Eindhoven University of Technology

MASTER

Collective cell dynamics in cancer metastasis

Debets, Vincent E.

Award date:
2019

Link to publication

Eindhoven University of Technology

Department of Applied Physics

University College London

Department of Physics and Astronomy

Collective Cell Dynamics in Cancer Metastasis

V.E. Debets

Supervisors:

Prof. dr. C. Storm
dr. L.M.C. Janssen
dr. A. Šarić

Committee:

dr. N.A. Kurniawan
dr. S. Tao

Final version

Eindhoven, March 2019

Abstract

After decades of study, there still remain many open questions about cancer metastasis. One of these involves the nature of tumor cells that give rise to the process. In contrast with conventional models in which single tumor cells seed metastases, recent experimental evidence points to tumor cell clusters migrating collectively as a significant contribution to metastases.

Since very little is known about the manner in which these clusters migrate and why they form a better seed for metastasis than single cells, our work attempts to further rationalise the difference between collective and single cell motion. For this we have modeled migration of single cells and small cell clusters using the cellular Potts model.

Our results demonstrate how cell-cell alignment can significantly improve migration for increasing cluster sizes and that the obtained results are consistent with both theoretical predictions and reproduce experimental oberservations in cell layers. We also show that an increased cluster size can result in enhanced durotaxis which is consistent with experimental results on collective durotaxis of monolayers. Additionally, the model indicates the importance of cell-cell adhesion when clusters experience steric hindrance.

Finally, to incorporate the effect of focal adhesions, we have implemented a binding potential motivated via Langevin Dynamics simulations and show how binding slows down cells but the effect on clusters or single cells is the same. We expect that a feedback between focal adhesions and cell polarity or the addition of extracellular matrix remodeling and degradation will further help elucidate collective cell migration.

Contents

1. Introduction

Together with cardiovascular diseases, cancer still remains the leading cause of death in the world by a significant margin [1]. It is troubling that the world health organisation WHO has predicted that cancer-related deaths will steadily rise until at least 2030 and that their prevalence over most other deadly diseases increases as well [2]. This emphasises the already pressing need for a better understanding of cancer.

Specifically, the vast majority of deaths due to cancer (approximately 90%) are a consequence of metastasis [3–6]. This is the detachment of cells from the primary tumor that traverse through the extracellular matrix (ECM), i.e. the 3D network which provides structural support to tissues, towards blood vessels; there, they enter and travel through the blood circulation system before eventually leaving it to proliferate in distant organs [4,5,7] (see fig. 1.1 for a schematic representation). In other words, disseminated tumor cells form a secondary tumor at a distant site in the body. The main problem for patients is that after metastases have been established, current therapies are often not effective in treating the metastatic disease and methods like chemotherapy or radiation therapy can come at the cost of much collateral damage to the body [5,7]. Additionally, malignant tumors start to metastasize early on in the development of the disease. This means that cells from a primary tumor might have already infiltrated distant organs before the primary tumor is even diagnosed or can do so quickly after the diagnosis [5]. Once a secondary tumor has formed at one these organs, it can also start to metastasise and the situation can rapidly worsen for a patient [8]. Taking all of this into account, it becomes evident why metastasis is the source of so many cancer-related deaths and why a better understanding of the whole process is paramount to prevent cancer from spreading and reaching a point where it becomes immediately life-threatening.

Unfortunately, after decades of research, tumor metastasis remains poorly understood and we are still far from providing prevention and a cure [4]. It has long been believed that only single tumor cells are able to reach distant organs and contribute to metastasis for the simple reason that multicellular clusters are too large to pass through narrow capillaries [8]. One of the most important conventional models of metastasis is the epithelial-mesenchymal transition (EMT) [9]. This describes the process in which an epithelial cell inside the tumor undergoes multiple biochemical changes in order to take on a mesenchymal phenotype which includes invasiveness, enhanced migratory abilities, loss of cell-cell adhesion structures and an elevated resistance to apoptosis (programmed cell death) [9,10]. This allows the cell to disseminate from the primary tumor and travel towards a distant organ. Once at the organ, the tumor cell can transform back via the reverse mesenchymal-epithelial transition (MET) in order to stabilise itself and enable it to grow into a secondary tumor. Since the first discovery of EMT with respect to tumor progression in 1986 [11], much research has been dedicated to the process and its involvement in metastasis [10,12,13].

However, recent experimental studies have highlighted significant contributions to metastasis of so called circulating tumor cell (CTC) clusters, i.e. (mostly) heteroge-
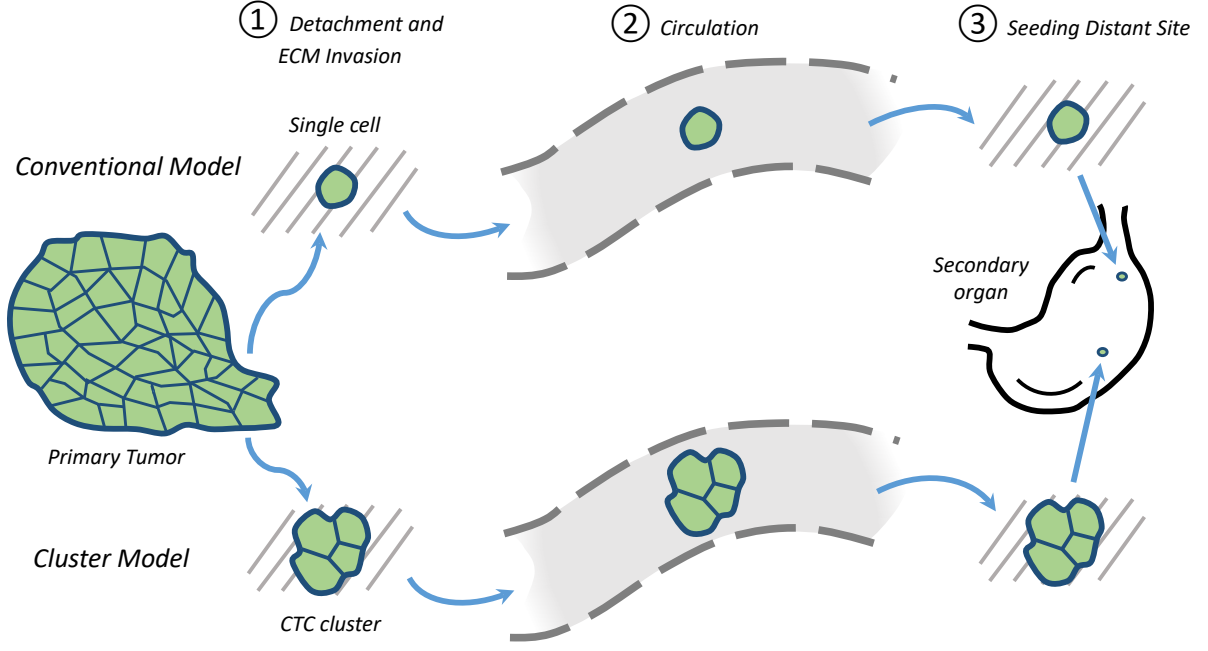
Figure 1.1: Schematic representation of the different steps involved in cancer metastasis. Image shows the conventional model (EMT) in which only single cells cause metastasis and the recently proposed idea that tumor cell clusters travelling as a whole from begin to end contribute the most to metastasis.

neous cell clusters typically consisting of approximately 2 to 20 cells that have detached from the primary tumor and are migrating towards a distant organ [3–5, 7, 8, 14–16]. Particularly, experiments studying spontaneous breast cancer in mice have for instance indicated that over 97% of all observed metastases originated from clusters, while other studies have suggested that CTC clusters may have at least a 50 times greater metastatic potential than individual CTCs [3, 7, 14, 16]. Interestingly, it has also been shown that a CTC cluster migrates as a single unit from the primary tumor to distant organs and the idea of CTCs assembling at some point during metastasis is improbable [7]. It thus seems that the original dismissal of cell clusters in metastasis proves to be too simplistic and that clusters apparently find a way to bypass narrow capillaries and other forms of steric hindrance. In fact, it is highly likely that cell clusters exhibit specfic advantages over single cells during the entire process of metastasis which allow them to harbor an increased metastatic potential, although it remains unclear what these precisely are. Additionally, studies in breast and pancreatic cancer have indicated that EMT appears not to be a necessary requirement for the establishment of metastasis, which implies that the contribution of EMT to metastasis might be more nuanced than originally thought [5]. In other words, it can motivate an even stronger shift in attention towards cell clusters.

Overall, the experimental findings have opened up a complete new field of study concerning relatively small cell clusters in relation to cancer metastasis and one that looks very promising and is only just beginning to be further rationalised. In order to explain the enhanced metastatic potential of CTC clusters a number hypotheses have already been brought forward including for example the cooperation of heterogeneous cell types within the CTC cluster, shielding from attacks by immune cells, and the protection from pressures and shear forces in the bloodstream [3, 8]. Still, much remains

unknown about the genesis, transit and the settlement of CTC clusters during metastasis [8].

It is clear that the whole process of metastatic colonization constitutes a number of very different aspects (detachment, invasion, circulation and seeding) which by themselves already represent rich areas of research. In this work we will specifically focus on the motion of cell clusters after they have detached from the primary tumor and seek to migrate towards a blood vessel, i.e. the invasion of cells through the ECM. Additionally, even if the crucial benefit of migrating as a cluster occurs in a different stage of metastasis, an improved understanding of cluster motion immediately after detachment can still contribute to the development of treatments that try to prevent cancer spread directly at its source.

Thus, the aim of this thesis is to, from a physical point of view and using a computational modeling approach, shed more light on the manner in which single cells and cell clusters migrate through the ECM, thereby elucidating whether there is a possible metastatic advantage for cell cluster migration. However, the ECM is a highly complex environment and a complete representation is difficult within a model framework. We will therefore use simplified representations of the ECM which seek to capture different characteristics of the environment a cell (cluster) faces during invasion. These are demonstrated in fig. 1.2. Initially, we represent the ECM as a homogeneous environment to exclusively study the effect of cell-cell alignment which has already implicated a benefit of migrating as a cluster in other biological contexts involving larger cluster sizes [17]. Next, we impose a gradient in stiffness to investigate durotaxis (transport up such a gradient), which has also been shown to favor cluster over single cell migration [18, 19]. Since cell clusters often have to pass through small openings in the ECM network, we also study the migration of clusters through a narrow pore, focusing on experimentally relevant parameters such as adhesion of the cell with other cells or the pore walls, and traction forces the cells generate when inside the pore [20, 21]. Cells not only pass through the network, but continuously bind (and unbind) to the ECM as well via so called focal adhesions (protein assemblies) [22, 23]. The final studied ECM environment includes these focal adhesions explicitly. Moreover, we want to infer the behavior of the focal adhesion by looking at it in more detail. Summarising, in this work we want to compare single cell to cluster motion in four distinct biologically relevant scenarios, which are shown in fig. 1.2, and observe when one is preferred over the other.

Specifically, we will use the cellular Potts model (CPM) to describe cell (cluster) motion. Because cells are able to actively migrate, we extend the CPM with an active matter theory; in particular, we describe the cells as (aligning) active Brownian particles (ABPs). Within this description alignment is introduced via a Vicsek-like model. Since stiffness has been experimentally linked to an increased cell speed and persistence, we model durotaxis in the CPM via linear gradients in these cell properties [18, 19, 24–26]. Moreover, to improve the realism of the simulations we eventually include a narrow pore and focal adhesions into the CPM via the inclusion of rigid obstacles and local time-dependent energetic bonds respectively. By performing more detailed Langevin Dynamics simulations of attachment of a cell membrane to the ECM, we can infer the functional dependence of this energetic bond.

The report will be structured in the following manner. We start with discussing the theory of active Brownian particles (ABPs), which will be extended by deriving the collective motion of fast-aligning ABP clusters using a velocity alignment potential. Then, we will introduce the cellular Potts model (CPM) and demonstrate how the theory of ABPs and cell attachment via focal adhesions are implemented in the CPM. Next, we discuss the implementation of the Langevin Dynamics simulations that have been per-
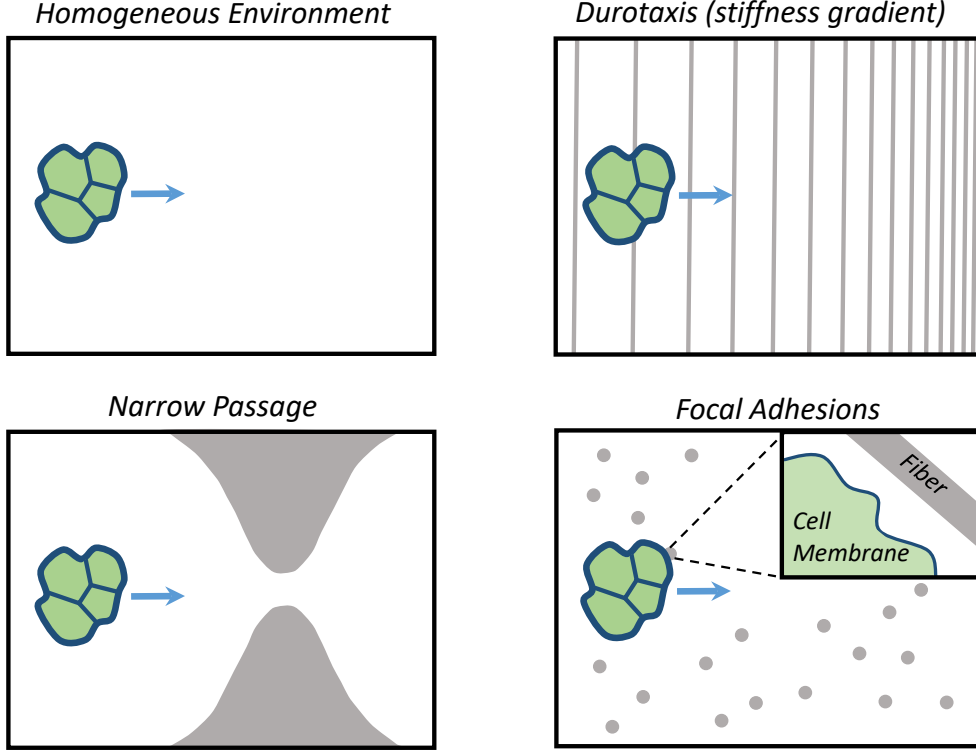
Figure 1.2: Schematic representation of the different ECM environments in which collective cell motion has been investigated with the cellular Potts model (CPM). The inset shows a visualisation of cell binding to a collagen fiber, which is examined with more detailed Langevin Dynamics simulations.

formed to extract a potential describing cell binding via focal adhesions; particularly, a description of the used cell membrane model and the interaction it has with a model collagen fiber are mentioned. The trajectories of single cells and cell clusters have afterwards been analysed and discussed for the four different considered biological scenarios. Finally, we will summarise the main findings and provide an outlook how the research can possibly be forwarded.

2. Theory

In this chapter we introduce the theoretical models which have been used to describe cell motion. Since biological cells are able to actively migrate, we will invoke so-called active matter theory for this description. The main results, eqs. (2.10), (2.20) and (2.37), will be confirmed with our cellular Potts model simulations.

2.1 Active Brownian Motion

Active particles are able to continuously consume energy and propel themselves with directed motion. This renders them intrinsically far from thermodynamic equilibrium. One of the most common models in the field of active matter is the so called active Brownian particles (ABP). These particles undergo Brownian motion, while they simultaneously self-propel with an absolute speed v_0 along their orientational axis, called the director $\mathbf{e}(t)$. The director in turn also evolves in time t due to Brownian noise and is parametrised by the polar angle $\phi(t) \in [0, 2\pi)$ in 2D and the spherical angles $\phi(t) \in [0, 2\pi)$ and $\theta(t) \in [0, \pi)$ in 3D (see fig. 2.1 for a visualisation of the angles), such that we have

$$\mathbf{e}(t) = \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \tag{2.1}$$

and

$$\mathbf{e}(t) = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{2.2}$$

respectively.

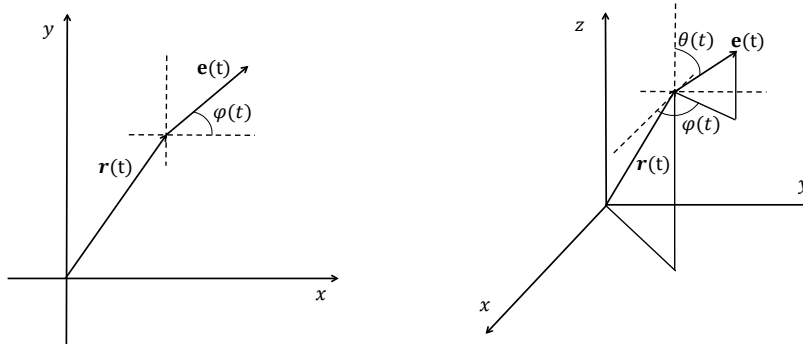**Figure 2.1:** Visualisation in 2D and 3D of the position $\mathbf{r}(t)$ of an ABP and the director $\mathbf{e}(t)$ that describes its direction of active motion. The latter has been described with the polar angle $\phi(t)$ and the spherical angles $\phi(t)$ and $\theta(t)$ respectively.

In the following subsections we will state the stochastic differential Langevin equations that describe active Brownian motion in both two and three dimensions for non-interacting particles and derive the mean square displacement and velocity autocorrelation function. These turn out to correspond to a persistent random walk. Afterwards, we will also demonstrate that in two dimensions the center-of-mass motion of two rapidly aligning identical active Brownian particles leads to similar results as the motion of a single particle with twice as large a persistence time and generalise this idea to N rapidly aligning particles. We finalise the section by discussing the case in which the aligning particles are non-identical.

2.1.1 Non-Interacting Particles

Two Dimensions

The time evolution of the position $\mathbf{r} = (x, y)$ of a 2D active Brownian particle can be captured by the following stochastic differential Langevin equation [27–29]

$$\frac{d\mathbf{r}(t)}{dt} = \mathbf{v}(t) = v_0 \mathbf{e}(t) + \sqrt{2D}\boldsymbol{\xi}(t), \tag{2.3}$$

while the dynamics of the angle ϕ (and thus $\mathbf{e}(t)$) are described by [27, 28]

$$\frac{d\phi}{dt} = \sqrt{\frac{2}{\tau}}\xi_\phi(t). \tag{2.4}$$

Here D denotes the diffusion coefficient that characterises translational Brownian motion, τ a rotational time scale, and $\boldsymbol{\xi} = [\xi_x, \xi_y]$ where ξ_α ($\alpha = x, y, \phi$) represents a independent white noise stochastic process with zero mean, $\langle \xi_\alpha(t) \rangle = 0$, and delta correlation, $\langle \xi_\alpha(t')\xi_\beta(t) \rangle = \delta(t' - t)\delta_{\alpha\beta}$. Note that eq. (2.4) essentially describes a random walk of the director $\mathbf{e}(t)$ on a unit circle.

Using techniques from stochastic calculus we can obtain the Fokker-Planck equation corresponding to eq. (2.3) and eq. (2.4) for the probability density function $P(\mathbf{r}, \phi, t)$ of finding a particle at position $\mathbf{r}$ with orientation ϕ at time t [30]. The result is given by

$$\frac{\partial P}{\partial t} = \frac{1}{\tau}\frac{\partial^2 P}{\partial \phi^2} + D\boldsymbol{\nabla}^2 P - v_0 \mathbf{e} \cdot \boldsymbol{\nabla} P \tag{2.5}$$

where $\mathbf{e}$ denotes the 2D director, eq. (2.1). Assuming the particle starts in the origin with a random uniformly distributed orientation we can formulate the initial condition $P(\mathbf{r}, \phi, 0) = \delta(\mathbf{r})/2\pi$. In order to find the mean square displacement (MSD) and from that the velocity autocorrelation function (VAC) we can take the Fourier-Laplace transform of eq. (2.5),

$$\frac{1}{\tau}\frac{\partial^2 \hat{P}}{\partial \phi^2} = \left[s + iv_0 \left(q_x \cos(\phi) + q_y \sin(\phi) \right) + D(q_x^2 + q_y^2) \right] \hat{P} - \frac{1}{2\pi} \tag{2.6}$$

where we adopted the notation $P(q_x, q_y, \phi, s) \equiv \hat{P}$ with q_x, q_y and s being the transformation variables of x, y and t respectively. The MSD is given by

$$\langle \mathbf{r}(t)^2 \rangle = \int \int \mathbf{r}^2 P(\mathbf{r}, \phi, t) d\phi d\mathbf{r}, \tag{2.7}$$

where we have assumed that the particle starts in the origin $\mathbf{r}(0) = \mathbf{0}$. Its Laplace transform is

$$\langle \mathbf{r}^2 \rangle (s) = -\int \left[\left(\frac{\partial^2 \hat{P}}{\partial q_x^2} \right)_0 + \left(\frac{\partial^2 \hat{P}}{\partial q_y^2} \right)_0 \right] d\phi, \tag{2.8}$$

6

with the subindex 0 denoting an evaluation of a function at $q_x = q_y = 0$, which will be used throughout. If we then set $q_x = q_y = 0$ in eq. (2.6) we find an ordinary differential equation which can be solved for $\hat{P}_0$. Furthermore, we can differentiate eq. (2.6) with respect to q_x or q_y and use the result for $\hat{P}_0$ in order to find explicit expressions for $(\frac{\partial \hat{P}}{\partial q_x})_0$ and $(\frac{\partial \hat{P}}{\partial q_y})_0$. By repeating this process once more we can also determine $(\frac{\partial^2 \hat{P}}{\partial q_x^2})_0$ and $(\frac{\partial^2 \hat{P}}{\partial q_y^2})_0$ which can be inserted in eq. (2.8) to give

$$\left\langle \mathbf{r}^2 \right\rangle (s) = \frac{2v_0^2 \tau}{s^2(1 + s\tau)} + \frac{4D}{s^2}. \tag{2.9}$$

For an extended derivation of this result the reader is referred to appendix A.1. Taking the inverse Laplace transform of eq. (2.9) we obtain

$$\left\langle \mathbf{r}(t)^2 \right\rangle = 2v_0^2 \tau^2 (e^{-t/\tau} + t/\tau - 1) + 4Dt, \tag{2.10}$$

which corresponds to a persistent random walk. In particular, the result shows that the MSD consists of a normal diffusive part characterised by the diffusion coefficient D and a persistent part which is characterised by a persistence time τ and an active diffusion coefficient $D_p \equiv v_0^2 \tau / 2$. Specifically, the motion is diffusive (MSD $\propto t$) with diffusion coefficient D at very short time scales $t \ll \tau$, ballistic (MSD $\propto t^2$) at intermediate time scales $t \approx \tau$ and again diffusive with an increased diffusion coefficient $D + D_p$ at long time scales $t \gg \tau$.

Finally, we can retrieve the VAC which is given by [31]

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle = \frac{1}{2} \frac{\partial^2}{\partial t^2} \left\langle \mathbf{r}^2 \right\rangle (t) = v_0^2 e^{-t/\tau} \tag{2.11}$$

where $\mathbf{v}(t)$ denotes the velocity of the particle. Note that this relation assumes the VAC to be both translationally invariant and symmetric in time.

Three Dimensions

In three dimensions the translational dynamics of an active Brownian particle are described in a similar matter as in two dimensions (eq. (2.3)), with the modification that we extend to three dimensions and therefore $\mathbf{r} = [x, y, z]$ and $\boldsymbol{\xi} = [\xi_x, \xi_y, \xi_z]$, while for $\mathbf{e}(t)$ we need to insert eq. (2.2). The Langevin dynamics of the angles becomes less trivial and is given by [32]

$$\frac{d\theta}{dt} = \sqrt{\frac{1}{\tau}} \xi_\theta(t) + \frac{1}{2\tau \tan(\theta)}, \tag{2.12a}$$

$$\frac{d\phi}{dt} = \sqrt{\frac{1}{\tau}} \frac{1}{\sin(\theta)} \xi_\phi(t), \tag{2.12b}$$

where we have introduced an additional independent stochastic process $\xi_\theta(t)$. This describes a random walk of the director $\mathbf{e}(t)$ on the unit sphere characterised by a rotational time scale τ.

Note that there is no spatial dependence in eq. (2.12), since we assume an isotropic environment and no interaction between particles. This means that we can also focus only on the probability density function of finding a particle with a specific orientation at a time t, i.e. $P(\theta, \phi, t)$. Applying techniques of stochastic calculus to eq. (2.12) we find that the corresponding Fokker-Planck equation reads [32]

$$\frac{\partial P(\theta, \phi, t)}{\partial t} = -\frac{\partial}{\partial \theta} \left(\frac{P}{2\tau \tan(\theta)} \right) + \frac{1}{2\tau} \frac{\partial^2 P}{\partial \theta^2} + \frac{1}{2\tau \sin^2(\theta)} \frac{\partial^2 P}{\partial \phi^2}. \tag{2.13}$$

It will become apparent that the so called marginal probability density $\tilde{P}(\theta, t)$, i.e. $P(\theta, \phi, t)$ integrated over ϕ, already suffices to retrieve the MSD and VAC. We therefore integrate eq. (2.13) over the polar angle ϕ to obtain its Fokker-Planck equation, which reads [32]

$$\frac{\partial \tilde{P}(\theta, t)}{\partial t} = -\frac{\partial}{\partial \theta} \left(\frac{\tilde{P}}{2\tau \tan(\theta)} \right) + \frac{1}{2\tau} \frac{\partial^2 \tilde{P}}{\partial \theta^2}. \tag{2.14}$$

Since the motion is isotropic we can assume that without loss of generality the initial direction of motion is always in z-direction [32] and we therefore have as an initial condition that $\tilde{P}(\theta, 0) = \delta(\theta)$ with $\delta(\theta)$ the Dirac delta function. Applying separation of variables and using the initial condition allows us to solve the Fokker-Planck equation to yield

$$\tilde{P}(\theta, t) = \frac{\sin(\theta)}{2} \sum_{l=0}^{\infty} (2l+1) P_l(\cos(\theta)) \exp(-l(l+1)t/2\tau), \tag{2.15}$$

where $P_l(x)$ are Legendre Polynomials [33]. For a more detailed derivation of this result the reader is referred to appendix A.2. Since the initial direction is in the z-direction we have $\mathbf{e}(0) = (0, 0, 1)$. The autocorrelation function of the director can then be calculated as

$$\langle \mathbf{e}(t) \cdot \mathbf{e}(0) \rangle = \int_0^{2\pi} \int_0^{\pi} \cos(\theta) P(\theta, \phi, t) d\theta d\phi = \int_0^{\pi} \cos(\theta) \tilde{P}(\theta, t) d\theta = \exp(-t/\tau), \tag{2.16}$$

where only the $l = 1$ term contributes to the integral.

We can now integrate the Langevin equation (eq. (2.3)) that describes the particle dynamics with respect to time t, i.e.

$$\mathbf{r}(t) - \mathbf{r}(0) = \int_0^t \left(v_0 \mathbf{e}(t') + \sqrt{2D} \boldsymbol{\xi}(t') \right) dt'. \tag{2.17}$$

Assuming that the particle starts in the origin, i.e. $\mathbf{r}(0) = \mathbf{0}$, we can find an expression for the MSD by taking the square and subsequently the ensemble average of eq. (2.17),

$$\langle \mathbf{r}(t)^2 \rangle = \int_0^t \int_0^t \left(v_0^2 \langle \mathbf{e}(t') \cdot \mathbf{e}(t'') \rangle + 2D \langle \boldsymbol{\xi}(t') \cdot \boldsymbol{\xi}(t'') \rangle \right) dt' dt'', \tag{2.18}$$

where we made use of the fact that the spatial stochastic processes are independent of the angular ones so that $\langle \mathbf{e}(t) \cdot \boldsymbol{\xi}(t) \rangle = \langle \mathbf{e}(t) \rangle \cdot \langle \boldsymbol{\xi}(t) \rangle = 0$.

Since we have $\langle \boldsymbol{\xi}(t') \cdot \boldsymbol{\xi}(t'') \rangle = 3\delta(t' - t'')$ (with $\delta(t)$ the Dirac delta function), and givien that the system is isotropic such that we have time translational invariance, we can simplify eq. (2.18) to

$$\langle \mathbf{r}(t)^2 \rangle = 2v_0^2 \int_0^t \int_0^{t'} \langle \mathbf{e}(t' - t'') \cdot \mathbf{e}(0) \rangle dt'' dt' + 6Dt. \tag{2.19}$$

By inserting eq. (2.16) and solving the integrals we arrive at

$$\langle \mathbf{r}(t)^2 \rangle = 2v_0^2 \tau^2 (e^{-t/\tau} + t/\tau - 1) + 6Dt. \tag{2.20}$$

This result is almost similar to the MSD for 2D active Brownian motion (and thus also describes a persistent random walk), except for the fact that the extra dimension adds a factor of two to the normal diffusive process ($4D \to 6D$). Also, note that

the v_0-dependent contribution to the MSD from the active motion is independent of the dimension; the active diffusion coefficient is therefore effectively smaller, i.e. $D_p \equiv v_0^2 \tau / 3$.

Finally, we can retrieve the VAC from the MSD as before via [31]

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle = \frac{1}{2} \frac{\partial^2}{\partial t^2} \langle \mathbf{r}(t)^2 \rangle = v_0^2 e^{-t/\tau} \tag{2.21}$$

which coincides with the 2D result [eq. (2.11)].

2.1.2 Alignment

Two Aligning Particles

We will now consider a 2D system of two identical active Brownian particles that try to align the direction of their velocities with each other. The translational equation of motion of both particles is still described in the same way as for a single non-interacting particle,

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t) = v_0 \mathbf{e}_i(t) + \sqrt{2D} \boldsymbol{\xi}_i(t), \tag{2.22}$$

where $i = 1, 2$ labels each particle and $\boldsymbol{\xi}_i = [\xi_{x_i}, \xi_{y_i}]$. A velocity alignment mechanism can then be included in the system via an angular dependent interaction potential $U(\phi_1, \phi_2)$, such that the rotational equation of motion becomes [34, 35]

$$\frac{d\phi_i}{dt} = -\eta \frac{\partial U}{\partial \phi_i} + \sqrt{\frac{2}{\tau}} \xi_{\phi_i}(t), \tag{2.23}$$

with η depicting a relaxation constant that controls how fast the alignment takes place. Analogous to spin systems, we can introduce an alignment mechanism via a potential defined by [34]

$$U(\phi_1, \phi_2) = -\lambda \cos(\phi_1 - \phi_2), \tag{2.24}$$

where λ denotes the coupling strength. This potential has a minimum when both particles point in the same direction ($\phi_1 = \phi_2$), while it exhibits a maximum when particles are pointing in the opposite direction ($\phi_1 = -\phi_2$). It should also be mentioned that in many descriptions of systems involving alignment, a cutoff radius is introduced such that the alignment potential only applies to particles being within each other's cut off radii. [34, 35]. However, relating the particles to cells that stick together (due to adhesion), we can assume that they always remain sufficiently close together to align with one another and therefore neglect this cutoff effect, which would otherwise complicate an analytical description dramatically. Inserting our alignment potential, eq. (2.24), into eq. (2.23) we retrieve the following angular evolution equations for both particles

$$\frac{d\phi_1}{dt} = -\eta \sin(\phi_1 - \phi_2) + \sqrt{\frac{2}{\tau}} \xi_{\phi_1}(t), \tag{2.25a}$$

$$\frac{d\phi_2}{dt} = -\eta \sin(\phi_2 - \phi_1) + \sqrt{\frac{2}{\tau}} \xi_{\phi_2}(t), \tag{2.25b}$$

where we have absorbed the constant λ into η.

We can define the center of mass of both particles as $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$, leading to the following time evolution

$$
\begin{aligned}
\frac{d\mathbf{R}(t)}{dt} = {} & \left(v_0 \cos\left(\frac{\phi_1 + \phi_2}{2}\right) \cos\left(\frac{\phi_1 - \phi_2}{2}\right) + \sqrt{\frac{D}{2}}(\xi_{x_1}(t) + \xi_{x_2}(t)) \right) \hat{\mathbf{x}} \\
& + \left(v_0 \sin\left(\frac{\phi_1 + \phi_2}{2}\right) \cos\left(\frac{\phi_1 - \phi_2}{2}\right) + \sqrt{\frac{D}{2}}(\xi_{y_1}(t) + \xi_{y_2}(t)) \right) \hat{\mathbf{y}},
\end{aligned}
\tag{2.26}
$$

where we have written out the vectors $\mathbf{e}_{1,2}$ and $\boldsymbol{\xi}_{1,2}$ explicitly with $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ denoting the unit vectors in the x- and y-direction respectively. Additionally, we have used the trigonometric sum-to-product rule to rewrite the sums of sines and cosines as a product. Let us now assume that alignment between the particles occurs very rapidly. This implies that the difference between both angles will remain very small, i.e. $|\phi_1 - \phi_2| \ll 1$. (Strictly speaking we assume that $|(\phi_1 - \phi_2) \bmod 2\pi| \ll 1$, but since we apply the approximation only to a cosine function additional factors of 2π do not alter the results). Retaining only terms up to first order in $\phi_1 - \phi_2$ allows us to approximate the equation of motion of the center of mass as

$$
\frac{d\mathbf{R}(t)}{dt} \approx \left(v_0 \cos(\phi_{cm}) + \sqrt{\frac{D}{2}}(\xi_{x_1}(t) + \xi_{x_2}(t)) \right) \hat{\mathbf{x}} + \left(v_0 \sin(\phi_{cm}) + \sqrt{\frac{D}{2}}(\xi_{y_1}(t) + \xi_{y_2}(t)) \right) \hat{\mathbf{y}}.
\tag{2.27}
$$

Here we have introduced $\phi_{cm} \equiv (\phi_1 + \phi_2)/2$ which describes the average direction of both particles. Combining both equations of eq. (2.25) thus yields for the orientation in the center-of-mass frame,

$$
\frac{d\phi_{cm}}{dt} \approx \sqrt{\frac{1}{2\tau}} (\xi_{\phi_1}(t) + \xi_{\phi_2}(t)),
\tag{2.28}
$$

where both sines have cancelled out.

Finally, since the zero mean stochastic processes ξ_α are independent and delta correlated, we can effectively replace a sum of these variables by a single one via $(\xi_\alpha(t) + \xi_\beta(t)) \rightarrow \sqrt{2}\xi_\gamma(t)$. The additional factor $\sqrt{2}$ ensures that the correlation will remain the same. We can implement the replacement in eq. (2.27) and eq. (2.28) (introducing new labels for ξ_α) which results in

$$
\frac{d\mathbf{R}(t)}{dt} \approx \left(v_0 \cos(\phi_{cm}) + \sqrt{D}(\xi_{x_{cm}}(t)) \right) \hat{\mathbf{x}} + \left(v_0 \sin(\phi_{cm}) + \sqrt{D}(\xi_{y_{cm}}(t)) \right) \hat{\mathbf{y}}.
\tag{2.29}
$$

and

$$
\frac{d\phi_{cm}}{dt} \approx \sqrt{\frac{1}{\tau}} \xi_{\phi_{cm}}(t).
\tag{2.30}
$$

Comparing these with eq. (2.3) and eq. (2.4) we can see that the equations of motion governing the center-of-mass motion are almost identical to the ones describing a single non-interacting active particle. The only required modification is that we replace $\tau \rightarrow 2\tau$ and $D \rightarrow D/2$. In other words, the motion is identical to a single non-interacting active Brownian particle which has twice the persistence time but only half the diffusion coefficient (assuming that the center of mass of the two particles starts in the origin, i.e. $\mathbf{R}(0) = \mathbf{0}$, and the average angle ϕ_{cm} is uniformly distributed). Furthermore, motivated by biological cells sticking together, we have assumed the particles to remain close together, which implies that the center-of-mass motion is approximately the same as the motion of both individual particles.

Generalisation to N Aligning Particles

Interestingly, the line of reasoning presented above can also be generalised to a system of N aligning identical particles. Assuming that each particle aligns with all the other particles via eq. (2.24), we can write the alignment potential for a set of N particles as

$$U(\{\phi_i\}) = -\lambda \sum_{j,k} \cos(\phi_j - \phi_k), \tag{2.31}$$

where $\{\phi_i\}$ denotes the set of N angles that describe the direction of motion for all particles i and the sum is taken over each possible particle combination j, k. This consequently results in the following angular evolution equation for particle i (again absorbing λ into η)

$$\frac{d\phi_i}{dt} = -\eta \sum_{j=1}^{N} \sin(\phi_i - \phi_j) + \sqrt{\frac{2}{\tau}} \xi_{\phi_i}(t). \tag{2.32}$$

Based on eq. (2.3), the center of mass of all particles, i.e. $\mathbf{R} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{r}_i$, obeys the following stochastic differential equation

$$\frac{d\mathbf{R}}{dt} = \sum_{i=1}^{N} \left(\left(\frac{v_0}{N} \cos(\phi_i) + \frac{\sqrt{2D}}{N} \xi_{x_i}(t) \right) \hat{\mathbf{x}} + \left(\frac{v_0}{N} \sin(\phi_i) + \frac{\sqrt{2D}}{N} \xi_{y_i}(t) \right) \hat{\mathbf{y}} \right). \tag{2.33}$$

Again we can assume that all the particles will align very fast and hence the difference between each pair of angles will remain very small, i.e. $|\phi_i - \phi_j| \ll 1$ for all i, j. (We actually assume $|(\phi_i - \phi_j) \bmod 2\pi| \ll 1$, but factors 2π do not influence the proposed approximation). However, instead of rewriting the sum of sine and cosine terms in eq. (2.33) with several trigonometic identities and using the fast-aligning approximation, we can also provide a more intuitive argument to simplify these sums. Since all involved angles are almost equal, all the directors $\mathbf{e}_i$ will point in almost the same direction. We can therefore approximate the sum of the N directors by N vectors which all point in the average direction of the directors. In other words we can replace $\sum_{i=1}^{N} \cos(\phi_i) \to N \cos(\phi_{cm})$ and $\sum_{i=1}^{N} \sin(\phi_i) \to N \sin(\phi_{cm})$ in eq. (2.33) where $\phi_{cm} \equiv \frac{1}{N} \sum_{i=1}^{N} \phi_i$. A visualisation of this approximation for $N = 2$ is shown in fig. 2.2.

Furthermore, realising that the zero mean stochastic noise terms ξ_α are independent and delta correlated, we can substitute a single stochastic variable instead of a sum of N stochastic variables, provided we add a factor $\sqrt{N}$ in front of it to have a consistent correlation. In total this allows us to simplify eq. (2.33) as

$$\frac{d\mathbf{R}}{dt} = \left(v_0 \cos(\phi_{cm}) + \sqrt{\frac{2D}{N}} \xi_{x_{cm}}(t) \right) \hat{\mathbf{x}} + \left(v_0 \sin(\phi_{cm}) + \sqrt{\frac{2D}{N}} \xi_{y_{cm}}(t) \right) \hat{\mathbf{y}}. \tag{2.34}$$

Furthermore, the time evolution of ϕ_{cm} can be formulated using eq. (2.32) giving

$$\frac{d\phi_{cm}}{dt} = \frac{1}{N} \sqrt{\frac{2}{\tau}} \sum_{i=1}^{N} \xi_{\phi_i}(t), \tag{2.35}$$

where again all the alignment terms cancel against each other. As already mentioned, we can replace the sum of the stochastic noise terms by a single one which yields

$$\frac{d\phi_{cm}}{dt} = \sqrt{\frac{2}{N\tau}} \xi_{\phi_{cm}}(t). \tag{2.36}$$
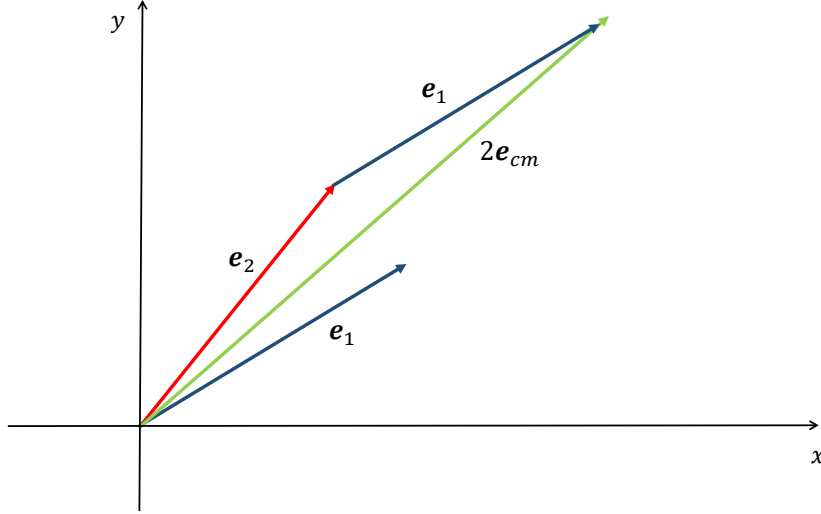
Figure 2.2: Visualisation of the approximation for $N = 2$ particles in which we replace the sum of both particle's directors $\mathbf{e}_{1,2} = (\cos(\phi_{1,2}), \sin(\phi_{1,2}))$ by two times the director with the average angle of both. This director is denoted $\mathbf{e}_{cm} = (\cos(\phi_{cm}), \sin(\phi_{cm}))$ with $\phi_{cm} = (\phi_1 + \phi_2)/2$. We have used $\phi_2 - \phi_1 = \pi/9$; it can be seen that the approximation is still reasonably accurate.

The resulting equations governing the motion of the center of mass (and thus of the entire cluster of N particles), i.e. eq. (2.34) and eq. (2.36), are again similar to the equations that describe a single active particle (eq. (2.3) and eq. (2.4)). The only difference lies in the fact that the persistence time increases linearly with the number of particles $\tau \to N\tau$, while the diffusion coefficient decreases with the number of particles $D \to D/N$. Summarising, we note that the center-of-mass motion of a cluster of N fast-aligning ABPs is characterised by a cluster persistence time τ_{cm}, passive diffusion coefficient D_{cm}, and active diffusion coefficient $D_{p,cm}$, which are given by

$$\tau_{cm} = N\tau, \quad D_{cm} = D/N, \quad D_{p,cm} = v_0^2 N\tau/2. \tag{2.37}$$

Based on the above discussion we conclude that fast alignment of more particles will increase the persistence of the total cluster of particles, allowing it to move more directionally, but at the cost of a decrease in the translational diffusion coefficient. However, if the motion is mostly dominated by the active motion ($D_p \gg D$) the decrease in the diffusion coefficient will hardly influence the overall motion. This implies that aligned clusters can, on average, cover more distance than single particles within a given timeframe, provided the particles are active enough, i.e. v_0 is sufficiently large. Relating the ABPs to CTC cells, we link this to an increase in metastatic success of the cluster.

It should also be stated that the argument presented above is not necessarily bound to the potential given in eq. (2.31). Aside from the requirement that the particles must align sufficiently quickly, we only need all alignment contributions to the time evolution of ϕ_{cm} to cancel out. In other words we require

$$-\eta \sum_{i=1}^{N} \frac{\partial U}{\partial \phi_i} = 0. \tag{2.38}$$

On a final note we should keep in mind that this process does eventually hit an upper

limit with respect to the number of particles N. Since we neglect a cutoff radius in the interaction potential, we assume all particles directly align with all other particles. This assumption will eventually break down when too many particles are involved.

Non-Identical Particles

Up until this point we have assumed that all particles involved are identical. We can extend the analysis to a set of N quickly aligning non-identical active particles in several ways. For instance we can let all particles travel at different velocities $v_0 \rightarrow v_{0,i}$ or make the degree of alignment explicitly particle-dependent. The latter implies that we substitute $\eta \rightarrow \eta_i$ and $\lambda \rightarrow \lambda_{j,k}$ (with $\lambda_{j,k} = \lambda_{k,j}$) in eq. (2.23) and eq. (2.31) respectively. However, we will refrain from exploring these options in too much detail. Firstly, letting the particles travel at different velocities will eventually lead to interparticle distances that are larger than the interaction range of the particles which will make the alignment between particles disappear. Since the proposed model neglects a cutoff radius, we assume all particles will always align, which makes it insufficient to describe this type of behavior. Secondly, introducing $\lambda_{j,k}$ will not change the equations governing the center-of-mass motion: all alignment terms are still canceling out, i.e. eq. (2.38) still applies. This gives the same results as discussed before. By introducing a particle-dependent relaxation constant η_i, however, eq. (2.38) will not be valid anymore. In particular letting $\eta \rightarrow \eta_i$ in eq. (2.32), we find that the time evolution of ϕ_{cm} obeys

$$\frac{d\phi_{cm}}{dt} = -\frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} (\eta_i - \eta_j) \sin(\phi_i - \phi_j) + \sqrt{\frac{2}{N\tau}} \xi_{\phi_{cm}}(t). \tag{2.39}$$

Nonetheless, for sufficiently small $(\phi_i - \phi_j)$ (or $(\phi_i - \phi_j) \bmod 2\pi$)), the sine term in eq. (2.39) will become negligible and we recover the same results as for identical particles, eq. (2.36)). Overall, a larger variety in relaxation constants thus also requires stronger alignment for the assumptions to be valid and retrieve the previously discussed results for identical particles.

The only unexplored option left is to have each particle move with different persistence times, i.e. we replace $\tau \rightarrow \tau_i$. This implies that instead of eq. (2.35) we now have

$$\frac{d\phi_{cm}}{dt} = \frac{\sqrt{2}}{N} \sum_{i=1}^{N} \sqrt{\frac{1}{\tau_i}} \xi_{\phi_i}(t). \tag{2.40}$$

We can again replace the sum of stochastic variables by a single stochastic variable but due to the factor $1/\sqrt{\tau_i}$ inside the sum this becomes less straightforward. Let us define $\beta(t) \equiv \sum_{i=1}^{N} \sqrt{\frac{1}{\tau_i}} \xi_{\phi_i}(t)$, for which it is clear that we have $\langle \beta(t) \rangle = 0$. The correlation function is given by

$$\langle \beta(t)\beta(t') \rangle = \left(\sum_{i=1}^{N} \frac{1}{\tau_i} \right) \delta(t - t'), \tag{2.41}$$

where we have used the fact that all stochastic processes $\xi_{\phi_i}(t)$ are independent and delta correlated. An inspection of eq. (2.41) shows that we can replace $\beta(t)$ by a single stochastic variable, provided we add the factor $\left(\sum_{i=1}^{N} \frac{1}{\tau_i} \right)^{1/2}$ in front of it. This is done

to keep the correlation the same as in eq. (2.41). Therefore, we can write

$$\frac{d\phi_{cm}}{dt} = \left(\frac{2\left(\sum_{i=1}^{N}\frac{1}{\tau_i}\right)}{N^2}\right)^{1/2} \xi_{\phi_{cm}}(t) \equiv \sqrt{\frac{2}{\tau_{cm}}}\xi_{\phi_{cm}}(t), \tag{2.42}$$

where we have defined the persistence time of the center-of-mass motion via $\tau_{cm} \equiv N^2\left(\sum_{i=1}^{N}\frac{1}{\tau_i}\right)^{-1}$. Note that when all persistence times are equal ($\tau_i = \tau$), we recover the linear increase of the persistence time with the number of particles ($\tau_{cm} = N\tau$) which makes it consistent with the previously found result.

It is now interesting to see how a variety in persistence times compares to the case in which all particles are identical and share the same persistence time. In fact it can be shown that for each set of N persistence times $\{\tau_i > 0\}$ we have (see appendix A.3)

$$N^2\left(\sum_{i=1}^{N}\frac{1}{\tau_i}\right)^{-1} \leq N\langle\tau\rangle, \tag{2.43}$$

with $\langle\tau\rangle = \frac{1}{N}\sum_{i=1}^{N}\tau_i$ the average persistence time of the set and the equal sign corresponding to a constant $\tau_i = \tau$. This result shows that spreading out the individual persistence times of particles (keeping a constant average) will always decrease the center-of-mass persistence time τ_{cm}. In other words, alignment is less effective in terms of increasing τ_{cm} when individual particles travel with a different persistence.

Finally, let us consider whether the inclusion of one less persistent particle among the aligning particles can cancel the benefits of alignment of the total cluster. Suppose we have N fast aligning particles, $N-1$ of which have an individual persistence time τ_l, and one particle has a smaller persistence time $\tau_s < \tau_l$. The persistence time of the center-of-mass motion is then given by

$$\tau_{cm} = N^2\left(\sum_{i=1}^{N}\frac{1}{\tau_i}\right)^{-1} = \frac{N^2\tau_s\tau_l}{(N-1)\tau_s + \tau_l}. \tag{2.44}$$

The effect of collective alignment will be canceled when $\tau_{cm} = \tau_l$, in that case, we find

$$\tau_s = \frac{\tau_l}{N^2 - N + 1}. \tag{2.45}$$

The obtained value for τ_s thus presents a critical value above which the cluster has a persistence time larger than τ_l, which means that the cluster moves more persistently than individual cells. Below this value we have $\tau_{cm} < \tau_l$ such that the cluster will move with less persistence than the individual cells. Note that for large N we have $\tau_s \sim \tau_l/N^2 \to 0$ and a single particle is not able to disturb the collective motion of a large cluster. However, since N is typically not large for CTC clusters, this effect is not negligible.

3. Numerical methods

In this chapter we first introduce the Cellular Potts Model (CPM) which has been used for numerical simulations of cell dynamics, and discuss the extensions to the originally proposed model that we have made use of throughout this report. We will introduce a cell membrane model which describes in more detail a part of the cell of approximately the same size as the resolution of the CPM. This is followed by a description of the numerical methods used to simulate the interaction between the introduced model membrane and a collagen fiber which in total will serve as our model system for the formation of a focal adhesion in time. The chapter is concluded by a discussion of several details of the implementation of the numerical focal adhesion model. However, for a full description the reader is referred to appendix B.

3.1 Cellular Potts Model

The cellular Potts model (CPM) has been introduced by Graner and Glazier in 1992 to model cell sorting and test the differential adhesion hypothesis [36–38]. Since then it has been successfully applied to a wide variety of biological phenomena involving for instance blood vessel network formation, cancer cell invasion, and collective cell motion [39–41]. Before we proceed and walk through the original model and the used extensions, we mention that in the descriptions of the CPM, we make use of several cell related properties such as cell center of mass for instance. These properties are fairly intuitive and in order to improve readability, formal definitions are not always provided throughout. For a list with all formal definitions of the cell properties the reader is referred to appendix C.

3.1.1 Original Model

The CPM is a variation on the classic Potts model [42] and consists of spins $\sigma(\mathbf{x}) \in \mathbb{Z}^{+,0}$ on a discrete square or cubic lattice (2D or 3D) with a lattice constant a whose sites are characterised by their position in space $\mathbf{x}$. Other lattices (hexagonal, triangular) are also possible, but in this thesis we will only focus on the square and cubic one due to their relatively easy computational implementation. Biological cells are represented as domains of identical spin $\sigma(\mathbf{x}) \in \mathbb{N}$ on the lattice, where the value of the spin σ effectively serves as a cell identity or label (see fig. 3.1 for a visualisation of the CPM cells). The medium or extracellular matrix (ECM) in or on which the cell resides, is normally denoted with $\sigma = 0$. Furthermore, the medium (or ECM) and each of the cells are assigned a biological cell type via a label $c(\sigma) \in \mathbb{Z}^{+,0}$. Again the label $c(0) = 0$ is generally reserved for the medium or ECM. Note that the original model places the cells in a homogeneous environment since all non-cellular spins are denoted by $\sigma = 0$ and $c(0) = 0$. However, in later simulations presented in this thesis the medium/ECM will consist of multiple elements for which the labels $\sigma, c \in \mathbb{Z}_{\leq 0}$ are reserved. This will
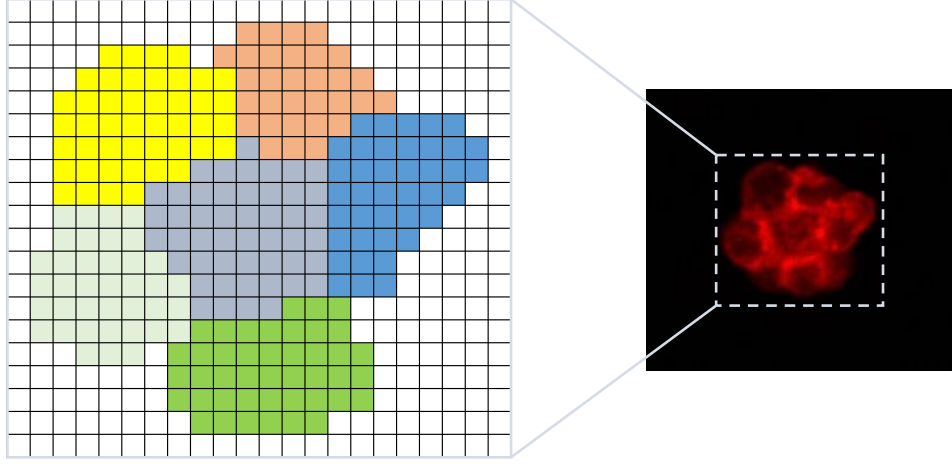
be discussed in more detail in section 3.1.4.

Figure 3.1: Visualisation in 2D of CPM cells representing an experimentally observed circulating tumor cell cluster (CTC cluster). Different colors represent different spins σ with the medium ($\sigma = 0$) shown in white. Experimental figure is taken from [16].

Cell movement can then be imposed on the system via a modified Metropolis Monte Carlo algorithm which operates with respect to a phenomenological Hamiltonian $\mathcal{H}$; the latter accounts for all physically relevant terms such as adhesion for instance. The algorithm consists of the following steps [43]:

(i) Randomly choose a lattice site i. Call this site the candidate site and let $\sigma(\mathbf{x})$ be its spin value.

(ii) Randomly select one of the neighbors of site i. Call this site the target site and let $\sigma(\mathbf{x}')$ be its spin value. If $\sigma(\mathbf{x}) = \sigma(\mathbf{x}')$ proceed to step (v), otherwise proceed to step (iii).

(iii) Calculate the change in Hamiltonian $\Delta\mathcal{H}$ upon changing the candidate site value from $\sigma(\mathbf{x})$ to the target site value $\sigma(\mathbf{x}')$.

(iv) Accept the spin-copy attempt with a probability $p(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) = \min(1, \mathrm{e}^{-\Delta\mathcal{H}/T})$.

(v) Increment the number of copy attempts, update relevant parameters and go back to step (i).

A visualisation of a spin-copy attempt is shown in fig. 3.2. The neighbors from which can be chosen at step (ii) are the ones directly adjacent to the site i (see figs. 3.3a and 3.3b). We employ an acceptance probability which follows a Boltzmann distribution with a simulation temperature T. This choice ensures that the average time evolution obeys an overdamped or Aristotelian force-velocity relation which we expect to be the case for real-life biological cells moving through the ECM [43]. Moreover, we should emphasise that the simulation temperature does not represent the actual temperature since it is too low to induce significant fluctuations to a biological cell [43] [44]. Note that this algorithm does not satisfy detailed balance and is not suited for systems in thermodynamic equilibrium [43]. Since biological cells are active particles, the simulated system is not expected to be in thermodynamic equilibrium and detailed balance is not a necessary requirement, though it can be regained by modifying the algorithm (see [43]).
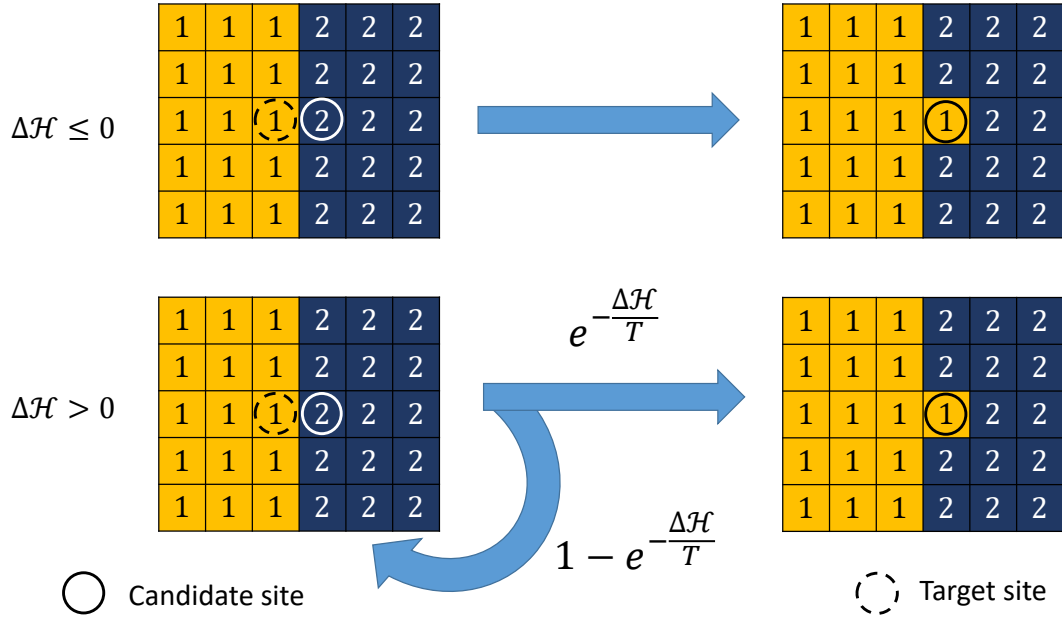
Figure 3.2: Visualisation in 2D of the spin-copy attempt used in the modified Metropolis Monte Carlo algorithm. If the proposed spin-copy lowers the Hamiltonian ($\Delta\mathcal{H} \leq 0$) we always change the configuration, but for an increasing Hamiltonian ($\Delta\mathcal{H} > 0$) the configuration can also stay the same.

The originally proposed Hamiltonian is defined in terms of a volume constraint term which ensures an approximate constant cell volume/area and a surface/line energy term to account for cell-cell adhesion [36, 37]:

$$\mathcal{H} = \mathcal{H}_{\text{volume}} + \mathcal{H}_{\text{adhesion}} = \lambda_v \sum_\sigma (V_\sigma - V_{\sigma,0})^2 + \sum_{\mathbf{x},\mathbf{x}'} \frac{J_{c(\sigma_\mathbf{x}),c(\sigma_{\mathbf{x}'})}(1 - \delta_{\sigma_\mathbf{x},\sigma_{\mathbf{x}'}})}{|\mathbf{x} - \mathbf{x}'|}, \quad (3.1)$$

where V_σ and $V_{\sigma,0}$ are the volume/area (number of lattice sites with $\sigma(\mathbf{x}) = \sigma$) and ideal volume/area of cell σ respectively, λ_v represents the strength of the volume constraint and the first sum is taken over all cell spins $\sigma > 0$. For the adhesion term the sum is taken over all neighboring sites $\mathbf{x}$, $\mathbf{x}'$ with $J_{i,j}(= J_{j,i})$ denoting the adhesion coefficient between cell types i and j, while $\delta_{i,j}$ corresponds to the Kronecker delta which ensures that only lattice site pairs of different cells contribute to the surface energy.

In comparison to the original Hamiltonian we add a factor $|\mathbf{x} - \mathbf{x}'|$ in the denominator of $\mathcal{H}_{\text{adhesion}}$ to account for the fact that a larger distance between neighbors gives a weaker contribution to the interface energy ($\mathcal{H}_{\text{adhesion}}$). Moreover, we are free to choose different neighborhoods around each lattice site which will contribute to the sum in $\mathcal{H}_{\text{adhesion}}$. In this thesis we have chosen the neighborhoods which are shown in figs. 3.3c and 3.3d for the 2D and 3D case, respectively, which from now we will refer to as the adhesion neighborhoods.

So far we have set up the manner in which a system is able to evolve within the CPM. The only thing it lacks in order to actually quantify cell dynamics is a time measure for the system. Normally, the Monte Carlo Step (MCS) is used to this end which is defined as N_l elementary Monte Carlo steps or copy attempts, with N_l the total number of sites in the lattice [36, 37, 44]. This method ensures that on average each lattice site
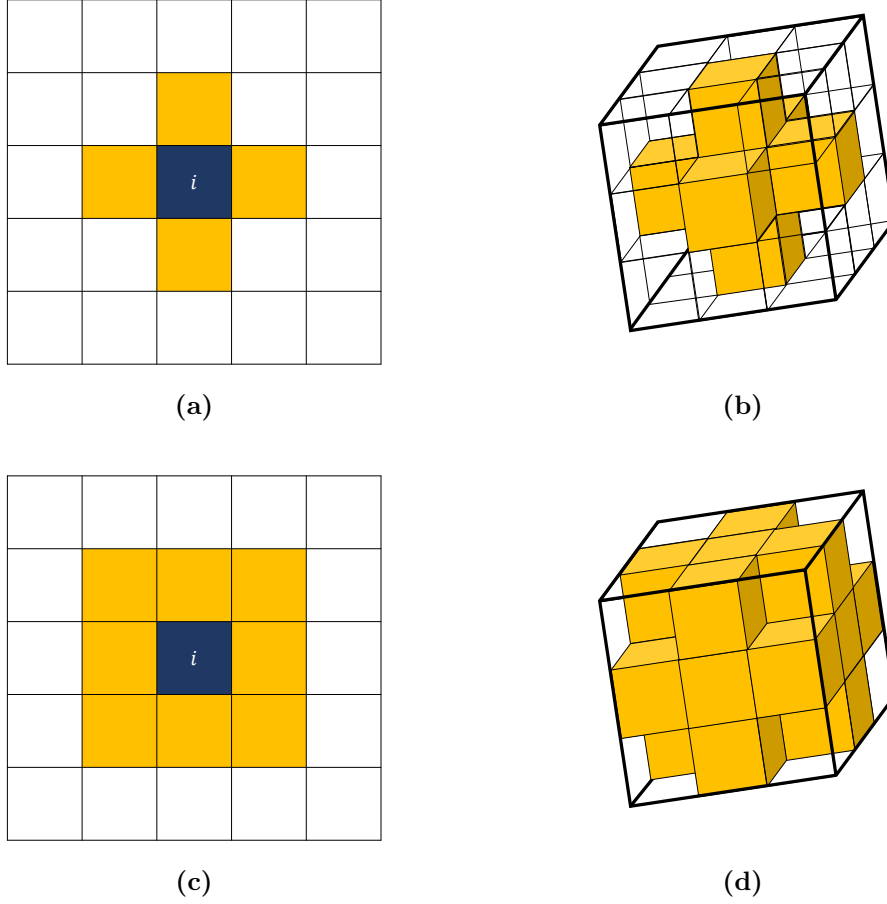
(a)

(b)

(c)

(d)

Figure 3.3: Visualisation in 2D and 3D of the adjacent neighbors from which one is randomly chosen in de modified Metropolis algorithm (a-b) and the neighbors that contribute to the adhesion term of the Hamiltonian $\mathcal{H}_{adhesion}$ (c-d). Neighborhoods are shown in orange and correspond to the lattice site i in the center of the grid (lattice site is not visible in 3D).

is updated once every MCS. The advantage of it is that we decouple the time step from the actual system size [44].

To save computation time we have slightly deviated from this definition of the time measure without changing the underlying principle. Since we seek to simulate small groups of cells or even single cells, they will mostly take up only a small portion of the entire lattice; that is, a large amount of the N_l copy attempts will involve medium/ECM ($\sigma = 0$) surrounded on all adjacent sides by medium/ECM which will not change the configuration of the system when attempting a spin-copy ($\sigma(\mathbf{x}) \rightarrow \sigma(\mathbf{x}')$). In order to skip these attempts and thereby save computation time, we make a list of all the (candidate) sites, which have at least one adjacent neighbor with a different spin (so they are able to change the system configuration), and randomly select one of these sites at step (i) in the modified Metropolis algorithm. Our time measure, the MCS, will then be defined as the amount of copy-attempts equal to the number of sites in the list. This ensures that on average each lattice site is updated once every MCS, which is consistent with the original definition of the MCS. We therefore expect no substantial difference in time evolution when using either definition of the MCS.

Note that this does require that we update the list of candidate sites after every update ($\sigma(\mathbf{x}) \rightarrow \sigma(\mathbf{x}')$) which implies that the amount of copy attempts a MCS consists of also changes after each time step. One must therefore keep in mind that when the

candidate list becomes too large relative to the lattice size (the system becomes too full), the computational gain of attempting fewer moves can be undone by the computational loss of updating the list at each attempt. For the simulations presented in this thesis, however, the newly proposed time measure has sped up simulation time significantly.

Finally, as mentioned earlier, the CPM in its original formulation has only been applied to model cell sorting. In order to adapt it to a broader variety of different biological systems and specify cell characteristics, the CPM has been extended over recent decades. The use of the Hamiltonian $\mathcal{H}$ allows for large flexibility since we can implement additional effects via extra energy terms including e.g. preferred anisotropic cell shapes. In the following subsections we will introduce the extensions to the CPM that have been used throughout this research.

3.1.2 Activity and Persistence

Cell dynamics in the CPM are controlled by the Hamiltonian $\mathcal{H}$. In its standard form presented in eq. (3.1) the dynamics will only arise from fluctuations in the cell volume and interfacial surface. As a result the cells do not experience any directional bias. In real life however, cells are able to actively move and experience biased motion. This can be an intrinsic characteristic of the cell (active cell motion) or guided via external signals in its environment. An example of the latter is the tendency for cells to bias their motion towards a concentration gradient often referred to as chemotaxis. It can be included in the CPM by incorporating an additional energy bias $\Delta\mathcal{H}_c$ in the change of the Hamiltonian $\Delta\mathcal{H}$ (calculated at step (iii) in the Metropolis algorithm) [46–48]:

$$\Delta\mathcal{H}_c = -\lambda_c \left(C(\mathbf{x}) - C(\mathbf{x}') \right), \tag{3.2}$$

where $C(\mathbf{x})$ is the local concentration of the chemicals that drive chemotaxis at position $\mathbf{x}$ and $\lambda_c > 0$ denotes the relative strength of the chemotactic motion. Since the direction of motion during a copy attempt is $\mathbf{x} - \mathbf{x}'$, we can see that moving up the gradient leads to a negative energy bias (making the move more favorable) and vice versa.

Similarly, one can also incorporate an additional energy bias $\Delta\mathcal{H}_p$ to model active motion of cells, which makes copy attempts that move the cell in its preferred internal direction more likely. We can implement the bias as [41, 45, 49]

$$\Delta\mathcal{H}_p = -\sum_{i=\sigma(\mathbf{x}),\sigma(\mathbf{x}')} \kappa_i \left(\Delta\mathbf{R}_i(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) \cdot \mathbf{p}_i \right). \tag{3.3}$$

Here $\mathbf{p}_i$ denotes the polarity (unit) vector of cell i, i.e. the direction in which the cell wants to move, $\Delta\mathbf{R}_i(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}'))$ is the center-of-mass displacement of cell i as a consequence of the proposed copy attempt and $\kappa_i > 0$ is the relative strength of active motion which effectively controls the speed of cell i. Note that κ_i can be chosen to depend on the local position of the attempt $\mathbf{x}$. This will be important when simulating durotaxis, where we let this value increase linearly in space to impose a cell speed gradient. Thus, when $\Delta\mathbf{R}_i(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}'))$ is parallel to its preferred direction, we have $\Delta\mathcal{H}_p < 0$ which makes the move more likely. In other words, there is bias along $\mathbf{p}_i$. We also mention that only cells are allowed to perform active motion and therefore the sum in eq. (3.3) only applies to $\sigma(\mathbf{x}), \sigma(\mathbf{x}') > 0$, i.e. we assume the ECM to be inactive.

Implementing cell activity via eq. (3.3) does not differentiate between the ways cells are able to actively move when surrounded by other cells or when they are surrounded by the medium or ECM. In general cells are able to actively move by pulling on their surroundings via extending and retracting structures like filopodia, lamellipodia or pseudopodia (actin filled protrusions of the cell membrane) [50, 51]. In the case of a cell-cell

contact these protrusions have been shown to be inhibited, hence the ability to pull is decreased [52–54]. Furthermore, we can imagine that when both cells mostly pull on each other (instead of a substrate in 2D for instance), they will move very little since their applied forces will balance (provided the cells are able to roughly generate the same amount of force which we can expect for cells of the same type). This notion is supported by recent experiments on cell monolayers and small cell clusters which seem to suggest that the traction forces the cells apply to move around are for the most part located at the cell cluster edges and less traction is visible between cells [18,19,55,56] (see fig. 3.4 for more details). We can therefore also assume that active motion will be more apparent and easier when cells are in contact with the medium or ECM, rather than places where they border on other cells. Taking this into account we can reformulate the activity energy bias as follows

$$\Delta\mathcal{H}_p = \begin{cases} -\kappa_{\sigma(\mathbf{x})}\Delta\mathbf{R}_{\sigma(\mathbf{x})}(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) \cdot \mathbf{p}_{\sigma(\mathbf{x})}, & \text{if } \sigma(\mathbf{x}) > 0 \wedge \sigma(\mathbf{x}') \leq 0 \\ -\kappa_{\sigma(\mathbf{x}')}\Delta\mathbf{R}_{\sigma(\mathbf{x}')}(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) \cdot \mathbf{p}_{\sigma(\mathbf{x}')}, & \text{if } \sigma(\mathbf{x}') > 0 \wedge \sigma(\mathbf{x}) \leq 0 \\ 0, & \text{otherwise} \end{cases} , \quad (3.4)$$

where we neglect the active energy bias for moves that involve two cells $(\sigma(\mathbf{x}), \sigma(\mathbf{x}') > 0)$. This implies that cells are only able to actively move at places where they are in contact with the medium/ECM; cells within the bulk (completely surrounded by other cells) do not perform active motion.

Defining $\Delta\mathcal{H}_p$ by eqs. (3.3) and (3.4) thus allows us to study two extremes in the manner in which cells can travel collectively: one where active forces are completely focused at cell cluster edges, and one where they are focused on every cell edge. Moreover, note that when simulating a single cell there is no distinction between both formulations of the active energy bias.
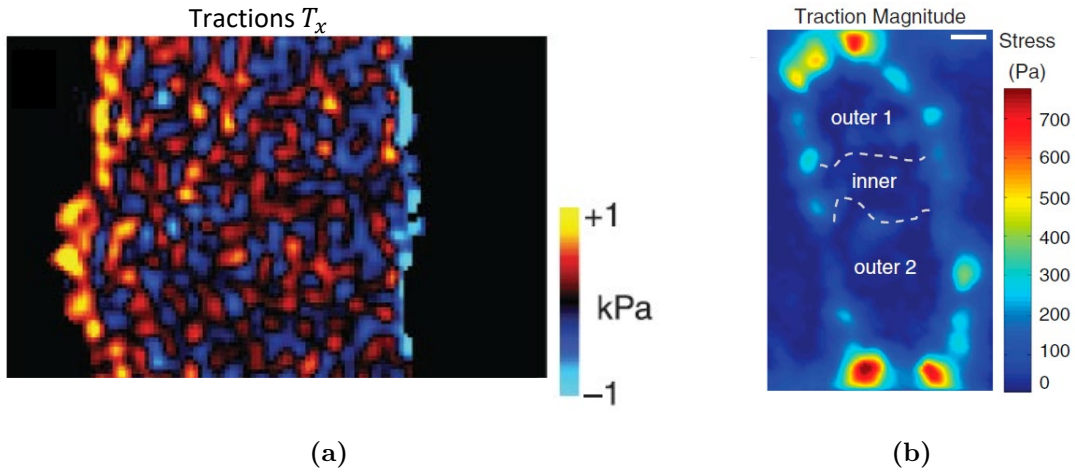
Figure 3.4: Heat maps of experimentally measured traction forces the cells exert on the underlying collagen coated substrate gel. Results correspond to a monolayer of cells (a) and a small cluster of three cells (b). Both images show that traction forces are more concentrated at cluster edges. Results are taken from [18, 55].

So far we can impose a direction of motion and velocity on the cells via $\mathbf{p}_i$ and κ_i respectively. Instead of travelling in a constant direction, however, isolated cells in the absence of signals often exhibit persistent motion in experiments. This means that the direction of motion remains roughly constant over a time scale characterised by the persistence time τ_p, but eventually reorientates into a different (random) direction. It

has been shown that single cell motility in 2D can be accurately described by a persistent random walk (PRW), i.e. eq. (2.10), which originates from self-propelled (active) persistent motion [57, 58]. Although this description is not valid anymore for 3D single cell motion, we note that the correct description in 3D involves an anisotropic persistent random walk model where two persistent random walks for a primary and nonprimary direction of motion of the environment are combined [58].

Overall it seems to be a logical starting point to implement the description of a persistent random walk into the CPM. Since we have determined that active Brownian motion leads to persistent random walks in both 2D and 3D and originates from self-propelled persistent motion (see section 2.1), we can draw a parallel between the cell polarity vector $\mathbf{p}_i$ which controls the direction of the cell velocity and the director $\mathbf{e}(t)$ describing the direction of particles performing active Brownian motion. Similar as before, we describe each polarity vector in terms of a polar angle ϕ_i in 2D using

$$\mathbf{p}_i(t) = \begin{pmatrix} \cos \phi_i \\ \sin \phi_i \end{pmatrix}, \tag{3.5}$$

and in terms of the spherical angles ϕ_i and θ_i in 3D using

$$\mathbf{p}_i(t) = \begin{pmatrix} \sin \theta_i \cos \phi_i \\ \sin \theta_i \sin \phi_i \\ \cos \theta_i \end{pmatrix}. \tag{3.6}$$

Discretising the Langevin dynamics of the relevant angles used for active Brownian motion, i.e. eqs. (2.4) and (2.12), we can find updating rules for the angles describing $\mathbf{p}_i$. In 2D this leads to

$$\phi_i(t + \Delta t) = \phi_i(t) + \Delta t \left(\sqrt{\frac{1}{2\tau}} \Gamma(\Delta t) \right), \tag{3.7}$$

and in 3D to

$$\theta_i(t + \Delta t) = \theta_i(t) + \Delta t \left(\sqrt{\frac{1}{\tau}} \Gamma(\Delta t) + \frac{1}{2\tau \tan(\theta_i(t))} \right), \tag{3.8a}$$

$$\phi_i(t + \Delta t) = \phi_i(t) + \Delta t \left(\sqrt{\frac{1}{\tau}} \frac{1}{\sin(\theta_i(t))} \Gamma(\Delta t) \right). \tag{3.8b}$$

Here Δt describes the timestep of the update which we set to 1 MCS. This means that after every MCS we update the involved angles using the above updating rules. The persistence time is again denoted with τ (given in units of MCS) and $\Gamma(\Delta t)$ is a random number drawn from a Gaussian distribution with zero mean and a variance equal to Δt, which represents a discretised version of the stochastic noise variable ξ_α. Note that the persistence time can also be made cell and position dependent: $\tau = \tau(\sigma, \mathbf{x})$. This will be required when simulating durotaxis, where we let this value increase linearly in space.

3.1.3 Vicsek Alignment

We have seen that with the update scheme given by eqs. (3.7) and (3.8), we can incorporate a model description of a persistent random walk into the CPM via reorientations of the polarity vectors $\mathbf{p}_i$. A physical interpretation of this vector can be sought in how the cytoskeleton is organised. It has for instance been suggested that polarity of a cell can be given by the vector connecting the cell nucleus to the centrosome [59]. More

generally, there at least seems to exist some intrinsically chiral structure in the cell that is responsible for the polarity [59]. The reorientation of the cell polarity over time can then for example be directly linked to a continuous reorganisation of this chiral structure which is known to occur for the cytoskeleton.

Although we will not go into further detail about the actual molecular details responsible for the reorientation of this chiral structure, it is reasonable to assume that when cells are in contact with each other they are able to influence the chiral structure (and thus the polarity vector $\mathbf{p}_i$) of the other cells. A manifestation of this interaction between cells is that when cells are in contact with each other they tend to align their respective polarities towards each other. We model this numerically with an adaptation of the well known Vicsek model [60, 61]. The implementation is done by extending the updating rules presented in eqs. (3.7) and (3.8) to

$$\phi_i(t + \Delta t) = \text{atan2}\left(P_y^i(t), P_x^i(t)\right) + \Delta t \left(\sqrt{\frac{1}{2\tau}}\Gamma(\Delta t)\right), \tag{3.9}$$

in 2D and

$$\theta_i(t + \Delta t) = \arccos\left(\frac{P_z^i(t)}{|\mathbf{P}^i(t)|}\right) + \Delta t \left(\sqrt{\frac{1}{\tau}}\Gamma(\Delta t) + \frac{1}{2\tau \tan(\theta_i(t))}\right), \tag{3.10a}$$

$$\phi_i(t + \Delta t) = \text{atan2}\left(P_y^i(t), P_x^i(t)\right) + \Delta t \left(\sqrt{\frac{1}{\tau}}\frac{1}{\sin(\theta_i(t))}\Gamma(\Delta t)\right), \tag{3.10b}$$

in 3D. Here we have defined

$$\mathbf{P}^i(t) \equiv \gamma \mathbf{p}_i(t) + \sum_{j(i)} \mathbf{p}_j(t), \tag{3.11}$$

(and its cartesian components $P_\alpha^i(t)$, $\alpha = x, y, z$) where the sum is taken over all cells j that are in direct contact with cell i (cell neighbors); with direct contact we mean that there is at least one site with spin i which has an adjacent neighbor (figs. 3.3a and 3.3b) with spin j. The trigonometric functions atan2 and arccos simply give the polar and spherical angles corresponding to the vector $\mathbf{P}^i(t)$ and effectively replace the variables $\phi_i(t)$ and $\theta_i(t)$ which are the polar and spherical angles corresponding to $\mathbf{p}_i(t)$. Note that using the cell neighbors slightly deviates from the original formulation of the Vicsek model which aligns all particles within an interaction radius.

In our model the polarity vector $\mathbf{p}_i$ thus takes the direction of $\mathbf{P}^i(t)$ after each time step and the weight factor γ controls the relative contribution of $\mathbf{p}_i$ to $\mathbf{P}^i(t)$ and is therefore a measure for the degree of alignment. In particular, we can notice that when $\gamma \to \infty$ only $\mathbf{p}_i(t)$ contributes to $\mathbf{P}^i(t)$ which implies that the alignment disappears and we regain eqs. (3.7) and (3.8). Also, when $\gamma = 1$ the polarity vector $\mathbf{p}_i$ takes the average direction of itself and its neighbors after each update and we have perfect alignment. This is how the alignment in the original Vicsek model is defined [60, 61].

Interestingly, it has been demonstrated that in the limit of fast angular relaxation, the 2D Vicsek model, i.e. eq. (3.9), is equivalent to the continuum description of angular alignment given by eq. (2.25) [34, 35]. This means that by implementing the Vicsek model we can draw a comparison between the numerical simulations and the theoretical results given in section 2.1.2.

3.1.4   Heterogeneous Environment

In its original formulation, the cells in the CPM are placed within or on a uniform medium (or ECM) which is labeled with a spin $\sigma = 0$ and a cell type $c(0) = 0$. Motivated by experimental cell trajectories, we have then introduced an active energy bias and random reorientations of the polarity vector $\mathbf{p}_\sigma$ to represent a total average effect that the ECM has on cell migration. In reality, however, the ECM in or on which cells reside is a highly complex environment consisting primarily of a network of macromolecules such as collagen. When migrating through the ECM, cells might have to squeeze through rigid and narrow parts of the network or have to degrade or remodel part of it. Additionally, cells are able to bind to the network by forming focal adhesions. In order to capture part of this complexity and study cell (cluster) motion in more detail, we seek to extend the description of the CPM environment in two ways.

Firstly, we include unmovable obstacles to represent rigid parts of the ECM which cells cannot overtake (degrade) and which are too stiff for them to remodel. These will be used to model narrow pores that cells have to squeeze through. In particular, we define the immutable CPM site to represent rigid materials which the cell cannot pass through. It is labeled with a spin $\sigma = -1$ and is assigned the cell type $c(-1) = -1$. The distinction between $\sigma = 0, -1$ rests in the fact that we immediately reject an attempt where the candidate site is an immutable site ($\sigma(\mathbf{x}) = -1$) such that it cannot be replaced and stays intact during the simulation. In comparison, we do proceed with moves where the target site is an immutable site ($\sigma(\mathbf{x}') = -1$). The immutable obstacle is, however, not allowed to grow and therefore we always propose to change the candidate site value ($\sigma(\mathbf{x})$) to zero when the target site value is $\sigma(\mathbf{x}') = -1$, corresponding to a local retraction of the cell. Furthermore, we neglect adhesion between both non-cellular elements ($\sigma = 0, -1$) so that we have $J_{-1,0} = 0$, but do allow for different adhesion coefficients between a cell and both non-cellular elements so that cells for instance adhere more easily to the immutable sites.

For the second extension we introduce focal adhesions into the CPM where the model cells bind to the environment (ECM) via a local energetic bond. For this we do not define an additional spin, but instead assign an adhesion label to each site on the lattice which is depicted as $\chi(\mathbf{x})$ and can take a value of either zero or one. The adhesion label then determines whether or not an energetic bond (focal adhesion) is formed when a cell takes over a site: if $\chi(\mathbf{x}) = 1$ a new bond is formed, otherwise no bond is formed. Therefore, we will refer to the sites for which $\chi(\mathbf{x}) = 1$ as binding sites.

To implement this behavior in the CPM, we include an additional term to the Hamiltonian, eq. (3.1), via

$$\mathcal{H} = \mathcal{H}_{\text{volume}} + \mathcal{H}_{\text{adhesion}} + \mathcal{H}_{\text{binding}}, \tag{3.12}$$

where we have introduced the total binding energy as $\mathcal{H}_{\text{binding}} = \sum_{\mathbf{x}} E(t_{\mathbf{x}})$ with the sum taken over all binding sites $\mathbf{x}$ which have a cellular spin $\sigma(\mathbf{x}) > 0$ so an actual bond has formed at these sites. The binding energy of each individual site is given by $E(t_{\mathbf{x}}) \leq 0$. This individual binding energy has a time dependence $t_{\mathbf{x}}$ which takes into account that a focal adhesion can develop over time due to receptors (integrins) in the cell membrane that assemble at the location of the bond thereby strengthening the bond. The variable $t_{\mathbf{x}}$ measures the time (in units of MCS) since the bond has first formed at site $\mathbf{x}$ keeping track of how old each bond is and how much the bond has matured. Also, note that the individual binding energy $E(t_{\mathbf{x}})$ is only a function of time which implies that we assume that the bond energy is independent of the location on the lattice and the same for each cell.

The added effect of $\mathcal{H}_{\text{binding}}$ to the simulation then comes from the fact that each time we propose a move where a cellular spin is copied to a binding site or is being replaced at it, we add or remove the relevant individual binding energies $E(t_{\mathbf{x}})$ which leads to shifts in $\mathcal{H}$. In other words detaching the cell from the binding site costs energy and can become more difficult over time, while attaching to the binding leads to an energy gain. Moreover, we want to highlight the three scenarios in which these processes can take place. At first, suppose we have an accepted spin-copy attempt that involves a binding site $\mathbf{x}$ with a cellular spin value of $\sigma(\mathbf{x}) > 0$ that is replaced by a target site (adjacent neighbor) which has a spin $\sigma(\mathbf{x}')$. If the target site is non-cellular ($\sigma(\mathbf{x}') \leq 0$) the cell at site $\mathbf{x}$ retracts and the bond is broken. This means that the relevant $E(t_{\mathbf{x}})$ is removed from the total binding energy $\mathcal{H}_{\text{binding}}$. However, when the target site represents a different cell $\sigma(\mathbf{x}') > 0$, the bond is broken but immediately replaced by a new one with the other cell. As a consquence we substitute $E(t_{\mathbf{x}})$ by $E(t_{\mathbf{x}} = 0)$, i.e. $t_{\mathbf{x}}$ is reset to zero and the bond has to mature again. Finally, a cellular target site can also copy its spin to a binding site $\mathbf{x}$ with non-cellular spin and create a new bond at this site without removing a previous one. As a result we add a new binding energy $E(t_{\mathbf{x}} = 0)$ to $\mathcal{H}_{\text{binding}}$.

Thus, we can interpret the proposed implementation as cells sliding over the binding sites, while they are able to continuously attach to and detach from them. This means that the implementation can only be applied to 2D CPM simulations for which it makes sense that cells are able to pass over or along a binding site which it occupies. Naturally, we can link it to cells that move on a substrate where they can attach to the subtrate at specific locations. However, it may also serve as a simplification of 3D cell motion inside a network of (mostly) collagen fibers, i.e. the actual ECM. The simulation then represents a plane section through a 3D structure where the binding sites represent adhesion sites of the network (fibers) on either side of the cells where they can form a focal adhesion.

On a final note, we mention that we refrain from providing an actual form of the individual binding energy as function of time $E(t_{\mathbf{x}})$ at this point. Instead we seek to find the functional form of this function via less coarse grained Langevin Dynamics simulations. This is discussed in more detail in section 3.2.

3.1.5 Connectivity

In the original CPM the possibility exists that one cell or the medium to fragment another cell into multiple pieces. A simple example of such a fragmentation is depicted in fig. 3.5a. When simulating for long times and at high enough temperature this can eventually lead to a large and unphysical amount of small fragments. Since we neglect cell division in our simulations, which circulating tumor cells are not expected to undergo when travelling to a secondary site [7, 14–16], we adopt the algorithm presented in [43]. This forbids cell fragmentation entirely via a local connectivity test. Additionally, the algorithm also ensures that the cells remain simply connected which is also more realistic for cells in non-pathological situations [43]. Moves in which cells will become multiply connected are therefore forbidden as well (see fig. 3.5b for such a move). The connectivity algorithm consists of the following two steps:

(1) Check the local connectivity of the candidate cell at site i (candidate site). If it is locally connected or $\sigma(\mathbf{x}) \leq 0$, proceed to the following step. If not, go to the final step of the algorithm.

(2) Check the local connectivity of the target cell at site i (target site). If it is locally

connected or $\sigma(\mathbf{x}') \leq 0$, proceed to the following step. If not, go to the final step of the algorithm.

Note that the candidate cell is the domain on the lattice with $\sigma = \sigma(\mathbf{x})$ and the target cell the domain on the lattice with $\sigma = \sigma(\mathbf{x}')$. Moreover, the connectivity algorithm only needs to be applied to cells and therefore if either the candidate or target site has non-cellular spin value ($\sigma \leq 0$) we always proceed. The connectivity constraint can then be implemented by inserting these two steps in between steps (ii) and (iii) of the original algorithm (see section 3.1.1).
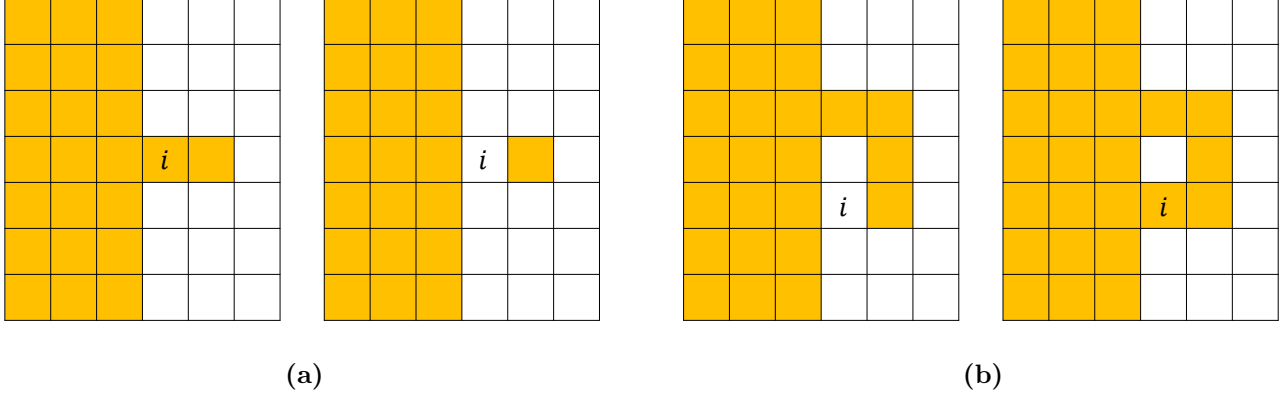

(a)   (b)

Figure 3.5: Schematic representation in 2D of (a) fragmentation of a cell and (b) simply connected cell becoming multiply connected by changing the spin at site i. The cell is shown in orange and the medium in white. Both moves are forbidden due to the imposed connectivity constraints. Note that if the white sites denote another cell, the second process also leads to fragmentation of this cell.

For the local connectivity test we will make use of the domains around the relevant candidate site i depicted in fig. 3.3. The test comes down to checking if all adjacent neighbors of site i (figs. 3.3a and 3.3b) with the same spin value as the tested cell (candidate or target) are connected to each other via paths in the adhesion neighborhood (figs. 3.3c and 3.3d) [43]. If this is the case, then the local connectivity is positively checked. The test effectively checks whether or not site i connects at least two local separate domains of the tested cell or not. With this information we can thus find out whether the candidate cell is locally cut in multiple pieces or if locally separate target cell domains are attached to each other by the attempt. Finally, the use of a local connectivity check saves much computing time in comparison to a global one in which we go over the entire cell and test whether it remains connected.

3.1.6 Shape Constraints

As shown in the Hamiltonian $\mathcal{H}$, we seek to model cell-cell attachment via the adhesion coefficient $J_{i,j}$. In particular, by setting the adhesion coefficient between cells ($J_{1,1}$) [letting all cells have the same type $c(\sigma > 0) = 1$] to a small enough value relative to the one between cells and the medium ($J_{0,1}$), it becomes energetically more favorable for cells to form a surface with other cells instead of with the medium. However, when this difference becomes too large or the cell-cell adhesion $J_{1,1}$ becomes negative, the cells will be able to easily create interfacial area with the other cells which can lead to a disintegration of the cell shape (see fig. 3.6a). This is clearly unphysical behavior. To prevent it from happening, we impose a shape constraint on the cells that forces the

cells to have a circular or spherical shape. We can interpret the constraint as a bending rigidity of the cells and formulate it in the form of an energy bias given by [62]

$$
\Delta \mathcal{H}_r = \left\{ \begin{array}{ll} \lambda_r \left(r_{\sigma(\mathbf{x})} - \left| \mathbf{x} - \mathbf{R}_{\sigma(\mathbf{x})} \right| \right), & \text{if } \sigma(\mathbf{x}) > 0 \wedge \sigma(\mathbf{x}') \leq 0 \\ -\lambda_r \left(r_{\sigma(\mathbf{x}')} - \left| \mathbf{x} - \mathbf{R}_{\sigma(\mathbf{x}')} \right| \right), & \text{if } \sigma(\mathbf{x}) \leq 0 \wedge \sigma(\mathbf{x}') > 0 \\ -\lambda_r \left(\left(r_{\sigma(\mathbf{x}')} - \left| \mathbf{x} - \mathbf{R}_{\sigma(\mathbf{x}')} \right| \right) - \left(r_{\sigma(\mathbf{x})} - \left| \mathbf{x} - \mathbf{R}_{\sigma(\mathbf{x})} \right| \right) \right), & \text{if } \sigma(\mathbf{x}) > 0 \wedge \sigma(\mathbf{x}') > 0 \end{array} \right. .
$$

$$(3.13)$$

Here λ_r denotes the relative strength of the constraint and r_σ is the preferred radius of cell σ so that its area or volume fits precisely in a circle or sphere respectively. $|\mathbf{x} - \mathbf{R}_\sigma|$ denotes the length of the vector that points from the center of mass of cell σ, i.e. $\mathbf{R}_\sigma$, to the location of the candidate site $\mathbf{x}$ and can be seen as a local cell radius at the site $\mathbf{x}$. Note that the function only applies to cells ($\sigma > 0$). The added effect of eq. (3.13) is shown in fig. 3.6.

We can explain the form of the energy bias by noting that during each attempt we want to replace the candidate site value $\sigma(\mathbf{x})$ by the value of the target site $\sigma(\mathbf{x}')$. This means that the candidate cell locally retracts at its location $\mathbf{x}$, while the target cell locally extends towards $\mathbf{x}$. The bias checks whether or not the extension or retraction moves the local cell radius ($|\mathbf{x} - \mathbf{R}_\sigma|$) towards or from the preferred radius of the cell r_σ. It then gives a negative energy bias for moves towards the preferred radius making them more favorable. The strength of the energy bias scales with the difference between the local and preferred cell radius; that is, when this difference is large, the cell is more deformed and is therefore more likely to move towards the preferred radius.
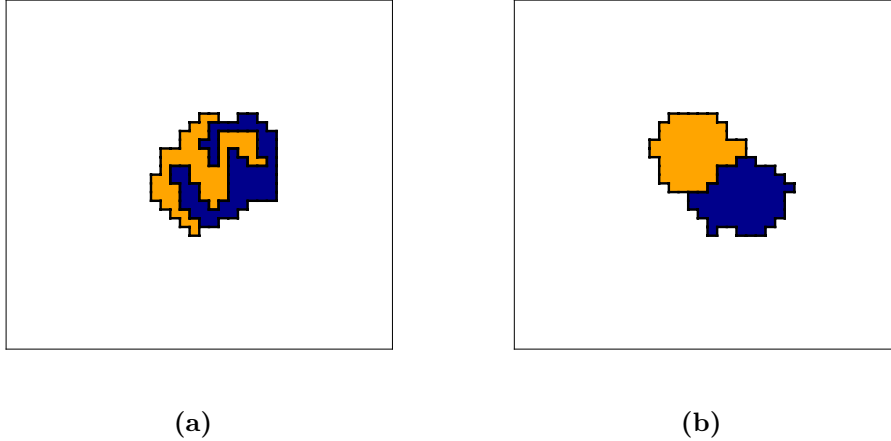

<div align="center">(a)   (b)</div>

Figure 3.6: Visualisation in 2D of the final cell configurations of a 2 cell cluster after 500 MCS. Simulations are performed for a standard set of parameters except for a zero adhesion coefficient between cells ($J_{1,1} = 0$). (a) Only the original Hamiltonian, eq. (3.1), and connectivity constraint have been used. (b) Same settings as (a) but including the circular shape constraint, eq. (3.13).

Finally, we mention that this constraint is not necessarily limited to a circular or spherical shape. By introducing an angular dependence in the preferred radius, other shapes can be obtained as well. Note that the input angles for the preferred radius are the ones that correspond to the vector $\mathbf{x} - \mathbf{R}_\sigma$. Moreover, the angles can even be taken relative to each cell's polarity vector $\mathbf{p}_i$, which can make it more consistent with the notion of an intrinsic chiral structure of the cell. For convenience, however, we will only impose a circular or spherical shape on the cells.

3.1.7 Initialization and Equilibration

Before each simulation we will also have to define a starting configuration for the cells. Because of computational convenience, we have begun each simulation by defining equal square or cubic domains (2D/3D) of unique spins $\sigma > 0$ representing the cells. The size of these domains, i.e. number of lattice sites they consist of, will define the ideal area/volume of the cells ($V_{\sigma,0}$), implying that we let each cell have the same size. These square/cubic domains are then placed next to each other in a predetermined configuration. This means that the cells will always start being attached to at least one other cell with one side of the square/cube and form only one cluster. A visualisation of such a starting configuration for both a single cell and a cluster of nine cells are shown in figs. 3.7a and 3.7b. Furthermore, placing the cells like this allows us to easily control the starting position of the center of mass of each cell $\mathbf{R}_\sigma$ and by extension the center of mass of the entire cluster $\mathbf{R}_c$.

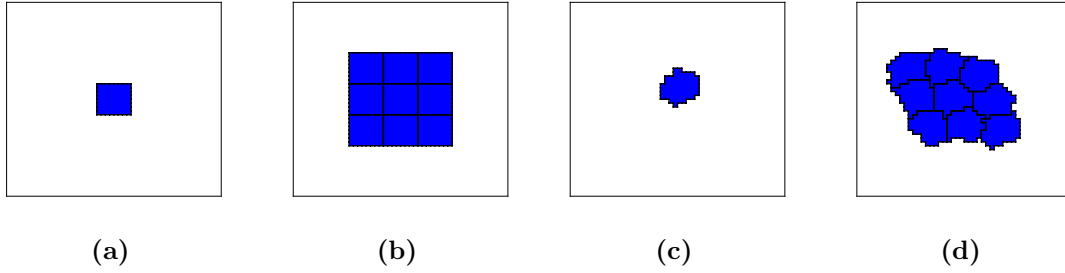| (a) | (b) | (c) | (d) |

Figure 3.7: Visualisation in 2D of the start configuration of a single cell and a small cluster of 9 identical cells before (a,b) and after (c,d) equilibration. Standard set of values have been used for the relevant parameters.

The system is then equilibrated by running a simulation for 500 MCS in which we only include the original Hamiltonian eq. (3.1) and the shape constraint eq. (3.13), while neglecting all other energy biases and interactions with binding sites. This is done to allow the cells to develop a more round and natural shape. After this time we redefine $\mathbf{R}_c = \mathbf{0}$ (which sets the coordinate system), assign (unless otherwise stated) polarity vectors drawn from a uniform distribution to each cell and let the actual simulation start from which the results can be analysed. An example in 2D of a single cell and a cluster of nine cells after equilibration are shown in figs. 3.7c and 3.7d.

3.2 Focal Adhesion

Focal adhesions are the machinery which cells employ to attach themselves to their surroundings. Specifically, they are supramolecular complexes made of integrin receptors which link the extracellular matrix (ECM) to the actin cytoskeleton of the cell [22, 23]. These complexes have assembled over time and typically have a size of the order $0.25 - 10$ μm^2 [63, 64]. Since in our simulations the resolution of the CPM will fall within the same order of magnitude as the size of a focal adhesion, we cannot describe the interaction of cells with a substrate or with the ECM explicitly. This means that we have to replace the actual interaction with an effective potential as mentioned in section 3.1.4. Instead of making a(n) (educated) guess about the precise form of this potential we seek to inform the potential from a less coarse grained and therefore more detailed simulation using Langevin Dynamics. In this simulation we will look at a model cell membrane consisting of both receptors and non-receptors which in time will attach to a cylinder fixed in space and floating above the membrane. This cylinder will represent the collagen fiber (binding site). Assuming the bond strength increases linearly with the amount of receptors attached to the cylinder (collagen fiber), the form of the potential can then be retrieved by tracking the amount of attached receptors in time.

3.2.1 Cell Membrane Model

The cell membrane is a two dimensional fluid which host a variety of lipid molecules and proteins. Since these lipids are amphylic particles, they self-assemble into mono- or bi-layer membranes when emersed in an aqueous solution [65]. This gives rise to the two-dimensional nature of the membrane that due to its fluid state has a continuously changing geometry. The most detailed approach of modelling a biological membrane is by performing an atomistic simulation where each individual atom or molecule in the system is taken into account [66]. However, in order to access larger length and time scales, which allows for better averaging, a coarse-grained model is preferred. In this case course-graining refers to the process in which a number of lipids is represented by a single bead, thereby decreasing the number of degrees of freedom and reducing computational costs. It remains, however, crucial that general membrane properties keep intact after course-graining. Additionally, to decrease computational costs even more, solvent molecules are treated only implicitly.

   In this thesis we have used the coarse-grained, solvent-free and fluid-like membrane model based on the work in [65] and [67] in order to describe the cell membrane which attaches to a collagen fiber. The model describes the membrane as a one-particle-thick fluid layer of spherical and axisymmetric particles. The axis of symmetry of each particle corresponds to the lateral direction of the lipid molecules of which it consists. This implies that the particles only carry five degrees of freedom, i.e. three translational and two rotational. The particles interact with each other via a pairwise interparticle potential given by

$$U_{ij}(\mathbf{r}_{ij}, \mathbf{n}_i, \mathbf{n}_j) = \begin{cases} u(r_{ij}) + \epsilon[1 - \phi(\hat{\mathbf{r}}_{ij}, \mathbf{n}_i, \mathbf{n}_j)], & r_{ij} < r_{min} \\ u(r_{ij})\phi((r_{ij}, \mathbf{n}_i, \mathbf{n}_j), & r_{min} < r_{ij} < r_c \end{cases}. \tag{3.14}$$

Here $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ (with definitions $r_{ij} = |\mathbf{r}_{ij}|$ and $\hat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/r_{ij}$) denotes the distance vector between the center positions of particle i ($\mathbf{r}_i$) and j ($\mathbf{r}_j$), while the unit vectors $\mathbf{n}_i$ and $\mathbf{n}_i$ represent the axes of symmetry of particle i and j respectively (see fig. 3.8 for more details). We define $r_{min} = 2^{1/6}\sigma$ as the distance at the minimum of the potential, which is based on the minimum of a 12-6 Lennard-Jones (LJ) potential, and $r_c = 2.6\sigma$

is the cut-off radius whose length is set to include up to second-neighbor interactions. Both distances are defined in terms of the length unit of the system σ, i.e. the diameter of the particles, and similarly the potential depth ϵ is described in terms of the energy unit of the system $k_B T$ with k_B the Boltzmann constant and T the temperature.

The total potential is separated in a distance dependent part $u(r_{ij})$ that forces particles to stick together and an orientation dependent part $\phi(\hat{\mathbf{r}}_{ij}, \mathbf{n}_i, \mathbf{n}_j)$ which substitutes the hydrophobic effects of the lipids. The former of these is described by a two-branch function given by

$$u(r_{ij}) = \begin{cases} \epsilon\left[\left(\frac{r_{min}}{r}\right)^4 - 2\left(\frac{r_{min}}{r}\right)^2\right], & r_{ij} < r_{min} \\ -\epsilon \cos^{2\zeta}\left[\frac{\pi(r-r_{min})}{2(r_c-r_{min})}\right], & r_{min} < r_{ij} < r_c \end{cases}, \tag{3.15}$$

where ϵ gives the depth of the potential. Furthermore, we can recognise that the repulsive branch ($r_{ij} < r_{min}$) is given by a 4-2 LJ type potential. The classical 12-6 LJ potential was found to not allow a fluid phase due to the steepness of the potential and therefore the less steep 4-2 LJ potential is preferred. The attractive branch ($r_{min} < r_{ij} < r_c$) is a cosine function that decays to zero at the cutoff radius r_c and whose exponent ζ determines how rapidly the function tends to zero (fig. 3.9). It is therefore a measure of the diffusivity of the particles.
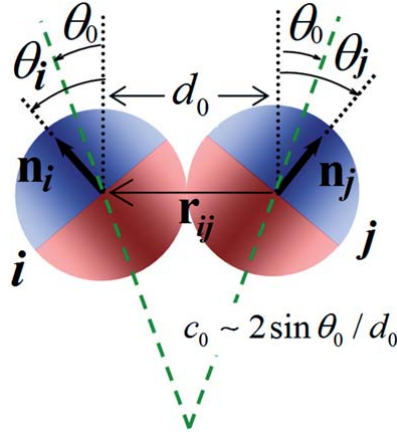

Figure 3.8: Schematic representation of two coarse-grained beads with all relevant vectors and angles displayed. $\theta_{i,j}$ denotes the angle that the the axes of symmetry $\mathbf{n}_{i,j}$ make with the line perpendicular to the connection vector $\mathbf{r}_{ij}$, whereas θ_0 depicts the spontaneous curvature angle with respect to the same line. Figure taken from [67].

The orientation dependent function is given by

$$\phi(\hat{\mathbf{r}}_{ij}, \mathbf{n}_i, \mathbf{n}_j) = 1 + \mu[b(\hat{\mathbf{r}}_{ij}, \mathbf{n}_i, \mathbf{n}_j) - 1], \tag{3.16}$$

where

$$b = (\mathbf{n}_i \times \hat{\mathbf{r}}_{ij}) \cdot (\mathbf{n}_j \times \hat{\mathbf{r}}_{ij}) + \sin(\theta_0)(\mathbf{n}_i - \mathbf{n}_j) \cdot \hat{\mathbf{r}}_{ij} - \sin^2(\theta_0). \tag{3.17}$$

Here θ_0 depicts the favorable interparticle orientations which determines the spontaneous curvature of the model. The function b reaches its maximum of 1, when $\theta_i = \theta_j = \theta_0$ (see fig. 3.8 for θ_0, θ_i and θ_j and fig. 3.9 for a plot of b). Maximizing b implies that the total potential U is minimized and therefore the relative orientation $\theta_i = \theta_j = \theta_0$ is most energetically favorable (fig. 3.9). This also allows us to interpet the parameter μ as a weight of the energy penalty when the particles are disoriented from θ_0 (see fig. 3.9)

relating it to bending rigidity of the model membrane. Finally, to clarify the functional dependence of b on its variables it is useful to confine all relevant vectors to a 2D plane (fig. 3.8). In this case b reduces to

$$b = \cos(\theta_i)\cos(\theta_j) + \sin(\theta_0)[\sin(\theta_i) + \sin(\theta_j)] - \sin^2(\theta_0). \qquad (3.18)$$

Thus, in total the model is described by a coarse-grained inter-particle interaction potential $U(\mathbf{r}_{ij}, \mathbf{n}_i, \mathbf{n}_j)$ in which the three potential parameters ζ, θ_0 and μ separately describe three key characteristics of the membrane; namely diffusivity, membrane curvature, and bending rigidity respectively.
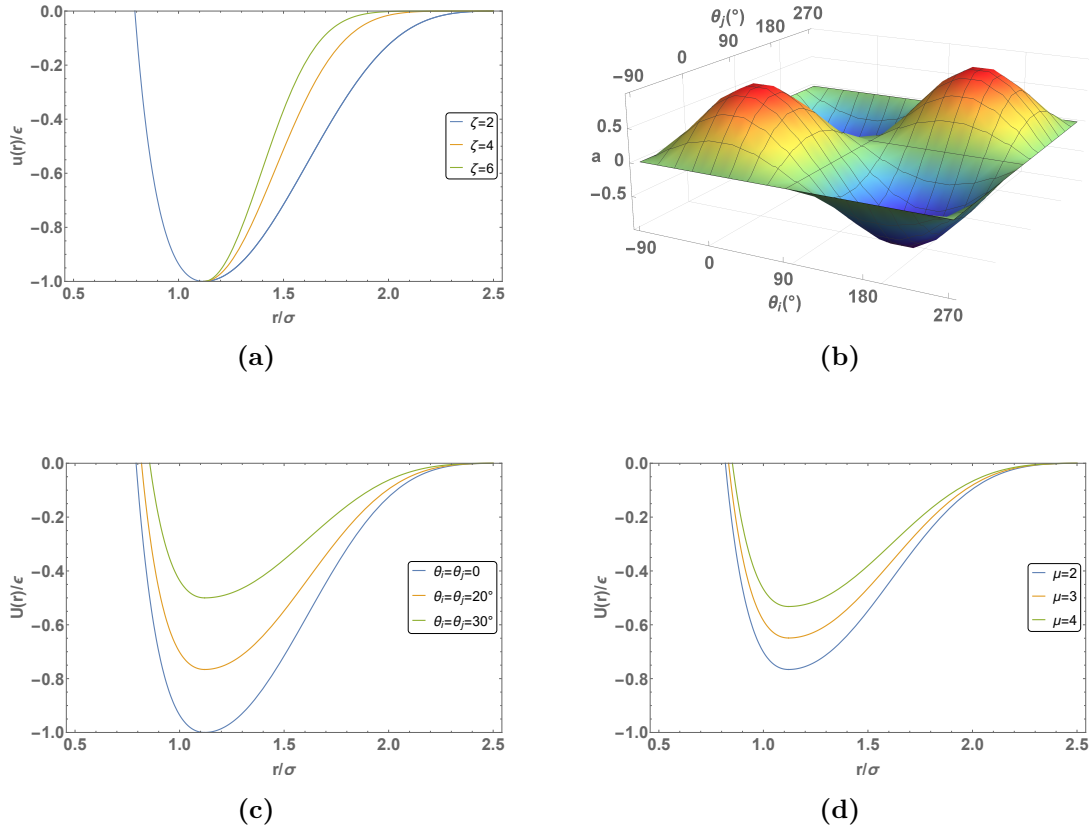

(a)

(b)

(c)

(d)

Figure 3.9: (a) Plot of the distance dependent function u as function for r/σ for different values of ζ. (b) 3D plot of orientation function b as a function of θ_i and θ_j with $\theta_0 = 0$. It shows the peak at $\theta_i = \theta_j = \theta_0 = 0$ and due to symmetry one at an equivalent configuration with $\theta_i = \theta_j = 180^o$. (c-d) Plots of the total interaction potential U with $\theta_0 = 0$ and $\zeta = 2$ as a function of r for different values of $\theta_{i,j}$, keeping $\mu = 2$ (c) and for different values of μ, keeping $\theta_i = \theta_j = 20^o$.

3.2.2 Fiber-Membrane Interaction

The introduced model above shows how the particles interact with each other to form a cell membrane. However, it still lacks a description of how the membrane particles will interact with the collagen fiber to form a focal adhesion. For simplicity, we model the collagen fiber as a rigid cylinder that floats above the membrane and remains fixed in space. To account for the excluded volume of the cylinder, we implement a repulsive,

truncated and shifted Lennard Jones $(12 - 6)$ potential which is given by [68]

$$U_{\text{LJ}}(r) = 4U_0 \left[\left(\frac{\sigma}{r + \Delta} \right)^{12} - \left(\frac{\sigma}{r + \Delta} \right)^6 + \frac{1}{4} \right], \qquad r < 2^{1/6}\sigma - \Delta. \qquad (3.19)$$

Here r is the distance between the edge of the cylinder and the center of a membrane bead, U_0 is a measure for the interaction strength and σ the length scale of the system (membrane particle diameter). Note that the factor $\Delta = 0.5\sigma$ is added to ensure that the potential already steeply rises (indicating a repulsive reaction) when the edge of a bead approaches the edge of the cylinder. The constant in U_{LJ} shifts the total potential so that it becomes zero at $r = 2^{1/6}\sigma + \Delta$. This has again been done to prevent jumps in potential energy when r passes the value of $r = 2^{1/6}\sigma + \Delta$ which will otherwise cause numerical problems with energy conservation.

Physically, a focal adhesion is formed via integrins in the cell membrane that are able to bind to the collagen fiber and eventually cluster around it. Of course the cell membrane consists of a large variety of components, but we will only distinguish between integrin and non-integrin particles. These will be indicated as receptors and non-receptors respectively. Non-receptors will only repel the collagen fiber when it comes too close (volume exclusion) and their interaction is solely described by eq. (3.19). On the other hand, receptors (integrin particles) will also try to bind to the collagen fiber which is described by a truncated and shifted Morse potential (see fig. 3.10 for a schematic representation)

$$U_{\text{morse}}(r) = D_0(e^{-2\alpha(r-r_0)} - 2e^{-\alpha(r-r_0)} - C_0), \qquad r \leq r_c. \qquad (3.20)$$

Here r is again the distance between the edge of the cylinder and the center of a membrane bead, D_0 indicates the attractive strength, α the width of the potential and $r_0 = 2^{1/6}\sigma - \Delta$ the point where the potential reaches its minimum. The value for r_0 has been chosen such the minima of both potentials coincide and the interaction due to the Morse potential becomes repulsive for $r < 2^{1/6}\sigma - \Delta$ as well. The constant $C_0 = e^{-2\alpha(r_c-r_0)} - 2e^{-\alpha(r_c-r_0)}$ shifts the potential to zero at the cut-off radius r_c.

Finally, it should be stated that describing the attraction with a Morse potential is preferred over a non-truncated Lennard-Jones potential since the width of the potential can also be varied via the parameter α. This allows for more control over the range on which receptor binding can take place, which is assumed to be an important aspect of the physical biological system.

3.2.3 Langevin Dynamics

The movement of a cell membrane attaching to a collagen fiber has been studied by means of Langevin dynamics simulations, which is a widespread approach to simulate particle dynamics in complex (biological) systems [69] [70]. Langevin dynamics effectively replaces solvent interactions by a friction force which slows down the particles and a stochastic force due to random collisions with the surrounding solvent particles. These effects keep the total temperature of the system constant and lead to the so called Langevin equation of motion for the position $\mathbf{r}_i$ of a particle i (membrane beads) [71]

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = -\nabla_i V(\mathbf{r}_1, ..., \mathbf{r}_N, t) - \lambda \frac{d\mathbf{r}_i}{dt} + \boldsymbol{\eta}_i(t) \qquad (3.21)$$

where m_i is the mass of the particle, ∇_i the gradient operator working on coordinates $\mathbf{r}_i$, λ the friction coefficient, V the total potential of the system (sum of the interaction
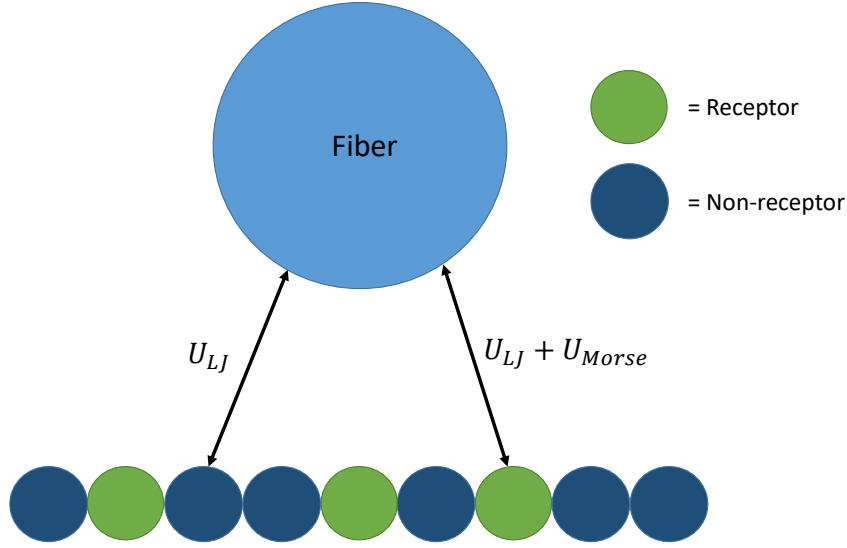
Figure 3.10: 2D (yz-plane) schematic representation of the interaction between the membrane and the collagen fiber. The mentioned potentials have been discussed in text.

potentials between all N particles) and $\boldsymbol{\eta}_i(t)$ a stochastic force. This force is a Gaussian-distributed random vector with zero mean and correlation given by $\langle \eta_i^\alpha(t)\eta_i^\alpha(t') \rangle = 2m_i k_B T \lambda \delta(t - t')$ for each of its cartesian components $\alpha = x, y, z$ [71].

By numerically solving the Langevin equations of motion for consecutive timesteps Δt, we obtain the trajectory for each membrane particle i. Generating multiple runs and using the retrieved trajectories, we are then able to determine the average number of receptors that have bound to the cylinder (collagen fiber) $N_{\text{bound}}(t)$, i.e. the receptors that are within one particle diameter σ of the cylinder surface, as a function of time t. As mentioned, we assume that the energy $E(t_{\mathbf{x}})$ of a binding (focal adhesion) site in the CPM (see section 3.1.4) is expected to scale linearly with the amount of bound receptors such that we have

$$E(t_{\mathbf{x}}) = \beta N_{\text{bound}}(t_{\mathbf{x}}), \tag{3.22}$$

with β a proportionality constant. This result states how the results we obtain from the Langevin Dynamics simulations are inserted into the CPM in order to improve the realism of the CPM environment and allow for more detailed simulations. Finally, we highlight that the time origin $t = 0$ for each Langevin simulation is defined at the timestep before the first receptor binds to the cylinder. This has been done because it can take different amounts of time before the first receptors binds to the cylinder and we are mainly interested in how the assembly of receptors at the cylinder takes place (maturation of the focal adhesion) and not how long it takes before it is initiated.

3.2.4   Numerical Details

After introducing the relevant potentials and equation of motion of the membrane particles, we will now provide some numerical details with respect to the precise implementation of the simulations. At the start of each simulation, we place a total of 7020 membrane particles on a hexagonal lattice in the xy-plane at $z = 0$ and randomly label a

chosen percentage of the membrane particles as receptors. The cylinder is placed above the membrane such that its center is placed along the x-axis from $[-35, 35]$ (slightly smaller than the simulation box domain) at $y = 0$ and $z = r_{cyl} + r_c$ with r_{cyl} the radius of the cylinder and r_c the cut-off radius of the attractive Morse potential (eq. (3.20)) such that the membrane lies just within the attractive range of the cylinder.

Numerical time integration is performed in the isoenthalpic-isobaric (NPH) ensemble for a total of $2 * 10^6$ timesteps which have been chosen at $\Delta t = 0.02\tau_{LJ}$ with $\tau_{LJ} = (m\sigma^2/k_B T)^{1/2}$ the time unit of the system (m denotes the mass of a membrane particle). Within this ensemble the pressure is kept fixed by means of a Nose-Hoover barostat [73] [74], while the volume (box size) fluctuates in time. This has been done to incorporate the fact that the cell membrane is continuously bending and stretching and thus a segment of the membrane, which has been modelled, is expected to change in size over time, but tension in the membrane is assumed to be constant leading to a constant pressure. Note that by performing a Langevin simulation we have also kept the temperature fixed.

Furthermore, we repeat that membrane interactions with the solvent above or cytoplasm below it is only modeled implicitly via a stochastic force. This replacement does, however, not take in to account that the fluids will exert a pressure on both sides of the membrane which we assume to be in equilibrium and therefore tries to keep the membrane at a fixed height. To account for this we have also applied a tethering force to all the particles in the membrane during the simulation that ensures that the center of mass of the entire membrane remains at $z = 0$. Additionally, rotational degrees of freedom have been included for the membrane particles, because the potential between them also depends on the orientation of the particles as well (see eq. (3.16)). Since the particles are axisymmetric only two rotational degrees of freedom need to be added.

Finally, we indicate that the Langevin simulations have been carried out using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) [72]. This program is able to numerically solve the Langevin equation for consecutive timesteps Δt by using the method described in [71]. For more details about the precise implementation the reader is referred to the used LAMMPS script, which has been added in appendix B.

4. Methods of Analysis

The results of the CPM to describe and quantify collective cell migration will (mostly) come in the form of the center-of-mass trajectories of the individual cells $\mathbf{R}_\sigma(t)$ and the cell cluster $\mathbf{R}_c(t)$ as function of time, while the Langevin Dynamics simulations give the trajectories of all particles in the membrane $\mathbf{r}_i(t)$. In this chapter we briefly discuss how we extract several important variables and functions from the trajectories which have been used for the discussions in the next chapter.

4.1 Mean (Square) Displacement

The first quantitive analysis of stochastic processes in either experiments or simulations is often carried out by calculating the mean square displacement (MSD), $\langle \mathbf{r}(t)^2 \rangle$, or simply the average position, $\langle \mathbf{r}(t) \rangle$. We can calculate both variables in several ways. The first method is by taking an ensemble average over all performed trajectories which translates to [75]

$$\langle \mathbf{r}(t_n)^m \rangle_{\text{ens}} = \frac{1}{N_T} \sum_{j=1}^{N_T} (\mathbf{r}_n^{(j)} - \mathbf{r}_0^{(j)})^m. \tag{4.1}$$

Here $t_n = n\Delta t$ gives the time, which is specified via an integer n and the timestep used in the simulation Δt, $\mathbf{r}_n^{(j)}$ denotes the center-of-mass position of a particle or cell at time $n\Delta t$ corresponding to trajectory j, N_T the number of trajectories and the variable $m = 1, 2$ specifies whether we calculate the average position or MSD. Note that each CPM simulation run leads to one trajectory for the cell cluster and hence ensemble averaging of the cell cluster comes down to averaging over each separate run. In comparison, for the Langevin Dynamics simulations mimicking focal adhesions, every run creates a large amount of trajectories for the membrane particles; one for every particle in the membrane. Thus, the averaging of membrane particles is done over all particles and runs.

In order to perform proper statistical averaging a large number of trajectories is required. Since we can only generate one trajectory per run for most CPM simulations, this can lead to high computational costs. To circumvent this problem one usually takes the temporal average over all possible discrete time intervals of particle trajectory assuming that the system is time-translational invariant. The advantage of this method is that a large number of points can be used for averaging. This results in [75]

$$\langle \mathbf{r}(t_n)^m \rangle_{\text{time}} = \frac{1}{(N-n)} \sum_{i=0}^{N-n} (\mathbf{r}_{i+n} - \mathbf{r}_i)^m, \tag{4.2}$$

with N is the total number of timesteps taken in the trajectory. We denote both types of averages differently and highlight that the theoretical average is always an ensemble average.

Furthermore, after temporal averaging an additional ensemble average can also be performed over all N_T trajectories in order to further improve the averaging and filter out the possible influence of the fluctuations in the starting configuration of the simulation. In total this combines eq. (4.1) and eq. (4.2) to give the time-ensemble-averaged mean (square) displacement [75]

$$\langle\langle\mathbf{r}(t_n)^m\rangle_{\text{time}}\rangle_{\text{ens}} = \frac{1}{N_T(N-n)} \sum_{j=1}^{N_T} \sum_{i=0}^{N-n} (\mathbf{r}_{i+n}^{(j)} - \mathbf{r}_i^{(j)})^m, \tag{4.3}$$

with all variables specified earlier. It should, however, be stressed that temporal averages only coincide with the ensemble average when the system is ergodic so that the temporal average converges towards the ensemble average when $N \to \infty$. This is not necessarily true for all systems.

After retrieving the MSD it is also convenient to calculate the so called instant diffusion coefficient. This is the diffusion coefficient that follows from assuming normal diffusion and can therefore be calculated via [76]

$$D_{\text{in}}(t_n) = \langle\mathbf{r}(t_n)^2\rangle/2dt_n, \tag{4.4}$$

where d denotes the spatial dimension. The instant diffusion coefficient proves to be a convenient tool to distinguish between anomalous (MSD $\propto t^\alpha$, $\alpha \neq 1$) and normal diffusive motion (MSD $\propto t$). In the case of anomalous diffusion it will either increase (superdiffusion/ballistic motion) or decrease (subdiffusion) over time, whereas the diffusion constant remains constant for normal diffusion. Additionally, it can demonstrate transitions between types of diffusion over the course of time.

The average displacement is normally used to show if a particle experiences drift over time. In order to quantify drift it is instructive to introduce the durotactic vector index [24, 77]

$$\mathbf{DI}(t_n) = \langle\mathbf{r}(t_n)\rangle/v_0 t_n, \tag{4.5}$$

where v_0 represents a reference speed against which the drift is compared, for example the speed of the particle. The index thus provides the fraction of the drift velocity of the particle relative to the chosen reference speed. Moreover, note that when we set $v_0 = 1$, eq. (4.5) simply gives the average drift velocity of the particle over time, which means that it is proportional to this velocity.

4.2   Velocity Autocorrelation Function

The retrieved trajectories also allow us to determine the particle velocity at time t which is calculated over a time period δ using [75, 78]

$$\mathbf{v}(t) = \frac{1}{\delta}\left(\mathbf{r}(t+\delta) - \mathbf{r}(t)\right). \tag{4.6}$$

To be able to further study underlying mechanisms of the particle's motion, one can calculate the normalised velocity autocorrelation function which is defined as [75, 78]

$$C_v^\delta(t) = \frac{\langle\mathbf{v}(t)\cdot\mathbf{v}(0)\rangle}{\langle\mathbf{v}(0)\cdot\mathbf{v}(0)\rangle}. \tag{4.7}$$

The involved averages can be taken in the same manner as described for the mean (square) displacement in section 4.1. However, for time averaging we have to take into

account that the velocity is defined over a time period δ which slightly limits the range over which can be time averaged. In particular, the time average is given by

$$\langle \mathbf{v}(t_n) \cdot \mathbf{v}(0) \rangle_{\text{time}} = \frac{1}{(N - n - k)} \sum_{i=0}^{N-n-k} \mathbf{v}(t_{i+n}) \cdot \mathbf{v}(t_i), \tag{4.8}$$

where we introduced the integer k which specifies the the time period $\delta_k = k\Delta t$ that is used in the definition of the velocity, i.e. eq. (4.6).

The time average can also be used to find the average absolute velocity $\langle v \rangle$ during a trajectory. Using the same adaptation as for eq. (4.8) we find

$$\langle v \rangle_{\text{time}} = \frac{1}{(N - k)} \sum_{i=0}^{N-k} \sqrt{\mathbf{v}(t_i) \cdot \mathbf{v}(t_i)}, \tag{4.9}$$

which, similar to the MSD, can be further ensemble averaged over multiple trajectories as well. Finally, in the case of more complex ECM-like environments, i.e. migration through pores, we will also consider mean passage times through the pore and fragmentation dynamics of the cell cluster.

Summarising, we have introduced several important statistical variables and how they can be obtained from cell (or particle) trajectories. In the following chapter these variables will be used to characterise and quantify (collective) cell motion, which allow us to investigate the difference between single cell and cluster migration and indicate when one is preferred over the other.

5. Results and Discussion

After setting up the simulation details and the different variables that can be obtained from them, this chapter will focus on analysing and discussing the results. We start by performing single cell CPM simulations to see how the numerical results compare to experimental single cell trajectories and ABP theory, and to establish which system parameters govern the observed dynamics. Next, we will compare single cell motion to the motion of small cell clusters, focusing on cell-cell alignment, durotaxis, and transport through a narrow channel. We then present the results of the Langevin Dynamics simulations in order to find the binding potential corresponding to the model focal adhesion. The chapter is finalised by comparing single cell motion to collective motion with explicit treatment of focal adhesions.

5.1 Single Cell Motion

Throughout this section we will discuss the results of both 2D and 3D CPM simulations of single cells that travel through a uniform medium without binding sites and analyse the cell center-of-mass trajectories $\mathbf{R}_\sigma(t)$ that follow from them. We are mainly interested in controlling the active (persistent) motion of the cells and comparing it to experimental cell motion in the absence of external signals. Therefore we have only varied the implemented persistence time τ and the strength of the active energy bias κ_1 (which we denote κ_p for convenience) in our simulations. Other parameters have been kept fixed. Specifically, for the 2D simulations we have set the simulation temperature $T = 1$, the ideal area of the cell $V_{1,0} = 64$, the volume and shape constraint strengths $\lambda_v = \lambda_r = 1$, and the cell-medium adhesion coefficient $J_{1,0} = J_{0,1} = 1$. In comparison, for the 3D simulations we have set $T = 1$, $V_{1,0} = 216$, $\lambda_v = 1$, $\lambda_r = 0.5$ and $J_{1,0} = J_{0,1} = 0.4$. We highlight that all the values of CPM parameters throughout the chapter have been chosen such that, consistent with experiments, cells have a typical size of $\sim 10\mu$m, speed of $\sim 50\mu$m/h and persistence time of ~ 1h [18, 24, 79]. Additionally, note that since each simulation involves only one cell, we have either spin $\sigma = 0, 1$ and cell type $c(\sigma) = 0, 1$ in the relevant subscripts of the parameters. Finally, we mention that all averages in this section have been retrieved via time-ensemble averaging over multiple trajectories and that we neglect external signals, which implies that chemotaxis has not been taken into account ($\lambda_c = 0$).

5.1.1 Two Dimensions

Let us first demonstrate how the strength of the active energy bias relates to the speed of an individual cell. To study this we have run 20 single cell simulations of 20000 MCS for different values of κ_p where we kept a constant cell polarity vector in the x-direction, $\mathbf{p}_1 = (1, 0)$. For clarity, we have visualised a part of one of the trajectories in fig. 5.1a. It shows that, as expected, the cell moves in the x-direction but does so slightly erratically

due to the fluctuations in cell shape and size. From the obtained trajectories we can find the average absolute velocity $\langle v \rangle$ by time and ensemble averaging. Since we seek to find the cell velocity due to the polarity, we choose a large value of $\delta = 1000$ MCS in the definition of the velocity (see eq. (4.6)). This is done to cancel out contributions of 'thermal' fluctuations in the cell shape that we expect to influence results on short time scales. It thus allows us to interpret the retrieved average absolute velocity $\langle v \rangle$ as the intrinsic speed of the cell that results solely from the active energy bias.

The resulting velocities as a function of κ_p have been plotted in fig. 5.1b. It can be seen that the average absolute velocity $\langle v \rangle$, i.e. the active cell speed, scales (almost) linearly with κ_p. This makes sense because in the overdamped limit the velocity should be proportional to the force or the gradient in the Hamiltonian and κ_p effectively controls the absolute value of the polarity contribution to this gradient. Thus, we can control cell speed relatively easily via κ_p.
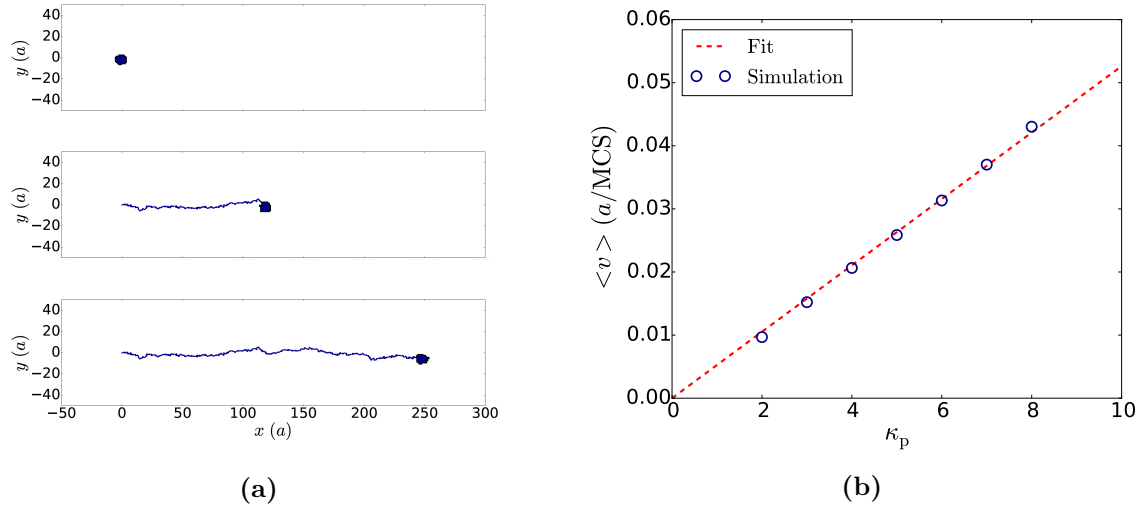

(a)   **(b)**

Figure 5.1: (a) Visualisation of a CPM cell and its trajectory where the cell experiences a constant polarity along the x-axis with a strength $\kappa_p = 5$. (b) Plot of the average cell velocity $\langle v \rangle$ as a function of the strength of the active energy bias κ_p. Linear fit to the CPM simulation results is shown as well.

Having established the relation between cell speed and the active energy bias, we now want to investigate how the theory of active Brownian motion maps onto the CPM. We therefore also allow the polarity vector to reorientate in time according to eq. (3.7) and have run 50 single cell simulations of 50000 MCS for different values of both κ_p and the persistence time τ. The resulting trajectories are similar to the one shown in fig. 5.1b except for the fact that the cell starts wiggle around more due to the changing polarity (see fig. D.2 for 2D persistent cell trajectories). We can then calculate the MSD for different values of τ and κ_p. To characterise the diffusive process we have plotted the instant diffusion coefficients D_{in} that follow from the calculated MSDs and fitted the results with a PRW. This is demonstrated in fig. 5.2. It can be seen that the results look well averaged and can be accurately fitted confirming that the MSD follows a PRW. By plotting the instant diffusion coefficient we can clearly notice the transition from an initial 'slow' diffusive process (constant D_{in}) via an intermediate ballistic regime (increasing D_{in}) to again a diffusive process with an increased diffusion coefficient in the long time limit. This behavior is fully consistent with the theoretical description of active Brownian motion.

Furthermore, each fit gives a persistence time τ_p (we add the subscript to distinguish

the fitted persistence time from the implemented one which is done throughout this chapter) and an active diffusion coefficient D_p (or active speed v_0) that characterise the motion of the cell. The resulting values of these parameters are plotted as a function of both τ and κ_p in fig. 5.3. The first notable result is that the persistence time τ which we implement in the CPM is (approximately) reproduced by the simulation data, i.e. by τ_p. Additionally, the value of the persistence time is independent of the polarity strength κ_p. Both results show that we can directly control the persistence time of an individual cell within a CPM simulation by changing τ.
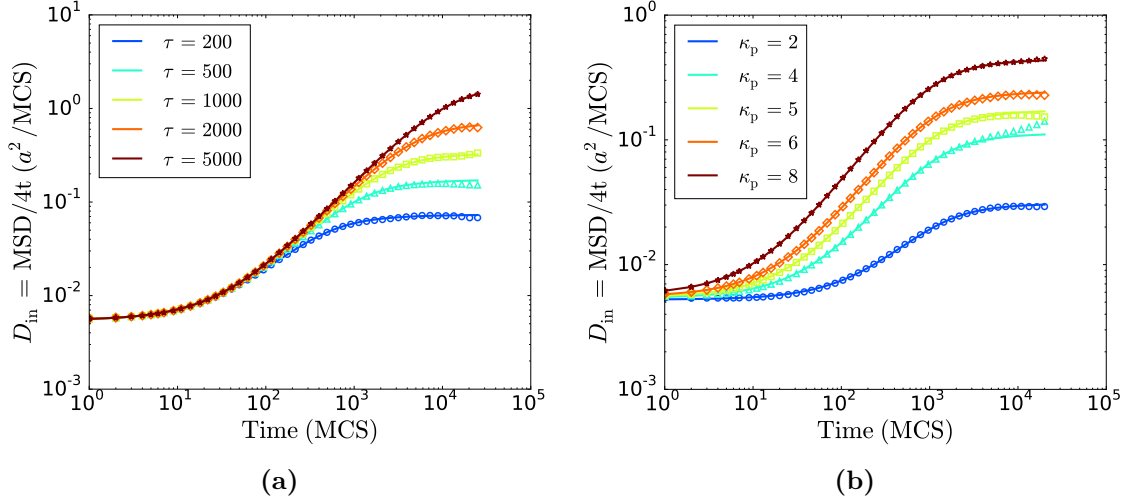

(a)   **(b)**

Figure 5.2: Plots of the calculated instant diffusion coefficients $D_{\mathrm{in}} = \mathrm{MSD}/4t$ (markers) which have been fitted using eq. (2.10) (lines). Results are obtained for (a) different implemented persistence times τ with a fixed value of $\kappa_p = 5$ and (b) different polarity strengths κ_p with a fixed value of $\tau = 500$ MCS.

Regarding the active diffusion coefficient D_p, we can see that it scales linearly with τ which is consistent with the definition of the active diffusion coefficient as introduced in section 2.1.1. Besides depending on the persistence time, the active diffusion coefficient should also scale quadratically with the speed of the relevant particle $D_p \propto v_0^2$. We have already demonstrated that the active cell speed $\langle v \rangle$ scales linearly with κ_p so we expect, linking $\langle v \rangle$ to the speed of an active Brownian particle v_0, that D_p needs to scale quadratically with κ_p as well. This notion is confirmed in fig. 5.3d. In particular by replacing v_0 with $\langle v \rangle$, the active diffusion coefficient is assumed to obey $D_p = \langle v \rangle^2 \tau / 2$. An inspection of figs. 5.1, 5.3c and 5.3d shows that the obtained values for D_p agree with this definition making all individual results consistent with each other as well.

The resulting fits also provide a value for the spatial diffusion coefficient D that describes the passive Brownian motion. The retrieved values for D remain, however, approximately the same upon changing both τ and κ_p. This makes sense since the passive motion is expected to originate from 'thermal' fluctuations in the cell volume and shape, which are not influenced by these parameters.

Finally, it has been experimentally shown and theoretically predicted that the persistent motion of cells (in the absence of signals) originates from self-propelled movement [57, 81]. An important hallmark of self-propelled movement on the underlying statistics of the trajectories is represented by the fact that the distribution of velocities, defined over a time period δ (see eq. (4.6)) of the same order as the persistence time, should show gaps in the center of the distribution [57, 81].

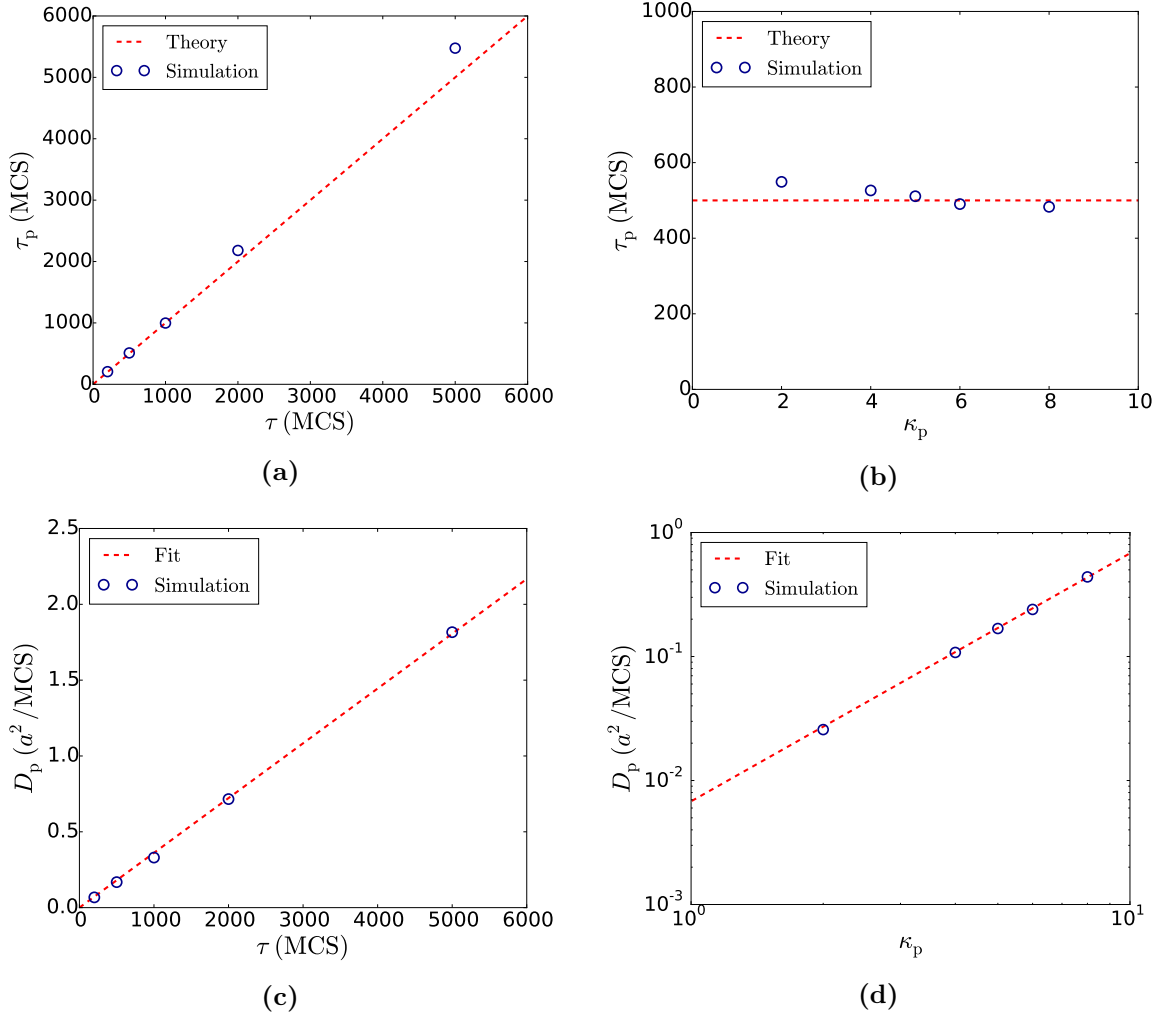To test if the implementation of active Brownian motion in the CPM also captures

39

Figure 5.3: Persistence time (a-b) and active diffusion coefficient D_p (c-d) obtained from fitting the calculated MSDs with eq. (2.10). Values are retrieved for different implemented persistence times τ and strengths κ_p, while fixing $\kappa_p = 5$ and $\tau = 500$ MCS respectively. Moreover, the results are compared to or fitted with the theoretical results of active Brownian motion.

the correct underlying statistics, we have plotted the distribution of velocity components v_x and v_y for different time periods δ in fig. 5.4. The results correspond to the 50 trajectories that have been obtained for $\kappa_p = 5$ and $\tau = 500$ MCS. These results demonstrate that for short time periods we obtain a Gaussian velocity distribution which can be explained from the fact that at short time intervals the cell jiggles around a little due to shape fluctuations but is not greatly influenced by the active energy bias. When increasing the time period δ to the same value as the persistence time, we can see the expected gap in the center of the velocity distribution appear. Specifically, at these time intervals the cells have mostly moved actively without changing direction yet, which implies that they have not been able to (possibly) reverse back to their original location. Thus, velocities relatively close to zero become less likely and we observe a gap in the distribution around the origin. For larger time periods δ beyond the persistence time, cells have been able fully reorientate their polarity direction several times. Therefore we expect that the cell displacements and therefore the velocity distribution becomes random again. This is confirmed by the Gaussian velocity distribution for $\delta = 3000$ MCS. On a final note, we highlight that the velocity distribution effectively represents the dis-

tribution of cell displacements at a time δ, which is due to the definition of the velocity in eq. (4.6); hence this distribution could have been equivalently used for the analysis of the displacements.

In total, the presented results confirm that 2D active Brownian motion can effectively be mapped onto the CPM and we can easily control the speed and persistence time of individual cells with κ_p and τ respectively. Additionally, the underlying distribution of velocities (or displacements) is consistent with experimental observations of persistent motion originating from self-propelled motion.
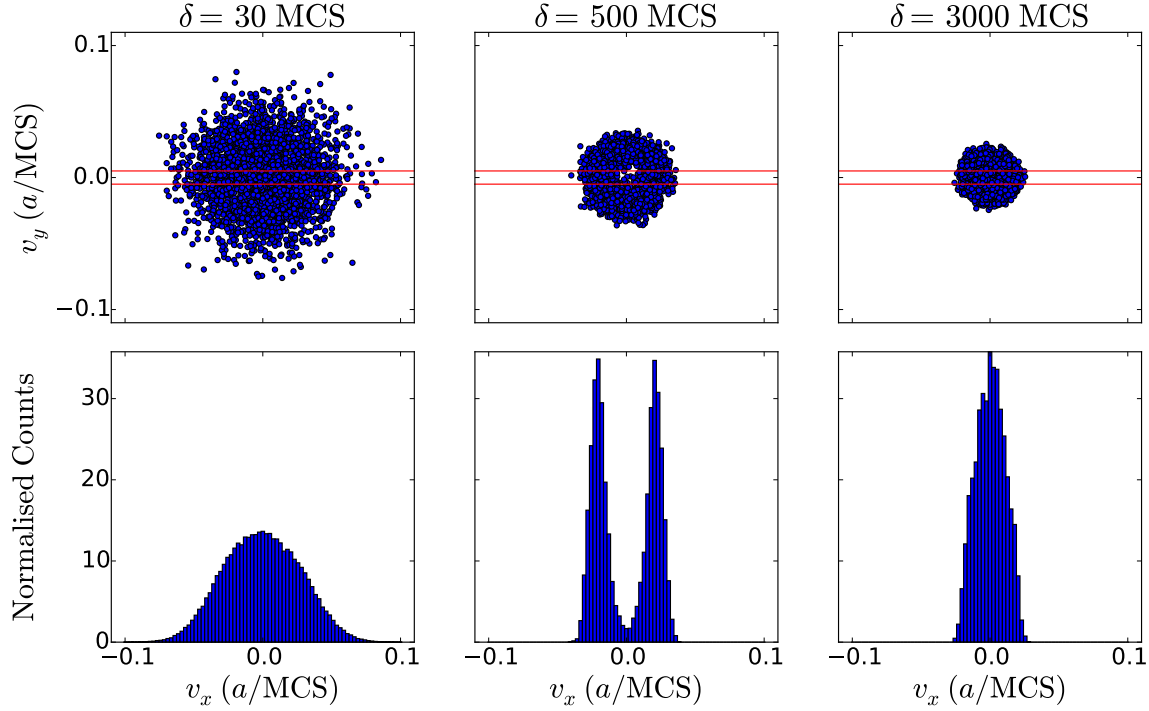

Figure 5.4: Plots of the distribution of the calculated cell velocities (v_x vs v_y) and histograms of the distribution of the v_x values that lie in between v_y values depicted by the red lines for different time periods δ. Results are obtained from trajectories for which we have used $\tau = 500$ MCS and $\kappa_p = 5$. To improve visibility we have only included a subset of the velocities in the upper panels.

5.1.2 Three Dimensions

Based on the effective implementation of 2D active Brownian motion into the CPM, we expect that extending the implementation to three dimensions will lead to similar results. To be able to confirm this, we have run 20 single cell simulations of 50000 MCS for different values of both κ_p and the implemented persistence time τ. Using the resulting trajectories we can again calculate the MSD for different values of τ and κ_p and by extension the corresponding instant diffusion coefficients D_{in}, which have been plotted and fitted with a PRW. This is shown in fig. 5.5. The plots are similar to the ones obtained in 2D (fig. 5.2) and we can clearly recognise the signature of a persistent random walk, i.e. the transition from a 'slow' diffusive process via an intermediate ballistic regime to a diffusive process with an increased diffusion coefficient at the long time limit. Moreover, the results are well averaged and can all be accurately fitted using eq. (2.20).
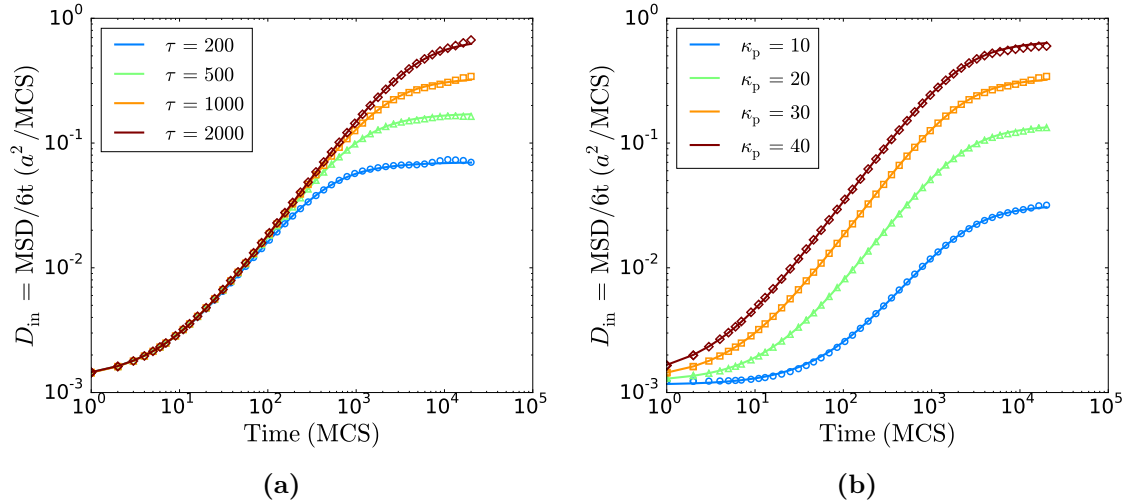
Figure 5.5: Plots of the calculated instant diffusion coefficients $D_{\text{in}} = \text{MSD}/6t$ (markers) which have been fitted using eq. (2.20) (lines). Results are obtained for (a) different implemented persistence times τ with a fixed value of $\kappa_p = 30$ and (b) different polarity strengths κ_p with a fixed value of $\tau = 1000$ MCS.

From the fits shown in fig. 5.5, we can retrieve the persistence time τ_p and the active and passive diffusion coefficients (D_p, D). Again, we will not go into much detail about the spatial diffusion coefficient D, which, as expected, remains (approximately) constant upon changing τ and κ_p. We attribute this to the fact that the values only influence the active motion of the cell. More interesting are the τ_p and D_p which have been plotted as a function of τ and κ_p in fig. 5.6. It can be seen that also in 3D the fitted persistence time τ_p is (almost) equal to the implemented one τ and that it is not influenced by variations in κ_p. Next we can notice the linear dependence $D_p \propto \tau$, which is consistent with the theoretical definition of the active diffusion coefficient and a quadratic dependence of D_p on the polarity strength κ_p. From similar simulations as in 2D we have determined that the average absolute cell velocity $\langle v \rangle$ scales approximately linearly with κ_p in 3D as well (see fig. D.1). Linking $\langle v \rangle$ to the speed of an active Brownian particle v_0 and realising that by definition $D_p \propto v_0^2$, the quadratic dependence is as expected. This implies that, in a similar way as for 2D, we can effectively implement active Brownian motion in 3D where the cell velocity and persistence time are controlled by κ_p and τ respectively.

Moreover, the velocity (or displacement) distribution is expected to behave similarly in 3D as in 2D, since it originates from the same type of motion, and is therefore not included. We mention that in experiments it has been shown that the velocity distribution is in fact different in 3D, which has been explained by using two persistence modes with distinct velocities and persistence times [58]. A proposed reason for this has been attributed to local digestion of the extracellular matrix which creates pathways where the cell can move easier and more persistently. This will, however, not be further pursued within the present work.

Summarising, we have demonstrated the succesful implementation of active (Brownian) motion into the CPM in both 2D and 3D. We will take this as the starting point for most further investigations that will seek to compare single cell motion to the motion of small cell clusters, since it allows for accurate control of individual cell properties. Finally, we want to highlight that this implementation could also provide a good model system for the study of a broader range of active (soft) materials and is not necessarily bound to a description of biological cells.
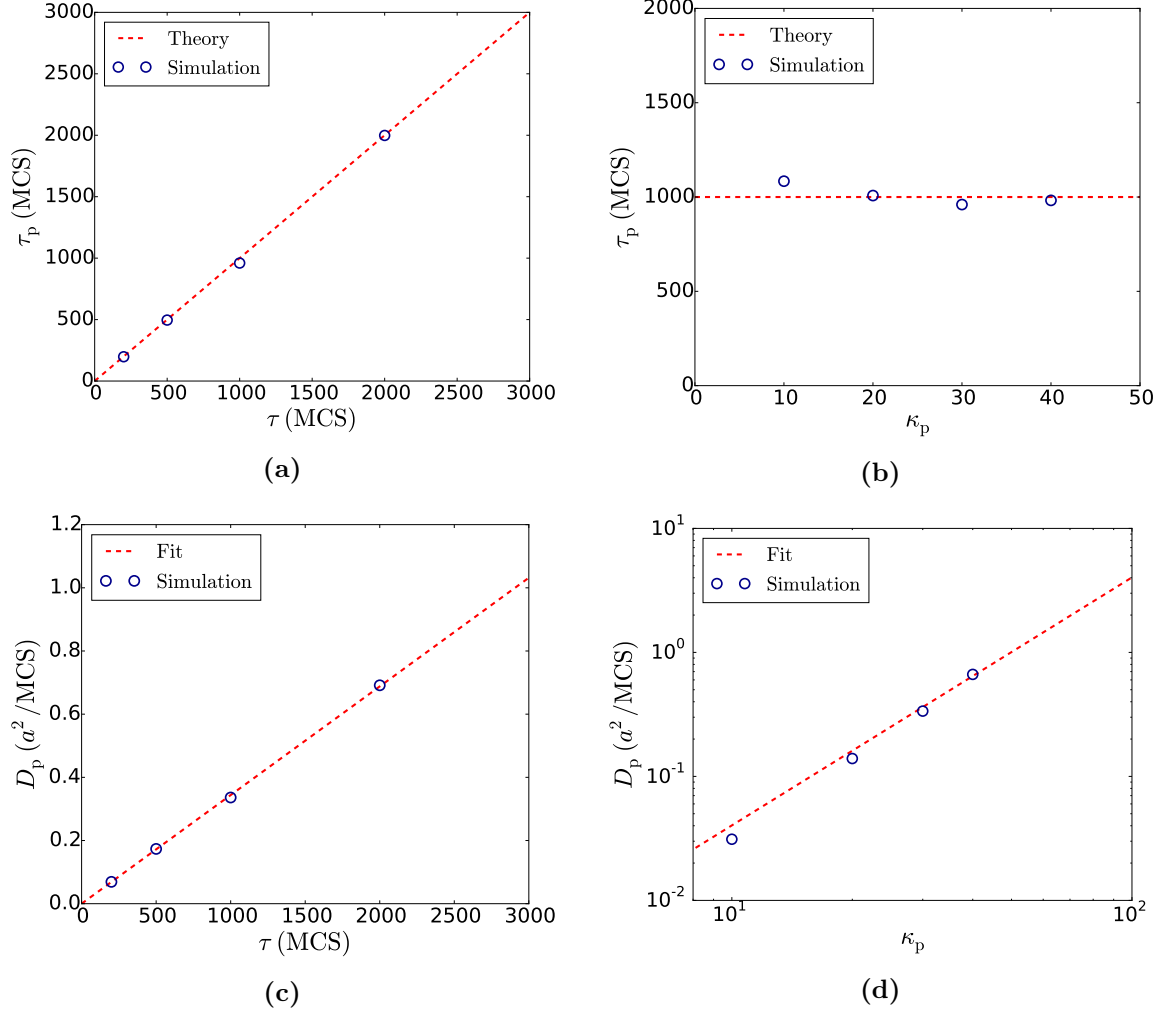
(a)

(b)

(c)

(d)

Figure 5.6: Persistence time (a-b) and active diffusion coefficient D_p (c-d) obtained from fitting the calculated MSDs with eq. (2.20). Values are retrieved for different implemented persistence times τ and strengths κ_p, while fixing $\kappa_p = 30$ and $\tau = 1000$ MCS respectively. Moreover, the results are compared to or fitted with the theoretical results of active Brownian motion.

43

5.2 Alignment

Having introduced an effective implementation for the active and persistent motion of individual cells, we now test how it translates to the motion of small cell clusters. Initially, we will refrain from modeling the ECM explicitly, i.e. we have no spins $\sigma < 0$ and neglect a binding potential $\mathcal{H}_{binding}$, and we assume that the interaction with the ECM only results in the cell's velocity and the reorientation of the polarity vector $\mathbf{p}_\sigma$ over time. This simplified model allows us to focus exclusively on cell-cell alignment, which we believe to be an important mechanism in order to have all circulating tumor cells (CTCs) within the cluster travel in the same direction. The alignment of cells with other cells in their vicinity has been pointed out in the context of wound healing where it has been shown that migrating cells in a confluent cell layer are influenced by the velocities of their neighbors [17]. As a result it is suggested that cells have a capacity for neighbor-induced migration alignment which has been demonstrated to increase the persistence of their motion [17].

However, the influence of cell-cell alignment on the migration of small CTC clusters has so far not been studied, let alone rationalised. Therefore we seek to study this influence within the CPM. In order to do so, we assume that in the context of CTC clusters neighbor-induced migration alignment is also present. Since in our case cell migration is primarily determined by the cell polarity $\mathbf{p}_\sigma$, we have modeled it using the adaptation on the Vicsek model introduced in section 3.1.3, which tends to align the polarity vectors of neighboring cells. As a consequence a parameter γ is added to the model which denotes the degree of alignment between cells, i.e. $\gamma = 1$ gives perfect alignment and $\gamma \to \infty$ results in no alignment [see eq. (3.11)].

In this section we will discuss the results of both 2D and 3D CPM simulations involving cell clusters consisting of variable number of cells that tend to align with each other by analysing the center-of-mass trajectories of the clusters $\mathbf{R}_c(t)$. For convenience and because we are mainly interested in the effects of alignment, we have taken all cells to be the same and thus of the same type, i.e. $c(\sigma > 0) = 1$. Our main control parameters are the number of cells in a cluster N_{cells} and the degree of alignment via γ; unless otherwise stated, we have kept the remaining parameters fixed. In particular we have (in the case of two values for a given parameter the first corresponds to 2D and second to 3D simulations): $T = 1$, $\lambda_v = 1$, $\lambda_c = 0$ (no chemotaxis), $V_{\sigma,0} = 64$, 216, $\lambda_r = 1$, 0.5, $J_{0,1} = 1.0$, 0.4, $J_{1,1} = 0.5$, 0.2 (strong cell-cell adhesion), $\kappa_\sigma = 5$, 30 (depicting it κ_p again) and $\tau = 500$ MCS. All retrieved averages are obtained via time-ensemble averaging and that all results are obtained for clusters have remained intact during the entire trajectory such that $\mathbf{R}_c(t)$ represents the translational time evolution of the entire cluster (and thus also of every individual cell).

5.2.1 Individual Cell Activity

As discussed in section 3.1.2 we can allow for active motion to arise at the edge of each individual cell or only at the edge of the cell cluster. We will start by considering the former, i.e. we use eq. (3.3) in the CPM. With this setting we have run sets of 30 (2D) and 20 (3D) simulations of 50000 MCS for cell clusters consisting of different amounts of cells (N_{cells}) which align with $\gamma = 1$ (perfect alignment). From the resulting trajectories we have calculated the MSD and the instant diffusion coefficient D_{in} for different N_{cells}. The latter has been plotted, including a fit with a PRW, in fig. 5.7. It can be seen that the MSDs look well averaged and can still be accurately fitted by a PRW. At first glance we can immediately notice that the effect of alignment looks to be similar for 2D and

3D cluster motion; the fact that on small time scales D_{in} decreases with N_{cells} can be attributed to the increased size of the cluster that results in weaker relative fluctuations in shape and size. We quantify this by plotting the fitted values of the passive diffusion
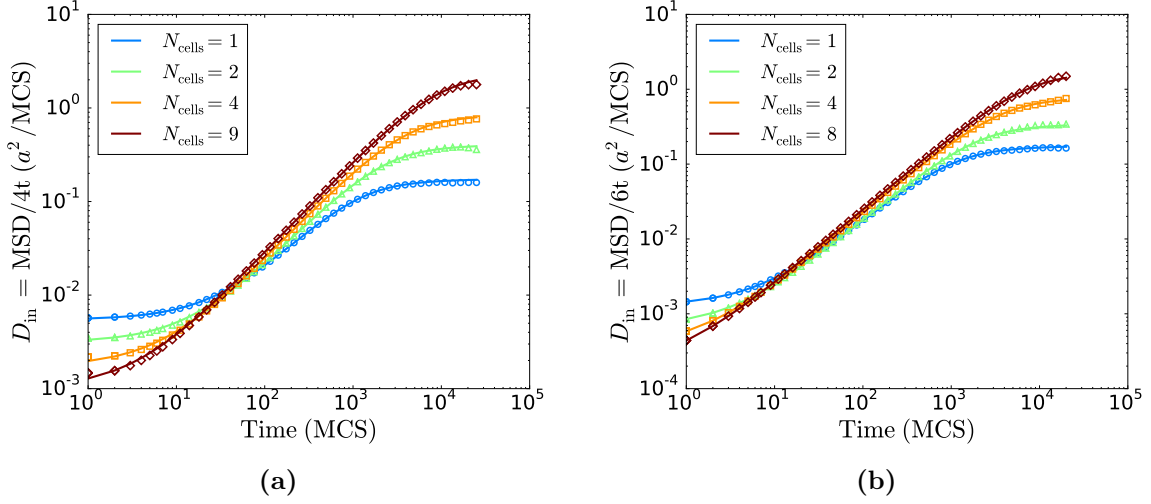

(a)   **(b)**

Figure 5.7: Plots of the calculated (a) 2D and (b) 3D instant diffusion coefficients $D_{\text{in}} = \text{MSD}/2dt$ for cell clusters consisting of a variable amount of N_{cells} cells (markers) which have been fitted with a PRW using eq. (2.10) and eq. (2.20) respectively (lines). Results correspond to $\gamma = 1$.

coefficient D in fig. 5.8. A power law fit of these results yields $D \propto N_{\text{cells}}^{-0.8}$ and $D \propto N_{\text{cells}}^{-0.9}$ for 2D and 3D motion respectively. These results are almost consistent with theory of 2D fast-aligning ABPs, which predicts a power of -1 [see eq. (2.37)]. The deviations are likely a result of the increased complexity of passive motion exhibited by the CPM cell. It remains, however, interesting that a single passive diffusive process in the theory almost captures this added complexity.
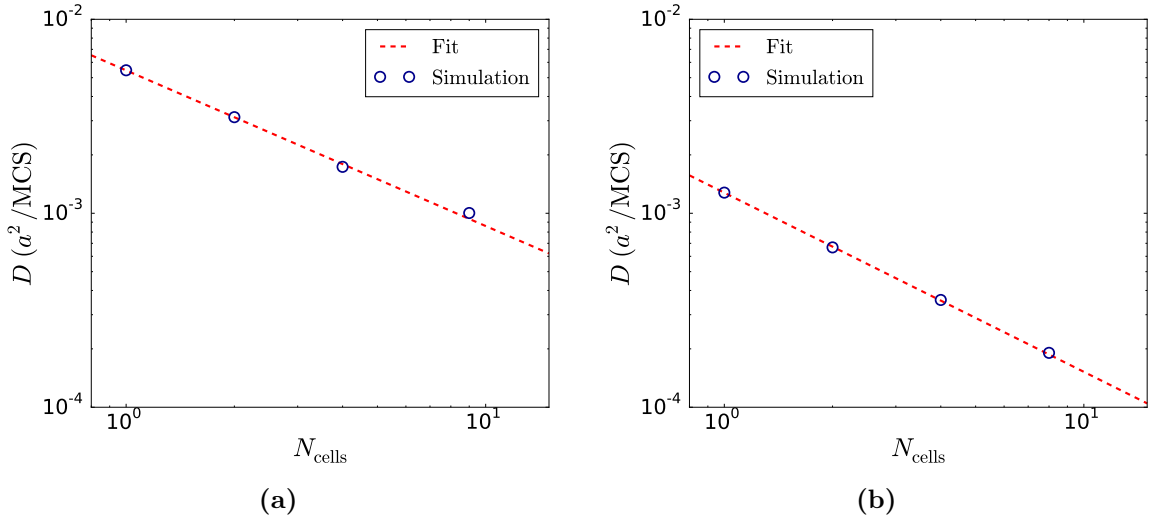

(a)   **(b)**

Figure 5.8: Passive diffusion coefficient D for (a) 2D and (b) 3D motion obtained from fitting the calculated MSDs with eq. (2.10) and eq. (2.20) respectively. Values are retrieved for clusters consisting of a variable number of N_{cells} cells and correspond to $\gamma = 1$. Additionally, results are fitted with a simple power law.

To further analyse the results, we have plotted the fitted cluster persistence time τ_p

and active diffusion coefficient D_p for both 2D and 3D cluster motion in fig. 5.9. It can be noted that both τ_p and D_p seem to increase approximately linearly with the number of cells in the cluster N_{cells} (deviations are most likely the result of insufficient averaging or small changes in the speed). Particularly, we can see that, in 2D, the obtained cluster persistence times are (almost) consistent with the theoretical results of active Brownian motion for N identical fast-aligning particles as discussed in section 2.1.2. The theory predicts a linear increase with the number of particles N (cells) and therefore we expect to find $\tau_p = N_{\text{cells}}\tau$, which is confirmed in fig. 5.9a. Interestingly, the same behavior is shown in 3D as well (see fig. 5.9c). This suggests that an extension of the theory of fast aligning active Brownian particles to 3D is likely to give similar results as the 2D model. Overall, it shows that within a more extended and detailed model setup the effect of alignment still results in a linear increasing persistence time with the amount of particles or in this case cells.
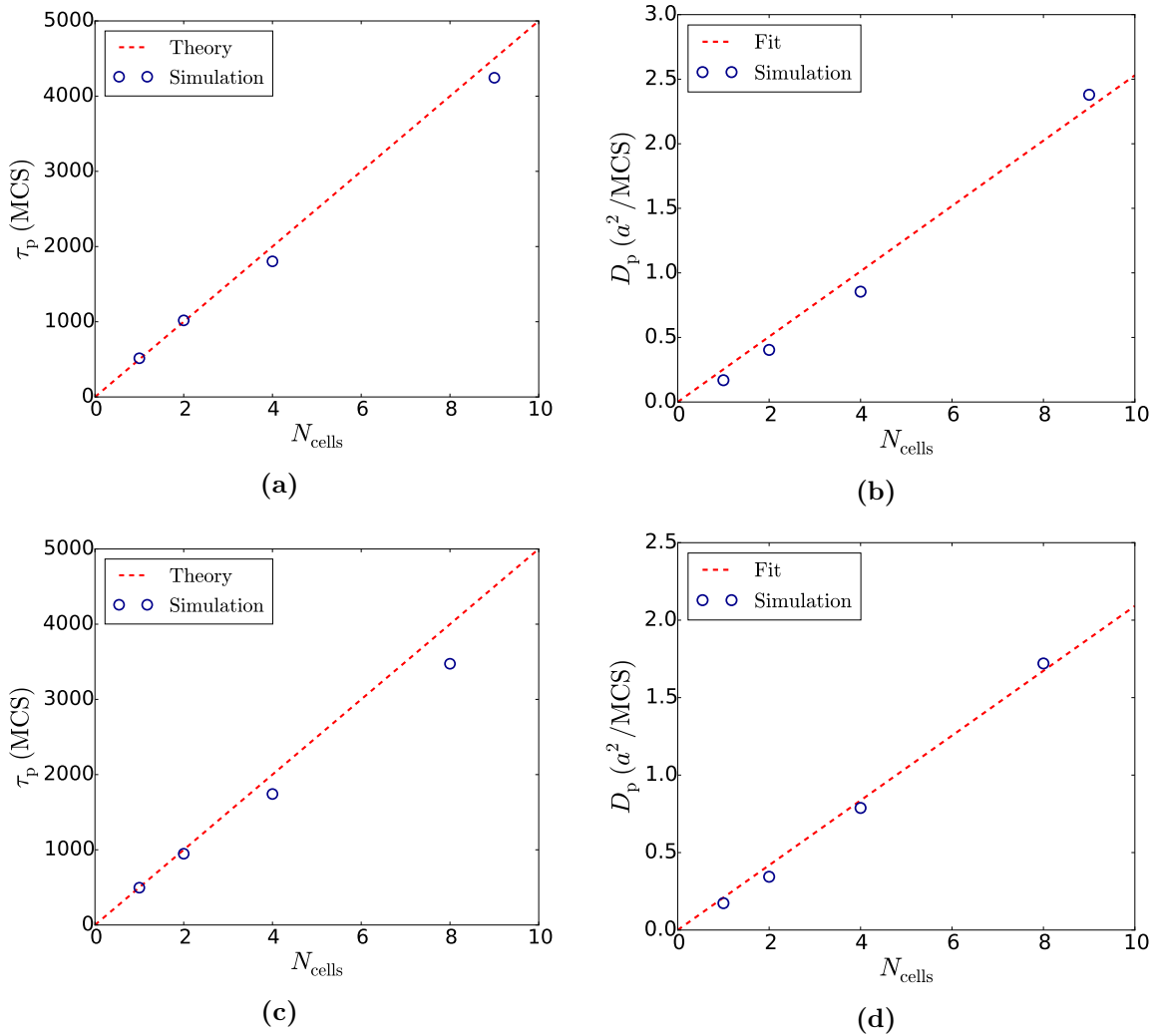

Figure 5.9: Persistence time and active diffusion coefficient D_p for (a-b) 2D and (c-d) 3D motion obtained from fitting the calculated MSDs with eq. (2.10) and eq. (2.20) respectively. Values are retrieved for clusters consisting of a variable number of N_{cells} cells and correspond to $\gamma = 1$. Moreover, the results are compared to or fitted with the theoretical results of 2D active Brownian motion for fast-aligning particles, i.e. eq. (2.37).

The linear increase in D_p with N_{cells} can also be explained by the theoretical results, which state that for fast alignment D_p should scale linearly with the cluster persistence time and thus by extension also with N_{cells}. This does assume that all cells (particles) travel with equal speed so that the cluster speed is approximately the same as that of a single cell independent of the cluster size. Cell clusters actually have a slightly larger active speed than single cells. This is demonstrated in fig. 5.10, which shows the average absolute velocity $\langle v \rangle$ as function of N_{cells} [calculated in a similar way as for a single cells (see section 5.1)]. However, the a power law fit yields $\langle v \rangle \propto N_{\text{cells}}^{0.1}$; hence there is only a weak dependence that does not influence the persistent motion of the cluster much. Moreover, we note that the results for D_p in 3D follow the behavior predicted for a 2D system again.

Up until now we have imposed very fast alignment on the cells by setting $\gamma = 1$. In order to test whether the degree of alignment significantly alters the results, we have run similar sets of simulations as mentioned earlier but instead weakened the alignment between cells by using $\gamma = 50$. The resulting trajectories allow us to calculate the MSD, D_{in} and the relevant fit parameters.
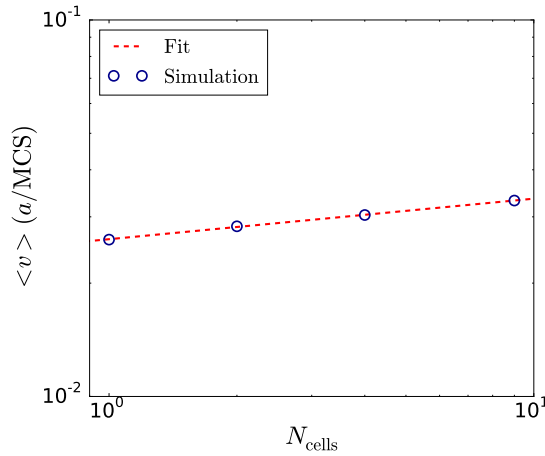

Figure 5.10: Plot of the average cell velocity $\langle v \rangle$ as a function of the number of cells in a cluster N_{cells}. Power law fit to the CPM simulation results is shown as well. Results correspond to an active energy strength $\kappa_p = 5$.

Surprisingly, the variables appear not to have changed noticeably upon this decrease in alignment and we find roughly the same results as for $\gamma = 1$ (see figs. D.3 and D.4). This demonstrates that the alignment still seems to be rapid enough to find the theoretically predicted results. However, if we keep on increasing γ the alignment will not be strong enough anymore. Inspecting several simulation visualisations has shown that when γ passes a critical value, the cluster of cells quickly falls apart into single cells and an analysis of the center-of-mass trajectory becomes meaningless. We can understand this by realising that in the case of no or weak alignment cells often want to travel in different directions (opposite polarity vectors) for long times and can then actively pull themselves loose from the other adjacent cells. Furthermore, the disintegration suggests that there exists a critical degree of alignment (γ value) beyond which alignment is not sufficiently strong to keep the cluster together. This idea is further pursued in the coming section 5.2.2 where we observe a sharp transition from an aligned towards a non-aligned cluster state at a critical value of γ.

Thus, within the proposed model a CPM cluster can greatly benefit from alignment by increasing the total persistence time. This means that a cluster has the ability to

travel more directionally towards targets like blood vessels. It can also suggest that when a cluster experiences an externally imposed polarity (through tracks in the ECM for instance), it might be harder to let the cluster deviate from such a track.

Finally, we want to test what the effect of fast alignment is for a heteregeneous cluster. Specifically, we want to know how one less persistent cell influences the motion of the total cluster. This can be interesting since individual cells show a variety of persistence times in experiments [58] and thus CTC clusters will likely consist of a hetergeneous mixture of cells as well. As before, we have run sets of 30 CPM simulations of 50000 MCS with $\gamma = 1$ (in 2D to be able to compare to theoretical results). However, instead of changing the cluster size, we have kept it fixed by setting $N_{\mathrm{cells}} = 4$ and varied the implemented 'small' persistence time (denoted τ_s) of one of the cells, while the 'large' persistence time of the remaining 3 cells (denoted τ_l) has been set at 1300 MCS.

From the simulation results we have calculated the MSDs for different τ_s which in turn have been fitted with a PRW. The fitted cluster persistence times have been plotted as a function of τ_s in fig. 5.11. We can see that the simulation data agrees well with the theoretically predicted persistence time of a fast-aligned cluster of $N = 4$ particles with one particle having a deviating persistence time, i.e. eq. (2.44). It shows that the benefits of alignment can become much smaller or even non-existent by for instance adding a cell with a small persistence to an existing cell cluster. We can understand this by realising that a small persistence time corresponds to a rapid reorientation of the cell's polarity. If the reorientation becomes to fast, the cell will drag along other cells towards this polarity as well, which results in a decrease of the cluster persistence.
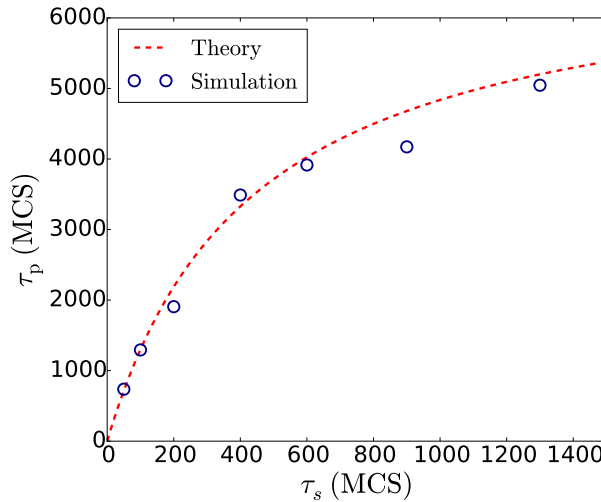

Figure 5.11: Plot of the fitted cluster persistence time τ_p, obtained from fitting calculated MSDs with eq. (2.10), as a function of the implemented small persistence time τ_s. Red dotted line denotes the theoretical prediction of fast-aligning active Brownian particles, i.e. eq. (2.44).

5.2.2 Collective Cell Activity

We will now seek to model the effects of cell-cell alignment in the other regime in which active forces only occur at the edge of the cell cluster, i.e. we implement eq. (3.4). As will become apparent, this formulation allows us to study cell-cell alignment over the entire range of γ values, since cells seem to stick together more easily when we evaluate an active energy bias only at the cluster edge. Based on the observed similarity between

2D and 3D aligned motion in the previous section, which is expected not to change upon changing the active energy bias, most presented results in this section correspond to 2D cell motion for which it is less computationally expensive to acquire data. For completeness, we do present some 3D results at the end of the section.

Two Dimensions

In order to study the influence of letting active forces only occur at the cluster edge, we have run sets of 30 simulations of 50000 MCS for cell clusters consisting of different amounts of cells (N_{cells}) which align with $\gamma = 1$ (perfect alignment) [see eq. (3.11)]. From the resulting trajectories we have calculated the MSD for different N_{cells}. The corresponding instant diffusion coefficient D_{in} has been plotted in fig. 5.12. The resulting curves can still be accurately fitted with a PRW and are well averaged, but are noticeably different from the ones obtained by implementing active motion at each cell edge (see fig. 5.7a). The only feature that remains approximately the same between both plots, is the value of D_{in} on very short time scales. This makes sense because on that scale the motion is not yet influenced by the active energy bias and the motion is purely a result of the shape and size fluctuations.
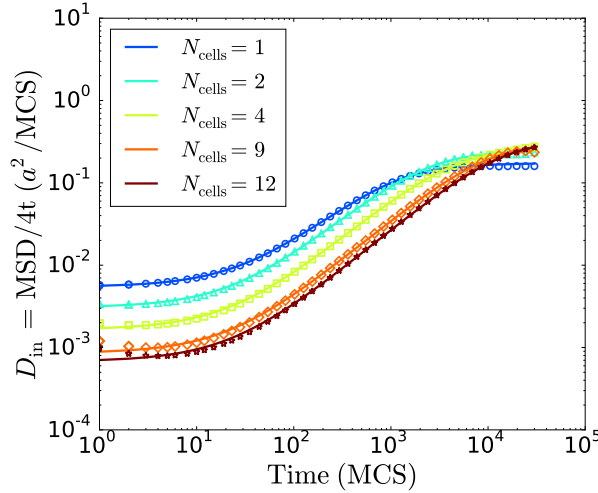

Figure 5.12: Plots of the calculated instant diffusion coefficients $D_{\text{in}} = \text{MSD}/2dt$ for cell clusters consisting of a variable amount of N_{cells} cells (markers) which have been fitted with a PRW using eq. (2.10) (lines). Results correspond to $\gamma = 1$.

To test how the results are affected by the redefinition of the active energy bias, we have plotted the fitted values for τ_p and D_{p} that follow from the calculated MSDs. This is shown in fig. 5.13. It can be seen that the persistence time of the cluster still scales linearly with the number of cells in the cluster and thus gives the same behavior as 2D active Brownian motion for fast-aligning particles. On the other hand we can notice that D_p shows a much weaker increase with the number of cells in the cluster. To illustrate this we have fitted the active diffusion coefficient which yields $D_p \propto N_{\text{cells}}^{0.2}$. In the theoretical description of active Brownian motion we have seen that $D_p \propto v_0^2 \tau_{cm}$ with v_0 the particle speed and τ_{cm} the persistence time of the cluster [see eq. (2.37)]. Since the fitted cluster persistence time increases linearly with N_{cells}, we can conclude that the active speed of the cells (and thus of the entire cluster) decreases with N_{cells}. We can understand this by imagining a cluster in which a cell in the center is completely surrounded by other cells. Because we do not allow for active motion along cell-cell

49

contacts, this cell is not able to perform active motion by itself and its velocity is purely a result of being dragged along by the other cells. This will slow down the cluster and leads to a smaller speed. In other words the more a cell is in contact with other cells, the less it can actively move by itself and the more it needs to be dragged along by the other cells. Since the amount of cell-cell contact relative to cell-medium contact increases with cluster size, this effect becomes more dominant for larger cell clusters and therefore we expect the decrease in the speed with N_{cells}. In comparison, we mention that even though enclosed cells do not perform active motion, they still align their polarity vector with their neighboring cells, which explains why the behavior of the cluster persistence time remains the same, i.e. it grows linearly with N_{cells}.
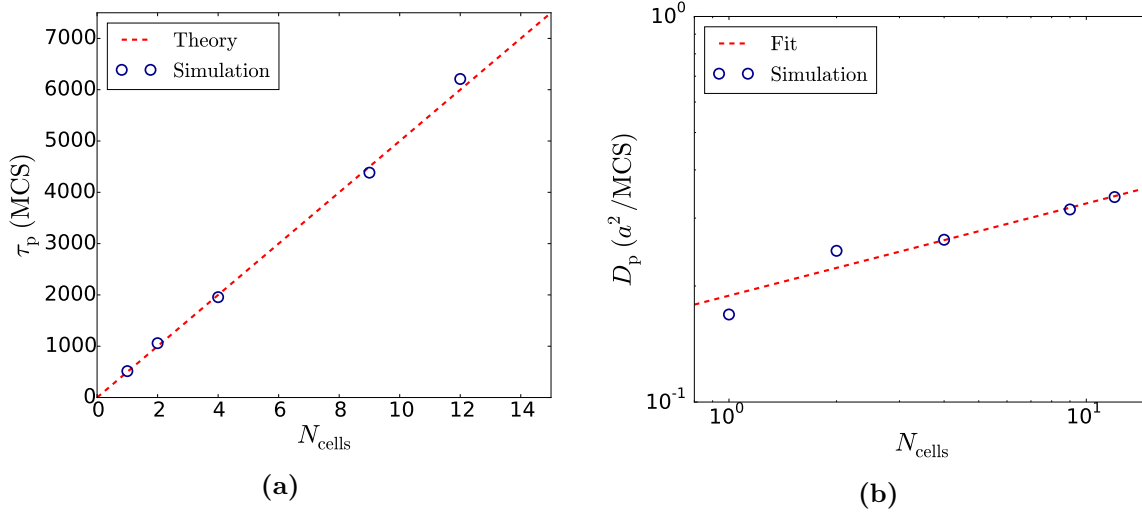

Figure 5.13: Persistence time and active diffusion coefficient D_p obtained from fitting the calculated MSDs with eq. (2.10). Values are retrieved for clusters consisting of a variable number of N_{cells} cells and correspond to $\gamma = 1$. Moreover, the results are (a) compared to theoretical results of 2D active Brownian motion for fast-aligning particles, i.e. eq. (2.37), and (b) fitted with a simple power-law function.

Interestingly, the decrease in velocity is not so strong that the active diffusion coefficient of the cluster decreases with N_{cells} as well. It seems that the linear increase of the cluster persistence time dominates over the decrease in cell speed, which is somewhat surprising because according to theory the velocity scales quadratically with the active diffusion coefficient and the persistence time only linearly. Thus, even in the limit where cells can only pull themselves forward where they are not in contact with other cells, fast alignment appears to improve the total diffusive ability of the cluster (larger D_p), but it does come with a tradeoff between increased persistence and slower cluster speed.

Besides leading to a smaller cell cluster speed, the fact that cells are not able to actively move where they are in contact with other cells also makes the cells less prone to detach from one another. It becomes harder to overcome the adhesion energy between cells if there is no active energy contribution. In fact, within the investigated parameter space and simulation time, the cluster still sticks together even without implementing any alignment. This allows us to study cluster motion over the entire range of γ values.

To quantify the effect of γ on the MSD, we have run sets of 30 simulations of 50000 MCS for a cell cluster consisting of $N_{\text{cells}} = 4$ cells with different γ values for each set. Moreover, we have obtained these results for two single cell persistence times, i.e. $\tau = 500$ MCS and $\tau = 1100$ MCS to test whether alignment is influenced by the persistence of an individual cell. Figure 5.14 shows the resulting instant diffusion coefficients D_{in} for a

subset of γ values. It can be seen that in all cases and thus also for very weak alignment the result can be accurately fitted with a PRW and all results are well averaged. We can
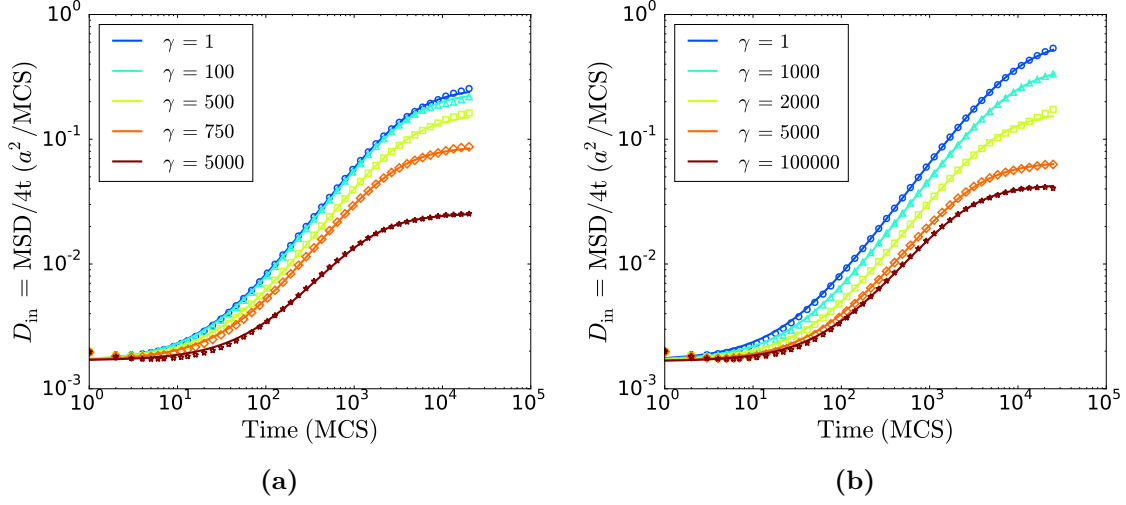

(a)   **(b)**

Figure 5.14: Plots of the calculated instant diffusion coefficients $D_{in} = \text{MSD}/4t$ for different values of γ (markers) which have been fitted with a PRW using eq. (2.10) (lines). Results correspond to (a) $\tau = 500$ MCS and (b) $\tau = 1100$ MCS.
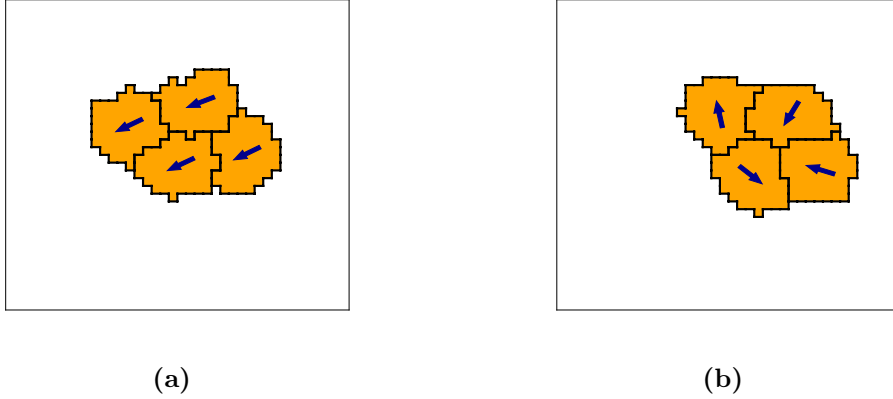

(a)   **(b)**

Figure 5.15: Snapshots of a CPM cluster consisting of $N_{cells} = 4$ cells halfway during a simulation. Arrows denote the polarity vector $\mathbf{p}_\sigma$ of each cell. Results correspond to $\tau = 500$ MCS, (a) $\gamma = 1$ and (b) $\gamma = 10000$.

also observe a transition in long time diffusive behavior upon weakening the alignment (increasing γ), which we can explore further by plotting the fitted values of the cluster persistence time and instant diffusion coefficient [fig. 5.16]. These plots demonstrate that the cluster persistence time drops fairly abruptly when passing a specific value of γ. Hence, there appears to exist a transition from a cooperative cluster state for which $\tau_p = N_{cells}\tau$ (cluster moves more persistently) and a non-cooperative cluster state where $\tau_p = \tau$ (no distinction between cluster persistence and the single cell one). These states are visualised in fig. 5.15. Furthermore, the transition point seems to depend on the individual persistence time of the cells, since the drop occurs for larger values of γ when we set the persistence time of an individual cell to a larger value. In other words, when the cell's polarity reorientates more slowly (larger τ) it requires less alignment effort

(larger critical γ value) for cells in the cluster to remain aligned. The assumption of fast alignment, which predicts a linear increase of cluster persistence time with the number of particles, thus seems to hold as long as the alignment effect is stronger than the individual cell reorientations so that all cells will always have their respective polarities in roughly the same direction.
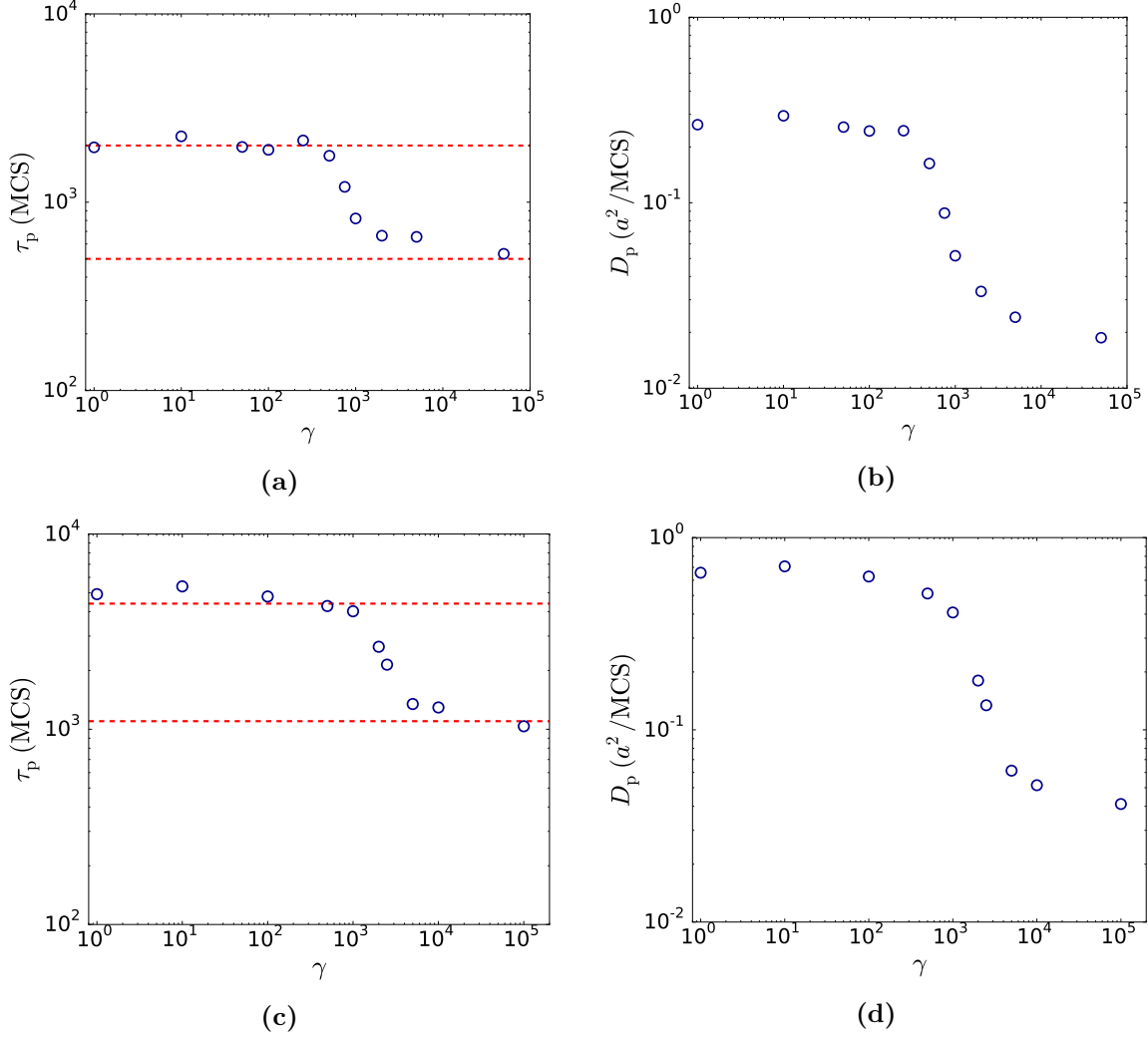

Figure 5.16: Cluster persistence time and active diffusion coefficient D_p obtained from fitting the calculated MSDs with eq. (2.10). Values correspond to (a-b) $\tau = 500$ MCS and (c-d) $\tau = 1100$ MCS. Red dotted lines in (a) and (c) serve as visual aide and denote the values $N_{\text{cells}}\tau$ (top line) and τ (bottom line).

The active diffusion coefficient D_p follows the same trend as the persistence time, i.e. a sudden drop from a cooperative cluster state to a non-cooperative state upon increasing γ past some critical value. Interestingly, the difference in D_p between both states (factor of ~ 10) is larger than the factor 4 (N_{cells}) observed for the persistence time. Realising that in the non-cooperative state individual cell polarities are not pointing in the same direction most of the time, we expect the speed of the cluster to be smaller in this state. This can explain why the diffusion coefficient decreases more dramatically, since it scales with both the persistence time and the speed. Thus, the degree of alignment can greatly influence the motion of a small cell cluster. Additionally, we mention that the observed transition occurs in an almost identical manner (including the critical γ value) for a

cluster consisting of $N_{cells} = 9$ cells (see fig. D.5) and is therefore not necessarily unique for different sized clusters.

Finally, we seek to explore the relationship between the transition point (critical γ value) and the individual persistence time of the cells in more detail. Instead of analysing cluster motion for different values of γ, we will now fix the degree of alignment (constant γ) and vary the individual persistence time τ of the cells that comprise the cluster. This means that we generate similar sets of simulations as discussed before, but for different values of τ, while keeping $\gamma = 500$ or $\gamma = 1500$. The resulting MSD and D_{in} are similar to the ones shown in fig. 5.14 (see fig. D.6) and have been fitted with a PRW. Since we are particularly interested in observing when a cluster is in a cooperative state or not, we have plotted the cluster persistence time divided over the single cell persistence time τ as a function of τ in fig. 5.17. This allows us to easily discriminate between both states, i.e. a value of 4 (N_{cells}) corresponds to a cooperative state and 1 corresponds to a non-cooperative state. Figure 5.17 reveals that the transition is also induced by increasing τ, which shows that the effect of alignment and single cell persistence appear to be linked. Moreover, for weaker alignment (larger γ) we can notice that the transition occurs for larger values of τ, which is consistent with the idea that weaker reorientations of the polarity vector require weaker alignment to keep the cluster in a cooperative state.
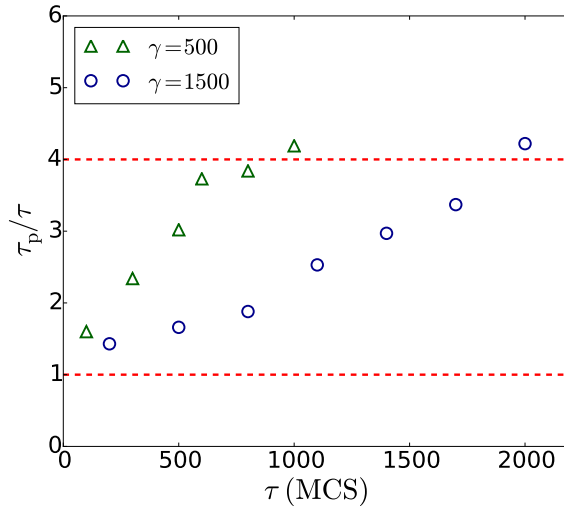

Figure 5.17: Plots of the cluster persistence time divided over the single cell persistence time τ as a function of τ for $\gamma = 500$ and 1500. Values are obtained from MSD fits with eq. (2.10). Red dotted lines serve as visual aide and denote the expected values for the cooperative (top) and non-cooperative state (bottom)

Interestingly, an increase in single cell persistence time has been experimentally associated with an increase in substrate stiffness [25, 26]. The obtained results can therefore suggest a link between the effectiveness of cell cluster migration and the stiffness of the environment that the cluster experiences.

Three Dimensions

As mentioned, we expect, based on earlier findings, that the effect of alignment will lead to similar behavior in 3D as in 2D. This has been confirmed by running sets of 3D simulations for different γ keeping a fixed single cell persistence τ and vice versa. These simulations have provided comparable results to the ones obtained in 2D.

However, we do want to highlight the effect of redefining the active energy bias,

i.e. no energy bias along cell-cell contacts, in 3D. We have argued how the velocity of a cell cluster in 2D decreases for larger cluster sizes due to an increase in the relative fraction of cell-cell contacts. Since this fraction does not necessarily have to be the same for a 3D cell cluster consisting of an equal amount of cells, we have performed sets of 3D simulations for cell clusters consisting of different amounts of cells (N_{cells}) which align with $\gamma = 1$ (perfect alignment). Using the simulation data, we calculate the MSD and the corresponding instant diffusion coefficient D_{in}. The latter has been plotted in fig. 5.18a and already reveals an increase in long term diffusive behavior for increasing N_{cells}. To inspect the precise details, we have fitted the MSD with a PRW. As expected the cluster persistence time increases linearly with N_{cells} (see fig. D.7). Moreover, we have plotted the fitted active diffusion coefficient D_p as function of N_{cells} in fig. 5.18b. It can be seen that D_p increases with N_{cells}, which implies that the effect of a linearly increasing persistence time dominates over the decrease in cluster speed in 3D as well. Moreover, we have fitted the active diffusion coefficient to a simple power law function, giving $D_p \propto N_{cells}^{0.5}$. Remembering that in 2D the exponent equals 0.2, we can conclude that the decrease in speed with N_{cells} is less significant in 3D as it is for 2D cluster motion. We can understand this by realising that in 3D the relative fraction of cell-cell contacts grows more slowly with the number of cells N_{cells}.
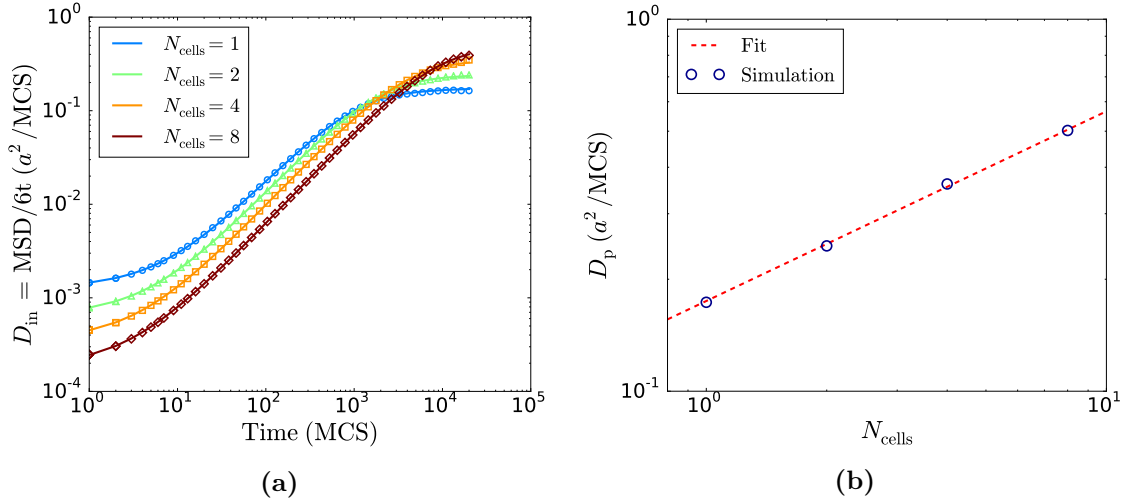

Figure 5.18: (a) plot of the calculated instant diffusion coefficients $D_{in} = \text{MSD}/4t$ for different values of N_{cells} (markers) which have been fitted with a PRW using eq. (2.10) (lines). (b) plot of the active diffusion coefficient D_p obtained from the fits shown in (a), which has been fitted with a simple power law. Results correspond to $\gamma = 1$.

5.3 Durotaxis

Up until this point we have taken the persistence time and active speed of single cells to remain constant over the entire domain, which implies that the stiffness of the ECM has been kept constant in the model. Now we extend the analysis to include durotaxis (migration up a stiffness gradient). Experiments have indicated that a multicellular cluster (cell monolayer) can exhibit much stronger durotaxis in comparison to its isolated constituents in the same circumstances, since a cluster experiences a much larger gradient in stiffness between its leading and trailing edge [18, 19]. This raises the question whether a circulating tumor cell (CTC) cluster also experiences stronger durotaxis than a single cell, possibly improving their metastatic potential. Experiments almost unanimously point out that an increased stiffness corresponds to an increased single cell persistence [25, 26] and simulations involving point particles have shown how a gradient in persistence time in itself is enough to generate durotactic motion [24]. Moreover, stiffer environments have also been linked to faster cell motion. When exposed to a stiffness gradient, cells on the stiffer edge of the cluster will thus move faster and more persistently in comparison to the cells on the softer edge and it is has been suggested that because this difference increases for larger clusters, durotaxis is enhanced for larger cell clusters [18, 19].

Motivated by these experiments, we now seek to model the motion of small cell clusters that experience a stiffness gradient in the CPM. Following the approach in [24], we model this by implementing a position dependent single cell persistence time $\tau(x)$ and cell speed, i.e. active energy bias $\kappa_\sigma(x)$ (denoted $\kappa_p(x)$ again), which, for convenience, are the same for all simulated cells and only depend on the position along the x-axis. It should be stressed that in the case of $\tau(x)$ x represents the x-component of the cell's center of mass $\mathbf{R}_\sigma$, while for $\kappa_p(x)$ it is the x-component of the position of the candidate site $\sigma(\mathbf{x})$. We then let $\tau(x)$ and $\kappa_p(x)$ increase linearly from a minimum value $\tau_{\min}$ and $\kappa_{p,\min}$ to a maximum value $\tau_{\max}$ and $\kappa_{p,\max}$, respectively, over a region $x \in [-w, w]$ around the origin (note that the origin is set by the cell cluster center of mass $\mathbf{R}_c$ so that the cluster always starts in the middle of the gradient). Beyond the gradient region the respective values remain constant such that $\tau(x) = \tau_{\min}$ and $\kappa_p(x) = \kappa_{p,\min}$ for $x \leq w$ and $\tau(x) = \tau_{\max}$ and $\kappa_p(x) = \kappa_{p,\max}$ for $x \geq w$. This means that the width w effectively controls the steepness of the gradient and that conform experimental setups the gradient only occupies part of the system [24].

In the remainder of this section we will discuss the results of 2D CPM simulations of single cells and clusters that experience the described gradient in τ and κ_p by analysing their center-of-mass trajectories $\mathbf{R}_c(t)$. Again all cells are of the same type $c(\sigma > 0) = 1$ and we keep (unless otherwise stated) most parameters fixed at: $T = 1$, $\lambda_v = 1$, $\lambda_c = 0$ (no chemotaxis), $V_{\sigma,0} = 64$, $\lambda_r = 1$, $J_{0,1} = 1.0$, $J_{1,1} = 0.5$ (strong cell-cell adhesion) and $\gamma = 1$ (cluster alignment is included). Finally, we note that averages have been obtained via ensemble averaging over approximately 10^4 trajectories to compute accurate averages, and in all cases cell clusters have remained intact.

5.3.1 Persistence Time Gradient

At first we have focused on a gradient in the persistence time so that $\kappa_p = 5$ remains constant over the simulation domain (constant cell speed). Based on a mapping of simulation parameters and units to typical experimental values we choose to vary the persistence time between $\tau_{\min} = 200$ MCS (~ 0.2h) and $\tau_{\max} = 2000$ MCS (~ 2.0h), which also implies that the gradient is in the positive x-direction. We have then run

single cell CPM simulations for different gradients $d\tau/dx$ to test whether the implementation of durotactic motion is consistent with earlier simulation work on point particles in [24]. Figure 5.19 shows the measured x-components of both the average displacement and the durotactic vector index, i.e. $\langle x(t) \rangle$ and $\mathrm{DI}_x(t)$ (see eqs. (4.1) and (4.5)), as a function of time. We mention that the single cell speed $\langle v \rangle$ (see fig. 5.1) has been used as reference speed for $\mathrm{DI}_x(t)$ and that the y-components have not been reported, since there is no gradient along this axis leading to their respective values being 0.

The results are consistent with the fact that a gradient in persistence time is enough to produce a single cell flux towards the stiff region of the domain (positive $\langle x \rangle$), thus causing durotaxis. Realising that DI_x represents the fraction of drift velocity relative to the active cell speed, we also observe an increase (on the investigated time scale) of the drift velocity for increasing gradients (larger values of DI_x). Additionally, note that DI_x peaks and afterwards seems to decrease in the long time limit, which is a result of cells leaving the gradient region. A mapping of the retrieved results to the ones obtained for point particles presented in [24] shows that they are also quantitatively the same, making our work fully consistent with literature.
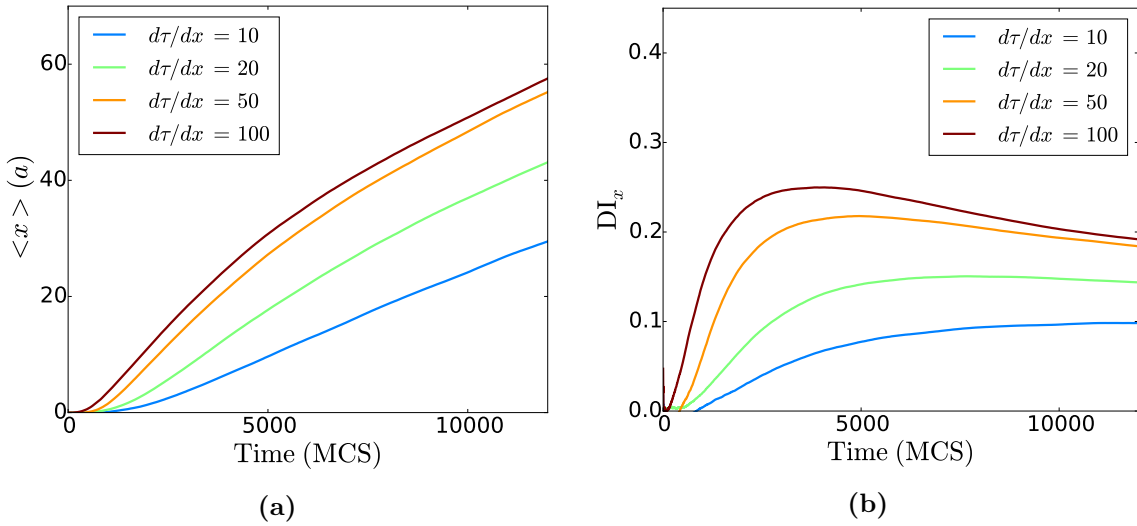

Figure 5.19: Plots of (a) the average displacement in the x-direction and (b) the x-component of the durotactic index as a function of time for a single cell experiencing a linear gradient in persistence time from $\tau_{\min} = 200$ MCS to $\tau_{\max} = 2000$ MCS. Results correspond to different gradients controlled by the width w of the gradient region.

We now seek to extend the durotaxis analysis to a setup involving a cell cluster. In the previous section we have shown that the persistence time of an aligned cluster scales linearly with the number of cells in the cluster N_{cells} and that when we allow an active energy bias at the cell edges [eq. (3.3)] the speed of the cell cluster remains approximately the same as for a single cell, i.e. the cluster moves in the same way as a single cell with an increased persistence time. We can place such a cluster in the same gradient region and not surprising this leads to stronger durotaxis (although DI_x takes longer to peak) for larger clusters (see fig. D.8), which can be explained by the fact that the gradient and absolute values of the cluster persistence time are effectively a factor N_{cells} larger.

In order to unambiguously test whether an improved durotaxis effect arises from the fact that cells on the softer side of the cluster move less persistently than on the stiff side, we have run simulations of a cell cluster (with active motion at all cell edges) placed

in a gradient between $\tau_{\min} = 200/N_{\text{cells}}$ MCS and $\tau_{\max} = 2000/N_{\text{cells}}$ MCS (keeping a constant width w). As a result the cluster as a whole experiences the same persistence time gradient in all cases, but the relative difference in single cell persistence on either side of the cluster increases. The resulting $\langle x(t) \rangle$ shown in fig. 5.20a indicate enhanced durotaxis for larger clusters. However, we have determined that cell clusters actually have a slightly larger active speed than single cells [see fig. 5.10]. Using these slightly increased speeds, we have then calculated and plotted DI_x in fig. 5.20b. It can be seen that in this case the results are not significantly different anymore, suggesting that, at least for small cell clusters, the actual size of a cluster in which all cells are fully aligned does not influence or enhance durotactic motion when it originates from only a persistence time gradient. Only the persistence time and the speed of the entire cluster appear to be a dominant factor.
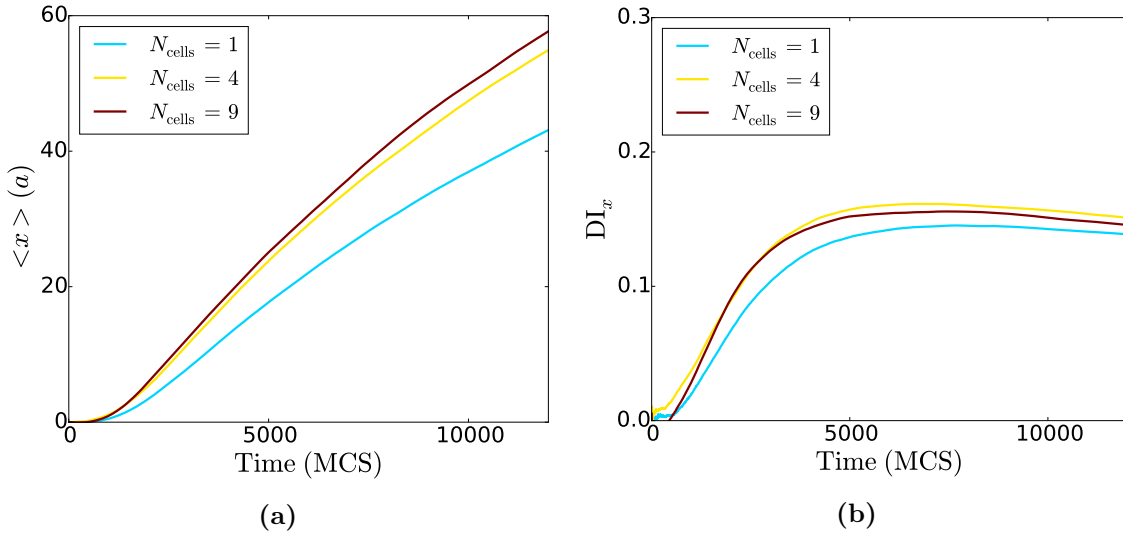

Figure 5.20: Plots of (a) the average displacement in the x-direction and (b) the x-component of the durotactic index as a function of time for aligned ($\gamma = 1$) cell clusters experiencing a linear gradient of $d\tau/dx = 20/N_{\text{cells}}$ MCS/a in single cell persistence time from $\tau_{\min} = 200/N_{\text{cells}}$ MCS to $\tau_{\max} = 2000/N_{\text{cells}}$ MCS. Different lines correspond to different number of cells within a cluster.

5.3.2 Cell Speed Gradient

Next we have implemented a gradient in cell speed, while keeping the persistence time of each cell fixed at $\tau = 1100$ MCS. In order to check whether this can also result in durotactic motion, we have run single cell CPM simulations for different gradients $d\kappa_p/dx$ that change from $\kappa_{p,\min} = 2.0$ towards $\kappa_{p,\max} = 8.0$. Using the single cell speed corresponding to $\kappa_p = 5.0$, i.e. in the middle of the gradient, as reference speed (see fig. 5.1), we have plotted both $\langle x(t) \rangle$ and DI_x in fig. 5.21. It can be noted that on average the cells also drift towards the stiff region (assuming that higher stiffness relates to a larger cell speed). Additionally, we observe an increase in drift velocity on the investigated timescale for increasing gradients in the cell speed. Exposing a cell to a persistence time gradient or cell speed gradient thus appears to give similar results, which suggests that the persistence length, i.e. the product of cell speed and persistence time, is a suitable parameter to determine the degree of single cell durotaxis.

As mentioned, a larger cluster size has been brought forward to explain collective durotaxis since it causes cells on the soft and stiff edge of the cluster to experience a big-
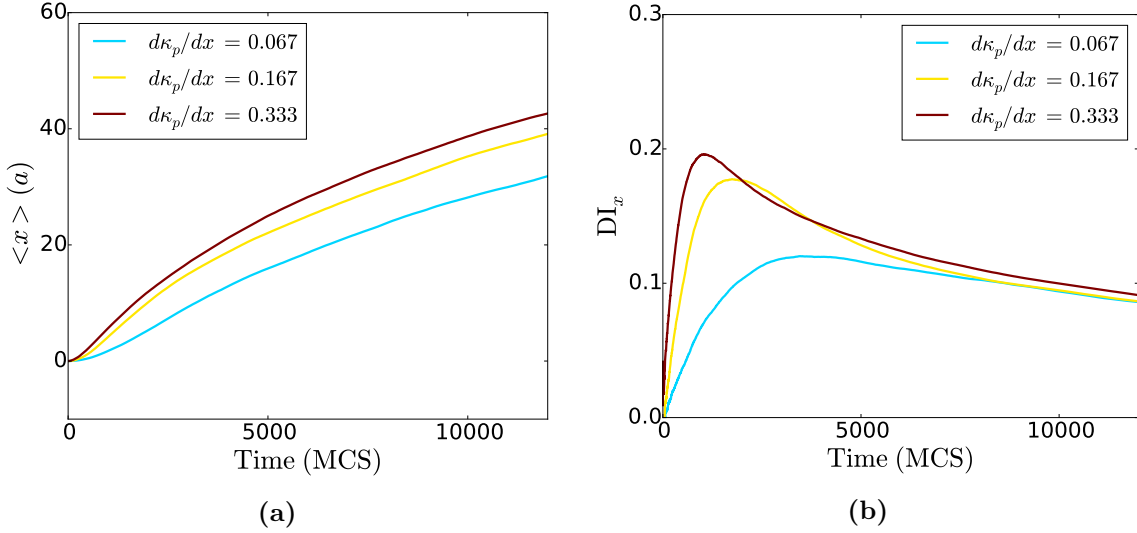
Figure 5.21: Plots of (a) the average displacement in the x-direction and (b) the x-component of the durotactic index as a function of time for a single cell experiencing a linear gradient in κ_p from $\kappa_{p,\min} = 2.0$ towards $\kappa_{p,\max} = 8.0$. Different lines correspond to different gradients controlled by the width w of the gradient region.

ger difference in cell speed. Because we model cell speed locally (using the x-coordinate of the candidate site $\sigma(\mathbf{x})$ for $\kappa_p(x)$) such a difference in cell speed between both edges is already used within our single cell model description. To test if increasing the distance between the edges improves durotaxis we will therefore simply increase the size of a single cell. In particular, we have run simulations where cells with different preferred cell areas $V_{0,\sigma}$ are placed in a fixed gradient between $\kappa_{p,\min} = 2.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ towards $\kappa_{p,\max} = 8.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ (keeping the width w constant) with $V_{\sigma,\mathrm{ref}} = 64$ denoting the reference preferred cell area which has been used beforehand. Note that the values of κ_p have thus been rescaled with the cell area so that all cells experience (approximately) the same cell speed gradient. This is a result of the active energy bias scaling with the center-of-mass displacement [see eq. (3.3)].

Interestingly, the average displacement in the x-direction $\langle x(t) \rangle$ resulting from the simulations (see fig. 5.22) shows that increasing the size has no influence on the durotaxis. We try to explain this by pointing out that in the model the entire cell at all times wants to move in a single direction (one polarity vector). Moreover, when we increase the size of a cell, it will experience a stronger energy bias on one side but simultaneously a weaker one on the other side. Combining these notions, we suggest that since extensions and retractions of the cell on both sides all seek to drive the cell in the direction of its polarity vector with approximately equal contributions (circular cell shape), the net cell motion is expected to be unaffected by changes in size. One contribution simply becomes stronger, while the other one becomes weaker. In other words, only the average κ_p value the cell experiences seems to contribute to the durotactic motion. Note that for cells with a more asymmetric shape, increasing the cell size might be able to enhance durotaxis.

In order to clarify the experimentally observed durotaxis for large cell clusters we have taken a closer look at the cell trajectories involved in durotaxis of a monolayer, which have revealed that on a substrate with uniform stiffness cells on the edge of the cluster on average tend to move outwards of the cluster [18]. This can be a consequence of cells not being able to actively move when is in contact with other cells, i.e. motion
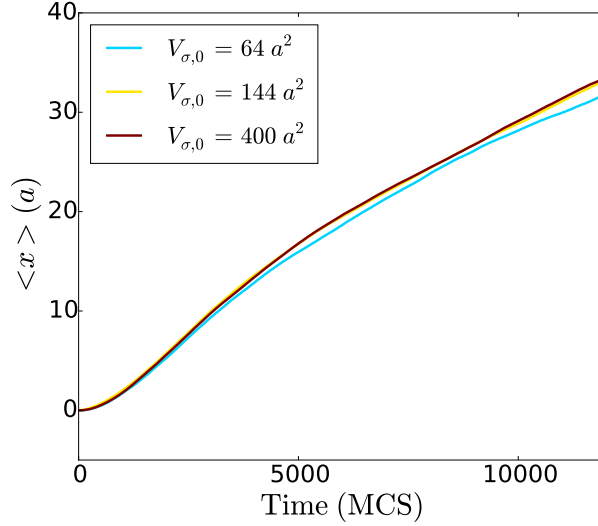
Figure 5.22: Plot of the average displacement in the x-direction as a function of time for single cells experiencing a linear gradient in κ_p between $\kappa_{p,\min} = 2.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ towards $\kappa_{p,\max} = 8.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ using a fixed width $w = 45a$. Results correspond to different preferred areas of the cell $V_{\sigma,0}$.

into the cluster is restricted due to there not being enough space. It is possible that in the case of durotaxis cells on the soft edge intrinsically still want to move outwards of the cluster but are dragged along by the faster cells on the stiff side of the membrane. In this scenario a larger difference in cell speed between both cells on the stiff and soft edge results in a stronger effect. This idea is also consistent with the experimentally observed requirement of cell-cell adhesions to induce collective durotaxis, since cells need to be dragged along [18].
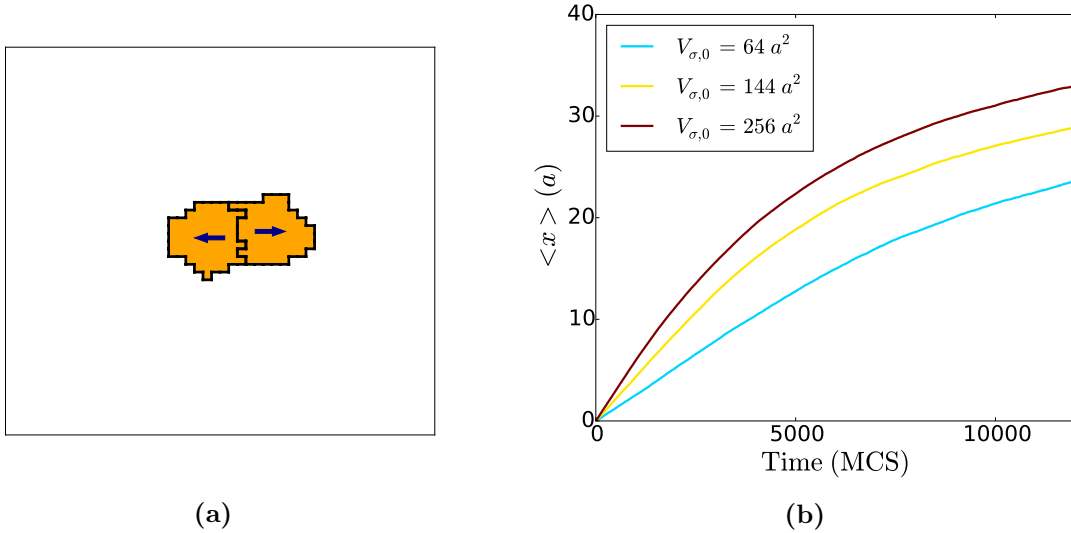

(a)

(b)

Figure 5.23: (a) Visualisation of a two cell cluster with opposite polarity vectors after equilibration inside a cell speed gradient. (b) Plot of the average displacement in the x-direction as a function of time for such a two cell cluster placed at the center of a linear gradient in κ_p between $\kappa_{p,\min} = 2.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ towards $\kappa_{p,\max} = 6.0\sqrt{V_{\sigma,0}/V_{\sigma,\mathrm{ref}}}$ using a fixed width $w = 20a$. Results are obtained for different preferred areas of the cell $V_{\sigma,0}$.

59

To model this effect, we have placed a cluster consisting of two cells that have a fixed polarity vector pointing in opposite directions along the x-axis, i.e. $\mathbf{p}_\sigma = (\pm 1, 0)$, in a gradient between $\kappa_{p,\text{min}} = 2.0\sqrt{V_{\sigma,0}/V_{\sigma,\text{ref}}}$ towards $\kappa_{p,\text{max}} = 6.0\sqrt{V_{\sigma,0}/V_{\sigma,\text{ref}}}$ (keeping the width w constant). A visualisation of the starting configuration (after equilibration) for $V_{\sigma,0} = 64$ is shown in fig. 5.23a. We mention that in order to have the cells stick together during the entire simulation we impose very strong adhesion by setting $J_{1,1} = -0.5$. Additionally, we only evaluate an active energy bias at the cluster edge (use eq. (3.4)).

The average displacement in the x-direction for different preferred cell areas $V_{\sigma,0}$ is plotted in fig. 5.23b. Interestingly, we can see that for this configuration increasing the individual cell size increases the strength of the observed durotaxis. Compared to the single cell case, it can be noted that there are now two competing effects, i.e. one cell wanting to migrate down the gradient and the other up the gradient. Due to the increasing size, the cell migrating up the gradient experiences a stronger active energy bias, while the cell migrating down the gradient experiences a weaker one. This means that the overall durotactic motion (migration up the gradient) is enhanced. This result can therefore suggest that when a CTC cluster is confined to a 1D domain or path, increasing the size of the cluster can improve its ability to follow a stiffness gradient. Note that this assumes that cells generally not tend to travel into the cluster, i.e. only the cells on opposite edges of the cluster actively migrate, and force transmission through such a cluster is sufficiently efficient. In the case of 2D (or 3D) motion this might be less evident, because cells have the ability to reorientate their polarity and reposition themselves relative to other cells.

5.4 Migration Through a Narrow Passage

We now seek to model the extracellular matrix (ECM) more explicitly. In cancer metastasis, cells typically have to pass through a large number of pores which are smaller than the cells [20]. Recent experimental work has highlighted the contribution of several parameters including cell-pore adhesion, nuclear volume and cell stiffness to the manner in which cells can pass through a pore which is smaller than its own size [20]. Furthermore, it has been suggested that cells need to generate large enough traction forces within the channel in order to pass through it [20]. These results have all been obtained for single cells passing through a narrow passage. Our aim, however, is to investigate how a circulating tumor cell (CTC) cluster passes through the pore relative to a single cell. It is therefore insightful to test whether several of the mentioned single cell parameters and the added cell-cell parameters have an influence on the cluster migrating through the narrow passage. Moreover, we are mainly interested in CTC clusters travelling through the 3D ECM, i.e. a network. Hence, all local passages within the ECM are expected to have a small length (approximately the size of a cell).

To study this, we have set up a simple simulation 'experiment' which is visualised in fig. 5.24. It involves starting with a cluster of N cells located to the left of a pore which is formed by two Gaussian-like shaped domains of immutable sites with spin $\sigma = -1$ and cell type $c(-1) = -1$. This implies that we take the pore to be immutable which is similar to experimental settings used in [20,21], but does not necessarily have to be true when cells travel through a relative soft matrix. The width of the pore equals 6 lattice sites such that it is smaller than a single cell diameter (approx. 9 lattice sites) but not so small that the influence of the cell nucleus becomes dominant; the latter effect is not addressed in our CPM. The length of the pore is set at 10 lattice sites so it is of the same size as a single cell. All cells then actively move at every cell edge [eq. (3.3)] with a constant polarity vector along the x-axis, i.e. $\mathbf{p}_\sigma = (1,0)$. This drives the cluster towards the pore and biases their motion to migrate through it. The moment the first cell enters the pore we start timing the passage of the cluster by setting the time to $t = 0$. We then time both the time it takes before the first and the last cell of the cluster have passed through the pore which we denote t_{first} and t_{last} respectively (see fig. 5.24). Additionally, we keep track of the number of fragments $N_{\text{fragments}}$ in which the cluster comes out on the other side of the pore. By averaging over many of these 'experiments' we can get a sense of the average time it takes to pass and to what degree a cluster is able to remain intact.

Before specifying the used parameters, we mention that the resulting simulations have been performed in 2D, whereas it is meant to provide insight about a 3D system, i.e. migration of cell clusters through a network with a small pore size. This has been done for computational convenience and mostly because we believe that similar simulations in 3D where a cell has to move through a circular or square opening for instance will provide comparable results. In other words, the general idea is that cells have to change their reference shape to pass through and we are mostly interested what the implications of that remodeling of the cell shape is on the way cell clusters move past the pore.

For convenience we choose all cells in the cluster to be equal. This means that all cells have the same cell type $c(\sigma > 0) = 1$ and experience (unless otherwise mentioned) the same strength of the active energy bias $\kappa_\sigma = 5$ which we denote as κ_p again. Specifically, we have fixed $T = 1$, $V_{\sigma,0} = 64$, $\lambda_v = 1$, $\lambda_r = 1$ and the cell-medium adhesion coefficient $J_{1,0} = J_{0,1} = 1$. The parameters that are left to specify are the ones we will vary. These are the adhesion coefficient between cells, $J_{1,1}$, and between a cell and the immutable obstacle, $J_{1,-1}$, which we call $J_{\text{cell}-\text{cell}}$ and $J_{\text{cell}-\text{pore}}$ throughout this section respectively.
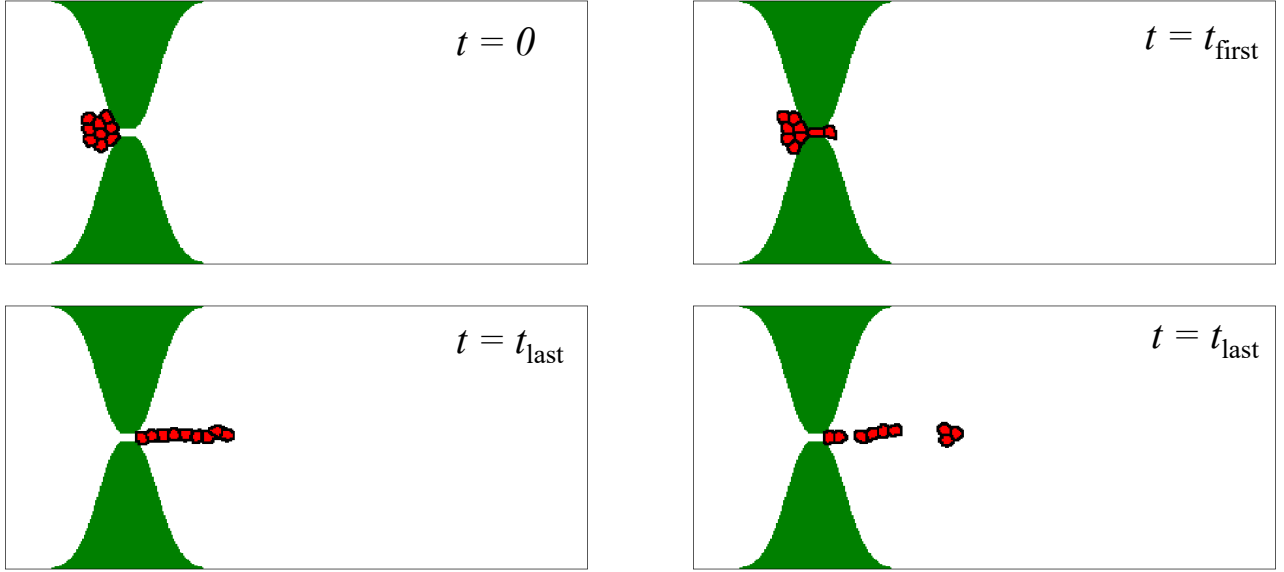
Figure 5.24: Visualisation of a cell cluster of $N = 9$ cells passing through narrow pore with a width $w = 6a$ (smaller than the single cell diameter) and a length $l = 10a$. The pore is formed by two domains of immutable lattice sites with $\sigma = -1$ shown in green. In the simulation the passage time for the first cell t_{first} and the last cell t_{last} are tracked. Furthermore, the cluster can pass as a whole or fragment in mutliple smaller clusters.

5.4.1 Cell-Cell Adhesion

We first consider the case of cluster migration without adhesion to the pore walls, $J_{\text{cell-pore}} = 1$. To this end, we have run a set of 200 simulation for a cluster of $N = 9$ cells where we have recorded t_{first}, t_{last} and the amount of fragments in which the cluster has fallen apart for different cell-cell adhesions $J_{\text{cell-cell}}$. The resulting times and fragment distributions have been plotted in fig. 5.25. We highlight that the smaller the value of $J_{\text{cell-cell}}$ the stronger the cells adhere to each other, since it requires less energy to form a cell-cell surface. Moreover, when $J_{\text{cell-cell}} = 2$ it is not energetically favorable anymore to form a cell-cell surface in comparison to both cells forming a surface with the medium ($\sigma = 0$) since we have set $J_{1,0} = 1$ [45]. This means that $J_{\text{cell-cell}} = 2$ denotes the threshold value below which cells adhere to each other.

It can be seen that the time it takes for the first cell to pass through does not alter much by changing the cell-cell adhesion. This makes sense since we expect that cell-cell adhesion will mostly play an important role in the ability of cells that just passed through the pore to drag along the cell behind them through it as well. We therefore predict more variety in the time it will take for the entire cluster to move past the pore. This is demonstrated in fig. 5.25a which shows that upon weakening the cell-cell adhesion the value of $\langle t_{\text{last}} \rangle$ rises significantly. Particularly, we can link this behavior to the degree of fragmentation of the cluster shown in fig. 5.25b. Notice that when $J_{\text{cell-cell}} = 0.0$ and the adhesion between cells is the strongest, the cluster always comes out as a single fragment, i.e. as a whole. This implies that all cells except for the first one are dragged along through the channel by another cell, making the passage of the entire cluster relatively fast. Additionally, for $J_{\text{cell-cell}} = 0.5$ most of the clusters ($\sim 90\%$) still

pass through as a whole and as a result the increase in $\langle t_{\text{last}} \rangle$ is relatively small. Upon further increase of $J_{\text{cell}-\text{cell}}$ we can see that the peak in the distribution shifts towards a larger amount fragments and eventually no clusters pass as a whole anymore. This translates itself to a strong increase in the average cluster passage time $\langle t_{\text{last}} \rangle$, especially in the region between $0.5 - 1.5$ as can be observed in fig. 5.25b. Note that the spread in the values of $\langle t_{\text{last}} \rangle$ also increases heavily which can be linked to the observed spread shown in the distribution of the number of fragments.
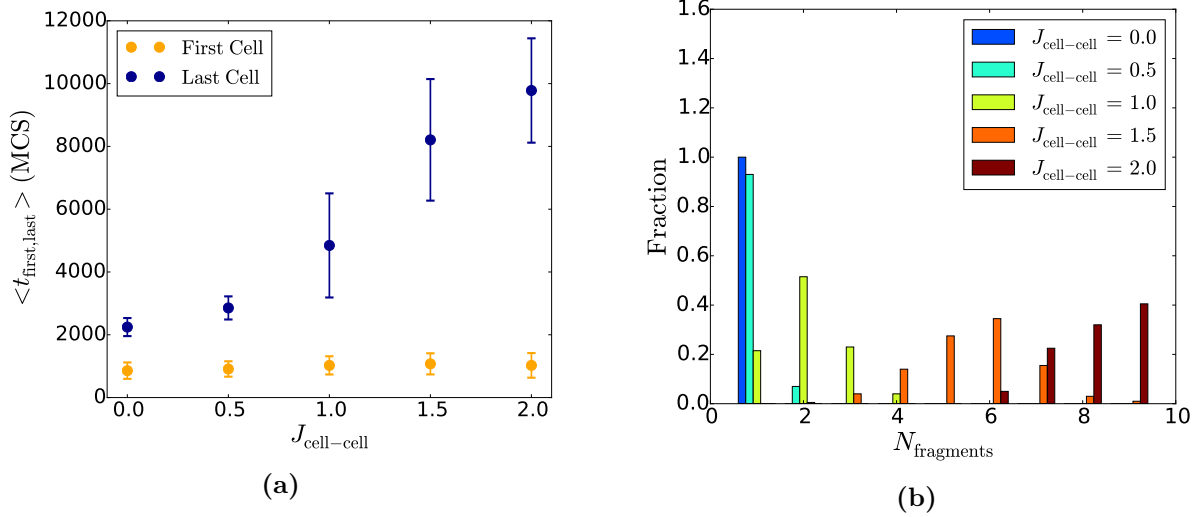

Figure 5.25: (a) Plot of the average time it takes for the first cell $\langle t_{\text{first}} \rangle$ and the last cell $\langle t_{\text{last}} \rangle$ of a $N = 9$ cell cluster to pass through a narrow passage as a function of the cell-cell adhesion coefficient $J_{\text{cell}-\text{cell}}$. (b) Plot of the distribution of fragments in which the cluster has split after passing through the pore for different $J_{\text{cell,cell}}$. Averages and distribution are obtained from 200 runs.

Furthermore, we can see that the time it takes before the cluster passes through the pore seems to flatten out a little when approaching $J_{\text{cell}-\text{cell}} = 2.0$. This can be understood from the fact that when cells do not adhere anymore, every cell will not experience the benefit of travelling as a cluster and essentially pass through like it would when migrating individually. The corresponding value for the passage time would then be given by $\langle t_{\text{last}} \rangle \approx N \langle t_{\text{first}} \rangle$ which is roughly consistent with the found results. Also, since the cluster completely splits up most of the times for $J_{\text{cell}-\text{cell}} = 2.0$ and the spread in the distribution of fragments decreases with respect to $J_{\text{cell}-\text{cell}} = 1.5$ for instance, the spread in the values of $\langle t_{\text{last}} \rangle$ is expected to decrease as well. This is confirmed in fig. 5.25.

Thus, within a simple model setup we have demonstrated the important that role cell-cell adhesion can play for the ability of a 'small' cell cluster to pass through a narrow pore. Strong adhesions can greatly improve the cluster's ability to pass through since cells can drag other cells along with them. An important factor therefore seems to be whether or not the adhesion is strong enough to allow the cluster to pass the pore in one piece so that cells can take advantage of this process. We can even imagine the extreme situation in which it is difficult to pass a really narrow pore, but due to strong adhesion cells can quickly move after the first one has passed so the effect of the hindrance is not so different for single cells in comparison to clusters.

5.4.2 Cell-Pore Adhesion

To mimic the effect of interactions with the ECM, we now seek to investigate the role of the cell adhesion to the pore walls. This means that we fix $J_{cell-cell} = 1$ such that in the case of no adhesion or repulsion to the pore, i.e. $J_{cell-pore} = J_{0,1} = 1$, the cluster mostly fragments in two pieces (see fig. 5.25b) and we can test whether changing $J_{cell-pore}$ keeps the cluster more intact or fragments it even more. Therefore, we have performed sets of 200 simulations for different $J_{cell-pore}$ values in which we let a cluster of $N = 9$ cells pass through the pore. The resulting average values of t_{first} and t_{last} and the distribution $N_{fragments}$ have been plotted in fig. 5.26.

It can be noted that $\langle t_{first} \rangle$ increases with $J_{cell-pore}$. Having the cells adhere less strongly to, or even repel from, the immutable obstacle makes it harder for cells to move through the pore. Since cells are almost completely in contact with the pore walls when they pass through, this configuration becomes energetically harder to reach when $J_{cell-pore}$ increases.
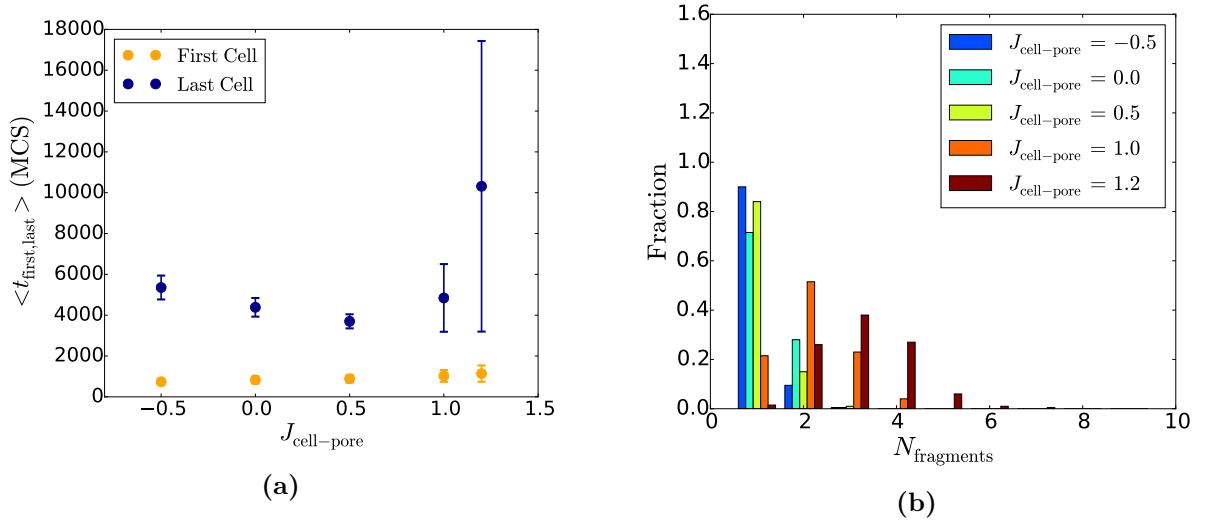

Figure 5.26: (a) Plot of the average time it takes for the first cell $\langle t_{first} \rangle$ and the last cell $\langle t_{last} \rangle$ of a $N = 9$ cell cluster to pass through a narrow passage as a function of the cell-pore adhesion coefficient $J_{cell-pore}$. (b) Plot of the distribution of fragments in which the cluster has split after passing through the pore for different $J_{cell,pore}$. Averages and distribution are obtained from 200 runs.

By contrast, the average time it takes before the entire cluster has passed through, $\langle t_{last} \rangle$, shows a less trivial dependence on $J_{cell-pore}$. There seems to exist an optimum amount of adhesion around a value of $J_{cell-pore} = 0.5$ for which the cluster passes through the pore the fastest. We cannot link this optimum directly to the distribution of fragments, which states that the cluster always passes through as a whole only for $J_{cell-pore} = -0.5$ even though this setting does not lead to the fastest passage of the cluster. The reason that the cluster appears to pass more slowly when adhesion becomes extremely strong can possibly be sought in the fact that cells remain attached to the obstacle after passing the site. This is shown in fig. 5.27a. By being forced to follow the Gaussian geometry of the pore, the velocity of the cells will slow down since the cell polarity is along the x-axis. Therefore we expect that it will become more difficult for the cells to drag along the cells behind them through the pore, resulting in the cluster passing more slowly even though the cluster remains intact.

The tendency of cells to move along the geometry of the obstacle for strong enough

adhesion can also explain why we observe a larger chance of cluster fragmentation for $J_{\text{cell}-\text{pore}} = 0.0$ in comparison to $J_{\text{cell}-\text{pore}} = 0.5$ (see fig. 5.26b). By visualising several simulations we find that for $J_{\text{cell}-\text{pore}} = 0.0$ cells are trying to follow the obstacle but that the adhesion to the pore is not sufficiently strong for all cells to do so. As a result, some fragments detach and move further into the medium whereas others remain attached to the pore wall, which is demonstrated in fig. 5.27b. For $J_{\text{cell}-\text{pore}} = 0.5$ we have observed that the adhesion is too weak to have the cells follow the pore geometry. Instead they always remain moving into the medium along the x-direction (see fig. 5.24), which prevents this source of fragmentation and as a result the probability of the cluster to pass as a whole is larger. Thus, it seems that the cluster passes the pore the fastest when adhesion helps single cells travel through more easily, but it is not strong enough to force the cells to follow the pore's geometry.

On the other hand, when cells are repelled by the obstacle, i.e. $J_{\text{cell}-\text{pore}} > J_{1,0} = 1$, we observe an especially sharp increase in the average time it takes the cluster to pass. This can be directly linked to the observed increase in $\langle t_{\text{first}} \rangle$ and the fact that the distribution of $N_{\text{fragments}}$ shifts to larger values and spreads out more. In other words, cells have more problems passing through and also more often need to do so without the help of the cell in front of them.

Summarising, we have demonstrated how cell adhesion to the channel it needs to pass through in our model system can lead to cells passing more easily while remaining more intact as well. Interestingly, when adhesion becomes too strong, cells are forced to follow the direction of the structure they adhere to which, within the investigated setup, limits the ability of a cluster to pass the pore.
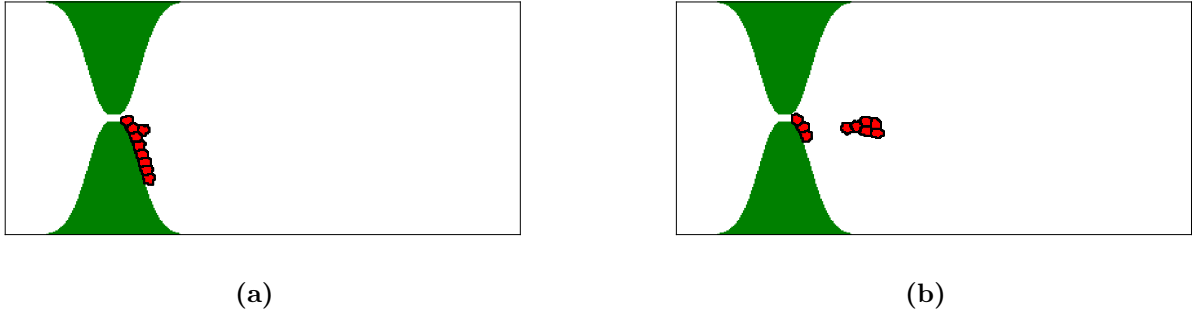

(a)   (b)

Figure 5.27: Visualisation of the final configuration resulting from a single simulation in which we let a $N = 9$ cell cluster pass a pore. Results correspond to (a) $J_{\text{cell}-\text{pore}} = -0.5$ and (b) $J_{\text{cell}-\text{pore}} = 0.0$. Other used parameters are discussed in text.

5.4.3 Traction Force

We now turn our attention to the influence of cell traction forces, i.e. the strength by which cells pull on the surrounding material that encloses it when inside the pore, on the total cluster passage. This has been suggested to play an important role on the manner in which single cells pass a narrow pore, since these forces are used to elevate the cell through the pore and into the medium again. In fact it appears that traction forces are part of an interplay of (cell) parameters that determine the velocity at which cells pass through the pore [20]. Furthermore, it has been shown that the confinement of the pore leads to cells moving in a highly persistent manner and with an increased velocity (factor ~ 2, 3) when passing through it [20,21]. It is reasonable to assume that

the amount of traction force the cell can generate relates to the increased velocity of the cell within the pore. Specifically, we assume that the cells exert stronger traction forces when inside the pore; while other factors such as the size of the nucleus have also been pointed out to influence the amount of increased velocity [20, 21], we will not consider these in the present work.

In order to model the traction forces we add a value of $\Delta\kappa_p$ to the existing polarity strength of each cell κ_p when the cell's center of mass $\mathbf{R}_\sigma$ resides within the pore. Note that in the proposed model we thus seek to describe stronger traction forces within the pore by increasing the polarity force acting on the cell. We have then performed 200 simulation 'experiments' for a cluster of $N = 9$ cells corresponding to different values of $\Delta\kappa_p$ keeping $J_{\text{cell}-\text{pore}} = 1.0$ and $J_{\text{cell}-\text{cell}} = 0.5$. The obtained average values of t_{first} and t_{last} and the distribution $N_{\text{fragments}}$ are plotted in fig. 5.28.
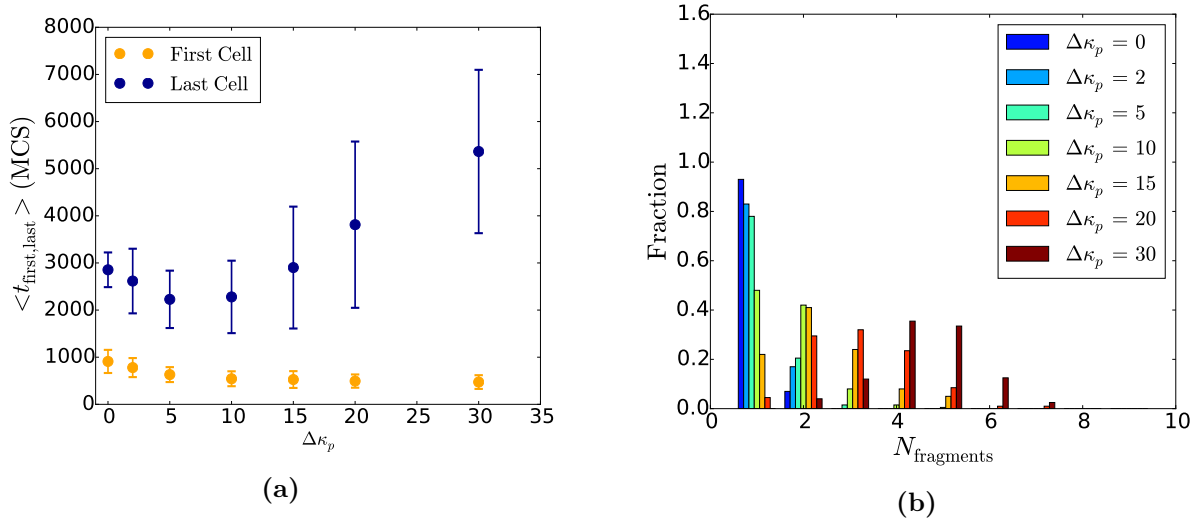

Figure 5.28: (a) Plot of the average time it takes for the first cell $\langle t_{\text{first}} \rangle$ and the last cell $\langle t_{\text{last}} \rangle$ of a $N = 9$ cell cluster to pass through a narrow passage as a function of additional polarity strength the cell retrieves within the pore $\Delta\kappa_p$. (b) Plots of the distribution of fragments in which the cluster has split after passing through the pore for different $\Delta\kappa_p$. Averages and distribution are obtained from 200 runs.

We can see that increasing the polarity strength within the pore results in the first cell passing through more quickly ($\langle t_{\text{first}} \rangle$ decreases), which seems logical since the cells can pull themselves forward with more force. Initially, the value of $\langle t_{\text{last}} \rangle$ follows the same trend, but upon further increasing $\Delta\kappa_p$ we can see that it eventually starts to rise significantly and as a result there appears to be an optimum value for $\langle t_{\text{last}} \rangle$ around $\Delta\kappa_p \sim 5-10$. This behavior can be connected to the fragmentation of the cell cluster. Specifically, we can notice that the optimum range of $\Delta\kappa_p$ also denotes the point at which the cell cluster is beginning to fragment. Further fragmentation seems to slow down the average passage of the cluster even more, which has already been linked to the fact that fewer cells are being dragged along by the cell in front of them. Moreover, we observe a larger spread in $\langle t_{\text{last}} \rangle$ when the spread in $N_{\text{fragments}}$ increases as well.

We can rationalise these results by realising that when the cells are able to pull themselves through the pore very strongly, it will happen so fast that cells are more likely to detach from the cells behind them; the latter do not experience the added force yet, leading to the formation of more fragments. In turn the leading cell will no longer facilitate the movement of the following cells, which slows down the average passage of the cluster. Interestingly, at least within the investigated range of $\Delta\kappa_p$, the added effect

of individual cells being able to pass through more rapidly as result of increased polarity force is thus overshadowed by the added fragmentation effect cells experience when a cell in front can drag them along.

Thus, we have demonstrated that although an additional force always improves the ability of a single cell to pass the pore, it does not necessarily mean that a cell cluster benefits from it as well. Whether this is actually the case seems to depend on the degree to which cell-cell adhesion improves cluster passage by having cells drag along other cells behind them and on the fact that the added force should not be too strong so that it leads to significant fragmentation of the cell cluster. Finally, on a general note, we want to mention that recent experimental work has also been performed to study the ability of CTC clusters to pass through a narrow vessel when it travels through the bloodstream [3]. In fact they have demonstrated that the clusters containing up to 20 cells are able to pass through a vessel smaller than a single cell [3]. The used simulation set up could also be used as a model system to study this scenario. However, this would require that we introduce inertia into the CPM, since the motion within a flow (the bloodstream) is not Aristotelian.

5.5 Focal Adhesion Formation

In order to provide more insight about cluster migration during metastasis we have chosen to simulate different aspects of the extracellular matrix (ECM), i.e. different environments. The final ECM environment in which we seek to compare single cell to cluster migration, is one that involves explicit focal adhesions, which are modeled by the addition of binding sites to the ECM medium (see section 3.1.4). As discussed in section 3.1.4, this requires a binding energy of a CPM site $E(t_{\mathbf{x}})$. In this section we will infer the binding energy, which has been used in the next section, from more detailed Langevin Dynamics simulations. In these simulations we zoom in on the formation of a single focal adhesion; in particular, we track the time evolution of receptors (integrins) in a model cell membrane and retrieve the average number of receptors that have bound to a cylinder (collagen fiber) $N_{\text{bound}}(t)$

Before proceeding, we mention that throughout this section we have kept the fraction of receptors, the volume exclusion strength in eq. (3.19) and the cutoff radius and width of the Morse potential [eq. (3.20)] fixed at 10%, $U_0 = 4.34 k_B T$, $r_c = 4.0\sigma$ and $\alpha = 1.0\sigma^{-1}$ respectively. Moreover, we have (unless otherwise stated) adopted similar membrane settings to the ones used in [67] for which they have demonstrated that the mechanical properties of the membrane fall within the experimental regime. Specifically, we have set $\mu = 3.0$, $\zeta = 4.0$ and $u_0 = 4.34 k_B T$ in eq. (3.14), yielding a bending rigidity of $\sim 20 k_B T$ and area compression modulus of $\sim 18 k_b T/\sigma^2$ [67]. To avoid confusion we state that in this section σ denotes the membrane particle diameter and not the CPM spin.

5.5.1 Normal Fiber

Since the cylinder is used to model a collagen fiber, we take its diameter to be consistent with the size of a collagen fiber. Realising that the typical thickness of the cell membrane (one particle diameter) is ~ 10 nm [83], we have set the cylinder radius at $r_{\text{cyl}} = 10\sigma$ to give the cylinder a typical diameter of a collagen fiber of ~ 200 nm [84–86]. We have then run 20 simulations and determined the average number bound receptors N_{bound} divided over the total number of receptors in the membrane N_{rec} as a function of time for different attractive strengths D_0 of the Morse potential. The results are shown in fig. 5.29.

It can be seen that the number of bound receptors initially increases and eventually saturates which is consistent with the idea of maturation of a focal adhesion. Moreover, the curves look very similar for different attractive strengths D_0. The only thing that noticeably changes for stronger receptor attraction to the cylinder is the saturation value of N_{bound}, which becomes larger. A possible explanation can be sought in a tradeoff between bending energy of the membrane and binding energy of the receptors. To have more receptors attach to the cylinder requires the membrane to wrap more around the cylinder and eventually this will reach an equilibrium which sets the saturation value of bound receptors.

In order to study the time evolution of N_{bound} in more detail, we seek to fit the obtained curves with a model function to infer a timescale of the maturation. A simple saturation curve is given by an exponential distribution given by

$$N_{\text{bound}}(t) = N_{\text{sat}} \left(1 - \exp(-\frac{t}{\tau_{\text{mat}}}) \right), \tag{5.1}$$

with N_{sat} the saturation number of bound receptors and τ_{mat} a time scale that describes the maturation. However, a fit with this function provides very inaccurate results on

short time scales. This suggests the existence of (at least) two processes involved in the growth of N_{bound}; to capture this we have fitted the curves with a bi-exponential function given by

$$N_{\text{bound}}(t) = N_{\text{sat}} \left(1 - w \exp(-\frac{t}{\tau_{\text{fast}}}) - (1-w)\exp(-\frac{t}{\tau_{\text{slow}}}) \right), \qquad (5.2)$$

with τ_{fast} and τ_{slow} two timescales describing the short and long time behavior, respectively, and w a weight factor that describes the contribution of each component. We can see that the curves can be very accurately fitted with this function. (Note that this is also a result of added fit parameters.)
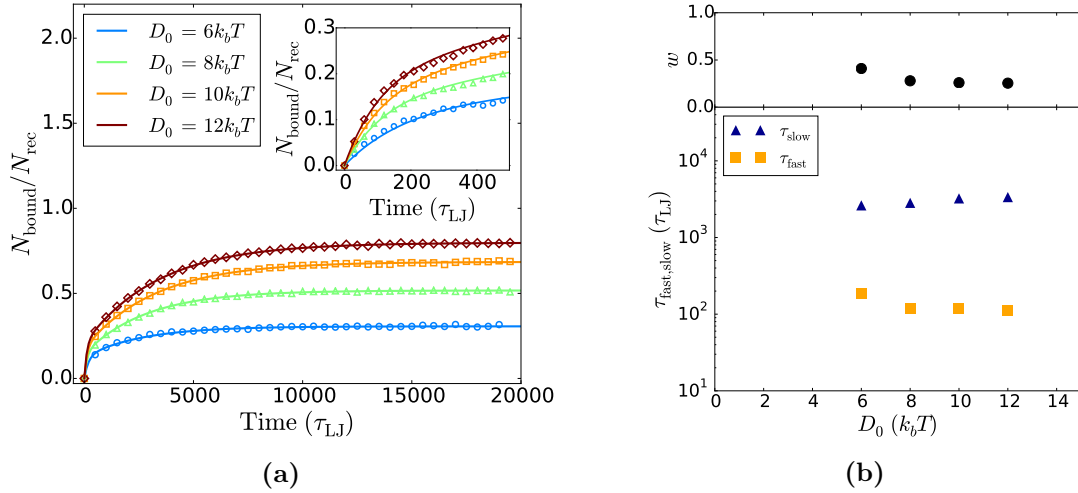

Figure 5.29: (a) Plot of the average fraction of receptors bound to the cylinder as a function of time obtained from simulations (markers) which have been fitted with eq. (5.2) (lines). Inset zooms in on short time behavior. (b) Plot of the bi-exponential parameters τ_{slow}, τ_{fast} and w resulting from the fits shown in (a).

To verify the existence of two distinct processes we have plotted the fitted time scales τ_{fast} and τ_{slow} in fig. 5.29b. It shows that these time scales differ by approximately one order of magnitude and that they are not significantly influenced by changes in D_0. Thus, there seem to exist distinct fast and slow processes, and both of them contribute to the maturation process since $w \neq 0, 1$. We expect, based on visualisations of the cell membrane [see fig. 5.30], that the fast process corresponds to the initial attachment of the membrane to the cylinder that lies above it, while the slow process represents the assembly of more and more receptors that have diffused from other parts in the membrane towards the contact area of the membrane with the cylinder.

To test this idea we have run a similar set of 20 simulations as before but decreased ζ to a value of 2.5. It has been shown that for this setting the membrane enters a more gel-like state and the diffusivity of membrane particles is significantly decreased [65,67]. Again we have determined $N_{\text{bound}}/N_{\text{rec}}$ for different attractive strengths of the Morse potential, plotted in fig. 5.31a. Following the reasoning of the hypothesis, we expect that the initial process is roughly unaffected by this change, while the second process should slow down. This is confirmed in fig. 5.31a, which shows that after an initial increase, i.e. the fast diffusive process, the amount of bound receptors increases less rapidly afterwards. Furthermore, a fit of the resulting curves with eq. (5.2) (see fig. 5.31) shows that τ_{slow} has increased, indicating the slowing down of the assembly of receptors around the cylinder.
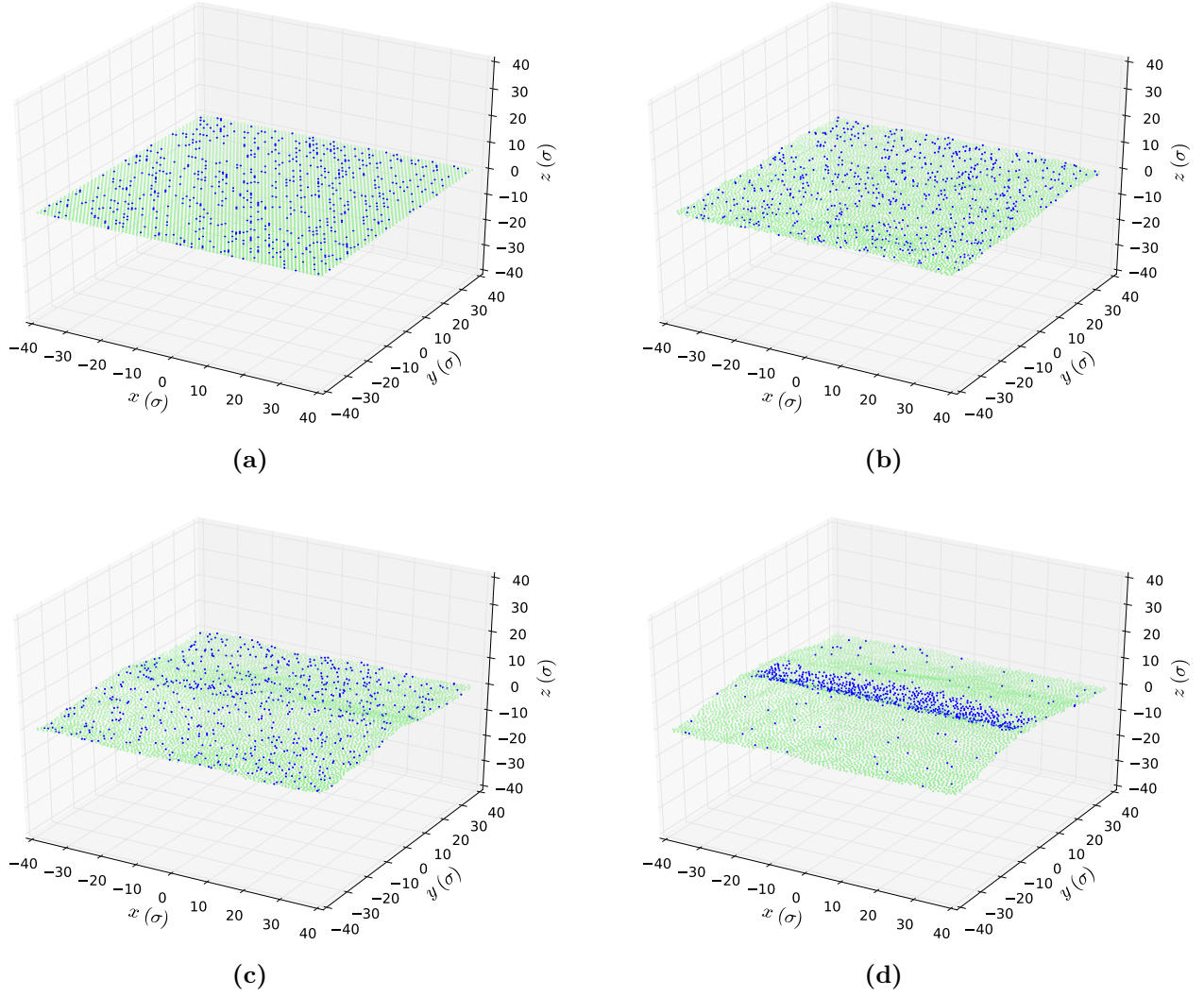
Figure 5.30: Visualisations of the cell membrane (a) in its initial configuration (b) before attachment, (c) directly after attachment, and (d) after assembly of receptors. Receptors are shown in blue. To improve visibility, we have not shown the cylinder, which floats above the membrane. Results correspond to normal membrane settings and $D_0 = 12 k_B T$.

We have also investigated the influence of the membrane flexibility by running similar simulations for $\mu = 2.4$ and $\mu = 6.0$ which changes the bending rigidity of the membrane within experimental extremes [65, 67]. The resulting curves of $N_{\text{bound}}/N_{\text{rec}}$ and fit parameters are approximately the same as the ones shown in fig. 5.29 (see figs. D.9 and D.10). This suggests that within the considered regime the membrane flexibility is not a dominant factor in the assembly of receptors.

Thus, we have identified two distinct processes (attachment and assembly) involved in the maturation of the model focal adhesion, which appear, under normal circumstances, to occur on time scales that differ by approximately one order of magnitude. The generic behavior of these processes seems to be unaffected by the attractive strength of the potential and membrane flexibility, while the assembly of receptors can be significantly slowed down by decreasing the diffusivity of membrane particles.
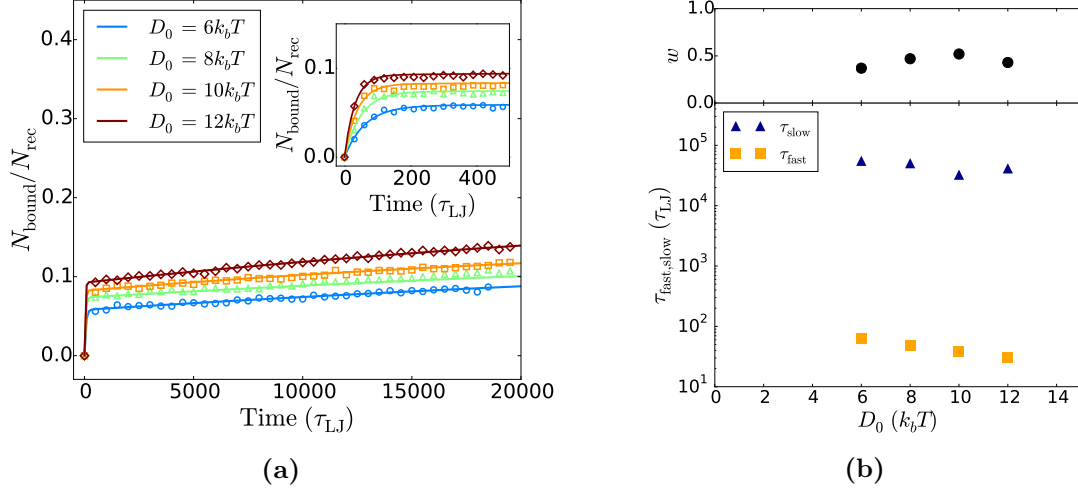
(a)   **(b)**

Figure 5.31: (a) Plot of the average fraction of receptors bound to the cylinder as a function of time obtained from simulations with a gel-like membrane by setting $\zeta = 2.5$ (markers) which have been fitted with eq. (5.2) (lines). Inset zooms in on short time behavior. (b) Plot of the bi-exponential parameters τ_{slow}, τ_{fast} and w resulting from the fits shown in (a).

5.5.2 Thin Fiber

Finally, to establish whether our choice for the cylinder size affects our results. In particular, we have set the cylinder radius at $r_{\text{cyl}} = 1\sigma$ such that the diameter is decreased by one order of magnitude. Running a set of 20 simulations for the standard membrane settings, we have determined $N_{\text{bound}}/N_{\text{rec}}$ as a function of time for different D_0 and plotted the results in fig. 5.32.

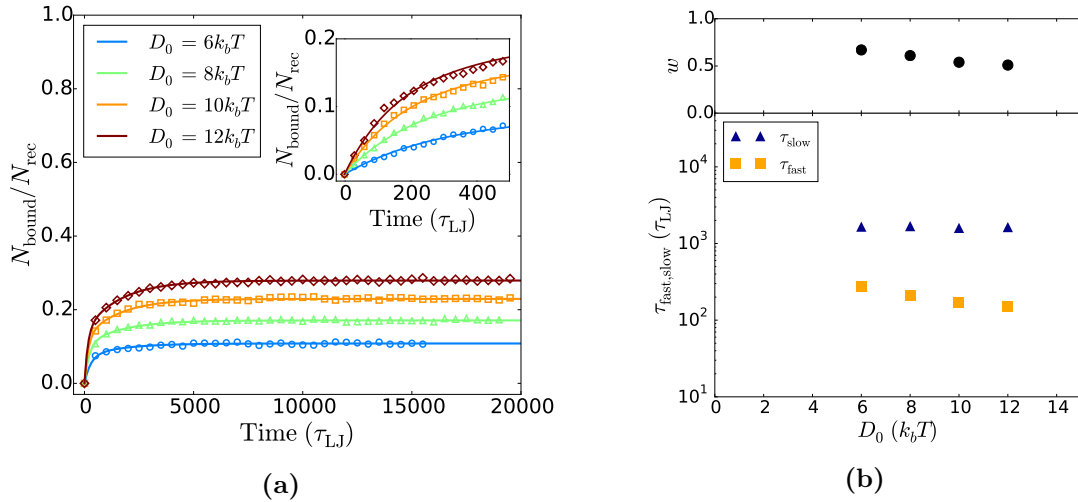**(a)**   **(b)**

Figure 5.32: (a) Plot of the average fraction of receptors bound to the cylinder as a function of time obtained from simulations (markers) which have been fitted with eq. (5.2) (lines). Inset zooms in on short time behavior. (b) Plot of the bi-exponential parameters τ_{slow}, τ_{fast} and w resulting from the fits shown in (a).

Inspecting fig. 5.32 shows that the generic behavior of the time evolution of $N_{\text{bound}}/N_{\text{ref}}$ for a smaller cylinder is roughly similar to the result of the original fiber size (fig. 5.29), and as a result can be accurately fitted with eq. (5.2) as well. The only differences are the smaller saturation values of N_{bound} and the fitted time scale for the assembly process

τ_{slow} appears to have decreased. We can understand these differences by noting that a smaller cylinder provides less space for the receptors to attach to, and that it requires more local bending energy (higher curvature) to wrap around the cylinder. Therefore fewer receptors attach to the cylinder. Moreover, since the diffusivity of the receptors remains the same, it will take less time to reach the saturation value as fewer particles are required. This can explain the smaller values of τ_{slow}.

In total, the observed behavior for the assembly of receptors at a rigid cylinder seems to behave rather robustly upon changing relevant membrane and cylinder parameters. Therefore, we will take it as inspiration for the binding energy in the CPM, i.e. $E(t_{\mathbf{x}})$. We highlight that this potential will still be a rough approximation, since it has been mentioned that at least some integrins attach to the ECM via so called catch bonds whose lifetime increases with increasing load [87], while we model binding integrins via a Morse potential, i.e. a slip bond whose lifetime decreases with increasing load.

5.6 Cell Binding to the Extracellular Matrix

In this section we want to compare the difference between single cells and clusters passing through a medium which includes randomly distributed binding (focal adhesion) sites. This is expected to be a more accurate description of the environment which can either represent a substrate where cells can locally bind at different locations or a simplification of cell motion through the 3D extracellular matrix (ECM) where the binding sites correspond to points in the network where cells can bind to it.

Specifically, we model a single cell and a cluster consisting of $N_{\text{cells}} = 9$ identical $(c(\sigma > 0) = 1)$ cells in an environment with no explicit ECM $(\sigma \geq 0)$, but with a fraction ρ of all sites randomly denoted as binding sites $\chi(\mathbf{x})$ where a focal adhesion can form. The binding energy (based on the results from the previous section) is given by

$$E(t_{\mathbf{x}}) = -E_0 \left(1 - w \exp(-\frac{t_{\mathbf{x}}}{\tau_{\text{fast}}}) - (1 - w) \exp(-\frac{t_{\mathbf{x}}}{\tau_{\text{slow}}}) \right). \tag{5.3}$$

Here E_0 denotes the saturated binding energy and we have fixed $w = 0.5$ for convenience. Since focal adhesions mature in typically ~ 1h [88] and the time scales inferred from the Langevin Dynamics simulations differ approximately one order of magnitude, we have set $\tau_{\text{fast}} = 50$ MCS and $\tau_{\text{slow}} = 500$ MCS. We let the cells drift through this environment by imposing a linear concentration gradient along the x-axis, $C(x) = x + C_0$ with C_0 an arbitrary constant, and using a chemotactic strength $\lambda_c = 0.5$ in eq. (3.2) (note that active motion due to polarity is ignored in this section). Additionally, the following parameters have been kept fixed at: $T = 1$, $\lambda_v = 1$, $V_{\sigma,0} = 64$, $\lambda_r = 1$, $J_{0,1} = 1.0$ and $J_{1,1} = 0.5$ (strong cell-cell adhesion).

To test the effect of the added focal adhesions on cell migration, we have then run sets of 200 2D CPM simulations for different energies E_0 and fractions (or focal adhesion densities) ρ. Using the obtained trajectories we have determined the average absolute velocity $\langle v \rangle$ of both the cell and the cluster by time-ensemble averaging and choosing a time period $\delta = 1000$ MCS in eq. (4.6) (again to neglect 'thermal' effects). The obtained values give an insight about how fast a cell can migrate through such an environment in comparison to a cell cluster and have been plotted in fig. 5.33.

The results demonstrate that by increasing either the strength of the individual bonds E_0 or the density of binding sites ρ the average velocity of the cell and the cluster decreases and the migration seems more hindered. This makes sense since more and stronger binding tends to keep the cell or cluster more in place and it becomes harder for the chemotactic gradient to pull them loose and drive them in the x-direction. Interestingly, we can observe a more sudden drop in average velocity upon increasing E_0, whereas it occurs more gradually when ρ becomes larger. This suggests that at some point the binding energy E_0 becomes sufficiently high that both a chemotactic bias and 'thermal' fluctuations almost cannot break the bond anymore which means that the cells are severely limited in their motion and the average velocity rapidly drops towards zero.

However, it can be seen that the dependence on these parameters is roughly the same for single cells and a small cell cluster. It seems that the migration of cells is not influenced by the fact that they travel as one cluster or not. This means that effects like dragging along cells that lag behind or cell-cell attachment do not significantly change the migration within the investigated regime. The main competition appears to be primarily between chemotaxis that wants to drive the cells forward and the focal adhesions that want to keep the cell in place. The only difference between single cell and cluster motion is the overall slower migration of a cell cluster. This is probably a consequence of the definition of the chemotactic energy bias $\Delta \mathcal{H}_c$, i.e. eq. (3.2).
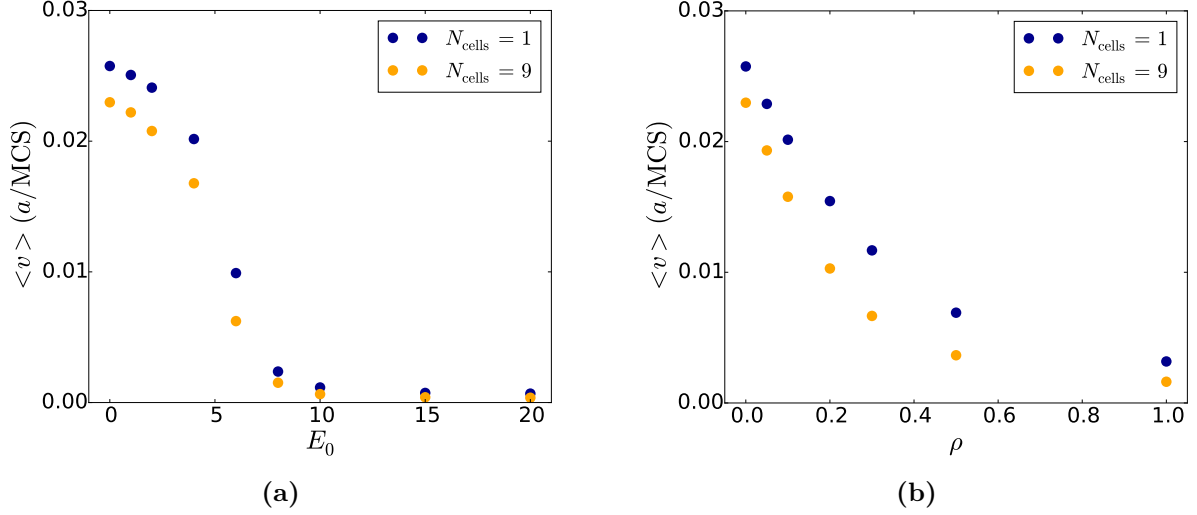
Figure 5.33: Plots of the average cell velocity $\langle v \rangle$ as a function of (a) the saturated binding energy E_0 (with $\rho = 0.1$) and (b) the relative density of binding sites ρ (with $E_0 = 4$) for a single cell and cluster of $N_{\text{cells}} = 9$ cells passing through a medium with randomly distributed binding (focal adhesion) sites.

Thus, we have demonstrated that the obtained binding energy of a focal adhesion can be implemented into the CPM resulting in a slowing down of both cells and clusters that becomes more dominant when the bond strength and the density of bonds increases. However, binding on itself appears to have similar effects on how a single cell or small cell cluster migrates. We expect that a difference becomes more apparent when the model is extended such that focal adhesions not only bind cells, but also influence their active motion. Specifically, since focal adhesions are not only used to bind cells to the ECM but also to pull them forwards and to reorientate their cytoskeletons, it might be interesting to add a cell polarity whose direction and strength is determined by the focal adhesions and degree of maturation of them. Moreover, it might be interesting to extend the focal adhesion model description to three dimensions. This includes for instance effects like the inability of cells to bind to focal adhesions at places where it is in contact with other cells.

6. Conclusion

In this thesis we have modeled and compared single versus collective cell motion in the context of circulating tumor cell (CTC) clusters using the cellular Potts model (CPM). The primary aim has been to gain more insight in why small CTC clusters demonstrate a dramatically enhanced metastatic potential. Here we will briefly state the most important findings and provide some suggestions for future research.

After a review of the theory of active Brownian motion, we have analysed the motion of clusters of fast-aligning active Brownian particles. Using a velocity alignment potential we have derived how fast alignment increases the persistence time of the cluster, yielding a linear scaling with the number of particles. If the cluster is active enough, this allows it to cover more distance than single particles, thus increasing metastatic success. The added effect of alignment is strongest for identical particles and can be counteracted by adding a particle that rapidly reorientates to the cluster.

Next, CPM simulations have been carried out to model cell motion. Single cell simulations have shown good agreement in terms of mean square displacement with the theory of active Brownian motion and with experimental results in 2D and in 3D. In particular, the cell speed and persistence time can be accurately controlled by an active energy bias and an implemented persistence time.

Extending the simulations to small cell clusters, we have examined the effect of cell-cell alignment by adding a Vicsek-like model to the CPM. Consistent with theory, alignment enhances the persistence of cell clusters and improves cluster motion, improving their metastatic ability. We have found that weakening the strength of alignment beyond a critical point results in a relatively rapid transition to a non-aligned state. This state is characterised by a similar persistence as for a single cell and a disintegration or significant slowing down of the cluster due to individual cells wanting to move in different directions. Additionally, we have shown that the strength of alignment required to have the cluster benefit from it, is strongly linked to the individual persistence time of cells, i.e. stronger individual persistence requires weaker alignment. Since single cell persistence is known to be enhanced by the stiffness of the environment, a cluster might only benefit from alignment when the ECM is stiff enough.

We have then focused on durotaxis, i.e. the migration up a stiffness gradient. To this end, we have imposed a linear persistence time or local cell speed gradient in the CPM. We have demonstrated that in a persistence time gradient there exists no durotactic benefit originating solely from an increased size of the cluster. Alignment, however, increases the persistence time of the cluster and so as a whole it experiences a stronger gradient which leads to improved durotaxis. In comparison, a cell cluster can benefit from only an increased size when we implement a local cell speed gradient and let cells on opposite edges tend to move in opposite directions. This is consistent with experimental results and suggests that the effect can be especially dominant when cells are confined to a 1D domain or path.

To study the ECM environment through which cells migrate more explicitly, we have

run simulations involving small cell clusters passing though a narrow passage. These have indicated how strong cell-cell adhesion can greatly decrease the average passage time of the cluster relative to each cell migrating on its own, limiting the disadvantage of migrating as a cluster. We have also demonstrated that a cluster passes most rapidly (and much faster relative to each cell migrating on its own) when cells adhere to the obstacle or experience an increased traction force inside the pore.

Finally, to introduce a more extended description of the extracellular matrix into the CPM, we have modeled the formation of a focal adhesion with more detailed Langevin Dynamics simulations. Specifically, we have tracked the number of membrane receptors attached to a rigid cylinder (collagen fiber) above the membrane. We have identified two distinct processes, attachment and assembly, involved in the maturation of the model focal adhesion, which occur over time scales that differ by approximately one order of magnitude. The saturation curves of the number of bound receptors are shown to behave robustly and their functional form has been implemented as a binding energy in the CPM to model sites where focal adhesions can form. Letting either cells or small clusters drift through an environment of such focal adhesion sites shows roughly the same results. This indicates that slowing down of the cells due to binding at focal adhesion sites on itself does not favor one over the other.

6.1 Outlook

We have demonstrated that already in a minimal model setup single cell and cluster migration can be significantly different and for certain settings the results favor migration in larger clusters. However, there still remain open questions. Firstly, we have captured all effects on the cell polarity from the environment by random reorientations. To improve the realism of the simulations, it can be interesting to let the polarity vector and possibly the active energy bias depend on the model focal adhesions the cell has formed or the physical obstacles it senses around itself. The effect of degradation or remodelling of the ECM by the cell has also not been included, while it can play an important role in especially 3D cell migration. To fully capture the effects of steric hindrance it seems necessary to explicitly model the cell nucleus by defining extra spins in the CPM. Additionally, the crucial benefit of migrating as a cluster in the process of metastasis might also occur within or when entering the bloodstream instead of in the ECM. This would require an extension of the CPM to an inertial system or the use of a different model. Finally, we mention that the CPM is a rather coarse-grained model on a discrete lattice, hence it is never fully accurate. This means that when very detailed descriptions of the cell are required, it might not be the most ideal model.

Bibliography

[1] "Causes of Death." [Online]. Available: https://ourworldindata.org/causes-of-death

[2] "The Global Burden of Disease." [Online]. Available: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf?ua=1

[3] S. H. Au et al., Clusters of circulating tumor cells traverse capillary-sized vessels, *PNAS*, 113(18), 4947-4952, 2016.

[4] Y. Hong, F. Fang, Q. Zhang, Circulating tumor cell clusters: What we know and what we expect (Review), *International Journal of Oncology*, 49, 2206-2216, 2016.

[5] J. Massagué, A.C. Obenauf, Metastatic colonization by circulating tumour cells, *Nature*, 529, 298-306, 2016.

[6] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, *CA Cancer J. Clin.*, 65, 5-29, 2015.

[7] K.J. Cheung, A.J. Ewald, A collective route to metastasis: Seeding by tumor cell clusters, *Science*, 352, 167-169, 2016.

[8] M. Giuliano et al., Perspective on Circulating Tumor Cell Clusters: Why It Takes a Village to Metastasize, *Cancer Research*, 78(4), 845-852.

[9] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: The next generation, *Cell*, 144(5), 646–674, 2011.

[10] M.W. Klymkowsky, P. Savagner, Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe, *Am. J. Pathol.*, 174, 1588–1593, 2009.

[11] G. Greenburg, E. Hay, Cytodifferentiation and tissue phenotypechange during transformation of embryonic lens epithelium to mesenchyme-like cells in vitro, *Dev. Biol.*, 115, 363–379, 1986.

[12] K. Polyak, R.A. Weinberg, Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits, *Nat. Rev. Cancer*, 9, 265–273, 2009.

[13] A. Singh, J. Settleman, EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer, *Oncogene* 29, 4741–4751, 2010.

[14] N. Aceto et al., Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis, *Cell*, 158, 1110-1122, 2014.

[15] R. Maddipati, B.Z. Stanger, Pancreatic Cancer Metastases Harbor Evidence of Polyclonality, *Cancer Discov.*, 5(10), 1086-1097, 2015.

[16] K.J. Cheung et al., Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters, *PNAS*, 113(7), 854-863, 2016.

[17] E. Lang et al., Coordinated collective migration and asymmetric cell division in confluent human keratinocytes without wounding, *Nature Communications*, 9, 1-15, 2018.

[18] R. Sunyer et al., Collective cell durotaxis emerges from long-range intercellular force transmission, *Science*, 353, 1157-1161, 2016.

[19] D. T. Tambe et al., Collective cell guidance by cooperative intercellular forces, *Nature Materials*, 10, 469-475, 2011.

[20] L. A. Lautscham et al., Migration in Confined 3D Environments Is Determined by a Combination of Adhesiveness, Nuclear Volume, Contractility, and Cell Stiffness, *Biophysical Journal*, 109, 900-913, 2015.

[21] C. Metzner et al., Superstatistical analysis and modelling of heterogeneous random walks, *Nature Communications*, 6, 1-8, 2015.

[22] C.S. Chen, J.L. Alonso, E. Ostuni, G.M. Whitesides, D.E. Ingber, Cell shape provides global control of focal adhesion assembly, *Biochemical and Biophysical Research Communications*, 307, 355-361, 2003.

[23] S.K. Sastry, K. Burridge, Focal Adhesions: A Nexus for Intracellular Signaling and Cytoskeletal Dynamics, *Experimental Cell Research*, 261, 25-36, 2000.

[24] E.A. Novikova, M. Raab, D.E. Discher, C. Storm, Persistence-driven durotaxis: Generic, directed motility in rigidity gradients, *Phys. Rev. Lett.*, 118(7), 1-5, 2017.

[25] M. Raab et al., Crawling from soft to stiff matrix polarizes the cytoskeleton and phosphoregulates myosin-II heavy chain, *J. Cell Biol.*, 199(4), 669-683, 2012.

[26] D. Missirlis, J.P. Spatz, Combined effects of PEG hydrogel elasticity and cell-adhesive coating on fibroblast adhesion and persistent migration, *Biomacromolecules*, 15(1), 195-205, 2014.

[27] D. Campos, V. Mendez, I. Llopis, Persistent random motion: Uncovering cell migration dynamics, *Journal of Theoretical Biology*, 267, 526-534, 2010.

[28] C. Bechinger et al., Active particles in complex and crowded environments, *Reviews of Modern Physics*, 88, 1-50, 2016.

[29] A. Zöttl, H. Stark, Emergent behavior in active colloids, *J. Phys.: Condens. Matter*, 28(25), 1-28, 2016.

[30] C.W. Gardiner, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, third ed. Springer, Berlin, 2004.

[31] M. Doi, *Soft Matter Physics*. Oxford University Press, New York, 2013.

[32] P. Grossmann, F. Peruani, M. Bär, A geometric approach to self-propelled motion in isotropic and anisotropic environments, *Eur. Phys. J. Special Topics*, 224, 1377-1394, 2015.

[33] K.F. Riley, M.P. Hobson, *Essential Mathematical Methods for the Physical Sciences*, Cambridge University Press, Cambridge, 2014.

[34] F. Peruani, A. Deutsch, M. Bär, A mean field theory for self-propelled particles interacting by velocity alignment mechanisms, *Eur. Phys. J. Special Topics*, 157(1), 111-122, 2008.

[35] F. Peruani, F. Ginelli, M. Bär, H. Chaté, Polar vs. apolar alignment in systems of polar self-propelled particles, *J. Phys.: Conf. Ser.*, 297(1), 1-8, 2013.

[36] F. Graner, J. A. Glazier, Simulation of biological cell sorting using a two-dimensional extended Potts model, *Phys. Rev. Lett.*, 69, 2013–2016, 1992.

[37] J. A. Glazier, F. Graner, Simulation of the differential adhesion driven rearrangement of biological cells, *Phys. Rev. E*, 47, 2128–2154, 1993.

[38] M.S. Steinberg, Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation, *Science*, 141, 401-408, 1963.

[39] C. Giverso, M. Scianna, L. Preziosi, N. Lo Buono and A. Funaro, Individual Cell-Based Model for In-Vitro Mesothelial Invasion of Ovarian Cancer, *Math. Model. Nat. Phenom.*, 5(1), 203-223, 2010.

[40] Merks, R. M. H., Brodsky, S. V., Goligorksy, M. S., Newman, S. A., and Glazier, J. A., Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Dev. Biol.* 289, 44–54, 2006.

[41] A. Szabo et al., Collective cell motion in endothelial monolayers, *Phys. Biol.* 7(4), 1-27, 2011.

[42] R.B. Potts, C. Domb, Some generalized order-disorder transformations, *Proceedings of the Cambridge Philosophical Society*, 48(1), 106-109, 1952.

[43] M. Durand, E. Guesnet, An efficient Cellular Potts Model algorithm that forbids cell fragmentation, *Computer Physics Communications*, 208, 54-63, 2016.

[44] A. Szabo, R.M.H. Merks, Cellular Potts modeling of tumor growth, tumor invasion, and tumor evolution, *Frontiers in Oncology*, 3, 1-12, 2013.

[45] A. Szabo et al., Invasion from a cell aggregate – the roles of active cell motion and mechanical equilibrium, *Phys. Biol.*, 9(1), 1-30, 2012.

[46] N.J. Savill, P. Hogeweg, Modelling morphogenesis: from single cells to crawling slugs, *J. Theor. Biol.* 184, 229–235, 1997.

[47] B.M. Rubenstein, L.J. Kaufman, The role of extra cellular matrix in glioma invasion: a cellular Potts model approach. *Biophys. J.* 95, 5661–5680, 2008.

[48] A.L. Bauer, T.L. Jackson, Y. Jiang, A cell-based model exhibiting branching and anastomosis during tumor-induced angiogenesis, *Biophys. J.* 92, 3105–3121, 2007.

[49] N. Guisoni, K.I. Mazzitello, L. Diambra, Modeling active cell movement with the Potts model, *Frontiers in Physics*, 6, 1-11, 2018.

[50] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, Molecular Biology of the Cell, Garland Science, New York, 2003.

[51] L. Bosgraaf, P.J.M. Van Haastert, The Ordered Extension of Pseudopodia by Amoeboid Cells in the Absence of External Cues, *PLoS ONE*, 4(4), 1-13, 2009.

[52] S.K. Schnyder, J.J. Molina, Y. Tanaka, R. Yamamoto, Collective motion of cells crawling on a substrate: roles of cell shape and contact inhibition, *Scientific Reports*, 7, 1-14, 2017.

[53] M. Abercrombie, Contact inhibition and malignancy, *Nature*, 281, 259-262, 1979.

[54] C. Carmona-Fontaine et al., Contact inhibition of locomotion in vivo controls neural crest directional migration, *Nature*, 456, 957-961, 2008.

[55] V. Maruthamuthu, B. Sabass, U.S. Schwarz, M.L. Gardel, Cell-ECM traction force modulates endogenous tension at cell–cell contacts, *PNAS*, 108(12), 4708-4713, 2011.

[56] J. Escribano, A hybrid computational model for collective cell durotaxis, *Biomechanics and Modeling in Mechanobiology*, 17, 1037-1052, 2018.

[57] L. Liang, S.F. Norrelykke, E.C. Cox, Persistent Cell Motion in the Absence of External Signals: A Search Strategy for Eukaryotic Cells, *PLoS ONE*, 3(5), 1-11, 2008.

[58] P.H. Wu, A. Giri, S.X. Sun, D. Wirtz, Three-dimensional cell migration does not follow a random walk, *PNAS*, 111(11), 3949-3954, 2014.

[59] J. Xu et al., Polarity reveals intrinsic cell chirality, *PNAS*, 104(22), 9296-9300, 2007.

[60] T. Vicsek et al., Novel Type of Phase Transition in a System of Self-Driven Particles, *Physical Review Letters*, 75(6), 1226-1229, 1995.

[61] E. Méhes, T. Vicsek, Collective motion of cells: from experiments to models, *Integr. Biol.*, 6, 831-854, 2014.

[62] Y.E. Sanchez-Corrales et al., Morphometrics of complex cell shapes: lobe contribution elliptic Fourier analysis (LOCO-EFA), *Development*, 145, 1-13, 2018.

[63] U. Horzum, B. Ozdil, D. Pesen-Okvur, Step-by-step quantitative analysis of focal adhesions, *MethodsX*, 1, 56-59, 2014.

[64] N.Q. Balaban et al., Force and focal adhesion assembly: a close relationship studied using elastic micropatterned substrates, *Nature Cell Biology*, 3, 466-472, 2001.

[65] H. Yuan et al, One-particle-thick, solvent-free, coarse-grained model for biological and biomimetic fluid membranes, *Physical Review E*, 82, 1-8, 2010.

[66] L. Zhang, K. Yu and A. Eisenberg, Ion-induced Morphological Changes in "Crew-Cut" Aggregates of Amphiphilic Block Copolymers, *Science*, 272, 1777-1779, 1996.

[67] H. Yuan, C. Huang and S. Zhang, Dynamic Shape Transformations of Fluid Vesicles, *Soft Matter*, 6, 4571-4579, 2010.

[68] User documentation LAMMPS, "pair style lj/sf command", [Online]. Available: http://lammps.sandia.gov/doc/pair_lj_expand.html

[69] I. Huopaniemi, K. Luo, T. Ala-Nissila and S. Ying, Langevin Dynamics Simulations of Polymer Translocation through Nanopores, *Journal of Chemical Physics*, 125(12), 1-8, 2006.

[70] N. Attig, K. Binder, H. Grubmueller and K. Kremer, Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells, *Computational Soft Matter: From Synthetic Polymers to Proteins*, 23, 1-28, 2004.

[71] T. Schneider and E. Stoll, Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions, *Physical Review B*, 17(3), 1302-1322, 1978.

[72] "LAMMPS." [Online]. Available: http://lammps.sandia.gov/

[73] D. Frenkel and B. Smit, *Understanding Molecular Simulation from Algorithms to Applications*. San Diego: Academic Press, 1996.

[74] S. Melchionna, G. Ciccotti and B.L. Holian, Hoover NPT dynamics for systems varying in shape and size, *Molecular Physics*, 78(3), 533-544, 1993.

[75] Y. Golan, E. Sherman, Resolving mixed mechanisms of protein subdiffusion at the T cell plasma membrane, *Nature Communications*, 8, 1-15, 2016.

[76] D. Chandler, *Introduction to Modern Statistical Mechanics*. Oxford University Press, New York, 1987.

[77] H.G. Othmer, S.R. Dunbar, W. Alt, Models of dispersal in biological systems, *J. Math. Biol.*, 26(3), 263-298, 1988.

[78] S.C. Weber et al., Analytical Tools to Distinguish the Effects of Localization Error, Confinement, and Medium Elasticity on the Velocity Autocorrelation Function, *Biophysical Journal*, 102, 2443-2450, 2012.

[79] M. Scianna, L. Preziosi, K. Wolf, A Cellular Potts Model Simulating Cell Migration on and in Matrix Environments, *Mathematical Bioscience and Engineering*, 10(1), 235-261, 2013.

[80] G.E. Uhlenbeck, L.S. Ornstein, On the theory of the Brownian motion, *Phys. Rev.*, 36, 823–841.

[81] U. Erdmann, W. Ebeling, L. Schimansky-Geier, F. Schweitzer, Brownian particles far from equilibrium, *Eur. Phys. J. B*, 15, 105–113, 2000.

[82] J. Brábek, C.T. Mierke, D. Rösel, P. Vesely, B. Fabry, The role of the tissue microenvironment in the regulation of cancer cell motility and invasion, *Cell Commun. Signal*, 8, 1-8, 2010.

[83] K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren, D.E. Engelman, Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins ratherthan cholesterol, *PNAS*, 101(12), 4083-4088, 2004.

[84] A. O. Brightman, B. P. Rajwa, J. E. Sturgis, M. E. McCallister, J. P. Robinson, S.L. Voytik-Harbin, Time-lapse confocal reection microscopy of collagen fibrillogenesis and extracellular matrix assembly in vitro, *Biopolymers*, 54, 222-234, 2000.

[85] C. B. Raub, V. Suresh, T. Krasieva, J. Lyubovitsky, J. D. Mih, A. J. Putnam, B. J. Tromberg, S. C. George, Noninvasive assessment of collagen gel microstructure and mechanics using multiphoton microscopy, *Biophys. J.*, 92, 2212-2222, 2007.

[86] B. A. Roeder, K. Kokini, J. E. Sturgis, J. P. Robinson, S. L. Voytik-Harbin, Tensile mechanical properties of three-dimensional type I collagen extracellular matrices with varied microstructure, *J. Biomech. Eng.*, 124, 214-222, 2002.

[87] E.A. Novikova, C. Storm, Contractile Fibers and Catch-Bond Clusters: a Biological Force Sensor?, *Biophysical Journal*, 105, 1336-1345, 2013.

[88] L.B. Smilenov, A. Mikhailov, R.J. Pelham, E.E. Mercantonio, G.G. Gundersen, Focal adhesion motility revealed in stationary fibroblasts, *Science*, 286, 1172–1174, 1999.

A. Derivations Chapter 2

In this appendix we present several extended derivations of results that have been used throughout chapter 2.

A.1 2D Mean Square Displacement

In order to derive the mean square displacement (MSD) for a 2D active Brownian particle travelling at a speed v_0 we start with the Fourier-Laplace transform of the corresponding Fokker-Planck equation introduced in section 2.1

$$\frac{1}{\tau}\frac{\partial^2 \hat{P}}{\partial \phi^2} = \left[s + iv_0 \left(q_x \cos(\phi) + q_y \sin(\phi) \right) + D(q_x^2 + q_y^2) \right] \hat{P} - \frac{1}{2\pi}. \tag{A.1}$$

Setting the wavevectors $q_x = q_y = 0$ (denoted by $_0$) results in

$$\frac{1}{\tau}\frac{\partial^2 \hat{P}_0}{\partial \phi^2} = s\hat{P}_0 - \frac{1}{2\pi}. \tag{A.2}$$

which can be solved to give

$$\hat{P}_0 = C_1 \exp(\sqrt{\tau s}\phi) + C_2 \exp(-\sqrt{\tau s}\phi) + \frac{1}{2\pi s}, \tag{A.3}$$

where the exponential terms represent the homogeneous solution with C_1 and C_2 integration constants. These constants can be determined by realising that a Fourier transform of the initial condition $(P(\mathbf{r}, \phi, 0) = \delta(\mathbf{r})/2\pi)$ results in $P(\mathbf{q}, \phi, 0) = 1/2\pi$ which is independent of the transformation variables $\mathbf{q} = [q_x, q_y]$. Using the initial value theorem, i.e. $\lim_{s \to \infty} sF(s) = f(0)$ for the Laplace transform $F(s)$ of an arbitrary function $f(t)$, we can formulate the condition [33]

$$\lim_{s \to \infty} s\hat{P}_0 = P(\mathbf{q} = 0, \phi, 0) = 1/2\pi, \tag{A.4}$$

from which we derive that $C_1 = C_2 = 0$. Thus, we discard the homogeneous solution and find that

$$\hat{P}_0 = \frac{1}{2\pi s}. \tag{A.5}$$

Next we can take the derivative with respect to q_x of eq. (A.1), insert eq. (A.5) and evaluate at $q_x = q_y = 0$ to find

$$\frac{1}{\tau}\frac{\partial^2 \hat{P}_{x,0}}{\partial \phi^2} = s\hat{P}_{x,0} + \frac{iv_0 \cos(\phi)}{2\pi s}. \tag{A.6}$$

(For convenience we adopt the notation $\hat{P}_{\alpha,0} \equiv \left(\frac{\partial \hat{P}}{\partial q_\alpha}\right)_0$ and $\hat{P}_{\alpha\alpha,0} \equiv \left(\frac{\partial^2 \hat{P}}{\partial q_\alpha^2}\right)_0$ with $\alpha = x, y$.) This equation can be solved to give

$$\hat{P}_{x,0} = \frac{-iv_0\tau \cos(\phi)}{2\pi s(1 + \tau s)}. \tag{A.7}$$

where again we have discarded the homogeneous solution to the differential equation (eq. (A.6)) due to the condition $\lim_{s\to\infty} s\hat{P}_{x,0} = 0$ which results from taking the derivative of eq. (A.4) with respect to q_x before evaluating at $q_x = q_y = 0$.

If we then take the second derivative with respect to q_x of eq. (A.1), insert both the results of eq. (A.5) and eq. (A.7) and evaluate at $q_x = q_y = 0$ we find

$$\frac{1}{\tau}\frac{\partial^2 \hat{P}_{xx,0}}{\partial \phi^2} = s\hat{P}_{xx,0} + \frac{v_0^2\tau \cos^2(\phi)}{\pi s(1 + s\tau)} + \frac{D\tau}{\pi s}, \tag{A.8}$$

which can be solved to give

$$\hat{P}_{xx,0} = -\frac{v_0^2\tau^2 \cos(2\phi)}{2\pi s(1 + s\tau)(4 + s\tau)} - \frac{v_0^2\tau}{2\pi s^2(1 + s\tau)} - \frac{D}{\pi s^2}. \tag{A.9}$$

Once more we have discarded the homogeneous solution as a result of the condition $\lim_{s\to\infty} s\hat{P}_{xx,0} = 0$ which follows from taking the second derivative with respect to q_x of eq. (A.4) before evaluating at $q_x = q_y = 0$.

A similar procedure can then be applied with respect to q_y instead of q_x resulting in

$$\hat{P}_{y,0} = \frac{-iv_0\tau \sin(\phi)}{2\pi s(1 + \tau s)}, \tag{A.10}$$

which can be used to find

$$\hat{P}_{yy,0} = \frac{v_0^2\tau^2 \cos(2\phi)}{2\pi s(1 + s\tau)(4 + s\tau)} - \frac{v_0^2\tau}{2\pi s^2(1 + s\tau)} - \frac{D}{\pi s^2}. \tag{A.11}$$

Finally, we can insert the retrieved expressions for $\hat{P}_{xx,0}$ and $\hat{P}_{yy,0}$ in eq. (2.8) which yields

$$\langle \mathbf{r}^2 \rangle (s) = -\int_0^{2\pi} \left(\hat{P}_{xx,0} + \hat{P}_{yy,0}\right) d\phi = \int_0^{2\pi} \left(\frac{v_0^2\tau}{\pi s^2(1 + s\tau)} - \frac{2D}{\pi s^2}\right) d\phi = \frac{2v_0^2\tau}{s^2(1 + s\tau)} + \frac{4D}{s^2}. \tag{A.12}$$

This result coincides with the one stated in eq. (2.9).

A.2   3D Probability Density Function

We intend to derive a solution for the marginal probability density function $\tilde{P}(\theta, t)$ of a point particle performing active Brownian motion with a speed v_0 in three dimensions. Starting with its Fokker-Planck equation given by eq. (2.14) we have

$$\frac{\partial \tilde{P}(\theta, t)}{\partial t} = -\frac{\partial}{\partial \theta}\left(\frac{\tilde{P}}{2\tau \tan\theta}\right) + \frac{1}{2\tau}\frac{\partial^2 \tilde{P}}{\partial \theta^2}. \tag{A.13}$$

Applying separation of variables by defining $\tilde{P}(\theta, t) = \Theta(\theta)T(t)$ and rearranging terms we can rewrite eq. (A.13) as

$$\frac{2\tau}{T(t)}\frac{dT(t)}{dt} = \frac{d^2\Theta}{d\theta^2} - \frac{1}{\tan(\theta)}\frac{d\Theta}{d\theta} + \frac{\Theta(\theta)}{\sin^2(\theta)}. \tag{A.14}$$

Since the left and right side of this equation are functions of only t or θ respectively, they can only be equal if both of them are constant. Denoting this constant $-l(l+1)$ yields a set of two ordinary differential equations:

$$\frac{dT(t)}{dt} + \frac{l(l+1)T(t)}{2\tau} = 0$$

$$\tag{A.15}$$

$$\frac{d^2\Theta}{d\theta^2} - \frac{1}{\tan(\theta)}\frac{d\Theta}{d\theta} + \left(\frac{1}{\sin^2(\theta)} + l(l+1)\right)\Theta(\theta) = 0$$

The top equation has a straightforward general solution given by

$$T(t) = C\exp\left(-\frac{l(l+1)}{2\tau}t\right), \tag{A.16}$$

with C a constant. Also, for the solution to remain finite at $t \to \infty$ we require $l \geq 0$.

For the second differential equation involving $\Theta(\theta)$ we substitute the variable $u \equiv \cos(\theta)$. This simplifies the equation to the following form

$$(1-u^2)\frac{d^2\Theta}{du^2} + \left(l(l+1) + \frac{1}{1-u^2}\right)\Theta(u). \tag{A.17}$$

Dividing the equation by $(1-u^2)^{1/2}$ and rewriting it in terms of the function $R(u) \equiv \Theta(u)/(1-u^2)^{1/2}$ finally results in

$$(1-u^2)\frac{d^2R}{du^2} - 2u\frac{dR}{du} + l(l+1)R(u) = 0, \tag{A.18}$$

in which we can recognise the Legendre differential equation [33]. The solution to this equation is given by

$$R_l(u) = a_l P_l(u) \qquad \text{for } l \in \mathbb{N}_0, \tag{A.19}$$

with $P_l(u)$ a Legendre polynomial of the first kind of degree l [33]. For a more extensive description of these functions and their properties the reader is for instance referred to [33]. Also, note that we already dismissed Legendre polynomials of the second kind (often denoted $Q_l(u)$) since these functions diverge for $u = \pm 1$ ($\theta = 0, \pi$). Converting back to our original angular function we find

$$\Theta_l(\theta) = a_l \sin(\theta) P_l(\cos(\theta)), \tag{A.20}$$

where we used that $(1-u^2)^{1/2} = \sin(\theta)$.

If we then combine both results, i.e. eq. (A.16) and eq. (A.20), and make use of the fact that Legendre polynomials form a complete set we can formulate a general solution for $\tilde{P}(\theta, t)$:

$$\tilde{P}(\theta, t) = \sin(\theta)\sum_{l=0}^{\infty} a_l P_l(\cos(\theta))\exp\left(-\frac{l(l+1)}{2\tau}t\right). \tag{A.21}$$

Here we absorbed the constant C from eq. (A.16) into the coefficients a_l from the Legendre polynomials. Letting the particle start in the z-direction at $t = 0$ we have

$$\tilde{P}(\theta, 0) = \sin(\theta) \sum_{l=0}^{\infty} a_l P_l(\cos(\theta)) = \delta(\theta), \tag{A.22}$$

with $\delta(\theta)$ the Dirac delta function. Mulitplying eq. (A.22) with $P_{l'}(\cos(\theta))$ and afterwards integrating both sides over the θ domain $[0, \pi)$ then results in

$$\frac{2a_{l'}}{2l' + 1} = 1, \tag{A.23}$$

where we have used that $P_l(1) = 1$ and the fact that Legendre polynomials are orthogonal so that $\int_{-1}^{1} P_l(x) P_{l'}(x) dx = 2\delta_{ll'}/(2l + 1)$ with $\delta_{ll'}$ the Kronecker delta [33]. Thus, we find that $a_l = (2l + 1)/2$. Inserting this in eq. (A.21) yields

$$\tilde{P}(\theta, t) = \frac{\sin(\theta)}{2} \sum_{l=0}^{\infty} (2l + 1) P_l(\cos(\theta)) \exp\left(-\frac{l(l+1)}{2\tau}t\right). \tag{A.24}$$

This result coincides with the one shown in eq. (2.15).

A.3 Cluster Persistence Time

We seek to prove that the persistence time of a cluster of N fast aligned active particles, i.e. $\tau_{cm} = N^2 \left(\sum_{i=1}^{N} \frac{1}{\tau_i}\right)^{-1}$, never exceeds the value it would obtain when each particle's persistence time equals the average persistence time of all of them. Specifically, we have to prove that for a set of N persistence times $\{\tau_i > 0\}$ the following relation always holds

$$N^2 \left(\sum_{i=1}^{N} \frac{1}{\tau_i}\right)^{-1} \leq \sum_{i=1}^{N} \tau_i. \tag{A.25}$$

Multiplying both sides with $\sum_{i=1}^{N} \frac{1}{\tau_i}$, which is a positive number, and reversing both sides we find

$$\left(\sum_{i=1}^{N} \tau_i\right)\left(\sum_{i=1}^{N} \frac{1}{\tau_i}\right) \geq N^2. \tag{A.26}$$

The product of two sums on the left hand side can now be rewritten to give (for $N \geq 2$)

$$\sum_{i=1}^{N} \frac{\tau_i}{\tau_i} + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\frac{\tau_i}{\tau_j} + \frac{\tau_j}{\tau_i}\right) \geq N^2, \tag{A.27}$$

Realising that the first term on the left hand side is simply N and taking it to the right hand side we have

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left(\frac{\tau_i}{\tau_j} + \frac{\tau_j}{\tau_i}\right) \geq N(N - 1). \tag{A.28}$$

Now we can notice that if $\frac{\tau_i}{\tau_j} + \frac{\tau_j}{\tau_i} = 2$ for all combinations i, j, the sum in eq. (A.28) will give $N(N - 1)$ corresponding to the equality sign. Thus, we can demonstrate that the

sum in eq. (A.28) is greater than or equal to $N(N-1)$ by showing that, independent of their respective values, each pair $\tau_{i,j}$ obeys

$$\frac{\tau_i}{\tau_j} + \frac{\tau_j}{\tau_i} \geq 2. \tag{A.29}$$

The above relation can be rewritten as

$$(\tau_i - \tau_j)^2 \geq 0, \tag{A.30}$$

and since τ_i is positive and real this will always be true. Thereby we have proved our original relation shown in eq. (A.25) and eq. (2.43).

B. LAMMPS Script

```
units           lj
atom_style      hybrid sphere dipole

dimension       3
boundary        p p p
processors  * * 1   # set the number of MPI processors in Lz direction to 1

read_data       "data_10.in"

group       mem1    type 1  # receptors fluid membrane
group       mem2    type 2  # inert beads fluid membrane
group   memall    type 1:2  # all membrane

set         group     mem1      dipole 0 0 1
set         group     mem2      dipole 0 0 1


# membrane particle parameters
variable        rc_global    equal    2.6
variable        rc           equal    2.6
variable        rmin         equal    1.12
variable        mu           equal    3.0
variable        zeta         equal    4.0
variable        eps          equal    4.34
variable        sigma        equal    1.00

variable        theta0_11    equal    0.0 #No spontaneous curvature
variable        theta0_22    equal    0.0
variable        theta0_12    equal    0.0

# Cylinder parameters (zcyl > r_actualcylinder = 10)
# The second cylinder describes volume exclusion via LJ126 potential,
# while the first cylinder describes an attraction via a Morse potential

variable rcyl2      equal    9.5 #Ensures cylinder has r=10
variable alpha      equal    1.0 #Width of the attractive potential
variable rcyl1      equal    9.5 #Ensures cylinder has r=10
variable zcyl       equal    14  #Membrane particles are 5
#sigma below the cylinder in the start configuration
```

```
variable r_attract    equal    4.5 #Range of the attractive Morse potential


# Tether parameters
variable  kmem equal 50  #Spring constant used to keep membrane at a fixed height

variable  rsp  equal    0    #Equilibrium distance spring

# only membrane gets velocity, start with Temp=3 to offset the potential
# energy when relaxed from hex lattice
variable        Nmem1        equal    "count(mem1)"
variable        Nmem2        equal    "count(mem2)"
if "${Nmem1} > 1" then "velocity    mem1 create   3.0  1 rot yes mom yes "
if "${Nmem2} > 1" then "velocity    mem2 create   3.0  1 rot yes mom yes "

# Use hybrid overlay for pair_style membrane and lj/expand
#pair_style      hybrid/overlay  membrane      ${rc_global} morse ${rc_global}
pair_style hybrid membrane ${rc_global}

pair_modify shift yes

# Initialise Membrane interactions to zero for all possible combinations
# of types, 1-1, 2-1, 2-2. * is just a shortcut for that
pair_coeff *       *         membrane    0.0   0.0   0.0   0.0   0.0   0.0   0.0
pair_coeff 1       1         membrane    ${eps}        ${sigma}       ${rmin}
${rc}     ${zeta}  ${mu}  ${theta0_11}
pair_coeff 2       2         membrane    ${eps}        ${sigma}       ${rmin}
${rc}     ${zeta}  ${mu}  ${theta0_22}
pair_coeff 1       2         membrane    ${eps}        ${sigma}       ${rmin}
${rc}     ${zeta}  ${mu}  ${theta0_12}

# Reduce the delay from default 10 to 2 to get rid of dangerous builds
neigh_modify    delay 1
neigh_modify    page 200000 one 20000

variable        dofsub       equal "count(mem1) + count(mem2)"
variable        dofsubrot    equal "count(memall)"
compute         cT           all    temp/sphere
compute_modify cT            extra  ${dofsub}
compute         cP           all    pressure cT
compute         cTrot    memall    temp/sphere dof rotate
compute_modify cTrot     extra  ${dofsubrot}

############## INTEGRATORS ############
#Define cylinder where the membrane interacts with
#(different potentials for receptors and nonreceptors so two walls)

region cyl1 cylinder x 0 ${zcyl} ${rcyl1} -35 35 side out

region cyl2 cylinder x 0 ${zcyl} ${rcyl2} -35 35 side out
```

```
fix wall1      mem1     wall/region  cyl1 morse 15.0 ${rmin} ${r_attract}
${alpha} #Morse potential with minimum at the same location as LJ126

fix wall2      memall  wall/region  cyl2  lj126 4.0 ${sigma} ${rmin}

#Define spring that keeps the membrane at a fixed height so that
#it does not encapsulate the cylinder

fix mheight     all      spring  tether  ${kmem}  NULL NULL 0 ${rsp}

#Integrate in the NPH ensemble and regulate temperature via a Langevin thermostat

fix fLANG      all      langevin 1.0 1.0 1 12341 zero yes

fix fNPH       all      nph/sphere  x 0.0 0.0 10 y 0.0 0.0 10 &
couple xy  update dipole
fix_modify fNPH    press cP

#fix_modify fNPH    press thermo_press
#fix fNVE      all nve/sphere update dipole
#fix          fCENTER membrane recenter INIT INIT INIT

#compute        avpe    all pe
#fix           fAVPE all ave/time 10 5000 50000 c_avpe ave one file avPe.avpe

# dump coordinates to file trj.dat. You can simply use vmd -lammpstrj trj.dat
#dump          fDump all custom 10000 "trj.dat" id type x y z
#dump cDump mem1 custom 1000 "cFolTrj.dat" x y z ix iy iz
#dump cDump2 mem1 custom 1000 "cFolTrj2.dat" x y z
dump   cUnfDump mem1 custom 1000 "cUnfTrj1.dat" xu yu zu
dump cUnfDump2 mem2 custom 1000 "cUnfTrj2.dat" xu yu zu
dump xyzDump all xyz 1000 output.xyz
#dump_modify    fDump sort id
thermo_style    custom step c_cT c_cTrot pe ke etotal xhi press c_cP

timestep       0.01
thermo         1000
run            2000000

write_data "lastconfig.out"
```

C. Cellular Potts Model Definitions

In the explanation of the standard cellular Potts model (CPM) and its extensions in section 3.1, we make use of several cell properties. These properties are fairly intuitive and are not explicitly defined throughout the section. In this appendix we will provide a formal definition of all relevant cell (cluster) properties. We repeat that cells are represented as domains of equal spin $\sigma(\mathbf{x}) \in \mathbb{N}$ on a lattice whose sites are characterised by their position $\mathbf{x}$ and that the spin σ labels each cell. The results are listed below.

– The volume or area of cell σ:

$$V_\sigma = \sum_{\mathbf{x}} \delta_{\sigma(\mathbf{x}),\sigma}, \tag{C.1}$$

where $\delta_{i,j}$ denotes the Kronecker delta and the sum is taken over all lattice sites $\mathbf{x}$

– The center of mass of cell σ:

$$\mathbf{R}_\sigma = \frac{1}{V_\sigma} \sum_{\sigma(\mathbf{x})=\sigma} \mathbf{x}, \tag{C.2}$$

where the sum is taken over all lattice sites $\mathbf{x}$ whose spin $\sigma(\mathbf{x})$ are equal to the cell's spin σ.

– Change in the center of mass of the candidate cell during a spin-copy attempt in which its value $\sigma(\mathbf{x})$ is replaced by the target site value $\sigma(\mathbf{x}')$:

$$\Delta \mathbf{R}_{\sigma(\mathbf{x})}(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) = \frac{\mathbf{R}_{\sigma(\mathbf{x})} - \mathbf{x}}{V_{\sigma(\mathbf{x})} - 1}, \tag{C.3}$$

where $\mathbf{R}_{\sigma(\mathbf{x})}$ and $V_{\sigma(\mathbf{x})}$ are the center of mass and volume of the candidate cell before the attempt.

– Change in the center of mass of the target cell during a spin-copy attempt in which its value $\sigma(\mathbf{x}')$ replaces the candidate site value $\sigma(\mathbf{x})$:

$$\Delta \mathbf{R}_{\sigma(\mathbf{x}')}(\sigma(\mathbf{x}) \to \sigma(\mathbf{x}')) = \frac{\mathbf{x} - \mathbf{R}_{\sigma(\mathbf{x}')}}{V_{\sigma(\mathbf{x}')} + 1}, \tag{C.4}$$

where $\mathbf{R}_{\sigma(\mathbf{x}')}$ and $V_{\sigma(\mathbf{x}')}$ are the center of mass and volume of the target cell before the attempt.

– The center of mass of a cluster of cells:

$$\mathbf{R}_c = \frac{1}{\sum_\sigma V_\sigma} \sum_\sigma V_\sigma \mathbf{R}_\sigma, \tag{C.5}$$

where the sum is taken over all cells σ in the cluster.

– The preferred radius of a cell in 2D:

$$r_\sigma = \left(\frac{V_\sigma}{\pi}\right)^{1/2} \tag{C.6}$$

– The preferred radius of a cell in 3D:

$$r_\sigma = \left(\frac{3V_\sigma}{4\pi}\right)^{1/3} \tag{C.7}$$

D. Additional Figures
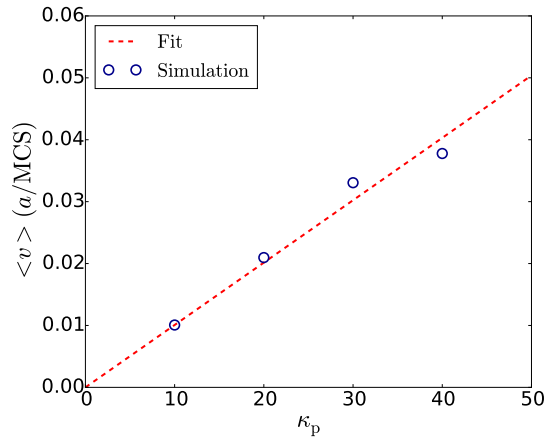
D.1 For Section 5.1

Figure D.1: Plots of the average cell velocity $\langle v \rangle$ over a time period $\delta = 500$ MCS as a function of the strength of the active energy bias κ_p. Results are obtained from a set of 20 3D single cell simulations of 20000 MCS with a constant polarity vector $\mathbf{p}$ and linearly fitted.

(a) $\tau = 500$ MCS, $\kappa_p = 5$

(b) $\tau = 2000$ MCS, $\kappa_p = 5$

(c) $\tau = 500$ MCS, $\kappa_p = 2$

(d) $\tau = 500$ MCS, $\kappa_p = 8$
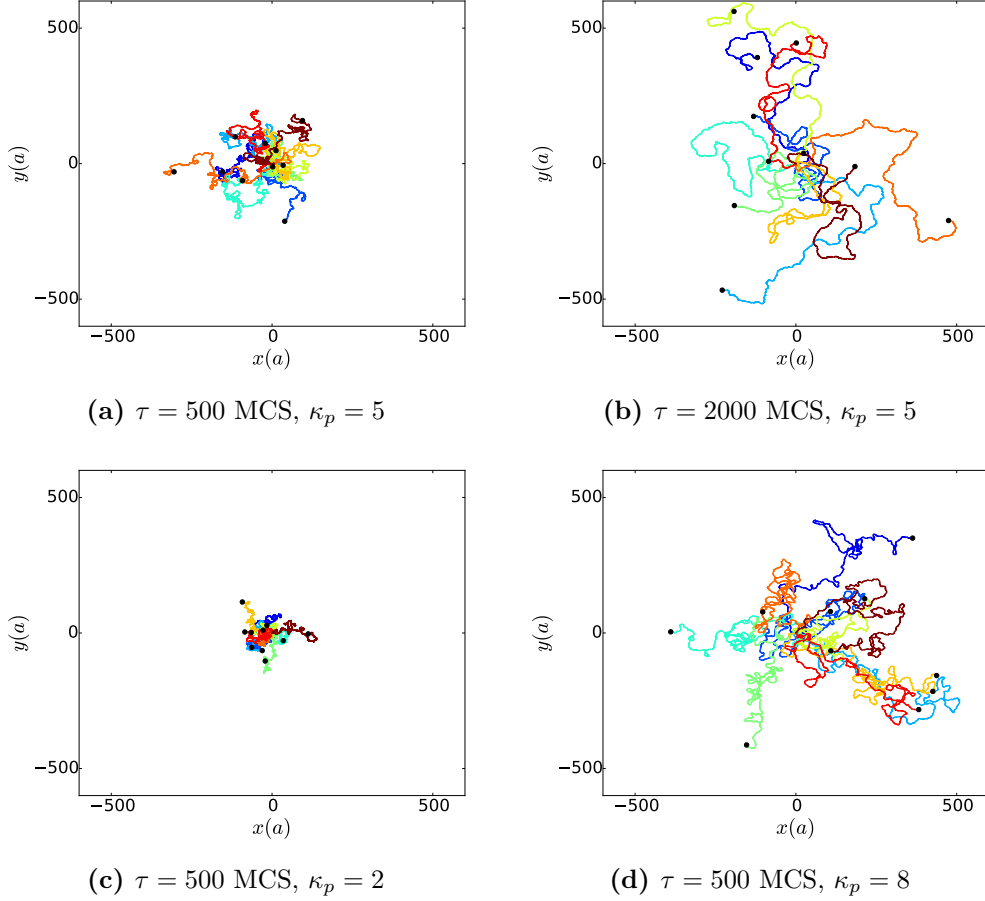
Figure D.2: Simulated trajectories of 10 2D CPM cells for different implemented persistence times τ and active energy bias strengths κ_p. Total simulation time is 50000 MCS and black dots denote the end point of the trajectory. Other CPM parameters are discussed in section 5.1.

D.2   For Section 5.2

(a)

(b)
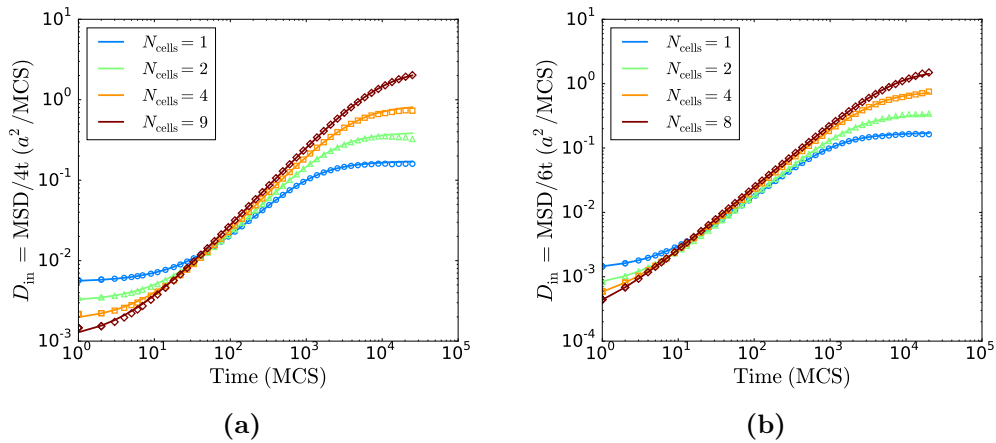
Figure D.3: Plots of the calculated (a) 2D and (b) 3D instant diffusion coefficients $D_{in} = $ MSD/$2dt$ for cell clusters consisting of variable numbers of N_{cells} cells (markers) which have been fitted with a PRW using eq. (2.10) and eq. (2.20) respectively (lines). Results correspond to $\gamma = 50$.
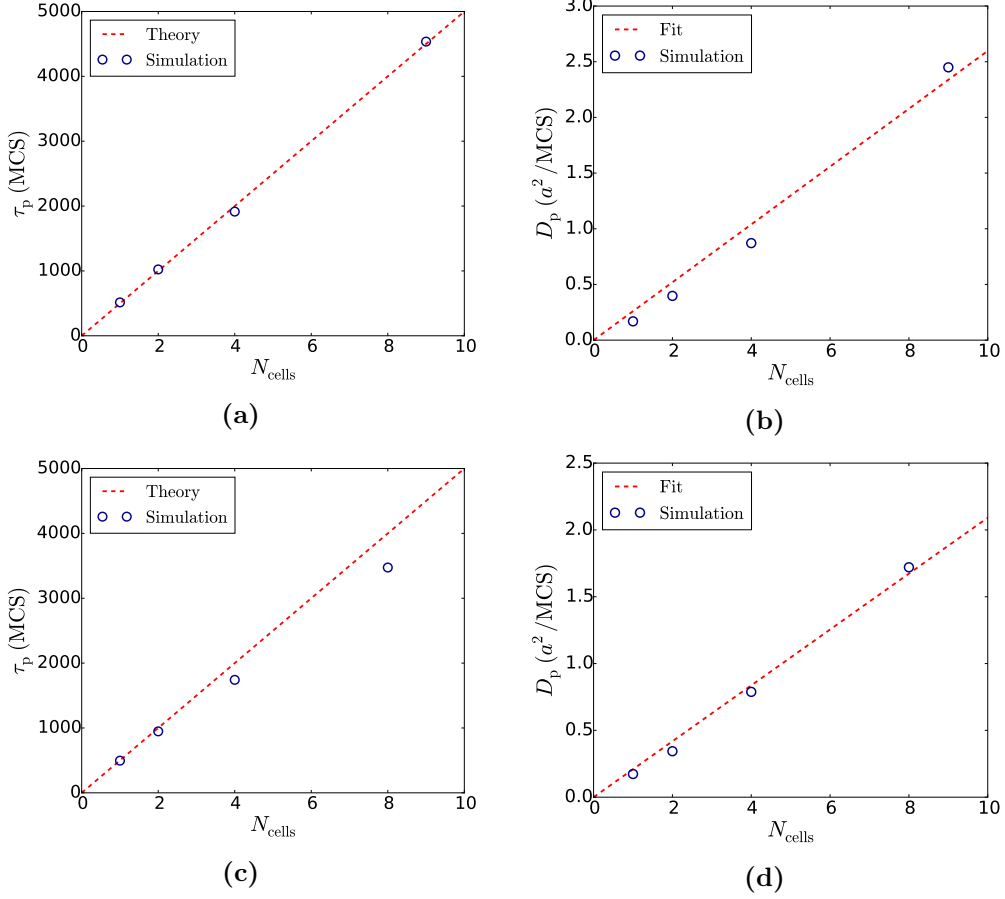
(a)

(b)

(c)

(d)

Figure D.4: Persistence time and active diffusion coefficient D_p for (a-b) 2D and (c-d) 3D motion obtained from fitting calculated MSDs with eq. (2.10) and eq. (2.20) respectively. Values are retrieved for clusters consisting of different number of N_cells cells and correspond to $\gamma = 50$. Moreover, the results are compared to or fitted with the theoretical results of 2D active Brownian motion for fast-aligning particles.
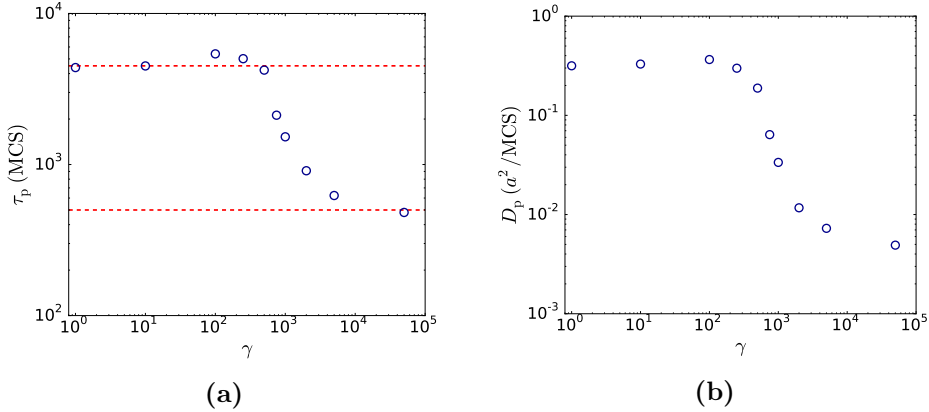

(a)

(b)

Figure D.5: Persistence time and active diffusion coefficient D_p obtained from fitting calculated MSDs with eq. (2.10). Values are retrieved for different alignment parameters γ and correspond to a cluster with $N_\text{cells} = 9$ cells and an individual persistence time $\tau = 500$ MCS. Red dotted lines in (a) serve as visual aide and denote the values $N_\text{cells}\tau$ (top line) and τ (bottom line).
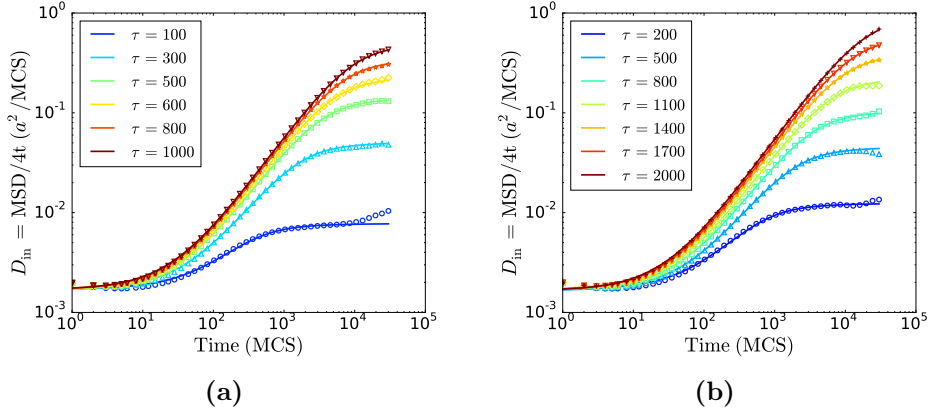
(a)   **(b)**

Figure D.6: Plots of the calculated instant diffusion coefficients $D_{in} = MSD/4t$ for different values of τ (markers) which have been fitted with a PRW using eq. (2.10) (lines). Results correspond to (a) $\gamma = 500$ MCS and (b) $\gamma = 1500$ MCS and a cluster consisting of $N_{cells} = 4$ cells.
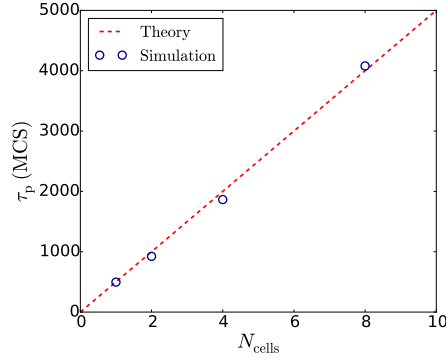

Figure D.7: Plot of the cluster persistence time τ_p obtained from the fits demonstrated in fig. 5.18a.

D.3   For Section 5.3

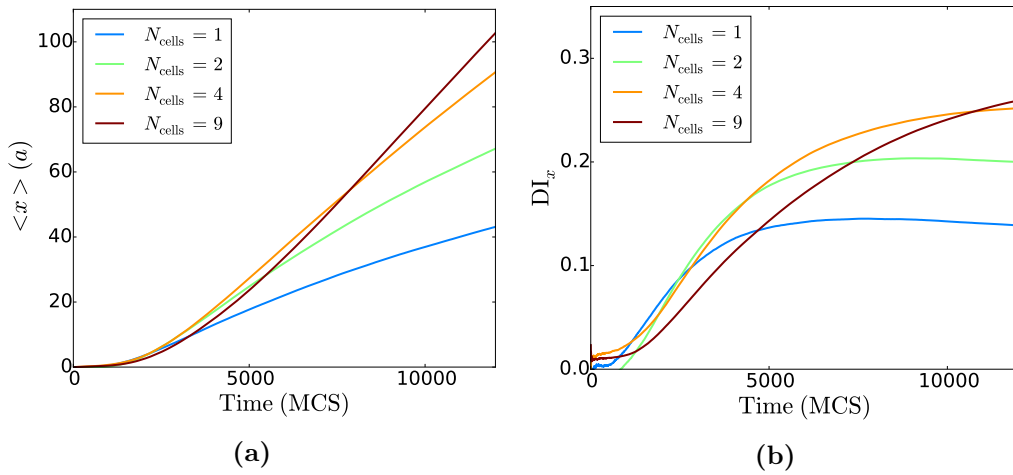**(a)**   **(b)**

Figure D.8: Plots of (a) the average displacement in the x-direction and (b) the x-component of the durotactic index as a function of time for aligned ($\gamma = 1$) cell clusters experiencing a linear gradient of $d\tau/dx = 20$ MCS/a in single cell persistence time from $\tau_{min} = 200$ MCS to $\tau_{max} = 2000$ MCS. Results correspond to different sized clusters.
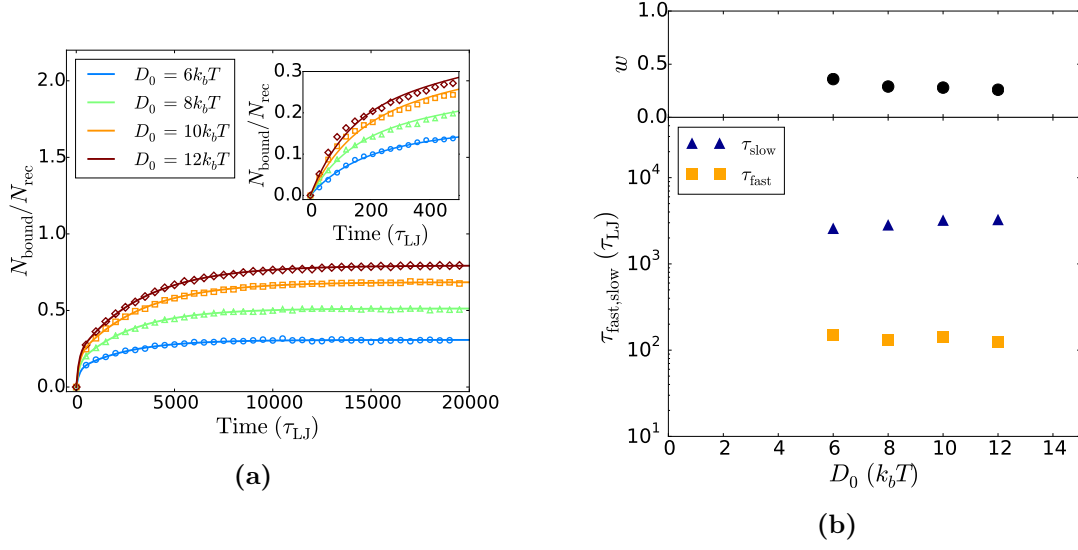
D.4 For Section 5.5

(a)

(b)

Figure D.9: (a) Plot of the average fraction of receptors bound to the cylinder as a function of time obtained from simulations (markers) which have been fitted with eq. (5.2) (lines). Inset zooms in on short time behavior. (b) Plot of the bi-exponential parameters τ_{slow}, τ_{fast} and w resulting from the fits shown in (a). Results correspond to flexible membrane settings $\mu = 2.4$.
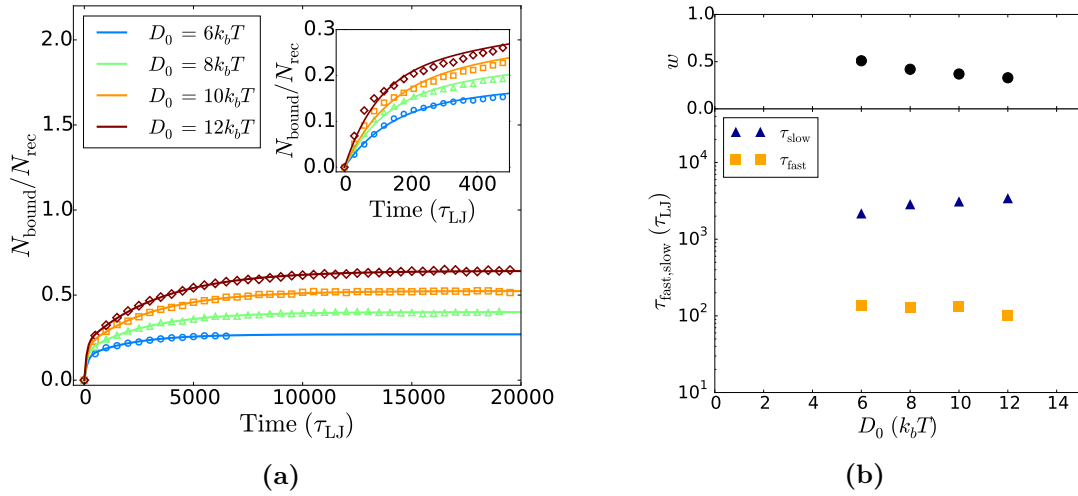

(a)

(b)

Figure D.10: (a) Plot of the average fraction of receptors bound to the cylinder as a function of time obtained from simulations (markers) which have been fitted with eq. (5.2) (lines). Inset zooms in on short time behavior. (b) Plot of the bi-exponential parameters τ_{slow}, τ_{fast} and w resulting from the fits shown in (a). Results correspond to stiff membrane settings $\mu = 6.0$.