



Dissertation

Multiscale Mathematical Modeling of Cell Migration: From Single Cells to Populations.

Aydar Uatay

Vom Fachbereich Mathematik der Technischen Universität
Kaiserslautern zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften
(Doctor rerum naturalium, Dr. rer. nat.)
genehmigte Dissertation

Gutachter:

Prof. Dr. Christina Surulescu, Technische Universität Kaiserslautern
Prof. Dr. Luigi Preziosi, Politecnico di Torino

Datum der Disputation: 24. Mai 2019

D 386

Acknowledgements

I am grateful to Prof. Dr. Christina Surulescu for supporting my pursuit of this work's topic and continuous guidance throughout my Ph.D. study.

I am thankful to the German Academic Exchange Service (DAAD) for their financial support throughout my doctoral studies.

I would like to thank my family for their encouragements, friends for motivation, colleagues for fruitful discussions, university faculty and staff for broadening my horizons and overall joyful experience during my stay at TU Kaiserslautern.

Abstract

Cell migration is essential for embryogenesis, wound healing, immune surveillance, and progression of diseases, such as cancer metastasis. For the migration to occur, cellular structures such as actomyosin cables and cell-substrate adhesion clusters must interact. As cell trajectories exhibit a random character, so must such interactions. Furthermore, migration often occurs in a crowded environment, where the collision outcome is determined by altered regulation of the aforementioned structures. In this work, guided by a few fundamental attributes of cell motility, we construct a minimal stochastic cell migration model from ground-up. The resulting model couples a deterministic actomyosin contractility mechanism with stochastic cell-substrate adhesion kinetics, and yields a well-defined piecewise deterministic process. The signaling pathways regulating the contractility and adhesion are considered as well. The model is extended to include cell collectives. Numerical simulations of single cell migration reproduce several experimentally observed results, including anomalous diffusion, tactic migration, and contact guidance. The simulations of colliding cells explain the observed outcomes in terms of contact induced modification of contractility and adhesion dynamics. These explained outcomes include modulation of collision response and group behavior in the presence of an external signal, as well as invasive and dispersive migration. Moreover, from the single cell model we deduce a population scale formulation for the migration of non-interacting cells. In this formulation, the relationships concerning actomyosin contractility and adhesion clusters are maintained. Thus, we construct a multiscale description of cell migration, whereby single, collective, and population scale formulations are deduced from the relationships on the subcellular level in a mathematically consistent way.

Zusammenfassung

Zellmigration spielt eine fundamentale Rolle für Embryogenese, Wundheilung, Immunüberwachung und Krankheitsverläufe, wie zum Beispiel bei Krebsmetastase. Damit die Migration stattfinden kann, müssen Zellstrukturen, wie z.B. Aktin-Myosin Stränge, mit Zell-Substrat-Adhäsionsclustern interagieren. Da Zelltrajektorien von stochastischer Natur sind, gilt dies ebenso für die Interaktionen. Zudem findet Zellmigration oft in dichtgedrängten Umgebungen statt, was das Ergebnis der Zellkollision durch die geänderte Regulation der oben genannten Strukturen beeinflusst.

Ausgehend von einigen fundamentalen Eigenschaften der Zellmotilität wird in dieser Arbeit ein minimales stochastisches Zellmigrationsmodell von Grund auf konstruiert. Das entwickelte Modell verbindet einen deterministischen Aktin-Myosin Kontraktionsmechanismus mit einer stochastischen Adhäsionskinematik von Zellsubstraten. Dies liefert einen wohldefinierten, stückweise stetigen, deterministischen Prozess. Die Signalpfade, welche die Kontraktion und Adhäsion regulieren, werden dabei ebenfalls beachtet. Das Modell wird dann erweitert, um auch Zellkollektiven Rechnung zu tragen. Numerische Simulationen der Migration einzelner Zellen sind in der Lage, mehrere experimentell beobachtete Sachverhalte zu reproduzieren, einschließlich anormaler Diffusion, taktischer Migration und „contact guidance“. Die Simulationsergebnisse für kollidierende Zellen entsprechen dem experimentell beobachteten Verhalten hinsichtlich kontaktinduzierter Veränderungen der Kontraktilität und Adhäsionsdynamik. Modulation der kollisionsinduzierten Reaktion der Zellen, kollektives Verhalten in Gegenwart externer Signale, sowie invasive und dispersive Migration können durch diesen Modellierungsansatz ebenfalls erklärt werden. Darüberhinaus wird aus dem Einzelzellmodell eine Formulierung auf der Populationsskala für die Migration nichtinteragierender Zellen hergeleitet, in der die Beziehung zwischen Aktin-Myosin Kontraktilitäts- und Adhäsionsclustern erhalten bleibt. Insgesamt wird eine mehrskalige Beschreibung der Zellmigration entwickelt, wobei Formulierungen für die Einzel-, Kollektiv- und Populationsskalen aus den Zusammenhängen auf der subzellulären Ebene mathematisch konsistent hergeleitet werden.

Acronyms

CIL	Contact inhibition of locomotion
ECM	Extracellular matrix
FA	Focal adhesion
GTP	Guanosine triphosphate
MSD	Mean squared displacement
PDMP	Piecewise deterministic Markov process
PIP	Phosphoinositides
SF	Stress fiber

Contents

1	Introduction	6
2	The Cell Migration Model	11
2.1	Overview	12
2.2	The migration cycle	17
2.2.1	Focal adhesion binding	18
2.2.2	Focal adhesion unbinding	18
2.2.3	Specifications	19
2.3	Focal adhesion events	21
2.3.1	Event times and probabilities	22
2.3.2	Cell motility and adhesion events	24
2.4	Motility process	25
2.4.1	Overview of PDMP	25
2.4.2	Cell motility and PDMP	28
3	Adhesion, Contraction, and Signaling	34
3.1	Focal adhesions	34
3.1.1	Unbinding rate	35
3.1.2	Binding dynamics	36
3.2	RhoGTPases	39
3.3	Stress fibers	40
3.4	Myosin motors	41
4	Numerical Simulations	48
4.1	Methods	49
4.1.1	Interarrival time generation	49
4.1.2	Sampling from the transition measure	51
4.2	Parameters	52
4.3	Results	53
4.3.1	Uniform environment	55
4.3.2	External cue gradient	60
4.3.3	Fibrillar architecture of ECM	65
4.3.4	Asymmetric contractility	69

5	Contact Inhibition of Locomotion	74
5.1	Binary collisions	75
5.2	Simulations	80
6	CIL in Collectives	86
6.1	Preliminaries	86
6.2	CIL and PDMP	87
6.3	Simulations	92
6.3.1	Methods	92
6.3.2	Results	93
7	Cell Populations	99
7.1	Probability law of a PDMP	99
7.2	The law of cell motility	102
7.3	Order reduction	105
7.3.1	Probabilities of FA states	106
7.3.2	Scaling and perturbation expansion	107
7.3.3	The diffusion tensor and the advection field	112
7.3.4	Summary of the population scale formulation	118
8	Conclusion and Outlook	119
A	Equations of cell motion	125
B	Data Analysis Methods	127

Chapter 1

Introduction

Cell migration is a fundamental biological process that is required for normal development and maintenance of an organism, but which can also contribute to its demise. In embryogenesis, coordinated migration of neural crest cells is required for formation of many organs and tissues in the body [116]. Migration of epithelial sheets and leukocytes is required for wound healing and immune response. During metastasis, which is the leading cause of death among cancer patients [18], tumor cells are able to become motile which enables them to migrate to distal organs.

Along with a variety of functions, there is also a multitude of cell migration modes. For example, spermatozoa and *E.coli* rely on flagellar activity to swim, leukocytes rely on rolling in the bloodstream, and epithelial sheets rely on tight coordination through intercellular junctions. The focus of this work, however, is on the crawling type of motion. Such motion, termed mesenchymal, is often described as a cyclical process, where morphological changes to the cell body are accompanied by coordinated interaction of adhesive contacts with the substrate to produce a crawling kind of movement [64]. The migration cycle comprise the following steps: protrusion of the membrane at the leading edge, assembly of adhesions at the front, and their release at the rear [1]. Following the last step, the cell body contracts, allowing the cell to move forward. Here, transmission of internally generated contractile forces onto an underlying substrate through strong adhesions is necessary to produce locomotion. In contrast, the amoeboid mode relies on more numerous, but weaker adhesions, and a more contractile cell body [90]. A cell, however, can transition from mesenchymal to amoeboid mode and vice versa, depending on external conditions [70].

Cell migration is a highly complex process involving interactions of different cellular structures, each contributing to a certain aspect of motility. This is reflected in numerous mathematical models, which highlight how a particular set of interactions translate into the observed migration patterns (see [49], [132] for reviews on cell motility models and [79] for a review on modeling of its critical components). For example, the contribution of, among other things, actin polymerization and myosin induced contractility on cell shape has been investigated in free boundary and phase-field models of steadily migrating cells in [16], [72], [86], [99], [106]. Models of cell migration in [24], [131] explored emergence of various motility modes due to mechanical coupling of intracellular components and the substrate. An essential role in mediating leading edge protrusions, adhesion formation and

contraction of the cell body is played by RhoGTPases - a family of signaling molecules [94], [97]. The dynamic interaction of its members (RhoA, Rac1, and Cdc42) and their effects on mechanical aspects of actin polymerization have been investigated in hybrid models by Marée et al. [73], [74], where agent-based modeling was utilized. A common feature of these models is that the resulting migration paths are not stochastic, while experimental observations suggest otherwise. To investigate random motility, stochastic models in [7], [31], [46], [101], [119] were proposed. There cell-substrate interaction has also been included and the underlying stochastic process is a Gaussian. However, there is evidence that the migration paths do not follow a Gaussian process [32], [67], [105]. To replicate the run-and-tumble behavior, a velocity jump process has been proposed in [87]. While stochastic models are able to generate random migration paths, their major drawback is that they do not account for interaction between cellular structures, which is, however, essential for locomotion. Furthermore, it is not clear how these interactions during the migration cycle give rise to random or biased motility.

In this work, we propose a minimal representation of a motile cell, based on which we construct a model of stochastic cell migration. We postulate that only (un)binding events of an adhesion in the migration cycle are random. Between the events, a cell performs a deterministic motion (e.g. locomotion if an adhesion is released). Based on this, we derive the forms of distributions, followed by the events and their interarrival times. We stress that we do not assume any prior knowledge about the distributions. Rather, we only consider major factors in adhesion dynamics, such as coupling with the contractile machinery and the surrounding environment. Thereby, we will obtain a piecewise deterministic Markov process [28], describing our cell migration model. In order to perform numerical simulations, we will also propose an efficient method to generate the trajectories. Assessment of the simulation results shows that our model is able to explain, among other things, the superdiffusive time scaling of the mean-squared displacement [32], [67], [70], biased migration in the presence of an external cue, contact guidance [95], and directed movement due to asymmetric contractility (and in the absence of guidance cues) [123], [129].

Motile cells often migrate in a crowded environment, wherein they collide and interact with other cells. Following a collision, a cell may continue its motion, crawling on top of another. Another collision outcome leads to cessation of movement, repolarization, and migration away from the collision site - a process termed contact inhibition of locomotion (CIL), first coined by Abercrombie and Heaysman [2]. Since then, its role in embryogenesis, cell dispersion and collective motion has been established [112]. It has also been shown to be implicated in aiding dissemination and invasion of cancer cells. Retaining such repulsive response between themselves allows cancerous cells to disperse, while its loss towards normal ones facilitates invasion. Thus, CIL response is divided into two different types: homotypic and heterotypic, occurring between cells of the same and different types, respectively. The heterotypic response also induces directed movement of otherwise non-responsive cells by chemotaxing ones [117] and can result in a “chase-and-run” kind of behavior [118]. Regardless of the CIL type that cells undergo, the outcome of a collision can vary: after cessation of movement, the cells may move in the same or opposite directions [30], [68], [102]. These experimental results also suggest that the outcome is determined

stochastically, and could be altered in the presence of a chemoattractant [68].

As CIL alters migratory behavior of colliding cells, its key mechanisms invariably involve those of freely migrating ones. In particular, RhoA is activated near the contact sites [17] and Rac1 away from it [103], [117]. The former leads to increased contractility of prospective rear; the latter to formation of protrusions and cell-substrate adhesions in the free edge [98]. Consequently, cell-cell and cell-substrate contacts in the rear are ruptured and the cell moves in the direction of the free edge. Thus, colliding cells follow the stereotypical steps of the migration cycle of freely migrating cells, while the CIL response is integrated therein.

As is the case with freely migrating cells, phase-field and particle based mathematical models have been developed to address CIL specifically, and more broadly, collective motion. For example, the phase-field model in [59] is able to reproduce the statistical outcomes of collisions reported in [102]. In another phase-field model [71], where alteration of actomyosin dynamics due to collisions was taken into account, collective migration emerges as a result of inelastic collisions. A vertex-based model in [78], explicitly including the Rac1/RhoA interplay, investigated the emergence of collective migration resulting from co-attraction and contact inhibition. Particle- and agent-based models in [26], [30], [133], were also able to simulate outcomes in agreement with their experimental observations. However, these studies highlight cell-cell interactions and model CIL as a stand-alone phenomenon, not integrated in the general migration mechanism of mesenchymal cells described above.

Encouraged by the ability of our single cell motility model to explain a variety of experimental observations, we sought to extend it in order to include the CIL response. In our extended model, the cells undergoing CIL cease movement upon contact, followed by increased actomyosin contractility around the collision site and elevated affinity for adhesion formation away from it. The extended model of a collection of cells is also described by a piecewise deterministic Markov model (PDMP). Unlike the case of non-interacting cells, here we have an “active” boundary, such that a jump occurs when the process hits it. This way we incorporate collisions into our model. Our simulation results are able to explain several experimental observations, such as modulation of CIL response in the presence of a chemotactic gradient [68], inducement of directed migration of non-chemotaxing cells due to CIL [117], and invasive migration in the presence and absence of heterotypic CIL, respectively. We also extend our numerical method to include the more general, but computationally expensive case of interacting cells.

One of the major findings of this work is that the motility of freely migrating and interacting cells can be explained in terms of mechanical coupling of cell cytoskeleton with cell-substrate adhesions in a mathematically consistent manner. Furthermore, the randomness of migration paths, which is a hallmark of cell motility, is explained by a constructed, rather than postulated, non-Gaussian stochastic process. To the best of our knowledge, this is also the first time that the migration cycle is (at least partially) explicitly taken into account as well.

The migration of cells is a multifaceted, multiscale phenomenon. This is especially prominent in the case of cancer spread. Its onset is defined by accumulation of catastrophic

mutations, ultimately followed by the invasion of cancer population into nearby tissue. To study this issue, models of migrating cell populations have been developed and analyzed in [77], [109], [110],[111], where the interactions between the different scales haven been described directly in a rather heuristic way. In these models, the migrating population is described by a reaction-diffusion-advection equation, coupled to a set of PDEs and ODEs reflecting the lower scale dynamics of transmembrane receptors, contractility, or acidity. Another approach to account for cell-ECM adhesions in migrating populations was presented in [33]. There, following and extending the approaches in [48], [56], the authors obtained an advection-diffusion equation via parabolic scaling of the velocity jump process. The forms of the corresponding advection field and diffusion tensor follow from the underlying model at the mesoscopic scale, rendering this approach more precise. However, the mesoscopic scale model itself stems from the underlying microscopic jumps process, as was first shown in [87] in the context of cell motility. Therefore, the population scale model in [33] is derived, rather than postulated, from the cell scale model of the velocity jump process.

Inspired by this approach, we also sought to derive a population scale model from our single cell motility description. Our formulation of the motility process as a PDMP allows us to derive the mesoscopic model. Observing the disparity of time scales at which the underlying processes of cell migration occur, aided by our detailed account of adhesion dynamics, and proceeding similarly as in [33], [48], [88], [89], we are able to obtain the population scale model. At this scale, unlike the studies in [33], [56], the coupling between cell-substrate adhesions and the internal contractile machinery is included, and unlike the studies in [77], [110] the coupling follows from the underlying motility model at the cell scale. Altogether, our work offers a more complete picture of cell migration as a multiscale phenomena. Because our cell motility model is able to explain a variety of experimental observations, our model at the population scale provides a more realistic way to assess invasion of cancer population.

This dissertation is organized as follows. In Chapter 2 we provide an overview of the migration cycle, the involved cellular structures, and how they relate to each other. We then introduce the minimal representation of a motile cell and describe the deterministic motion of a cell between the events of the migration cycle. Thereafter we construct a stochastic model of adhesion events, which signify transitions of the cycle stages. Combining the deterministic and the stochastic components, we formulate the motility process as a PDMP. In Chapter 3 we discuss in greater detail the dynamics of adhesion clusters and the contractile apparatus, and how the regulatory role of RhoGTPases signaling pathways can be integrated into our model. In Chapter 4 we present a numerical method for simulating the cell motility process, and the simulation results as well. We extend our single cell migration model to account for CIL in Chapter 5. Here we investigate the mechanism of CIL by considering a special case of binary collisions and simulations thereof. A general treatment of CIL involving cell collectives is presented in Chapter 6. Here we generalize the single cell PDMP to encompass colliding cells and present an appropriate numerical method. The simulation results are presented as well. In Chapter 7 we deduce a population migration model of non-interacting cells from the single cell model. We also reduce the

order of the population scale model by introducing a few simplifying assumptions. Finally, we present our conclusions and suggestions for future work in Chapter 8.

Most of the contents in this dissertation are based on the author's works in [120], [121]. In particular, Chapters 2, Section 3.1, Chapter 4, and Appendix B are based on [120]; Chapters 5 and 6 are based on [121].

Chapter 2

The Cell Migration Model

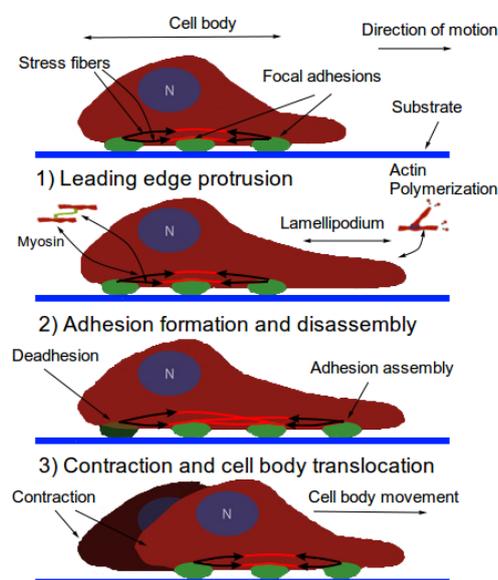


Figure 2.1: Schematic diagram of the cell migration cycle and the implicated cellular structures. Actin polymerization at the front pushes the membrane allowing protrusions to form. Then, adhesions assemble at the front and disassemble at the rear. Finally, deadhesion and cell contraction produce locomotion, pulling the body forward. The black arrows overlaying the stress fibers show the inwardly directed contractile forces. Modified from [69].

The cell migration cycle begins with protrusion of the leading edge as a result of actin polymerization (Figure 2.1). The polymerization process in lamellipodia is mediated by the Arp2/3 complex, which acts downstream of signaling pathways responsible for cell polarization [97]. Next, the protrusions are stabilized due to formation of focal adhesions (FAs) in the lamellae (region behind the lamellipodia), which link the actin cytoskeleton to the extracellular matrix (ECM). An FA is a multiprotein integrin-based adhesion cluster,

which matures in a RhoGTPase dependent manner [94]. Furthermore, FA maturation depends on the applied tension, and occurs concomitantly with actomyosin bundle formation [41], [130], and hence colocalizes with the bundle ends [108]. These bundles, called stress fibers (SFs), generate contractile forces due to non-muscle myosin II motors. A more detailed description of FAs and SFs is postponed until Sections 3.1 and 3.3. Due to increased tension at cell rear, FAs rupture. Finally, deadhesion leads to cell body translocation due to cytoskeletal contraction.

In order to construct the mathematical model, we make the following observations. First, FA unbinding leads to remodeling of the SF configuration (and of the entire cytoskeleton) and to cell movement, whereas assembly of new FAs leads to restructuring only. Second, FA events need not occur in the order described above. Several adhesions might be assembled (disassembled) before deadhesion (adhesion) occurs. Note also that while the contractile machinery is important, the dynamic instability of adhesions is what drives the migratory process, for stable FAs prevent retraction. Thus, we consider only interactions of SFs and FAs. Moreover, we do not consider the actin polymerization process and simplify the migration cycle down to two steps: after FA assembly occurs, a cell does not move, but reconfigures SFs; after disassembly, a cell does both. Neglecting the polymerization process and the reduction to binding/unbinding events can be justified by the fact that one of the major consequences of the leading edge protrusions is promotion of FA assembly. Because the repolarization of migrating cells occurs frequently as an outcome of intricate biochemical activity, then, in order to keep the model tractable, we do not explicitly model cell polarity. Instead, (de)adhesion frequency is indicative of (rear)front.

This chapter is based on the work [120] by the author.

2.1 Overview

Consider the situation in Figure 2.2. The disk represents a cell. Let the radius be R_{cell} and let the position of the center at time t be $\mathbf{x}(t) \in \mathbb{R}^2$. Suppose there are M equally spaced adhesion sites $\mathbf{x}_i(t) \in R_{cell}\mathbb{S}^1$, $i = 1, \dots, M$ on a cell circumference with constant relative distance. Let $\mathbf{Y}(t) \in \{0, 1\}^M$ be a vector of focal adhesion states at time t , i.e. $Y_i(t) = 0, 1$ correspond to unbound and bound FA at node i , respectively.

Since the traction stresses are oriented inward, transmitted to ECM by FAs, and generated by contractile SFs, then the FAs on the circumference must be one of the ends of SFs [80], [84], [108], [104]. Suppose the other end of all SFs at time t is at the position $\mathbf{x}_n(t) \in \Omega_{cell} := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq R_{cell}^2\}$ (in a cell's reference frame with origin at \mathbf{x}), i.e. all SFs are connected at \mathbf{x}_n . Since stress fibers behave like Hookean springs on extension, but readily buckle under compression [83], then, inspired by Guthardt Torres et

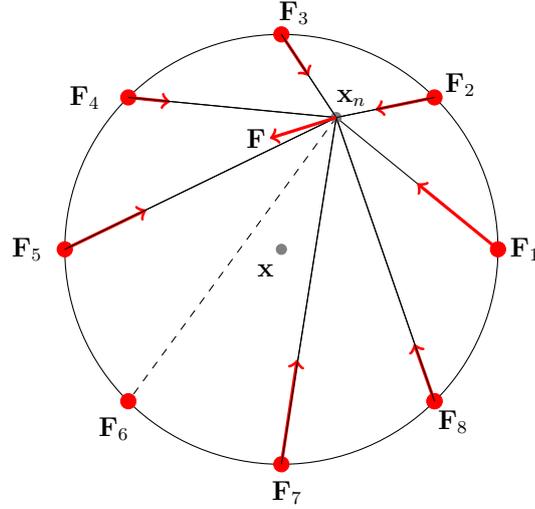


Figure 2.2: Schematic representation of a cell with $M = 8$ focal adhesions. Solid black lines represent stress fibers while red bullets represent focal adhesions. Red arrows indicate the direction and magnitude of applied traction force \mathbf{F}_i , $i = 1, \dots, 8$. The dashed line and the corresponding red bullet represent an absent stress fiber and unbound focal adhesion, respectively. The central red arrow indicates the net force \mathbf{F} on \mathbf{x}_n .

al. [47], the force \mathbf{F}_i at focal adhesion i is given by:

$$\mathbf{F}_i = \begin{cases} \left(T_i + EA \frac{L_i - L_0}{L_0} \right) \mathbf{e}_i, & L_0 < L_i \\ T_i \mathbf{e}_i, & L_c \leq L_i \leq L_0 \\ \frac{L_i - L_c + \delta}{\delta} T_i \mathbf{e}_i, & L_c - \delta \leq L_i < L_c \\ 0 & L_i < L_c - \delta, \end{cases} \quad (2.1.1)$$

where T_i is the magnitude of the contractile force due to myosin motors, EA is the one-dimensional Young's modulus, L_0 and L_c are, respectively, rest and critical lengths, $L_i = \|\mathbf{x}_n - \mathbf{x}_i\|$, $\mathbf{e}_i = \frac{\mathbf{x}_n - \mathbf{x}_i}{L_i}$ is the unit vector along the i^{th} SF, and δ is a small positive constant. The first case in (2.1.1) is due to the Hookean behavior of SFs upon extension and myosin tension generation. Furthermore, stress fiber laser ablation experiments [54], [61], [100] revealed that the initial instantaneous response (elastic behavior due to the SF length dependence in the first case) is followed by slower contraction due to myosin activity (force dependence on T_i) in the remaining portion of the fiber (Figure 2.3). Combined with stress fiber buckling, we obtain the second case in (2.1.1). Deguchi et al. [29] also found that SF contraction ceased after reaching a certain critical length. This implies that $F_i = 0$ when $L_i < L_c - \delta$. For technical reasons we assume F_i is piecewise continuous - hence the last cases in (2.1.1). We also assumed for simplicity that the myosin generated force T_i may vary between SFs, but is otherwise constant.

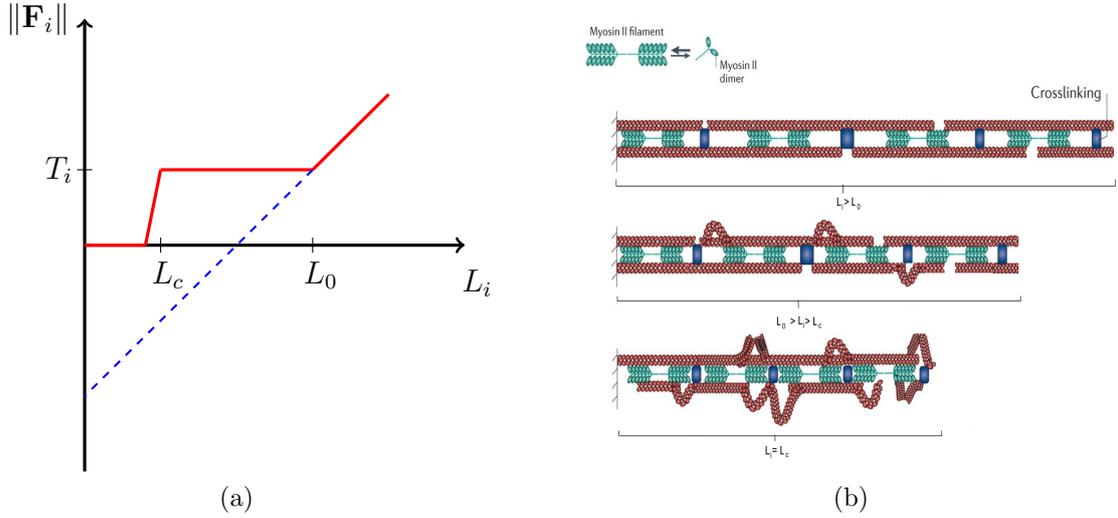


Figure 2.3: (a) Magnitude of \mathbf{F}_i in red. The blue dashed line corresponds to the profile of \mathbf{F}_i along \mathbf{e}_i if we were to treat the fiber as a Hookean spring with the spring constant EA/L_0 . (b) Schematic representation of stress fiber contraction. As the fiber contracts below the rest length L_0 , buckling occurs. As myosin mediated contraction causes the fiber to contract below the rest length L_0 , buckling occurs due to lack of resistance to compression. Below the critical length L_c , the fiber ceases to contract due to vanishing interfilament distance. Modified from [83].

Since $\mathbf{x}_i(t) \in R_{cell}\mathbb{S}^1$ and FA sites are equally spaced, then in polar coordinates we have:

$$\mathbf{x}_i(t) = R_{cell}(\cos(\theta_i), \sin(\theta_i))^T, \quad \theta_i(t) := \theta(t) + (i-1)\frac{2\pi}{M},$$

where θ_i, θ are the polar positions of the i^{th} and the first FAs, respectively. Note that it is sufficient to keep track of θ , due to equidistant FAs. Thus, $\mathbf{F}_i = \mathbf{F}_i(\mathbf{x}_n, \theta)$.

Since the total force at \mathbf{x}_n is

$$\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) := - \sum_{i=1}^M Y_i \mathbf{F}_i(\mathbf{x}_n, \theta), \quad (2.1.2)$$

then, assuming negligible inertial effects (due to the viscous nature of cytoplasm) and constant \mathbf{Y} :

$$\beta_{cell} \dot{\mathbf{x}}_n = \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}), \quad (2.1.3)$$

where β_{cell} is the drag coefficient in the cytoplasm.

The representation of a cell in such a way is justified, for the following reasons:

- The traction stresses are largely applied on the cell periphery and their magnitude decays rapidly towards the center [80], [104] (see Figures 2.4, 2.6). Thus, the cell

body SF ends are at or near mechanical equilibrium. Since contractile forces are generated by SFs, then a cell body SF end must be balanced by all other SFs (due to the equilibrium). Hence, it is reasonable to have a single connecting node of radial SFs which is either at mechanical equilibrium (for stationary cells) or tends to it.

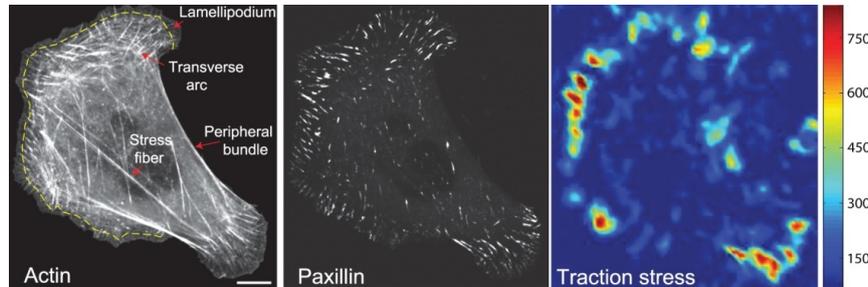


Figure 2.4: Fluorescence (left and center) and traction force (right) microscopy images of a fibroblast [104]. In the left image the actin cytoskeleton is shown, while in the center FA-associated paxillin. Note the correspondence between stress fibers, adhesions, and traction applied on the substrate.

- Since motile cells assume a wide variety of cell shapes and continuously remodel their actin cytoskeleton, one can view this representation as a cell shape normalization (it is implicitly assumed that a cell volume remains constant). That is, Figure 2.2 depicts a cell and forces applied on FAs normalized to a circle. Möhl et al. [80] applied shape normalization technique to a timelapse series data of migrating keratinocytes and demonstrated that this allows consistent analysis of FA dynamics, actin flow and traction forces (Figure 2.6). In view of their results, a particular cell traction force map and FA configuration normalized to a circle can be effectively captured by our representation.
- Paul et al. [91] demonstrated in their active cable network model, that combined with force application originating from nuclear region on FAs by star-like SF arrangement, results in cells acquiring morphologies typical for motile cells (Figure 2.5). Since the distribution of forces applied on FAs affects their rupturing, then it also influences the motion of a cell (due to retraction). Since we are primarily interested in cell migration it is justified to assume that this architecture represents a realistic situation. Furthermore, Oakes et al. [85] found that modeling SFs embedded in contractile networks, where only SFs actively contract, yields a behavior mimicking the experimental results - the cytoskeletal flow was directed along the stress fibers. In the same study, the authors concluded that it is appropriate to treat an SF as a 1D viscoelastic contractile element, which also justifies neglecting inertia in (2.1.2).

Note that $\mathbf{x}_n(t) \in \Omega_{cell}$ for $t > 0$, as proved below.

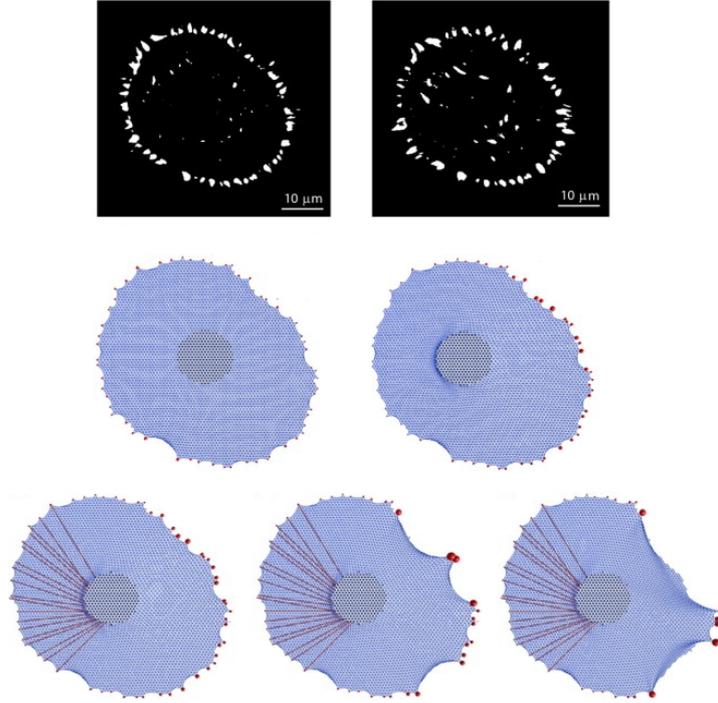


Figure 2.5: Experimental and simulation results of a non-motile cell in [91]. (Top and middle) Experimental and simulated images of a prestressed cell. Focal adhesions are represented in white. (Bottom) Simulation of internally generated contractile force from perinuclear region.

Proposition 2.1.1. *Let $\mathbf{x}_n \in \partial\Omega_{cell}$, $\theta \in [0, 2\pi)$, and $\mathbf{Y} \in \{0, 1\}^M$ be arbitrary. Let \mathbf{n} be the outward unit normal at \mathbf{x}_n . Then $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \mathbf{n} \leq 0$ with equality sign if and only if $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) = 0$.*

Proof. Suppose $\exists \mathbf{x}_n \in \partial\Omega_{cell}$, $\exists \theta \in [0, 2\pi)$, and $\exists \mathbf{Y} \in \{0, 1\}^M$ such that

$$\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \mathbf{n} = \sum_{i=1}^M Y_i (-\mathbf{F}_i(\mathbf{x}_n, \theta)) \cdot \mathbf{n} > 0. \quad (2.1.4)$$

Since $\mathbf{x}_n \in \partial\Omega_{cell}$, then in polar coordinates we have $\mathbf{x}_n = R_{cell}(\cos(\hat{\theta}), \sin(\hat{\theta}))^T$ for some $\hat{\theta} \in [0, 2\pi)$. Note that $\forall i$ such that $\theta_i \notin [\hat{\theta} - \pi/2, \hat{\theta} + \pi/2]$ we have:

$$Y_i (-\mathbf{F}_i(\mathbf{x}_n, \theta)) \cdot \mathbf{n} = Y_i \|\mathbf{F}_i(\mathbf{x}_n, \theta)\| (-\mathbf{e}_i \cdot \mathbf{n}) = \frac{Y_i \|\mathbf{F}_i(\mathbf{x}_n, \theta)\|}{L_i} (\mathbf{x}_i - \mathbf{x}_n) \cdot \mathbf{n} \leq 0,$$

since $\mathbf{x}_i \cdot \mathbf{n} < 0$ and $\mathbf{x}_n \cdot \mathbf{n} > 0$. Thus $\exists i \in \{1, \dots, M\}$ such that $\theta_i \in [\hat{\theta} - \pi/2, \hat{\theta} + \pi/2]$, $Y_i = 1$ and $(-\mathbf{F}_i(\mathbf{x}_n, \theta)) \cdot \mathbf{n} > 0$. Otherwise the inequality (2.1.4) does not hold and we are done. But then,

$$(\mathbf{x}_i - \mathbf{x}_n) \cdot \mathbf{n} > 0 \Rightarrow (\cos(\theta_i), \sin(\theta_i))^T \cdot \mathbf{n} > 1,$$

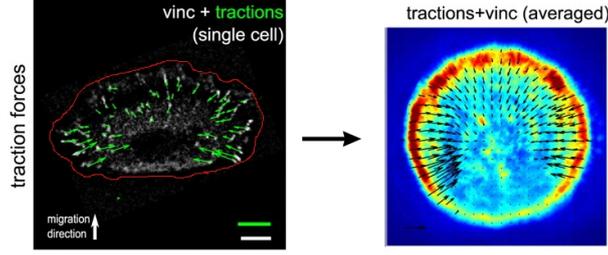


Figure 2.6: Shape normalization applied to migrating keratinocytes. Note that vinculin is an adapter protein that is recruited to FAs. Heat map on the right figure indicates FA activity; black arrows show magnitude and direction of traction forces. Source: [80]

which is false, since both are unit vectors. Therefore, $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \mathbf{n} \leq 0$ for arbitrary $\mathbf{x}_n \in \partial\Omega_{cell}$ and $\mathbf{Y} \in \{0, 1\}^M$.

Suppose $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \mathbf{n} = 0$. Then, we have (by the above considerations):

$$\sum_{i'} Y_{i'} \mathbf{F}_{i'}(\mathbf{x}_n, \theta) \cdot \mathbf{n} = \sum_i Y_i (-\mathbf{F}_i(\mathbf{x}_n, \theta)) \cdot \mathbf{n} \geq 0,$$

where i' is such that $\theta_{i'} \notin [\hat{\theta} - \pi/2, \hat{\theta} + \pi/2]$ and i is such that $\theta_i \in [\hat{\theta} - \pi/2, \hat{\theta} + \pi/2]$. If we have equality, then obviously $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) = 0$. If there is inequality, then $\exists i$ such that $\theta_i \in [\hat{\theta} - \pi/2, \hat{\theta} + \pi/2]$, $Y_i = 1$ and $(-\mathbf{F}_i(\mathbf{x}_n, \theta)) \cdot \mathbf{n} > 0$, which is also false (see above).

Therefore, $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) = 0$. ■

Corollary 2.1.2. *Let $\mathbf{x}_n(0) \in \Omega_{cell}$ be arbitrary and let $\theta(t)$ be given. Suppose $\mathbf{x}_n \in C^1([0, \infty))$ is a solution of (2.1.3). Then, $\mathbf{x}_n(t) \in \Omega_{cell} \forall t > 0$ and $\forall \mathbf{Y}(t) \in \{0, 1\}^M$.*

Proof. Due to (2.1.3) it suffices to show that $\forall \mathbf{x}_n \in \partial\Omega_{cell}$ we have $\mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \mathbf{n} \leq 0$, which follows from Proposition 2.1.1. ■

2.2 The migration cycle

Recall that during the migration cycle, deadhesion leads to cell body translocation, while adhesion binding does not. In both cases actomyosin contractility leads to reconfiguration of the cytoskeleton. Here we show how our cell representation can describe the reconfiguration and cell body motion following binding and unbinding events.

Without loss of generality assume that an event occurred at $t = 0$. Let $\tau > 0$ be the time of the next adhesion event, be it binding or unbinding. Let $\mathbf{Y}(0) \in \{0, 1\}^M$, $\mathbf{x}(0) \in \mathbb{R}^2$, and $\mathbf{x}_n(0) \in \Omega_{cell}$ be arbitrary. Then, $\mathbf{Y}(t) = \text{const.}$ for $t \in [0, \tau)$. We assume $\theta(t = 0) = 0$.

2.2.1 Focal adhesion binding

Following an FA binding, we suppose that a cell becomes stationary (i.e. the cell centroid remains constant). However, a newly formed FA and the associated SF lead to cytoskeletal reshaping. Thus, we have the following system of ODEs for $t \in [0, \tau)$:

$$\begin{aligned}\dot{\mathbf{x}} &= 0 \\ \dot{\mathbf{x}}_n &= \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \\ \dot{\theta} &= 0.\end{aligned}\tag{2.2.1}$$

2.2.2 Focal adhesion unbinding

Following an unbinding event, the cytoskeletal contraction leads to cell body movement. Due to the circular geometry, the contractile forces induce both rotational and translational motion.

Note that the bound focal adhesions are able to slide for short distances [80]. Oakes et al. [85] found that the cytoskeleton behaves like an elastic solid on timescales up to one hour. Provided the time τ between adhesion events is small enough, the following is justified.

The force \mathbf{F} along the radial vector $\hat{\mathbf{r}}(\mathbf{x}_n)$ is acting on the cell center, thereby inducing translational motion (see Figure 2.7). On the other hand, the rotational motion is produced due to \mathbf{F} acting along the tangential vector $\hat{\boldsymbol{\varphi}}(\mathbf{x}_n)$. The radial and tangential components of the force \mathbf{F} are given by:

$$\begin{aligned}F_r &:= \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \\ F_\varphi &:= \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\boldsymbol{\varphi}}(\mathbf{x}_n),\end{aligned}$$

where $\mathbf{x}_n = (x_{n,1}, x_{n,2})$ and

$$\hat{\mathbf{r}}(\mathbf{x}_n) = \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|}, \quad \hat{\boldsymbol{\varphi}}(\mathbf{x}_n) = \left(-\frac{x_{n,2}}{\|\mathbf{x}_n\|}, \frac{x_{n,1}}{\|\mathbf{x}_n\|} \right)^T. \tag{2.2.2}$$

The characteristic Reynolds number Re is given by

$$Re = \frac{\rho \cdot s \cdot L}{\nu} \sim 10^{-6} - 10^{-4},$$

where we assumed the surrounding fluid is water (with corresponding values for density ρ and viscosity ν , and that characteristic cell speed s and size L are $0.1 - 1 \mu\text{m}/\text{s}$, $L =$

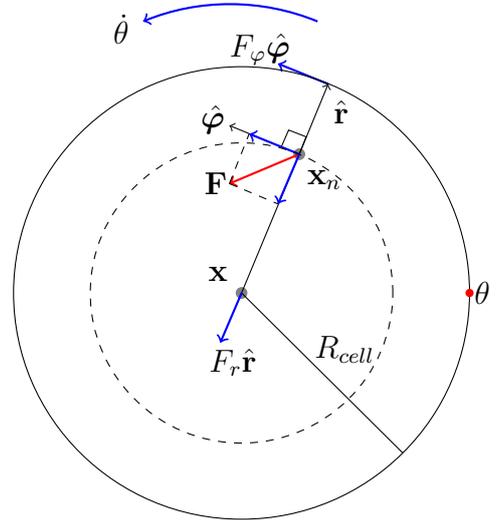


Figure 2.7: Force diagram showing transmission of internally generated contractile forces into translational and rotational motion. $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\varphi}}$ are radial and angular unit vectors, respectively. $\dot{\theta}$ is the angular velocity, \mathbf{F} is a net contractile force, F_r and F_φ are radial and tangential components of \mathbf{F} , \mathbf{x} and R_{cell} are the cell center and radius, respectively.

10 – 50 μm , respectively. Since Re is small, we neglect inertia and obtain (see Appendix A for details):

$$\begin{aligned}\dot{\mathbf{x}} &= \beta_{ECM}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \hat{\mathbf{r}}(\mathbf{x}_n) \\ \dot{\mathbf{x}}_n &= \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \\ \dot{\theta} &= \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\boldsymbol{\varphi}}(\mathbf{x}_n)\end{aligned}\quad (2.2.3)$$

where β_{ECM} and β_{rot} are, respectively, translational and rotational drag coefficients in the ECM.

2.2.3 Specifications

It is convenient to transform the system above into nondimensional form. In order to do so, we define the following scales:

- The spatial and cell length scales are defined by the cell radius R_{cell} .
- The time scale is defined by some constant t_e , which is of the same order as the FA lifetime (see Section 3.1 for details).
- The force scale is defined by the characteristic force F_b .

The constants are to be specified later. Whence we define the new variables:

$$\tilde{\mathbf{x}} := \frac{\mathbf{x}}{R_{cell}}, \quad \tilde{\mathbf{x}}_n := \frac{\mathbf{x}_n}{R_{cell}}, \quad \tilde{t} := \frac{t}{t_e}$$

and transform \mathbf{F}_i from (2.1.1):

$$\tilde{\mathbf{F}}_i := \frac{\mathbf{F}_i}{F_b} = \begin{cases} \left(\tilde{T}_i + \widetilde{EA} \frac{\tilde{L}_i - \tilde{L}_0}{\tilde{L}_0} \right) \tilde{\mathbf{e}}_i, & \tilde{L}_0 < \tilde{L}_i \\ \tilde{T}_i \tilde{\mathbf{e}}_i, & \tilde{L}_c \leq \tilde{L}_i \leq \tilde{L}_0 \\ \frac{\tilde{L}_i - \tilde{L}_c + \tilde{\delta}}{\tilde{\delta}} \tilde{T}_i \tilde{\mathbf{e}}_i, & \tilde{L}_c - \tilde{\delta} \leq \tilde{L}_i < \tilde{L}_c \\ 0, & \tilde{L}_i < \tilde{L}_c - \tilde{\delta}, \end{cases}\quad (2.2.4)$$

where

$$\begin{aligned}\tilde{L}_i &= \frac{L_i}{R_{cell}}, & \tilde{L}_0 &= \frac{L_0}{R_{cell}}, & \tilde{L}_c &= \frac{L_0}{R_{cell}}, & \tilde{\delta} &= \frac{\delta}{R_{cell}}, \\ \tilde{T}_i &= \frac{T_i}{F_b}, & \widetilde{EA} &= \frac{EA}{F_b}, & \tilde{\mathbf{e}}_i &= \frac{\tilde{\mathbf{x}}_n - \tilde{\mathbf{x}}_i}{\tilde{L}_i}, & \tilde{\mathbf{x}}_i &= \frac{\mathbf{x}_i}{R_{cell}}.\end{aligned}$$

Note that we have $\tilde{\mathbf{x}}_n \in \tilde{\Omega}_{cell} := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ and $\tilde{\mathbf{x}}_i \in \mathbb{S}^1$.

Let

$$\tilde{\mathbf{F}} := \mathbf{F}/F_b, \quad \tilde{\beta}_{cell} := \frac{R_{cell}}{t_e F_b} \beta_{cell}, \quad \tilde{\beta}_{ECM} := \frac{R_{cell}}{t_e F_b} \beta_{ECM}, \quad \tilde{\beta}_{rot} := \frac{1}{t_e R_{cell} F_b} \beta_{rot}.$$

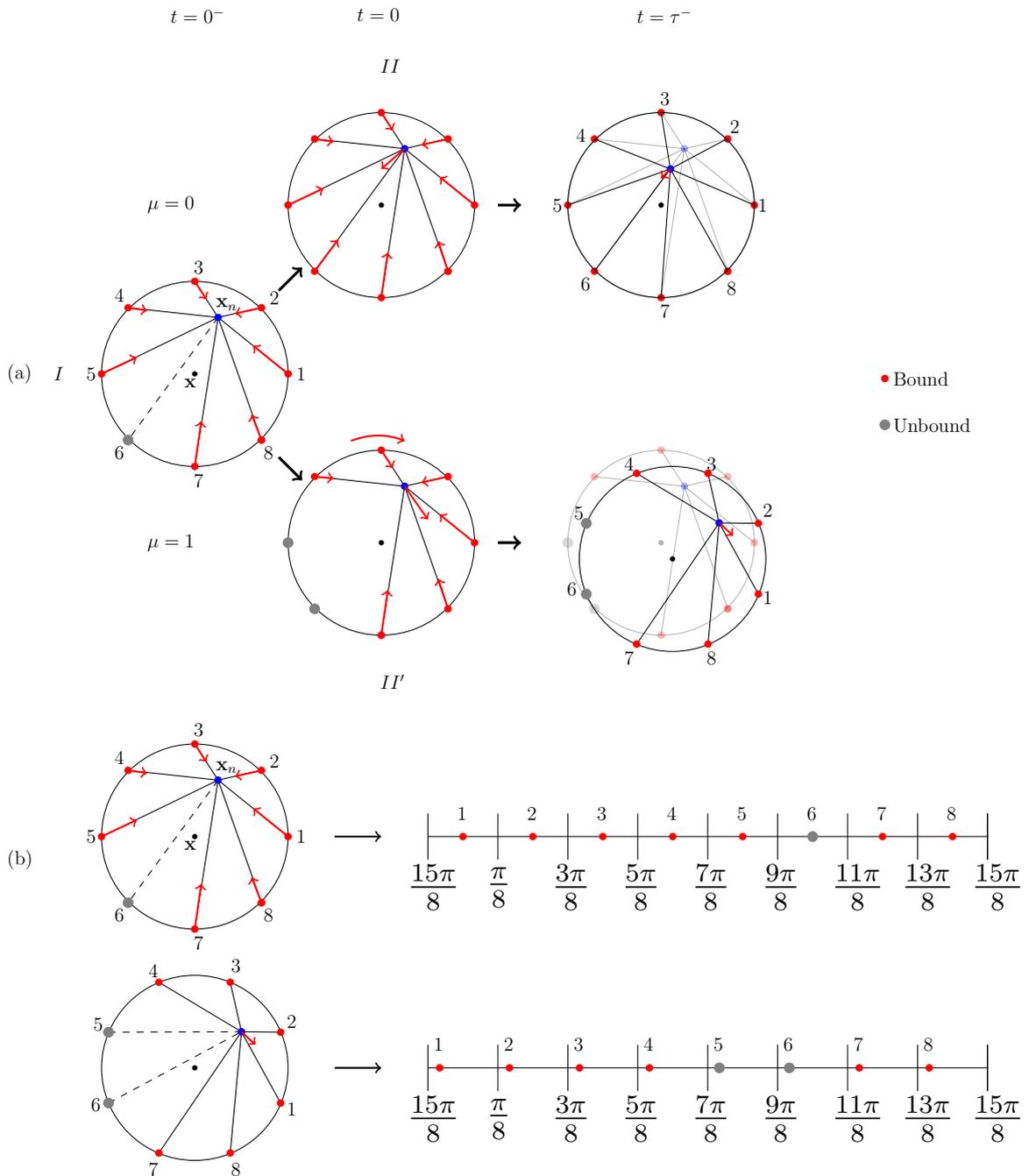


Figure 2.8: (a) Schematic representation of the migration cycle between adhesion events. Suppose that just before an event occurs at time $t = 0$, the cell is in state I . If at time $t = 0$ (de)adhesion occurs, the cell jumps into state the $(II')II$ and the system evolves according to (2.2.5) until the next event occurs at time $t = \tau$, after which the cycle begins anew. The scenarios can be characterized as “run” and “tumble” phases in the bottom and top panels, respectively. (b) Schematic representation of the FA positions projected on cell's circumference at $t = 0^-$ and $t = \tau^-$ in the top and bottom panels, respectively.

To complete the specification of cell kinematics between adhesion events, we introduce another discrete variable $\mu(t) \in \{0, 1\}$, which reflects asymmetry of the cell response to binding/unbinding:

$$\mu = \begin{cases} 1, & \text{if the last event was unbinding} \\ 0, & \text{if the last event was binding.} \end{cases}$$

Then, plugging in (2.2.1), (2.2.3) the rescaled quantities and dropping tildes, it follows that between the FA events the following ODE system holds for $t \in [0, \tau)$:

$$\begin{aligned} \dot{\mathbf{x}} &= \mu \beta_{ECM}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \hat{\mathbf{r}}(\mathbf{x}_n) \\ \dot{\mathbf{x}}_n &= \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \\ \dot{\theta} &= \mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\boldsymbol{\varphi}}(\mathbf{x}_n). \end{aligned} \quad (2.2.5)$$

The migration cycle can be described as follows (see Figure 2.8). Suppose that just before an event occurs, the cell is in state I with arbitrary $\mu \in \{0, 1\}$ at time $t \rightarrow 0^-$. If at time $t = 0$ (de)adhesion occurs, the cell jumps into the state $(II')II$ with a corresponding value of μ , and the system evolves according to (2.2.5) until the next event occurs at time $t = \tau$, after which the cycle begins anew. The scenarios can be characterized as “run” and “tumble” phases in the bottom and top panels of Figure 2.8, respectively. However, as mentioned in the introduction, we postulate that FA events occur randomly. It follows that the time of the next event τ is also random. In the following we elaborate on how to determine the corresponding distributions.

Remark. Our assumption on constant relative distance between FA sites stems from two slightly weaker assumptions: 1) total number of adhesion sites (occupied and unoccupied) is constant; 2) there is a neighborhood around each adhesion site, in which no other site is present, and the size of this neighborhood is the same (and constant) for each site. Figure 2.8 how it reflects on their peripheral motion. This assumption implies that in each line segment of size $2\pi/M$ (with $M = 8$) there is only one FA site present, which may correspond to bound (in red) or unbound FA (in gray).

2.3 Focal adhesion events

In the previous section we constructed a model of cell motion between FA events. Following [44], here we construct a stochastic model describing the random adhesion/deadhesion events and their arrival times. The discussion here differs from the standard approach of the Gillespie algorithm in [44], as we do not assume that the propensity functions vary inappreciably between the reactions. Moreover, it provides a connection to the theory of PDMPs, as the forms of the objects, necessary to define a piecewise deterministic process, follow from the derivations here.

2.3.1 Event times and probabilities

Since there are M FAs and since each FA can participate in only two reactions (binding and unbinding), then there are $2M$ total possible reactions. We adopt the following convention for enumerating reactions: reaction j corresponds to a *binding* reaction of the FA site $i = (j + 1)/2$ if j is odd; otherwise reaction j corresponds to an *unbinding* reaction of the FA site $i = j/2$. Let $\mathbf{Y}(t)$ be defined as before.

Let $a_j(\mathbf{y}, t)dt$ be the probability, given $\mathbf{Y}(t) = \mathbf{y} \in \{0, 1\}^M$, $\mu(t)$, $(\mathbf{x}(t), \mathbf{x}_n(t), \theta(t))$ and time t , that reaction j will occur in the time interval $[t, t + dt)$. For clarity, we suppress here the dependence of the rate $a_j(\mathbf{y}, \cdot)$ on $(\mathbf{x}(\cdot), \mathbf{x}_n(\cdot), \theta(\cdot))$ and $\mu(\cdot)$. We assume that the rate a_j satisfies the following:

$$a_j(\mathbf{y}, t) = \begin{cases} 0, & \text{if } j \text{ is odd and } y_{(j+1)/2} = 1 \\ 0, & \text{if } j \text{ is even and } y_{j/2} = 0 \\ \neq 0, & \text{else.} \end{cases} \quad (2.3.1)$$

That is, if the FA is (un)bound, the probability of the (un)binding reaction is zero; if the FA is (un)bound, the probability of (binding) unbinding is nonzero. This implies that $a_j(\mathbf{y}, t) \neq 0$ for at least one $j \in \{1, \dots, 2M\}$, since for each FA site $i \in \{1, \dots, M\}$, either a_{2i-1} or a_{2i} is nonzero.

Lemma 2.3.1. *Let $\mathbf{Y}(t) = \mathbf{y}$. Then the probability that no FA event occurs in the time interval $[t, t + dt)$ is $1 - \sum_{j=1}^{2M} a_j(\mathbf{y}, t)dt + o(dt)$.*

Proof. Using the definition of a_j , the probability that reaction j does not happen is $1 - a_j(\mathbf{y}, t)dt$. Then, the probability that no FA reaction occurs is:

$$\prod_{j=1}^{2M} (1 - a_j(\mathbf{y}, t)dt) = 1 - \sum_{j=1}^{2M} a_j(\mathbf{y}, t)dt + o(dt).$$

■

Let $K(\tau, j|t, \mathbf{y})d\tau$ be the probability, given $\mathbf{Y}(t) = \mathbf{y}$ and $(\mathbf{x}(t), \mathbf{x}_n(t), \theta(t))$ at time t , that the *next* reaction will occur in the time interval $[t + \tau, t + \tau + d\tau)$ and will be reaction j . Here, again, we suppress for clarity the dependence on $\mathbf{x}, \mathbf{x}_n, \theta$.

Proposition 2.3.2. *Let $\tau > 0$ and $\mathbf{Y}(t) = \mathbf{y}$. Then,*

$$K(\tau, j|t, \mathbf{y}) = a_j(\mathbf{y}, t + \tau) \exp \left(- \int_t^{t+\tau} \sum_{j'=1}^{2M} a_{j'}(\mathbf{y}, \tau') d\tau' \right).$$

Proof. Let $P(\tau|t, \mathbf{y})$ denote the probability that no reaction occurs in the time interval

$[t, t + \tau)$, given \mathbf{y} (and $\mathbf{x}, \mathbf{x}_n, \theta$) at time t . Then, by Lemma 2.3.1:

$$P(\tau + d\tau | t, \mathbf{y}) = P(\tau | t, \mathbf{y}) \left(1 - \sum_{j=1}^{2M} a_j(\mathbf{y}, t + \tau) d\tau + o(d\tau) \right) \Rightarrow$$

$$\frac{P(\tau + d\tau | t, \mathbf{y}) - P(\tau | t, \mathbf{y})}{d\tau} = -P(\tau | t, \mathbf{y}) \sum_{j=1}^{2M} a_j(\mathbf{y}, t + \tau) + P(\tau | t, \mathbf{y}) \frac{o(d\tau)}{d\tau}.$$

Letting $d\tau \rightarrow 0$ we obtain the following ODE:

$$\frac{d}{d\tau} P(\tau | t, \mathbf{y}) = -P(\tau | t, \mathbf{y}) \sum_{j=1}^{2M} a_j(\mathbf{y}, t + \tau).$$

Since $P(0 | t, \mathbf{y}) = 1$, the solution $P(\tau | t, \mathbf{y})$ is given by:

$$P(\tau | t, \mathbf{y}) = \exp \left(- \int_t^{t+\tau} \sum_{j=1}^{2M} a_j(\mathbf{y}, \tau') d\tau' \right).$$

We have then:

$$K(\tau, j | t, \mathbf{y}) = P(\tau | t, \mathbf{y}) a_j(\mathbf{y}, t + \tau) = a_j(\mathbf{y}, t + \tau) \exp \left(- \int_t^{t+\tau} \sum_{j'=1}^{2M} a_{j'}(\mathbf{y}, \tau') d\tau' \right). \quad (2.3.2)$$

■

Let $K_{time}(\tau | t, \mathbf{y}) d\tau$ be the probability that the next reaction will occur in the time interval $[t + \tau, t + \tau + d\tau)$, given $\mathbf{Y}(t) = \mathbf{y}$ and $(\mathbf{x}(t), \mathbf{x}_n(t), \theta(t))$ at time t .

Let $K_{index}(j | \tau, t, \mathbf{y})$ be the probability that the index of the next reaction is j given $\mathbf{Y}(t) = \mathbf{y}$, $(\mathbf{x}(t), \mathbf{x}_n(t), \theta(t))$ at time t and given that the reaction will occur at time $t + \tau$.

By elementary probability theory (using the definition of conditional probability), we know that

$$K(\tau, j | t, \mathbf{y}) d\tau = K_{index}(j | \tau, t, \mathbf{y}) K_{time}(\tau | t, \mathbf{y}) d\tau.$$

Due to equation (2.3.2), we see that:

$$K_{index}(j | \tau, t, \mathbf{y}) = \frac{a_j(\mathbf{y}, t + \tau)}{a_0(\mathbf{y}, t + \tau)}$$

$$K_{time}(\tau | t, \mathbf{y}) = a_0(\mathbf{y}, t + \tau) \exp \left(- \int_t^{t+\tau} a_0(\mathbf{y}, \tau') d\tau' \right), \quad (2.3.3)$$

where

$$a_0(\mathbf{y}, t) = \sum_{j=1}^{2M} a_j(\mathbf{y}, t),$$

and $a_0 \neq 0$ due to (2.3.1).

Obviously,

$$\sum_{j=1}^{2M} K_{index}(j|\tau, t, \mathbf{y}) = 1$$

$$\int_0^\infty K_{time}(\tau|t, \mathbf{y}) d\tau = 1.$$

Thus, if T is (random) time until the next reaction, then its probability density function given by K_{time} , its survival function $S(s)$ is given by (without loss of generality, suppose that $t = 0$):

$$\mathbb{P}(T > s) = S(s) = \exp\left(-\int_0^s a_0(\mathbf{y}, \tau') d\tau'\right), \quad (2.3.4)$$

and its (cumulative) distribution function is given by $1 - S(s)$.¹ Note that the distribution of a random variable is uniquely determined by its distribution function.

Using the proof of Proposition 2.3.2 one has the following:

Proposition 2.3.3. *Let $\tau > 0$ and let $\hat{K}(\tau|t, \mathbf{y})$ be the probability of more than one FA event occurring in the time interval $[t + \tau, t + \tau + d\tau)$, given the state of the system at time t . Then $\hat{K}(\tau|t, \mathbf{y}) = o(d\tau)$ as $d\tau \rightarrow 0$.*

Proof. By the proof of Proposition 2.3.2:

$$\hat{K}(\tau|t, \mathbf{y}) = P(\tau|t, \mathbf{y})o(d\tau),$$

since, following the definition of a_j , the probability of more than one reaction occurring in time interval $[t, t + d\tau)$ is $o(d\tau)$. ■

Proposition 2.3.3 implies that we can neglect the case when more than one FA event occurs at the event time. Thus, an FA event (binding or unbinding) unambiguously correspond to a switch in motility state. If this were not the case and the probability of two FA events at the same time were not negligible, then binding and unbinding of distinct FAs could occur simultaneously. Since the cell becomes motile after unbinding only, simultaneous events could lead to ambiguity in determining the motile state of the cell.

2.3.2 Cell motility and adhesion events

With the results of the previous section we can now formally state the cyclical mesenchymal cell motility model as a stochastic process (see Figure 2.8).

Let $t = 0$, $\mathbf{x}(0)$, $\mathbf{x}_n(0)$, $\theta(0)$, $\mu(0)$ be given and $\mathbf{Y}(0) = \mathbf{y}^0$.

- The time \mathcal{T}_1 of the FA event is chosen such that $\mathbb{P}(\mathcal{T}_1 > s) = S(s)$.

¹One can check this by differentiating the distribution function, given by $1 - S(s)$, with respect to s .

- The system evolves according to (2.2.5) for $t \in [0, \mathcal{T}_1)$.
- At time $t = \tau$, the index j of the FA event is chosen with probability $K_{index}(j|\mathcal{T}_1, 0, \mathbf{y}^0)$. \mathbf{Y} and μ jump to new values:

$$\mathbf{Y}(t = \tau) = \begin{cases} \mathbf{y}^0 + \widehat{\mathbf{e}}_i, & i = (j + 1)/2, \text{ if } j \text{ is odd} \\ \mathbf{y}^0 - \widehat{\mathbf{e}}_i, & i = j/2, \text{ else} \end{cases},$$

$$\mu(t = \tau) = \begin{cases} 0, & \text{if } j \text{ is odd} \\ 1, & \text{else} \end{cases},$$

where $\widehat{\mathbf{e}}_i \in \mathbb{R}^M$ is the standard basis vector. Note that due to (2.3.1), we always have $\mathbf{Y}(t = \tau) \in \{0, 1\}^M$, since the probability of (un)binding of (un)bound FA is zero.

- The cycle starts anew with initial time $t = \mathcal{T}_1$ and initial values of other variables at this time: starting at $t = \mathcal{T}_1$ we choose the time \mathcal{T}_2 of the FA event such that

$$\mathbb{P}(\mathcal{T}_2 > s|\mathcal{T}_1) = \exp\left(-\int_{\mathcal{T}_1}^{\mathcal{T}_1+s} a_0(\mathbf{y}, \tau')d\tau'\right).$$

- The system evolves according to (2.2.5) for $t \in [\mathcal{T}_1, \mathcal{T}_1 + \mathcal{T}_2)$ and so on.

One sees that the cyclical process described above is a Markov process, since the evolution of the system depends only on the current state. This completes the formal specification of the model. In the following we will show that this process is a well-defined piecewise deterministic Markov process.

2.4 Motility process

In this section we briefly overview a class of piecewise deterministic processes, first introduced by Davis [27]. We then show how the deterministic equations, describing motion between stochastic focal adhesion events, can be combined to yield a well-defined piecewise deterministic Markov process (PDMP).

2.4.1 Overview of PDMP

Let A be countable and let $\Gamma \subset \mathbb{R}^d$ be open. Let $\mathbf{X}_t \in \Gamma$ and let $\mathbf{H}_\nu : \Gamma \rightarrow \mathbb{R}^d$ for $\nu \in A$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space, where Ω is a sample space, \mathcal{F} is a σ -algebra on Ω , $(\mathcal{F}_t)_{t \geq 0}$ is a (natural) filtration, and \mathbb{P} is a probability measure. Let $E := \{(\nu, \boldsymbol{\xi}) : \nu \in A, \boldsymbol{\xi} \in \Gamma\}$ and let (E, \mathcal{E}) be a Borel space. For details see [Chapter 2 in [28]].

We can define the piecewise deterministic process on the state space (E, \mathcal{E}) (for a more detailed general treatment see Davis [28]) by the following objects¹:

¹Here we first provide a constructive definition. The verification of the conditions and their explicit representation corresponding to our case of cell motility are postponed for the sake of clearer exposition.

I Some vector fields $(\mathbf{H}_\nu, \nu \in A)$ such that for all $\nu \in A$ there exists a unique global solution $\mathbf{X}_t \in \Gamma$ to the following equation:

$$\begin{aligned} \frac{d}{dt} \mathbf{X}_t &= \mathbf{H}_\nu(\mathbf{X}_t) \\ \mathbf{X}_0 &\in \Gamma. \end{aligned} \quad (2.4.1)$$

Let $\phi_\nu : [0, \infty) \times \Gamma \rightarrow \Gamma$ denote the flow corresponding to Equation (2.4.1), i.e.

$$\phi_\nu(t, \mathbf{X}_0) = \mathbf{X}_t.$$

II A measurable function $a_0 : E \rightarrow \mathbb{R}_+$ such that the function $s \mapsto a_0(\nu, \phi_\nu(s, \mathbf{X}_0))$ is integrable.

III A transition measure $Q : \mathcal{E} \times E \rightarrow [0, 1]$, such that for fixed $C \in \mathcal{E}$, $(\nu, \xi) \mapsto Q(C; (\nu, \xi))$ is measurable for $(\nu, \xi) \in E$, and $Q(\cdot; (\nu, \xi))$ is a probability measure for all (ν, ξ) on (E, \mathcal{E}) .

Let $(\nu^0, \mathbf{X}^0) \in E$ at time $t = 0$ be given. Let a survival function S be defined similarly as in equation (2.3.4):

$$S(t, (\nu, \mathbf{X})) := \exp\left(-\int_0^t a_0(\nu, \phi_\nu(s, \mathbf{X})) ds\right). \quad (2.4.2)$$

Let \mathcal{T}_1 be the first jump time such that

$$\mathbb{P}(\mathcal{T}_1 > t \mid (\nu^0, \mathbf{X}^0)) = S(t, (\nu^0, \mathbf{X}^0)),$$

and let (ν^1, \mathbf{X}^1) be distributed according to the probability law $Q(\cdot, \phi_{\nu^0}(\mathcal{T}_1, \mathbf{X}^0))$. Then, the motion of (ν_t, \mathbf{X}_t) for $t \leq \mathcal{T}_1$ is given by:

$$(\nu_t, \mathbf{X}_t) = \begin{cases} (\nu^0, \phi_{\nu^0}(t, \mathbf{X}^0)), & t < \mathcal{T}_1, \\ (\nu^1, \mathbf{X}^1), & t = \mathcal{T}_1. \end{cases}$$

At time $t = \mathcal{T}_1$ the next jump time \mathcal{T}_2 is distributed such that

$$\mathbb{P}(\mathcal{T}_2 - \mathcal{T}_1 > s \mid (\nu_{\mathcal{T}_1}, \mathbf{X}_{\mathcal{T}_1})) = S(s, (\nu_{\mathcal{T}_1}, \mathbf{X}_{\mathcal{T}_1})).$$

The value of the process at the jump time \mathcal{T}_2 is determined by the measure $Q(\cdot, \phi_{\nu_{\mathcal{T}_1}}(\mathcal{T}_2, \mathbf{X}_{\mathcal{T}_1}))$ and the process continues in a similar way. Thus, we have a well-defined piecewise deterministic process [28].

Remark. We assume that the jumps occur only at the random times determined by the survival function (2.4.2), i.e. only at the FA event times. In a general PDMP, one could consider the case when a jump also occurs if \mathbf{X}_t hits the boundary of its domain (see [28]). We include this feature in our treatment of colliding cells.

Theorem 2.4.1 ([27]). *The process $(\nu_t, \mathbf{X}_t)_{t \geq 0}$ is a homogeneous Markov process.*

Proof. Let \mathcal{T}_k be the k^{th} jump time. Let $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$. Then, by construction of the process, we have:

$$\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > s) = S(s, (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k})).$$

In particular, for $s > 0$, we have:

$$\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > s \mid \mathcal{F}_t) = \mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > s \mid (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k})) = S(s, (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k})), \quad (2.4.3)$$

since, given that the jump occurred at time $t = \mathcal{T}_k$, the next jump time is distributed according to the information given at $t = \mathcal{T}_k$ by construction, i.e. the state of the process $(\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k})$.

We then have the following:

$$\begin{aligned} \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid \mathcal{F}_t) &= \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid \mathcal{F}_t, \mathcal{T}_{k+1} > t) \\ &= \mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > t + s - \mathcal{T}_k \mid \mathcal{F}_t, \mathcal{T}_{k+1} - \mathcal{T}_k > t - \mathcal{T}_k) \\ &= \frac{\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > t + s - \mathcal{T}_k \mid \mathcal{F}_t)}{\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > t - \mathcal{T}_k \mid \mathcal{F}_t)} \\ &= \frac{\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > t + s - \mathcal{T}_k \mid (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k}))}{\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > t - \mathcal{T}_k \mid (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k}))} \\ &= \exp\left(-\int_0^s a_0(\nu_{\mathcal{T}_k}, \phi_{\nu_{\mathcal{T}_k}}(u + t - \mathcal{T}_k, \mathbf{X}_{\mathcal{T}_k})) du\right). \end{aligned}$$

The first equality is due to $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$ and so $\mathbb{P}(\{\mathcal{T}_{k+1} > t\}) = 1$. The third equality is due to $\{\mathcal{T}_{k+1} - \mathcal{T}_k > t + s - \mathcal{T}_k\} \cap \{\mathcal{T}_{k+1} - \mathcal{T}_k > t - \mathcal{T}_k\} = \{\mathcal{T}_{k+1} - \mathcal{T}_k > t + s - \mathcal{T}_k\}$ and the definition of conditional probability. The fourth is due to equations (2.4.2), (2.4.3).

Since $\nu_t = \nu_{\mathcal{T}_k}$, then it follows from the semigroup property of the flow ϕ :

$$\phi_{\nu_{\mathcal{T}_k}}(u + t - \mathcal{T}_k, \mathbf{X}_{\mathcal{T}_k}) = \phi_{\nu_t}(u, \phi_{\nu_t}(t - \mathcal{T}_k, \mathbf{X}_{\mathcal{T}_k})) = \phi_{\nu_t}(u, \mathbf{X}_t).$$

Therefore, we have:

$$\begin{aligned} \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid \mathcal{F}_t) &= \exp\left(-\int_0^s a_0(\nu_{\mathcal{T}_k}, \phi_{\nu_{\mathcal{T}_k}}(u + t - \mathcal{T}_k, \mathbf{X}_{\mathcal{T}_k})) du\right) \\ &= \exp\left(-\int_0^s a_0(\nu_t, \phi_{\nu_t}(u, \mathbf{X}_t)) du\right) \\ &= \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid (\nu_t, \mathbf{X}_t)). \end{aligned} \quad (2.4.4)$$

Moreover, by construction of the process (and due to $\nu_t = \nu_{\mathcal{T}_k}$) we have for $C \in \mathcal{E}$:

$$\begin{aligned} \mathbb{P}((\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid \mathcal{F}_t, \mathcal{T}_{k+1}) &= Q(C, \phi_{\nu_t}(\mathcal{T}_{k+1} - t, \mathbf{X}_t)) \\ &= \mathbb{P}((\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid (\nu_t, \mathbf{X}_t), \mathcal{T}_{k+1}), \end{aligned} \quad (2.4.5)$$

since the new state is determined by the information given at the jump time. Thus, due to equations (2.4.4), (2.4.5) and by the definition of conditional probability we have for $s > 0$ and $C \in \mathcal{E}$

$$\begin{aligned}
& \mathbb{P}(\mathcal{T}_{k+1} > t + s, (\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid \mathcal{F}_t) \\
&= \mathbb{P}((\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid \mathcal{F}_t, \mathcal{T}_{k+1}) \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid \mathcal{F}_t) \\
&= \mathbb{P}((\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid (\nu_t, \mathbf{X}_t), \mathcal{T}_{k+1}) \mathbb{P}(\mathcal{T}_{k+1} > t + s \mid (\nu_t, \mathbf{X}_t)) \\
&= \mathbb{P}(\mathcal{T}_{k+1} > t + s, (\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \in C \mid (\nu_t, \mathbf{X}_t)), \tag{2.4.6}
\end{aligned}$$

i.e. the future value of the process depends only on the current value and not on the history of the process. Therefore, $(\nu_t, \mathbf{X}_t)_{t \geq 0}$ is a homogeneous Markov process. \blacksquare

Remark. The process is in fact a strong Markov process, i.e. the Markov property holds for all stopping times, and not only for fixed ones. See [Theorem 25.5, Chapter 2 in [28]].

2.4.2 Cell motility and PDMP

In this section we show that the cyclical cell motility model described in Section 2.3.2 is a well-defined PDMP.

One can show that $\mathbf{F}_i(\mathbf{x}_n, \theta)$ satisfies the Lipschitz condition for $(\mathbf{x}_n, \theta) \in \Omega_{cell} \cup [0, 2\pi) := D_{cell}^1$. Furthermore, one can show that $\beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y})$ and $\mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\varphi}(\mathbf{x}_n)$, given by

$$\begin{aligned}
\beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) &= -\beta_{cell}^{-1} \sum_{i=1}^M Y_i \mathbf{F}_i(\mathbf{x}_n, \theta) \\
\mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\varphi}(\mathbf{x}_n) &= -\mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \sum_{i=1}^M Y_i \mathbf{F}_i(\mathbf{x}_n, \theta) \cdot \hat{\varphi}(\mathbf{x}_n)
\end{aligned}$$

also satisfy the Lipschitz condition for $(\mathbf{x}_n, \theta) \in D_{cell}$ and arbitrary $\mu \in \{0, 1\}$, $\mathbf{Y} \in \{0, 1\}^M$.

Proposition 2.4.2. *Let $\mathbf{x}(0) = \mathbf{x}_0$, $(\mathbf{x}_n(0), \theta(0)) \in D_{cell}$. Let $\mu \in \{0, 1\}$, $\mathbf{Y} \in \{0, 1\}^M$. Then there exists a unique solution of the system*

$$\begin{aligned}
\dot{\mathbf{x}} &= \mu \beta_{ECM}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \hat{\mathbf{r}}(\mathbf{x}_n) \\
\dot{\mathbf{x}}_n &= \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \\
\dot{\theta} &= \mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\varphi}(\mathbf{x}_n), \tag{2.4.7}
\end{aligned}$$

for $t > 0$.

¹The restriction to the interval $[0, 2\pi)$ is due to the periodic dependence on θ in the definition of \mathbf{x}_i . See Section 2.1

Proof. Note that since the evolution of \mathbf{x} is decoupled from the other two equations, it is sufficient to prove the claim for the following subsystem:

$$\begin{aligned}\dot{\mathbf{x}}_n &= \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \\ \dot{\theta} &= \mu \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\varphi}(\mathbf{x}_n)\end{aligned}\quad (2.4.8)$$

Since the right hand side of this system is Lipschitz on D_{cell} and $(\mathbf{x}_n(0), \theta(0)) \in D_{cell}$, then there exists a unique solution of the subsystem (2.4.8) for time $t \leq t_{D_{cell}}$, where $t_{D_{cell}} = \inf \{t^* > 0 \mid \mathbf{x}_n(t^*) \notin \Omega_{cell}\}$ is the exit time from D_{cell} . By Corollary 2.1.2, we see that $t_{D_{cell}} = \infty$. ■

Let $A := \{1, 2, \dots, 2^{M+1}\}$ and let $\alpha : A \rightarrow \{0, 1\} \times \{0, 1\}^M$ be a bijection. This is simply a mapping such that $\alpha(\nu) = (\mu, \mathbf{Y}) \in \{0, 1\} \times \{0, 1\}^M$ corresponds to a particular cell motion and FA states (recall that the former can either be moving or stationary).

Let $(\mathbf{x}(0), \mathbf{x}_n(0), \theta(0)) \in \Gamma := \mathbb{R}^2 \times \Omega_{cell} \times [0, 2\pi)$ and denote $\mathbf{X}_t = (\mathbf{x}(t), \mathbf{x}_n(t), \theta(t))$. Moreover, let $\mathbf{H}_\nu : \Gamma \rightarrow \mathbb{R}^5$ be such that

$$\mathbf{H}_\nu(\mathbf{X}) := \begin{pmatrix} \alpha_\mu(\nu) \beta_{ECM}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \alpha_{\mathbf{Y}}(\nu)) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \hat{\mathbf{r}}(\mathbf{x}_n) \\ \beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \theta, \alpha_{\mathbf{Y}}(\nu)) \\ \alpha_\mu(\nu) \beta_{rot}^{-1} \|\mathbf{x}_n\| \mathbf{F}(\mathbf{x}_n, \theta, \alpha_{\mathbf{Y}}(\nu)) \cdot \hat{\varphi}(\mathbf{x}_n) \end{pmatrix}, \quad (2.4.9)$$

where $\alpha(\nu) = (\alpha_\mu(\nu), \alpha_{\mathbf{Y}}(\nu)) = (\mu, \mathbf{Y}) \in \{0, 1\} \times \{0, 1\}^M$.

Let the probability $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and state space (E, \mathcal{E}) be defined as in the previous section.

We now specify the objects (I,II,III) described in Section 2.4.1.

I By Proposition 2.4.2 we see that for all $\nu \in A$, there exists a unique global solution to (2.4.1).

II Note that in our case the rate function a_0 is given by (recalling Section 2.3.1)¹:

$$a_0(\nu, \mathbf{X}_t) = a_0(\alpha_{\mathbf{Y}}(\nu), \mathbf{X}_t) = \sum_{j=1}^{2M} a_j(\alpha_{\mathbf{Y}}(\nu), \mathbf{X}_t). \quad (2.4.10)$$

Thus, for the integrability condition to be satisfied, we assume that each probability rate function a_j is integrable along the solution of equation (2.4.1). An exact form of the rates a_j satisfying this condition will be given in the subsequent section. Note that a_0 is nonzero, which follows from (2.3.1).

¹We abuse the notation introduced in Section 2.3.1: $a_j(\alpha_{\mathbf{Y}}(\nu), t) = a_j(\alpha_{\mathbf{Y}}(\nu), \mathbf{X}_t) = a_j(\nu, \mathbf{X}_t)$ for $j = 0, \dots, 2M$.

III In our case, the transition measure $Q(\cdot; (\nu, \boldsymbol{\xi}))$ is given by (recalling Section 2.3):

$$\begin{aligned} Q(\{\eta\} \times d\boldsymbol{\xi}'; (\nu, \boldsymbol{\xi})) &= \delta_{\boldsymbol{\xi}}(d\boldsymbol{\xi}') \sum_{j=1}^M \delta_{\alpha_{\mu}(\eta), 0} \frac{a_j^+(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \boldsymbol{\xi})}{a_0(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \boldsymbol{\xi})} \delta_{\alpha_{\mathbf{Y}}(\eta), 1} \prod_{i \neq j}^M \delta_{\alpha_{\mathbf{Y}}(\eta)_i, \alpha_{\mathbf{Y}}(\nu)_i} \\ &\quad + \delta_{\alpha_{\mu}(\eta), 1} \frac{a_j^-(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \boldsymbol{\xi})}{a_0(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \boldsymbol{\xi})} \delta_{\alpha_{\mathbf{Y}}(\eta), 0} \prod_{i \neq j}^M \delta_{\alpha_{\mathbf{Y}}(\eta)_i, \alpha_{\mathbf{Y}}(\nu)_i}, \end{aligned} \quad (2.4.11)$$

where δ is the Kronecker delta function, $\delta_{\boldsymbol{\xi}}(\cdot)$ is the Dirac measure at $\boldsymbol{\xi}$, $a_j^+ = a_{2j-1}$ and $a_j^- = a_{2j}$ correspond to, respectively, the binding and unbinding probability rates at FA site j , and $\boldsymbol{\alpha}_{\mathbf{Y}}(\cdot)_i$ is the i^{th} component of the vector $\boldsymbol{\alpha}_{\mathbf{Y}}(\cdot)$.

The justification for choosing the functions above stems from our deductions in Section 2.3.1. In particular, the rate function a_0 in (II) is due to (2.3.3): the probability density function of the jump time \mathcal{T}_{k+1} , given that $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$, is given by:

$$\begin{aligned} K_{time}(\cdot | t, \boldsymbol{\alpha}_{\mathbf{Y}}(\nu_t)) &= a_0(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu_t), \mathbf{X}_t) \exp\left(-\int_0^{\cdot} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) ds\right) \\ &= a_0(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu_{\mathcal{T}_k}), \mathbf{X}_t) \exp\left(-\int_0^{\cdot} a_0(\nu_{\mathcal{T}_k}, \phi_{\nu_{\mathcal{T}_k}}(s, \mathbf{X}_t)) ds\right), \end{aligned}$$

which corresponds to the survival function given by (2.4.2).

We now turn our attention to the measure Q in (2.4.11). The components of \mathbf{X}_t do not jump, and vary continuously in time, i.e. if \mathcal{T}_k is the jump time, then $\mathbf{X}_{T_k^-} = \mathbf{X}_{\mathcal{T}_k}$ (see Section 2.3.2), hence the Dirac measure $\delta_{\boldsymbol{\xi}}(\cdot)$ at $\boldsymbol{\xi}$ in (2.4.11). Clearly, such transition of the continuous component \mathbf{X}_t of the PDMP at a jump time is probabilistically independent of the transition of the discrete component ν . Hence we have the product of the Dirac measure with the sum, which is a discrete measure for the transition of the discrete component.

By Proposition 2.3.3, there is only one FA event at a jump time. Hence, for the probability of transition $\nu \rightarrow \eta$ to be nonzero, the vectors of FA states $\boldsymbol{\alpha}_{\mathbf{Y}}(\nu)$ and $\boldsymbol{\alpha}_{\mathbf{Y}}(\eta)$ must differ only by one component. Consider the following example to illustrate the jump mechanism.

Example. Let $M = 4$, $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$ and suppose \mathcal{T}_{k+1} , \mathbf{X}_t are given. Let $\nu, \eta \in A$ be such that $\boldsymbol{\alpha}(\nu) = (\mu_{\nu}, \mathbf{Y}_{\nu})$ and $\boldsymbol{\alpha}(\eta) = (\mu_{\eta}, \mathbf{Y}_{\eta})$, where

$$\mathbf{Y}_{\nu} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{Y}_{\eta} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Then, by equation (2.3.3), the probability of the transition $\nu \rightarrow \eta$ is given by:

$$\begin{aligned} \delta_{\mu\eta,0} K_{index}(3|\mathcal{T}_{k+1} - t, t, \mathbf{Y}_\nu) &= \delta_{\mu\eta,0} \frac{a_3(\boldsymbol{\alpha}_Y(\nu), \phi_\nu(\mathcal{T}_{k+1} - t, \mathbf{X}_t))}{a_0(\boldsymbol{\alpha}_Y(\nu), \phi_\nu(\mathcal{T}_{k+1} - t, \mathbf{X}_t))} \\ &= \delta_{\alpha_\mu(\eta),0} \frac{a_3(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})}{a_0(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})} \\ &= \delta_{\alpha_\mu(\eta),0} \frac{a_2^+(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})}{a_0(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})}. \end{aligned}$$

Clearly, the transition $\mathbf{Y}_\nu \rightarrow \mathbf{Y}_\eta$ corresponds to the binding event at FA site 2, explaining the Kronecker delta term (see Sections 2.2.3, 2.3.2). Now, consider the sum in (2.4.11) for this example. We see that

$$\prod_{i \neq j}^M \delta_{\alpha_Y(\eta)_i, \alpha_Y(\nu)_i} \neq 0, \text{ for } j = 2 \text{ only.}$$

We therefore obtain

$$Q(\{\eta\} \times d\boldsymbol{\xi}'; (\nu, \mathbf{X}_{\mathcal{T}_{k+1}})) = \delta_{\mathbf{X}_{\mathcal{T}_{k+1}}} (d\boldsymbol{\xi}') \delta_{\alpha_\mu(\eta),0} \frac{a_2^+(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})}{a_0(\boldsymbol{\alpha}_Y(\nu), \mathbf{X}_{\mathcal{T}_{k+1}})} \delta_{\alpha_Y(\eta)_{2,1}}.$$

Note that if at time t the vector of FA states is given by \mathbf{Y}_ν , then there are M possible FA state vectors into which a transition can occur with nonzero probability:

$$\left\{ \begin{array}{l} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \end{array} \right\}.$$

Note that the new state $\eta \neq \nu$ and the probability that the discrete component of the process jumps to an identical state is null.

Similarly as with the rate function a_0 , we can derive equation (2.4.11) from the principles we established before.

Proposition 2.4.3. *The transition probability measure $Q(\cdot, (\nu, \boldsymbol{\xi}))$ is given by equation (2.4.11) for each $(\nu, \boldsymbol{\xi}) \in E$.*

Proof. Let $(\nu, \boldsymbol{\xi}) \in E$, $\{\eta\} \times d\boldsymbol{\xi}' \in \mathcal{E}$. Let $(N, \boldsymbol{\Xi})$ and $(N_-, \boldsymbol{\Xi}_-)$ be E -valued random variables before and after the jumps. Then,

$$\begin{aligned} Q(\{\eta\} \times d\boldsymbol{\xi}'; (\nu, \boldsymbol{\xi})) &= \mathbb{P}((N, \boldsymbol{\Xi}) \in \{\eta\} \times d\boldsymbol{\xi}' | (N_-, \boldsymbol{\Xi}_-) = (\nu, \boldsymbol{\xi})) \\ &= \mathbb{P}(\{\eta\} \times d\boldsymbol{\xi}' | (\nu, \boldsymbol{\xi})), \end{aligned}$$

where we omitted the random variables for notational convenience. Then we have:

$$\mathbb{P}(\{\eta\} \times d\boldsymbol{\xi}' | (\nu, \boldsymbol{\xi})) = \mathbb{P}(d\boldsymbol{\xi}' | \{\eta\}, (\nu, \boldsymbol{\xi})) \mathbb{P}(\{\eta\} | (\nu, \boldsymbol{\xi})) = \delta_\xi(d\boldsymbol{\xi}') \mathbb{P}(\eta | (\nu, \boldsymbol{\xi})),$$

since $\Xi = \Xi_-$ a.s., by construction of the process. Since α is a bijection, we have¹

$$\begin{aligned}\mathbb{P}(\eta | (\nu, \xi)) &= \mathbb{P}((\alpha_\mu(\eta), \alpha_Y(\eta)) | (\alpha_\mu(\nu), \alpha_Y(\nu)), \xi) \\ &= \mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | \alpha_Y(\nu), \xi),\end{aligned}$$

since, by construction of the cell motility process, the new FA state $\alpha_Y(\eta)$ is determined independently of whether a cell was moving or not (represented by $\alpha_\mu(\nu) \in \{0, 1\}$) and the new motility state $\alpha_\mu(\eta)$ is determined only by which FA event took place (binding or unbinding), regardless of whether a cell was previously moving or not.

Note that when a jump occurs, then, by Proposition 2.3.3, one and only one of $j = 1, \dots, 2M$ possible (binding and unbinding) FA events occurs. Thus, for $j, j' \in \{1, \dots, 2M\}$ and $j \neq j'$ the events “reaction j occurs” and “reaction j' occurs” are mutually exclusive. We then have, by using the definition of conditional probability:

$$\mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | \alpha_Y(\nu), \xi) = \sum_{j=1}^{2M} \mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | j, \alpha_Y(\nu), \xi) \mathbb{P}(j | \alpha_Y(\nu), \xi),$$

where

- $\mathbb{P}(j | \alpha_Y(\nu), \xi)$ is the probability that the FA event j occurs, given $\alpha_Y(\nu)$ and ξ .
- $\mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | j, \alpha_Y(\nu), \xi)$ is the probability of a jump into cell state $(\alpha_\mu(\eta), \alpha_Y(\eta))$, given $\alpha_Y(\nu)$ and ξ , and that the FA event j occurred.

Let $j \in \{1, \dots, M\}$ and $j^+ = 2j - 1$ and $j^- = 2j$. Due to (2.3.3) we have:

$$\mathbb{P}(j^\pm | \alpha_Y(\nu), \xi) = \frac{a_j^\pm(\alpha_Y(\nu), \xi)}{a_0(\alpha_Y(\nu), \xi)}. \quad (2.4.12)$$

Furthermore,

$$\begin{aligned}\mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | j^+, \alpha_Y(\nu), \xi) &= \delta_{\alpha_\mu(\eta), 0} \delta_{\alpha_Y(\eta) j, 1} \prod_{i \neq j}^M \delta_{\alpha_Y(\eta) i, \alpha_Y(\nu) i} \\ \mathbb{P}(\alpha_\mu(\eta), \alpha_Y(\eta) | j^-, \alpha_Y(\nu), \xi) &= \delta_{\alpha_\mu(\eta), 1} \delta_{\alpha_Y(\eta) j, 0} \prod_{i \neq j}^M \delta_{\alpha_Y(\eta) i, \alpha_Y(\nu) i}.\end{aligned}$$

¹If $(H, \mathbf{Z}) \in E$ is a random variable, then, due to α being a bijection:

$$\{\omega \in \Omega | H(\omega) = \eta(\omega)\} = \{H = \eta\} = \{\alpha^{-1}(H) = \alpha^{-1}(\eta)\} = \{\alpha(H) = \alpha(\eta)\}.$$

Therefore,

$$\begin{aligned}
& \sum_{j=1}^{2M} \mathbb{P}(\alpha_\mu(\eta), \boldsymbol{\alpha}_Y(\eta) | j, \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \mathbb{P}(j | \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \\
&= \sum_{j=1}^M \mathbb{P}(\alpha_\mu(\eta), \boldsymbol{\alpha}_Y(\eta) | j^+, \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \mathbb{P}(j^+ | \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \\
&\quad + \mathbb{P}(\alpha_\mu(\eta), \boldsymbol{\alpha}_Y(\eta) | j^-, \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \mathbb{P}(j^- | \boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi}) \\
&= \sum_{j=1}^M \delta_{\alpha_\mu(\eta), 0} \frac{a_j^+(\boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi})}{a_0(\boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi})} \delta_{\alpha_Y(\eta)_j, 1} \prod_{i \neq j}^M \delta_{\alpha_Y(\eta)_i, \alpha_Y(\nu)_i} \\
&\quad + \delta_{\alpha_\mu(\eta), 1} \frac{a_j^-(\boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi})}{a_0(\boldsymbol{\alpha}_Y(\nu), \boldsymbol{\xi})} \delta_{\alpha_Y(\eta)_j, 0} \prod_{i \neq j}^M \delta_{\alpha_Y(\eta)_i, \alpha_Y(\nu)_i}.
\end{aligned}$$

■

Chapter 3

Adhesion, Contraction, and Signaling

In this chapter we discuss the cellular structures (FAs and SFs) involved in the cell motility model. We also describe how the signaling feedback of RhoGTPases plays a role in regulating FA and SF dynamics, and elaborate on myosin motor dynamics.

3.1 Focal adhesions

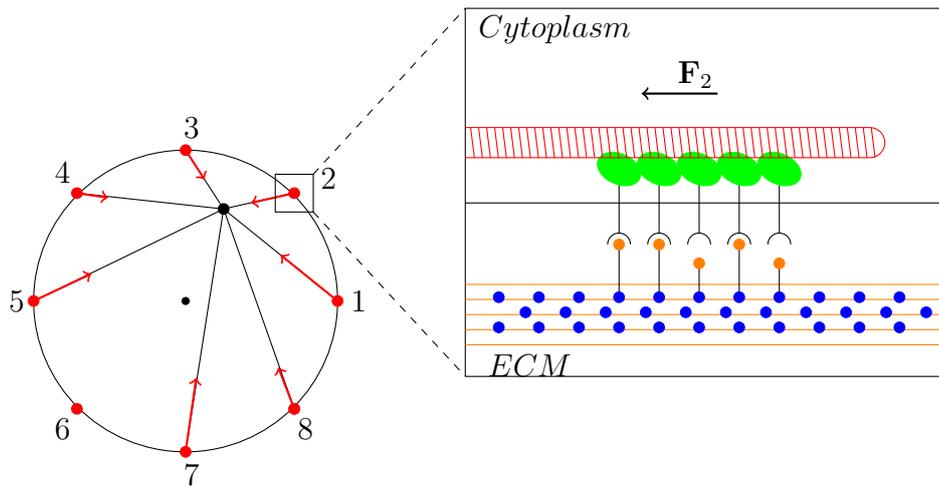


Figure 3.1: Schematic representation of an FA. SF is represented in red, adapter complexes linking integrins with the SF in green. The middle and the rightmost bonds are open, the rest are closed. Soluble cues are represented in blue.

A simplified schematic of an FA is shown in Figure 3.1. As mentioned in the introduction, FA is a multiprotein cluster consisting of transmembrane receptors (integrins), and numerous other regulatory and adapter protein complexes, which link the receptors with the cytoskeleton [15], [43]. An FA provides mechanical and chemical feedback between a cell and its surrounding environment, which gives rise to various types of tactic migration. For example, durotaxis and haptotaxis occurs in response to a perceived gradient in the

ECM stiffness and density, while chemotaxis in response to a chemical one. Regardless of the kind of (un)biased motility, in the migration cycle FA disassociates from the ECM due to mechanical causes (applied tension by contractile fibers), while association is both mechanically and chemically regulated.

While we introduced the probability rates of binding and unbinding events, we have not yet fully specified them. Here, we provide functional forms of propensity functions a_j^+ and a_j^- , and also provide the associated parameter estimates.

This section is based on the work by the author in [120].

3.1.1 Unbinding rate

Consider the unbinding rate a_j^- of FA adhesion site $j \in \{1, \dots, M\}$ and let $\mathbf{y} \in \{0, 1\}^M$, $\boldsymbol{\xi} = (\mathbf{x}, \mathbf{x}_n, \theta) \in \Gamma$. Following Bell [12], the bond disassociation rate under applied force is given by:

$$a_j^-(\mathbf{y}, \boldsymbol{\xi}) = k_{off}^0 e^{\|\mathbf{F}_j(\mathbf{x}_n, \theta)\|/F_b} y_j, \quad (3.1.1)$$

where k_{off}^0 is the FA disassociation rate under no load, \mathbf{F}_i is the force applied at the FA, given by equation (2.1.1), and $F_b = 5.5nN$ is a characteristic force scale. The last factor y_j simply indicates that only bound FAs can unbind (thus satisfying equation (2.3.1)). Clearly, the function in (3.1.1) is integrable. Here we neglect the fact that the force may be applied at the FA (and consequently at the transmembrane receptors) at an angle and assume for tractability of the model that it is applied normally to the FA (hence consider only magnitude).

Remark. In the context of cell migration and within the framework of our model, we only consider FA disassembly on the cell periphery (including the lamellae). The primary cause of such FA unbinding has mechanical, rather than biochemical nature due to the cell contractile mechanism applying load to the adhesions. Although it is known that the Rho family of GTPases (in particular its member RhoA) mediates disassembly of FAs, their effect is indirect. For example the activity of myosin motors, which generate contractile forces in SFs, is regulated by RhoA [96], hence the force dependence of the unbinding rate a_j^- . Recalling the definition of \mathbf{F}_i in equation (2.1.1), we can include such indirect biochemical mediation by considering mediators of the force T_i . In order to keep the model tractable, here we omit the interaction between RhoA and myosin motors (see Section 3.3 for details on SF contractility).

Note that force is required for formation of FAs and in its absence they disassociate [41], [114], [130]. It is thought that this is due to the adapter complexes undergoing conformational changes, which exposes cryptic binding sites and promotes assembly through activation of protein recruiting signaling pathways. Thus, disassociation of adapter complexes under no force implies that the cluster becomes a collection of independent integrin bonds. The average lifetime T_{life} of such a cluster is given by [34]:

$$T_{life} = \frac{1}{k_0^1 + k_{on}^1} \left[H_N + \sum_{n=1}^N \left(\frac{k_{on}^1}{k_0^1} \right)^n \frac{1}{n} \binom{N}{n} \right],$$

where N is the number of bonds in a cluster, H_N is the N th harmonic number, k_0^1 and k_{on}^1 are, respectively, unbinding and binding rates of integrins under no load. Note that in the absence of a load, (re-)binding of individual integrins is an independent event, which bears no relation to the FA, since tension is required for an FA to form and sustain itself. We then estimate:

$$k_{off}^0 = 1/T_{life} |_{k_{on}^1=0} = k_0^1/H_N.$$

Li et al. [66] found that $k_0^1 = 0.012s^{-1}$ for $\alpha_5\beta_1$ -integrin binding to fibronectin. For $N = 10^3 - 10^4$ we estimate that $k_{off}^0 = 0.05s^{-1} - 0.1s^{-1}$. Balaban et. al. [9] found that focal adhesions apply a constant stress of $\sim 5.5nN/\mu m^2$ over an area of $\sim 1\mu m^2$ on an elastic substrate. Thus, we take $F_b = 5.5nN$.

Recall that FA disassociation leads to cell body translocation. It follows that the rate of FA unbinding determines the time scale of cell migration t_e (see Section 2.2.3) and hence $t_e = k_{off}^0$.

3.1.2 Binding dynamics

Consider the binding probability rate a_j^+ of the FA adhesion site $j \in \{1, \dots, M\}$ and let $\mathbf{y} \in \{0, 1\}^M$, $\boldsymbol{\xi} = (\mathbf{x}, \mathbf{x}_n, \theta) \in \Gamma$. The probability rate a_j^+ is given by:

$$a_j^+(\mathbf{y}, \boldsymbol{\xi}) = k_{on,j}(\boldsymbol{\xi})(1 - y_j),$$

where $k_{on,j} : \Gamma \rightarrow \mathbb{R}^+$ is the effective binding rate at FA site j . The last term $(1 - y_j)$ simply indicates that only unbound FAs can bind. Whereas unbinding can be viewed effectively as a bond rupturing under applied tension, a binding reaction, or focal adhesion assembly and maturation, is a highly regulated process. Due to the complexity of the FA assembly process, we focus on two major mediators of FA formation: Rac1 activity and contractile forces.

Rac1 dependence

For simplicity, we assume that the probability of FA formation is directly proportional to local Rac1 concentration. Consider the case of chemotactic cell migration. Leading edge protrusions preferentially form in the direction of a chemoattractant. Since Rac1 is required for formation of lamellipodium and FA formation [96], then local Rac1 activity correlates with the concentration of the chemical cues. Conversely, local Rac1 activity negatively correlates with chemorepellent. Moreover, Rac1 itself is activated by integrin-ECM bonds [93], suggesting that there is a positive feedback loop.

Let $Q_{cue} : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ denote the concentration of a cue in the spatial domain and let $q : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ denote the Q_{cue} dependent concentration of Rac1. Clearly, q is an increasing function for the case of an attractant and a decreasing function for a repellent. Then,

$$k_{on,j}(\boldsymbol{\xi}) \propto q(Q_{cue}(\mathbf{x} + \mathbf{x}_j)), \quad (3.1.2)$$

where we recall that \mathbf{x}_j is the position of j^{th} FA site.

For example, we can take $Q_{cue}(\mathbf{x})$ to be the density of a chemoattractant at $\mathbf{x} \in \mathbb{R}^2$ and take $q(x) = x$. Then, the probability of binding is simply proportional to the chemoattractant density¹. If we let Q_{cue} to denote the density of the ECM and take $q(x) = x$, then the binding rate is proportional to the local density of the ECM, which corresponds to the haptotactic migration. Thus, both chemo- and haptotaxis can be taken into account. The simultaneous effect of chemo- and haptotactic cues on the binding probability can be considered by modifying the form of q . For example, if we let $q(x, y) = w_{chemo}x + w_{hapto}y$, where x, y are the densities of a chemoattractant and the ECM, respectively, then, by varying the weights w_{chemo} , w_{hapto} , we can examine the effects of both cues on the cell migration.

Force dependence

Note that the enlargement of nascent adhesions² is concurrent with their maturation into focal adhesions. Thus, enlargement and maturation are synonymous. While the initial stage of adhesion growth is force-independent [23], maturation occurs in a force-dependent manner [41], [114], [130].

However, during such a force-dependent maturation, the positive correlation between the adhesion size and the applied tension exists only in the initial stages of maturation. As FAs increase in size, the force dependence plateaus [114].

That is, the study by Stricker et. al. [114] showed that for mature FAs there is no correlation between applied force and FA size. One can consider an adhesion site as mature when its size reaches $\sim 1\mu m^2$ (see e.g. [9], [108]).

Choi et. al. [23] showed that nascent adhesions assemble at a rate of $\sim 1.3\text{min}^{-1} = 0.021\text{s}^{-1}$ reaching a size of $\sim 0.2\mu m^2$. Furthermore, it was shown that the formation of these adhesions is independent of fibronectin density [23], stiffness [23] and myosin II activity [4],[23],[125].

Let k_{on}^0 be the force-independent FA maturation rate, which we take $k_{on}^0 = 0.01\text{s}^{-1}$. We now want to find a function that could represent force dependence of FA maturation rate. Denote this function $k_{force} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$.

It satisfies the following:

- $k_{force}(0) = k_{on}^0$, i.e. if there is no force applied, the rate is k_{on}^0 .

¹This corresponds to a bi-molecular reaction rate, which depends on the *number* of one of the reactants (FA) and on the *concentration* (chemoattractant) of the other.

²The enlargement of nascent adhesions occurs by clustering of substrate bound integrins and sequestering of the adapter proteins.

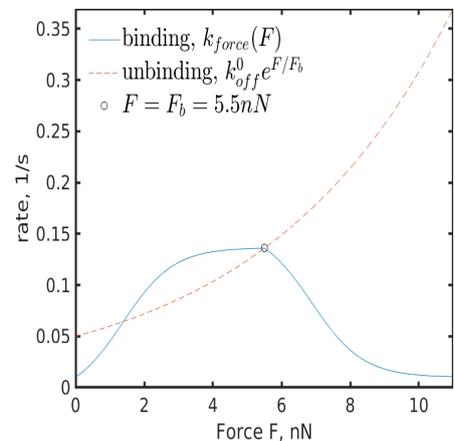


Figure 3.2: Force dependence of unbinding and binding rates.

- If the applied force is below the characteristic force F_b , then k_{force} is greater than the unbinding rate, i.e. it is more likely that an FA increases in size than that it ruptures.
- If the characteristic force F_b is applied, the rate k_{force} should equal the unbinding rate, i.e. we assume that there is a dynamic equilibrium in some sense.
- If the applied force is larger than F_b , then the unbinding rate dominates the binding rate. Note that as FA increases in size, the force dependence plateaus [114]. Thus, k_{force} should plateau around F_b . We also assume that for large applied forces k_{force} plateaus at k_{on}^0 , since exceeding loads rupture integrin bonds frequently and impede stable maturation.

The following form of k_{force} satisfies the conditions above:

$$k_{force}(F) = \begin{cases} \frac{k_{off}^0 e + k_{on}^0}{1 + \exp(-\gamma_1(F - F_1^*)/F_b)} + k_{on}^0 - \epsilon, & F \leq F_b \\ \frac{k_{off}^0 e + k_{on}^0}{1 + \exp(\gamma_2(F - F_2^*)/F_b)} + k_{on}^0, & \text{else} \end{cases}, \quad (3.1.3)$$

where $F_1^* = F_b/4$ and $F_2^* = 5F_b/4$ are the midpoints of the sigmoid functions (see Figure 3.2). The values of γ_1, γ_2 and ϵ can be found as follows:

Suppose $F \leq F_b$. Then,

$$\gamma_1 = -\frac{F_b}{F - F_1^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{k_{force}(F) - k_{on}^0 + \epsilon} - 1 \right).$$

Since $k_{force}(0) = k_{on}^0$ and $k_{force}(F_b) = k_{off}^0 e$, then:

$$\begin{aligned} \gamma_1 &= \frac{F_b}{F_1^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{\epsilon} - 1 \right) \\ \gamma_1 &= -\frac{F_b}{F_b - F_1^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{k_{off}^0 e - k_{on}^0 + \epsilon} - 1 \right). \end{aligned}$$

It follows that ϵ is given as the solution of the following equation:

$$\frac{F_b}{F_1^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{\epsilon} - 1 \right) + \frac{F_b}{F_b - F_1^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{k_{off}^0 e - k_{on}^0 + \epsilon} - 1 \right) = 0.$$

Similarly, since $k_{force}(F_b) = k_{off}^0 e$, we find:

$$\gamma_2 = \frac{F_b}{F_b - F_2^*} \log \left(\frac{k_{off}^0 e + k_{on}^0}{k_{off}^0 e - k_{on}^0} - 1 \right).$$

The values of γ_1, γ_2 , and ϵ are fixed for a value of F_b given in Section 4.2.

Remark. Chan and Odde [19] showed that adhesion site dynamics depends on substrate stiffness. In particular, they showed that for a stiff substrate the transmembrane bonds rupture more frequently, compared to the case with softer substrate under the same load, since the critical load is reached faster. This mechanism provides means for a cell to assess the surrounding rheology. Within the context of our model, this means that the force F_b is smaller for the stiffer substrate, thus increasing the disassociation rate for the same load (see (3.1.1)). Consequently, the force dependent binding rate k_{force} also changes for the stiffer ECM. In this case, the curves in Figure 3.2 will shift to the left. Therefore, our model provides an opportunity to include mechanosensitivity of migrating cells by considering the relevant dynamics for individual FAs.

Therefore, the binding propensity rate a_j^+ of an adhesion $j \in \{1, \dots, M\}$ is given by:

$$a_j^+(\mathbf{y}, \boldsymbol{\xi}) = q(Q_{cue}(\mathbf{x} + \mathbf{x}_j))k_{force}(\|\mathbf{F}_j(\mathbf{x}_n, \theta)\|)(1 - y_j). \quad (3.1.4)$$

3.2 RhoGTPases

Although the activity of Rac1 and RhoA regulates FA assembly and SF contractility, in this work we do not consider their effects on cell migration in detail. However, in this section we briefly describe how the effects of Rac1-RhoA feedback can be incorporated in a simplified manner into our model. The following is based on the mathematical models of RhoGTPase signaling feedback in [50]. There, the spatially dependent dynamics and interaction of RhoGTPase members (Rac1, Cdc42, and RhoA) with membrane lipids (phosphoinositides PIP1, PIP2, PIP3) were studied. Similar models of RhoGTPases crosstalk were also used in [59] and [73] to model single cell migration and contact inhibition of locomotion.

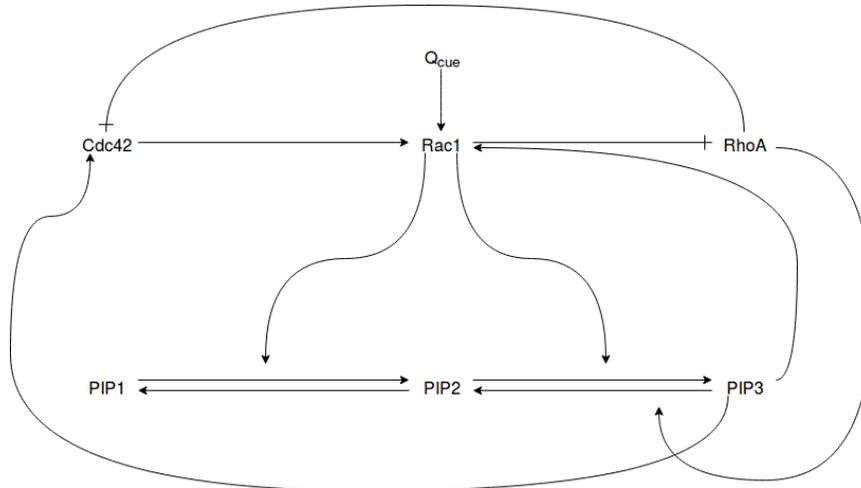


Figure 3.3: Schematic diagram of the interactions between RhoGTPases and PIs considered in [50].

Schematic diagram of the feedback loops between RhoGTPases and phosphoinositides (PIs) corresponding to Model 4 in [50] is shown in Figure 3.3. It follows from the results

there that in steady state conditions, and in the absence of diffusion and feedback with PIs, we have:

$$\begin{aligned} R &= \bar{R} \frac{\alpha}{\delta_R} C + \bar{R}(\hat{I}_R + Q_{cue}) \\ \rho &= \frac{\bar{\rho}}{\delta_\rho} \frac{\hat{I}_\rho}{1 + \left(\frac{R}{a_2}\right)^n} \\ C &= \frac{\bar{C}}{\delta_C} \frac{\hat{I}_C}{1 + \left(\frac{R}{a_1}\right)^n}, \end{aligned} \quad (3.2.1)$$

where R, ρ, C denote the concentrations of Rac1, RhoA, Cdc42, and the rest are constant parameters. Here, the notation is the same as in [50], except that we use Q_{cue} to denote the concentration of a cue. Notice that Rac1 depends linearly on Q_{cue} , which could justify taking $q(x) = x$ in (3.1.2) and letting the proportionality constant to absorb the other parameters in the above equation for Rac1. Moreover, we also see that the (local) concentration of Rac1 depends on Q_{cue} . Therefore, in (3.1.2) we can take into account the effect of an external cue on Rac1 mediated FA association rate, albeit in a simple, linear manner. From (3.2.1), we can also see how an external cue can affect RhoA concentrations, due to the inhibition of the latter by Rac1. As the activation of myosin motors is mediated by RhoA [124], the above considerations provide means by which the RhoGTPase dynamics can be included in an extended model. In the following sections, we briefly describe the mechanism of SF contraction by myosin motors and the effects of RhoA thereon.

3.3 Stress fibers

As mentioned in Section 2.1, we treat SFs as “active” contractile elements, whereby their mechanical properties are influenced by force generating non-muscle myosin IIA molecular motors. They are organized in minifilaments of 10-30 motors [122], where the group is able to apply a contractile force only when there is a distance between the filaments (see [83] and Figure 2.3b). As a motor produces a power stroke (see Figure 3.4), it slides the filamentous actin (F-actin), contracting the fiber and reducing the separation between the minifilaments. If the distance between them vanishes, a power stroke cannot slide the F-actin as the front of the motor is jammed by a crosslinking protein or another minifilament.

The length of a myosin mini-filament is $\sim 0.3\mu m$ [115] and the interfilament distance is $\sim 1\mu m$ in an uncontracted fiber [5]. Thus, the proportion of the minifilaments to the initial, uncontracted SF length is $\frac{0.3}{1+0.3} = 0.23$. Assuming vanishing interfilament distance, it follows that an SF can contract up to 23%. Interestingly, Deguchi et. al. [29] found that stress fibers contract, on average, to 20% of their original length. As stress fibers generally span more than half of a cell, and since it was found that there is a preexisting strain [29] in them, we estimate $L_0 = 1.1R_{cell}$ and $L_c = 0.2R_{cell}$ in (2.1.1).

If we take $R_{cell} = 25\mu m$, and assuming an SF has at rest the length of R_{cell} , we can estimate the number of myosin minifilaments in a fiber to be $25\mu m/1.3\mu m \approx 20$. As

each motor produces a force of $2 - 10pN$ [38], [81], [83], and there are 10-30 motors in minifilament, we then estimate the (constant) tension $T_i = 4nN$ in i^{th} SF due to myosin motors. As the characteristic force at an FA is $F_b = 5.5nN$ (Section 3.1.1), and assuming a preexisting strain in an SF was 0.1 when this measurement was taken, and since $T_i = 4nN$, we then estimate the one dimensional Young's modulus $EA = 15nN$. We obtain this estimate from the first case in (2.1.1):

$$F_b = T_i + EA \frac{L_i - L_0}{L_0}.$$

In the discussion of these macroscopic (in relation to its constituents) mechanical properties of contractile SFs, we did not take into account their regulators. In particular, how do RhoGTPases come into play? In the following, we model mediation of myosin activity by RhoA.

3.4 Myosin motors

A single non-muscle myosin II is a non-processive motor with a low duty ratio (fraction of time spent on an actin filament). However, organization in mini-filaments increases the duty ration significantly, generating appreciable contractile forces [3], [57]. In order to address the cooperative behavior, cross-bridge models were studied in [3], [128] (see Figure 3.4). In [3] it was also found that organization in an ensemble reduces stochastic fluctuations in the total amount of force produced, which may explain why contraction of an SF occurs smoothly.

According to our estimates in the previous section, there are $\sim 10^3$ motors applying force on an FA. Thus, due to the number and size of the motors, stochastic effects may be prominent. Here, we are rather interested in aggregate quantities, such as the effective number of active motors producing tension T_i , and how RhoA mediates it. We reduce the complexity of the model by Albert et al. [3] and recast it as a chemical, rather than mechanochemical, stochastic system with a substrate, enzyme complex and a product (Figure 3.5). Furthermore, we do not take into account the cooperative behavior of small ensembles considered in [3], which arises, in part, due to dependence of the transition rates on mechanical factors. In some sense, we are investigating the dynamics on a larger scale (thousands vs. tens of motors), by considering an individual motor, bypassing the intermediate scale. The parameter values used here were introduced in [126].

Regulation of the cross-bridge cycle by RhoA occurs indirectly through binding ROCK (Rho-associated coil kinase) and subsequently activating it [36]. Since RhoA binds ROCK, then, assuming abundance of ROCK, the numbers of (active) RhoA and ROCK molecules are equal. Thus, we consider regulation of the cycle by ROCK, and use it as a proxy to infer the influence of RhoA on motor activity. Active ROCK phosphorylates myosin light chains both directly and indirectly [124]. Only with phosphorylated light chain, the myosin head is able to bind to actin, allowing transition into the wb state (Figure 3.4, 3.5). Hence, the transition of ub into the wb state can be treated as binding, thus seeing

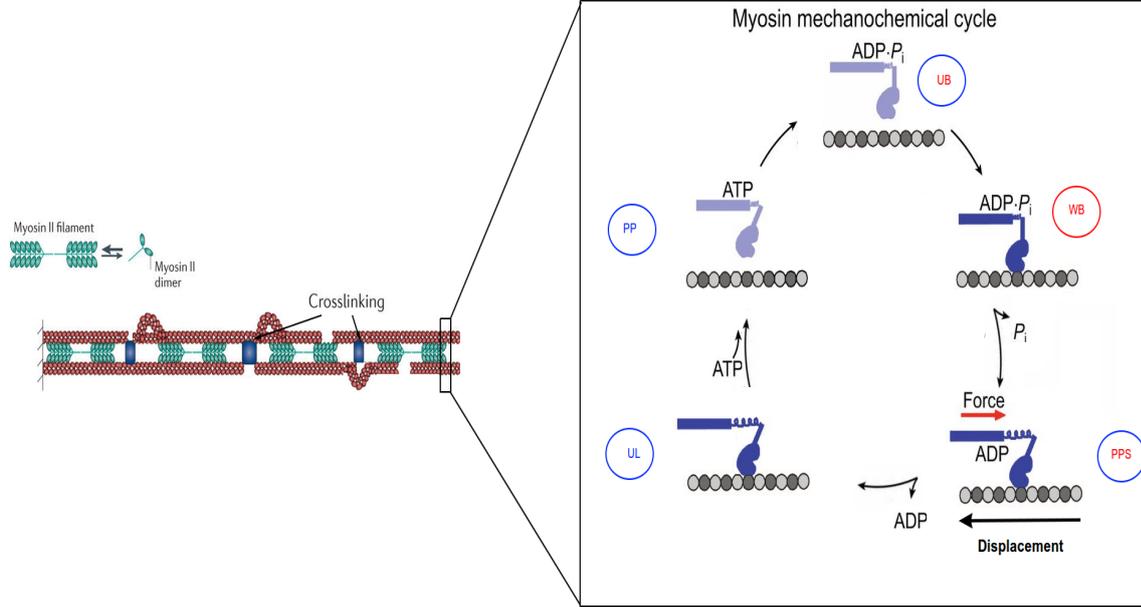


Figure 3.4: Schematic representation of the myosin motor cross-bridge cycle (right panel). The mechanochemical model of four states in blue is developed by Walcott et al. [128], while the model with three states in red is developed by Albert et al. [3]. In the unbound (primed) state ub , the myosin head is loaded with ADP and P_i . The head binds to actin and transitions into the weakly bound state wb . With release of P_i the motor produces a power stroke and transitions into the post power stroke state pps . After releasing ADP, the motor is in the unloaded state ul . Binding ATP terminates the actin bond and motor transitions into pre-primed state pp . ATP hydrolysis causes conformational change and returns the head to the unbound (primed) state ub . Modified from images in [60] and [83].

it as enzyme complex forming reaction with forward rate $k_f^M = 40s^{-1}$ and reverse rate $k_{ub}^M = 2s^{-1}$. For simplicity, we take k_f^M constant. However, effects of ROCK levels can be included later. Although the forward and reverse rates in the wb to pps transition are equal in the three state cross-bridge model by Albert et al. [3], the reaction is driven towards the pps state with a free energy difference $\Delta E \simeq -60pN \text{ nm}$. Thus, the ratio of the probability of this subsystem to be in the wb state to probability to be in the pps state is $e^{\Delta E/k_B T} = 3 \cdot 10^{-7}$, where k_B is the Boltzmann constant and T represents the room temperature. Hence, we treat the wb to pps transition as a product forming irreversible reaction with rate $k_{cat}^M = 10^3s^{-1}$, where the product is the motor activation. Finally, the active motor in the state pps reverts back to the inactive state ub with the irreversible rate $k_r^M = 80s^{-1}$ (as the recovery stroke and the return to the primed state require ATP hydrolysis).

Let N_T be the (constant) total number of the motors in an SF. Let N and $N_0 := N_T - N$ be the numbers of active (i.e. in the pps state) and the inactive (i.e. in the ub state) motors, respectively. Let E be the number of free ROCK molecules. Since in our model (see Figure 3.5) the ub to wb transition is a bimolecular reaction, then $k_f^M := \hat{k}_f^M E/V$, where \hat{k}_f^M is a

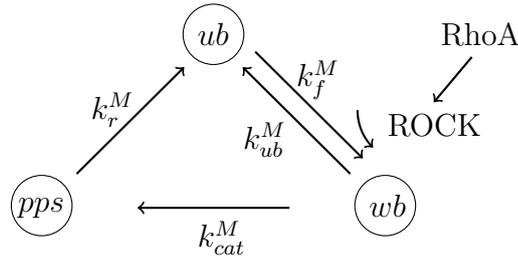


Figure 3.5: Reduced chemical model with enzymatic activity. The substrate ub binds ROCK with the forward rate k_f^M and the reverse rate k_{ub}^M , and forms the complex wb . The complex wb forms the product pps with rate k_{cat}^M and reverts back with rate k_r^M .

constant and V is the volume of the system (see Gillespie [44] for details on the forms of reaction rates). Then, the motor activation reaction chain is the following:



where $[N_0E]$ is an enzyme complex. We then have the following:

- $\hat{k}_f^M N_0E/V$ and $k_{ub}^M[N_0E]$ are the rates of complex formation and disassociation, respectively.
- $k_{cat}^M[N_0E]$ and $k_r^M N$ are the rates of product formation (i.e. myosin activation) and disassociation (i.e. deactivation), respectively.

Since the rate constant of the product formation is significantly larger than other rates considered here, the complex $[N_0E]$ is short-lived. Note that, as expected the analysis of the original three state model [3] also, revealed that the weakly bound state is a short-lived transient state. Therefore, in a quasi-steady-state approximation we have the following:

$$\hat{k}_f^M N_0E/V = k_{ub}^M[N_0E] + k_{cat}^M[N_0E].$$

This is reminiscent of classical, deterministic Michaelis-Menten kinetics approximation of enzyme dynamics. One could view the equation above as a conservation law: the number of molecules “flowing into” the wb state equals the number of molecules “flowing out”.

Let $E_0 := E + [N_0E]$ be the total number of (active) ROCK molecules. Using the equation above, we have:

$$[N_0E] = \frac{\hat{k}_f^M E_0}{V} \cdot \frac{N_0}{k_{ub}^M + k_{cat}^M + \frac{\hat{k}_f^M}{V} N_0}.$$

Remark. Assuming the value of the rate constant k_f^M was deduced with $E = E_0$, and since the Michaelis-Menten kinetics approximation assumes that substrate concentration

is significantly larger than enzyme concentration, then we estimate (with $E_0 \sim 10 - 100$):

$$\hat{k}_f^M / V = k_f^M / E_0 \simeq 40 \cdot 10^{-2} s^{-1} = 40 \cdot 10^{-1} s^{-1},$$

as the number of myosin motors is on the order of $\sim 10^3$ (see above).

Let $b^+, b^- : \{0, \dots, N_T\} \rightarrow \mathbb{R}_+$ be the rates of myosin activation and deactivation, respectively. Then, we have:

$$\begin{aligned} b^+(N) &= k_{cat}^M \frac{\hat{k}_f^M E_0}{V} \cdot \frac{N_T - N}{k_{ub}^M + k_{cat}^M + \frac{\hat{k}_f^M}{V} (N_T - N)} \\ b^-(N) &= k_r^M N. \end{aligned} \tag{3.4.1}$$

Note that the total ROCK concentration is E_0/V . Since it is proportional to the RhoA concentration ρ , we can use (3.2.1) to relate b^+ to ρ . These two rates specify our reduced two-state chemical model.

Steady state approximation

Let $p(t, N)$ be the probability that there are N active motors at time t . Then, p satisfies the following master equation:

$$\begin{aligned} \frac{\partial}{\partial t} p(t, N) &= b^+(N-1)p(t, N-1) + b^-(N+1)p(t, N+1) \\ &\quad - [b^+(N) + b^-(N)] p(t, N), \end{aligned} \tag{3.4.2}$$

where we set

$$\begin{aligned} b^+(-1) &\equiv 0, \quad p(t, -1) \equiv 0, \\ b^-(N+1) &\equiv 0, \quad p(t, N+1) \equiv 0, \end{aligned}$$

so that

$$\begin{aligned} \frac{\partial}{\partial t} p(t, 0) &= b^-(1)p(t, 1) - b^+(0)p(t, 0) \\ \frac{\partial}{\partial t} p(t, N_T) &= b^+(N_T-1)p(t, N_T-1) - b^-(N_T)p(t, N_T), \end{aligned} \tag{3.4.3}$$

which simply corresponds to the reflecting boundary conditions, as we assumed that the total number of motors in an SF is constant.

Following [Chapter VI in [53]], we introduce for convenience a ‘‘step operator’’ denoted by the symbol \mathbb{W} , which we define by its action on $p(t, N)$:

$$\mathbb{W}p(t, N) = p(t, N+1), \quad \mathbb{W}^{-1}p(t, N) = p(t, N-1).$$

It has the following useful properties:

$$\mathbb{W}\mathbb{W}^{-1} = \mathbb{W}^{-1}\mathbb{W} = I \quad (3.4.4)$$

$$\sum_{N=0}^{N_T-1} g(t, N)\mathbb{W}f(t, N) = \sum_{N=1}^{N_T} f(t, N)\mathbb{W}^{-1}g(t, N), \quad (3.4.5)$$

where I is identity operator and f, g are arbitrary.

Remark. Multiplying (3.4.2) by N , summing over the range of N values, and using (3.4.4), we obtain:

$$\begin{aligned} \frac{d}{dt} \langle N \rangle &= \sum_{N=0}^{N_T} N (\mathbb{W}^{-1} - I) b^+(N) p(t, N) + \sum_{N=0}^{N_T} N (\mathbb{W} - I) b^-(N) p(t, N) \\ &= \sum_{N=0}^{N_T} b^+(N) p(t, N) (\mathbb{W} - I) N + \sum_{N=0}^{N_T} b^-(N) p(t, N) (\mathbb{W}^{-1} - I) N \\ &= \langle b^+(N) \rangle - \langle b^-(N) \rangle, \end{aligned}$$

where $\langle \cdot \rangle$ denotes the average. Using Taylor's expansion around $\langle N \rangle$ we have:

$$\begin{aligned} b^\pm(N) &= b^\pm(\langle N \rangle) + (b^\pm)'(\langle N \rangle)(N - \langle N \rangle) \\ &\quad + \frac{1}{2}(b^\pm)''(\langle N \rangle)(N - \langle N \rangle)^2 + \dots, \end{aligned}$$

from which it follows:

$$\langle b^\pm(N) \rangle = b^\pm(\langle N \rangle) + \frac{1}{2}(b^\pm)''(\langle N \rangle)(\langle N^2 \rangle - \langle N \rangle^2) + \dots$$

Neglecting fluctuations around the mean $\langle N \rangle$, we obtain a more familiar ODE:

$$\frac{d}{dt} \langle N \rangle = b^+(\langle N \rangle) - b^-(\langle N \rangle).$$

From this, we can obtain the usual ODE governing Michaelis-Menten enzyme reactions for *concentration*, rather than *number* of reactants. Using (3.4.1) and dividing by V the equation above, we have:

$$\frac{d}{dt} [N] = k_{cat}^M [E_0] \frac{[N_T] - [N]}{\frac{k_{ub}^M + k_{cat}^M}{k_f^M} + [N_T] - [N]} - k_r^M [N],$$

where $[\cdot] = \frac{\langle \cdot \rangle}{V}$ denotes the reactant concentration. This shows that our treatment of the two-state chemical model of myosin activation yields the expected macroscopic law for Michaelis-Menten kinetics (under suitable conditions).

From (3.4.2) it follows that the stationary distribution $p_s(N)$ is given by:

$$\begin{aligned} 0 &= (\mathbb{W} - I) b^-(N) p_s(N) + (\mathbb{W}^{-1} - I) b^+(N) p_s(N) \\ &= (\mathbb{W} - I) [b^-(N) p_s(N) - \mathbb{W}^{-1} b^+(N) p_s(N)]. \end{aligned}$$

Thus, the term in the square brackets is independent of N :

$$b^-(N) p_s(N) - \mathbb{W}^{-1} b^+(N) p_s(N) = \text{Const.}, \quad \forall N \in \{1, \dots, N_T\}.$$

In particular, the equation above holds for $N = 1$. Using (3.4.3) and the definition of \mathbb{W} :

$$0 = b^-(1) p_s(1) - b^+(0) p_s(0) = \text{Const.}$$

Thus, the following holds $\forall N \in \{0, \dots, N_T\}$:

$$b^-(N) p_s(N) = b^+(N-1) p_s(N-1),$$

which is a statement of detailed balance. Applying the relation recursively:

$$p_s(N) = \frac{b^+(N-1) b^+(N-2) \cdots b^+(0)}{b^-(N) b^-(N-1) \cdots b^-(1)} p_s(0), \quad (3.4.6)$$

where $p_s(0)$ can be obtained by noting that p_s is a probability distribution:

$$\begin{aligned} 1 &= \sum_{N=0}^{N_T} p_s(N) = p_s(0) + \sum_{N=1}^{N_T} \frac{b^+(N-1) b^+(N-2) \cdots b^+(0)}{b^-(N) b^-(N-1) \cdots b^-(1)} p_s(0) \\ \Rightarrow \frac{1}{p_s(0)} &= 1 + \sum_{N=1}^{N_T} \frac{b^+(N-1) b^+(N-2) \cdots b^+(0)}{b^-(N) b^-(N-1) \cdots b^-(1)}. \end{aligned}$$

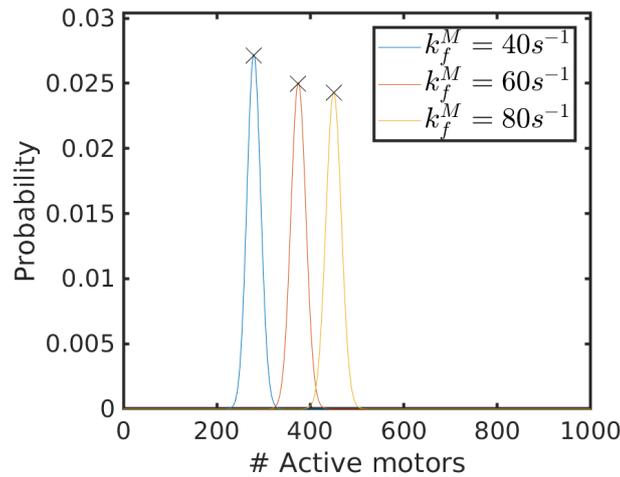


Figure 3.6: Stationary distribution p_s with $N_T = 10^3$ and varying values of k_f^M . The values of $\langle N \rangle$ and $p_s(\langle N \rangle)$ are marked by a black cross.

The distribution p_s is plotted in Figure 3.6, where an increasing value of k_f^M corresponds to elevated concentration of active RhoA. We also see that the distribution is peaked around the average $\langle N \rangle$, given by:

$$\langle N \rangle = \sum_{N=0}^{N_T} N p_s(N).$$

Therefore, given the concentration of RhoA, we can estimate the contractile force exerted by myosin motors T_i in i^{th} SF as $T_i = \langle N \rangle F_{myo}$, where F_{myo} is force applied by a single motor.

We have thus devised means by which the signaling feedback between RhoGTPases can be incorporated in our model, namely, regulation of FA dynamics by Rac, and mediation of SF contractility by RhoA, where the activity of the GTPases are related by (3.2.1).

Chapter 4

Numerical Simulations

In this chapter we present the numerical method used to simulate the model presented in Chapter 2, assess its parameters, and discuss the results of the simulations. Recall that the model contains interdependent stochastic and deterministic components. However, it is not immediately clear how to sample from the corresponding probability distributions, since they depend on the solution of the ODE system describing the deterministic components.

A crucial measure of model validity is its ability to explain a variety of experimental observations. In order to compare them with our simulation results, we must also evolve our system for comparable time scales. For example, to assess such quantities as velocity or average displacement, the time scale of such observations is typically around many hours to allow for significant displacements to occur. However, recall that the time scale of our model is set by the frequency of FA unbinding events, which is typically expressed in minutes. Thus, we must evolve our system for relatively long times, during which its form will change many times, as each FA event modifies the ODE system governing the deterministic components. Moreover, we must also simulate the trajectories of many cells independently in order to obtain a good measure of averaged quantities. Therefore, the implemented method must be efficient. Hence, we opt for C++ implementation of the method. We also utilize the parallelization functionality of the deal.II environment [10] to generate multiple independent cell trajectories concurrently. The method presented here will also serve as a basis to simulate a population of colliding cells. It is also applicable to a general piecewise deterministic process without an “active” boundary¹, and where the deterministic component varies continuously (otherwise, only a slight modification is necessary). Furthermore, a user is free to choose specific subroutine methods, such as ODE time stepping, quadrature, or root finding schemes.

This chapter is based on the work by the author in [120].

¹Jumps occur only at the random event times, and not when the deterministic part attains a certain value in a set.

4.1 Methods

We employ the Algorithm 1 to simulate the trajectories. In order to generate the interarrival time Δ_k , we need to solve for τ in the following equation:

$$1 - \exp\left(-\int_t^{t+\tau} a_0(\nu_t, \phi_\nu(s, \mathbf{X}_t)) ds\right) = u,$$

where $u \sim U(0, 1)$ is uniformly distributed on the interval $(0, 1)$. This is equivalent to finding the root of the following:

$$f(\tau) := \int_t^{t+\tau} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) ds + \ln(1 - u) = 0, \quad (4.1.1)$$

Notice that the evaluation of the integral by a quadrature rule requires computing the solution $\mathbf{X}_{t+s} = \phi_{\nu_t}(s, \mathbf{X}_t)$ up to time s , where s is a quadrature point. Moreover, using an iterative method to solve (4.1.1) requires computing the integral at each iteration. Therefore, it is imperative to devise an efficient method to sample from the distribution, given by (4.1.2).

Algorithm 1 Simulation of the PDMP

1. Set the initial conditions $(\nu_0, \mathbf{X}_0) \in A \times \Gamma$ and $t = \mathcal{T}_0 = 0$

2. For $k = 0, 1, \dots$

Generate interarrival time $\Delta_k = \mathcal{T}_{k+1} - \mathcal{T}_k$, whose distribution function is given by:

$$\mathbb{P}(\Delta_k \leq \tau) = 1 - \exp\left(-\int_t^{t+\tau} a_0(\nu_t, \phi_\nu(s, \mathbf{X}_t)) ds\right) \quad (4.1.2)$$

Compute $\mathbf{X}_{t+s^-} := \phi_\nu(s, \mathbf{X}_t)$

Set $t = \mathcal{T}_{k+1} = \mathcal{T}_k + \Delta_k$

Generate $(\nu_t, \mathbf{X}_t) \sim Q(\cdot; (\nu_{t^-}, \mathbf{X}_{t^-}))$

In the following, we propose a general method to generate the next event time.

4.1.1 Interarrival time generation

Let $\mathcal{T}_k \leq t < \mathcal{T}_{k+1}$ and let $\mathbf{G}(\cdot; h) : \Gamma \rightarrow \Gamma$ represent a numerical method to solve the ODE system

$$\frac{d}{dt} \mathbf{X}_t = \mathbf{H}_{\nu_t}(\mathbf{X}_t)$$

for a single time step of size h . That is, $\mathbf{X}_{t+h} = \mathbf{G}(\mathbf{X}_t; h)$ is the numerical solution of the above ODE system at time $t + h$.

Let $[\mathcal{T}_k, \mathcal{T}_{k+1}]^2 \ni (s', t') \mapsto A_0(t', s') := \int_{s'}^{t'} a_0(\nu_t, \phi_{\nu_t}(u, \mathbf{X}_t)) du$ denote the integrated rate function.

The method to find the root τ of equation (4.1.1) is given in Algorithm 2. First, in steps (1-22), we find the upper bound τ_{max} by solving the ODE system for n steps with step size h and store the solution, the computed rate a_0 , and the integrated rate A_0 at these time steps (see also Figure 4.1 for illustration). Then, for any $\tau \leq \tau_{max}$ we can compute $A_0(t + \tau, t)$ upon using the stored value of A_0 at time $t + \tau_i$:

$$\begin{aligned} A_0(t + \tau, t) &= \int_t^{t+\tau} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) ds \\ &= \int_t^{t+\tau_i} a_0(\nu_t, \mathbf{X}_{t+s}) ds + \int_{t+\tau_i}^{t+\tau_i+\tau-\tau_i} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) ds \\ &= A_0(t + \tau_i, t) + \int_{t+\tau_i}^{t+\tau_i+h_i} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) ds, \end{aligned}$$

where $i = \lfloor \frac{\tau}{h} \rfloor$, $\tau_i = ih$ is the time point at which the solution is computed, and $h_i = \tau - \tau_i$. Note that an equidistant mesh is not required and any mesh can be used. To compute the last integral in the expression above, we interpolate the integrand using the stored values $a_0(\nu_t, \mathbf{X}_{t+\tau_i})$ and $a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}})$. Let

$$I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h) := \int_{t+\tau_i}^{t+\tau_i+h_i} a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t))$$

denote the approximation of the integral using the interpolated integrand. We can use the following interpolation methods for $t + \tau_i < s < t + \tau_{i+1}$:

1. Piecewise constant

Forward: $a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) = a_0(\nu_t, \phi_{\nu_t}(\tau_i, \mathbf{X}_t)) = a_0(\nu_t, \mathbf{X}_{t+\tau_i})$.

$$I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h) = h_i a_0(\nu_t, \mathbf{X}_{t+\tau_i}). \quad (4.1.3)$$

Backward: $a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) = a_0(\nu_t, \phi_{\nu_t}(\tau_{i+1}, \mathbf{X}_t)) = a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}})$

$$I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h) = h_i a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}). \quad (4.1.4)$$

2. Average: $a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) = \frac{1}{2} (a_0(\nu_t, \mathbf{X}_{t+\tau_i}) + a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}))$.

$$I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h) = \frac{1}{2} h_i (a_0(\nu_t, \mathbf{X}_{t+\tau_i}) + a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}})). \quad (4.1.5)$$

3. Piecewise linear:

$$a_0(\nu_t, \phi_{\nu_t}(s, \mathbf{X}_t)) = (s - t - \tau_i) \frac{a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}) - a_0(\nu_t, \mathbf{X}_{t+\tau_i})}{h} + a_0(\nu_t, \mathbf{X}_{t+\tau_i}).$$

$$I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h).$$

$$= h_i a_0(\nu_t, \mathbf{X}_{t+\tau_i}) \left(1 - \frac{h_i}{h}\right) + \frac{h_i^2}{2h} a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}). \quad (4.1.6)$$

Thus, $f(\tau)$ can be evaluated using equations (4.1.3)-(4.1.6):

$$f(\tau) = A_0(t + \tau_i, t) + I(a_0(\nu_t, \mathbf{X}_{t+\tau_i}), a_0(\nu_t, \mathbf{X}_{t+\tau_{i+1}}); h_i, h) + \ln(1 - u). \quad (4.1.7)$$

Using the interpolations above, we can now employ any root finding method with bracket-

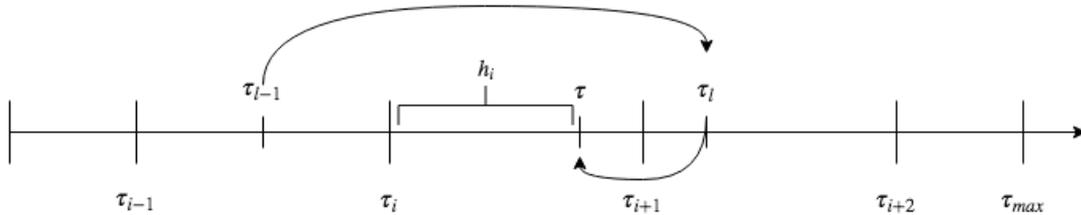


Figure 4.1: Illustration of the interarrival time computation. Subscripts i indicate the time points at which the solution of the ODE system, a_0 and A_0 are stored. Subscripts l indicate iterates of a root finding method, where the value of f in (4.1.1) is computed using the stored values. See text for details.

ing to solve (4.1.1), such as bisection or Brent's method. Once the root is found, we simply advance the ODE system for one time step as described in Steps (28-33) of the Algorithm 2.

Note that we solve the ODE system for $n + 1 = \tau_{max}/h + 1$ steps. We also solve for the interarrival time τ primarily by using a look-up table, since the evaluation of f in (4.1.7) requires the stored values of a_0 and A_0 . Moreover, obtaining a relatively sharp upper bound τ_{max} does not yield a large computational overhead, since one simply can start the Algorithm 2 with a small n, m . Consequently, choosing an initial guess close to the sharp upper bound for a root finding method results in a faster convergence. In case of thinning methods (see [14] or [65] for adaptive methods), after each rejection one needs to solve the ODE system for time period that is, on average, approximately the same as τ_{max} (in the best case scenarios for both methods, i.e. when the bound τ_{max} and the bound for the rate function in thinning methods are sharp). Of course, these arguments apply when the computational cost of solving the ODE system is relatively large.

4.1.2 Sampling from the transition measure

Given the time t of the next event and \mathbf{X}_t we need to sample from the transition distribution $Q(\cdot, (\nu_{t-}, \mathbf{X}_{t-}))$. Recalling Section 2.3.2 and the proof of Proposition 2.4.3, in order to sample from the transition measure it is sufficient to simulate the index $j \in \{1, \dots, 2M\}$ of the next reaction, since the continuous component of the process does not jump. The discrete distribution of the next reaction index is given by equation (2.4.12):

$$\mathbb{P}(j | \alpha_{\mathbf{Y}}(\nu_{t-}, \mathbf{X}_{t-})) = \frac{a_j(\alpha_{\mathbf{Y}}(\nu_{t-}, \mathbf{X}_{t-}))}{a_0(\alpha_{\mathbf{Y}}(\nu_{t-}, \mathbf{X}_{t-}))}.$$

To simulate from the discrete distribution one can use the fairly efficient Vose Alias Method [127].

Algorithm 2 Event time computation

```

1: procedure INITIALIZATION
2:   Input: Time  $t = \mathcal{T}_k$ ;  $(\nu_t, \mathbf{X}_t)$ ; time step  $h$  and ODE method  $\mathbf{G}$ ;  $n, m \in \mathbb{N}$ .
3:   Set  $s_0 := 0, n_0 := 0$ , create  $List_{a_0}$ ,  $List_{A_0}$ , and  $List_{\mathbf{X}}$ .
4:   Append  $List_{a_0} \leftarrow a_0(\nu_t, \mathbf{X}_t)$ ,  $List_{A_0} \leftarrow 0$ ,  $List_{\mathbf{X}} \leftarrow \mathbf{X}_t$ .
5:   Set  $\tau_{max} := nh$ ,  $s_0 := n_0h$ .
6:   Set initial condition  $\mathbf{X}_s \leftarrow List_{\mathbf{X}}[last]$ .
7:   Set  $A_0^0 := List_{A_0}[last]$ .
8:   for  $i = 1$  to  $n$  do
9:      $s_i := t + s_0 + ih$ .
10:    Compute  $\mathbf{X}_{s_i} := \mathbf{G}(\mathbf{X}_s; h)$  and  $a_0(\nu_t, \mathbf{X}_{s_i})$ .
11:    Compute  $A_0(s_i; t + s_0)$  with quadrature points  $s_j, j = 0, \dots, i$ .
12:     $A_0(s_i; t) := A_0^0 + A_0(s_i; t + s_0)$ .
13:    Append  $List_{a_0} \leftarrow a_0(\nu_t, \mathbf{X}_{s_i})$ ,  $List_{A_0} \leftarrow A_0(s_i; t)$ ,  $List_{\mathbf{X}} \leftarrow \mathbf{X}_{s_i}$ .
14:  end for
15:  Generate  $u \sim U(0, 1)$ .
16:  if  $List_{A_0}[last] < -\ln(1 - u)$  then
17:     $n := n + m$ .
18:     $n_0 = n_0 + m$ .
19:    go to 5.
20:  end if
21:  Output:  $\tau_{max}$ ,  $List_{a_0}$ ,  $List_{A_0}$ , and  $List_{\mathbf{X}}$ .
22: end procedure
23: procedure EVALUATION OF  $f$ 
24:   Input:  $\tau$ ; time step  $h$ ;  $List_{a_0}$ ,  $List_{A_0}$ ; Integrated interpolation method  $I$ .
25:   Set  $i := \lfloor \frac{\tau}{h} \rfloor$  and  $h_i = \tau - ih$ .
26:   Output:  $f(\tau) = List_{A_0}[i] + I(List_{a_0}[i], List_{a_0}[i + 1]; h_i, h) + \ln(1 - u)$ 
27: end procedure
28: procedure EVENT TIME
29:   Find the root  $\tau$  of  $f(\tau) = 0$  using 23 and a root finding method.
30:    $i := \lfloor \frac{\tau}{h} \rfloor$ .
31:   Compute  $\mathbf{X}_{t+\tau} := \mathbf{G}(List_{\mathbf{X}}[i]; \tau - ih)$ .
32:   Output:  $\tau$ ,  $\mathbf{X}_{t+\tau}$ 
33: end procedure

```

4.2 Parameters

Here we present our estimates of the remaining model parameters - the drag coefficients. The values of other parameters are estimated in Sections 3.1, 3.3.

Using Stokes' Law for drag in the low Reynolds number regime, the drag coefficient β_{ECM} can be estimated as:

$$\beta_{ECM} = 6\pi\eta_{ECM}R_{cell},$$

where η_{ECM} is the dynamic viscosity of the environment. Assuming that the viscosity η_{ECM} is higher than that of water, and taking into account that the contact between cell surface and the substrate increases the effective friction, we estimate $\beta_{ECM} \approx 10 - 10^3 \frac{N \cdot s}{m^2} \times R_{cell}$. Similarly, due to the low Reynolds number, the rotational drag coefficient β_{rot} is given by [63]:

$$\beta_{rot} = 8\pi\eta_{ECM}R_{cell}^3 \approx 10 - 10^3 \frac{N \cdot s}{m^2} \times R_{cell}^3.$$

In order to obtain estimates for the drag coefficient β_{cell} one needs to have an estimate of the cytoplasm viscosity. Assessing the effective cytoplasmic viscosity of migrating cells is a challenging task, since the viscoelastic properties of the cytoskeleton (which, among other things, consists of polymer networks) are highly dynamic due to constant remodeling and spatiotemporal mediation of the rheology by various signaling pathways. Particularly, the actin network bundle size and cross-linkers influence the viscoelastic properties [42]. Furthermore, the effective viscosity experienced by an experimental probe (or a protein) in polymer solutions depends not only on the type of material properties of the fluid, but also on the size of a probe ¹ [52]. Inferring that the body being displaced in the cell due to contractile fibers is the nucleus with radius $R_{nucleus}$, we estimate:

$$\beta_{cell} = 6\pi\eta_{cell}R_{nucleus} \approx 10 - 10^2 \frac{N \cdot s}{m^2} \times R_{nucleus},$$

where η_{cell} is the cytoplasm viscosity.

4.3 Results

Here we show the results of simulating cell trajectories under different scenarios, which represent various experimental settings, namely:

1. Uniform environment with no cues.
2. Non-uniform environment with external cue gradient and uneven myosin motor activity within a cell.
3. Striped ECM architecture.

¹Using a small molecule as a probe, the cytoplasm viscosity was found to be $\simeq 2 - 3 \times 10^{-4} Pa \cdot s$ [75]. With larger probes, the viscosity was found to be $\simeq 2 - 4 \times 10^{-2} Pa \cdot s$ [52] and $\simeq 5 \times 10^{-2} Pa \cdot s$ [6]

Parameter	Value	Value	Parameter	Value	Value
T_i	$4nN$	0.72	R_{cell}	$25\mu m$	1
EA	$15nN$	2.72	L_0	$27.5\mu m$	1.1
F_b	$5.5nN$	1	L_c	$5\mu m$	0.2
k_{off}^0	$0.05s^{-1}$	1	β_{rot}	$1.56 \times 10^{-11} N \cdot s \cdot m$	5.68
k_{on}^0	$0.01s^{-1}$	0.2	β_{ECM}	$5 \times 10^{-4} \frac{N \cdot s}{m}$	0.11
$R_{nucleus}$	$5\mu m$	0.2	β_{cell}	$5 \times 10^{-3} \frac{N \cdot s}{m}$	1.14

Table 4.1: Parameters used for simulation and their relative size with respect to spatial, temporal, and force scales. See Section 2.2.3 for details.

Note that the total number of adhesion sites M is a free parameter, which differs from cell to cell. Nevertheless, it is a crucial parameter, determining whether the motility type is amoeboid or mesenchymal. Amoeboid motility is characterized by a large number of weak adhesions, high turnover, and more contractile cell body. On the other hand, mesenchymal migration relies on fewer, but stronger focal adhesions with slower turnover and lower overall contractility. The cells with the former motility type are faster and have more diffusive motion [70], [90]. Note that if $a_j^\pm \sim O(1)$, then the rate function is $a_0 \sim O(M)$. Therefore, adhesion events occur more frequently for increasing M , implying higher adhesion turnover. Thus, by varying M , our model is capable of explaining this particular aspect of the difference between the two migration types.

The rationale for simulating $M = 8, 16, 32$ FAs is the following. Note that the number of cell-substrate adhesions is higher than the number of focal adhesions we chose for our simulations. However, not all adhesions are directly involved in translocating the cell body, during which large traction forces are applied to the substrate through focal adhesions (which are fewer in number than immature, less stable focal complexes/points and nascent adhesions). Moreover, detachment of focal adhesions that leads to translocation, is primarily the result of contractile tension applied by ventral stress fibers, as opposed to transverse arcs and dorsal stress fibers [55]. The latter two have primarily structural role, while the former is fundamental to rear retraction [22], [55]. Thus, the number of focal adhesions that are directly involved in cell body translocation is controlled by the number of the associated ventral stress fibers, which are also the most significant source of traction force applied to the substrate due to large tension within them [55], [108]. Although reports of ventral stress fiber numbers are elusive, visual inspection of the fluorescence images in, for example, [51], [62], [108] (or any other appropriate study) suggests that simulations with the chosen number of (ventral) fibers (and focal adhesions) is realistic. Moreover, diameter of focal adhesions $d \sim 1 - 5\mu m$ [41]. Assuming that the separation between focal adhesions is comparable to their size, and taking the cell radius to be $R_{cell} = 25\mu m$ (as

in our simulations), we see that the upper range of possible number of adhesions on the cell circumference is $2\pi R_{cell}/2d \approx 16 - 80$. We reiterate that this number is an estimate of focal adhesions attached to ventral stress fibers, and it underestimates the *total* number of focal adhesions that a cell employs, since significant number of them are attached to other types of stress fibers and may also be present within the cell body and not on its periphery. We performed simulations with $M = 64$ focal adhesions and did not find any added insight.

For all scenarios we employ the same initial conditions for all cells. Namely, at $t = 0$ the conditions are:

- \mathbf{x} is at the origin, \mathbf{x}_n is uniformly distributed on a circle with radius R_{cell} , and $\theta = 0$.
- Each FA is in (un)bound state and each cell is in moving state with probability $1/2$.

We simulate trajectories of $n_{cell} := 56$ cells for time $t_{end} := 600min$. We divided the time interval into $n_{time} := 5000$ intervals, at the end of which we recorded the cell centroid positions \mathbf{x} . For details on the data analysis, see Appendix B.

4.3.1 Uniform environment

The results of the simulation with uniform spatial cue Q_{cue} are presented in Figure 4.2. Due to the absence of spatial variation of Q_{cue} , we take $q = 1$ in equation (3.1.4).

The cell trajectories with varying number of adhesion sites, depicted in Figure 4.2 (a-c), show no clear trend and resemble those of a Brownian motion. Indeed, fitting the mean-squared displacement $msd(t)$ to the curve $\widehat{msd}(t) = \beta_0 t^{\bar{\beta}}$ (see Appendix B for details), we see that the exponent is $\bar{\beta} \approx 1$ for the three cases (see Figure 4.2 (d-f)). This suggests that the cell motion has diffusive characteristics in this scenario. In Figure 4.2 (g-i) we see the simulated distribution of speeds. The average speeds s_{av} and the parameters $\beta_0, \bar{\beta}$ are shown in Table 4.2. We see that as M increases, the cell motion becomes progressively faster and more diffusive¹, which is expected for a dominantly amoeboid type of motility. Because $\bar{\beta} \sim 1$, we can estimate the diffusion coefficient $D = \beta_0/4$. The obtained values are lower, but within an order of magnitude estimated by Liu et. al. [70], who found that $D \approx 2.7\mu m^2/min$. Interestingly, the gamma distribution gives a very good fit to the simulated data for various values of M , suggesting that cell speeds are gamma distributed. Moreover, the obtained values of the average speed s_{av} fall in the range reported by Liu et. al. [70], who found the speeds to be in the range from $0.2\mu m/min$ to $7\mu m/min$. Although there are very high speeds observed in Figure 4.2 (i), which seem to be outliers, speeds as high as $25\mu m/min$ have been observed [40]. As expected, angle histogram plots of normalized velocities show no bias in any particular direction in Figure 4.2 (j-l). Along with time scaling of the squared displacement, persistence of motion can be measured by directionality ratio (distance between cell centroids divided by path length) and velocity autocorrelation [45]. Expectedly, Figure 4.3 illustrates that the cells strongly deviate from

¹Since $\bar{\beta} \sim 1$, the slope β_0 is a measure of diffusivity. See Appendix B.

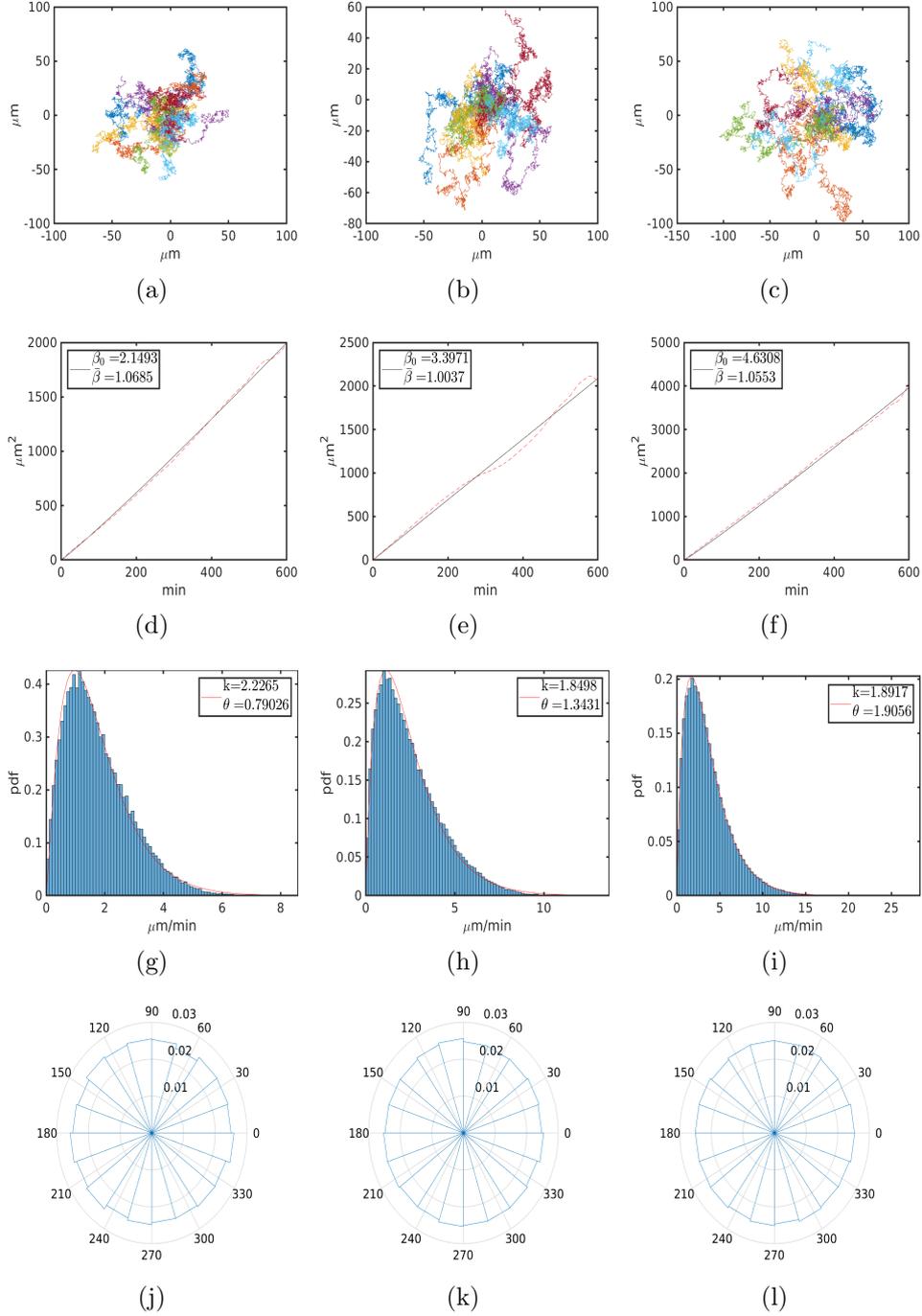


Figure 4.2: Simulation results with $M = 8, 16, 32$ adhesions in the first, second, and third columns respectively. (a-c) Trajectories of 13 cell centroids $\mathbf{x}(t)$. (d-f) Mean-squared displacements $msd(t)$ (red, dash) and fitted $\widehat{msd}(t)$ (black, solid) with parameters β_0 and $\bar{\beta}$ (see text for details). The unit of β_0 is $\mu\text{m}^2/\text{min}^{\bar{\beta}}$. (g-i) Histograms of speed probability density functions and fitted density function of gamma distribution (red) with parameters k and θ . (j-l) Relative frequency histogram of normalized velocities.

straight-path migration (Figure 4.3 (left); see also time average of the directionality ratio \bar{r} in Table 4.2) and the deviation directions are uncorrelated in time (Figure 4.3 (right)). The rapid decay in Figure 4.3 (right) also suggests that correlations in time become stationary very fast. The increased oscillations in Figure 4.3 (right) towards the end of the observation window are due to decreased number of observations (see Appendix B).

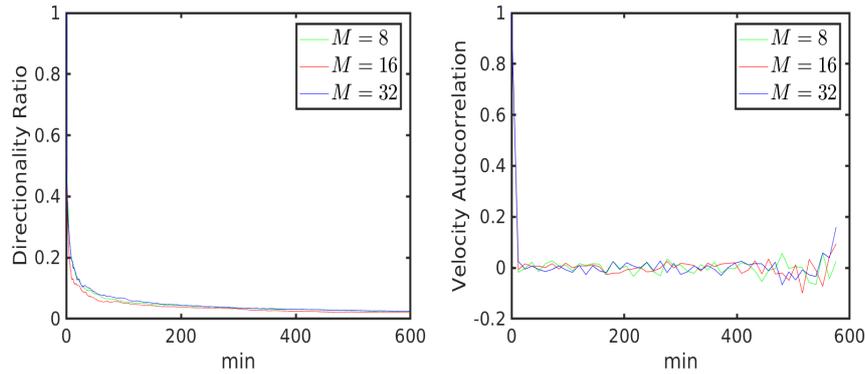


Figure 4.3: Directionality ratio (left) and velocity autocorrelation (right) for $M = 8$ (green), 16 (red), 32 (blue).

M	8	16	32
$s_{av}, \mu m/min$	1.7595	2.4845	3.6047
$\bar{\beta}, 1$	1.0683	1.0035	1.0552
$\beta_0, \mu m^2/min^{\bar{\beta}}$	2.1493	3.3971	4.6308
$\bar{r}, 1$	0.0483	0.0408	0.0497

Table 4.2: Parameters obtained from the simulations.

Although our results show in Figure 4.2 (d-f) the mean-squared displacement scales diffusively (i.e. linearly) with time, this is not consistent with the reported results. For example, Dietrich et. al. [32], Liang et. al. [67], and Liu et. al. [70] showed that the displacement scales superdiffusively. In these studies it was experimentally found that the time scaling went as $\sim t^{\bar{\beta}}$, where $\bar{\beta} \approx 1.2-1.5$. The primary reason why, in our case, we have diffusive behavior is that, given a certain state of adhesion sites, there is a complete circular symmetry of the probability rates a_j^{\pm} with respect to \mathbf{x}_n variable. Due to this symmetry, then, the probability of moving in one direction is exactly the same as the probability of moving in the opposite direction if we rotate \mathbf{x}_n by π radians. Hence, somewhat reminiscent of a random walk, we obtain a diffusive time scaling of the squared displacement. Moreover, this symmetry of the probability rates is somewhat idealistic, since it implies that the

signaling activity relevant for adhesion dynamics is homogeneous within a cell. One of the ways to break this symmetry, is to multiply each binding probability rate a_j^+ by $1 + u$, where $u \sim U(-\delta, \delta)$ is uniformly distributed on the interval $(-\delta, \delta)$ with $\delta \in (0, 1)$. Then, on average, the rates are unmodified¹. This way, we not only simulate a non-homogeneous binding rate (and hence, for example, non-homogeneous Rac1 activity) within a cell, but also simulate otherwise completely identical copies of cells. Such a modification, where we do not explicitly apply a directed, predefined bias can be referred to as chemokinesis [92].

M	8			16			32		
δ	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
$\bar{\beta}, 1$	0.9859	1.3184	1.4084	1.0086	1.3505	1.5581	1.1014	1.4299	1.5639
$s_{av},$ $\mu m/min$	1.7656	1.7768	1.7557	2.4918	2.5009	2.5021	3.5818	3.5735	3.6104
$\beta_0,$ $\mu m^2/min\bar{\beta}$	3.0846	0.6934	0.5534	3.4517	0.7103	0.3543	4.1257	0.9536	0.5716
$\bar{r}, 1$	0.0452	0.0519	0.0597	0.0440	0.0513	0.0587	0.0522	0.05	0.0623

Table 4.3: Parameters obtained from the simulations with varying δ .

The effect of modifying the rates a_j^+ with $\delta = 0.05, 0.1, 0.15$ can be seen in Figure 4.4. The cell trajectories, depicted in Figure 4.4 (a-c), show that the motion consists of periods with relatively regular path intermingled with highly irregular and random movement. In Figure 4.4 (d-f) we see that the rate modification leads to a superdiffusive time scaling of the mean-squared displacement, as the exponent $\bar{\beta}$ becomes larger than one and falls within the experimentally observed range of values [32], [67], [70]. Moreover, we see that as δ increases, so does $\bar{\beta}$, and the increase of the latter is more pronounced for a larger number of adhesion sites M (see also Table 4.3). This is due to the fact that as each adhesion site is modified independently, the variance of the modified rates of a cell grows with the number of FAs, which corresponds to increased cell polarization, and hence more prominent persistent motion resulting in higher values of $\bar{\beta}$. However, the distribution of speeds for the corresponding values of M is virtually identical to the case with the unmodified probability rates (Figure 4.4 (g-i) and Table 4.3). The uniform distribution of normalized velocities also remained unchanged (Figure 4.4 (j-l)). These results suggest that in the absence of spatial cues, the distribution of speeds for a given adhesiveness (represented by the total number of adhesions M) remains invariant under symmetry breaking of adhesion binding, while the diffusion type (normal vs. anomalous) does not. Thus, the adhesion number and its turnover is a major determinant of the cell speed, which is consistent with [90].

¹The multiplication factor $1 + u$ for each $j = 1, \dots, M$ of every cell is computed at the beginning of simulations and is held fixed thereafter.

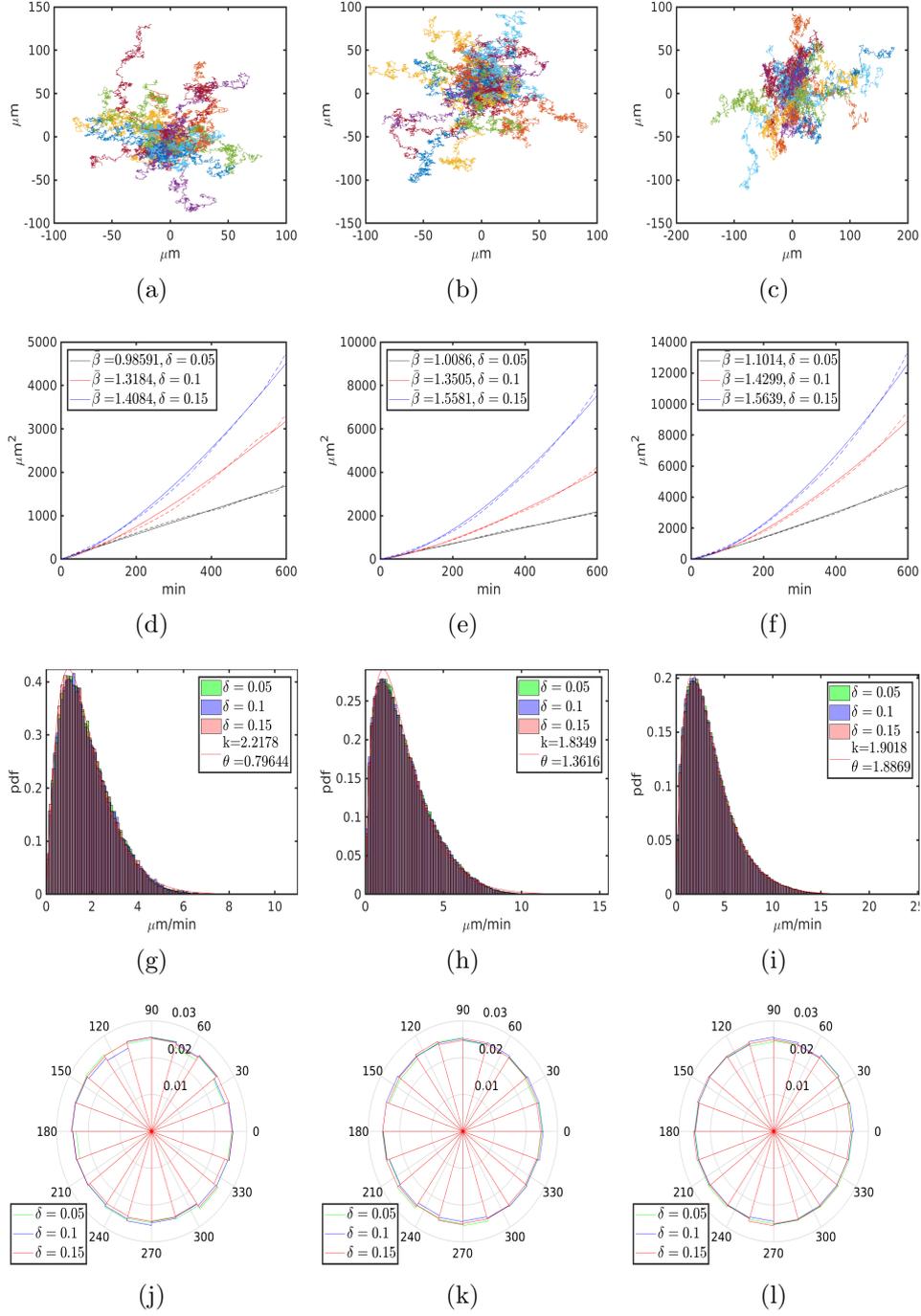


Figure 4.4: Simulation results with $M = 8, 16, 32$ adhesions in the first, second, and third columns respectively, and with various values of δ . (a-c) Trajectories of 27 cell centroids $\mathbf{x}(t)$ with $\delta = 0.1$. (d-f) Mean-squared displacements $msd(t)$ (solid) and fitted $\widehat{msd}(t)$ (dash) with $\delta = 0.05$ (black), 0.1 (red), 0.15 (blue). (g-i) Superimposed histograms of speed probability density functions and fitted density function of gamma distribution (solid red) with average parameters k and θ (see text for details). (j-l) Superimposed histogram of relative frequency of normalized velocities.

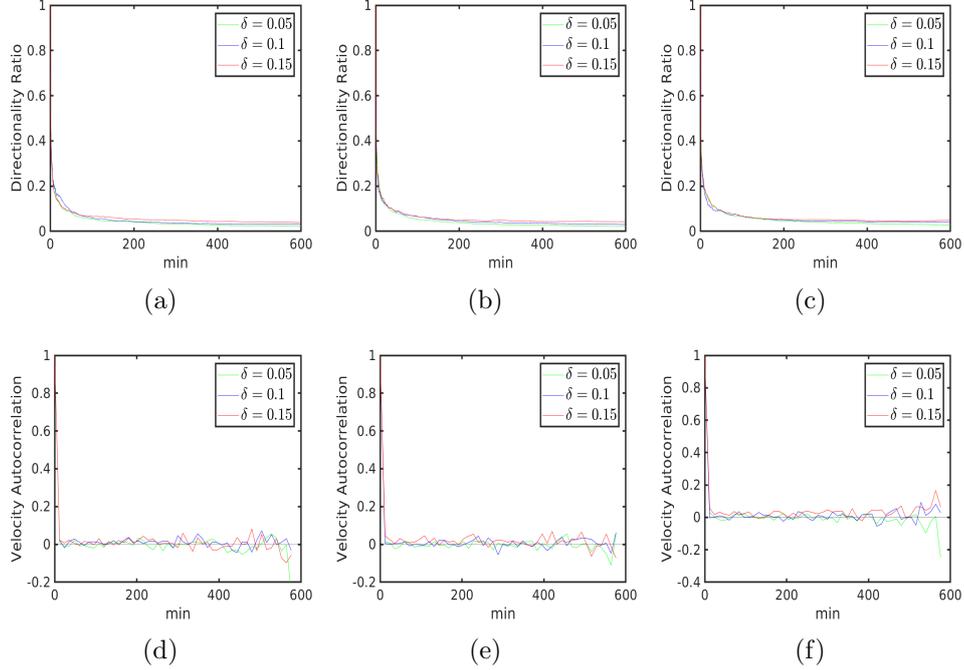


Figure 4.5: Persistence of motion for cells with $M = 8, 16, 32$ adhesions in the first, second, and third columns, respectively, and with $\delta = 0.05$ (green), 0.1 (blue), 0.15 (red). (a-c) Directionality ratio. (d-f) Velocity autocorrelation.

Note that the increased values of $\bar{\beta}$ indicate that the cells explore a larger surface area [45]. However, other indicators of motion persistence are not affected significantly (Figure 4.5), although migration paths become slightly straighter, as indicated by increased values of \bar{r} (Table 4.3). These results suggest that symmetry breaking of adhesion binding may allow cells to explore larger area without introducing velocity correlations (Figure 4.5(d-f)).

As cell polarization is required for migration even in the absence of external signals, it is not surprising that our results show that an imbalance of adhesion formation within a cell leads to experimentally observed superdiffusive scaling of the squared displacement [32],[67], [70]. Nevertheless, this highlights a potential mechanism of anomalous diffusion. In the following, we examine whether our model gives biologically consistent results in the case of externally induced polarization.

4.3.2 External cue gradient

We first investigate how cell trajectories are varied in the presence of an external cue gradient. If a cue, for example, is a chemoattractant, then it is well known that adhesion formation in a cell is biased in the direction of the attractant. Thus, to simulate such biased migration, we let the functions Q_{cue} and q to have the following form (recall equation

(3.1.4)):

$$Q_{cue}(\mathbf{x}) = \begin{cases} 1 + \delta_E x_2, & \text{if } x_2 \geq 0 \\ 1, & \text{else} \end{cases}$$

$$q(Q_{cue}(\mathbf{x})) = Q_{cue}(\mathbf{x}),$$

where δ_E represents the gradient magnitude and x_2 is the second component of \mathbf{x} . Here, for simplicity we took the identity function for q and a linear cue gradient in the y coordinate. This cue can represent, for example, density of ECM or concentration of a chemoattractant. Thus, we simulate, respectively, hapto- or chemotactic migration.

M	8			16			32		
δ_E	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
$\bar{\beta}, 1$	1.2551	1.5051	1.52	1.3405	1.6963	1.7545	1.5427	1.7722	1.8569
$s_{av},$ $\mu\text{m}/\text{min}$	1.8136	1.9133	2.0365	2.5235	2.5972	2.6089	3.5819	3.4218	3.3074
$\beta_0,$ $\mu\text{m}^2/\text{min}^{\bar{\beta}}$	1.0697	0.4625	0.6845	1.0312	0.3496	0.3654	0.8319	0.3412	0.2242
$\bar{r}, 1$	0.0523	0.0607	0.0693	0.053	0.08	0.097	0.0726	0.1	0.1223

Table 4.4: Parameters obtained from the simulations with varying δ_E .

In the presence of a cue gradient, we see that the cell trajectories, shown in Figure 4.6 (a-c), exhibit a clear trend in the direction of an increasing concentration. The corresponding plots of the mean-squared displacements show the superdiffusive time scaling in Figure 4.6 (d-f), with the exponent $\bar{\beta} > 1$ for all cases. Notice that as the number of adhesion sites M increases, so does $\bar{\beta}$ for the same δ_E (see Table 4.4). Together with the trajectory plots in Figure 4.6, our results suggest that in the presence of an external gradient, the taxis becomes more prominent and a cell more sensitive to a cue for increasing number of FAs. Moreover, comparing with the case of a uniform environment, we see that although the amoeboid motility is more diffusive in the absence of external cues, it is also more regular and directed when a cue gradient is present (see Tables 4.2, 4.3 vs. Table 4.4 and Figures 4.2, 4.4(a-c) vs. 4.6(a-c)). In Figure 4.6 (g-i) we see that the evolution of time-averaged exponents $\beta_{av}(t)$ (see Appendix B) have three phases. Following the rapid increase in the first phase, there is a gradual decrease in the rate of change in the second phase, followed by stabilization of $\beta_{av}(t)$ at $\bar{\beta}$. Curiously, a similar behavior has also been observed by Dieterich et al. [32].

The distribution of speeds again remained invariant and the average speeds are very close to the cases with no external cues (see Table 4.4). However, the frequency of normalized velocities (see Figure 4.7 (d-f)) show, as expected, that the cell velocities are aligned

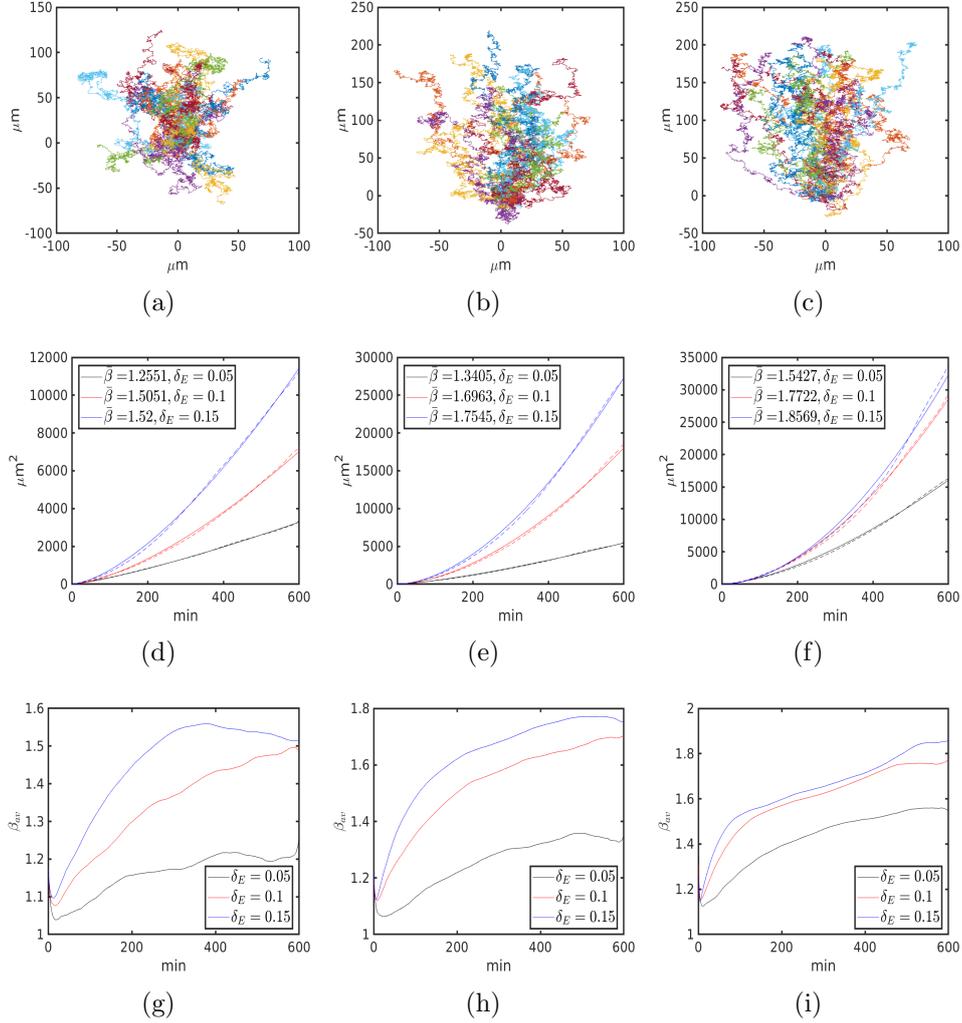


Figure 4.6: Simulation results with $M = 8, 16, 32$ adhesions in the first, second, and third columns respectively, and with various values of δ_E . (a-c) Trajectories of 27 cell centroids $\mathbf{x}(t)$ with $\delta_E = 0.1$. (d-f) Mean-squared displacements $\text{msd}(t)$ (solid) and fitted $\widehat{\text{msd}}(t)$ (dash). (g-i) Time-averaged exponents $\beta_{av}(t)$

with the cue gradient. Accordingly, we see that persistent motion emerges: directionality ratio increases compared to unbiased migration (Table 4.4) and the velocities become correlated (Figure 4.8(d-f)). We also observe that an external signal has a stronger impact on motion persistence for higher number of adhesions due to relative increases of \bar{r} and the degree of velocity autocorrelation. Recall that in the presence of, for example, a chemotactic cue, a cell polarizes so that its adhesion dynamics is aligned with the gradient. In particular, adhesions are preferentially formed at the front (where the chemoattractant concentration is larger), and preferentially ruptured at the back. We can see in Figure 4.7 (g-i), that our simulation results reproduce such polarized dynamics: the ratio of binding

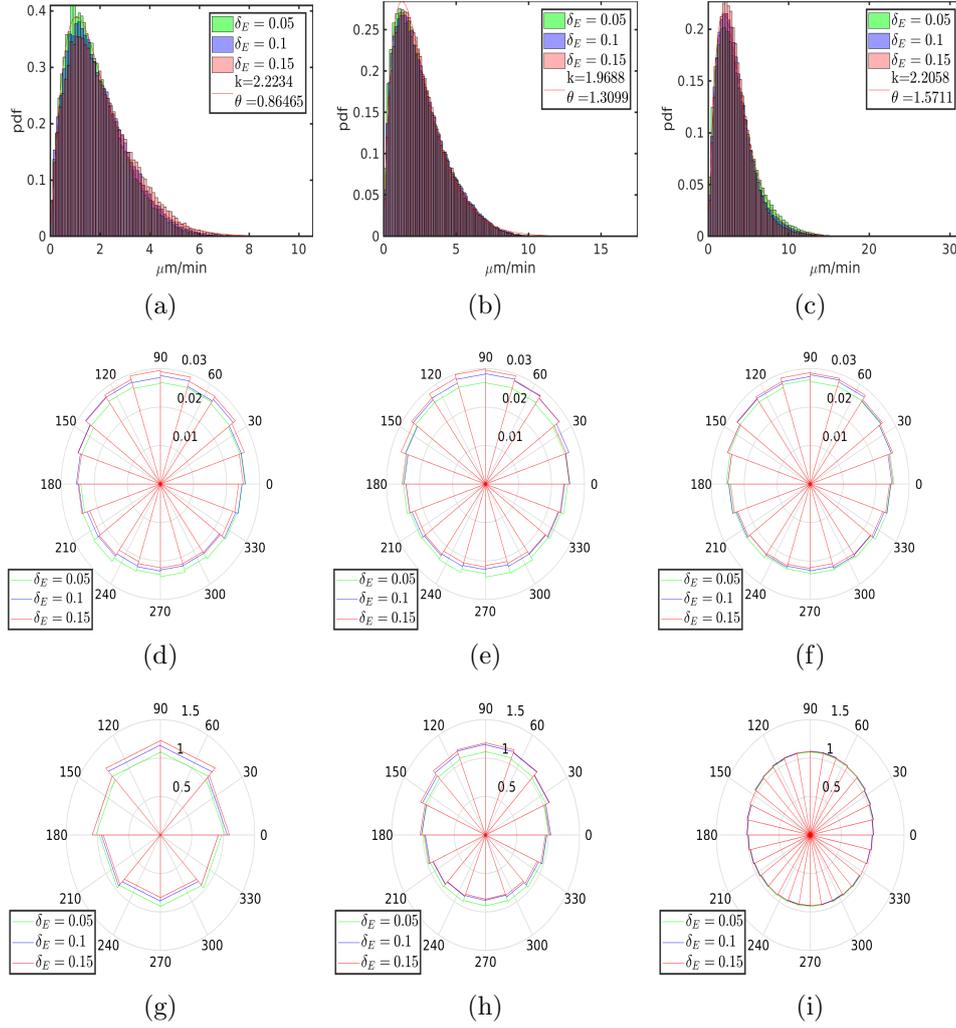


Figure 4.7: Superimposed histograms of speeds, velocities and adhesion events with $M = 8, 16, 32$ adhesions in the first, second, and third columns respectively, and with various values of δ_E . (a-c) Speed probability density functions and fitted density function of gamma distribution (solid red) with average parameters k and θ (see text for details). (d-f) Relative frequency of normalized velocities. (g-i) Ratio of the number of binding to unbinding events in each sector, such that any given time, only one adhesion site is in each sector

to unbinding events is larger (smaller) than unity in the northern (southern) part of the cells, where the cue is stronger (weaker) relative to the cell centroid. Also, for a smaller number of adhesion sites, the effects of increasing the cue gradient have more noticeable effect on the ratios of events (see 4.7 (g-i)). This is simply due to the reduced density of adhesion sites, which leads to larger relative difference in the concentration of the cue between them. From Figure 4.9 we can assess the effect of an external cue Q_{cue} on the the binding rate a_i^+ (omitting the force dependence for clarity), since the rate is proportional to Q_{cue} .

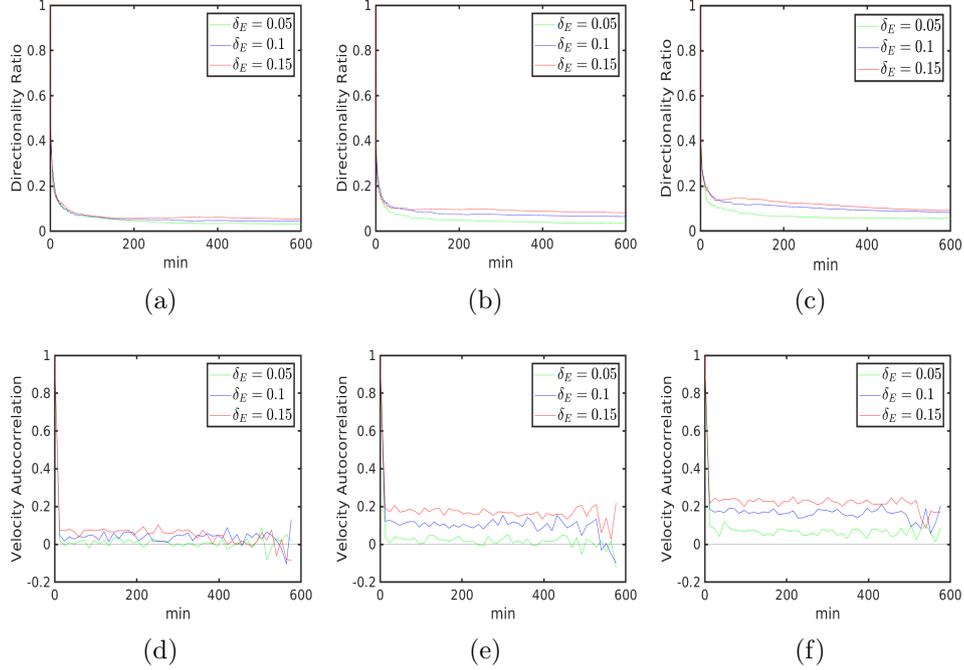


Figure 4.8: Persistence of motion for cells with $M = 8, 16, 32$ adhesions in the first, second, and third columns, respectively, and with $\delta_E = 0.05$ (green), 0.1 (blue), 0.15 (red). (a-c) Directionality ratio. (d-f) Velocity autocorrelation.

Together with Figure 4.9, the simulations illustrate that directed tactic migration, resulting from biased adhesion formation, follows from the *local* information about the external cue. That is, the spatial dependence of the FA binding rate is solely due to the local concentration of an external cue (see (3.1.4)) and no central mechanism for gradient determination was utilized to bias adhesion formation. Consequently, migration along the gradient of an external cue is achieved without its explicit “computation” by the cell.

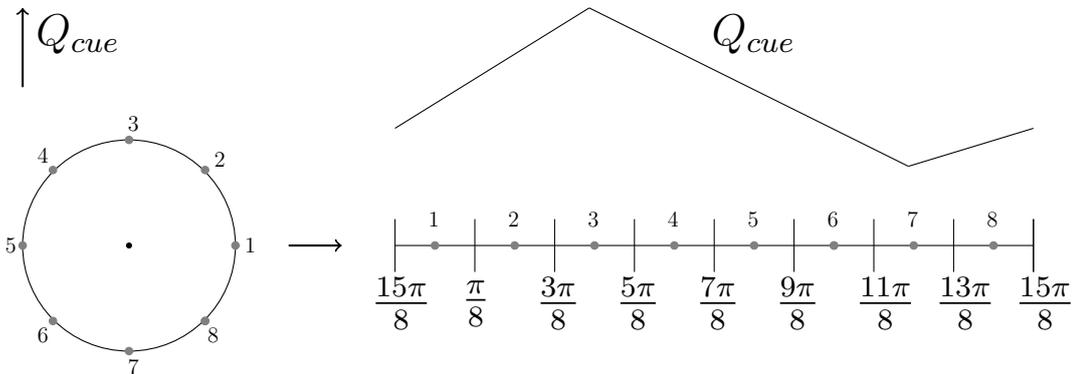


Figure 4.9: Concentration of an external cue projected on the cell’s circumference. Gray bullets represent FA sites.

Along with external cue, force dependence of the binding rate is also important for directed migration and without it, the cells do not exhibit biased migration (data not shown). Figure 4.10 illustrates how the dependence fits into the migration cycle (recall Figure 2.1). For the directed migration to occur, at the time of FA disassociation \mathbf{x}_n must be preferentially in the rear (Figure 4.10 step 2). After FA unbinding the increased force at the rear FAs due to extended SFs promotes binding there (Figure 4.10 step 3). Note that since cell body translocation occurs only after an unbinding event, formation of new FA in the prospective rear of the cell does not lead to backwards movement. Also, due to the external signal more FAs tend to be at the front than at the rear.

Thus, the pulling force exerted by the front on the rear tends to be larger than the opposite and hence the cell moves preferentially in the direction of the gradient. Without the signal, of course, movement becomes unbiased, as shown in the previous section. This suggests that the SF length dependence of the forces (see (2.1.1)) and the force dependence of the FA binding rate (see (3.1.3)) are necessary for directed migration resulting from biased adhesion formation in the presence of an external signal.

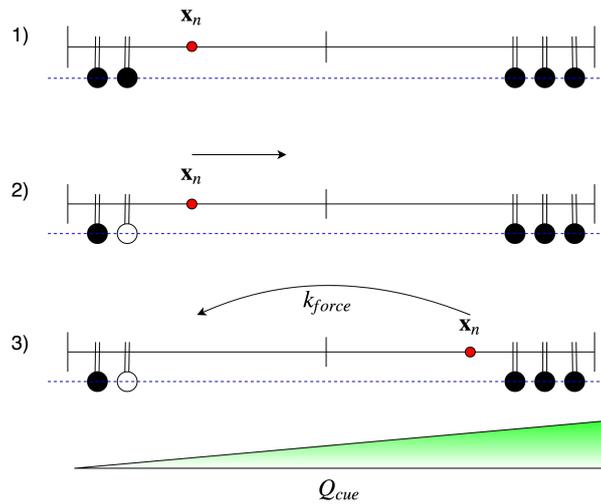


Figure 4.10: The force dependence of the binding rate and the biased adhesion formation during the migration cycle. Side view schematic of the cell is illustrated, where (un)bound FAs are shown as (white)black circles. 1) Initial configuration. 2) Unbinding leads to cell translocation and motion of \mathbf{x}_n within the cell. 3) Increased force on the cell rear (due to its dependence on SF extension) promotes FA association due to force dependence k_{force} of the binding rate (see Section 3.1.2), after which the cycle begins anew.

4.3.3 Fibrillar architecture of ECM

The ECM topography is another important determinant of directed cell migration. In particular, the spatial distribution of the ECM fibers guides the motility by inducing cell shape alignment along the adhesive cues, resulting in characteristic directed movement along the fiber tracts [92]. Such guided migration is called contact guidance [92], [95]. Ramirez-San

Juan et. al. [95] showed that contact guidance can be modulated by micrometer scale variations of interfiber spacing (see Figure 4.11). Inspired by this study, we simulate how subcellular scale fiber spacing influences cell motility, and whether such ECM architecture yields migration patterns characteristic of contact guidance.

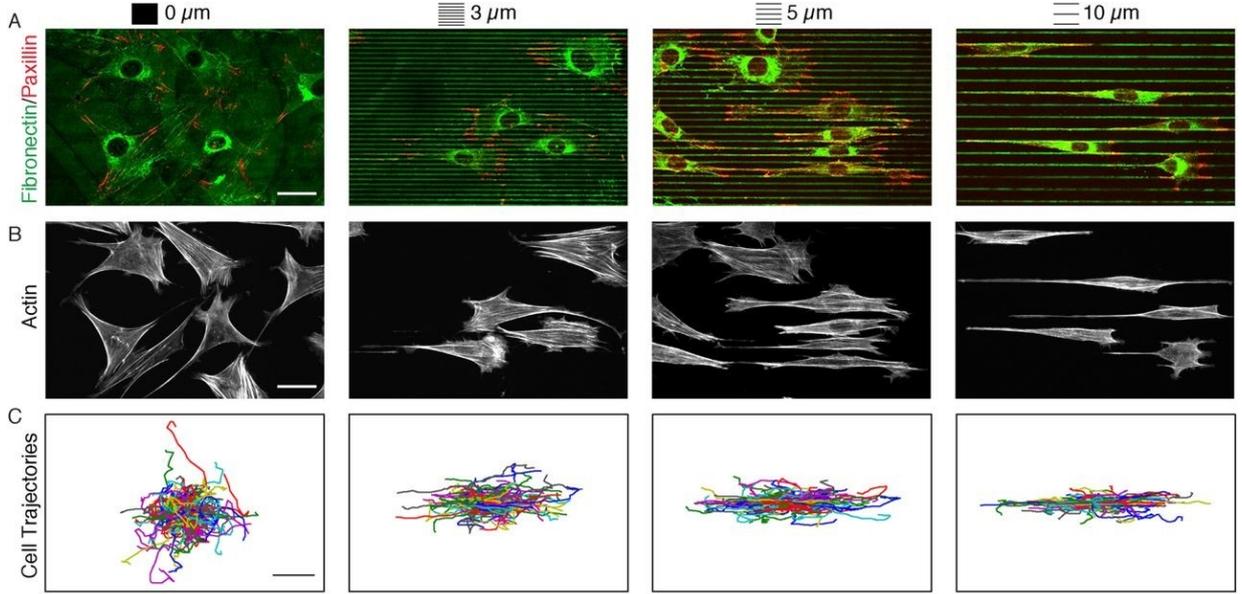


Figure 4.11: A result of the study in [95]. The interfiber distance is indicated at the top of each column. Each stripe is $2\mu m$ wide. (A,B) Fluorescence images of fibronectin (indicating the ECM), paxillin (indicating cell-substrate adhesions), and actin. (C) Cell trajectories corresponding to each setting. Scale bar $30\mu m$.

Similar to the case with an external cue gradient, the functions Q_{cue} and q have the following form:

$$Q_{cue}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_{\delta_G} \\ 0.01, & \text{else} \end{cases}$$

$$q(Q_{cue}(\mathbf{x})) = Q_{cue}(\mathbf{x}),$$

where Ω_{δ_G} represents the stripe pattern, $\delta_G = 0.15, 0.25, 0.35$ represents the spacing between stripes such that the distances between them is $\delta_G R_{cell}$ (Figure 4.12). The stripe width is taken to be $0.25 R_{cell}$. Similarly as in [95], these dimensions are chosen so that a cell is spread on multiple stripes.

The simulation results, shown in Figure 4.13, indicate that the cell motility has characteristics of contact guidance. Namely, the trajectories show preferential horizontal cell movement (Figure 4.13 (a-c)), and the displacements are aligned with the fiber pattern (Figure 4.13 (d-f)). However, increasing the spacing does not simply lead to a greater adhesion alignment along the horizontal direction, as can be observed in Figure 4.13 (g-i). Rather, it is the combination of the ECM pattern and the radial position of FAs that gives

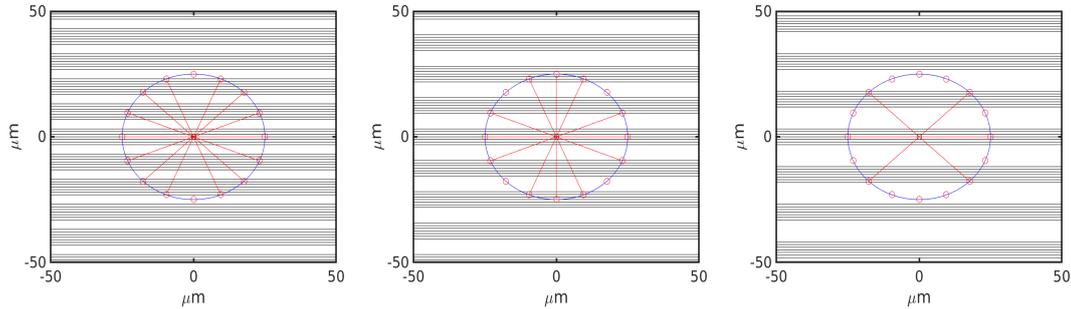


Figure 4.12: Stripe pattern with $\delta_G = 0.15, 0.25, 0.35$ on, respectively, left, middle, and right plots. A cell is illustrated such that each FA on a stripe is bound and cell center \mathbf{x} coincides with \mathbf{x}_n at the origin.

rise to, for example, definite x-shaped adhesion binding patterns (Figure 4.13 (h)). Such binding (and unbinding) pattern leads to fluctuating movement along northwest-southeast and northeast-southwest axis, with the resulting net migration pattern shown in Figure 4.13 (b). Similarly, the binding pattern shown in Figure 4.13 (i) with more frequent events along the equator corresponds to a mixture of diagonal and horizontal movements (Figure 4.13 (c)), as larger interfiber spacing precludes FA binding at the poles and facilitates adhesion along, as well as across the stripes in x-shaped pattern (see also Figure 4.12 (right) for illustration of a characteristic FA configuration). On the other hand, smaller spacing also leads to horizontal movement, but with more frequent vertical displacement across the stripes (Figure 4.13). These results are also in line with conclusions made in [58], where it was found that adhesion alignment determines contact guidance. We also found that the average speeds were lower than in previous scenarios (Tables 4.3, 4.4): $1.52\mu\text{m}/\text{min}$, $0.94\mu\text{m}/\text{min}$, and $0.87\mu\text{m}/\text{min}$ corresponding to, respectively, $\delta_G = 0.15, 0.25, 0.35$. Interestingly, the average speeds reported in [95] were $\sim 0.6\mu\text{m}/\text{min}$, although in that study the speeds were nearly constant for varying fiber pattern.

In Figure 4.14 we illustrate the characteristic adhesions pattern and the profiles of the FA binding rate corresponding to ECM architecture in Figure 4.12 (right). Assuming that there is a mechanical equilibrium for simplicity, we see that the adhesion pattern on the cell's periphery reflect the structure of the environment, since low values of Q_{cue} translate into low probability of focal adhesion binding. Alternatively, if the cell is positioned as in Figure 4.14 (bottom, left), then the adhesion pattern is modified accordingly. Thus, we see that our assumption about constant relative distance of FAs does not preclude the characteristic cell adhesion patterns to reflect environmental inhomogeneities (see also Figure 4.9).

Altogether, our simulations of contact guidance are, for the most part, consistent with the observations reported in the literature. In particular, we obtain the expected guidance of cell movement (Figure 4.13 (a-f)) and the geometric constraint of adhesion sites (Figure 4.13 (g-i)) by the fibrillar ECM pattern, in agreement with [92], [95].

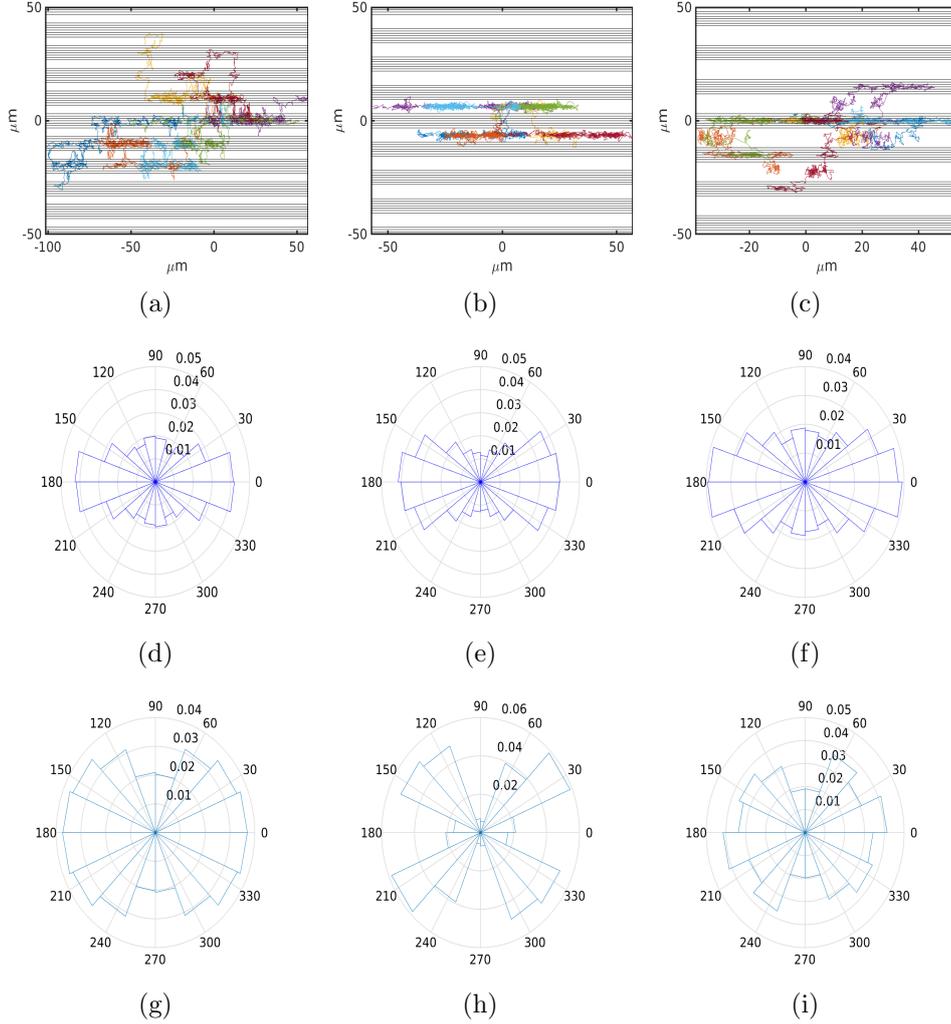


Figure 4.13: Simulation results with $M = 16$, and $\delta_G = 0.15, 0.25, 0.35$ in first, second, and third columns, respectively. (a-c) Trajectories of 7 cells and the striped ECM pattern. (d-f) Relative frequencies of normalized velocities. (g-i) Relative frequencies of binding events in each of the 16 cell sectors.

Nevertheless, since our model does not explicitly take into account morphological changes in cell shape (recall that in our model cell shape is normalized to a circle; see Section 2.1) and since cell shape control is essential to contact guidance [92], [95], increasing the interfiber distance does not necessarily lead to greater alignment of cell migration along the ECM fibers in our simulations¹. Moreover, in the case when the total number of adhesion sites is very low, the stripes are too narrow, and the separation between them is large, then it might occur that all adhesion sites “miss” the stripes, although the cell is spread over multiple stripes. In this case, the probability that any adhesion binds to the

¹The values of the guidance parameter G , defined as in [95] (See Appendix B for details), were found to be 0.64, 0.70, 0.64, corresponding to, respectively, $\delta_G = 0.15, 0.25, 0.35$.

substrate is low, which is not biologically consistent. To remedy these shortcomings, the model needs to be extended in order to accommodate strong changes to cell morphology.

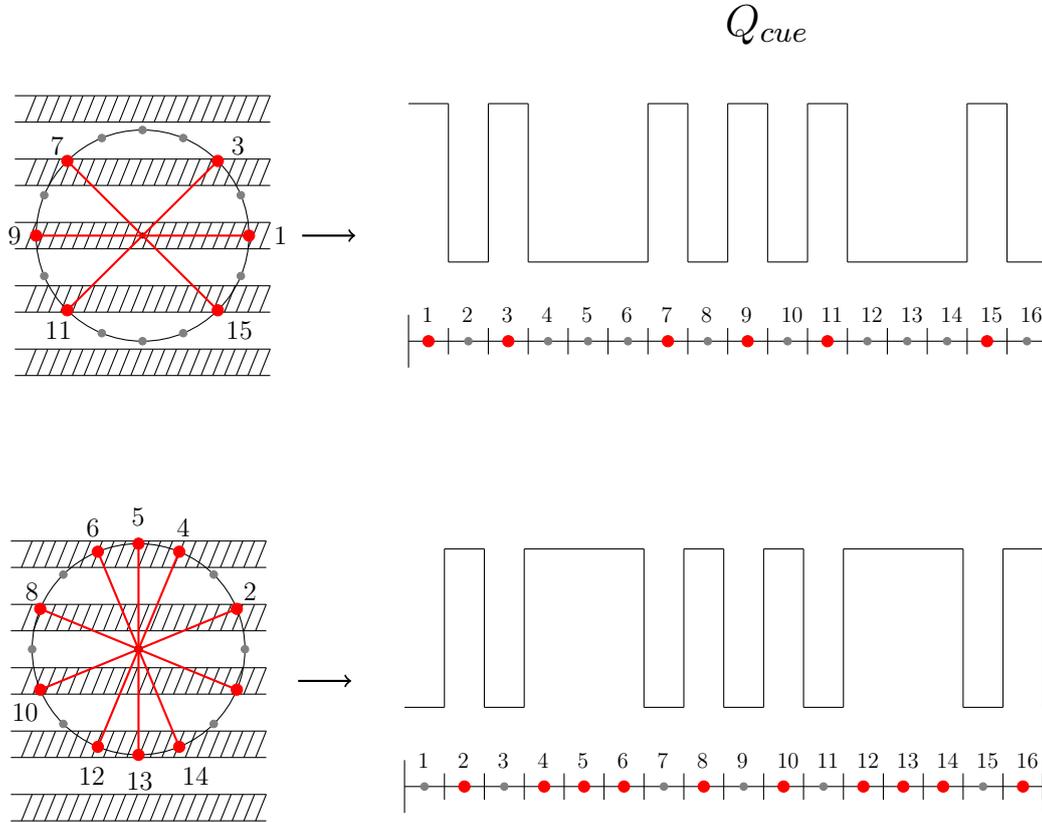


Figure 4.14: Profiles of an adhesion pattern and an external cue, projected on cell's circumference. Bound and unbound focal adhesions are depicted as red and gray circles, respectively. Stress fibers are also colored in red.

4.3.4 Asymmetric contractility

We now investigate how cell motility is influenced by asymmetrical contractile forces in a cell. Along with preferential adhesion formation, due to, for example, a chemotactic gradient, the formation of cell rear by increased actomyosin contractile activity serves as an alternative mechanism by which a directed migration can be induced in the absence of such gradient [25]. In particular, local stimulation of contractility leads to directed motility in the direction opposite to the stimulated area, even in the absence of response to chemotactic stimuli [129]. Here we show that our model is also capable of capturing such directed movement, triggered by breaking myosin mediated contractile symmetry.

Recall that T_i in equation (2.1.1) denotes the force generated by myosin motors at an adhesion site i . Instead of taking it constant, we let it vary with the radial position of an

FA. Namely, let $T_i : [0, 2\pi) \rightarrow \mathbb{R}_+$ be defined as:

$$T_i(\theta) = \begin{cases} (1 + \delta_{myo})T_{i_0}, & \text{if } \pi < \theta + (i-1)\frac{2\pi}{M} < 2\pi \\ T_{i_0}, & \text{else} \end{cases},$$

where $\delta_{myo} = 0.35, 0.40, 0.45$, T_{i_0} is the constant value used in previous simulations, and $\theta + (i-1)\frac{2\pi}{M}$ is the radial position of the i^{th} FA. Thus, contractile forces south of cell equator are larger by 35%, 40%, 45% for corresponding values of δ_{myo} . We should, therefore, expect in our simulations that the northern part of a cell becomes the front due to the imposed contractile symmetry breaking, and that cells will move accordingly (see Figure 4.15 for illustration).

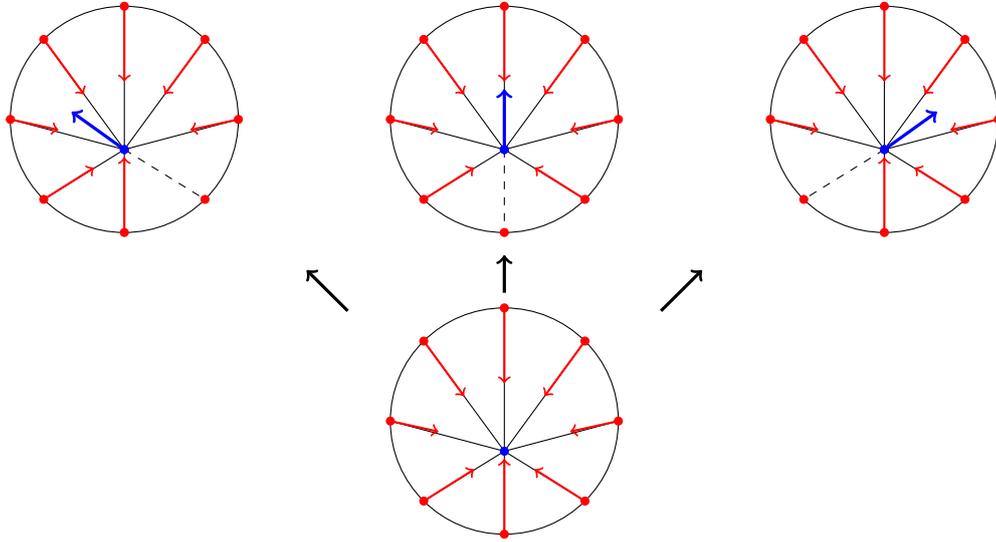


Figure 4.15: A schematic representation of asymmetric contractility. (Bottom row) Increased contractility causes \mathbf{x}_n (blue circle) to shift south of the otherwise equilibrium point in the center. (Top row) Preferential unbinding of FAs south of equator leads to directed movement indicated by blue arrows. See also Figure 2.8 (II') for a schematic representation of cell motility in case of an unbinding event.

Indeed, Figure 4.16 (a-c) shows, as expected, the trajectories of cells maintaining north-south polarity corresponding to, respectively, front and rear. Since the asymmetry of myosin forces remained during the simulations, the cell's north-south polarity also persisted, resulting in the cell movement that was highly directed along this axis, consistent with [129]. Consequently, we obtain higher values of $\beta_{av}(t)$, as shown in Figure 4.16 (d-f). In particular, for $\delta_{myo} = 0.45$, we see that the time scaling of the mean squared displacement is close to ballistic (see Table 4.5 for values of β). Moreover, increasing the number of FAs leads to more polarized, directed migration. As in the previous cases, neither speed averages (Table 4.5) nor their distribution (data not shown) changed significantly for a given number of focal adhesions. Interestingly, for $\delta_{myo} = 0.35$, the binding is relatively more

frequent in the rear (i.e. south of equator) than in the front, and unbinding is relatively more frequent in the front (i.e. north of equator) than in the back (Figure 4.16) (g-i). This suggests, then, that cells were preferentially moving in the southern direction. However, as can be seen in Figure 4.17, this is not the case. Although movements southwards are more frequent in this situation (due to the above-mentioned event frequencies), the speeds are lower than northward movements: the ratios of the average speeds directed north to the average speeds directed south were found to be 1.0165, 1.0181, and 1.0858 corresponding to, respectively, $M = 8, 16, 32$. The net effect is northward movement. For higher values of δ_{myo} , we see that the unbinding is, expectedly, more frequent in the rear, while binding is preferentially in the front.

M	8			16			32		
δ_{myo}	0.35	0.40	0.45	0.35	0.40	0.45	0.35	0.4	0.45
$\bar{\beta}, 1$	1.3072	1.6325	1.7759	1.1311	1.7905	1.8524	1.4650	1.8892	1.9353
$s_{av},$ $\mu m/min$	1.6230	1.5639	1.5103	2.3742	2.3798	2.3516	3.5861	3.7414	3.7701
$\beta_0,$ $\mu m^2/min\bar{\beta}$	0.7767	0.3659	0.2493	2.3088	0.2893	0.3037	1.0713	0.9894	1.2787

Table 4.5: Parameters obtained from the simulations with varying δ_{myo} .

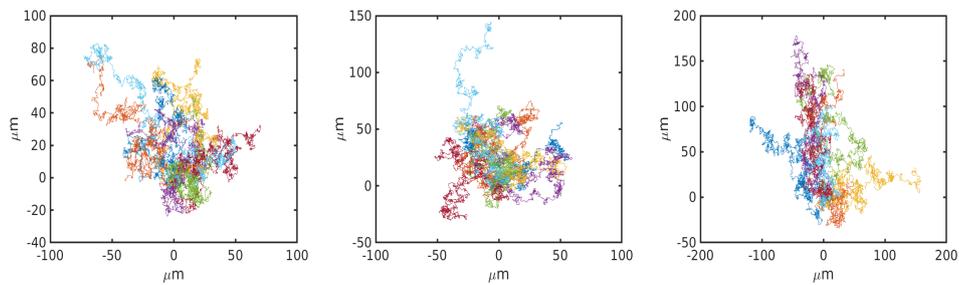


Figure 4.17: Simulated trajectories with $M = 8, 16, 32$ adhesions with $\delta_{myo} = 0.35$ on left, middle, and right plots, respectively.

These adhesion frequency patterns also illustrate the significance of the force dependence of the FA binding rate. Recalling Figure 3.2, we see that, for $\delta_{myo} = 0.4, 0.45$ (corresponding to $T_i = 1.018F_b, 1.054F_b$), the binding rate dominates unbinding north of equator due to greater SF extension (see Figure 4.15 for an illustration) leading to increased contractile force. Since the expected adhesion pattern is reversed for $\delta_{myo} = 0.35$

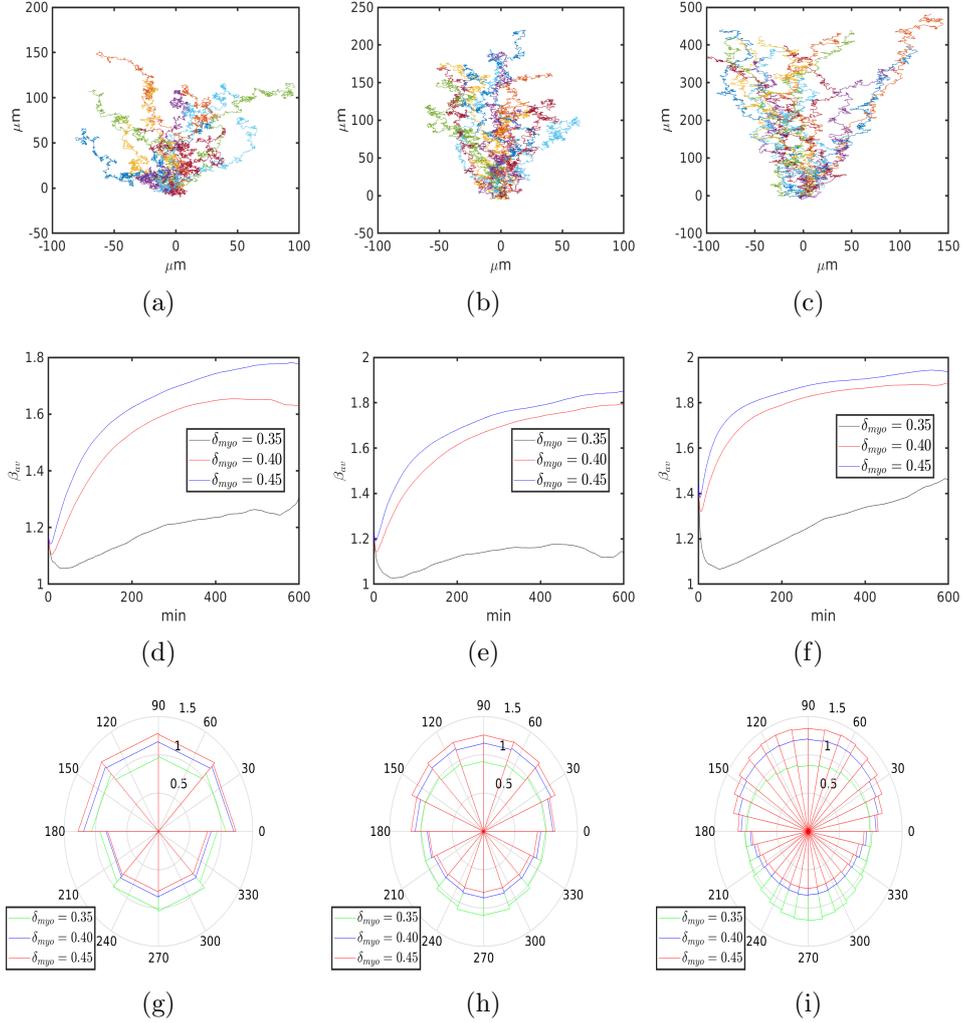


Figure 4.16: Simulation results with $M = 8, 16, 32$ adhesions in the first, second, and third columns respectively, and with various values of δ_{myo} . (a-c) Trajectories of 13 cells with $\delta_{myo} = 0.4$. (d-f) Time-average exponents β_{av} . (g-i) Ratio of the number of binding to unbinding events in each sector.

(corresponding to $T_i = 0.981F_b$) and yet the cells migrate northwards, it may suggest that there is a threshold value of δ_{myo} , above which cells can migrate in a certain direction solely by asymmetric contractility, and/or below which cells must additionally bias adhesion formation to do so.

This prompted us to investigate whether varying mechanical properties of SFs can yield the expected adhesion pattern for lower degree of asymmetry, corresponding to $\delta_{myo} = 0.35$. Specifically, we varied the buckling length L_0 and the stiffness EA such that $x = (1 + \delta_x)x^0$ corresponds to the modified value of the parameter $x \in \{L_0, EA\}$, where x^0 correspond to the default values given in Section 3.3. In Figure 4.18(a) we see that reducing the buckling

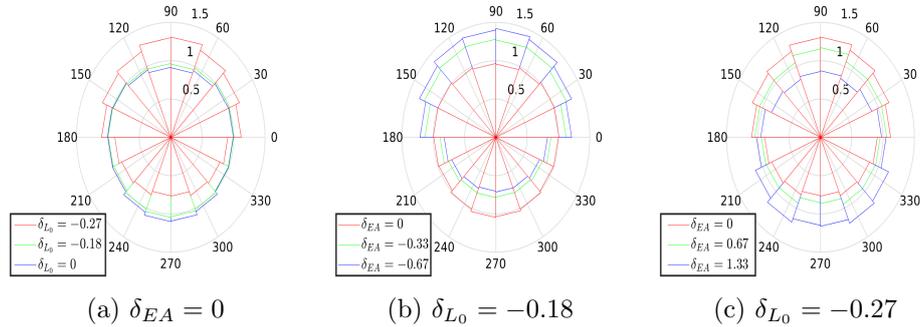


Figure 4.18: Ratios of the number of binding to unbinding events in each sector with varying buckling length and stiffness of SFs. (a) The effect of reducing the buckling length L_0 with fixed and stiffness value EA . (b-c) The effects of varying stiffness EA for reduced buckling length.

length L_0 by 27% leads to the expected adhesion pattern, while reducing it by 18% leaves it largely unchanged. However, decreasing and increasing stiffness when $\delta_{L_0} = -0.18, -0.27$, respectively, leads to the opposite results (Figure 4.18(b,c)). This suggests that if SFs are less prone to buckling and less stiff, lower degree of myosin induced contractile asymmetry may be required to drive directed migration.

Remark. Another way to induce contractile asymmetry is, for example, to decrease the myosin force T_i north of the cell's equator. Then, again, the south of the cell equator is more contractile. However, the simulated trajectories show southward directed movement (data not shown), contrary to what we should expect. Therefore, merely inducing contractile asymmetry is not sufficient. For the expected directed migration to occur, there must be a local increase of contractile forces above some critical level in the prospective cell rear, rather than a local decrease of contractility in the prospective front. Interestingly, Yam et al. [129] were able to initiate directed movement by increasing local actomyosin contraction, while locally decreasing the contractile activity did not lead to migration initiation. More recently, Shellard et al. [107] showed that directed collective cell migration of neural crest cells requires greater contractility at the rear of the clump.

Chapter 5

Contact Inhibition of Locomotion

We now turn our attention to collections of colliding cells. As discussed in the introduction, the collision may lead to contact inhibition of locomotion (CIL), whereby the colliding cells undergo repulsive interaction.

CIL can be divided into the following sequence of stages (Figure 5.1). First, after collision, the movement ceases and cadherin mediated cell-cell contacts are formed. Second, in the vicinity of the contact protrusions collapse and actomyosin contractility is enhanced, as a result of Rac1 inhibition and RhoA activation [98]. Their activity away from the collision site is altered in the opposite manner (Figure 5.2). Finally, the cells move away from each other.

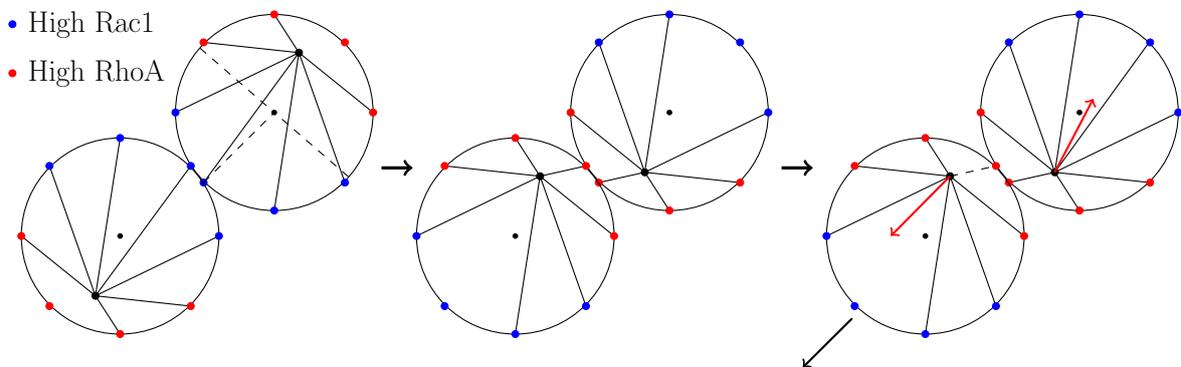


Figure 5.1: Schematic representation of CIL stages. (Left) Cells moving towards each other collide. Blue circles indicate regions of higher protrusion and FA binding activity, characteristic of cell front. Red circles indicate regions of increased contractility, characteristic for the cell rear. A diametric dashed line indicates a cell half from the point of contact, which is shown by the radial dashed line. (Middle) After the collision, both cells cease to move and repolarize, such that polarity is reflected along the former diametric dashed line. (Right) The migration cycle restarts with modified affinities for adhesion formation/rupturing. The dashed line indicates a ruptured FA. The cell on the left starts moving (ruptured FA), while that on the right remains stationary (newly formed FA).

Within the context of our cell motility model, CIL has the following consequences:

first, cessation of movement after the collision causes the cells to jump into the non-motile state. Second, activation of Rac1 leads to increased FA binding affinity away from cell-cell contacts [103] and activation of RhoA enhances myosin generated contractile forces in SFs around the collision site [98] (recall our discussion on mediation of RhoGTPases on FAs and SFs in Chapter 3).

Remark. In Chapter 3 we discussed how an external cue, like a chemoattractant, can influence FA dynamics through mediation of Rac1 (see in particular Section 3.2). Thus, the effects of CIL and an external signal can overlap. Motivated by the study in [68], we will explore this scenario in our simulations (see also Figure 5.4).

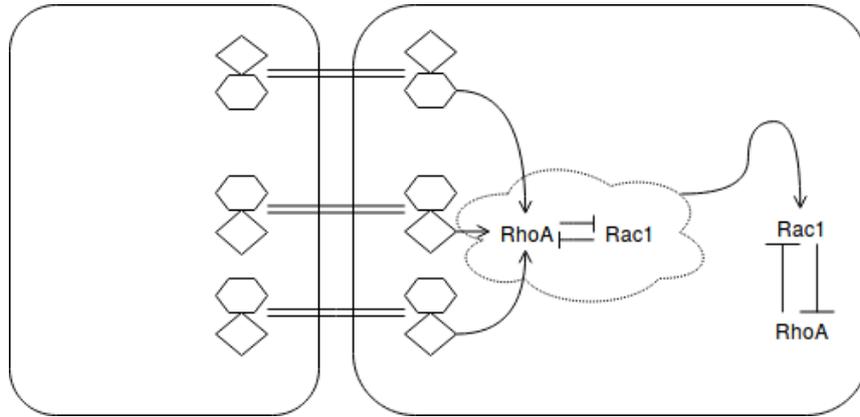


Figure 5.2: RhoGTPase switch following cell-cell contact during CIL. Intercellular bonds represent adhesion complexes like cadherins, EphA, or Frizzled transmembrane receptors. They recruit and activate intermediary agents, represented by diamonds and hexagons, which subsequently activate RhoA near the contacts. Repolarization of RhoGTPase activity occurs, in part due to antagonism between Rac1 and RhoA. See [98] for a review.

In the following sections we will integrate the CIL mechanism into our previously developed model. In this chapter we will consider binary collisions to illustrate our model of the CIL mechanism. Such binary collisions, confined to one-dimensional lanes, correspond to the experimental setup in [30], [68], [102], which we also consider in the numerical simulations here. Treatment of CIL in a general, two-dimensional setting is postponed to the subsequent Chapter 6.

This chapter is based on the work by the author in [121].

5.1 Binary collisions

Let $\bar{C}^i(t) \in \{0, 1\}$ denote the collision state¹ at time t and $\bar{\Phi}^i(t) \in [0, 2\pi)$ be the polar angle where the last contact of cell i occurred², $i \in \{1, 2\}$. Let the variables $\mu^i, \mathbf{Y}^i, \mathbf{X}^i$,

¹By collision state we mean that a cell is in contact with some other cell: $\bar{C}(t)^i = 1$ if it is in contact, and $\bar{C}(t)^i = 0$ if it is not.

² $\bar{\Phi}^i$ is constant until the next collision occurs.

corresponding to cell $i \in \{1, 2\}$ be defined as before. Let $u_j : [0, 2\pi) \times [0, 2\pi) \times \{0, 1\} \rightarrow \{0, 1\}$, $j = 1, \dots, M$, be given by:

$$u_j(\theta^i, \bar{\Phi}^i, \bar{C}^i) = \begin{cases} 1, & \bar{\Phi}^i - \frac{\pi}{2} \leq \theta^i + (j-1)\frac{2\pi}{M} \leq \bar{\Phi}^i + \frac{\pi}{2} \text{ and } \bar{C}^i = 1 \\ 0, & \text{else.} \end{cases} \quad (5.1.1)$$

This function indicates whether j^{th} FA is in the vicinity¹ of the cell-cell contact site, provided there is one.

As mentioned above, collisions lead to increased actomyosin contractility around the collision site. Thus, recalling equation (2.1.1), the tension due to myosin motors T_j is modified as follows:

$$T_j \rightarrow T_j(1 + \delta_{myo}u_j(\theta^i, \bar{\Phi}^i, \bar{C}^i)), \quad j = 1, \dots, M, \quad i = 1, 2, \quad (5.1.2)$$

where $\delta_{myo} > 0$ is a parameter that signifies the increase in myosin generated force due to increased RhoA activity. We then have $\mathbf{F}_j \rightarrow \mathbf{F}_j(\mathbf{X}^i, \bar{\Phi}^i, \bar{C}^i)$ and $\mathbf{F} \rightarrow \mathbf{F}(\mathbf{Y}^i, \mathbf{X}^i, \bar{\Phi}^i, \bar{C}^i)$ (see (2.1.1) and (2.1.2)). The propensity function a_j^+ is modified as follows:

$$a_j^+(\mathbf{Y}^i, \mathbf{X}^i) \rightarrow a_j^+(\mathbf{Y}^i, \mathbf{X}^i, \bar{\Phi}^i, \bar{C}^i)(1 + \delta_+(1 - u_j(\theta^i, \bar{\Phi}^i, \bar{C}^i))), \quad (5.1.3)$$

where $\delta_+ > 0$ is a parameter that signifies the increase in FA association rate due to increased Rac1 activity away from a contact site. Similarly, we also modify a_j^- :

$$a_j^-(\mathbf{Y}^i, \mathbf{X}^i) \rightarrow a_j^-(\mathbf{Y}^i, \mathbf{X}^i, \bar{\Phi}^i, \bar{C}^i)(1 - \delta_-(1 - u_j(\theta^i, \bar{\Phi}^i, \bar{C}^i))),$$

where $\delta_- \in [0, 1]$. Note that the dependence of a_j^\pm on $\bar{C}^i, \bar{\Phi}^i$ is due to its dependence on \mathbf{F}_j (see Section 3.1 for the form of a_j^\pm). If $\delta_- = 1$, this implies that FAs away from a contact site do not disassociate. Thus, if a cell moves, it does so necessarily away from a collision site. That is, for $\delta_- = 1$ cells do not crawl on top of one another.

For clarity, we introduce the following shorthand notation:

$$\begin{aligned} a_j^{\pm, i}(\cdot) &= a_j^\pm(\mathbf{Y}^i(\cdot), \mathbf{X}^i(\cdot), \bar{\Phi}^i(\cdot), \bar{C}^i(\cdot)) \\ a_0^2(\cdot) &= \sum_{i=1}^2 a_0(\mathbf{Y}^i(\cdot), \mathbf{X}^i(\cdot), \bar{\Phi}^i(\cdot), \bar{C}^i(\cdot)). \end{aligned}$$

Then, if \mathcal{T}_k is the time of k^{th} event, we have (see Section 6.1 for the derivation):

$$\mathbb{P}(\mathcal{T}_{k+1} - \mathcal{T}_k > \tau | \{\mathbf{Y}^i(\mathcal{T}_k), \mathbf{X}^i(\mathcal{T}_k), \bar{\Phi}^i(\mathcal{T}_k), \bar{C}^i(\mathcal{T}_k)\}_{i=1,2}) = \exp\left(-\int_0^\tau a_0^2(\mathcal{T}_k + s) ds\right), \quad (5.1.4)$$

¹By vicinity we simply mean within $\frac{\pi}{2}$ angle from the contact angle $\bar{\Phi}^i$. Here we assumed that RhoGTPases activity is modified in half of a cell.

and

$$\mathbb{P}(j^{\pm,i}|\mathcal{T}_{k+1}) = \frac{a_j^{\pm,i}(\mathcal{T}_{k+1}^-)}{a_0^2(\mathcal{T}_{k+1}^-)}, \quad (5.1.5)$$

where $\mathbb{P}(j^{\pm,i}|\mathcal{T}_{k+1})$ is the probability of binding/unbinding of j^{th} FA of cell i , given the FA event time \mathcal{T}_{k+1} . Note that between two events, \mathbf{X}^i evolves according to (2.2.5). Also, the event time \mathcal{T}_k needs not be the time when an FA reaction occurred. It is possible that at time \mathcal{T}_k a collision occurred. In this case \mathbf{Y}^i is unchanged, but $\mu^i, \bar{C}^i, \bar{\Phi}^i$ jump to new values. Figure 5.3 illustrates how cell collisions are incorporated into the cell motility model.

- (I) Suppose an FA event occurred at time $t = \mathcal{T}_{k-1}$ and cell 1 is stationary ($\mu^1 = 0$), while cell 2 is moving ($\mu^2 = 1$). Suppose \mathcal{T}_k^* is given according to (5.1.4). The evolution of \mathbf{X}^i is given by (2.2.5) until a collision occurs at time $t = t^* < \mathcal{T}_k^*$. Then $\mathcal{T}_k = t^*$.
- (II) Due to the collision, both cells become stationary ($\mu^i = 0$), the collision states and the contact angles jump to new values: $\bar{C}^1 = \bar{C}^2 = 1$ and $\bar{\Phi}^1 = 0, \bar{\Phi}^2 = \pi$ for cell 1 and 2, respectively. Then, \mathbf{X}^i follows (2.2.5) until time $t = \mathcal{T}_{k+1}$, given by (5.1.4).
- (III) At time $t = \mathcal{T}_{k+1}$ an FA event, determined by (5.1.5), occurs. If an adhesion event occurs in cell i , \mathbf{Y}^i changes accordingly, a new FA event time is found, the ODE system proceeds until this time and we are back at the same stage (III). Suppose a deadhesion event occurs. If it was in cell 1(2), then $\mu^1(\mu^2)$ jumps to a new value and cell 1(2) moves until time $t = \mathcal{T}_{k+2}$ of the next event.
- (IV, IV') Following FA rupturing in cell 2, $\mu^2 = 1$ and $\bar{C}_2 = 0$, corresponding to scenario (IV). Likewise, for an FA rupturing in cell 2, $\mu^2 = 1$ and $\bar{C}_2 = 0$, corresponding to scenario (IV'). The collision state switches since the cells are no longer in contact. In both cases, the other cell is unaffected and continues its motion.
- (V, V') Suppose the next FA event at time $t = \mathcal{T}_{k+3}$ occurred in the previously unaffected cell. Then, its collision state \bar{C}^i jumps to a new value, which is zero in this case.

There are two implicit assumptions we made. First, a cell state changes only when a collision or an FA event occurs. Second, an FA event in a cell only changes the state of a cell in which it occurred. Thus, cell 1 and 2 continue their motion away from the collision site in (IV) and (IV'), respectively, unaffected by what happened in the other cell. In particular \bar{C}^1 and \bar{C}^2 in (IV) and (IV'), respectively, remain the same, since at the onset of post-collision motion in (III) the cells are still in contact. When an event occurs in (V) and (V'), the corresponding collision states are switched as cells are no longer in contact, while the other cells continue their motion. Note that whether cells move in the same or opposite directions after collisions is determined stochastically in our model, which is in line with [30], [68], [102].

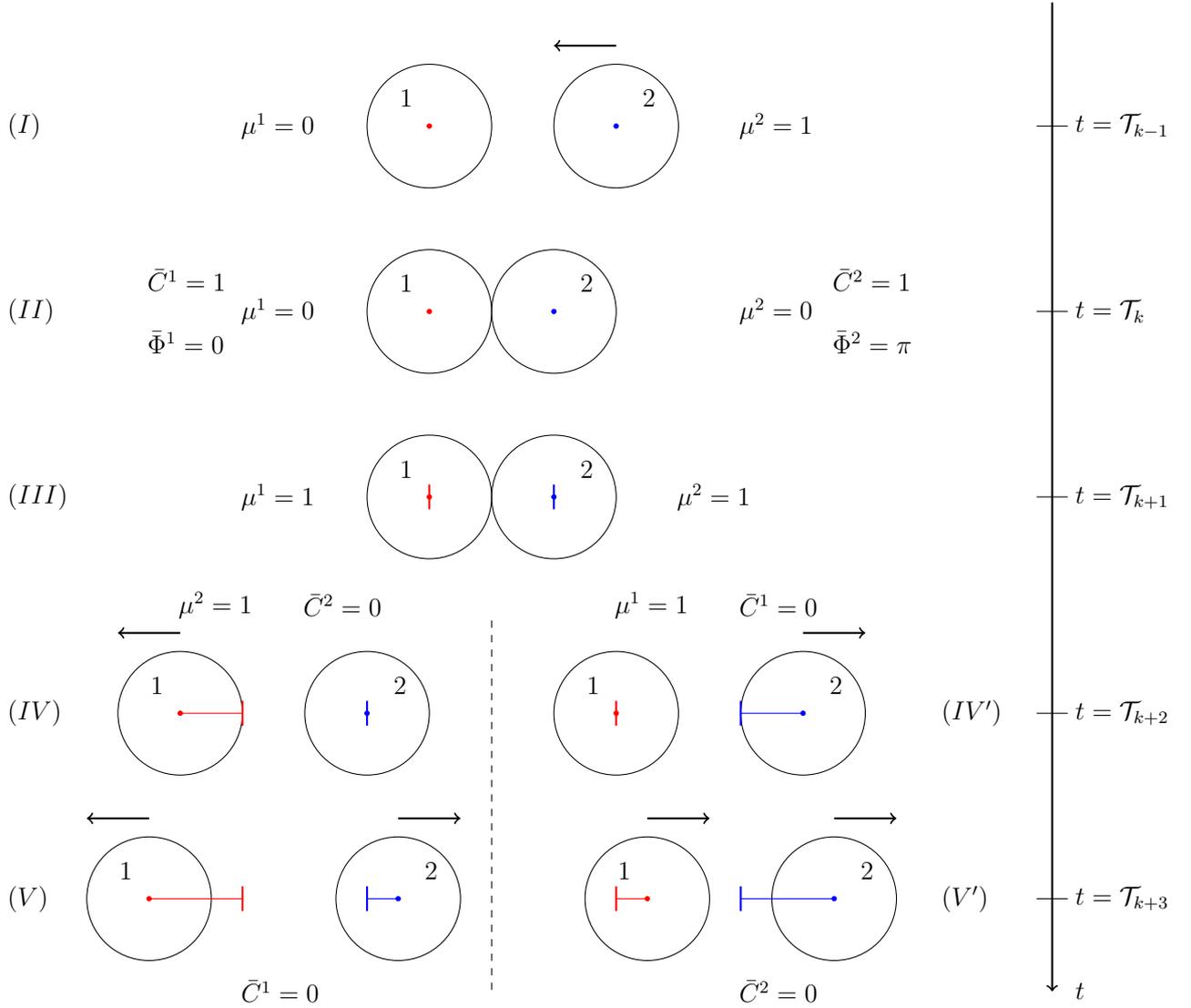


Figure 5.3: Schematic representation of binary collisions. (I) Cell 2 moves in the direction of cell 1. The centroids \mathbf{x}^i are indicated by blue and red dots, respectively. (II) Collided cells become stationary. (III) An event occurs at time $t = \mathcal{T}_{k+1}$. Vertical bars indicate centroid positions at the collision time. An FA unbinding event in cell 1 or 2 leads to outcome (IV) or (IV'), respectively. An adhesion event leads back to (III). (IV, IV') Another event occurs at time $t = \mathcal{T}_{k+2}$. FA rupturing in cell 2 or 1 leads to movement in the opposite (V) or the same directions (V'), respectively, until the next FA event occurs at time $t = \mathcal{T}_{k+3}$.

Remark. We assume, more generally, that cell interactions occur solely by collisions and that there is no coupling of cells before or after they interact. That is, neither the equations of motion (2.2.5) between the events, nor the probabilities (5.1.5) of FA events in a cell depend on the state of another cell. This can be justified by the results in [30], where it was found that CIL response in cells is statistically independent.

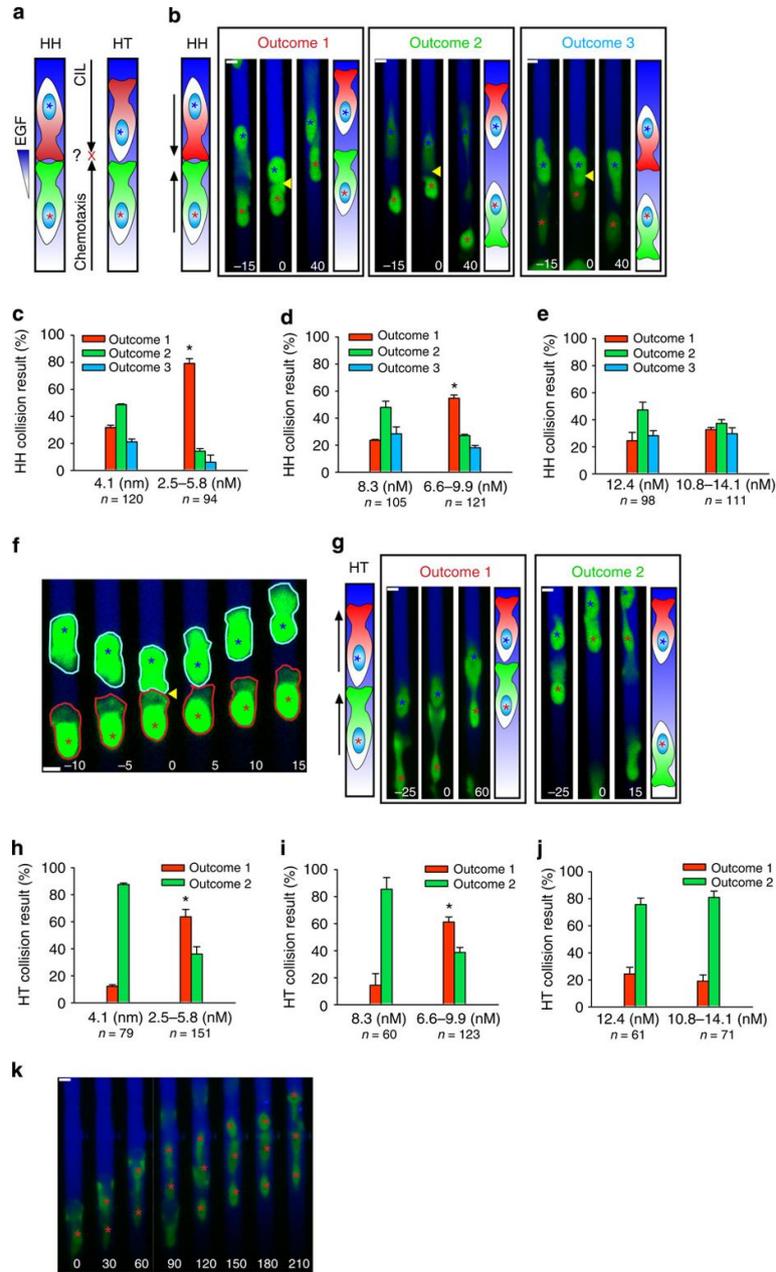


Figure 5.4: A result of the study in [68], showing modulation of CIL response in the presence of an external signal (EGF in this case), and motivating a similar scenario in our simulations. (a) Schematic showing how the repulsive response may conflict with chemotaxis. (b,g) Examples of post collision response in head-to-head and head-to-tail collisions, respectively. Collisions occur at time $t = 0$. (c-e) and (h-j) Probabilities for each outcome in head-to-head and head-to-tail collisions, respectively. In each figure, the corresponding probabilities in the case of uniform and increasing EGF concentrations are shown. (f) Example of head-to-head collisions, where protrusions of the trailing cell are not affected by the cell-cell contact. Collision occurs at time $t = 0$. (k) A group of cells, streaming up the EGF gradient. Time is in minutes and the scale bar is $10\mu m$.

5.2 Simulations

The illustration in Figure 5.3 depicts binary collisions in one-dimensional tracks. As noted in [30], [68], [102], this setup allows for a more efficient study of the CIL mechanism. In particular, it allows for unambiguous quantification of collision outcomes for measuring the CIL response. As in [30], [68], we classify the outcomes into two categories. Namely, outcome 1 and 2 leading to cells moving in the opposite and the same directions, respectively, as illustrated in Figure 5.3 (V, V'). In order to investigate these outcomes, we introduce the following quantities:

- The distance between the cell centroids $d(t) := |x_1^1(t + t_*) - x_1^2(t + t_*)|$ at time t after the first cell collision, where x_1^i is the x -component of \mathbf{x}^i , $i = 1, 2$ and t_* is the time of the collision.
- Define $d^i(t) := x_1^i(t + t_*) - x_1^i(t_*)$, illustrated in Figure 5.3 as the difference between the red (blue) dot and red (blue) vertical bar.

Note that restriction to movement in lanes implies that the first equation in (2.2.5) is modified as follows:

$$\dot{\mathbf{x}} = \mu\beta_{ECM}^{-1}\hat{\mathbf{e}}_1 \cdot (\mathbf{F} \cdot \hat{\mathbf{r}}\hat{\mathbf{r}}),$$

where $\hat{\mathbf{e}}_1 = (1, 0)^T$, i.e. the cells move in horizontal direction only.

Consider Figure 5.3 (V, V'). If the cells are moving in opposite directions (Figure 5.3 (V)), then $d^1(t)$ and $d^2(t)$ have opposite signs - negative and positive, respectively. If the two cells are moving in the positive (negative) x -direction, then $d^i(t) > 0$ ($d^i(t) < 0$), for $i = 1, 2$. Note that while $d(t)$ is used as a readout of CIL in [102]¹, where its increase with time was used as an indication that cells are moving in opposite directions and hence undergoing CIL, $d^i(t)$ allows to distinguish between outcome 1 and 2. Moreover, increasing $d(t)$ might simply indicate that one cell is faster than the other, while both are moving in the same direction.

Note that we introduced three new parameters in addition to the single cell motility model, namely, δ_{myo} , δ_+ , and δ_- . Their magnitude indicates the strength of CIL repolarization signal upon collision. Below we perform numerical simulations with varying values of δ_{myo} , δ_+ in the absence of an external cue, and in the presence of a chemotactic gradient with varying strength (mimicking the experimental setup in [68], see Figure 5.4) and fixed δ_{myo} , δ_+ . For each scenario we simulate 64 pairs of cells for 20 hours of simulation time. Initially, the distance between the cell centroids is $2.4R_{cell}$, and the initial values for the x -components of the centroids are $x_1^1(0) = 1$ and $x_1^2(0) = 3.4$ for cell 1 and 2, respectively. Here, we also set $\delta_- = 0$, as we would like to explore hallmarks of CIL (contraction of the leading edge and FA activation away from it) specifically in the absence of volume exclusion. The initial conditions for other variables and parameter values are taken as in Section 4.3. We also set the number of adhesion sites $M = 16$.

¹In [102] the distance between cell nuclei, rather than cell centroids, was measured.

Remark. Among other factors, the collision outcome depends on whether it was a head-to-tail or a head-to-head collision [30], [68] (see Figure 5.4). Here, we analyze the outcomes in terms of effects CIL has on FA dynamics and SF contractility.

Absence of an external cue

Here we investigate three scenarios corresponding to three pairs of values for δ_{myo} and δ_+ . Similar values were used in Section 4.3 to simulate directed movement.

Parameters	S1	S2	S3
δ_{myo}	0.3	0.4	0.5
δ_+	0.1	0.2	0.3

Table 5.1: Parameter values corresponding to three scenarios S1-S3.

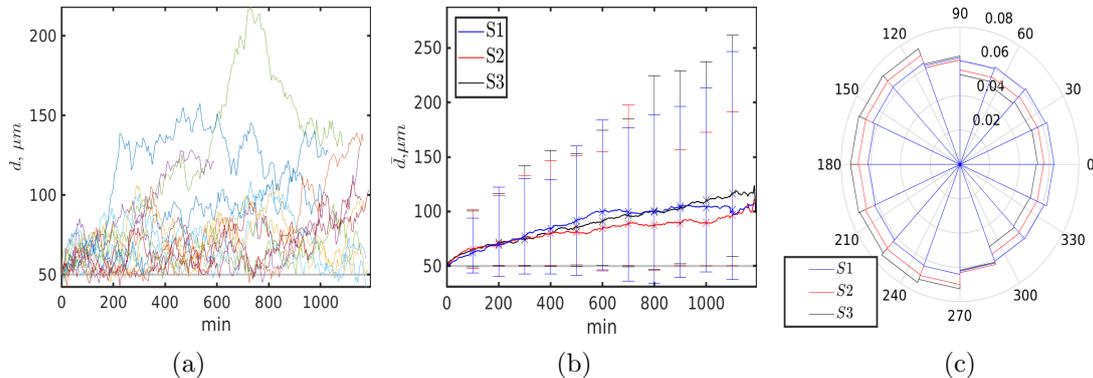


Figure 5.5: The opaque horizontal line indicates the distance of $2R_{cell}$, i.e. the cells in contact. (a) Cell centroid distance d of 14 cell pairs corresponding to scenario S2. (b) Ensemble averages \bar{d} for each scenario. The corresponding error bars indicate ensemble minimum and maximum. (c) Relative frequency of binding events of cells with $\bar{C}^i = 1$ and colliding at 0° . Each sector corresponds to a single FA counting counterclockwise. Cell 2 is accounted for by reflection about south-north axis.

Since in our model the cells are not treated as hard spheres, it is possible that some overlaps may occur (Figure 5.5a,b). However, the slight overlap is followed by an increase in d and separation (Figure 5.5a). Although the average distances \bar{d} are similar (Figure 5.5b), increasing δ_{myo} and δ_+ leads to a stronger response: the minimum of d is consistently lower for S1 compared to S3 (Figure 5.5b) and FA formation away from the collision site is more frequent for S3 (Figure 5.5c). Notice that the cells need not obey the volume exclusion principle for eventual separation to occur and the stronger response in S3 implies that the separation can be modulated by modifying contractility and FA formation.

Since increasing d only suggests that the cells are separating, we examined their relative direction of motion after collision (Figure 5.6). Ensemble averages of d^i in Figure 5.6d show

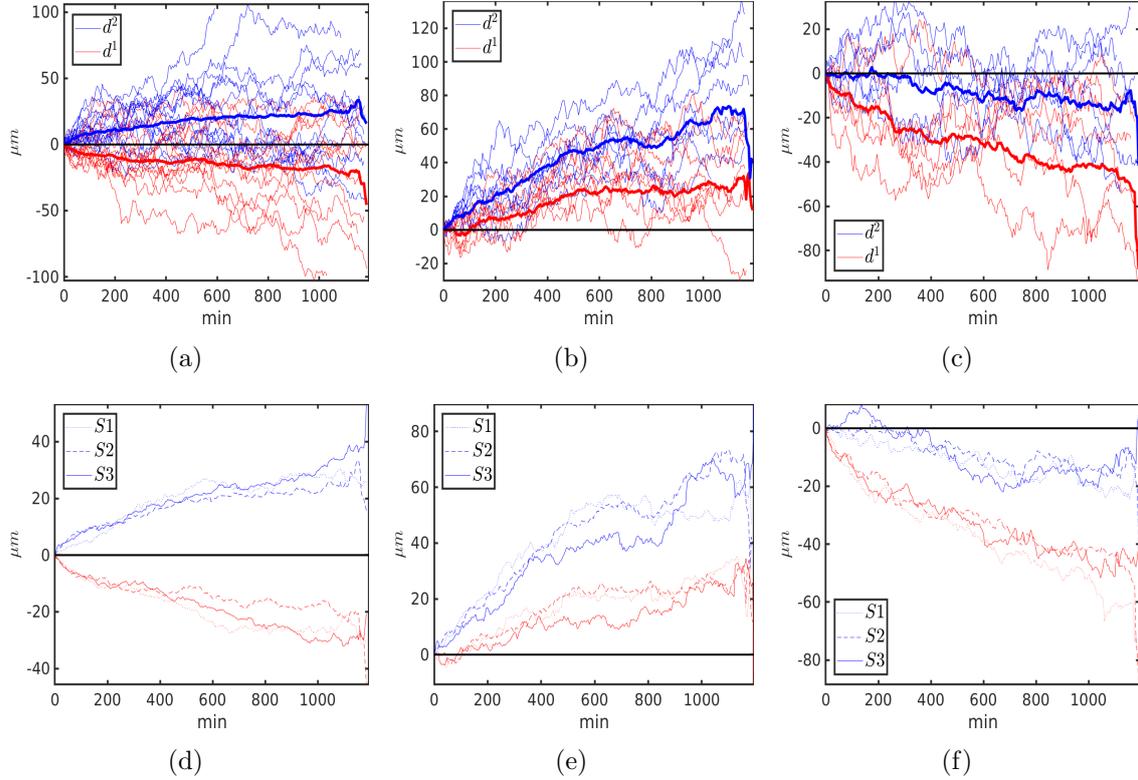


Figure 5.6: (a-c) The differences d^i corresponding to scenario S2. Thick lines represent the corresponding ensemble averages. (a) A sample of 14 pairs. (b) and (c) The differences d^i , whose averages over time are positive and negative, respectively. Samples of 8 and 6 pairs are shown, respectively. (d-f) Ensemble averages of d^i for each scenario. Blue and red colored plots correspond to cells 1 and 2, respectively. (e) and (f) d^i with positive and negative times averages.

that following collisions, the movement in the opposite directions is prevalent, which is in line with results in [30], [68]. It may also occur that cells follow one another after collision, as indicated by positive and negative time averages of d^1 and d^2 (Figure 5.6b,c). The ensemble averages in Figure 5.6(d-f) do not show a strong difference between the scenarios S1-S3. This suggests that varying the strength of cell response to collision does not have a significant effect on the relative direction of migration after the collision. In our simulations, 56% of collided pairs moved in the opposite directions, compared to $\sim 65\%$ in [68].

Remark. The collision times t_* for each simulated pair are different. Thus, the number of cells at time t after the collision time t_* varies, and reduces towards the terminal time. This skews the values for ensemble averages and causes the abrupt changes in Figure 5.6.

Note that a freely migrating cell before collision is equally likely to move in either direction, as indicated by a rapid decay of normalized velocities to zero in Figure 5.7a. How fast does a cell become freely migrating after a collision? Figure 5.7b shows a much slower decay of the normalized velocities for the three scenarios. This suggests that either

there are frequent follow up collisions after the first one, resulting in cell 1(2) moving left(right), or collisions lead to persistent movement in the opposite direction. It must be the latter, since in light of our results, cells separate (Figure 5.5b) and move away from each other (Figure 5.6d). Thus, in our model transient perturbations in cell motility lead to persistent, but decaying, alterations in migration dynamics. This is unexpected, since the collision state \bar{C}^i of a cell is switched off after separation, i.e. the cell migrates freely. However, studies in [68] and [102] indicate that cells continue to move in opposite directions even after separation occurs.

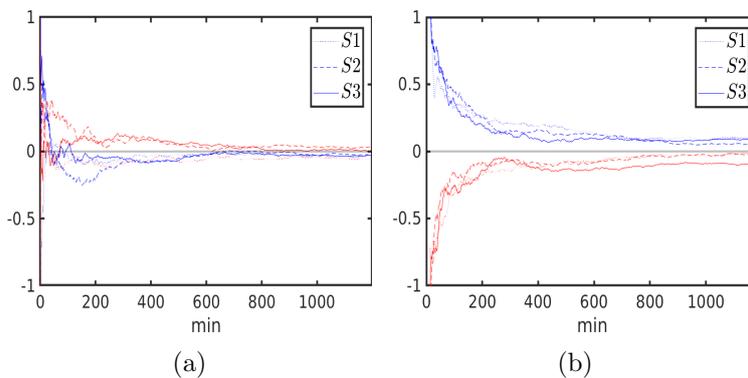


Figure 5.7: (a) and (b) Ensemble averages of normalized velocities before and after collisions, respectively. Plots in red and blue correspond to cells 1 and 2, respectively. In (a) non-colliding pairs have also been accounted for.

Presence of a chemotactic gradient

We now explore how collision outcomes are affected in the presence of a chemotactic gradient, as experimental evidence in [68] suggest that CIL response is modulated by the strength of the external signal. As before, we suppose that $a_j^+ \propto Q_{cue}$, i.e. the binding probability of the j^{th} FA is proportional to the (local) concentration of chemoattractant Q_{cue} at the position of the FA. We assume that Q_{cue} has the following form:

$$Q_{cue}(\mathbf{x}) = \begin{cases} 1 + \delta_E x_1, & \text{if } x_1 > 0 \\ 1, & \text{else} \end{cases},$$

where \mathbf{x} is the position of an FA (in units of R_{cell}) in the lab reference frame, and $\delta_E > 0$ indicates strength of the signal, i.e. there is a chemotactic gradient in the positive x -direction. We also take $\delta_{myo} = 0.4$, $\delta_+ = 0.2$.

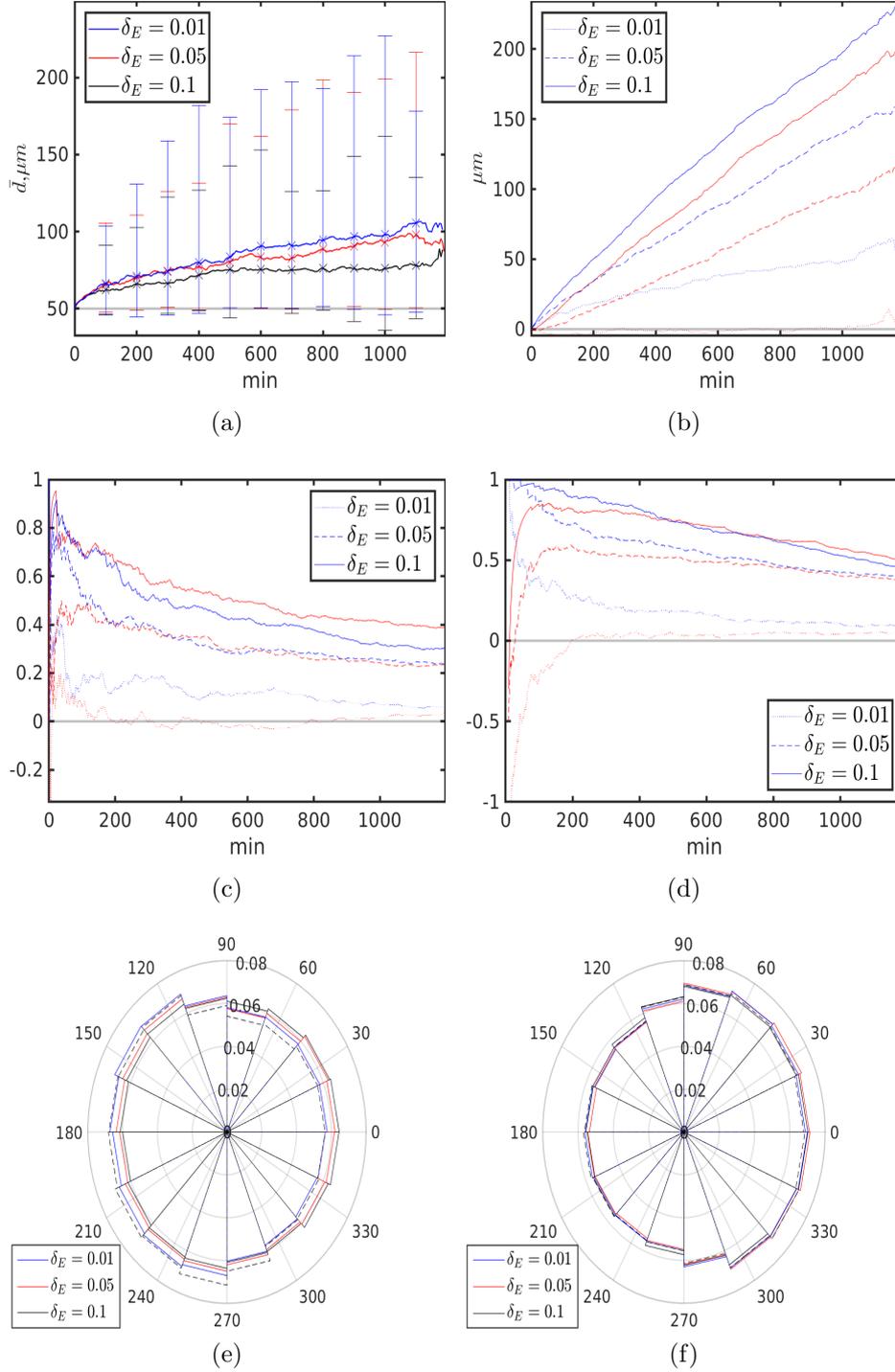


Figure 5.8: The effect of varying chemotactic signal strength δ_E . (a) Ensemble averages of cell-cell distances. (b) Ensemble averages of d_1 (red) and d_2 (blue). (c) and (d) Ensemble averages of normalized velocities before and after collisions. (e) and (f) Relative frequency of binding events after collision of cells 1 and 2. Dashed lines correspond to $\delta_E = 0$.

The influence of a chemotactic signal on CIL can be seen in Figure 5.8. We see that increasing the signal strength reduces average cell-cell separation (Figure 5.8a). Although the difference between the averages is slight (relative to cell size), the variance (as indicated by the error bars) of cell-cell distances is noticeably smaller for the case of the strongest signal. Moreover, after the collision, cells tend to move in the same direction following the signal, as shown in Figure 5.8b. This, together with what appears to be a plateauing of cell-cell distance (Figure 5.8a), suggests emergence of collective movement. Observe that reducing the signal strength leads to reduced propensity of cells to move in the same direction, in line with the results reported in [68]. Note that in [68] three scenarios with different EGF concentrations were explored. There, the gradients of EGF concentration were kept constant at $3.3nM$ per length of the lane. However, the relative changes in EGF concentrations were $\frac{5.5nM-2.2nM}{2.2nM} = 1.6$, $\frac{9.9nM-6.6nM}{6.6nM} = 0.5$, $\frac{14.1nM-10.8nM}{10.8nM} = 0.3$ and reduced relative changes led to diminished alteration of a typical CIL response, which our simulations show as well.

Motion alignment is not immediate, as the amount of time during which cells move in the opposite directions after collision depends on the gradient magnitude (Figure 5.8e), compared to a rapid velocity alignment of uncollided pairs (Figure 5.8b).

We also see that the effect on adhesion dynamics of cells to the left and to the right of a collision point is different (Figure 5.8e,f). If the CIL signal in a cell and the chemotactic gradient are in the opposite directions, the affinity of FA association away from the contact reduces with increasing gradient strength (Figure 5.8e). However, if the signals are aligned, the FA binding dynamics does not appear to be significantly modified (Figure 5.8f). This suggests that in relation to adhesion dynamics, the chemotactic cue either reduces CIL response or has little to no effect. Interestingly, in [117] it was shown that elevated Rac1 activity (and hence enhanced adhesion to a substrate) away from the contact site (and in a free edge) is primarily due to cell-cell contacts, rather than to a chemoattractant.

Chapter 6

CIL in Collectives

Here we construct a general motility model with N colliding cells. To do so, we proceed as in Sections 2.3-2.4. In particular, we first provide a formal derivation of the survival function for the next event time and the distribution of the next event index for N cells (the special case of which is given in (5.1.4)-(5.1.5)). Then we formulate our motility model as a piecewise deterministic Markov process (see [28] for a comprehensive treatment) including collisions. Finally, we present the numerical method used to simulate the collisions in an unconfined setting, and the results thereof.

This chapter is based on the work by the author in [121].

6.1 Preliminaries

Let N be the number of cells and let μ^i, \mathbf{Y}^i , $i = 1, \dots, N$, be defined as in the previous section, and let $\boldsymbol{\mu} = (\mu^1, \dots, \mu^N)$. Let $\mathbf{C}^i(t) \in \{0, 1\}^N$ denote the collision state of cell i at time t with other cells:

$$C_k^i = \begin{cases} 1, & \text{if cell } i \text{ is in collision state with cell } k \\ 0, & \text{else,} \end{cases}$$

where $k = 1, \dots, N$ and we assume that $C_i^i = 0$. Let $\boldsymbol{\Phi}^i(t) \in [0, 2\pi)^N$ denote the vector of collision angles of cell i with other cells, such that $\Phi_i^i(t) = 0$. For $N = 2$ in Section 5.1, for example, we have $\mathbf{C}^1(\mathcal{T}_k) = (0, 1)$ and $\boldsymbol{\Phi}^2(\mathcal{T}_k) = (\pi, 0)$. For ease of notation, let $\mathbf{X}^i = (\mathbf{x}^i, \mathbf{x}_n^i, \theta^i, \boldsymbol{\Phi}^i, \mathbf{d}^i)$, where \mathbf{d}^i is defined in (6.2.3), and $\mathbf{A} = (\mathbf{A}^1, \dots, \mathbf{A}^N)$ for $\mathbf{A} \in \{\mathbf{Y}, \mathbf{C}, \mathbf{X}\}$.

Since there are N cells and $2M$ possible reactions for each cell (binding and unbinding of an FA), then there are $2MN$ possible reactions among all cells. Let $a_{j'}(\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t))dt$ be the probability, given $\mathbf{X}(t)$ and $\mathbf{A}(t) = \mathbf{a}$, for $\mathbf{A} \in \{\boldsymbol{\mu}, \mathbf{Y}, \mathbf{C}\}$ and $\mathbf{a} \in \{\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}\}$, that a reaction $j' = 1, \dots, 2MN$ will occur in the time interval $[t, t + dt)$.

Finally, let $K_{time}(\tau|t, \bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t))d\tau$ be the probability that a reaction occurs in the time interval $[t + \tau, t + \tau + d\tau)$ and let $K_{index}(j'|t, \tau, \bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t))$ be the probability of reaction j' , given that it occurs at time $t + \tau$. Applying Lemma 2.3.1 and Proposition 2.3.2

we have the following (as in (2.3.3)):

$$\begin{aligned} K_{time}(\tau|t, \bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t)) \\ = a_0^N(\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t + \tau)) \exp\left(-\int_t^{t+\tau} a_0^N(\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t + \tau'))\tau'\right) \end{aligned} \quad (6.1.1)$$

and

$$K_{index}(j'|t, \tau, \bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t)) = \frac{a_{j'}(\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t + \tau))}{a_0^N(\bar{\boldsymbol{\mu}}, \mathbf{y}, \mathbf{c}, \mathbf{X}(t + \tau))}, \quad (6.1.2)$$

where $a_0^N = \sum_{j'=1}^{2MN} a_{j'}$. Here, we adopt the following convention:

- A reaction j' occurs in cell i if $i = \lfloor \frac{j'-1}{2M} \rfloor + 1$.
- A reaction j' corresponds to a binding reaction of j^{th} FA if $2j - 1 = j' \bmod N$, and to an unbinding reaction of j^{th} FA if $2j = j' \bmod N$.

Thus, $a_j^{+,i} = a_{2j-1+2M(i-1)}$ and $a_j^{-,i} = a_{2j+2M(i-1)}$ correspond, respectively, to binding and unbinding probability rates of the j^{th} FA of cell i . For an example utilizing the above, see the special case with $N = 2$ in Section 5.1.

6.2 CIL and PDMP

Let $A := \{1, \dots, 2^{N+MN+N^2}\}$ and let $\boldsymbol{\alpha} : A \rightarrow \{0, 1\}^N \times \{0, 1\}^{MN} \times \{0, 1\}^{N^2}$ be a bijection. This is a mapping such that $\boldsymbol{\alpha}(\nu) = (\boldsymbol{\mu}, \mathbf{Y}, \mathbf{C})$ corresponds to motility, FA, and collision states of N cells.

Let $\nu \in A$ and $\boldsymbol{\alpha}(\nu) = (\boldsymbol{\mu}, \mathbf{Y}, \mathbf{C})$. Let $D_\nu^{i,k} \subset \mathbb{R}$, $D_\nu^i \subset \mathbb{R}^N$, $D_\nu \subset \mathbb{R}^{N^2}$ be defined as:

$$D_\nu^{i,k} := \begin{cases} (-\infty, 1], & \text{if } C_k^i = 0 \text{ and } i, k \in I \\ (-\infty, \infty), & \text{else} \end{cases}, \quad (6.2.1)$$

$$D_\nu^i := \prod_{k=1}^N D_\nu^{i,k}, \quad D_\nu := \prod_{i=1}^N D_\nu^i, \quad (6.2.2)$$

where $I \subset \{1, \dots, N\}$ is the index set of cells exhibiting CIL. Let $\mathbf{d}^i : [0, \infty) \rightarrow D_\nu^i$, $i = 1, \dots, N$ be defined as:

$$d_k^i(t) := \exp\left(2R_{cell}^2 - \frac{1}{2}\|\mathbf{x}^i(t) - \mathbf{x}^k(t)\|^2\right), \quad k = 1, \dots, N, \quad (6.2.3)$$

and let $\mathbf{d} := (\mathbf{d}^1, \dots, \mathbf{d}^N) \in D_\nu$. This particular form of d_k^i is chosen since it satisfies the following requirements, which we impose on d_k^i :

- d_k^i must be a measure of distance between cells i and k , such that it attains a unique value when the cells are in contact (in our case the value is one), and such that a certain range of values correspond to the case when the cells overlap.
- d_k^i must be bounded and continuously differentiable.

Depending on the form of d_k^i , D_k^i must be modified accordingly.

Let $\mathbf{X}^i := (\mathbf{x}^i, \mathbf{x}_n^i, \theta^i, \Phi^i, \mathbf{d}^i) \in \mathbb{R}^2 \times \Omega_{\text{cell}} \times [0, 2\pi) \times D_\nu^i \times [0, 2\pi)^N := E_\nu^i$. For convenience of notation, we define $\alpha_{\mathbf{A}}^i(\nu) := \mathbf{A}^i$, where $\mathbf{A} \in \{\boldsymbol{\mu}, \mathbf{Y}, \mathbf{C}\}$. We also extend the definition of u_j in (5.1.1):

$$u_j(\mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)) := \begin{cases} 1, & \Phi_k^i - \frac{\pi}{2} \leq \theta^i + (j-1)\frac{2\pi}{M} \leq \Phi_k^i + \frac{\pi}{2} \text{ and } C_k^i = 1, \\ 0, & \text{else,} \end{cases}$$

for some $k = 1, \dots, N$ and where $j = 1, \dots, M$. Then, we have:

$$\begin{aligned} T_j &\rightarrow T_j(1 + \delta_{\text{myo}} u_j(\mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu))) \\ \mathbf{F}_j &\rightarrow \mathbf{F}_j(\mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)) \\ \mathbf{F} &\rightarrow \mathbf{F}(\alpha_{\mathbf{Y}}^i(\nu), \mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)). \end{aligned}$$

Let $E_\nu := \prod_{i=1}^N E_\nu^i$ and define $\mathbf{H}_\nu^i : E_\nu \rightarrow \mathbb{R}^{5+2N}$ as:

$$\frac{d}{dt} \mathbf{X}^i = \begin{pmatrix} \alpha_{\boldsymbol{\mu}}^i(\nu) \beta_{ECM}^{-1} \mathbf{F}(\alpha_{\mathbf{Y}}^i(\nu), \mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)) \cdot \hat{\mathbf{r}}(\mathbf{x}_n^i) \hat{\mathbf{r}}(\mathbf{x}_n^i) \\ \beta_{\text{cell}}^{-1} \mathbf{F}(\alpha_{\mathbf{Y}}^i(\nu), \mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)) \\ \alpha_{\boldsymbol{\mu}}^i(\nu) \beta_{\text{rot}}^{-1} \|\mathbf{x}_n^i\| \mathbf{F}(\alpha_{\mathbf{Y}}^i(\nu), \mathbf{X}^i, \alpha_{\mathbf{C}}^i(\nu)) \cdot \hat{\boldsymbol{\varphi}}(\mathbf{x}_n) \\ \mathbf{0} \\ -(\mathbf{x}^i - \mathbf{x}^1) \cdot (\dot{\mathbf{x}}^i - \dot{\mathbf{x}}^1) d_1^i \\ \vdots \\ -(\mathbf{x}^i - \mathbf{x}^N) \cdot (\dot{\mathbf{x}}^i - \dot{\mathbf{x}}^N) d_N^i \end{pmatrix} := \mathbf{H}_\nu^i(\mathbf{X}). \quad (6.2.4)$$

This is simply an ODE system that governs the evolution of \mathbf{X}^i between events (see Section 2.4.2 for the case with single cell migration). Note that Φ^i changes only when collisions occur, and is constant at all other times. For a collection of N cells, we then have:

$$\begin{aligned} \frac{d}{dt} \mathbf{X}_t &= \mathbf{H}_\nu(\mathbf{X}_t) \\ \mathbf{X}_0 &= \mathbf{Z} \in E_\nu \end{aligned} \quad (6.2.5)$$

where $\mathbf{H}_\nu : E_\nu \rightarrow \mathbb{R}^{5N+2N^2}$ and $\mathbf{H}_\nu := (\mathbf{H}_\nu^1, \dots, \mathbf{H}_\nu^N)$. One can also show that there exists a unique solution to (6.2.5) for fixed $\nu \in A$, by using Proposition 2.4.2 and noting that \mathbf{H}_ν^i is Lipschitz continuous.

Let $\phi_\nu : \mathbb{R}_+ \times E_\nu \rightarrow E_\nu$ be the flow corresponding to (6.2.5). Note that a cell i collides with a cell k , if $d_k^i \in \partial D_\nu^{i,k} = \{1\}$ for some $\nu \in A$ such that $C_i^k = 0$. Thus, the boundary

of E_ν plays an important role in addressing the collisions. Let ∂E_ν denote the boundary of E_ν , and define $\partial^* E_\nu, \Gamma^*$ as:

$$\begin{aligned}\partial^* E_\nu &:= \{\mathbf{X} \in \partial E_\nu : \phi_\nu(t, \mathbf{Z}) = \mathbf{X}, (t, \mathbf{Z}) \in \mathbb{R}_+ \times E_\nu\} \\ \Gamma^* &:= \{(\nu, \mathbf{X}) : \nu \in A, \mathbf{X} \in \partial^* E_\nu\}.\end{aligned}$$

Let $E := \{(\nu, \mathbf{X}) : \nu \in A, \mathbf{X} \in E_\nu\}$ and define $t^* : E \rightarrow \mathbb{R}_+$ as:

$$t^*(\nu, \mathbf{X}) = \inf\{t > 0 : \phi_\nu(t, \mathbf{X}) \in \partial^* E_\nu\}.$$

Here, t^* is simply the next collision time, given the state of the system $(\nu, \mathbf{X}) \in E$. Let $a_0^N : E \rightarrow \mathbb{R}_+$ be defined as above:

$$a_0^N(\nu, \mathbf{X}) = \sum_{j=1}^{2MN} a_j(\nu, \mathbf{X}),$$

where for ease of notation we write $a_0^N(\nu, \mathbf{X}) = a_0^N(\boldsymbol{\alpha}(\nu), \mathbf{X})$ and $a_j(\nu, \mathbf{X}) = a_j(\boldsymbol{\alpha}(\nu), \mathbf{X})$ for $(\nu, \mathbf{X}) \in E, j = 1, \dots, NM$. We define

$$\mathcal{E} = \{B : B = \{(\nu, \mathbf{X}) : \nu \in A, \mathbf{X} \in \tilde{E}_\nu\}, \tilde{E}_\nu \in \mathcal{E}_\nu\},$$

where \mathcal{E}_ν denotes the Borel sets of E_ν . As in Section 2.4.1, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space and let (E, \mathcal{E}) be the state space of the stochastic process. Define a transition measure $Q : \mathcal{E} \times E \cup \Gamma^* \rightarrow [0, 1]$.

We now have all the ingredients to specify and construct a piecewise deterministic process of cell motility including collisions. These objects satisfy the following standard conditions given in (24.8), Section 24 in [28]:

1. The vector fields $(\mathbf{H}_\nu, \nu \in A)$, given by (6.2.4), governing the system's evolution between events are such that there exists a unique global solution to (6.2.5).
2. An intensity function a_0^N , determining the arrival times of FA events, such that $s \mapsto a_0^N(\nu, \phi_\nu(s, \mathbf{X}))$ is integrable for $(\nu, \mathbf{X}) \in E$.
3. A transition measure Q (to be specified below), determining the system's state after an event, such that $(\nu, \mathbf{X}) \mapsto Q(B, (\nu, \mathbf{X}))$ is measurable for fixed $B \in \mathcal{E}$; $Q(\cdot, (\nu, \mathbf{X}))$ is a probability measure for $(\nu, \mathbf{X}) \in E$; $Q(\{(\nu, \mathbf{X})\}, (\nu, \mathbf{X})) = 0$ for $(\nu, \mathbf{X}) \in E$.
4. $\mathbb{E}[N_t] < \infty$, where N_t denotes the number of jumps in $[0, t]$:

$$N_t(\omega) = \sum_k \mathbf{1}_{\{t \geq \mathcal{T}_k\}}(\omega), \quad \omega \in \Omega,$$

where \mathcal{T}_k is the k^{th} event time.

We have already seen above that the conditions 1 and 2 are satisfied (each summand in the definition of a_0^N is integrable - see Section 3.1). Below, we elaborate on the conditions 3 and 4.

Let $(\nu, \mathbf{X}) \in E \cup \Gamma^*$. Then, we have:

$$Q(\{\eta\} \times d\mathbf{X}', (\nu, \mathbf{X})) = \mathbb{P}(\{\eta\} \times d\mathbf{X}' \mid (\nu, \mathbf{X}) \in \Gamma^*) + \mathbb{P}(\{\eta\} \times d\mathbf{X}' \mid (\nu, \mathbf{X}) \notin \Gamma^*).$$

The first and the second terms on the right are, respectively, transition probabilities given that a collision or an FA event occurred. Using Proposition 2.4.3, we have:

$$\begin{aligned} & \mathbb{P}(\{\eta\} \times d\mathbf{X}' \mid (\nu, \mathbf{X}) \notin \Gamma^*) \\ &= \delta_{\mathbf{X}}(d\mathbf{X}') \times \sum_{i=1}^N \left[\sum_{j=1}^M \delta_{\alpha_{\mu}^i(\eta), 0} \frac{a_j^{+,i}(\nu, \mathbf{X})}{a_0^N(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{Y}}^i(\eta)_j, 1} \prod_{j' \neq j}^M \delta_{\alpha_{\mathbf{Y}}^i(\eta)_{j'}, \alpha_{\mathbf{Y}}^i(\nu)_{j'}} \right. \\ & \quad \left. + \delta_{\alpha_{\mu}^i(\eta), 1} \frac{a_j^{-,i}(\nu, \mathbf{X})}{a_0^N(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{Y}}^i(\eta)_j, 0} \prod_{k \neq j}^M \delta_{\alpha_{\mathbf{Y}}^i(\eta)_k, \alpha_{\mathbf{Y}}^i(\nu)_k} \right] \\ & \quad \times \prod_{k \neq i}^N \delta_{\alpha_{\mu}^k(\eta), \alpha_{\mu}^k(\nu)} \prod_{j'=1}^M \delta_{\alpha_{\mathbf{Y}}^k(\eta)_{j'}, \alpha_{\mathbf{Y}}^k(\nu)_{j'}} \prod_{l=1}^N \delta_{\alpha_{\mathbf{C}}^k(\eta)_l, \alpha_{\mathbf{C}}^k(\nu)_l} \\ & \quad \times \prod_{k \neq i}^N \left[\delta_{\alpha_{\mathbf{C}}^i(\eta)_k, 1} \mathbf{1}_{\mathbb{R}^+ \setminus \{0\}} (d_k^i - 1) + \delta_{\alpha_{\mathbf{C}}^i(\eta)_k, 0} \mathbf{1}_{\mathbb{R}^- \cup \{0\}} (d_k^i - 1) \right]. \end{aligned}$$

The first line indicates that components of \mathbf{X} do not jump at an FA event time. The next two lines reflect the fact that an FA event changes the motility state and the state of one adhesion site. The fourth line corresponds to the fact that an FA event in a cell does not affect other cells. The last line indicates that the collision state of a cell is determined according to cell-cell distances at the time of an FA event.

Define the following for (ν, \mathbf{X}) :

$$\begin{aligned} B_{(\nu, \mathbf{X})} &:= \{(m, l) \in \{1, \dots, N\}^2 : d_m^m = d_l^l = 1, \alpha_{\mathbf{C}}^m(\nu)_l = \alpha_{\mathbf{C}}^l(\nu)_m = 0\} \\ B_{(\nu, \mathbf{X})}^c &:= \{1, \dots, N\}^2 \setminus B_{(\nu, \mathbf{X})}, \end{aligned}$$

i.e. tuples of cell indices that have collided, and the remaining pairs, respectively. Let $\mathbf{b} : \Gamma^* \rightarrow \mathbb{R}^{5N+2N^2}$ and $\hat{\Phi} : E \rightarrow [0, 2\pi)^N$ be given by:

$$\begin{aligned} b_i(\nu, \mathbf{X}) &:= (\mathbf{x}^i, \mathbf{x}_n^i, \theta^i, \hat{\Phi}^i(\nu, \mathbf{X}), \mathbf{d}^i) \\ \hat{\Phi}_k^i(\nu, \mathbf{X}) &:= \begin{cases} \Phi_k^i, & \text{if } (i, k) \in B^c(\nu, \mathbf{X}) \\ \hat{\varphi}(\mathbf{x}^i, \mathbf{x}^k), & \text{else,} \end{cases} \end{aligned}$$

where $\hat{\varphi}(\mathbf{x}^i, \mathbf{x}^k)$ is the polar angle at which a contact between cells i and k occurred. Then,

we have:

$$\begin{aligned} \mathbb{P}(\{\eta\} \times d\mathbf{X}' | (\nu, \mathbf{X}) \in \Gamma^*) &= \delta_{\mathbf{b}(\nu, \mathbf{X})}(d\mathbf{X}') \prod_{(m,l) \in B(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{C}}^m(\eta)_l, 1} \delta_{\alpha_{\mu}^m(\eta), 0} \\ &\times \prod_{k=1}^{NM} \delta_{\alpha_{\mathbf{Y}}(\eta)_l, \alpha_{\mathbf{Y}}(\nu)_l} \prod_{(m,l) \in B_{(\nu, \mathbf{X})}^c} \delta_{\alpha_{\mathbf{C}}^m(\eta)_l, \alpha_{\mathbf{C}}^m(\nu)_l} \delta_{\alpha_{\mu}^m(\eta), \alpha_{\mu}^m(\nu)}. \end{aligned}$$

The first line on the right reflects that at the time of collision, the contact angles, collision, and motility states jump to new values. The second line indicates that the FA, collision, and motility states of other cells are unaffected.

Regardless of the event type (FA binding/unbinding vs. collision), the post-jump and pre-jump states are never the same. Thus, $Q(\{(\nu, \mathbf{X})\}, (\nu, \mathbf{X})) = 0$ and the condition 3 above is satisfied.

Suppose the process hits the boundary, i.e. there is a collision between cells i and k for some $i, k \in \{1, \dots, N\}$, and the transition $\Gamma^* \ni (\nu, \mathbf{X}) \rightarrow (\eta, \mathbf{X}')$ occurred. Then, the new collision state is $C_k^i = 1$ and hence $\partial D_{\eta}^{i,k} = \emptyset$ (see Section 6.1). Therefore, $(\eta, \mathbf{X}') \in E$. It follows by Proposition (24.6) in [28] that the above condition 4 is satisfied.

The motion of the extended process with collisions is constructed in the same manner as described in Section 2.4, except that the survival function S and the transition measure Q are now given by (6.2.6) - (6.2.7):

$$S(t) = \begin{cases} \exp\left(-\int_0^t a_0^N(\nu_0, \phi_{\nu_0}(s, \mathbf{X}_0)) ds\right), & t < t^*(\nu_0, \mathbf{X}_0) \\ 0, & t \geq t^*(\nu_0, \mathbf{X}_0). \end{cases} \quad (6.2.6)$$

$$\begin{aligned} &Q(\{\eta\} \times d\mathbf{X}', (\nu, \mathbf{X})) \\ &= \delta_{\mathbf{X}}(d\mathbf{X}') \times \sum_{i=1}^N \left[\sum_{j=1}^M \delta_{\alpha_{\mu}^i(\eta), 0} \frac{a_j^{+,i}(\nu, \mathbf{X})}{a_0^N(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{Y}}^i(\eta)_j, 1} \prod_{j' \neq j}^M \delta_{\alpha_{\mathbf{Y}}^i(\eta)_{j'}, \alpha_{\mathbf{Y}}^i(\nu)_{j'}} \right. \\ &\quad \left. + \delta_{\alpha_{\mu}^i(\eta), 1} \frac{a_j^{-,i}(\nu, \mathbf{X})}{a_0^N(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{Y}}^i(\eta)_j, 0} \prod_{k \neq j}^M \delta_{\alpha_{\mathbf{Y}}^i(\eta)_k, \alpha_{\mathbf{Y}}^i(\nu)_k} \right] \\ &\quad \times \prod_{k \neq i}^N \delta_{\alpha_{\mu}^k(\eta), \alpha_{\mu}^k(\nu)} \prod_{j'=1}^M \delta_{\alpha_{\mathbf{Y}}^k(\eta)_{j'}, \alpha_{\mathbf{Y}}^k(\nu)_{j'}} \prod_{l=1}^N \delta_{\alpha_{\mathbf{C}}^k(\eta)_l, \alpha_{\mathbf{C}}^k(\nu)_l} \\ &\quad \times \prod_{k \neq i}^N \left[\delta_{\alpha_{\mathbf{C}}^i(\eta)_k, 1} \mathbf{1}_{\mathbb{R}^+ \setminus \{0\}}(d_k^i - 1) + \delta_{\alpha_{\mathbf{C}}^i(\eta)_k, 0} \mathbf{1}_{\mathbb{R}^- \cup \{0\}}(d_k^i - 1) \right]. \\ &+ \delta_{\mathbf{b}(\nu, \mathbf{X})}(d\mathbf{X}') \prod_{(m,l) \in B(\nu, \mathbf{X})} \delta_{\alpha_{\mathbf{C}}^m(\eta)_l, 1} \delta_{\alpha_{\mu}^m(\eta), 0} \\ &\quad \times \prod_{k=1}^{NM} \delta_{\alpha_{\mathbf{Y}}(\eta)_l, \alpha_{\mathbf{Y}}(\nu)_l} \prod_{(m,l) \in B_{(\nu, \mathbf{X})}^c} \delta_{\alpha_{\mathbf{C}}^m(\eta)_l, \alpha_{\mathbf{C}}^m(\nu)_l} \delta_{\alpha_{\mu}^m(\eta), \alpha_{\mu}^m(\nu)}. \end{aligned} \quad (6.2.7)$$

Homotypic and heterotypic CIL

In order to take into account mixed populations with different CIL response, we only need to slightly modify the definition of D_k^i in (6.2.1). Let $I_1 \subset \{1, \dots, N\}$, $I_2 \subset \{1, \dots, N\}$, be index sets of cells with and without CIL, respectively, such that $I_1 \cap I_2 = \emptyset$. Then:

$$D_\nu^{i,k} := \begin{cases} (-\infty, 1], & \text{if } C_k^i = 0, i, k \in I_1 \text{ or } i, k \in I_2 \\ (-\infty, \infty), & \text{else,} \end{cases}$$

$$D_\nu^i := \prod_{k=1}^N D_\nu^{i,k}, \quad D_\nu := \prod_{i=1}^N D_\nu^i.$$

Thus, only members of the same group undergo CIL. Here, in the absence of heterotypic CIL we effectively rule out collisions between members of different groups.

6.3 Simulations

6.3.1 Methods

To simulate the constructed process we employ Algorithm 3 presented below.

Algorithm 3 Simulation of the PDMP

1. Set $(\nu_0, \mathbf{X}_0) \in E$ and $t = \mathcal{T}_0 = 0$, $k = 0$.
 2. Generate interarrival time $\hat{\Delta}_k$ using Algorithm 2 applied to the ODE system (6.2.5) and the survival function (6.2.6).
 3. Find $\mathbf{X}_{\mathcal{T}_k + \hat{\Delta}_k} = \phi_{\nu_{\mathcal{T}_k}}(\hat{\Delta}_k, \mathbf{X}_{\mathcal{T}_k})$ and $\hat{B}_{(\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k + \hat{\Delta}_k})}$. Set $\Delta_k = \hat{\Delta}_k$.
 4. If $\hat{B}_{(\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k + \hat{\Delta}_k})} \neq \emptyset$ (Collision)

$$\Delta_k = \min \left\{ s > 0 : d_l^m(\mathcal{T}_k + s) = d_m^l(\mathcal{T}_k + s) = 1, (m, l) \in \hat{B}_{(\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k + \hat{\Delta}_k})} \right\}$$
 5. Set $\mathcal{T}_{k+1} := \mathcal{T}_k + \Delta_k$
 $(\nu_{\mathcal{T}_{k+1}}, \mathbf{X}_{\mathcal{T}_{k+1}}) \sim Q(\cdot, (\nu_{\mathcal{T}_k}, \mathbf{X}_{\mathcal{T}_k + \Delta_k}))$
 $k := k + 1$
-

Here, we use our previously developed method in Section 4.1 to simulate a general piecewise deterministic process. However, we now need to take into account collisions as well. To do so, we define

$$\hat{B}_{(\nu, \mathbf{X})} := \{(m, l) \in \{1, \dots, N\}^2 : d_l^m = d_m^l \geq 1, \boldsymbol{\alpha}_{\mathbf{C}}^m(\nu)_l = \boldsymbol{\alpha}_{\mathbf{C}}^l(\nu)_m = 0\}. \quad (6.3.1)$$

Note that if $(m, l) \notin \hat{B}_{(\nu_t, \mathbf{x}_t)}$ and $(m, l) \in \hat{B}_{(\nu_t, \mathbf{x}_{t+s})}$, then this implies that a collision between cells m and l occurred in the time interval $[t, t + s]$.

After initialization in Step 1 of the algorithm below, we find the interarrival time $\hat{\Delta}_k$ of the next FA event in Step 2 using Algorithm 2. Then, in Step 3 we evolve the ODE system (6.2.5) and identify the cells, which collided in this time period. For each colliding pair, we find their collision time s , and their minimum in Step 4. The collision time $s \in (0, \hat{\Delta}_k]$ for $(m, l) \in \hat{B}_{(\nu_{\mathcal{T}_k}, \mathbf{x}_{\mathcal{T}_k + \hat{\Delta}_k})}$ is the root of

$$f(s) = f(\mathbf{X}_{\mathcal{T}_k + s}) = d_l^m(\mathcal{T}_k + s) - 1 = 0. \quad (6.3.2)$$

Note that after Step 3, the solution \mathbf{X}_t of the ODE system (6.2.5) is available at the time points $t = \mathcal{T}_k + s_i$, where $i = 0, \dots, n$ and $s_n = \hat{\Delta}_k$. Thus,

$$f(s) = f(\mathbf{X}_{\mathcal{T}_k + s}) = f(\phi_{\nu_{\mathcal{T}_k}}(s - s_i, \mathbf{X}_{\mathcal{T}_k + s_i})).$$

Therefore, evaluation of (6.3.2) needed for a root finding method¹ amounts to advancing the ODE system for a single time step of size $s - s_i$. This way, the amount of extra computations needed to find the collision time is minimized, which yields increasing computational savings as the number of cells N increases. Finally, in Step 5 we set the time of the next event \mathcal{T}_{k+1} and update the system according to the event occurred.

This method can be used to efficiently simulate an arbitrary PDMP, where solving an ODE system is expensive and the boundary hitting time is finite. In our case, one can improve the numerical performance by observing that solving the N -cell ODE system (6.2.5) can be parallelized by solving N ODE subsystems governing the evolution of $(\mathbf{x}^i, \mathbf{x}_n^i, \theta^i)$ independently, and then computing d_k^i given by (6.2.3). Thus, the computation time needed to generate the interarrival time $\hat{\Delta}_k$ can be reduced drastically. Each simulation scenario below required roughly 60 hours of the computation time with parallelization on four cores.

Finally, to sample from the transition measure in Step 5, one can use Vose's method as in the case for single cell migration (see Section 4.1.2).

6.3.2 Results

We now simulate our model in an unconfined 2D setting and investigate the effect of CIL on chemotaxing and non-chemotaxing cells. As was shown in Section 5.2, taking $\delta_- = 0$ may lead to overlapping cells. Since in a general 2D setting a cell might have a contact with multiple cells at the same time (see Figure 6.1), it is possible (for arbitrary values of δ_{myo} , δ_+ , and δ_-) that multiple cells overlap each other. Note that cells undergoing CIL do not crawl on top of each other. Thus, for cells undergoing CIL we take $\delta_{myo} = 0.4$, $\delta_+ = 0.2$, $\delta_- = 1$, and for cells not exhibiting it we take $\delta_{myo} = \delta_+ = \delta_- = 0$.

We also explore the interplay between CIL and chemotaxis in a heterogeneous population of cells. Namely, we investigate the effect of CIL on a mix of cells responsive and non-responsive to an external cue. For chemotaxing cells we take $\delta_E = 0.05$.

¹Any root finding method with bracketing could be employed.

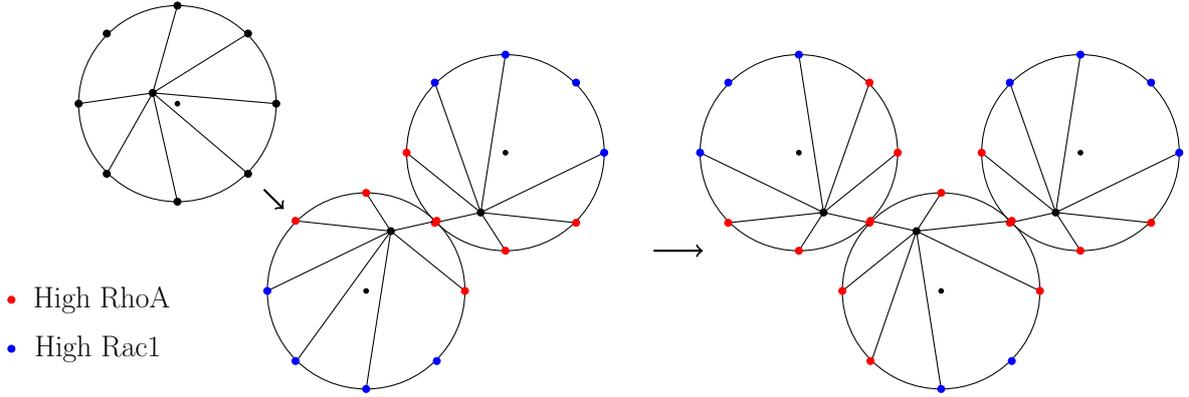


Figure 6.1: Schematic representation of CIL for non-binary collisions.

In the following, we simulate 36 cells and evolve them for 20 hours, such that initially the cells are positioned as in Figure 6.2, and the distance between the centroids of neighboring cells is $2.4R_{cell}$. All other parameter values are the same as in Section 5.2.

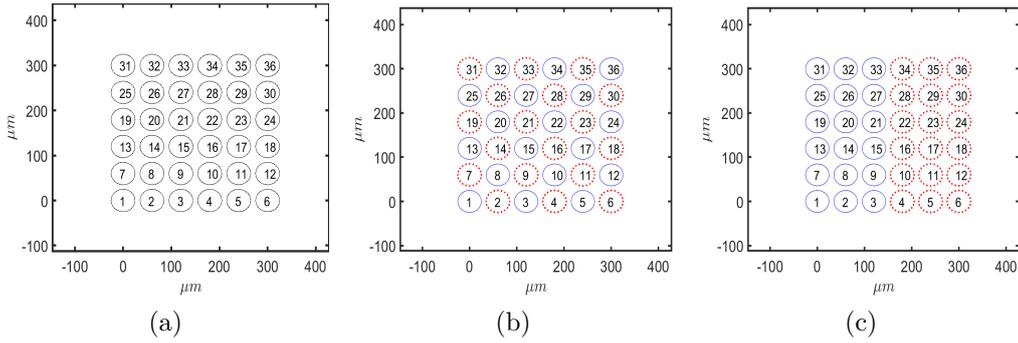


Figure 6.2: Initial configuration of cells. (a) Homogeneous population. (b,c) Heterogeneous populations of chemotaxing (blue, solid) and non-chemotaxing (red, dotted) cells.

Homogeneous population

Simulation results for a homogeneous population of chemotaxing and non-chemotaxing cells are shown in Figure 6.3. We see that the biased migration of chemotaxing cells occurs in a cluster-like manner. In contrast, we see that the non-chemotaxing cells disperse randomly, such that the center of mass deviates very little as compared to cell dimensions ($R_{cell} = 25\mu m$). Note that the motion of randomly migrating cells exhibits a superdiffusive character (Figure 6.3h), as indicated by fitting the mean-squared displacement to the curve $\widehat{msd}(t) = \beta_0 t^{\bar{\beta}}$ (see Appendix B for details). In Section 4.3.1, it was shown that non-interacting cells¹ exhibit normal diffusive behavior ($\bar{\beta} \sim 1$) in the absence of any source

¹But otherwise identical, as the parameter values are the same.

of asymmetry affecting FA dynamics. Here, since the exponent $\bar{\beta}$ corresponding to non-chemotaxing cells is larger than one, we see that cell-cell collisions also lead to anomalous diffusion as $\bar{\beta} > 1$. Comparing chemotaxing cells, we also see that $\bar{\beta}$ increases if cells collide with one another (in Section 4.3.2 for the same value of δ_E). Thus, we see that the average displacement increases due to CIL, despite the fact that motion ceases upon contact.

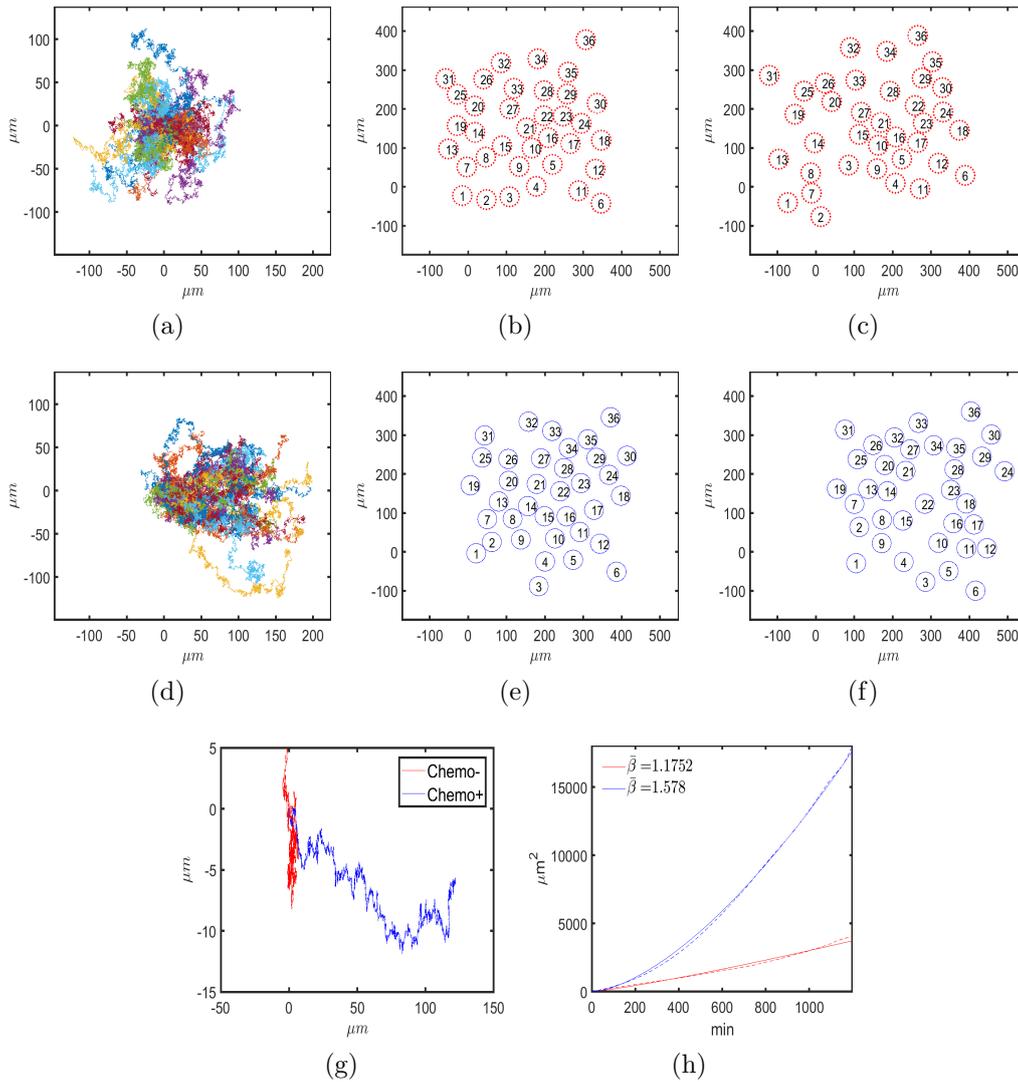


Figure 6.3: Simulation results for a homogeneous population of non-chemotaxing (top row) and chemotaxing cells (middle row). (a,d) Centered trajectories. (b,e) and (c,f) Cell positions at $t = 600 \text{ min}$, 1200 min , respectively. (g) Centered trajectories of the cluster centers of mass. (h) Mean-squared displacement (dashed) of chemotaxing (blue) and non-chemotaxing (red) cells, and the fit $\widehat{msd}(t)$.

It has been hypothesized that superdiffusive motion is optimal for searching a target

source, that itself diffuses [11], [35]. Thus, cancer cells that acquire ability to undergo homotypic CIL can find a diffusing source (e.g. VEGF) more efficiently and hence facilitate tumor progression. Interestingly, it has also been hypothesized that homotypic CIL facilitates dispersion of cancer cells [76], [112].

Inhomogeneous population

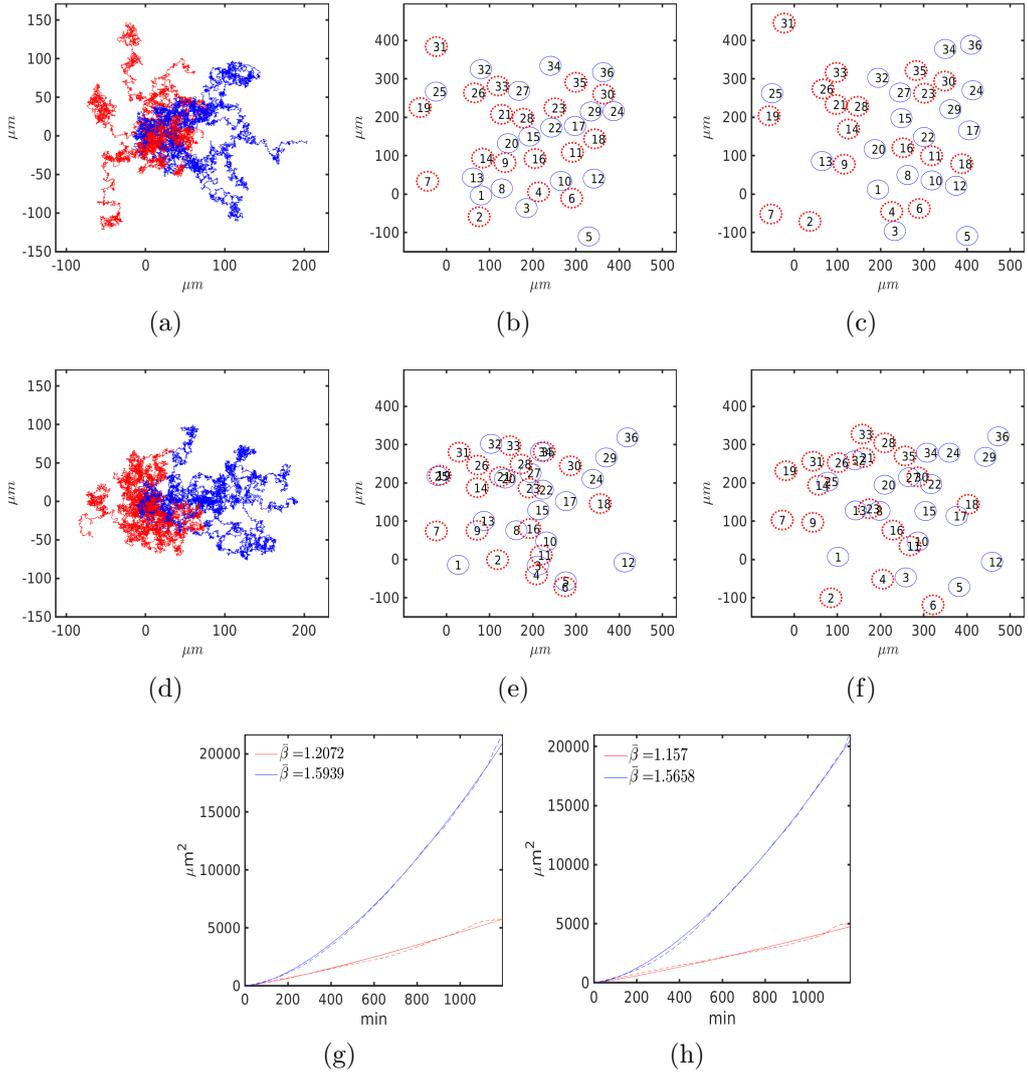


Figure 6.4: Simulation results for the mixed population with (top row) and without (middle row) heterogeneous CIL. Initially, cells are positioned as in Figure 6.2b. (a,d) Centered trajectories of 9 chemotaxing (blue) and non-chemotaxing (red) cells. (b,e) and (c,f) Positions of chemotaxing (blue, solid) and non-chemotaxing (red, dotted) cells at $t = 600\text{min}$ in (b,e) and at $t = 1200\text{min}$ in (c,f). (g,h) Mean squared displacements of chemotaxing (blue) and non-chemotaxing (red) cells with (g) and without heterotypic CIL (h).

We now explore the effects of heterotypic CIL between populations of chemotaxing and non-chemotaxing cells (Figure 6.2b,c). Here, cells always exhibit CIL when they collide with the members of the same group.

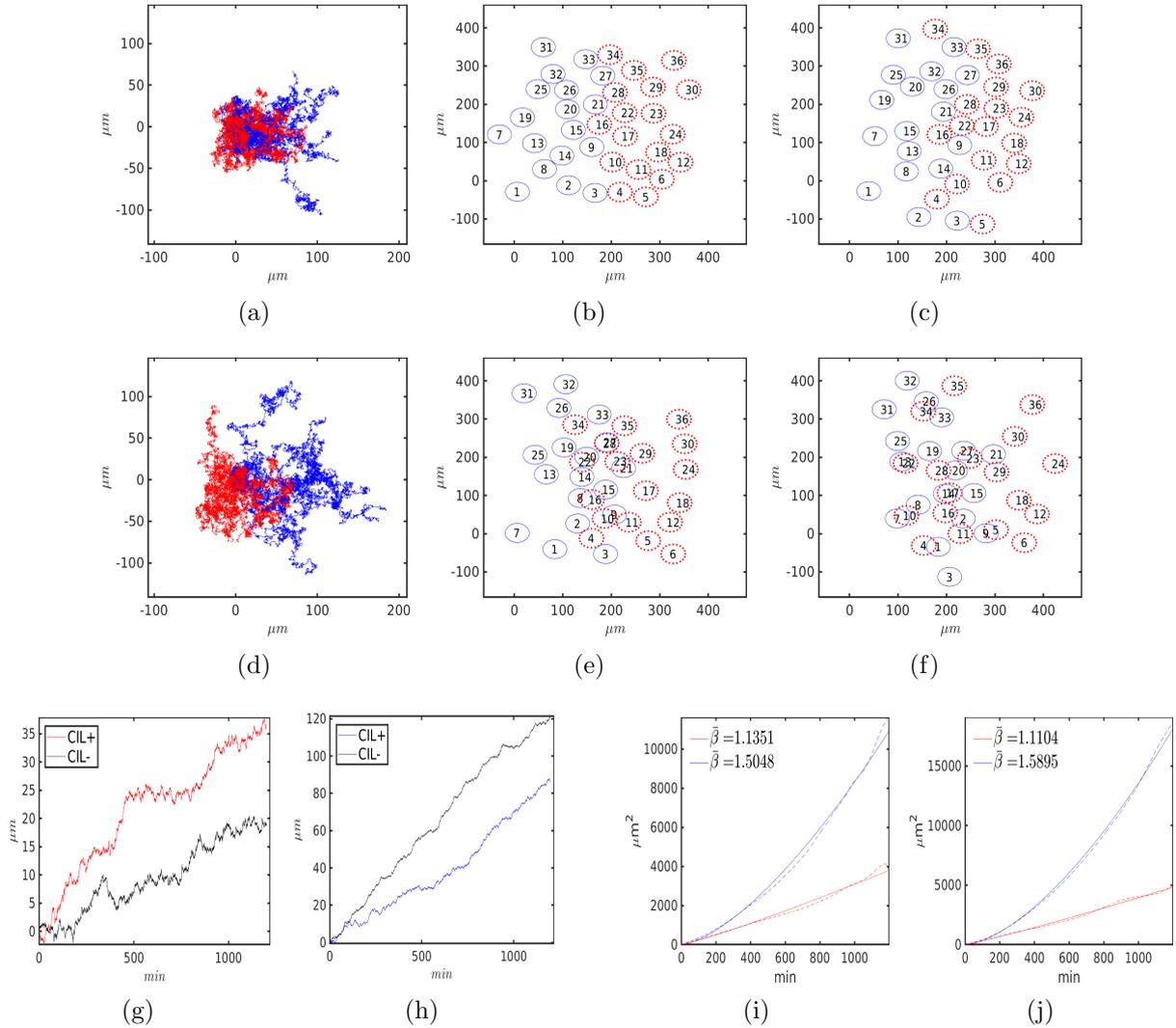


Figure 6.5: Simulation results for the separated population with (top row) and without (middle row) heterogeneous CIL. Initially, cells are position as in Figure 6.2c. (a,d) Centered trajectories of 9 chemotaxing (blue) and non-chemotaxing (red) cells. (b,e) and (c,f) Positions of chemotaxing (blue, solid) and non-chemotaxing (red, dotted) cells at $t = 600min$ in (b,e) and at $t = 1200min$ in (c,f). (g,h) x components of non-chemotaxing (g) and chemotaxing (h) cells' centers of mass with (colored) and without (black) heterotypic CIL. (i,j) Mean squared displacements of chemotaxing (blue) and non-chemotaxing (red) cells with (i) and without (j) heterotypic CIL.

When evenly mixed (Figure 6.2b), we see that heterotypic CIL does not have a sig-

nificant impact on chemotaxing or non-chemotaxing cells (Figure 6.4), as the behavior of each subgroup resembles the case with the corresponding homogeneous populations. This suggests that in a disordered population of cells, homotypic, but not heterotypic CIL facilitates directed migration of cells (as $\bar{\beta} \approx 1.34$ in freely chemotaxing cells, see Section 4.3.2). Nevertheless, notice that in this unclustered configuration, the chemotaxing cells are able to push their way out, leading to dispersion of the surrounding cells akin to billiard balls (Figure 6.4a,d): centered trajectories of the non-responsive cells show higher dispersion due to the repulsive interaction with the chemotaxing cells, who must push out the non-responsive cells to achieve the observed directed migration when such interaction is present. Clustering cells according to their responsiveness to an external cue, however, lead to a qualitatively different outcome. If responsive and non-responsive cells are separated as in Figure 6.2c, we see a cluster-like interaction when heterotypic CIL is present (Figure 6.5): the dividing line between the groups remains discernible for a long time (Figure 6.5b,c), which is not the case when the heterotypic CIL is absent (Figure 6.5e,f). This indicates that the initial clustering (Figure 6.2c) is conserved due to heterotypic interaction. Unlike the case of evenly mixed cells, we see that the dispersion of the non-chemotaxing cells is not as prominent (Figure 6.4a vs. Figure 6.5a), and the chemotaxing cells do not push out the non-responsive ones. In fact, we observe that the latter are being displaced in a sheet-like manner by the responsive cells. A similar behavior was observed in [118], although in that study the non-chemotaxing cells were themselves the source of a chemoattractant. Such displacement induces the non-chemotaxing cells to align their motion with the direction of an external cue (Figure 6.5g), although the effect of heterotypic CIL is slight. On the other hand, we see that directed migration of the chemotaxing cells is impeded (Figure 6.5h), which is also reflected in the reduced average displacement (Figure 6.5i). Altogether, these results suggest that the role of heterotypic CIL varies with the distribution of the cell population: it may either facilitate dispersion (Figure 6.4) or induce directed motion in otherwise randomly migrating cells (Figure 6.5). Its loss, however, is beneficial for tactic migration irrespective of spatial configuration.

Chapter 7

Cell Populations

In this chapter we deduce the population scale description from the single cell migration described in Chapter 2. Our deductions largely follow the procedure used to derive the Fokker-Planck equation describing the evolution of the probability density of particle's position and velocity, which follow the Langevin equation. However, unlike the case of a Brownian particle following a diffusion process, our cell motility model is a non-diffusion process with piecewise deterministic paths. Nevertheless, we can also describe the cell's internal state and its position in terms of a probability density function and its governing equation. Assuming that cells are not interacting with one-another, we thus obtain a population scale description of cell migration. We also demonstrate how the resulting equation can characterize migration on a spatial scale, much larger than a cell.

The content of this chapter directly follows from Chapter 2, and so do all the objects described there as well as the notation.

7.1 Probability law of a PDMP

In this section we formally show how a connection between an arbitrary PDMP and a deterministic description of the process can be established. Except the newly introduced ones, all notation used here follows from Section 2.4. We also follow Sections 14, 24-26 in [28].

We slightly abuse the notation by adopting the following convention. Let X_t denote the state of the PDMP with the state space (E, \mathcal{E}) . That is, $X_t = (\nu_t, \mathbf{X}_t) \in E$.

Let $p : \mathbb{R}_+ \times E \times \mathbb{R}_+ \times \mathcal{E}$ be the *transition function*, defined by

$$p(s, x, t, C) := \mathbb{P}(x_t \in C \mid x_s = x),$$

and satisfying the following properties:

1. For fixed s, t, x , the mapping $C \mapsto p(s, x, t, C)$ is a probability measure on (E, \mathcal{E}) .
2. For fixed s, t, C , the mapping $x \mapsto p(s, x, t, C)$ is \mathcal{E} -measurable.
3. For all s, x, C , $p(s, x, s, C) = \mathbf{1}_C(x)$.

4. For $s \leq t$, $x \in E$ and $C \in \mathcal{E}$ we have

$$\mathbb{E}[\mathbf{1}_B(X_t) \mid \mathcal{F}_s] = p(s, X_s, t, B) \quad \text{a.s.}$$

Note that since the process is Markovian, we have:

$$\mathbb{E}[f(X_t) \mid \mathcal{F}_s] = \mathbb{E}[f(X_t) \mid X_s], \quad (7.1.1)$$

for all $f \in B(E)$, where $B(E)$ is the set of bounded measurable functions $f : E \rightarrow \mathbb{R}$. For example, property 4 above becomes:

$$\mathbb{E}[\mathbf{1}_C(X_t) \mid \mathcal{F}_s] = \mathbb{E}[\mathbf{1}_C(X_t) \mid X_s] = p(s, X_s, t, C) \quad \text{a.s.}, \quad (7.1.2)$$

which is just another way of saying that the future of the process depends on its current value only, and is independent of its past history. Due to property 1 above, we have:

$$\mathbb{E}[f(X_t) \mid X_s] = \int_E f(y)p(s, X_s, t, dy) \quad \text{a.s.} \quad (7.1.3)$$

Let $r \leq s \leq t$. Since $\mathcal{F}_r \subset \mathcal{F}_s$, we have the ‘‘tower property’’ of conditional expectation:

$$\mathbb{E}[f(X_t) \mid \mathcal{F}_r] = \mathbb{E}[\mathbb{E}[f(X_t) \mid \mathcal{F}_s] \mid \mathcal{F}_r] \quad \text{a.s.} \quad (7.1.4)$$

Thus we have:

$$\begin{aligned} p(r, x, t, C) &= \mathbb{E}[\mathbf{1}_C(X_t) \mid X_r = x] = \mathbb{E}[\mathbf{1}_C(X_t) \mid \mathcal{F}_r] \\ &= \mathbb{E}[\mathbb{E}[\mathbf{1}_C(X_t) \mid \mathcal{F}_s] \mid \mathcal{F}_r] \\ &= \int_E \int_E \mathbf{1}_C(y)p(s, z, t, dy)p(r, x, s, dz) \\ &= \int_E p(s, z, t, C)p(r, x, s, dz), \end{aligned}$$

where the first line is due to (7.1.2), the second is due to (7.1.4), and the third line is due to (7.1.2) and (7.1.3). The relation above is called Chapman-Kolmogorov equation.

Suppose that the process is time-homogeneous, i.e.:

$$p(s, x, t, C) = p(0, x, t - s, C),$$

for $s \leq t$. That is, only the time difference is relevant. No generality is lost, since a process with time-inhomogeneous transition function can be transformed into a homogeneous one by restarting the process appropriately. In particular, one could consider the case with time-dependent external environment (e.g. a diffusing signal). Abusing the notation, the transition function p becomes:

$$p(t, x, C) = p(0, x, t, C).$$

Thus, the Chapman-Kolmogorov equation becomes:

$$p(t + s, x, C) = \int_E p(t, z, C)p(s, x, dz). \quad (7.1.5)$$

For $t \geq 0$, let an operator $P_t : B(E) \rightarrow B(E)$ be defined as

$$(P_t f)(x) = \mathbb{E}[f(X_t) \mid X_0 = x].$$

Clearly, $(P_0 f)(x) = f(x)$, i.e. $P_0 = I$, where I is an identity operator. We also have the following relation:

$$\begin{aligned} (P_{t+s} f)(x) &= \mathbb{E}[f(X_{t+s}) \mid X_0 = x] = \int_E f(y)p(s+t, x, dy) \\ &= \int_E \int_E f(y)p(s, z, dy)p(t, x, dz) = \int_E (P_s f)(z)p(t, x, dz) \\ &= (P_t \circ P_s f)(x), \end{aligned}$$

where the second and third equalities are due to (7.1.3) and (7.1.5), respectively. Therefore, the family P_t defines a semigroup, provided $B(E)$ is a Banach space equipped with a supremum norm.

Associated with the semigroup P_t is infinitesimal generator $\mathcal{L} : \mathcal{D}(\mathcal{L}) \rightarrow B(E)$, defined by

$$\mathcal{L}f = \lim_{t \downarrow 0} \frac{P_t f - f}{t},$$

where $\mathcal{D}(\mathcal{L}) \subset B(E)$ is the set of functions such that the above limit exists. Let $\hat{x} \in E$. Then, using the definition of P_t and the Chapman-Kolmogorov equation (7.1.5), we have the following:

$$\begin{aligned} \int_E ((P_\tau f)(x) - f(x))p(t, \hat{x}, dx) &= \int_E \int_E f(y)p(\tau, x, dy)p(t, \hat{x}, dx) - \int_E f(x)p(t, \hat{x}, dx) \\ &= \int_E f(x) [p(t + \tau, \hat{x}, dx) - p(t, \hat{x}, dx)]. \end{aligned}$$

Dividing both sides by τ , and taking the limit, we obtain:

$$\int_E f(x) \frac{\partial}{\partial t} p(t, \hat{x}, dx) = \int_E \mathcal{L}f(x)p(t, \hat{x}, dx), \quad (7.1.6)$$

for $f \in \mathcal{D}(\mathcal{L})$. Considering, formally, the adjoint \mathcal{L}^* of \mathcal{L} acting on probability measures:

$$\begin{aligned} \int_E f(x) \frac{\partial}{\partial t} p(t, \hat{x}, dx) &= \int_E f(x) \mathcal{L}^* p(t, \hat{x}, dx) \\ \Rightarrow \int_E f(x) \left[\frac{\partial}{\partial t} p(t, \hat{x}, dx) - \mathcal{L}^* p(t, \hat{x}, dx) \right] &= 0, \end{aligned}$$

or equivalently

$$\frac{\partial}{\partial t} p = \mathcal{L}^* p, \quad p(0, \hat{x}, \cdot) = \delta_{\hat{x}}(\cdot). \quad (7.1.7)$$

We are now interested in the following quantity:

$$\mu(t, \cdot) := \mathbb{P}(X_t \in \cdot) = \int_E p(t, \hat{x}, \cdot) \mu(0, d\hat{x}),$$

i.e. the law of the process X_t , given the initial distribution $\mu(0, \cdot)$. Instead of relying on (7.1.7), we can directly find how $\mu(t, \cdot)$ evolves with time. Multiplying (7.1.6) with $\mu_0(d\hat{x})$ and integrating, we obtain:

$$\begin{aligned} \int_E \int_E f(x) \frac{\partial}{\partial t} p(t, \hat{x}, dx) \mu(0, d\hat{x}) &= \int_E \int_E \mathcal{L} f(x) p(t, \hat{x}, dx) \mu(0, d\hat{x}) \\ \iff \int_E f(x) \frac{\partial}{\partial t} \mu(t, dx) &= \int_E \mathcal{L} f(x) \mu(t, dx), \end{aligned}$$

or equivalently (as above)

$$\frac{\partial}{\partial t} \mu = \mathcal{L}^* \mu. \quad (7.1.8)$$

Now that we have an abstract relation describing how the law of the Markov process X_t evolves, in the following section we will see how it applies to the cell motility process.

7.2 The law of cell motility

Recalling Section 2.4.2 and the definitions *I, II, III* of the objects necessary to specify a PDMP, we see that the conditions (24.8), Section 24 in [28] are satisfied. Namely, for $x = (\nu, \mathbf{X}) \in E$, we have the following:

- I The vector fields $\mathbf{X} \mapsto \mathbf{H}_\nu(\mathbf{X})$, given by (2.4.9), are Lipschitz continuous and there exists a unique solution to (2.4.1) due to Proposition 2.4.2.
- II $a_0 : E \rightarrow \mathbb{R}_+$ is a measurable function such that $t \mapsto a(\nu, \phi_\nu(t, \mathbf{X}))$ is integrable for all $t \geq 0$, where a_0 is given by (2.4.10) and ϕ_ν is the flow corresponding to (2.4.1).
- III The transition measure $Q : \mathcal{E} \times E \rightarrow [0, 1]$, given by (2.4.11), is such that for a fixed $C \in \mathcal{E}$, $(\nu, \boldsymbol{\xi}) \mapsto Q(C; (\nu, \boldsymbol{\xi}))$ is measurable for $(\nu, \boldsymbol{\xi}) \in E$, and $Q(\cdot; (\nu, \boldsymbol{\xi}))$ is a probability measure for all $(\nu, \boldsymbol{\xi})$ on (E, \mathcal{E}) . Moreover, it holds that $Q(\{x\}; x) = 0$.
- IV $\mathbb{E}[N_t] < \infty$ for every starting point $x \in E$ of the process, where

$$N_t(\omega) = \sum_k \mathbf{1}_{\{t \geq \tau_k\}}(\omega), \quad \omega \in \Omega,$$

is the number of occurred FA events at time t .

The last condition above is satisfied by Proposition 24.6 in [28], since for the case of single cell migration, we do not have an “active” boundary. That is, events do not occur when the process attains a certain value, as is the case with cell-cell collisions (see Chapter 6). For details on a general PDMP, see Chapter 2 in [28].

Then, by Theorem 26.14 in [28] the generator \mathcal{L} is given by:

$$\mathcal{L}g(x) = \mathbf{H}_\nu(\mathbf{X}) \cdot \nabla_{\mathbf{X}}g(\nu, \mathbf{X}) + a_0(x) \int_E (g(y) - g(x)) Q(dy; x), \quad (7.2.1)$$

for $x = (\nu, \mathbf{X}) \in E$, $g \in \mathcal{D}(\mathcal{L})$. Informally, the first term on the right hand side of the equation above corresponds to the “generator” of the deterministic process in the absence of FA jumps. Correspondingly, the second term is the generator of the FA jump process in the absence of any deterministic motion between the jumps.

We now turn our attention to the adjoint \mathcal{L} acting on the measure $\mu(t, \cdot)$, which are related by (7.1.8). For $g \in \mathcal{D}(\mathcal{L})$ we have:

$$\begin{aligned} \int_E \mathcal{L}g(x) \mu(t, dx) &= \int_E \mathbf{H}(x) \cdot \nabla g(x) \mu(t, dx) \\ &\quad + \int_E \int_E a_0(x) (g(y) - g(x)) Q(dy; x) \mu(t, dx), \end{aligned} \quad (7.2.2)$$

where, for clarity, we denoted $\mathbf{H}(x) = \mathbf{H}_\nu(\mathbf{X})$, $\nabla g(x) = \nabla_{\mathbf{X}}g(\nu, \mathbf{X})$ for $x = (\nu, \mathbf{X}) \in E$. Suppose that the measure $\mu(t, \cdot)$ has density $f(t) : E \rightarrow [0, \infty)$ with respect to measure λ , where $f(t)$ is measurable $\forall t > 0$, and λ is a product of Lebesgue and counting measures. That is, for $x \in E$:

$$\mu(t, dx) = f(t, x) \lambda(dx),$$

where $\lambda(dx) = d\mathbf{X}dI(\nu)$, and I is a counting measure. Consider the first term on the right hand side in (7.2.2):

$$\begin{aligned} \int_E \mathbf{H}(x) \cdot \nabla g(x) f(t, x) \lambda(dx) &= \int_E \mathbf{H}_\nu(\mathbf{X}) \cdot \nabla_{\mathbf{X}}g(\nu, \mathbf{X}) f(t, \nu, \mathbf{X}) d\mathbf{X}dI(\nu) \\ &= - \int_E g(\nu, \mathbf{X}) \nabla_{\mathbf{X}} \cdot (\mathbf{H}_\nu(\mathbf{X}) f(t, \nu, \mathbf{X})) d\mathbf{X}dI(\nu), \end{aligned} \quad (7.2.3)$$

where we used integration by parts and assumed that the density $f(t, \cdot)$ decays sufficiently fast at the boundary.

Recalling the form of the transition measure Q in (2.4.11), let $q : A \times E \rightarrow [0, 1]$ be such that

$$Q(\{\eta\} \times d\mathbf{X}'; (\nu, \mathbf{X})) = \delta_{\mathbf{X}}(d\mathbf{X}') q(\eta; \nu, \mathbf{X}),$$

i.e. $q(\eta; \nu, \mathbf{X})$ is the probability that the transition $\nu \rightarrow \eta$ occurs, given the state of the system $(\nu, \mathbf{X}) \in E$ before the jump. Due to (2.4.11), q is given by:

$$q(\eta; \nu, \mathbf{X}) = \sum_{j=1}^M \delta_{\alpha_\mu(\eta), 0} \frac{a_j^+(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \mathbf{X})}{a_0(\boldsymbol{\alpha}_{\mathbf{Y}}(\nu), \mathbf{X})} \delta_{\alpha_{\mathbf{Y}}(\eta), j, 1} \prod_{i \neq j}^M \delta_{\alpha_{\mathbf{Y}}(\eta), i, \alpha_{\mathbf{Y}}(\nu)_i}$$

$$+ \delta_{\alpha_\mu(\eta),1} \frac{a_j^-(\boldsymbol{\alpha}_Y(\nu), \mathbf{X})}{a_0(\boldsymbol{\alpha}_Y(\nu), \mathbf{X})} \delta_{\boldsymbol{\alpha}_Y(\eta)_j,0} \prod_{i \neq j}^M \delta_{\boldsymbol{\alpha}_Y(\eta)_i, \boldsymbol{\alpha}_Y(\nu)_i}. \quad (7.2.4)$$

We then have the following:

$$\begin{aligned} \int_E \int_E a_0(x)g(y)Q(dy; x)\mu(t, dx) &= \int_E \int_E a_0(y)g(x)Q(dx; y)\mu(t, dy) \\ &= \int_E \int_E g(\nu, \mathbf{X})a_0(\eta, \mathbf{X}')\delta_{\mathbf{X}'}(d\mathbf{X})q(\nu; \eta, \mathbf{X}')f(t, \eta, \mathbf{X}')d\mathbf{X}'dI(\eta)dI(\nu) \\ &= \int_E g(\nu, \mathbf{X}) \sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f(t, \eta, \mathbf{X})d\mathbf{X}dI(\nu), \end{aligned}$$

where $x = (\nu, \mathbf{X})$ and $y = (\eta, \mathbf{X}')$. Similarly, the following holds:

$$\begin{aligned} \int_E \int_E a_0(x)g(x)Q(dy; x)\mu(t, dx) &= \int_E \int_E g(\nu, \mathbf{X})a_0(\nu, \mathbf{X})\delta_{\mathbf{X}}(d\mathbf{X}')q(\eta; \nu, \mathbf{X})f(t, \nu, \mathbf{X})d\mathbf{X}'dI(\eta)dI(\nu) \\ &= \int_E g(\nu, \mathbf{X}) \sum_{\eta} a_0(\nu, \mathbf{X})q(\eta; \nu, \mathbf{X})f(t, \nu, \mathbf{X})d\mathbf{X}dI(\nu). \end{aligned}$$

Together with (7.2.3), (7.2.2) becomes:

$$\begin{aligned} \int_E \mathcal{L}g(x)\mu(t, dx) &= \int_E \mathcal{L}g(x)f(t, x)\lambda(dx) \\ &= \int_E g(\nu, \mathbf{X}) [-\nabla_{\mathbf{X}} \cdot (\mathbf{H}_\nu(\mathbf{X})f(t, \nu, \mathbf{X}))] \\ &\quad + \left[\sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f(t, \eta, \mathbf{X}) - a_0(\nu, \mathbf{X})q(\eta; \nu, \mathbf{X})f(t, \nu, \mathbf{X}) \right] d\mathbf{X}dI(\nu) \\ &= \int_E g(x)\mathcal{L}^*f(t, x)\lambda(dx) = \int_E g(x)\mathcal{L}^*\mu(t, dx). \end{aligned}$$

Finally, due to (7.1.8), we have:

$$\begin{aligned} \frac{\partial}{\partial t}f(t, \nu, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot (\mathbf{H}_\nu(\mathbf{X})f(t, \nu, \mathbf{X})) &= -a_0(\nu, \mathbf{X})f(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f(t, \eta, \mathbf{X}) \end{aligned} \quad (7.2.5)$$

Remark. For a slightly more general case, suppose the transition measure $Q(dx; y)$ has density $q(x; y)$ with respect to the measure $\lambda(dx)$. Then, we would obtain:

$$\frac{\partial}{\partial t} f(t, x) + \nabla \cdot (\mathbf{H}(x)f(t, x)) = -a_0(x)f(t, x) + \int_E a_0(y)q(x; y)f(t, y)dy$$

Although (7.2.5) is intuitive and could be obtained heuristically, here we showed that it is deducible from the piecewise deterministic model of cell motility we constructed in Chapter 2.

Assuming we have a population of non-interacting cells, then (7.2.5) also describes the evolution of cell population density. That is, $\int_C f(t, \nu, \mathbf{x}, \mathbf{x}_n, \theta)d\mathbf{x}$ is the number density of cells in region C of space having state $\nu, \mathbf{x}_n, \theta$ at time t (recalling that $\mathbf{X} = (\mathbf{x}, \mathbf{x}_n, \theta)$, see Subsection 2.4.2).

Remark. In this interpretation, (7.2.5) is often said to describe the population on a mesoscopic scale, while our model in Chapter 2 describes it on microscopic scale (for more examples, see [33], [56] and references therein). However, interpretations may vary in literature. For example, in [13] the mesoscopic scale refers to a collection of tens of cells.

We are primarily interested in the total number of cells in a particular region of space, regardless of their internal state. Computing such quantity directly from (7.2.5) is unfeasible due to the dimensionality of the problem (recall that $\nu \in \{1, \dots, 2^{M+1}\}$, where M is the number of FAs - see Section 2.4.2). In the following section we introduce a few simplifying assumptions, yielding a more tractable problem.

7.3 Order reduction

We assume the following:

- **(A1)** Separation of time scales: FA binding/unbinding occur on the fast time scale, cell movement and contractility on the same, intermediate scale, and appreciable changes to the density function occur on the slowest time scale.
- **(A2)** The probabilities of finding a cell in motile or non-motile states are equal due to **(A1)**.
- **(A3)** Cell rotations are negligible, or equivalently $\beta_{rot} \gg 1$. Thus, for simplicity, we take $\theta = 0$ (recall 2.1 for details).

Since a cell body translocation requires appropriate contraction of SFs, it is justified to assume that the movement and contractility occur on the same time scale. Note that mesenchymal cells migrate at a speed of several micrometers per minute [70], [90] (see also the simulation results in Section 4.3). Thus, a cell covers an appreciable distance (relative to its size) on a scale of hours¹. Recalling our discussions in Sections 3.1, 4.2-4.3.1, we also see that the FA events occur on a time scale of seconds. This justifies the assumption

¹Characteristic cell size is $\sim 10 - 100\mu m$.

(A1). Recall that in our model FA disassociation leads to cell body translocation, while after an FA binding event the cell becomes non-motile. Since both types of events occur on the fast time scale (due to (A1)), it is reasonable to assume that a cell can be found in either state with equal likelihood, which justifies assumption (A2). The third assumption above is a simplification, which can easily be relaxed without affecting what follows next.

Due to (A3) and (2.1.2), \mathbf{F} is independent of θ , and $\dot{\theta} = 0$ regardless of the motility state. In view of (A2), we are now interested in how the density of the moving cells evolves. Abusing the notation, we let $\boldsymbol{\alpha} : \{1, \dots, 2^M\} \rightarrow \{0, 1\}^M$ be a bijection, such that $\boldsymbol{\alpha}(\nu) = \mathbf{Y}$ corresponds to a particular FA state (recall Section 2.4.2), and let $\mathbf{X} = (\mathbf{x}, \mathbf{x}_n)$. Thus, the density of the (moving) cells $f(t, \nu, \mathbf{X})$ follows (7.2.5), where we again abuse the notation:

$$\mathbf{H}_\nu(\mathbf{X}) = \left(\frac{\beta_{ECM}^{-1} \mathbf{F}(\mathbf{x}_n, \boldsymbol{\alpha}(\nu)) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) \hat{\mathbf{r}}(\mathbf{x}_n)}{\beta_{cell}^{-1} \mathbf{F}(\mathbf{x}_n, \boldsymbol{\alpha}(\nu))} \right), \quad (7.3.1)$$

$$\begin{aligned} q(\eta; \nu, \mathbf{X}) &= \sum_{j=1}^M \frac{a_j^+(\boldsymbol{\alpha}(\nu), \mathbf{X})}{a_0(\boldsymbol{\alpha}(\nu), \mathbf{X})} \delta_{\boldsymbol{\alpha}(\eta)_j, 1} \prod_{i \neq j}^M \delta_{\boldsymbol{\alpha}(\eta)_i, \boldsymbol{\alpha}(\nu)_i} \\ &\quad + \frac{a_j^-(\boldsymbol{\alpha}(\nu), \mathbf{X})}{a_0(\boldsymbol{\alpha}(\nu), \mathbf{X})} \delta_{\boldsymbol{\alpha}(\eta)_j, 0} \prod_{i \neq j}^M \delta_{\boldsymbol{\alpha}(\eta)_i, \boldsymbol{\alpha}(\nu)_i}, \end{aligned} \quad (7.3.2)$$

and $\nu \in \{1, \dots, 2^M\}$. Note that due to (A2), the terms involving the cell motility state are omitted in (7.3.1)-(7.3.2) (compare with (2.4.9) and (7.2.4), respectively). Before proceeding, we take a closer look at the FA dynamics.

7.3.1 Probabilities of FA states

Let $\hat{\pi}(t, \mathbf{y} | \mathbf{X})$ be the probability that at time t the vector of FA states is $\mathbf{y} \in \{0, 1\}^M$, given \mathbf{X} . Following our discussion on FA dynamics in Section 3.1, we see that, given $\mathbf{X} = (\mathbf{x}, \mathbf{x}_n)$, binding and unbinding of a single FA are independent of time and of the state of other FAs (see also (3.1.1), (3.1.4)). This implies that

$$\hat{\pi}(t, \mathbf{y} | \mathbf{X}) = \prod_{i=1}^M \hat{\pi}_i(t, y_i | \mathbf{X}),$$

where $\hat{\pi}_i(t, y | \mathbf{X})$ is the conditional probability that the state of the i^{th} FA is $y \in \{0, 1\}$. It is governed by the following master equation:

$$\frac{\partial}{\partial t} \hat{\pi}_i(t, y = 1 | \mathbf{X}) = a_i^+(\mathbf{X}) \hat{\pi}_i(t, y = 0 | \mathbf{X}) - a_i^-(\mathbf{X}) \hat{\pi}_i(t, y = 1 | \mathbf{X}),$$

where $\hat{\pi}_i(t, y | \mathbf{X})$ is the probability that the i^{th} FA is bound for $y = 1$, or unbound for $y = 0$, a_i^+ and a_i^- are the corresponding binding and unbinding probability rates, respectively (see (3.1.1), (3.1.4) for their forms). Obviously it holds that

$$\hat{\pi}_i(t, y = 1 | \mathbf{X}) + \hat{\pi}_i(t, y = 0 | \mathbf{X}) = 1.$$

Using the constraint above, the steady state distribution of a single FA is given by:

$$\begin{aligned}\hat{\pi}_i(y = 1 \mid \mathbf{X}) &= \frac{a_i^+(\mathbf{X})}{a_i^-(\mathbf{X})} \left(1 + \frac{a_i^+(\mathbf{X})}{a_i^-(\mathbf{X})}\right)^{-1} \\ \hat{\pi}_i(y = 0 \mid \mathbf{X}) &= \left(1 + \frac{a_i^+(\mathbf{X})}{a_i^-(\mathbf{X})}\right)^{-1},\end{aligned}\quad (7.3.3)$$

and hence

$$\hat{\pi}(\mathbf{y} \mid \mathbf{X}) = \prod_{i=1}^M \hat{\pi}_i(y \mid \mathbf{X}) \quad (7.3.4)$$

is the steady state distribution of M FAs.

Given \mathbf{X} , consider now a jump process, which a collection of M FAs follow. Define $\pi(t, \nu \mid \mathbf{X}) := \hat{\pi}(t, \boldsymbol{\alpha}(\nu) \mid \mathbf{X})$, where π is the conditional probability that at time t the process has state ν corresponding to the state of FAs given by $\boldsymbol{\alpha}(\nu)$. This is well defined, since $\boldsymbol{\alpha}$ is a bijection. This jump process is described by the following master equation:

$$\frac{\partial}{\partial t} \pi(t, \nu \mid \mathbf{X}) = -a_0(\nu, \mathbf{X}) \pi(t, \nu \mid \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu; \eta, \mathbf{X}) \pi(t, \eta \mid \mathbf{X}),$$

where $q(\nu; \eta, \mathbf{X})$ is given by (7.3.2). Consequently, the steady state distribution of such process is simply $\pi(\nu \mid \mathbf{X}) = \hat{\pi}(\boldsymbol{\alpha}(\nu) \mid \mathbf{X})$, i.e.

$$0 = -a_0(\nu, \mathbf{X}) \pi(\nu \mid \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu; \eta, \mathbf{X}) \pi(\eta \mid \mathbf{X}). \quad (7.3.5)$$

7.3.2 Scaling and perturbation expansion

We rescale the time as $\hat{t} = \epsilon t$. Dropping the hat, (7.2.5) becomes:

$$\begin{aligned}\epsilon \frac{\partial}{\partial t} f(t, \nu, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\nu}(\mathbf{X}) f(t, \nu, \mathbf{X})) \\ = \frac{1}{\epsilon} \left[-a_0(\nu, \mathbf{X}) f(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu; \eta, \mathbf{X}) f(t, \eta, \mathbf{X}) \right].\end{aligned}\quad (7.3.6)$$

The factor ϵ^{-1} on the right hand side of the equation above stems from our assumption **(A1)** about the separation of time scales (see also its justification): FA events occur on a time scale (seconds) that is much faster than changes represented by the left hand side (hours and minutes).

Using a Hilbert expansion $f = f_0 + \epsilon f_1 + \epsilon^2 f_2 + \dots$, plugging it in the equation above, and collecting the terms of equal order in ϵ :

ϵ^{-1} :

$$0 = -a_0(\nu, \mathbf{X}) f_0(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu; \eta, \mathbf{X}) f_0(t, \eta, \mathbf{X}). \quad (7.3.7)$$

ϵ^0 :

$$\begin{aligned} \nabla_{\mathbf{X}} \cdot (\mathbf{H}_\nu(\mathbf{X})f_0(t, \nu, \mathbf{X})) \\ = -a_0(\nu, \mathbf{X})f_1(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f_1(t, \eta, \mathbf{X}). \end{aligned} \quad (7.3.8)$$

ϵ^1 :

$$\begin{aligned} \frac{\partial}{\partial t} f_0(t, \nu, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot (\mathbf{H}_\nu(\mathbf{X})f_1(t, \nu, \mathbf{X})) \\ = -a_0(\nu, \mathbf{X})f_2(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f_2(t, \eta, \mathbf{X}). \end{aligned} \quad (7.3.9)$$

Note that by the definition of conditional density, we have:

$$f(t, \nu, \mathbf{X}) = \pi(t, \nu \mid \mathbf{X})\rho(t, \mathbf{X}),$$

where $\rho(t, \mathbf{X})$ is the marginal probability density, i.e.:

$$\begin{aligned} \rho(t, \mathbf{X}) &= \sum_{\nu} f(t, \nu, \mathbf{X}) = \sum_{\nu} f_0(t, \nu, \mathbf{X}) + \epsilon \sum_{\nu} f_1(t, \nu, \mathbf{X}) + \epsilon^2 \sum_{\nu} f_2(t, \nu, \mathbf{X}) + \dots \\ &= \rho_0(t, \mathbf{X}) + \epsilon\rho_1(t, \mathbf{X}) + \epsilon^2\rho_2(t, \mathbf{X}) + \dots \end{aligned}$$

Recalling (7.3.5), we see that

$$f_0(t, \nu, \mathbf{X}) = \pi(\nu \mid \mathbf{X})\rho_0(t, \mathbf{X}) \quad (7.3.10)$$

satisfies (7.3.7).

Let $\mathcal{Q} : \mathbb{R}^2 \times \Omega_{cell} \rightarrow \mathbb{R}^{2^M \times 2^M}$, $\mathbf{X} \mapsto \mathcal{Q}(\mathbf{X})$ be the matrix-valued function¹, whose entries $\mathcal{Q}_{\nu\eta}(\mathbf{X})$ are given by:

$$\mathcal{Q}_{\nu\eta}(\mathbf{X}) = \begin{cases} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X}), & \text{if } \nu \neq \eta \\ -a_0(\nu, \mathbf{X}), & \text{if } \nu = \eta. \end{cases}$$

Assuming $\rho_0 \neq 0$, it follows from (7.3.7) that

$$\mathcal{Q}\boldsymbol{\pi} = 0, \quad (7.3.11)$$

where the elements of the vector $\boldsymbol{\pi} = \boldsymbol{\pi}(\mathbf{X}) \in \mathbb{R}^{2^M}$ are given by

$$\boldsymbol{\pi}_\nu(\mathbf{X}) = \pi(\nu \mid \mathbf{X}).$$

Note that $\boldsymbol{\pi}$ is the nonzero unique solution of (7.3.11) due to (7.3.3)-(7.3.4).

¹Recalling Section 2.2.3, Ω_{cell} is a unit disk.

Remark. From (7.3.11) we see that, for fixed \mathbf{X} , $\boldsymbol{\pi}(\mathbf{X})$ is a nonzero eigenvector corresponding to a zero eigenvalue of the matrix $\mathcal{Q}(\mathbf{X})$. Since the spectrum of \mathcal{Q}^T is identical to that of \mathcal{Q} , and \mathcal{Q}^T is given by

$$\mathcal{Q}_{\nu\eta}^T(\mathbf{X}) = \begin{cases} a_0(\nu, \mathbf{X})q(\eta; \nu, \mathbf{X}), & \text{if } \nu \neq \eta \\ -a_0(\nu, \mathbf{X}), & \text{if } \nu = \eta, \end{cases}$$

then by the Gershgorin's Theorem we see that the eigenvalues are, for fixed \mathbf{X} , in the set

$$\{z \in \mathbb{C} : |z - (-a_0(\nu, \mathbf{X}))| \leq a_0(\nu, \mathbf{X}), 1 \leq \nu \leq 2^M\},$$

because

$$\sum_{\eta} a_0(\nu, \mathbf{X})q(\eta; \nu, \mathbf{X}) = a_0(\nu, \mathbf{X}).$$

Thus, all of $2^M - 1$ nonzero eigenvalues have negative real part.

Consider now (7.3.8). We can rewrite it in the following form:

$$\widehat{\mathbf{g}}_0 = \mathcal{Q}\widehat{\mathbf{f}}_1,$$

where the entries $\widehat{g}_{0,\nu}$ and $\widehat{f}_{1,\nu}(\mathbf{X})$ of the vectors $\widehat{\mathbf{g}}_0(\mathbf{X})$, $\widehat{\mathbf{f}}_1(\mathbf{X}) \in \mathbb{R}^{2^M}$, respectively, are given by:

$$\begin{aligned} \widehat{g}_{0,\nu}(\mathbf{X}) &= \nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\nu}(\mathbf{X})f_0(t, \nu, \mathbf{X})) \\ \widehat{f}_{1,\nu}(\mathbf{X}) &= f_1(t, \nu, \mathbf{X}). \end{aligned}$$

Although \mathcal{Q} is not invertible due to (7.3.11), the restriction $\mathcal{Q}|_{\ker(\mathcal{Q})^\perp} : \ker(\mathcal{Q})^\perp \rightarrow \text{ran}(\mathcal{Q})$ has an inverse $(\mathcal{Q}|_{\ker(\mathcal{Q})^\perp})^{-1} =: \mathcal{F} : \text{ran}(\mathcal{Q}) \rightarrow \ker(\mathcal{Q})^\perp$, where $\ker(\mathcal{Q})^\perp$ is an orthogonal complement of the kernel $\ker(\mathcal{Q}) = \text{span}(\{\boldsymbol{\pi}\})$. We then have:

$$\widehat{\mathbf{f}}_1 = \mathcal{F}\widehat{\mathbf{g}}_0 \iff f_1(t, \nu, \mathbf{X}) = \sum_{\eta} \mathcal{F}_{\nu\eta}(\mathbf{X})\nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\eta}(\mathbf{X})f_0(t, \eta, \mathbf{X})),$$

where $\mathcal{F}_{\nu\eta}$ is an entry of the matrix \mathcal{F} . Plugging the expression for f_1 in (7.3.9), we obtain:

$$\begin{aligned} \frac{\partial}{\partial t} f_0(t, \nu, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot \left(\mathbf{H}_{\nu}(\mathbf{X}) \sum_{\eta} \mathcal{F}_{\nu\eta}(\mathbf{X})\nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\eta}(\mathbf{X})f_0(t, \eta, \mathbf{X})) \right) \\ = -a_0(\nu, \mathbf{X})f_2(t, \nu, \mathbf{X}) + \sum_{\eta} a_0(\eta, \mathbf{X})q(\nu; \eta, \mathbf{X})f_2(t, \eta, \mathbf{X}). \end{aligned}$$

Plugging in the expression for f_0 , given by (7.3.10), and summing over ν , yields:

$$\frac{\partial}{\partial t} \rho_0(t, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot \left(\sum_{\nu} \mathbf{H}_{\nu}(\mathbf{X}) \sum_{\eta} \mathcal{F}_{\nu\eta}(\mathbf{X})\nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\eta}(\mathbf{X})\boldsymbol{\pi}(\eta | \mathbf{X})\rho_0(t, \mathbf{X})) \right) = 0.$$

Expanding and rearranging the terms in the parenthesis gives:

$$\frac{\partial}{\partial t}\rho_0(t, \mathbf{X}) + \nabla_{\mathbf{X}} \cdot (\mathbb{D}(\mathbf{X})\nabla_{\mathbf{X}}\rho_0(t, \mathbf{X})) + \nabla_{\mathbf{X}} \cdot (\mathbf{u}(\mathbf{X})\rho_0(t, \mathbf{X})) = 0, \quad (7.3.12)$$

where

$$\begin{aligned} \mathbb{D}(\mathbf{X}) &= \sum_{\nu} \sum_{\eta} \mathbf{H}_{\nu}(\mathbf{X}) \otimes \mathbf{H}_{\eta}(\mathbf{X}) \mathcal{F}_{\nu\eta}(\mathbf{X}) \pi(\eta | \mathbf{X}) \\ \mathbf{u}(\mathbf{X}) &= \sum_{\nu} \sum_{\eta} \mathbf{H}_{\nu}(\mathbf{X}) \nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\eta}(\mathbf{X}) \pi(\eta | \mathbf{X})) \mathcal{F}_{\nu\eta}(\mathbf{X}). \end{aligned} \quad (7.3.13)$$

Although it is possible to compute \mathcal{F} in general (e.g. by using singular value decomposition), if either or both of the below specified conditions **(C1)**, **(C2)** are satisfied, then we can obtain explicit solution f_1 in (7.3.8). Given ν', t, \mathbf{X} , these conditions are:

- **(C1)** $\sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu'; \eta, \mathbf{X}) f_1(t, \eta, \mathbf{X}) = 0$.
- **(C2)** $\sum_{\nu} a_0(\nu, \mathbf{X}) q(\nu'; \nu, \mathbf{X}) f_1(t, \nu, \mathbf{X}) = \sum_{\nu} q(\nu'; \nu, \mathbf{X}) \sum_{\eta} a_0(\eta, \mathbf{X}) q(\nu; \eta, \mathbf{X}) f_1(t, \eta, \mathbf{X})$.

Obviously, **(C1)** implies **(C2)**.

Remark. The right hand side in **(C2)** can be rewritten as

$$\sum_{\eta} a_0(\eta, \mathbf{X}) \left[\sum_{\nu} q(\nu'; \nu, \mathbf{X}) q(\nu; \eta, \mathbf{X}) \right] f_1(t, \eta, \mathbf{X}).$$

If it were true that

$$\sum_{\nu} q(\nu'; \nu, \mathbf{X}) q(\nu; \eta, \mathbf{X}) = q(\nu'; \eta, \mathbf{X}),$$

then **(C2)** would be satisfied. Unfortunately, in our case the transitivity property does not hold. For example, if $M = 2$, then the transitions $1 \leftrightarrow 2, 1 \leftrightarrow 3, 2 \leftrightarrow 4, 3 \leftrightarrow 4$ corresponding to the transitions of FA states

$$\begin{aligned} \boldsymbol{\alpha}(1) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \boldsymbol{\alpha}(2), & \boldsymbol{\alpha}(1) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \boldsymbol{\alpha}(3) \\ \boldsymbol{\alpha}(2) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \boldsymbol{\alpha}(4), & \boldsymbol{\alpha}(3) &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \boldsymbol{\alpha}(4) \end{aligned}$$

occur with nonzero probability. However, the probability of $1 \leftrightarrow 4$ transition is zero, since only one FA event (binding or unbinding) occurs at the time of transition (see Section 2.3.1 and (7.3.2)). Thus,

$$0 < \sum_{\nu=1}^4 q(4; \nu, \mathbf{X}) q(\nu; 1, \mathbf{X}) \neq q(4; 1, \mathbf{X}) = 0.$$

Nevertheless, for problems where the transitivity property is satisfied, the procedure described here can be applied, as we do not extensively rely on the specifics of our cell motility model. Of course, the transitivity property need not hold for **(C2)** to be true.

If **(C1)** holds, then from (7.3.8) we have:

$$f_1(t, \nu, \mathbf{X}) = -\frac{1}{a_0(\nu, \mathbf{X})} \nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\nu}(\mathbf{X}) f_0(t, \nu, \mathbf{X})). \quad (7.3.14)$$

Suppose **(C2)** is true. Multiplying (7.3.8) by $q(\nu'; \nu, \mathbf{X})$ and summing over ν , we obtain:

$$\begin{aligned} \sum_{\nu} q(\nu'; \nu) g(\nu) &= -\sum_{\nu} a_0(\nu) q(\nu'; \nu) f_1(\nu) + \sum_{\nu} q(\nu'; \nu) \sum_{\eta} a_0(\eta) q(\nu; \eta) f_1(\eta) \\ &= 0, \end{aligned} \quad (7.3.15)$$

where $g(\nu) = \nabla \cdot (\mathbf{H}_{\nu} f_0(\nu))$ and for clarity we omit the dependence on t and \mathbf{X} . From (7.3.8), we obtain:

$$f_1(\nu) = -\frac{1}{a_0(\nu)} g(\nu) + \sum_{\eta} \frac{a_0(\eta)}{a_0(\nu)} q(\nu; \eta) f_1(\eta), \quad (7.3.16)$$

with $g(\nu)$ as above.

It follows from (7.3.15) that $f_1(\nu) = -\frac{1}{a_0(\nu)} g(\nu)$ solves (7.3.16), i.e. f_1 is given by (7.3.14). Therefore, \mathbb{D} and \mathbf{u} are simplified to

$$\begin{aligned} \mathbb{D}(\mathbf{X}) &= -\sum_{\nu} \frac{1}{a_0(\nu, \mathbf{X})} \mathbf{H}_{\nu}(\mathbf{X}) \otimes \mathbf{H}_{\nu}(\mathbf{X}) \pi(\nu | \mathbf{X}) \\ \mathbf{u}(\mathbf{X}) &= -\sum_{\nu} \frac{1}{a_0(\nu, \mathbf{X})} \mathbf{H}_{\nu}(\mathbf{X}) \nabla_{\mathbf{X}} \cdot (\mathbf{H}_{\nu}(\mathbf{X}) \pi(\nu | \mathbf{X})). \end{aligned} \quad (7.3.17)$$

The employed procedure to obtain (7.3.12) from (7.2.5) is analogous to the one used in [33], [48], [88], [89]. Moreover, the form of our drift-diffusion equation (7.3.12) is similar to the one obtained in [33]. This is, in part, due to the fact that multiplying (7.3.6) by ϵ effectively introduces “diffusion”-like scaling with respect to time and “space” variable \mathbf{X} , while the gain and the loss terms on the right hand side remain zeroth order in ϵ . Also, in both cases the unscaled equations originate from jump processes that are special cases of piecewise deterministic Markov processes. Whereas the turning operator in [33], [48], [88], [89] directly acts on the velocity variable, here, informally, it acts on the subcell level determinants of the cell velocity (see (7.2.5)), thereby inducing the changes in the cell velocity. Specifically, the (de)adhesion events lead to velocity alteration, i.e. turning. Consequently, the turning rate, represented by a_0 , depends on \mathbf{X} and the state of cell-substrate adhesions $\boldsymbol{\alpha}(\nu)$. The dependence of the turning rate on the receptor state is also taken into account in [33], [56]. In contrast to [33], [56], however, the turning rate a_0 reflects the individual state of adhesions (bound and unbound), along with the effects of

contractile mechanism driving cell locomotion (dependence on $\mathbf{X} = (\mathbf{x}, \mathbf{x}_n)$). Thus, a closer look at the subcellular effects on adhesion dynamics allows us to deduce a more precise description of such feature of cell motility as the turning rate.

There are also stark differences in the underlying physical model. In [33] the authors consider fast moving, amoeboid type of migration with constant speeds on the order $\sim 2\mu\text{m}/\text{s}$ (the speeds in our case are on the scale of several $\mu\text{m}/\text{h}$). In their multi-scale model, the directional changes are modulated by the local fiber orientation and cell-substrate interaction through integrin-ECM bonds. In contrast, the underlying model in our case describes mesenchymal type of migration. In our model, we explicitly include the mechanical coupling between the prominent cellular structures and how this interaction results in cell locomotion. Here we see that the details of interactions occurring on cell and subcell levels are included in population scale description: the diffusion tensor \mathbb{D} and the advection field \mathbf{u} incorporate the information about the low level interactions (see (7.3.13)). In particular, the dynamic changes in the cell's contractile machinery ($\mathbf{H}_\nu(\mathbf{X})$), and its effects on FA kinetics ($\pi(\nu | \mathbf{X})$) are reflected in \mathbb{D} and \mathbf{u} . Note that the dependence of \mathbb{D} and \mathbf{u} on the probability rates a_i^\pm (see e.g. (7.3.3)) correspond to a rather localized sensing mechanism in our model (see e.g. (3.1.4)): the FAs respond to the local cue rather than the local cue gradient. Therefore, (7.3.12) contains only the myopic diffusion terms \mathbb{D} and \mathbf{u} , unlike the macroscopic scale equation in [33] where “genuine” taxis terms are also present. This is not surprising, because the simulations of the cell scale model in Subsection 4.3.2 show that the tactic migration emerges from taking into account the local information about an external cue and without a centralized gradient sensing mechanism. Moreover, the simulations in Subsection 4.3.3 show that such “myopic sensing” can resolve the fine details of the underlying ECM architecture, which is, consequently, incorporated in \mathbb{D} and \mathbf{u} .

In Section 4.3 we saw that the surrounding environment affects cell motility by modifying FA dynamics. Recalling our discussion in Section 3.2 on coupling RhoGTPase signaling pathways with FA dynamics, we see that the procedure above is capable of connecting various processes on distinct spatiotemporal scales that are involved in cell locomotion. For example, since RhoA mediates myosin forces (see Section 3.3) and hence contractile force in an SF (see (2.1.1)), then the effects of RhoA manifest themselves on the population level by its influence on \mathbf{H}_ν (see (7.3.1)). Analogously, since Rac1 mediates FA dynamics (see Section 3.2), then its effects on the population level are reflected in the distribution π (see (7.3.3)-(7.3.4)).

7.3.3 The diffusion tensor and the advection field

In the following we assume that either of the conditions **(C1)** and **(C2)** hold. Thus, the focus of the discussion here is to compute the diffusion tensor \mathbb{D} and the advection field \mathbf{u} , whose formulas are given by (7.3.17).

Making the transformation $\{1, \dots, 2^M\} \ni \nu \rightarrow \boldsymbol{\alpha}(\nu) = \mathbf{y} \in \{0, 1\}^M$, we have the

following:

$$\mathbf{H}_\nu(\mathbf{X}) \rightarrow \mathbf{H}_y(\mathbf{X}) = \underbrace{\begin{bmatrix} \beta_{ECM}^{-1} \mathcal{R}(\mathbf{x}_n) \mathbf{B}(\mathbf{x}_n) & 0 \\ 0 & \beta_{cell}^{-1} \mathbf{B}(\mathbf{x}_n) \end{bmatrix}}_{:=\mathbf{C}(\mathbf{x}_n)} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}, \quad (7.3.18)$$

where

$$\begin{aligned} \mathbf{B}(\mathbf{x}_n) &= [-\mathbf{F}_1(\mathbf{x}_n) \mid -\mathbf{F}_2(\mathbf{x}_n) \mid \cdots \mid -\mathbf{F}_M(\mathbf{x}_n)] \\ \mathcal{R}(\mathbf{x}_n) &= \begin{bmatrix} \hat{r}_1^2(\mathbf{x}_n) & \hat{r}_1(\mathbf{x}_n) \hat{r}_2(\mathbf{x}_n) \\ \hat{r}_1(\mathbf{x}_n) \hat{r}_2(\mathbf{x}_n) & \hat{r}_2^2(\mathbf{x}_n) \end{bmatrix}, \end{aligned}$$

since $\mathbf{H}_\nu(\mathbf{X})$ is given by (7.3.1), $\mathbf{F}(\mathbf{x}_n, \boldsymbol{\alpha}(\nu) = \mathbf{y}) = \sum_{i=1}^M (-\mathbf{F}_i(\mathbf{x}_n)) y_i$ due to (2.1.2)¹ and $\theta = 0$, and $\hat{\mathbf{r}}(\mathbf{x}_n)$ is given by (2.2.2). Then,

$$\begin{aligned} \mathbb{D}(\mathbf{X}) &= - \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y}, \mathbf{X})} \mathbf{C}(\mathbf{x}_n) \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \otimes \mathbf{C}(\mathbf{x}_n) \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \hat{\pi}(\mathbf{y} \mid \mathbf{X}) \\ &= -\mathbf{C}(\mathbf{x}_n) \left[\sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y}, \mathbf{X})} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \hat{\pi}(\mathbf{y} \mid \mathbf{X}) \right] \mathbf{C}^T(\mathbf{x}_n), \end{aligned} \quad (7.3.19)$$

where, as in Section 2.4.2, we abuse the notation and set $a_0(\mathbf{y}, \mathbf{X}) = a_0(\nu, \mathbf{X})$. The square matrix in the brackets has a block structure, where each block consists of the matrix $\mathbf{K} \in \mathbb{R}^{M \times M}$, defined as (omitting the dependence on \mathbf{X} for clarity)

$$\mathbf{K} := \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{y} \mathbf{y}^T \hat{\pi}(\mathbf{y}),$$

with off-diagonal entries K_{ij} given by

$$\begin{aligned} K_{ij} &= \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} y_i y_j \hat{\pi}(\mathbf{y}) = \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} y_i y_j \hat{\pi}_i(y_i) \hat{\pi}_j(y_j) \prod_{k \neq i,j} \hat{\pi}_k(y_k) \\ &= \hat{\pi}_i(y_i = 1) \hat{\pi}_j(y_j = 1) \sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_i, y_j = 1}} \frac{1}{a_0(\mathbf{y})} \prod_{k \neq i,j} \hat{\pi}_k(y_k), \end{aligned}$$

where the second equality follows from (7.3.4), and the diagonal entries are given by

$$K_{ii} = \hat{\pi}_i(y_i = 1) \sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_i = 1}} \frac{1}{a_0(\mathbf{y})} \prod_{k \neq i} \hat{\pi}_k(y_k).$$

¹It is straightforward to extend the discussion below for the case when \mathbf{F}_i depends on \mathbf{x} , i.e. when, for example, myosin force generation depends on an external cue. In this case, $\mathbf{C} = \mathbf{C}(\mathbf{X})$

We can either compute the sums directly, or approximate them in the following way.

Recalling our discussion in Section 2.4.2 (in particular (2.4.10)), we can rewrite $a_0(\mathbf{y}, \mathbf{X})$ as:

$$a_0(\mathbf{y}, \mathbf{X}) = \sum_{m=1}^M a_m^+(\mathbf{X})(1 - y_m) + a_m^-(\mathbf{X})y_m = \mathbf{a}^+(\mathbf{X}) \cdot \mathbf{e} + (\mathbf{a}^-(\mathbf{X}) - \mathbf{a}^+(\mathbf{X})) \cdot \mathbf{y},$$

where $a_m^\pm(\mathbf{X})$ is defined as in Section 7.3.1, $\mathbf{a}^\pm(\mathbf{X}) \in \mathbb{R}^M$ is a vector with entries $a_m^\pm(\mathbf{X})$, and $\mathbf{e} = [1, \dots, 1]^T \in \mathbb{R}^M$. Note that if $\mathbf{Y} \in \{0, 1\}^M$ is a random variable with (conditional) distribution function $\hat{\pi}$, then

$$\begin{aligned} \bar{\mu} &:= \mathbb{E}[a_0(\mathbf{Y})] = \mathbf{a}^+ \cdot \mathbf{e} + (\mathbf{a}^- - \mathbf{a}^+) \cdot \mathbb{E}[\mathbf{Y}] = \mathbf{a}^+ \cdot \mathbf{e} + (\mathbf{a}^- - \mathbf{a}^+) \cdot \bar{\mathbf{g}} \\ &= \mathbf{a}^+ \cdot \mathbf{e} + \sum_{m=1}^M (a_m^- - a_m^+) \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-}\right)^{-1}, \\ \sigma^2 &:= \text{Var}[a_0(\mathbf{Y})] = \mathbb{E}[(a_0(\mathbf{Y}) - \mathbb{E}[a_0(\mathbf{Y})])^2] = \text{Var}[(\mathbf{a}^- - \mathbf{a}^+) \cdot \mathbf{Y}] \\ &= \sum_{m=1}^M (a_m^- - a_m^+)^2 \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-}\right)^{-2}, \end{aligned}$$

where

$$\bar{\mathbf{g}} = \begin{bmatrix} \hat{\pi}_1(y_1 = 1) \\ \hat{\pi}_2(y_2 = 1) \\ \vdots \\ \hat{\pi}_M(y_M = 1) \end{bmatrix}.$$

Noting that $a_0 > 0$ (and hence $\bar{\mu} > 0$), expanding $1/a_0(\mathbf{Y})$ in Taylor series around $\bar{\mu}$, and taking expectation, we get:

$$\begin{aligned} \mathbb{E}\left[\frac{1}{a_0(\mathbf{Y})}\right] &= \frac{1}{\bar{\mu}} + \sum_{n=1}^{\infty} (-1)^n \frac{n}{n!} \bar{\mu}^{-(n+1)} \mathbb{E}[(a_0(\mathbf{Y}) - \bar{\mu})^n] \\ &= \frac{1}{\bar{\mu}} + \frac{1}{\bar{\mu}^3} \sigma^2 + o(\sigma^2). \end{aligned}$$

The second equality follows from the fact that for $n \geq 1$ we have

$$\begin{aligned} \mathbb{E}[|a_0(\mathbf{Y}) - \bar{\mu}|^n] &= \mathbb{E}\left[\left|\sum_{m=1}^M (a_m^- - a_m^+) (Y_m - \bar{g}_m)\right|^n\right] \\ &\leq \mathbb{E}\left[\left(\sum_{m=1}^M |a_m^- - a_m^+|\right)^n\right] = \|\mathbf{a}^- - \mathbf{a}^+\|_1^n \end{aligned}$$

due to the triangle inequality, and $|Y_m - \bar{g}_m| \leq 1$ a.s., since $Y_m \in \{0, 1\}$ and $\bar{g}_m \leq 1$.

Similarly as above, we have:

$$\begin{aligned}
\bar{\mu}_{ij} &:= \mathbb{E}[a_0(\mathbf{Y}) \mid Y_i, Y_j = 1] \\
&= \mathbf{a}^+ \cdot \mathbf{e} + \sum_{\substack{m=1 \\ m \neq i, j}}^M (a_m^- - a_m^+) \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-}\right)^{-1} + \sum_{m=i, j} (a_m^- - a_m^+) \\
\sigma_{ij}^2 &:= \text{Var}[a_0(\mathbf{Y}) \mid Y_i, Y_j = 1] \\
&= \sum_{\substack{m=1 \\ m \neq i, j}}^M (a_m^- - a_m^+)^2 \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-}\right)^{-2}, \tag{7.3.20}
\end{aligned}$$

and

$$\sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_i, y_j = 1}} \frac{1}{a_0(\mathbf{y})} \prod_{k \neq i, j} \hat{\pi}_k(y_k) = \mathbb{E}[a_0^{-1}(\mathbf{Y}) \mid Y_i, Y_j = 1] = \frac{1}{\bar{\mu}_{ij}} + \frac{1}{\bar{\mu}_{ij}^3} \sigma_{ij}^2 + o(\sigma_{ij}^2).$$

Therefore, we can approximate the entries K_{ij} as

$$K_{ij} = \begin{cases} \bar{\mathbf{g}}_i \bar{\mathbf{g}}_j (\bar{\mu}_{ij}^{-1} + \bar{\mu}_{ij}^{-3} \sigma_{ij}^2), & \text{if } i \neq j \\ \bar{\mathbf{g}}_i (\bar{\mu}_{ii}^{-1} + \bar{\mu}_{ii}^{-3} \sigma_{ii}^2), & \text{if } i = j. \end{cases} \tag{7.3.21}$$

Consequently, \mathbb{D} is given by

$$\mathbb{D} = -\mathbf{C} \begin{bmatrix} \mathbf{K} & \mathbf{K} \\ \mathbf{K} & \mathbf{K} \end{bmatrix} \mathbf{C}^T. \tag{7.3.22}$$

We now turn our attention to the advection field $\mathbf{u}(\mathbf{X})$, which is given by (7.3.17). Under the transformation $\nu \rightarrow \boldsymbol{\alpha}(\nu)$, \mathbf{u} becomes:

$$\begin{aligned}
\mathbf{u} &= - \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \nabla \cdot (\mathbf{H}_{\mathbf{y}} \hat{\pi}(\mathbf{y})) = - \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \otimes \mathbf{H}_{\mathbf{y}} \nabla \hat{\pi}(\mathbf{y}) \\
&\quad - \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \nabla \cdot \mathbf{H}_{\mathbf{y}} \hat{\pi}(\mathbf{y}).
\end{aligned}$$

Due to (7.3.4), we have:

$$\nabla \hat{\pi}(\mathbf{y}) = \sum_{k=1}^M \nabla \hat{\pi}_k(y_k) \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m),$$

and thus

$$\sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \otimes \mathbf{H}_{\mathbf{y}} \nabla \hat{\pi}(\mathbf{y}) = \sum_{\mathbf{y} \in \{0,1\}^M} \sum_{k=1}^M \frac{1}{a_0(\mathbf{y})} \mathbf{C} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \mathbf{C}^T \nabla \hat{\pi}_k(y_k) \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m)$$

$$\begin{aligned}
&= \sum_{k=1}^M \sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_k=0}} \frac{1}{a_0(\mathbf{y})} \mathbf{C} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \mathbf{C}^T \nabla \hat{\pi}_k(y_k = 0) \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m) \\
&+ \sum_{k=1}^M \sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_k=1}} \frac{1}{a_0(\mathbf{y})} \mathbf{C} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \mathbf{C}^T \nabla \hat{\pi}_k(y_k = 1) \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m) \\
&= \mathbf{C} \sum_{k=1}^M \left[\sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_k=0}} \frac{1}{a_0(\mathbf{y})} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m) \right] \mathbf{C}^T \nabla \hat{\pi}_k(y_k = 0) \\
&+ \mathbf{C} \sum_{k=1}^M \left[\sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_k=1}} \frac{1}{a_0(\mathbf{y})} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix}^T \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m) \right] \mathbf{C}^T \nabla \hat{\pi}_k(y_k = 1).
\end{aligned}$$

The terms in the brackets above are similar to those in the brackets of (7.3.19). Here we also have block matrices, where each block consists of matrices $\mathbf{K}^{k,0}$ or $\mathbf{K}^{k,1}$ corresponding to the terms with $y_k = 0$ and $y_k = 1$, respectively, and which are given by

$$\mathbf{K}^{k,\tilde{y}} := \sum_{\substack{\mathbf{y} \in \{0,1\}^M \\ y_k=\tilde{y}}} \frac{1}{a_0(\mathbf{y})} \mathbf{y} \mathbf{y}^T \prod_{\substack{m=1 \\ m \neq k}}^M \hat{\pi}_m(y_m),$$

where $\tilde{y} \in \{0,1\}$. The computation of the entries $\mathbf{K}^{k,\tilde{y}}$ proceeds in the same manner as above, except that the expectations in (7.3.20) are also conditioned on $Y_k = \tilde{y}$. Therefore, the entries of $\mathbf{K}^{k,\tilde{y}}$ can be approximated as follows:

$$K_{ij}^{k,0} = \begin{cases} \bar{\mathbf{g}}_i \bar{\mathbf{g}}_j \left((\bar{\mu}_{ij}^{k,0})^{-1} + (\bar{\mu}_{ij}^{k,0})^{-3} (\sigma_{ij}^k)^2 \right), & \text{if } i \neq j \text{ and } k \neq i, j \\ \bar{\mathbf{g}}_i \left((\bar{\mu}_{ii}^{k,0})^{-1} + (\bar{\mu}_{ii}^{k,0})^{-3} (\sigma_{ii}^k)^2 \right), & \text{if } i = j \text{ and } k \neq i, j \\ 0, & \text{if } k = i \text{ or } k = j, \end{cases} \quad (7.3.23)$$

where

$$\begin{aligned}
\bar{\mu}_{ij}^{k,0} &= \mathbf{a}^+ \cdot \mathbf{e} + \sum_{\substack{m=1 \\ m \neq i,j,k}}^M (a_m^- - a_m^+) \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-} \right)^{-1} + \sum_{m=i,j} (a_m^- - a_m^+) \\
(\sigma_{ij}^k)^2 &= \sum_{\substack{m=1 \\ m \neq i,j,k}}^M (a_m^- - a_m^+)^2 \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-} \right)^{-2},
\end{aligned}$$

and

$$K_{ij}^{k,1} = \begin{cases} \bar{\mathbf{g}}_i \bar{\mathbf{g}}_j \left((\bar{\mu}_{ij}^{k,1})^{-1} + (\bar{\mu}_{ij}^{k,1})^{-3} (\sigma_{ij}^k)^2 \right), & \text{if } i \neq j \text{ and } k \neq i, j \\ \bar{\mathbf{g}}_i \left((\bar{\mu}_{ii}^{k,1})^{-1} + (\bar{\mu}_{ii}^{k,1})^{-3} (\sigma_{ii}^k)^2 \right), & \text{if } i = j \text{ and } k \neq i, j \\ \bar{\mathbf{g}}_j \left((\bar{\mu}_{kj}^{k,1})^{-1} + (\bar{\mu}_{kj}^{k,1})^{-3} (\sigma_{kj}^k)^2 \right), & \text{if } k = i \\ \bar{\mathbf{g}}_i \left((\bar{\mu}_{ik}^{k,1})^{-1} + (\bar{\mu}_{ik}^{k,1})^{-3} (\sigma_{ik}^k)^2 \right), & \text{if } k = j, \end{cases} \quad (7.3.24)$$

where

$$\bar{\mu}_{ij}^{k,1} = \mathbf{a}^+ \cdot \mathbf{e} + \sum_{\substack{m=1 \\ m \neq i, j, k}}^M (a_m^- - a_m^+) \frac{a_m^+}{a_m^-} \left(1 + \frac{a_m^+}{a_m^-} \right)^{-1} + \sum_{m=i, j, k} (a_m^- - a_m^+).$$

Thus,

$$\sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \otimes \mathbf{H}_{\mathbf{y}} \nabla \hat{\pi}(\mathbf{y}) = \mathbf{C} \sum_{k=1}^M \sum_{\tilde{y}=0}^1 \mathbf{K}^{k, \tilde{y}} \mathbf{C}^T \nabla \hat{\pi}_k(\tilde{y}).$$

Finally,

$$\begin{aligned} \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{H}_{\mathbf{y}} \nabla \cdot \mathbf{H}_{\mathbf{y}} \hat{\pi}(\mathbf{y}) &= \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \mathbf{C} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \nabla \cdot \left(\mathbf{C} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \right) \hat{\pi}(\mathbf{y}) \\ &= \mathbf{C} \sum_{\mathbf{y} \in \{0,1\}^M} \frac{1}{a_0(\mathbf{y})} \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \otimes \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \hat{\pi}(\mathbf{y}) \nabla \cdot \mathbf{C} \\ &= \mathbf{C} \begin{bmatrix} \mathbf{K} & \mathbf{K} \\ \mathbf{K} & \mathbf{K} \end{bmatrix} \nabla \cdot \mathbf{C} \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbf{u} &= -\mathbf{C} \sum_{k=1}^M \sum_{\tilde{y}=0}^1 \begin{bmatrix} \mathbf{K}^{k, \tilde{y}} & \mathbf{K}^{k, \tilde{y}} \\ \mathbf{K}^{k, \tilde{y}} & \mathbf{K}^{k, \tilde{y}} \end{bmatrix} \mathbf{C}^T \nabla \hat{\pi}_k(\tilde{y}) - \mathbf{C} \begin{bmatrix} \mathbf{K} & \mathbf{K} \\ \mathbf{K} & \mathbf{K} \end{bmatrix} \nabla \cdot \mathbf{C} \\ &= -\mathbf{C} \sum_{k=1}^M \begin{bmatrix} \mathbf{K}^{k,0} - \mathbf{K}^{k,1} & \mathbf{K}^{k,0} - \mathbf{K}^{k,1} \\ \mathbf{K}^{k,0} - \mathbf{K}^{k,1} & \mathbf{K}^{k,0} - \mathbf{K}^{k,1} \end{bmatrix} \mathbf{C}^T \nabla \hat{\pi}_k(0) - \mathbf{C} \begin{bmatrix} \mathbf{K} & \mathbf{K} \\ \mathbf{K} & \mathbf{K} \end{bmatrix} \nabla \cdot \mathbf{C} \end{aligned}$$

since, using the formula for $\hat{\pi}$ in (7.3.3), we compute

$$\begin{aligned} \nabla \hat{\pi}(0) &= -\frac{a_k^- \nabla a_k^+ - a_k^+ \nabla a_k^-}{(a_k^-)^2} \left(1 + \frac{a_k^+}{a_k^-} \right)^{-2} \\ \nabla \hat{\pi}(1) &= -\nabla \hat{\pi}(0). \end{aligned} \quad (7.3.25)$$

7.3.4 Summary of the population scale formulation

To summarize, the population scale formulation of the cell migration model is the following. Let $\Omega \subset \mathbb{R}^2$ be the spatial domain of interest, and $\Omega_{cell} = \{\mathbf{x}_n \in \mathbb{R}^2 : \|\mathbf{x}_n\|^2 \leq 1\}$ be the domain, representing the cell (in the nondimensional form, recall Sections 2.1, 2.2.3). Let $\rho(t, \mathbf{X})$ denote the density of cells at time t , where $\mathbf{X} = (\mathbf{x}, \mathbf{x}_n) \in \Omega \times \Omega_{cell}$. Then, ρ satisfies

$$\frac{\partial}{\partial t} \rho(t, \mathbf{X}) - \nabla_{\mathbf{X}} \cdot (\mathbb{D}(\mathbf{X}) \nabla_{\mathbf{X}} \rho(t, \mathbf{X})) + \nabla_{\mathbf{X}} \cdot (\mathbf{u}(\mathbf{X}) \rho(t, \mathbf{X})) = 0, \quad (7.3.26)$$

subject to some initial condition

$$\rho(0, \mathbf{X}) = \hat{\rho}_0(\mathbf{X}), \quad \forall \mathbf{X} \in \Omega \times \Omega_{cell}$$

and the no-flux boundary condition

$$(\mathbb{D}(\mathbf{X}) \nabla_{\mathbf{X}} \rho(t, \mathbf{X}) + \mathbf{u}(\mathbf{X}) \rho(t, \mathbf{X})) \cdot \hat{\mathbf{n}} = 0, \quad \forall \mathbf{X} \in \partial(\Omega \times \Omega_{cell}),$$

where $\hat{\mathbf{n}}$ is an outward unit normal. The no-flux boundary condition has twofold physical significance. First, cells do not leave the spatial domain Ω . Second, it reflects the fact that in our cell scale model, the node \mathbf{x}_n remains in the cell (see Section 2.1 and in particular Corollary 2.1.2).

The coefficients in (7.3.26) are given by

$$\mathbb{D}(\mathbf{X}) = \mathbf{C}(\mathbf{x}_n) \begin{bmatrix} \mathbf{K}(\mathbf{X}) & \mathbf{K}(\mathbf{X}) \\ \mathbf{K}(\mathbf{X}) & \mathbf{K}(\mathbf{X}) \end{bmatrix} \mathbf{C}^T(\mathbf{x}_n),$$

where for convenience we flipped the sign in front of the term on the right in (7.3.22), and

$$\begin{aligned} \mathbf{u}(\mathbf{X}) = & - \mathbf{C}(\mathbf{x}_n) \sum_{k=1}^M \sum_{\tilde{y}=0}^1 \begin{bmatrix} \mathbf{K}^{k,\tilde{y}}(\mathbf{X}) & \mathbf{K}^{k,\tilde{y}}(\mathbf{X}) \\ \mathbf{K}^{k,\tilde{y}}(\mathbf{X}) & \mathbf{K}^{k,\tilde{y}}(\mathbf{X}) \end{bmatrix} \mathbf{C}^T(\mathbf{x}_n) \nabla \hat{\pi}_k(\tilde{y} | \mathbf{X}) \\ & - \mathbf{C}(\mathbf{x}_n) \begin{bmatrix} \mathbf{K}(\mathbf{X}) & \mathbf{K}(\mathbf{X}) \\ \mathbf{K}(\mathbf{X}) & \mathbf{K}(\mathbf{X}) \end{bmatrix} \nabla \cdot \mathbf{C}(\mathbf{x}_n) \end{aligned}$$

with \mathbf{C} given by (7.3.18), \mathbf{K} given by (7.3.21), $\mathbf{K}^{k,\tilde{y}}$ given by (7.3.23) and (7.3.24), and $\hat{\pi}_k$ given by (7.3.3). Note that since \mathbf{K} is symmetric and has strictly positive entries (for $a_i^+ > 0$), then \mathbb{D} is positive definite.

Note that the population formulation presented here is not truly ‘‘macroscale’’, as it still involves the \mathbf{x}_n variable, which is not macroscopic, as it closely relates to cellular scale dynamics (e.g. extension of SFs). The variable \mathbf{x}_n can be seen as a supplementary structure variable, so that solving the population scale problem (7.3.26) provides the information about the population density not only with respect to time and space, but also with respect to cellular configuration of SFs. The latter could inform about the evolution of cell density with respect to intracellular stress distribution.

Chapter 8

Conclusion and Outlook

Cell migration is a multiscale phenomenon, where dynamic interactions of multiple, possibly heterogeneous, agents on each scale give rise to a dynamic system on a higher level. Several such systems can themselves interact with each other, yielding yet another organization on a higher level. For example, interactions of focal adhesions with stress fibers give rise to cell migration. In turn, multiple migrating cells can interact and develop group behavior. Although the whole is more than the sum of its parts, understanding each part and how they combine is essential to understanding the whole. In this work, we constructed a multiscale model of migrating cells, where the constituent parts at the subcell, cell, and cell group levels were combined to obtain a mathematically consistent description of cell migration at an individual, collective, and population scales (Figure 8.1).

To describe single cell migration, we constructed a stochastic model using a minimal representation of cellular structures, essential for crawling, such as stress fibers and focal adhesions (Chapter 2). Using this representation, and observing that FA assembly and disassembly events of the migration cycle lead to different migratory outcomes, we obtained the equations describing deterministic cell motion between the random occurrence of FA events. After introducing the probabilities rates of FA binding and unbinding, we obtained the remaining necessary objects to define a piecewise deterministic Markov process: the distribution of interarrival times and of the next FA event. Note that the forms of these distributions have been derived, rather than simply postulated.

To complete the description of single cell migration, we discussed the interactions of the subcell level constituents of FAs and SFs (Chapter 3). We showed how their mediation by RhoGTPases can be taken into account. For example, by simplifying the cross-bridge model in [3], we found a relation between local RhoA levels and tension applied by myosin motors. We also demonstrated that one could use the previously developed model of RhoGTPases signaling pathways by [50] to relate local Rac1 activity with FA association rate. Altogether, we provided a model of mechanical (force dependence on adhesion binding/unbinding) and chemical interaction (through RhoGTPase mediation) of FAs and SFs.

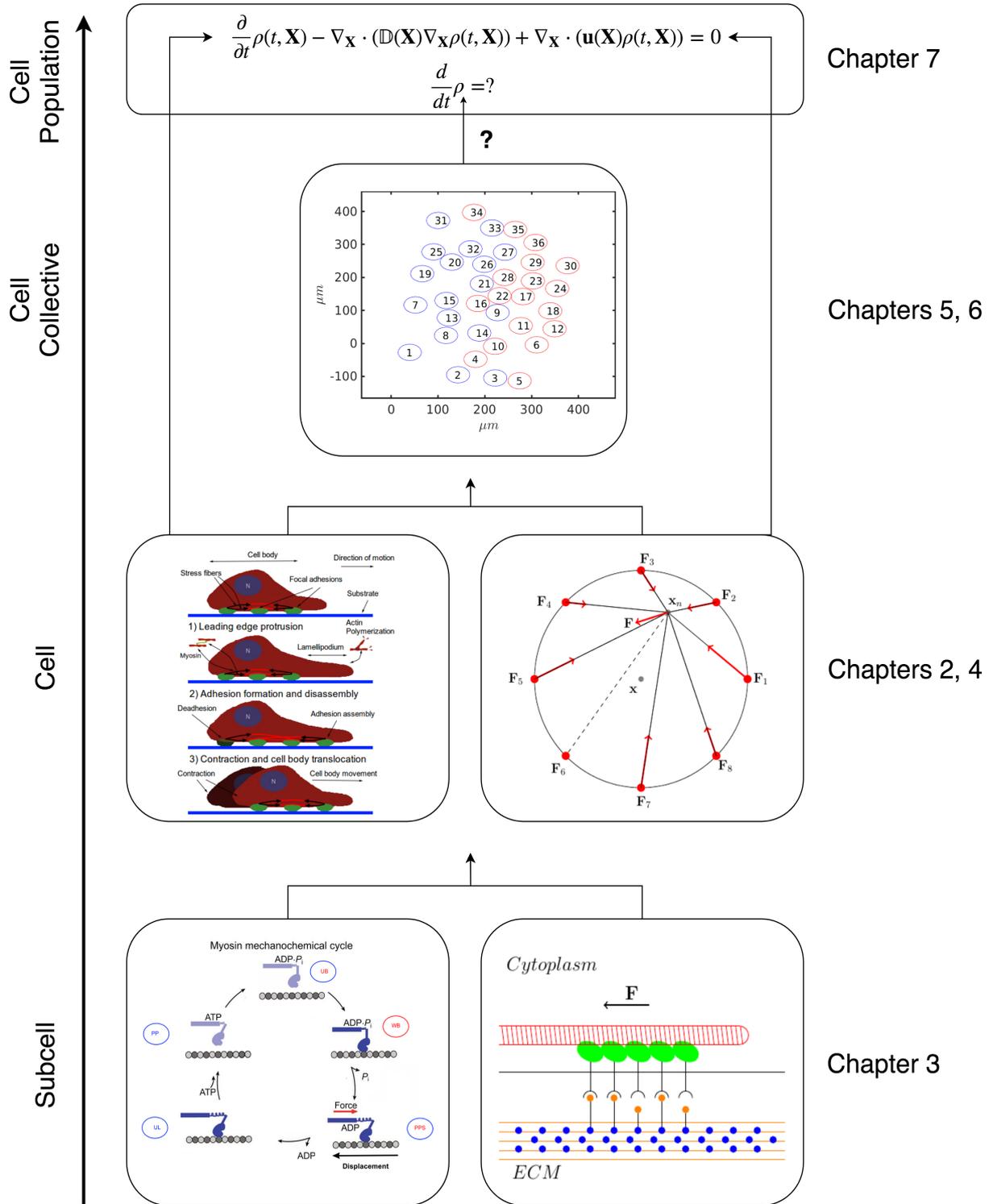


Figure 8.1: Multiscale description of cell migration. Synopsis of this thesis.

Having specified the coupling between SFs and FAs, as well as between the cellular environment and FAs, we performed numerical simulations (Chapter 4). We showed that our model is able to reproduce experimental observations, such as: superdiffusive scaling of the mean-squared displacement [32], [67], [70] (Figure 4.4); biased motility in the presence of external cue (Figure 4.6); contact guidance [95] (Figure 4.13). In these cases, the obtained results followed solely due to asymmetric, dynamic instability of FAs in direct response to environmental stimuli. Specifically, it is only the biased FA assembly rate that drives biased cell motility along the cue gradient or the fiber tracts (Figures 4.7 and 4.13 (d-i)). That is, preferred velocities were not imposed or chosen in any way, but simply followed as a consequence of front-rear polarity, as the cell front is characterized by preferential FA binding and the rear by unbinding.

Another characteristic of directed migration is the asymmetric contraction of actomyosin bundles. By increasing the force generation of myosin motors in the prospective rear, we obtained directed movement in the opposite direction (Figure 4.17). Here asymmetric FA dynamics (and so front-rear polarity) was also obtained, but as a consequence of locally induced contractile activity, consistent with [129].

Our simulation results in various settings suggest that the cell speeds follow a gamma distribution (Figures 4.2 (g-i), 4.4 (g-i), 4.7 (a-c)). Furthermore, the number of adhesion sites seems to be a determinant of the gamma distribution, as its parameters are similar under different settings and given number of FAs. These results suggest that cell speeds are independent of biased FA formation, i.e. the bias only alters the directionality and not the speed. It is also interesting to see a correlation between the number of adhesion sites and diffusivity (Table 4.2), as well as average speed (Tables 4.2-4.5). Note that faster and diffusive amoeboid movement is characterized by an increased number of weaker adhesions with high turnover and contractility [90]. Thus, the aforementioned correlation is also consistent with experimental observations. We note that our model is not fit to take into account motility strongly relying on cell shape control, which is required, for example, in highly mobile cells. However, the simulations reproduce migration along fiber tracts, where cell reshaping takes place [95]. Our results suggest, then, that adhesion along the tracts is sufficient to produce such migration patterns.

Although the model of the internal contractile machinery driving cell locomotion and cytoskeletal remodeling is simple, the resulting numerical simulations explain several experimental observations. Moreover, the cyclical nature of cell motility is captured with our piecewise deterministic model. Note also that the FA dynamics in our model is independent of a particular cellwise contractility mechanism: only local mechanochemical interactions are relevant.

Following our account of cell migration on the individual scale, we considered the scale of cell collectives (Chapters 5, 6). We extended our model to account for contact inhibition of locomotion arising as a result of cell-cell collisions. Here, the cells, exhibiting CIL response, alter cell-substrate adhesions dynamics and SF contractility following contact with another cell. Mathematically, the model is described by a piecewise deterministic process, whereby collisions occur when some deterministic components (cell-cell distances) reach a corresponding value, and cell motility itself emerges due to mechanochemically

mediated stochastic adhesion dynamics. Consequently, the outcome of a collision is also determined stochastically, as reported in [30], [68], [102].

Mimicking the experimental setup in [68], we simulated binary collisions between cells migrating while confined to a 1D lane. In this setting, we did not invoke the volume exclusion principle, and showed that a CIL response can be explained solely due to increased cell-substrate adhesion away from the collision site and increased actomyosin contractility in its vicinity. Although cell overlaps occur, we see that by strengthening the CIL response we can reduce its occurrence (Figure 5.5b). Our results also show that an external cue can modulate CIL response, in line with [68]. Specifically, typical CIL response (Figure 5.6) can be overridden by chemotaxis (Figure 5.8) if post collision velocity is not aligned with the chemotactic gradient. Regardless of an external signal presence, our simulations suggest that transient perturbations in cell motility lead to persistent alterations of migration dynamics (Figure 5.7, 5.8c,d).

In an unconfined setting, we simulated the effects of homo- and heterotypic CIL. We found that homotypic CIL leads to increased cell displacement of chemotaxing and non-chemotaxing cells (Figure 6.3h). We also found that the spatial configuration of heterogeneous cells can have an impact on how heterotypic CIL affects migration of cells. In a disordered population it can facilitate the dispersion of randomly migrating cells (Figure 6.4), while letting directed migration be unhindered. When separated into groups, our simulations suggest that directed movement can be induced in non-chemotaxing cells (Figure 6.5), as reported in [117]. Altogether, simulations in the unconfined setting suggest that homotypic, but not heterotypic CIL, is advantageous for dispersive and invasive migration of cells. It has been speculated that such CIL behavior is responsible for the initial spread of cancer cells [76], [112].

To simulate the models on the cell and the cell collective scales, we developed appropriate numerical methods (Sections 4.1, 6.3.1). These methods reflect an important distinction between the two PDMPs describing cell migration in the absence and presence of interactions. In the latter, more general case, jumps can also occur when the deterministic component attains a certain value. Consequently, the method in Section 6.3.1 can be applied to a general PDMP¹. However, the special case represents a large enough class of problems for the method presented in Sections 4.1 to be of practical use.

We were also able to deduce a population scale model of non-interacting migrating cells from the single cell scale model (Chapter 7). While we employed a well-known procedure to obtain an equation governing the law of the cell motility process, it was necessary to utilize an order reduction method to reduce the dimensionality of the problem. Analogous to the approaches in [33], [48], [88], [89], we used perturbation expansion and time rescaling to obtain a low order approximation of the problem. Such approximation was possible because our constructive approach to modeling single cell migration allowed us to exploit the detailed knowledge about the dynamics of FAs. Consequently, the population scale equation (7.3.26) still contains some details about subcell and cell dynamics. This

¹With slight modification of a routine used to determine whether the jump occurs when the process attains a certain value.

equation also has a supplementary structure variable, which influences spatiotemporal evolution of the population. Thus, one can view (7.3.26) as a structured population model (see [8] and references therein for numerous such models). However, it is not often that such class of models are obtained through derivation from elementary interactions. Thus, our constructive approach and perturbation expansion provides an example of a derived structured population model.

There are numerous shortcomings at each scale of our model, which we plan to improve in the future work. For example, we could include the role of RhoGTPase signaling pathways in regulating FA association rate and myosin motor force generation (as discussed in Sections 3.2, 3.3). Because of its essential role, accounting for the Rac1 signaling would also enable us to explore an essential feature of the migration cycle - actin polymerization and protrusion formation. Thus, due to the role of RhoGTPases in regulating adhesion dynamics, contractility, and actin polymerization [94], [96], [97], the modifications at the subcell level will allow us to consider a more biologically consistent description at the cell level. In our future work, we also plan to address probably the most significant shortcoming of our model at the cell scale - disregard of cellular shape changes. One possible avenue of addressing this issue is through a phase field description of our cell, similar to the works in [16], [72], [86], [99], [106]. Another possibility is adopting the vertex-based model of a cell [39], as was done in [78] in the context of cell motility. a more complex contractility apparatus can be described via active cable network model [47] and a more detailed account of mechanical forces (e.g. protrusions due to actin polymerization) can be done as in [24]. Together with models of RhoGTPases signaling pathways [49], [50], [82], the most significant drawbacks of our approach (including rigid rotation of the SF structure) can be overcome. The presented framework of piecewise deterministic motility process can also be extended to three-dimensional setting, as neither the event interarrival time distribution, nor the transition measure rely on the particular features of migration in a plane. Overall, our plans for future work at the cell scale involve adding layers of complexity on top of the minimal model presented here.

Guided by the study in [30], in our model of CIL at the scale of cell collectives we assumed that the CIL response between colliding cells is transient and mutually independent. That is, immediately following the collision, cell dynamics and FA event probabilities in both cells are decoupled. However, there is evidence that a mechanical coupling is established prior to repulsion [113]. Moreover, some cells exhibiting homotypic CIL tend to disperse and reaggregate into small clusters, which increases their chemotactic efficiency [117]. Thus, addressing mechanical coupling by including cell-cell adhesions represents one of the avenues for future work, whereby collective migration could be further investigated. Including RhoGTPase signaling pathways constitutes another major improvement of our CIL model, since these are affected in a contact dependent manner during CIL [98]. Therefore, addressing the above mentioned shortcomings of the model at the subcellular scale will also help us to refine our model at the higher scales.

We also leave an in-depth investigation of the population scale model for future work. In particular, performing numerical simulations will allow us to gain an insight on how the interactions between cell-substrate adhesions and actomyosin contractility translate into

population migration. Although the numerical solution of (7.3.26) might be challenging due to its dimensionality, the method in [37] provides a possibility to overcome it. Another major issue left to be addressed is how to obtain a population scale model of interacting cells directly from the scale of cell collectives. While we could apply the procedure in Sections 7.1, 7.2 to the PDMP of interacting cells in 6.2, it is not clear how to proceed further in order to obtain a population scale description. As in our case otherwise independent cells interact through collisions only, a potential route for investigating the issue is by drawing analogies with the kinetic theory, similar to the studies in [20], [21].

Despite of its shortcomings, our minimal model can explain various experimental observations about cell motility at the scales of individuals and groups. Together with our population scale model, in this work we achieved a genuine multiscale description of cell migration, where the information about the lower scales is retained and carried over to higher scales in a mathematically consistent manner.

Appendix A

Equations of cell motion

In our model, using common, the lab's reference frame will yield the same governing relations, because the involved forces are determined by relative position of cellular structures. Below, we show why this is the case and provide a more detailed explanation regarding the equation of motions for $\mathbf{x}, \mathbf{x}_n, \theta$ presented in Section 2.2.2. The material here is based on the work [120] by the author.

Let $\mathbf{x}'_n = \mathbf{x} + \mathbf{x}_n$ and $\mathbf{x}'_i = \mathbf{x} + \mathbf{x}_i$, where primes indicate the corresponding variables in the lab's reference frame (recall that \mathbf{x}_i is the position of the i^{th} FA in the cell's reference frame). Then, in this frame, the length of the i^{th} SF L'_i and the unit vector \mathbf{e}'_i along it are given by

$$\begin{aligned} L'_i &= \|\mathbf{x}'_n - \mathbf{x}'_i\| = \|\mathbf{x}_n - \mathbf{x}_i\| = L_i \\ \mathbf{e}'_i &= \frac{\mathbf{x}'_n - \mathbf{x}'_i}{L'_i} = \frac{\mathbf{x}_n - \mathbf{x}_i}{L_i} = \mathbf{e}_i, \end{aligned}$$

respectively. Thus, $\mathbf{F}'_i(\mathbf{x}'_n, \theta) = \mathbf{F}_i(\mathbf{x}_n, \theta)$, where \mathbf{F}'_i is the force applied by the i^{th} SF at the i^{th} FA. Note that the force at \mathbf{x}'_n (or \mathbf{x}_n) due to the i^{th} SF is $-\mathbf{F}'_i(\mathbf{x}'_n, \theta)$ (or $-\mathbf{F}_i(\mathbf{x}_n, \theta)$) by action-reaction principle. Therefore, net force \mathbf{F}' at \mathbf{x}'_n is $\mathbf{F}'(\mathbf{x}'_n, \theta, \mathbf{Y}) = -\sum_{i=1}^M Y_i \mathbf{F}'_i = -\sum_{i=1}^M Y_i \mathbf{F}_i = \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y})$. Neglecting inertia, we have

$$\beta_{cell} \dot{\mathbf{x}}_n = \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) = \mathbf{F}'(\mathbf{x}'_n, \theta, \mathbf{Y}).$$

Now, let us examine the equations of motion after FA unbinding, stated in Section 2.2.2, but in the lab's reference frame. In this frame, the radial unit vector $\hat{\mathbf{r}}'(\mathbf{x}'_n)$ from the cell center \mathbf{x} is given by (see Figure 2.7 in the manuscript for illustration)

$$\hat{\mathbf{r}}'(\mathbf{x}'_n) = \frac{\mathbf{x}'_n - \mathbf{x}}{\|\mathbf{x}'_n - \mathbf{x}\|} = \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|} = \hat{\mathbf{r}}(\mathbf{x}_n).$$

Analogously, the tangential unit vector $\hat{\boldsymbol{\varphi}}'(\mathbf{x}'_n)$ is given by

$$\hat{\boldsymbol{\varphi}}'(\mathbf{x}'_n) = \left(-\frac{x'_{n,2} - x_2}{\|\mathbf{x}'_n - \mathbf{x}\|}, \frac{x'_{n,1} - x_1}{\|\mathbf{x}'_n - \mathbf{x}\|} \right)^T = \left(-\frac{x_{n,2}}{\|\mathbf{x}_n\|}, \frac{x_{n,1}}{\|\mathbf{x}_n\|} \right)^T = \hat{\boldsymbol{\varphi}}(\mathbf{x}_n).$$

Note that the tangential component F'_φ of the force \mathbf{F}' at \mathbf{x}'_n induces rotational motion, while the radial component F'_r of the force \mathbf{F}' at \mathbf{x}'_n induces translational motion. These components are given by:

$$\begin{aligned} F'_\varphi &= \mathbf{F}'(\mathbf{x}'_n, \theta, \mathbf{Y}) \cdot \hat{\boldsymbol{\varphi}}'(\mathbf{x}'_n) = \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\boldsymbol{\varphi}}(\mathbf{x}_n) = F_\varphi \\ F'_r &= \mathbf{F}'(\mathbf{x}'_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}'(\mathbf{x}'_n) = \mathbf{F}(\mathbf{x}_n, \theta, \mathbf{Y}) \cdot \hat{\mathbf{r}}(\mathbf{x}_n) = F_r. \end{aligned}$$

Neglecting rotational inertia, we then have

$$\begin{aligned} \beta_{rot} \dot{\theta} &= \|\mathbf{x}'_n - \mathbf{x}\| F'_\varphi(\mathbf{x}'_n, \theta, \mathbf{Y}) \\ &= \|\mathbf{x}_n\| F_\varphi(\mathbf{x}_n, \theta, \mathbf{Y}) \end{aligned}$$

where the right hand side in the first (second) line is the torque due to tangential component of the force \mathbf{F}' (\mathbf{F}) at \mathbf{x}'_n (\mathbf{x}_n). Because of the low Reynolds number, we also have

$$\begin{aligned} \beta_{ECM} \dot{\mathbf{x}} &= F'_r(\mathbf{x}'_n, \theta, \mathbf{Y}) \hat{\mathbf{r}}'(\mathbf{x}'_n) \\ &= F_r(\mathbf{x}_n, \theta, \mathbf{Y}) \hat{\mathbf{r}}(\mathbf{x}_n), \end{aligned}$$

due to the translational motion induced by the radial component of the force \mathbf{F}' at \mathbf{x}'_n .

In the common reference frame, the following system of ODEs holds after unbinding (using the definition of \mathbf{x}'_n):

$$\begin{aligned} \beta_{ECM} \dot{\mathbf{x}} &= F'_r(\mathbf{x}'_n, \theta, \mathbf{Y}) \hat{\mathbf{r}}'(\mathbf{x}'_n) \\ \beta_{cell} \dot{\mathbf{x}}'_n &= \beta_{cell} \dot{\mathbf{x}} + \mathbf{F}'(\mathbf{x}'_n, \theta, \mathbf{Y}) \\ \beta_{rot} \dot{\theta} &= \|\mathbf{x}'_n - \mathbf{x}\| F'_\varphi(\mathbf{x}'_n, \theta, \mathbf{Y}), \end{aligned}$$

which is equivalent to (2.2.3). Using the common reference frame becomes even less convenient when we formulate and analyze our stochastic process of cell motility. Moreover, our approach in the main text does not contradict the formulation with the single reference frame, and is equivalent to it.

Appendix B

Data Analysis Methods

The material here is based on the work [120] by the author.

Given that the time interval $[0, t_{end}]$ is divided into n_{time} subintervals of equal length Δt and given the positions $\mathbf{x}^i(t_j)$ of cell i at the time points $t_j := j\Delta t$, $j = 0, \dots, n_{time}$, the mean-squared displacement $msd_i(t_j)$ of cell $i \in \{1, \dots, n_{cells}\}$ over a time interval of length t_j is given by:

$$msd_i(t_j) := \frac{1}{n_{time} - j} \sum_{k=1}^{n_{time}-j} \|\mathbf{x}^i(t_{j+k}) - \mathbf{x}^i(t_k)\|^2, \quad (\text{B.0.1})$$

where $j = 1, \dots, n_{time} - 1$ and n_{cells} is the total number of cells. Then, the mean-squared displacement $msd(t)$ of an ensemble of cells over time interval of length t_j is defined by:

$$msd(t) := \frac{1}{n_{cells}} \sum_{i=1}^{n_{cells}} msd_i(t_j). \quad (\text{B.0.2})$$

Remark. In general, the (time-averaged) mean-squared displacement $\langle d^2(t, T) \rangle$ of a particle trajectory $\mathbf{x}(t)$ at the time t , time endpoint T is formally defined as:

$$\langle d^2(t, T) \rangle = \frac{1}{T-t} \int_0^{T-t} \|\mathbf{x}(s+t) - \mathbf{x}(t)\|^2 ds. \quad (\text{B.0.3})$$

For an ergodic process, we have:

$$\lim_{T \rightarrow \infty} \langle d^2(t, T) \rangle = \langle \mathbf{x}^2(t) \rangle,$$

where $\langle \mathbf{x}^2(t) \rangle$ is formally defined as:

$$\langle \mathbf{x}^2(t) \rangle = \int \mathbf{x}^2 P_t(d\mathbf{x}),$$

and $P_t(d\mathbf{x})$ is the probability measure of the underlying stochastic process at time t . That is, for an ergodic process, and for sufficiently long times, the time average equals the

phase space average. However, our cell motility process needs not be ergodic and hence, using a quadrature to evaluate the integral in equation (B.0.3), we obtain time average displacement in equation (B.0.1). To smooth out trajectory-to-trajectory fluctuations, we then average the displacements over all trajectories in equation (B.0.2).

For a diffusive motion we expect that $msd(t) \sim t^{\beta(t)}$ with $\beta(t) \approx 1$, while for a ballistic motion $\beta(t) \approx 2$. Since $msd(0) = 0$, we can estimate the exponent $\beta(t)$ for $t \in [\Delta t, t_{end} - \Delta t]$ from the simulated data as:

$$\beta(t) = \frac{d \ln msd(t)}{d \ln t}.$$

Although averaging reduces fluctuations, it does not eliminate them completely. Thus, computing the derivative above will yield a result that may oscillate wildly, which we want to avoid. Then, in order to investigate how β varies with time, we define the time average $\beta_{av}(t)$ over the interval $[\Delta t, t]$ as:

$$\beta_{av}(t) := \frac{1}{t - \Delta t} \int_{\Delta t}^t \beta(s) ds = \frac{1}{t - \Delta t} \left(s \ln msd(s) \Big|_{\Delta t}^t - \int_{\Delta t}^t \ln msd(s) ds \right),$$

where $t \in [2\Delta t, t_{end} - \Delta t]$, and we used integration by substitution and by parts. Then, $\bar{\beta} := \beta_{av}(t_{end} - \Delta t)$ estimates the time scaling of msd over the whole time interval. To assess how well $\bar{\beta}$ reflects the scaling of msd , we define the following function $\widehat{msd}(t) := \beta_0 t^{\bar{\beta}}$, where β_0 is found by minimizing the square error:

$$\begin{aligned} \min_{\beta_0} \frac{1}{2} \sum_{j=1}^{n_{time}-1} \left(\beta_0 t_j^{\bar{\beta}} - msd(t_j) \right)^2 &\Rightarrow \\ \beta_0 &= \frac{\sum_{j=1}^{n_{time}-1} msd(t_j) t_j^{\bar{\beta}}}{\sum_{j=1}^{n_{time}-1} t_j^{2\bar{\beta}}}. \end{aligned}$$

Letting $\bar{\beta} = \beta_{av}(t - \Delta t)$ to assess time scaling of $msd \sim t^{\beta(t)}$ is more accurate than the standard methods used for Brownian motion, since it also takes into account time dependence of the exponent. Also, our stochastic model has no Gaussian component. We refer to, Section 4.3 for comparisons between msd and \widehat{msd} , which show that the former well approximates the latter.

Note that because binding events can occur, a cell needs not have moved between the two time points t_j and t_{j+1} . Thus, the speed between the consecutive time points may be zero for many time points, which would give an inaccurate statistical assessment of cell speeds. In order to estimate the speeds of a cell i we use the following procedure:

First, we find l_i , given by:

$$l_i := \min \{ l \in \mathbb{N} : \mathbf{x}^i(t_{l+k}) \neq \mathbf{x}^i(t_k), 0 \leq k < n_{time}, l+k \leq n_{time} \}.$$

Then we find the set of speeds S_i as:

$$S_i := \left\{ s \in \mathbb{R}^+ : s = \frac{\|\mathbf{x}^i(t_{(k+1)l_i}) - \mathbf{x}^i(t_{kl_i})\|}{l_i \Delta t}, k \in \mathbb{N}, (k+1)l_i \leq n_{time} \right\}.$$

This simply means that to compute speeds we only use a (minimal) time interval, such that a cell i is guaranteed to change its position. The total set of speeds for n_{cells} is $S := \cup_{i=1}^{n_{cells}} S_i$. The average speed s_{av} is then defined as an arithmetic average:

$$s_{av} := \frac{1}{|S|} \sum_{s \in S} s.$$

The set of normalized velocities V_i (or, alternatively, displacements) of cell i is given by:

$$V_i := \left\{ \mathbf{v} \in \mathbb{R}^2 : \mathbf{v} = \frac{\mathbf{x}^i(t_{(k+1)l_i}) - \mathbf{x}^i(t_{kl_i})}{\|\mathbf{x}^i(t_{(k+1)l_i}) - \mathbf{x}^i(t_{kl_i})\|}, k \in \mathbb{N}, (k+1)l_i \leq n_{time} \right\},$$

and the total set of normalized velocities V is given by $V := \cup_{i=1}^{n_{cells}} V_i$.

The directionality ratio $r_i(t_j)$ of cell i over a time interval of length t_j is given by:

$$r_i(t_j) := \frac{\sum_{k=1}^j \|\mathbf{x}^i(t_k) - \mathbf{x}^i(t_{k-1})\|}{\|\mathbf{x}^i(t_j) - \mathbf{x}^i(t_0)\|}.$$

The population and the time averages of the directionality ratio are given by, respectively:

$$r(t_j) = \frac{1}{n_{cell}} \sum_{i=1}^{n_{cell}} r_i(t_j)$$

$$\bar{r} = \frac{1}{n_{time}} \sum_{j=1}^{n_{time}} r(t_j).$$

Velocity autocorrelation $v_{ac}^i(t_j)$ of cell i over a time interval of length t_j is given by:

$$v_{ac}^i(t_j) := \frac{1}{n_{time} - j} \sum_{k=1}^{n_{time}-j} \frac{\mathbf{v}^i(t_{j+k}) \cdot \mathbf{v}^i(t_k)}{\|\mathbf{v}^i(t_{j+k})\| \|\mathbf{v}^i(t_k)\|},$$

where $j = 1, \dots, n_{time} - 1$ and $\mathbf{v}^i(t_k) = (\mathbf{x}^i(t_k) - \mathbf{x}^i(t_{k-1})) / \Delta t$. Velocity autocorrelation of the population $v_{ac}(t)$ is simply the arithmetic average of each cell's velocity autocorrelation. To compute v_{ac} we used the time step of $12min$, whereas for all other quantities involving time dependence (e.g. *msd*) we used the time step of $0.12min$.

We define the guidance parameter $G \in [0, 1]$ similarly as in [95]:

$$G := \frac{1}{|\Theta|} \sum_{\theta \in \Theta} g(\theta),$$

where $\Theta := \cup_{i=1}^{n_{cells}} \Theta_i$. The set Θ_i of angles between a displacement vector of cell i and the ECM fibers is defined by

$$\Theta_i := \left\{ \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] : \theta = \arcsin \left(\frac{x_2^i(t_{(k+1)l_i}) - x_2^i(t_{kl_i})}{\|\mathbf{x}^i(t_{(k+1)l_i}) - \mathbf{x}^i(t_{kl_i})\|} \right), k \in \mathbb{N}, (k+1)l_i \leq n_{time} \right\},$$

where x_2^i is the y -component of \mathbf{x}^i . The function $g : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow \{0, 1\}$ is given by

$$g(\theta) = \begin{cases} 1, & \text{if } |\theta| \leq \pi/4 \\ 0, & \text{else} \end{cases}.$$

Thus, G increases when the displacements are aligned with the horizontal axis.

Bibliography

- [1] M. Abercrombie. The croonian lecture, 1978 - the crawling movement of metazoan cells. *Proceedings of the Royal Society of London B: Biological Sciences*, 207(1167):129–147, 1980.
- [2] M. Abercrombie and J. E. Heaysman. Observations on the social behaviour of cells in tissue culture: II. monolayering of fibroblasts. *Experimental Cell Research*, 6(2):293 – 306, 1954.
- [3] P. J. Albert, T. Erdmann, and U. S. Schwarz. Stochastic dynamics and mechanosensitivity of myosin II minifilaments. *New Journal of Physics*, 16(9):093019, 2014.
- [4] A. Y. Alexandrova, K. Arnold, S. Schaub, J. M. Vasiliev, J.-J. Meister, A. D. Bershadsky, and A. B. Verkhovsky. Comparative dynamics of retrograde actin flow and focal adhesions: formation of nascent adhesions triggers transition from fast to slow flow. *PLoS ONE*, 3(9):e3234, 2008.
- [5] Y. Aratyn-Schaus, P. W. Oakes, and M. L. Gardel. Dynamic and structural signatures of lamellar actomyosin force generation. *Molecular Biology of the Cell*, 22(8):1330–1339, 2011.
- [6] D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler, and D. Heinrich. Temporal analysis of active and passive transport in living cells. *Physical Review Letters*, 101:248103, Dec 2008.
- [7] C. Arriemerlou and T. Meyer. A local coupling model and compass parameter for eukaryotic chemotaxis. *Developmental Cell*, 8(2):215–227, 2005.
- [8] P. Auger, P. Magal, and S. Ruan. *Structured population models in biology and epidemiology*, volume 1936. Springer, 2008.
- [9] N. Q. Balaban, U. S. Schwarz, D. Riveline, P. Goichberg, G. Tzur, I. Sabanay, D. Mahalu, S. Safran, A. Bershadsky, L. Addadi, et al. Force and focal adhesion assembly: a close relationship studied using elastic micropatterned substrates. *Nature Cell Biology*, 3(5):466–472, 2001.
- [10] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33(4):24/1–24/27, 2007.

- [11] F. Bartumeus, J. Catalan, U. L. Fulco, M. L. Lyra, and G. M. Viswanathan. Optimizing the encounter rate in biological interactions: Lévy versus Brownian strategies. *Physical Review Letters*, 88:097901, Feb 2002.
- [12] G. I. Bell. Models for the specific adhesion of cells to cells. *Science*, 200(4342):618–627, 1978.
- [13] G. B. Blanchard, A. G. Fletcher, and L. J. Schumacher. The devil is in the mesoscale: Mechanical and behavioural heterogeneity in collective cell movement. In *Seminars in Cell & Developmental Biology*. Elsevier, 2018.
- [14] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet. The bouncy particle sampler: A non-reversible rejection-free Markov chain Monte Carlo method. *ArXiv e-prints*, Oct. 2015.
- [15] K. Burridge and M. Chrzanowska-Wodnicka. Focal adhesions, contractility, and signaling. *Annual Review of Cell and Developmental Biology*, 12(1):463–519, 1996.
- [16] B. A. Camley, Y. Zhao, B. Li, H. Levine, and W.-J. Rappel. Crawling and turning in a minimal reaction-diffusion cell motility model: coupling cell shape and biochemistry. *Physical Review E*, 95:012401, Jan 2017.
- [17] C. Carmona-Fontaine, H. K. Matthews, S. Kuriyama, M. Moreno, G. A. Dunn, M. Parsons, C. D. Stern, and R. Mayor. Contact inhibition of locomotion in vivo controls neural crest directional migration. *Nature*, 456(7224):957, 2008.
- [18] C. L. Chaffer and R. A. Weinberg. A perspective on cancer cell metastasis. *Science*, 331(6024):1559–1564, 2011.
- [19] C. E. Chan and D. J. Odde. Traction dynamics of filopodia on compliant substrates. *Science*, 322(5908):1687–1691, 2008.
- [20] A. Chauviere, T. Hillen, and L. Preziosi. Modeling cell movement in anisotropic and heterogeneous network tissues. *Networks and Heterogeneous Media*, 2(2):333, 2007.
- [21] A. Chauviere, T. Hillen, and L. Preziosi. Modeling the motion of a cell population in the extracellular matrix. *Discrete and Continuous Dynamical Systems*, pages 250–59, 2007.
- [22] W. T. Chen. Mechanism of retraction of the trailing edge during fibroblast movement. *Journal of Cell Biology*, 90(1):187–200, 1981.
- [23] C. K. Choi, M. Vicente-Manzanares, J. Zareno, L. A. Whitmore, A. Mogilner, and A. R. Horwitz. Actin and α -actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nature Cell Biology*, 10(9):1039–1050, 2008.

- [24] C. A. Copos, S. Walcott, J. C. del Álamo, E. Bastounis, A. Mogilner, and R. D. Guy. Mechanosensitive adhesion explains stepping motility in amoeboid cells. *Biophysical Journal*, 112(12):2672 – 2682, 2017.
- [25] L. P. Cramer. Forming the cell rear first: breaking cell symmetry to trigger directed cell migration. *Nature Cell Biology*, 12:628 – 632, 2010.
- [26] J. R. Davis, C.-Y. Huang, J. Zanet, S. Harrison, E. Rosten, S. Cox, D. Y. Soong, G. A. Dunn, and B. M. Stramer. Emergence of embryonic pattern through contact inhibition of locomotion. *Development*, 139(24):4555–4560, 2012.
- [27] M. H. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(3):353–388, 1984.
- [28] M. H. A. Davis. *Markov models and optimization*. Chapman and Hall, 1993.
- [29] S. Deguchi, T. Ohashi, and M. Sato. Tensile properties of single stress fiber isolated from cultured vascular smooth muscle cells. *Journal of Biomechanics*, 39(14):2603–2610, 2006.
- [30] R. A. Desai, S. B. Gopal, S. Chen, and C. S. Chen. Contact inhibition of locomotion probabilities drive solitary versus collective cell migration. *Journal of The Royal Society Interface*, 10(88), 2013.
- [31] R. B. Dickinson and R. T. Tranquillo. A stochastic model for adhesion-mediated cell random motility and haptotaxis. *Journal of Mathematical Biology*, 31(6):563–600, 1993.
- [32] P. Dieterich, R. Klages, R. Preuss, and A. Schwab. Anomalous dynamics of cell migration. *Proceedings of the National Academy of Sciences*, 105(2):459–463, 2008.
- [33] C. Engwer, T. Hillen, M. Knappitsch, and C. Surulescu. Glioma follow white matter tracts: a multiscale DTI-based model. *Journal of Mathematical Biology*, 71(3):551–582, 2015.
- [34] T. Erdmann and U. S. Schwarz. Stability of adhesion clusters under constant force. *Physical Review Letters*, 92(4):108102, 2004.
- [35] C. L. Faustino, L. R. da Silva, M. G. E. da Luz, E. P. Raposo, and G. M. Viswanathan. Search dynamics at the edge of extinction: Anomalous diffusion as a critical survival state. *EPL (Europhysics Letters)*, 77(3):30002, 2007.
- [36] J. Feng, M. Ito, Y. Kureishi, K. Ichikawa, M. Amano, N. Isaka, K. Okawa, A. Iwamatsu, K. Kaibuchi, D. J. Hartshorne, and T. Nakano. Rho-associated kinase of chicken gizzard smooth muscle. *Journal of Biological Chemistry*, 274(6):3744–3752, 1999.

- [37] F. Filbet and C.-W. Shu. Discontinuous Galerkin methods for a kinetic model of self-organized dynamics. *Mathematical Models and Methods in Applied Sciences*, 28(06):1171–1197, 2018.
- [38] J. T. Finer, R. M. Simmons, and J. A. Spudich. Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature*, 368(6467):113–119, 1994.
- [39] A. Fletcher, M. Osterfield, R. Baker, and S. Shvartsman. Vertex models of epithelial morphogenesis. *Biophysical Journal*, 106(11):2291 – 2304, 2014.
- [40] P. Friedl, P. Noble, E. Shields, and K. Zänker. Locomotor phenotypes of unstimulated CD45RA^{high} and CD45RO^{high} CD4⁺ and CD8⁺ lymphocytes in three-dimensional collagen lattices. *Immunology*, 82(4):617, 1994.
- [41] M. L. Gardel, I. C. Schneider, Y. Aratyn-Schaus, and C. M. Waterman. Mechanical integration of actin and adhesion dynamics in cell migration. *Annual review of cell and developmental biology*, 26:315–333, 2010.
- [42] M. L. Gardel, M. T. Valentine, J. C. Crocker, A. R. Bausch, and D. A. Weitz. Microrheology of entangled F-actin solutions. *Physical Review Letters*, 91:158302, Oct 2003.
- [43] B. Geiger, J. P. Spatz, and A. D. Bershadsky. Environmental sensing through focal adhesions. *Nature Reviews. Molecular Cell Biology*, 10(1):21, 2009.
- [44] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1-3):404–425, 1992.
- [45] R. Gorelik and A. Gautreau. Quantitative and unbiased analysis of directional persistence in cell migration. *Nature protocols*, 9(8):1931, 2014.
- [46] A. Groh and A. K. Louis. Stochastic modelling of biased cell migration and collagen matrix modification. *Journal of Mathematical Biology*, 61(5):617–647, 2010.
- [47] P. Guthardt Torres, I. B. Bischofs, and U. S. Schwarz. Contractile network models for adherent cells. *Physical Review Letters E*, 85:011913, Jan 2012.
- [48] T. Hillen and K. J. Painter. Transport and anisotropic diffusion models for movement in oriented habitats. In *Dispersal, individual movement and spatial ecology*, pages 177–222. Springer, 2013.
- [49] W. R. Holmes and L. Edelstein-Keshet. A comparison of computational models for eukaryotic cell shape and motility. *PLoS Computational Biology*, 8(12):e1002793, 2012.
- [50] W. R. Holmes, B. Lin, A. Levchenko, and L. Edelstein-Keshet. Modelling cell polarization driven by synthetic spatially graded Rac activation. *PLoS Computational Biology*, 8(6):e1002366, 2012.

- [51] P. Hotulainen and P. Lappalainen. Stress fibers are generated by two distinct actin assembly mechanisms in motile cells. *The Journal of Cell Biology*, 173(3):383–394, 2006.
- [52] T. Kalwarczyk, N. Ziębacz, A. Bielejewska, E. Zaboklicka, K. Koynov, J. Szymański, A. Wilk, A. Patkowski, J. Gapiński, H.-J. Butt, and R. Holyst. Comparative analysis of viscosity of complex liquids and cytoplasm of mammalian cells at the nanoscale. *Nano Letters*, 11(5):2157–2163, 2011.
- [53] N.G. van Kampen. *Stochastic processes in physics and chemistry*. Elsevier, 1992.
- [54] E. Kassianidou, A. C. Brand, U. S. Schwarz, and S. Kumar. Geometry and network connectivity govern the mechanics of stress fibers. *Proceedings of the National Academy of Sciences*, 114(10):2622–2627, 2017.
- [55] E. Kassianidou and S. Kumar. A biomechanical perspective on stress fiber structure and function. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1853(11):3065–3074, 2015.
- [56] J. Kelkel and C. Surulescu. A multiscale approach to cell migration in tissue networks. *Mathematical Models and Methods in Applied Sciences*, 22(03):1150017, 2012.
- [57] S. Klumpp and R. Lipowsky. Cooperative cargo transport by several molecular motors. *Proceedings of the National Academy of Sciences*, 102(48):17284–17289, 2005.
- [58] K. E. Kubow, V. D. Shuklis, D. J. Sales, and A. R. Horwitz. Contact guidance persists under myosin inhibition due to the local alignment of adhesions and individual protrusions. *Scientific Reports*, 7(1):14380, 2017.
- [59] D. A. Kulawiak, B. A. Camley, and W.-J. Rappel. Modeling contact inhibition of locomotion of colliding cells migrating on micropatterned substrates. *PLoS Computational Biology*, 12(12):e1005239, 2016.
- [60] F. J. Kull and S. A. Endow. Force generation by kinesin and myosin cytoskeletal motor proteins. *Journal of Cell Science*, 126(1):9–19, 2013.
- [61] S. Kumar, I. Z. Maxwell, A. Heisterkamp, T. R. Polte, T. P. Lele, M. Salanga, E. Mazur, and D. E. Ingber. Viscoelastic retraction of single living stress fibers and its impact on cell shape, cytoskeletal organization, and extracellular matrix mechanics. *Biophysics Journal*, 90(10):3762–3773, 2006.
- [62] M. Kuragano, T. Q. Uyeda, K. Kamijo, Y. Murakami, and M. Takahashi. Different contributions of nonmuscle myosin iia and iib to the organization of stress fiber subtypes in fibroblasts. *Molecular biology of the cell*, 29(8):911–922, 2018.
- [63] L. D. Landau and E. M. Lifshitz. *Fluid Mechanics*, volume 6. Butterworth-Heinemann, 1987.

- [64] D. A. Lauffenburger and A. F. Horwitz. Cell migration: a physically integrated molecular process. *Cell*, 84(3):359–369, 1996.
- [65] P. A. Lewis and G. S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistics (NRL)*, 26(3):403–413, 1979.
- [66] F. Li, S. D. Redick, H. P. Erickson, and V. T. Moy. Force measurements of the $\alpha5\beta1$ integrin–fibronectin interaction. *Biophysical Journal*, 84(2):1252–1262, 2003.
- [67] L. Li, S. F. Nørrelykke, and E. C. Cox. Persistent cell motion in the absence of external signals: A search strategy for eukaryotic cells. *PLoS ONE*, 3(5):1–11, 05 2008.
- [68] B. Lin, T. Yin, Y. I. Wu, T. Inoue, and A. Levchenko. Interplay between chemotaxis and contact inhibition of locomotion determines exploratory cell migration. *Nature Communications*, 6:6619, 2015.
- [69] J. A. Lippincott-Schwartz. Steps of cell crawling. <https://science.nichd.nih.gov/confluence/display/sob/Cell+Motility>, 2018. [Online; accessed 15.09.2018].
- [70] Y.-J. Liu, M. Le Berre, F. Lautenschlaeger, P. Maiuri, A. Callan-Jones, M. Heuzé, T. Takaki, R. Voituriez, and M. Piel. Confinement and low adhesion induce fast amoeboid migration of slow mesenchymal cells. *Cell*, 160(4):659–672, 2015.
- [71] J. Löber, F. Ziebert, and I. S. Aranson. Collisions of deformable cells lead to collective migration. *Scientific Reports*, 5:9172, 2015.
- [72] A. Manhart, D. Oelz, C. Schmeiser, and N. Sfakianakis. An extended filament based lamellipodium model produces various moving cell shapes in the presence of chemotactic signals. *Journal of Theoretical Biology*, 382:244–258, 2015.
- [73] A. F. Marée, A. Jilkine, A. Dawes, V. A. Grieneisen, and L. Edelstein-Keshet. Polarization and movement of keratocytes: a multiscale modelling approach. *Bulletin of Mathematical Biology*, 68(5):1169–1211, 2006.
- [74] A. F. M. Marée, V. A. Grieneisen, and L. Edelstein-Keshet. How cells integrate complex stimuli: The effect of feedback from phosphoinositides and cell shape on cell polarization and motility. *PLoS Computational Biology*, 8(3):1–20, 03 2012.
- [75] A. M. Mastro, M. A. Babich, W. D. Taylor, and A. D. Keith. Diffusion of a small molecule in the cytoplasm of mammalian cells. *Proceedings of the National Academy of Sciences*, 81(11):3414–3418, 1984.
- [76] R. Mayor and C. Carmona-Fontaine. Keeping in touch with contact inhibition of locomotion. *Trends in Cell Biology*, 20(6):319 – 328, 2010.

- [77] G. Meral, C. Stinner, and C. Surulescu. On a multiscale model involving cell contractivity and its effects on tumor invasion. *Discrete & Continuous Dynamical Systems - B*, 20(1):189–213, 2015.
- [78] B. Merchant, L. Edelstein-Keshet, and J. J. Feng. A Rho-GTPase based model explains spontaneous collective migration of neural crest cell clusters. *Developmental Biology*, 2018.
- [79] A. Mogilner. Mathematics of cell motility: have we got its number? *Journal of Mathematical Biology*, 58(1-2):105, 2009.
- [80] C. Möhl, N. Kirchgessner, C. Schäfer, B. Hoffmann, and R. Merkel. Quantitative mapping of averaged focal adhesion dynamics in migrating cells by shape normalization. *Journal of Cell Science*, 125(1):155–165, 2012.
- [81] J. E. Molloy, J. E. Burns, J. Kendrick-Jones, R. T. Tregear, and D. C. White. Movement and force produced by a single myosin head. *Nature*, 378(6553):209–212, 1995.
- [82] Y. Mori, A. Jilkin, and L. Edelstein-Keshet. Wave-pinning and cell polarity from a bistable reaction-diffusion system. *Biophysical Journal*, 94(9):3684–3697, 2008.
- [83] M. Murrell, P. W. Oakes, M. Lenz, and M. L. Gardel. Forcing cells into shape: the mechanics of actomyosin contractility. *Nature Reviews Molecular Cell Biology*, 16(8):486–498, 2015.
- [84] P. Oakes, S. Banerjee, M. Marchetti, and M. Gardel. Geometry regulates traction stresses in adherent cells. *Biophysical Journal*, 107(4):825 – 833, 2014.
- [85] P. W. Oakes, E. Wagner, C. A. Brand, D. Probst, M. Linke, U. S. Schwarz, M. Glotzer, and M. L. Gardel. Optogenetic control of RhoA reveals zyxin-mediated elasticity of stress fibers. *Nature Communications*, 8:15817, 2017.
- [86] D. Oelz and C. Schmeiser. *How do cells move? Mathematical modeling of cytoskeleton dynamics and cell migration. Cell mechanics: from single scale-based models to multiscale modeling*. Chapman and Hall, 2010.
- [87] H. G. Othmer, S. R. Dunbar, and W. Alt. Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26(3):263–298, 1988.
- [88] H. G. Othmer and T. Hillen. The diffusion limit of transport equations II: Chemotaxis equations. *SIAM Journal on Applied Mathematics*, 62(4):1222–1250, 2002.
- [89] K. Painter and T. Hillen. Mathematical modelling of glioma growth: The use of Diffusion Tensor Imaging (DTI) data to predict the anisotropic pathways of cancer invasion. *Journal of Theoretical Biology*, 323:25 – 39, 2013.

- [90] K. Paňková, D. Rösel, M. Novotný, and J. Brábek. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cellular and Molecular Life Sciences*, 67(1):63–71, 2010.
- [91] R. Paul, P. Heil, J. P. Spatz, and U. S. Schwarz. Propagation of mechanical stress through the actin cytoskeleton toward focal adhesions: model and experiment. *Biophysical Journal*, 94(4):1470–1482, 2008.
- [92] R. J. Petrie, A. D. Doyle, and K. M. Yamada. Random versus directionally persistent cell migration. *Nature Reviews. Molecular Cell Biology*, 10(8):538–549, 2009.
- [93] L. S. Price, J. Leng, M. A. Schwartz, and G. M. Bokoch. Activation of Rac and Cdc42 by integrins mediates cell spreading. *Molecular Biology of the Cell*, 9(7):1863–1871, 1998.
- [94] M. Raftopoulou and A. Hall. Cell migration: Rho GTPases lead the way. *Developmental Biology*, 265(1):23–32, 2004.
- [95] G. R. Ramirez-San Juan, P. W. Oakes, and M. L. Gardel. Contact guidance requires spatial control of leading-edge protrusion. *Molecular Biology of the Cell*, 28(8):1043–1053, 2017.
- [96] A. J. Ridley. Rho GTPases and cell migration. *Journal of Cell Science*, 114(15):2713–2722, 2001.
- [97] A. J. Ridley, M. A. Schwartz, K. Burridge, R. A. Firtel, M. H. Ginsberg, G. Borisy, J. T. Parsons, and A. R. Horwitz. Cell migration: integrating signals from front to back. *Science*, 302(5651):1704–1709, 2003.
- [98] A. Roycroft and R. Mayor. Molecular basis of contact inhibition of locomotion. *Cellular and Molecular Life Sciences*, 73(6):1119–1130, 2016.
- [99] B. Rubinstein, K. Jacobson, and A. Mogilner. Multiscale two-dimensional modeling of a motile simple-shaped cell. *Multiscale Modeling & Simulation*, 3(2):413–439, 2005.
- [100] R. J. Russell, S.-L. Xia, R. B. Dickinson, and T. P. Lele. Sarcomere mechanics in capillary endothelial cells. *Biophysical Journal*, 97(6):1578–1585, 2009.
- [101] J. Satulovsky, R. Lui, and Y.-l. Wang. Exploring the control circuit of cell migration by mathematical modeling. *Biophysical Journal*, 94(9):3671–3683, 2008.
- [102] E. Scarpa, A. Roycroft, E. Theveneau, E. Terriac, M. Piel, and R. Mayor. A novel method to study contact inhibition of locomotion using micropatterned substrates. *Biology Open*, 2(9):901–906, 2013.
- [103] E. Scarpa, A. Szabó, A. Bibonne, E. Theveneau, M. Parsons, and R. Mayor. Cadherin switch during EMT in neural crest cells leads to contact inhibition of locomotion via repolarization of forces. *Developmental Cell*, 34(4):421 – 434, 2015.

- [104] U. S. Schwarz and M. L. Gardel. United we stand integrating the actin cytoskeleton and cell matrix adhesions in cellular mechanotransduction. *Journal of Cell Science*, 125(13):3051–3060, 2012.
- [105] D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg. Cell motility as persistent random motion: Theories from experiments. *Biophysical Journal*, 89(2):912 – 931, 2005.
- [106] D. Shao, W.-J. Rappel, and H. Levine. Computational model for cell morphodynamics. *Physical Review Letters*, 105(10):108104, 2010.
- [107] A. Shellard, A. Szabó, X. Trepát, and R. Mayor. Supracellular contraction at the rear of neural crest cell groups drives collective chemotaxis. *Science*, 362(6412):339–343, 2018.
- [108] J. R. D. Soiné, C. A. Brand, J. Stricker, P. W. Oakes, M. L. Gardel, and U. S. Schwarz. Model-based traction force microscopy reveals differential tension in cellular actin bundles. *PLoS Computational Biology*, 11(3):e1004076, 2015.
- [109] C. Stinner, C. Surulescu, and G. Meral. A multiscale model for pH-tactic invasion with time-varying carrying capacities. *IMA Journal of Applied Mathematics*, 80(5):1300–1321, 2014.
- [110] C. Stinner, C. Surulescu, and A. Uatay. Global existence for a go-or-grow multiscale model for tumor invasion. *Mathematical Methods in the Applied Sciences*, 26(11):2163–2201, 2016.
- [111] C. Stinner, C. Surulescu, and M. Winkler. Global weak solutions in a pde-ode system modeling multiscale cancer cell invasion. *SIAM Journal on Mathematical Analysis*, 46(3):1969–2007, 2014.
- [112] B. Stramer and R. Mayor. Mechanisms and in vivo functions of contact inhibition of locomotion. *Nature Reviews Molecular Cell Biology*, 18(1), 2017.
- [113] B. Stramer, S. Moreira, T. Millard, I. Evans, C.-Y. Huang, O. Sabet, M. Milner, G. Dunn, P. Martin, and W. Wood. Clasp-mediated microtubule bundling regulates persistent motility and contact repulsion in drosophila macrophages in vivo. *Journal of Cell Biology*, 189(4):681–689, 2010.
- [114] J. Stricker, Y. Aratyn-Schaus, P. W. Oakes, and M. L. Gardel. Spatiotemporal constraints on the force-dependent growth of focal adhesions. *Biophysical Journal*, 100(12):2883–2893, 2011.
- [115] T. M. Svitkina, I. G. Surguchova, A. B. Verkhovsky, V. I. Gelfand, M. Moeremans, and J. De Mey. Direct visualization of bipolar myosin filaments in stress fibers of cultured fibroblasts. *Cell Motility and the Cytoskeleton*, 12(3):150–156, 1989.

- [116] A. Szabó and R. Mayor. Mechanisms of neural crest migration. *Annual Review of Genetics*, 52:43–63, 2018.
- [117] E. Theveneau, L. Marchant, S. Kuriyama, M. Gull, B. Moepps, M. Parsons, and R. Mayor. Collective chemotaxis requires contact-dependent cell polarity. *Developmental Cell*, 19(1):39–53, 2010.
- [118] E. Theveneau, B. Steventon, E. Scarpa, S. Garcia, X. Trepast, A. Streit, and R. Mayor. Chase-and-run between adjacent cell populations promotes directional collective migration. *Nature Cell Biology*, 15(7):763, 2013.
- [119] R. T. Tranquillo, D. A. Lauffenburger, and S. Zigmond. A stochastic model for leukocyte random motility and chemotaxis based on receptor binding fluctuations. *The Journal of Cell Biology*, 106(2):303–309, 1988.
- [120] A. Uatay. A stochastic modeling framework for single cell migration: coupling contractility and focal adhesions. *arXiv: 1810.11435*, Oct. 2018.
- [121] A. Uatay. A mathematical model of contact inhibition of locomotion: coupling contractility and focal adhesions. *arXiv:1901.02386*, 2019.
- [122] A. B. Verkhovsky and G. G. Borisy. Non-sarcomeric mode of myosin ii organization in the fibroblast lamellum. *The Journal of Cell Biology*, 123(3):637–652, 1993.
- [123] A. B. Verkhovsky, T. M. Svitkina, and G. G. Borisy. Self-polarization and directional motility of cytoplasm. *Current Biology*, 9(1):11–S1, 1999.
- [124] M. Vicente-Manzanares, X. Ma, R. S. Adelstein, and A. R. Horwitz. Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nature Reviews. Molecular Cell Biology*, 10(11):778–790, 2009.
- [125] M. Vicente-Manzanares, J. Zareno, L. Whitmore, C. K. Choi, and A. F. Horwitz. Regulation of protrusion, adhesion dynamics, and polarity by myosins IIA and IIB in migrating cells. *The Journal Of Cell Biology*, 176(5):573–580, 2007.
- [126] A. Vilfan and T. Duke. Instabilities in the transient response of muscle. *Biophysical Journal*, 85(2):818–827, 2003.
- [127] M. D. Vose. A linear algorithm for generating random numbers with a given distribution. *IEEE Transactions on Software Engineering*, 17(9):972–975, 1991.
- [128] S. Walcott, D. M. Marshaw, and P. E. Debold. Mechanical coupling between myosin molecules causes differences between ensemble and single-molecule measurement. *Biophysical Journal*, 103(3):501–510, 2012.

-
- [129] P. T. Yam, C. A. Wilson, L. Ji, B. Hebert, E. L. Barnhart, N. A. Dye, P. W. Wiseman, G. Danuser, and J. A. Theriot. Actin–myosin network reorganization breaks symmetry at the cell rear to spontaneously initiate polarized cell motility. *Journal of Cell Biology*, 178(7):1207–1221, 2007.
- [130] R. Zaidel-Bar, M. Cohen, L. Addadi, and B. Geiger. Hierarchical assembly of cell–matrix adhesion complexes. *Biochemical Society Transactions*, 32(3):416–420, 2004.
- [131] J. Zhu and A. Mogilner. Comparison of cell migration mechanical strategies in three-dimensional matrices: a computational study. *Interface focus*, 6(5):20160040, 2016.
- [132] F. Ziebert and I. S. Aranson. Computational approaches to substrate-based cell motility. *npj Computational Materials*, 2:16019, 2016.
- [133] J. Zimmermann, B. A. Camley, W.-J. Rappel, and H. Levine. Contact inhibition of locomotion determines cell–cell and cell–substrate forces in tissues. *Proceedings of the National Academy of Sciences*, 113(10):2660–2665, 2016.

List of Figures

2.1	Migration cycle	11
2.2	Schematic representation of a cell	13
2.3	Schematic representation of an SF and its force profile	14
2.4	Fluorescence and traction force microscopy images	15
2.5	Active cable network	16
2.6	Shape normalization illustration	17
2.7	Transmission of internally generated forces	18
2.8	Schematic representation of the migration cycle between adhesion events.	20
3.1	Schematic representation of an FA	34
3.2	Force dependence of unbinding and binding rates.	37
3.3	Schematic diagram of the interactions between RhoGTPases and PIs	39
3.4	Myosin cross-bridge cycle	42
3.5	Reduced chemical model of cross-bridge cycle	43
3.6	Stationary distribution of active myosin motors	46
4.1	Illustration of the interarrival time computation	51
4.2	Simulations with varying number of FAs only	56
4.3	Directionality ratio in uniform environment	57
4.4	Simulations with varying binding bias	59
4.5	Persistence of motion with varying δ	60
4.6	Simulations with external cue gradient	62
4.7	Simulation with external cue gradient	63
4.8	Persistence of motion for varying δ_E	64
4.9	Concentration of an external cue projected on the cell's circumference.	64
4.10	The force dependence of the binding rate and the biased adhesion formation during the migration cycle.	65
4.11	A result of the study in [95]	66
4.12	A cell spread on different stripe patterns	67
4.13	Simulations of contact guidance	68
4.14	Profiles of an adhesion pattern and an external cue, projected on cell's circumference	69
4.15	A schematic representation of asymmetric contractility	70

4.17	Simulations with asymmetric contractility and varying number of FAs . . .	71
4.16	Simulations of asymmetric contractility	72
4.18	Ratios of the number of binding to unbinding events in each sector with varying buckling length and stiffness of SFs.	73
5.1	Representation of CIL	74
5.2	RhoGTPases switch schematics	75
5.3	Representation of binary collisions	78
5.4	A result of the study in [68]	79
5.5	CIL simulations in confined setting	81
5.6	CIL simulations in confined setting	82
5.7	CIL simulations in confined setting	83
5.8	CIL simulations on confined setting with an external signal.	84
6.1	Schematic representation of CIL for non-binary collisions.	94
6.2	Initial configurations of cells in unconfined setting	94
6.3	Simulation results for a homogeneous population of non-chemotaxing and chemotaxing cells	95
6.4	Simulation results of heterotypic CIL in the mixed configuration	96
6.5	Simulation results of heterotypic CIL in the setting with distinct separation	97
8.1	Multiscale description of cell migration.	120

List of Tables

4.1	Parameters used for simulation	54
4.2	Parameters obtained from the simulations.	57
4.3	Parameters obtained from the simulations with varying δ	58
4.4	Parameters obtained from the simulations with varying δ_E	61
4.5	Parameters obtained from the simulations with varying δ_{myo}	71
5.1	Parameter values corresponding to three scenarios S1-S3.	81

Wissenschaftlicher Werdegang

Mai 2008	Abitur Lyceum №1 Astana, Kasachstan
August 2008 - Mai 2012	Bachelor of Science , Mathematik Rochester Institute Of Technology Rochester, NY, USA
Oktober 2013 - Oktober 2015	Master of Science , Mathematik Technische Universität Kaiserslautern Kaiserslautern, Deutschland
Oktober 2015 - Mai 2019	Dr. rer. nat. , Mathematik Technische Universität Kaiserslautern Kaiserslautern, Deutschland

Academic Curriculum Vitae

May 2008	High School Diploma Lyceum №1 Astana, Kazakhstan
August 2008 - May 2012	Bachelor of Science , Mathematics Rochester Institute Of Technology Rochester, NY, USA
October 2013 - October 2015	Master of Science , Mathematics Technische Universität Kaiserslautern Kaiserslautern, Germany
October 2015 - May 2019	Ph.D. , Mathematics Technische Universität Kaiserslautern Kaiserslautern, Germany