific active particles [62–67]. It is also possible predict the mean squared displacement with MCT instead of the intermediate scattering function [60]. Furthermore, there are methods to improve the MCT-equation by delaying the factorization of the four-point density correlation, which is called generalized MCT [68, 69]. We briefly return to these version of mode coupling theory in Subsections 7.1 and 7.2. In the next subsection we discuss examples of successes and failures of mode coupling theory.

2.3.3 Successes and failures

Because of the (uncontrolled) approximations used in the derivation of mode coupling theory, it is almost a surprise that it gives accurate predictions. A lot of research is done to verify these MCT-predictions or to understand the cases in which the predictions do not hold (see among others Refs. [41, 42, 49, 50, 58–61, 65, 70–77]). In this subsection, we summarize some of these successes and failures of mode coupling theory to illustrate its applicability and shortcomings. Most examples focus on the glass transition with temperature as control parameter and we expect the conclusions to hold for other control parameters as well, although this is still an open question.

Successes

- MCT is able to predict a sharp glass transition, even for small changes in $S(k)$. Figure 16 shows results for the intermediate scattering function from MCT-calculations with a transition to a glassy state at $T = T_c$.
- In predicting the glass transition, MCT accounts for a caging effect which results in a plateau in the intermediate scattering function. In Subsection 2.3.1 is explained that this caging is included via the non-linear feedback mechanism of the memory function and Figure 16 already shows the growth of a plateau. The MCT-predictions for the ISF of Brownian particles at different effective temperatures is given in the right panel of Figure 17, while the left panel shows the corresponding ISFs calculated via Equation 22. Both graphs show similar caging behavior, which verifies the MCT-predictions for this system.
- MCT predicts the scaling relations for the α- and β-regime as given in Figure 15 correctly for temperatures close to the glass transition, analytically for the schematic MCT-equation and numerically for the full MCT-equation [59, 72, 75]. For example, in MCT the stretched exponential of the α-relaxation as observed in experiments and simulations is caused by the fact that each $F(\boldsymbol{k},t)$ decays with a different exponent for different $\boldsymbol{k}$ and all these $F(\boldsymbol{k},t)$ are coupled.

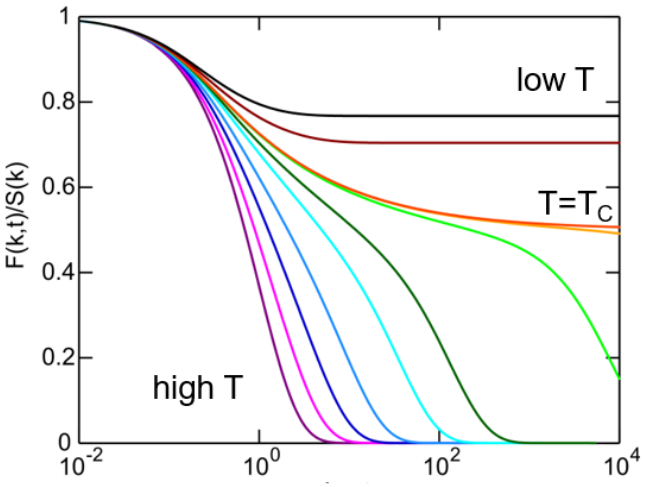

Figure 16: Glass transition predicted by MCT (source: L.M.C. Janssen).

- MCT predicts a time temperature super position (TTSP) in the asymptotic limit $T \to T_c$. This TTSP states that the shape of the intermediate scattering function in the late β-relaxation and early α-relaxation regimes does not depend on temperature. This TTSP is also seen in experiments and simulations, although only for a specific temperature range (see Figure 18).
- Besides predicting a divergence of the relaxation time at the glass transition (see Figure 16), MCT also gives the specific form of this divergence, namely via the power law $\tau_\alpha(T) = C(T - T_c)^{-\gamma}$ with C and γ fitting parameters. This power law is also seen in experiments and simulations close but not to close to the glass transition (see Figure 20).
- MCT predicts the wave vector dependence of the non-ergodicity parameter reasonably well [75, 78]. This is illustrated in Figure 19 where the non-ergodicity parameter as a function of the wave vector q is given for a bidisperse Lennard Jones-model. The particles are labeled A and B, which results in three different combinations for the graphs of the non ergodicity parameter. The direct results from simulations (dots in Figure 19 and the MCT-predictions (dashed line) of the MCT-equation as given in Subsection 2.3.1 are similar (while for the solid line an improved MCT-equation is used).
- MCT can predict non trivial glassy re-entrance effects in specific systems. Examples are a system consisting of hard spheres with short-range attraction [79] or a system with ultrasoft colloids that have a soft repulsive potential [80]. Ultrasoft colloids are in a fluid-like state at low density and a glassy state can be reached via increasing the density, while an even further increase might fluidize the system again. This re-entrance into the glassy state is predicted correctly with MCT, as shown in Figure 21 where the pink line represents a transition from a fluid to a glass state or vice versa in MD-simulations and the orange line in mode coupling theory. Despite the mismatch in height of the lines, the re-entrance is visible in both simulations and the theory.
- Mode coupling theory has turned out to be exact for certain spin-glass models [49, 81, 82].

Failures

- Although MCT predicts a glass transition, the transition temperature T_c is not predicted correctly and is typically 20 to 30 % higher than the real glass temperature T_g in experiments or simulations. Due to the fact that MCT is a mean-field theory and neglects the spatial fluctuations in a system, it overestimates the tendency vitrify. To compare the MCT-predictions with direct calculations of, for example, the relaxation time, one normally scales the temperatures with the corresponding glass transition T_c or T_g. Thus, relaxation times are compared for the same ϵ, with $\epsilon = (T - T_c)/T_c$ in MCT and $\epsilon = (T - T_g)/T_g$ in direct calculations.
- Although MCT predicts a sharp glass transition, this sharp transition normally does not exist in experiments and simulations because hopping processes cause the intermediate scattering function to decay to zero. These hopping processes are not included in the mode coupling theory as derived in Subsection 2.3.1 preventing the ISF to decay to zero. However, there are attempts to incorporate these hopping events in the theory [42].

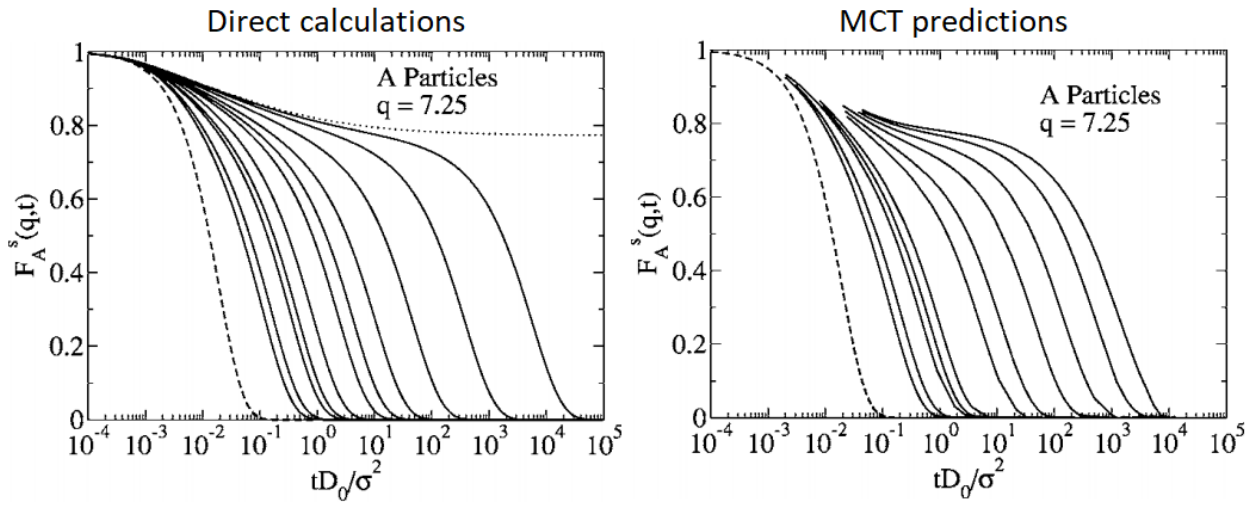

Figure 17: Dynamical slowdown in a glass-forming fluid predicted by MCT [61].

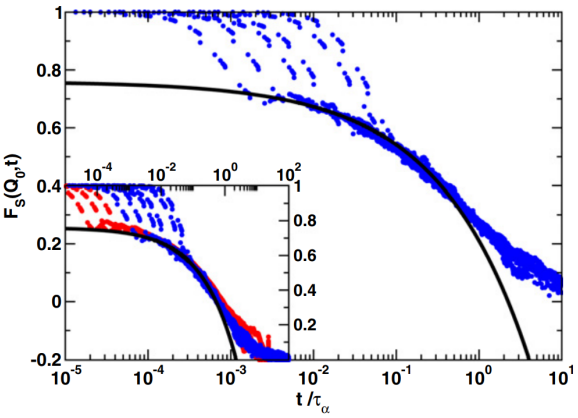

Figure 18: Illustration of the TTSP for a molecular dynamics simulation of supercooled water. The blue lines are ISFs of different temperatures scaled to their relaxation time and they all fall on the same master curve defined by the solid black line [65].

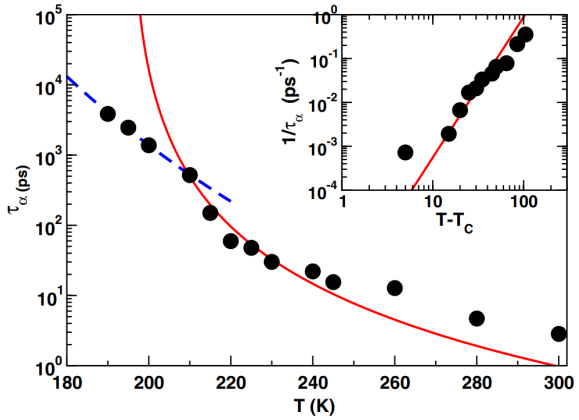

Figure 20: Relaxation times as a function of the temperature for a molecular dynamics simulation of supercooled water. The dots represent relaxation times obtained via direct calculations of the ISF according to Equation 22 and the red line is a fit of the MCT power-law behavior. The inset shows the inverse relaxation time as a function of $T - T_c$ [65].

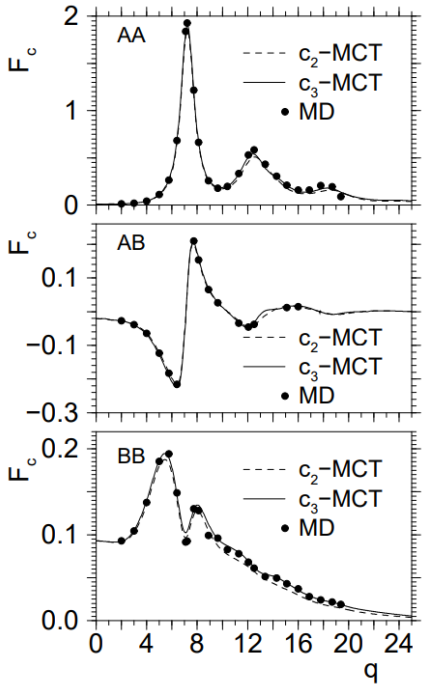

Figure 19: Non-ergodicity parameters for a bi-disperse mixture of Lennard-Jones particles, labeled with A and B [41].

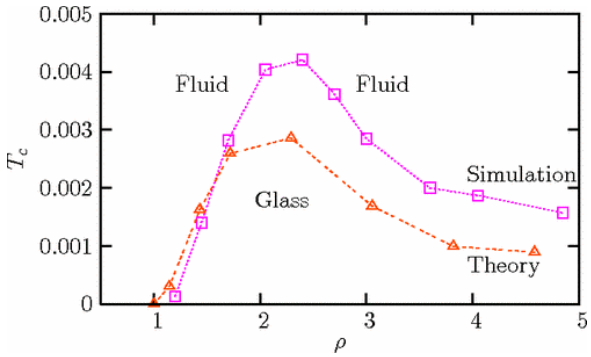

Figure 21: Predictions (orange line) of the re-entrance into a fluid state in a system with ultrasoft colloids and the corresponding glass transition obtained from direct calculations of MD-simulations (pink line) [80].

- The list of successes states that mode coupling theory predicts the scaling laws of for the α- and β-regime correctly and also predicts a power law divergence of the relaxation time. Nevertheless, these predictions are only valid in a certain regime which makes MCT limitedly applicable. Furthermore, the fitting parameters of these scaling laws are predicted to be constant in MCT, while in reality they are temperature dependent.
- Because MCT is a mean-field theory, it does not predict dynamic heterogeneity in supercooled liquids, while this is one of the main characteristics of the glass transition. One of the consequences is that MCT does not predict a breakdown of the Stokes-Einstein relation, which is thought to be caused by this dynamic heterogeneity [49, 83, 84].
- MCT is not able to predict the fragility of a material [77]. This is illustrated by the fact that MCT always predict a power-law divergence of the relaxation time, while strong glass formers exhibit an exponential form (see Figure 10).

This list of successes and failures makes clear that mode coupling theory is not a perfect theory. Nevertheless, there does not exist a perfect theory to describe glassy dynamics and the power of MCT lies in the fact that it is the only first principle-based framework available for glass-forming liquids. Thus, we use MCT but keep in mind the examples of its successes and failures. In the next subsection we discuss how we can transfer the knowledge derived from glassy physics regarding MCT and correlation functions to confluent cell layers.

2.4 Extending the theory to confluent cell layers

We derived the correlation functions in Subsection 2.2 because they serve as a basis for mode coupling theory. Additionally, it is also interesting to investigate these functions on its own, as it is not widely evaluated in confluent cell layers (see examples in Refs. [85, 86]). The terminology of jamming and the jamming transition are not well established yet in biophysics and a profound understanding is necessary to do so. Correlation functions give a rich set of parameters concerning dynamic behavior and can therefore help to establish this jamming and the jamming transition in confluent cell layers.

To the best of our knowledge, mode coupling theory has never been applied to models for confluent cell layers, although closest are the MCT-analyses of the self-propelled particle models in Refs. [63, 67, 87–89]. In the next section we will talk in detail about possible models for confluent cell layers, but the two main differences of these models with respect to the molecular liquids or Brownian particles for which we derived mode coupling theory in respectively Subsections 2.3.1 and 2.3.2 are the potential and the motility or internal activity of the cells. Interestingly, MCT does not depend on the explicit form of the potential and consequently another potential should not make a difference, although the validity of the approximations in the MCT-derivation might change. However, the internal activity is a more complex story. Although active particles still undergo kinetic arrest and caging and the resulting glass transition shares important similarities with its passive counterpart [90, 91], studies have shown that this glass transition in active materials is profoundly affected by the self-propulsion of the particles [87, 91–95]. The question is how to take this activity into account in mode coupling theory and we propose two methods to do so. The first possibility is assuming that the activity-induced change in the static structure factor fully contains the effect of this activity on the dynamics and that therefore the mode coupling theory equations as given in Subsection 2.3.2 stay unaltered. The second option is explicitly incorporating the activity in the derivation of the mode coupling theory which results in an MCT-equation that depends on the form of the activity. This second method is hard to apply to experimental data as the form of the active force is unknown in real cell cultures. Hence, we apply the first method in our research and shortly look into the second method in Subsection 7.1.

In this section we have discussed the principles and theories from glassy physics that we are going to apply on confluent cell layers. In the next section we will return to these confluent cell layers and describe how we can model them.

3 Modelling confluent cell layers

Epithelial cells form confluent cell layers that serve as a barrier between different media or cell types. An example is the skin epithelium that separates a human body from the outer world. Other epithelial cell layers can be found in organs and blood vessels. Confluence literally translates to 'moving together' and in the context of cell layers it means that the complete layer consists of cells without any gaps between them. In this research these confluent epithelial cell layers are analyzed in simulations and experiments. Properties of these cell layers are discussed in Subsection 3.1 and different possibilities to model these cells are described in Subsection 3.2. We use two models in particular and these are described in detail in Subsection 3.3 and 3.4, with the computational details we use in this research in Subsection 3.5.

3.1 Properties of epithelial cells

In Section 1 it is explained that the motion and especially the collective motion of epithelial cells play an important role in biological processes. Cells can crawl over the surface spontaneously or guided by external signals of their neighbors or their environment. Figure 22 gives a simplified procedure of how cells move. Actin is a multi-functional protein that plays an important role in this process. The front edge of a moving cell consists of lamellipodium actin filaments and this is extended by actin polymerization to move. The motor protein myosin causes the other parts of the cell body to retract. Sufficient friction is needed between the cell and its substrate and neighbors for a net movement. Thus the cell's motion depends on a complex force balance between these protrusion, retraction and friction forces. Furthermore, these forces are influenced by biological signals. This crawling of cells and how it is influenced by internal and external factors is an active field of research [[96–98]]. Nevertheless, we only focus on simple coarse-grained models to describe these epithelial cells, which can already characterize the collective cell motility surprisingly well. Before looking into these models, three relevant properties of these epithelial cells needs to be clarified first, namely the representation of the cell layer in a two-dimensional plane, their packing structure and their self propulsion.

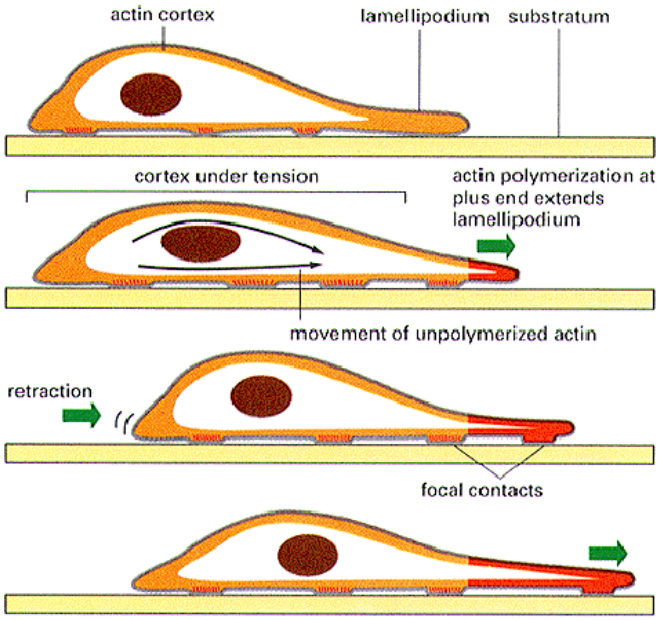

Figure 22: A simplified sketch of the crawling of a cells. Protrusion due to actin polymerization in the direction of the movement and retraction of the other parts of the cell play an important role [99].

Although epithelial cells can form monolayers and multiple layers, we focus only on monolayers. One representation of such a monolayer is given in Figure 23, where the cells have a columnar shape. Most movements in this monolayer take place parallel to the plane and the height of the cells do not variate much. Therefore, we can treat the cell layers as a two-dimensional system.

Epithelial cell layers evolve towards a honeycomb or hexagonal packing structure (see Figure 24) [102], which is the optimal structure of a partition of a plane into cells of equal area regarding the ratio of the perimeter to the squared area [103] . The honeycomb structure consists of equally sized hexagons with this ratio equal to 3.72 (see Appendix D.3). For all other configurations that can tile the plane into equal areas, like pentagrams

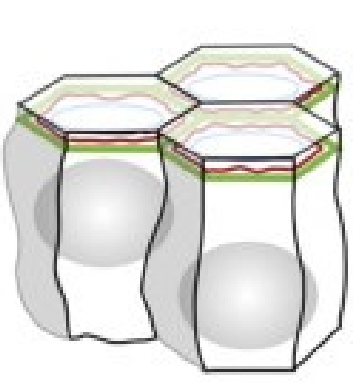

Figure 23: A schematic representation of a columnar epithelial monolayer [100].

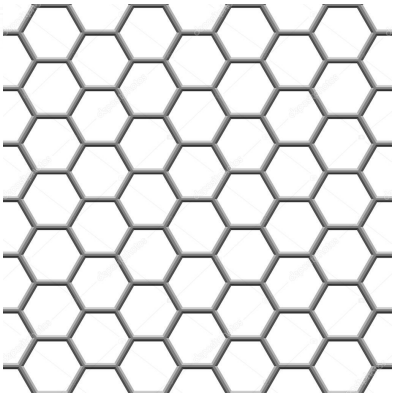

Figure 24: A honeycomb or hexagonal packing structure.

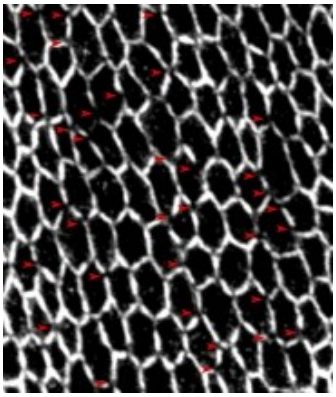

Figure 25: A honeycomb packing structure in the matured Drosophila epithelium [101].

and triangles, this ratio is larger. An example of a honeycomb structure in biology is the matured Drosophila epithelium in Figure 25. However, a layer is often driven away from this optimal packing due to cellular divisions, apoptosis and the activity of the cells [14]. This results in a cellular arrangement consisting of polydisperse polygons, but still with a preference for hexagons [104, 105].

The activity of cells is often represented by a self-propulsion force with a specific polarity. This polarity mimics the spatial asymmetry between the front and rear of the cells and defines the direction of movement (see also Figure 22) [106–108]. Cells consume energy in the form of adenosine triphosphate (ATP) to move and because of this energy consumption cell models belong to active matter physics. This addition of an active force can suppress the glass or jamming transition, as the amount of driving energy in the system increase [94, 95]. But in practice, it turns out to be very difficult to predict whether activity fluidizes or solidifies a system [95]. In the next section, possible models are described that takes into account these properties of confluent cell layers.

3.2 Model classes for confluent cell layers

Among many different models, cellular Potts models, self-propelled particle (SPP) models and vertex models are the most popular classes to describe confluent cell layers [1, 96, 109–111] and in models of all these three classes features of the glass transition have been demonstrated recently [108, 112–115]. The three models are illustrated in Figure 26 and explained below.

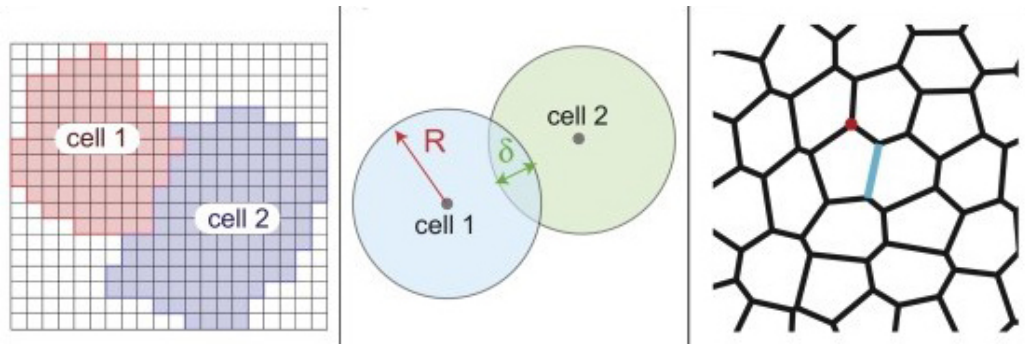

Figure 26: Illustration of three different cell models: the cellular Potts model (left panel), the self-propelled particle model (center panel) and the vertex model (right panel).

Cellular Potts model

The cellular Potts model [107, 116–118] is one of the earliest successful cell models and in this model cells are represented by grid points on a lattice (left panel of Figure 26). Each lattice point has a cell id σ and all points with the same σ define one cell. An example of a coarse-grained energy function in this model is [117]

$$E = \sum_{a,b//a} J_{a,b}(1 - \delta_{\sigma(a),\sigma(b)}) + \sum_i \lambda(A_i - A_{i,0})^2, \tag{49}$$

where lattice points are indexed with a and b and cells with i. The first term counts the interaction energy between neighboring cells, where $b /\!\!/ a$ means that grid point a shares a border with grid point b and $J_{a,b}$ is the interaction strength between those two grid points. The δ-function indicates that two sites belonging to the same cell do not contribute to the interaction energy. The second term of Equation 49 penalizes deformations of the area A_i of cell i from its target area $A_{i,0}$, with a springlike penalty constant λ. In cellular Potts models a Monte-Carlo algorithm is applied to find the configuration that minimizes the total interaction energy of Equation 49. Simply said, the algorithm chooses randomly a lattice point and tries to switch its id σ to the id from a neighboring point. This trial move is accepted with a specific probability depending on the energy difference between the previous and the new configuration; the move is always accepted if this difference is negative. This procedure is repeated until some threshold is reached and the system is in a low energy state. Motility can be included in the cellular Potts model by adding some specific self-propulsion force. A glass transition in the cellular Potts model was found by Chang and Marenduzzo [112], where the control parameters are the interaction energy $J_{a,b}$ and the strength of the self-propulsion.

Self-propelled particle model

In self-propelled particle models [9, 106, 119–123], each cell is represented by its center and some parameters that describe its shape (see center panel of Figure 26). For circular or spherical shaped particles this shape parameter is the radius R and this means that in general the cell layer is not confluent, although there are some extensions to the SPP-model that incorporate confluence. According to Newton's first law, the overdamped equation of motion for cell i is

$$\boldsymbol{F}_i^{\text{friction}} + \boldsymbol{F}_i^{\text{act}} + \boldsymbol{F}_i^{\text{int}} = 0, \tag{50}$$

with $\boldsymbol{F}_i^{\text{friction}}$ the friction force, $\boldsymbol{F}_i^{\text{act}}$ the force due to the self-propulsion of the cells and $\boldsymbol{F}_i^{\text{int}}$ a force due to the interaction between cells. This self-propulsion force can have many forms, but in general contains a stochastic term. A simple version of the friction force is used that is appropriate to low Reynolds number,

$$\boldsymbol{F}_i^{\text{friction}} = -\gamma \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t}, \tag{51}$$

where γ is a friction constant. The interaction force $\boldsymbol{F}_i^{\text{act}}$ between cells depends on an interaction potential $U(\boldsymbol{r}_1, ..., \boldsymbol{r}_N)$ of which the specific from can be chosen. This potential includes a short-range repulsion arising from an aversion of cell-cell overlap mimicking cortical tension and cystoskeletal rigidity and an attraction due to a cell's adhesion to its neighbors. Based on Equation 50, the resulting equation of motion is

$$\gamma \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = -\boldsymbol{\nabla}_{\boldsymbol{r}_i} U + \boldsymbol{F}_i^{\text{act}}, \tag{52}$$

which is a stochastic differential equation and called a Langevin equation. This equation can be solved numerically using molecular dynamics simulations. As Equation 50 is overdamped because of the neglect of the acceleration term, this is often called overdamped Langevin or Brownian dynamics. A glass or jamming transition in the SPP-model can be controlled by changing the packing density or the strength of the self-propulsion [114].

Vertex models

Vertex models [100, 115, 124–130], first used in 1980 [131], describe a confluent tissue as a polygonal tilling of space (see right panel of Figure 26), inspired by the packing properties described in Subsection 3.1. The degrees of freedom in the vertex models are the vertices of the polygons, which allows for complex, non-convex cell shapes. The model uses a coarse-grained version of the interaction energy E, for example,

$$E_i = K_{A_i}(A_i - A_{0,i})^2 + \xi_i P_i^2 + \gamma_i P_i, \tag{53}$$

where A_i and P_i are the area and perimeter of cell i. $A_{0,i}$, K_{A_i}, ξ_i and γ_i are cell properties that take into account the cortical elasticity, cortical surface tension, bulk incompressibility and cell-cell adhesion. The details of this interaction energy are explained below in Subsection 3.3. Equilibrium is based upon a force balance at each vertex. It is possible to solve vertex models using Monte Carlo simulations as described for the cellular Potts model [132], but more often overdamped Langevin dynamics is used similar to the self-propelled particle models. In the SPP-model the equation of motion (Equation 52) is defined for the cell centers, while this equation depends on the positions $\boldsymbol{r}_v$ of the vertices v in the vertex model:

$$\gamma \frac{\mathrm{d}\boldsymbol{r}_v}{\mathrm{d}t} = -\boldsymbol{\nabla}_{\boldsymbol{r}_v} E + \boldsymbol{F}_v^{\text{act}}, \tag{54}$$

where $\boldsymbol{F}_v^{\text{act}}$ is again a self-propulsion force with a stochastic term, but applied on vertex v. The jamming transition of this model is studied by Bi *et al.* [115] and is governed by a parameter that sets the ratio of the preferred cell perimeter to the preferred area.

Cellular Potts models, self-propelled particle models and vertex models all have their pros and cons and which one is preferred depends on the simulation objectives. For example, cellular Potts and vertex models require more computational time than SPP-models, but are better at representing changes of the cell shapes. Vertex models are generally restricted to confluent cell layers, while SPP-models are normally not applicable to confluent cell layers. Potts models can handle both confluent and non-confluent tissues, but makes it more difficult to include a self-propelled force. Because confluence and self-propulsion are two important properties of epithelial cell layers, a vertex models seems the best choice for our research. However, the way vertex models include the active force is not intuitive because the self-propulsion force of a cell is divided over its vertices [128]. A more natural way is a force that acts on the cell centers instead of the vertices, just as in the SPP-model. Bi *et al.* [108] bridge this gap and study a hybrid between the vertex and the SPP-model. This model is called a self-propelled Voronoi model (SPV) because the cell shapes are obtained from Voronoi tesselations.

In the SPV-model the seeds of a Voronoi tesselation, also called Voronoi centers, are the degrees of freedom [133, 134] and the location of the seed of cell i is given by $\boldsymbol{r}_i$. The Voronoi cell belonging to seed i is defined as

$$\mathcal{V}_i = \{\boldsymbol{r} \in \mathcal{R} : \mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_i) \leq \mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_j) \quad \forall j \neq i\}, \tag{55}$$

where $\mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_i) = (\boldsymbol{r} - \boldsymbol{r}_i)^2$ is the squared Euclidean distance between $\boldsymbol{r}$ and $\boldsymbol{r}_i$. This results in a tiling where all points $\boldsymbol{r}$ belonging to Voronoi cell i are closer to seed i than to all other seeds j. A Voronoi tesselation is constructed by drawing perpendicular bisectors of the lines connecting two neighboring seeds. An example of such a tesselation is given in Figure 27, where the seeds are black dots and the Voronoi vertices and edges are respectively the red dots and lines. The black tiling is called a Delaunay triangulation and is the dual graph of the Voronoi tesselation.

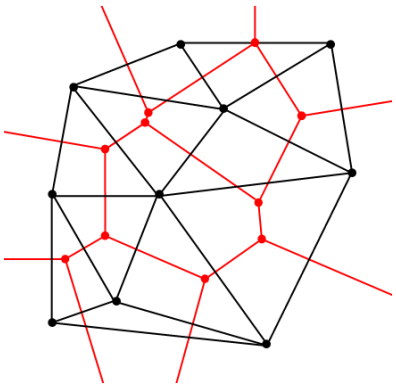

Figure 27: Seeds (black dots) and the corresponding Voronoi tiling (red lines and dots) and Delaunay triangulation (black lines).

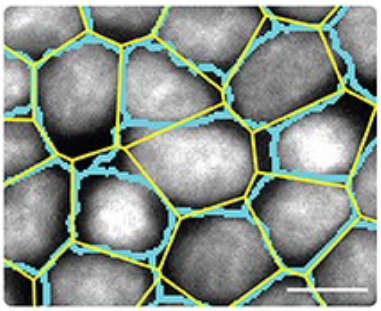

Figure 28: Differences between the cell boundaries of MDCK obtained via imaging methods (blue) and a Voronoi tesselation of the nuclei (yellow) [86].

Representing a cell layer by a Voronoi tiling is first proposed by Honda [135], and later shown to hold in different situations. For example, Kaliman *et al.* [86] show that the cell boundaries of MDCK cells obtained via a Voronoi tesselation of the nuclei only has an error in cell area and perimeter of ten to fifteen percent with respect to boundaries obtained from imaging techniques. This agreement in cell boundaries for MDCK-cells is shown in Figure 28. Due to this success, several cell models exists that implement a Voronoi tesselation [121, 132, 136–138], but we focus only on the SPV-model and possible extensions.

Compared to the original vertex models, using the seeds of a Voronoi tesselation simplifies the process of handling cell dynamics and makes it easier to add an active force. Furthermore, the Voronoi model has fewer degrees of freedom in comparison with the vertex models as shown in Figure 29, where the red dots represent the degrees of freedom. Note furthermore that the variation of cell shapes in the SPV-model is less than in regular vertex models because the resulting structure has to be a Voronoi tesselation which does not exist for most of the configurations of the vertex model. Based on its successes, the SPV-model is now used and extended in several research groups [128, 138–140].

In our research we used two version of the SPV-model, namely the original version and an extension.. We call these models respectively the Voronoi Rotational Diffusion (VRD) model and the Voronoi Vicsek model (VV) for reasons that become clear in the next subsections. We only use the 2D version of these models, but extensions to 3D are possible [141]. We neglect cell proliferation and cell death in our models, although this could be included as well [142]. Furthermore, we assume that each cell is identical regarding cell properties such as the preferred shape, the adhesion with other cells, and the form of the self propulsion. In the next subsection we explain the VRD-model in detail and in Subsection 3.4 we focus on the VV-model.

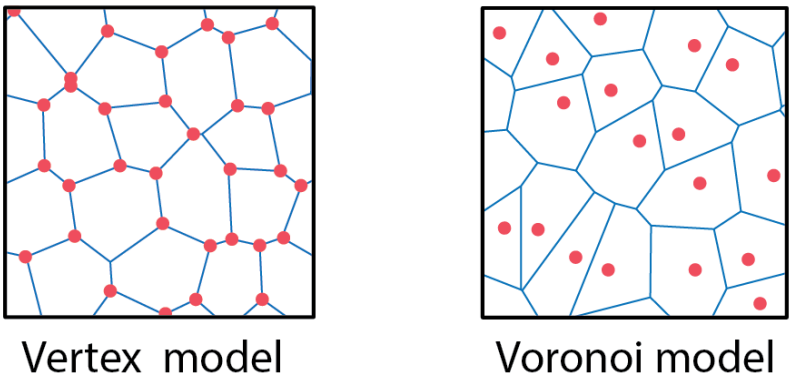

Figure 29: Degrees of freedom (red dots) in a vertex model (left panel) and a Voronoi model (right panel).

3.3 Voronoi Rotational Diffusion model

The VRD-model is derived from the SPV-model originally proposed by Bi *et al.* [108] and we discuss the details of this model in Subsection 3.3.1 and the relevant literature in Subsection 3.3.2.

3.3.1 Description of the VRD-model

The coarse-grained interaction energy in the VRD-model is similar to the energy in regular vertex models given by Equation 53 and is for each cell equal to [14, 108, 115, 121, 128]

$$E_i = K_{A_i}(A_i - A_{0,i})^2 + \xi_i P_i^2 + \gamma_i P_i, \tag{56}$$

where the first term results from a combination of three-dimensional cell incompressibility and the monolayer's resistance to height fluctuations and the cell bulk elasticity. A_i is the actual cell area (light green area in Figure 30), $A_{0,i}$ the preferred cell area and K_A is a penalty constant for deviations from a cell's preferred cell area. The second term in Equation 56, quadratic in the perimeter, models the active contractility of the actin-myosin subcellular cortex that strives to decrease the perimeter P_i (dark green boundary in Figure 30) of the cell with an elastic constant ξ_i. The last term in Equation 56, linear in the perimeter, represents a competition between the cortical tension and the cell-cell adhesion which results in a net line tension given by γ_i. This net line tension γ_i is positive when the cortical tension is larger than the cell-cell adhesion and negative in the opposite case.

We will only use the derivative of the energy with respect to the degrees of freedom and can therefore simplify the energy term of Equation 56. In Appendix D.4 it is shown that the preferred area $A_{0,i}$ does not influence these derivatives [140] and that we can include a preferred perimeter $P_{0,i}$ to rewrite the last two terms of Equation 56. While doing this, we change from ξ_i to $K_{P,i}$ and from $P_{0,i}$ to $-\frac{\gamma_i}{2\xi_i}$. Furthermore, we assume that all cells have identical properties such that $K_{A,i} = K_A$, $K_{P,i} = K_P$ and $P_{0,i} = P_0$. The rewritten interaction energy for a system with N cells is

$$E = \sum_i^N K_A(A_i - A_0)^2 + K_P(P_i - P_0)^2, \tag{57}$$

where $A_0 = A_{\text{tot}}/N$, with A_{tot} the total area of the confluent cell layer. Simply said, each cell strives to obtain a preferred area and perimeter and each deviation from these values results in a quadratic, spring-like, penalty. This energy term in Equation 57 results in an effective mechanical interaction force $\boldsymbol{F}_i^{\text{int}} = -\boldsymbol{\nabla}_i E$.

In addition to this interaction force, we add self-propelled motility to the cells by assigning a polarity vector $\hat{\boldsymbol{n}}_i = (\cos(\theta_i), \sin(\theta_i))$ along which each cell exerts a self-propulsion force with constant magnitude $\frac{v_0}{\mu}$, in which μ is the mobility and equal to the inverse of the friction constant γ and v_0 is a velocity defining the strength of the self-propulsion force. In the VRD-model, the polarity angle θ_i undergoes random rotational diffusion defined by

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \eta_i(t), \quad \text{with} \quad \langle \eta_i(t)\eta_j(t')\rangle = 2D_r\delta(t-t')\delta_{ij}, \tag{58}$$

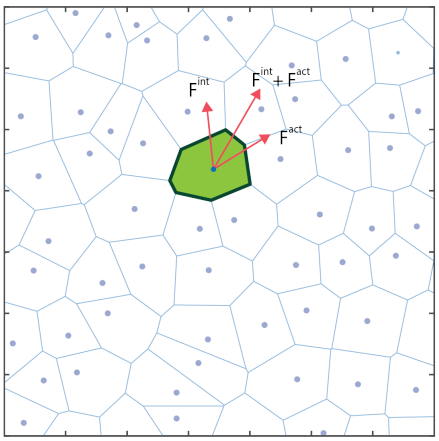

Figure 30: Defining the cell area (light green) and perimeter (dark green). Furthermore, the forces acting on the cell are drawn, consisting of the interaction force F^{int} and an active force F^{act}. The sum of F^{int} and F^{act} defines the direction of movement of the cell.

such that η_i represents white noise with zero mean and variance $2D_r$.

The overdamped equation of motion for regular vertex models is given by Equation 54 and by implementing the specific form of the interaction and self-propulsion force we obtain the equation of motion in the VRD-model:

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \mu \boldsymbol{F}_i^{\mathrm{int}} + v_0 \hat{\boldsymbol{n}}_i. \tag{59}$$

We make this model dimensionless by using $r_0 = 1/\sqrt{A_0}$ as unit of time and $\tau_0 = \frac{1}{\mu K_A A_0}$ as unit of length and we define k_a, a_0, a_i, k_p, p_0 and p_i to be the dimensionless versions of their counterparts written with capitals. Three important independent dimensionless parameters remain in the VRD-model, namely p_0, v_0 and D_r. p_0 determines the preferred shape of the cells and a small value pushes the cells towards hexagons, while a large value makes the cells more elongated. v_0 determines the strength of the active force with respect to the influence of the interaction potential. D_r determines the persistence of the active force and gives rise to a persistence time scale $\frac{1}{D_r}$. For small values of D_r the noise is very small and the persistence time is long, while for large values of D_r the dynamics are governed by simple Brownian motion as the persistence timescale $\frac{1}{D_r}$ is shorter than the timescale on which the interaction potential interacts.

Equations 57 - 59 are solved using overdamped Langevin dynamics with a modified version of *CellGPU* [143]. This algorithm is explained in Appendix E.1 and in Subsection 3.5 we briefly state our simulation details. In the next subsection we will explain some studies about the VRD-model (called the SPV-model in literature) and focus on what is known about the jamming phenomenon in this model.

3.3.2 Jamming of the VRD-model in literature

Bi *et al.* [108] were the first to study the jamming phenomenon in the VRD-model; the mean squared displacement they found for different values of p_0 is given in Figure 31. For short times the slope is close to two on a log-log plot indicating that the cells are free to move and the motion is ballistic. For long times and large p_0 the slope is about one indicating that the motion is diffusive. For lower values of p_0 we see a plateau arising, which is an indication of jamming and happens because cells are caged by their neighbors and their movements are hindered. Bi *et al.* [108] define the jamming transition using a (somewhat arbitrarily) threshold for the diffusion coefficient D_{eff} which is given by the diffusion coefficient D in Equation 2 divided by the diffusion coefficient D_0 of an isolated cell (see Appendix D.5),

$$D_0 = \frac{v_0^2}{2D_r}. \tag{60}$$

This dimensionless effective diffusion D_{eff} is given by

$$D_{\mathrm{eff}} = \lim_{t\to\infty} \frac{2D_r \left\langle \Delta \boldsymbol{r}(t)^2 \right\rangle}{4v_0^2 t}. \tag{61}$$

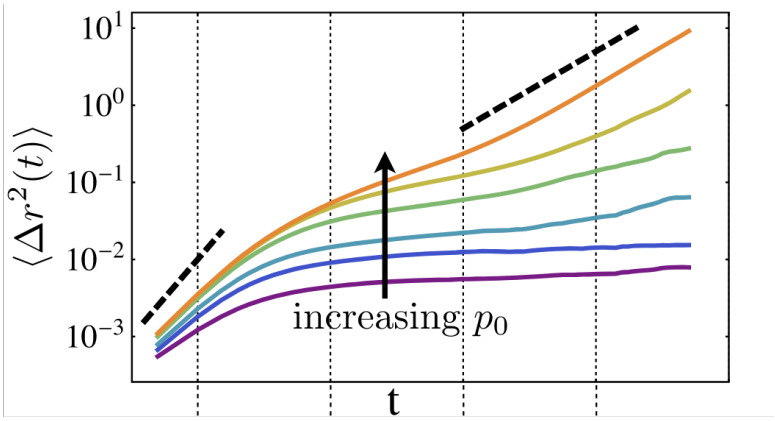

Figure 31: Mean squared displacements as a function of p_0 for the VRD-model. When p_0 decreases a plateau arises, indicating caging and jamming [108].

and the layer is said to be jammed when $D_{\text{eff}} \leq 0.001$. The resulting jamming phase diagram from Bi *et al.* [108] is given in Figure 32. The VRD-model has a jammed state when p_0 is smaller than some threshold. In this case the cortical tension dominates over cell-cell adhesion which makes the energy barriers for local cell rearrangements and motions high. The system behaves like a solid. This critical value of p_0 depends on v_0, because increasing v_0 makes it easier for cells to overcome the energy barriers and move around and this results in a jamming transition at smaller p_0. According to this phase diagram, there are two ways to escape from a jammed state: namely via increasing p_0 or via increasing v_0. Figure 32 shows furthermore examples of the layer structure and the trajectories in an unjammed and a jammed state. In the unjammed state the structure is disordered and the cells have irregular shapes. The cell trajectories show that cells migrate over long distances. In the jammed state cells are forming a very structured and almost hexagonal packing with approximately equal cell shapes. Furthermore, cells travel less and stay around their original position. Besides v_0 and p_0 the jamming transition can also be controlled by the noise-term D_r. This gives rise to the jamming phase diagram in Figure 33. When v_0 is large enough, decreasing D_r can unjam the layer due to a more persistent motion which helps cells to escape from the cage of their neighbors and fluidizes the system.

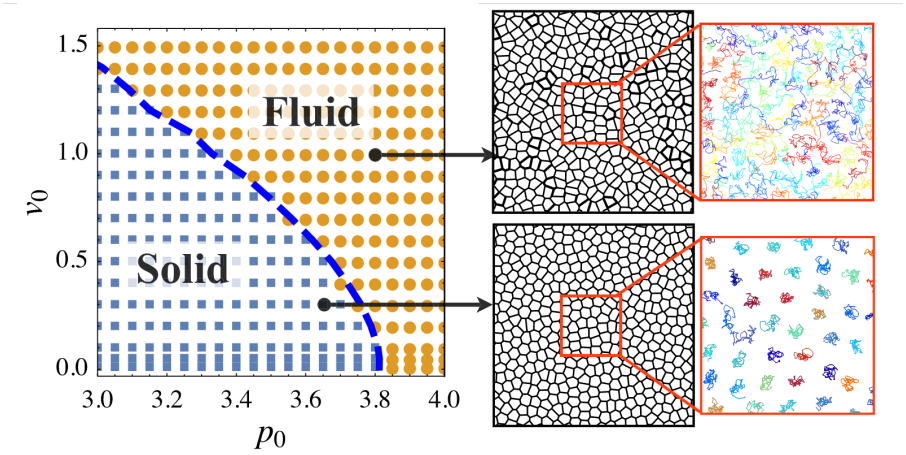

Figure 32: A jamming phase diagram depending on p_0 and v_0. The insets show the structure and trajectories for a jammed and an unjammed state [108].

Strikingly, Bi *et al.* [108] found that the critical value of $D_{\text{eff}} = 0.001$ coincides with a critical value for the average resulting dimensionless perimeter q , given by

$$q = \frac{1}{N} \sum_i p_i. \tag{62}$$

The critical threshold below which a state is jammed is 3.81 [10, 108]. This value is equal to q in a pentagonal packing and just above the value in a hexagonal packing, which is the packing structure that minimizes q (see Subsection 3.1). Figure 32 intuitively shows that the perimeter q is different in a jammed and an unjammed state. The jamming threshold $q = 3.81$ is verified in experiments with HBEC-cells, where a decrease in q towards

3.81 is accompanied by a slowing down of dynamics in the cell layer. Nevertheless, the universality of using q to define jamming remains questionable. For example, Sussman *et al.* [113] show in a slightly different model than the SPV-model that the line $q = 3.81$ does not completely coincide with standard measures to determine the glass or jamming transition.

In the VRD-model, the amount of constraints equals exactly the number of degrees of freedom. Each cell has two constraints, namely the area and the perimeter, and also two degrees of freedom because it can move in two perpendicular directions. Such a marginal model normally does not have a jamming or glass transition and Sussman and Merkel [144] prove that this is indeed not the case. Thus, a jamming or unjamming transition in the SPV-model is only possible in the presence of an active force and therefore we do not investigate the athermal or non-active case in this thesis.

Our analysis of the VRD-model is given in Section 4. In the next subsection we describe our Voronoi Vicsek-model, which is an extension to the VRD-model.

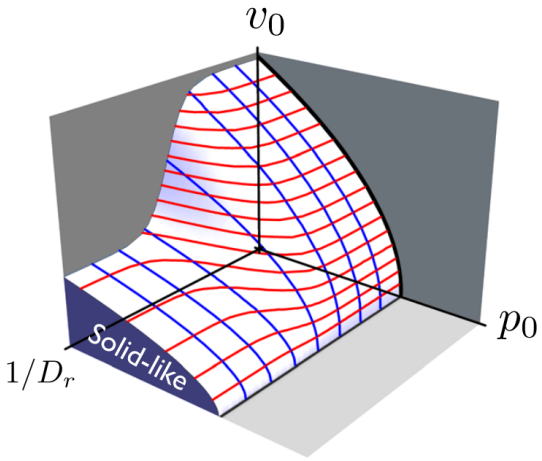

Figure 33: A three-dimensional jamming phase diagram showing three ways to unjam: increasing the velocity v_0, increasing the preferred perimeter p_0 and decreasing the noise D_r [115].

3.4 Voronoi Vicsek model

In the VRD-model the active force undergoes rotational diffusion, but there are many other ways to implement an active force. As far as we know, only two other version are analyzed in literature. One that includes translational noise and one that includes an alignment mechanism of a cell's polarity with its own direction of movement. In this subsection, we add a force that incorporates a Vicsek-like alignment mechanism in which cells align with their neighbors and in Subsection 7.1 we list other ideas.

The Voronoi Vicsek model is motivated by experiments that show that a change in connectivity between cells results in a change in their cooperation [146]. In the VV-model the SPV-model is combined with a Vicsek-like alignment [145, 147]. The Vicsek alignment is first used in the Vicsek model to explain collective motion in phenomena such as the swarming of birds or bacterial colony growth. The original Vicsek model is very basic and only takes into account a self-propulsion force and alignment of this force with its neighbors. How many neighbors are taken into account depends on the Vicsek radius R_V: all neighbors j that are within a distance R_V from particle i contribute to the polarity of particle i. This results in the following equation of motion in the Vicsek model

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = v_0 \hat{\boldsymbol{n}}_i, \tag{63}$$

with v_0 the velocity of a particle and $\hat{\boldsymbol{n}}_i = (\cos(\theta_i), \sin(\theta_i))$ the polarity of a particle satisfying

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \langle \theta_j \rangle_{|\boldsymbol{r}_i - \boldsymbol{r}_j| < R_V} + \eta_i(t), \quad \text{with} \quad \langle \eta_i(t)\eta_j(t') \rangle = 2D_r \delta(t - t')\delta_{ij}, \tag{64}$$

with $\eta_i(t)$ white noise with mean zero and variance $2D_r$ and $\langle \theta_j \rangle_{|\boldsymbol{r}_i - \boldsymbol{r}_j| < R_V}$ the average angle of all neighbors within a radius R_V. Besides R_V, the density is also an independent parameter in the Vicsek model. The

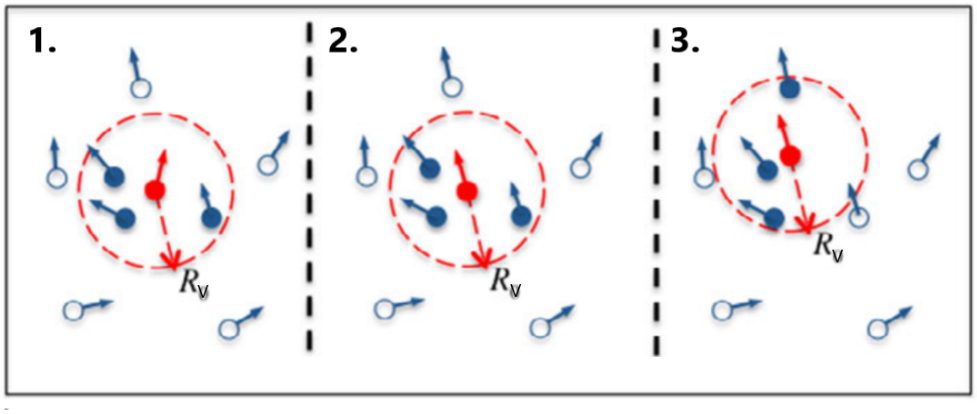

Figure 34: Voronoi alignment in three steps. In the first step the neighbors within R_V of the red particle are determined (blue solid particles). In the second step, the red particle aligns its polarity with these neighbors and some noise is added. In the last step the particle moves in the direction of its polarization, which might give the particle new neighbors[145].

Vicsek-alignment is illustrated in Figure 34. In the second panel, the red particle aligns it motion with its neighbors within R_V (blue solid particles) and some noise is added. In the third panel the red particle moves which results in new neighbors. The original Vicsek model shows two phases, namely a disordered phase for low density, strong noise or small R_V and an ordered phase for high density, low noise level or high R_V. Both phases are shown in Figure 35.

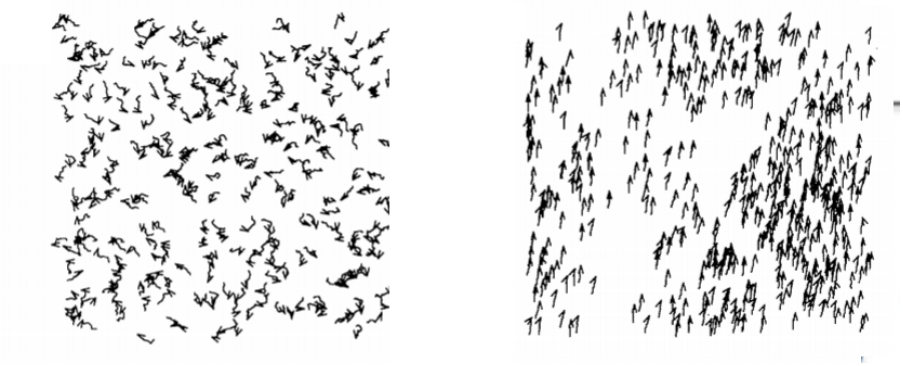

Figure 35: Examples of a disordered (left panel) and an ordered phase (right panel) in the Vicsek model [147].

Different extensions of this Voronoi model are developed, for example by adding an interaction potential [106, 148] and we make an extension that incorporates the interaction potential of the SPV-model, resulting in the Voronoi Vicsek model. The details of this model are described in Subsection 3.4.1 and in Subsection 3.4.2 we look at alignment in comparable models in literature.

3.4.1 Description of the VV-model

The interaction energy in the VV-model is identical to the interaction energy in the VRD-model given by Equation 57 and is in dimensionless parameters given by

$$E = \sum_{i}^{N} k_a(a_i - a_0)^2 + k_p(p_i - p_0)^2, \tag{65}$$

and the equation of motion is again

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \mu \boldsymbol{F}_i^{\mathrm{int}} + v_0 \hat{\boldsymbol{n}}_i. \tag{66}$$

The time evolution of the polarization angle θ_i is different in the VV-model, namely

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \frac{1}{\tau_v}\left\langle \phi_j(t) - \theta_i(t) \right\rangle_{0<|\boldsymbol{r}_i-\boldsymbol{r}_j|<R_V} + \eta_i(t), \tag{67}$$

where τ_V is a persistence time for the cell-cell alignment, $\phi_j(t)$ is the angle associated with the instantaneous velocity vector of neighboring cell j at time t, defined by $\begin{pmatrix}\cos\phi_j(t)\\ \sin\phi_j(t)\end{pmatrix} = \frac{\mathrm{d}\boldsymbol{r}_j/\mathrm{d}t}{|\mathrm{d}\boldsymbol{r}_j/\mathrm{d}t|}$, the brackets $\langle ... \rangle_{0<|\boldsymbol{r}_i-\boldsymbol{r}_j|<R_V}$ denote an average over all neighboring cells j that are within a distance R_V of cell i and $\eta_i(t)$ is a white noise term with mean zero and variance $2D_r$. Thus, in the VV-model each cell tends to align its direction of self-propulsion with the velocity direction of its neighbors over a characteristic time and in the presence of noise, which mimics the connectivity between cells according to our hypothesis. Note that the velocity direction ϕ of its neighbors is used, and not the polarity θ as is common in Vicsek models. The velocity direction ϕ is the sum of the polarity θ and the displacements due to the interaction energy E and the reason for the use of ϕ is the assumption that cells only feel the direction of total movement of a neighboring cell and are not able to distinguish between the movement due to self-propulsion and the movement due to the interaction energy. A second difference with the original Vicsek-alignment given by Equation 64 is that we exclude the alignment of the polarity of a cell with its velocity and only take into account the velocity vectors of its neighbors.

The dynamics in the VV-model is thus regulated by a competition between the cells' tendency to minimize the energy by achieving a target cell geometry and the cells' tendency to align their velocities through the Vicsek mechanism with stochastic Brownian noise. The VRD-model contains three independent parameters, p_0, v_0 and D_r and in the VV-model two parameters are added: R_V and τ_V. Equations 65 till 67 dictates the full dynamics of the VV-model and are solved numerically using Langevin dynamics with a modified version of *CellGPU* [143] (see Appendix E.1) with simulation details specified in Subsection 3.5. In the next subsection we discuss some studies of models that are comparable to the VV-model and focus on alignment.

3.4.2 Alignment and jamming in comparable models in literature

The VV-model reduces to existing models in two limits. When $\tau_v \to \infty$, cells are unable to align on any finite time scale and the model reduces to the VRD-model described in Subsection 3.3 where the direction of the active force undergoes only simple Brownian rotation. The VRD-model is also obtained when $R_V = 0$, because the velocity direction of neighboring cells is never taken into account. Secondly, in the limit of $\mu = 0$ we essentially recover the original Vicsek model for point-like self-propelled particles. The only difference is that we do not take into account the own polarity vector of a particle i in calculating the average angle of its neighbors $\langle ... \rangle_{0<|\boldsymbol{r}_i-\boldsymbol{r}_j|<R_V}$, while the original Vicsek force does include this vector.

The VRD-model already exhibits collective behavior without any mechanism regulating it [108]. In the vicinity of the jamming transition collective motion arises spontaneously because it is the only way for cells to overcome energy barriers for rearrangements. There are also studies that add to the SPP and SPV-models a subtle alignment rule, in which the polarity of a cell aligns to the direction of motion of the cell [106, 117, 149, 150]. Since the motion of each cell is affected by neighboring cells, such a mechanism can indirectly induce polarity alignment of neighboring cells. Nevertheless, it neglects the direct alignment of a cell with its neighbors, and is therefore fundamentally different from the neighbor-alignment in the VV-model. The advantage of the neighbor-alignment over the self-alignment is the possibility to tune the alignment with the Vicsek radius. To the best of our knowledge, this Vicsek-like alignment is never used in combination with the SPV-model [108]. There exist, however, SPP models that include a Vicsek alignment [1, 119, 145, 148].

The model most comparable to the VV-model is the extension of the SPV-model that includes self-alignment [149, 151] and therefore we have a look at its results. In this model, Equation 67 changes to

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = -J\sin(\theta_i(t) - \phi_i(t)) + \eta_i(t), \tag{68}$$

with η_i white noise with zero mean and variance $2D_r$. The angles ϕ_i and θ_i are just as in Equation 67 respectively the velocity angle and the polarization angle. Thus in this flocking model, a cell aligns its polarity with its own movement. Giavazzi *et al.* [149] find that by controlling the alignment interaction J and the shape-index p_0 four different phases can be obtained: a stationary solid, a stationary liquid, a solid flock and a liquid flock (see Figure 36). The stationary solid and liquid states are the same states as in the VRD-model and they are distinguished by the critical threshold $D_{\text{eff}} = 0.001$. Also the solid and liquid flocking states are distinguished by this threshold, while an additional parameter that defines the degree of alignment is used to distinguish the stationary state from a flocking state. The flocking state is defined by a finite value of an alignment parameter and in the stationary state this parameter goes to zero. The left panel of Figure 36 shows a phase diagram based on the shape index p_0 and the alignment interaction J. The center and right panel illustrate the differences

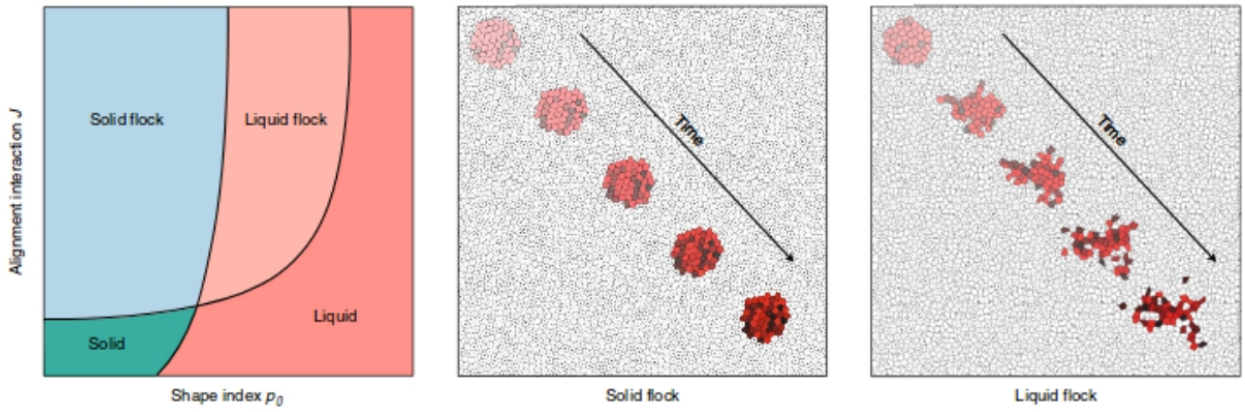

Figure 36: Four phases in an SPV-model with self-alignment (left panel): a liquid, a solid, a liquid flock and a solid flock. The center and right panel show respectively the time evolution of a solid flock and a liquid flock [152].

between a solid flock and a liquid flock. Besides the existence of two extra phases, a key result of this phase diagram is that alignment promotes solidification. When the alignment interaction J is higher, the critical p_0 that distinguish a solid from a liquid also becomes higher. The expectation is that the VV-model shows similar behavior, and this is investigated in Subsection 6.

Several parameters are used in literature to define alignment. The original Vicsek model uses [145]

$$\mathrm{VP} = \frac{1}{Nv_0}|\sum_{i=1}^{N} \boldsymbol{v_i}|, \tag{69}$$

where VP stands for Vicsek Parameter, $|...|$ denotes the size of a vector, v_0 is the input velocity and v_i the instantaneous velocity of each cell or particle. A disadvantage of the VP is that v_0 cannot be obtained in experiments, and therefore makes it hard to compare simulations with experiments. The alignment index (AI) and instantaneous order parameter (IOP) are two other possibilities in which the use of v_0 is avoided [151]:

$$\mathrm{AI} = \frac{1}{N}\sum_{i=1}^{N} \frac{\boldsymbol{v}_i \cdot \langle \boldsymbol{v} \rangle}{|\boldsymbol{v}_i||\langle \boldsymbol{v} \rangle|}, \tag{70}$$

where $\langle ... \rangle = \frac{1}{N}\sum_{i=1}^{N} ...$ and thus $\langle \boldsymbol{v} \rangle$ the average instantaneous velocity vector. The IOP is only slightly different than the AI and is given by

$$\mathrm{IOP} = \frac{|\langle \boldsymbol{v} \rangle|^2}{\langle |\boldsymbol{v}|^2 \rangle}. \tag{71}$$

An important phenomenon in flocking is that the correlation distances are long because particles far apart move on average in the same direction. How far this correlation goes is probed by the correlation length, which we define here as the $5\sigma - L_{\mathrm{corr}}$,

$$L_{\mathrm{corr}} = \{\max r : \mu(\Phi_r) + 5\frac{\sigma_r(\Phi_r)}{N_{\Phi_r)}} \leq \frac{\pi}{2}\}, \tag{72}$$

where Φ_r is the set of N_{Φ_r} differences in velocity angles ϕ_{ij} between cell i and j, given that the distance $r_{ij} = |\boldsymbol{r}_j - \boldsymbol{r}_i|$ between cell i and j lies between $r - \epsilon$ and $r + \epsilon$, with ϵ a value depending on the simulation settings. In the next subsection we explain our simulation and analysis details.

3.5 Computational details and measurement procedures

In this subsection we define our computational details for the VRD- and VV-model as well as a short explanation of the experimental data we use. Furthermore, we define our measurement methods and corresponding

non-trivial details. A complete overview of the algorithms and measurement details are given in respectively Appendices E and F.

VRD-model

We simulate the VRD-model with a modified version of *CellGPU* [143] as explained in detail in Appendix E.1. We set $N = 1000$ and use a step size of $\Delta t = 0.01\,\tau_0$ (see Appendix G.2 for a verification). We define the glass or jamming transition as when the relaxation time exceeds $10^4\,\tau_0$, which is an arbitrary value but common in glassy physics. Therefore, we initialize the system with 10^6 initializing steps $n_{\text{initsteps}}$ (see Appendix G.3 for a verification). Thereafter, we perform $10^6 + 10^5$ simulation steps n_{steps}, where 10^5 is the number of time steps needed for the analysis of dynamic correlation functions. We average the results over ten independent runs, which is a trade-off between calculation costs and reducing the statistical error. Furthermore, for every run calculations of the static functions are averaged over 1000 snapshots at a frequency of 10^4 time steps and the dynamic functions are averaged over 100 time steps at a frequency of 10^2 time steps. We set $k_a = 1.0\,\tau_0^{-2}$, $k_p = 1.0\,r_0/\tau_0^2$, $\mu = 1.0\,\tau_0$, $a_0 = 1.0\,r_0^2$ and $D_r = 1.0\,\tau_0^{-1}$. We vary v_0 between $0.1\,r_0/\tau_0$ and $10\,r_0/\tau_0$ and we vary p_0 between $3.0\,r_0$ and $4.5\,r_0$.

We measure the MSD, the average perimeter q, the RDF, the SSF, the coherent and incoherent ISF and properties that can be derived from these functions such as the diffusion coefficient and the relaxation time, of which the latter is defined as the time after which the ISF reaches 0.1. The measurement procedures are given in detail in Appendix F.1. Important notes are that the average motion of all cells is subtracted from the individual displacements (see Appendix G.4) to calculate the MSD and coherent and incoherent ISF and that the coherent and incoherent ISF are evaluated at the location $\overline{k}$ of the first peak in the SSF (see in Appendix G.5 that using $k = 6.75\,r_0$ does not change the observations).

The MCT-algorithm to predict the coherent ISF is explained in Appendix E.3 and the system specific input consists of the static structure factor $S(k)$, the self-diffusion coefficient D_0 (given by Equation 60) and the density ρ, which is always $1.0\,r_0^{-2}$ in our models. We modify the static structure factor by setting $S(k) = 0$ for the smallest k-values and use a cut-off of $k = 30\,r_0^{-1}$ (see Appendix F.1). The initial time step size is set to $\Delta t = 10^{-6}\tau_{\text{MCT}}$, in which τ_{MCT} is a time unit of MCT (see Subsection 2.3.3) and we use $N_T = 64$. Our maximum simulation time is $t_{\max} = 10^{20}\tau_{\text{MCT}}$. Furthermore, we use a threshold for the memory kernel defined by $\epsilon = 10^{-20}$ and this threshold is always reached within 20 steps. Adaptations to the MCT-algorithm for the incoherent ISF are explained in Appendix E.3.7 and we use the same input and algorithm settings as for the coherent ISF except that this coherent ISF predicted with MCT serves as extra input.

Experimental data

We use experimental data obtained from an MDCK culture and the exact procedure how this data is obtained is described by Atia *et al.* [40]. Images were recorded every 3 minutes for 54 hours using phase microscopy (see supplementary video 1 of Atia *et al.* [40] for a movie). However, we will only use the data from the time that confluence is reached, which is after 33.5 hours. We define the waiting time as the time that has passed since confluence is reached. Each image has a size of 1024 x 1024 pixels, with one pixel equal to $0.89\mu m \times 0.89\mu m$. The images were processed to find the nucleus of each cell using a method that locates around 95 percent of the nuclei (see Appendix G.7 for an analysis of the effect of this accuracy).

We perform the same analysis on these experimental data than on the data obtained from the VRD-model, but there are a few differences between the experimental and the simulated data which we have to take into account. First, the experimental data shows a variation of cell density over time resulting in a varying number of cells in the FOV. To account for this, we scale all lengths at waiting time T with $r_0(T)$ according to

$$r_0(T) = \sqrt{\frac{A}{N(T)}}, \tag{73}$$

with A the area of the snapshot image in pixels and $N(t)$ the number of cells in the snapshot at time t (see Appendix G.9 for an analysis of the effect of this scaling). Secondly, the experimental data is not in a steady state which means that we cannot average over many snapshots to reduce the noise. Therefore, we assume the changes in the system to be negligible during one hour such that we can take the average over all snapshots within this hour. We apply a running average where we define the value of a quantity at waiting time T as the average value of that quantity between waiting time T and $T + 1\,\text{hour}$ (see Appendix G.8 for an analysis with different waiting times). Thirdly, we do not have periodic boundary conditions and lastly, the positions can only be determined with an accuracy of one pixel. In Appendix F.2 we explain in more detail how these four properties of the experimental data effect the measurement procedures.

VV-model

We simulate the VV-model with a modified version of *CellGPU* [143] and explain in Appendix E.1.3 how this algorithm differs from the implementation of the VRD-model. Similar as in the VRD-model, we set $N = 1000$ and use a step size of $\Delta t = 0.01\,\tau_0$. Differently than in the VRD-model, we let the system relax without self-propulsion force before we start measuring. At $t = 0.0\tau_0$ we turn on the self-propulsion force and the system relaxes to a new steady state. Therefore, we initialize the system only for $t = 200\,\tau_0$ to save time (see Appendix G.3 for a verification). We always set $k_a = 1.0\,\tau_0^{-2}$, $k_p = 1.0\,r_0/\tau_0^2$, $\mu = 1.0\,\tau_0$ and $a_0 = 1.0\,r_0^2$.

For the analysis of the alignment in the VV-model we execute 15000 time steps and vary five parameters: R_V, p_0, v_0, D_r and τ_V. We vary one parameter at a time for three different base settings of these five parameters, called Settings A, B and C, given in Table 1. Setting A has a combination of v_0, p_0 and D_r that belongs just to the unjammed regime of the VRD-model (see Figure 32). To verify the results of Setting A, we also analyze Setting B and C. In Setting B p_0 and v_0 are changed such that the VRD-model would result in a jammed state and in Setting C, more collective behavior is added by increasing R_V and decreasing τ_V. We vary R_V from $0.0\,r_0$ till $5.0\,r_0$, p_0 from $3.0\,r_0$ till $5.0\,r_0$, v_0 from $0.0\,r_0/\tau_0$ till $10\,r_0/\tau_0$, D_r from $0.0\,\tau_0^{-1}$ till $10\,\tau_0^{-1}$ and τ_V from $0.01\,\tau_0$ till $100\,\tau_0$. To determine these parameter ranges we have taken into account in which regime of the parameters the dynamical behavior changes most and what the limitations set by our choice of step size are. We average the results over 30 independent runs, which is again a trade-off between computational costs and reducing the statistical error. We measure the alignment index (see Subsection G.6 for a comparison with VP and IOP), the average perimeter q, the average speed $\overline{v}$ over time $\Delta t' = 1.0\,\tau_0$, and the correlation length L_{corr}. The measurement procedures are given in detail in Appendix F.3.

For the analysis of jamming in the VV-model we use an additional 10^6 initialization time steps with the active force turned on. Thereafter, we perform $10^6 + 10^5$ simulation steps n_{steps}, where 10^5 is the number of time steps needed for the analysis of dynamic correlation functions. We average the results over ten independent runs and for every run calculations of the static functions are averaged over 1000 snapshots at a frequency of 10^4 time steps and the dynamic functions are averaged over 100 time steps at a frequency of 10^2 time steps. We additionally set $D_r = 1.0\,\tau_0^{-1}$, $v_0 = 0.25\,r_0/\tau_0$ and $\tau_V = 1.0\,\tau_0$. We vary R_V between $0.0\,r_0$ and $5.0\,r_0$ and p_0 between $3.0\,r_0$ and $4.0\,r_0$. We measure the RDF, SSF, coherent and incoherent ISF and their corresponding relaxation time in a similar way as in the VRD-model. In addition, we also measure the non-ergodicity parameter and the alignment index (see Appendix F.3).

Table 1: Definition of the three different parameter settings A, B and C.

	$R_V(r_0)$	$p_0(r_0)$	$v_0(\tau_0/r_0)$	$D_r(\tau_0^{-1})$	$\tau_V(\tau_0)$	
Setting A	1.5	3.8	0.5	1.0	1.0	*Basic*
Setting B	1.5	**3.6**	**0.1**	1.0	1.0	*Further jammed*
Setting C	**2.5**	3.8	0.5	1	**0.1**	*More collectivity*

4 Results of the Voronoi rotational diffusion model

In this section we present the results of the Voronoi rotational diffusion (VRD) model concerning correlation functions and mode coupling theory. As explained in Subsection 3.3.1, the Voronoi rotational diffusion model is defined by the following equation of motion for cell i:

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \mu \boldsymbol{F}_i^{\mathrm{int}} + v_0 \hat{\boldsymbol{n}}_i, \tag{74}$$

where the interaction force $\boldsymbol{F}^{\mathrm{int}}$ is given by

$$\boldsymbol{F}_i^{\mathrm{int}} = -\nabla_{\boldsymbol{r}_i} E = -\nabla_{\boldsymbol{r}_i} \left(\sum_i^N k_a (a_i - a_0)^2 + k_p (p_i - p_0)^2 \right), \tag{75}$$

and the time evolution of the polarity vector $\hat{\boldsymbol{n}}_i = (\cos(\theta_i), \sin(\theta_i))$ is given by

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \eta_i(t), \quad \text{with} \quad \langle \eta_i(t) \eta_j(t') \rangle = 2 D_r \delta(t - t') \delta_{ij}. \tag{76}$$

Thus, the dynamics of a cell is governed by the tendency of a cell to obtain its preferred shape and by a self-propulsive motion affected by random rotational diffusion. The key parameters are the preferred perimeter p_0, the magnitude of the self-propulsive motion v_0 and the random rotational diffusion D_0.

We start with a short exploration of this VRD-model in Subsection 4.1, which is merely a verification of Bi *et al.* [108] and serves as a phenomenological description. In Subsection 4.2 and Subsection 4.3 we look at respectively structure and dynamics in the VRD-model via correlation functions. Mode coupling theory makes predictions about the dynamics and in Subsection 4.4 we verify these predictions by comparing the MCT-results with the relaxation time obtained via direct calculations, the diffusion coefficient and the average perimeter q. We end with summarizing the main conclusions in Subsection 4.5.

4.1 Phenomenological description

The VRD-model has a jamming transition from a fluid-like state to a solid-like state as shown by Bi *et al.* [108]. One way to cross this transition is by varying the preferred perimeter p_0. Figure 37 shows the cell shapes and trajectories for three different values of p_0 corresponding to a jammed state, a state near the jamming transition and an unjammed state according to Figure 32. Cells form almost hexagons in the jammed state because a hexagonal tiling is the most ideal tiling of space (see Subsection 3.2) and cells stay around their original positions due to the caging of their neighbors. Near the jamming transition the variation in cell shapes becomes larger, but the preferred hexagonal shape is still visible. The cell trajectories become longer as well, although the cells do not migrate far. In an unjammed state cell shapes are far from hexagonal and exhibit a lot of variation. The boundaries between cells become longer and the distances between cell seeds (Voronoi centers) variate more. Furthermore, cells migrate over long distances. The only cause of the unjamming transition as illustrated in Figure 37 is the change in the preferred perimeter p_0, which is too small to be reached in the jammed state causing cells to have an almost hexagonal shape and large in the unjammed state resulting in long boundaries and highly irregular shapes. Thus, there is a critical preferred perimeter p_0 at which a confluent cell layer undergoes a jamming transition.

The location of jamming or unjamming transition is also influenced by the two other free parameters v_0 and D_r. An increase in the self-propelled velocity v_0 might fluidize a system because a higher velocity makes it easier for cells to escape from the cage of their neighbors and, consequently, cells migrate further. Bi *et al.* [108] also show that increasing the persistence of the self-propelled force by decreasing the rotational noise D_r might lead to unjamming because a persistent motion also makes it easier for cells to escape from their neighbors. Because the effect of the parameters p_0, v_0 and D_r on the location of the jamming transition is studied extensively by Bi *et al.* [108], we will not look further into the details of the jamming phase diagram of the VRD-model in this thesis. Instead, we are going to focus on the structural and dynamical differences around the jamming transition and the predictions of mode coupling theory.

In doing this, we also incorporate more extreme values of p_0 than shown in Figure 37 and Figure 38 shows how this influences the cell shapes. When p_0 is very small ($p_0 = 3.0\, r_0$) cells can never reach their preferred perimeter, which results in a high stress level in the layer (note that the minimum value of the average perimeter q is given

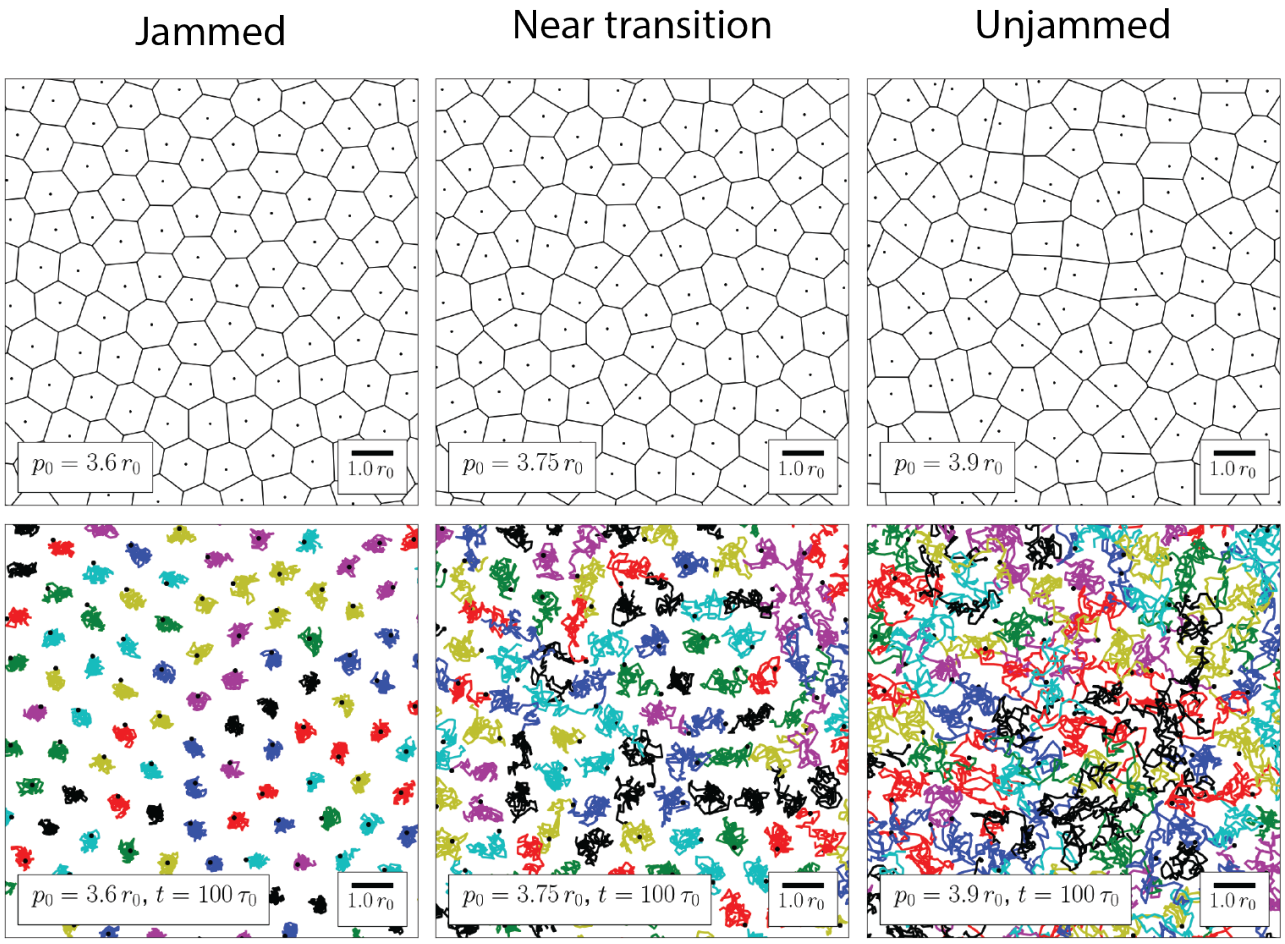

Figure 37: The shapes (top panels) and trajectories (bottom panels) of a subset of the cells with from left to right $p_0 = 3.6\,r_0$, $p_0 = 3.75\,r_0$ *and* $p_0 = 3.9\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$).

by 3.72, according to Subsection D.3). Cells still are preferably hexagonal for these small p_0 values, but the variation in cell sizes is higher. When p_0 is very large ($p_0 = 4.5\,r_0$), cells tend to increase the boundaries with their neighbors resulting in more elongated shapes. The resulting Voronoi tesselations are not stable because they change very rapidly in time and it is questionable whether this is still a representation of real cells. The reason that we still include these extreme values in our analysis is that we also regard the VRD-model as a toy model to apply mode coupling theory to, besides its usual implementation to describe biological processes in confluent cell layers. Including these extreme p_0-values results in a larger variation in structural and dynamical properties. In the next subsection we look at the structural properties of the VRD-model in detail.

4.2 Analysis of the structure

The vectorial radial distribution function for four different values of p_0 and $v_0 = 0.25\,r_0/\tau_0$ is given in Figure 39 and the corresponding static structure factor in Figure 40. For this value of v_0 the jamming transition takes place around $p_0 = 3.75\,r_0$ according to Figure 32 in Subsection 4.3. Below the jamming transition the radial distribution function shows long-range order (right upper panel of Figure 39), which is also reflected in high peak values of the static structure factor (right upper panel of Figure 40). When we go to even lower values of p_0, the long-range order disappears and the system is less structured (left upper panels of Figures 39 and 40). Cells can never obtain such a low values of p_0 and this results in a high level of stress in the system, which we call elastic frustration. This frustration hinders the forming of a hexagonal pattern and results in a less structured RDF and SSF for very low values of p_0. Just above the jamming transitions (left bottom panels of Figures 39 and40), there are still signs in the RDF and the SSF of a long-range order, although much weaker than on the other side of the jamming transition. This structural arrangement and the long-range order vanish when we go to even higher p_0 (right bottom panels of Figures 39 and 40). A hexagonal packing structure is no longer optimal, which results in a higher variety of cell shapes and therefore a less structured layer.

The vectorial radial distribution function and static structure factor look isotropic, which means that they show identical behavior for all $\boldsymbol{r}$- and $\boldsymbol{k}$-vectors with the same magnitude. Nevertheless, for highly structured systems

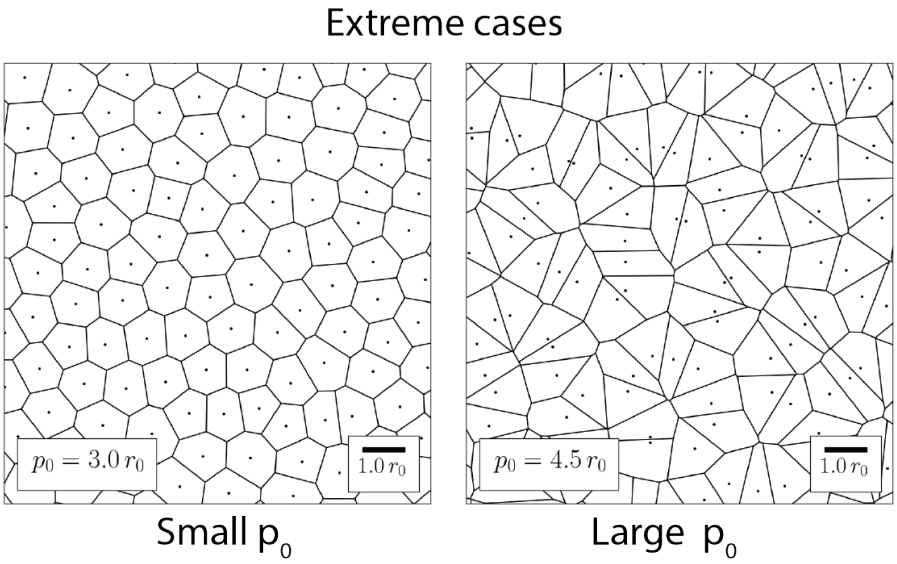

Figure 38: The cell shapes for the smallest ($p_0 = 3.0\,r_0$, left panel) and largest preferred perimeter ($p_0 = 4.5\,r_0$, right panel) that we investigate in this thesis ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$).

around the jamming transition the radial distribution function and the static structure factor show an hexagonal pattern reflecting the hexagonal order of the layer. The graphs of Figures 39 and 40 for $p_0 = 3.6\,r_0$ already hints towards this order, but it is not clear visible here because the data is averaged over ten independent runs. Therefore, in Appendix H.1 the hexagonal pattern in the RDF and SSF is shown for a single run at $p_0 = 3.7\,r_0$, which is even closer to the jamming transition than $p_0 = 3.6\,r_0$ in Figures 39 and 40. Despite this anisotropy, we will from now on only look at the scalar version of the RDF and SSF and assume that by averaging over several runs the effect of this anisotropy is negligible.

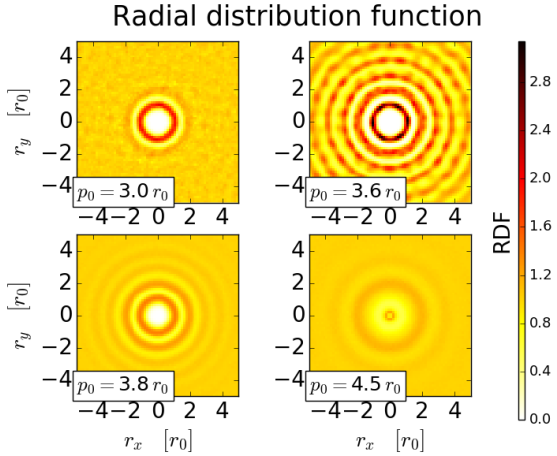

Figure 39: Vectorial radial distribution function for $p_0 = 3.0\,r_0$, $p_0 = 3.6\,r_0$, $p_0 = 3.8\,r_0$ and $p_0 = 4.5\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$).

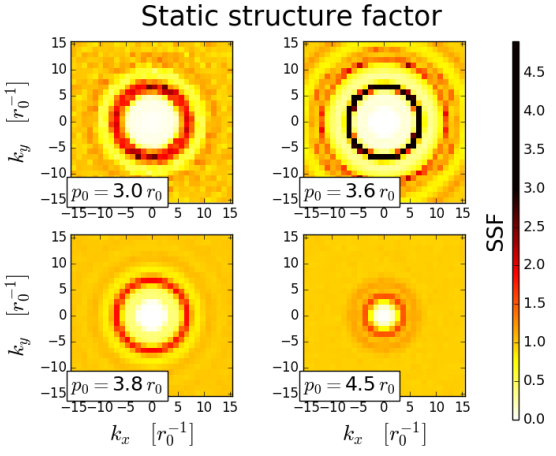

Figure 40: Vectorial static structure factor for $p_0 = 3.0\,r_0$, $p_0 = 3.6\,r_0$, $p_0 = 3.8\,r_0$ and $p_0 = 4.5\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$).

These scalar versions of the RDF and SSF are shown in Figure 41 and they also reflect the structured order for values of p_0 around the jamming transition and the decrease in this order for extremer values of p_0. A structural order is characterized by higher peaks and the existence of peaks and valleys for large values of r or k in respectively the RDF and SSF, while these peaks are smaller and the RDF and SSF at large value of r or k are equal to one for less structured states. The location of the first peak of the RDF represents the nearest neighbor distance and is for values of $p_0 < 3.8\,r_0$ slightly larger than $1.0\,r_0$, which is the closest packing for hard spheres. Because the cells in the VRD-model are confluent, their nearest neighbor distance is a bit higher and are comparable to the nearest neighbor distance in a hexagonal grid (see Subsection 3.1 with a first peak in the RDF around $r = 1.07\,r_0$. For values of r smaller than this first peak the RDF goes to zero implying mutual exclusion. Furthermore, the location of the first peak in the SSF for $p_0 < 3.8r_0$ corresponds to the inverse length scale of the first peak in the RDF. What stands out in Figure 41 is the shift in the first peak when p_0 is

large. The mutual exclusion at short range disappears as cell align and have their Voronoi centers closer to each other to become more elongated to satisfy the higher preferred perimeter. In addition, cells that are not aligned have seeds that are further apart. This variance in distances between cell seeds results in a broadening of the peaks in the RDF and SSF. This transition from mutual exclusion to cell seeds sticking to each other happens around $p_0 = 4.0\,r_0$, which is above the jamming transition and we will investigate the effect of this transition on the dynamics further in Subsection 4.3. The almost linearly decay of the SSF to zero when k goes to zero is known as hyperuniformity. This hyperuniformity is also seen in certain random packings of spheres [153], and might suggest that the cell layers in the VRD-model are tiled in a similar disordered manner for large but not too large values of p_0. For these too large values ($p_0 > 4.25\,r_0$) the growth in the first peak of the RDF suggests a preference for cell seeds sticking to each other instead of a completely random order.

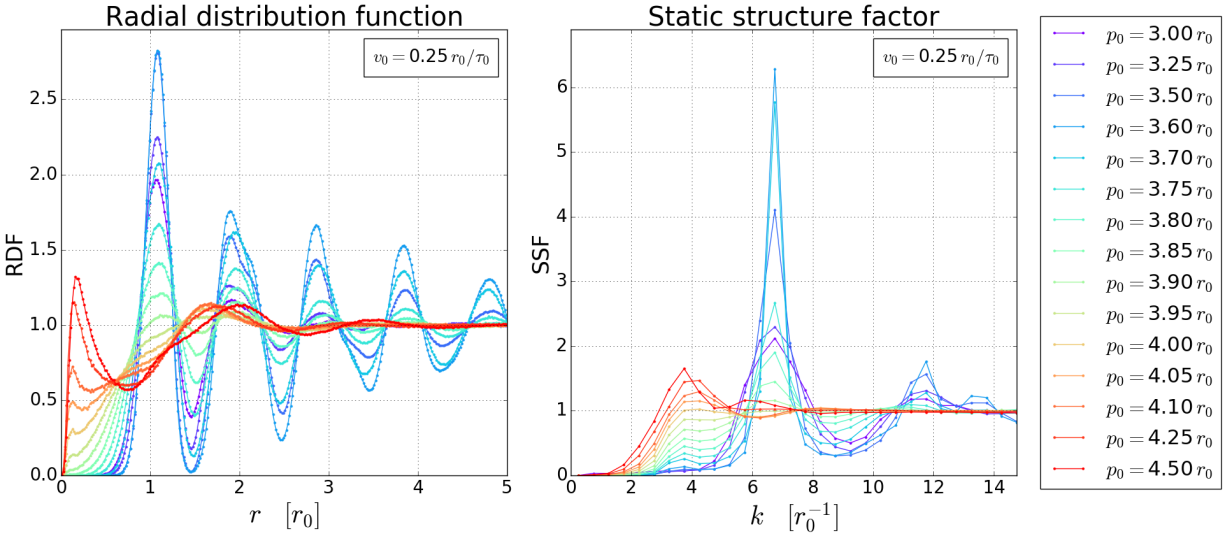

Figure 41: The scalar radial distribution function (left panel) and static structure factor (right panel) for various values of p_0 between $3.0\,r_0$ and $4.5\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$).

Although the RDF and SSF are different for other values of v_0 (see Appendix H.2 for $v_0 = 0.1\,r_0/\tau_0$, $v_0 = 0.5\,r_0/\tau_0$ and $v_0 = 1.0\,r_0/\tau_0$), the functions look qualitatively similar upon a shift in the value of p_0. Our results for the RDF and SSF are comparable to Li [154], which uses the same model but neglects the self-propulsion force corresponding to $v_0 = 0.0\,r_0/\tau_0$. The changes in structure seen in this subsection are encouraging for applying mode coupling theory as MCT predicts changes in dynamics based on the changes in structure. Before applying MCT, we investigate first the dynamics of the VRD-model in the next subsection such that we can verify these MCT-predictions.

4.3 Analysis of the dynamics

The mean squared displacement for simulations of the VRD-model with $v_0 = 0.25\,r_0/\tau_0$ and various p_0 values is given in Figure 42 and looks similar to the graph of Bi *et al.* [108] in Figure 31. The motion is ballistic on a short time scale with a slope of the MSD on a log-log plot equal to two and changes to a diffusive motion on long time scales for large p_0 values with a slope equal to one. For values of $p_0 > 3.75\,r_0$ a plateau arises due to the caging of cells by their neighbors, which clearly indicates a jamming transition. The effective diffusion coefficient obtained from the MSD (Equation 1) is one of the parameters to quantify the jamming transition. The left upper panel of Figure 44 shows this effective diffusion coefficient as a function of p_0 for four different values of v_0. Below the black horizontal line $D_{\text{eff}} = 10^{-3}$ the system is conventionally said to be jammed.

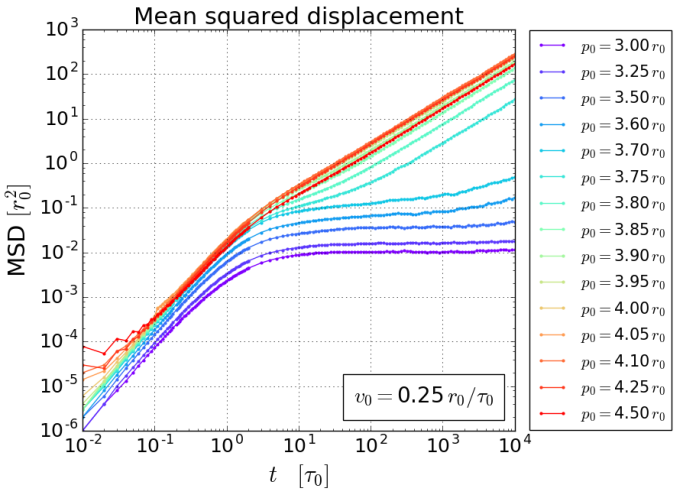

Figure 42: The mean squared displacement for various values of p_0 between $3.0\, r_0$ and $4.5\, r_0$ ($v_0 = 0.25\, r_0/\tau_0$, $D_r = 1.0\, \tau_0^{-1}$).

This jamming transition is also visible in the coherent and incoherent intermediate scattering function shown in Figure 43. For $p_0 > 3.75\, r_0$ the ISFs fall back to zero within $10^4\, \tau_0$, while for small p_0 a plateau arises and the value at $t = 10^4\, \tau_0$ is nonzero. The non-ergodicity parameter is larger when p_0 is smaller, which indicates that the cells migrate less, despite the lower structural order observed observed in the RDF and SSF for $p_0 = 3.0\, r_0$ in comparison to slightly higher values. The coherent and incoherent ISF show similar behavior for different values for v_0 (see Appendix H.2 for $v_0 = 0.1\, r_0/\tau_0$, $v_0 = 0.5\, r_0/\tau_0$ and $v_0 = 1.0\, r_0/\tau_0$).

A second way to quantify the jamming transition is via the relaxation time τ_α obtained via the coherent or incoherent ISF. These relaxation times are shown in respectively the left and right center panel of Figure 44 as a function of p_0 for four different values of v_0. A system is said to be jammed when the relaxation times exceeds $10^4\, \tau_0$. Because we only simulate until $t = 10^4\, \tau_0$, we cannot obtain relaxation times larger than $10^4\, \tau_0$ and this is indicated with dashed lines in the graphs of Figure 44.

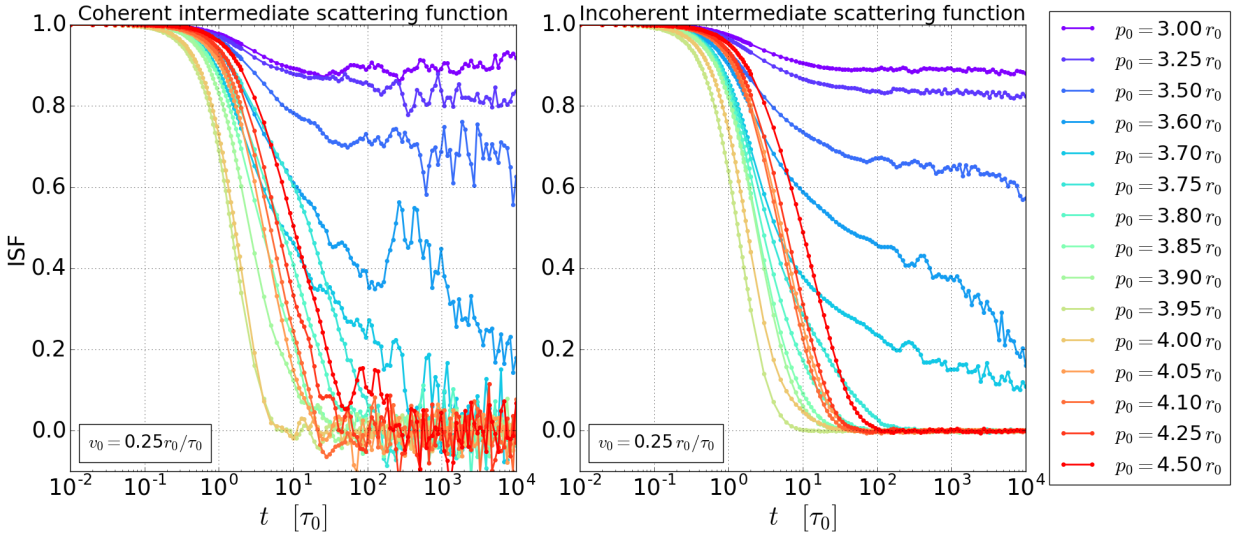

Figure 43: The coherent intermediate scattering function (left panel) and incoherent intermediate scattering function (right panel) for various values of p_0 between $3.0\, r_0$ and $4.5\, r_0$ ($v_0 = 0.25\, r_0/\tau_0$, $D_r = 1.0\, \tau_0^{-1}$).

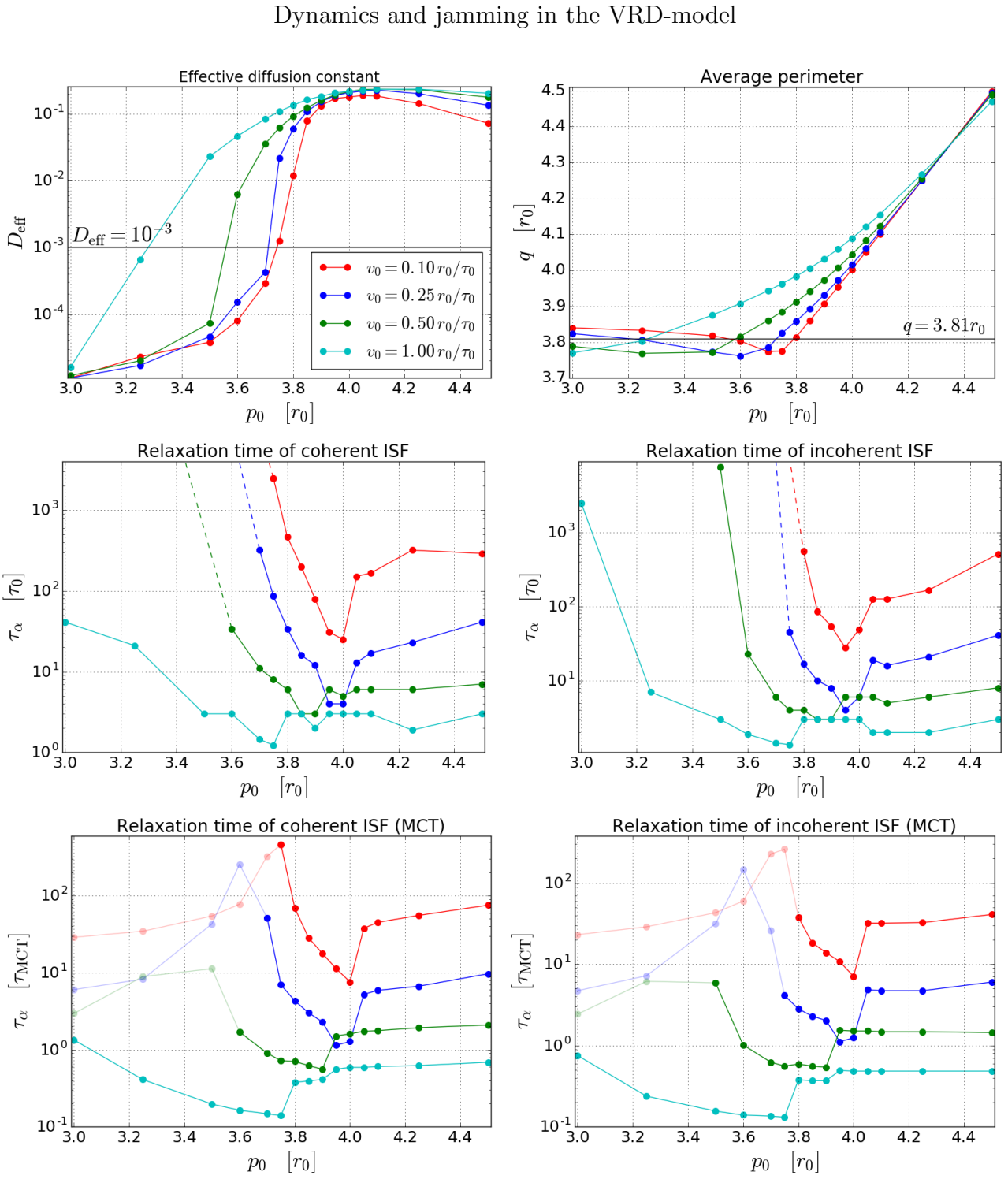

Figure 44: The effective diffusion coefficient (left upper panel), the average perimeter q (right upper panel), the relaxation time of the coherent ISF (left center panel), the relaxation time of the incoherent ISF (right center panel), the mode coupling predictions for the relaxation time of the coherent ISF (left bottom panel) and the mode coupling theory predictions for the relaxation time of the incoherent ISF (right bottom panel) for $v_0 = 0.1\, r_0/\tau_0$, $v_0 = 0.25\, r_0/\tau_0$, $v_0 = 0.5\, r_0/\tau_0$ and $v_0 = 1.0\, r_0/\tau_0$, various values of p_0 between $3.0\, r_0$ and $4.5\, r_0$ and $D_r = 1.0\, \tau_0^{-1}$.

Both D_{eff} and τ_α indicate that below a certain p_0 a system is jammed in agreement with our observations in Subsection 4.1. Around this jamming transition cells become caged by their neighbors and this results in the arising of a plateau in both the MSD and the ISF. This plateau is included in the values of respectively D_{eff} and τ_α, which makes these parameters suitable for defining the jamming transition. The value of p_0 where the jamming transition takes places decreases when v_0 increases in agreement with the jamming diagram in Figure 33. Nevertheless, the exact value of p_0 at which the jamming transition takes place differs for the two methods used to define jamming. One example is that D_{eff} defines $p_0 = 3.0\,r_0$ and $p_0 = 3.25\,r_0$ as jammed, while τ_α indicates that the system is unjammed. The jamming transition is a continuous transition, meaning that there is no clear boundary at which this transition occurs and Figure 44 shows that the location of the jamming transition depends on the method you choose to define this transition and the threshold value for these methods. For example, we use the thresholds $D_{\text{eff}} < 10^{-3}$ and $\tau_\alpha < 0.1\tau_0$ to define jamming and different values will change the location of the jamming transition. Besides indicating the jamming transition, D_{eff} and τ_α also show a slowing down of the dynamics for $p_0 > 4.1\,r_0$. This coincides with the transition from mutual exclusion to cell seeds sticking together as explained in Subsection 4.2. For these large values of p_0, the structural order partly returns to the confluent cell layers. However, the form of the structure is based on elongated aligned cells instead of hexagonal shapes.

Bi *et al.* [108] discovered a third method to define the jamming transition, which is based on the average perimeter q. This q is shown in the right upper panel of Figure 44 for the same combinations of v_0 and p_0 as D_{eff} and τ_α are calculated. A system is said to be jammed when $q < 3.81\,r_0$. According to our data, this threshold only holds for a certain range of p_0. When p_0 is far away from the value where jamming takes place, q is not a valid criterion anymore. For example, q predicts an unjamming transition for $p_0 < 3.5\,r_0$ or $p_0 < 3.25\,r_0$ for respectively $v_0 = 0.1\,r_0/\tau_0$ or $v_0 = 0.25\,r_0/\tau_0$, while the MSD and the ISFs shows even further slowing down of the dynamics for these values. In addition, the slowing down for large values of p_0 is also not captured by q.

When temperature controls a glass transition, the fragility can be deduced from a plot of the relaxation time as a function of the temperature (see Figure 10). Because in the VRD-model the relaxation time is a non-monotonous function of p_0, it is hard to define whether the VRD-model is comparable to a strong or fragile glass former. However, the comparison might be possible with v_0 or D_r as control parameter because both influence the self-diffusivity D_0 which plays a similar role in the VRD-model as the temperature in glasses. Furthermore, our results for v_0 and the results of Bi *et al.* [108] for D_r suggests that D_{eff} and τ_α depend monotonously on these parameters.

In the next subsection we investigate the dynamics predicted by MCT and compare them with the results discussed in this subsection. To do so, we refer to the intermediate scattering function and relaxation times presented in this subsection as "direct calculations" and their counterparts obtained via MCT as "MCT-predictions".

4.4 MCT-predictions and verifications

We calculate the predictions of the coherent and incoherent intermediate scattering function via the mode coupling theory equations given by respectively Equations 41 and 46. Recall that these MCT-equations do not explicitly take into account the activity of the cells, but that we assume that the activity is captured via its influence on the static structure factor. Figure 45 shows the results for $v_0 = 0.25\,r_0/\tau_0$ and several values of p_0, where the static structure factor of Figure 41 is used as input. The lines for all investigated values of p_0 fall back to zero and there is no clear evidence for a jamming transition as was seen in the graphs of the ISF obtained via direct calculations (Figure 43). MCT can predict the existence of a plateau in the ISF as was seen in Subsection 2.3.3, so it is noteworthy that it does not do so in the VRD-model. Normally, the caging behavior of the MCT-predictions is caused by strong values of the vertices in the MCT-equation. These vertices are larger when the peaks in the static structure factor are larger. So apparently, the static structure factor is not peaked enough for MCT to pick up this caging behavior.

Despite the absence of the plateau in the MCT-predictions of the ISF, the rate at which the ISFs decrease vary for different values of p_0 and this can be measured with the relaxation time. The left and right bottom panel of Figure 44 show the relaxation times obtained from respectively the predicted coherent and incoherent ISF. The data points corresponding to combinations of p_0 and v_0 of which the direct calculations of the relaxation times give a jammed state are made transparent. For all other data points we see a remarkably good qualitative agreement between the MCT-predictions and the computed values via direct calculation of the relaxation time. Starting from the jamming transition and increasing p_0 first results in a decrease of the relaxation time and thereafter an increase. Mode coupling theory predicts this non-monotonous behavior correctly as well as the location of the valley in the relaxation time. Furthermore, MCT predicts correctly that the relaxation time

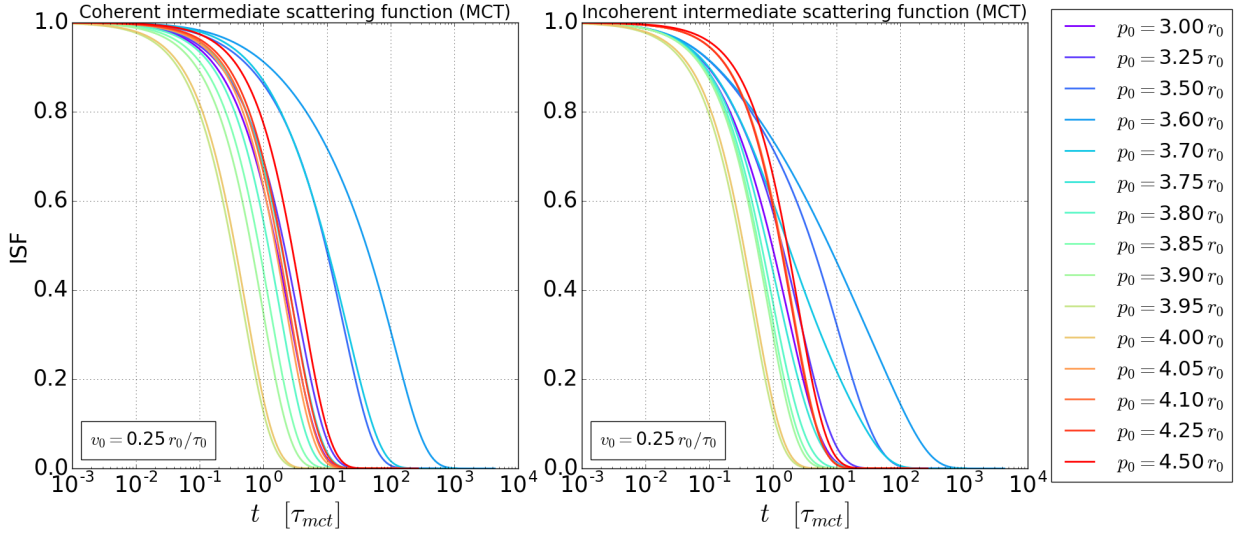

Figure 45: The mode coupling predictions for the coherent (left panel) and incoherent ISF (right panel) with the SSFs of Figure 41 as input.

decreases when the velocity increases. Remarkably, it even predicts that this does not hold for $p_0 = 3.8\, r_0$ and $p_0 = 4.0\, r_0$ where the relaxation time of $v_0 = 0.5\, r_0/\tau_0$ is larger than the relaxation time of $v_0 = 0.25\, r_0/\tau_0$.

MCT does not predict the quantitative value of the relaxation time correctly, which is a well-known failure of mode coupling theory (see Subsection 2.3.3). Interestingly, MCT predicts smaller relaxation times than they actually are while it normally predicts slower dynamics. This prediction of too slow dynamics is explained by the fact that mean field theory generally overestimate the tendency to vitrify due to the lack of ergodicity-restoring fluctuations, and it is yet unknown why this does not hold in our case. One might think that this is due to not taking into account the active force explicitly in the mode coupling theory. However, the results of Bi *et al.* [108] and the analysis in this section suggests that the active force fluidizes the system and therefore would lead to smaller relaxation times. Consequently, including this in MCT would probably not increase the predicted relaxation times. Another explanation for the fact that MCT predicts too fast dynamics could be our assumption that the layer is monodisperse because we use monodisperse values for p_0 and a_0. Nevertheless, multidispersity arises spontaneously in the layer and this could influence the static structure factor which might lead to an incorrect input in the monodisperse version of the mode coupling theory. A comparable situation is found by applying mode coupling theory on polymers, where a monodisperse MCT predicts too fast dynamics and improvements are obtained by applying multidisperse MCT [155]. It is therefore worthwhile to test whether applying multidisperse MCT predicts slower dynamics in the VRD-model too. As a first test the static structure factors corresponding to different particle size can be compared and differences in these SSFs might encourage a multidisperse MCT approach. Besides the multidispersity, the size of the cells also changes in time and this might also influence the dynamics. Szamel [156] presents a mode coupling theory that includes size swaps in a binary mixture and show that these swaps speed up the dynamics. Nevertheless, at first view this might contradict our results as we are looking for a mechanism in mode coupling theory that slows down the dynamics.

Another failure is that mode coupling theory predicts a decrease in relaxation times in cases where the system is jammed and becomes further jammed, shown by the transparent parts of the left and right bottom panels of Figure 44. One explanation for this discrepancy is that the static structure factor used as input for the MCT-calculations is not correct. Because for these parameter settings it takes longer than $t = 10^4\, \tau_0$ for the layer to relax and we only initialize for $t = 10^4\, \tau_0$, the system is not yet in a steady state when we start measuring the static structure factor. Nevertheless, this discrepancy might also be a failure of mode coupling theory.

The peak height in the static structure factor is a measure of how structured a system is. One might therefore ask whether a very simple approach that predicts the relaxation time only based on the height of the first peak gives equally good results as mode coupling theory. If predictions of such a simple ad hoc approach can be verified as well, this would suggest that the verification of the MCT-predictions as done in this section is not significant for the validity of mode coupling theory. However, the peak values of the SSF in Appendix H.3 for the same parameter combinations as in Figure 44 show that the MCT-predictions are not simply a reflection

of these peaks. Thus, MCT does not only take into account the value of the SSF at its highest peak, but also uses the SSF at other k-values or modes as its name already suggests.

We only applied a standard version of mode coupling theory and the fact that this already gives verifiable predictions gives hope that modified versions might improve the results even further. The use of a multidisperse version of MCT is already suggested and another potential modification is attempting to improve the approximation of the four-point density correlations functions by delaying the factorization using generalized mode coupling theory (see Subsection 7.2). The projection of the memory kernel in the standard MCT-derivation of Subsection 2.3.1 on the pair densities is based on pair interactions, while the VRD-model incorporates multi-particle interactions. Including higher order density correlations via so-called generalized MCT might therefore result in a better incorporation of these higher order particle interactions and an improvement in the MCT-predictions.

A different approach to improve MCT is to include the active force explicitly in the mode coupling theory equations, as we only take into account the activity via the static structure factor. In Subsection 7.1 we look further into active versions of MCT. On a side note, this assumption to only include the activity via the SSF might be interesting to test by verifying the MCT predictions for different noise levels. We expect the MCT predictions to become better when the noise level increases and worse when the noise level decreases, because in the second case the persistence of the active force increases the impact of this activity on the dynamics.

Based on the literature study in Subsection 2.3.3, MCT predicts scaling laws for the α- and β-relaxation of the intermediate scattering function as well for the relaxation time as a function of the temperature. Now we have given a first proof of the applicability of standard-MCT, an interesting next step is to test whether these scaling laws are also predicted correctly in the VRD-model. A first test is to try whether the scaling laws given in Figure 15 can be fitted to the intermediate scattering functions given in Figure 43. To test the power-law scaling for the relaxation time similar to Figure 20, the self-diffusivity D_0, depending on v_0 and D_r, can be used instead of the temperature (according to Equation 48). This self-diffusivity can also be used for testing the equivalent of the TTSP as illustrated in Figure 18. MCT does not analytically predict scaling laws with p_0 as control parameter as the MCT-equation does not explicitly depend on p_0. Moreover, as far as we know it is never investigated whether a scaling law for the relaxation time and an equivalent of the TTSP exists for p_0. The existence of these scaling laws in the VRD-model could be investigated first before testing these whether MCT predicts these laws numerically.

4.5 Conclusions

The jamming transition in the VRD-model is characterized by an increase in structure probed by the radial distribution function and the static structure factor, caused by the tendency of cells to reach a hexagonal shape, and by an increase in relaxation time probed by the intermediate scattering function. This makes these functions a powerful tool to investigate the jamming transition. Further away from the jamming transition, the correlation functions show that the dynamics is governed by different length scales upon varying p_0. The long-range order decreases by decreasing p_0 to very small values due to elastic frustration, while the dynamics slow down further. By increasing p_0 to very high values, a transition arises from a disordered state towards a more structured state in which cells get elongated, align and have cell seeds arbitrarily close to each other, which also results in a slowing down of the dynamics. Although a larger v_0 decreases the structure according to the RDF and SSF and increases the relaxation time, the qualitative results about the jamming transition do not depend on v_0. For further research, it might be interesting to investigate the effect of the rotational diffusion D_r on the correlation functions as well.

Mode coupling theory predicts the relaxation times qualitatively correctly when the system is unjammed or reaches the jamming transition. The advantage of MCT is that it only needs structural information as input, which can be obtained from only one snapshot, while direct calculations of the relaxation time require long simulation runs during $t > 10^4\,\tau_0$. Another method that provides information about the dynamics using only one snapshot is based on the average perimeter q. This method can only be applied close to the jamming transition and because we showed that the MCT-analysis also works for high p_0-values it is broader applicable.

Despite its successes in simulated confluent cell layers, MCT also has its shortcomings. In the jammed regime MCT does not predict a jammed state and overall the relaxation times predicted by MCT are too short. For future research we advise to look into possible extensions of mode coupling theory that might reduce these shortcomings. These extensions include multidisperse MCT, generalized MCT and active MCT.

A big advantage of MCT is that it is a theory based on first principles and therefore has a strong physical foundation. This makes it possible to use the theory to predict how the dynamics changes when the structure or input parameters changes. MCT also predicts scaling laws for the α- and β-relaxation as well as the relaxation

time of the intermediate scattering function, which might be interesting to test in further research. Because MCT is based on the fact that the dynamics is governed by the structure, this research shows that this statement is also valid for confluent cell layers.

Concluding, we did not find any evidence that MCT might not be applied to the VRD-model and this supports the applicability of MCT and the analogy between confluent cell layers and molecular fluids or granular particles. Because these molecular fluids or granular particles can undergo a glass or jamming transition, this underpins the existence of an equivalent transition in confluent cell layers. In the next section we use experimental data instead of simulated data to perform the same analyses as done in this section.

5 Results of the analysis of experimental data

In this section we extend our analysis to experimental data and we follow the exact same analysis procedure as for the VRD-model in Section 4. The experimental data was obtained from a maturing MDCK-layer and snapshots were taken every three minutes until three hours after confluence. In Subsection 5.1 we start with a first exploration of the data resulting in a phenomenological description, in Subsection 5.2 we study the structure and in Subsection 5.3 we study the dynamics. The mode coupling predictions and verifications are discussed in Subsection 5.4 and we end with the main conclusions in Subsection 5.5.

5.1 Phenomenological description

The top panels of Figure 46 show the nuclei after 0 and 1911 min waiting time, which is our maximum measurement time, and Voronoi tessellations are drawn with these nuclei as seeds. Many processes influence the aging of the cell layer, such as proliferation, cell death and cell motion. One of the results is the significant increase in density upon aging and this is shown for the complete measurement period in Figure 47. To be able to compare the system at different densities, we scale all length scales with the square root of the density such that at all times the scaled density is equal to one according to Equation 73. Although time dependent, this scaling is similar to the scaling in the VRD-model, where the square root of the average area of each cell is used. The growth in density and the biological processes working in the layer also influence the dynamics in the system. Although it is an interesting question what the effect of each process is on the structure and dynamics, we do not need to take this into account in our analyses because correlation functions and mode coupling theory are generic tools with which the structure and dynamics can be analyzed regardless of the processes regulating it.

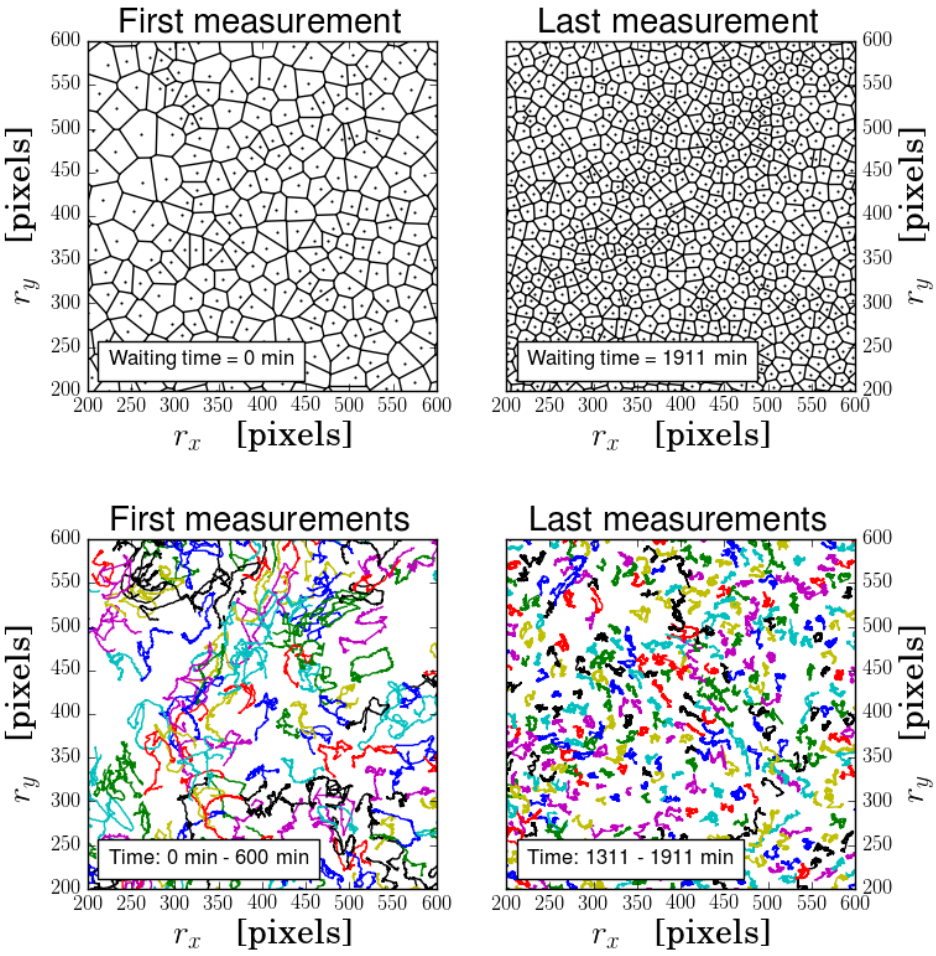

Figure 46: Nuclei and corresponding Voronoi tessellation at the moment confluence is reached (left top panel) and at the last measurement time (right top panel). The left and right bottom panel show trajectories at respectively the start and end of the measurement time period.

Both snapshots in Figure 46 show structural heterogeneity in cell shapes and a difference in the dynamics cannot be derived from these images only. Nevertheless, the trajectories of the cells during 600 min in the bottom panels of Figure 46 suggests a slowing down of the dynamics upon aging. At later times the trajectories become shorter and cells migrate less, which might be a sign that the system approaches a jamming transition. Note that the open spaces in the graphs of Figure 46 do not mean that there are no cells at these positions, but that the nuclei of these cells could not be identified during the entire period of 600 minutes. In the next subsection we investigate the structural properties of the data in detail.

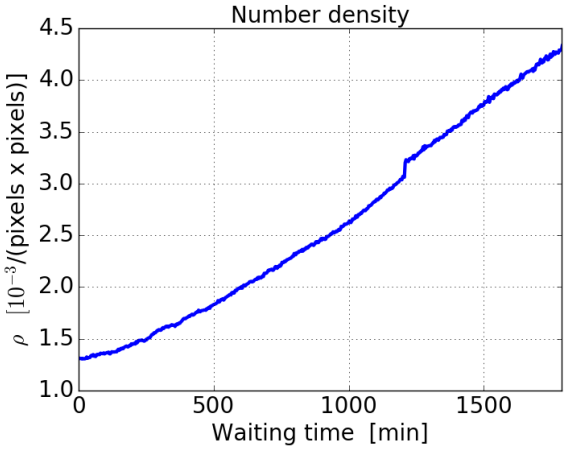

Figure 47: The number density of the cell layer increases during aging.

5.2 Analysis of the structure

Figures 48 and 49 show respectively the radial distribution function and the static structure factor after four different waiting times. Both figures show that the structure increases as a function of the waiting time. The length scale of mutual exclusion in the RDF grows as well and the first peak and valley become clearly visible. Also in the SSF the first peak and valley grow during aging. Because the vectorial RDF and SSF show that the system is nearly isotropic, we can look at the scalar version of the RDF and SSF to see the growth of the peaks and valleys more clearly in respectively Figures 50 and 51.

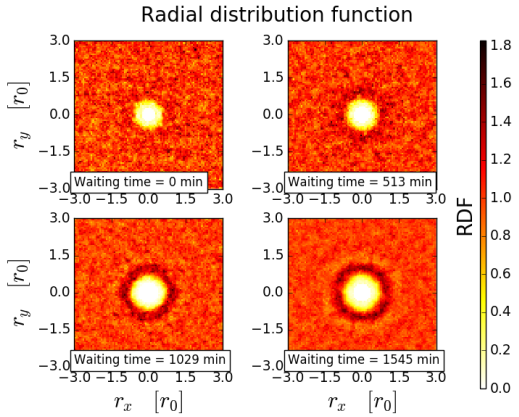

Figure 48: Vectorial radial distribution function after waiting 0, 513, 1029 and 1545 min.

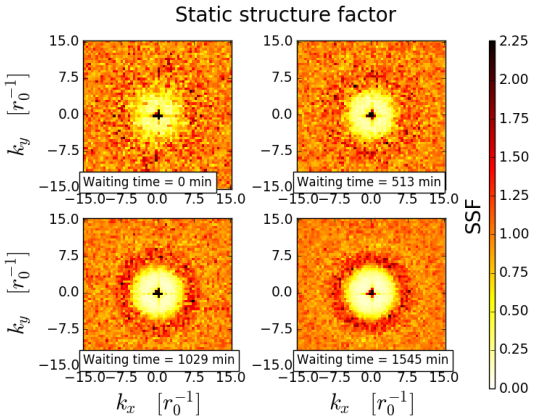

Figure 49: Vectorial static structure factor after waiting 0, 513, 1029 and 1545 min.

Although Figures 50 and 51 suggest so, longer waiting times do not always result in a more structured layer. In Figures 52 and 53 the height of the first peak of respectively the RDF and SSF are shown as a function of the waiting time and they show a noisy increase. This noise might be partly caused by the measurement methods, but it might also have a physical relevance. Because many different biological processes are at work in the cell layer, it is possible that the layer becomes temporarily less structured before increasing its structural order again. Despite the increase in first peak height, the maximum height in Figures 52 and 53 is still very low

in comparison to jammed or glassy materials (see for example the RDF and SSF of the VRD-model in Figure 41).

For a classical system with hard spheres the first peak height of the SSF grows according to a power-law as a function of the density [157], while other systems might exhibit a different growth function. Determining the function according to which these peak values grow in confluent cell layers might give insights into which systems these cell layers are comparable to and might help understanding the physics governing the dynamics in these layers. Unfortunately, the noise in the data prevents determining this function as is shown in Appendix I.1. Besides, we show in Appendix I.2 that the peak height of the RDF and SSF can both be used as a method to define when confluence is reached.

The structural changes upon aging indicated by the increase in the height and narrowing in the width of the peaks and valleys in the RDF and SSF motivate a mode coupling theory analysis. But before looking at the MCT-predictions for the dynamics, we first investigate the dynamics in the next section.

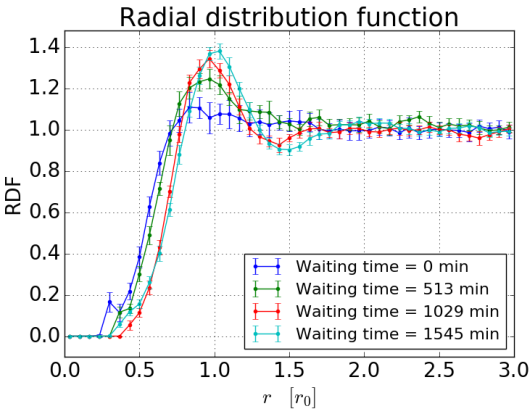

Figure 50: Scalar radial distribution function after waiting 0, 513, 1029 and 1545 min.

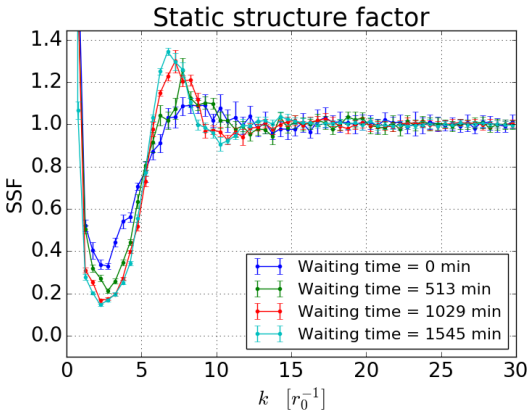

Figure 51: Scalar static structure factor after waiting 0, 513, 1029 and 1545 min.

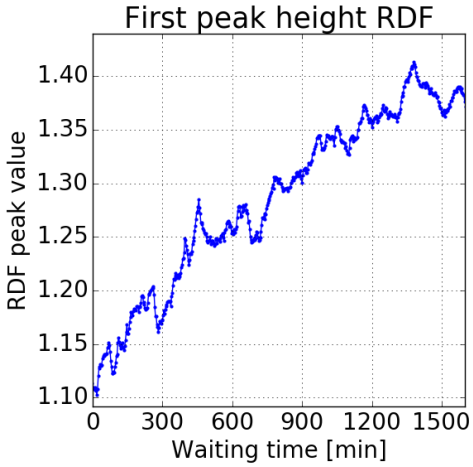

Figure 52: Height of the first peak of the radial distribution function as a function of the waiting time.

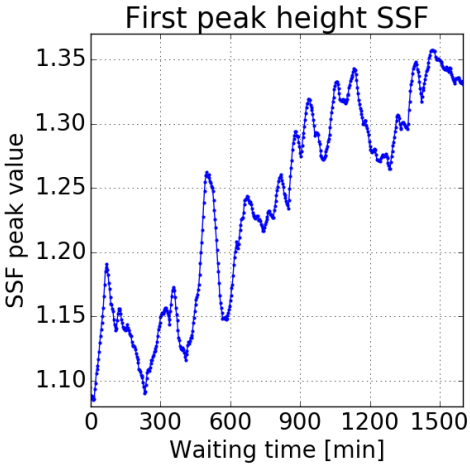

Figure 53: Height of the first peak of the static structure factor as a function of the waiting time.

5.3 Analysis of the dynamics

The mean squared displacement after four different waiting times is given in Figure 54. The graphs at different waiting times have a similar form in which the slope on the log-log plot is around 0.5 for times smaller than 10 min, then increases to a value around 1.3 and after 100 min weakens again. This form of the MSD is not comparable to typical MSDs for liquids and glasses (see Figure 11) where the slope is 2 at small times reflecting ballistic motion and 1 or smaller at large times reflecting diffusive motion. Several reasons might be possible for

this different behavior of the MSD in the experimental data. First, the ballistic regime might be at times smaller than 3 minutes and therefore not visible in our data. Furthermore, because of the many biological processes at play, the layer might possess migration phenomenon different than diffusion. And lastly, at large times the state of the system might have changed a lot resulting in different dynamics and therefore a weakening of the slope in the MSD.

Despite the odd behavior of the MSD for all waiting times, we see that the quantitative behavior differs depending on the waiting time. This can be measured via the diffusion coefficient and is displayed in the left top panel of Figure 55. This diffusion coefficient shows a slowing down of the dynamics upon aging, indicating that the system becomes increasingly jammed. Because we do not know the self-diffusion coefficient, we cannot calculate the effective diffusion coefficient as is given in the left top panel of Figure 44 for the VRD-model. For this reason, we do not know how close we are to the jamming transition by only looking at the the diffusion coefficient. Nevertheless, the absence of a plateau in the MSD indicates that we are still far away.

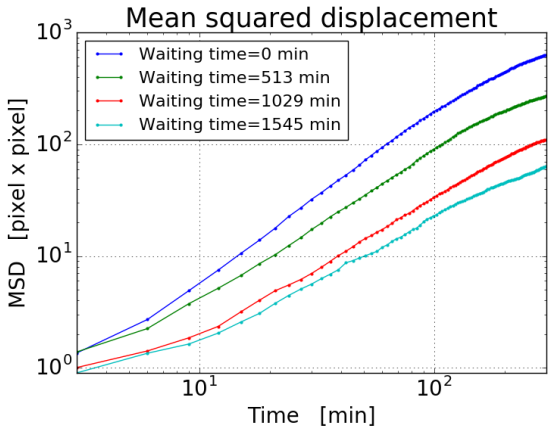

Figure 54: Mean squared displacement waiting 0, 513, 1029 and 1545 min.

The same observations can be made from the coherent and incoherent intermediate scattering function given in respectively Figure 56 and 57, where longer waiting times show a slowing down in the dynamics but a plateau stays absent. The relaxation times obtained for the coherent and incoherent ISF are respectively given in the left center and right center panel of Figure 55. Despite the existence of noise, which might or might not be caused by biological processes within the cell layer, the relaxation times grow steadily as a function of the waiting time.

A third parameter to define jamming is the average perimeter q proposed by Bi *et al.* [108] and is given in the right top panel of Figure 55 where the error bars define the standard deviation in the perimeter. This is not an error or uncertainty in the measurement of q according to Atia *et al.* [40], but are physically relevant as the standard deviation decreases upon reaching the jamming transition. Similar to the diffusion coefficient and the relaxation time, the results for the average perimeter q also suggests a slowing down of the dynamics upon waiting as q decreases in time, but the system is still far away from the critical threshold $q = 3.81\, r_0$. Furthermore, we see that the Voronoi tessellation results in slightly different cell shapes because the average perimeter obtained from the Voronoi tessellation is slightly different than the value directly obtained from the cell images. Although a good tool for cell models, this shows that Voronoi tessellations have to be used with care.

Besides identifying the dynamics in simulations and experiments, correlation functions and especially the relaxation time can also be used to compare the simulation model with experiments. Because many biological processes are at work during aging of the cell layer, it is hard to identify how each individual process influences the dynamics of the system. Effects that might slow down the dynamics include the increase in density, the decrease in the ability of cells to move on their own and the increase in cell-cell adhesion. The last two are accounted for in the VRD-model via decreasing respectively v_0 and p_0. In future experimental work it might be interesting to test these hypotheses by executing experiments under different circumstances that change the ability of cells to move or changes the cell-cell adhesion. With the VRD-model, predictions can be made about how this influences the relaxation times observed in experiments and thus enhance or undermine the agreement between the VRD-model and experiments.

In the next subsection we apply MCT to predict the dynamics and verify these predictions by comparing them with the results of this subsection.

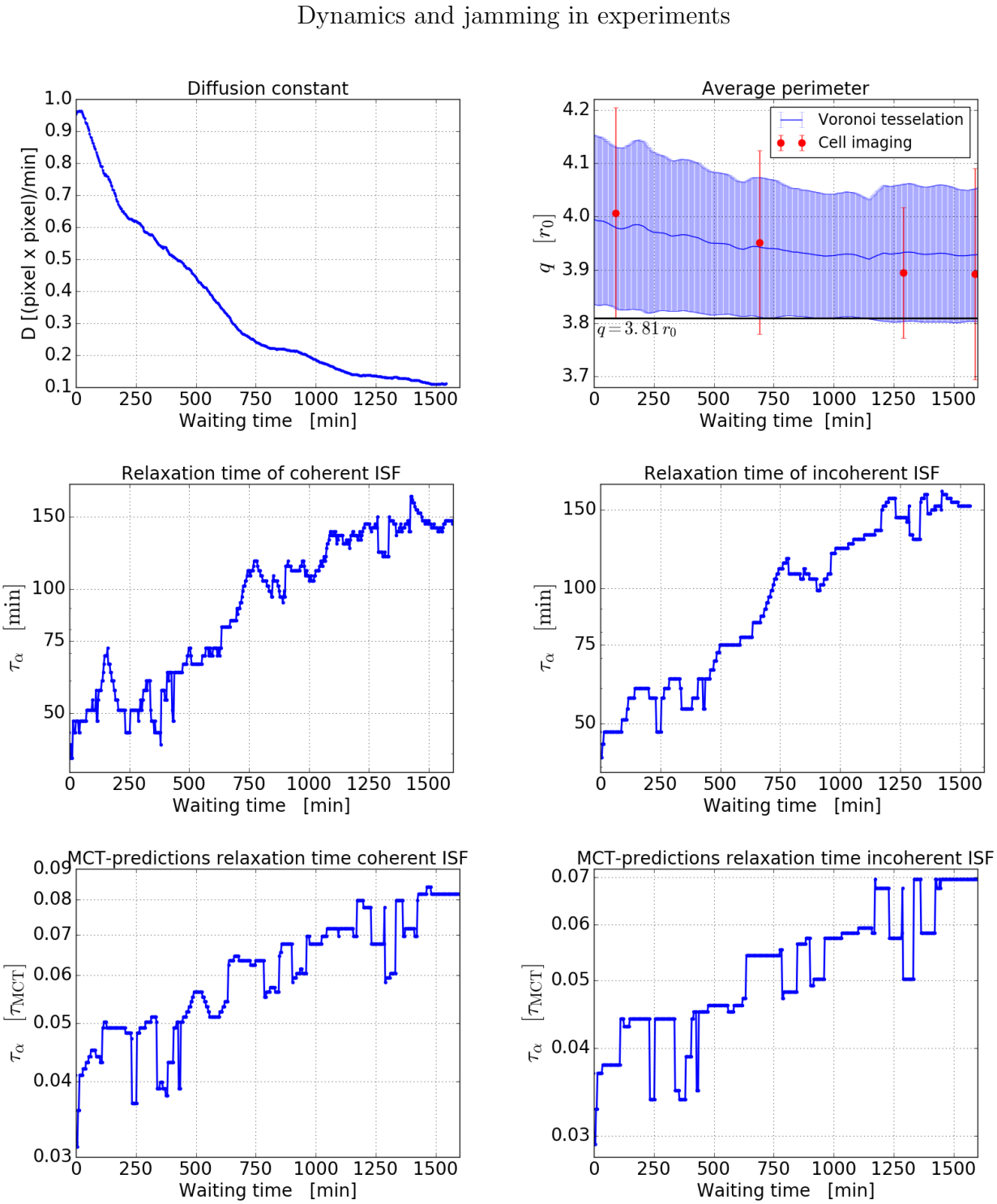

Figure 55: The diffusion coefficient (left upper panel), the average perimeter q (right upper panel), the relaxation time of the coherent ISF (left center panel), the relaxation time of the incoherent ISF (right center panel), the mode coupling predictions for the relaxation time of the coherent ISF (left bottom panel) and the mode coupling theory predictions for the relaxation time of the incoherent ISF (right bottom panel) as a function of the waiting time.

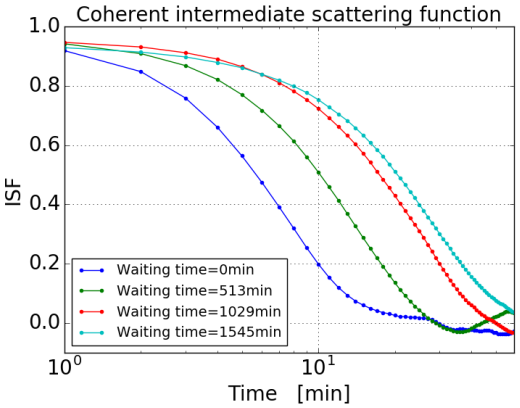

Figure 56: Coherent intermediate scattering function after waiting 0, 513, 1029 and 1545 min.

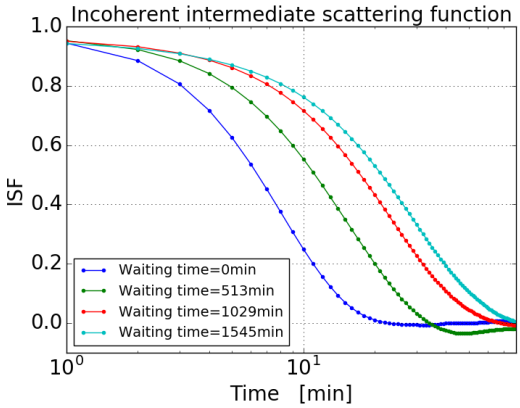

Figure 57: Incoherent intermediate scattering function after waiting 0, 513, 1029 and 1545 min.

5.4 MCT-predictions and verifications

Figures 58 and 59 give the mode coupling theory predictions for respectively the coherent and incoherent intermediate scattering function at the same waiting times as used in Figures 56 and 57. Just as for the MCT-predictions of the ISFs for the VRD-model in Figure 45, we do not see a plateau arising. Because we are still far away from the jamming transition and the peaks in the SSF (Figure 53) are relatively small, the vertices in the MCT-equation are not strong enough to predict a large caging effect. However, direct calculations of the ISFs in the experimental data (Figures 56 and 57) do not show this plateau either and the MCT-predictions are in this respect consistent with the direct calculations.

From these predictions we obtain the predicted relaxation times of the coherent and incoherent ISF in respectively the left and right bottom panels of Figure 55. These graphs show that the growth in the relaxation time as a function of the waiting time is captured correctly by MCT. Because we are far from the jamming transition, we cannot scale the relaxation times to compare the predicted and calculated values quantitatively as is done normally (see Subsection 2.3.1).

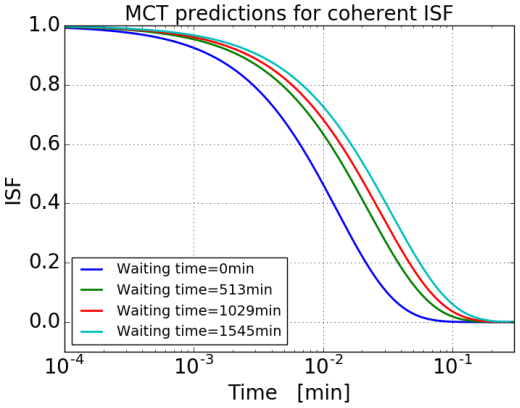

Figure 58: Mode coupling theory predictions for the coherent ISF.

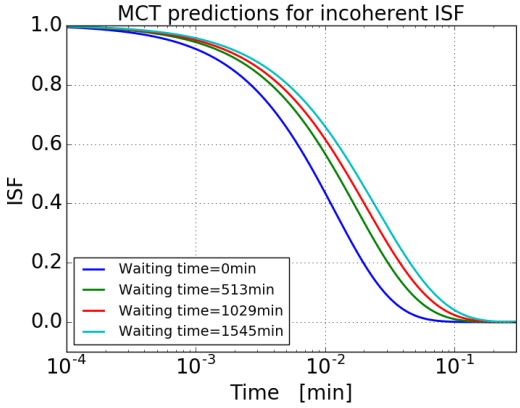

Figure 59: Mode coupling theory predictions for the incoherent ISF.

Although MCT predicts an increase in the relaxation time upon aging, these predictions contain large fluctuations. This is not a surprise, because the static structure factor, used as input for the MCT, also contains large fluctuations as illustrated by the peak heights in Figure 53. Furthermore, the direct calculations of the relaxation times in the center panels of Figure 55 show fluctuations too. Therefore, these fluctuations are not a failure of MCT and might even represent real fluctuations in the state of the cell layer. On a side note, the fluctuations in the SSF and the MCT predictions are not similar, meaning that a peak or valley in the SSF does not always coincides with a peak or valley in the MCT predictions and vice versa. This shows that MCT does not simply take into account the peak height of the SSF but the complete SSF to predict dynamics, just as was explained for the VRD-model in Subsection 4.4.

Summarizing, all graphs in Figure 55 show a slowing down of the dynamics when the layer ages and the mode coupling theory predictions for the relaxation time perfectly satisfies these observations. This provides us with a new way to measure dynamics in experiments based on only one or a few snapshots of the system. This method is advantageous with respect to methods that use the diffusion coefficient or relaxation time based on experimental measurements during a long period. Such long measurements are usually expensive and difficult or are even impossible because the state of the cell layer continuously changes.

So far we have neglected the self-diffusion coefficient D_0 because we cannot obtain its value from the experimental data. In the mode coupling theory we inserted $D_0 = 1.0\, r_0^2/\tau_0$ and a different value only scales the time according to Appendix D.2. However, it is likely that D_0 changes as well during proliferation. In Figures 60 and 61 we divided the relaxation time predictions of MCT by the relaxation time obtained via direct calculations for respectively the coherent and incoherent static structure factor. Assuming MCT predicts the relaxation time correctly up to a constant, this ratio should be a measure for the diffusion coefficient D_0. Unfortunately, we cannot test this because D_0 is unknown, but it is plausible that the self-diffusivity D_0 of cells decreases when the cell layer ages in agreement with the predictions of Figures 58 and 59. It is tempting to use this ratio as a measure for D_0 because direct measurement of its value is impossible, but this has to be done with great caution for two reasons. First the ratio also depends on a friction constant μ governing the short term dynamics and its value is unknown too. And secondly, this approach is based on one of the main failures of MCT, namely the prediction of the exact time scales (see Subsection 2.3.1 and the VRD-results of Subsection 4.4).

For the VRD-model several adaptations to mode coupling theory that possibly improve the predictions are summarized, including multidisperse MCT, generalized MCT and active MCT for self-propulsive systems. Cells are definitely self-propulsive, but it is unknown how this active force can be described in experimental data. One possibility is a force exhibiting only rotational diffusion as in the VRD-model [10] and another possibility is a force which aligns with its neighbors as will be discussed in Section 6. It would be very useful if MCT could be used to make predictions about the nature and properties of the active force in experiments. One approach would be to assume and test several MCT-predictions for different active forces and based on a comparison of the accuracy of the predictions conclusions can be made. Nevertheless, given the current accuracy of MCT there are a lot of steps to be taken first before MCT can possibly be used to make these predictions and it is doubtful whether MCT would be suitable for this at all.

Although adapting the MCT-equations to improve the predictions is very valuable, we think that the most interesting next step is to repeat the MCT-analysis with experimental data of different cell lines or under different circumstances. This would be especially useful in data that show a larger variation in dynamics and approach or even cross the jamming transition. This results in a much larger growth of the relaxation time and it would strengthen our conclusions of the applicability of standard mode coupling theory on experimental data of confluent cell layers.

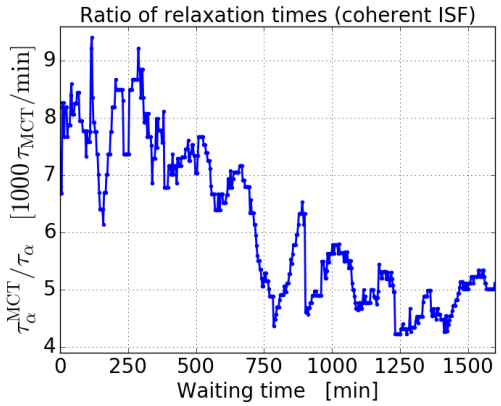

Figure 60: Ratio of predicted to calculated relaxation times of the coherent ISF.

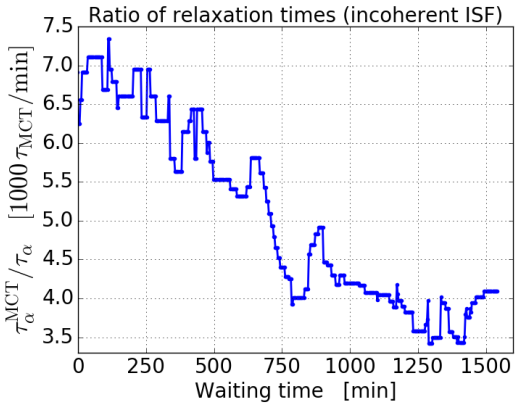

Figure 61: Ratio of predicted to calculated relaxation times of the incoherent ISF.

5.5 Conclusions

The radial distribution function and static structure factor show an increase in structural order in the confluent cell layer during aging. Despite this growth, the structural order is still low in comparison with hexagonal

ordered states that are reached in the VRD-model and the rich variation in the RDF and SSF that the VRD-model displays is not obtained. From the mean squared displacement and the intermediate scattering function we derive that this increase in structural order is accompanied by a slowing down of the dynamics, but that we are still far away from the jamming transition. Analysis of the average perimeter also suggests this slowing down. Which biological and physical processes are responsible for the change in dynamics is a complex question, but the fact that we can still investigate this dynamics without knowing the effects of these underlying processes is a powerful property of the correlation functions and mode coupling theory.

In this region far before the jamming transition, mode coupling theory captures the qualitative growth in relaxation times correctly. This verification of MCT on experimental data is in line with the verification of MCT in the VRD-model in the same region before the jamming transition (Section 4) and we can draw the same conclusions. Firstly, our analyses extend the applicability of mode coupling theory to a new field of physics. And secondly, this validity of MCT provides us with a theory that describes the physics of confluent cell layers and strengthens the analogy between confluent cell layers and non-living materials approaching a glass or jamming transition.

Besides this theoretical significance, MCT provides us as well with an alternative tool to get insights into the dynamics of a cell layer based on only one snapshots instead of many snapshots during a long period of time. This is a big advantage because in experiments it is often very expensive to measure during a long time window. Moreover, it might even be impossible to calculate the diffusion coefficient or relaxation times based on these long measurements because the state of living materials is continuously changing.

In this research, we only tested the MCT predictions in a limited range of relaxation times. Therefore, we suggest to repeat our analyses in the future using experimental data of cells that reach a further jammed state to verify our conclusions. In the next and last result section, we switch back to simulations and investigate the VV-model in which a different active force is incorporated than in the VRD-model.

6 Results of the Voronoi Vicsek model

In Subsection 3.4 we proposed a new model which combines the Self-Propelled Voronoi model [108] with a Vicsek-like alignment [145]. The equation of motion for this model is given by

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \mu \boldsymbol{F}_i^{\mathrm{int}} + v_0 \hat{\boldsymbol{n}}_i, \tag{77}$$

where the interaction force $\boldsymbol{F}^{\mathrm{int}}$ is given by

$$\boldsymbol{F}_i^{\mathrm{int}} = -\nabla_{\boldsymbol{r}_i} E = -\nabla_{\boldsymbol{r}_i} \left(\sum_i^N k_a (a_i - a_0)^2 + k_p (p_i - p_0)^2 \right), \tag{78}$$

and the time evolution of the polarity vector $\hat{\boldsymbol{n}}_i = (\cos(\theta_i), \sin(\theta_i))$ is given by

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \frac{1}{\tau_v} \left\langle \phi_j(t) - \theta_i(t) \right\rangle_{0<|\boldsymbol{r}_i - \boldsymbol{r}_j|<R_V} + \eta_i(t) \quad \text{with} \quad \langle \eta_i(t) \eta_j(t') \rangle = 2 D_r \delta(t - t') \delta_{ij} \tag{79}$$

with ϕ_j the angle of the velocity vector of cell j. This last equation distinguishes the VV-model from the VRD-model that does not contain the neighbor-alignment term.

The dynamics in the VV-model is regulated by the competition between the cell's tendency to obtain its preferred shape and to align its polarity with the velocity of its neighbors under influence of stochastic noise. The VV-model adds the Vicsek radius R_V and the Vicsek persistence time τ_V to the free parameters of the VRD-model, which are p_0, v_0 and D_r. The main purpose to investigate the VV-model is to see whether tuning of R_V could be a mechanism to influence the connectivity between cells and therefore their alignment in a similar way as observed in experiments. We first investigate the effect of R_V on the alignment of the cells in Subsection 6.1 and in Subsection 6.2 we perform a broad parameter analysis where we investigate how the five free parameters influence the alignment and the structure. In Subsection 6.3 we compare the results of the Voronoi Vicsek model regarding alignment to experiments that have inspired this model. Besides influencing the alignment, the Vicsek radius also effects the dynamics and jamming properties and this is demonstrated in Subsection 6.4. We end with summarizing the main conclusions in Subsection 6.5.

6.1 Phenomenological description

Figure 62 illustrates the effect of the Vicsek radius R_V on the velocity field (left) and the trajectories (right) of the cells. The Vicsek radius is $0.0\,r_0$, $1.0\,r_0$ and $3.0\,r_0$ for respectively the top, center and bottom panels, and in the left top corners of the velocity field the size of this Vicsek radius is illustrated with a circle. All velocity vectors ϕ that are within $\pi/4$ from the average velocity vector $\phi_{\mathrm{avg}} = \frac{1}{N}\sum_{i=1}^{N} \phi_i$ are marked green in the velocity field graphs. The velocity vectors that deviate between $\pi/4$ and $\pi/2$ from the average velocity vector are orange and all other vectors are marked red. The time is chosen such that the system has reached a steady state and waiting longer does not alter the alignment. We see that there is no long-range alignment for $R_V = 0.0\,r_0$, for which the VV-model is equal to the VRD-model. When we increase R_V to 1, we see the number of orange velocity vectors increases indicating that some alignment arises in the layer. For $R_V = 3.0\,r_0$ the velocity field turns green, which means that there is significant long-range alignment. This difference in alignment as a function of R_V influences the trajectories of the cells, as shown in the right graphs. The trajectories of 10 percent of the cells are drawn during $50\,\tau_0$. For $R_V = 0.0\,r_0$ the trajectories do not show alignment and all have different forms. For $R_V = 1.0\,r_0$, the cells move in one preferred direction, but the fluctuations are high. Ultimately, for $R_V = 3.0\,r_0$ the cells are highly aligned and they flow all in the same direction.

Figure 64 gives the alignment index AI as a function of time for several values of R_V. All lines start at 0 because the system is relaxed without an active force. At time $t = 0\,\tau_0$ the VV-force is turned on and if R_V is large enough, cells start to align until a certain value is reached. This value and the time it takes to reach it depends on R_V and we look further into this in Subsection 6.2.1. Figure 63 shows how this velocity alignment takes place for $R_V = 3.0\,r_0$. At three different times the velocity field is given and as time evolves (from left to right panel) the velocity vectors become more correlated indicating a higher alignment.

Now we have seen qualitatively that R_V can indeed be used to tune the alignment, we investigate how this alignment is influenced quantitatively by R_V and the other four free parameters in the next subsection.

Velocity field and trajectories as a function of the Vicsek radius

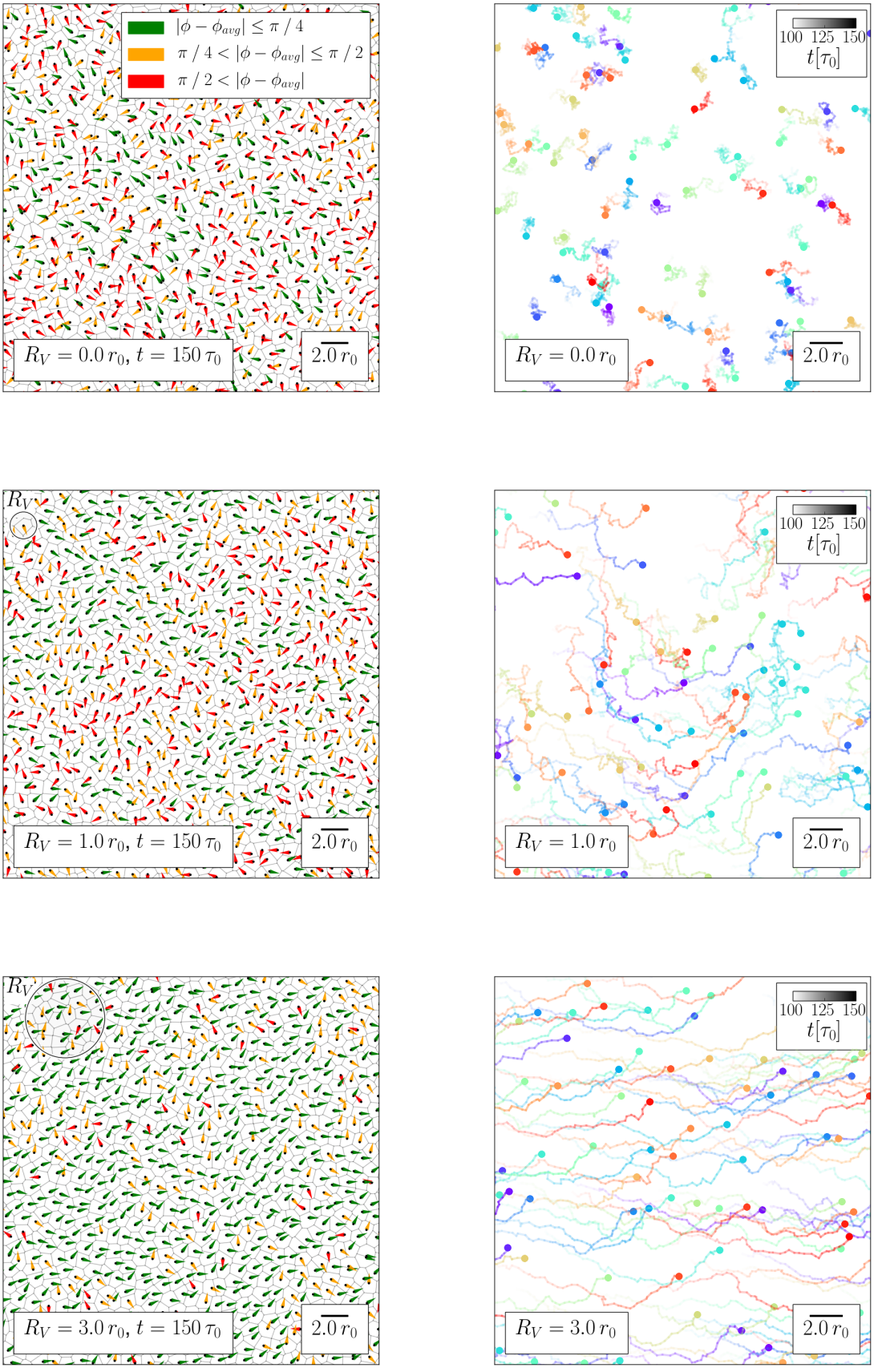

Figure 62: The velocity field (on the left side) and the trajectories (on the right side) for three different values of R_V: $0.0\, r_0$ (top), $1.0\, r_0$ (center) and $3.0\, r_0$ (bottom). For the trajectories, only ten percent of the cells are drawn. The other parameters are given by Setting A in Table 1.

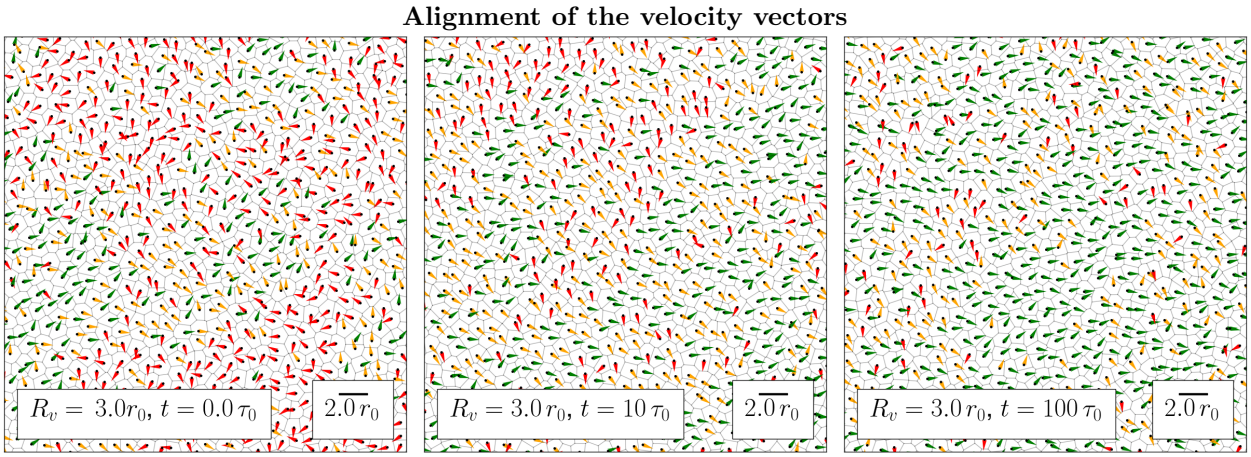

Figure 63: Velocity field at three different times: $t = 0.0\,\tau_0$ (left), $t = 50.0\,\tau_0$ (center) and $t = 100.0\,\tau_0$ (right) for $R_V = 3.0\,r_0$ and Setting A. The legend is defined in the left top panel of Figure 62

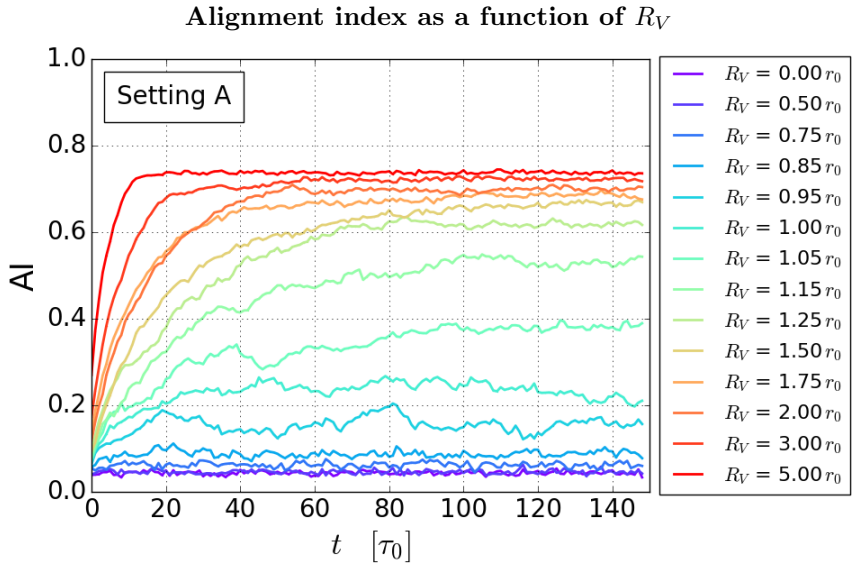

Figure 64: The effect of R_V on the alignment rate and the alignment value at $t \to \infty$ for Setting A in Table 1 (with R_V varying between $0.0\,r_0$ and $5.0\,r_0$).

6.2 Parameter analysis

In this subsection we focus on the effect of the five free parameters on the value of the alignment index at $t \to \infty$, the alignment time τ_{AI}, the average perimeter and the average velocity $\overline{v}$ at $t \to \infty$. We use Setting A given in Subsection 3.5 and check our conclusions with the two other Settings B and C. The parameter values of these settings are repeated in Table 2.

Table 2: Definition of the three different parameter settings A, B and C.

	$R_V(r_0)$	$p_0(r_0)$	$v_0(\tau_0/r_0)$	$D_r(\tau_0^{-1})$	$\tau_V(\tau_0)$	
Setting A	1.5	3.8	0.5	1.0	1.0	*Basic*
Setting B	1.5	**3.6**	**0.1**	1.0	1.0	*Further jammed*
Setting C	**2.5**	3.8	0.5	1	**0.1**	*More collectivity*

6.2.1 Effect of the Vicsek radius

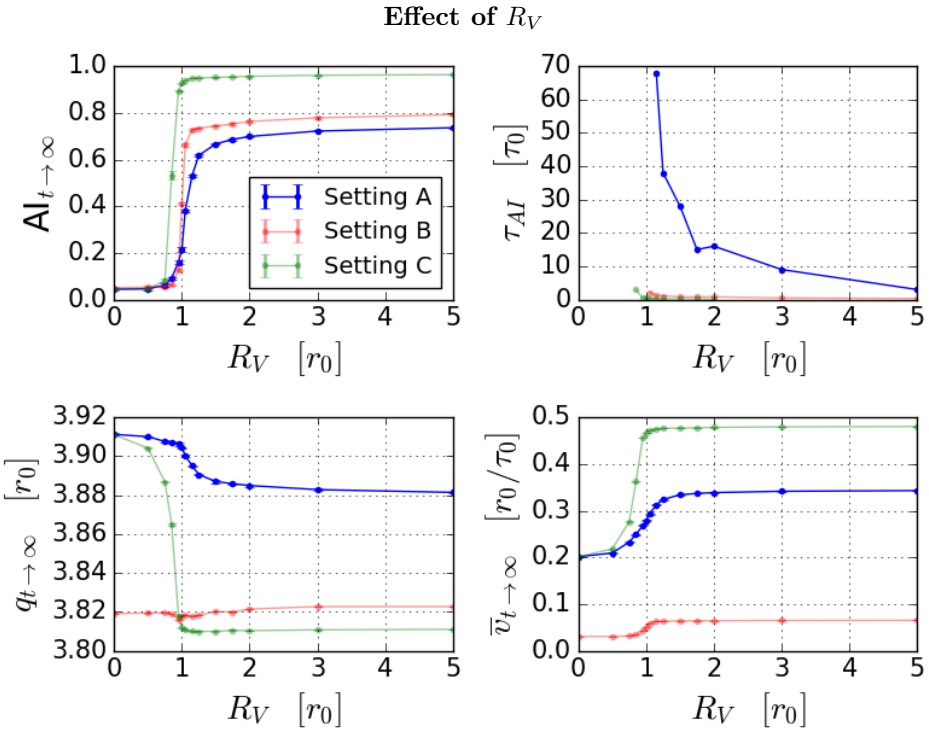

Figure 65: The influence of R_V on the alignment index (left top panel), the alignment time (right top panel), the average perimeter (let bottom panel) and the speed (right bottom panel) for Setting A, B and C in Table 2 and R_V varying between $0.0\,r_0$ and $5.0\,r_0$.

The left top panel of Figure 65 shows the effect of R_V on the long-time alignment index. There is no alignment for small R_V because there are no or a only few neighbors within the Vicsek radius. Around $R_V = 1.0\,r_0$ the alignment increases, almost switch-like, to a high value of alignment of which the specific value depends on the other parameters. In a hexagon tiling the distance between two neighbors is $1.07\,r_0$ (see Appendix D.3) and it is therefore worth mentioning that the transition in AI takes place around $R_V = 1.0\,r_0$. This means that the layer is driven away from an optimal packing structure causing some cells to deform and have seeds closer to each other. When R_V is increased further the value of AI does not change much, but the alignment rate is much faster as indicated by τ_{AI} in the right top panel of Figure 65. The reason for this increase in alignment rate is that each cell aligns its polarity with more neighbors.

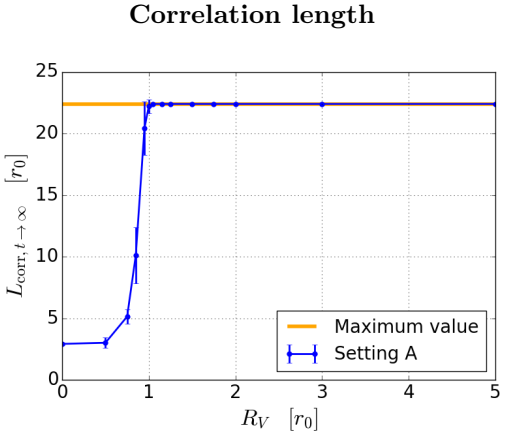

Figure 66: Correlation length as a function of R_V for Setting A in Table 2 and R_V varying between $0.0\,r_0$ and $5.0\,r_0$.

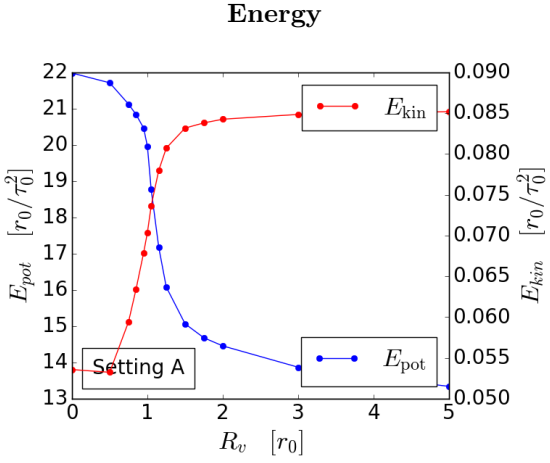

Figure 67: Potential and kinetic energy as a function of R_V for Setting A in Table 2 and R_V varying between $0.0\,r_0$ and $5.0\,r_0$.

The switch-like behavior as a function of R_V is also visible in the correlation length $L_{\text{corr},t\to\infty}$ in Figure 66. The correlation length goes to $22.4\,r_0$ when R_V is around one, which indicates that all cells are correlated as this is the maximum distance between two particles in a square box with periodic boundaries and area $A = 1000\,r_0^2$ (orange line). When the correlation length is smaller but nonzero, different clusters exists in the system in which the velocity is correlated but a long-distance correlation is lacking. The fact that the correlation length goes to its maximum possible value very rapidly suggests that there are finite-size effects at play. It might therefore be good in further research to verify our results in larger system sizes. Because the correlation length shows always similar behavior as AI in our analysis, we will not look at the correlation length in the remainder of this analysis anymore and focus on the AI.

An increase in R_V promotes collective behavior and a flow arises when R_V is large enough and cells migrate over large distances. This is reflected in an increase in the average speed $\overline{v}$ around $R_V = 1.0\,r_0$ (right bottom panel of Figure 65) and in a larger average squared displacement (Figure 68). Note that this squared displacement is unequal to the MSD, because the average displacement of the cellular collective is not subtracted. The reason for these large displacements is that aligned cells have the ability to move efficiently in a flock in which the configuration barely changes. As a consequence, the displacements are mostly regulated by the aligned self-propelled forces and barely disrupted by the interaction force. This results in an increase in the kinetic energy, while the potential energy decreases because cells are getting closer to their preferred shape (see Figure 67).

The left bottom panel of Figure 65 shows q as a function of the Vicsek radius and show that for Setting A and C q decreases around $R_V = 1.0\,r_0$. The average perimeter even reaches $q = 3.81\,r_0$ in Setting C, which is the jamming criterion according to Bi *et al.* [108]. These results are in line with the conclusions of Giavazzi *et al.* [149] that flocking promotes solidification. Their explanation for the decrease in q when the alignment increases is as follows. When there is no alignment, the motion is Brownian. This gives rise to an effective temperature that can help cells to overcome energy barriers. When R_V increases and cells align more, this effective temperature decreases resulting in more strongly caged cells as it becomes more difficult to overcome energy barriers. In Subsection 6.4 we look further into the effect of R_V on the dynamics and jamming and will see that the dependence of jamming on R_V is more complex than q suggests. In conclusion, around $R_V = 1.0\,r_0$ the alignment increase from almost zero to a finite value depending on the other parameters and increasing R_V to larger values increases the alignment rate but does not increase the alignment further.

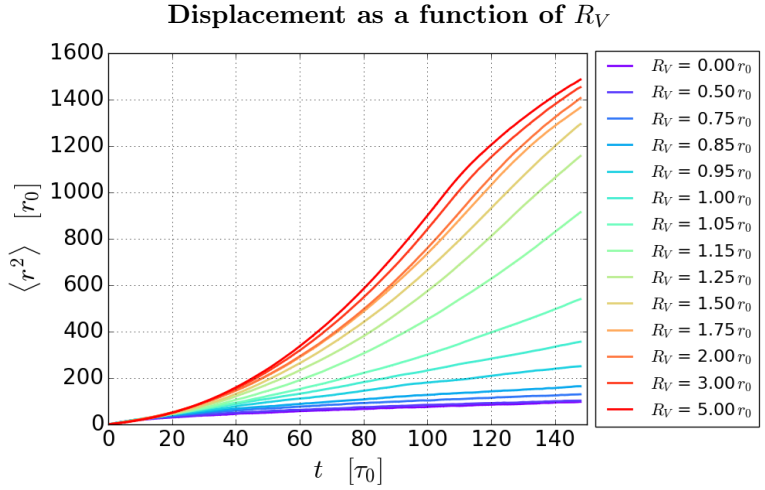

Figure 68: R_V influences the mean squared displacement (where the average collective displacement is not subtracted) for Setting A given in Table 2 and R_V varying between $0.0\,r_0$ and $5.0\,r_0$.

6.2.2 Effect of the preferred perimeter

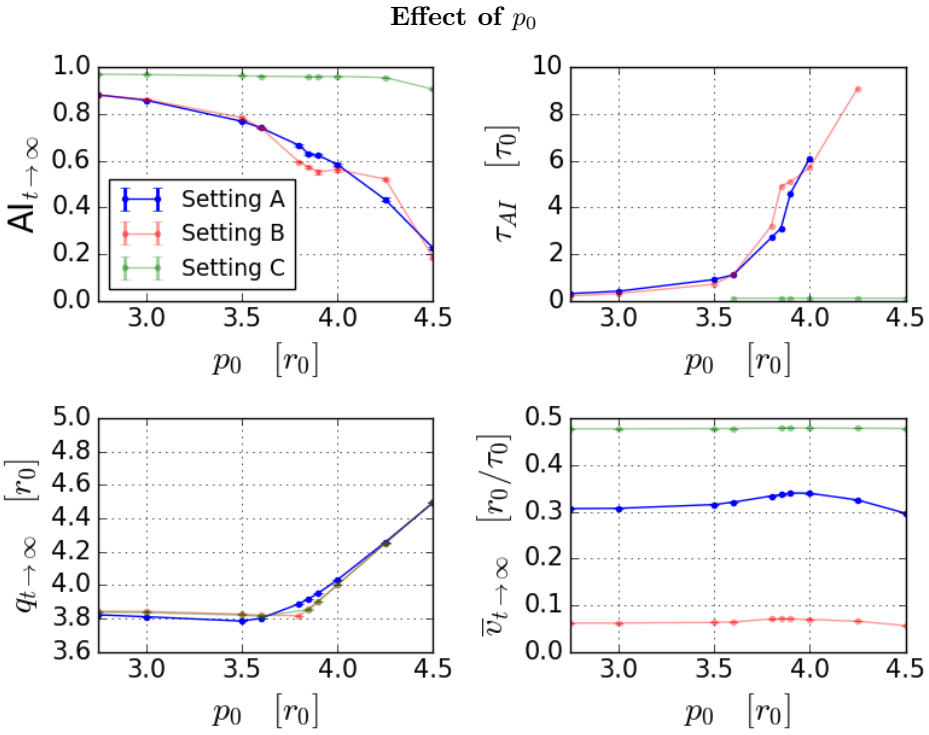

Figure 69: The influence of p_0 on the alignment index (left top panel), the alignment time (right top panel), the average perimeter (let bottom panel) and the speed (right bottom panel) for Setting A, B and C in Table 2 and p_0 varying between $2.75\,r_0$ and $4.5\,r_0$.

The left top panel of Figure 69 shows that decreasing p_0 leads to a higher alignment. This is accompanied by respectively an increase or decrease in the alignment rate according to the right panel of Figure 69. Most interesting is the increase from $p_0 = 3.5\,r_0$ till $p_0 = 4.0\,r_0$ as that is the most physical regime (see Subsection 4.1). In Section 4 is already showed that increasing p_0 promotes fluidification in the VRD-model and the graph of q in the left bottom panel of Figure 69 shows that this also holds in the VV-model at and above the jamming transition. When a system becomes more fluid-like, cells are moving more and more randomly and this opposes alignment. Due to this smaller alignment, the speed of the cells also slightly decreases (left bottom panel of Figure 69). In Subsection 6.2.1 we showed that flocking promotes solidification and these results also suggests the opposite statement that solidification promotes flocking. For $p_0 > 4.2\,r_0$ another effect arises which counteracts velocity alignment, namely that the seeds of the cells are sticking to other seeds (see right panel of Figure 38). This can results in clustering of cell seeds in which cells do not have neighbors within R_V outside their cluster, which causes the alignment to be only local and the alignment index to decrease further. This effect is slightly reduced by the fact that cells belonging to different clusters still influence each other via the interaction potential E and consequently the alignment index has still a finite value. In conclusion, increasing the preferred perimeter decreases the alignment and increases the alignment rate.

6.2.3 Effect of the velocity

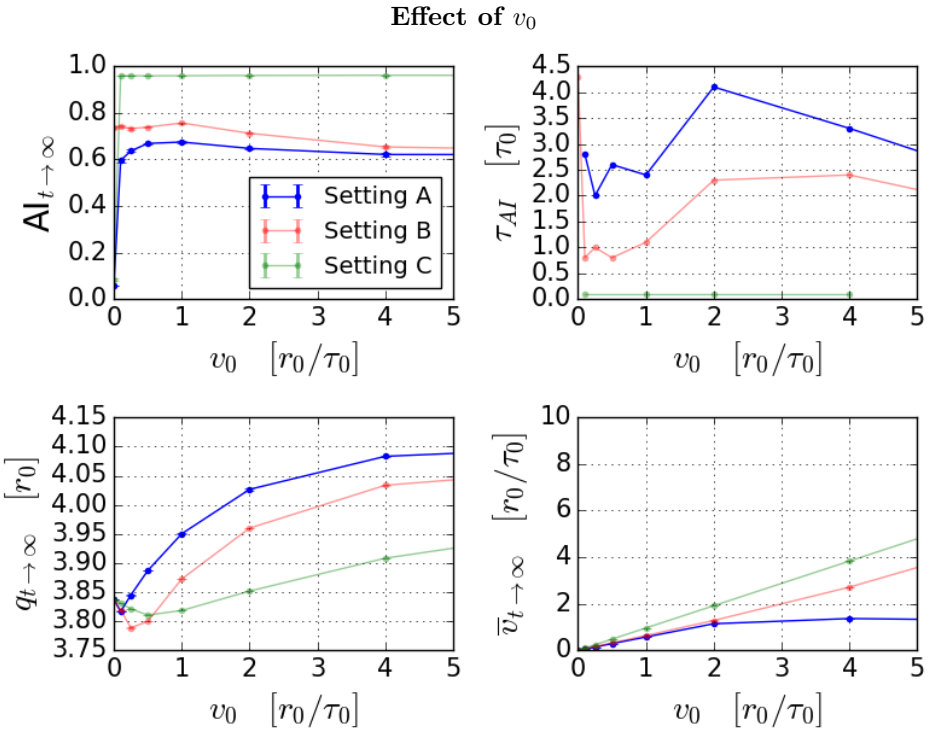

Figure 70: The influence of v_0 on the alignment index (left top panel), the alignment time (right top panel), the average perimeter (let bottom panel) and the speed (right bottom panel) for Setting A, B and C in Table 2 and v_0 varying between $0.0\,r_0/\tau_0$ and $5.0\,r_0/\tau_0$.

The left panel of Figure 70 shows the effect of v_0 on the alignment. For $v_0 = 0.0\,r_0/\tau_0$ the self-propulsion force is absent and no alignment takes place. Setting A and B has a maximum in the alignment around $v_0 = 1.0\,r_0/\tau_0$ for the following reason. There is always an interplay between the self-propulsion force that tends to align the cells and the interaction force that tends to optimize the shape of the particles. At the maximum of AI, this self-propulsion force is strongest in relation to the interaction force. For smaller values of v_0, the self-propulsion force becomes weaker and for larger values of v_0 the interaction force becomes stronger as the cell shapes are more disrupted by the larger displacements due to the self-propulsion force. These less optimal cell shapes for large v_0 is shown in the left bottom panel of Figure 70 where q grows when v_0 grows (for $v_0 > 1.0\,r_0/\tau_0$). The interaction force tries to recover the more optimal shapes and this causes cells to move in different directions than their alignment and this accounts for the decrease in AI for $R_V > 1\,r_0$ (in Setting A and B). For very high velocities ($v_0 > 4.0\,r_0/\tau_0$) the alignment does not decrease further while the alignment rate increases (right top panel of Figure 70). We expect this faster alignment to be caused by the fact that cells see more different neighbors due to the higher instantaneous velocity. The right bottom panel of Figure 70 shows that for Setting A a higher v_0 does not result in a higher velocity while it does so for Setting B and C. The reason for this opposite behavior of Setting A is that the persistence in the movement of the cells is lower because the system is more fluid-like than in the other settings.

The left bottom panel of 70 shows furthermore that the Vicsek force can lower q for small v_0 in comparison to the average perimeter at $v_0 = 0.0\,r_0/\tau_0$. Activation of the self-propulsion force lowers the energy barriers for cell rearrangements as more pathways become dynamically accessible. This can help the cells to optimize their cell shape, which raises the energy barriers again and causing cells to be trapped and eventually, when the caging is strong enough, to be jammed when $q < 3.81\,r_0$. When v_0 increases further, the self-propulsion force can help the cells to escape their cage again and this results in an increase in q. Sussman and Merkel [144] find that there is no unjamming transition in the athermal limit of the SPV-model (see Subsection 3.3.2), where athermal means that there is no active force via setting $v_0 = 0.0\,r_0/\tau_0$, while the results of Bi *et al.* [108] for small values had suggested that there is a transition. In Figure 70 we also see how the athermal limit differs from higher values of v_0, in line with the conclusions of Sussman and Merkel [144] for the SPV-model to which the VV-model reduces in the athermal limit. In conclusion, alignment only takes place when the velocity v_0 has a finite value. Increasing v_0 does not change the alignment much, but changes the alignment rate non-monotonously.

6.2.4 Effect of the rotational diffusion

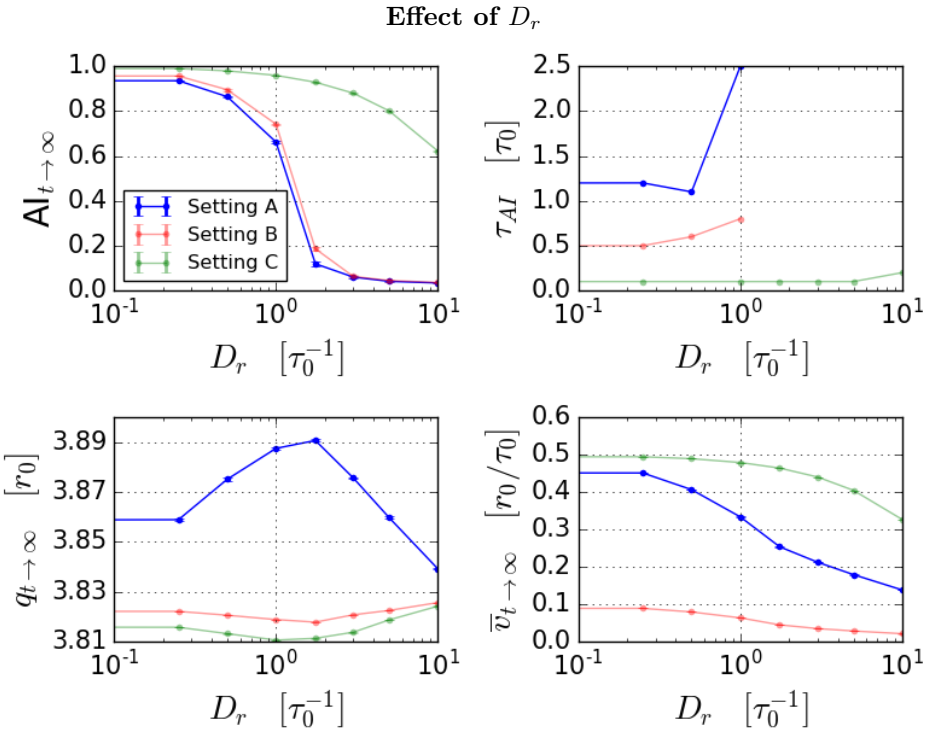

Figure 71: The influence of D_r on the alignment index (left top panel), the alignment time (right top panel), the average perimeter (let bottom panel) and the speed (right bottom panel) for Setting A, B and C in Table 2 and D_r varying between $0.1\,\tau_0^{-1}$ and $10\,\tau_0^{-1}$.

The left top panel of Figure 71 shows the effect of D_r on the alignment index. For small values of the noise, the alignment goes to almost 1. AI does not become exactly 1 because of the influence of the interaction potential. When D_r increases the alignment index increases and the rate decreases (right top panel of Figure 71). At larger D_r cells have less time to align before becoming disrupted by rotational diffusion and this results in a smaller AI as well. The resulting value of AI is an interplay between D_r and the Vicsek persistence time (see Subsection 6.2.5), because a smaller persistence time causes cells to align faster and counteracts the effect of an increase in noise and vice versa. Setting C has a smaller τ_V and a higher R_V than Setting A and this results in a higher AI. The left bottom panel of Figure 71 shows a decrease in velocity upon increasing D_r because cell motions are less persistent.

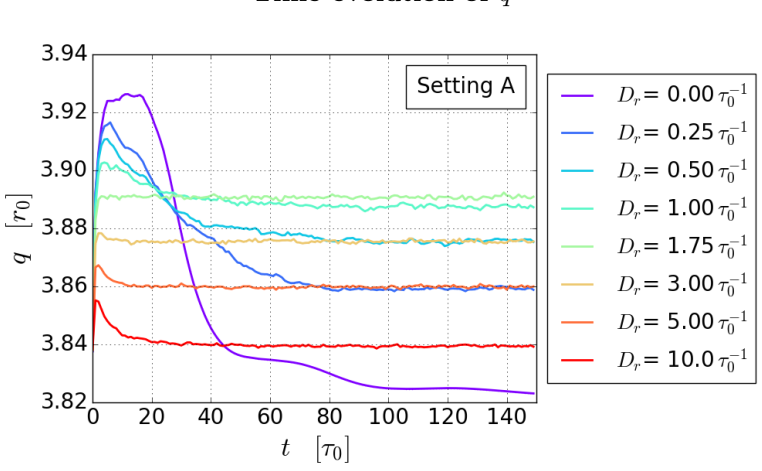

Figure 72: Time evolution of q as a function of D_r for Setting A in Table 2 and D_r varying between $0.0\,\tau_0^{-1}$ and $10\,\tau_0^{-1}$.

The left bottom panel of 71 shows q as a function of D_r and the results differ for the different settings. Focusing on Setting A we can distinguish two regimes. For the regime until roughly $D_r = 1.0\,\tau_0^{-1}$ an increase in the rotational diffusion results in an increase in q and increasing D_r in the regime $D_r > 1\,\tau_0^{-1}$ has an opposite effect. These differences in behavior can be explained with Figure 72, which shows the time evolution of q. At $t = 0.0\,\tau_0$ the active force is turned on and this always results in an increase of q because the cells are not aligned yet. The smaller D_r the bigger this increase because of more persistent motion of the cells resulting in larger shape deformations. At large D_r in comparison to other inverse time scales, the motion is more Brownian and this promotes a smaller q-value. But what stands out in Figure 72 is that for small D_r the perimeter eventually reaches a smaller value than for the large values of D_r. All cells align for small D_r and move as one flock in which the self-propulsion does no longer influence the shape of each cell and the interaction force has no competition in optimizing the shape anymore. This non-monotonous behavior of q shows that there are two similar states possible regarding the shape, but with different dynamics. The cellular collective is stationary in one state and flowing in the other state. This is the same distinction as Giavazzi *et al.* [149] make between a flocking liquid and a stationary liquid and is illustrated by Malinverno *et al.* [151] in Figure 36. In conclusion, an increase in D_r decreases the alignment while the alignment rate decreases and has a non-monotonous effect on the average perimeter q.

6.2.5 Effect of the Vicsek persistence time

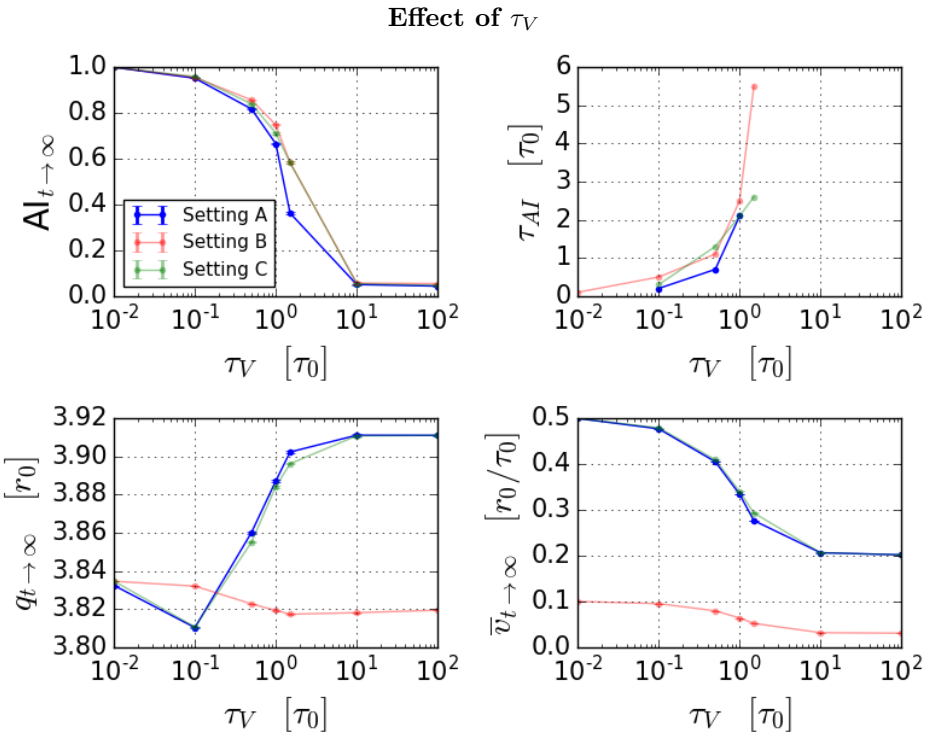

Figure 73: The influence of τ_V on the alignment index (left top panel), the alignment time (right top panel), the average perimeter (let bottom panel) and the speed (right bottom panel) for Setting A, B and C in Table 2 and D_r varying between $0.01\,\tau_0$ and $100\,\tau_0$.

The left top panel of Figure 73 shows the effect of τ_V on the alignment index. When τ_V is the smallest time scale in the system, the alignment is almost one. The alignment decreases when τ_V increases because other time scales disrupting the alignment become more important, like the time scale of rotational diffusion D_r^{-1}. The alignment rate decreases together with the alignment index as shown in the right top panel of Figure 73. Also the speed decreases (right bottom panel of Figure 73) because cells move more randomly resulting in a smaller net displacement.

The effect of τ_V on q in left bottom panel of Figure 73 differs for the three Settings A, B and C. Focusing on Setting A shows that in general q decreases when τ_V decreases and vice versa, except for very small values of τ_V. This confirms the statement in Subsection 6.2.1 that flocking promotes solidification and the same arguments hold in this case. Tuning τ_V is a way to reach the jamming transition according to the threshold $q = 3.81\,r_0$. q suggests an unjammed state for very small τ_V when there are barely random movements that can help cells to

decrease their perimeter. Nevertheless, analysis of the diffusion coefficient or relaxation time has to be executed to determine whether this is really an unjammed state or that q is not a good measure in this regime.

The effect of changing D_r in Figure 71 looks similar to the effect of changing τ_V in Figure 73. Although the physical origin of both parameters is different, they both influence the persistence of the cells' motion. Furthermore, the combination of D_r and τ_V is dimensionless. This raises the question whether the product $D_r\tau_V$ can be treated as one control variable; further analysis is needed to verify this hypothesis. In conclusion, an increase in τ_V decreases the alignment while the alignment rate decreases, which is similar to the effect of D_r.

6.3 Comparison to experiments

The Voronoi Vicsek model is inspired by experiments of Lång *et al.* [146] that show that a depletion of calcium can decreases the velocity alignment of cells and in this subsection we explain how these experimental results can be compared to our VV-model. In these experiments, HaCaT keratinocytes cells derived from human epidermis are used. After starvation for 48 hours in different calcium concentrations, they are activated with epidermal growth factor (EGF). Before looking at the effect of calcium, the effect of this EGF is examined. Normally, quiescent epidermal cells get activated when they get in touch with blood serum. The serum signals the cells that they have to close a wound and also provide nutrition to do so. The function of EGF is to mimic the effect of the serum as is shown in the left and right top panels of Figure 74. Both the speed (left panel) and the alignment (right panel) increase when the serum is added and addition of EGF shows similar behavior. In our simulations the addition of EGF is represented by turning on the self-propulsion force by giving v_0 a finite value. This is shown in Figure 75, where the purple line has $v_0 = 0\, r_0/\tau_0$ and the red line $v_0 = 0.5\, r_0/\tau_0$.

Experimental results for speed and alignment

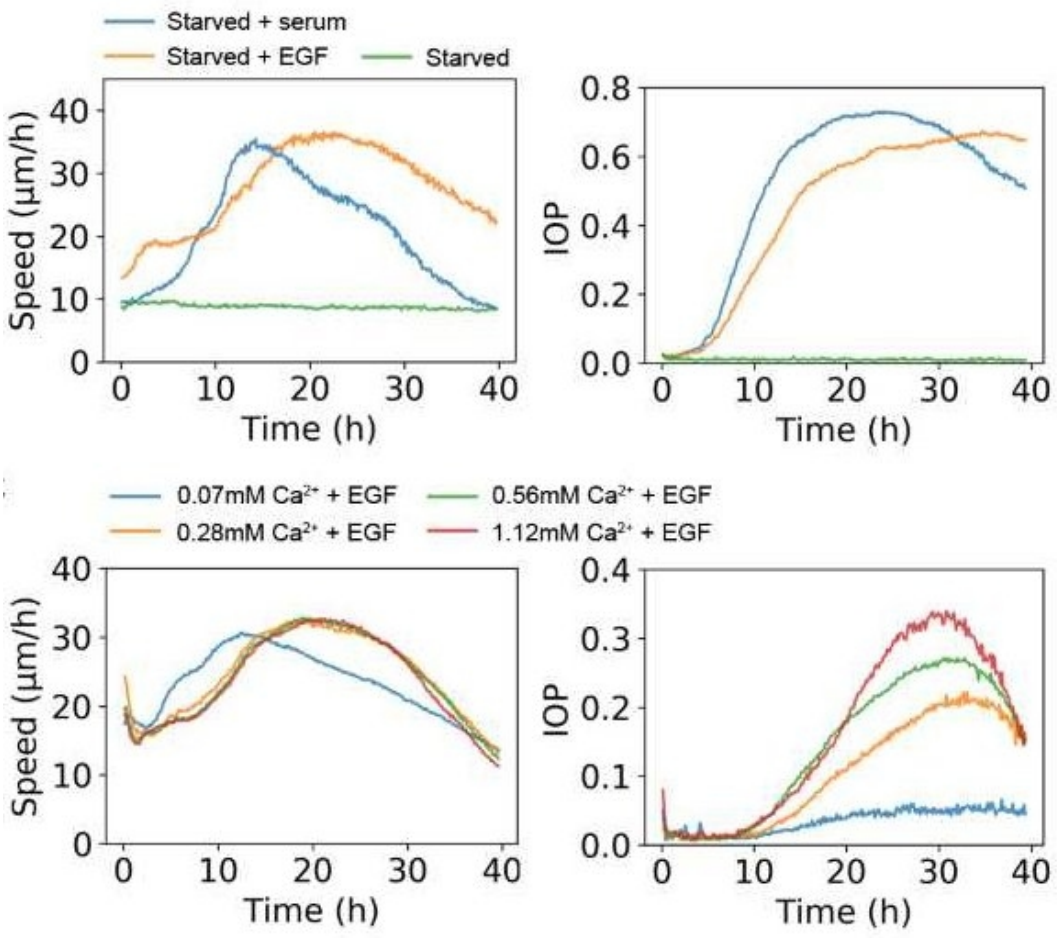

Figure 74: Experimental results on HaCaT keratinocytes cells. Serum and EGF have similar effects on the speed (left top panel) and the IOP (right top panel). The calcium concentration does not influence the speed (left bottom panel), but does influence the IOP (right bottom panel) [146].

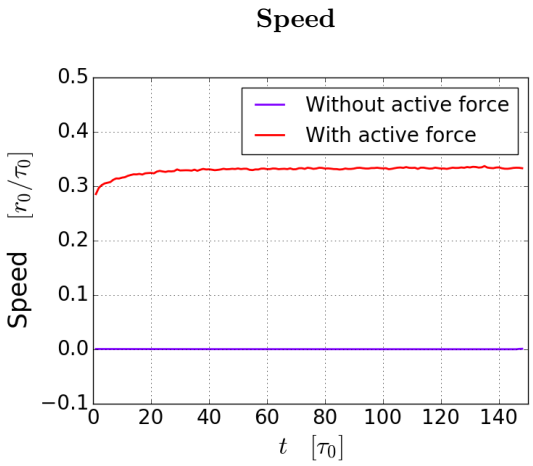

Figure 75: The speed in absence ($v_0 = 0.0\, r_0/\tau_0$) and presence ($v_0 = 0.5\, r_0/\tau_0$) of a self-propulsion force (and other parameters given by Setting A in Table 2).

The effect of calcium in experiments is examined by exposing the cells to a low calcium concentration prior to starvation. It is known that calcium plays a critical role in promoting cell-cell adhesion and depleting it reduces the cell-cell connectivity. Lång *et al.* [146] shows that depletion of calcium decreases the collective behavior (right bottom panel of Figure 74), while it does barely influence the speed (right bottom panel of Figure 74). It is yet unknown how cells align their velocities and our hypothesis is that the mechanism behind this alignment can be described with the VV-model. In this model the depletion of calcium is represented by a decrease of the Vicsek radius R_V and Figure 64 shows indeed this intended effect of the Vicsek radius on AI. Although Appendix G.6 showed that the AI and IOP behave similarly, we show in the right panel of Figure 76 the IOP for clarification (and during a shorter time period). Because the IOP depends also on other parameters than R_V, we cannot derive to which R_V the experimental data corresponds. However, it is likely that only a subsets of the R_V values used in Figure 76 is captured by the experimental data and the left panel of Figure 76 shows that for a small range of R_V the speed barely changes.

In our simulations v_0 is increased instantaneously at time $t = 0.0\, \tau_0$ causing an immediate increase in alignment and speed in Figure 76. However, the experimental results of Figure 74 show a much slower build up of the alignment. Our hypothesis is that cells require some time to assimilate the EGF and to respond to the new situation. A possible way to improve our simulations might be incorporating this response time by making the parameters v_0 and R_V time-dependent such that at small values of t these parameters are still very small and they increase gradually to their finite value. We expect this adaptation to slow down the alignment rate in the right panel of Figure 76. Furthermore, the experimental results in Figure 74 show that after some finite time the alignment and speed decreases again. One hypothesis is that this is caused by the depletion of EGF or other biological substances necessary for the cells to move and this might be imitated by tuning down v_0 and R_V again. Additionally, it is also interesting to see whether the alignment in experiments can increase further when cells are being fed during a longer period.

In addition to changing R_V, there are many ways to tune the alignment of the cell layer and to get identical situations as shown in Subsection 6.2. Nevertheless, not all ways are experimentally available. R_V has its representation in biological experiments by varying calcium, but it is questionable how easy it is to change, for example, D_r or τ_V in experiments. Furthermore, p_0 is thought to be a function of the cell-cell adhesion (see Subsection 3.3.1) and an decrease of the adhesion is reflected by a decrease in p_0. This results in an increase in the alignment of the VV-model according to 6.2.2, while in the experiments of Lang [146] a decrease in adhesion causes a decrease in alignment. This contradiction does not show that the VV-model is incorrect, but that only taking into account the effect of adhesion on p_0 and neglecting R_V is not sufficient to describe the alignment. Moreover, in this light it is an interesting question whether it is sufficient to mimic the effect of calcium by only tuning R_V and not by tuning a set of multiple parameters like R_V and p_0 at the same time. Nevertheless, our results show that by only varying R_V the main characteristics of the experiments with calcium are captured.

Now we have shown that the VV-model can qualitatively account for the effect of calcium in experiments, it is interesting to see whether the model can also provide quantitative information about the experiments such as with how many cells a cell interacts on average, which is measured by R_V. Because the alignment depends on the combination of all parameters in the VV-model, this requires first more knowledge about the values of these parameters in experiments.

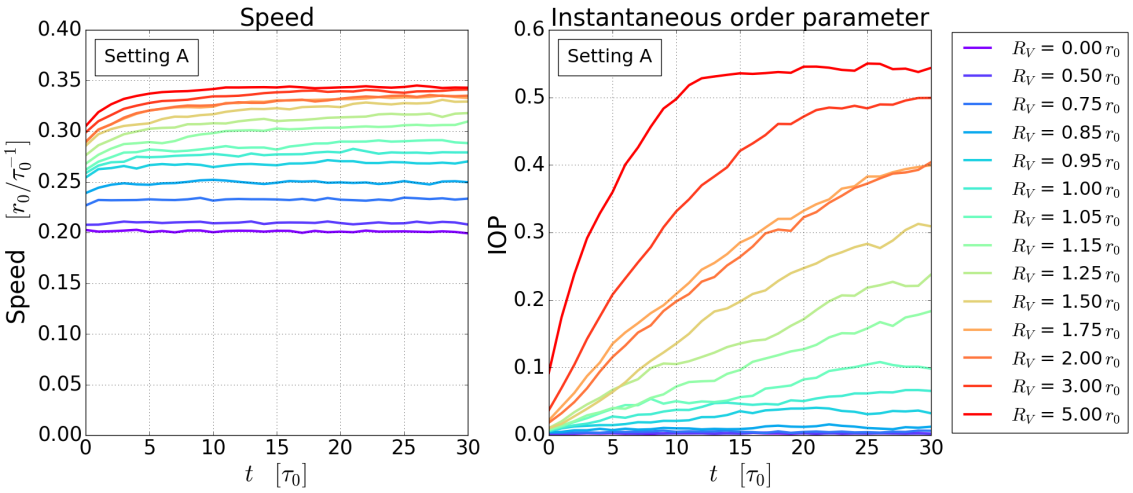

Figure 76: The speed (left panel) and IOP (right panel) as a function of time for Setting A in Table 2 and R_V varying between $0.0\,r_0$ and $5.0\,r_0$.

6.4 Jamming analysis

In addition to tuning alignment, the Vicsek radius can also be used to tune dynamics in the VV-model. In analogy with our analysis of the VRD-model and experimental data, we demonstrate this using correlation functions. We have chosen to use the Vicsek radius R_V together with the preferred perimeter p_0 as control parameters, because we have shown in the previous subsections that R_V is a powerful parameter to tune alignment in the VV-model and in Subsection 4.3 that p_0 dominates the dynamical behavior in the VRD-model. In Subsection 6.4.1 we show how R_V effects the intermediate scattering function and how this relates to alignment. The combination of jamming and alignment results in a rich phase diagram, which we discuss in Subsection 6.4.2.

6.4.1 Effect of the Vicsek radius on the dynamics

In Figure 77 the coherent and incoherent intermediate scattering function are drawn for several values of the Vicsek radius R_V, with the preferred perimeter p_0 set to $3.55\,r_0$ and the velocity v_0 to $0.25\,r_0/\tau_0$, corresponding to a jammed state in the VRD-model (see Figure 33). This jammed state of the VRD-model corresponds to the bold purple line in Figure 77, where $R_V = 0.0\,r_0$. When the Vicsek radius is increased to $1.0\,r_0$ the layer becomes more fluid-like, which is shown by the bold cyan line. Increasing the Vicsek radius even further slows down the dynamics again as shown by, for example, the bold red line for $R_V = 3.0\,r_0$. The alignment index corresponding to the values of R_V at which the ISFs are evaluated is drawn in the left bottom panel of Figure 77 and we conclude that the layer is most fluid around the transition point between zero and a finite alignment. This fluidification caused by R_V is not limited to the jammed regime as we show in Appendix J.1 where we evaluate the intermediate scattering function for $p_0 = 3.8\,r_0$, corresponding to an unjammed state in the VRD-model (see Figure 33). This intermediate scattering function decays faster when the Vicsek radius is around $1.0\,r_0$ in comparison with smaller or higher values of the Vicsek radius.

The fluidification or increase in dynamics caused by the Vicsek radius can be understood as follows. A finite value of R_V adds persistence to the motion of the cells due to alignment of the cell with its neighbors. This persistence counteracts the Brownian noise in the polarity vector, which is given by D_r, and the effect of the neighbor-alignment can therefore be seen as an effective decrease in the Brownian noise. Bi *et al.* [108] show that such a decrease can make a layer more fluid-like in the VRD-model, because a persistence in the motion of cells helps to escape from caging by their neighbors and therefore lowers the energy barriers for structural rearrangements. However, when R_V is so high that all cells are aligned in the same direction, the layer moves essentially as a flock. In this flock the self-propulsion force does not help the cells to escape from the cage of their neighbors because they move all cells in the same direction and consequently the energy barriers for structural rearrangements increase again, resulting in a jammed state. These three phases are illustrated in Figure 78 where the trajectories of the cells are drawn in the comoving frame. The left panel shows a so-called stationary

flock for $R_V = 0.0\,r_0$ and the right panel a stationary solid for $R_V = 3.0\,r_0$. In the center panel ($R_V = 1.0\,r_0$) a transition between these two states is captured in which individual cells travel over longer distances, indicating a more fluid-like state.

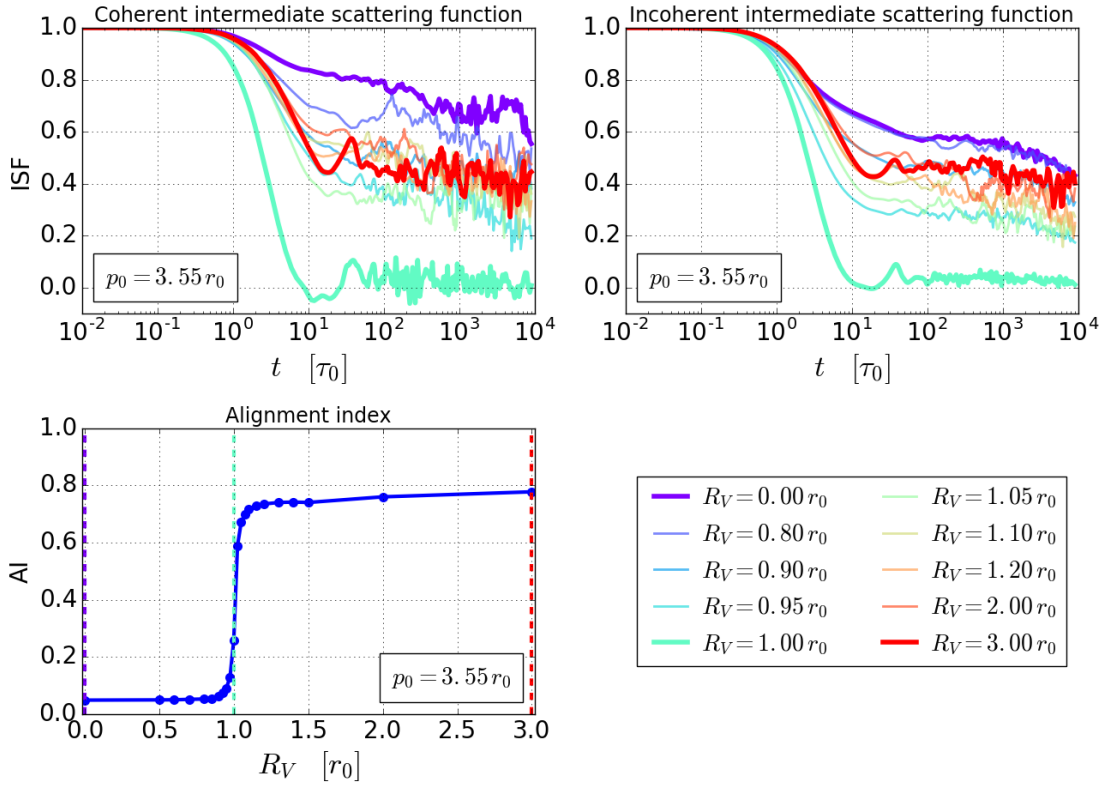

Figure 77: The coherent ISF (top left panel), the incoherent ISF (top right panel) and the alignment index (left bottom panel) for $p_0 = 3.55\,r_0$ and R_V varying from $0.0\,r_0$ till $3.0\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).

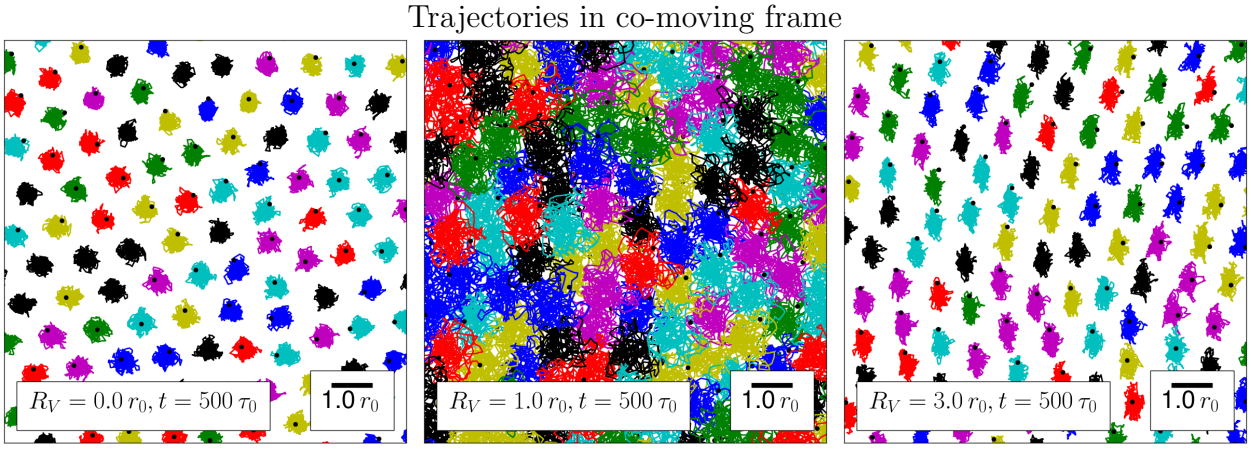

Figure 78: Trajectories in the co-moving frame for $R_V = 0.0\,r_0$ (left panel), $R_V = 1.0\,r_0$ (center panel) and $R_V = 3.0\,r_0$ (right panel) with $p_0 = 3.55\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).

This increase of the dynamics by R_V does not correspond to an increase of the average perimeter q according to Figure 68 in Subsection 6.2.1, which underlines the limited applicability of q as a measure for the dynamics in the VV-model. In Appendix J.2 we repeat this analysis of q but for the same parameters as in Figure 77. The reason that q does not captures the increase in dynamics around R_V is that this change is accompanied by an increase in structure as probed by the RDF and SSF (see Appendix J.2). This is in contrast with our observations in the VRD-model, where we concluded that an increase in structure indicates a slowing down of the dynamics. Apparently, the structural rearrangements in the VV-model for a value of R_V around $1.0\,r_0$ also

make lower energy configurations dynamically accessible. This would imply that the standard MCT-analyses that we have applied in the VRD-model and on experimental data is not applicable in the VV-model with R_V as control parameter and this calls for further research to active MCT in which the Vicsek-alignment is explicitly taken into account.

6.4.2 Jamming and alignment phase diagram

To quantitatively describe the change in dynamics upon varying R_V we measure two different parameters for the unjammed and jammed regime. In the unjammed regime, we look at the relaxation time similar as done in the VRD-model. This relaxation time is not defined in the jammed regime as the ISFs never falls back to zero and therefore we use the non-ergodicity parameter (NEP), which we define as the height of the ISF at $t = 10^4\,\tau_0$. The relaxation time and the NEP are shown in respectively the top left and top right panel in Figure 79. We have chosen to evaluate the incoherent ISF because Figure 77 indicates that this function is less influenced by noise than its coherent counterpart. Because both ISFs shows similar behavior in our analysis of the VRD- and VV-model, we do not expect that the coherent ISF would give different results in the VV-model. The relaxation time and the non-ergodicity parameter show the increase in dynamics caused by the Vicsek radius for all evaluated values of the preferred perimeter. For these preferred perimeters the fastest dynamics is always observed around $R_V = 1.0\,r_0$, but the exact value shifts to lower R_V when p_0 decreases, which is in line with the shift in transition point from zero to a finite alignment as shown by the AI in the bottom left panel of Figure 79. Several effects may cause this shift in transition point such as a larger variety in interparticle distances and more motion in the cell layer for higher values of p_0. The effect of the latter can be understood by the fact that the interaction potential highly influences the velocity vector in an unjammed state (high p_0), which counteracts the Vicsek-like alignment in which cells align their polarity to the velocity of their neighbors. In appendix J.3, we show zoomed-out versions of the graphs in Figure 79, which includes values of R_V ranging from $0.0\,r_0$ and $5.0\,r_0$ and we observe that the interesting region regarding dynamics and alignment is indeed located around $R_V = 1.0\,r_0$ as Figure 79 suggests.

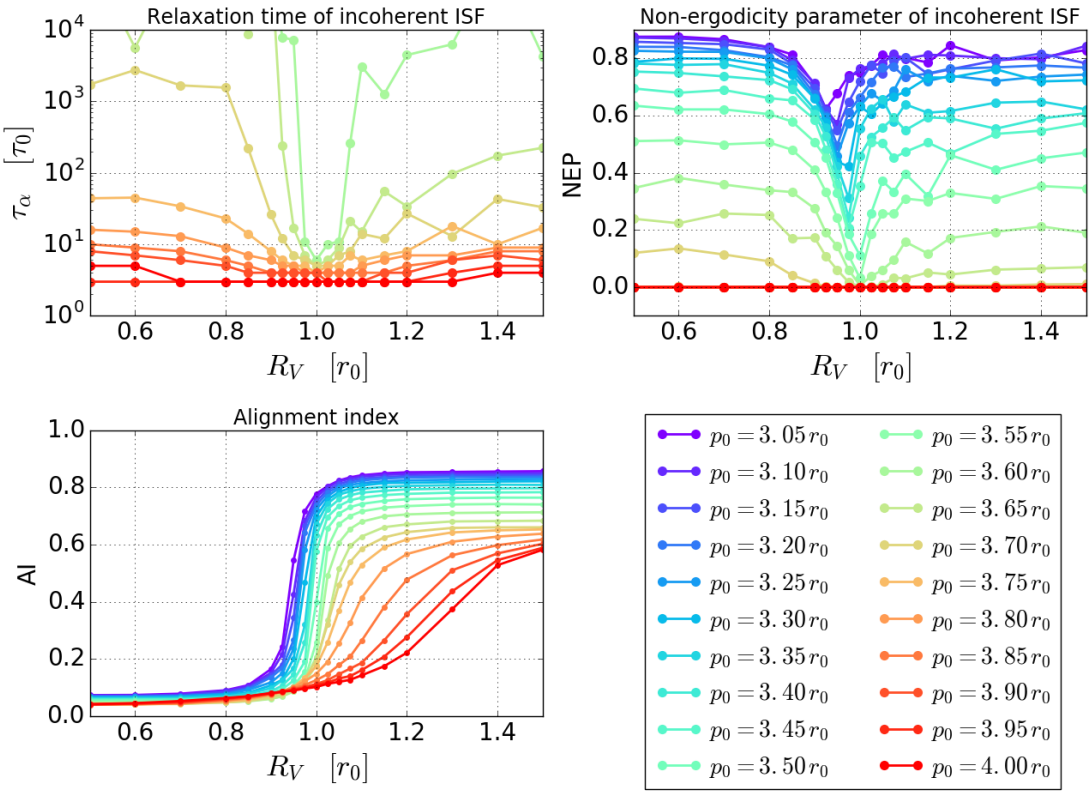

Figure 79: The relaxation time of the incoherent ISF (left top panel), the non-ergodicity parameter of the incoherent ISF (right top panel), the alignment index (left bottom panel) for R_V varying between $0.5\,r_0$ and $1.5\,r_0$ and p_0 varying between $3.0\,r_0$ and $4.0\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).

Based on the jamming and alignment captured in Figure 79 we obtain the phase diagrams in Figure 80, where an interpolation according to the tricontour function of *python 2.7.12* is used. The jamming phase diagram consists again of two separate plots for the unjammed (left top panel) and jammed regime (right top panel), based on respectively the relaxation time and the non-ergodicity parameter of the incoherent ISF. Furthermore, the alignment phase diagram (left bottom panel) is again based on the alignment index. The combination of jamming and alignment results in four different states, which we summarize in the sketch in the right bottom panel of Figure 80. We have chosen the contour line NEP=0.2 to distinguish between an unjammed state and a jammed state, because the NEP is less influenced by the noise in the ISF than the relaxation time as we can average the NEP during a short time period $t' << 10^4\tau_0$ at the end of the measurement period. In addition, we have chosen the value 0.2 as threshold instead of lower values to reduce the influence of the noise as well. To distinguish between an aligned and unaligned state we have chosen the contour line AI=0.2, also for the reason that lower values of AI might be under influence of noise and, besides, that contour lines corresponding to higher values are close to the line of AI=0.2, at least for $p_0 > 1.0\, r_0$.

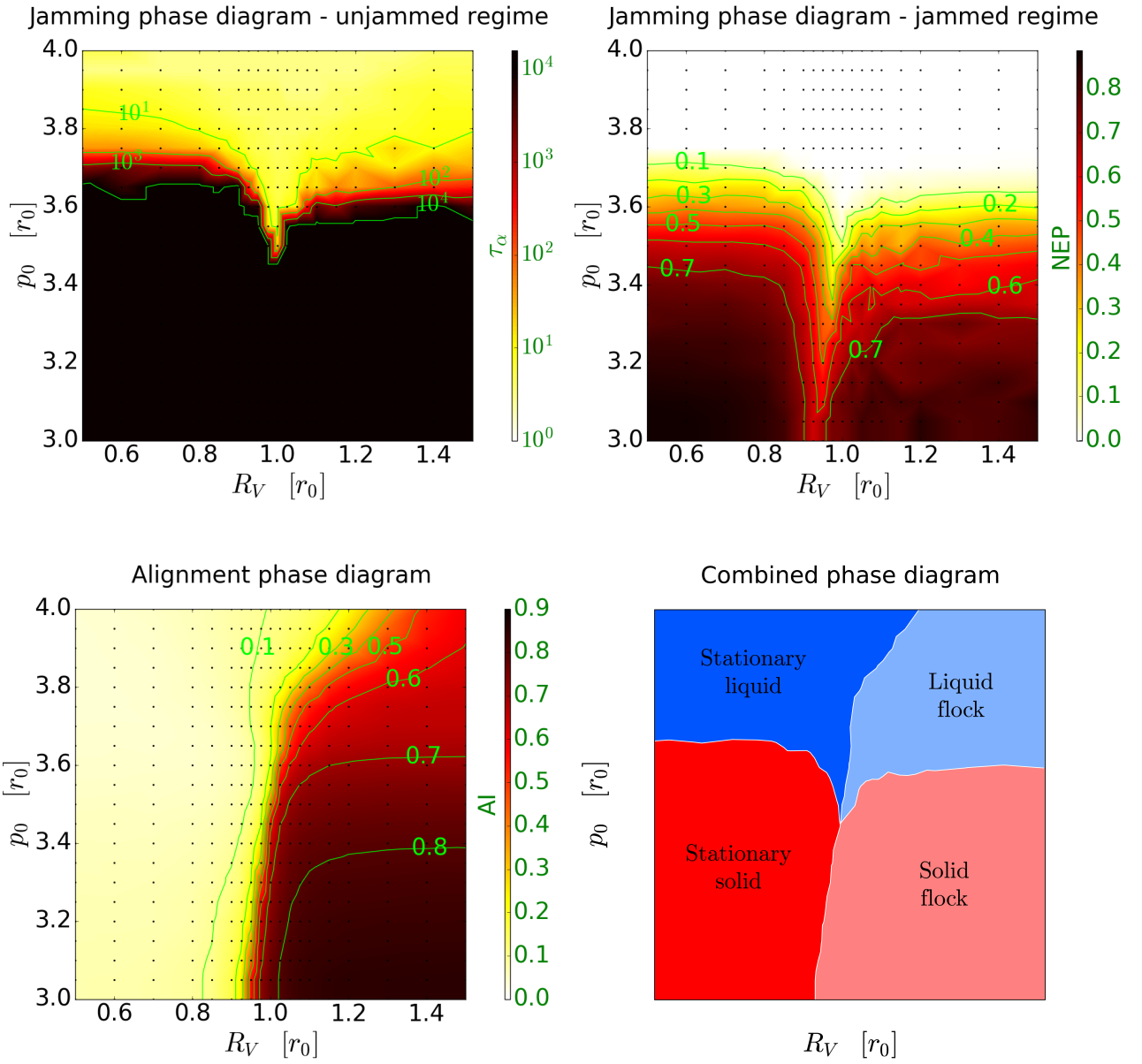

Figure 80: Jamming phase diagram of the unjammed regime according to the relaxation time of the incoherent ISF (top left panel), jamming phase diagram of the jammed regime according to the non-ergodicity parameter of the incoherent ISF (top right panel), alignment phase diagram (bottom left panel) and a combined phase diagram for jamming and alignment based on NEP=0.2 and AI=0.2.

Thus, the phase diagram of the VV-model shows four phases in which the Vicsek radius dominates the flock forming and the preferred perimeter controls the dynamics. Interesting behavior occurs around the meeting point between these four states, where a solid state is suppressed. These four states are identical to the four states observed in the SPV-model with self-alignment (see Giavazzi *et al.* [149] and given in Figure 36), although the dependence of the states on the alignment differs in both models. For example, in the VV-model with R_V as control parameter is not observed that flocking promotes solidification, which is one of the main conclusions of Giavazzi *et al.* [149]. This could be a difference in the definition of the jamming transition as we evaluated the intermediate scattering function and Giavazzi *et al.* [149] applied the effective diffusion coefficient obtained from the mean squared displacement, but can also mark a fundamental difference between self-alignment controlled with an interaction strength and neighbor-alignment controlled with the Vicsek radius. Because the Vicsek persistence time τ_V determines the interaction strength of the neighbor-alignment in the VV-model, an additional comparison in which τ_V is tuned in the VV-model might give further insights in the similarities and differences between the two models with different alignment mechanisms. Examining this effect of the Vicsek persistence time τ_V results in an extra axis in the phase diagram given in the left bottom panel of Figure 80. Also the velocity v_0 or the rotational noise D_r can be added, resulting in different opportunities to make a rich three-dimensional phase diagram for the VV-model, similar to the phase diagram of the VRD-model in Figure 33. This is left to further research.

6.5 Conclusions

We have shown that the alignment in the VV-model can be tuned with the Vicsek radius R_V, in line with the intended behavior of this model. In addition, the other four free parameters also influence the alignment, which results in numerous options to reach similar states. Increasing the Vicsek radius R_V, decreasing the preferred perimeter p_0, tuning the velocity v_0 towards $1.0\, r_0/\tau_0$, decreasing the rotational diffusion D_r or increasing the Vicsek persistence time τ_V leads in general to a larger alignment. Furthermore, a larger alignment is usually accompanied by a faster alignment rate. Analyses of the average perimeter q suggest that in most cases flocking promotes solidification and vice versa.

However, further investigation of the dynamics and jamming in the VV-model using correlation functions suggest that this connection between flocking and solidification is more complex when R_V is used as control parameter. A jammed state without neighbor-alignment might become unjammed when the Vicsek radius is increased to a value around $1.0\, r_0$ and a further increase might re-solidify the layer. In contrast to the VRD-model, this fluidification or increase in dynamics is accompanied by an increase in structure. Consequently, the average perimeter q fails to capture this dynamic behavior, which makes a mode coupling theory analysis even more interesting. However, an applicability of MCT would be highly non-trivial as MCT in general predicts that the dynamics slows down when the structure increases.

The combination of alignment and jamming results in a rich phase diagram consisting of four phases: a stationary fluid, a stationary liquid, a solid flock and a liquid flock. For further research we suggest to evaluate this phase diagram further and extend the diagram to three-dimensions using one of the other free parameters (τ_V, v_0 or D_r) as a third axis. We expect that an increase in v_0 or a decrease in D_r might result in an unjamming transition in analogy with the VRD-model. Furthermore, we expect that a decrease in τ_V might slow down the dynamics in analogy with the SPV-model with self-alignment (Figure 36).

The VV-model suggests a mechanism of cell alignment in experiments with EGF and varying calcium concentration. The addition of EGF is reflected by a nonzero value of the velocity v_0 and an increase or decrease in the calcium concentration by respectively an increase or decrease in R_V. This is physically explained by the fact that cell are less connected under low calcium concentrations in experiments which results in fewer neighbors that influence a cell's alignment in the VV-model. Experiments and model show different behavior on short times and on long times, respectively due to the adaptation time and the depletion of EGF and calcium in experiments. It might therefor be interesting for further research to connect the simulations and experiments for these periods as well.

Before continuing to the final conclusions of our research, we explain how this research can be improved and extended in the next section.

7 Future research possibilities

The main goal of this thesis has been to understand the collective behavior in confluent cell layers better, but the research has also lead to many ideas for improvements and extensions. Some of these possibilities are already mentioned in this thesis and in this section we summarize them together with some extra ideas we have gained. One of the main interesting pathways is to investigate the activity further by investigating different forms of the active force in cell models and by applying active mode coupling theory, which we will both discuss in Subsection 7.1. In Subsection 7.2 we list other research possibilities regarding model extensions, structure and dynamic heterogeneity, MCT, the VV-model and experiments. We conclude in Subsection 7.3 by looking at the bigger picture concerning both mode coupling theory in general and jamming and collective behavior in confluent cell layers.

7.1 Activity and active MCT

It is clear that cells are self-propulsive, but it is unknown how this active force behaves in confluent cell layers. In this thesis we investigated two implementations of this self-propulsion, namely via a force whose polarity undergoes rotational diffusion (VRD-model) and via a force whose polarity aligns with its neighbors (VV-model). However, the list of possibilities are endless and the differences in dynamic behavior between the VRD-model and the VV-model suggest that other implementations of the active force might give rise to other interesting dynamical phenomena. Below we highlight one specific implementation of the active force based on biology and two implementations for which an active version of mode coupling theory is already studied in a self-propelled particle model (SPP). These SPP-models have a different interaction potential than the Voronoi potential in our cell models, but the fact that MCT is independent of the specific form of this potential implies that an active MCT-analysis can be relatively easily implemented in these latter two cell models. This active MCT might improve the results as until now we have only applied standard MCT in which the activity is only taken into account via the static structure factor. We first explain these possible cell models below before explaining the active versions of MCT further.

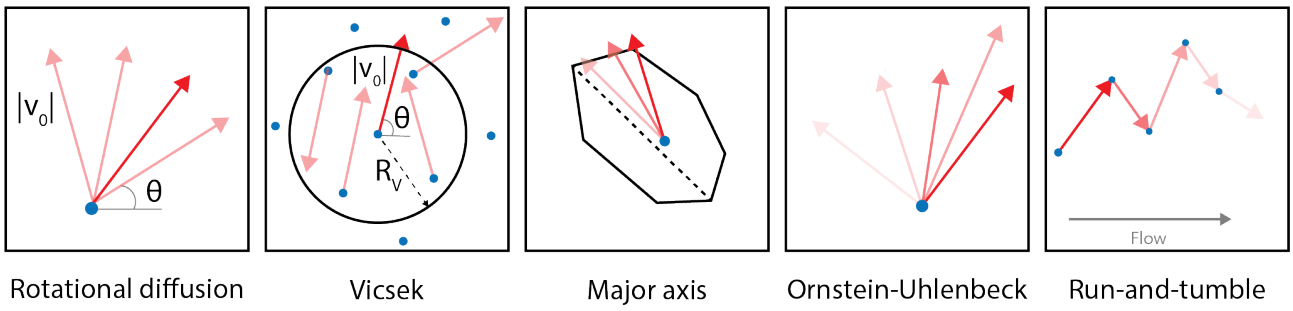

Figure 81: Five implementations of the active force.

The active force in the VRD-model, the VV-model and three new models are sketched in Figure 81 and defined as follows.

- Voronoi rotational diffusion model: the direction of the force changes according to Brownian noise, while the size of the active force is constant. This is our basis model.
- Voronoi Vicsek model: the direction of the force alters via a Vicsek-like alignment and Brownian noise, while the size of the active force is constant. In this Vicsek-like alignment the direction of the force changes towards the average direction of all particles that are within a specific radius. This model is investigated in Section 6.
- Voronoi major axis model: the active force aligns with the major axis of the cell, while the size of the active force stays constant. This model takes into account the biological interpretation that cells generally move along the direction in which the cytoskeleton is stretched. In simple Voronoi models, the behavior of the cytoskeleton is not explicitly included. Arguably the simplest means to account for cytoskeletal orientation is to assume that the cytoskeleton is stretched along the major axis of the cell.
- Voronoi Ornstein-Uhlenbeck model: the active force changes due to an effective temperature in the presence of a persistence time. This type of active force is already investigated in cell models [1, 96], but not yet in combination with the Voronoi interaction potential.

- Voronoi run-and-tumble model: the active force changes according to a white noise term in both the direction and in the magnitude and an extra shear flow is added. Run-and-tumble models are a good way to model single cells and it is interesting to investigate whether it has a counterpart in confluent cell layers.

For the remainder of this subsection we focus on active mode coupling theory and especially on the versions derived for an Ornstein-Uhlenbeck and run-and-tumble model. In general, to derive an active MCT-equation, the Smoluchowski operator $\mathcal{S}$ in Equation 39 changes according to the form of the active force. This implies additional assumptions in the derivation of the mode coupling theory and gives rise to different routes to derive the MCT-equation, which results in the existence of different active MCT-versions for similar models.

Refs. [62] and [63] both proposed a mode coupling theory for Ornstein-Uhlenbeck particles. Szamel *et al.* [62] (see also Szamel [75]) make the assumption that the particle positions evolve on a time scale much larger than the time scale governing the evolution of the activity vector, which means that the persistence length is much smaller than the caging length, and they make a projection onto a local steady state to be able to assume quasi-equilibrium. The resulting MCT-equation needs the Fourier transform of the velocity correlations as an extra input. A completely different route is followed by Feng and Hou [63] and includes an effective diffusion coefficient and a modification of the static structure factor, both dependent on the form of the active force. Despite the differences in approach, it turns out that the effective diffusion coefficient plays the same role as the velocity correlations in Refs. [62, 75]. The MCT-predictions of Szamel *et al.* [62] give a non-monotonous behavior of the relaxation time based on persistence time, which is in agreement with simulations. Feng and Hou [63] conclude from their MCT-predictions that the location of the jamming transition also depends on the persistence time and on the magnitude of the self-propulsion force, which is in agreement with simulations as well.

Farage and Brader [64] derived a mode coupling theory for run-and-tumble particles. They apply a so-called diffusion limit in which they assume that a single, non-interacting particle behaves effectively as a passive particle but with a higher effective diffusion coefficient, which is comparable to the assumption made by Szamel *et al.* [62]. In essence, this approach leads to the removal of explicit rotational freedom and results in an MCT-equation similar to the standard version in Subsection 2.3.1 but with an activity dependent prefactor for the memory kernel and frequency term. Farage and Brader [64] show that this version of MCT predicts a monotonous shift in the glass transition by varying the velocity as seen in simulations.

Perhaps the most complete version of active MCT is derived by Liluashvili *et al.* [65] for active Brownian particles that move with a constant self-propulsion speed in a random direction, subject to translational and rotational Brownian noise. In their derivation they treat both the translational and rotational degrees of freedom on an equal footing instead of making a reduction to a near-equilibrium state. The rotational degrees of freedom are now explicitly coupled to the translational motion and the static velocity correlations are not needed as extra input. The MCT-analysis of Liluashvili *et al.* [65] results in a phase diagram dependent on the velocity of the self-propulsion force, the rotational noise and the packing fraction.

Other studies of active mode coupling theory include the approach of Kranz *et al.* [66] for driven granular fluids and the study of Nandi and Gov [67] for an Ornstein-Uhlenbeck process, but we will not discuss them further.

Despite the many assumptions that are done in the several derivations for active MCT, the different theories all give some accurate predictions about, for example, the relaxation time depending on the characteristics of the active force. This makes it highly relevant for us to look further into these active MCT-versions. Nevertheless, more research is needed to investigate the different approaches and assumptions made in the active MCT derivations described above. This can hopefully lead to an active MCT-version for the VRD-model that can improve the results presented in this paper for the VRD-model and to an adequate active MCT-analysis for the VV-model. In the next subsection we summarize further research extensions and improvements.

7.2 Further research extensions and improvements

In Subsection 7.2.1 we explain possibilities to extend the cell models that we have used and in Subsection 7.2.2 we list options to delve further into structural and dynamical heterogeneity. Possibilities to verify, extend and improve our MCT-analysis are discussed in Subsection 7.2.3. In Subsection 7.2.4 we explain how the Voronoi Vicsek model can be further investigated and in Subsection 7.2.5 we give suggestions for new experiments based on our theoretical analyses.

7.2.1 Model extensions

In this subsection we list several possibilities for extending the models used in this research, but first we make one remark about the current model. In our research, we have assumed that the Voronoi centers represent cell nuclei and we have calculated the correlation functions and applying mode coupling theory using these Voronoi centers, which sounds reasonable as in the VRD-model the equation of motion is solved for these Voronoi centers as degrees of freedom. Nevertheless, each cell also possesses another characteristic point, namely the centroid. The centroid, also called the center of mass or the geometric center, has equal cell volume on all sides and might represent the nucleus of the cell as well. Therefore it might be useful to investigate how the correlation functions change when using the centroid instead of the Voronoi centers. Especially in cases where p_0 is high and the Voronoi centers are sticking together (see right panel of Figure 38), the centroids might give different results including a weakening in the vanishing of the mutual exclusion. Furthermore, there is no strong reason to assume that the relaxation times for the seeds and the centroids are the same, which makes it useful to look at the MCT-predictions for the centroids as well.

We have used a simple Voronoi potential in our VRD- and VV-model, because this is sufficient to study the glassy behavior of confluent cell layers, but adapting this potential can give rise to interesting research. For example, Sussman *et al.* [130] add an extra interfacial tension between different cells in a bidisperse mixture, mimicking the dislike of cells to share boundaries with a different type of cell. Other possible additions are terms that take into account the substrate adhesion or the attraction or repulsion of the nuclei. Using a more complicated potential, similar analysis can be performed as done in this thesis to study the changes in structure and dynamics. Because mode coupling theory does not take into account the form of the potential explicitly, it would be useful to verify our MCT-results using different potentials as this does not change the MCT-equation.

Another modification to the VRD- and VV-model is changing the Voronoi tesselation defined by Equation 55 by adapting the distance function $\mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_i)$ using weights depending on cell i. One way to do this is using the Euclidean squared distance with subtracted weights w_i called the difference method,

$$\mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_i) = (\boldsymbol{r} - \boldsymbol{r}_i)^2 - w_i^2, \tag{80}$$

and a second method is dividing the Euclidean distance by the weights, called the quotient method,

$$\mathcal{D}(\boldsymbol{r}, \boldsymbol{r}_i) = \frac{(\boldsymbol{r} - \boldsymbol{r}_i)^2}{w_i^2}. \tag{81}$$

Figure 82 shows how the two methods changes the Voronoi tesselation, with in the left panel the difference method and in the right panel the quotient method. The green stars define the Voronoi centers and the green circles have a radius equal to the weight. In the quotient method, the edge between cell i and j is at equidistant from the circles of cell i and j. This relation is more difficult in the difference method and results in piecewise spherical edges. These non-linear boundaries might provide a better way to describe real cellular systems than the straight edges that we are using in our VRD- and VV-model. However, we do not expect that these adapted Voronoi tesselation changes the dynamics fundamentally and this adaption would therefore mainly serve to increase the understandability of cell models in describing real cell layers.

In our models we have neglected cell proliferation and death, while this has a significant influence on the behavior of cell systems and specifically on the glassy behavior [12, 142]. Cell divisions can fluidize a layer as it can decrease the energy barriers by adding movements to the system, while it can also cause self-driven jamming in confined organisms because of an increase in density. Cell death fluidizes the system because the living cells obtain more free space to migrate to, which results in lower boundaries for structural rearrangements. Cell division and death can be incorporated in simulation models by adding a probability that a cell divides or dies in each time step. This is not taken into account in the mode coupling theory equations and it might therefore be interesting to investigate its effect on the mode coupling theory predications. Note that when the division rate is not equal to the rate at which cells die, the density changes and this makes it hard to calculate the static structure factor and the intermediate scattering function. This is the same difficulty as we faced in the experimental data in Section 5, where we solved it by normalizing the inter-particle distances with the square root of the density. A similar approach could be applied in a VRD-model where cell divisions and death are included and might reveal interesting dynamic behavior.

Glassy behavior is also observed in three-dimensional cellular systems [141, 158] and it is possible to extend our research to 3D systems. This can be done for the VRD- and VV-model by applying a 3D-version of the interaction potential, given by

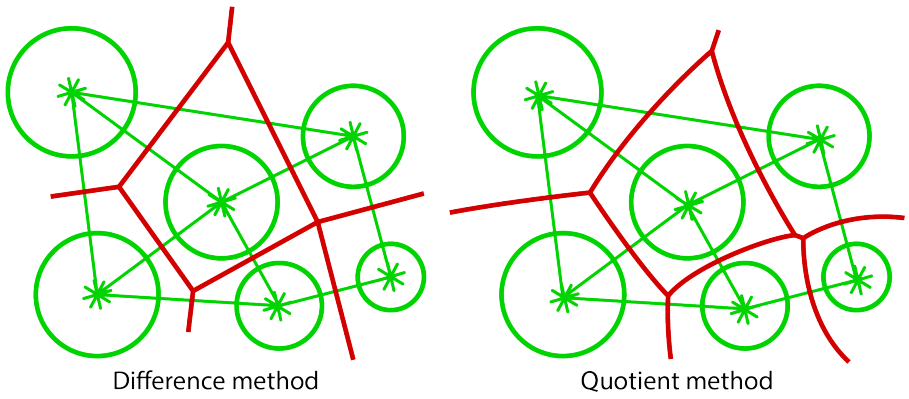

Figure 82: Illustration of the difference method (left panel) and the quotient (right panel) to define the edges in a Voronoi tesselation, where the green stars are Voronoi centers, the green circles have a radius equal to the weight and the red lines define the resulting Voronoi tesselation [137].

$$E_i = K_V(V_i - V_0)^2 + K_A(A_i - A_0)^2, \tag{82}$$

in which a deviation from a cell's volume V_i from its preferred value V_0 has a penalty constant K_V and a deviation from a cell's preferred area has a penalty constant K_A. Although the equation of motions for the molecular dynamics is more complicated in 3D than in 2D, the simulation procedures are similar. Because the correlation functions and MCT can easily be extended to 3D [49, 159, 160], the analyses done in this thesis all have their analogy in 3D.

In Subsection 3.2, the Cellular Potts models, the SPP models and the vertex models are explained and in all three classes a jamming transition and associated glassy behavior are observed. Therefore, the three classes all have the potential for a mode coupling theory analysis. We only applied MCT on one subversion of the vertex model class in this thesis and it might be useful to check our conclusions by applying MCT on other cell models as well. Although the models use different energy potentials and simulation procedures, the fundamentals and the resulting dynamics are similar. Verification of MCT in these different models could enhance the validity of MCT in cellular layers and the concept that cells can be described as simple glasses.

7.2.2 Strutural and dynamic heterogeneity investigations

Dynamic heterogeneity is one of the main characteristics of the glass transition (see Subsection 2.1) and is observed in many experiments of confluent cellular layers that exhibit a jamming transition (see Section 1). The two-point density correlation functions (described in Subsection 2.2) are insufficient to describe spatially heterogeneous dynamics, and therefore we have to go to the next non-trivial, higher order function which is the time dependent four-point density correlation function $G_4(\boldsymbol{r}_1, \boldsymbol{r}_2, t)$ [161, 162] for positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, given by

$$G_4(\boldsymbol{r}_1, \boldsymbol{r}_2, t) = \langle \delta\rho(\boldsymbol{r}_1, 0)\delta\rho(\boldsymbol{r}_1, t)\delta\rho(\boldsymbol{r}_2, 0)\delta\rho(\boldsymbol{r}_2, t)\rangle - \langle \delta\rho(\boldsymbol{r}_1, 0)\delta\rho(\boldsymbol{r}_1, t)\rangle\langle \delta\rho(\boldsymbol{r}_2, 0)\delta\rho(\boldsymbol{r}_2, t)\rangle_t, \tag{83}$$

with $\rho(\boldsymbol{r}, t)$ given by Equation 10, $\Delta\rho(\boldsymbol{r}, t) = \rho(\boldsymbol{r}, t) - \langle\rho\rangle$ and $\langle\rho\rangle = N/V$. Extending our analysis using correlation functions with $G_4(t)$ or its analogy in Fourier space to probe dynamic heterogeneity might give additional insights in the dynamics in confluent cell layers. Furthermore, being able to describe this dynamics with mode coupling theory might be even more valuable. Standard MCT is not suitable for this as it is a mean-field theory. Biroli *et al.* [163] derive an inhomogeneous mode coupling theory equation for a three-point correlation function dependent on two different wave vectors and time, which can be used as an approximation for the four-point correlation function. This inhomogeneous MCT is derived by applying an external field and assuming that the fluctuations caused by this external field are similar to the internal fluctuations of the system. To the best of our knowledge, there do not exist MCT versions that predicts $G_4(\boldsymbol{r}_1, \boldsymbol{r}_2, t)$ or its counterpart in Fourier space and the search for an MCT-version that predicts $G_4(\boldsymbol{r}_1, \boldsymbol{r}_2, t)$ will be a challenging and interesting task.

Besides dynamic heterogeneity, the layer may also contain structural heterogeneity. An example is the spontaneous arise of cell size heterogeneity, which might be one of the reason that MCT predicts too fast dynamics as suggested in Subsection 4.4. Although the input parameters would suggest a monodisperse system, when the cell shapes do not match their preferred shape, a variety of cell shapes and sizes exist in the layer. This influences

the monodisperse static structure factor which, consequently, captures less structure and this might result in the prediction of too fast dynamics according to monodisperse MCT. By binning the cells according to their cell sizes and calculating the corresponding static structure factors, we can apply a multidisperse MCT [60, 61] in which a matrix equation is solved for different intermediate scattering functions corresponding to the different static structure factors for different cell sizes and this might improve the quantitative MCT-predictions.

Structural heterogeneity can also explicitly be added to the cell layer and one of the main reasons to do so is to prevent crystallization, which is unnecessary in our models as the active force already ensures that crystallization does not occur. Nevertheless, another reason to look at multidisperse mixtures is the fact that real cell are never monodisperse. A natural way to incorporate multidispersity might be applying a Gaussian distribution for the cell areas a_i and adapting the preferred perimeters p_i such that the ratio $p_i/\sqrt{a_i}$ is the same for every cell i. Cell size heterogeneity can also be added via a weighted Voronoi tesselation as explained in Subsection 7.3. A simpler way to incorporate multidispersity is using a bidisperse mixture in which there are only two different preferred areas a_i and preferred perimeters p_i, mimicking two different cell types. A more complicated way to add cell size heterogeneity is applying a weighted Voronoi tesselation as explained in Subsection 7.3. The bidisperse mixture is investigated by Sussman *et al.* [113] and they show that the relaxation time and the mean squared displacement are comparable between the monodisperse and bidisperse case. Consequently, mode coupling theory should give similar predictions. However, both cell types have a different characteristic length scale and the radial distribution function and the static structure factor changes when multidispersity is applied. This makes it useful to verify whether MCT still predicts similar dynamics as that would contributes to the validation of MCT.

7.2.3 Additional verifications and extensions of MCT

In Subsections 7.1 till 7.3 we have already suggested that MCT in confluent cell layers can be further verified and improved by using centroids instead of Voronoi seeds, using adapted or different cell models and adding cell size heterogeneity. Furthermore, we suggested to extend our MCT-analysis with active MCT, inhomogeneous MCT and MCT for multidisperse mixtures. In this subsection we will summarize other ideas for verifying and extending mode coupling theory in confluent cell layers.

To start with, we can vary the rotational diffusion constant in the VRD-model as this is the only free parameter we have held constant. Bi *et al.* [108] already showed that an increase in D_r shifts the jamming transition to a higher velocity and a larger preferred perimeter p_0. This implies that the persistence of the self-propelled motion helps cells to escape from the cage of their neighbors and can therefore fluidize the layer. First, it might be interesting to investigate how D_r changes the structural and dynamical correlation functions and secondly, how this influences the mode coupling theory predictions. We expect that these predictions become more accurate when the rotational diffusion increases and less accurate when this decreases, because a low noise level increases the persistence of the active force and consequently its effect on the motion of the cells. Therefore, not taking into account this activity explicitly in the MCT-equations may have a higher impact on the predictions when D_r is relatively small.

We only tested the MCT-predictions for the value of k at the first peak of the static structure factor. At this inverse length scale the intermediate scattering function shows the slowest dynamics and therefore the broadest range of relaxation times for different values of the input parameters p_0 and v_0. Nevertheless, comparing the mode coupling theory predictions for the relaxation times at different k-values with the direct calculations of the corresponding relaxation times could give additional insights in the validity of MCT. Normally, the ISF at k-values corresponding to the peaks in the SSF show longer relaxation times than at k-values corresponding to valleys and we expect MCT to predict these peaks and valleys as well, which would be in line with the predictions of the non-ergodicity parameter in Figure 19.

Furthermore, we have only tested the MCT-predictions for the coherent and incoherent intermediate scattering function and their relaxation times, but MCT can also make predictions about the mean squared displacement [60]. Because the background and assumptions of the MCT-derivation is still the same, the MCT-predictions for the MSD should give similar agreement as MCT-predictions for the ISF. Our conclusions about the applicability of MCT could therefore be substantiated by verifying these predictions for the MSD.

Finally, we can test further predictions of mode coupling theory such as the scaling behavior in the α- and β-regime as given in Figure 15, the scaling of the relaxation time and an equivalent of the time temperature superposition, as explained in Subsection 2.3.3.

We conclude this subsection with a promising extension of mode coupling theory. Standard mode coupling theory makes use of the approximation of a four-point density correlation function into a product of two-point densities (see Equation 33) and it is unknown beforehand whether this is a valid approximation, although the accuracy of

MCT-predictions suggests its legitimation. Nevertheless, it is tempting to try to improve this approximation and one possibility is to apply generalized mode coupling theory (GMCT) [68, 69]. In this extension, a new equation of motion is made for the four-point density correlations on the left hand side of Equation 33 which results in an unknown six-point density correlation. This can be approximated or again solved via an equation of motion resulting in an unknown eight-point density. Because GMCT still uses uncontrolled approximations, it is not guaranteed that it improves the results with respect to standard MCT. However, studies have shown that GMCT indeed improves the MCT-predictions qualitatively and quantitatively [71, 157]. A reason why GMCT might give improvements for the VRD-model is that the pair density correlations in standard-MCT are motivated by a pair-interaction potential, while the VRD-model possesses a multi-particle potential. The inclusion of higher order density correlations might account for this and this can possibly improve the quantitative results of MCT in the unjammed regime of the VRD-model and extend the applicability of MCT to the jammed regime.

7.2.4 Further investigations for the VV-model

Based on our research of the VV-model, the most interesting step to take is to investigate the jamming phase diagram further. This might lead to a better understanding of the non-monotonous behavior of the relaxation time and corresponding structural changes upon varying the Vicsek radius R_V. Furthermore, jamming as a function of the other free parameters such as the velocity and persistence time of the self-propulsion force might be useful to map out. This will probably lead to a rich three-dimensional jamming phase diagram for the VV-model, similar to the jamming diagram in Figure 33 for the VRD-model. In the remainder of this subsection we list several other possibilities to improve and extend the research of the VV-model.

It might be useful to scale up our simulations to larger cell numbers as it is likely that our results are influenced by finite size effects. Only 1000 cells are used in our simulation and Figure 66 shows how easily the maximum correlation length of $L_{\text{corr}} = 22.4\, r_0$ is reached. This maximum correlation length increases when the number of cells in the simulation increases and we expect that the time it takes to reach this length increases as well. As a result, there might be parameter combinations for which the $N = 1000$ case reaches complete correlation while this is not reached using much larger cell numbers and this would result in different collective behavior depending on the system size. In addition, we expect that swirls of comoving cells arise when the correlation length is smaller than the maximum distance between two cells in the system, similar as observed in experiments [146]. We expected that this swirl-forming is completely suppressed in the $N = 1000$ case and increasing the numbers of cells gives the opportunity to investigate whether the VV-model can capture this behavior. One method to test for swirls in a system is by dividing the simulation box in different fields of view (FOV) such that in each FOV the cells are aligned. Differences in the direction of the average velocity vector in each frame of field indicates the existence of swirls.

Another feature of the experiments that we did not capture in our VV-model is the slow increase of alignment at the beginning of the measurement period and the decay of the alignment at the end. We expect that this is respectively caused by the fact that cells in experiments require some time to process the EGF, which delays the alignment, and that the EGF is depleted causing cells to starve again, which cancels the alignment. Our suggestion to mimic these effects in simulations of the VV-model is to implement a time-dependent Vicsek radius and velocity that both start at zero, increase slowly to a certain steady value and after some time decreases back to zero. This might enhance the agreement between the experiments of E. Lang [146] and the VV-model.

On a different note, the phase diagram of the VV-model consists of the same phases as a comparable cell model with a self-alignment mechanism in which the director of a cell aligns with its own velocity vector (see Subsection 3.4.2). In both models flocking is seen and in both models the jamming transition can be shifted by tuning the self-propulsion force. However, we use the Vicsek radius as control parameter in the VV-model and in the model with self-alignment the interaction strength is used. This marks one difference between both models as the self-alignment model does not account for a length scale that can be tuned to change alignment. In addition, we expect that similarities between both models can be found by comparing the effect of the Vicsek persistence time in the VV-model and the alignment interaction strength as both determine a time scale over which alignment takes place. Therefore, we suggest for further research to study the similarities and differences between these models in more detail and, finally, draw conclusions about the alignment mechanisms in real cells.

There are five free parameters in the VV-model that can be tuned, but our results in Subsection 6.2 suggest that not all parameters are independent. The similarities between the effect of D_r and τ_V on the alignment and the fact that their product is dimensionless, gives rise to the hypothesis that we can combine them to one parameter $D_r\tau_V$ such that we can study the effect of this new parameter on alignment and jamming instead of studying the effect of both parameters independently. This can be tested by checking whether simulations for the same value of $D_r\tau_V$ with different combinations of D_r and τ_V hold the same results. Besides the combination of D_r and

τ_V, there might also be other groups of parameters that can be combined to one dimensionless parameter. For example, p_0 and R_V both have a typical length scale and the ratio of the two values is independent. Although the alignment index as a function of p_0 and R_V^{-1} does not show similar behavior, both decrease when the parameter increases. This investigation has been left to future research.

Our simulations are done with periodic boundary conditions, mimicking the large number of cells in experiments. However, in experiments also hard walls or free boundaries are investigated [164] and it is interesting to investigate how our VV-model behaves under these conditions. When hard walls are applied, we expect a circular migration when the system size is large enough and when a free border is applied, we expect the velocity of the cells to align in the direction of this border. Furthermore, we expect the system to show slower dynamics close to a hard boundary and faster dynamics at a free boundary. Consequently, the form of the boundaries might therefore fluidify or solidify a confluent cell layer. However, due to the requirement that the layer has to stay confluent and the finite value of the preferred cell area, we do not expect complete diffusion through the free wall.

7.2.5 Experimental suggestions

Although our work has been completely theoretical, it gives rise to several suggestions and possibilities for experiments, which we will explain here. To start with, it would be very useful to repeat our MCT-analysis of Section 5 using cells that exhibit a broader range of relaxation times and become more jammed. Our current analysis of MDCK-cells already show an increase in structure and slowing down of the dynamics upon aging. However, these effects are small because the MDCK-cells never reach a jammed state and to verify our MCT-conclusion we should look at cell systems that jam further to see a larger variation in relaxation times. Furthermore, from our results we can only say that the relaxation time increase during aging, but we cannot make conclusions about how these relaxation times increases because the results are too noisy in comparison with the range of relaxation times that are found. One of the candidate cell lines that reaches a further jammed state are HBEC cells and consequently they show a broader range of relaxation times according to Atia *et al.* [10].

The comparison of the mode coupling theory predictions with direct calculations of the intermediate scattering function for experimental data suggests that the self-diffusion coefficient decreases upon aging in MDCK-cells (see graphs 60 and 61). This prediction can be verified by executing experiments on single, isolated MDCK-cells to determine its diffusion coefficient. In order to do this, cell division has to be suppressed and this alters the experimental circumstances. This might influence the self-diffusivity of the cell, but hopefully still could give qualitative insights in the behavior of the self-diffusion coefficient upong aging.

Furthermore, the analysis of the VRD-model predicts how changes in experimental conditions might alter the relaxation time of the intermediate scattering function. For example, decreasing the ability of cells to move on its own or decreasing the cell-cell adhesion slows down the dynamics according to simulations of the VRD-model and decreasing the self-diffusivity slows down the dynamics as well according to our MCT-analysis. All these predictions could be tested by designing experimental setups in which the conditions can be adjust such that, for example, the cell-cell adhesion or the motility changes.

Using a completely different setup, namely one according to E. Lång *et al.* [146], we can test predictions obtained from the VV-model. For example, it is an interesting question what the cause of the neutralization of the alignment is after some finite time. One hypothesis is that this is due to the depletion of EGF and this suggests that by constantly adding EGF the alignment can be increased even further than the values reported in Lång *et al.* [146].

7.3 The bigger picture

The physics of glasses is a complex field of science and one of the theories that seek to describe its dynamics is mode coupling theory, which is the only framework entirely based on first principles. However, MCT is not an exact theory because its derivation involves uncontrolled approximations and yet, it predicts the dynamic behavior of supercooled liquids remarkably well using only static information. By investigating different versions of MCT and applying MCT to different fields of science, the aim is to better understand why MCT-predictions are accurate in some cases and in other cases not. This understanding can eventually lead to an even broader applicability and further extensions to the theory. We contribute to this research by applying MCT on confluent cell layers and thus extending MCT to a completely different field of research, namely that of living matter. There are no fundamental reasons to a priori believe that the behavior of cells can be described with standard MCT, especially because it does not incorporate the active force explicitly, but the observed correlation between

structure and dynamics in confluent cell layers already suggested this applicability of MCT. The fact that MCT indeed describes the observed dynamical behavior at least qualitatively correctly, even for the simplest version of MCT, is promising and suggests that MCT might be a suitable framework for a predictive theory of cell dynamics. The research possibilities listed prior in this section can help to develop this framework further and, in addition, can lead to a better understanding of MCT, a broader applicability of MCT and further improvements of the MCT-predictions in general.

Our analysis of MCT on confluent cell layers and the prior analysis of correlation functions contributes to the understanding of glassy behavior in these layers and enhances the analogy between cells and supercooled liquids, and our model extension that includes neighbor-alignment sheds light on the observed alignment during collective migration. Understanding the jamming phenomenon and collective behavior of confluent cell layers is necessary to answer challenging questions such as what is the physics governing jamming, what is the role of jamming in biological processes and is jamming a cause or an effect in these processes. Furthermore, can we influence the jamming or unjamming in biological processes, what is the relationship of unjamming to collective migration and what is the physics behind cell alignment? Solving these questions is essential to understand biological processes such as embryonic development, wound healing, cancer metastasis and asthma better and this understanding might lead to the development of new and better medical diagnosis and interventions in the human body.

8 Final conclusions

For our final conclusions we return to the research questions that motivated our research and answer them based on the knowledge we have gained.

1. What do static and dynamic correlation functions obtained from glassy physics teach us about the jamming phenomenon in epithelial cell layers in experiments and simulations?

Using correlation functions we were able to analyze experimental and simulated data through a lens that is not yet widely applied in biological physics. This gave us additional understanding about the fluid-like to solid-like transition in confluent cell layers and enables us to investigate this transition in the same light as the glass transition in molecular liquids or the jamming transition in particulate systems. Specifically, this jamming transition in confluent cell layers is accompanied by an increase in structure probed by the radial distribution function and the static structure factor and by a slowing down of the dynamics as seen in the intermediate scattering function. When approaching the jamming transition a plateau arises in the intermediate scattering function, which signifies the caging of a cell by its neighbors. This is a characteristic of jamming and is reflected in the relaxation time of the intermediate scattering function, which can therefore be used as a parameter to define when or to what extent a system is jammed. In the VRD-model, we have used the preferred perimeter and the velocity of the self-propulsion force as control parameters to reach the jamming transition. We observed that a jammed state is reached below a threshold value of the preferred perimeter and that this critical value decreases with increasing velocity. This may be understood from the fact that small values of the preferred perimeter drive cells to a hexagonal packing in which the energy barriers for cellular rearrangements are high, while increasing the velocity makes it easier for cells to overcome these barriers. In addition to the behavior at the jamming transition, the correlation functions teach us about interesting behavior for values of the preferred perimeter further away from this transition, like the decrease in structural order for very small values or the tendency to obtain a differently ordered state with elongated cells for very large values. Hence, correlation functions are a broadly applicable analysis method and will also be worthwhile for investigating and comparing extensions to the VRD-model, like the VV-model. In experimental data of confluent cell layers, correlation functions show a growth in structure and a slowing down of the dynamics upon aging similar as seen in the VRD-model upon reaching the jamming transition, although over a much smaller range of dynamics and further away from this jamming transition. Many complex biological processes govern this aging behavior, and correlation functions offer a generic tool to analyze structural and dynamical changes without complete knowledge of these contributing processes.

2. By applying mode coupling theory on epithelial cells, can we broaden the applicability of MCT, extend the analogy between glasses and cells, and, ultimately, define a new analysis framework for epithelial cells?

Mode coupling theory predicts the non-monotonous behavior of the relaxation time of the incoherent intermediate scattering correctly up to a constant in our simulations of the VRD-model. Furthermore, MCT captures the slowing down of dynamics upon aging correctly in experiments. These results verify that, at least in the unjammed regime, MCT can be applied to confluent cell layers and we have thus found a new field of science for MCT-analyses. Moreover, the applicability of MCT solidifies the analogy between cells and glass-forming liquids and shows that MCT has the potential to be a suitable theory that describes the dynamics of confluent cell layers. These cell layers exhibit complex collective motion and a theory can help to better understand this dynamics and gives the opportunity to make predictions. Because in MCT the dynamics is fully explained by its structure, we learned that dynamical changes in cell layers are governed by structural changes. Furthermore, MCT predict certain scaling laws for the ISF and its relaxation time, and the applicability of MCT gives the opportunity to test these scaling laws in confluent cell layers. Besides its theoretical relevance, MCT provides us with a new analysis method in which we can get insights into the dynamics of the system based on only one snapshot. This is advantageous as measuring over long times is very expensive in simulations and experiments and might even be impossible in experiments due to continuous changes in the state of cells in the layer. In spirit, our results confirm previous findings that a specific structural measure, namely the ratio of cell perimeter to area, can capture the jamming and unjamming dynamics. However, our results and those of other published work also suggest that this measure only applies close to the transition. MCT is therefore more broadly applicable but appears to require some adaptations to make more accurate predictions. We did not find agreement of MCT with direct calculations of the dynamics in the jammed regime and also failed to predict the correct time scaling, but extensions of MCT might give improvements and even remove these shortcomings. These possible extensions include generalized MCT and an active MCT version that incorporates the self-propulsion force explicitly. This defines an interesting pathway for further research. Another suggestion

for further research is to apply MCT to experimental data that reach a state that is more jammed in comparison with the data we used in order to verify the MCT-predictions experimentally on a broader range of dynamics.

3. How does a model for epithelial cells behave in terms of alignment and jamming when a principle regulating collective motion is explicitly included?

By combining the Voronoi potential with a Vicsek-like alignment, the VV-model explicitly regulates collective motion via neighbor-alignment. The Vicsek radius, the preferred perimeter, the velocity and the persistence time of the self-propulsion force all influence the alignment in the VV-model, resulting in several regulatory opportunities to tune the collective behavior of cells. Specifically, the Vicsek radius and the self-propulsive velocity may be interpreted directly in experimental terms; we find that experiment and theory map well onto each other when we interpret an increase in the EGF fueling the motion as an increase in the velocity in the VV-model, and an increase in calcium (which regulates the cell-cell contacts and thereby increases interconnectivity in the layer) as an increase in the Vicsek radius. Furthermore, we found that the switching on of the Vicsek-alignment via the Vicsek radius can fluidize a jammed layer and can solidify the layer again upon further increasing the Vicsek radius. Fluidization happens for intermediate values of the Vicsek radius because of a higher persistent motion of the cells, while for a high value of the Vicsek radius all cells move in the same direction as a solid flock. The combination of jamming and alignment leads to a rich phase diagram in which four phases can be distinguished, namely a stationary solid, a stationary liquid, a solid flock and a liquid flock. For further research, we suggest to investigate this phase diagram further and extend it with other control parameters such as the velocity or persistence time of the self-propulsion force. Furthermore, this interesting jamming behavior motivates an MCT-analysis for which an active version has to be derived that matches the active force of the VV-model.

Concluding, via correlation functions and mode coupling theory obtained from glassy physics and an extension of an existing cell model we gained a better understanding of the collective behavior of confluent cell layers regarding structure, jamming and alignment and contribute to a field of research that might eventually lead to new and better medical diagnosis and interventions in the human body.

References

[1] V. Hakim and P. Silberzan, "Collective cell migration: A physics perspective," *Reports on Progress in Physics*, vol. 80, no. 7, p. 076 601, 2017.

[2] G. Danuser, J. Allard, and A. Mogilner, "Mathematical modeling of eukaryotic cell migration: Insights beyond experiments," *Annual review of cell and developmental biology*, vol. 29, pp. 501–528, 2013.

[3] J. Travis, "Mysteries of the cell," *Science*, vol. 334, no. 6059, pp. 1046–1046, 2011, ISSN: 0036-8075. DOI: 10.1126/science.334.6059.1046-a.

[4] T. E. Angelini, E. Hannezo, X. Trepat, M. Marquez, J. J. Fredberg, and D. A. Weitz, "Glass-like dynamics of collective cell migration," *Proceedings of the National Academy of Sciences*, vol. 108, no. 12, pp. 4714–4719, 2011.

[5] D. K. Vig, A. E. Hamby, and C. W. Wolgemuth, "Cellular contraction can drive rapid epithelial flows," *Biophysical Journal*, vol. 113, no. 7, pp. 1613–1622, 2017.

[6] E.-M. Schoetz, M. Lanio, J. A. Talbot, and M. L. Manning, "Glassy dynamics in three-dimensional embryonic tissues," *Journal of The Royal Society Interface*, vol. 10, no. 89, p. 20 130 726, 2013.

[7] K. D. Nnetu, M. Knorr, S. Pawlizak, T. Fuhs, and J. A. Käs, "Slow and anomalous dynamics of an mcf-10a epithelial cell monolayer," *Soft Matter*, vol. 9, no. 39, pp. 9335–9341, 2013.

[8] S. Pawlizak, A. W. Fritsch, S. Grosser, D. Ahrens, T. Thalheim, S. Riedel, T. R. Kießling, L. Oswald, M. Zink, and M. L. Manning, "Testing the differential adhesion hypothesis across the epithelial- mesenchymal transition," *New Journal of Physics*, vol. 17, no. 8, p. 083 049, 2015.

[9] S. Garcia, E. Hannezo, J. Elgeti, J.-F. Joanny, P. Silberzan, and N. S. Gov, "Physics of active jamming during collective cellular motion in a monolayer," *Proceedings of the National Academy of Sciences*, vol. 112, no. 50, pp. 15 314–15 319, 2015.

[10] J.-A. Park, J. H. Kim, D. Bi, J. A. Mitchel, N. T. Qazvini, K. Tantisira, C. Y. Park, M. McGill, S.-H. Kim, J. Gweon Bomi Notbohm, R. Steward Jr, S. Burger, S. H. Randell, A. T. Kho, D. T. Tambe, C. Hardin, S. A. Shore, E. Israel, D. A. Weitz, D. J. Tschumperlin, E. P. Henske, S. T. Weiss, M. L. Manning, J. P. Butler, J. M. Drazen, and J. J. Fredberg, "Unjamming and cell shape in the asthmatic airway epithelium," *Nature materials*, vol. 14, no. 10, p. 1040, 2015.

[11] C. Hardin, K. Rajendran, G. Manomohan, D. T. Tambe, J. P. Butler, J. J. Fredberg, R. Martinelli, C. V. Carman, and R. Krishnan, "Glassy dynamics, cell mechanics, and endothelial permeability," *The Journal of Physical Chemistry B*, vol. 117, no. 42, pp. 12 850–12 856, 2013.

[12] M. Delarue, J. Hartung, C. Schreck, P. Gniewek, L. Hu, S. Herminghaus, and O. Hallatschek, "Self-driven jamming in growing microbial populations," *Nature physics*, vol. 12, no. 8, p. 762, 2016.

[13] M. Sadati, A. Nourhani, J. J. Fredberg, and N. Taheri Qazvini, "Glass-like dynamics in the cell and in cellular collectives," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, vol. 6, no. 2, pp. 137–149, 2014.

[14] J.-A. Park, L. Atia, J. A. Mitchel, J. J. Fredberg, and J. P. Butler, "Collective migration and cell jamming in asthma, cancer and development," *Journal of Cell Science*, vol. 129, no. 18, pp. 3375–3383, 2016.

[15] D. T. Tambe, C. C. Hardin, T. E. Angelini, K. Rajendran, C. Y. Park, X. Serra-Picamal, E. H. Zhou, M. H. Zaman, J. P. Butler, D. A. Weitz, J. J. Fredberg, and X. Trepat, "Collective cell guidance by cooperative intercellular forces," *Nature materials*, vol. 10, no. 6, p. 469, 2011.

[16] L. Oswald, S. Grosser, D. M. Smith, and J. A. Käs, "Jamming transitions in cancer," *Journal of Physics D: Applied Physics*, vol. 50, no. 48, p. 483 001, 2017.

[17] X. Trepat and J. J. Fredberg, "Plithotaxis and emergent dynamics in collective cellular migration," *Trends in cell biology*, vol. 21, no. 11, pp. 638–646, 2011.

[18] D. Bi, J. Zhang, B. Chakraborty, and R. P. Behringer, "Jamming by shear," *Nature*, vol. 480, no. 7377, p. 355, 2011.

[19] J. P. Garrahan, "Dynamic heterogeneity comes to life," *Proceedings of the National Academy of Sciences*, vol. 108, no. 12, pp. 4701–4702, 2011.

[20] D. J. Montell, "Border-cell migration: The race is on," *Nature reviews Molecular cell biology*, vol. 4, no. 1, p. 13, 2003.

[21] D. Dormann and C. J. Weijer, "Chemotactic cell movement during development," *Current opinion in genetics & development*, vol. 13, no. 4, pp. 358–364, 2003.

[22] G. Danuser, J. Allard, and A. Mogilner, "Mathematical modeling of eukaryotic cell migration: Insights beyond experiments," *Annual review of cell and developmental biology*, vol. 29, pp. 501–528, 2013.

[23] M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin, and P. Silberzan, "Collective migration of an epithelial monolayer in response to a model wound," *Proceedings of the National Academy of Sciences*, vol. 104, no. 41, pp. 15 988–15 993, 2007.

[24] R. A. Clark and D. Margolis, "The molecular and cellular biology of wound repair," *Archives of Dermatology*, vol. 132, no. 12, p. 1531, 1996.

[25] A. Brugués, E. Anon, V. Conte, J. H. Veldhuis, M. Gupta, J. Colombelli, J. J. Muñoz, G. W. Brodland, B. Ladoux, and X. Trepat, "Forces driving epithelial wound healing," *Nature physics*, vol. 10, no. 9, p. 683, 2014.

[26] M. Basan, J. Elgeti, E. Hannezo, W.-J. Rappel, and H. Levine, "Alignment of cellular motility forces with tissue flow as a mechanism for efficient wound healing," *Proceedings of the National Academy of Sciences*, vol. 110, no. 7, pp. 2452–2459, 2013.

[27] A. Wells, *Cell motility in cancer invasion and metastasis*. Springer Science & Business Media, 2006, vol. 8.

[28] C. H. Stuelten, C. A. Parent, and D. J. Montell, "Cell motility in cancer invasion and metastasis: Insights from simple model organisms," *Nature Reviews Cancer*, 2018.

[29] K. D. Nnetu, M. Knorr, J. Käs, and M. Zink, "The impact of jamming on boundaries of collectively moving weak-interacting cells," *New Journal of Physics*, vol. 14, no. 11, p. 115 012, 2012.

[30] A. W. Lambert, D. R. Pattabiraman, and R. A. Weinberg, "Emerging biological principles of metastasis," *Cell*, vol. 168, no. 4, pp. 670–691, 2017.

[31] J. Notbohm, S. Banerjee, K. J. Utuje, B. Gweon, H. Jang, Y. Park, J. Shin, J. P. Butler, J. J. Fredberg, and M. C. Marchetti, "Cellular contraction and polarization drive collective cellular motion," *Biophysical journal*, vol. 110, no. 12, pp. 2729–2738, 2016.

[32] M. Merkel and M. L. Manning, "Using cell deformation and motion to predict forces and collective behavior in morphogenesis," in *Seminars in cell & developmental biology*, Elsevier, vol. 67, 2017, pp. 161–169.

[33] J. J. Fredberg, "Power steering, power brakes, and jamming: Evolution of collective cell-cell interactions," *Physiology*, vol. 29, no. 4, pp. 218–219, 2014, PMID: 24985324. DOI: 10.1152/physiol.00026.2014.

[34] M. Sadati, N. T. Qazvini, R. Krishnan, C. Y. Park, and J. J. Fredberg, "Collective migration and cell jamming," *Differentiation*, vol. 86, no. 3, pp. 121–125, 2013.

[35] A. J. Liu and S. R. Nagel, "The jamming transition and the marginally jammed solid," *Annual Review of Condensed Matter Physics*, vol. 1, no. 1, pp. 347–369, 2010.

[36] V. Trappe, V. Prasad, L. Cipelletti, P. Segre, and D. A. Weitz, "Jamming phase diagram for attractive particles," *Nature*, vol. 411, no. 6839, p. 772, 2001.

[37] V. Vitelli and M. Van Hecke, "Soft materials: Marginal matters," *Nature*, vol. 480, no. 7377, p. 325, 2011.

[38] X. Trepat, M. R. Wasserman, T. E. Angelini, E. Millet, D. A. Weitz, J. P. Butler, and J. J. Fredberg, "Physical forces during collective cell migration," *Nature Physics*, vol. 5, no. 6, p. 426, 2009.

[39] E. H. Zhou, F. D. Martinez, and J. J. Fredberg, "Cell rheology: Mush rather than machine," *Nature materials*, vol. 12, no. 3, p. 184, 2013.

[40] L. Atia, D. Bi, Y. Sharma, J. A. Mitchel, B. Gweon, S. A. Koehler, S. J. DeCamp, B. Lan, J. H. Kim, R. Hirsch, A. F. Pegoraro, K. H. Lee, J. R. Starr, D. A. Weitz, A. C. Martin, J.-A. Park, J. P. Butler, and J. J. Fredberg, "Geometric constraints during epithelial jamming," *Nature Physics*, p. 1, 2018.

[41] W. Kob, "Supercooled liquids, the glass transition, and computer simulations," *ArXiv preprint cond-mat/0212344*, 2002.

[42] L. Berthier and G. Biroli, "Theoretical perspective on the glass transition and amorphous materials," *Reviews of Modern Physics*, vol. 83, no. 2, p. 587, 2011.

[43] P. G. Debenedetti and F. H. Stillinger, "Supercooled liquids and the glass transition," *Nature*, vol. 410, no. 6825, p. 259, 2001.

[44] C. A. Angell, "Formation of glasses from liquids and biopolymers," *Science*, vol. 267, no. 5206, pp. 1924–1935, 1995.

[45] D. Kennedy and C. Norman, "What don't we know?" *Science*, vol. 309, no. 5731, pp. 75–75, 2005.

[46] J. P. Boon and S. Yip, *Molecular hydrodynamics*. Courier Corporation, 1980.

[47] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids*. Elsevier, 1990.

[48] U. Balucani and M. Zoppi, *Dynamics of the liquid state*. Clarendon Press, 1995, vol. 10.

[49] D. R. Reichman and P. Charbonneau, "Mode-coupling theory," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 05, P05013, 2005.

[50] W. Götze and L. Sjögren, "Comments on the mode coupling theory for structural relaxation," *Chemical physics*, vol. 212, no. 1, pp. 47–59, 1996.

[51] E. Leutheusser, "Dynamical model of the liquid-glass transition," *Physical Review A*, vol. 29, no. 5, p. 2765, 1984.

[52] U. Bengtzelius, W. Gotze, and A. Sjolander, "Dynamics of supercooled liquids and the glass transition," *Journal of Physics C: Solid state Physics*, vol. 17, no. 33, p. 5915, 1984.

[53] P. Mayer, K. Miyazaki, and D. R. Reichman, "Cooperativity beyond caging: Generalized mode-coupling theory," *Physical review letters*, vol. 97, no. 9, p. 095 702, 2006.

[54] H. Mori, "Transport, collective motion, and brownian motion," *Progress of theoretical physics*, vol. 33, no. 3, pp. 423–455, 1965.

[55] J. Barrat, W. Gotze, and A. Latz, "The liquid-glass transition of the hard-sphere system," *Journal of Physics: Condensed matter*, vol. 1, no. 39, p. 7163, 1989.

[56] G. Szamel and H. Löwen, "Mode-coupling theory of the glass transition in colloidal systems," *Physical Review A*, vol. 44, no. 12, p. 8215, 1991.

[57] W. Götze, "Recent tests of the mode-coupling theory for glassy dynamics," *Journal of Physics: Condensed matter*, vol. 11, no. 10A, A1, 1999.

[58] L. Berthier, G. Biroli, J.-P. Bouchaud, W. Kob, K. Miyazaki, and D. Reichman, "Spontaneous and induced dynamic fluctuations in glass formers. i. general results and dependence on ensemble and dynamics," *The Journal of chemical physics*, vol. 126, no. 18, p. 184 503, 2007.

[59] M. K. Nandi, A. Banerjee, S. Sengupta, S. Sastry, and S. M. Bhattacharyya, "Unraveling the success and failure of mode coupling theory from consideration of entropy," *The Journal of Chemical Physics*, vol. 143, no. 17, p. 174 504, 2015.

[60] F. Weysser, A. M. Puertas, M. Fuchs, and T. Voigtmann, "Structural relaxation of polydisperse hard spheres: Comparison of the mode-coupling theory to a langevin dynamics simulation," *Physical review E*, vol. 82, no. 1, p. 011 504, 2010.

[61] T. Franosch, M. Fuchs, W. Götze, M. R. Mayr, and A. Singh, "Asymptotic laws and preasymptotic correction formulas for the relaxation near glass-transition singularities," *Physical Review E*, vol. 55, no. 6, p. 7153, 1997.

[62] G. Szamel, E. Flenner, and L. Berthier, "Glassy dynamics of athermal self-propelled particles: Computer simulations and a nonequilibrium microscopic theory," *Physical Review E*, vol. 91, no. 6, p. 062 304, 2015.

[63] M. Feng and Z. Hou, "Mode coupling theory for nonequilibrium glassy dynamics of thermal self-propelled particles," *Soft matter*, vol. 13, no. 25, pp. 4464–4481, 2017.

[64] T. Farage and J. Brader, "Dynamics and rheology of active glasses," *ArXiv preprint arXiv:1403.0928*, 2014.

[65] A. Liluashvili, J. Ónody, and T. Voigtmann, "Mode-coupling theory for active brownian particles," *Physical Review E*, vol. 96, no. 6, p. 062 608, 2017.

[66] W. T. Kranz, M. Sperl, and A. Zippelius, "Glass transition for driven granular fluids," *Physical review letters*, vol. 104, no. 22, p. 225 701, 2010.

[67] S. K. Nandi and N. S. Gov, "Nonequilibrium mode-coupling theory for dense active systems of self-propelled particles," *Soft matter*, vol. 13, no. 41, pp. 7609–7616, 2017.

[68] G. Szamel, "Colloidal glass transition: Beyond mode-coupling theory," *Physical review letters*, vol. 90, no. 22, p. 228 301, 2003.

[69] W. Van Megen and S. Underwood, "Glass transition in colloidal hard spheres: Measurement and mode-coupling-theory analysis of the coherent intermediate scattering function," *Physical Review E*, vol. 49, no. 5, p. 4206, 1994.

[70] H. Cummins, G. Li, Y. Hwang, G. Shen, W. Du, J. Hernandez, and N. Tao, "Dynamics of supercooled liquids and glasses: Comparison of experiments with theoretical predictions," *Zeitschrift für Physik B Condensed Matter*, vol. 103, no. 3, pp. 501–519, 1997.

[71] L. M. Janssen and D. R. Reichman, "Microscopic dynamics of supercooled liquids from first principles," *Physical review letters*, vol. 115, no. 20, p. 205 701, 2015.

[72] W. Kob, "The mode-coupling theory of the glass transition," *ArXiv preprint cond-mat/9702073*, 1997.

[73] L. M. Janssen, P. Mayer, and D. R. Reichman, "Relaxation patterns in supercooled liquids from generalized mode-coupling theory," *Physical Review E*, vol. 90, no. 5, p. 052 306, 2014.

[74] H. Ding, M. Feng, H. Jiang, and Z. Hou, "Nonequilibrium glass transition in mixtures of active-passive particles," *ArXiv preprint arXiv:1506.02754*, 2015.

[75] G. Szamel, "Theory for the dynamics of dense systems of athermal self-propelled particles," *Physical Review E*, vol. 93, no. 1, p. 012 603, 2016.

[76] S. Karmakar, C. Dasgupta, and S. Sastry, "Growing length scales and their relation to timescales in glass-forming liquids," *Annual Review of Condensed Matter Physics*, vol. 5, no. 1, pp. 255–284, 2014.

[77] L. Janssen, "Mode-coupling theory of the glass transition: A primer," *ArXiv preprint arXiv:1806.01369*, 2018.

[78] W. Goetze and L. Sjoegren, "Beta-relaxation at the glass transition of hard-spherical colloids," *Physical Review A*, vol. 43, no. 10, p. 5442, 1991.

[79] K. N. Pham, A. M. Puertas, J. Bergenholtz, S. U. Egelhaaf, A. Moussaıd, P. N. Pusey, A. B. Schofield, M. E. Cates, M. Fuchs, and W. C. Poon, "Multiple glassy states in a simple model system," *Science*, vol. 296, no. 5565, pp. 104–106, 2002.

[80] L. Berthier, A. J. Moreno, and G. Szamel, "Increasing the density melts ultrasoft colloidal glasses," *Physical Review E*, vol. 82, no. 6, p. 060 501, 2010.

[81] J.-P. Bouchaud, L. Cugliandolo, J. Kurchan, and M. Mézard, "Mode-coupling approximations, glass theory and disordered systems," *ArXiv preprint cond-mat/9511042*, 1995.

[82] J.-P. Bouchaud, L. F. Cugliandolo, J. Kurchan, and M. Mezard, "Out of equilibrium dynamics in spin-glasses and other glassy systems," *Spin glasses and random fields*, pp. 161–223, 1998.

[83] B. Cichocki and W. Hess, "On the memory function for the dynamic structure factor of interacting brownian particles," *Physica A: Statistical Mechanics and its Applications*, vol. 141, no. 2-3, pp. 475–488, 1987.

[84] M. Fuchs, W. Gotze, I. Hofacker, and A. Latz, "Comments on the alpha -peak shapes for relaxation in supercooled liquid," *Journal of physics: Condensed matter*, vol. 3, no. 26, p. 5047, 1991.

[85] D. Chen, W. Y. Aw, D. Devenport, and S. Torquato, "Structural characterization and statistical-mechanical model of epidermal patterns," *Biophysical Journal*, vol. 111, no. 11, pp. 2534–2545, 2016.

[86] S. Kaliman, C. Jayachandran, F. Rehfeldt, and A.-S. Smith, "Limits of applicability of the voronoi tessellation determined by centers of cell nuclei to epithelium morphology," *Frontiers in physiology*, vol. 7, p. 551, 2016.

[87] E. Flenner, G. Szamel, and L. Berthier, "The nonequilibrium glassy dynamics of self-propelled particles," *Soft Matter*, vol. 12, no. 34, pp. 7136–7149, 2016.

[88] G. Szamel, E. Flenner, and L. Berthier, "Glassy dynamics of athermal self-propelled particles: Computer simulations and a nonequilibrium microscopic theory," *Physical Review E*, vol. 91, no. 6, p. 062 304, 2015.

[89] E. Feenberg, *Theory of quantum fluids*. Elsevier, 2012, vol. 31.

[90] L. Berthier and J. Kurchan, "Non-equilibrium glass transitions in driven and active matter," *Nature Physics*, vol. 9, no. 5, p. 310, 2013.

[91] L. Berthier, "Nonequilibrium glassy dynamics of self-propelled hard disks," *Physical review letters*, vol. 112, no. 22, p. 220 602, 2014.

[92] L. Berthier, E. Flenner, and G. Szamel, "How active forces influence nonequilibrium glass transitions," *New Journal of Physics*, vol. 19, no. 12, p. 125 006, 2017.

[93] R. Ni, M. A. C. Stuart, and M. Dijkstra, "Pushing the glass transition towards random close packing using self-propelled hard spheres," *Nature communications*, vol. 4, p. 2704, 2013.

[94] R. Mandal, P. J. Bhuyan, M. Rao, and C. Dasgupta, "Active fluidization in dense glassy systems," *Soft Matter*, vol. 12, no. 29, pp. 6268–6276, 2016.

[95] L. Berthier, E. Flenner, and G. Szamel, "How active forces influence nonequilibrium glass transitions," *New Journal of Physics*, vol. 19, no. 12, p. 125 006, 2017.

[96] B. A. Camley and W.-J. Rappel, "Physical models of collective cell motility: From cell to tissue," *Journal of physics D: Applied physics*, vol. 50, no. 11, p. 113 002, 2017.

[97] B. Ladoux and R.-M. Mège, "Mechanobiology of collective cell behaviours," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 12, p. 743, 2017.

[98] R. Mayor and S. Etienne-Manneville, "The front and rear of collective cell migration," *Nature reviews Molecular cell biology*, vol. 17, no. 2, p. 97, 2016.

[99] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential cell biology*. Garland Science, 2013.

[100] A. G. Fletcher, M. Osterfield, R. E. Baker, and S. Y. Shvartsman, "Vertex models of epithelial morphogenesis," *Biophysical journal*, vol. 106, no. 11, pp. 2291–2304, 2014.

[101] A.-K. Classen, K. I. Anderson, E. Marois, and S. Eaton, "Hexagonal packing of drosophila wing epithelial cells by the planar cell polarity pathway," *Developmental cell*, vol. 9, no. 6, pp. 805–817, 2005.

[102] D. W. Thompson, "On growth and form.," *On growth and form.*, 1942.

[103] T. C. Hales, "The honeycomb conjecture," *Discrete & Computational Geometry*, vol. 25, no. 1, pp. 1–22, 2001.

[104] M. Basan, J. Elgeti, E. Hannezo, W.-J. Rappel, and H. Levine, "Alignment of cellular motility forces with tissue flow as a mechanism for efficient wound healing," *Proceedings of the National Academy of Sciences*, vol. 110, no. 7, pp. 2452–2459, 2013.

[105] M. C. Gibson, A. B. Patel, R. Nagpal, and N. Perrimon, "The emergence of geometric order in proliferating metazoan epithelia," *Nature*, vol. 442, no. 7106, p. 1038, 2006.

[106] B. Szabo, G. Szöllösi, B. Gönci, Z. Jurányi, D. Selmeczi, and T. Vicsek, "Phase transition in the collective migration of tissue cells: Experiment and model," *Physical Review E*, vol. 74, no. 6, p. 061 908, 2006.

[107] A. Szabó, R. Ünnep, E. Méhes, W. Twal, W. Argraves, Y. Cao, and A. Czirók, "Collective cell motion in endothelial monolayers," *Physical biology*, vol. 7, no. 4, p. 046 007, 2010.

[108] D. Bi, X. Yang, M. C. Marchetti, and M. L. Manning, "Motility-driven glass and jamming transitions in biological tissues," *Physical Review X*, vol. 6, no. 2, p. 021 011, 2016.

[109] E. Méhes and T. Vicsek, "Collective motion of cells: From experiments to models," *Integrative biology*, vol. 6, no. 9, pp. 831–854, 2014.

[110] G. W. Brodland, "Computational modeling of cell sorting, tissue engulfment, and related phenomena: A review," *Applied Mechanics Reviews*, vol. 57, no. 1, pp. 47–76, 2004.

[111] P. Pathmanathan, J. Cooper, A. Fletcher, G. Mirams, P. Murray, J. Osborne, J. Pitt-Francis, A. Walter, and S. Chapman, "A computational study of discrete mechanical tissue models," *Physical biology*, vol. 6, no. 3, p. 036 001, 2009.

[112] M. Chiang and D. Marenduzzo, "Glass transitions in the cellular potts model," *Europhysics Letters*, vol. 116, no. 2, p. 28 009, 2016.

[113] D. M. Sussman, M. Paoluzzi, M. C. Marchetti, and M. L. Manning, "Anomalous glassy dynamics in simple models of dense biological tissue," *EPL (Europhysics Letters)*, vol. 121, no. 3, p. 36 001, 2018.

[114] D. Levis and L. Berthier, "From single-particle to collective effective temperatures in an active fluid of self-propelled particles," *EPL (Europhysics Letters)*, vol. 111, no. 6, p. 60 006, 2015.

[115] D. Bi, J. Lopez, J. Schwarz, and M. L. Manning, "A density-independent rigidity transition in biological tissues," *Nature Physics*, vol. 11, no. 12, p. 1074, 2015.

[116] A. J. Kabla, "Collective cell migration: Leadership, invasion and segregation," *Journal of The Royal Society Interface*, rsif20120448, 2012.

[117] F. Graner and J. A. Glazier, "Simulation of biological cell sorting using a two-dimensional extended potts model," *Physical review letters*, vol. 69, no. 13, p. 2013, 1992.

[118] J. A. Glazier and F. Graner, "Simulation of the differential adhesion driven rearrangement of biological cells," *Physical Review E*, vol. 47, no. 3, p. 2128, 1993.

[119] J. M. Belmonte, G. L. Thomas, L. G. Brunnet, R. M. De Almeida, and H. Chaté, "Self-propelled particle model for cell-sorting phenomena," *Physical Review Letters*, vol. 100, no. 24, p. 248 702, 2008.

[120] S. Henkes, Y. Fily, and M. C. Marchetti, "Active jamming: Self-propelled soft particles at high density," *Physical Review E*, vol. 84, no. 4, p. 040 301, 2011.

[121] B. Li and S. X. Sun, "Coherent motions in confluent cell monolayer sheets," *Biophysical journal*, vol. 107, no. 7, pp. 1532–1541, 2014.

[122] N. Sepúlveda, L. Petitjean, O. Cochet, E. Grasland-Mongrain, P. Silberzan, and V. Hakim, "Collective cell motion in an epithelial sheet can be quantitatively described by a stochastic interacting particle model," *PLoS computational biology*, vol. 9, no. 3, e1002944, 2013.

[123] S. Soumya, A. Gupta, A. Cugno, L. Deseri, K. Dayal, D. Das, S. Sen, and M. M. Inamdar, "Coherent motion of monolayer sheets under confinement and its pathological implications," *PLoS computational biology*, vol. 11, no. 12, e1004670, 2015.

[124] S. Alt, P. Ganguly, and G. Salbreux, "Vertex models: From cell mechanics to tissue morphogenesis," *Philosophical Transactions of the Royal Society B*, vol. 372, no. 1720, p. 20 150 520, 2017.

[125] D. Bi, J. H. Lopez, J. Schwarz, and M. L. Manning, "Energy barriers and cell migration in densely packed tissues," *Soft Matter*, vol. 10, no. 12, pp. 1885–1890, 2014.

[126] T. Nagai and H. Honda, "A dynamic cell model for the formation of epithelial tissues," *Philosophical Magazine B*, vol. 81, no. 7, pp. 699–719, 2001.

[127] A. G. Fletcher, J. M. Osborne, P. K. Maini, and D. J. Gavaghan, "Implementing vertex dynamics models of cell populations in biology within a consistent computational framework," *Progress in biophysics and molecular biology*, vol. 113, no. 2, pp. 299–326, 2013.

[128] D. L. Barton, S. Henkes, C. J. Weijer, and R. Sknepnek, "Active vertex model for cell-resolution description of epithelial tissue mechanics," *PLoS computational biology*, vol. 13, no. 6, e1005569, 2017.

[129] R. Farhadifar, J.-C. Röper, B. Aigouy, S. Eaton, and F. Jülicher, "The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing," *Current Biology*, vol. 17, no. 24, pp. 2095–2104, 2007.

[130] D. M. Sussman, J. Schwarz, M. C. Marchetti, and M. L. Manning, "Soft yet sharp interfaces in a vertex model of confluent tissue," *Physical review letters*, vol. 120, no. 5, p. 058 001, 2018.

[131] H. Honda and G. Eguchi, "How much does the cell boundary contract in a monolayered cell sheet?" *Journal of theoretical biology*, vol. 84, no. 3, pp. 575–588, 1980.

[132] G. Lan and T. Su, "Overcrowding drives the unjamming transition of gap-free monolayers," in *APS March Meeting Abstracts*, 2017.

[133] G. L. Dirichlet, "Über die reduction der positiven quadratischen formen mit drei unbestimmten ganzen zahlen.," *Journal für die reine und angewandte Mathematik*, vol. 40, pp. 209–227, 1850.

[134] R. Marcelpoil and Y. Usson, "Methods for the study of cellular sociology: Voronoi diagrams and parametrization of the spatial relationships," *Journal of Theoretical Biology*, vol. 154, no. 3, pp. 359–369, 1992.

[135] H. Honda, "Description of cellular patterns by dirichlet domains: The two-dimensional case," *Journal of Theoretical Biology*, vol. 72, no. 3, pp. 523–543, 1978.

[136] G. Schaller and M. Meyer-Hermann, "Multicellular tumor spheroid in an off-lattice voronoi-delaunay cell model," *Physical Review E*, vol. 71, no. 5, p. 051 910, 2005.

[137] M. Bock, A. K. Tyagi, J.-U. Kreft, and W. Alt, "Generalized voronoi tessellation as a model of two-dimensional cell tissue dynamics," *Bulletin of mathematical biology*, vol. 72, no. 7, pp. 1696–1731, 2010.

[138] E. Teomy, D. A. Kessler, and H. Levine, "Confluent and non-confluent phases in a model of cell tissue," *ArXiv preprint arXiv:1803.03962*, 2018.

[139] Y.-W. Li and M. P. Ciamarra, "Role of cell deformability in the two-dimensional melting of biological tissues," *Physical Review Materials*, vol. 2, no. 4, p. 045 602, 2018.

[140] X. Yang, D. Bi, M. Czajkowski, M. Merkel, M. L. Manning, and M. C. Marchetti, "Correlating cell shape and cellular stress in motile confluent tissues," *Proceedings of the National Academy of Sciences*, vol. 114, no. 48, pp. 12 663–12 668, 2017.

[141] T. A. Sharp, M. Merkel, M. L. Manning, and A. J. Liu, "Statistical properties of 3d cell geometry from 2d slices," *ArXiv preprint arXiv:1802.09131*, 2018.

[142] D. Matoz-Fernandez, K. Martens, R. Sknepnek, J. Barrat, and S. Henkes, "Cell division and death inhibit glassy behaviour of confluent tissues," *Soft matter*, vol. 13, no. 17, pp. 3205–3212, 2017.

[143] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, "A structural approach to relaxation in glassy liquids," *Nature Physics*, vol. 12, no. 5, p. 469, 2016.

[144] D. M. Sussman and M. Merkel, "No unjamming transition in a voronoi model of biological tissue," *Soft matter*, vol. 14, no. 17, pp. 3397–3403, 2018.

[145] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical review letters*, vol. 75, no. 6, p. 1226, 1995.

[146] E. Lång, A. Połeć, A. Lång, M. Valk, P. Blicher, A. Rowe, K. Tønseth, C. Jackson, T. Utheim, L. Janssen, J. Eriksson, and S. Ove Bøe, "Coordinated collective migration and asymmetric cell division in sheets of cultured human keratinocytes," *Nature Communications, accepted*, 2018.

[147] F. Ginelli, "The physics of the vicsek model," *The European Physical Journal Special Topics*, vol. 225, no. 11-12, pp. 2099–2117, 2016.

[148] T. Vicsek and A. Zafeiris, "Collective motion," *Physics Reports*, vol. 517, no. 3-4, pp. 71–140, 2012.

[149] F. Giavazzi, M. Paoluzzi, M. Macchi, D. Bi, G. Scita, M. L. Manning, R. Cerbino, and M. C. Marchetti, "Flocking transitions in confluent tissues," *Soft matter*, vol. 14, no. 18, pp. 3471–3477, 2018.

[150] B. A. Camley and W.-J. Rappel, "Velocity alignment leads to high persistence in confined cells," *Physical Review E*, vol. 89, no. 6, p. 062 705, 2014.

[151] C. Malinverno, S. Corallino, F. Giavazzi, M. Bergert, Q. Li, M. Leoni, A. Disanza, E. Frittoli, A. Oldani, E. Martini, T. Lendenmann, G. Deflorian, G. V. Beznoussenko, D. Poulikakos, K. H. Ong, M. Uroz, X. Trepat, D. Parazzoli, P. Maiuri, W. Yu, A. Ferrari, R. Cerbino, and G. Scita, "Endocytic reawakening of motility in jammed epithelia," *Nature materials*, vol. 16, no. 5, p. 587, 2017.

[152] X. Trepat and E. Sahai, "Mesoscale physical principles of collective cell organization," Jul. 2018.

[153] S. Torquato and F. H. Stillinger, "Local density fluctuations, hyperuniformity, and order metrics," *Physical Review E*, vol. 68, no. 4, p. 041 113, 2003.

[154] X. Li, A. Das, and D. Bi, "Biological tissue-inspired tunable photonic fluid," *Proceedings of the National Academy of Sciences*, p. 201 715 810, 2018.

[155] S.-H. Chong and M. Fuchs, "Mode-coupling theory for structural and conformational dynamics of polymer melts," *Physical review letters*, vol. 88, no. 18, p. 185 702, 2002.

[156] G. Szamel, "A theory for the dynamics of glassy mixtures with particle size swaps," *ArXiv preprint arXiv:1805.02753*, 2018.

[157] L. M. Janssen, P. Mayer, and D. R. Reichman, "Relaxation patterns in supercooled liquids from generalized mode-coupling theory," *Physical Review E*, vol. 90, no. 5, p. 052 306, 2014.

[158] E.-M. Schoetz, M. Lanio, J. A. Talbot, and M. L. Manning, "Glassy dynamics in three-dimensional embryonic tissues," *Journal of The Royal Society Interface*, vol. 10, no. 89, p. 20 130 726, 2013.

[159] M. Bayer, J. M. Brader, F. Ebert, M. Fuchs, E. Lange, G. Maret, R. Schilling, M. Sperl, and J. Wittmer, "Dynamic glass transition in two dimensions," *Physical Review E*, vol. 76, no. 1, p. 011 508, 2007.

[160] W. Götze, "The essentials of the mode-coupling theory for glassy dynamics," *Condensed Matter Physics*, 1998.

[161] N. Lačević, F. W. Starr, T. Schrøder, and S. Glotzer, "Spatially heterogeneous dynamics investigated via a time-dependent four-point density correlation function," *The Journal of chemical physics*, vol. 119, no. 14, pp. 7372–7387, 2003.

[162] S. C. Glotzer, "Spatially heterogeneous dynamics in liquids: Insights from simulation," *Journal of Non-Crystalline Solids*, vol. 274, no. 1-3, pp. 342–355, 2000.

[163] G. Biroli, J.-P. Bouchaud, K. Miyazaki, and D. R. Reichman, "Inhomogeneous mode-coupling theory and growing dynamic length in supercooled liquids," *Physical review letters*, vol. 97, no. 19, p. 195 701, 2006.

[164] J. H. Kim, X. Serra-Picamal, D. T. Tambe, E. H. Zhou, C. Y. Park, M. Sadati, J.-A. Park, R. Krishnan, B. Gweon, J. P. Millet Emil Butler, X. Trepat, and J. J. Fredberg, "Propulsion and navigation within the advancing monolayer sheet," *Nature materials*, vol. 12, no. 9, p. 856, 2013.

Acknowledgements

I would like to thank Liesbeth Janssen and Kees Storm for their supervision making it possible to fulfill my graduation project. I highly appreciate their time and input during this research and while writing this thesis. In addition, I would like to thank all other members of TPS for making me part of the group and motivating me during my research, as well as helping me with research- and software-related questions.

Furthermore, I would like to thank all people from the Fredberg lab for making me feel very welcome in their group, giving insights into their research and advising me on my research. Especially, I would like to thank Jim Butler, Jeffrey Fredberg, Emil Millet, Yoav Green and all others who attended our numerous meetings and discussions about mode coupling theory, as well as Stephan Koehler and Stephen DeCamp for making it possible to extend my analyses to experimental data.

Lastly, I would like to thank Sjak Jansen for supporting me during the whole year and especially for helping me with software-related issues and improving the representation of this thesis.

Contents Appendix

A List of symbols

Subsection 2.1

T_m	Melting temperature
T_g	Glass temperature
$\mathrm{MSD}(t)$	Mean squared displacement at time t
N	Number of particles
t	Time
$\boldsymbol{r}_i$	Position of particle i
i	Particle index
D	Diffusion coefficient
d	Dimensionality
τ_α	Relaxation time
η	Viscosity
k_B	Boltzmann constant
T	Temperature
r	Radius

Subsection 2.2.1

$A(\boldsymbol{r}, \boldsymbol{R}(t))$	A dynamic variable dependent on location $\boldsymbol{r}$ and particle configuration $\boldsymbol{\mathcal{R}}(t)$ at time t
$\boldsymbol{r}$	Position
$\boldsymbol{R}(t)$	Particle configuration
$\hat{A}_{\boldsymbol{k}}(t)$	Fourier transform of dynamic variable $A(\boldsymbol{r}, \boldsymbol{R}(t))$
$\boldsymbol{k}$	Wave vector
$\boldsymbol{R}_i(t)$	Position of particle i at time t
$\boldsymbol{\Gamma}$	Phase-space
$\boldsymbol{P}$	Momentum configuration
$\boldsymbol{P}_i$	Momentum of particle i
$f_{\mathrm{eq}}(\boldsymbol{\Gamma})$	Equilibrium probability of $\boldsymbol{\Gamma}$
β	Thermodynamic beta
$H(\boldsymbol{\Gamma})$	Hamiltonian for $\boldsymbol{\Gamma}$
$C(\Delta\boldsymbol{r}, t)$	Correlation function in real space
τ	Time difference
$C(\boldsymbol{k}, t)$	Correlation function in Fourier space
V	Two dimensional volume
$\rho(\boldsymbol{r}, \boldsymbol{R}(t))$	Local density function in real space
$\hat{\boldsymbol{\rho}}_k(t)$	Local density function in Fourier space
$G(\Delta\boldsymbol{r}, \tau)$	Van Hove correlation function
ρ_0	Average number density
j	Particle index
$g(\Delta\boldsymbol{r})$	Radial distribution function
$F(\boldsymbol{k}, \tau)$	Intermediate scattering function
$S(\boldsymbol{k})$	Static structure factor

Subsection 2.2.2

N^{ss}	Number of snapshots
m	Snapshot index
$R_{ij}(t)$	Distance between particle i and j
Δr	Magnitude of $\Delta\boldsymbol{r}$
$p(\Delta\boldsymbol{r})$	Probability of finding a particle at a vector $\Delta\boldsymbol{r}$ from another particle
n	Number density
$\overline{\Delta r}$	Location of the first peak of $g(\Delta\boldsymbol{r})$
k	Magnitude of $\boldsymbol{k}$
σ	Diameter
$\overline{k}$	Location of the first peak of $S(\boldsymbol{k})$
$F_s(k, \tau)$	Incoherent or self intermediate scattering function
$F_d(k, \tau)$	Distinct intermediate scattering function

Subsection 2.3.1

$\boldsymbol{A}(t)$	Set of phase space variables
$\boldsymbol{r}_i(t)$	Position of particle i at time t
$\boldsymbol{p}_i(t)$	Momentum of particle i at time t
$\mathcal{H}$	Hamilton operator
$\mathcal{L}$	Liouvillian operator
$\mathcal{P}$	Projection operator

f	A function
$\boldsymbol{\Omega}$	Frequency matrix
$\boldsymbol{M}$	Memory kernel
$\boldsymbol{f}(t)$	Fluctuation force
$\boldsymbol{C}(t)$	Correlation matrix
m	Mass
$j^L_{\boldsymbol{k}}$	Longitudinal density current
$R_{\boldsymbol{k}}(t)$	Random force
$M^0(\boldsymbol{k},t)$	Short-time memory kernel
$M^{\mathrm{MCT}}(\boldsymbol{k},t)$	Long-time memory kernel
A	A constant
$V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}$	Vertex dependent on wave vectors $\boldsymbol{q}$ and $\boldsymbol{k}$
$\boldsymbol{q}$	Wave vector
$c(k)$	Direct correlation function
c	A spring constant

Subsection 2.3.2

$P(\boldsymbol{\Gamma},t)$	Probability density
$\mathcal{S}$	Smoluchowski operator
D_0	Self-diffusion coefficient
$M_s^{\mathrm{MCT}}(\boldsymbol{k},t)$	Memory kernel for the incoherent ISF
$V^s_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}$	Vertex dependent on wave vectors $\boldsymbol{q}$ and $\boldsymbol{k}$ for the incoherent ISF

Subsection 2.3.3

C	A fitting parameter
γ	A fitting parameter
ϵ	Reduced temperature

Subsection 3.2

E	Interaction energy
a	Grid point index
b	Grid point index
$J_{a,b}$	Interaction energy between grid point a and b
$\sigma(a)$	Cell id of grid point a
λ	Penalty constant
A_i	Area of cell i
$A_{i,0}$	Target area of cell i
$\boldsymbol{F}_i^{\mathrm{friction}}$	Friction force on cell i
$\boldsymbol{F}_i^{\mathrm{act}}$	Active or self-propulsion force of cell i
$\boldsymbol{F}_i^{\mathrm{int}}$	Interaction force on cell i
γ	Friction constant
U	Interaction potential
E_i	Energy of cell i
K_{A_i}	Height elasticity of cell i
ξ_i	Elasticity constant of cell i
γ_i	Net line tension of cell i
P_i	Perimeter of cell i
v	Vertex index
$\boldsymbol{r}_v$	Position of the vertex v
$\mathcal{V}_i$	Voronoi cell i

Subsection 3.3.1

$P_{0,i}$	Preferred perimeter of cell i
K_{P_i}	Penalty constant for cell i for deviation from the preferred perimeter
A_{tot}	Total area of a layer
$\boldsymbol{n}_i$	Polarity vector of cell i
θ_i	Angle of the polarity vector of cell i
v_0	Velocity due to the self-propulsion force
μ	Mobility
η_i	White noise term for cell i
D_r	Rotational diffusion coefficient
r_0	Unit of length
τ_0	Unit of time
k_a	Dimensionless penalty for deviations from preferred area
a_i	Dimensionless area of cell i
a_0	Dimensionless preferred area

k_p	Dimensionless penalty for deviations from preferred perimeter
p_0	Dimensionless perimeter of cell i
p_i	Dimensionless preferred perimeter

Subsection 3.3.2

D_0	Diffusion coefficient of a free cell
D_{eff}	Effective diffusion coefficient

Subsection 3.4.1

R_V	Vicsek radius
τ_V	Vicsek persistence time
ϕ_i	Angle of the velocity vector of cell i

Subsection 3.4.2

J	Alignment interaction strength
τ	Reorientation time
N	Number of cells
VP	Vicsek parameter
$\boldsymbol{v}_i$	Velocity vector of cell i
AI	Alignment index
$\boldsymbol{v}_{\text{avg}}$	Average instantaneous velocity of all cells
IOP	Instantaneous order parameter
L_{corr}	Correlation length
r	Distance
μ	Mean
σ	Standard deviation
N_{ϕ_r}	Number of velocity angles at distance r
ϵ	Analysis parameter

Subsection 3.5

Δt	Step size
$n_{\text{initsteps}}$	Number of initialization steps
n_{steps}	Number of steps
$t_{\max}$	Maximum simulation time
$\overline{v}$	Average speed over time Δt

Subsection 7.3

w_i	Weight factor for cell i
$G_4(\boldsymbol{r}_1, \boldsymbol{r}_2, t)$	Time dependent four-point density for positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in real space
K_V	Penalty constant for deviation from a cells preferred volume
V_i	Volume of cell i
V_0	Preferred volume

Subsection C.1

$O(t)$	A time-dependent function
$v_i^{\boldsymbol{k}}$	Velocity of particle i in the direction of vector $\boldsymbol{k}$
$R_{\boldsymbol{k}}$	Random force
$\phi(\boldsymbol{r})$	Potential in real space
$\phi_{\boldsymbol{k}}$	Potential in Fourier space
$\mathcal{P}_2$	Projection operator

Subsection C.2

$\delta\rho_{\boldsymbol{k},s,i}$	Density of particle i
$j^L_{\boldsymbol{k},s,i}$	Longitudinal current of particle i
$R_{s,\boldsymbol{k}}$	Random force belonging to one particle
$\mathcal{P}_2^s$	Projection operator

Subsection D.2

α	Scaling parameter

Subsection D.3

α	Angle
β	Angle
h	height
a	Edge size
A_{tri}	Area of triangle

Subsection D.4

μ	A constant

Subsection D.5

n	An integer

Subsection E.1.1

x	Cartesian coordinate
y	Cartesian coordinate

Subsection E.1.2

$\boldsymbol{h}_{ijk}$	Circumcenter of the triangle given by $\boldsymbol{r}_i$, $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$
k	Cell seed index
R	Radius of a circumcircle
α	Weight factor
β	Weight factor
γ	Weight factor
d	Scaling factor
μ	Carthesian coordinate
ν	Carthesion coordinate
g	Cell seed index
l_{ij}	Length of the shared edge between cell i and j in a Voronoi tesselation
$\hat{\boldsymbol{n}}_{ij}$	Unit vector perpendicular to l_{ij}

Subsection E.2

s	Snapshot index
H	Histogram
Δh	Bin size
h_x	Bin index
h_y	Bin index
S_{temp}	Temporary function
N_w	Number of wave vectors
w	Wave vector index
$Y_k^{\cos}$	A function depending on the cosine of wave vector w
$Y_k^{\sin}$	A function depending on the sine of wave vector w

Subsection E.3.1

T_m	Time period indexed with m
N_T	Number of time segments in one period
ϵ	Algorithm parameter
A_k	Function dependent only dependent on wave vector k
B_k^t	Function dependent dependent on wave vector k and times before t
C_k	Function dependent only dependent on wave vector k
D_k^t	Function dependent dependent on wave vector k and times before t
$M^{\text{new}}(k,t)$	New estimate for $M(k,t)$
$M^{\text{old}}(k,t)$	Old estimate for $M(k,t)$

Subsection E.3.3

ϕ	Polar coordinate
$\boldsymbol{p}$	Wave vector

Subsection E.3.5

$A(\tau)$	Arbitrary time-dependent function
$B(\tau)$	Arbitrary time-dependent function
T_j	Time period indexed with j
j	Time index
n_1	Integer
n_2	Integer

Subsection F.2

T	Waiting time
Q	Area of FOV
P	Area smaller than FOV
L_c	Length defining the difference between Q and P
N_P	Number of cells in p

B Derivation of static and dynamic correlation functions in detail

In Subsection 2.2 the static and dynamic correlation functions and their origin are explained and in this standalone appendix section we give a full derivation of these correlation functions. In analogy with Subsection 2.2, we first derive the formal definitions of the radial distribution function, the static structure factor, the Van Hove correlation function and the intermediate scattering function in Appendix subsection B.1 and in Appendix subsection B.2 we derive how these functions can be measured in simulated or experimental data.

B.1 Correlation functions in theory

We start with determining the correlation function and density function in general. We consider a system with N point particles in a volume V and define the volume to be 2D in agreement with our research. We consider a dynamical variable $A(\boldsymbol{r}, \boldsymbol{R}(t))$ that is a function of position $\boldsymbol{r}$ and the space configuration of the particles, given by $\boldsymbol{R}(t) \equiv \{\boldsymbol{R}_1(t), \boldsymbol{R}_2(t), ..., \boldsymbol{R}_N(t)\}$ with $\boldsymbol{R}_i(t)$ the position of particle i at time t. The thermodynamic average (or ensemble average) of $A(\boldsymbol{r}, \boldsymbol{R}(t))$ is defined as

$$\langle A(\boldsymbol{r}, \boldsymbol{R}(t))\rangle \equiv \int \mathrm{d}\boldsymbol{\Gamma} f_{\mathrm{eq}}(\boldsymbol{\Gamma}) A(\boldsymbol{r}, \boldsymbol{R}(t)), \tag{84}$$

where $\boldsymbol{\Gamma}$ is the $4N$-dimensional phase space coordinate $\boldsymbol{\Gamma} \equiv \{\boldsymbol{R}_1, ..., \boldsymbol{R}_N, \boldsymbol{P}_1, ..., \boldsymbol{P}_N\} \equiv \{\boldsymbol{R}, \boldsymbol{P}\}$ with $\boldsymbol{P}_i$ the momentum of particle i and $f_{\mathrm{eq}}(\boldsymbol{\Gamma})$ is the equilibrium probability distribution of the positions and momenta of the particles. Because we assume a canonical system, the distribution function $f_{\mathrm{eq}}(\boldsymbol{\Gamma})$ is given by

$$f_{\mathrm{eq}}(\boldsymbol{\Gamma}) \equiv \frac{e^{-\beta H(\boldsymbol{\Gamma})}}{\int \mathrm{d}\boldsymbol{\Gamma}' e^{-\beta H(\boldsymbol{\Gamma}')}}, \tag{85}$$

with $\beta = (k_B T)^{-1}$ and H the Hamiltonian depending on the momenta $\boldsymbol{P}_i$ and interparticle distances $|\boldsymbol{R}_j - \boldsymbol{R}_i|$ of all particles i and j. The fluctuation in $A(\boldsymbol{r}, \boldsymbol{R}(t))$ is given by

$$\delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \equiv A(\boldsymbol{r}, \boldsymbol{R}(t)) - \langle A(\boldsymbol{r}, \boldsymbol{R}(t))\rangle . \tag{86}$$

The spatial Fourier transform of $A(\boldsymbol{r}, \boldsymbol{R}(t))$ is given by ($\hat{}$ indicates a function in Fourier space and $\boldsymbol{k}$ is a wave vector):

$$\hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) \equiv \int \mathrm{d}\boldsymbol{r} A(\boldsymbol{r}, \boldsymbol{R}(t)) e^{i\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{87}$$

and the inverse Fourier transform of $\hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t))$ gives $A(\boldsymbol{r}, \boldsymbol{R}(t))$ again:

$$A(\boldsymbol{r}, \boldsymbol{R}(t)) \equiv \frac{1}{(2\pi)^2} \int \mathrm{d}\boldsymbol{r} \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{88}$$

The fluctuations of the local density function in Fourier space as given in Equation 86 are

$$\delta\hat{A}_{\boldsymbol{k}}(t) = \int \mathrm{d}\boldsymbol{r} \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) - \left\langle \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) \right\rangle , \tag{89}$$

with $\left\langle \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) \right\rangle$ the Fourier transform of $\langle A(\boldsymbol{R}(t))\rangle$.

Having defined the density fluctuation functions in real and Fourier space, respectively $\delta A(\boldsymbol{r}, \boldsymbol{R}(t))$ and $\delta\hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t))$, we can now define correlation functions for these variables. The general space-time correlation function of $A(\boldsymbol{r}, \boldsymbol{R}(t))$ is

$$C(\boldsymbol{r}', t'; \boldsymbol{r}, t) \equiv \langle \delta A(\boldsymbol{r}', \boldsymbol{R}(t')) \delta A(\boldsymbol{r}, \boldsymbol{R}(t))\rangle , \tag{90}$$

with $\langle \quad \rangle$ again the ensemble average defined in Equation 84. After assuming spatially and temporal invariance, Equation 90 is only a function of the difference in space and time, respectively $\Delta \boldsymbol{r}$ and τ (assuming $t \geq t$):

$$C(\boldsymbol{r}-\boldsymbol{r}', t-t') = \langle \delta A(\boldsymbol{r}', \boldsymbol{R}(t') \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \rangle = \langle \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \delta A(\boldsymbol{r}+\Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \rangle \equiv C(\Delta \boldsymbol{r}, \tau), \tag{91}$$

where $\Delta \boldsymbol{r}$ and τ are space and time differences.

The second correlation function that we will need is the spatial Fourier transform of Equation 91:

$$C(\boldsymbol{k}, \tau) = \int \mathrm{d}(\Delta \boldsymbol{r}) C(\Delta \boldsymbol{r}, \tau) e^{i \boldsymbol{k} \cdot (\Delta \boldsymbol{r})} = \frac{1}{V} \left\langle \delta A_{\boldsymbol{k}}(\boldsymbol{R}(t)) \delta A^*_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \right\rangle, \tag{92}$$

where A^* is the complex conjugate of A. The proof of the second step is given in Appendix D.1.

To determine structure functions in real space we choose the variable $A(\boldsymbol{r}, \boldsymbol{R}(t))$ to be the local density function $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ defined by (with δ is the Dirac delta-function):

$$\rho(\boldsymbol{r}, \boldsymbol{R}(t)) \equiv \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)), \tag{93}$$

such that $\int_V \mathrm{d}\boldsymbol{r} \rho(\boldsymbol{r}, \boldsymbol{R}(t))$ is the total number of particles N in volume V. This formula measures whether at time t there is a particle at a vector $\boldsymbol{r}$ from the origin ($\lim_{\epsilon \to 0} \int_{\boldsymbol{r}-\boldsymbol{\epsilon}}^{\boldsymbol{r}+\boldsymbol{\epsilon}} \mathrm{d}\boldsymbol{r}' \rho(\boldsymbol{r}', \boldsymbol{R}(t)) = 1$) or not ($\lim_{\epsilon \to 0} \int_{\boldsymbol{r}-\boldsymbol{\epsilon}}^{\boldsymbol{r}+\boldsymbol{\epsilon}} \mathrm{d}\boldsymbol{r}' \rho(\boldsymbol{r}', \boldsymbol{R}(t)) = 0$). The density function depends on how the origin is chosen and is therefore called local.

By taking a spatial Fourier transform of Equation 93 we get the local density in Fourier space $\hat{\rho}_{\boldsymbol{k}}(t)$ and this will be our implementation of $\hat{A}_{\boldsymbol{k}}(t)$. This Fourier transform is given by

$$\hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \int \mathrm{d}\boldsymbol{r} \rho(\boldsymbol{r}, \boldsymbol{R}(t)) e^{i \boldsymbol{k} \cdot \boldsymbol{r}} = \sum_{i=1}^{N} e^{i \boldsymbol{k} \cdot \boldsymbol{R}_i(t)}, \tag{94}$$

where the wave vector $\boldsymbol{k}$ has a magnitude given by $k = \frac{2\pi}{\lambda}$ with λ the wavelength. The density in Fourier space given by Equation 94 decomposes the real density function given by Equation 93 into the waves it is build of. This can be visualized as follows. Assuming we choose the origin random, then $\hat{\rho}_{\boldsymbol{k}}(t) = 1$ when there is a particle i for which $\boldsymbol{k} \cdot \boldsymbol{R}_i(t) = 2\pi n$, with n an integer. This means that a wave with wave vector $\boldsymbol{k}$ arrives at particle i after n full cycles, because the position of that wave is then given by $\frac{2\pi n}{\boldsymbol{k}} = \boldsymbol{R}_i(t)$. If $\boldsymbol{k} \cdot \boldsymbol{R}_i(t) \neq 2\pi n$ the contribution of particle i to $\rho_{\boldsymbol{k}}(t)$ is complex and the real value $\Re(\hat{\rho}_{\boldsymbol{k}})(\boldsymbol{R}(t)) = \cos(\boldsymbol{k} \cdot \boldsymbol{R}_i(t))$ measures the value of a wave with wave vector k, starting at the origin with value 1, at the position of particle i. So the combination of the functions for all these wave vectors $\boldsymbol{k}$ defines the local density in Fourier space. From Equation 94 we derive the following relationship

$$\hat{\rho}^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \hat{\rho}_{-\boldsymbol{k}}(\boldsymbol{R}(t)). \tag{95}$$

According to Equations 86 and 89 the fluctuation functions in the local density in real and Fourier space are respectively

$$\delta \rho(\boldsymbol{r}, \boldsymbol{R}(t)) = \rho(\boldsymbol{r}, \boldsymbol{R}(t)) - \langle \rho(\boldsymbol{r}, \boldsymbol{R}(t)) \rangle = \rho(\boldsymbol{r}, \boldsymbol{R}(t)) - \rho_0, \tag{96}$$

where $\rho_0 \equiv \langle \rho(\boldsymbol{r}, \boldsymbol{R}(t)) \rangle$ is the average density, and

$$\delta \hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) - \langle \hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) \rangle = \hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t)) - \hat{\rho}_0, \tag{97}$$

where $\hat{\rho}_0 = \langle \rho_{\boldsymbol{k}}(\boldsymbol{R}(t)) \rangle$ is the Fourier transform of the average density.

Using the density fluctuation function in real space given by Equation 96 and the correlation function in real space given by Equation 91, we determine the Van Hove distribution function $G(\Delta \boldsymbol{r}, t)$ in Appendix Subsection B.1.1 and the radial distribution function $g(\Delta \boldsymbol{r})$, a static equivalent of $G(\Delta \boldsymbol{r}, t)$, in Appendix Subsection B.1.2. Using the density fluctuation function in Fourier space given by Equation 97 and the correlation function in Fourier space given by Equation 92, we determine the intermediate scattering function $F(\boldsymbol{k}, \tau)$ in Appendix Subsection B.1.3 and the static structure factor $S(\boldsymbol{k})$, a static equivalent of $F(\boldsymbol{k}, \tau)$, in Appendix Subsection B.1.4.

B.1.1 Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$

The Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$, also called the dynamic density correlation function, is defined as the correlation function (Equation 91) for the local density fluctuations (Equation 96) divided by ρ_0 for normalization purposes :

$$
\begin{aligned}
G(\Delta \boldsymbol{r}, \tau) &\equiv \frac{1}{\rho_0} C(\Delta \boldsymbol{r}, \tau) && (98)\\
&= \frac{1}{\rho_0} \left\langle \delta\rho(\boldsymbol{r}, \boldsymbol{R}(t)) \delta\rho(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \right\rangle && (99)\\
&= \frac{1}{\rho_0} \left\langle \left(\rho(\boldsymbol{r}, \boldsymbol{R}(t)) - \rho_0\right) \left(\rho(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\Delta t)) - \rho_0\right) \right\rangle && (100)\\
&= \frac{1}{\rho_0} \left\langle \rho(\boldsymbol{r}, \boldsymbol{R}(t)) \rho(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \right\rangle - \rho_0 && (101)\\
&= \frac{1}{\rho_0} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j=1}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t+\tau) \right\rangle - \rho_0, && (102)
\end{aligned}
$$

where in Equation 101 is used that the system is translational and time invariant.

B.1.2 Radial distribution function $g(\Delta \boldsymbol{r})$

The pair correlation function $h(\Delta \boldsymbol{r})$ is a static variant of the the Van Hove distribution function (Equation 102) where we divide again by ρ_0 for normalization purposes and subtract the correlation of a particle with itself $(i = j)$ as this gives a δ-peak at the origin $\Delta \boldsymbol{r} = 0$,

$$
\begin{aligned}
h(\Delta \boldsymbol{r}) &\equiv \frac{1}{\rho_0} \left(G(\Delta \boldsymbol{r}, 0) - \delta(\Delta \boldsymbol{r})\right) && (103)\\
&= \frac{1}{\rho_0^2} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t) \sum_{j \neq i}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t)) \right\rangle - 1. && (104)
\end{aligned}
$$

To obtain the radial distribution function $g(\Delta(\boldsymbol{r}))$ we add the trivial contribution of an ideal gas to the pair correlation function $h(\Delta(\boldsymbol{r})$:

$$
\begin{aligned}
g(\Delta \boldsymbol{r}) &\equiv h(\Delta \boldsymbol{r}) + 1 && (105)\\
&= \frac{1}{\rho_0^2} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j \neq i}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t)) \right\rangle. && (106)
\end{aligned}
$$

B.1.3 Intermediate scattering function $F(\boldsymbol{k}, \tau)$

The intermediate scattering function $F(\Delta \boldsymbol{r}, \tau)$, also called the dynamic correlation function in Fourier space, is defined as the correlation function (Equation 92) of the local density fluctuations in Fourier space (Equation 97) divided by ρ_0 for normalization purposes:

$$F(\boldsymbol{k},\tau) \equiv \frac{1}{\rho_0} C(\boldsymbol{k},\tau) \tag{107}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \langle \delta\hat{\rho}^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) \delta\hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \rangle \tag{108}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \langle (\hat{\rho}_{-\boldsymbol{k}}(\boldsymbol{R}(t)) - \hat{\rho}_0)\,(\hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) - \hat{\rho}_0) \rangle \tag{109}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \langle \hat{\rho}_{-\boldsymbol{k}}(\boldsymbol{R}(t)) \hat{\rho}_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \rangle - \frac{\hat{\rho}_0}{V} \tag{110}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \left\langle \sum_{i=1}^{N} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}_i(t)} \sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau)} \right\rangle - \frac{\hat{\rho}_0}{V} \tag{111}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t))} \right\rangle - \frac{\hat{\rho}_0}{V}. \tag{112}$$

Because $\hat{\rho}_0$ is the Fourier transform of a constant and therefore causes a δ-function at $\boldsymbol{k} = 0$, this term will be neglected and we redefine the intermediate scattering function as

$$F(\boldsymbol{k},\tau) \equiv \frac{1}{\rho_0}\frac{1}{V} \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t))} \right\rangle. \tag{113}$$

B.1.4 Static structure factor $S(\boldsymbol{k})$

The static structure factor $S(\boldsymbol{k})$ is the static variant of the intermediate scattering function given by Equation 113:

$$S(\boldsymbol{k}) \equiv F(\boldsymbol{k}, 0) \tag{114}$$

$$= \frac{1}{\rho_0}\frac{1}{V} \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))} \right\rangle. \tag{115}$$

Note that we did not divide $F(\boldsymbol{k},\tau)$ by ρ_0 to get $S(\boldsymbol{k})$ as we did with $G(\Delta\boldsymbol{r},\tau)$ to get $g(\Delta\boldsymbol{r})$.

B.2 Correlation functions in measurements

In the previous appendix subsection we have derived the general formulas for four correlation functions and in this appendix subsection we will define how we can measure these structure functions in data obtained from experiments or simulations. In theory, $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ is known for the complete phase-space $\boldsymbol{R}$ at time t and for every position $\boldsymbol{r}$. However, in measurements the value of $\rho(\boldsymbol{r}, \boldsymbol{R}(t))$ is only known at points $\boldsymbol{r}$ where and times t when measured. So, to calculate an estimate of the ensemble average given by Equation 84, we use two modifications: one concerning $\boldsymbol{r}$ and one concerning t.

We assume that the data consists of N^{ss} snapshots m of the system with volume V at different moments in time. Furthermore, we will assume that the system is ergodic, which means that the fraction of time the system is in a certain microscopic state is equal to the a priori probability to find the system in this state. This means that the ensemble average of a system can be obtained by measuring the time average, which we will do by calculating the unweighted average over N_{ss} snapshots. A second modification to calculate the ensemble average is done by averaging over the whole volume (area in 2D) of a snapshot. For each snapshot m we only have one configuration $\boldsymbol{R}$ but by varying $\boldsymbol{r}$ over the whole volume, we average over all possible origins. Because the interparticle distances $|\boldsymbol{R}_j - \boldsymbol{R}_i|$ and therefore the Hamiltonian $H(\boldsymbol{\Gamma})$ and the value of $f_{\text{eq}}(\boldsymbol{\Gamma})$ are independent of the location of the origin, we can again calculate an unweighted average but now over all possible positions of the origin. Combining this with the averaging over time, we define a new ensemble average that we can actually measure:

$$\langle \quad \rangle = \int \mathrm{d}\boldsymbol{\Gamma} f_{\mathrm{eq}}(\boldsymbol{\Gamma}) \approx \left\langle \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \quad \right\rangle_t \equiv \frac{1}{N^{ss}} \sum_{m=1}^{N^{ss}} \frac{1}{V} \int \mathrm{d}\boldsymbol{r}, \tag{116}$$

where $\langle \quad \rangle_t$ means an unweighted average over all N^{ss} snapshots.

Using this new ensemble average we can calculate the average density ρ_0

$$\rho_0 = \langle \rho(\boldsymbol{r}, \boldsymbol{R}(t)) \rangle = \frac{1}{N^{ss}} \sum_{m=1}^{N^{ss}} \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \rho(\boldsymbol{r}, \boldsymbol{R}(t)) = \frac{1}{N^{ss}} \sum_{m=1}^{N^{ss}} \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) = \frac{N}{V}, \tag{117}$$

and its Fourier transform $\hat{\rho}_0$ is given by

$$\hat{\rho_0} = \int \mathrm{d}\boldsymbol{r} \rho_0 e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = (2\pi)^2 \rho_0 \delta(\boldsymbol{k}). \tag{118}$$

Using the new ensemble average of Equation 116, we will calculate the measurable versions of the correlation functions defined in Appendix Section B.1. Following a more logical order, we start with the radial distribution function in Appendix Subsection B.2.1 followed by the Van Hove distribution function in Appendix Subsection B.2.2. In Appendix Subsection B.2.3 we discuss the static structure factor and in Appendix Subsection B.2.4 the intermediate scattering function.

B.2.1 Radial distribution function $g(\Delta \boldsymbol{r})$

Applying the bracket implementation defined in Equation 116 to the radial distribution function $g(\Delta \boldsymbol{r})$ given in Equation 105, we obtain

$$g(\Delta \boldsymbol{r}) = \frac{1}{\rho_0^2} \left\langle \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j \neq i}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t)) \right\rangle \tag{119}$$

$$\approx \frac{1}{\rho_0^2} \left\langle \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t) \sum_{j \neq i}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t)) \right\rangle_t \tag{120}$$

$$= \frac{1}{\rho_0 N} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta\left(\Delta \boldsymbol{r} - (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))\right) \right\rangle_t . \tag{121}$$

Besides $g(\Delta \boldsymbol{r})$ being a correlation function of local densities, Equation 121 shows another interpretation. $g(\Delta \boldsymbol{r})$ is namely an average of N density functions $\rho(\boldsymbol{r})$, one for each particle i, and divided by the result of this average for an ideal gas, namely ρ_0. So the radial distribution function measures how many particles there are on average at vector $\Delta \boldsymbol{r}$ from a particle with respect to this number in an ideal gas and is therefore proportional to the probability $p(\Delta \boldsymbol{r})$ of finding a particle at position $\Delta \boldsymbol{r}$, given that there is a particle at the origin.

In case the system is isotropic (identical in all directions) we can change from the vector $\Delta \boldsymbol{r}$ to its magnitude $|\Delta \boldsymbol{r}| = \Delta r$:

$$g(\Delta r) = \frac{1}{\rho_0} \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j \neq i}^{N} \delta(r - R_{ij}(t)) \right\rangle_t , \tag{122}$$

with $R_{ij}(t) = |\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t)|$. The average number of particles n in a shell of thickness $\mathrm{d}(\Delta r)$ at distance Δr from a particle is given by $\mathrm{d}n(\Delta r) = \rho_0 g(\Delta r) \mathrm{d}^2(\Delta r)$.

B.2.2 Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$

Applying the bracket implementation defined in Equation 116 to the Van Hove distribution function $G(\Delta \boldsymbol{r}, \tau)$ given in Equation 102, we obtain

$$G(\Delta \boldsymbol{r}, \tau) = \frac{1}{\rho_0} \left\langle \sum_{j=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{i=1}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t+\tau)) \right\rangle - \rho_0 \tag{123}$$

$$\approx \frac{1}{\rho_0} \left\langle \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{R}_i(t)) \sum_{j=1}^{N} \delta(\boldsymbol{r} + \Delta \boldsymbol{r} - \boldsymbol{R}_j(t+\tau)) \right\rangle_t - \rho_0 \tag{124}$$

$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\Delta \boldsymbol{r} - (\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t))) \right\rangle_t - \rho_0. \tag{125}$$

The Van Hove distribution function is a dynamic variant of the radial distribution function and is proportional to the probability of finding a particle at position $\Delta \boldsymbol{R}$ at time $t + \tau$ given that there is a particle at the origin at time t.

Again, in case of an isotropic system we can write

$$G(\Delta r, \tau) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(r - |\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t)|) \right\rangle_t . \tag{126}$$

B.2.3 Static structure factor $S(\boldsymbol{k})$

Applying the bracket implementation defined in Equation 116 to the static structure factor $S(\boldsymbol{k})$ given in Equation 115, we obtain

$$S(\boldsymbol{k}) = \frac{1}{\rho_0} \frac{1}{V} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k} \cdot (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))} \right\rangle \tag{127}$$

$$\approx \frac{1}{\rho_0} \left\langle \frac{1}{V} \int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k} \cdot (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))} \right\rangle_t \tag{128}$$

$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{i\boldsymbol{k} \cdot (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))} \right\rangle_t \tag{129}$$

$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \cos\left(\boldsymbol{k} \cdot (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))\right) + i \sin\left(\boldsymbol{k} \cdot (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))\right) \right\rangle_t \tag{130}$$

$$= \frac{1}{N} \left\langle \left(\sum_{i=1}^{N} \cos\left(\boldsymbol{k} \cdot \boldsymbol{R}_i(t)\right) \right)^2 + \left(\sum_{j=1}^{N} \sin\left(\boldsymbol{k} \cdot \boldsymbol{R}_j(t)\right) \right)^2 \right\rangle_t . \tag{131}$$

where we have used that the ensemble average of the sinus, an uneven function, vanishes. $S(\boldsymbol{k})$ measures how much the structure of the system contains a periodicity matching the wave vector $\boldsymbol{k}$.

In case of an isotropic system we can change from the wave vector $\boldsymbol{k}$ to its magnitude $k = |\boldsymbol{k}|$ to get

$$S(k) = \frac{1}{N} \left\langle \left(\sum_{i=1}^{N} \cos\left(\boldsymbol{k} \cdot \boldsymbol{R}_i(t)\right) \right)^2 + \left(\sum_{j=1}^{N} \sin\left(\boldsymbol{k} \cdot \boldsymbol{R}_j(t)\right) \right)^2 \right\rangle_t \tag{132}$$

The static structure factor can also be calculated by taking the Fourier transform of the radial distribution function and to derive this we start with rewriting Equation 129 to

$$S(\boldsymbol{k}) = \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t)-\boldsymbol{R}_i(t))}\right\rangle_t \tag{133}$$

$$= 1 + \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j\neq 1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t)-\boldsymbol{R}_i(t))}\right\rangle_t \tag{134}$$

$$= 1 + \frac{1}{N}\left\langle \int \mathrm{d}(\Delta\boldsymbol{r}) \sum_{i=1}^{N}\sum_{j\neq 1}^{N} \delta(\Delta\boldsymbol{r} - (\boldsymbol{R}_j(t) - \boldsymbol{R}_i(t))) e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})}\right\rangle_t \tag{135}$$

$$= 1 + \rho_0 \int \mathrm{d}(\Delta\boldsymbol{r})(g(\Delta\boldsymbol{r}) - 1)e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})} + (2\pi)^2\rho_0\delta(\boldsymbol{k}). \tag{136}$$

Note that we have used the pair correlation function $h(\Delta\boldsymbol{r})$ because the Fourier transform of $g(\Delta\boldsymbol{r})$ does not converge ($\lim_{\Delta\boldsymbol{r}\to\infty} g(\Delta\boldsymbol{r}) = 1$ while $\lim_{\Delta\boldsymbol{r}\to\infty} h(\Delta\boldsymbol{r}) = 0$). The non-converging term arises now as a δ-peak at $\boldsymbol{k} = 0$.

When we again assume isotropy, we can define the static structure factor $S(k)$ as a function of the radial distribution function $g(\Delta r)$. In doing this, we neglect the δ-term, making $S(k)$ indetermined for $k = 0$.

$$S(k) = 1 + \rho_0 \int \mathrm{d}(\Delta r) e^{ik\cdot\Delta r}(g(\Delta r) - 1) \qquad (k \neq 0) \tag{137}$$

$$= 1 + \rho_0 \int \mathrm{d}(\Delta r)(\cos(k\cdot\Delta r) + i\sin(k\cdot\Delta r))(g(\Delta r) - 1) \qquad (k \neq 0) \tag{138}$$

$$= 1 + \rho_0 \int_0^\infty \mathrm{d}(\Delta r)\Delta r \int_0^{2\pi} \mathrm{d}\theta \left(\cos(k(\Delta r)\cos\theta) + i\sin(k(\Delta r)\cos\theta)\right)(g(\Delta r) - 1) \qquad (k \neq 0) \tag{139}$$

$$= 1 + \rho_0 \int_0^\infty \mathrm{d}(\Delta r)\Delta r \int_0^{2\pi} \mathrm{d}\theta \left(\cos(k(\Delta r)\cos\theta)\right)(g(\Delta r) - 1) \qquad (k \neq 0) \tag{140}$$

$$= 1 + 2\pi\rho_0 \int_0^\infty \mathrm{d}(\Delta r)\Delta r J_0(k\Delta r)(g(\Delta r) - 1) \qquad (k \neq 0). \tag{141}$$

where is used that $\int_0^{2\pi} \mathrm{d}\theta \sin(k(\Delta r)\cos\theta) = 0$ and $J_0(k\Delta r)$ is a Bessel function of the first kind.

B.2.4 Intermediate scattering function $F(\boldsymbol{k}, \tau)$

Applying the bracket implementation defined in Equation 116 to the intermediate scattering function $F(\boldsymbol{k}, \tau)$ given in Equation 113, we obtain

$$F(\boldsymbol{k}, \tau) = \frac{1}{\rho_0}\frac{1}{V}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau)-\boldsymbol{R}_i(t))}\right\rangle \tag{142}$$

$$\approx \frac{1}{\rho_0}\left\langle \frac{1}{V}\int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau)-\boldsymbol{R}_i(t))}\right\rangle_t \tag{143}$$

$$= \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} \cos(\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t))) + i\sin(\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau) - \boldsymbol{R}_i(t)))\right\rangle_t \tag{144}$$

$$= \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right) + \left(\sum_{i=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right)\right\rangle_t, \tag{145}$$

where we have used that the ensemble average of sinus, an uneven function, vanishes. The intermediate scattering function measures how much of the structure of a system is still present after some time τ. Again we can define an isotropic version of the intermediate scattering function

$$F(k,\tau) = \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right) + \left(\sum_{i=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right)\right\rangle_t . \tag{146}$$

The intermediate scattering function can also be calculated by taking the Fourier transform of the van Hove correlation function, and to derive this we start with rewriting Equation 146 to

$$F(\boldsymbol{k},\tau) = \frac{1}{N}\left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\boldsymbol{k}\cdot(\boldsymbol{R}_j(t+\tau)-\boldsymbol{R}_i(t))}\right\rangle_t \tag{147}$$

$$= \frac{1}{N}\left\langle \int \mathrm{d}(\Delta\boldsymbol{r})\sum_{i=1}^{N}\sum_{j=1}^{N}\delta\left(\Delta\boldsymbol{r}-(\boldsymbol{R}_j(t+\tau)-\boldsymbol{R}_i(t))\right)e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})}\right\rangle_t \tag{148}$$

$$= \int \mathrm{d}(\Delta\boldsymbol{r})(G(\Delta\boldsymbol{r},\tau)e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})}. \tag{149}$$

When we again assume isotropy this results in

$$F(k,\tau) = \int \mathrm{d}(\Delta r)(G(\Delta r,\tau)e^{ik\cdot(\Delta r)} \tag{150}$$

$$= \int \mathrm{d}(\Delta r)(\cos(k\cdot\Delta r) + i\sin(k\cdot\Delta r))(G(\Delta r,\tau) \tag{151}$$

$$= \int_0^\infty \mathrm{d}(\Delta r)\Delta r\int_0^{2\pi}\mathrm{d}\theta\left(\cos(k(\Delta r)\cos\theta) + i\sin(k(\Delta r)\cos\theta)\right)(G(\Delta r,\tau) \tag{152}$$

$$= \int_0^\infty \mathrm{d}(\Delta r)\Delta r\int_0^{2\pi}\mathrm{d}\theta\left(\cos(k(\Delta r)\cos\theta)\right)(G(\Delta r,\tau) \tag{153}$$

$$= 2\pi\int_0^\infty \mathrm{d}(\Delta r)\Delta r J_0(k\Delta r)G(\Delta r), \tag{154}$$

where is used that $\int_0^{2\pi}\mathrm{d}\theta\sin(k(\Delta r)\cos\theta) = 0$ and and $J_0(k\Delta r)$ is a Besselfunction of the first kind.
The intermediate scattering function can be split up in a self $F_s(k,\tau)$ and a distinct part $F_d(k,\tau)$ as follows

$$F(k,\tau) \tag{155}$$

$$= \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right) + \left(\sum_{i=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\right)\left(\sum_{j=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right)\right\rangle_t \tag{156}$$

$$= \frac{1}{N}\left\langle \sum_{i=1}^{N}\left(\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t+\tau))\right) + \sum_{i=1}^{N}\left(\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t+\tau))\right)\right\rangle_t \tag{157}$$

$$+ \frac{1}{N}\left\langle \sum_{i=1}^{N}\left(\cos(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\sum_{j\neq i}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau)\right) + \sum_{i=1}^{N}\left(\sin(\boldsymbol{k}\cdot\boldsymbol{R}_i(t))\sum_{j\neq i}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{R}_j(t+\tau))\right)\right\rangle_t \tag{158}$$

$$= F_s(k,\tau) + F_d(k,\tau), \tag{159}$$

where $F_s(k,\tau)$ describes the correlation of the position of a particle at time $t+\tau$ with the position of the same particle at time t.

C Derivation of mode coupling theory for glass-forming liquids in detail

In Subsection 2.3.1 the derivation of the standard mode coupling theory for glass-forming fluids is sketched and in this appendix section we give the full derivation based on Reichman and Charbonneau [49]. In Appendix Subsection C.2 we explain how the derivation changes for the incoherent instead of the coherent intermediate scattering function.

C.1 Derivation of the standard MCT-equation for the coherent ISF

Consider an N-particle system with $\boldsymbol{A}(\boldsymbol{r}_1(t), ..., \boldsymbol{r}_N(t), \boldsymbol{p}_1(t), ..., \boldsymbol{p}_N(t))$ a set of phase-space variables dependent on the positions $\boldsymbol{r}_i(t)$ and momenta $\boldsymbol{p}_i(t)$ of all particles i at time t, which we shortly denote with $\boldsymbol{A}(t)$. The derivation of $\boldsymbol{A}(t)$ to time can be written as

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = \sum_i \left(\frac{\delta \boldsymbol{A}(t)}{\delta \boldsymbol{r}_i} \frac{\delta \boldsymbol{r}_i}{\delta t} - \frac{\delta \boldsymbol{A}(t)}{\delta \boldsymbol{p}_i} \frac{\delta \boldsymbol{p}_i}{\delta t} \right) \equiv \{\boldsymbol{A}(t), \mathcal{H}\} \equiv i\mathcal{L}\boldsymbol{A}(t), \tag{160}$$

where the Hamilton equations are used, given by

$$\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{r}_i}, \quad \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_i}. \tag{161}$$

Furthermore, $\{ \quad , \quad \}$ defines the classical Poisson brackets, which for arbitrary functions A and B is expressed as

$$\{A, B\} \equiv \sum_i \left(\frac{\delta A}{\delta \boldsymbol{r}_i} \cdot \frac{\delta B}{\delta \boldsymbol{p}_i} - \frac{\delta A}{\delta \boldsymbol{p}_i} \cdot \frac{\delta B}{\delta \boldsymbol{r}_i} \right), \tag{162}$$

and the Liouvillian operator $\mathcal{L}$ is given by

$$i\mathcal{L} \equiv -\sum \frac{\partial}{\partial \boldsymbol{r}_i} \frac{\partial \boldsymbol{r}_i}{\partial t} + \frac{\partial}{\partial \boldsymbol{p}_i} \frac{\partial \boldsymbol{p}_i}{\partial t}. \tag{163}$$

Starting with this Liouvillian equation, mode coupling theory can be derived in three steps. First, the Mori-Zwanzig formalism is applied to rewrite the equation of motion for an arbitrary correlation function of $\boldsymbol{A}(t)$ to a generalized Langevin equation (Appendix Subsection C.1.1). Secondly, a specific implementation of $\boldsymbol{A}(t)$ is used to obtain the equation of motion for the incoherent intermediate scattering function (Appendix Subsection C.1.2). This equation of motion depends on a memory kernel that cannot be solved exactly and we approximate this memory kernel in Appendix Subsection C.1.3. We give the resulting MCT-equation in Appendix Subsection C.1.4.

On a notational note, we use $\boldsymbol{A} \equiv \boldsymbol{A}(\tau)$ and $\boldsymbol{A}(t) \equiv \boldsymbol{A}(\tau + t)$ to distinguish between the two values at an arbitrary time τ and at a time t later. We apply the same definitions for other variables like the fluctuation force $\boldsymbol{f}$, the density ρ, etc. Furthermore, pay attention to two different uses of i, namely as the imaginary unit and as index for particles.

C.1.1 Application of the Mori-Zwanzig formalism to obtain a generalized Langevin Equation

The first step in the derivation of mode coupling theory is to apply the Mori-Zwanzig formalism [54], which is a widely used formalism to rewrite an equation of motion based on a separation of time scales. The reason we apply this principle for supercooled liquids is that they exhibit fluctuations on a microscopic time scale while other dynamical processes such as the relaxation occur on a time scale that can be many magnitudes of orders larger. We are only interested in these so-called slow variables and we assume that fast variables influence these slow variables only via a fluctuation force. With the Mori-Zwanzig principle we split the slow variables from the fast variables. To do so, we define a projection operator $\mathcal{P}$ that projects a function f along the direction of the slow variable $\boldsymbol{A}$,

$$\mathcal{P}f \equiv \frac{(\boldsymbol{A}, f)}{(\boldsymbol{A}, \boldsymbol{A})}\boldsymbol{A}, \tag{164}$$

where $(..., ...)$ denotes the ensemble average of a dyadic product, given by $\langle \boldsymbol{A}^* \boldsymbol{A} \rangle$. Integrating Equation 160 gives

$$\boldsymbol{A}(t) = e^{i\mathcal{L}t}\boldsymbol{A}, \tag{165}$$

and taking again the derivative to t gives

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = e^{i\mathcal{L}t} i\mathcal{L}\boldsymbol{A} = e^{i\mathcal{L}t}(\mathcal{P} + 1 - \mathcal{P}) i\mathcal{L}\boldsymbol{A}, \tag{166}$$

because we are free to insert $(\mathcal{P} + 1 - \mathcal{P})$. Working out Equation 166 (with applying Equation 165) gives

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = i\boldsymbol{\Omega} \cdot \boldsymbol{A}(t) + e^{i\mathcal{L}t}(1 - \mathcal{P}) i\mathcal{L}\boldsymbol{A}, \quad \text{with} \quad i\boldsymbol{\Omega} = (\boldsymbol{A}, i\mathcal{L}\boldsymbol{A}) \cdot (\boldsymbol{A}, \boldsymbol{A})^{-1}, \tag{167}$$

where $\boldsymbol{\Omega}$ is called the frequency matrix.

We split $e^{i\mathcal{L}t}$ into a normal and orthogonal part as follows

$$e^{i\mathcal{L}t} = e^{i\mathcal{L}t} O(t) + e^{i(1-\mathcal{P})\mathcal{L}t}, \tag{168}$$

because $1 - P$ is orthogonal to P and we derive $O(t)$ by first deriving Equation 168 to t. This results in

$$i\mathcal{L}e^{i\mathcal{L}t} = i\mathcal{L}e^{i\mathcal{L}t}O(t) + e^{i\mathcal{L}t}\frac{\mathrm{d}O(t)}{\mathrm{d}t} + i(1-\mathcal{P})\mathcal{L}e^{i(1-\mathcal{P})\mathcal{L}t} \tag{169}$$

$$i\mathcal{L}(e^{i\mathcal{L}t}O(t) + e^{i(1-\mathcal{P})\mathcal{L}t}) = i\mathcal{L}e^{i\mathcal{L}t}O(t) + e^{i\mathcal{L}t}\frac{\mathrm{d}O(t)}{\mathrm{d}t} + i(1-\mathcal{P})\mathcal{L}e^{i(1-\mathcal{P})\mathcal{L}t} \tag{170}$$

$$e^{i\mathcal{L}t}\frac{\mathrm{d}O(t)}{\mathrm{d}t} = i\mathcal{P}\mathcal{L}e^{i(1-\mathcal{P})\mathcal{L}t}. \tag{171}$$

With $O(0) = 0$ obtained from Equation 168, the integral of Equation 171 is given by

$$O(t) = i\int_0^t \mathrm{d}\tau e^{-i\mathcal{L}\tau}\mathcal{P}\mathcal{L}e^{i(1-\mathcal{P})\mathcal{L}\tau}. \tag{172}$$

Filling Equation 172 in in Equation 168 and multiplying by $i(1 - \mathcal{P})\mathcal{L}\boldsymbol{A}$ gives

$$e^{i\mathcal{L}t} i(1-\mathcal{P})\mathcal{L}\boldsymbol{A} = \int_0^t \mathrm{d}\tau e^{i\mathcal{L}(t-\tau)} i\mathcal{P}\mathcal{L}\boldsymbol{f}(\tau) + \boldsymbol{f}(t), \tag{173}$$

with $\boldsymbol{f}(t)$ the fluctuating force defined by

$$\boldsymbol{f}(t) \equiv e^{i(1-\mathcal{P})\mathcal{L}t} i(1-\mathcal{P})\mathcal{L}\boldsymbol{A}, \tag{174}$$

and $\boldsymbol{f}(t)$ is orthogonal to $\boldsymbol{A}$. Therefore, we can write

$$(\boldsymbol{A}, i\mathcal{L}\boldsymbol{f}(t)) = i(\mathcal{L}\boldsymbol{A}, \boldsymbol{f}(t)) = i((1-\mathcal{P})\mathcal{L}\boldsymbol{A}, \boldsymbol{f}(t)) = -(\boldsymbol{f}, \boldsymbol{f}(t)). \tag{175}$$

Filling in Equation 173 together with Equation 175 into Equation 167 gives for the equation of motion of $\boldsymbol{A}(t)$:

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = i\boldsymbol{\Omega} \cdot \boldsymbol{A}(t) - \int_0^t \mathrm{d}\tau e^{i\mathcal{L}(t-\tau)}(\boldsymbol{A}, \mathcal{P}\mathcal{L}\boldsymbol{f}(t))(\boldsymbol{A}, \boldsymbol{A})^{-1}\boldsymbol{A} + \boldsymbol{f}(t), \tag{176}$$

which can be rewritten with Equation 175 to

$$\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = i\boldsymbol{\Omega}\cdot\boldsymbol{A}(t) - \int_0^t \mathrm{d}\tau M(t)\cdot\boldsymbol{A}(t-\tau) + \boldsymbol{f}(t), \tag{177}$$

where the memory function $M(t)$ is defined as

$$M(t) \equiv (\boldsymbol{f},\boldsymbol{f})\cdot(\boldsymbol{A},\boldsymbol{A})^{-1}. \tag{178}$$

The correlation matrix of $\boldsymbol{A}(t)$ in Fourier space is given by (see Subsection 2.2.1)

$$C(t) \equiv \langle \boldsymbol{A}^*(\tau)\boldsymbol{A}(t+\tau)\rangle = (\boldsymbol{A}(\tau),\boldsymbol{A}(t+\tau)), \tag{179}$$

which we will write as $C(t) \equiv (\boldsymbol{A},\boldsymbol{A}(t))$ according to our notation definition explained in Appendix Subsection C.1 and the assumption of time invariance. By multiplying Equation 177 from the left side by $\boldsymbol{A}^*$ and taking the ensemble averages gives for the equation of motion of the correlation matrix $\boldsymbol{C}(t)$:

$$\frac{\mathrm{d}\boldsymbol{C}(t)}{\mathrm{d}t} = i\boldsymbol{\Omega}\cdot\boldsymbol{C}(t) + \int_0^t \mathrm{d}\tau \boldsymbol{M}(t)\cdot\boldsymbol{C}(t-\tau), \tag{180}$$

because $\left\langle \boldsymbol{A}^*\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}\right\rangle = \left\langle \frac{\mathrm{d}(\boldsymbol{A}^*\boldsymbol{A}(t))}{\mathrm{d}t}\right\rangle$, $\langle \boldsymbol{A}^*\boldsymbol{A}(t-\tau)\rangle = \boldsymbol{C}(t-\tau)$ and $\langle \boldsymbol{A}^*\boldsymbol{f}(t)\rangle = 0$.

C.1.2 Specific implementation of $\boldsymbol{A}(t)$ to derive the equation of motion for the coherent ISF

Because our goal is to obtain an equation of motion for the intermediate scattering function, which depends on density fluctuations, we choose $\boldsymbol{A}(t)$ to be these density fluctuations and their derivative, called the longitudinal current $j^L_{\boldsymbol{k}}$:

$$\boldsymbol{A} = \begin{bmatrix}\delta\rho_{\boldsymbol{k}} & j^L_{\boldsymbol{k}}\end{bmatrix}, \tag{181}$$

with (see Appendix Subsection B.2)

$$\delta\rho_{\boldsymbol{k}} = \sum_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} - (2\pi)^2\rho\delta(\boldsymbol{k}) \tag{182}$$

$$j^L_{\boldsymbol{k}} = \frac{1}{m}\sum_i (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}. \tag{183}$$

The correlation function $\boldsymbol{C}$ for this implementation of $\boldsymbol{A}(t)$ is

$$\boldsymbol{C}(t) = \langle \boldsymbol{A}^*\boldsymbol{A}(t)\rangle = \begin{bmatrix}\langle\delta\rho_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\rangle & \langle\delta\rho_{-\boldsymbol{k}}j^L_{\boldsymbol{k}}(t)\rangle \\ \langle j^L_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\rangle & \langle j^L_{-\boldsymbol{k}}j^L_{\boldsymbol{k}}(t)\rangle\end{bmatrix}, \tag{184}$$

which is only valid for $\boldsymbol{k}\neq 0$ because we have neglected the δ-term in Equation 182.

Before rewriting the equation of motion in Equation 180 for this specific implementation of $\boldsymbol{A}(t)$, we start with deriving some equations and identities for the correlation functions of the density and the longitudinal current.

Using Equation 22 we obtain

$$\langle\delta\rho_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\rangle = NF(\boldsymbol{k},t). \tag{185}$$

The right top and left bottom entry of Equation 184 are equal, which is derived as follows

$$\left\langle j^L_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\right\rangle = \left\langle j^L_{-\boldsymbol{k}}(t_1)\delta\rho_{\boldsymbol{k}}(t_2)\right\rangle = \frac{-1}{ik}\frac{\mathrm{d}}{\mathrm{d}t_1}\left\langle\delta\rho_{-\boldsymbol{k}}(t_1)\delta\rho_{\boldsymbol{k}}(t_2)\right\rangle \tag{186}$$

$$= \frac{-1}{ik}\frac{\mathrm{d}}{\mathrm{d}t_1}\left\langle\delta\rho_{-\boldsymbol{k}}(-t_2)\delta\rho_{\boldsymbol{k}}(-t_1)\right\rangle = \left\langle\delta\rho_{-\boldsymbol{k}}(-t_2)j^L_{\boldsymbol{k}}(-t_1)\right\rangle \tag{187}$$

$$= \left\langle\delta\rho_{-\boldsymbol{k}}j^L_{\boldsymbol{k}}(t)\right\rangle, \tag{188}$$

and given by (using Equation 22)

$$\left\langle j^L_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\right\rangle = \left\langle \delta\rho_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t)\right\rangle = \frac{1}{ik}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \delta\rho_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\right\rangle = \frac{N}{ik}\frac{\mathrm{d}}{\mathrm{d}t}F(\boldsymbol{k},t). \tag{189}$$

Furthermore, at time $t = 0$ the right bottom entry of Equation 184 is

$$\left\langle j^L_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}\right\rangle = \left\langle \frac{1}{m^2}\sum_i (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)^2 + \frac{1}{m^2}\sum_{j,i\neq j}(\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)(\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_j)e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_j-\boldsymbol{r}_i)}\right\rangle \tag{190}$$

$$= \left\langle N(v_i^{\boldsymbol{k}})^2\right\rangle = \frac{Nk_BT}{m}, \tag{191}$$

where $v_i^{\boldsymbol{k}}$ is the velocity of particle i in the direction of $\boldsymbol{k}$. To derive this equation we have used that the ensemble average is zero when $i \neq j$ and that $\left\langle \frac{1}{2}m(v_i^{\boldsymbol{k}})^2\right\rangle = \frac{1}{2}k_BT$ according to the equipartition theorem.

The ensemble average of a variable with its derivative is always zero because

$$\left\langle \boldsymbol{A}^*\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \boldsymbol{A}^*\boldsymbol{A}\right\rangle - \left\langle \frac{\mathrm{d}\boldsymbol{A}^*}{\mathrm{d}t}\boldsymbol{A}^*\right\rangle = 0, \tag{192}$$

because the ensemble average is constant in time and $\langle A^*B\rangle = \langle B^*A\rangle$ for arbitrary observables A and B. The time derivative of the right top and left bottom entry of Equation 184 are equal according to Equation 188 and given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle j^L_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \delta\rho_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t)\right\rangle = \left\langle \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{1}{ik}\delta\rho_{-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}(t)\right)\right\rangle = \frac{N}{ik}\frac{\mathrm{d}^2F(\boldsymbol{k},t)}{\mathrm{d}t^2}. \tag{193}$$

Furthermore, the following relationship holds (using Equation 192)

$$\left\langle j^L_{-\boldsymbol{k}}\frac{\mathrm{d}\delta\rho_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{ik}\left\langle \delta\rho_{-\boldsymbol{k}}\frac{\mathrm{d}\delta\rho_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle\right) + \left\langle \delta\rho_{-\boldsymbol{k}}\frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle = \left\langle \delta\rho_{-\boldsymbol{k}}\frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle, \tag{194}$$

$$\tag{195}$$

and both terms can be written as

$$\left\langle j^L_{-\boldsymbol{k}}\frac{\mathrm{d}\delta\rho_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle = \left\langle \delta\rho_{-\boldsymbol{k}}\frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t}\right\rangle \tag{196}$$

$$= \left\langle \frac{1}{m}\sum_i(\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}\sum_j i(\boldsymbol{k}\cdot\frac{\mathrm{d}\boldsymbol{r}_j}{\mathrm{d}t})e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j}\right\rangle \tag{197}$$

$$= \frac{i}{m^2}\sum_{i,j}\left\langle (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}(\boldsymbol{k}\cdot\boldsymbol{p}_j)e^{i\boldsymbol{k}\cdot\boldsymbol{r}_j}\right\rangle \tag{198}$$

$$= \frac{i}{m^2}\sum_i\left\langle (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)(\boldsymbol{k}\cdot\boldsymbol{p}_i)\right\rangle \tag{199}$$

$$= \frac{ik}{m}\sum_i\left\langle m(v_i^{\boldsymbol{k}})^2\right\rangle = \frac{iNkk_BT}{m}. \tag{200}$$

To obtain the equation of motion for the intermediate scattering function, we solve the left bottom entry of the matrix Equation 180 using the relations given by Equations 185 till 200. First we derive the expression of the left bottom entry of the three parts separately, defined by

$$\underbrace{\frac{\mathrm{d}\boldsymbol{C}(t)}{\mathrm{d}t}}_{\text{Part 1}} = \underbrace{i\boldsymbol{\Omega}\cdot\boldsymbol{C}(t)}_{\text{Part 2}} + \underbrace{\int_0^t \mathrm{d}\tau\boldsymbol{M}(t)\cdot\boldsymbol{C}(t-\tau))}_{\text{Part 3}}. \tag{201}$$

The bottom left corner of Part 1 is given by $\frac{N}{ik}\frac{\mathrm{d}^2F(\boldsymbol{k},t)}{\mathrm{d}t^2}$ according to Equation 193. For the bottom left corner of Part 2, we first write down the explicit form of the frequency matrix $\boldsymbol{\Omega}$:

$$
\begin{aligned}
i\boldsymbol{\Omega} &= \left\langle \boldsymbol{A}^* \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} \right\rangle \cdot \langle \boldsymbol{A}^* \boldsymbol{A} \rangle^{-1} && (202)\\
&= \begin{bmatrix} \left\langle \delta\rho_{-\boldsymbol{k}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{k}}}{\mathrm{d}t} \right\rangle & \left\langle \delta\rho_{-\boldsymbol{k}} \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} \right\rangle \\ \left\langle j^L_{-\boldsymbol{k}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{k}}}{\mathrm{d}t} \right\rangle & \left\langle j^L_{-\boldsymbol{k}} \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{d\mathrm{d}} \right\rangle \end{bmatrix} \cdot \langle \boldsymbol{A}^* \boldsymbol{A} \rangle^{-1} && (203)\\
&= \begin{bmatrix} 0 & \frac{iNkk_BT}{m} \\ \frac{iNkk_BT}{m} & 0 \end{bmatrix} \cdot \boldsymbol{C}^{-1} && (204)\\
&= \begin{bmatrix} 0 & \frac{iNkk_BT}{m} \\ \frac{iNkk_BT}{m} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{NS(\boldsymbol{k})} & 0 \\ 0 & \frac{m}{Nk_BT} \end{bmatrix} && (205)\\
&= \begin{bmatrix} 0 & ik \\ \frac{ikkBT}{mS(\boldsymbol{k})} & 0 \end{bmatrix}, && (206)
\end{aligned}
$$

where Equations 192, 194 and 200 are used for Equation 204. Furthermore, in Equation 205 the correlation function at time $t = 0$ is used given by (using Equations 185, 191 and 192)

$$
C(0) = \begin{bmatrix} NS(\boldsymbol{k}) & 0 \\ 0 & \frac{Nk_BT}{m} \end{bmatrix}, \tag{207}
$$

because $S(\boldsymbol{k}) = F(\boldsymbol{k}, 0)$ and Equations 185 and 191 are used. This results in:

$$
i\boldsymbol{\Omega} \cdot \boldsymbol{C}(t) = \begin{bmatrix} 0 & ik \\ \frac{ikk_BT}{mS(\boldsymbol{k})} & 0 \end{bmatrix} \cdot \begin{bmatrix} \langle \delta\rho_{-\boldsymbol{k}} \delta\rho_{\boldsymbol{k}}(t) \rangle & \left\langle \delta\rho_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t) \right\rangle \\ \left\langle j^L_{-\boldsymbol{k}} \delta\rho_{\boldsymbol{k}}(t) \right\rangle & \left\langle j^L_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t) \right\rangle \end{bmatrix}, \tag{208}
$$

which gives for the left bottom term of Part 2: $\frac{ikk_BTN}{mS(\boldsymbol{k})} F(\boldsymbol{k}, t)$.

For Part 3 of Equation 201 we need to rewrite the memory kernel, given by $(\boldsymbol{f}, \boldsymbol{f}(t)) \cdot (\boldsymbol{A}, \boldsymbol{A})^{-1}$. First, the fluctuation force is given by (see Equation 174)

$$
\begin{aligned}
\boldsymbol{f} &= (1 - \mathcal{P}) \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} && (209)\\
&= \frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} - i\Omega\boldsymbol{A} && (210)\\
&= \begin{bmatrix} \delta \frac{\mathrm{d}\rho_{\boldsymbol{k}}}{\mathrm{d}t} \\ \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} \end{bmatrix} - \begin{bmatrix} 0 & ik \\ \frac{ikk_BT}{mS(\boldsymbol{k})} & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta\rho_{\boldsymbol{k}} \\ j^L_{\boldsymbol{k}} \end{bmatrix} && (211)\\
&= \begin{bmatrix} 0 \\ \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} - \frac{ikk_BT}{mS(\boldsymbol{k})} \delta\rho_{\boldsymbol{k}} \end{bmatrix} \equiv \begin{bmatrix} 0 \\ R_{\boldsymbol{k}} \end{bmatrix}, && (212)
\end{aligned}
$$

where for Equation 210 is used that: $\mathcal{P}\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} = i\Omega\boldsymbol{A}$. This can be verified by multiplying Equation 167 from the left by $\mathcal{P}$ to get

$$
\begin{aligned}
\mathcal{P}\frac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t} &= i\mathcal{P}\boldsymbol{\Omega} \cdot \boldsymbol{A}(t) + \mathcal{P}e^{i\mathcal{L}t}(1 - \mathcal{P})i\mathcal{L}\boldsymbol{A} && (213)\\
&= i\Omega A(t), && (214)
\end{aligned}
$$

because $\boldsymbol{\Omega}$ is already along the direction of $\boldsymbol{A}$. According to Equation 174 the time dependence of the fluctuation force is given by

$$
R_{\boldsymbol{k}}(t) = e^{i(1-\mathcal{P}\mathcal{L}t} R_{\boldsymbol{k}}, \tag{215}
$$

and the conjugate $R^*_{\boldsymbol{k}}(t)$ is $R_{-\boldsymbol{k}}(t)$.

The memory matrix is given by

$$\boldsymbol{M}(t) = \left\langle \begin{bmatrix} 0 \\ R_{-\boldsymbol{k}} \end{bmatrix} \cdot \begin{bmatrix} 0 & R_{\boldsymbol{k}}(t) \end{bmatrix} \right\rangle \cdot \langle \boldsymbol{A}^* \boldsymbol{A} \rangle^{-1} \tag{216}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & \langle R_{-\boldsymbol{k}} R_{\boldsymbol{k}}(t) \rangle \end{bmatrix} \cdot \begin{bmatrix} \frac{k}{NS(\boldsymbol{k})} & 0 \\ 0 & \frac{m}{Nk_BT} \end{bmatrix} \tag{217}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & \frac{m\langle R_{0\boldsymbol{k}} R_{\boldsymbol{k}}(t) \rangle}{Nk_BT} \end{bmatrix} \cdot \tag{218}$$

This results in the following expression for Part 3 in the equation of motion given by Equation 201:

$$\int_0^t \mathrm{d}\tau \boldsymbol{M}(t) \cdot \boldsymbol{C}(t-\tau) = \int_0^t \mathrm{d}\tau \begin{bmatrix} 0 & 0 \\ 0 & \frac{m\langle R^*_{\boldsymbol{k}} R_{\boldsymbol{k}}(t) \rangle}{Nk_BT} \end{bmatrix} \cdot \begin{bmatrix} \langle \delta\rho_{-\boldsymbol{k}} \delta\rho_{\boldsymbol{k}}(t-\tau) \rangle & \left\langle \delta\rho_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t-\tau) \right\rangle \\ \left\langle j^L_{-\boldsymbol{k}} \delta\rho_{\boldsymbol{k}}(t-\tau) \right\rangle & \left\langle j^L_{-\boldsymbol{k}} j^L_{\boldsymbol{k}}(t-\tau) \right\rangle \end{bmatrix}, \tag{219}$$

of which the left bottom entry is given by

$$\frac{m}{Nk_BT} \int_0^t \mathrm{d}\tau \left\langle R_{-\boldsymbol{k}} R_{\boldsymbol{k}}(\tau) \right\rangle \left\langle j^L_{-\boldsymbol{k}} \delta\rho_{\boldsymbol{k}}(t-\tau) \right\rangle = \frac{m}{ikk_BT} \int_0^t \mathrm{d}\tau \left\langle R_{-\boldsymbol{k}} R_{\boldsymbol{k}}(\tau) \right\rangle \frac{\mathrm{d}}{\mathrm{d}t} F(\boldsymbol{k}, t-\tau). \tag{220}$$

Combining the expressions for Part 1, 2 and 3 gives for the left bottom entry of Equation 201

$$\frac{\mathrm{d}^2 F(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{k^2 k_B T}{mS(\boldsymbol{k})} F(\boldsymbol{k},t) + \int_0^t \mathrm{d}\tau M(\boldsymbol{k},\tau) \frac{\mathrm{d}}{\mathrm{d}t} F(\boldsymbol{k}, t-\tau), \tag{221}$$

with the memory kernel $M(\boldsymbol{k},t)$ given by

$$M(\boldsymbol{k},t) = \frac{m}{Nk_BT} \left\langle R_{-\boldsymbol{k}} R_{\boldsymbol{k}}(t) \right\rangle. \tag{222}$$

This is the exact equation of motion for the coherent intermediate scattering function $F(\boldsymbol{k},t)$. Unfortunately, $\langle R_{-\boldsymbol{k}} R_{\boldsymbol{k}}(t) \rangle$ cannot be solved exactly and therefore needs to be approximated.

C.1.3 Approximation of the memory kernel

We apply four approximations to rewrite the memory kernel of Equation 222. We start with splitting $M(\boldsymbol{k},t)$ in a part containing fast and slow modes, which we call respectively $M^0(\boldsymbol{k},t)$ and $M^{\mathrm{MCT}}(\boldsymbol{k},t)$:

$$M(\boldsymbol{k},t) = M^0(\boldsymbol{k},t) + M^{\mathrm{MCT}}(\boldsymbol{k},t). \tag{223}$$

The *first* approximation is to approximate the fast part $M^0(\boldsymbol{k},t)$ with a δ-function around the origin times a constant α,

$$M^0(\boldsymbol{k},t) \approx \alpha\delta(t), \tag{224}$$

because this only reflects the behavior of the intermediate scattering function at short times and we are interested in longer times, which makes the exact form at short times irrelevant. Now we focus on $M^{\mathrm{MCT}}(\boldsymbol{k},t)$ and we can derive that hidden in the random force $R_{\boldsymbol{k}}$ is a pair of densities. First, rewriting the random force $R_{\boldsymbol{k}}$ gives

$$R_{\boldsymbol{k}} = \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} - ik\frac{k_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}} \tag{225}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{m}\sum_i (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i) e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\right) - ik\frac{k_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}} \tag{226}$$

$$= \frac{1}{m}\sum_i \left(\hat{\boldsymbol{k}}\cdot\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t}\right) e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} + \frac{i}{m^2}\sum (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i)^2 e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} - ik\frac{k_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}}, \tag{227}$$

where $\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t}$ is a force which we can approximate using potential ϕ as follows

$$\sum_i \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} \sim -\sum_i \sum_{j\neq i} \nabla\phi(|\boldsymbol{r}_i - \boldsymbol{r}_j|) = \sum_{\boldsymbol{q}} i\boldsymbol{q}\phi_{\boldsymbol{q}}\delta\rho_{\boldsymbol{q}}\delta\rho_{-\boldsymbol{q}}, \tag{228}$$

where in the last equation is used that $\phi(\boldsymbol{r}) = \sum_{\boldsymbol{k}} \phi_{\boldsymbol{k}} e^{i\boldsymbol{r}\cdot\boldsymbol{k}}$. This dependency of the pair of densities in the fluctuation force is remarkable as we first have removed the slow modes $\delta\rho_{\boldsymbol{k}}$ from it and now see that this force still depends on it. Therefore, the *second* approximation is to project the random force onto its slow modes using the projection operator $\mathcal{P}_2$

$$\mathcal{P}_2 \equiv \sum_{\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3,\boldsymbol{q}_4} A_{\boldsymbol{q}_1,\boldsymbol{q}_2} \left\langle A^*_{\boldsymbol{q}_3,\boldsymbol{q}_4} ... \right\rangle \left\langle A^*_{\boldsymbol{q}_1,\boldsymbol{q}_2} A_{\boldsymbol{q}_3,\boldsymbol{q}_4} \right\rangle^{-1} \quad \text{with} \tag{229}$$

$$A_{\boldsymbol{q}_1,\boldsymbol{q}_2} = \delta\rho_{\boldsymbol{q}_1}\delta\rho_{\boldsymbol{q}_2}, \tag{230}$$

with $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$ and $\boldsymbol{q}_4$ wave vectors. The the projection $\mathcal{P}_2$ of $R_{\boldsymbol{k}}$ is

$$\mathcal{P}_2 R_{\boldsymbol{k}} = \sum_{\boldsymbol{q}_1,\boldsymbol{q}_2} V_{\boldsymbol{k}}(\boldsymbol{q}_1,\boldsymbol{q}_2)\delta\rho_{\boldsymbol{q}_1}\delta\rho_{\boldsymbol{q}_2}, \tag{231}$$

with

$$V_{\boldsymbol{k}}(\boldsymbol{q}_1,\boldsymbol{q}_2) = \sum_{\boldsymbol{q}_3,\boldsymbol{q}_4} \langle \delta\rho_{-\boldsymbol{q}_3}\delta\rho_{-\boldsymbol{q}_4} R_{\boldsymbol{k}} \rangle \cdot \langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{-\boldsymbol{q}_2}\delta\rho_{\boldsymbol{q}_3}\delta\rho_{\boldsymbol{q}_4} \rangle^{-1}. \tag{232}$$

The numerator $\langle \delta\rho_{-\boldsymbol{q}_3}\delta\rho_{-\boldsymbol{q}_4} R_{\boldsymbol{k}} \rangle$ is given by

$$\langle \delta\rho_{-\boldsymbol{q}_3}\delta\rho_{-\boldsymbol{q}_4} R_{\boldsymbol{k}} \rangle = \left\langle \delta\rho_{-\boldsymbol{q}_3}\delta\rho_{-\boldsymbol{q}_4} (\frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} - i\frac{kk_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}}) \right\rangle \tag{233}$$

$$= \left\langle \delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} (\frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} - i\frac{kk_BT}{mS(\boldsymbol{k})}\delta\rho_{\boldsymbol{k}}) \right\rangle \tag{234}$$

$$= -\left\langle \frac{\mathrm{d}\delta\rho_{-\boldsymbol{q}}}{\mathrm{d}t}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} j^L_{\boldsymbol{k}} \right\rangle - \left\langle \delta\rho_{-\boldsymbol{q}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}}{\mathrm{d}t} j^L_{\boldsymbol{k}} \right\rangle - \frac{ikk_BT}{mS(\boldsymbol{k})} \langle \delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} \rangle, \tag{235}$$

where in Equation 234 is used that the ensemble average is only nonzero when $-\boldsymbol{q}_3 - \boldsymbol{q}_4 + \boldsymbol{k} = 0$. Furthermore, in Equation 235 we have used that

$$\left\langle \delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} \frac{\mathrm{d}j^L_{\boldsymbol{k}}}{\mathrm{d}t} \right\rangle = \frac{\mathrm{d}}{\mathrm{d}t} \langle \delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} j^L_{\boldsymbol{k}} \rangle - \left\langle \frac{\mathrm{d}\delta\rho_{-\boldsymbol{q}}}{\mathrm{d}t}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} j^L_{\boldsymbol{k}} \right\rangle - \left\langle \delta\rho_{-\boldsymbol{q}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}}{\mathrm{d}t} j^L_{\boldsymbol{k}} \right\rangle \tag{236}$$

$$= -\left\langle \frac{\mathrm{d}\delta\rho_{-\boldsymbol{q}}}{\mathrm{d}t}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} j^L_{\boldsymbol{k}} \right\rangle - \left\langle \delta\rho_{-\boldsymbol{q}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}}{\mathrm{d}t} j^L_{\boldsymbol{k}} \right\rangle. \tag{237}$$

We can write

$$\left\langle \frac{\mathrm{d}\delta\rho_{-\boldsymbol{q}}}{\mathrm{d}t}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}} j^L_{\boldsymbol{k}} \right\rangle = -i \left\langle \sum_i (\boldsymbol{q}\cdot\boldsymbol{p}_i) e^{-i\boldsymbol{q}\cdot\boldsymbol{r}_i} \sum_j e^{i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{r}_j} \sum_l (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_l) e^{i\boldsymbol{k}\cdot\boldsymbol{r}_l} \right\rangle \tag{238}$$

$$= -i \left\langle \sum_{i,j} (\boldsymbol{q}\cdot\boldsymbol{p}_i) e^{i(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{r}_i} e^{i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{r}_j} (\hat{\boldsymbol{k}}\cdot\boldsymbol{p}_i) \right\rangle \tag{239}$$

$$= -i \left\langle \sum_{i,j} e^{i(\boldsymbol{k}-\boldsymbol{q})\cdot\boldsymbol{r}_i} e^{i(\boldsymbol{q}-\boldsymbol{k})\cdot\boldsymbol{r}_j} \right\rangle \frac{k_BT}{m} (\boldsymbol{k}\cdot\hat{\boldsymbol{q}}) \tag{240}$$

$$= -i(\boldsymbol{q}\cdot\hat{\boldsymbol{k}}) \frac{k_BT}{m} NS(\boldsymbol{q}-\boldsymbol{k}), \tag{241}$$

where we have used that the ensemble average is zero when $i \neq l$ in Equation 239, the equipartition theorem in Equation 240 and the definition of the density correlation given by Equation 15. In analogy with Equation 241 we have

$$\left\langle \delta\rho_{-\boldsymbol{k}} \frac{\mathrm{d}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}}{\mathrm{d}t} j_{\boldsymbol{q}}^{L} \right\rangle = -i((\boldsymbol{q}-\boldsymbol{k})\cdot\hat{\boldsymbol{k}})\frac{k_B T}{m} N S(\boldsymbol{q}). \tag{242}$$

The *third* approximation is rewriting the third term in Equation 235 according to the convolution approximation [55] as

$$-\frac{ikk_BT}{mS(\boldsymbol{k})}\left\langle \delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}}\right\rangle \approx -\frac{ikk_BT}{mS(\boldsymbol{k})} N S(\boldsymbol{q})S(\boldsymbol{k})S(\boldsymbol{q}-\boldsymbol{k}). \tag{243}$$

The *fourth* approximation is used for the denominator of Equation 232, where we factorize a four-point density term into products of two-point density terms. This gives,

$$\sum_{\boldsymbol{q}_3,\boldsymbol{q}_4} \left\langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{-\boldsymbol{q}_2}\delta\rho_{\boldsymbol{q}_3}\delta\rho_{\boldsymbol{q}_4}\right\rangle \tag{244}$$

$$\approx \sum_{\boldsymbol{q}_3,\boldsymbol{q}_4} \left(\left\langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{-\boldsymbol{q}_2}\right\rangle \left\langle \delta\rho_{\boldsymbol{q}_3}\delta\rho_{\boldsymbol{q}_4}\right\rangle + \left\langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{\boldsymbol{q}_3}\right\rangle \left\langle \delta\rho_{-\boldsymbol{q}_2}\delta\rho_{\boldsymbol{q}_4}\right\rangle + \left\langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{\boldsymbol{q}_4}\right\rangle \left\langle \delta\rho_{-\boldsymbol{q}_2}\delta\rho_{\boldsymbol{q}_3}\right\rangle\right) \tag{245}$$

$$= 2\left\langle \delta\rho_{-\boldsymbol{q}_1}\delta\rho_{\boldsymbol{q}_1}\right\rangle \left\langle \delta\rho_{-\boldsymbol{q}_2}\delta\rho_{\boldsymbol{q}_2}\right\rangle \tag{246}$$

$$= 2N^2 S(\boldsymbol{q}_1)S(\boldsymbol{q}_2). \tag{247}$$

Because of translational variance, only the differences of the wave vectors $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ matters and therefore we can define $\boldsymbol{q}_1 \to \boldsymbol{q}$ and $\boldsymbol{q}_2 \to \boldsymbol{k}-\boldsymbol{q}$. This gives for Equation 231

$$\mathcal{P}_2 R_{\boldsymbol{k}} = \sum_{\boldsymbol{q}} V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}. \tag{248}$$

Combining Equations 235, 241, 242, 243 and 247 to fill in Equation 232 gives for the vertices

$$V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}} = V_{\boldsymbol{k}}(\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}) \tag{249}$$

$$= \frac{ik_BT}{2mN}\left\{\frac{\boldsymbol{q}\cdot\hat{\boldsymbol{k}}}{S(\boldsymbol{q})} + \frac{(\boldsymbol{k}-\boldsymbol{q})\cdot\hat{\boldsymbol{k}}}{S(\boldsymbol{q}-\boldsymbol{k})} - \boldsymbol{k}\right\} \tag{250}$$

$$= -\frac{ik_BT\rho}{2mN}\left\{\boldsymbol{q}\cdot\hat{\boldsymbol{k}}c(\boldsymbol{k}) + (\boldsymbol{k}-\boldsymbol{q})\cdot\hat{\boldsymbol{k}}c(\boldsymbol{k}-\boldsymbol{q})\right\}, \tag{251}$$

with $c(\boldsymbol{k}) = \frac{1}{\rho}(1-\frac{1}{S(\boldsymbol{k})})$ the direct correlation function. The memory kernel of Equation 222 can now be written as

$$M(\boldsymbol{k},t) = \frac{m}{Nk_BT}\langle R_{-\boldsymbol{k}}R_{\boldsymbol{k}}(t)\rangle \tag{252}$$

$$\approx A\delta(t) + \frac{m}{Nk_BT}\left\langle \mathcal{P}_2^* R_{-\boldsymbol{k}} e^{i(1-\mathcal{P})\mathcal{L}t}\mathcal{P}_2 R_{\boldsymbol{k}}\right\rangle \tag{253}$$

$$= A\delta(t) + \frac{m}{Nk_BT}\left\langle \sum_{\boldsymbol{q}'} V^*_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}'}\delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}} e^{i(1-\mathcal{P})\mathcal{L}t}\sum_{\boldsymbol{q}} V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}\delta\rho_{\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}\right\rangle \tag{254}$$

$$= A\delta(t) + \frac{m}{Nk_BT}\sum_{\boldsymbol{q},\boldsymbol{q}'} V^*_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}'}V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}\left\langle \delta\rho_{-\boldsymbol{q}'}\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}} e^{i(1-\mathcal{P})\mathcal{L}t}\delta\rho_{\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}\right\rangle \tag{255}$$

$$= A\delta(t) + \frac{m}{Nk_BT}\sum_{\boldsymbol{q},\boldsymbol{q}'} V^*_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}'}V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}\langle \delta\rho_{-\boldsymbol{q}'}\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}\delta\rho_{\boldsymbol{q}}(t)\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\rangle \tag{256}$$

$$\approx A\delta(t) + \frac{2N^2m}{Nk_BT}\sum_{\boldsymbol{q},\boldsymbol{q}'} V^*_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}'}V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}F(\boldsymbol{q},t)F(\boldsymbol{k}-\boldsymbol{q},t)(\delta_{\boldsymbol{q}',\boldsymbol{q}} + \delta_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}}) \tag{257}$$

$$= A\delta(t) + \frac{2N^2m}{Nk_BT}\sum_{\boldsymbol{q}} |V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2F(\boldsymbol{k},t)F(\boldsymbol{k}-\boldsymbol{q},t) \tag{258}$$

$$= A\delta(t) + \frac{2N^2mV}{4Nk_BT\pi^2}\int \mathrm{d}\boldsymbol{q}|V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2F(\boldsymbol{k},t)F(\boldsymbol{k}-\boldsymbol{q},t) \tag{259}$$

$$= A\delta(t) + \frac{NmV}{2k_BT\pi^2}\int \mathrm{d}\boldsymbol{q}|V_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2F(\boldsymbol{k},t)F(\boldsymbol{k}-\boldsymbol{q},t), \tag{260}$$

where we have neglected the term $1-\mathcal{P}$ in the exponent in Equation 255. Furthermore, for Equation 257 we have again used the Guassian approximation (*fourth* approximation) to rewrite a four-point density correlation function into a product of two two-point density functions (see Equation 247) as follows

$$\sum_{\boldsymbol{q},\boldsymbol{q}'}\langle \delta\rho_{-\boldsymbol{q}'}\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}\delta\rho_{\boldsymbol{q}}(t)\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\rangle \tag{261}$$

$$\approx \sum_{\boldsymbol{q},\boldsymbol{q}'}\left(\langle\delta\rho_{-\boldsymbol{q}'}\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}\rangle\langle\delta\rho_{\boldsymbol{q}}(t)\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\rangle + \langle\delta\rho_{-\boldsymbol{q}'}\delta\rho_{\boldsymbol{q}}(t)\rangle\langle\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\rangle + \langle\delta\rho_{-\boldsymbol{q}}\delta\rho_{\boldsymbol{k}-\boldsymbol{q}}(t)\rangle\langle\delta\rho_{\boldsymbol{q}'-\boldsymbol{k}}\delta\rho_{\boldsymbol{q}}(t)\rangle\right) \tag{262}$$

$$\approx \sum_{\boldsymbol{q},\boldsymbol{q}'} 2N^2F(\boldsymbol{q},t)F(\boldsymbol{k}-\boldsymbol{q},t)(\delta_{\boldsymbol{q}',\boldsymbol{q}} + \delta_{\boldsymbol{q}',\boldsymbol{k}-\boldsymbol{q}}) \tag{263}$$

Furthermore, in Equation 260 we have switch from a discrete sum to a continuous integral via $\sum_{\boldsymbol{q}} \to \frac{V}{(2\pi)^2}\int \mathrm{d}\boldsymbol{q}$.

C.1.4 Final mode coupling theory equation

Using the approximation of the memory kernel of Equation 260, the mode coupling theory equation given by Equation 221 becomes

$$\frac{\mathrm{d}^2F(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{k^2k_BT}{mS(\boldsymbol{k})}F(\boldsymbol{k},t) + \int_0^t \mathrm{d}\tau M(\boldsymbol{k},t-\tau)\frac{\delta F(\boldsymbol{k},\tau)}{\delta\tau}, \tag{264}$$

via the coordinate transformation $t-\tau\to\tau$, with

$$M(\boldsymbol{k},t) = M^0(\boldsymbol{k},t) + \frac{\rho k_BT}{8\pi^2m}\int \mathrm{d}\boldsymbol{q}|\tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}}|^2F(\boldsymbol{k},t)F(\boldsymbol{k}-\boldsymbol{q},t), \tag{265}$$

with

$$\tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} = \boldsymbol{k}\cdot\hat{\boldsymbol{q}}c(\boldsymbol{q}) + (\boldsymbol{k}-\boldsymbol{q})\cdot\hat{\boldsymbol{q}}c(\boldsymbol{k}-\boldsymbol{q}), \tag{266}$$

with

$$c(\boldsymbol{k}) \equiv \frac{1}{\rho}(1 - \frac{1}{S(\boldsymbol{k})}). \tag{267}$$

When the systemic is isotropic, we can rewrite the above equations according to $F(\boldsymbol{k},t) \to F(k,t)$, $M(\boldsymbol{k},t) \to M(k,t)$ and $S(\boldsymbol{k}) \to S(k)$.

C.2 Adaptations to the MCT-derivation for the incoherent ISF

The derivation for the incoherent MCT follows the same steps as the derivation for the coherent MCT described in Appendix Subsection C.2. Below we explain which equations are changing in the derivative. We start with adapting the specific implementation of $\boldsymbol{A}$ of Equation 181, which is now given by

$$\boldsymbol{A} = \begin{bmatrix} \delta\rho_{\boldsymbol{k},s,i} & j^L_{\boldsymbol{k},s,i} \end{bmatrix}, \tag{268}$$

with

$$\delta\rho_{\boldsymbol{k},s,i} = e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \tag{269}$$

$$j^L_{\boldsymbol{k},s,i} = \frac{1}{m}\boldsymbol{r}_i \cdot \boldsymbol{p}_i e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}. \tag{270}$$

Because we are only looking at correlation functions including $\delta\rho_{\boldsymbol{k},s,i}$ and $j^L_{\boldsymbol{k},s,i}$ and these correlation functions are defined as ensemble averages which include taking the sum over all particles i, we omit the subscript i from now on. Following a similar derivation as in Appendix Subsection C.1.2 gives for the equation of motion of the incoherent intermediate scattering function

$$\frac{\mathrm{d}^2F_s(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{k^2k_BT}{mS(\boldsymbol{k})}F_s(\boldsymbol{k},t) + \frac{m}{Nk_BT}\int_0^t \mathrm{d}\tau \left\langle R_{s,-\boldsymbol{k}}R_{s,\boldsymbol{k}}(\tau)\right\rangle \frac{\mathrm{d}}{\mathrm{d}t}F_s(\boldsymbol{k},t-\tau), \tag{271}$$

with the random force $R_{s,\boldsymbol{k}}(t)$ given by

$$R_{s,\boldsymbol{k}}(t) = \frac{\mathrm{d}j^L_{\boldsymbol{k},s}}{\mathrm{d}t} - \frac{ikk_BT}{m}\delta\rho_{\boldsymbol{k},s}. \tag{272}$$

Instead of projecting the fluctuation force on the product of the densities, we project it on the product of the self-density and the collective density. Thus, the projection operator $\mathcal{P}_2$ is now

$$\mathcal{P}_2^s \equiv \sum_{\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3,\boldsymbol{q}_4} A_{\boldsymbol{q}_1,\boldsymbol{q}_2}\left\langle A^*_{\boldsymbol{q}_3,\boldsymbol{q}_4}...\right\rangle\left\langle A^*_{\boldsymbol{q}_1,\boldsymbol{q}_2}A_{\boldsymbol{q}_3,\boldsymbol{q}_4}\right\rangle^{-1} \quad \text{with} \tag{273}$$

$$A_{\boldsymbol{q}_1,\boldsymbol{q}_2} = \delta\rho_{\boldsymbol{q}_1,s}\delta\rho_{\boldsymbol{q}_2}. \tag{274}$$

Applying the same four approximation as explained in Appendix Subsection C.1.3 results in the following MCT-equation for the incoherent intermediate scattering function:

$$\frac{\mathrm{d}^2F_s(\boldsymbol{k},t)}{\mathrm{d}t^2} + \frac{k^2k_BT}{m}F_s(\boldsymbol{k},t) + \int_0^t \mathrm{d}\tau M_s(\boldsymbol{k},t-\tau))\frac{\mathrm{d}F_s(\boldsymbol{k},\tau)}{\mathrm{d}\tau}, \tag{275}$$

with

$$M_s(\boldsymbol{k},t) = M^0(\boldsymbol{k},t) + \frac{\rho k_BT}{8\pi^2m}\int \mathrm{d}\boldsymbol{q}|V^s_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}}|^2F(\boldsymbol{q},t)F_s(\boldsymbol{k}-\boldsymbol{q},t), \tag{276}$$

with

$$V^s_{\boldsymbol{q},\boldsymbol{k}-\boldsymbol{q}} = (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(\boldsymbol{q}), \tag{277}$$

and $c(\boldsymbol{q})$ the direct correlation function.

D Other mathematical derivations and proofs

In this appendix section we give the mathematical proofs and derivations of non-trivial mathematical equations that we have used in this thesis. We start with a proof of the correlation theorem in Appendix Subsection D.1 and a derivation of the time-diffusion scaling in Appendix Subsection D.2. In Appendix Subsection D.3 we derive some properties of the honeycomb tiling and in Appendix Subsection D.4 we explain how the Voronoi interaction potential can be simplified. Finally, we derive the equation for the self-diffusion coefficient in the VRD- and VV-model in Appendix Subsection D.5.

D.1 Proof of correlation theorem

The correlation theorem is used in Equation 92 in Appendix B.1 to derive a general correlation function in Fourier space and in this Appendix subsection we derive this theorem. Hence, we want to prove that the Fourier of the correlation function can be written as

$$\int \mathrm{d}(\Delta \boldsymbol{r}) C(\Delta \boldsymbol{r}, \tau) e^{i\boldsymbol{k}\cdot\Delta\boldsymbol{r}} = \frac{1}{V} \left\langle \delta A^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) \delta A_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \right\rangle , \tag{278}$$

given the following definitions and identities:

$$C(\Delta \boldsymbol{r}, \tau) = \langle \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \delta A(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \rangle \tag{279}$$

$$\hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \int \mathrm{d}\boldsymbol{r} A(\boldsymbol{r}, \boldsymbol{R}(t)) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{280}$$

$$A(\boldsymbol{r}, \boldsymbol{R}(t)) = \frac{1}{(2\pi)^2} \int \mathrm{d}\boldsymbol{r} \hat{A}_{\boldsymbol{k}}(\boldsymbol{R}(t)) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{281}$$

$$\hat{A}^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) = \hat{A}_{-\boldsymbol{k}}(\boldsymbol{R}(t)). \tag{282}$$

So,

$$\int \mathrm{d}(\Delta \boldsymbol{r}) C(\Delta \boldsymbol{r}, t) e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})} \tag{283}$$

$$= \int \mathrm{d}(\Delta \boldsymbol{r}) \left\langle \delta A(\boldsymbol{r}, \boldsymbol{R}(t)) \delta A(\boldsymbol{r} + \Delta \boldsymbol{r}, \boldsymbol{R}(t+\tau)) \right\rangle e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})} \tag{284}$$

$$= \int \mathrm{d}(\Delta \boldsymbol{r}) \left\langle \frac{1}{(2\pi)^4} \int \mathrm{d}\boldsymbol{k}_1 A_{\boldsymbol{k}_1}(\boldsymbol{R}(t)) e^{-i\boldsymbol{k}_1\cdot\boldsymbol{r}} \int \mathrm{d}\boldsymbol{k}_2 A_{\boldsymbol{k}_2}(\boldsymbol{R}(t+\tau)) e^{-i\boldsymbol{k}_2\cdot(\boldsymbol{r}+\Delta\boldsymbol{r})} \right\rangle e^{i\boldsymbol{k}\cdot(\Delta\boldsymbol{r})} \tag{285}$$

$$\tag{286}$$

$$= \frac{1}{(2\pi)^4} \frac{1}{V} \int \mathrm{d}\boldsymbol{r}_1 \int \mathrm{d}\boldsymbol{r}_2 \left\langle \int \mathrm{d}\boldsymbol{k}_1 \int \mathrm{d}\boldsymbol{k}_2 A_{\boldsymbol{k}_1}(\boldsymbol{R}(t)) A_{\boldsymbol{k}_2}(\boldsymbol{R}(t+\tau)) e^{-i\boldsymbol{k}_1\cdot\boldsymbol{r}_1} e^{-i\boldsymbol{k}_2\cdot\boldsymbol{r}_2} \right\rangle e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)} \tag{287}$$

$$= \frac{1}{(2\pi)^4} \frac{1}{V} \left\langle \int \mathrm{d}\boldsymbol{k}_1 \int \mathrm{d}\boldsymbol{k}_2 A_{\boldsymbol{k}_1}(\boldsymbol{R}(t)) A_{\boldsymbol{k}_2}(\boldsymbol{R}(t+\tau)) \int \mathrm{d}\boldsymbol{r}_1 \int \mathrm{d}\boldsymbol{r}_2 e^{-i(\boldsymbol{k}_1-\boldsymbol{k})\cdot\boldsymbol{r}_1} e^{-i(\boldsymbol{k}_2+\boldsymbol{k})\cdot\boldsymbol{r}_2} \right\rangle \tag{288}$$

$$= \frac{1}{V} \left\langle \int \mathrm{d}\boldsymbol{k}_1 \int \mathrm{d}\boldsymbol{k}_2 A_{\boldsymbol{k}_1}(\boldsymbol{R}(t)) A_{\boldsymbol{k}_2}(\boldsymbol{R}(t+\tau)) \delta(\boldsymbol{k}-\boldsymbol{k}_1) \delta(\boldsymbol{k}+\boldsymbol{k}_2) \right\rangle \tag{289}$$

$$= \frac{1}{V} \left\langle \delta A_{-\boldsymbol{k}}(\boldsymbol{R}(t)) \delta A_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \right\rangle \tag{290}$$

$$= \frac{1}{V} \left\langle \delta A^*_{\boldsymbol{k}}(\boldsymbol{R}(t)) \delta A_{\boldsymbol{k}}(\boldsymbol{R}(t+\tau)) \right\rangle , \tag{291}$$

where we have used the coordinate transformation $\boldsymbol{r}_1 = \boldsymbol{r} + \Delta\boldsymbol{r}$ and $\boldsymbol{r}_2 = \boldsymbol{r}$, which changes the integral $\int \mathrm{d}(\Delta\boldsymbol{r})$ to $\frac{1}{V} \int \mathrm{d}\boldsymbol{r}_1 \int \mathrm{d}\boldsymbol{r}_2$.

D.2 Time-diffusion scaling in the MCT-equation

In this appendix subsection we show that a change in the self-diffusion coefficient D_0 only results in a rescaling of the time in mode coupling theory. The mode coupling theory equation is given by (see Equation 41)

$$\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} + \Omega^2 F(k,t) + \int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau}, \tag{292}$$

where $\Omega^2 = k^2 D_0/S(k)$ and $M(k,t)$ is linear in D_0. By multiplying the equation with a scaling constant α we obtain

$$\alpha\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} + \alpha\Omega^2 F(k,t) + \alpha\int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau}. \tag{293}$$

We make the following substitution: $t \to \alpha t'$, which results in

$$\tau = \alpha\tau' \tag{294}$$

$$\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} = \frac{1}{\alpha}\frac{\mathrm{d}F(k,\alpha t')}{\mathrm{d}t'} \tag{295}$$

$$\int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} = \alpha\int_0^{\alpha t'} \mathrm{d}\alpha\tau' M(k,\alpha t'-\alpha\tau')\frac{1}{\alpha}\frac{\mathrm{d}F(k,\alpha\tau')}{\mathrm{d}\tau'} \tag{296}$$

$$= \int_0^{\alpha t'} \mathrm{d}\tau' M(k,\alpha t'-\alpha\tau')\frac{\mathrm{d}F(k,\alpha\tau')}{\mathrm{d}\tau'}. \tag{297}$$

Filling this in in Equation 293 results in

$$\frac{\mathrm{d}F(k,\alpha t')}{\mathrm{d}t'} + \alpha\Omega^2 F(k,\alpha\tau') + \alpha\int_0^{\alpha t'} \mathrm{d}\tau' M(k,\alpha t'-\alpha\tau')\frac{\mathrm{d}F(k,\alpha\tau')}{\mathrm{d}\tau'}. \tag{298}$$

Keeping in mind that both Ω^2 and $M(k,t)$ are linear in D_0, this shows that $D_0 \to \alpha D_0$ leads to $F(k,t) \to F(k,\alpha t)$ resulting in a time rescaling of $t \to \alpha t$. Hence, the rescaled relaxation time for αD_0 is given by $\frac{\tau_\alpha}{\alpha}$.

D.3 Honeycomb tiling properties

A honeycomb or hexagonal tiling minimizes the ratio of perimeter to the squared area [103]. In this appendix subsection we show that this ratio is 3.7224 and that the ratio of the interparticle distance to the squared area is 1.0745 in a hexagonal tiling.

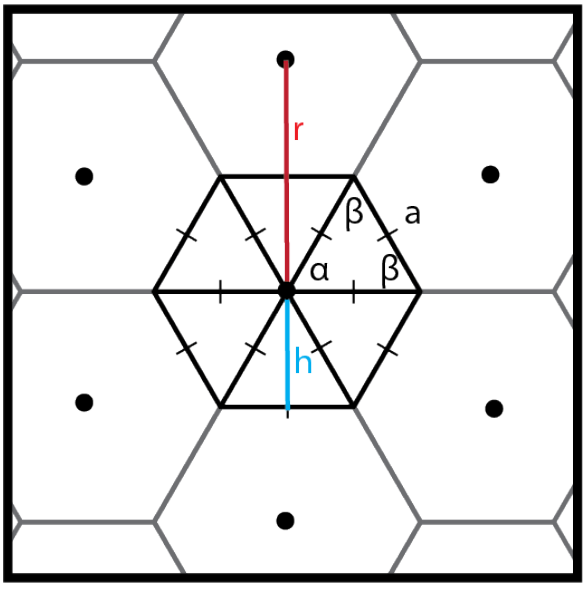

Figure 83: Calculation of the perimeter to squared area ratio and the interparticle distance to squared area ratio in a hexagonal grid.

The central hexagon in Figure 83 can be divided in six triangles with identical shape and size due to symmetry reasons. The angle near the center is called α and obeys the formula $6\alpha = 2\pi \rightarrow \alpha = \frac{1}{3}\pi$. Looking at one triangle, we see that both angles near the border are equal due to symmetry, which results in $\alpha + 2\beta = \pi \rightarrow \beta = \frac{1}{3}\pi$. Thus, each triangle is equilateral and we define the length of the edges with a. The height of the triangle h (blue line in Figure 83) is given by $h = \sqrt{a^2 - \frac{1}{4}a^2} = \frac{\sqrt{3}}{2}a$ which results in an area of $A_{\text{tri}} = \frac{1}{2}ah = \sqrt{3}4a^2$. The hexagon consists of 6 triangles and its area is $A = 6A_{\text{tri}}$. The perimeter of the pentagons is $6a$ and this results in the following ratio

$$\frac{P}{\sqrt{A}} = \frac{6a}{\sqrt{\frac{3\sqrt{3}}{2}a^2}} = 3.7224... \tag{299}$$

The distance between the centers of two neighboring hexagons is $2h$ (see red line in 83) and the ratio with respect to the squared area is

$$\frac{2h}{\sqrt{A}} = \frac{\frac{\sqrt{3}}{2}a}{\sqrt{\frac{3\sqrt{3}}{2}a^2}} = 1.0745... \tag{300}$$

D.4 Simplification of the Voronoi interaction potential

In Subsection 3.3.1 we have simplified the Voronoi potential, starting from the total interaction energy in a system with N cells, given by

$$E = \sum_{i=1}^{N} K_A(A_i - A_0)^2 + \xi P_i^2 + \gamma P_i, \tag{301}$$

where we already have assumed that all cells are identical such that $K_{A_i} = K_A$, $A_{0,i} = A_0$, $\xi_i = \xi$ and $\gamma_i = \gamma$. In this appendix subsection we rewrite this integration to a form that is easier interpretable. The interaction force on a cell i is a function of the derivative of the energy with respect to the position of cell i and therefore we are free to add or subtract terms to the interaction energy that do not depend on the position of cell i. With this in mind, we rewrite

$$E = \sum_{i=1}^{N} K_A(A_i - A_0)^2 + \xi(P_i^2 + \frac{\gamma}{\xi}P_i) \tag{302}$$

$$= \sum_{i=1}^{N} K_A(A_i - A_0)^2 + \xi(P_i^2 + \frac{\gamma}{2\xi})^2 - (\frac{\gamma}{2\xi})^2 \tag{303}$$

$$\sim \sum_{i=1}^{N} K_A(A_i - A_0)^2 + \xi(P_i + \frac{\gamma}{2\xi})^2 \tag{304}$$

$$= \sum_{i=1}^{N} K_A(A_i - A_0)^2 + K_P(P_i - P_0)^2, \tag{305}$$

where in the last line is used that $K_P = \xi$ and $-\frac{\gamma}{2\xi} = P_0$, and the last equation is similar to Equation 57 in Subsection 3.3.1.

Next step is to show that the derivative of the interaction energy with respect to the position of cell i is independent of the choice of A_0 and therefore we can choose A_0 to be equal to the average cell area $A_{\text{avg}} = \frac{A_T}{N}$, with A_T the total area of the system. To show this, assume that we choose a preferred area $A_0 = A_{\text{avg}} + \mu$, in which μ is an arbitrary positive or negative constant.

$$E = \sum_{i=1}^{N} K_A(A_i - A_{\text{avg}} - \mu)^2 + K_P(P_i - P_0)^2 \tag{306}$$

$$= \sum_{i=1}^{N} K_A((A_i - A_{\text{avg}})^2 - 2\mu(A_i - A_{\text{avg}}) + \mu^2) + K_P(P_i - P_0)^2 \tag{307}$$

$$= \sum_{i=1}^{N} K_A((A_i - A_{\text{avg}})^2 + K_P(P_i - P_0)^2 - \sum_{i=1}^{N} 2K_A\mu(A_i - A_{\text{avg}}) + \sum_{i=1}^{N} K_A\mu^2 \tag{308}$$

$$= \sum_{i=1}^{N} K_A(A_i - A_{\text{avg}})^2 + K_P(P_i - P_0)^2 - K_A\mu A_T + K_A\mu A_T + NK_A\mu^2 \tag{309}$$

$$= \sum_{i=1}^{N} K_A(A_i - A_{\text{avg}})^2 + K_P(P_i - P_0)^2 + NK_A\mu^2 \tag{310}$$

$$\sim \sum_{i=1}^{N} K_A(A_i - A_{\text{avg}})^2 + K_P(P_i - P_0)^2, \tag{311}$$

where we have used in the last step that the derivative of $NK_A\mu^2$ to the position of a cell i is zero.

D.5 Self-diffusion coefficient in the VRD- and VV-model

In this appendix subsection we show that the diffusion coefficient of a free particle with velocity v_0 and rotational diffusion with diffusion coefficient D_r is given by $D_0 = \frac{v_0}{2D_r}$. We start with the equation of motion for a free self-propulsive particle under influence of rotational noise:

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = v_0\boldsymbol{n} \tag{312}$$

with $\boldsymbol{n} = (\cos\theta, \sin\theta)$ and

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \eta, \tag{313}$$

with $\langle\eta(t)\eta(t')\rangle = 2D_r\delta(t - t')$. By integration Equation 312 we obtain

$$\boldsymbol{r}(t) = \boldsymbol{r}(0) + \int_0^t \mathrm{d}t_1 v\boldsymbol{n}(t_1). \tag{314}$$

By choosing $\boldsymbol{r}(0) = 0$ we obtain for the mean squared displacement

$$\left\langle \boldsymbol{r}^2(t)\right\rangle = \left\langle \int_0^t \mathrm{d}t_1 v_0\boldsymbol{n}(t_1) \int_0^t \mathrm{d}t_2 v_0\boldsymbol{n}(t_2)\right\rangle \tag{315}$$

$$= v_0^2 \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \left\langle \boldsymbol{n}(t_1)\cdot\boldsymbol{n}(t_2)\right\rangle \tag{316}$$

$$= v_0^2 \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \left\langle \cos\theta(t_1)\cos\theta(t_2) + \sin\theta(t_1)\sin\theta(t_2)\right\rangle \tag{317}$$

$$= v_0^2 \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 \left\langle \cos\left(\theta(t_1) - \theta(t_2)\right)\right\rangle. \tag{318}$$

$$\tag{319}$$

Using $\Delta t_n = \frac{t_2 - t_1}{n}$ with n an integer, we rewrite the term in the integral as follows

$$\langle \cos(\theta(t_1) - \theta(t_2))\rangle = \left\langle \mathcal{R}\left(e^{i(\theta(t_2)-\theta(t_1))}\right)\right\rangle \tag{320}$$

$$= \left\langle \mathcal{R}\left(e^{i(\theta(t_2)-(\theta(t_2-\Delta t)+(\theta(t_2-\Delta t)-(\theta(t_2-2\Delta t)+(\theta(t_2-2\Delta t)-\ldots-(\theta(t_2-(n-1)\Delta t)+(\theta(t_2-(n-1)\Delta t)-(\theta(t_2-n\Delta t))}\right)\right\rangle \tag{321}$$

$$= \left\langle \mathcal{R}\left(e^{i(\theta(t+\Delta t)-\theta(t))}\right)^n\right\rangle \tag{322}$$

$$= \langle \cos(\theta(t+\Delta t_n) - \theta(t))\rangle^n \tag{323}$$

$$= \lim_{n\to\infty} \langle \cos(\theta(t+\Delta t_n) - \theta(t))\rangle^n \tag{324}$$

$$= \lim_{n\to\infty} \left(1 - \frac{1}{2}\left\langle(\theta(t+\Delta t_n) - \theta(t))^2\right\rangle\right)^n \tag{325}$$

$$= \lim_{n\to\infty} (1 - D_r|\Delta t_n|)^n \tag{326}$$

$$= e^{-D_r|t_2-t_1|}, \tag{327}$$

where in Equation 325 is used that $\cos(\theta) \approx 1-\frac{1}{2}\theta^2$ for small θ and in Equation 327 that $\lim_{n\to\infty}(1+\frac{ax}{n})^n = e^{ax}$, with a a constant. Filling Equation 327 in in the mean squared displacement of Equation 318 gives

$$\left\langle \boldsymbol{r}^2(t)\right\rangle = v_0^2 \int_0^t \mathrm{d}t_1 \int_0^t \mathrm{d}t_2 e^{-D_r|t_2-t_1|} \tag{328}$$

$$= v_0^2 \left(\int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 e^{-D_r|t_2-t_1|} + \int_0^t \mathrm{d}t_1 \int_{t_1}^t \mathrm{d}t_2 e^{-D_r|t_2-t_1|}\right) \tag{329}$$

$$= \frac{v_0^2}{D_r}\left(\int_0^t \mathrm{d}t_1 (1 - e^{-D_r t_1}) + \int_0^t \mathrm{d}t_1 (1 - e^{-D_r t_2})\right) \tag{330}$$

$$= \frac{2v_0^2}{D_r}\left(t - \frac{1 - e^{-D_r t}}{D_r}\right). \tag{331}$$

With the MSD we can compute the diffusion coefficient D_0 as follows

$$D_0 = \lim_{t\to\infty} \frac{\left\langle \boldsymbol{r}^2(t)\right\rangle}{4t} \tag{332}$$

$$= \lim_{t\to\infty} \frac{\frac{2v_0^2}{D_r}\left(t - \frac{1-e^{-D_r t}}{D_r}\right)}{4t} \tag{333}$$

$$= \frac{v_0^2}{2D_r}. \tag{334}$$

Because the Vicsek-like alignment in the VV-model does not influence the MSD for an isolated particle, Equation 332 is also the self-diffusion coefficient of the VV-model.

E Algorithms

In this appendix section we explain the algorithms that we have used in this thesis. In Appendix Subsection E.1 we give the algorithm for the VRD-model and the adaptations in this algorithm for the VV-model. In Appendix Subsection E.2 the algorithms for the static and dynamic correlation functions are given and in Appendix Subsection E.3 the MCT-algorithm for the coherent ISF and the adaptations in the algorithm for the incoherent ISF are explained.

E.1 Algorithms for the VRD-model and VV-model

We explain the algorithm of the VRD in Appendix Subsection E.1.1. Hardest part of this algorithm is to implement the derivation of the Voronoi interaction potential and this is explained in Appendix Subsection E.1.2. In Appendix Subsection E.1.3 we describe how we have adapted the VRD-algorithm for the VV-model.

E.1.1 Overview of the algorithm

The VRD-model is defined by Equations 57 till 59 in Subsection 3.3.1 and we solve these equations using molecular dynamics. In a molecular dynamics simulation Newton's equations of motion are solved, and the trajectories of the particles in configuration space are obtained. In order to do so, we rewrite the equations of motions for $\boldsymbol{r}_i$ and θ_i using Euler's method to

$$\boldsymbol{r}_i(t+\Delta t) = \boldsymbol{r}_i(t) - \mu \Delta t \boldsymbol{\nabla}_i E + v_0 \Delta t \hat{\boldsymbol{n}}_i, \tag{335}$$

and

$$\theta_i(t+\Delta t) = \theta_i(t) + \eta_i, \quad \text{with} \quad \langle \eta_i(t)\eta_j(t')\rangle = 2D_r \Delta t \delta(t-t')\delta_{ij}, \tag{336}$$

with Δt the size of a time step. The interaction potential E stays unaltered and is in dimensionless form given by

$$E = \sum_i^N k_a(a_i - a_0)^2 + k_p(p_i - p_0)^2. \tag{337}$$

The algorithm we have used to solve Equations 335 till 337 is sketched in Figure 84 and each individual step is briefly described below.

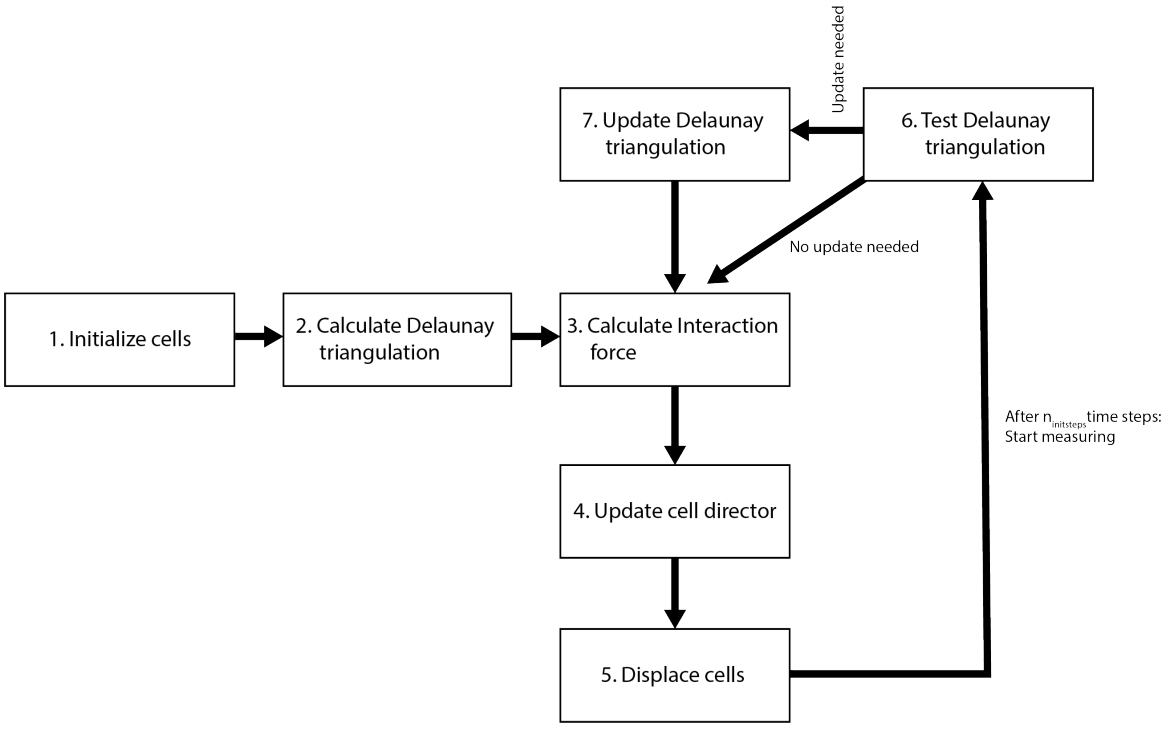

Figure 84: Molecular dynamics algorithm for the VRD-model.

1. Initialization
We start with N cells in a square box of size $\sqrt{N} \times \sqrt{N}$ with periodic boundary conditions and choose the position of the seeds $\boldsymbol{r}_i$ randomly. Different algorithms can be used for the initial position of the seeds to speed up the process to reach a steady state, like random sequential addition, but we do this completely random. Furthermore, each cell is assigned a preferred area a_0 and perimeter p_0 and because these values are identical for all cells the system is monodisperse.

2. Calculate Delaunay triangulation
The Delaunay triangulation is the dual graph of the Voronoi tesselation, see Figure 27. The Voronoi seeds (red dots in Figure 85) are the vertices in a Delaunay triangulation and the edge (black and green lines) are drawn such that the circumcircle (purple) of any triangle (green) does not contain any seeds. There is only one triangulation that satisfies this condition.

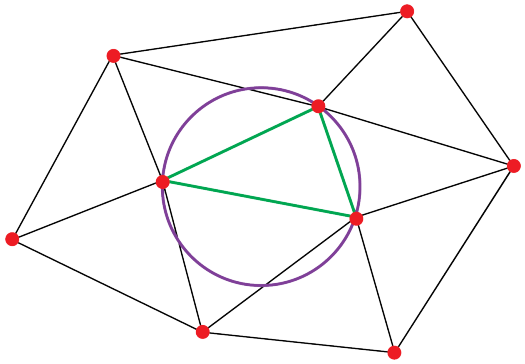

Figure 85: Delaunay triangulation of the seeds (red) with the circumcircle (purple) for one Delaunay triangle (green).

3. Calculate the interaction force
The interaction force on cell i depends on the derivative of the interaction energy E_i of cell i and on the derivative of the energy E_j of its nearest neighbors NN(i). Nearest neighbors are vertices in the Delaunay triangulation connected via only one edge. Thus, the interaction force on cell i is given by

$$\boldsymbol{F}_i = -\boldsymbol{\nabla}_i E = -\begin{pmatrix} \frac{\partial E_i}{\partial r_{i,x}} + \sum_{j \in \mathrm{NN}(i)} \frac{\partial E_j}{\partial r_{i,x}} \\ \frac{\partial E_i}{\partial r_{i,y}} + \sum_{j \in \mathrm{NN}(i)} \frac{\partial E_j}{\partial r_{i,y}} \end{pmatrix}, \tag{338}$$

with x and y Cartesian coordinates. The derivation and expressions for the derivatives are quite complex and are therefore given in the next appendix subsection.

4. Update the director of the self-propulsion force
A random number is drawn from a standard normal distribution and multiplied by the standard deviation of the noise: $\sqrt{2\Delta t D_r}$. This random number is added to the current angle of the director.

5. Update the positions
Since all terms of Equation 335 are known, the displacements for all cells i are executed.

6. Test Delaunay triangulation
The displacements do not always disrupt the Delaunay triangulation and we test this first to save time. This test is done by checking for each Delaunay triangle if its circumcircle contains a seed.

7. Update Delaunay triangulation
If there are Delaunay triangles that contain a seed inside its circumcircle, the triangulation is incorrect. For each incorrect triangle, the triangulation is locally repaired. One possible method is using an edge-flipping algorithm in which one of the edges of the incorrect triangulation flips to connect two other seeds.

One time step contains of step 3 till 7 and we let the system reach a steady state by performing $n_{\mathrm{initsteps}}$ time steps before we start measuring the system.

The VRD-algorithm is implemented in *CellGPU* [84] and we use a modified version for our simulations. *CellGPU* is an open source software program to simulate regular vertex and SPV-models. The SPV-model is implemented with a hybrid CPU/GPU-approach because the GPU highly reduces the calculation costs by doing parallel calculations. The largest time saving is obtained in step 3 where the interaction forces on the cells are calculated parallel due to the GPU-implementation.

E.1.2 Implementation of the Voronoi interaction potential

In this appendix subsection we derive the equation for the derivative of the interaction potential given in Equation 338. But before looking into this derivatives, we need to define some of the properties of a Delaunay triangle. An example of a Delaunay triangle is given in Figure 86. The three vertices $\boldsymbol{r}_i$, $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$ belong to the three cells i, j and k. The Voronoi vertex $\boldsymbol{h}_{ijk}$ inside the triangle is given by the crossing of the perpendicular bisectors of the three edges $\boldsymbol{r}_{ij}$, $\boldsymbol{r}_{jk}$ and $\boldsymbol{r}_{ki}$, where $\boldsymbol{r}_{ij}$ is the vector $\boldsymbol{r}_j - \boldsymbol{r}_i$. By drawing lines from $\boldsymbol{h}_{ijk}$ to $\boldsymbol{r}_i$, $\boldsymbol{r}_j$ and $\boldsymbol{r}_j$, we obtain three times two equal triangles resulting in the following property:

$$|\boldsymbol{h}_{ijk} - \boldsymbol{r}_i| = |\boldsymbol{h}_{ijk} - \boldsymbol{r}_j| = |\boldsymbol{h}_{ijk} - \boldsymbol{r}_k| = R \tag{339}$$

in which R the radius of the circumcircle. By solving Equation 339 using Mathematica we obtain the coordinates of the vertex $\boldsymbol{h}_{ijk}$. After rewriting in a neat way, the vertex $\boldsymbol{h}_{ijk}$ becomes

$$\boldsymbol{h}_{ijk} = \alpha \boldsymbol{r}_i + \beta \boldsymbol{r}_j + \gamma \boldsymbol{r}_k, \tag{340}$$

with

$$\alpha = |\boldsymbol{r}_{jk}|^2 (\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{ik})/d \tag{341}$$
$$\beta = |\boldsymbol{r}_{ik}|^2 (\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj})/d \tag{342}$$
$$\gamma = |\boldsymbol{r}_{ij}|^2 (\boldsymbol{r}_{ik} \cdot \boldsymbol{r}_{jk})/d \tag{343}$$
$$d = 2(r_{ij,x} r_{kj,y} - r_{ij,y} r_{kj,x})^2. \tag{344}$$

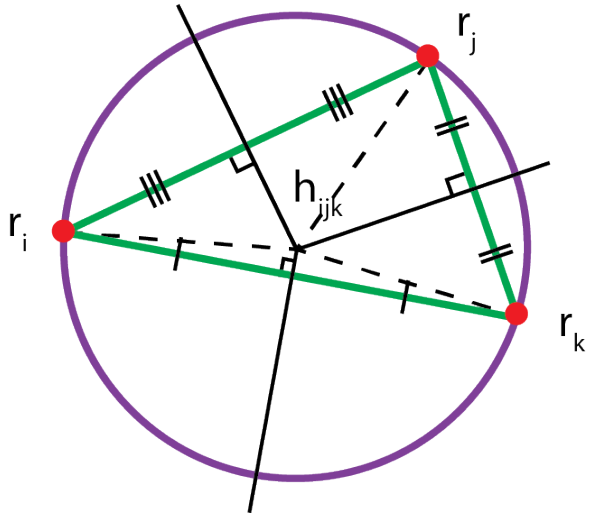

Figure 86: Delaunay triangle connecting cells i, j and k and its circumcircle and circumcenter.

The dimensionless interaction potential E is given by (see Equation 57)

$$E = \sum_i^N k_a (a_i - a_0)^2 + k_p (p_i - p_0)^2, \tag{345}$$

and to calculate the force $\boldsymbol{F}_i$ on cell i, we need to take into account cell i and its neighboring cells $\mathrm{NN}(i)$,

$$\boldsymbol{F}_i = -\boldsymbol{\nabla}_i E = -\begin{pmatrix} \frac{\partial E_i}{\partial r_{i,x}} + \sum_{j \in \mathrm{NN}(i)} \frac{\partial E_j}{\partial r_{i,x}} \\ \frac{\partial E_i}{\partial r_{i,y}} + \sum_{j \in \mathrm{NN}(i)} \frac{\partial E_j}{\partial r_{i,y}} \end{pmatrix}. \tag{346}$$

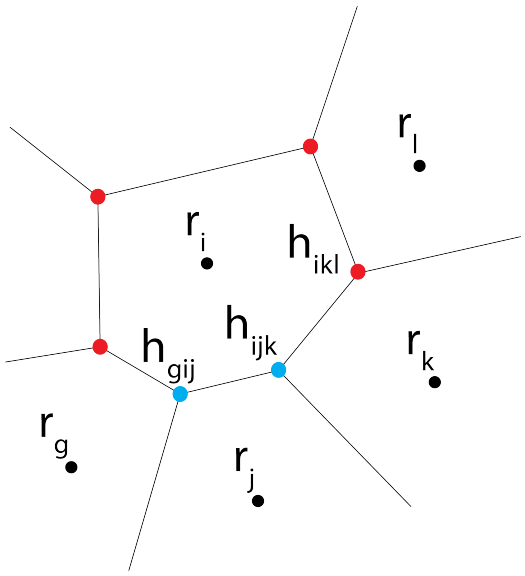

Figure 87: Nomenclatuur and vertices (red and blue) that have to be taken into account to calculate the total interaction force on cell i.

The energy term is a function of the perimeter and area of a cell and this is calculated easiest via the vertices. By using the chain rule, the partial derivatives in Equation 346 can be rewritten as a function of the vertices. For example, the derivative of the energy of cell j to the position of cell i depends on the two shared vertices of cell i and j, namely $\boldsymbol{h}_{ijk}$ and $\boldsymbol{h}_{gij}$ (blue dots in Figure 87), because all other vertices of cell j do not depend on the location of cell i. Thus,

$$\frac{\partial E_j}{\partial r_{i\mu}} = \sum_{\nu} \left(\frac{\partial E_j}{\partial h_{ijk,\nu}} \frac{\partial h_{ijk,\nu}}{\partial r_{i,\mu}} + \frac{\partial E_j}{\partial h_{gij,\nu}} \frac{\partial h_{gij,\nu}}{\partial r_{i,\mu}} \right), \tag{347}$$

in which μ and ν both represent Cartesian coordinates. To calculate the derivative of the energy of cell i to the position of cell i, all the vertices of cell i have to be taken into account (see red and blue dots in Figure 87). The total force on cell i is for Cartesian coordinate μ is given by

$$\begin{aligned} \boldsymbol{F}_{i,\mu} = &- \sum_{j \in \mathrm{NN}(i)} \sum_{k>j, k \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)} \sum_{\nu} \left(\frac{\partial E_i}{\partial h_{ijk,\nu}} \frac{\partial h_{ijk,\nu}}{\partial r_{i,\mu}} \right) \\ &- \sum_{j \in \mathrm{NN}(i)} \sum_{k \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)} \sum_{\nu} \left(\frac{\partial E_j}{\partial h_{ijk,\nu}} \frac{\partial h_{ijk,\nu}}{\partial r_{i,\mu}} + \frac{\partial E_j}{\partial h_{gij,\nu}} \frac{\partial h_{gij,\nu}}{\partial r_{i,\mu}} \right), \end{aligned} \tag{348}$$

with ν representing the Cartesian coordinates x and y. The last step is defining the derivative of the energy of a cell to one of its vertices, given by $\frac{\partial E_i}{\partial h_{ijk,\nu}}$, and the derivative of a vertex to the position of one of its adjacent cells $\frac{\partial h_{ijk,\nu}}{\partial r_{i,\mu}}$. We start with the derivative of the energy, which is rewritten with Equation 345 to

$$\frac{\partial E_i}{\partial h_{ijk,\nu}} = 2k_a(a_i - a_0) \frac{\partial a_i}{\partial h_{ijk,\nu}} + 2k_p(p_i - p_0) \frac{\partial p_i}{\partial h_{ijk,\nu}}, \tag{349}$$

where the derivative of the area to one of the vertices is given by

$$\frac{\partial a_i}{\partial h_{ijk,\nu}} = \frac{1}{2}(l_{ij}\hat{n}_{ij,\nu} + l_{jk}\hat{n}_{jk,\nu}) \tag{350}$$

where l_{ij} is the length of the edge between cell i and j, given by $|\boldsymbol{h}_{gij} - \boldsymbol{h}_{ijk}|$ with $g \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)$, $k \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)$ and $l \neq g$. The vector $\hat{n}_{ij}$ is the unit vector perpendicular to the edge between i and j given by $\frac{(\boldsymbol{h}_{gij} - \boldsymbol{h}_{ijk})_{\perp}}{|\boldsymbol{h}_{gij} - \boldsymbol{h}_{ijk}|}$. The factor $1/2$ in Equation 350 is because the derivative is taken to one of the vertices of the edge, while the position of the other vertex stays fixed, resulting in a triangular-like displacement of the edge.

The derivative of the perimeter to one of the vertices is given by

$$\frac{\partial p_i}{\partial h_{ijk,\nu}} = \frac{(h_{ijk,\nu} - h_{gij,\nu})}{|\boldsymbol{h}_{ijk} - \boldsymbol{h}_{gij}|} + \frac{(h_{ijk,\nu} - h_{ikl,\nu})}{|\boldsymbol{h}_{ijk} - \boldsymbol{h}_{ikl}|}, \tag{351}$$

where again $g \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)$, $l \in \mathrm{NN}(i) \wedge \mathrm{NN}(j)$ and $l \neq g$.

For the derivative of the vertex $\boldsymbol{h}_{ijk}$ to $\boldsymbol{r}_i$, we go back to the formula of the vertex in Equation 340 and notice that $\alpha + \beta + \gamma = 1$ (Mathematica). Now we choose $\boldsymbol{r}_i$ as the origin and rewrite the equation to

$$\boldsymbol{h}_{ijk} = \alpha \boldsymbol{r}_i + \frac{1}{d}(d\beta \boldsymbol{r}_{ij} + d\gamma \boldsymbol{r}_{ik}) \quad \text{with} \quad \boldsymbol{r}_i = \boldsymbol{0}, \tag{352}$$

where $\boldsymbol{r}_{ij}$ is the vector from j to i. Next step is to take the derivative to $\boldsymbol{r}_i$ as follows

$$\frac{\partial h_{ijk,\nu}}{\partial r_{i,\mu}} = \frac{\partial \alpha r_{i,\nu}}{\partial r_{i,\mu}} + \frac{\partial \frac{1}{d}(d\beta r_{ij,\nu} + d\gamma r_{ik,\nu})}{\partial r_{ik,\mu}} \qquad \text{with} \quad r_{i,\mu} = 0, \tag{353}$$

$$= \alpha \delta_{\mu\nu} + \frac{1}{d}(r_{ij,\nu}\frac{\partial(\beta d)}{\partial r_{i,\mu}} + r_{ik,\nu}\frac{\partial(\beta d)}{\partial r_{i,\mu}} - \frac{1}{d}\frac{\partial d}{\partial r_{i,\mu}}) \tag{354}$$

$$= (1 - \gamma - \delta)\delta_{\mu\nu} + \frac{1}{d}(r_{ij,\nu}\frac{\partial(\beta d)}{\partial r_{i,\mu}} + r_{ik,\nu}\frac{\partial(\beta d)}{\partial r_{i,\mu}} - \frac{1}{d}\frac{\partial d}{\partial r_{i,\mu}}). \tag{355}$$

The derivatives in Equation 355 are now the derivatives of Equations 341 till 344 and given by

$$\frac{\partial(\beta d)}{\partial r_{i,\mu}} = 2r_{ik,\mu}(\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{jk}) + |\boldsymbol{r}_{ik}|^2 r_{jk,\mu} \tag{356}$$

$$\frac{\partial(\gamma d)}{\partial r_{i,\mu}} = -2r_{ij,\mu}(\boldsymbol{r}_{ik} \cdot \boldsymbol{r}_{jk}) - |\boldsymbol{r}_{ij}|^2 r_{jk,\mu} \tag{357}$$

$$\frac{\partial d}{\partial r_{i,\mu}} = 4(r_{ij,x}r_{kj,y} - r_{ij,y}r_{kj,x})(r_{jk,x}\delta_{y\mu} - r_{jk,y}\delta_{x\mu}). \tag{358}$$

Thus, the force $\boldsymbol{F}_i$ on cell i is calculated by Equation 348, after filling in Equations 349 till 351 and 355 till 358.

E.1.3 Adaptations for the VV-algorithm

The algorithm for the VV-model is similar to the algorithm shown in Figure 89 for the VV-model, only with a change in Step 4. In this step the cell director is updated according to Equation 67, which is rewritten using Euler's method to

$$\theta_i(t + \Delta t) = \theta_i(t) + \frac{\Delta t}{\tau_v} \langle \phi_j(t) - \theta_i(t) \rangle_{0<|\boldsymbol{r}_i - \boldsymbol{r}_j|<R_V} + \eta_i(t), \tag{359}$$

with $\eta_i(t)$ a random number with mean zero and variance $2D_r\Delta t$. Equation 359 is implemented in the CellGPU-code as follows. First all neighbors that are within R_V from cell i are determined and because the algorithm stores a neighbor list this can be done efficiently using an iterative algorithm. First the algorithm determines which direct neighbors j that share an edge with cell i are at a distance less than R_V and we call them first neighbors. Next the direct neighbors of the first neighbors are checked and we call them second neighbors if they are within R_V from cell i. Then, the direct neighbors of the second neighbors are checked, and this continues until no more direct neighbors are found within R_V from cell i. Then, the average angle $\langle \phi \rangle_{R_V}$ of the velocity vectors of all neighbors is calculated and a random number from a standard normal distribution is drawn and multiplied with $\sqrt{2D_r\Delta t}$. The new director is now calculated using Equation 359.

E.2 Algorithms for the correlation functions

In this appendix subsection we summarize the algorithms that we used to calculated the scalar and vectorial radial distribution function, the scalar and vectorial static structure factor and the coherent and incoherent intermediate scattering function. N is the number of particles, and N^{ss} the number of snapshots that are taken into account. $\boldsymbol{r}_{ij}(s) = (r_{ij,x}, r_{ij,y}) = \boldsymbol{r}_j(s) - \boldsymbol{r}_i(s)$ is the vector from particle i to particle j in snapshots s and $r_{ij}(s)$ is its magnitude.

The algorithm for calculating the vectorial radial distribution function is given by Algorithm 1. The algorithm makes a two dimensional histogram H with bin size Δh of all interparticle distances in the x- and y-direction. The algorithm for calculating the scalar radial distribution function is given by Algorithm 2. This algorithm makes a one dimensional histogram H with bin size Δh of all interparticle distances.

Algorithm 1 Radial distribution function - vector

for all N^{ss} snapshots s **do**
  for all N particles i **do**
    for all N particles $j \neq i$ **do**
      Add 1 to histogram entry $H(h_x, h_y)$ where the index h_x is given by round $\left(\frac{r_{ij,x}(s)}{\Delta h}\right)$ with Δh the bin spacing and h_y is given by round $\left(\frac{r_{ij,y}(s)}{\Delta h}\right)$
    end for
  end for
end for
Calculate $g(r_{h,x}, r_{h,y}) = \frac{H(h_x,h_y)}{N^{ss}N\rho_0 h^2}$ with $r_{h,x} = (h_x + \frac{1}{2})\Delta h$ and $r_{h,y} = (h_y + \frac{1}{2})\Delta h$

Algorithm 2 Radial distribution function - scalar

for all N^{ss} snapshots **do**
  for all N particles i **do**
    for all N particles $j \neq i$ **do**
      Add 1 to histogram entry $H(h)$ where the index h is given by round $\left(\frac{r_{ij}(s)}{\Delta h}\right)$ with Δh the bin spacing
    end for
  end for
end for
Calculate $g(r_h) = \frac{H(h)}{N^{ss}N\rho_0 V(h)}$ with $r_h = (h + \frac{1}{2})\Delta h$ and $V(h)$ the 2D volume of the bin given by $V(h) = 4\pi((h+1)^2 - h^2)\Delta h$

The algorithm for calculating the vectorial static structure factor via the direct method is given by Algorithm 3 and for the scalar static structure factor by Algorithm 4. First N_w wave vectors $\boldsymbol{k} = (k_x, k_y)$, indexed with w, are created satisfying $k_x = \frac{2\pi n_x}{L}$ and $k_y = \frac{2\pi n_y}{L}$ and having magnitude k. We calculate the contribution of each particle for each wave vector and add this to a temporary function S_{temp}. In the last step, we bin the results of $S(\boldsymbol{k})$ or $S(k)$ to reduce the statistical error.

Algorithm 3 Static structure factor - vector

Make N_w wave vectors matching the periodicity of the box
for all N_{ss} snapshots s **do**
  for all N_w wave vectors w **do**
    for all N particles i **do**
      Calculate $S_{\text{temp}}(k_{w,x}, k_{w,y}) += (\sum_{i=1}^{N} \cos(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s))^2 + (\sum_{i=1}^{N} \sin(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s)))^2$
    end for
  end for
end for
Calculate the average static structure factor $S(k_x, k_y) = \frac{S_{\text{temp}}(k_x,k_y)}{NN_{ss}}$
Bin $S(k_{h,x}, k_{h,y})$ by taking the average $S(k_x, k_y)$ with $k_{h,x} - \frac{\Delta h}{2} \leq k_x < k_{h,x} + \frac{\Delta h}{2}$ and $k_{h,y} - \frac{\Delta h}{2} \leq k_y < k_{h,y} + \frac{\Delta h}{2}$

Algorithm 4 Static structure factor - scalar

Make N_w wave vectors matching the periodicity of the box
for all N_{ss} snapshots s **do**
  for all N_w wave vectors w **do**
    for all N particles i **do**
      Calculate $S_{\text{temp}}(k_w) += (\sum_{i=1}^{N} \cos(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s))^2 + (\sum_{i=1}^{N} \sin(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s)))^2$
    end for
  end for
end for
Calculate the average static structure factor $S(k) = \frac{S_{\text{temp}}(k)}{NN_{ss}}$
Bin the average static structure factor $S(k_h)$ by taking the average $S(k)$ with $k_h - \frac{\delta h}{2} \leq k < k_h + \frac{\delta h}{2}$

The algorithm for calculating the coherent intermediate scattering function $F(\overline{k}, \tau)$ is given by Algorithm 5. First $\overline{k}$ where $S(k)$ has its first peak is determined and a set of wave vectors W is made that fall within a bin of size Δh around $\overline{k}$ and satisfy the periodic boundary conditions. Next, we calculate for each time τ smaller than the maximum time $\tau_{\max}$ the intermediate scattering function. The algorithm for the incoherent intermediate scattering function $F_s(\overline{k}, \tau)$ is given by Algorithm 6.

Algorithm 5 Coherent intermediate scattering function

Define the set of wave vectors W satisfying the boundary conditions of the box and satisfying $k_w - \frac{\Delta h}{2} \leq \overline{k} < k_w + \frac{\Delta h}{2}$ with $\overline{k} = \arg\max(S(k))$, using Algorithm 4
for all $N^{ss} + \tau_{\max}$ snapshots s **do**
  for all N_w wave vectors w **do**
    Calculate $Y_w^{\cos}(s) = \sum_{i=1}^{N} \cos(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s))$ and $Y_w^{\sin}(s) = \sum_{i=1}^{N} \sin(\boldsymbol{k}_w \cdot \boldsymbol{r}_i(s))$
  end for
end for
for $\tau < \tau_{\max}$ **do**
  for all N^{ss} snapshots s **do**
    for all N_w wave vectors w **do**
      $F_{\text{temp}}(\tau) += Y_w^{\cos}(s) Y_w^{\cos}(s+\tau) + Y_w^{\sin}(s) Y_w^{\sin}(s+\tau)$
    end for
  end for
  Calculate the intermediate scattering function $F(\overline{k}, \tau) = \frac{F_{\text{temp}}(\tau)}{N_{ss} N N_w}$
end for

Algorithm 6 Incoherent intermediate scattering function

Define the set of wave vectors W satisfying the boundary conditions of the box and satisfying $k_w - \frac{\Delta h}{2} \leq \overline{k} < k_w + \frac{\Delta h}{2}$ with $\overline{k} = \arg\max(S(k))$, using Algorithm 4
for $\tau < \tau_{\max}$ **do**
  for all N^{ss} snapshots s **do**
    for all N_w wave vectors w **do**
      for all N particles i **do**
        $F_{\text{s,temp}}(\tau) += \cos(\boldsymbol{r}_i(\tau) \cdot \boldsymbol{k}_w)$
      end for
    end for
  end for
  Calculate the intermediate scattering function $F_s(\overline{k}, \tau) = \frac{F_{\text{s,temp}}(\tau)}{N_{ss} N N_w}$
end for

E.3 Algorithms for mode coupling theory

In this appendix subsection we explain the algorithm that we have used to in our mode coupling theory analysis, which is an adapted version of the algorithm written by L.M.C. Janssen. In Appendix Subsection E.3.1 we give a general overview of this algorithm and in Appendix Subsection E.3.2 till E.3.6 we give the details of the algorithm. We conclude in Appendix Subsection E.3.7 with the adaptations in the MCT-algorithm for the incoherent ISF.

E.3.1 Overview of the algorithm

The MCT-equation for the coherent intermediate scattering function is given by Equation 41. Because our simulations are overdamped, we use the overdamped version of the MCT-equation by neglecting the second order term. Furthermore, we normalize $F(k,t)$ via dividing by $F(k,0) = S(k)$ and set the short term memory kernel to $A\delta(t)$ with $A = 1$. Thus, we solve the following MCT-equation (see Appendix Subsection E.3.2):

$$\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} + \Omega^2 F(k,t) + \int_0^t \mathrm{d}\tau M^{\mathrm{MCT}}(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} = 0, \tag{360}$$

with $\Omega^2 = \frac{k^2 D_0}{S(k)}$ and

$$M^{\mathrm{MCT}}(k,t) = \frac{\rho D_0}{8\pi^2}\int \mathrm{d}\mathbf{q} \mid \tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} \mid^2 F(q,t)S(q)F(\mid \boldsymbol{k}-\boldsymbol{q} \mid,t)S(\mid \boldsymbol{k}-\boldsymbol{q} \mid), \tag{361}$$

with

$$\tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} = (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(q) + \hat{\boldsymbol{k}}\cdot(\boldsymbol{k}-\boldsymbol{q})c(\mid \boldsymbol{k}-\boldsymbol{q} \mid), \tag{362}$$

with

$$c(k) = \frac{1}{\rho}\left(1 - \frac{1}{S(k)}\right). \tag{363}$$

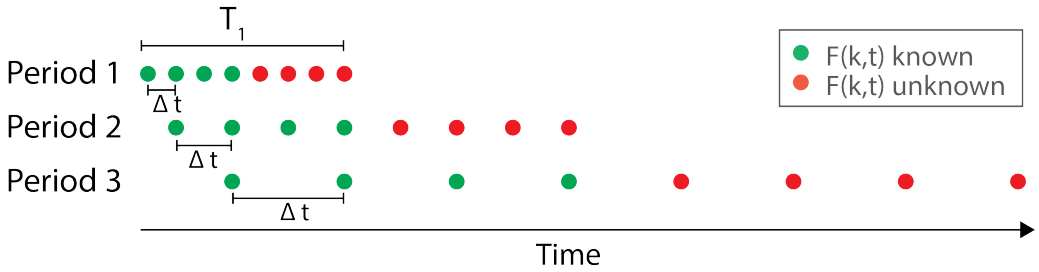

Figure 88: Illustration of the integration pattern used by Fuchs et al. [84], where in each period the step size Δt is doubled.

To solve this first-order integro-differential equation in time for many orders of magnitude efficiently, we apply the algorithm specified by Fuchs *et al.* [84] and illustrated in Figure 88. We start with Period 1 of length T_1 and divide the period in $4N_T$ equal segments of $\Delta t = T_1/4N_T$ ($N_T = 2$ in Figure 88). We assume that the value of $F(k,t)$ for the first $2N_T$ segments is known (for example via a Taylor expansion) and based on these points we calculate the value at the last $2N_T$ segments of Period 1. Then, we double the period to $T_2 = 2T_1$ and again divide the period in $4N_T$ equal segments with $\Delta t = T_2/4N_T$. Again the value of $F(k,t)$ for the first $2N_T$ segments is known and we can calculate $F(k,t)$ for the last $2N_T$ segments. This procedure is iterated until the maximum time $t_{\max}$ is reached.

The overview of the full MCT-algorithm to calculate $F(k,t)$ is given in Figure 89 and we discuss the main steps below.

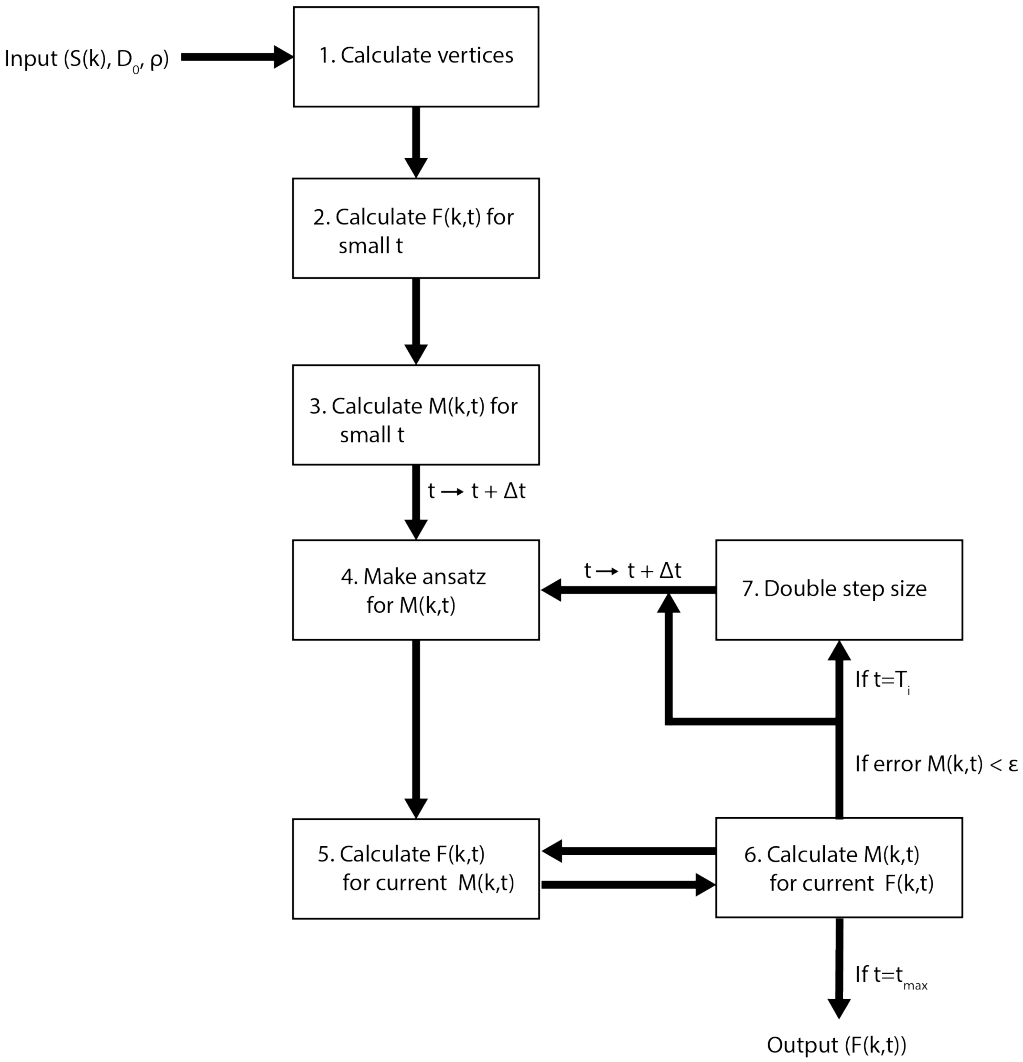

Figure 89: Schematic overview of the MCT-algorithm to calculate the coherent intermediate scattering function.

1. Input
The scalar static structure factor, the self-diffusion coefficient D_0 and the density ρ serve as input for the MCT-algorithm.

2. Calculate vertices
The vertices in Equation 362 do not depend on time, and therefore only need to be calculated once. The formula for calculating these vertices is derived in Appendix Subsection E.3.3.

3. Calculate $F(k,t)$ for small t
Because we are not interested in what happens at short-time scales, we approximate $F(k,t)$ for small t by using a Taylor expansion, given in Appendix Subsection E.3.4.

4. Calculate $M(k,t)$ for small t
Knowing $F(k,t)$ for small t, we can calculate $M(k,t)$ for the same values of t using Equation 361. To do this, we change the integral in Equation 361 to a sum over a discrete set of k-values at which the $S(k)$ is known and this rewritten equation is given in Appendix Subsection E.3.3.

5. Make an ansatz for $M(k,t)$
To calculate $F(k,t+\Delta t)$ we solve the following equation for all k-values until the cut-off value of $S(k)$:

$$A_k F(k,t+\Delta t) + B_k^t = D_k M(k,t+\Delta t) + C_k^t, \tag{364}$$

where A_k and D_k are functions independent of t, and B_k^t and C_k^t depend only on times smaller than the current time t. The explicit expressions of A_k, B_k^t, C_k^t and D_k are given in Appendix Subsection E.3.5. There are n_k equations of Equation 364 with n_k the number of wave vectors, and $M(k,t)$ depends on all of them. Therefore

we need to solve Equation 364 for $M(k, t+\Delta t)$ and $F(k, t+\Delta t)$ at the same time using an iterative method. To do this, we start with the ansatz $M(k, t+\Delta t) = M(k, t)$.

6. Calculate $F(k,t)$ for current $M(k,t)$
With the current $M(k, t+\Delta t)$ (either known from the ansatz or previous iterations), we calculate $F(k, t+\Delta t)$ using Equation 364.

7. Calculate $M(k,t)$ for current $F(k,t)$
With Equation 361, we calculate a new value for $M^{\text{new}}(k, t+\Delta t)$ based on the current value of $F(k, t+\Delta t)$. The relative error is defined as

$$\text{Error} = \frac{M^{\text{new}}(k, t+\Delta t) - M^{\text{old}}(k, t+\Delta t)}{M^{\text{old}}(k, t+\Delta t)}, \tag{365}$$

with $M^{\text{old}}(k, t+\Delta t)$ the previous estimate of $M(k, t+\Delta t)$). When this relative error is larger than ϵ, the algorithm returns to Step 5 to calculate a new estimate for $F(k, t+\Delta t)$. Step 5 and 6 are iterated until the error is smaller than ϵ and the algorithm proceeds.

8. Double the step size
When $F(k,t)$ has been calculated for $2N_T$ time points within the current time Period T_i, we double the step size and calculate $F(k,t)$ for the next $2N_T$ time points. How the discrete values of $F(k,t)$, $M(k,t)$ and their integrals are updated for $\Delta t \to 2\Delta t$ is described in Appendix E.3.6.

E.3.2 Rewrite of the MCT-equation

In deriving the mode coupling theory, we have removed the fast modes from the memory kernel (see Subsection 2.3.1) because we are only interested in the long-term dynamics. We include these fast modes by adding $A\delta(t)$ to the memory kernel in the isotropic version of the MCT-equation as follows

$$\frac{\mathrm{d}F^2(k,t)}{\mathrm{d}t^2} + \Omega^2 F(k,t) + \int_0^t \mathrm{d}\tau \underbrace{M(k,t-\tau)}_{\underbrace{A\delta(t)}_{\text{fast modes}} + \underbrace{M^{\text{MCT}}(k,t-\tau)}_{\text{slow modes}}} \frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} = 0, \tag{366}$$

where the superscript MCT is used to emphasize that it represents the memory kernel from mode coupling theory and $\Omega^2 = \frac{k^2 D_0}{S(k)}$. Equation 366 is rewritten to

$$\frac{\mathrm{d}F^2(k,t)}{\mathrm{d}t^2} + A\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} + \Omega^2 F(k,t) + \int_0^t \mathrm{d}\tau M^{\text{MCT}}(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} = 0. \tag{367}$$

Because our simulations are based on an overdamped equation of motion, we neglect the second-order derivative. Furthermore, we set $A = 1$, because a change in A corresponds to a change in D_0, which only scales the time axis in the MCT-results (see Subsection D.2). This results in

$$\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} + \Omega^2 F(k,t) + \int_0^t \mathrm{d}\tau M^{\text{MCT}}(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} = 0. \tag{368}$$

Because we want to calculate the normalized $F(k,t)$, we divide $F(k,t)$ by $F(k,0) = S(k)$. Equation 368 stays the same, but the memory kernel changes to

$$M(k,t) = \frac{\rho D_0}{8\pi^2} \int \mathrm{d}\boldsymbol{q} \mid \tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} \mid^2 F(q)S(q,t)F(\mid \boldsymbol{k}-\boldsymbol{q} \mid,t)S(\mid \boldsymbol{k}-\boldsymbol{q} \mid). \tag{369}$$

This results in the MCT-equation specified by Equations 360 - 363.

E.3.3 Discretization of $M(k,t)$ and calculation of the vertices (step 1)

We only know the static structure factor for discrete k-values given by $(i+\frac{1}{2})\Delta k$, with i an integer and Δk the spacing between grid points. Therefore, we transform the integral over $\mathrm{d}\boldsymbol{q}$ in Equation 369 to a sum dependent only on the magnitude of the wave vectors. First we rewrite Equation 369 using polar coordinates ($\int \mathrm{d}\boldsymbol{q}... = \int_0^\infty \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}qq$) to

$$M(k,t) = \frac{\rho D_0}{8\pi^2} \int_0^\infty q\mathrm{d}q \int_0^{2\pi} \mathrm{d}\phi \mid \tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} \mid^2 S(q)S(\mid \boldsymbol{k}-\boldsymbol{q} \mid)F(q,t)F(\mid \boldsymbol{k}-\boldsymbol{q} \mid,t) \tag{370}$$

$$= \frac{2\rho D_0}{8\pi^2} \int_0^\infty q\mathrm{d}q \int_0^{\pi} \mathrm{d}\phi \mid \tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} \mid^2 S(q)S(\mid \boldsymbol{k}-\boldsymbol{q} \mid)F(q,t)F(\mid \boldsymbol{k}-\boldsymbol{q} \mid,t) \tag{371}$$

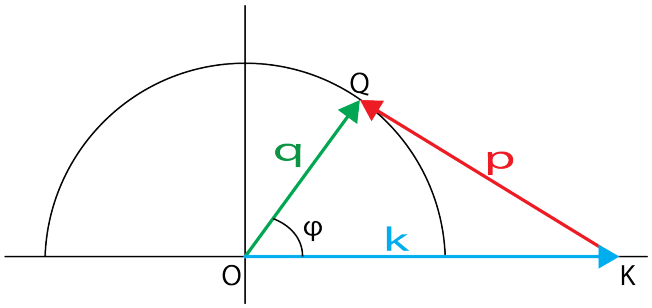

Figure 90: Defining the vector $\boldsymbol{p} = \boldsymbol{k} - \boldsymbol{q}$.

We choose k to be in the x-direction such that the angle between $\boldsymbol{k}$ and $\boldsymbol{q}$ is given by ϕ (see Figure 90). The circle shows all possible $\boldsymbol{q}$-vectors with the same magnitude. We define a new vector $\boldsymbol{p}$ from $\boldsymbol{k}$ to $\boldsymbol{q}$ such that $p = |\boldsymbol{k}-\boldsymbol{q}|$. We apply the cosine rule on the triangle OQK given in Figure 90 for angle ϕ:

$$\cos\phi = \frac{q^2+k^2-p^2}{2qk}, \tag{372}$$

and the derivative is

$$\sin\phi \mathrm{d}\phi = \frac{2p\mathrm{d}p}{2kq} = \frac{p\mathrm{d}p}{kq}, \tag{373}$$

because only p depends on ϕ.

Furthermore,

$$\sin\phi = \sqrt{(1-\cos(\phi)^2} = \sqrt{1-\frac{1}{4q^2k^2}(q^2+k^2-p^2)} = \frac{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}}{2qk}. \tag{374}$$

Now $q\mathrm{d}\phi$ is

$$q\mathrm{d}\phi = \frac{p\mathrm{d}p}{k\sin\phi} = \frac{2pq\mathrm{d}p}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}}, \tag{375}$$

and with this equation for $q\mathrm{d}\phi$ we can rewrite the memory kernel given in Equation 371 as

$$M(k,t) = \frac{2\rho D_0}{8\pi^2} \int_0^\infty \mathrm{d}q \int_{|k-q|}^{|k+q|} \mathrm{d}p \frac{2pq}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}} \mid \tilde{V}_{p,q} \mid^2 S(q)S(p)F(q)F(p,t), \tag{376}$$

where we have used that $|\boldsymbol{k}-\boldsymbol{q}| \leq p \leq |\boldsymbol{k}+\boldsymbol{q}|$. We transform the integrals into sums to get

$$M(k,t) = \frac{4\rho D_0}{8\pi^2}(\Delta_k)^2 \sum_{q=0}^{q_{\max}} \sum_{p=|k-q|}^{|k+q|} \frac{pq}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}} \mid \tilde{V}_{p,q} \mid^2 S(q)S(p)F(q)F(p,t), \tag{377}$$

where we have used that q has a maximum value $q_{\max}$ because $S(k)$ is cut off.

The vertices can be rewritten to

$$\begin{aligned}
\tilde{V}_{p,q} &= (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(q) + (\hat{\boldsymbol{k}}\cdot\boldsymbol{p})c(p) && (378)\\
&= (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(q) + (\hat{\boldsymbol{k}}\cdot \mid \boldsymbol{k}-\boldsymbol{q} \mid)c(p) && (379)\\
&= \frac{1}{2k}\left((k^2+q^2-(k^2-2k\cdot q+q^2))C(q) + (2k^2-2k\cdot q-q^2+q^2)C(p)\right) && (380)\\
&= \frac{1}{2k}\left((k^2+q^2-p^2)C(q) + (k^2-q^2+p^2)C(p)\right). && (381)
\end{aligned}$$

This results in the following equation for the total memory kernel

$$M(k,t) = \frac{\rho D_0}{8\pi^2 k^2}(\Delta_k)^2 \sum_{q=0}^{q_{\max}} \sum_{p=|k-q|}^{|k+q|} \frac{pq\left((k^2+q^2-p^2)C(q) + (k^2-q^2+p^2)C(p)\right)^2}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}} S(q)S(p)F(q,t)F(p,t), \tag{382}$$

in which the the summation over q is over the same equidistant grid as k.

The only time dependency in the memory kernel comes from $F(q,t)F(p,t)$, thus $Z(k,p)$, defined by

$$Z(k,p) = \frac{\rho D_0}{8\pi^2 k^2}(\Delta_k)^2 \frac{pq\left((k^2+q^2-p^2)C(q) + (k^2-q^2+p^2)C(p)\right)^2}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}} S(q)S(p), \tag{383}$$

only has to be calculated once during the execution of the algorithm.

E.3.4 Calculation of $F(k,t)$ and $M(k,t)$ for small t (step 2 and 3)

We use a Taylor expansion to calculate $F(k,t)$ for small values of t:.

$$F(k,t) = F(k,0) + \frac{\mathrm{d}F(k,0)}{\mathrm{d}t}(t-0) + \frac{1}{2}\frac{\mathrm{d}^2F(k,0)}{\mathrm{d}t^2}(t-0)^2 \tag{384}$$

From Equation 368 we know

$$\frac{\mathrm{d}F(k,0)}{\mathrm{d}t} = -\Omega^2 F(k,0) = -\Omega^2, \tag{385}$$

and by taking the derivative of Equation 367 we obtain

$$\frac{\mathrm{d}^2F(k,0)}{\mathrm{d}t^2} + \Omega^2\frac{\mathrm{d}F(k,0)}{\mathrm{d}t} + M(k,0)\frac{\mathrm{d}F(k,0)}{\mathrm{d}t} = 0, \tag{386}$$

which gives

$$\frac{\mathrm{d}^2F(k,0)}{\mathrm{d}t^2} = \Omega^4 F(k,0) + \Omega^2 M(k,0)F(k,0) = \Omega^2(\Omega^2 + M(k,0)). \tag{387}$$

Thus the formula of $F(k,t)$ for small t is

$$F(k,t) = 1 - \Omega^2 t + \Omega^2(\Omega^2 + M(k,0))t^2. \tag{388}$$

Now we know $F(k,t)$ for small t, we can compute $M(k,t)$ for the same values of t using Equation 382.

E.3.5 Iterative solving for $F(k,t)$ and $M(k,t)$ (step 4, 5 and 6)

Assume we want to calculate $F(k,t_i)$ at time t_i, then we first rewrite Equation 382 in terms that depend only on $t < t_i$ and $M(k,t_i)$. We start with rewriting $\int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}t}$ and break these integral into two parts at an arbitrary point t_2

$$\int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} \tag{389}$$

$$= \int_0^{t2} \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} + \int_{t2}^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F(k,\tau)}{\mathrm{d}\tau} \tag{390}$$

$$= M(k,t-\tau)F(k,\tau)\Big|_0^{t_2} - \int_0^{t2} \mathrm{d}\tau \frac{\mathrm{d}M(k,t-\tau)}{\mathrm{d}\tau}F(k,\tau) + \int_{t-t2}^0 \mathrm{d}\tau M(k,\tau)\frac{\mathrm{d}F(k,t-\tau)}{\mathrm{d}\tau}, \tag{391}$$

where in the last formula we have integrated the first integral of Equation 390 by parts and have applied the coordinate transformation $\tau \to t-\tau$ on the second integral.

We break both integrals into integrals of length δt, the first in n_1 parts and the second in n_2 parts:

$$\ldots \quad = M(k,t-\tau)F(k,\tau)\Big|_0^{t_2} - \sum_{j=1}^{n_1}\int_{t_{j-1}}^{t_j} \mathrm{d}\tau \frac{\mathrm{d}M(k,t-\tau)}{\mathrm{d}\tau}F(k,\tau) - \sum_{j=1}^{n2}\int_{t_{j-1}}^{t_j} \mathrm{d}\tau M(k,\tau)\frac{\mathrm{d}F(k,t-\tau)}{\mathrm{d}\tau}. \tag{392}$$

Both integrals have the form $\int_{t_{j-1}}^{t_j} \mathrm{d}\tau \frac{\mathrm{d}F(A(\tau)B(\tau)}{\mathrm{d}\tau}$, which we can approximate with (by using the trapezoidal rule for the integral)

$$\int_{t_{j-1}}^{t_j} \mathrm{d}\tau \frac{\mathrm{d}A(\tau)B(\tau)}{\mathrm{d}\tau} \approx \frac{[A(t_j)-A(t_{j-1})]}{\Delta t}\int_{t_{j-1}}^{t_j} \mathrm{d}\tau B(\tau) \tag{393}$$

$$\approx [A(t_j)-A(t_{j-1})\mathcal{I}[B(t_j)], \tag{394}$$

with $\mathcal{I}[B(t_j)]$ times Δt an approximation for the integral $\int_{t_{j-1}}^{t_j} \mathrm{d}\tau B(\tau)$. This integral is in the first period T_1 given by $\mathcal{I}[B(t_j)] = \frac{1}{2}[B(t_{j-1}) + B(t_j)]$ and in next periods calculated when the step size is doubled in step 7.

We obtain

$$\begin{aligned}\int_0^t \mathrm{d}\tau M(k,t-\tau)\frac{\mathrm{d}F((k,\tau)}{\mathrm{d}\tau} &= M(k,t-t_2)F(k,t_2) - M(k,t)F(k,0) \\ &- \sum_{j=1}^{n_1}[M(t-t_j)-M(t-t_{j-1})]\mathcal{I}[F(t_j)] - \sum_{j=1}^{n_2}[F(t-t_j)-F(t-t_{j-1})]\mathcal{I}[M(t_j)].\end{aligned} \tag{395}$$

Furthermore, we approximate $\frac{\mathrm{d}F(k,t)}{\mathrm{d}t}$ by

$$\frac{\mathrm{d}F(k,t)}{\mathrm{d}t} \approx \frac{1}{2\Delta t}F(k,t_{i-2}) - \frac{2}{\Delta t}F(k,t_{i-1}) + \frac{3}{2\Delta t}F(k,t_i). \tag{396}$$

Now we have all the ingredients to rewrite Equation 368 for time t_i as follows

$$\begin{aligned}\frac{1}{2\Delta t}F(k,t_{i-2}) - \frac{2}{\Delta t}F(k,t_{i-1}) + \frac{3}{2\Delta t}F(k,t_i) + \Omega^2 F(k,t_i) + M(k,t_{i-i2})F(k,t_{i2}) - M(k,t_i)F(k,0) \\ - \sum_{j=1}^{i2}[M(t_{i-j})-M(t_{i-j+1})\mathcal{I}][F(t_j)] - \sum_{j=1}^{i-i2}[F(t_{i-j})-F(t_{i-j+1})\mathcal{I}[F(t_j)] = 0,\end{aligned} \tag{397}$$

where we have used $t_2 = i2 = i/2$. From this it follows that $n_1 = i2$ and $n_2 = i - i2$.

Equation 397 has the following form

$$A_k F(k,t_i) + B_k^{t_i} = D_k M(k,t_i) + C_k^{t_i}, \tag{398}$$

where A_k, $B_k^{t_i}$, $C_k^{t_i}$ and D_k are as defined by

$$A_k = \frac{3}{2\Delta t} + \Omega^2 + \mathcal{I}[M(t_1)] \tag{399}$$

$$B_k^{t_i} = \frac{1}{2\Delta t} F(k,t_{i-2}) - \frac{2}{\Delta t} F(k,t_{i-1}) \tag{400}$$

$$C_k^{t_i} = -M(k,t_{i-i2}) F(k,t_{i2}) + F(k,t_{i-1}) \frac{1}{2}[M(t_0) + M(t_1)] + M(k,t_{i-1}) \frac{1}{2}[F(t_0) + F(t_1)] \tag{401}$$

$$+ \sum_{j=2}^{i2} [M(t_{i-j}) - M(t_{i-j-1})] \frac{1}{2}[F(t_{j-1}) + F(t_j)] + \sum_{j=2}^{i-i2} [F(t_{i-j}) - F(t_{i-j+1})] \frac{1}{2}[M(t_{j-1}) + M(t_j)] \tag{402}$$

$$D_k = 1 - \frac{1}{2}\mathcal{I}[F(t_1)]. \tag{403}$$

$$\tag{404}$$

Thus, for the current estimate of $M(k,t_i)$, we calculate an estimate for $F(k,t_i)$ using Equation 398 with Equation 399 till 403. With this $F(k,t_i)$, we can calculate a new $M(k,t_i)$ using Equation 369. This procedure is repeated until

$$\frac{M^{\text{new}}(k,t+\Delta t) - M^{\text{old}}(k,t+\Delta t)}{M^{\text{old}}(k,t+\Delta t)} < \epsilon, \tag{405}$$

with ϵ a small algorithm parameter of which the value can be chosen.

E.3.6 Step size doubling (step 7)

When $i = N_T$ and $F(k,t_i)$ and $M(k,t_i)$ are calculated, the step size is doubled to $2\Delta t$. Let i denote the numbering in the current period, and j the numbering in the next period, we now define a mapping for $F(k,t_i)$, $\mathcal{I}[F(t_i)]$, $M(k,t_i)$ and $\mathcal{I}[M(t_i)]$ to the corresponding functions dependent on t_j.

For $1 \leq j \leq N/2$, we know the values of $F(k,t_j)$ and $M(k,t_j)$ and the mapping is

$$2i \to j \tag{406}$$

$$F(k,t_{2i}) \to F(k,t_j) \tag{407}$$

$$M(k,t_{2i}) \to M(k,t_j) \tag{408}$$

$$\tag{409}$$

The mapping of $\mathcal{I}[F(t_i)]$ and $\mathcal{I}[M(t_i)]$ is more complicated and is different for $1 \leq j \leq N/4$ and $N/4+1 \leq j \leq N/2$. For the first $N/4$ points we write

$$\frac{1}{2}(\mathcal{I}[F(k,t_{2i-1})] + \mathcal{I}[F(k,t_{2i})]) \to \mathcal{I}[F(k,t_j)] \tag{410}$$

$$\frac{1}{2}(\mathcal{I}[M(k,t_{2i-1}] + \mathcal{I}[M(k,t_{2i}]) \to \mathcal{I}[M(k,t_j)]. \tag{411}$$

This can be understood as follows for $F(k,t)$ and for $M(k,t)$ the same explanation holds. $\mathcal{I}[F(k,t_j]$ is the average value of $F(k,t)$ over the interval $t_{j-1} \leq t \leq t_j$. In the previous period, this interval can be split up in $t_{2i-2} \leq t \leq t_{2i-1}$ and $t_{2i-1} \leq t \leq t_{2i}$, and the average values of these intervals are respectively $\mathcal{I}[F(k,t_{2i-1}]$ and $\mathcal{I}[F(k,t_{2i}]$.

For $N/4+1 \leq j \leq N/2$, we do not know the corresponding values of $\mathcal{I}[F(k,t_{2i}]$ and $\mathcal{I}[M(k,t_{2i}]$ and therefore we use the following approximation

$$\frac{1}{6}[F(k,t)_{2j} + 4F(k,t_{2j-1}) + F(k,t_{2j-2})] \rightarrow \mathcal{I}F(k,t_j) \tag{412}$$

$$\frac{1}{6}[M(k,t)_{2j} + 4M(k,t_{2j-1}) + M(k,t_{2j-2})] \rightarrow \mathcal{I}M(k,t_j). \tag{413}$$

E.3.7 Changes in the MCT-algorithm for the incoherent ISF

The work flow of the algorithm for the incoherent intermediate scattering function is the same as for the coherent intermediate scattering function in Figure 89, except that the coherent intermediate scattering function is needed as an extra input and therefore has to be calculated first. Some of the equations in the algorithm changes and this is discussed in this Appendix subsection. The mode coupling theory equation that we are solving is given by Equation 46 and becomes after normalizing with $F(k,0)$ and taking the overdamped limit:

$$\frac{\mathrm{d}F_s(k,t)}{\mathrm{d}t} + \Omega^2 F_s(k,t) + \int_0^t \mathrm{d}\tau M_s^{\mathrm{MCT}}(k,t-\tau)\frac{\mathrm{d}F_s(k,\tau)}{\mathrm{d}\tau} = 0, \tag{414}$$

with $\Omega^2 = k^2 D_0$ and

$$M_s^{\mathrm{MCT}}(k,t) = \frac{\rho D_0}{4\pi^2}\int \mathrm{d}\boldsymbol{q} \mid \tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} \mid^2 F(q,t)S(q)F_s(\mid \boldsymbol{k}-\boldsymbol{q} \mid,t), \tag{415}$$

with

$$\tilde{V}_{\boldsymbol{k}-\boldsymbol{q},\boldsymbol{q}} = (\hat{\boldsymbol{k}}\cdot\boldsymbol{q})c(q), \tag{416}$$

with

$$c(k) = \frac{1}{\rho}(1 - \frac{1}{S(k)}). \tag{417}$$

Following the derivations of the algorithm for the coherent intermediate scattering function, we make the following adaptations. The memory kernel becomes

$$M_s(k,t) = \frac{\rho D_0}{4\pi^2 k^2}(\Delta_k)^2 \sum_{q=0}^{q_{\max}} \sum_{p=|k-q|}^{|k+q|} \frac{pq\left((k^2+q^2-p^2)C(q)\right)^2}{\sqrt{4k^2q^2-(q^2+k^2-p^2)^2}} S(q)F(q,t)F_s(p,t). \tag{418}$$

The formula for $F_s(k,t)$ for small t is

$$F_s(k,t) = 1 - \Omega^2 t + \Omega^2(\Omega^2 + M_s(k,0))t^2 \tag{419}$$

and the formula to solve $M_s(k,t_i)$ and $F_s(k,t_i)$ simultaneously becomes

$$A_k F_s(k,t_i) + B_k^{t_i} = D_k M_s(k,t_i) + C_k^{t_i}, \tag{420}$$

with A_k, $B_k^{t_i}$, $C_k^{t_i}$ and D_k unaltered.

F Detailed explanation of the measurement methods

In Subsection 3.5 we already summarized the methods we used to analyze the simulated and experimental data and in this appendix subsection we explain the measurement procedures in detail for the VRD-model, the experimental data and the VV-model in respectively Appendix Subsection F.1, Appendix Subsection F.2 and Appendix Subsection F.3.

F.1 VRD-model (Section 4)

For the VRD-model we measure the mean squared displacement and diffusion constant, the average perimeter q, the scalar and vectorial radial distribution function, the scalar and vectorial static structure factor and the coherent and incoherent intermediate scattering function and their relaxation times. Below we list the details of these measurements.

- *Mean squared displacement and diffusion coefficient*: We use Equation 1 to calculate the MSD with the transformation $\boldsymbol{r}(t) \rightarrow \boldsymbol{r}'(t) = \boldsymbol{r}(t) - \langle \boldsymbol{r}(t) \rangle$ with $\langle \boldsymbol{r}(t) \rangle$ the average displacement of all cells in time t. The reason for the subtraction of this average displacement is that the average director is nonzero due to finite-size effects and this results in a non-physical displacement for the complete system for which we make a correction (see Appendix G.4). The effective diffusion coefficient is derived from the MSD via Equation 61, where we use $t_{\max} = 10^4\,\tau_0$, because this is our maximum measurement time.
- *Average perimeter q*: We use Equation 62 to calculate q.
- *Vectorial and scalar radial distribution function*: The vectorial RDF is given by Equation 17 and the scalar version by Equation 18 and algorithms to calculate them are respectively Algorithm 1 and Algorithm 2 in Appendix E.2. For the vectorial RDF we calculate a two-dimensional histogram for the interparticle vectors $\boldsymbol{r}$ with two times 70 bins between $-5.0\,r_0$ and $5.0\,r_0$ and for the scalar version we calculate a one-dimensional histogram for the interparticle distance $|\boldsymbol{r}|$ with 300 bins between $0.0\,r_0$ and $5.0\,r_0$.
- *Vectorial and scalar static structure factor*: There are two ways to calculate the static structure factor, namely via a Fourier transform and using direct calculations. Appendix G.1 shows that both ways give similar results and we choose to use the direct way. The formula to calculate the vectorial static structure factor directly is given by Equation 20 and the scalar version by Equation 21 and their algorithms are respectively Algorithm 3 and Algorithm 4 in Appendix E.2. Because we are measuring in a finite system of size LxL with periodic boundary conditions, we choose the wave numbers in the x- and y-direction (respectively k_x and k_y) to be multiples of $\frac{2\pi}{L}$. Although this is not necessary in a system with low structural order where the influences of the periodic boundaries on the structure can be neglected, in a highly structured system this is important because the structure is periodically in L and therefore only waves that fit exactly in the box can probe this structure. Furthermore, wavelengths larger than L are impossible, and therefore the results of $S(k)$ for $k < \frac{2\pi}{L}$ are unphysical. Thus, we calculate the static structure factor in Equation 20 and Equation 21 for respectively $\boldsymbol{k} = (k_x, k_y)$ and $k = |(k_x, k_y)|$ with $k_x = \frac{2\pi n_x}{L}$ and $k_y = \frac{2\pi n_y}{L}$, where we use $-\frac{15L}{2\pi} \leq n_x \leq \frac{15L}{2\pi}$ and $-\frac{15L}{2\pi} \leq n_y \leq \frac{15L}{2\pi}$. We bin the results for the vectorial static structure factor with bins of size $1.0\,r_0^{-1}$ and for the scalar static structure factor with bins of size $0.5\,r_0^{-1}$.
- *Coherent and incoherent intermediate scattering function and their relaxation times:* Both ISFs can also be calculated via a Fourier transform or via direct calculations. Following on the static structure factor, we use the direct way for the ISFs too, which is given by Equation 22 and Equation 24 for respectively the coherent and incoherent ISF. Similar as for the MSD we change from $\boldsymbol{r}(t)$ to $\boldsymbol{r}'(t) = \boldsymbol{r}(t) - \langle \boldsymbol{r}(t) \rangle$ to correct for the average displacement of the complete layer (see Appendix G.4). We implement Algorithm 5 and Algorithm 6 of Appendix E.2 for respectively the coherent and incoherent ISF in which we only take into account wave vectors that are within $0.25\,r_0^{-1}$ of the peak of the static structure factor $\overline{k}$. This results in different $\overline{k}$-values for different values of p_0 and v_0. Another option is to always analyze the ISF at the same value of k, but this has as drawback that there are combinations of p_0 and v_0 at which one probes a valley instead of a peak in the SSF. For this reason we choose the first method to determine k. Nevertheless, in Appendix G.5 we checked that both methods give similar results. We define the relaxation time as the time it takes until the coherent or incoherent ISF reaches a value of 0.1.
- *Mode coupling theory*: The MCT-algorithm to predict the coherent ISF is explained in Appendix E.3.1 and the system specific input consists of the static structure factor $S(k)$, the self-diffusion coefficient D_0 (given by Equation 60) and the density ρ, which is always $1.0\,r_0^{-2}$ in our models. Because for small values of k the static structure factor goes to N due to the finite size and periodic boundary conditions, we modify

the static static structure factor by setting $S(k) = 0$ for the smallest k-values. Furthermore, we cut-off the static structure factor at $k = 30\, r_0^{-1}$ because in all simulations the SSF is steady at a value of one at this k-value. Our MCT-algorithm settings are as follows. The initial time step size is set to $\Delta t = 10^{-6}\tau_{\text{MCT}}$, in which τ_{MCT} is a time unit of MCT (see Subsection 2.3.3) and we use $N_T = 64$. Our maximum simulation time is $t_{\max} = 10^{20}\tau_{\text{MCT}}$, but in practice we see that $F(k,t)$ has reached zero within machine accuracy far before $t_{\max}$. Furthermore, we use as threshold for the memory kernel $\epsilon = 10^{-20}$ and this threshold is always reached within 20 steps. Adaptations to the MCT-algorithm for the ISF are explained in Appendix E.3.7 and we use the same input and algorithm settings as for the coherent ISF except that this coherent ISF predicted with MCT serves as extra input. The output of the MCT-algorithm is $F(k,t)$ (or $F_s(k,t)$) and we define the predicted relaxation time τ_α to be the time at which the predicted $F(\bar{k},t)$ reaches 0.1 with $\bar{k}$ the position of the first peak in the static structure factor.

F.2 Experimental data (Section 5)

We perform the same analysis in the experimental data as in the VRD-model, which includes the mean squared displacement and diffusion constant, the average perimeter q, the scalar and vectorial radial distribution function, the scalar and vectorial static structure factor and the coherent and incoherent intermediate scattering function and their relaxation times. As explained in Subsection 3.5 there are a few differences between the experimental and the simulated data which we have to take into account and we explain them in more detail here.

First, the experimental data show cell proliferation, cell death, cells escaping the FOV and cells entering the FOV, which all leads to a variation of cell density over time resulting in a varying number of cells in the FOV. To account for this, we scale all lengths at waiting time T with $r_0(T)$ according to

$$r_0(T) = \sqrt{\frac{A}{N(T)}}, \tag{421}$$

with A the area of the snapshot image in pixels and $N(t)$ the number of cells in the snapshot at time t. Although time dependent, this scaling is similar to the scaling in the VRD-model, where the square root of the average area of each cell is used. We investigate the effect of this scaling in Appendix G.9.

Secondly, the experimental data is not in a steady state which means that we cannot average over many snapshots to reduce the noise. Nevertheless, we assume the changes in the system to be negligible during one hour such that we can take the average over all snapshots within this hour. We apply a running average where we define the value of a quantity at waiting time T as the average value of that quantity between waiting time T and $T + 1$ hour. In Appendix G.8 we show how much the noise depends on this averaging time.

Thirdly, we do not have periodic boundary conditions and lastly, the positions can only be determined with an accuracy of one pixel. How this influences the procedure to calculate the correlation functions and other analysis parameters is explained below, where we summarize the measurement methods as well.

- *Mean squared displacement and diffusion coefficient*: The mean squared displacement is calculated via Equation 1 and there is no need to subtract the average displacement as is done in the VRD-model. We measure the MSD during 300 min and calculate the diffusion coefficient by Equation 2 with $t_{\max} = 300$ min. Because fast moving cells that escape the FOV are not taken into account, the calculated MSD and diffusion coefficient might be slightly too low.
- *Average perimeter q*: The perimeter is calculated both using the boundaries from the phase image and using the Voronoi tessellation. In the latter, cells within $2\, r_0$ from the boundary are not taken into account because these Voronoi cells are incorrect due to the non periodic boundary conditions. Error bars define the standard deviation in the average perimeter.
- *Vectorial and scalar radial distribution function*: The vectorial radial distribution function is calculated via

$$g(\boldsymbol{r}) = \frac{1}{\rho_0 N_P} \left\langle \sum_{i=1, i\in P}^{N} \sum_{j\neq i, j\in Q}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_{ij}(t)) \right\rangle_t, \tag{422}$$

  where Q is the complete area of the FOV and P the inner square within L_c from the boundary. N_P is the number of cells within this area P. This is a simple modification to Equation 17 to account for the non-periodic boundary conditions. Nevertheless, noise might be further reduced by taking the cells

near the boundary (cells in $Q - P$) into account in the sum of i using a weight factor. The scalar radial distribution function of Equation 18 is modified in a similar way as Equation 422. A second modification to $g(\boldsymbol{r})$ and $g(r)$ is made by changing the area in the denominator of Algorithm 1 and 2 (respectively given by h^2 and $V(h)$) to the number of grid points within that area divided by the grid point density. This is done because the previous method results in a wrong normalization due to the discrete values of the positions of the nuclei. We set the bin size to $\frac{3}{45}\, r_0$, which is chosen such that for the highest density the unscaled bin size is still larger than 1 pixel. Furthermore, we use $L_c = 3\, r_0$. We take the average over all RDFs within one hour weighted by the number of cells and the errorbars define the standard deviation of the RDF within this hour.

- *Vectorial and scalar static structure factor*: The vectorial and scalar static structure factor are calculated via respectively Equation 20 and 21 and the algorithms are respectively Algorithms 3 and 4. We tested in Appendix G.10 if it makes a difference to take into account the discrete values of the position of the nuclei in the choice of vectors $\boldsymbol{k}$ at which the SSF is calculated. This modification means that we only use $\boldsymbol{k}$- or $|\boldsymbol{k}|$-values that satisfy $k_x = \frac{2\pi}{n_{x,1}} n_{x,2}$ and $k_y = \frac{2\pi}{n_{y,1}} n_{y,2}$ with $n_{x,1}$, $n_{x,2}$, $n_{y,1}$ and $n_{y,2}$ integers. However, we find that these wave vectors did not reduce the noise and therefore we apply the same method to choose the vectors $\boldsymbol{k}$ as used for the VRD-model described in Appendix Subsection F.1. Furthermore, we use a bin size of $0.5\, r_0^{-1}$, which is a trade-off between noise reduction and preservation of the shape of the SSF. We take the average over all SSFs within one hour weighted by the number of cells and the errorbars define the standard deviation of the SSFs within this hour.
- *Coherent and incoherent intermediate scattering function and their relaxation times*: The coherent and incoherent intermediate scattering function are respectively given by Equation 22 and 24 and there is again no need to subtract the collective displacement. In analogy to the VRD-model, we calculate the coherent and incoherent ISF at the location $\overline{k}$ of the first peak in the SSF. The algorithms are respectively given by Algorithm 5 and 6, where we use the same sampling of wave vectors as explained for the static structure factor. To calculate the coherent ISF starting at time t, we scale all positions with $r_0(t)$ given by Equation 421. For the incoherent ISF starting at time t, we only take into account all cells that stay in the FOV during the whole period of which the ISF is measured. This means that we neglect some cells that escape the FOV or enter the FOV, proliferate, die or get missed in the nuclei detection. We tested this approach of only taking into account the cells that stay for the coherent ISF as well and Appendix G.11 shows that the results are similar to taking into account all detected cells at time t. We calculate the coherent and incoherent ISF over a time period of $\tau = 240$ min and use a bin size of $0.5\, r_0^{-1}$.
- *Mode coupling theory*: The MCT predictions for the coherent and incoherent ISF and their corresponding relaxation times are obtained in a similar way as described in F.1 for the VRD-model and we use the same algorithm settings. The static structure factor $S(k)$ is used as input and we set $D_0 = 1\, r_0^2/\tau_0$ as we do not know its value D_0. Therefore we have to keep in mind that the scaling of the relaxation time τ_α is arbitrarily as a different D_0 results in a different scaling (see Appendix D.2).

F.3 VV-model (Section 6)

For the parameter analysis in the VV-model, we measure alignment, the correlation length L_{corr}, the average perimeter q and the average speed. Below we list the details of these measurements.

- *Alignment*: In Subsection 3.3.2 we have defined three alignment parameters, namely the Vicsek parameter (VP), the alignment index (AI) and the instantaneous order parameter (IOP). In Appendix G.6 is shown that these parameters give similar results and therefore we only look at one of them. Because the implementation of VP given in Equation 69 cannot be extended to experiments where v_0 is unknown and AI shows a broader range of values than IOP (Equation 71), we choose to use AI which is given by Equation 70 and we emphasize that we the instantaneous velocity $\boldsymbol{v}_i$ is used. We measure AI as a function of time and look as well at the steady state value $\text{AI}_{t\to\infty}$ and the alignment time τ_{AI}, which is defined as the time it takes to reach AI= 0.5 and is undefined if AI never reaches this threshold.
- *correlation length L_{corr}*: The correlation length L_{corr} is defined in Equation 72 and calculated as follows. The maximum distance between two particles in a box of size $\sqrt{1000} \times \sqrt{1000}$ is $22.36\, r_0$ and we divide this into fifty bins. For every particle combination i and j we calculate the difference in velocity angle ϕ_{ij} and the distance r_{ij}. This angle ϕ_{ij} is then added to the bin where r_{ij} to belongs. After doing this for all runs, we calculate the mean μ and the standard deviation σ of the list of angles belonging to a bin. The correlation length is now defined as the largest bin for which $\mu + 5\sigma/N_{\phi_r} \leq \pi/2$, with N_{ϕ_r} the number of angles in that bins.

- *Average perimeter q*: We use Equation 62 to calculate q and use this parameter to get insights into the dynamics of the system, because other methods, such as the diffusion coefficient D_{eff} and the relaxation time τ_α of the ISF, have high computational costs. However, it is not guaranteed that q gives correct information about the dynamics as we already saw in Section 4 and therefore we keep in mind that the conclusions have to be verified with the relaxation time or diffusion coefficient.
- *Average speed*: The instantaneous velocity might not give a correct interpretation of the velocity of the cells. For example, when a cell moves at time t a distance $v\Delta t$ in one direction and at time $t + \Delta t$ the same distance in the opposite direction, the instantaneous velocity has a finite value v while the cells are barely moving. Therefore we define another measure for the speed, namely

$$\overline{v} = \frac{1}{N} \sum_{i=1}^{N} \frac{|\boldsymbol{r}_i(t + \Delta t') - \boldsymbol{r}_i(t)|}{\Delta t'}, \tag{423}$$

  where $\Delta t' >> \Delta t$ and we choose $\Delta t' = 1.0\,\tau_0$.

For the jamming analysis in the VV-model we measure the alignment index as described above and the scalar RDF, scalar SSF and the coherent and incoherent ISF and their relaxation times in a similar way as described for the VRD-model in Appendix Subsection F.1. In addition, we also measure the non-ergodicity parameter, which we define as the average value of the ISF between $t = 0.9 \cdot 10^4 \tau_0$ and $t = 1.0 \cdot 10^4 \tau_0$.

G Supportive data to verify the measurement choices

In this appendix section we verify the most important simulation and measurement choices. We start with comparing the Fourier and direct method to calculate the static structure factor in Appendix Subsection G.1. In Appendix Subsections G.2 and G.3 we test respectively the step size and the initialization time used in the VRD- and VV-model. In both models we subtracted the collective displacement to calculate the MSD and the ISF and in Appendix Subsection G.4 we explain why. In Appendix Subsection G.5 we investigate how the choice of k at which the ISF is probed influences the results and in Appendix Subsection G.6 we compare three alignment parameters for the VV-model. In Appendix Subsection G.7 till G.11 we investigate the effect of several properties of the experimental data and choices made to analysis them, including the accuracy of the nuclei detection, the averaging, the scaling, the discrete grid for the cell positions and the tracking of nuclei.

G.1 Fourier and direct calculations of the SSF and ISF

There are two ways to calculate the static structure factor, namely via a Fourier transform of the radial distribution function (Equation 21) or via direct calculations (Equation 141). Figure 91 shows the SSF via direct and Fourier calculations for three different parameter settings in the VRD-model, representing a jammed sate, a state near the jamming transition and an unjammed state. Differences between both methods are the peak at small k for the direct method, which is an artifact of the algorithm, and the oscillations of the Fourier results for small k due to oscillations of $g(r)$ about 1 for big r. Figure 92 shows the SSF via direct and Fourier calculations for two different snapshots of the experimental data used in Section 5. One of the snapshots is chosen in the beginning of the time window and one at the end. Although the noise is higher than for simulated data, the direct and Fourier method give similar results too. From these figures we conclude that we can use both methods to calculate the static structure factor. Because the direct way is normally applied in literature, we choose to use this method as well. In addition, the intermediate scattering function can also be calculated via a direct way or via a Fourier transform and because both methods hold no difference for the SSF, we assume that this statement is also valid for the ISF and we therefore use the direct method in analogy with the SSF.

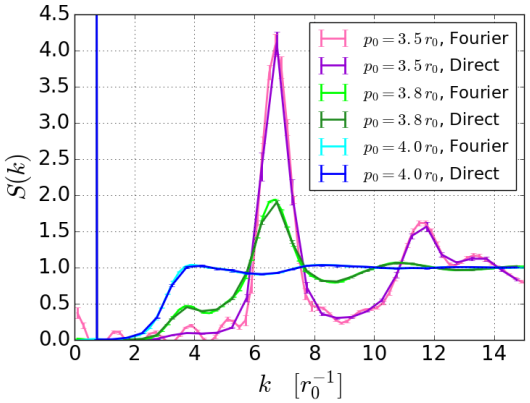

Figure 91: Comparison between the direct way and the Fourier transform to calculate the SSF in the VRD-model. We use $v_0 = 0.25\,r_0/\tau_0$ and three values of p_0, namely $3.5\,r_0$, $3.8\,r_0$ and $4.0\,r_0$.

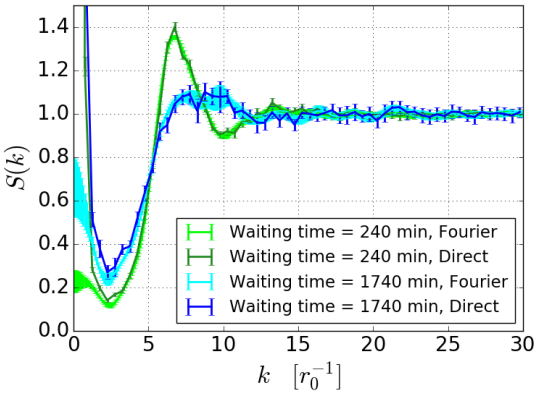

Figure 92: Comparison between the direct way and the Fourier transform to calculate the SSF for experimental data after 240 and 1740 minutes of waiting time.

G.2 Simulation settings VRD- and VV-model: step size

In MD simulations it is important to choose the time step size wisely because a small step size results in long simulation times and a big step size might lead to incorrect results. In Figure 93 we show the effect of the choice of the step size on the MSD in two situations for the VRD-model. In the first situation (left panel) p_0 is chosen large such that the dynamics is fast and in the second situation (right panel) p_0 is chosen such that the system is in a jammed state and the dynamics is slow. Because the results for $\Delta t = 0.001\,\tau_0$, $\Delta t = 0.1\,\tau_0$ and $\Delta t = 1.0\,\tau_0$ are not averaged over multiple runs to save time, errorbars are not shown.

We see that $\Delta t = 1.0\,\tau_0$ gives incorrect results, while the lines for the other three step sizes lay close to each other. When the dynamics is fast (left panel), $\Delta t = 0.1\,\tau_0$ gives a slightly higher MSD than $\Delta t = 0.01\,\tau_0$ and $\Delta t = 0.001\,\tau_0$, while this difference is not visible in the jammed state (right panel). Because we want to apply the same step size in every simulation, we choose $\Delta t = 0.01\,\tau_0$ in our simulations.

We assume that the effect of the step size in the VV-model is similar to the VRD-model, and we have tested the extreme cases (results not shown). From this we concluded that we have to be cautious in choosing the minimum values of τ_V and the maximum value of D_r, because both extremes are limited by the choice of Δt.

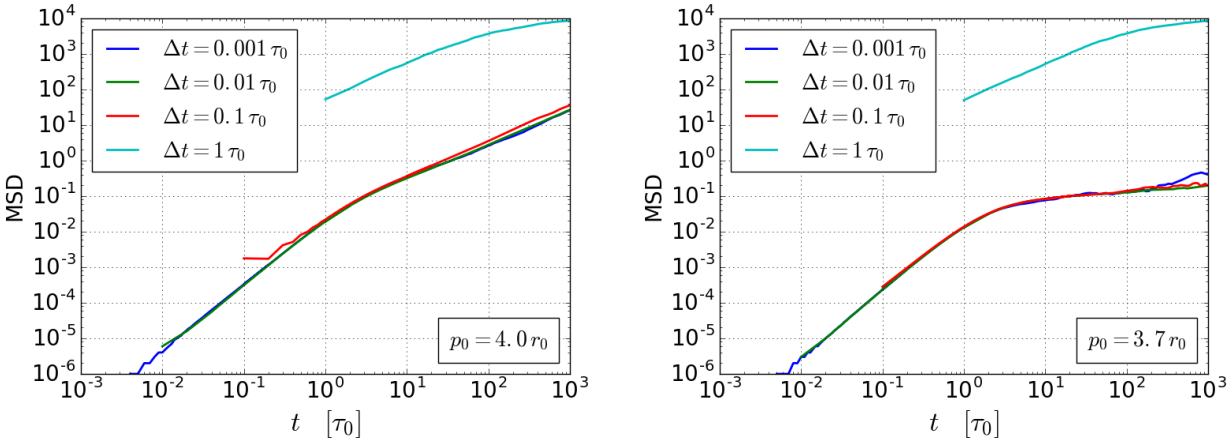

Figure 93: MSD for $\Delta t = 0.001\,\tau_0$, $\Delta t = 0.01\,\tau_0$, $\Delta t = 0.1\,\tau_0$ and $\Delta t = 1.0\,\tau_0$ and with $p_0 = 4.0\,r_0$ in the left panel and $p_0 = 3.7\,r_0$ in the right panel ($v_0 = 0.25\,r_0/\tau_0$).

G.3 Simulation settings VRD- and VV-model: initialization time

Just as the time step size, the initialization time also has to be chosen wisely, because the simulations are very expensive when this time is too long and the system might not be in a steady state yet when the initialization time is too short. Figure 94 shows the potential and the kinetic energy for the VRD-model with the same two parameter settings as the MSD in Figure 93 in the previous appendix subsection. The potential energy E_p is the same as in Equation 65 and given by

$$E_p(t) = \sum_{i=1}^{N} k_a(a_i(t) - a_0)^2 + k_p(p_i(t) - p_0)^2 \tag{424}$$

and the kinetic energy E_k is defined as

$$E_k(t) = \frac{1}{2}\sum_{i=1}^{N} |\boldsymbol{v}_i(t)|^2, \tag{425}$$

with $\boldsymbol{v}_i(t)$ the instantaneous velocity at time t. We use an initialization time of $10^4\,\tau_0$ for the VRD-model and this is shown by the red line in Figure 94. Because for both settings the potential and kinetic energy have reached their minimum before the red line, we conclude that in these cases an initialization period of $10^4\,\tau_0$ is sufficient.

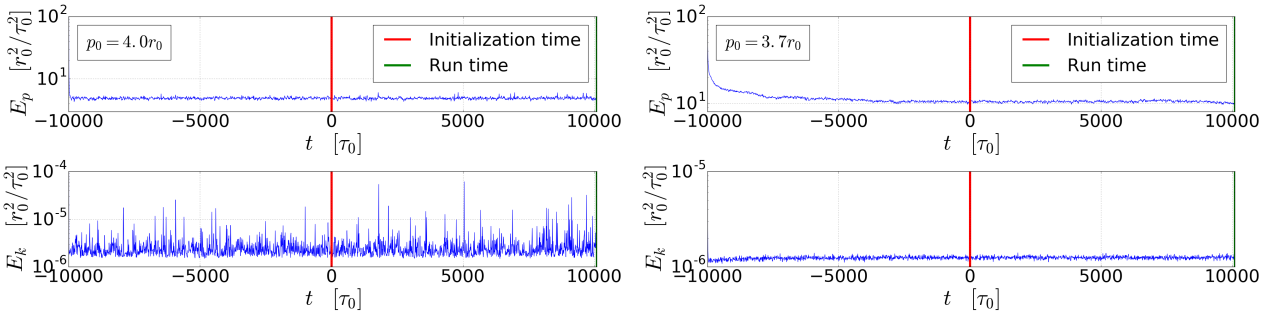

Figure 94: The potential (top panels) and kinetic energy (bottom panels) as a function of the time, with $p_0 = 4.0$ in the left two panels and $p_0 = 3.7\, r_0$ in the right two panels ($v_0 = 0.25\, r_0/\tau_0$). The red line defines the end of the initialization period.

For the alignment analysis in the VV-model we use a shorter initialization period, namely $t = 200\, \tau_0$, to save time. In the VV-model we turn on the self-propelled force after the initialization period, which causes a jump in both the potential and the kinetic energy. Therefore, we assume that the fact that we did not reach the absolute minimum in energy after the initialization period does not influence the results. In Figure 95 the potential and kinetic energy for Setting A and varying R_V are shown and the jump in energies after the initialization period is clearly observed. However, there are exceptions in which the activation of the self-propulsion force can lower the potential energy and in these situations a longer initialization time is used.

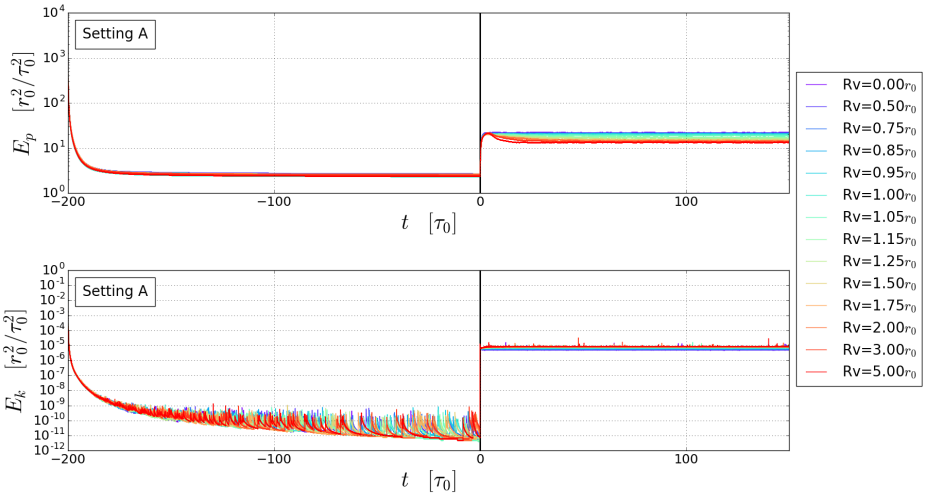

Figure 95: The time evolution of the potential (top) and kinetic energy (bottom) in the VV-model for Setting A in Table 1 and R_V varying from $0.0\, r_0$ till $5.0\, r_0$. The black line defines the end of the initialization period.

G.4 Correction for collective displacement in the VRD- and VV-model

To calculate the mean squared displacement and the intermediate scattering function in the VRD- and VV-model, we subtract the collective displacement from the individual displacements. The reason for this is that due to finite size effects the system moves as a collective in addition to the individual particle movements. Figure 99 shows an example of the trajectories of several cells in a highly jammed system ($p_0 = 3.5\, r_0$ and $v_0 = 0.1\, r_0/\tau_0$). The similarity in the trajectories of these cells is caused by the finite size effects and can be explained as follows. The directors of all particles are independently distributed with values between -1 and 1 and standard deviation σ in the x-direction and in the y-direction. According to the central limit theorem, the sum of the values in one of the Cartesian directions approaches a Gaussian distribution with mean $n\mu = 0$ and standard deviation $\sigma\sqrt{N}$. This results in an average director in the x- and y-direction with standard deviation $\sigma/\sqrt{N}$, which is finite when N is finite. Hence, this average director moves the cellular collective in the average direction of the director. The effect of this movement decreases when system size N increases.

In Figure 97 the influence of this collective movement is shown for the coherent intermediate scattering function. Only in the left panel the collective motion is subtracted and we see that we obtain glassy states for p_0 small

enough. While the right panel suggests that there are no glassy states because of the effect of the collective movement.

In the VV-model it is obvious that not subtracting the collective behavior highly influences the intermediate scattering function and the mean squared displacement because in aligned situations the cells can move as a flock. When we do not subtract the collective behavior the relaxation time or diffusion coefficient might tell us that a layer is unjammed in cases that we define as a jammed flock.

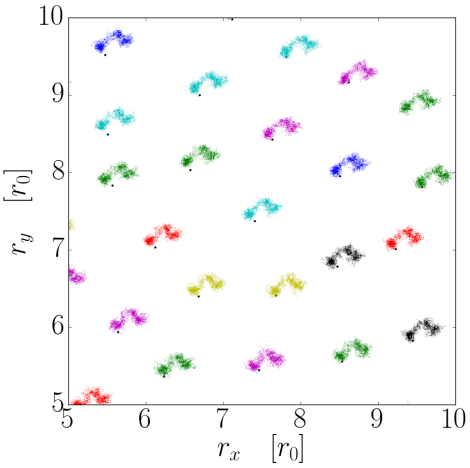

Figure 96: The trajectories of several cells during $t = 10^3\,\tau_0$ for a highly jammed state ($p_0 = 3.5\,r_0$ and $v_0 = 0.1\,r_0/\tau_0$).

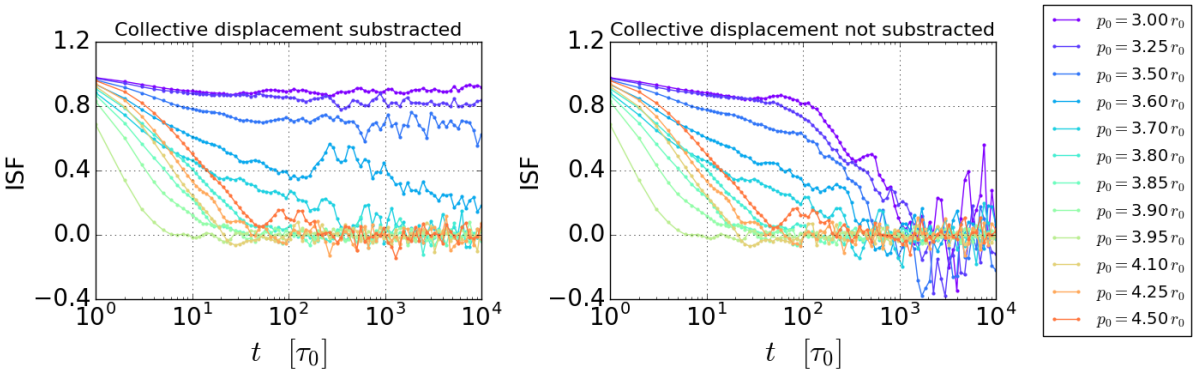

Figure 97: The differences in the coherent intermediate scattering function when corrected for the collective displacements (left panel) and not corrected (right panel) for $v_0 = 0.25\,r_0/\tau_0$ and p_0 varying from $3.0\,r_0$ till $4.5\,r_0$.

G.5 Effect of the choice of the k-value for the ISF

To evaluate the intermediate scattering function we need to choose a k-value at which we probe this ISF. Normally, the location $\overline{k}$ of the peak in the static structure factor is chosen because the relaxation time is longest for $\overline{k}$. However, Figure 41 shows that this location of the peak highly depends on p_0 in the VRD-model and this give rise to two ways to choose k for our analysis. The first method is to define the peak value $\overline{k}$ for every p_0 separately and calculate $F(\overline{k}, t)$. The second method is to choose one k in general and probe, for example, $F(6.75, t)$ for all parameter settings. The relaxation times calculated for both methods are shown in Figure 98 and 99 for respectively the coherent and incoherent intermediate scattering function. For higher p_0 values the relaxation times are different, but the trends are the same. Therefore, we choose to use $F(\overline{k}, t)$ because of the physical motivation to choose k as the location of the first peak of the SSF.

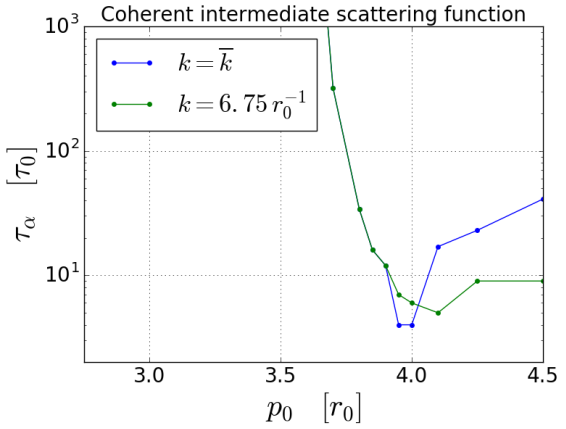

Figure 98: The relaxation time obtained for the coherent intermediate scattering for two different ways to choose k, with $v_0 = 0.25\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$.

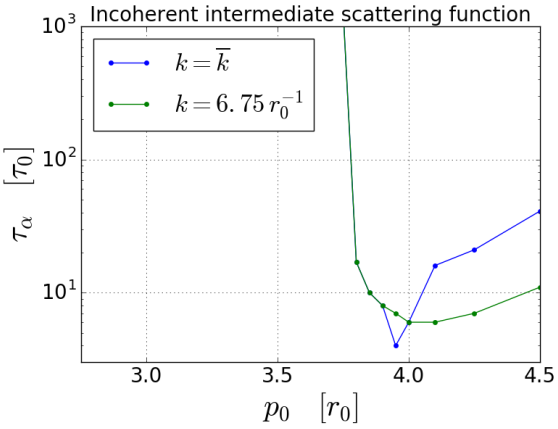

Figure 99: The relaxation time obtained for the incoherent intermediate scattering for two different ways to choose k, with $v_0 = 0.25\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$.

G.6 Similar behavior of the three alignment parameters in the VV-model

In Subsection 3.4.2 three different alignment parameters are defined that are used in literature. In this appendix subsection we show that in the VV-model all three alignment parameters give similar behavior except the Vicsek parameter in some cases. Figure 100 shows this comparable behavior as a function of five parameters for Setting *A*. The only large difference is seen when v_0 is used as control parameter in the third graph of Figure 100. The Vicsek parameter decreases to zero for large values of v_0, while the other alignment parameters stay at a finite value. The reason for this deviation of VP is that v_0 is explicitly taken into account in the formula of VP (see Equation 69). This explicit use of v_0 is a drawback of VP because this value cannot be determined in experiments and this reduces our choice to either AI or IOP. Because the range of the AI-values is broader than that of the IOP-values, we choose to use this alignment index in our analysis.

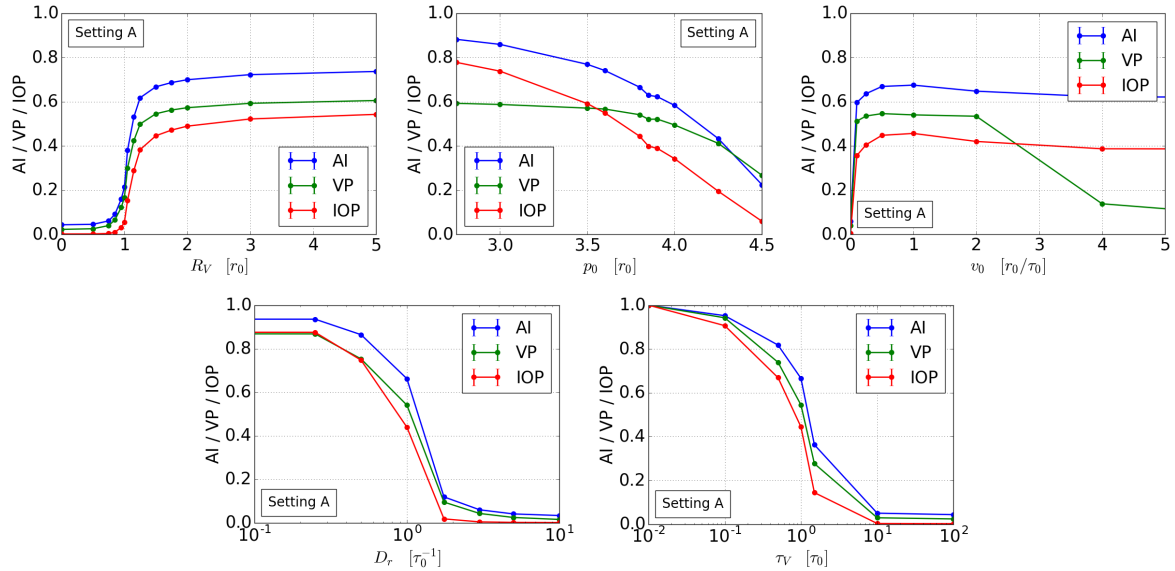

Figure 100: The three different alignment parameters (AI, VP and IOP) as a function of the five different input parameters for Setting A in Table 1.

G.7 Experimental data: effect of the accuracy of nuclei detection

In the currently used imaging method to analyze the snapshots of the MDCK-cells not all nuclei are detected and by comparing the located nuclei with the raw images, the percentage of missed nuclei has an estimated value of 5. To test how our results depend on this accuracy we evaluate in Figure 101 the relaxation time of the

coherent ISF in cases were an additional 1, 2, 5 or 10 percent of the cells stay undetected, where the undetected cells are chosen randomly from the data set. All five graphs show the growth in relaxation time during aging, however the fluctuations become less when more cells are detected. This suggests that we can reduce some noise by improving the nuclei detection up to one hundred percent.

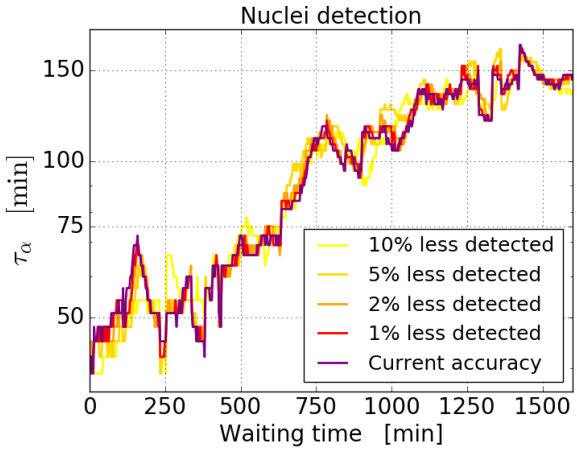

Figure 101: Relaxation time of the coherent intermediate scattering function for the current accuracy of the nuclei detection and for 1 %, 2 %, 5 % and 10 % less nuclei detected.

G.8 Experimental data: effect of averaging

In our analysis of the experimental data, we chose to use an average of twenty snapshots, which equals one hour. In general, averaging over a higher number decreases the noise, but in the case of experiments on living cells the system changes significantly when averaging over a too long period and a trade-off has to be made. Figure 102 shows the relaxation time for the coherent intermediate scattering function averaged over 1, 10, 20 and 40 snapshots and suggests that 20 snapshots is a good trade-off.

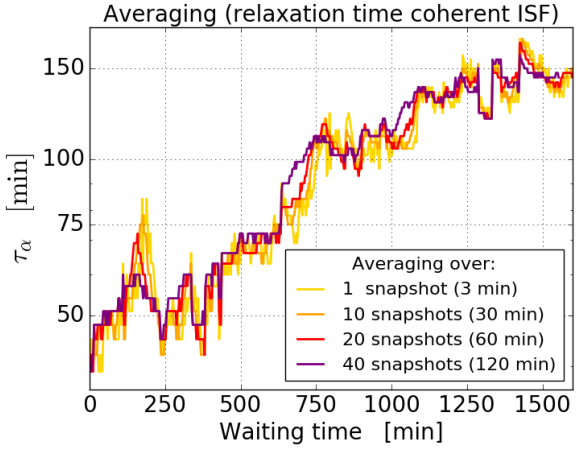

Figure 102: Relaxation time of the coherent intermediate scattering function averaged over 1, 10, 20 and 40 snapshots.

G.9 Experimental data: effect of scaling according to the density

In our analysis of the experimental data in Section 5 we have scaled the data according to the density to be able to compare the data at different waiting times and densities. However, for the analysis of the dynamics it should not make a difference whether we apply this scaling as long as we are probing at the same inverse length scale k. This can be seen, for example, in the formula for the coherent intermediate scattering function:

$$F(k,\tau) = \frac{1}{N}\left\langle \left(\sum_{i=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{r}_i(t))\right)\left(\sum_{j=1}^{N}\cos(\boldsymbol{k}\cdot\boldsymbol{r}_j(t+\tau))\right) + \left(\sum_{i=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{r}_i(t))\right)\left(\sum_{j=1}^{N}\sin(\boldsymbol{k}\cdot\boldsymbol{r}_j(t+\tau))\right)\right\rangle_t , \tag{426}$$

where we scale $\boldsymbol{r}$ with $r_0(t)$ and $\boldsymbol{k}$ with $r_0^{-1}(t)$. This results in a product that does not depend on the scaling factor. In Figure 103 we show the relaxation time of the coherent intermediate scattering function for both the unscaled and scaled method. The trend is comparable, but the exact values are different. This is due to the fact that the $\boldsymbol{k}$-sampling and bin size slightly differs in both methods. In the next subsection, we look further into the effect of this $\boldsymbol{k}$-sampling.

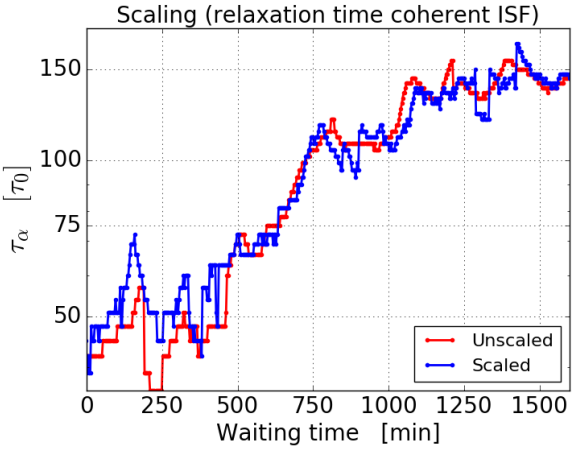

Figure 103: Difference between the relaxation time of the coherent ISF when we do or do not apply a scaling according to the density.

G.10 Experimental data: effect of the discrete grid of the cell positions

The accuracy of the data is limited by the pixel size and we take this into account by calculating the correlation functions in real space, but neglect this discrete grid by the correlation functions in Fourier space. In Figure 104 we have compared both methods for the relaxation time of the coherent intermediate scattering function to test whether this neglect is legitimate. Both graphs show noise and the data points are slightly different, but the overall trend is equal. From this we conclude that the neglect of this discrete grid does not influence the results regarding the experimental data in Section 5. However, the graph in Figure 104 suggests that we might reduce some of the noise when we average over even more $\boldsymbol{k}$-vectors than we did to obtain the results in Section 5.

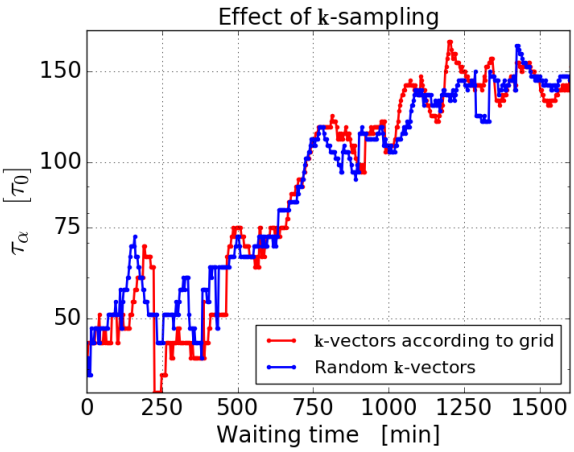

Figure 104: Difference between the relaxation time of the coherent ISF when we do or do not take into account the discrete grid for the positions of the cells.

G.11 Experimental data: effect of nuclei tracking on the coherent ISF

For the incoherent intermediate scattering function we only take into account the cells that are in the field of view during 300 minutes. However, this cell tracking is not necessary when calculating the coherent intermediate scattering function and this gives us the choice whether we take into account all cells or only the cells we use for the incoherent ISF. Figure 105 shows the relaxation time of the coherent ISF for both methods. The results are comparable and we choose to take into account all cells.

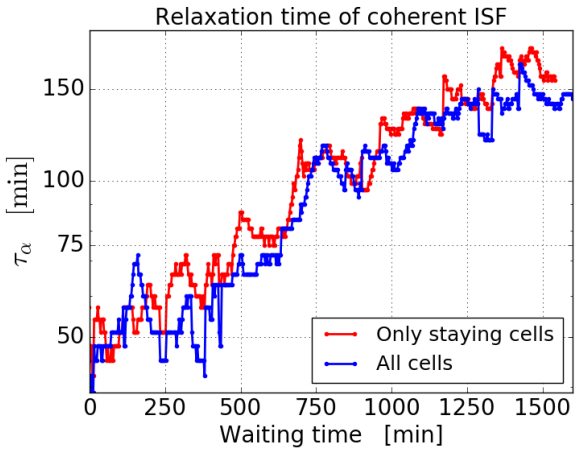

Figure 105: Difference between the relaxation time of the ISF when all cells are taken into account or only the cells that stay in the field of view during 300 minutes.

H Additional results for the VRD-model in Section 4

In this appendix section we show some additional results for the VRD-model. First, in Appendix Subsection H.1 we show that the layer exhibits a hexagonal structure when values of p_0 and v_0 are chosen close to the jamming transition. In Appendix Subsection H.2 we give the RDF, SSF and coherent and incoherent ISF for $v_0 = 0.1\, r_0/\tau_0$, $v_0 = 0.5\, r_0/\tau_0$ and $v_0 = 1.0\, r_0/\tau_0$ and in Appendix Subsection H.3 we give the peak height of the SSF as a function of p_0 and v_0.

H.1 A hexagonal structure near the jamming transition

When the VRD-model is highly jammed, we can observe a high structural order in the layer in which the cells are almost hexagons. Because we average over ten runs and the orientation of the hexagons is different in each run, this hexagonal pattern is not clearly visible in Figures 39 and 40. However, Figures 106 and 107 show respectively the RDF and SSF for only one run and this clearly indicates this hexagonal structure for $p_0 = 3.7\, r_0$ and $v_0 = 0.25\, r_0/\tau_0$.

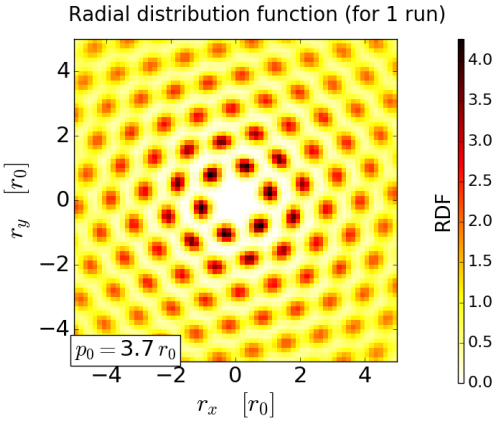

Figure 106: The radial distribution function for $p_0 = 3.7\, r_0$ and $v_0 = 0.25\, r_0/\tau_0$ shows a hexagonal pattern.

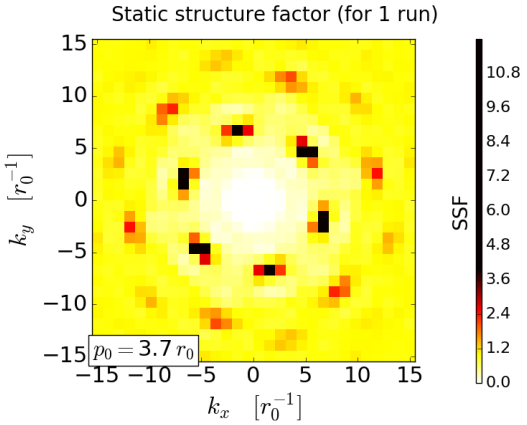

Figure 107: The static structure factor for $p_0 = 3.7\, r_0$ and $v_0 = 0.25\, r_0/\tau_0$ shows a hexagonal pattern.

H.2 RDF, SSF and ISF for $v_0 = 0.1\, r_0/\tau_0$, $v_0 = 0.5\, r_0/\tau_0$ and $v_0 = 1.0\, r_0/\tau_0$

In Subsection 4.2 and 4.3 we have respectively investigated the structural and dynamical changes upon varying p_0 for $v_0 = 0.25\, r_0/\tau_0$. In this appendix subsection we show the RDF, SSF and ISFs for three other values of the velocity ($v_0 = 0.1\, r_0/\tau_0$ in Figure 108, $v_0 = 0.5\, r_0/\tau_0$ in Figure 109 and $v_0 = 1.0\, r_0/\tau_0$ in Figure 110) and conclude that an increase in v_0 can decrease the structural order and shifts the jamming transition to lower values of p_0. Furthermore, we conclude that the qualitative changes upon varying p_0 in the RDF, SSF and ISFs are similar for different values of v_0. The graphs in this appendix subsection are used to obtain the direct calculations of the relaxation times in Figure 44.

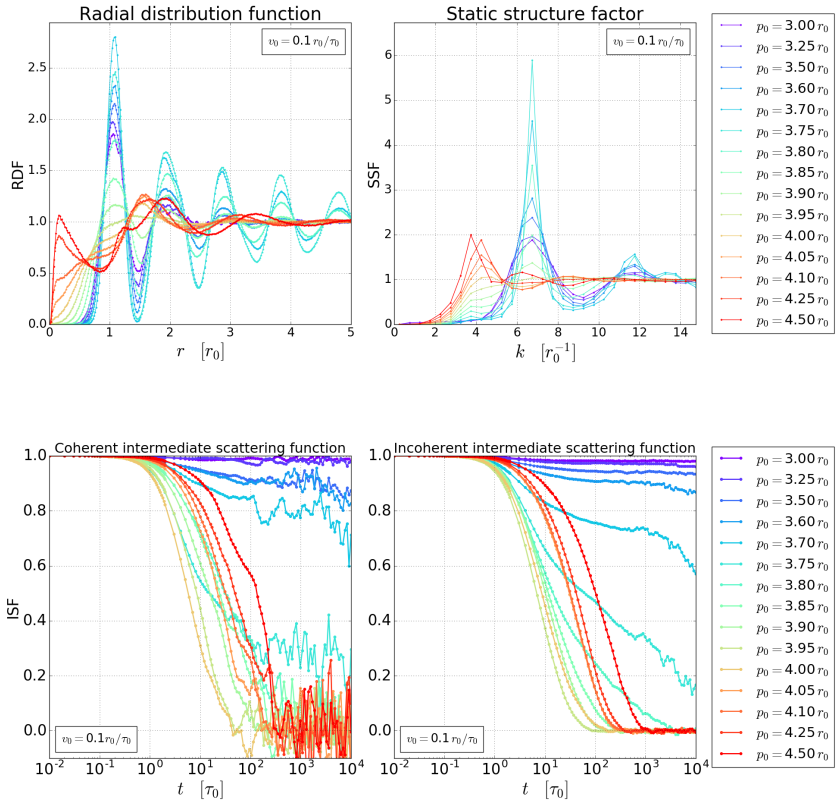

Figure 108: Correlation functions for $v_0 = 0.1\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$.

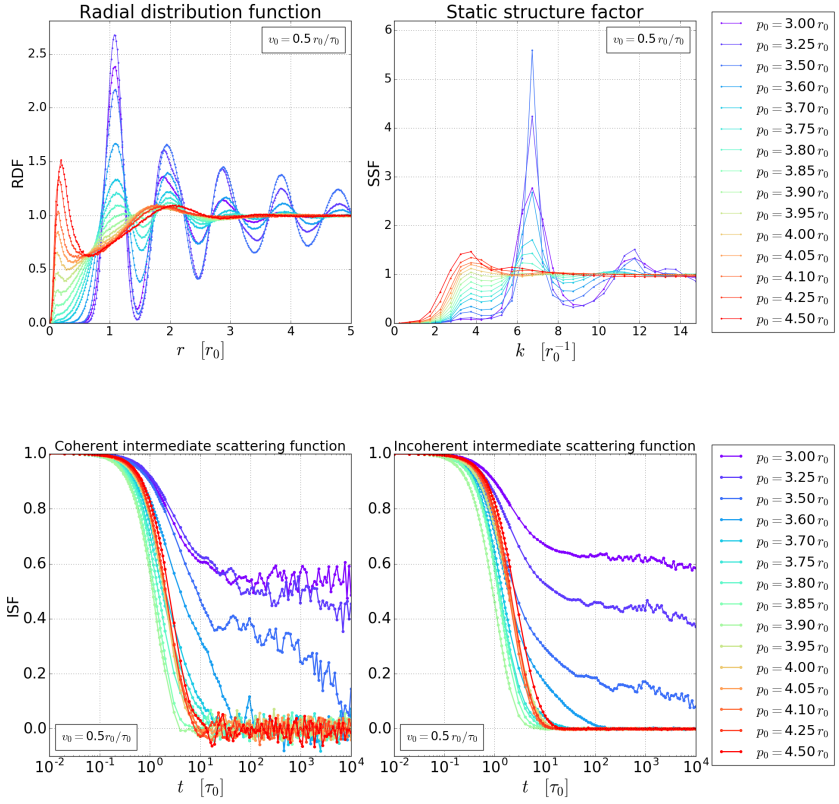

Figure 109: Correlation functions for $v_0 = 0.25\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$.

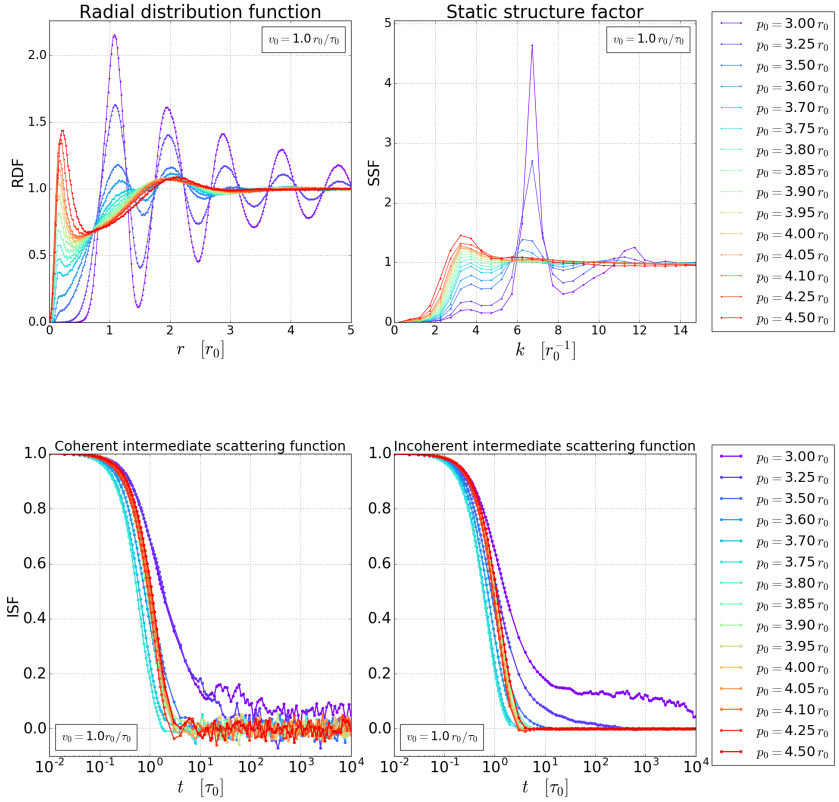

Figure 110: Correlation functions for $v_0 = 1.0\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$.

H.3 Peak height of the static structure factor

Figure 111 shows the height of the first peak of the SSF as a function of p_0 and v_0 in the VRD-model to compare with MCT. These values are derived from the SSFs given in Figures 41 (in Subsection 4.2), 108, 109, and 110 (in the previous appendix subsection), and the comparison with the MCT-predictions is discussed in Subsection 4.4.

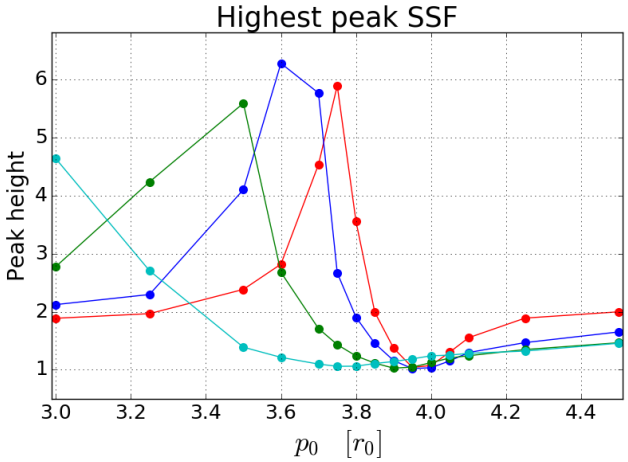

Figure 111: Peak height of SSF for different $v_0 = 0.1\, r_0/\tau_0$, $v_0 = 0.25\, r_0/\tau_0$, $v_0 = 0.5\, r_0/\tau_0$ and $v_0 = 1.0\, r_0/\tau_0$ and p_0 varying from $3.0\, r_0$ till $4.5\, r_0$

I Additional results for the experimental data in Section 5

In this appendix section we give two additional analysis regarding the experimental data in Section 5. First, in Appendix Subsection I.1 we look at the height of the first peak as a function of the density and second, in Appendix Subsection I.2 we explain how we can determine confluence based on the RDF and SSF.

I.1 Peak height versus density

In Subsection 5.2 we raised the questions whether we can compare the experimental layer to physical systems based on the growth of the height of the first peak as a function of the density. The graph for this first peak height is given in Figure 112 and we conclude that we cannot deduce how this peak height grows due to the noise in comparison with the range of peak heights. Therefore, this analysis has to be postponed until improvements of the experimental data are available.

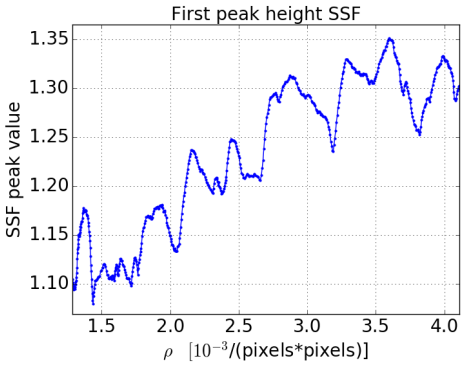

Figure 112: Height of the first peak of the SSF as a function of the density.

I.2 Looking before confluence

In Subsection 5.2 we stated that the RDF and SSF can be used to determine when confluence is reached and in this appendix subsection we explain this further. Figures 113 and 114 show respectively the radial distribution function and scalar function from the moment measurements are started, which is 2000 minutes before confluence is reached. The RDF peak value has its minimum at the moment confluence is reached and the reason for the higher peak values before confluence can be understood as follows. Different isolated islands exists in a preconfluent layer and only the interparticle distances of cells on the same island contribute to the RDF on short and medium length scales. Because of the gaps in the cell layer, the cell number density on the island is higher than the calculated average density of the layer. The radial distribution function is normalized by the average density, which is too small on the islands and therefore results in values of the RDF that are too high. The SSF, however, is not normalized by the density and the peak height grows barely before confluence. After confluence is reached the peak height grows steadily and this transition point can also be used to determine confluence.

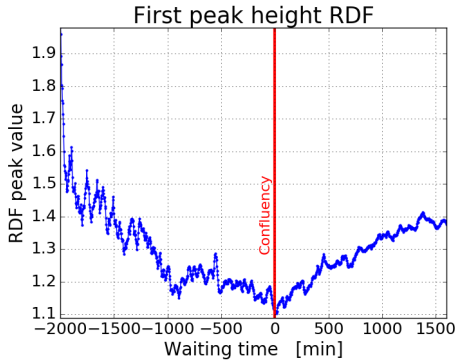

Figure 113: Height of the first peak of the RDF during the complete measurement period.

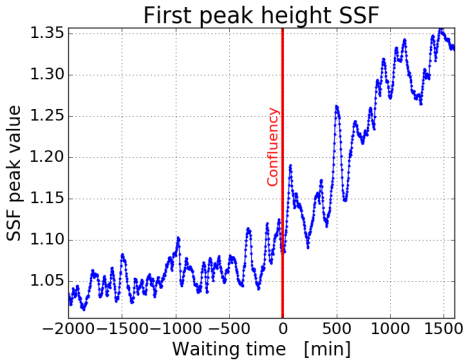

Figure 114: Height of the first peak of the SSF during the complete measurement period.

J Additional results for the VV-model in Section 6

In this appendix section we present additional data that show the effect of the Vicsek radius R_V together with the preferred perimeter p_0 on the structure and dynamics in the VV-model. Appendix Subsection J.1 shows how the coherent and incoherent ISF changes upon varying R_V for a value of p_0 corresponding to an unjammed state in the VRD-model. In Appendix Subsection J.2 we show the structural changes in the layer upon varying R_V for a value of p_0 corresponding to a jammed state and in Appendix Subsection J.3 we show the effect of extremer values of R_V on the jamming and alignment in the layer.

J.1 Effect of the Vicsek radius on the dynamics in the unjammed regime

In Figure 115 the coherent and incoherent ISF (top left and right panel) are drawn for $p_0 = 3.8\,r_0$ and $v_0 = 0.25\,r_0/\tau_0$, which corresponds to an unjammed state in the VRD-model according to Figure 33. From these graphs we conclude that in the unjammed regime the relaxation time can be tuned with the Vicsek radius, in line with the observations for the jammed regime in Figure 33 in Subsection 6.4.1. Around $R_V = 1.0\,r_0$ the decay in the ISFs is fastest, which is around the transition point from almost no to a finite alignment (left bottom panel of Figure J.1), similar as in the jammed regime.

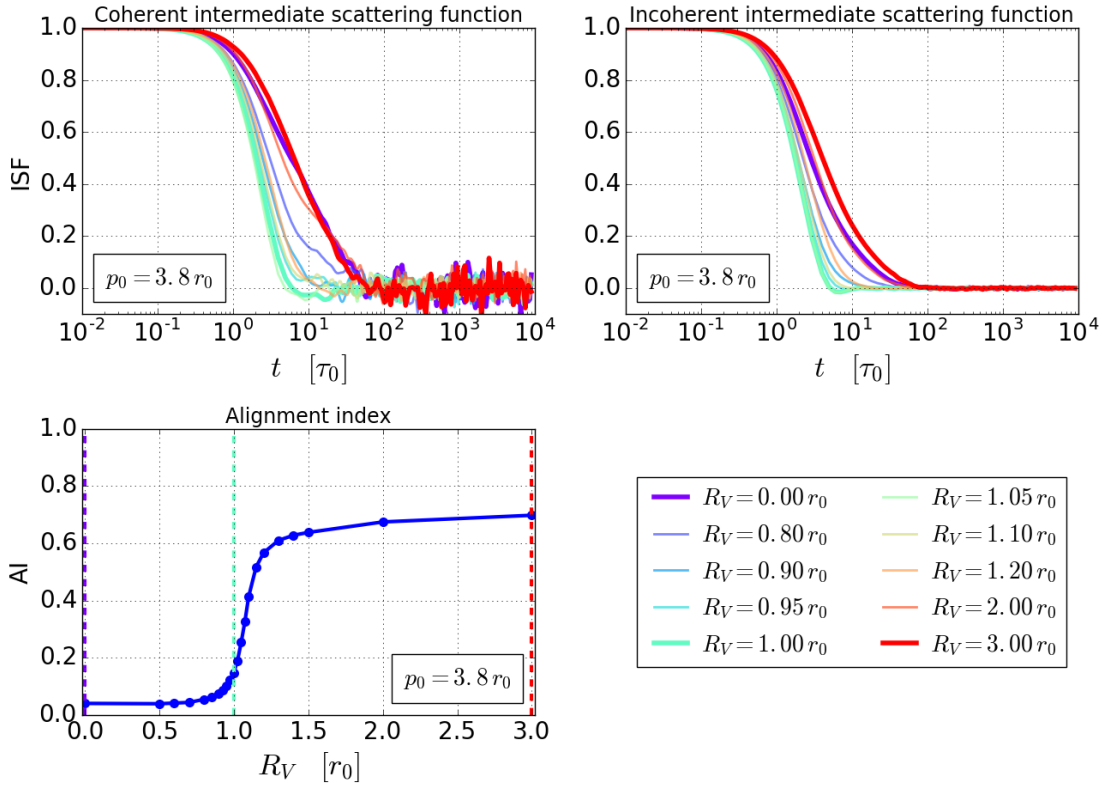

Figure 115: The coherent ISF (top left panel), the incoherent ISF (top right panel) and the alignment index (left bottom panel) for $p_0 = 3.8\,r_0$ and R_V varying from $0.0\,r_0$ till $3.0\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).

J.2 Effect of the Vicsek radius on the structure

In the left top panel of Figure 116 the average perimeter as a function of the Vicsek radius is drawn for $p_0 = 3.55\,r_0$,$v_0 = 0.25\,r_0/\tau_0$, which corresponds to a jammed state in the VRD-model. This average perimeter has its minimal value around $R_V = 1.0\,r_0$ and in Subsection 6.4.1 is shown that the dynamics is fastest around this value of R_V. This increase in structural order is also seen in the RDF (bottom left panel of Figure 116) and the SSF (bottom right panel), where the peaks for $R_V = 1.0\,r_0$ are highest. This increase in structure for faster dynamics is completely opposite to what is observed in the VRD-model (see Subsection 4.3).

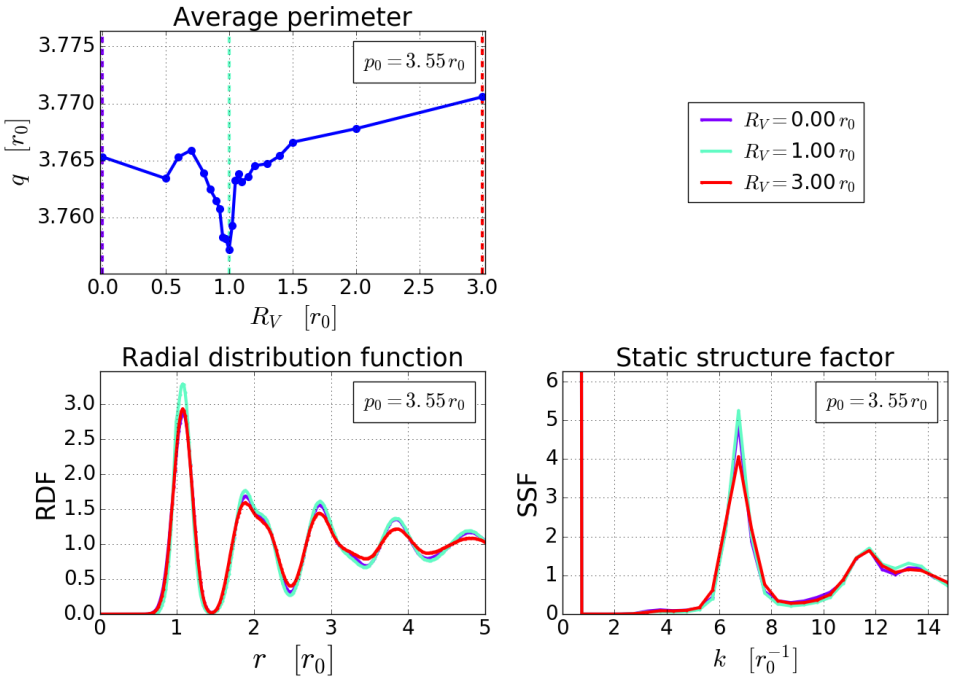

Figure 116: The average perimeter (top left panel) for R_V varying between $0.0\,r_0$ and $3.0\,r_0$, the RDF (left bottom panel) and the SSF (right bottom panel) for $R_V = 0.0\,r_0$, $R_V = 1.0\,r_0$ and $R_V = 3.0\,r_0$ ($p_0 = 3.55\,r_0$,$v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).

J.3 Jamming and alignment for extreme values of the Vicsek radius

Figure 117 is a zoomed-out version of Figure 79 and shows that the interesting regime regarding jamming and dynamics is around $R_V = 1.0\,r_0$ for all evaluated values of p_0.

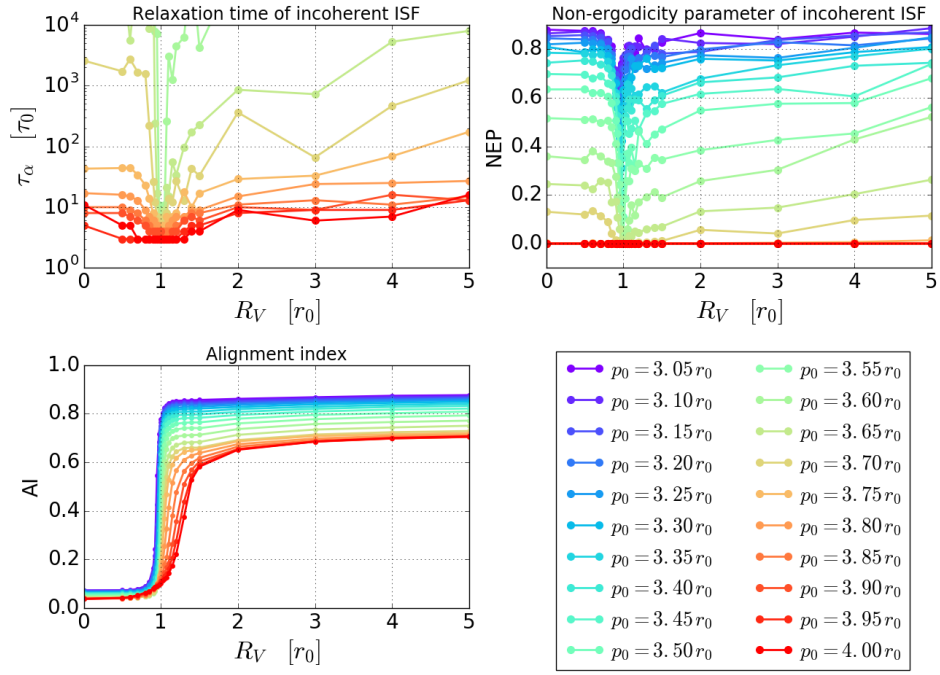

Figure 117: The relaxation time (left top panel) and the non-ergodicity parameter (right top panel) obtained from the incoherent ISF and the alignment index (left bottom panel) for p_0 varying from $3.0\,r_0$ till $4.0\,r_0$ and R_V varying from $0.0\,r_0$ till $5.0\,r_0$ ($v_0 = 0.25\,r_0/\tau_0$, $D_r = 1.0\,\tau_0^{-1}$ and $\tau_V = 1.0\,\tau_0$).