DISSERTATION

COMPUTATIONAL METHODS FOR THE ANALYSIS OF CELL MIGRATION AND MOTILITY

Submitted by

Eric Colton Havenhill

Department of Mechanical Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2024

Doctoral Committee:

Advisor: Soham Ghosh

Paul Heyliger
Kirk McGilvray
Jianguo Zhao

Copyright by Eric Colton Havenhill 2024

All Rights Reserved

ABSTRACT

COMPUTATIONAL METHODS FOR THE ANALYSIS OF CELL MIGRATION AND MOTILITY

Collective cell migration (CCM) is necessary for many biological processes, such as in the formation or regeneration of tissue, fibroblast movement in wound healing, and the movement of macrophages and neutrophils in the body's immune response, to name a few. CCM is commonly modeled with PDEs, however these equations usually model the population density, rather than the displacement field describing the movement of any arbitrary cell. One unknown aspect of this movement is the various methods that cells use to facilitate communication to each other. Chemical communication plays a substantial role in directed cell movement, however, other mechanical methods, such as the propagation of stresses through a shared substrate to neighboring cells and cell behavior in a crowded environment, also play an important role which is less understood. The quantification of the kinematic and dynamic characteristics in CCM would present several novel advancements in understanding the collective cell behavior.

First, the dynamic mode decomposition (DMD) framework is utilized. DMD allows for the recovery of a dynamic system, in the form of an ODE or PDE, by sampling the states of a system. In the context of the cell migration, the displacements of fibroblasts during a scratch-wound assay are obtained, which result in a governing PDE through the DMD process. This PDE is used in conjunction with modern optimal control theory to develop a 2D and 3D trajectory for the migration of controllable cells to a target.

On an individual level, with the hybrid use of modern static structural optimization and simple non-linear control, a cell's cytoskeleton during migration can be studied, providing for the quantification of the traction force exerted on the substrate. The results of this analysis are compared with stress and structural optimization models in ANSYS and FEBio, which uses the finite ele-

ment method, so that a reasonable range of these stresses during CCM can be provided. To further study the individual mechanics of cell migration, the proposed hybrid model is extended to a fully dynamic model which predicts the cytoskeletal stress fiber formations over time that require the minimal amount of material with the use of optimal control theory. The results of this research could provide useful applications in many real-world situations, from the generating of a trajectory for microrobots during drug delivery to the study of the collective migration of organisms including cells.

ACKNOWLEDGEMENTS

I'd like to thank my advisor Dr. Soham Ghosh for the opportunities that he has given to me by his giving me a place in his research group. His guidance has been incredibly valuable and his trust to let me pursue my own intellectual interests in the field are very much appreciated. I'd like to thank my former lab mates and colleagues, including but not limited to Dr. Manoochehr Rabiei, Dr. Vatsal Joshi, Josh Thompson, Omar Sanchez, Gurashish Singh, and many others; their feedback, support, and good times we have shared during this challenging time is cherished. I would also like to thank my current lab mates in the Cellular Engineering and Mechanobiology Lab, including but not limited to Lauren Monroe, Kanita Hrustanovic, and Dr. Samantha Kaonis; their help in improving and incorporating my work, specifically as it relates to biological applications, is appreciated. I am grateful for the support and encouragement from my family over the years, including my mother Debora Havenhill and father Richey Havenhill. Their inspirational work ethic and help has been monumental in my success. I'd like to also thank my committee members, Dr. Paul Heyliger, Dr. Kirk McGilvray, and Dr. Jianguo Zhao, for their time and expertise in reviewing my work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Overview

This dissertation presents several advancements in the modeling of cell migration, giving insight to both collective and individual cell movement. This progress is largely achieved through the combination of image-based microscopy with a mathematical framework for the discovery of the rules of dynamical systems, called Dynamic Mode Decomposition (DMD) [1], modern optimal control theory, nonlinear control theory, and finite element analysis.

1.2 Interests in Collective Cell Migration

Cell migration plays an important role in many biological functions, such as during the formation or regeneration of tissue, osteoclast movement in bone formation, embryogenesis, fibroblast movement in wound healing, the movement of macrophages and neutrophils in the immune response, cancer metastasis, tumor invasion, and sustaining life [2–4]. CCM is also critical in tissue engineering applications where the cells must infiltrate through a scaffold and bypass obstacles to reach an intended target. CCM can be extended to larger biological systems, such as herd migration in different organisms from bacteria to large animals [5]. Therefore, the development of a predictive biophysical model of the characteristics and key parameters of migrating entities would be beneficial to the scientific community.

Mathematical equations, such as the diffusion equation and the Fisher-Kolmogorov-Petrovskii-Piskunov, shown in Eq. (1.1), are being used to model collective cell migration, where ρ represents the population density, D is the diffusivity, λ is the proliferation rate, and K is the carrying capacity density [6–8]. Most models in the literature using the density as the field variable of consideration in the mathematical model, rather than the individual position of a cell over time. Making progress to close this gap in mathematical modeling could give important contribution to the biomechanics

community.

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} + \lambda\rho(1 - \frac{\rho}{K}) \tag{1.1}$$

1.3 Dynamic Model Discovery

In the literature, Dynamic Mode Decomposition (DMD) is a method that has been developed and used to obtain the governing dynamical equations of a system, both ODEs and PDEs, for many well-known systems such as the Van der Pol oscillator and Burgers' equation [1] with only the use of spatio-temporally obtained data obtained through sampling the system. While the DMD process hasn't been applied to biological system as frequently as other systems studied in scientific and engineering fields, the success of this method in these other fields serves as encouraging evidence that this framework will work for biological systems as well.

Additionally, research have been done in the literature that integrates DMD and control theory [9]. Some studies have integrated cell migration with optimal control [10]. Here, the authors combine optimal control theory with phase field equations of equilibrium to model cell shape and deformation. Further studies, including those done in this dissertation, will need to look at multicellular systems as well as consider alternative models that include the various components within the cell's cytoskeleton.

1.4 Experimental Methods for Cell Migration

Directed cell migration happens through means including haptotaxis, chemotaxis, topotaxis, and galvanotaxis [11]. Internally, cell migration happens through an expansion, protrusion-adhesion-deadhesion-contraction cycle with the extra-cellular matrix, which is often just the substrate [4, 12–14]. In order to study the forces that are generated by the cell during these processes, experimental methods have been developed in order to quantify some of the dynamic properties of cell migration. One such experimental method is known as deformation microscopy, which is a methodology that leverages conventional imaging and an automated hyperelastic warping algorithm to investigate strain history, deformation dynamics, and changes in structural heterogeneity within the interior of

cells and cell nuclei [15]. Another popular experimental method for determining the forces during cell migration is known as traction force microscopy [16]. This process can produce stress maps during cell migration by observing the displacement of beads immersed in an extra-cellular matrix (ECM) of known material properties during migration and mapping these deformations to stresses through constitutive laws and stress-strain relations.

1.5 FEM-based Models for Cell Migration

Physical systems, specifically cellular systems, can be and are frequently modeled by PDEs. Analytical methods exist that can solve many types of PDEs, such as through separation of variables, for example. However, non-linearities and complicated boundary conditions are a couple of issues that can quickly make analytically-based solutions infeasible. For such situations, numerical methods are sought. Several methods exist for numerically solving ODEs and PDEs, such as the finite difference method, finite volume method, and finite element method. The finite element method, developed largely over the 20^{th} century, has become the most popular and lucrative method for solving such equations in engineering and physics. For cellular systems, finite element-based models have been used in order to quantify substrate stress [17], model cell motility [10, 18], focus on the mechanical forces, using a reaction-diffusion model that emphasizes cell protrusion, adhesion, and contraction with the substrate [2, 14, 18–20]. Literature on cell migration shows that dynamic tensile forces drive cell migration [21]. Cell to cell communication through the shared substrate, an important aspect of collective cell migration, could possibly be better quantified with the use of sophisticated finite element models.

1.6 Modern Optimization

Trajectory optimization, which traces its roots to Bernoulli's brachistrochone problem, started in the late 17^{th} century, has laid the foundation for modern optimal control that was developed in the 1950's. Optimal control is a powerful and flexible framework that has been hugely successful for the controlling of complicated dynamical systems, from robotic and aerospace systems to

controlling the spread of a pandemic [22]. For biological systems at the cellular level, practical applications of optimal control exist in the field of medical treatment services that involve magnetic micro-robots, such as drug delivery with a micro-robot [23].

Optimization extends past human-based engineering and includes optimizations in nature. For example, swimming in schools helps fish to swim more efficiently [24, 25] and birds fly together in a V-shape to reduce each bird's individual drag force. This type of interaction between members of a system is key to achieving an optimal behavior. In microscopic applications, one questions whether or not the cells communicate with each other through the substrate, which acts as a kind of mediator for cell-to-cell communication, in order to collectively move in an efficient manner. One can use optimal control to develop a metric for how well the systems perform under under a variety of conditions.

1.6.1 Structural Optimization

One important component of cell migration is the means by which cells communicate mechanically. The quantification of the traction forces during cell migration serve as valuable information for cell mechanosensing and understanding the driving components to cell migration [26]. The stress fibers that are produced as a part of the cytoskeleton serve as the key structure in moving the nucleus during migration. The stress fiber orientation has been studied in the literature [27, 28]. Different predictive models have been created to study the formation of these stress fibers as the cell ambulates, as well as the forces that these stress fibers exert on the substrate. The method proposed in this dissertation suggests that the use of modern structural optimization advancements could hold the key to quantifying these unknowns. The structural optimization problems are posed similarly to the optimal control formulation shown in Eq. (1.2), with similar methods of solving. The quantification of these traction forces and orientation and formation of the stress fibers during migration would be beneficial to the scientific community.

1.6.2 Optimal Control

Mathematically, optimal control problems are posed as in Eq. (1.2), where the cost functional J to be minimized is the Bolza cost function including the Mayer term, $\Phi[\cdot]$, and the Lagrange term, L$[\cdot]$, which is subjected to constraints on the system, where $\mathbf{f}(\mathbf{x}(t), t)$, $\phi(\mathbf{x}_1, t_1, \mathbf{x}(t_f), t_f)$, and $\mathbf{C}(\mathbf{x}(t), t)$ represent the dynamics, boundary constraints, and path constraints, respectively. The state vector, $\mathbf{x}$, is constructed based on the context of the problem. For example, if the motion of only one migrational cell is of consideration, then $\mathbf{x} \in \mathbb{R}^{1\times 1}$. If the motion of i cells is of consideration, then $\mathbf{x} \in \mathbb{R}^{i\times 1}$.

$$
\begin{aligned}
\underset{\mathbf{u}}{\text{minimize}} \quad & J = \Phi(\mathbf{x}(t_1), t_1, \mathbf{u}(t_f), t_f) + \int_{t_1}^{t_f} L(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), t)dt \\
\text{subject to:} \quad & \text{Dynamic Constraint:} \quad \dot{\mathbf{x}}(t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{0} \\
& \text{BCs:} \quad \phi(\mathbf{x}(t_1)), t_1, \mathbf{x}(t_f), t_f) = \mathbf{0} \\
& \text{Path Constraint:} \quad \mathbf{C}(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{0}
\end{aligned}
\tag{1.2}
$$

Optimal control problems are commonly solved with a variety of methods shown in Fig. 1.1: indirect, direct, and hybrid methods. Indirect methods seek to formulate the necessary and sufficient conditions of optimality for Eq. (1.2) with the use of calculus of variations. This system of PDE are solved analytically. Indirect methods will be impractical to solve such a complicated optimization problem that will be posed here; therefore, direct or hybrid methods will be the appropriate approach. So-called hybrid methods work by first obtaining the same necessary and sufficient conditions as the indirect method, however the resulting PDE are solved numerically, rather than analytically, through methods such as with shooting methods. Direct methods do not seek to formulate the necessary and sufficient conditions. Instead, direct methods will seek to discretize Eq. (1.2) first, and then optimize the resulting formulation numerically. While the hybrid methods tend to be more accurate than the direct methods, the hybrid methods tend to be harder to pose and solve. Additionally, direct collocation methods tend to be better for problems that involve complicated control and path constraints. Due to these intricacies of the problems posed in this dissertation, direct collocation is used.

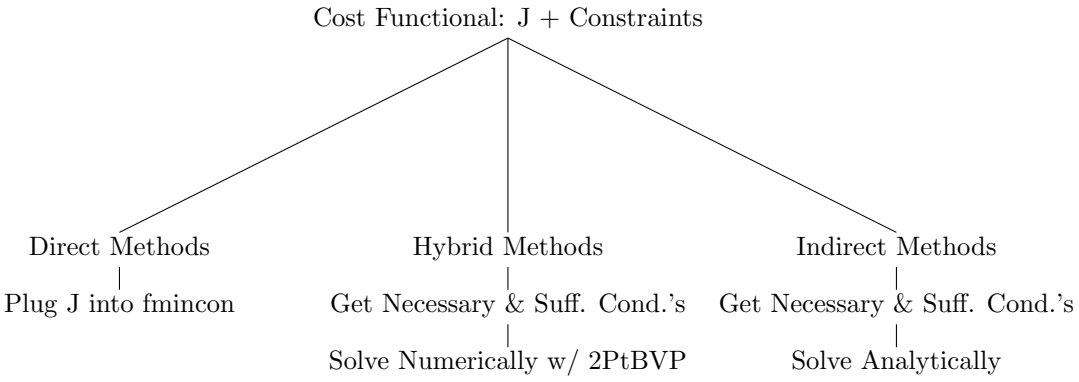


Figure 1.1: Tree Diagram of Optimal Control Methodologies

1.6.3 Direct Collocation Method - Optimization Through Dynamic Programming

The optimization problem formulated in Eq. (1.2) is a continuous-time formulation, which is commonly transcribed into a discretized set of equations that approximates the original formulation as part of the direct collocation method [29]. First, the integral in the cost functional to be minimized can be approximated with a quadrature formulation, shown in Eq. (1.3), where n represents the number of nodes that comes from the temporal discretization; this is commonly done with the trapezoidal method, however, more sophisticated methods are also commonly used for increased accuracy. The integral has been successfully converted into an algebraic expression.

$$\int_{t_1}^{t_f} L(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), t)dt \approx \sum_{j=1}^{n} L(\mathbf{x}(t_j), \dot{\mathbf{x}}(t_j), \mathbf{u}(t_j), t_j)\Delta t_j \tag{1.3}$$

The dynamics are also discretized, which is commonly done with a finite differencing scheme such as the one presented in Eq. (A.11); however, the finite volume method, finite element method, or other numerical methods can alternatively be used. For practical implementation, MATLAB's ode45 is used with Runge-Kutta numerical integration for improved stability and accuracy purposes. A rudimentary illustration of the transcription of the ODE dynamics into a system of algebraic equations that approximates the dynamics is shown in Eq. (1.4).

$$\dot{\mathbf{x}}(t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{0} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & \ddots & \ddots \end{bmatrix} \begin{Bmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \end{Bmatrix} + \begin{Bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{Bmatrix} = \Delta t \begin{Bmatrix} f(x_2) \\ f(x_3) \\ f(x_4) \\ \vdots \end{Bmatrix} \tag{1.4}$$

The Lagrangian cost will need to be selected; this cost is typically chosen based on the context of the problem. The expressions for different common cost formulations are given in Eq. (1.5), which will seek to minimize the final time, energy expended, error, and action, respectively. Note that T and V in these equations represent the kinetic and potential energies, respectively. In situations where no clear Lagrangian term is to be chosen, selecting a minimum energy approach in the context of biological applications may be well suited, however, more testing is needed to demonstrate this.

$$L = \begin{cases} 1 & \text{Min. Time} \\ \frac{1}{2}\dot{\mathbf{x}}^T M \dot{\mathbf{x}} & \text{Min. Energy} \\ \mathbf{x}_f - \mathbf{x} & \text{Min. Error} \\ T - V & \text{Min. Action} \end{cases} \tag{1.5}$$

After the application of the quadrature scheme, Bellman's Principle of Optimality can be applied, which state that the optimal trajectory between two nodes also lies on the optimal trajectory for the entire trajectory; this will allow the construction of a dynamic optimization problem from many consecutive static optimization problems. MATLAB's fmincon function optimizer is commonly used to solve for the minimum of the resulting cost subject to the system algebraic constraint equations, which will solve the static optimization problems with the use of algorithms such as interior-point, trust-region-reflective, sqp, or others [30].

Chapter 2

Specific Aims

The overarching goal of this research was to mathematically quantify cell migration, both internal to the cell as well as with a group of cells, and many of the parameters that play an important role during this process. In order to find answers to these unknowns, the following specific aims were formulated.

2.1 Quantify Kinematic Parameters of Cell Migration

The primary aim was to apply the DMD framework that has been previously developed to collective cell migrational systems with the goal of obtaining a mathematical expression with the displacement as the field variable pertaining to the movement of the cells during migration.

2.2 Build a Trajectory Optimization Framework for Biological Systems

The primary aim was to use trajectory optimization techniques that are common in the optimal control community to develop an optimal trajectory for the study of cell migration in biological functions like the immune response as well as with the use of microrobots for therapeutic treatment. The goal is to integrate the dynamic models that were discovered with the DMD process with modern optimal control theory.

2.3 Predict Cytoskeleton Structure During Migration

The primary aim was to use modern structural optimization techniques to predict the cytoskeletal stress fiber formations during cell migration using both a static and dynamic formulation for the governing equations of equilibrium.

2.4 Quantify the Forces Throughout the Cytoskeleton Structure During Migration

The primary aim was to use modern optimal control and nonlinear control theory in conjunction with structural optimization to predict the forces in the cell's stress fibers during migration for comparison with the forces obtained with other methods common in the field.

Chapter 3

Optimization-Based Synthesis with Directed Cell Migration

Collective behavior of biological agents such as herds of organisms and cells is a fundamental feature in the systems biology and in the emergence of new phenomena in the biological environment. Collective cell migration under a physical or chemical cue is an example of this fundamental phenomenon where individual cell migration is driven by the collective behavior of the neighboring cells and vice versa. The goal of this research is to discover the mathematical rules of collective cell migration using experimental data and testing the predictive nature of the models in independent experimental data. Such insight is made possible in this work with the hybrid use of dynamic mode decomposition (DMD) and optimal control theory. Both single and also multi-cellular systems are investigated, including obstacle courses, using this framework. The result of this work show how cells collectively behave during their migration and opens the possibility of designing robotic cells for possible therapeutic purpose where the cell trajectory can be controlled.

3.1 Introduction

Collective cell migration is a fundamental event in nature for many biological processes including the formation and regeneration of tissues. Specific examples include collective movement of cells during the developmental stage, fibroblast migration in wound healing, movement of macrophages and neutrophils in the immune response and cancer metastasis [4, 31]. Collective migration is also critical in tissue engineering applications where the cells must infiltrate through a scaffold and bypass obstacles such as extracellular matrix fibers to reach an intended target. The conceptual framework of collective cell migration can be extended to larger biological contexts, such as herd migration in different organisms from bacteria to large animals. Therefore, the devel-

opment of a predictive biophysical model of the characteristics and key parameters of migrating entities can be applied to a large array of scientific problems that extend beyond cell migration.

Currently existing approaches to model cell migration include the potential/phase-field/gradient methods which uses the population density as the field variable that limits our ability to predict the kinematic variables such as cell position and velocity [32–34]. Here, the prediction of such kinematic variables is made from the knowledge that cells have finite resources to achieve the goal of migrating from one place to another. For example, studies suggest that leukocytes, a type of blood cell, migrate efficiently with a minimum energy expenditure [33]. Accordingly, a hybrid utilization of dynamic mode decomposition (DMD) and optimal control theory is sought to develop a new predictive biophysical model for cell migration that relies on the efficient use of cell's resources.

Optimal control frequently appear in a wide variety of scientific problems, such as robotic trajectory generation, satellite orbit transfer problems, and the regulation of the spread of disease in a population. Commonly, the goal of such problems is to find a control input that will drive the states to a particular value while also minimizing the control effort, time, or some other function. For example, in a recent study [22] the authors used masks, social distancing, etc. as the controls to minimize the number of people infected and hospitalized with coronavirus during the COVID-19 pandemic while minimizing parameters such as economic impact.

Within the context of the present work, a framework for predicting collective cell migration has been developed, which is a ubiquitous problem in biology. One of the many ideals of such a collective cell migration system is the most efficient migration of individual cells in the crowded environment. Additionally, the same framework is applied to design robotic systems for drug delivery, such as a microrobot, in a cellular environment for therapeutic applications as described in several previous studies [35–37]. Such a robotic drug delivery system has a minimal impact on the surrounding tissue as the microrobot traverses the Extracellular Matrix (ECM). For both situations, any optimal control signal should include a constraint that penalizes excessive contact within multiple cells or between cells with fibers in the ECM. The obstacle avoidance of the microrobots

or cells with the surrounding fibers/cells relies upon an accurate prediction of the cell environment, which arises from the Dynamic Mode Decomposition (DMD) of the cells. An inequality constraint prevents the penetration of the cells or robotic entities with the surrounding structures as well as pushing the cells off their natural trajectory.

To achieve the goals stated above, a governing ordinary differential equation (ODE) for a single cell, or partial differential equation (PDE) for a collection of cells in tandem is developed and used to model cell migration in the direction defined by a gradient. A hybrid model, based on optimal control, is used in conjunction with the associated ODE or PDE for a synthesis allowing for the study of cell migration in a two-dimensional space.

In literature, DMD has been used to obtain the governing dynamic equations, both ODEs and PDEs, for many well-known systems such as the Van der Pol oscillator and Burgers' equation [1]. This serves as encouraging evidence that this same methodology can be applied in biological systems. To illustrate this, consider the diffusion equation in Eq. (3.1), which is a parabolic PDE that is common in multiple fields of physics to model phenomena such as thermal, chemical, and particle diffusion. Similar equations, such as the diffusion equation with a different diffusion coefficient and the Fisher-Kolmogorov equation, were preiously used to model collective cell migration [6–8]. Artificial data can be generated with MATLAB's 'pdepe' solver for the diffusion equation presented in Eq. (3.1) with the associated boundary conditions (BC) and initial condition (IC) in Eqs. (3.2) and (3.3), respectively.

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{3.1}$$

$$u(0,t) = 0 \qquad u(1,t) = 2cos(4t) \tag{3.2}$$

$$u(x,0) = 2sin(\frac{5\pi}{2}x) \tag{3.3}$$

The solution of Eqs. (3.1-3.3) can be visualized via a surface plot Fig. 3.1. The recovered PDE is also shown for comparison. The successful reconstruction of the governing equations of

motion for these dynamic systems provides the rationale that the same methodology can be applied to biological systems, specifically for collective cell migration.

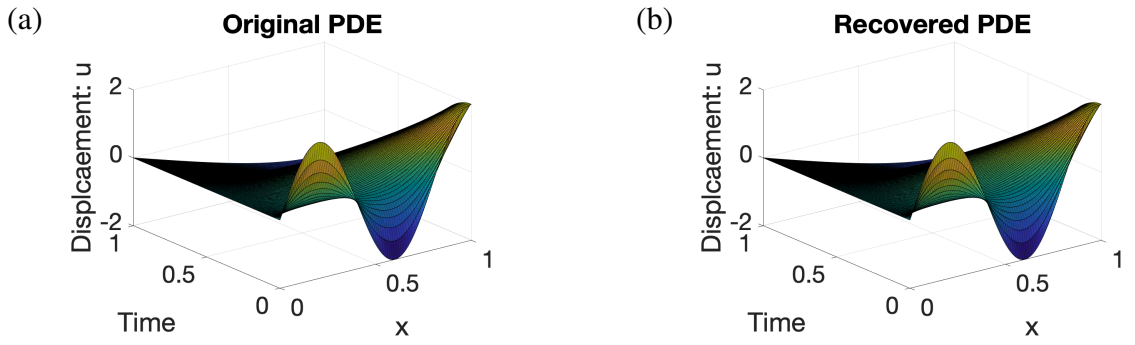


Figure 3.1: (a) Surface plot generated from Eqs. (3.1-3.3). (b) Surface plot generated from the recovered PDE and Eqs. (3.2-3.3).

3.2 Experimental Data

Experimental data of cell migration was generated utilizing the traditional scratch wound assay by culturing a monolayer of murine fibroblast cell line (NIH 3T3) in an 8-well plate (ibidi 80821), as described in our recent work [38]. A scratch was applied using a pipette tip on the confluent layer of cells. Subsequently, time-lapse images of the cell migration during the wound closure were recorded using a Zeiss LSM 980 confocal microscope with environmental chamber which mimics the physiologically relevant conditions such as temperature, humidity and carbon di-oxide level. The automated cell tracking was done on nucblue stained nuclei to collect quantitative data of the cell migration using TrackMate in Fiji [39]. The image sequence files were utilized to track hundreds of cells in the field of view. The TrackMate coordinate system and a sample data set are shown in Fig. 3.2, where the positive x-direction is horizontal and pointed to the right and the positive y-direction is vertical and pointed downwards. Cells show mostly directed migration along the x axis, with some deviations where they partially move in the y direction as well. Modes of directed cell migration include a gradient and based on the specific mechanism they are named as haptotaxis, chemotaxis, topotaxis, and galvanotaxis [11]. From the nature of this experiment,

the cells are assumed to migrate in the direction toward the center of the scratch due to a chemical attractant.

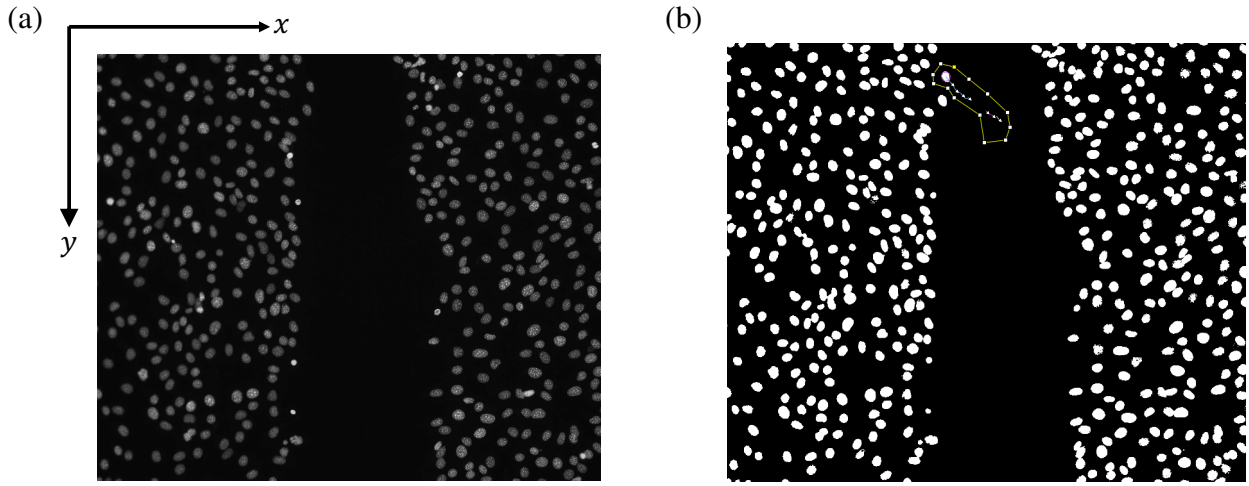


Figure 3.2: (a) Live cell imaging and automated tracking before wound closure shown of murine fibroblasts for a sample control group illustrating the scratch wound assay. (b) Tracked cell trajectory for an individual cell of interest after thresholding the image.

3.3 Dynamic Equations

The methodology for discovering the dynamics is outlined in [1] and reproduced in Appendices A.1, A.2, A.3, A.4, and A.5 in order to apply in the cell systems presented in this work. The formulation starts with Eq. (3.4), where u is the displacement, the A-matrix consists of the library of candidate dynamic functions, and ξ consists of the coefficients of the functions in the A-matrix. The larger the library of terms in the A-matrix is, from Eqs. (A.19) and (A.23), the more robust the model can be, e.g. $1, O^1, O^2, ..., u', u'', ...$ Solving this equation will allow the formulation of a differential equation that serves as the governing equation for the cell migration in the x-direction. The system solution can be obtained by solving the matrix system shown in Eq. (3.4).

$$\left\{ \begin{array}{c} | \\ \xi \\ | \end{array} \right\} = A^{+} \left\{ \begin{array}{c} | \\ \frac{\partial u}{\partial t} \\ | \end{array} \right\} \tag{3.4}$$

Note that in Eq. (3.4), A^+ represents the appropriate inverse of choice. The sparse solution for ξ was obtained in MATLAB. A sparse or parsimonious solution is preferred since it allows for the identification of the dominant terms that are driving the dynamics. An argument from intuition suggests that the recovered equation should likely contain the fewest number of terms since this is the structure that is found in many important governing PDEs such as the Korteweg–De Vries, diffusion, reaction-diffusion, Burgers, Schrödinger, Navier-Stokes, Kuramoto–Sivashinsky, wave, Laplace, and Poisson equations. Thresholding to achieve sparsity can be done to eliminate some of the more insignificant elements of ξ. A first approach to thresholding is outlined in the process shown as follows:

1. Assume $\boldsymbol{\xi} \in \mathbb{R}^{\zeta \times 1}$, where ζ is large. This is equivalent to assuming a large number of the potential functions in Eq. (A.2).

2. Solve Eq. (3.4) for $\boldsymbol{\xi}$ using the Moore-Penrose pseudoinverse for a least squares solution.

3. Inspect elements of $\boldsymbol{\xi}$ whose magnitude is smaller than a specified threshold value, typically a small number defined by the user.

4. Remove the terms from the A-matrix in Eq. (A.2) that correspond with the small $\boldsymbol{\xi}$ value from the thresholding.

5. Solve Eq. (3.4) with the new A-matrix.

6. Repeat this process until no terms in $\boldsymbol{\xi}$ are below the defined threshold.

There exists additional thresholding options for eliminating these small terms if the method stated above is undesirable. The Lasso algorithm can be used for automatically thresholding the solution of Eq. (3.4). Another option is to use MATLAB's "\" to solve the system of equations. If a scenario arisess that A is a non-square matrix, MATLAB's "\" will resort to a QR decomposition solver, which will promote sparsity. The results in this paper were obtained using this last process due to ease of implementation.

3.3.1 Single-Cell System

With the formulation stated above, the methodology is ready to be applied to the single-cell and multi-cell systems. The dynamical equations of the single-cell system are shown in Eqs. (A.25), (A.26), (A.28), and (A.30). The plots of the trajectory of the cell for the different dynamic systems is shown in Fig. A.2.

3.3.2 Multi-Cell System

A method is needed that discovers the partial differential equations of a collective cellular system of the form shown in Eq. (3.5).

$$\frac{\partial u}{\partial t} = N(u, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}, ..., x, t) \tag{3.5}$$

The dynamics are recovered as shown in Eq. (3.6), where the terms with relatively small coefficients are omitted such that the dominant terms are shown.

$$\frac{\partial u}{\partial t} \approx \underbrace{0.0085\frac{\partial u}{\partial x}}_{\text{Convection}} + \underbrace{1.1635\frac{\partial^2 u}{\partial x^2}}_{\text{Diffusion}} \tag{3.6}$$

Eq. (3.6) takes the form of the convection-diffusion equation; however, the sign of the convection coefficient is opposite to the sign of the convection coefficient in the standard convection-diffusion equation. The solution suggests that the convective term serves as a model for some amount of migration of the cells in the direction opposite to the scratch, which is consistent with experimental data where a bulk retraction of the cell layer is observed; the diffusion term serves as a model for the migration of the cells towards the scratch due to a chemical gradient.

a posteriori Dirichlet BCs

The initial condition (IC) for all of the simulations of this PDE is given in Eq. (3.7).

$$u(x, 0) = 0 \tag{3.7}$$

The boundary conditions (BC) are obtained from TrackMate from the trajectories of the cells at the two boundaries. For implementation, the data is interpolated to fit a smooth polynomial with the structure given in Eqs. (3.8) and (3.9), where the value of ν is chosen according to the user's discretion.

$$u(x = x_1, t) = \sum_{\mu=1}^{\nu} \alpha_\mu t^\mu \tag{3.8}$$

$$u(x = x_m, t) = \sum_{\mu=1}^{\nu} \beta_\mu t^\mu \tag{3.9}$$

The data of the original experiment is shown in Fig. 3.3. Comparing the surface plot of the original data with a surface plot of the data generated by the DMD framework shows an overall similar shape.

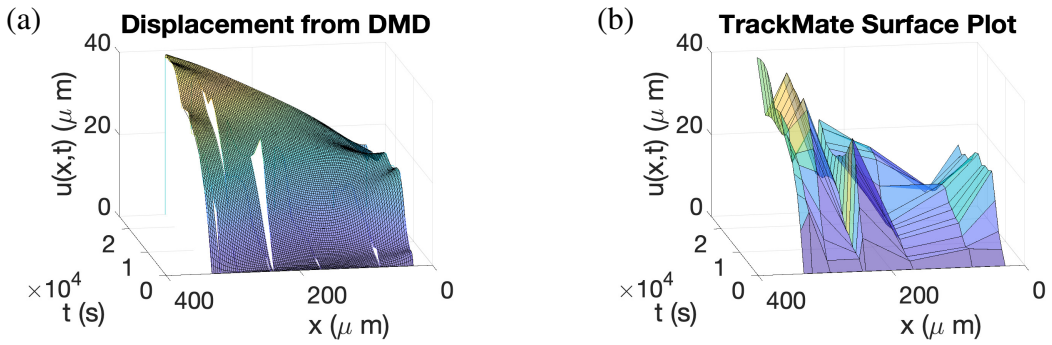


Figure 3.3: (a) Surface plot of Eq. (3.6), subject to the BC in Eqs. (3.8) and (3.9) and IC in Eq. (3.7). Waterfall plots of the trajectories obtained from the TrackMate data are included. (b) Surface plot of the original data. The final time used in the simulation is the same as in the experiment.

Velocity-Based Dirichlet BCs

Dirichlet BC can be constructed based on the first few time points of the cell trajectory. The construction of the Dirichlet boundary condition relies upon the assumption that one can obtain the trajectory of the cells at least upon the completion of the first time step, specifically in the form of the time derivative of the position. The first-order time-derivative approximation of the displacement, shown in Eq. (3.10), can be constructed with a forward-biased stencil, where δ

represents a difference approximation to a derivative [40].

$$(\delta_t u_i)^j = \frac{u_i^{j+1} - u_i^j}{\Delta t} \approx \frac{\partial u_i^j}{\partial t}\Big|_{x_i,t_j} \tag{3.10}$$

Eq. (3.10) is utilized when only data from the initial time step is available. However, more advanced stencils can be used for a more accurate approximation of the first derivative if data from multiple time steps are accessible. Examples of these advanced stencils are provided in Eqs. (A.31), and (A.32). For all of the stencils, the values of x_i are x_1 for the left boundary (wall) and x_m for the right boundary (scratch-wound interface); $j = 1$ in all cases.

In addition to the derivatives, the development of an approximation of the amount of time that it takes for the cells at the scratch wound interface to reach the center of the scratch, as shown in Eq. (3.11), can be developed. For two representative data sets, the final times of the simulation as predicted by Eq. (3.11) are an order of magnitude larger than the final time of the actual experiment.

$$t_n = \frac{u(x_m, t_n)}{\dot{x}_m(t_1)} \tag{3.11}$$

The numerator in Eq. (3.11) can be experimentally determined once the scratch in the scratch wound assay is created and imagedd. Note that for the simulations in Figs. 3.4 - 3.6, the final time of the actual experiment is used for the simulation in order to draw a comparison. Using either Eqs. (3.10), (A.31), or (A.32), one can develop the boundary conditions that are given in Eq. (3.12). Note that this same expression is used for the BC at the scratch-wound interface and the cell at the well's wall; hence, $x_i = x_1$ (at the wall) and $x_i = x_m$ (at the scratch-wound interface).

$$u(x_i, t) = \frac{\partial u_i^j}{\partial t}\Big|_{(x_i,t)} t \qquad i = 1, m \tag{3.12}$$

The approach outlined here is anticipated to be the most versatile of all of the approaches, due to the fact that it does not need the information of the global position of cells. The results of this method are shown in Fig. 3.4.

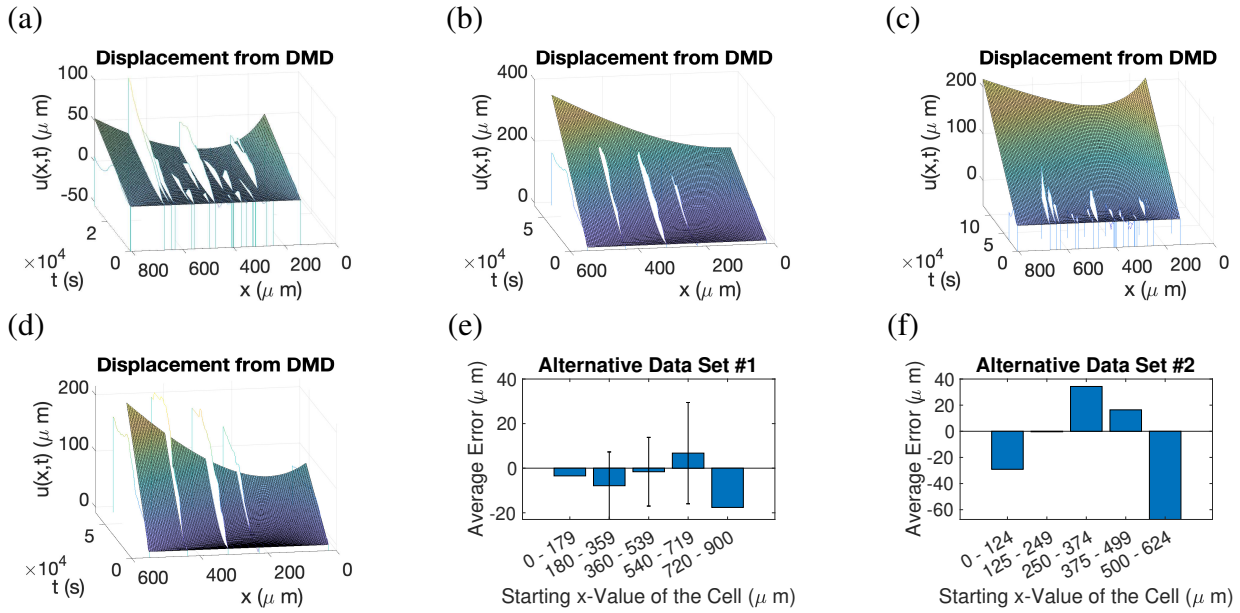


Figure 3.4: Surface plots of Eq. (3.6), subject to the BC in Eq. (3.12) and IC in Eq. (3.7). Waterfall plots of the trajectories obtained from the TrackMate data are included. For the simulation: (a) The 1^{st} data set is used with experiment's final time. (b) The 1^{st} data set is used with experiment's predicted final time, found from Eq. (3.11). (c) The 2^{nd} data set is used with experiment's final time. (d) The 2^{nd} data set is used with experiment's predicted final time, found from Eq. (3.11). (e) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (a). (f) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (b).

Zero-Displacement and Velocity-Based Dirichlet BC

The BC at the right end can be constructed from Eq. (3.12). The BC at the left end will be assumed to be a zero-displacement BC. The goal here is to provide BCs that will allow the prediction of the cell motion by using the trajectory of the cells during the first time step. The first boundary condition, given in Eq. (3.13), is obtained by assuming the trajectory of the immediate interior cell to the cell positioned at $x = x_1$ (at the wall).

$$u(x_1, t) = 0 \tag{3.13}$$

The simulation results of this approach are shown in Fig. 3.5. It should be noted that for some situations, the BC presented in Eq. (3.13) can be problematic. For simulations containing a cell whose location with the lowest x-value, x_1, is at the wall of the well, then Eq. (3.13) is the

appropriate BC. However, for cases where the lowest x-value, x_1, is far removed from the wall, then the BC representing zero displacement should not be applied.

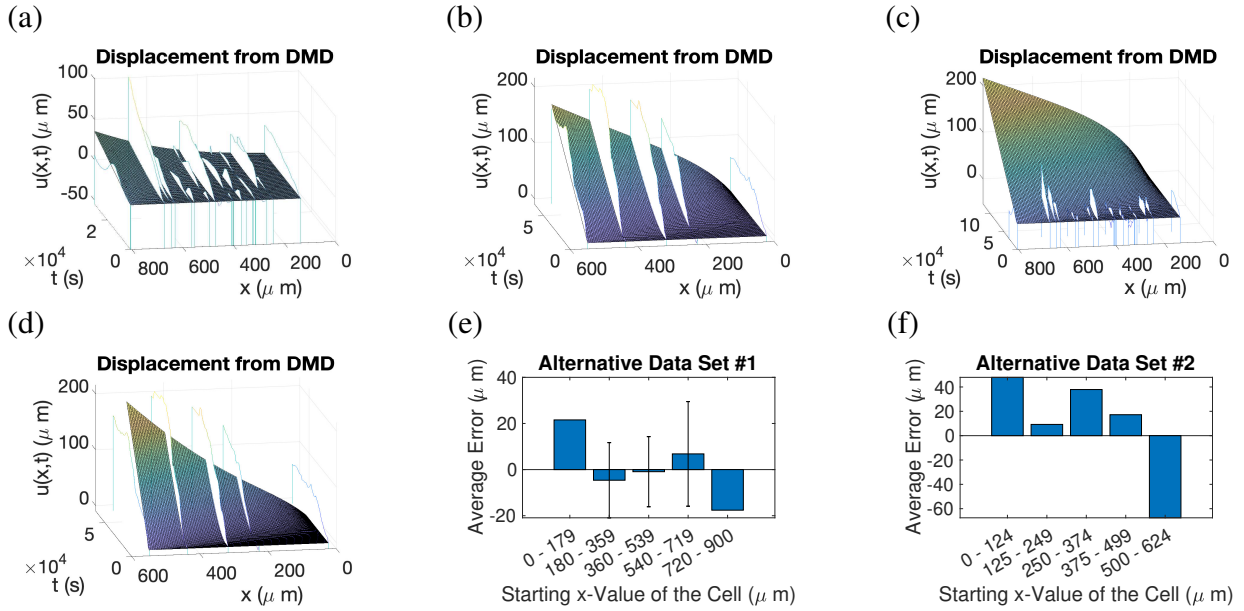


Figure 3.5: Surface plots of Eq. (3.6), subject to the BC in Eqs. (3.12) and (3.13) and IC in Eq. (3.7). Waterfall plots of the trajectories obtained from the TrackMate data are included. For the simulation: (a) The 1^{st} data set is used with experiment's final time. (b) The 1^{st} data set is used with experiment's predicted final time, found from Eq. (3.11). (c) The 2^{nd} data set is used with experiment's final time. (d) The 2^{nd} data set is used with experiment's predicted final time, found from Eq. (3.11). (e) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (a). (f) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (b).

Mixed (Dirichlet/Neumann) BCs

The BC for the right-hand side from Eq. (3.12) is used again. The second BC is constructed from the spatial derivative of Eq. (3.13), as shown in Eq. (3.14). This BC is generated as an assumption that the neighboring cell to the x_1 location will approximately move the same as the one positioned at the x_1 location.

$$\left.\frac{\partial u}{\partial x}\right|_{(x_1,t)} = 0 \tag{3.14}$$

The results of this approach are shown in Fig. 3.6. This approach is anticipated to serve as the second most robust approach after the velocity based Diricchlet BCs.

Error bars have been added to Figs. 3.4 - 3.6, where the error is defined pointwise as the difference between the DMD prediction and the experimental data. The error is averaged over time by computing the difference between the PDE and experimental data displacement over time for the respective amount of cells used; 20 cells for the first alternative data set and 5 cells for the second alternative data set. The errors are partitioned and grouped together based on their corresponding cell's starting x-value for graphing purposes. Standard deviation bars of the average error in each partition are added to the bar graphs. Some of the error bars do not have standard deviation because there was only one cell to be analyzed in that range of x-values.

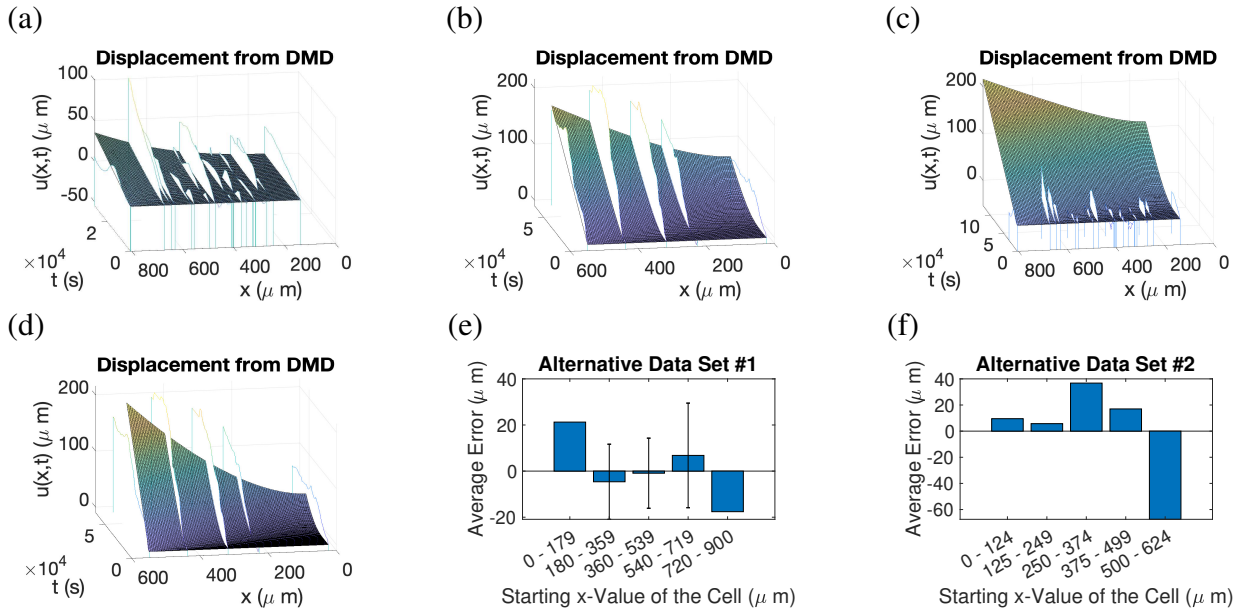


Figure 3.6: Surface plots of Eq. (3.6), subject to the BC in Eqs. (3.12) and (3.14) and IC in Eq. (3.7). Waterfall plots of the trajectories obtained from the TrackMate data are included. For the simulation: (a) The 1^{st} data set is used with the experiment's final time. (b) The 1^{st} data set is used with the experiment's predicted final time, found from Eq. (3.11). (c) The 2^{nd} data set is used with the experiment's final time. (d) The 2^{nd} data set is used with the experiment's predicted final time, found from Eq. (3.11). (e) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (a). (f) Error bars of the difference in the PDE's prediction with respect to the experimental displacement in (b).

3.4 Optimization

The movement of the cells in the direction towards the scratch was studied in the previous sections. In order to study the motion of cells perpendicular to the direction of the scratch, which is of particular interest when investigating the 2D and 3D trajectories of the cells subjected to obstacles such as other cells, an understanding through an optimization framework is sought. Optimal control relies upon this mathematical framework in order to construct optimal trajectories.

3.4.1 Direct Collocation Method - Optimization Through Dynamic Programming

The optimization problem is formulated in Eq. (5.1), where the cost functional J to be minimized is the Bolza cost function including the Mayer term, $\Phi(\cdot)$, and the Lagrange term, $\mathrm{L}(\cdot)$, which is subjected to constraints on the system, where $\mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t)$, $\phi(\mathbf{x}_1, t_1, \mathbf{x}(t_f), t_f)$, and $\mathbf{C}(\mathbf{x}(t), \mathbf{u}(t), t)$ represent the dynamics, boundary constraints, and path constraints, respectively. The state vector, $\mathbf{x}$, is constructed based on the context of the problem. For example, if the motion of only one cell is considered, then $\mathbf{x} \in \mathbb{R}^{2\times 1}$ in order to include both the x and y positions. If the motion of i cells is considered, then $\mathbf{x} \in \mathbb{R}^{2i\times 1}$. For the optimization performed on the cells in Figs. 3.7 and A.3, there is no control input, u, in the traditional sense of optimal control theory. However, one of the states, particularly the Cartesian coordinate y, is a parameter of the system that can be adjusted in order to minimize the value of the cost, which is sometimes referred to as a *decision variable*. For the cells in Fig. 3.8, both the x and y coordinates serve as the decision

variables. Therefore, the term "optimal control" is used for this optimization.

$$\begin{aligned}
\underset{\mathbf{x},\dot{\mathbf{x}},t}{\text{minimize}} \quad & J = \Phi(\mathbf{x}(t_1), t_1, \mathbf{x}(t_f), t_f) \\
& \qquad + \int_{t_1}^{t_f} L(\mathbf{x}(t), \dot{\mathbf{x}}(t), \mathbf{u}(t), t) dt \\
\text{subject to:} \quad & \dot{\mathbf{x}}(t) - \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{0} \\
& \phi(\mathbf{x}(t_1)), t_1, \mathbf{x}(t_f), t_f) = \mathbf{0} \\
& \mathbf{C}(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{0}
\end{aligned} \tag{3.15}$$

Note that the formulation presented in Eq. (5.1) is a continuous-time formulation, which means that it needs to be converted into an approximate formulation for the direct collocation through transcription into a nonlinear program. First, the integral in the cost functional to be minimized can be approximated with a quadrature formulation. For example, using the trapezoidal method, the integral given in Eq. (5.1) can be converted into an algebraic expression. The quadrature conversion used in the simulations presented in this study is presented in Eq. (5.9).

$$\sum_{j=1}^{n} L(\mathbf{x}(t_j), \dot{\mathbf{x}}(t_j), \mathbf{u}(t_j), t_j) \Delta t_j \tag{3.16}$$

The dynamics was discretized, done with a finite differencing scheme such as the one presented in Eq. (A.11); however, for more accurate results, MATLAB's ode45 was used to numerically integrate the equations. MATLAB's fmincon function solved the resulting system of algebraic equations; therefore, solved the static optimization problems with the use of algorithms such as interior-point, trust-region-reflective, sqp, or others [30]. Finally, once applying this quadrature scheme, Bellman's Principle of Optimality is applied, which states that the optimal trajectory between two nodes also lies on the optimal trajectory for the entire trajectory. This allows the construction of a dynamic optimization problem from many consecutive static optimization problems.

Using this method, several simulations are performed and included as follows. Simulations without obstacles are included Fig. A.3. Simulations with obstacles are presented in 3.4.1 and 3.4.1.

Static Obstacles

For the trajectories generated here, the Lagrangian was chosen such that the minimum-energy solution is obtained. This is done to seek a trajectory for the moving cell that resembles a natural system since the agents in biological systems tend to expend the least amount of energy. With the initial position of the cell and the displacement data obtained from solving the PDE in Eq. (3.6) subject to the IC and BC of choice, DMD can be used to generate an ODE that describes the motion of the cell's x-directional movement. An example ODE resulting from this process is shown in Eq. (3.17).

$$
\begin{aligned}
\underset{y,\dot{y},t}{\text{minimize}} \quad & J = \int_{t_0}^{t_f} \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) dt \\
\text{subject to:} \quad & \dot{x} \approx (5.5 * 10^{-4}) \frac{\mu m}{s} + (5.6 * 10^{-7} \frac{\mu m}{s^2}) t \\
& \qquad - (6.2 * 10^{-11} \frac{\mu m}{s^3}) t^2 \\
& \qquad + (1.6 * 10^{-15} \frac{\mu m}{s^4}) t^3 \\
& \dot{y} \leq 0.0033 \frac{\mu m}{s} \\
& r_{obs,k} \leq \sqrt{((x - x_{obs,k})^2 + (y - y_{obs,k})^2)} \\
& x_1(0s) \approx 309 \mu m \qquad x_1(t_f) \approx 332 \mu m \\
& x_2(0s) \approx 476 \mu m \qquad x_2(t_f) \approx 495 \mu m \\
& t_0 = 0s \qquad t_f = \text{Free}
\end{aligned}
\tag{3.17}
$$

The optimization could be performed in the presence of obstacles, specifically static cells. The simulation for this scenario is shown in Fig. 3.7. While the results presented in Fig. 3.7 partially match the experimental trajectory, the correlation is not perfect. As is the nature with systems experiencing stochastic movement, the migration of any one particular cell can be difficult

to predict exactly. In this case, a cell is shown that has a very large migrational displacement as found from the experiment, whereas the DMD does not predict that it would have been so large. This highlights one of the limitations of such an approach; however, simulations can still be constructed with a high amount of accuracy since most cells would not be outliers.

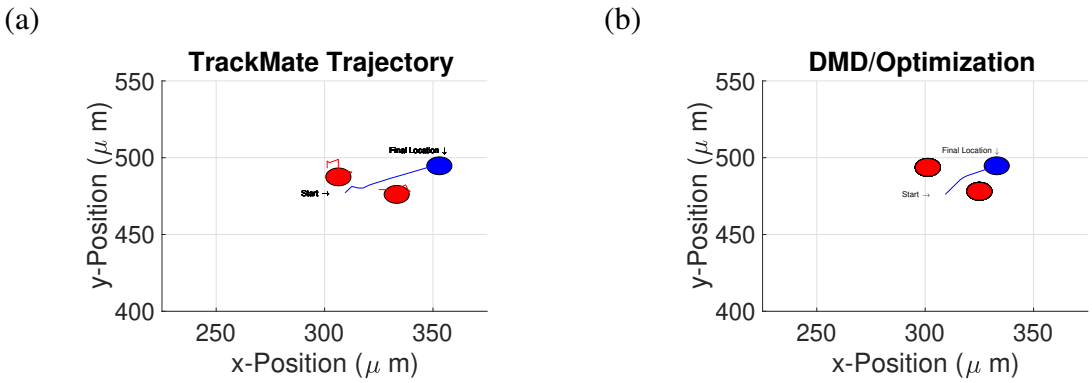


Figure 3.7: (a) Actual trajectory of the cell of consideration, as obtained from TrackMate. (b) The hybrid DMD-optimization model, resulting in an optimal trajectory, with the discovered dynamics from Eq. (3.17).

Dynamic Obstacles

From the experimental data derived cell trajectories, it was evident that some cells barely move while other migrate fast. Those barely mobile cells are treated as dynamic obstacles. The simulations in Fig. 3.7 assume that the obstacles are static, and hence the obstacles don't move through the entirety of the simulation. Using the results from the previous sections, a more realistic simulation is included for comparison to the static case. The other cells are assumed to move according to the PDE discovered from the DMD. The controlled cells move according to an optimization policy, specifically a minimum energy approach. Each cell is allowed to arrive at the goal at their own final time; they do not have to arrive at the goal location at the same time. The results of different cases of dynamic optimizations with single cell or two cells in 2D and 3D settings are shown in Fig. 3.8. Similar simulations have been done [34] with the use of a potential field approach; however, such approaches do not use an optimization-based framework.

3.5 Discussion

This paper demonstrates a new framework for the discovery of the rules of cell migration through the synthesis of a hybrid approach with DMD and optimal control. The mathematical model frames the cell migration in terms of a predictive displacement field rather than population density, which is the approach practiced by most current models. Like many PDEs in physics, the discovered PDE only relies upon a few terms to reasonably model the dynamics. While many terms were allowed to be in the dynamics, only a few terms appear in the result.

The DMD allows for the successful description of the one-dimensional movement of a single cell and multiple cells. This differs from a linear interpolation of the migrational data since it relies upon existing spatio-temporal data and thus cannot predict the global level population behavior. While the governing equation shown in Eq. (3.6) contains a diffusion equation that represents the migration of the cells towards the scratch, a convection term is also present that represents bulk migration in the direction opposite to the scratch, possibly attributed to the bulk retraction of the cell monolayer during the migration. When simulated, the PDE will have the same IC as presented in Eq. (3.7); however, the BC will vary depending on the context of the experiment. Eq. (3.12) can be used when the global position of the cell furthest from the scratch is unknown. Eqs. (3.12) and (3.13) can be used when the cell furthest from the scratch is known to be located at the well wall or in the context of tissue - far from the wound in the context of wound healing where no directed cell movement occurs. Eq. (3.13) and (3.14) can be used when the other methods fail due to limitations in calculating the time derivatives at the beginning of the experiment. It should be noted that there are several limitations and challenges to the application of DMD. While the DMD can be used with a noisy system, such as the small Brownian motion of a cell during migration, the recovered governing equation will not predict such random fluctuations. Another limitation of DMD of this experiment relates to the availability of a fine spatio-temporal grid where the finite differencing is performed. These limitations arise from the number of cells when taking the spatial derivative and the number of time points due to the allowable sampling frequency of a cell system with a microscope relevant for temporal differentiation. Due to these constraints on the system, the

solution of the PDE shown in Fig. 3.3 is imperfect, whereas the recovered system in Fig. 3.1 much more closely resemble the original system which is free from noise. For cell migration systems, constructing a governing equation with this DMD process is not always possible, especially when the sampling frequency of the imaging is too sparse or when a very limited number of cells exist in the field of view.

The DMD technique is independent of the cell type and the only input that it requires is a time lapse imaging, due to the "top-down" nature of the DMD framework. For any clinical or experimental imaging setting, the only requirement for this technique to be successful is knowledge of the trajectory of the organism's cells over time, which comes from the DMD, and it is solely limited by the quality of the experimental data. Besides, the method is agnostic of the cause of migration such as any physical or chemical gradient. Although this work relies on the data from a scratch wound assay where the cells migrate by a chemical gradient, the methodology can be applied to cells moving under any physical gradient or even though cells move without any gradient which is common for aggressive cancer cells. However, for a reliable and successful implementation of this technique, higher resolution in the time points, both experimentally and in the optimization simulation, will be required. Additionally, the method cannot account for the cell division by mitosis process.

Using a hybrid DMD/optimal control-based model for a single cell with nearby obstacles, an optimal path in two-dimensions was developed. The simulated cell in Fig. 3.7 does not move as far in the x-direction as the cell from the experiment due to the large displacement of this particular cell in the experiment; actual cell's movement is abnormally large. A minimum-energy path is still generated with a notably similar trajectory to the experimental cell's trajectory, which demonstrates the strength of this technique.

The optimization was further performed to predict the optimal path in a dynamic environment in two dimensions for one cell or multiple cells. This optimization shows the capability of predicting the optimal path of a cell in two-dimensions with respect to the other cells in the population

that are following the path defined by the DMD. Observing the control policy generated by the optimization, a bang-bang control type is obtained for the cells that are controlled.

The initialization for the states of the optimization is very important for obtaining meaningful results for methods like those presented in this paper. The final solution of the optimization can be sensitive to the initial guess; with a poor guess, the solution can converge to an inferior minimum. For this paper, the cells for which the optimal trajectory was generated were given an initial guess for their trajectory which is formed from a linear interpolation of their initial position to their goal location. Since the cells are modeled as first-order systems here, this is a reasonable assumption. For refinement, h-methods were used to attenuate the error and thereby converge the solution to the optimal trajectory; p-methods were not considered due to the capabilities of numerically integrating with MATLAB's ode45.

The hybrid DMD/optimization method is limited due to a high computational cost. Since the time to complete the simulation on a 2018 MacBook Pro with 16 GB of RAM is approximately one hour and 24 hours for the simulations presented in Fig. 3.8 (e) and (f), respectively, this method cannot be used quickly in real-time in the current state. However, with a more powerful computer it can be possible. Since the computational cost rises significantly with the number of nodes, one possible way to alleviate this long simulation time is to perform a simulation with a fewer number of nodes and then use a method that is not computationally expensive, such as gradient descent with potential functions, in order to guarantee an obstacle avoidance.

A simulation of a moving cell governed by second order dynamics was not included due to the need to quantify the force bounds with which the cells exert themselves on the substrate. Data from this experiment provided velocity only; however, future works will focus on the quantification of cell traction forces and the system might need more simulations for refinement.

Depending on the context, either the single cell migration or collective cell migration can be biologically more relevant, and the framework presented here can be applied accordingly. Single cell migration is more common in leukocyte movement whereas cancer cells move en masse. While the study of both single and multicellular migration can be insightful which is possible with the

present framework, the results suggest that the technique provides more robust and reliable prediction of cell trajectory in multicellular system. The experimental data shows that the migration of the cells is strongly dependent upon the BCs and IC of the system as the biological situation demands. Also, the cells are subject to random microscopic fluctuations, which can make the trajectory of any individual cell challenging; however, the prediction of the overall global pattern of the collective migration can be made.

Practical implementations of such a model exist in the field of therapies that involve magnetic micro-robots, such as drug delivery with a micro-robot [23].Research has been done to develop automated control and path planning of micro-robots [41]; it is expected that the research outlined in this paper will allow for the solving of a similar problem but from a different approach. Many cutting edge technologies for cell therapy and drug delivery employs robots that are externally controlled by electrical and magnetic fields, that include cargo transportation, cell manipulation, toxic substance removal, and micromanipulation [37, 42, 43]. For example, Huang et al reported a novel approach with magnetized cartilage ECM-derived scaffolds for osteoarthritis therapy, utilizing human bone marrow mesenchymal stem cells and decellularized porcine cartilage ECM [44]. Additionally, the use of degradable microrobots has been shown to enable the precise delivery of engineered stem cells for cancer therapy, demonstrating therapeutic efficacy in preclinical tests [45]. Although these technologies are at infancy, the control framework described in this work is needed for their eventual use. The optimization framework presented in this paper will also open possibilities into research related to studying cell migration at the individual level, which can be utilized by systems biologists.

3.6 Conclusion

This work demonstrates a new framework for the discovery of the rules of cell migration through the synthesis of a hybrid approach with Dynamic Mode Decomposition and optimal control. The mathematical model frames the cell migration in terms of a predictive displacement field, rather than on population density, which is the approach practiced by current models. Like many

PDEs in physics, the discovered PDE only relies upon a few terms to reasonably model the dynamics. While many terms were allowed to be in the dynamics, only a few terms appear in the result. The model is utilized to predict cell migration trajectory in cells that migrate alone, and in a complex environment where extracellular matrix obstacles and other cells are present. The framework can be utilized to discover the fundamental physics of cell migration and also for developing actuated microrobots that can be used for targeted cell or drug delivery.

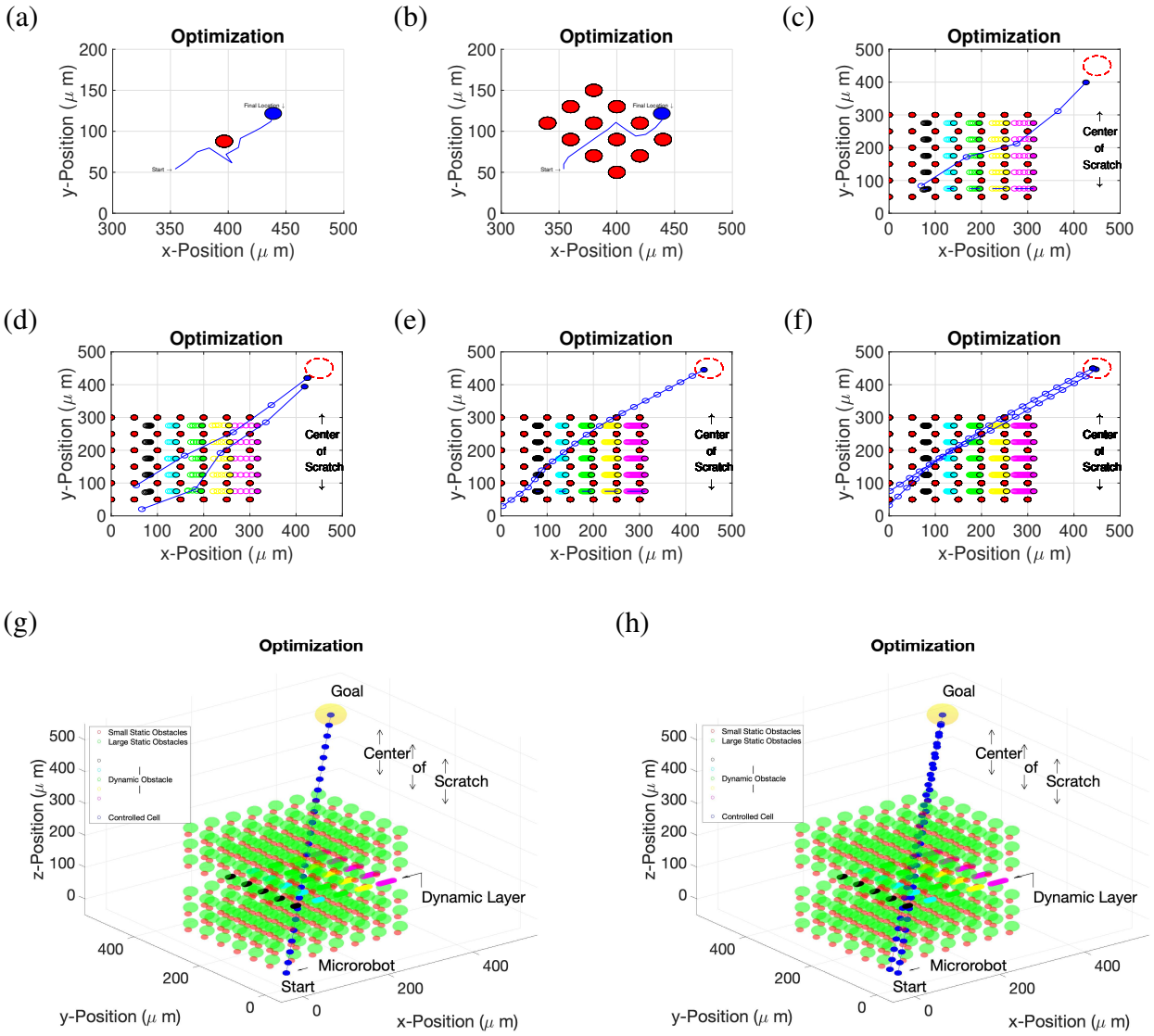


Figure 3.8: (a) An individual cell's trajectory to a prescribed final location. (b) An individual cell's trajectory through an obstacle course of static obstacles. (c), (d), (e) and (f) An individual cell (c and e) or two cells (d and f) traversing through a 2D obstacle course of both moving and non-moving obstacles, where (c) and (d) are simulated with 5 collocation nodes and (e) and (f) are simulated with 20 collocation nodes. (g) an individual cell and (h) two cells traversing through a 3D obstacle course of both moving and non-moving obstacles. For the 2D simulations in (a - f), the red dots represent static obstacles or cells, the blue dot(s) represents the microrobot(s), and all other dots represent moving obstacles whose motion is governed by the DMD. For the 3D simulations in (g - h), the red and green dots represent static obstacles, the blue dot(s) represents the microrobot(s), and all other dots represent moving obstacles whose motion is governed by the DMD.

Chapter 4

Mechanical Analysis of Cell Migration Using Hybrid Structural Optimization

Cell migration requires the dynamic formation and dissolution of mechanical structures inside the cytoplasm. Stress fibers are made of F-actin during cell migration driven by the strategic localization of focal adhesion complexes at the cell-substrate interface. The nucleus is also strategically positioned in the cell during the migration and the stress fibers wrap around the nucleus possibly to carry the nucleus with the cell. Cell migration is energetically demanding and should require strategic utilization of resources such as the F-actin stress fiber formation at specific locations so that they generate enough force by actomyosin contraction at the cell-matrix adhesion sites for a directed movement. In this work we propose a structural optimization based biophysical model to predict the strategic localization and sizes of F-actin fibers that supports the nucleus and the cytoplasm during migration. With the use of a nonlinear controller via a Newton-Euler-based model of the generated design, we further quantified the force in the stress fibers during migration, with results close to those obtained through experimental methods such as traction force microscopy. The predicted force decreases for a cell that migrates slowly due to a pharmacological perturbation. Such quantification of forces only require the information of the trajectory of the cell that can be readily obtained from time lapse microscopy. With novel microscopy techniques emerging, such biophysical model framework can be combined with traction force microscopy data to achieve unprecedented mechanical information inside and outside cells during migration, which is otherwise not possible by experiments only.

4.1 Introduction

Cell migration is a fundamental biological phenomenon. Prokaryotes migrate as a single organism to find resources and in this process the cell-cell communication might be critical if the

cell is the part of a colony. In eukaryotes, cells migrate to achieve a systemic goal such as tissue formation and regeneration. Immune cells migrate to rescue tissues from infection and cancer cells migrate to metastasize from a primary location to a secondary location [4, 31]. In some situations cells migrate alone and based upon the need, cells migrate together showing some specific attributes of collective cell migration where mechanical and chemical communication between cells occur [46,47]. For example, a few leader cells are shown to drive the follower cells in the collective cell migration and the follower cells take turn to be the leaders when the previous leaders run out of the energy [48,49]. Mechanically, cells can communicate with each other via the compliant matrix thus having possible effects in the collective cell behavior [50], such as the speed of individual cell migration. Traction force generated by cells are shown to be different when cells are away from each other compared to when they are closer [51], thus possibly contributing to cell migration by cell-cell mechanical communication.

Structures inside the cell such as F-acin protein in the cytoskeleton and the focal adhesion complex at the cell-matrix interface are critical for the energetically demanding cell migration process. On an individual level, cell movement can be divided into a few stages: protrusion of the leading edge of the cell, adhesion of the leading edge with the matrix (in 3D) or substrate (in 2D), release of the cell from the matrix or the subsrate at the cell body and rear, and cytoskeletal contraction to pull the cell forward [52]. The force from the actin polymerization drives the cell membrane forward - the protrusion stage. Subsequently the polymerized F-actin connects to the substrate at focal adhesion points. The contraction, where the bulk of the trailing cytoplasm is drawn forward, is executed by the protein myosin. Thus, the actomyosin contraction and the resultant cytoskeletal forces are mechanically important factors in cell migration. Experimental image of a migrating cell shows such F-actin cytoskeletal structures, called stress fibers Fig. 4.1. Cytoskeletal forces in the stress fibers are related to the traction force created by the cell at the cell-matrix interface. The quantification of the traction forces during cell migration serve as valuable information for cell mechanosensing and understanding the driving mechanical components to cell migration [26]. Because cell migration is energetically demanding requiring the cells continuously

forming and dissolving structures, it is anticipated that structures are optimally created inside the cell for most efficient utilization of the resources expending least amount of energy.

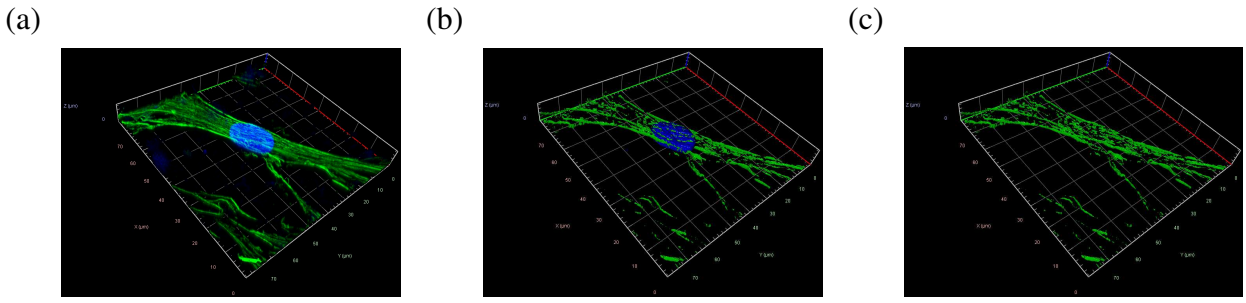


Figure 4.1: Confocal microscopy images of a migrating murine fibroblast (NIH 3T3 cell). (a) 3D reconstruction of the raw image stack showing the actin (green) and nucleus (blue). (b) Segregated image shows the actin stress fiber with the nucleus and (c) showing the actin stress fibers only.

The nucleus also plays a critical role during the cell migration as recent studies indicate [38]. During the migration, the nucleus must be moved with the cell which possibly dictates the location of the centrosome during cell migration [53]. The stress fibers that are produced as a part of the dynamic cytoskeleton serve as the structure in 'carrying' the nucleus during migration. Previous studies have investigated the stress fiber orientation [27, 28], and different predictive models have been created to study the formation of the stress fibers as the cell moves, as well as the forces that these stress fibers exert on the substrate [54–56]. However a critical knowledge gap exists regarding the understanding of cell migration - (1) where the stress fibers are strategically formed that meets the cell's demand to move forward with the nucleus and (2) what forces those stress fibers create to move the cell forward along with the nucleus? Answering these questions utilizing experiments will require high resolution live microscopy combined with stress calculation inside and outside cells which is currently limited. Biophysical model can be utilized to fill this knowledge gap, which is the primary focus of the present study.

4.2 Materials and Methods

4.2.1 Modeling of a 3 $\times$ 3 cell array under resting state using FEM

To quantify the stress distribution in cells a 3 by 3 cell array in two dimensionm was created in FEBio. Tetrahedral elements were used to mesh the combined model of the cytoplasm and the nucleus. The cytoplasm of the cells are made with a neo-Hookean material, with $\rho = 1g/\mu m^3$, $E = 0.5kPa$ and $\nu = 0.35$ for the density, Young's modulus and Poisson's ratio, respectively. A prescribed isotropic active contraction was included in the material property, with a contractile stress value of $T_0 = 0.01kPa$. The nuclei of the system are made with a neo-Hookean material, with $\rho = 1g/\mu m^3$, $E = 1.2kPa$ and $\nu = 0.35$ for the density, Young's modulus and Poisson's ratio, respectively. The substrate of the system is made with a neo-Hookean material, with $\rho = 1g/\mu m^3$, $E = 5kPa$ and $\nu = 0.35$ for the density, Young's modulus and Poisson's ratio, respectively.

4.2.2 Methodology to quantify the distribution of cytoskeletal material in cell

Modern optimization theory has revolutionized various engineering disciplines by enhancing capabilities in tasks such as generative design of structural systems for improved efficiency, such as minimizing mass while maintaining structural integrity, and facilitating dynamic path planning for enhanced system performance. Utilizing calculus of variations, both static and dynamic optimization problems are commonly formulated. These problems can be addressed through different methods, namely indirect, direct, and hybrid approaches. Indirect methods aim to establish and solve the necessary and sufficient conditions analytically, which is suitable for simpler systems but becomes complex rapidly. Hybrid methods, on the other hand, numerically solve these conditions, often employing shooting methods. Direct methods eschew formulating the necessary and sufficient conditions, instead opting to discretize the system and employ minimization algorithms for solution-seeking. In this paper's optimization implementation, direct methods are employed.

To begin the optimization, the problem will first need to be posed within a specific framework. Eq. 5.1 shows an example of how fully dynamic optimization problems are typically posed [29]. The cost J to be minimized is the Bolza cost that includes the Mayer term, $\Phi(\cdot)$, and the Lagrange term, $\mathrm{L}(\cdot)$, which is subjected to constraints on the system, where $\boldsymbol{\varepsilon}(\cdot)$, $\boldsymbol{\beta}(\cdot)$, and $\boldsymbol{\Pi}(\cdot)$ are the vectors that represent the equations of equilibrium, also referred to as the dynamics, boundary constraints, and path constraints, respectively. Contrast $\boldsymbol{\varepsilon}$ with $\boldsymbol{\epsilon}$, where the latter represents the strain tensor.

$$\begin{aligned}
\underset{\boldsymbol{\Delta}=\{\mathbf{S},\mathbf{F}\}^T}{\text{minimize}} \quad & J = \Phi(\mathbf{U}(\mathbf{X},t_1),t_1,\mathbf{U}(\mathbf{X},t_f),t_f) \\
& \qquad + \int_{t_1}^{t_f} L(\mathbf{U}(\mathbf{X},t),\dot{\mathbf{U}}(\mathbf{X},t),\mathbf{F}(t),t)dt \\
\text{subject to:} \quad & \text{EoE} \quad \ddot{\mathbf{U}}(\mathbf{X},t) - \boldsymbol{\varepsilon}(\mathbf{U},\mathbf{U}',\mathbf{U}'',\boldsymbol{\Delta},\mathbf{E},t) = \mathbf{0} \\
& \text{BCs:} \quad \boldsymbol{\beta}(t_1,\mathbf{U}(t_1)),t_f,\mathbf{U}(t_f)) = \mathbf{0} \\
& \text{PC:} \quad \boldsymbol{\Pi}(\mathbf{U}(\mathbf{X},t),\mathbf{F},t) = \mathbf{0}
\end{aligned} \tag{4.1}$$

In Eq. 5.1 for the case of the 1D stress fibers the vector $\mathbf{U}(\mathbf{X},t) \in \mathbb{R}^{2\times 1}$ is the vector representing the displacement field of the system, as shown in Eq. 4.2, where $\mathbf{X}$ is the vector containing the global coordinates of the system, i.e. $\mathbf{X} = \{X,Y,Z\}^T$. Contrast the $\mathbf{U}$ with $\mathbf{u}^{(p)} \in \mathbb{R}^{2\times 1}$ which represents the local displacement of the pth stress fiber of an element in the finite element method (FEM) - based construction. The vector $\mathbf{x}^{(p)} \in \mathbb{R}^{2\times 1}$ represents the spatial coordinates expressed in the local coordinate system of the pth stress fiber.

$$\mathbf{U}(\mathbf{X},t) = \mathbf{X}(t) - \mathbf{X}(t_0) \tag{4.2}$$

The vector $\mathbf{F} \in \mathbb{R}^{2\times 1}$ represents is the open-loop control, which in this context is a distributed load that causes the contractile or tensile stress fiber forces from the chemical reaction of the nucleotide adenosine triphosphate (ATP). The vector $\mathbf{f}^{(p)} \in \mathbb{R}^{2\times 1}$ is reserved for the distributed load across

the p^{th} stress fiber. $\mathbf{S} \in \mathbb{R}^{2p\times 1}$ is the vector containing a certain set of states of the system; for this paper, these states are the cross-sectional areas and lengths of the stress fibers, grouped as defined in Eq. 4.3.

$$\mathbf{S} = \{\mathbf{A} \dots \boldsymbol{\ell}\}^T \tag{4.3}$$

$\mathbf{A} \in \mathbb{R}^{p\times 1}$ is the vector that contains all p cross-sectional areas of the stress fibers. The vector $\boldsymbol{\ell} \in \mathbb{R}^{p\times 1}$ contains the lengths of the stress fibers. The vector $\mathbf{E} \in \mathbb{R}^{p\times 1}$ represents the modulus of elasticity of all p stress fibers. In this paper, the cost in Eq. 5.1 will be proportional to the collective mass of the stress fibers, specifically the volume of the stress fibers, due to the finite resources from the biological limitations of the cell. The volume is formulated as the product of the length and cross-sectional areas of the stress fibers. The vector $\boldsymbol{\Delta}$ represents the parameters of the system that can be adjusted in order to minimize the value of J, in the literature called *decision variables*, which are formulated depending on the context of the problem.

4.2.3 Static Structural Optimization

The equations of equilibrium in Eq. 5.1 take the form of a partial differential equation (PDE), specifically in a structure similar to the wave equation, however for some systems, such as the ones presented here in Eq. 4.4 and later in the Newton-Euler model, the equation of equilibrium is an ordinary differential equation (ODE), rather than a PDE. The full equation of motion for an axial bar is given in Eq. 4.4, where A, E, and ρ represent the cross-sectional area, modulus of elasticity, and mass-density of the stress fibers; f represents a distributed load on the stress fiber. This equation establishes the relationship that the forces in the stress fibers have on the remaining decision variables.

$$-\frac{\partial}{\partial x}(AE\frac{\partial u}{\partial x}) + \rho A\frac{\partial^2 u}{\partial t^2} = f \tag{4.4}$$

For statically-modeled systems, the time-dependent terms, such as $\ddot{\mathbf{U}}$ or $\ddot{\mathbf{u}}$, will not appear in the static formulations, however the spatially-dependent terms will still appear. Static structural optimization problems are commonly posed in some variation of the way shown in Eq. 5.1, where

the Lagrange term is left null and the Mayer term is commonly chosen as a function that represents the compliance, so as to maximize the stiffness, commonly to minimize the displacement of the structure.

One-Dimensional Elements

The cellular system of consideration with stress fibers consists of one-dimensional elements, no cytoplasm contribution, and modeling the nucleus as a point mass is illustrated in Fig. 4.3, where both the protruding and non-protruding stances are considered.

The optimization problem is posed as shown in Eq. 4.5, where the objective is to minimize the collective mass of the stress fibers. Note that the lengths of the stress fibers are not included in $\mathbf{S}$ since this is a static optimization problem and hence the lengths are fixed. It is assumed that the same value for the maximum allowable stress is the same in each stress fiber. The governing differential equation for the displacement of an axial bar is given in the constraints. Buckling constraints are also considered in this model, where the inequality constrains the axial load in the bar to be less than the critical buckling load. Note that the lowest modeshape of buckling is considered due to the fact that the fibroblasts move very slowly and as such the higher frequency buckling modes will not be expected at any point. The boundary conditions (BCs) are Dirichlet and defined such that the displacements at the nodes located on the ECM are zero.

$$\begin{aligned} \underset{\mathbf{\Delta}=\mathbf{A}}{\text{minimize}} \quad & J = \boldsymbol{\ell}^T \mathbf{A} \\ \text{subject to:} \quad & f^{(p)} = -\frac{d}{dx}(A^{(p)} E^{(\mathrm{p})} \frac{du^{(\mathrm{p})}}{dx}) \\ & |\sigma^{(\mathrm{p})}| \leq \sigma^{(p_{allow})} \\ & f^{(p)} \leq \pi^2 \frac{E^{(p)} I^{(p)}}{(L^{(\mathrm{p})})^2} \\ & \text{BCs} \end{aligned} \tag{4.5}$$

Three-Dimensional Elements

For a higher-dimensional model, particularly in three dimensions, the structural optimization problem can be posed in Eq. 4.6, where J is the cost function to be minimized and m is the mass of the system. The governing equations of equilibrium, stress-strain relations, and strain-displacement relations that act as a constraint on the system are expressed in indicial form, where i, j, and k correspond to the x, y, and z components. For a planar model, i = 1,2 and j = 1,2. For a three-dimensional model, i = 1,2,3 and j = 1,2,3.

$$\begin{aligned} \underset{\mathbf{\Delta}=m}{\text{minimize}} \quad & J = m^2 \\ \text{subject to:} \quad & \sigma_{ij,j} + f_i = 0 \\ & \sigma_{ij} = C_{ijkl}\epsilon_{kl} \\ & \epsilon = \frac{1}{2}(u_{i,j} + u_{j,i}) \\ & \sigma \leq \sigma_{allowable} \\ & \text{Euler Buckling} \\ & \text{BCs} \end{aligned} \tag{4.6}$$

4.2.4 FEM-Based Transcription Method

The formulations presented in Eq. 5.1, and hence Eqs. 4.5 and 4.6, are continuous-time formulations, which means that they will have to be converted into an approximate discrete formulation for the direct collocation through transcription into a nonlinear program, similarly to the process in [57]. The transcription will result in a system of equations that take the structure of Eq. 4.7, where $\mathcal{M}$, $\mathcal{K}$, $\boldsymbol{\mathcal{G}}$, and $\boldsymbol{\mathcal{F}}$ represent the mass and stiffness matrices, respectively, and the gravitational force and other force vectors, respectively.

$$\mathcal{M}\ddot{\mathbf{U}} + \mathcal{K}\mathbf{U} + \boldsymbol{\mathcal{G}} = \boldsymbol{\mathcal{F}} \tag{4.7}$$

Using this method, several solutions to the structural optimization are obtained and included in this paper. Here, the structural optimization was solved in MATLAB with the use of 'fmincon'.

1D FEM-Based Transcription Method

Only the intermediate region of the cell that is bounded by the nucleus and the cell membrane, and not the nucleus was optimized. The connection of the stress fibers (SF) to the extra-cellular matrix (ECM) and the nucleus is modeled with a pin-joint connection. The connection of the stress fibers to the nucleus is modeled at the center of the nucleus for simplicity. The goal was to minimize the total mass of the stress fibers. This optimization formulation is chosen because the cell has finite resources with which to construct the actin stress fibers, hence we should expect the cell to seek the usage of the minimum mass possible for the construction of each stress fiber.

The equations in Eq. 4.5 were re-expressed, through transcription, before the optimal solution was obtained with MATLAB's fmincon. This process involves the converting of the ODE into an approximately equivalent set of algebraic equations, which can be done with several methods; here, the finite element method is used. For a small system like this, the stiffness matrix can be easily obtained and assembled ; the building block for each stress fiber's contribution is given in Eq. 5.23. It was assumed that the modulus of elasticity will be the same for all of the stress fibers. Note that this matrix is formulated from the use of linear shape functions. Complete details of the finite element-based transcription are reported later 4.2.5.

The reduced matrix, which is obtained upon applying the boundary conditions of zero displacements at all of the nodes located at the substrate side, was used in Eq. 4.8. The buckling criteria was re-expressed in terms of the displacements and cross-sectional areas. The stress constraints was re-expressed in terms of the displacements. Note that $|\sigma_p| \leq \sigma_{allow}$ is mathematically the same as $-\sigma_{allow} \leq \sigma_p \leq \sigma_{allow}$; the latter formation was used so that the stress constraint could be inserted into fmincon. Values for the material properties, particularly the breaking force, cross-sectional radius, and the modulus of elasticity of the stress fibers were assumed as 377 nN, 100 nm, 1.45 MPa, respectively [58]. Therefore, the ultimate tensile strength of the stress fibers was assumed to be 12 MPa. The stress fiber length could vary in length, from 10 to 100 μm;

typical stress fiber lengths are around 100 μm [59]. Also, the stress fiber length was $\geq 20\mu m$ [28]. The cross-sectional area of the stress fibers can range from 0.001 μm^2 to 0.04 μm^2, whereas the focal adhesion cross-sectional areas can vary from approximately 0.75 μm^2 to approximately 4.5 μm^2 [59]. Values of the traction forces during cell motion range from roughly 0 nN to roughly 120 nN [26]; 120 nN was used in the simulation. The optimization problem posed in Eq. 4.5 was converted into a roughly equivalent optimization problem, shown in Eq. 4.8.

$$
\begin{aligned}
&\underset{\boldsymbol{\Delta}=\mathbf{A}}{\text{minimize}} \quad J = \boldsymbol{\ell}^T \mathbf{A} \\
&\text{subject to:} \quad \begin{Bmatrix} F_a \\ 0 \end{Bmatrix} = \sum_{p=1}^{7} \begin{bmatrix} \frac{A^{(p)}E^{(p)}}{L^{(p)}}(c^{(p)})^2 & \frac{A^{(p)}E^{(p)}}{L^{(p)}}(c^{(p)})^2(s^{(p)})^2 \\ \frac{A^{(p)}E^{(p)}}{L^{(p)}}(c^{(p)})^2(s^{(p)})^2 & -\frac{A^{(p)}E^{(p)}}{L^{(p)}}(c^{(p)})^2 \end{bmatrix} \begin{Bmatrix} u_{8x} \\ u_{8y} \end{Bmatrix} \\
&\qquad c_p u_{8x} + s_p u_{8y} \leq \frac{\pi A_p}{4L_p} \\
&\qquad -\sigma_{p_{allow}} \leq \frac{E_p}{L_p}(c_p u_{7x} + s_p u_{7y}) \leq \sigma_{p_{allow}}
\end{aligned}
\tag{4.8}
$$

4.2.5 FEM Formulation - 1D Element

A numerical solution was obtained with the use of the finite element method. The weak form was obtained as shown in Eq. 5.10, where $\zeta(x)$ represents the weighting function. Note that the superscript denoting the p^{th} element is not shown; the finite element formulation is the same for each stress fiber, so the superscript here is omitted.

$$
0 = \int_{x_p}^{x_q} \zeta(x)(-\frac{\partial}{\partial x}(AE\frac{\partial u}{\partial x}) + \rho A\frac{\partial^2 u}{\partial t^2} - f)dx \tag{4.9}
$$

Regardless of the order of the time derivative, there is no integration by parts in time. The final expression for the semi-discrete weak form is shown in Eq. 4.10.

$$
0 = \int_{x_p}^{x_q} (AE\frac{\partial \zeta}{\partial x}\frac{\partial u}{\partial x} + \rho A\zeta\frac{\partial^2 u}{\partial t^2} - \zeta f)dx - \Big[-\zeta(AE\frac{\partial \zeta}{\partial x})\Big]_{x_p}^{x_q} \tag{4.10}
$$

The displacements and weighted residuals are defined in terms of the shape functions as shown in Eqs. 4.11 and 4.12.

$$\zeta = \Psi_i \qquad\qquad u = \sum_{j=1}^{2} u_j \Psi_j \tag{4.11}$$

$$\Psi_1 = (\frac{x_q - x}{h}) \qquad\qquad \Psi_2 = (\frac{x - x_p}{h}) \tag{4.12}$$

Inserting these equations into the former expression for the semi-discrete weak form will result in the system of ODEs in Eq. 4.13.

$$\Big[m\Big]\Big\{\ddot{u}\Big\} + \Big[k\Big]\Big\{u\Big\} = \Big\{F\Big\} \tag{4.13}$$

The values of the entries of the mass matrix, stiffness matrix, and force vector are given in Eqs. 4.14-4.16.

$$m_{ij} = \int_{x_p}^{x_q} \rho A \Psi_i \Psi_j dx \tag{4.14}$$

$$k_{ij} = \int_{x_p}^{x_q} AE \frac{d\Psi_i}{dx} \frac{d\Psi_j}{dx} dx \tag{4.15}$$

$$F_i = \underbrace{-\Big[\Psi_i (AE \frac{d\Psi_i}{dx})\Big]_{x_p}^{x_q}}_{P_i} + \int_{x_p}^{x_q} \Psi_i f dx \tag{4.16}$$

The system was converted using the parent-coordinate formulation in order to formulate the mass matrix. This is shown for the mass matrix in Eqs. 4.17 and 4.18.

$$m_{11} = \int_0^h \rho A (1 - \frac{s}{h})(1 - \frac{s}{h}) ds = \frac{\rho A h}{3} \tag{4.17}$$

$$m_{12} = \int_0^h \rho A (1 - \frac{s}{h})(\frac{s}{h}) ds = \frac{\rho A h}{6} \tag{4.18}$$

Assembling this results in the mass matrix for a single element as shown in Eq. 4.19.

$$m_{ij} = \frac{\rho A h}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{4.19}$$

4.2.6 FEM Formulation - Rotated Element

For the purposes of the combined stress fiber/nucleus system, the finite element formulations were expressed in the global X, Y, Z system, as opposed to their local coordinate systems, in order to formulate the global mass and stiffness matrices as well as the global gravitational and force vectors. Performing the coordinate system transformation resulted in the system of equations for the static case shown in Eq. 4.20, where m and n represent the left and right nodes on the one-dimensional finite element, respectively. The stiffness matrix of the rotated bar element is k_p and $c_p = cos(\theta_p)$, $s_p = sin(\theta_p)$.

$$\begin{Bmatrix} F_{mx} \\ F_{my} \\ F_{nx} \\ F_{ny} \end{Bmatrix} = \underbrace{\frac{A^{(p)}E^{(p)}}{L^{(p)}} \begin{bmatrix} (c^{(p)})^2 & c^{(p)}s^{(p)} & -(c^{(p)})^2 & -c^{(p)}s^{(p)} \\ c^{(p)}s^{(p)} & (s^{(p)})^2 & -c^{(p)}s^{(p)} & -(s^{(p)})^2 \\ -(c^{(p)})^2 & -c^{(p)}s^{(p)} & (c^{(p)})^2 & c^{(p)}s^{(p)} \\ -c^{(p)}s^{(p)} & -(s^{(p)})^2 & c^{(p)}s^{(p)} & (s^{(p)})^2 \end{bmatrix}}_{k^{(p)}} \begin{Bmatrix} u_{mx} \\ u_{my} \\ u_{nx} \\ u_{ny} \end{Bmatrix} \tag{4.20}$$

4.2.7 FEM Formulation - Combined System

The finite element formulations for each element are used to assemble the reduced global system of equations in a succinct manner that arise from the application of the boundary conditions. The matrix $\mathcal{M}$ in Eq. 4.21 is known as the reduced mass matrix of the discretized system, which is full rank and invertible and obtained after applying the boundary conditions. This is contrasted with the global mass matrix, which is not full rank and hence not invertible.

$$\mathcal{M} = \begin{bmatrix} m_n + \sum_{p=1}^{7} m_{22}^{(p)} cos^2(\theta_p) & \sum_{p=1}^{7} m_{22}^{(p)} cos(\theta_p) sin(\theta_p) \\ \sum_{p=1}^{7} m_{22}^{(p)} cos(\theta_p) sin(\theta_p) & m_n + \sum_{p=1}^{7} m_{22}^{(p)} sin^2(\theta_p) \end{bmatrix} \tag{4.21}$$

The matrix $\mathcal{K}$ in Eq. 4.22 is known as the reduced stiffness matrix of the discretized system, which like the mass matrix, is full rank, invertible, and obtained after applying the boundary con-

ditions. Like the mass matrix, this can be seen in contrast with the global stiffness matrix, which is not full rank and not invertible.

$$\mathcal{K} = \begin{bmatrix} \sum_{p=1}^{7} k_{22}^{(p)} cos^2(\theta_p) & \sum_{p=1}^{7} k_{22}^{(p)} cos(\theta_p) sin(\theta_p) \\ \sum_{p=1}^{7} k_{22}^{(p)} cos(\theta_p) sin(\theta_p) & \sum_{p=1}^{7} k_{22}^{(p)} sin^2(\theta_p) \end{bmatrix} \tag{4.22}$$

The vector $\boldsymbol{\mathcal{G}}$ in Eq. 4.23 represents the contribution due to gravitational forces that arise when considering gravity in the vertical direction with the nucleus' being modeled as a point mass at the end of the stress fibers, where m_n represents the mass of the nucleus.

$$\boldsymbol{\mathcal{G}} = \begin{Bmatrix} 0 \\ m_n g \end{Bmatrix} \tag{4.23}$$

Lastly, the $\boldsymbol{\mathcal{F}}$ in Eq. 4.24 represents the forces produced by the stress fibers that support the nucleus and carry the nucleus along the desired trajectory.

$$\boldsymbol{\mathcal{F}} = \begin{Bmatrix} \sum_{p=1}^{7} cos(\theta^{(p)}) \Big(f^{(p)} x_q^{(p)} + \dfrac{f^{(p)}((x_p^{(p)})^2 - (x_q^{(p)})^2)}{2h^{(p)}} \Big) \\ \sum_{p=1}^{7} sin(\theta^{(p)}) \Big(- f^{(p)} x_p^{(p)} + \dfrac{f^{(p)}((x_q^{(p)})^2 - (x_p^{(p)})^2)}{2h^{(p)}} \Big) \end{Bmatrix} \tag{4.24}$$

3D FEM-Based Transcription Method

The optimization problem posed in Eq. 4.6 can be converted into an approximately equivalent form, by expressing the constraint equations as a system of algebraic equations; one possible way to do this is through the finite element method. By applying the boundary conditions, the reduced matrix system, with corresponding stiffness matrix K, can be obtained. Subsequently, the optimization problem can be posed in Eq. 4.25, where $\mathbf{U}$, $\mathbf{V}$, and represent the global horizontal

displacement, vertical displacements, and volume of the system, respectively.

$$\begin{aligned} \underset{\mathbf{\Delta}=}{\text{minimize}} \quad & J = m^2 \\ \text{subject to:} \quad & K \begin{Bmatrix} \{\mathbf{U}\} \\ \{\mathbf{V}\} \end{Bmatrix} = \begin{Bmatrix} F \end{Bmatrix} \\ & \sigma \leq \sigma_{allowable} \\ & \text{Euler Buckling} \\ & \text{BCs} \end{aligned} \tag{4.25}$$

4.3 Results

4.3.1 Finite element method predicts significantly different stress distribution in the cell based on cell-substrate contact model

Mechanical stress and strain in the cell can be modeled utilizing the Finite Element Method based techniques. Previous works modeled such mechanical quantities in a resting cell utilizing a material model in the cell and the nucleus, along with a suitable boundary condition such as the traction forces [60]. One approach is to define a traction force field as a continuous field variable at the cell surface. A second approach is to define the traction force at individual points where the cell matrix adhesion is formed. Traction force is the result of cell contractile force which is not common in modeling cell mechanics but can capture more fundamental aspects of the physical forces in cells compared to the traction force. In the present work we applied a contractile stress inside cells to model the resting mechanics inside the cells using FEBio, a nonlinear finite element analysis software created specifically for biomechanics and biophysics applications [61]. The goal is to investigate the stress distribution inside cells.

In our model we incorporated the provisions to investigate the stress and strains inside the cell, the nucleus and the substrate in the three-dimensional setting. Nine cells were arranged in a 3 $\times$ 3 array with the distance between cells being a possible parameter that can be changed. For

visualization purpose only a planar slice is shown 4.2. In the first case (a), continuous surface contact was assumed and in the second case (b), specific contact areas were modeled utilizing 'feet' under the cells to mimic the focal adhesion, similar to the model considered in [62]. The results reveal that when the "feet" are included, the stress is concentrated largely at the end regions of the cell, with smaller amount of stress in the middle, as opposed to the continuous distribution observed when the cell makes a continuous flush contact with the substrate.

Multiple simulations were performed, where the contractile stress T_0 inside the cell was perturbed in order to investigate the effect of different contractile stress values on the stress magnitude field shown in the color plot in Fig. 4.2. Such parametric study changes the magnitude of the stress field inside the cell, but it does not change the shape of the stress field, therefore results from multiple simulations are not included since the purpose of this particular model is to look at the shape of the stress distribution that arises from the cell's contraction during migration, not the magnitude. Overall this results indicate that only specific areas of the cell experience an elevated amount of stress and they are close to the outer periphery of the cell whereas inner locations undergo lower amount of stress. The F-actin tread milling is a mechanical stress dependent process therefore the stress fibers are most likely formed at those adhesion points near the cell periphery and connecting the cell across with fiber formations as shown in the images from the experimental data in Fig. 4.1.

(a) (b)

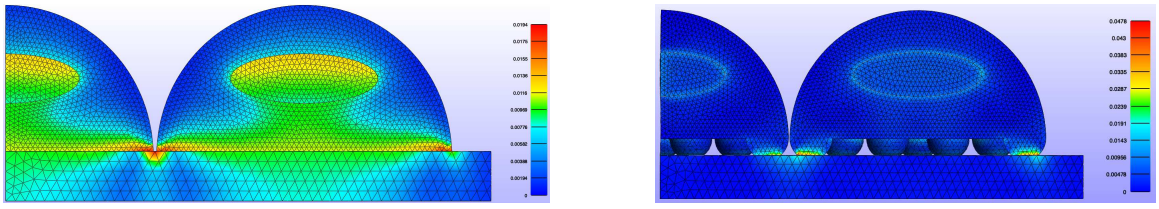


Figure 4.2: Distribution of 'stress magnitude' inside the cell, nucleus and substrate under resting state modeled with cell contractility inside the cytoplasm (a) with the cell modeled under perfect contact with the substrate and (b) with focal adhesion points that contact the substrate at specific locations marked by the 'feet'.

4.3.2 Stress fiber thickness is optimally defined by cells at strategic locations during migration as revealed by a planar cell model

A handful of stress fibers are formed during cell migration that build the mechanical structure to carry the cell forward as shown in Fig. 4.1. The nucleus is connected to the cytoplasm via the F-actin fibers and the nucleus also moves with the cell in a coordinated manner [63]. Therefore, it is expected that the F-actin fibers carry the nucleus. This leads us to the question - how only a few stress fibers are mechanically able to carry the nucleus. This is logical though because the F-actin fiber formation is energetically demanding and cells will try to optimize the stress fiber formation at strategic locations and numbers with a goal to carry the cell and the nucleus forward.

Table 4.1: The values of the parameters of the simplified model in Fig. 4.3.

	$\mathbf{SF_p}$	$\boldsymbol{\theta_p}$	$\mathbf{L_p}$ (μm)	$\mathbf{E_p}$ (MPa)	$\mathbf{A_p^*}$ ($\mu\mathbf{m^2}$)	$\mathbf{E_p}$ (MPa)	$\mathbf{A_p^*}$ ($\mu\mathbf{m^2}$)	$\mathbf{E_p}$ (MPa)	$\mathbf{A_p^*}$ ($\mu\mathbf{m^2}$)
Symmetric									
	1	15°	≈ 100.5		$4.5 * 10^6$		$4.5 * 10^6$		$4.5 * 10^6$
	2	40°	≈ 40.4		$0.75 * 10^6$		$0.75 * 10^6$		$0.75 * 10^6$
	3	65°	≈ 28.7	|	$0.75 * 10^6$	|	$0.75 * 10^6$	|	$0.75 * 10^6$
	4	90°	26	0.5	$3.160 * 10^6$	1.45	$3.14 * 10^6$	2.4	$3.135 * 10^6$
	5	115°	≈ 28.7	|	$0.75 * 10^6$	|	$0.75 * 10^6$	|	$0.75 * 10^6$
	6	140°	≈ 40.4		$0.75 * 10^6$		$0.75 * 10^6$		$0.75 * 10^6$
	7	165°	≈ 100.5		$0.45 * 10^6$		$4.5 * 10^6$		$4.5 * 10^6$
Protruding									
	1	90°	26		$3.143 * 10^6$		$3.142 * 10^6$		$3.140 * 10^6$
	2	102.5°	≈ 26.6		$0.75 * 10^6$		$0.75 * 10^6$		$0.75 * 10^6$
	3	115°	≈ 28.7	|	$0.75 * 10^6$	|	$0.75 * 10^6$	|	$0.75 * 10^6$
	4	127°	≈ 32.8	0.5	$0.75 * 10^6$	1.45	$0.75 * 10^6$	2.4	$0.75 * 10^6$
	5	140°	≈ 40.4	|	$0.75 * 10^6$	|	$0.75 * 10^6$	|	$0.75 * 10^6$
	6	152°	≈ 56.3		$0.45 * 10^6$		$4.5 * 10^6$		$4.491 * 10^6$
	7	165°	≈ 100.5		$0.45 * 10^6$		$4.5 * 10^6$		$4.4499 * 10^6$

Therefore we attempted to predict the thick stress fiber formations at specific focal adhesion site locations by solving an optimization problem that seeks to minimize the collective mass of the stress fibers. The model of the cell is shown in Fig. 4.3 where the model is similar to some used in the literature [62], specifically with the connections that the stress fibers make with the

substrate and the nucleus being modeled as pin joints. The top panel shows the configuration of the cell at resting (a) and protruding states (b). The thin lines represent the possible locations of the stress fibers. Subsequently the optimization problem was posed and executed to calculate the requirement and thickness of the fibers, sufficient to carry the nucleus subjected to a force F_a. in this case, F_a is chosen to be 120 nN, the magnitude of the force was derived from the literature [26]. The numerical values of the lengths of the stress fibers and their orientations are given in Table 4.1.

MATLAB's 'fmincon' was used in order to optimize the system under such mechanical loading demand which provides the optimal cross-sectional areas of the stress fibers. The values of the areas from the optimization, as well as the parameters of the model, are presented in Table 4.1. Note that the $(\cdot)^*$ represents the optimal solution. After thresholding, we get that $A^{(2)} \approx A^{(3)} \approx A^{(5)} \approx A^{(6} \approx 0$, which leaves just the contribution from stress fibers 1,4, and 7. The resulting system is shown in Fig. 4.3 c and d. The thicknesses of the fibers have been scaled for visualization purposes. It should be noted that F_a can be varied within the neighborhood of this value, however the distribution of the three principal stress fibers remain the same, suggesting that an exact value of this force is not critical. Overall this results agree with the experimental observation that only a few thick stress fibers are formed in the cell to carry the nucleus forward. A three dimensional counterpart of this model is also framed and presented in Fig. A.4 in the appendix.

4.3.3 Prediction of traction forces from the cell trajectory

Dynamic model to predict the cell trajectory

The results presented so far were obtained by applying a static structural optimization on the cell's cytoskeleton. However, a more sophisticated model is desired that accounts for a dynamic system by utilizing the trajectory of the nucleus. The trajectory was used to quantify the position, velocity, and acceleration of the nucleus, so dynamic equations were generated using a technique we previously described [57], that utilizes the dynamic mode decomposition (DMD) method. Equations for such trajectories that represent position, velocity and acceleration are presented below in Eq. 4.26. The cell (or nucleus) trajectories were derived from our previous experimental

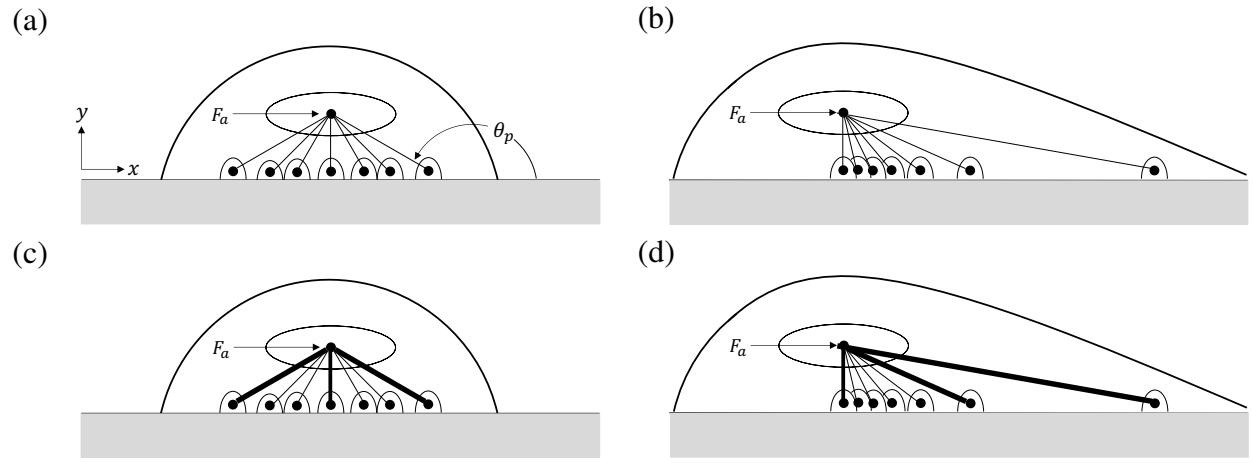


Figure 4.3: (a) Model of the simplified cell with the possible locations of stress fibers and associated focal adhesion points shown for (a) the symmetric case and (b) the protruding case. The cell with the stress fibers, with thicknesses shown to scale after the static structural optimization, and associated focal adhesion points presented in (c) before protrusion and (d) during protrusion. For complete description of stress fiber areas after optimization refer to Table 4.1.

work [38], where murine embryonic fibroblast line NIH 3T3 cells were visualized during migration in a scratch wound assay. We utilized the data from an experimental perturbation where a drug Trichostatin-A lowers the cell migration velocity by inhibiting the chromatin condensation [38].

$$\text{Desired Trajectory} \begin{cases} \text{Non-treated} \begin{cases} q_{8_d}(t) = 7E^{-6} + 4.859E^{-9}t - 3.400E^{-13}t^2 + 1.348E^{-17}t^3 - 2.092E^{-22}t^4 \\ \dot{q}_{8_d}(t) = 5E^{-9} - 7E^{-13}t + 4E^{-17}t^2 - 8E^{-22}t^3 \\ \ddot{q}_{8_d}(t) = -6.759E^{-13} + 8.051E^{-17}t - 25.054E^{-22}t^2 \end{cases} \\ \text{Drug-Treated} \begin{cases} q_{8_d}(t) = 7E^{-6} + E^{-8}(-0.1440t + 5.9793E^{-5}t^2 - 3.7387E^{-9}t^3 + 6.9E^{-14}t^4) \\ \dot{q}_{8_d}(t) = E^{-8}(-0.1397 + 1.1959E^{-4}t + -1.1216E^{-8}t^2 + 2.76E^{-13}t^3) \\ \ddot{q}_{8_d}(t) = E^{-8}(1.1959E^{-4} - 2.2432E^{-8}t + 8.28E^{-13}t^2) \end{cases} \end{cases} \tag{4.26}$$

In order to illustrate the resemblance between the experimental data obtained from ImageJ's TrackMate [39] and the resulting DMD process, the trajectories of both for the control and drug-

treated (TSA treated) cells are shown in Fig. 4.4. Procedures towards developing a desired trajectory that will be used in the control selection can take multiple approaches; here, the DMD-based trajectory is used for it's higher C^k continuity, where here k represents the k^{th} derivative, compared to the experimental trajectory. A smoother desired trajectory function helps in eliminating any jerky movement of the system as well as it allows the easier formulation of the desired kinematic trajectories.

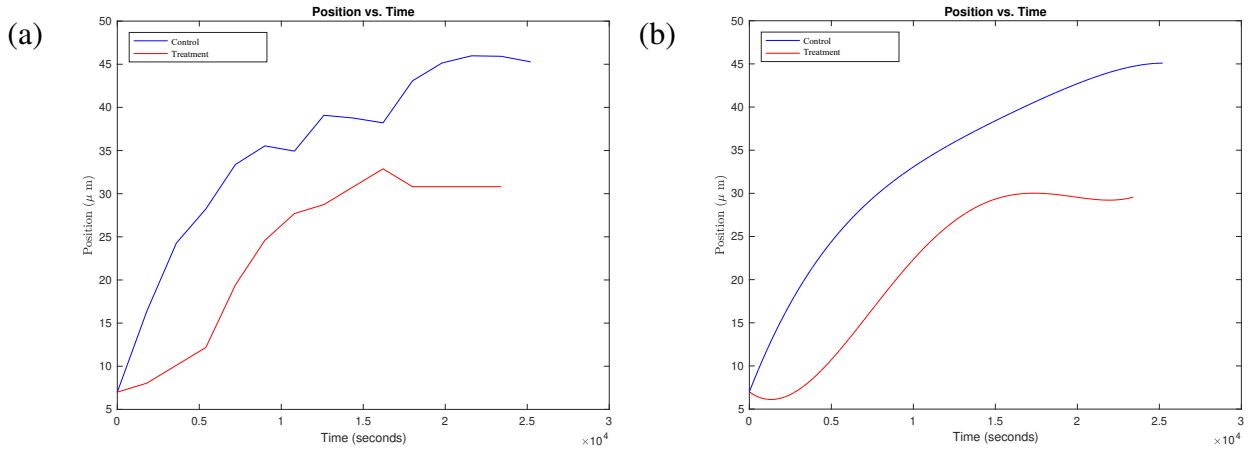


Figure 4.4: (a) Plot of the experimental data where the starting x-values of both plots have been adjusted to start at 7 μm for trajectory comparison and (b) the plot of the approximated trajectory obtained through dynamic mode decomposition based on the experimental data.

Newton-Euler hybrid-based model to represent the stress fibers as robotic elements

One approach towards obtaining a dynamic model is to construct a multi-link Newton-Euler based "robotic-like" system as a model for the cell. The benefit of such an approach includes the ability to apply well-established control theory methods to this system. The "robotic-like" model of the cell is shown in Fig. 4.5. The coordinate directions are defined by the inertial reference frame $\mathbf{N}$, where the orthonormal unit vectors that define the basis of the *right-handed* spatial coordinate system are $\hat{\mathbf{N}}_1$, $\hat{\mathbf{N}}_2$, and $\hat{\mathbf{N}}_3$.

The dimensions of all of the lengths of the rigid bodies in this robotic representation of the cell, as well as all of the initial values of the generalized coordinates, are shown in Table 4.2, where L_n represents the length of the n^{th} element and q_n represents the n^{th} generalized coordinate. The size

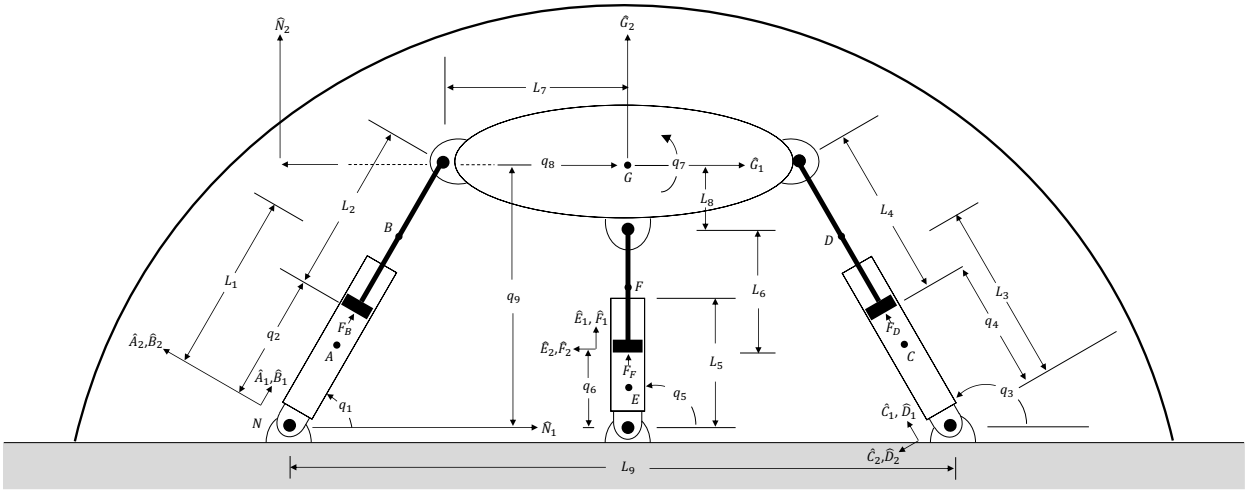


Figure 4.5: Coordinate system diagram of the "robotic-like" model of the cell with three actuator-based stress fibers. The connections to the actuators at both the substrate and nucleus are modeled as pin joints. The generalized coordinates representing translational or rotational position are represented by $q_1 ... q_9$. The body-attached frames for each rigid body are $\mathbf{A} = [\hat{A}_1, \hat{A}_2, \hat{A}_3,]$, ... , $\mathbf{G} = [\hat{G}_1, \hat{G}_2, \hat{G}_3,]$, where $\hat{A}_1$,...,$\hat{G}$ represent the orthogonal set of unit vectors defining each frame. Relevant lengths of the rigid bodies are given as L_1,...,L_9. All bodies, including the substrate, are assumed to be completely rigid.

of the nucleus of a fibroblast ranges between 500 μm^3 to 1000 μm^3 [64]; the size of the nucleus in this work is assumed to be 513.13 μm^3. The mass of the nucleus is assumed as 2.29 ng [65]; the mass of the cell is assumed to be 22.9 ng.

Table 4.2: The parameters of the three-link "robotic" model of the cell shown in Fig. 4.5, at its initial states.

white **n**	$\mathbf{L_n}$	$\mathbf{q_n(t_0)}$
1	20 μm	$90°$
2	20 μm	6 μm
3	20 μm	$\approx 150.58°$
4	20 μm	$\approx$ 32.926 μm
5	20 μm	$\approx 134.90°$
6	20 μm	$\approx$ 12.759 μm
7	7 μm	$0°$
8	2.5 μm	7 μm
9	60 μm	26 μm

Several classical mechanics methods exist that can produce the governing differential equations of motion, such as with the Newton-Euler equations, Lagrange's method, D'Alembert's principle,

Kane's method, etc. Here, Kane's method was used with the help of Autolev motion analysis software [66] in order to generate the equations of motion. The governing equations for the system are represented in joint space in Eq. 4.27, and are subject to the kinematic constraints on the system obtained from the development of kinematic chains shown in Eq. 4.28.

$$A(\mathbf{Q})\ddot{\mathbf{Q}} + \mathbf{B}(\mathbf{Q}, \dot{\mathbf{Q}}) + \mathbf{G}(\mathbf{Q}) = \mathbf{\Gamma} = \begin{bmatrix} 0 \\ G_T^T \end{bmatrix} \mathbf{\Upsilon} \tag{4.27}$$

$$\mathbf{c}(\mathbf{Q}) = \mathbf{0} \tag{4.28}$$

In the equations above, $\mathbf{Q}$, $\dot{\mathbf{Q}}$, and $\ddot{\mathbf{Q}} \in \mathbb{R}^9$ are the vectors containing the generalized coordinates that define the configuration of the system, velocities, and accelerations, respectively, $A(\mathbf{Q}) \in \mathbb{R}^{9\times9}$ is the inertial matrix of the system, $\mathbf{B} \in \mathbb{R}^{9\times1}$ is a vector containing the Coriolis and centrifugal terms, $\mathbf{G} \in \mathbb{R}^{9\times1}$ is the matrix containing the terms due to gravity, $\mathbf{\Gamma} \in \mathbb{R}^{9\times1}$ is the vector containing the actuator force and torque values, $G_T \in \mathbb{R}^{9\times3}$ is the transmission matrix, and $\mathbf{\Upsilon} \in \mathbb{R}^{3\times1}$ is the vector containing only the actuator forces. It is noteworthy that the equations stated above are second-order, nonlinear differential equations that describe the motion of the robotic model of the stress fibers in the cell. The degrees of freedom (DOF), sometimes referred to as mobility, of the system can be determined with the Chebychev–Grübler–Kutzbach criterion in Eq. 4.29, where R, J_1, and J_2 represent the number of rigid bodies, joints of type 1, and joints of type 2, respectively.

$$\text{DOF} = 3(R-1) - 2J_1 - J_2 = 3 \tag{4.29}$$

The kinematic constraints on the system, represented by the vector $\mathbf{c}(\mathbf{Q}) \in \mathbb{R}^{6\times1}$, are formulated here from vector loops or kinematic chains of the configuration.

Nonlinear control of the configuration space equations to quantify the forces in stress fibers

Once the robotic model of the stress fibers were built, we quantified the forces in the stress fibers. To build a classical controller, various methods can be utilized. Here, the equations in Eqs. 4.27 and 4.28 were combined through constraint embedding in Autolev and expressed in a single

set of equations, as represented in Eq. 4.30, where $\mathbf{q}$, $\dot{\mathbf{q}}$, and $\ddot{\mathbf{q}} \in \mathbb{R}^{3\times1}$ are the vectors containing the generalized coordinates, generalized velocity, and generalized acceleration that remain after the constraint embedding. The matrix $\mathbf{a}(\mathbf{q}) \in \mathbb{R}^{3\times3}$ represents the reduced mass matrix, $\mathbf{b}(\mathbf{q}) \in \mathbb{R}^{3\times1}$ is the vector containing the Coriolis and centrifugal terms, $\mathbf{g}(\mathbf{q}) \in \mathbb{R}^{3\times1}$ is the vector containing the gravitational terms, and $g_T^T \in \mathbb{R}^{3\times3}$ is the reduced transmission matrix.

$$\mathbf{a}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = g_T^T \boldsymbol{\Upsilon} \tag{4.30}$$

An operational space controller was developed in order to specify the constraint on the movement of the nucleus. Typically, the Jacobian matrix needs to be formulated in order to convert between textcolorredconvert between does not sound right - use some other phrase the joint space and the operational space, if the coordinate of consideration is rotational. However the generalized coordinate representing the nucleus is already expressed in the translational coordinate system, so no Jacobian is needed. The computed torque method was used in order to develop a controller that drives the movement of the system. We selected the following control in order to linearize the dynamics in Eq. 4.31,

$$\boldsymbol{\Upsilon} = (g_T^T)^{-1}\Big(\mathbf{b}(\boldsymbol{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) + a(\mathbf{q})\boldsymbol{\Upsilon}^*\Big) \tag{4.31}$$

where:

$$\boldsymbol{\Upsilon}^* = \ddot{\mathbf{q}}_d + K_v(\dot{\mathbf{q}}_d - \dot{\mathbf{q}}) + K_p(\mathbf{q}_d - \mathbf{q}_d) \tag{4.32}$$

which gives the resulting error dynamics are shown in Eq. 4.33, where K_p and K_v are the proportional and derivative gains.

$$(\ddot{\mathbf{q}}_d - \ddot{\mathbf{q}}) + K_v(\dot{\mathbf{q}}_d - \dot{\mathbf{q}}) + K_p(\mathbf{q}_d - \mathbf{q}_d) = \mathbf{0} \tag{4.33}$$

The values of the gains K_p and K_v are selected at the discretion of the user. These gain values will determine if the error dynamics behave in an overdamped, critically damped, or underdamped fashion. The gain selection in Eq. 4.34 guarantees that the the error dynamics will behave in

a critically damped manner; the proportional gain is selected to be $K_p = 1$. This is a common heuristic model of control that doesn't necessarily guarantee optimal solution.

$$K_v = 2\sqrt{K_p} \tag{4.34}$$

It can be noted that this type of controller can be demonstrated to asymptotically converge the system to the desired trajectory with the use of a Lyupanov-based analysis, even with a variety of gains. To show this, a Lyupanov candidate function for this system is proposed in Eq. 4.35, where E and $\dot{E}$ represent the error, defined as $E \triangleq q_d - q$ and time derivative of the error, defined as $\dot{E} \triangleq \dot{q}_d - \dot{q}$, respectively.

$$V = \frac{1}{2}\dot{E}^T\dot{E} + \frac{1}{2}E^T K_p E \tag{4.35}$$

Note that the gain matrix, K_p, is a positive-definite matrix. The candidate function V satisfies the criteria for a Lyupanov function, namely V is zero *iff* E and $\dot{E}$ are zero and V is greater than zero *iff* E and $\dot{E}$ are not zero. Differentiating Eq. 4.35 gives Eq. 4.36.

$$\dot{V} = \ddot{E}^T\dot{E} + \dot{E}^T K_p E = -\dot{E}K_v\dot{E} \leq 0 \tag{4.36}$$

The expression in Eq. 4.36 is zero only when E and $\dot{E}$ are zero, and negative when E and $\dot{E}$ are nonzero, therefore asymptotic convergence of the system to the desired position is guaranteed with the use of any positive-definite gain matrix K_v.

The biological insight received from this model is stated below and in Fig. 4.6. Trajectory of the control cell (that moves faster) and the TSA treated cell (that moves slower) are shown in Figure 6 (a) and (c) respectively. The cells were tracked for almost 7 hours. It is clear that the control cell moved a larger distance along the x axis (straight line) over the same total tracking time. The y component is zero and rotation (θ) is also zero because the nucleus was moving in an approximated straight line. The nonlinear controller is able to converge the position (black line) to the desired trajectory (purple line) in both cases, which is consistent with the Lyupanov-based prediction. The

force in the individual controllers (stress fibers) change over time but not drastically, which is expected. The distance travelled by the cells in this situation is fairly small - between 30 microns (TSA) to 47 (control) microns which is almost same as the length of cell. Therefore, if we consider a full cycle of cell migration consisting protrusion, contraction and stabilization, the force in individual fibers won't change much during this timeframe. The comparison of the values of forces in the stress fiber compare remarkably closely with the experimental data derived from traction force microscopy, which reports the traction force in the range of tens of nanonewtons [26].

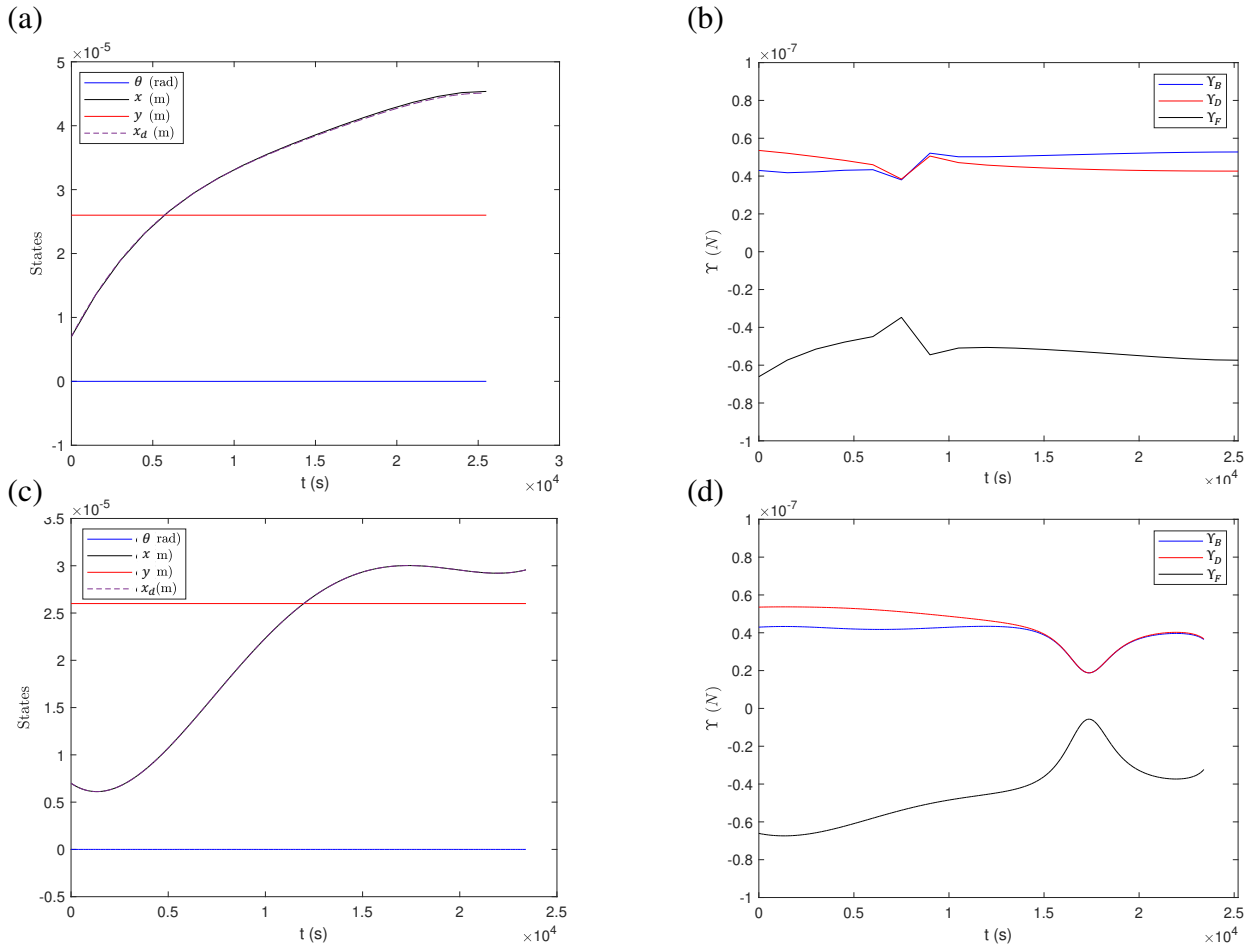


Figure 4.6: (a) Trajectory of the states of the control cell. (b) Actuator forces from the nonlinear controller of the system for the control cell, where the forces in the left, middle, and right actuators start at 42.97 nN, -66.10 nN, and 53.56 nN; the actuator forces, from left to right, finish at 52.75 nN, -57.41 nN, and 42.62 nN. (c) Trajectories of the states cell with TSA treatment. (d) Actuator forces from the nonlinear controller of the system with TSA treatment, where the forces in the left, middle, and right actuators start at 42.97 nN, -66.10 nN, and 53.56 nN; the actuator forces, from left to right, finish at 36.45 nN, -32.32 nN, and 36.76 nN.

4.4 Discussion

The goal of this research is twofold: to exploit the use of modern optimization techniques in order to predict the orientation and size of F-actin stress fibers networks within the cytoskeleton and to quantify the forces in the stress fibers during cell migration. The migration event was posed as an optimization problem that allows for the cell to achieve a desired motion in a particular direction, subject to a finite actin mass supply constraint. The assembly of the cell's internal organization at this level during migration is currently unavailable due to the limitations of even the most cutting edge microscopy. Computationally derived quantification of the stress in fibers, that can provide the information of the traction forces is also achieved using the present framework. Only the cell trajectory is the input in this framework which is experimentally obtainable by time lapse microscopy. This can be especially useful when cells are moving in clusters as in those situations quantification of the traction force at the single cell level will require many assumptions. Therefore, to investigate the cell-to-cell communication via mechanosensing in a dense environment we can use the current framework which only relies on the trajectory tracking, which is much easier using time lapse microscopy. However, it should be noted that the method provided in this study is not meant to replace the traction force microscopy, rather to complement that technique to provide the force information both inside and outside the cell.

The optimal location and size of the stress fibers were revealed in this study as shown in Figs. 4.2 and 4.3. Although the study was done on a vertical cross-section of the cell (4.3) to quantify the size of the stress fiber, it still captures the requirement of the cell to create only a few thick stress fibers which matches the experimental observation. A more detailed three-dimensional model with the same technical framework as shown in Fig. A.4 also captures the same biological insight. Additionally, the location of stress fibers are concentrated at the periphery of cells which is captured by the high stress concentration in specific locations as shown in Fig. 4.2.

The quantification of stress fiber forces not only resemble the previously described experimental data, it also shows that the control cells which move faster has a higher value of stress fiber force compared to the TSA treated cell which moves slower. The computational framework only

relied on the cell trajectory that was derived from the experiment and the Dynamic Mode Decomposition. The benefit of such an analytical biophysical model as compared to experimental models (traction force microscopy), is that the parameters of the model can be adjusted based on the cell type to obtain the forces in stress fibers.

The framework presented here also exists for dynamic structural optimization problems, however current models are commonly centered around dynamic optimization from a frequency perspective through the augmentation of a frequency constraint [67], such as the one in Eq. 5.7, where K, ω_r, and M represent the global stiffness matrix, r^{th} eigenvalue, and global stiffness matrix, that imposes an eigenvalue problem-type constraint, in place of the equation of equilibrium in Eq. 5.1 [68]. The solution more closely resembles that of a static optimization problem solved in this paper, since the geometry is not designed to change over time. Further improvement of this framework will allow us to predict the redistribution of the cytoskeletal material during the dynamic process of cell migration

$$\Big([K] - \omega_r^2[M]\Big)\mathbf{U}_r = \mathbf{0} \tag{4.37}$$

The results presented in this paper set the framework for simulating more sophisticated cell migration behavior. Future work will focus on using the complete equations of equilibrium, as opposed to the reduced static equations, which will allow for a more robust model of cell migration due to the dynamic nature and more elaborate model. This paper also lays the groundwork for allowing more sophisticated three-dimensional model of cytoskeletal material distribution, as presented in Fig. A.4. Future work can also add additional constraints to the robotic Newton-Euler model of the cell, specifically by adding frictional or adhesion constraints at the stress fiber attachment points, rather than assuming that these points remain in contact through the entire motion.

4.5 Conclusions

The outcome of this paper demonstrates a methodology for the creation of a biophysical model that estimates the stress fiber forces and hence it can insinuate the traction forces during cell migration. Additionally it predicts the position and size of the cell's stress fibers during this process. Therefore, with the power of structural optimization, nonlinear control, and the finite element method, a novel dynamic model of cell migration is achieved, with expected impacts in the field of biomechanics, cell migration and applied optimization sciences.

Chapter 5

A Mechanical Analysis of Cell Locomotion - a Dynamic Structural Optimization Approach

The goal of this research is to exploit the use of modern structural optimization to suggest the shape of a biological cell's cytoskeletal actin stress fiber orientation and focal adhesion sites during locomotion that allows the cell to achieve maximum motility in its desired migrational direction. Such a biophysical model would allow the prediction of cellular motion during migration and ambulation, which is otherwise unobtainable due to the limitations of modern microscopy. The proposed model also allows quantification of the traction forces during migration, which is a crucial part of cell-to-cell communication via mechanosensing.

5.1 Introduction

Collective migration plays an important role in the movement of many biological systems, from bacteria to large animals [5]. The migration of cells is crucial for many biological processes, including morphogenesis during embryonic development, tissue and skeletal cell generation, fibroblasts during wound healing, and the movement of immune cells during a response to infection [4, 31]. Mechanical forces regulate a wide range of these biological activities, and the field of mechanobiology seeks to understand the mechanisms by which cells sense and respond to mechanical signals. A couple of large mechanical factors of collective cell migration that affect cell-to-cell communication include the sensing of the propagated stress through a shared compliant extra-cellular matrix (ECM) (in 3D) or substrate (in 2D) [50] and the proximity of the neighboring cells to one another [51].

Internal to the cell, the cytoskeletal structure plays an important role in cell locomotion due to the force generation that is created to drive the cell in the direction of its motion caused by the chemical reaction of the nucleotide adenosine triphosphate (ATP). An illustration of the internal

components of a simplified 2D cell is given in Fig. 5.1, where Ω_n, Γ_n, Ω_c, Γ_c, Ω_{sf}, and Γ_{sf} represent the domain and the bounding surface of the nucleus, cytoplasm, and stress fibers, respectively. These internal components play an important role in cell locomotion during migration, which can

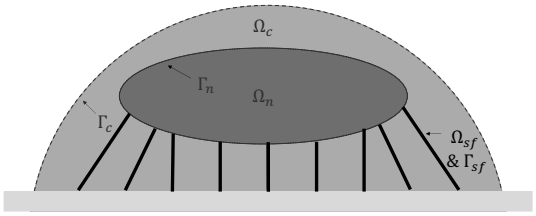


Figure 5.1: The continuous simplified 2-D model of the cell.

be divided into a few general stages: protrusion of the leading edge of the cell, adhesion of the leading edge, deadhesion at the cell body and rear, and cytoskeletal contraction to pull the cell forward [52]. The force from the actin polymerization drives the plasma membrane forward creating the lamellipodium, called protrusion. Tension builds that moves the cell forward. A protein called actin in the cytoskeleton connects to the substrate at focal adhesional points, serving as a scaffold to support the inertial components and cell shape. Myosin, a motor protein, attaches to the filamentous actin (F-actin) and draws the bulk of the trailing cytoplasm during the contraction stage, although myosin can create both tensile and contractile forces in stress fibers [69].

An in-depth study of the internal mechanics of the cell will allow the quantification of the traction forces generated by the stress fibers, which in turn play an important role in collective cell migration via mechanical communication, called mechanosensing. A method known as traction force microscopy seeks to experimentally determine the traction forces during migration by observing the displacement field of beads embedded in an ECM or substrate with known material properties during migration and later calculating the stress fields through a known constitutive law. Experimental methods are limited largely due to sampling thresholds of the microscope in order to avoid greatly disturbing the cell system. One way to combat these limitations are with the use of analytical mathematical methods to predict stress fiber orientation, such as in [27, 28]. Mathematical methods, particularly with the use of optimal control, are commonly used to solve a wide

variety of difficult problems, including applications in aerospace engineering and robotics; they have also been used in biological applications, such as regulating the spread of a disease within a population, as in [22], where the authors used decision variables like social distancing and masks to limit the spread of the COVID-19 virus while also minimizing economic impact. Limited research has been done on applying this optimization framework to study structurally dynamic systems, particularly those capable of *dynamic generative design*, where the geometry is capable of changing over time, which in this case is made possible from a cell's growing limbs as it ambulates. This lack of research is possibly due to the recent development of the field of optimal control theory and limited applications to a system capable of dynamic generative design, however the novel approach presented in this paper could lead to models seeking to study cell migration and morphogenesis in more complex environments.

5.2 Methods

Eq. (5.1) shows an example of how dynamic optimization problems are typically posed [29]. The cost J to be minimized is the Bolza cost that includes the Mayer term, $\Phi(\cdot)$, and the Lagrange term, $\mathrm{L}(\cdot)$, which is subjected to constraints on the system, where $\boldsymbol{\varepsilon}(\cdot)$, $\boldsymbol{\beta}(\cdot)$, and $\boldsymbol{\Pi}(\cdot)$ are the vectors that represent the equations of equilibrium, also referred to as the dynamics, boundary constraints, and path constraints, respectively. Contrast $\boldsymbol{\varepsilon}$ with $\boldsymbol{\epsilon}$, where the latter represents the strain tensor. The selection of J is context-dependent. Previous studies have chosen to minimize energy [70] and stress fiber volume [71] for cell migrational problems. For the study presented here, the volume of the stress fibers is minimized due to finite resources of the cell F-actin and myosin motor proteins; this is the same logic as in [71].

$$
\begin{aligned}
\min_{\boldsymbol{\Delta}=\{\mathbf{S},\mathbf{F}\}^T} \quad & J = \Phi(\mathbf{U}(\mathbf{X},t_1),t_1,\mathbf{U}(\mathbf{X},t_f),t_f) \\
& \quad + \int_{t_1}^{t_f} L(\mathbf{U}(\mathbf{X},t),\dot{\mathbf{U}}(\mathbf{X},t),\mathbf{F}(t),t)dt \\
\text{subject to:} \quad & \text{EoE:} \quad \ddot{\mathbf{U}}(\mathbf{X},t) - \boldsymbol{\varepsilon}(\mathbf{U},\mathbf{U}',\mathbf{U}'',\boldsymbol{\Delta},\mathbf{E},t) = \mathbf{0} \\
& \text{BC:} \quad \boldsymbol{\beta}(t_1,\mathbf{U}(t_1)),t_f,\mathbf{U}(t_f)) = \mathbf{0} \\
& \text{PC:} \quad \boldsymbol{\Pi}(\mathbf{U}(\mathbf{X},t),\mathbf{F},t) = \mathbf{0}
\end{aligned}
\tag{5.1}
$$

The simplified 2D model presented in Fig. 5.1 is further simplified to the model shown in Fig. 5.2 and all variables presented in Eq. (5.1) will be related in terms of the variables in this diagram. In the research presented in this paper, the cell's stress fibers are modeled as 1D bar elements and are not formed at the nuclear envelope; instead, the nucleus is modeled as a point-mass with the stress fibers connecting to the nucleus at this point. The components of the focal adhesion that help to link the stress fiber and the substrate are well documented [16] and are composed of an integrin node with ligand receptor bonding [12]. Simplified models are considered, such as in [27], which models the focal adhesions with multiple connected springs. The model presented here is similar to others used in the literature, such as in [62]; specifically, the stress fibers are simply connected to the substrate and nucleus with a pin-joint connection.

In Eq. 5.1 for the case of the 1D stress fibers the vector $\mathbf{U}(\mathbf{X},t) \in \mathbb{R}^{2\times 1}$ is the vector representing the global displacement field of the system, as defined in Eq. (5.2), where $\mathbf{X}$ is the vector containing the global coordinates of the system, i.e. $\mathbf{X} = \{X,Y,Z\}^T$.

$$
\mathbf{U}(\mathbf{X},t) = \mathbf{X}(t) - \mathbf{X}(t_0) \tag{5.2}
$$

The displacement field of the e^{th} element expressed in the global coordinates is shown in Eq. (5.3).

$$
\mathbf{U}^{(e)}(\mathbf{X}^{(e)},t) = \mathbf{X}^{(e)}(t) - \mathbf{X}^{(e)}(t_0) \tag{5.3}
$$

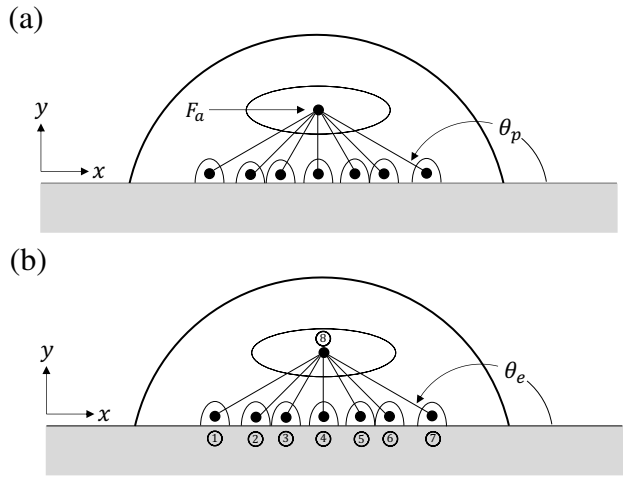


Figure 5.2: (a) Simplified model of the cell composed of 1D elements and (b) The global nodes of each element.

Contrast these displacement expressions with the displacement of the e^{th} stress fiber, expressed in the local coordinate systems, denoted as $\mathbf{u}^{(e)} \in \mathbb{R}^{2\times1}$, is defined in Eq.(5.4), where the vector $\mathbf{x}^{(e)} \in \mathbb{R}^{2\times1}$ represents the spatial coordinates expressed in the local coordinate system of the e^{th} stress fiber.

$$\mathbf{u}^{(e)}(\mathbf{x}^{(e)}, t) = \mathbf{x}^{(e)}(t) - \mathbf{x}^{(e)}(t_0) \tag{5.4}$$

Additionally, $u_n^{(e)} \in \mathbb{R}^{1\times1}$ represents the displacement of the n^{th} node of the e^{th} element. In the context of this paper, linear shape functions will be used to approximate the displacement field; as such, the stress fibers will contain only two nodes. This relationship is defined in Eq. (5.5).

$$\mathbf{u}_n^{(e)} = \begin{Bmatrix} u_1^{(e)} \\ u_2^{(e)} \end{Bmatrix} \tag{5.5}$$

The vector $\mathbf{F} \in \mathbb{R}^{2\times1}$ represents the open-loop control, which in this context is a distributed load that causes the contractile or tensile fiber forces from the reaction of ATP. The vector $\mathbf{f}^{(e)} \in \mathbb{R}^{2\times1}$ is reserved for the distributed load across the e^{th} stress fiber. $\mathbf{S} \in \mathbb{R}^{p\times1}$ is the vector containing a certain set of states of the system that are allowed to be changed in the optimization design process; for this paper, these states are the cross-sectional areas of the stress fibers, $\mathbf{A} \in \mathbb{R}^{e\times1}$. The

vectors $\boldsymbol{\ell} \in \mathbb{R}^{e\times 1}$ and $\mathbf{E} \in \mathbb{R}^{e\times 1}$ contain the lengths and elastic modulus of the stress fibers, respectively. In this paper, the cost in Eq. 5.1 will be related to the collective mass of the stress fibers, specifically the volume of the stress fibers, due to the finite resources from the biological limitations of the cell, where the volume is formulated as the product of the length and cross-sectional areas of the stress fibers. The vector $\boldsymbol{\Delta}$, in the literature called *decision variables*, represents the parameters of the system that can be adjusted to minimize the value of J, which are formulated depending on the context of the problem; here, the decision variables are the cross-sectional areas and forces generated by the stress fibers. The lengths are geometrically constrained while the nucleus moves its prescribed path; hence, they are not design variables.

The equations of equilibrium given in Eq. (5.1) are written in terms of partial differential equations (PDE). Note: For other systems, such as those governed by ordinary differential equations, the dynamics will take the form of one or multiple ordinary differential equations (ODE). Shorthand notation of $(\cdot)'$, $(\cdot)''$, $\dot{(\cdot)}$, $\ddot{(\cdot)}$ represent the first and second-order spatial and temporal partial derivatives, respectively. Also, note that for statically modeled systems the time-dependent terms, *i.e.* $\ddot{\mathbf{U}}$, will not appear in the static formulations; however, the spatially-dependent terms will appear.

5.2.1 Static Structural Optimization

Static structural optimization problems are commonly posed in some variation as shown in Eq. (5.6), where the Mayer term is commonly chosen as a function that represents the compliance (to maximize the stiffness) or as a function to minimize the displacement of the structure; the Lagrange term is left null due to the static nature of the problem [67]. The stress bounds are included such that the stress, $\sigma^{(e)}$, of any individual stress fiber does not exceed the material property threshold for tension and compression. Euler buckling constraints are included since the stress fibers are

allowed to be in compression, where $I^{(e)}$ represents the moment of inertia of the stress fiber.

$$\begin{aligned} \min_{\mathbf{A}} \quad & J = \boldsymbol{\ell}^T \mathbf{A} \\ \text{subject to:} \quad & f^{(p)} = -\frac{d}{dx}(A^{(p)} E^{(p)} \frac{du^{(p)}}{dx}) \\ & |\sigma^{(p)}| \leq \sigma^{(p_{allow})} \\ & F^{(p)} \leq \pi^2 \frac{E^{(p)} I^{(p)}}{(L^{(p)})^2} \\ & \text{BCs} \end{aligned} \tag{5.6}$$

5.2.2 Dynamic Structural Optimization

Similar formulations exist for so-called dynamic structural optimization problems; however, current models are commonly centered around dynamic optimization from a cyclical response or frequency perspective through the augmentation of a frequency constraint, such as the one in Eq. (5.7), where K, ω_i, and M represent the global stiffness matrix, i^{th} eigenvalue, and global mass matrix, that imposes an eigenvalue problem-type constraint [68].

$$\Big([K] - \omega_i^2 [M]\Big)\mathbf{U}_i = \mathbf{0} \tag{5.7}$$

The approach towards using Eq. (5.7) more closely resembles that of a static structural optimization problem, since the geometry is not designed to change over time. In other words, this approach does not include the ability for a dynamic generative design of the structure, such as the changing cytoskeletal structure of a cell during migration. The dynamic structural optimization problem solved in this paper is formulated in Eq. (5.8), where in place of a dynamic frequency constraint, the *true* dynamics can be used, where the model dynamic equation is used where A, E, and ρ represent the cross-sectional area, modulus of elasticity, and mass-density of the stress fibers; f represents a distributed load on the stress fiber.

$$
\begin{aligned}
\min_{\mathbf{A}} \quad & J = \int_{t_0}^{t_f} \boldsymbol{\ell}^T \mathbf{A} dt \\
\text{subject to:} \quad & -\frac{\partial}{\partial x}(A^{(e)} E^{(e)} \frac{\partial u^{(p)}}{\partial x}) + \rho^{(e)} A^{(e)} \frac{\partial^2 u^{(e)}}{\partial t^2} = f^{(e)} \\
& |\sigma^{(e)}| \leq \sigma^{(e_{allow})} \\
& F^{(e)} \leq \pi^2 \frac{E^{(e)} I^{(e)}}{(L^{(e)})^2} \\
& \text{Boundary Conditions}
\end{aligned}
\tag{5.8}
$$

5.2.3 Collocation Method

Optimization problems with cost functions similar to the one presented in Eq. (5.6) that are subject to algebraic constraints are commonly solved numerically with solvers like MATLAB's *fmincon* [30]. Note: The formulations presented in Eqs. (5.1), (5.8), and (5.6) contain non-algebraic equations either in the constraints, cost, or both, which means that to solve the optimization problem numerically with a solver like *fmincon*, the problem will be converted into an approximate discrete formulation for the direct collocation through transcription into a nonlinear program. First, the integral in the cost functional to be minimized can be approximated with a quadrature formulation. For example, using the trapezoidal method, the integral given in Eq. (5.1) can be converted into an algebraic expression. The quadrature conversion used in the simulations presented in this paper are given in Eq. (5.9).

$$
\begin{aligned}
&\int_{t_1}^{t_f} L(\mathbf{U}(\mathbf{X},t), \dot{\mathbf{U}}(\mathbf{X},t), \mathbf{F}(t), t) dt \\
&\qquad \approx \sum_{j=1}^{n} (\mathbf{X},t), \dot{\mathbf{U}}(\mathbf{X},t), \mathbf{F}(t), t) \Delta t_j
\end{aligned}
\tag{5.9}
$$

The dynamics will also be discretized, which can be done with a finite differencing scheme such as Euler forward. MATLAB's *fmincon* function will solve the resulting system of algebraic equations; hence, it will solve the static optimization problems using algorithms such as interior-point, trust-region-reflective, sqp, or others. Finally, once applying this quadrature scheme, we can apply Bellman's Principle of Optimality, which states that the optimal trajectory between two

nodes also lies on the optimal trajectory for the entire trajectory; this allows the construction of a dynamic optimization problem from many consecutive static optimization problems. Using this method, a solution to the optimization problem can be obtained.

5.2.4 FEM Transcription

The transcription process for the optimization problem presented in Eq. (5.8) will first involve the conversion of the governing equation of equilibrium, here a PDE, into an approximately equivalent set of ODEs. This spatial discretization can be done with several methods; here, the finite element method will be used and shown. Next, the resulting ODEs can be temporally discretized into a roughly equivalent system of algebraic equations; this conversion is commonly done with the finite difference method; however, for better results, MATLAB's *ode45* will be used to numerically integrate the equations that will serve as the dynamic constraints on the system. The finite element system formulation process is shown in this subsection.

Build Element Coefficient Matrices

The equation of equilibrium has already been given in Eq. (5.8) and the cell domain has been discretized into finite elements as shown in Fig. 5.2. The next step is to use the weighted residual method, where $\zeta(x)$ is the weighting function and x refers to the local coordinate of the element of consideration. The weak form of the equilibrium equation is shown in Eq. (5.10), where x_p and x_q represent the beginning and ending nodes of the element, respectively. Note that the superscript $(\,\cdot\,)^{(e)}$ has been omitted from Eq. (5.10-5.18) since the derivation presented is for only one element; starting at Eq. (5.20) the superscript indicating a particular element of the system is reintroduced.

$$0 = \int_{x_p}^{x_q} \zeta(x)(-\frac{\partial}{\partial x}(AE\frac{\partial u}{\partial x}) + \rho A\frac{\partial^2 u}{\partial t^2} - f)dx \tag{5.10}$$

Regardless of the order of the time derivative, there is no integration by parts in time. The final expression for the semi-discrete weak form is shown in Eq. (5.11).

$$0 = \int_{x_p}^{x_q} (AE\frac{\partial \zeta}{\partial x}\frac{\partial u}{\partial x} + \rho A\zeta\frac{\partial^2 u}{\partial t^2} - \zeta f)dx$$

$$-\Big[-\zeta(AE\frac{\partial \zeta}{\partial x})\Big]_{x_p}^{x_q} \tag{5.11}$$

The displacement and weighting functions in Eq. (5.11) are approximated with the shape functions in Eq. (5.13). The shape functions are chosen to be linear shape functions; however, more sophisticated functions could be used if desired.

$$\zeta = \Psi_i \qquad\qquad u = \sum_{j=1}^{2} u_j \Psi_j \tag{5.12}$$

$$\Psi_1 = (\frac{x_q - x}{h}) \qquad\qquad \Psi_2 = (\frac{x - x_p}{h}) \tag{5.13}$$

Using the shape functions in Eq. (5.13), the so-called semi-discrete form of the system can be obtained as a system of ODEs, shown in Eq. (5.14),

$$\Big[m\Big]\Big\{\ddot{u}\Big\} + \Big[k\Big]\Big\{u\Big\} = \Big\{F\Big\} \tag{5.14}$$

where generally the [m] and [k] matrices are called the mass and stiffness matrices, respectively, and { F } refers to the point forces and distributed load on the system. These terms are defined explicitly using *indicial form* in Eqs. (5.15-5.16)

$$m_{ij} = \int_{x_p}^{x_q} \rho A\Psi_i\Psi_j dx \;=\; \frac{\rho Ah}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{5.15}$$

$$k_{ij} = \int_{x_p}^{x_q} AE\frac{d\Psi_i}{dx}\frac{d\Psi_j}{dx}dx \;=\; \frac{AE}{L}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{5.16}$$

$$F_i = \underbrace{-\left[\Psi_i(AE\frac{d\Psi_i}{dx})\right]_{x_p}^{x_q}}_{P_i} + \int_{x_p}^{x_q} \Psi_i f dx \tag{5.17}$$

By evaluating the bounds in the first term of the right-hand side of Eq. (5.17), F_i can be further expressed as in Eq. (5.18). Through a Hookean constitutive law, the functions inside the vector on the right-hand side can be easily seen to be equivalent to any point loads on the system applied at the nodal points. The elements in this vector will be considered in this sense during the formulation of the equilibrium equations while forming the FBD of the system.

$$F_i = \begin{Bmatrix} AE\frac{du}{dx}\Big|_{x_p} \\ -AE\frac{du}{dx}\Big|_{x_q} \end{Bmatrix} + \int_{x_p}^{x_q} \Psi_i f dx \tag{5.18}$$

F_i can be further expressed by evaluating the integral in the above term with the substitution of the shape functions given in Eq. (5.13) into the integrand in Eq. (5.18). The anti-derivative is evaluated and the result is given in Eq. (5.19).

$$F_i = \begin{Bmatrix} AE\frac{du}{dx}\Big|_{x_p} \\ -AE\frac{du}{dx}\Big|_{x_q} \end{Bmatrix} + \begin{Bmatrix} f^{(e)}x_q^{(e)} - \frac{f^{(e)}}{2h^{(e)}}\left[\left(x_q^{(e)}\right)^2 - \left(x_p^{(e)}\right)^2\right] \\ \frac{f^{(e)}}{2h^{(e)}}\left[\left(x_q^{(e)}\right)^2 - \left(x_p^{(e)}\right)^2\right] - fx_p^{(e)} \end{Bmatrix} \tag{5.19}$$

Lastly, the system of equations in Eq. (5.14) will need to be transformed from their local coordinate formulation into the global coordinate system counterparts. To do this transformation, the transformation matrix, T, is first defined in Eq. (5.20), which relates a vector, in this case the displacement expressed in the local coordinate system, in terms of the displacement in the global coordinate system, where $U^{(e)}$ and $V^{(e)}$ represent the displacements in the global X and

Y-directions, respectively.

$$\underbrace{\begin{bmatrix} U_1^{(e)} \\ V_1^{(e)} \\ U_2^{(e)} \\ V_2^{(e)} \end{bmatrix}}_{\mathbf{U}^{(e)}} = \underbrace{\begin{bmatrix} cos(\theta^{(e)}) & 0 \\ sin(\theta^{(e)}) & 0 \\ 0 & cos(\theta^{(e)}) \\ 0 & sin(\theta^{(e)}) \end{bmatrix}}_{T^T} \underbrace{\begin{bmatrix} u_1^{(e)} \\ u_2^{(e)} \end{bmatrix}}_{\mathbf{u}^{(e)}} \tag{5.20}$$

Eq. (5.20) can be rearranged to isolate the local displacements. The result is shown in Eqs. (5.21).

$$\underbrace{\begin{bmatrix} cos(\theta^{(e)}) & sin(\theta^{(e)}) & 0 & 0 \\ 0 & 0 & cos(\theta^{(e)}) & sin(\theta^{(e)}) \end{bmatrix}}_{T} \begin{bmatrix} U_1^{(e)} \\ V_1^{(e)} \\ U_2^{(e)} \\ V_2^{(e)} \end{bmatrix} = \begin{bmatrix} u_1^{(e)} \\ u_2^{(e)} \end{bmatrix} \tag{5.21}$$

Note: The same transformation can be used for the other vectors in this system of equations, particularly with the force and acceleration vectors, to express them in their global vector counterparts. This transformation is used to re-express Eq. (5.14) into the form shown in Eq. (5.22).

$$\underbrace{T^T\,[m]\,T}_{M}\,\ddot{\mathbf{u}}_N + \underbrace{T^T\,[k]\,T}_{K}\,\mathbf{u}_N = \underbrace{T^T\,\mathbf{f}_i^{(e)}}_{\mathbf{F}_i^{(e)}} \tag{5.22}$$

Finally, the building block for each stress fiber's stiffness contribution expressed in the global coordinate system is given in Eq. (5.23), where $c^{(e)} = cos(\theta^{(e)})$, $s^{(e)} = sin(\theta^{(e)})$, and $k^{(e)}$ is the stiffness matrix of the rotated bar element. It will be assumed that the modulus of elasticity will be the same for all the stress fibers. The global mass matrix for the e^{th} stress fiber, $M^{(e)}$, takes a similar structure.

$$K^{(e)} = \underbrace{\frac{A^{(e)}E^{(e)}}{L^{(e)}} \begin{bmatrix} (c^{(e)})^2 & c^{(e)}s^{(e)} & -(c^{(e)})^2 & -c^{(e)}s^{(e)} \\ c^{(e)}s^{(e)} & (s^{(e)})^2 & -c^{(e)}s^{(e)} & -(s^{(e)})^2 \\ -(c^{(e)})^2 & -c^{(e)}s^{(e)} & (c^{(e)})^2 & c^{(e)}s^{(e)} \\ -c^{(e)}s^{(e)} & -(s^{(e)})^2 & c^{(e)}s^{(e)} & (s^{(e)})^2 \end{bmatrix}}_{k^{(e)}} \tag{5.23}$$

Assemble the Element Equations

The analysis considered previously assembled equilibrium equations from the finite element method for each stress fiber and formed the framework for combining all stress fibers into a single system of equations. The contribution of the nucleus will also need to be considered to get the final equations of equilibrium of the cell system. To do this, first, the free-body diagram of the nucleus is considered in Fig. 5.3.

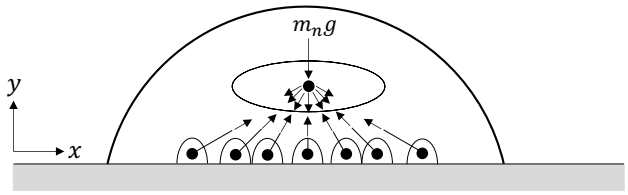


Figure 5.3: The free-body diagram of the nucleus with internal and external forces shown.

The paper focuses on the dynamics of the nucleus, and Newton's laws are used to describe its behavior, as indicated in the equations shown in Eqs. (5.24) and (5.25) to find its governing equations. Equations regarding the rotation of the nucleus, specifically Euler's law, are not included since the nucleus is modeled as a point-mass; hence, the rotational inertia is zero.

$$\sum F_X = m_{nuc}\ddot{U}_8 = -P_{8x}^{(1)} - ... - P_{8x}^{(7)} \tag{5.24}$$

$$\sum F_Y = m_{nuc}\ddot{V}_8 = -P_{8y}^{(1)} - ... - P_{8y}^{(7)} - m_n g \tag{5.25}$$

Impose the BC's

Next, the boundary conditions must be imposed to solve the system of element equations. The boundary conditions for all the nodes are Dirichlet and given in Eqs. (5.26) and (5.27).

$$U_1 = U_2 = \ , \ \ldots \ , \ = U_7 = 0 \quad \forall \quad t \tag{5.26}$$

$$V_1 = V_2 = \ , \ \ldots \ , \ = V_7 = 0 \quad \forall \quad t \tag{5.27}$$

By imposing the boundary conditions, the singular global system of equations formulated before is reduced to the so-called *reduced global system of equations*, which are non-singular. The *reduced global system of equations* are shown in Eq. (5.28).

$$M \begin{Bmatrix} \ddot{U}_8 \\ \ddot{V}_8 \end{Bmatrix} + K \begin{Bmatrix} U_8 \\ V_8 \end{Bmatrix} + \mathbf{G} = \mathbf{F} \tag{5.28}$$

The matrices M and K in Eq. (5.28) are known as the *reduced global mass and stiffness matrices*, respectively. The vectors $\mathbf{G}$ and $\mathbf{F}$ represent the contributions due to gravitational forces and the forces produced by the stress fibers, respectively. Note: The summation terms appear due to the combination of the contributions of each of the seven stress fibers.

$$M = \begin{bmatrix} m_n + \sum_{e=1}^{7} m_{22}^{(e)} cos^2(\theta^{(e)}) & \sum_{e=1}^{7} m_{22}^{(e)} cos(\theta^{(e)}) sin(\theta^{(e)}) \\ \sum_{e=1}^{7} m_{22}^{(e)} cos(\theta^{(e)}) sin(\theta^{(e)}) & m_n + \sum_{e=1}^{7} m_{22}^{(e)} sin^2(\theta^{(e)}) \end{bmatrix} \tag{5.29}$$

$$K = \begin{bmatrix} \sum_{e=1}^{7} k_{22}^{(e)} cos^2(\theta^{(e)}) & \sum_{e=1}^{7} k_{22}^{(e)} cos(\theta^{(e)}) sin(\theta^{(e)}) \\ \sum_{e=1}^{7} k_{22}^{(e)} cos(\theta^{(e)}) sin(\theta^{(e)}) & \sum_{e=1}^{7} k_{22}^{(e)} sin^2(\theta^{(e)}) \end{bmatrix} \tag{5.30}$$

$$\mathbf{G} = \begin{Bmatrix} 0 \\ m_n g \end{Bmatrix} \tag{5.31}$$

$$\mathbf{F} = \begin{Bmatrix} \sum_{e=1}^{7} cos(\theta^{(e)}) \Big(f^{(e)} x_q^{(e)} + \frac{f^{(e)}((x_p^{(e)})^2 - (x_q^{(e)})^2)}{2h^{(e)}} \Big) \\ \sum_{e=1}^{7} sin(\theta^{(e)}) \Big(- f^{(e)} x_p^{(e)} + \frac{f^{(e)}((x_q^{(e)})^2 - (x_p^{(e)})^2)}{2h^{(e)}} \Big) \end{Bmatrix} \tag{5.32}$$

Solve the Reduced Equations (Impose the ICs)

The equations will be used in the collocation method for the optimization, so the initial conditions will be presented in Eqs. (5.33) and (5.34) that will be used in the optimization.

$$U_1(t=0) = U_2(t=0) =, ..., = U_8(t=0) = 0 \tag{5.33}$$

$$V_1(t=0) = V_2(t=0) =, ..., = V_8(t=0) = 0 \tag{5.34}$$

Conversion of Inequality Constraints

The inequality constraints given in Eq. (5.8) imposing the stress and buckling constraints will be need to be re-expressed in terms of the displacement to be used in *fmincon*, as shown in Eqs. (5.35-5.36). Note: $|\sigma^{(e)}| \leq \sigma^{(e_{allow})}$ is mathematically the same as $-\sigma^{(e_{allow})} \leq \sigma^{(e)} \leq \sigma^{(e_{allow})}$; the latter formation will be used so that the stress constraint can be inserted into fmincon.

$$c^{(e)}U_8 + s^{(e)}V_8 \leq \frac{\pi A^{(e)}}{4L^{(e)}} \tag{5.35}$$

$$-\sigma^{(p_{allow})} \leq \frac{E^{(e)}}{L^{(e)}}(c^{(e)}U_7 + s^{(e)}V_7) \leq \sigma^{(p_{allow})} \tag{5.36}$$

5.3 Results

With the framework presented, one can predict the dynamic structural optimization of the cell's cytoskeletal structure during migration using material constants and other known parameters from the literature. The size of the nucleus of a fibroblast ranges between 500 μm^3 to 1000 μm^3 [64]; the size of the nucleus in this work is assumed to be 513.13 μm^3. The mass of the nucleus is assumed as 2.29 ng [65]; the cell mass is assumed to be 22.9 ng.

As is the nature of static and dynamic optimization problems with the methods in this paper, both static and dynamic, initial guesses of the solution must be provided in the solver. For a static system, only one guess is required per state. For the dynamic system, n guesses are required, where n represents the number of nodal points in time, which is tied to the temporal democrati-

zation process. As a general rule, a better guess will help the optimizer converge on the optimal solution faster with more accuracy. Several methods exist for generating the starting point for the states; one common process is to generate a preliminary trajectory with classical control, nonlinear control, or other methods as the initial guess for the optimal solution. The optimization serves as a refinement for the solution obtained from the other methods. The kinematic variables of cell migration, particularly the position and velocity of the cell, which in this context quantify the desired U_8, $\dot{U}_8$, and $\ddot{U}_8$, have been quantified in previous studies [70]. The force bounds of the stress fibers from previous studies suggest forces in the range of 0 to 120 nN [26, 71]. These values will be useful in supplying an initial guess of the kinematic and dynamic unknowns surrounding cell migration for the optimizer.

It is assumed that the value for the maximum allowable stress is the same in each stress fiber. Values in the literature for the material properties, particularly the breaking force, cross-sectional radius, the modulus of elasticity, and the density of the stress fibers are given as 377 nN, 100 nm, 1.45 MPa, and 40 mg/mL, respectively [58, 72]. Therefore, the ultimate tensile strength of the stress fibers could be assumed to be 12 MPa, however, there is evidence of a fluid-based models for the mechanical properties of the cell undergoing dynamic strain, therefore the value of the elastic modulus is chosen to be near zero [73]. The stress fiber length can vary from 10 to 100 μm; typical stress fiber lengths are around 100 μm [59]. The stress fiber length is $\geq 20\mu m$ [28]. Some studies show that the cross-sectional area of the stress fibers can range from 0.001 μm^2 to 0.04 μm^2, whereas focal adhesion cross-sectional areas can vary from approximately 0.75 μm^2 to approximately 4.5 μm^2 [59]. Additional studies show that the modulus of elasticity and cross-sectional areas of the focal adhesions are approximately 1.45 MPa for smooth muscle cells and 300 kPa for endothelial cells assuming uniform structure and 0.05 μm^2, respectively [74]. The results of the 1D optimization with these parameters are shown in Figs. 5.4 and 5.5.

The results of the optimization present some interesting findings. The forces spike approximately double the values compared to those obtained using a nonlinear controller [71]. The optimization results show that the stress fibers are heavily concentrated under the nucleus during the

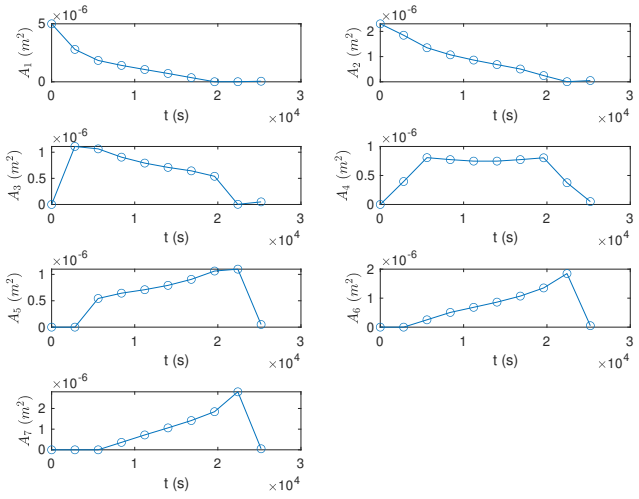


Figure 5.4: The areas from the dynamic structural optimization.

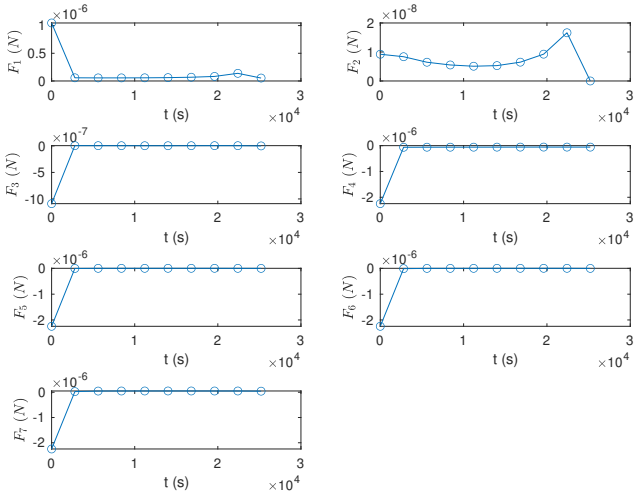


Figure 5.5: The forces from the dynamic structural optimization.

entire migration, presumably to support the weight of the nucleus. Large stress fiber forces are present under the nucleus and on the stress fibers on the protruding end of the cell, which are used to pull the cell forward. For the reader's convenience, an illustration of the stress fiber formations is illustrated in Fig. 5.6.

5.4 Discussion

The results of the optimization simulation are consistent with the expectations from experimental methods such as traction force microscopy and [71]. This is not surprising since [71] uses

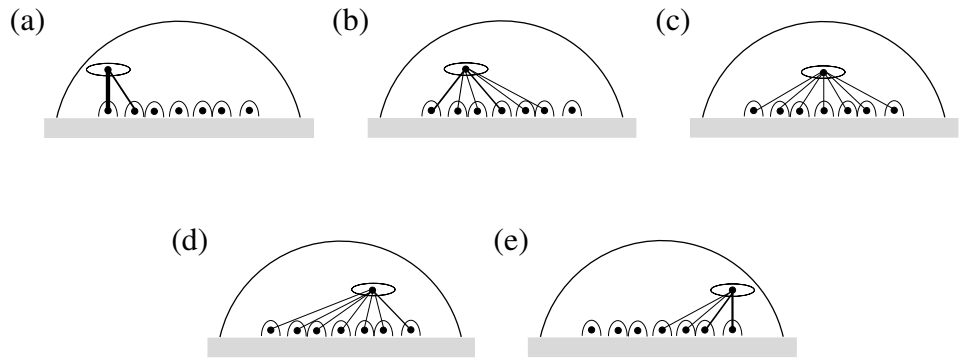


Figure 5.6: An illustration of the results from Fig. 5.6.

an optimization framework as well; the optimal force and areas serve as a form of refinement of the system controlled by the nonlinear controller in [71].

The model presented in this paper has several limitations, nonetheless. First, more sophisticated constraints could be imposed on the system to make the system more realistic, specifically biochemical constraints that further limit the forces in the stress fibers to ATP synthesis. Second, the model posed in this paper uses only 1D elements, which lacks the nuances of higher-dimensional models. Higher-dimensional optimization models can be constructed using the same framework presented in Eq. (5.1), shown in Eq. (5.37) in indicial form, where i and j = 1, and 2 for a 2D model and 1, 2, and 3 for a 3D model. In this formulation, a viscoelastic constitutive model can be used to model the cell cytoplasm while the stress fibers are still modeled with a Hookean

constitutive law.

$$\begin{aligned}
\min_{\mathbf{\Delta}=\{\mathbf{S},\mathbf{F}\}^T} \quad & J = \Phi(\mathbf{U}(\mathbf{X},t_1),t_1,\mathbf{U}(\mathbf{X},t_f),t_f) \\
& \quad + \int_{t_1}^{t_f} L(\mathbf{U}(\mathbf{X},t),\dot{\mathbf{U}}(\mathbf{X},t),\mathbf{F}(t),t)dt
\end{aligned}$$

$$\begin{aligned}
\text{subject to:} \quad & \text{EoE:} \quad \sigma_{ij,j}^{(e)} + f_i^{(e)} = \rho^{(e)} u_{i,tt}^{(e)} \\
& \text{CL:} \quad \begin{cases} \sigma_{ij}^{(e)} = C_{ijkl}^{(e)} \epsilon_{kl}^{(e)} \\ \sigma_{ij}^{(e)} + A_{ijkl}^{(e)} \sigma_{lk,t}^{(e)} = B_{ijkl}^{(e)} \epsilon_{kl,t}^{(e)} \end{cases} \\
& \text{SSR:} \quad \epsilon_{kl}^{(e)} = \frac{1}{2}(u_{i,j}^{(e)} + u_{j,i}^{(e)}) \\
& \text{BCs:} \quad \boldsymbol{\beta}(t_1,\mathbf{U}(t_1)),t_f,\mathbf{U}(t_f)) = \mathbf{0} \\
& \text{PC:} \quad \mathbf{\Pi}(\mathbf{U}(\mathbf{X},t),\mathbf{F},t) = \mathbf{0}
\end{aligned} \tag{5.37}$$

5.5 Conclusion

The framework presented in this paper seeks to accomplish two goals: to estimate the optimal traction forces of the cell on the substrate and the optimal positioning of the cell's stress fibers during migration. The outcome of this paper successfully demonstrated a methodology to create a novel biophysical model built on an optimization framework that predicts the forces within a reasonable range compared to data from other methods. The model presented here can be extended to higher-dimensional complex ECM environments to study cell morphogenesis during this process.

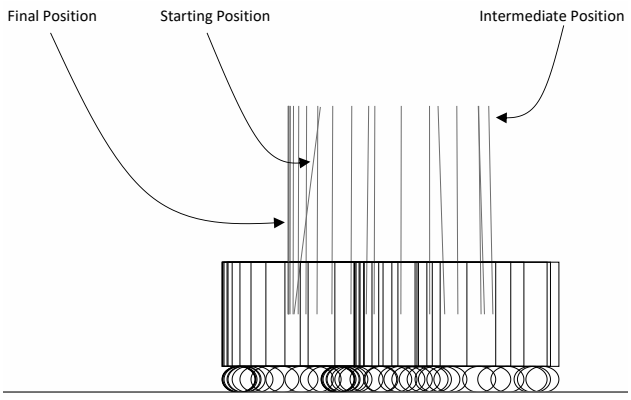
Chapter 6

Future Work

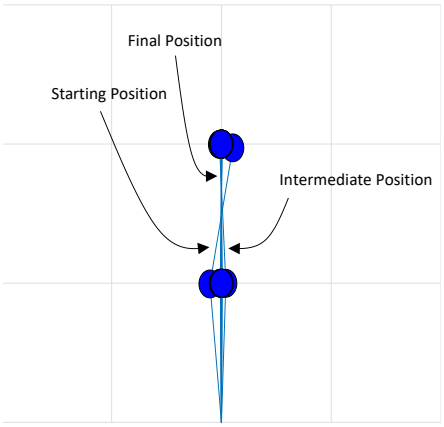
Extensions to the research presented in this dissertation could manifest in multiple ways. Future projects could also provide solution to the same problems that are posed in this dissertation using new methods.

6.1 Newton-Euler Optimal Model Extension

Future work can look at the construction of an optimal control policy for the Newton-Euler model of the robot that was obtained from a static structural optimization that was presented in this dissertation. Similar underactuated systems are presented as follows in order to illustrate the capability of developing an optimal control for simpler mechanical systems.



(a) Stroboscopic visualization of the cart-pendulum trajectory generated from a linear quadratic regulator (LQR) - based controller.



(b) Optimal trajectory of the acrobot.

Figure 6.1: Optimal trajectories, with minimum energy-based functionals, used to stabilize the underactuated systems of interest.

6.2 Cell Migration with Finite Element Models

One of the major goals in the field of cell migration is to quantify the stresses in the shared substrate of a group of cells that are produced while a cell or multiple cells migrate. Finite element models could provide useful insight to the quantification of the stress, strain, and force field variables.

6.2.1 Static Finite Element Models

Some static finite element-based models in FEBio have been produced, shown in Fig. 6.2, that seek to quantify the stress that is propagated during migration to neighboring cells. The contact between the cell and the substrate is modeled with a sticky contact. The contact between the interface of the nucleus and the remainder of the cell is tied elastic. The cell material is made of an isotropic active contraction with a neo-Hookean constitutive model.

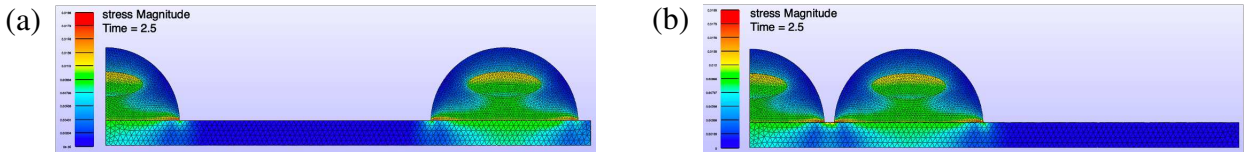
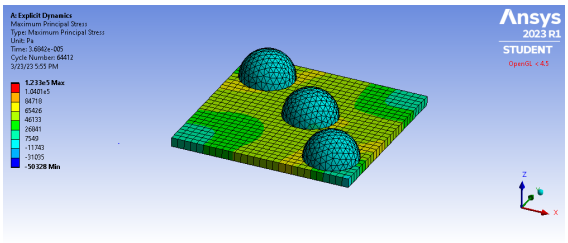


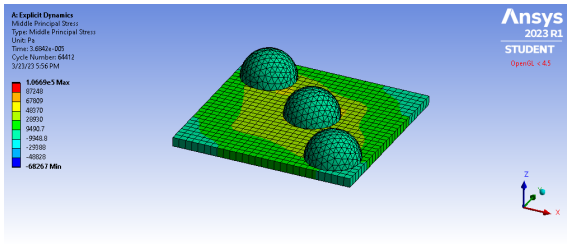
Figure 6.2: Stress magnitude distribution of contracting cells (a) with a large separation distance and (b) with a small separation distance.

6.2.2 Dynamic Finite Element Models

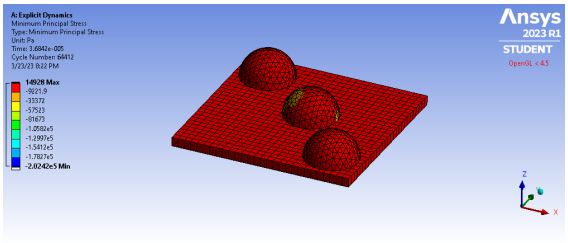
The results from the previous sections that developed a predictive model for a cell's spatio-temporal motion will be used here with the dynamic finite element method-based solver from the Ansys software package. In Fig. 6.3, a model of cells on a substrate is shown, with the effective stress distribution. In literature, the nucleus, and specifically its stiffness, is studied in order to determine the effects that it has on the migrational characteristics of the cell, specifically the speed of the cell [4]. Future work could explore this question by using the framework presented here.



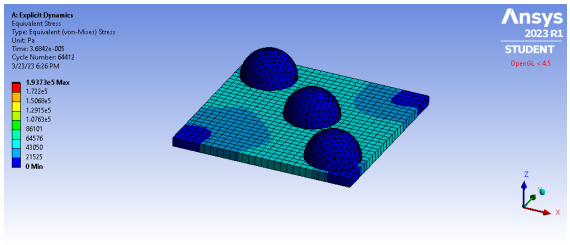
(a) Maximum principal stress during migration.



(b) Middle principal stress during migration.



(c) Mimimum principal stress during migration.



(d) Von-Mises stress during migraiton.

Figure 6.3: Principal stresses and von-Mises equivalent stress during migration found using a dynamic model in ANSYS.

6.3 Dynamic Cytoskeletal Optimization

The dynamic structural optimization presented in this dissertation has some limitations. The framework presented can be improved with more sophisticated constitutive laws, using higher-dimensional elements, and possibly from using better numerical methods in place of the finite element-based transcription.

6.3.1 Viscoelastic Model

A Maxwellian viscoelastic model is shown in (6.1), where η represents the damping coefficient and E represents the modulus of elasticity. Such a viscoelastic model could be used in place of the Hookean constitutive law currently used in this dissertation.

$$\frac{\eta}{E}\dot{\sigma} + \sigma = \eta\dot{\epsilon} \tag{6.1}$$

6.3.2 Higher-Dimensional Elements

One major limitation is that the model presented in this dissertation is only made from 1D elements. More sophisticated models, such as those with the nuances that come with 2D and 3D elements, will produce more accurate structural optimization results. A 2D finite element model can look like the one shown in Fig. 6.4.

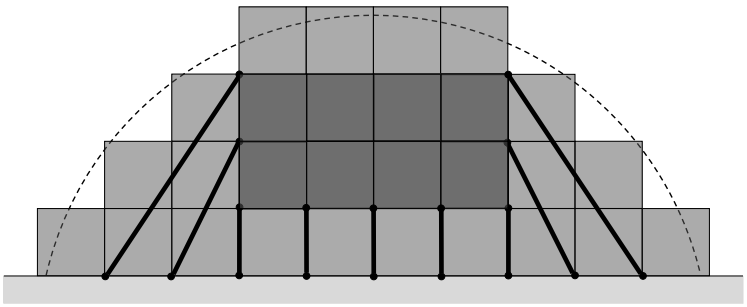


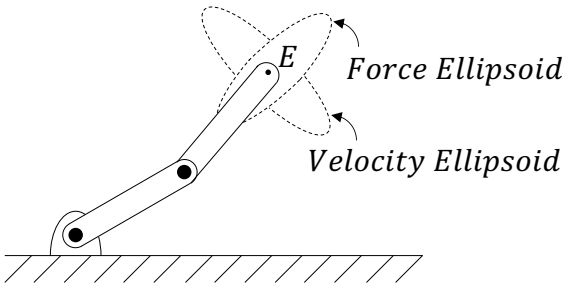
Figure 6.4: Higher-dimensional finite element model composed of 1D stress fibers and 2D cytoplasmic plate elements.

6.3.3 Material Point Method

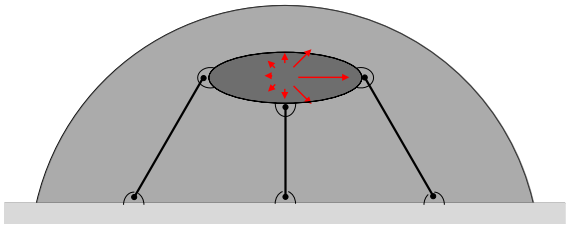
One of the other major limitations of the current optimization framework presented in this dissertation relates to the numerical scheme that is used to discretize the governing equation of equilibrium, expressed as a differential equation, into a system of algebraic equations, through a process called transcription. The current finite element formulation relies upon a small deformation assumption for accurate results. Future work could focus on the use of other numerical methods for solving differential equations, such as the material point method (MPM) which uses a mesh-free approach. The MPM is very useful for solving systems with very large displacements, since it avoids mesh tangling, unlike the finite element method [75].

6.4 Dynamic Capability Equations

In robotics literature, the formation of the force and velocity ellipsoids and their corresponding eigenvalues/eigenvectors using the principal axis theorem are a common way to estimate the capabilities of a robotic manipulator, shown in Fig. 6.5a.



(a) Force and velocity ellipsoids of the two-link robotic manipulator.



(b) Arrows indicating maximum capability of a particular dynamic quantity of interest.

Figure 6.5: (a) Hypothetical force and velocity ellipsoids for the robotics two-link manipulator, based on the configuration and actuator saturation limits and (b) Directional dependence of a hypothetical case of the capabilities, as determined from Eq. (6.2).

To further develop models for cell migration, specifically with respect to the cytoskeletal formation during locomotion, one avenue of analysis is through the so-called Dynamic Capability Equations [76], which combine the analysis from the force and velocity ellipsoids with the contact bound and dynamics constraints. This results in a system of inequalities shown in Eq. 6.2, that form a hypersurface that quantify the performance bounds on a dynamic system. The hypersurface represents the limitations on the system that arise from actuator saturation limits as well as frictional slip limits. Algorithms can be used to search this system of inequalities for the maximum of a particular quantity, such as force, acceleration, or velocity, which are known as the *directional DCE*. A simple hypothetical illustration of the maximum values of any of these parameters is illustrated in Fig. 6.5b. This knowledge has been harnessed in order to generate a gait for hopping robots [77], where the governing equations take the form of the manipulator equations, which can be obtained through the Newton-Euler method, Lagrange's method, D'Alembert's principle,

Kane's method, etc.

$$\mathbb{A} \begin{Bmatrix} |\dot{v}| \\ |\dot{\omega}| \end{Bmatrix} + \mathbb{F} \begin{Bmatrix} |f| \\ |m| \end{Bmatrix} \leq \mathbb{T} \tag{6.2}$$

Future work could focus on an extension of the dynamic capability equations for the cellular systems with the use of flexible system dynamics, such as with those presented in this dissertation, as opposed to the rigid-body models used in previous literature. The resulting equations can be useful in analyzing the dynamic performance of the cell along its path, specifically through a motion analysis that uses the worst-case direction information to avoid low performance configurations.

Bibliography

[1] J. Nathan Kutz, Steven L. Brunton, and Joshua L. Proctor Bingni W. Brunton. *Dynamic mode decomposition: Data-driven modeling of complex systems*. Society for Industrial and Applied Mathematics, 2016.

[2] Laura Murphy and Anotida Madzvamuse. A moving grid finite element method applied to a mechanobiochemical model for 3d cell migration. *Applied Numerical Mathematics*, pages 336–359, 2020.

[3] Jieling Zhao. *A Dynamic Finite Element Cellular Model and Its Application on Cell Migration*. IntechOpen, Rijeka, 2020.

[4] Adrian Moure and Hector Gomez. Dual role of the nucleus in cell migration on planar substrates. *Biomech Model Mechanobiol*, pages 1491–1508, 2020.

[5] O. Witkowski and T. Ikegami. Emergence of swarming behavior: Foraging agents evolve collective motion based on signaling. *PLoS ONE*, 11(4), 4 2016.

[6] Li Chen, Kevin Painter, Christina Surulescu, and Anna Zhigun. Mathematical models for cell migration: a non-local perspective. *Philosophical Transactions of the Royal Society B*, 2020.

[7] Rasa Giniūnaitė, Ruth E. Baker, and Philip K. Maini Paul M. Kulesa. Modelling collective cell migration: neural crest as a model paradigm. *J Math Biol*, 80:481–504, 2020.

[8] Matthias L. Zorn, Anna-Kristina Marel, Felix J. Segerer, and Joachim O. Rädler. Phenomenological approaches to collective behavior in epithelial cell migration. *Biochimica et Biophysica Acta (BBA) -Molecular Cell Research*, 1853:3143–3152, 2015.

[9] Brunton S. L. Kutz J. N. Proctor, J. L. Dynamic mode decom- position with control. *SIAM J. Appl. Dyn. Sys.*, 15:142–161, 2016.

[10] Konstantinos N. Blazakis. *Computational methods for investigating cell motility with applications to neutrophil cell migration*. PhD thesis, University of Sussex, 2015.

[11] Shuvasree SenGupta, Carole A. Parent, and James E. Bear. The principles of directed cell migration. *Nat Rev Mol Cell Biol*, 22:529–547, 2021.

[12] Min-Cheol Kim, Devin M. Neal, Roger D. Kamm, and H. Harry Asada. Dynamic modeling of cell migration and spreading behaviors on fibronectin coated planar substrates and micropatterned geometries. *PLOS Computational Biology*, 9(2):1–13, 02 2013.

[13] Y. Sakamoto, S. Prudhomme, and M.H. Zaman. Modeling of adhesion, protrusion, and contraction coordination for cell migration simulations. *Journal of Mathematical Biology*, pages 267—-302, 2014.

[14] Javier Satulovsky, Roger Lui, and Yu li Wang. Exploring the control circuit of cell migration by mathematical modeling. *Biophysical Journal*, pages 3671–3683, 2008.

[15] Soham Ghosh, Benjamin Seelbinder, Jonathan T. Henderson, Ryan D. Watts, Adrienne K. Scott, Alexander I. Veress, and Corey P. Neu. Deformation microscopy for dynamic intracellular and intranuclear mapping of mechanics with high spatiotemporal resolution. *Cell Reports*, 27(5):1607–1620.e4, 2019.

[16] Andrea Zancla, Pamela Mozetic, Monica Orsini, Giancarlo Forte, and Alberto Rainer. A primer to traction force microscopy. *J Biol Chem*, 298(5), 2022.

[17] Lorenzo Santoro, Lorenzo Vaiani, Antonio Boccaccio, Luciano Lamberti, Lorenzo Lo Muzio, Andrea Ballini, and Stefania Cantore. Finite element modeling of cells adhering to a substrate: An overview. *Applied Sciences*, 14(6), 2024.

[18] Charles M. Elliott, Bjorn Stinner, and Chandrasekhar Venkataraman. Modelling cell motility and chemotaxis with evolving surface finite elements. *Journal of The Royal Society Interface*, pages 3027—-3044, 2012.

[19] Eric J. Campbell and Prosenjit Bagchi. A computational model of amoeboid cell motility in the presence of obstacles. *Soft Matter*, pages 5741–5763, 2018.

[20] M.P. Neilson, J.A. Mackenzie, S.D Webb, and R.H. Insall. Modelling cell movement and chemotaxis using pseudopod-based feedback. *Society for Industrial and Applied Mathematics*, page 1035–1057, 2011.

[21] Nikolce Gjorevski, Alexandra S. Piotrowski, Victor D. Varner, and Celeste M. Nelson. Dynamic tensile forces drive collective cell migration through three-dimensional extracellular matrices. *Scientific Reports*, 5, 2015.

[22] Shahriar Seddighi Chaharborj, Sarkhosh Seddighi Chaharborj, Jalal Hassanzadeh Asl, and Pei See Phang. Controlling of pandemic covid-19 using optimal control theory. *Results in Physics*, 26:104311, 2021.

[23] Hyoryong Lee, Kim Dong-in, Su hyun Kwon, and Sukho Park. Magnetically actuated drug delivery helical microrobot with magnetic nanoparticle retrieval ability. *ACS Applied Materials & Interfaces*, 13:19633–19647, 2021.

[24] Philip Ball. Fluid interactions help fish in a school swim faster. https://physics.aps.org/articles/v11/46, May 2018.

[25] Jennifer Sargent. Why fish don't need to be 'schooled' in swimming. https://www.scientificamerican.com/article/why-fish-dont-need-to-be-schooled-in-swimming/, September 2013.

[26] N. S. Gov. Traction forces during collective cell motion. *HFSP Journal*, 3(4):223–227, August 2009.

[27] Rumi De. A general model of focal adhesion orientation dynamics in response to static and cyclic stretch. *Communications Biology*, 1, 2018.

[28] Pirentis A Suki B Stamenović D, Lazopoulos KA. Mechanical stability determines stress fiber and focal adhesion orientation. *Cell Mol Bioeng*, 2(4):475–485, 2009 Dec 1.

[29] Matthew Kelly. An introduction to trajectory optimization: how to do your own direct collocation. *SIAM Review*, 59:849–904, 01 2017.

[30] The MathWorks Inc. Find minimum of constrained nonlinear multivariable function - matlab fmincon. https://www.mathworks.com/help/optim/ug/fmincon.html#busog7r-options, 2023.

[31] Laura Murphy and Anotida Madzvamuse. A moving grid finite element method applied to a mechanobiochemical model for 3d cell migration. *Applied Numerical Mathematics*, 158:336–359, 2020.

[32] Adrian Moure and Hector Gomez. Phase-field modeling of individual and collective cell migration. *Arch. Comput. Methods Eng.*, 28:311–344, 2021.

[33] J. Renkawitz, Kopf A., Stopp J., and et al. Nuclear positioning facilitates amoeboid migration along the path of least resistance. *Nature*, 568:546–550, 2019.

[34] Inbal Hecht, Herbert Levine, Wouter-Jan Rappel, and Eshel Ben-Jacob. "self-assisted" amoeboid navigation in complex environments. *PLoS One*, 6, 2011.

[35] Hong S Park H. Choi A, Javius-Jones K. Cell-based drug delivery systems with innate homing capability as a novel nanocarrier platform. *Int J Nanomedicine*, 18:509–525, 2023.

[36] Lucas M. Bush, Connor P. Healy, Shwan B. Javdan, Jonathan C. Emmons, and Tara L. Deans. Biological cells as therapeutic delivery vehicles. *Trends in Pharmacological Sciences*, 42(2):106–118, 2021.

[37] Hao Yang, Xiangpeng Li, Yong Wang, Gang Feng, and Dong Sun. Control of single-cell migration using a robot-aided stimulus-induced manipulation system. *IEEE/ASME Transactions on Mechatronics*, 22(2):815–825, 2017.

[38] Jack Forman, Brady Hine, Samantha Kaonis, and Soham Ghosh. Chromatin remodeling and epigenetic modifications in planar cell migration. *BioRxiv*, 2023.

[39] Dmitry Ershov, Minh-Son Phan, Joanna W. Pylvänäinenand, Stéphane U. Rigaudand, Laure Le Blanc, Arthur Charles-Orszagand, James R. W. Conway, Romain F. Laine, Nathan H. Roy, Daria Bonazzi, Guillaume Duménil, Guillaume Jacquemet, and Jean-Yves Tinevez. Trackmate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nature Methods*, 19:829–832, 2022.

[40] Harvard Lomax, Thomas H. Pulliam, and David W. Zingg. *Fundamentals of Computational Fluid Dynamics*. Springer, 2011.

[41] Xiaoqing Tang, Yuke Li, Xiaoming Liu, Dan Liu, Zhuo Chen, and Tatsuo Arai. Vision-based automated control of magnetic microrobots. *Micromachines*, 13:337, 2022.

[42] Honglin Shen, Shuxiang Cai, Zhen Wang, Zhixing Ge, and Wenguang Yang. Magnetically driven microrobots: Recent progress and future development. *Materials Design*, 227:111735, 2023.

[43] Junyang Li, Xiaojian Li, Tao Luo, Ran Wang, Chichi Liu, Shuxun Chen, Dongfang Li, Jianbo Yue, Shuk han Cheng, and D. Sun. Development of a magnetic microrobot for carrying and delivering targeted cells. *Science Robotics*, 3(19):eaat8829, 2018.

[44] Huang Hanjin, Junyang Li, Cheng Wang, Liuxi Xing, Hui Cao, Chang Wang, Chung Leung, Zongze Li, Yue Xi, Hua Tian, Feng Li, and Dong Sun. Using decellularized magnetic microrobots to deliver functional cells for cartilage regeneration. *Small (Weinheim an der Bergstrasse, Germany)*, 20:e2304088, 11 2023.

[45] Tanyong Wei, Jiang Liu, Dongfang Li, Shuxun Chen, Yachao Zhang, Junyang Li, Lei Fan, Zhangyan Guan, Chung-Mau Lo, Lidai Wang, Kwan Man, and Dong Sun. Development of magnet-driven and image-guided degradable microrobots for the precise delivery of engineered stem cells for cancer therapy. *Small*, 16, 09 2020.

[46] Xaviar Trepat, Michael R. Wasserman, Thomas E. Angelini, Emil Millet, David A. Weitz, James P. Butler, and Jeffrey J. Fredberg. Physical forces during collective cell migration. *Nature Physics*, 5:426–430, 2009.

[47] Abhinava K. Mishra, Joseph P. Campanale, James A. Mondo, and Denise J. Montell. Cell interactions in collective cell migration. *Development*, 146:1–7, 2019.

[48] Lei Qin, Dazhi Yang, Weihong Yi, Huiling Cao, and Guozhi Xiao. Roles of leader and follower cells in collective cell migration. *Molecular Biology of the Cell*, 32:1267–1272, 2021.

[49] Jian Zhang and Cynthia A. Reinhart-King. Analysis of energy-driven leader-follower hierarchy during collective cancer cell invasion. *Cell Migration in Three Dimensions*, pages 247–262, 2023.

[50] Christopher D. Davidson, Firaol S. Midekssa, Samuel J. DePalma, William Y. Wang Jordan L. Kamen, Danica Kristen P. Jayco, Megan E. Wieger, and Brendon M. Baker. Mechanical intercellular communication via matrix-borne cell force transmission during vascular network formation. *Advanced Science*, 11, 2023.

[51] Cynthia A. Reinhart-King, Micah Dembo, , and Daniel A. Hammer. Cell-cell mechanical communication through compliant substrates. *Biophysical Journal*, 95:6044–6051, 2008.

[52] R. Ananthakrishnan and A. Ehrlicher. The forces behind cell movement. *Int J Biol Sci.*, pages 303–317, 2007.

[53] Jian Zhang and Yu li Wang. Centrosome defines the rear of cells during mesenchymal migration. *Molecular Biology of the Cell*, 28, 2017.

[54] Jaakko I Lehtimäki, Eeva Kaisa Rajakylä, Sari Tojkander, and Pekka Lappalainen. Generation of stress fibers through myosin-driven reorganization of the actin cortex. *eLife*, 10, 2020.

[55] Keith Burridge and Christophe Guilluy. Focal adhesions, stress fibers and mechanical tension. *Experimental Cell Research*, 343, 2016.

[56] Steven Munevar, Yu li Wang, and Micah Dembo. Traction force microscopy of migrating normal and h-ras transformed 3t3 fibroblasts. *Biophysical Journal*, 80, 2001.

[57] Eric Havenhill and Soham Ghosh. Optimization-based synthesis with directed cell migration. *bioRxiv*, 2024.

[58] Shinji Deguchi, Toshiro Ohashi, and Masaaki Sato. Tensile properties of single stress fibers isolated from cultured vascular smooth muscle cells. *Journal of Biomechanics*, 39(14):2603–2610, 2006.

[59] Geiger B Livne A. The inner workings of stress fibers-from contractile machinery to focal adhesions and back. *J Cell Sci*, 129(7):1293–304, 2016 Apr 1.

[60] Marita L. Rodriguez, Patrick J. McGarry, and Natan J. Sniadecki. Review on cell mecchanics: Experimental and modeling approaches. *Applied Mechanics Reviews - Transactions of the ASME*, 65, 2013.

[61] Steve A. Maas, Benjamin J. Ellis, Gerard A. Ateshian, and Jeffrey A. Weiss. FEBio: finite elements for biomechanics. *Journal of Biomechanical Engineering*, 134(1):011005, 02 2012.

[62] Jaques S. Milner, Matthew W. Grol, Kim L. Beaucage, S. Jeffrey Dixon, and David W. Holdsworth. Finite-element modeling of viscoelastic cells during high-frequency cyclic strain. *Journal of Functional Biomaterials*, 3(1):209–224, 2012.

[63] Francisco J Calero-Cuenca, Catia S Janota, and Edgar R Gomes. Dealing with the nucleus during cell migration. *Current Opinion in Cell Biology*, 50, 2018.

[64] Floriane Poignant, Ianik Plante, Zarana S. Patel, Janice L. Huff, and Tony C. Slaba. Geometrical properties of the nucleus and chromosome intermingling are possible major parameters

of chromosome aberration formation. *International Journal of Molecular Sciences*, 23(15), 2022.

[65] David Martínez-Martín, Gotthold Fläschner, Benjamin Gaub, Sascha Martin, Richard Newton, Corina Beerli, Jason Mercer, Christoph Gerber, and Daniel J. Müller. Inertial picobalance reveals fast mass fluctuations in mammalian cells. *Nature*, 550:500–505, 2017.

[66] David A. Levinson and Thomas R. Kane. *AUTOLEV - A new approach to multibody dynamics*, pages 81–102. Springer Berlin Heidelberg, Berlin, Heidelberg, 1990.

[67] Peter W. Christensen and Anders Klarbring. *An introduction to structural optimization*. Springer.

[68] Ramana Grandhi. Structural optimization with frequency constraints - a review. *AIAA Journal*, 31(12):2296–2303, 1993.

[69] M. Murrell, P. Oakes, M Lenz, and et al. Forcing cells into shape: the mechanics of actomyosin contractility. *Nature Reviews Molecular Cell Biology*, 16:486–498, 07 2015.

[70] Eric Havenhill and Soham Ghosh. Optimization-based synthesis with directed cell migration. *bioRxiv*, 2024.

[71] Eric Havenhill and Soham Ghosh. Mechanical analysis of cell migration using hy- brid structural optimization. *bioRxiv*, 2024.

[72] Daniel Ch. Haspinger, Sandra Klinge, and Gerhard A. Holzapfel. Numerical analysis of the impact of cytoskeletal actin filament density alterations onto the diffusive vesicle-mediated cell transport. *PLOS Computational Biology*, 17(5):1–26, 05 2021.

[73] Tenghu Wu and James J. Feng. A biomechanical model for fluidization of cells under dynamic strain. *Biophysical Journal*, 108:43–52, 01 2015.

[74] Elena Kassianidou and Sanjay Kumar. A biomechanical perspective on stress fiber structure and function. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1853(11, Part B):3065–3074, 2015. Mechanobiology.

[75] Deborah Sulsky. The material-point method. https://math.unm.edu/~sulsky/research/mpm.html, September 1998.

[76] A. Bowling and O. Khatib. The dynamic capability equations: a new tool for analyzing robotic manipulator performance. *IEEE Transactions on Robotics*, 21(1):115–123, 2005.

[77] S. Harmeyer and A. Bowling. Autonomous gait generation using acceleration capability analysis. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1319–1325, 2005.

[78] Nathan Kutz. Model discovery | amath 563 inferring structure of complex systems. https://faculty.washington.edu/kutz/am563/page1/page5/am563.html, 2016.

Appendix A

Supplementary Mathematics

A.1 DMD

In order to illustrate these concepts, a popular dynamic system of consideration from the field of physics, the Van der Pol Oscillator, will be considered in Eq. (A.1), where $\dot{y}$, $\ddot{y}$, and μ represent the time derivatives of the state, second time derivative of the state, and coefficient, respectively.

$$\ddot{y} - \mu(1 - y^2)\dot{y} + y = 0 \tag{A.1}$$

A synthetic data set of this system can be generated in MATLAB, allowing the production of spatiotemporal data of the two states of the system, $x_1 = y$ and $x_2 = \dot{y}$. Using the methodology already mentioned in the literature, as will be reproduced in this paper, the ODE can be recovered [78]. The A-matrix in Eq. A.2 is used, which contains a library of functions that are assumed to participate in the system's dynamics.

$$A = \begin{bmatrix} x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3, & x_2^2x_1 & x_1^2x_2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{A.2}$$

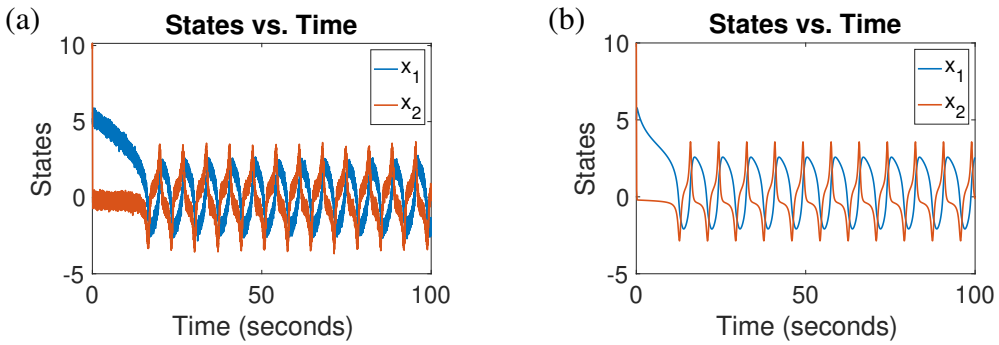


Figure A.1: (a) Plot of the states of the Van der Pol oscillator with noise added to the system. (b) Recovered states of the Van der Pol oscillator recovered dynamics from the DMD process.

The plot of the original dynamics compared to the recovered dynamics are shown in Fig. A.1.

A.2 Data Set

A.2.1 Single-Cell System

The data obtained from TrackMate is arranged in an arbitrary order, where n represents the number of time points. An example of this is shown in Eq. (A.3).

$$\mathbf{x} = \begin{Bmatrix} x_1(t=t_1) & x_1(t=t_n) & \dots & x_1(t=t_2) \end{Bmatrix}^T \tag{A.3}$$

A.2.2 Multi-Cell System

The original data is gathered in a matrix similar to the following, where the columns are arranged in an arbitrary order. The i^{th} column represents the x-values of the trajectory of the i^{th} biological cell in the system. The j^{th} row of a i^{th} column represents the i^{th} biological cell's x-value at the j^{th} time point. An example of this is shown in Eq. (A.4).

$$\begin{matrix} & \text{Cell \# 1} & \text{Cell \# 5} & \text{Cell \# 2} & \text{Cell \# 9} & \\ \texttt{x\_cell} = & \begin{bmatrix} x_1(t_1) & x_5(t_1) & x_2(t_1) & x_9(t_1) & \dots \\ x_1(t_2) & x_5(t_2) & x_2(t_2) & x_9(t_2) & \dots \\ x_1(t_3) & x_5(t_3) & x_2(t_3) & x_9(t_3) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1(t_n) & x_5(t_n) & x_2(t_n) & x_9(t_n) & \dots \end{bmatrix} & \begin{matrix} \text{t} = t_1 \\ \text{t} = t_2 \\ \text{t} = t_3 \\ \\ \text{t} = t_n \end{matrix} \end{matrix} \tag{A.4}$$

For the multi-cellular case, m represents the number of cells and n represents the number of time points for each cell. In general, the positions are monotonically increasing as in Eq. (A.5), and for all strictly increasing time as in Eq. (A.6).

$$x_1 < x_2 < x_3 < \cdots < x_m \tag{A.5}$$

$$t_1 < t_2 < t_3 < \ldots t_n \tag{A.6}$$

A.3 Dynamic Model Discovery - Sol'n. Formulation

A.3.1 Single-Cell System

Organize the Data

The data of the positional points over time will be organized into the vector that is shown in Eq. (A.7). The subscript on the x-value represents the biological cell's number. The same process for the y-direction will be repeated for this cell but is not shown here to avoid redundancy.

$$\mathbf{x} = \begin{Bmatrix} x_1(t = t_1) & x_1(t = t_2) & \ldots & x_1(t = t_n) \end{Bmatrix}^T \tag{A.7}$$

A.3.2 Multi-Cell System

Organize the Data

The matrix of positional vs. temporal data for each cell is given in Eq. (A.8).

$$x = \begin{array}{c} \\ x \text{ changes} \downarrow \end{array} \overbrace{\begin{bmatrix} x_1(t_1) & \ldots & x_1(t_n) \\ \vdots & \ddots & \vdots \\ x_m(t_1) & \ldots & x_m(t_n) \end{bmatrix}}^{t \text{ changes} \rightarrow} \tag{A.8}$$

The matrix of displacement data is given in Eq. (A.9).

$$u = \begin{array}{c} \\ x \text{ changes} \downarrow \end{array} \overbrace{\begin{bmatrix} u(x_1, t_1) & \ldots & u(x_1, t_n) \\ \vdots & \ddots & \vdots \\ u(x_m, t_1) & \ldots & u(x_m, t_n) \end{bmatrix}}^{t \text{ changes} \rightarrow} \tag{A.9}$$

This is the displacement matrix that corresponds to the displacement field of the cells, where the displacement is defined as the difference between the current x-position of the biological cell and

the original x-position of the cell, as shown in Eq. (A.10), where $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$.

$$u(x_i, t_j) \triangleq x_i(t_j) - x_i(t_1) \tag{A.10}$$

A.4 Finite Difference Method

Both spatial and temporal derivatives of the data will need to be calculated since they will be going into the matrix in Eq. (A.23). Finite differencing is used to calculate the derivatives; however, it should be noted that more sophisticated differentiation techniques, such as total variation of derivatives, can be used in order to improve the model. For this paper, simple differencing schemes, such as Euler backward and the central difference method, as opposed to higher-order stencils, are used to calculate the first and second-order derivatives, respectively. Although they are simple stencils, they are sufficient to perform the differentiation.

A.4.1 Single-Cell System

The finite differencing is outlined in the multi-cell system solution so the process is not repeated in order to avoid redundancy.

A.4.2 Multi-Cell System

First-Order Differentiation - Time Derivation

The finite difference differentiation scheme for a first-order time derivative is given in Eq. (A.11).

$$\dot{u}(x_i, t_j) \triangleq \frac{\partial u_i^j}{\partial t} \approx \frac{u_i^j - u_i^{j-1}}{\Delta t} \tag{A.11}$$

For implementation in this paper, the differentiation stencil expressed in Eq. (A.11) will be expressed equivalently in matrix form in Eq. (A.12).

$$\begin{Bmatrix} \frac{\partial u_i^2}{\partial t} \\ \frac{\partial u_i^3}{\partial t} \\ \vdots \\ \frac{\partial u_i^n}{\partial t} \end{Bmatrix} = \frac{1}{\Delta t} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & \ddots & \ddots \end{bmatrix} \begin{Bmatrix} u_i^2 \\ u_i^3 \\ \vdots \\ u_i^n \end{Bmatrix}$$

$$+ \frac{1}{\Delta t} \begin{Bmatrix} -u_i^1 \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \tag{A.12}$$

The time derivatives are calculated for each biological cell throughout its entire trajectory, which will result in a differencing matrix, shown in Eq. (A.12). After each cell's trajectory is numerically differentiated with respect to time, they are assembled into the matrix given in Eq. (A.13), where $\dot{u} \in \mathbb{R}^{m \times n-1}$.

$$\dot{u} = \begin{array}{c} \\ x \text{ changes} \downarrow \end{array} \begin{array}{c} \overrightarrow{t \text{ changes}} \\ \begin{bmatrix} \dot{u}(x_1, t_2) & \dots & \dot{u}(x_1, t_n) \\ \vdots & \ddots & \vdots \\ \dot{u}(x_m, t_2) & \dots & \dot{u}(x_m, t_n) \end{bmatrix} \end{array} \tag{A.13}$$

However, since using higher-order derivatives in time, the last column from the matrix is omitted to maintain compatible dimensions. For the same reason, the first and last rows will also be omitted for the spatial derivatives. Therefore, an entire matrix for the time-derivative of u is obtained, $\dot{u} \in \mathbb{R}^{m-2 \times n-2}$, that takes the form given in Eq. (A.14).

$$\dot{u} = \begin{array}{c} \\ x \text{ changes} \\ \downarrow \end{array} \overbrace{\begin{bmatrix} \dot{u}(x_2, t_2) & \dots & \dot{u}(x_2, t_{n-1}) \\ \vdots & \ddots & \vdots \\ \dot{u}(x_{m-1}, t_2) & \dots & \dot{u}(x_{m-1}, t_{n-1}) \end{bmatrix}}^{t \text{ changes}} \tag{A.14}$$

First-Order Differentiation - Spatial Derivation

The finite difference differentiation scheme using Euler backward for a first-order spatial derivative, which assumes that the initial condition is known, is given in Eq. (A.15).

$$u'(x_i, t_j) \triangleq \frac{\partial u_i^j}{\partial x} \approx \frac{u_i^j - u_{i-1}^j}{\Delta x} \tag{A.15}$$

For implementation in this paper, the differentiation stencil expressed in Eq. (A.15) will be expressed equivalently in matrix form in Eq. (A.16).

$$\begin{Bmatrix} \frac{\partial u_2^j}{\partial x} \\ \frac{\partial u_3^j}{\partial x} \\ \vdots \\ \frac{\partial u_m^j}{\partial x} \end{Bmatrix} = \frac{1}{\Delta x} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & \ddots & \ddots \end{bmatrix} \begin{Bmatrix} u_2^j \\ u_3^j \\ \vdots \\ u_m^j \end{Bmatrix}$$

$$+ \frac{1}{\Delta x} \begin{Bmatrix} -u_1^j \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \tag{A.16}$$

Second-order differentiation - Spatial Derivation

The finite difference differentiation scheme for the central difference method for a second-order derivative is given in Eq. (A.17).

$$u''(x_i, t_j) \triangleq \frac{\partial^2 u_i^j}{\partial x^2} \approx \frac{u_{i-1}^j - 2u_i^j + u_{i+1}^j}{\Delta x^2} \tag{A.17}$$

For implementation in this paper, the differentiation stencil expressed in Eq. (A.17) will be expressed equivalently in matrix form in Eq. (A.18).

$$\begin{Bmatrix} \frac{\partial^2 u_2^j}{\partial x^2} \\ \frac{\partial^2 u_3^j}{\partial x^2} \\ \vdots \\ \frac{\partial^2 u_{m-1}^j}{\partial x^2} \end{Bmatrix} = \frac{1}{(\Delta x)^2} \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & \ddots & 0 \\ 0 & \ddots & \ddots & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} \begin{Bmatrix} u_2^j \\ u_3^j \\ \vdots \\ u_{m-1}^j \end{Bmatrix}$$

$$+ \frac{1}{(\Delta x)^2} \begin{Bmatrix} u_1^j \\ 0 \\ \vdots \\ u_m^j \end{Bmatrix} \tag{A.18}$$

A.5 DMD Formulation

A.5.1 Single-Cell System

The finite differencing techniques for the singular cell will be similar to the multi-cellular system. Both systems use a finite differencing scheme to approximate the derivative at particular nodes, which will be needed in Eq. (A.20).

$$\begin{Bmatrix} | \\ \frac{\partial \mathbf{x}}{\partial t} \\ | \end{Bmatrix} = A \begin{Bmatrix} | \\ \xi \\ | \end{Bmatrix} \tag{A.19}$$

$$\begin{Bmatrix} | & | \\ \frac{dx}{dt} & \frac{dy}{dt} \\ | & | \end{Bmatrix} = \underbrace{\begin{bmatrix} | & | & | & | & | \\ 1 & O & O^2 & O^3 & t \\ | & | & | & | & | \end{bmatrix}}_{A} \begin{Bmatrix} | & | \\ \xi_x & \xi_y \\ | & | \end{Bmatrix} \tag{A.20}$$

$$\dot{x} = f_x(\cdot) \tag{A.21}$$

$$\dot{y} = f_y(\cdot) \tag{A.22}$$

A form of matrix A was originally shown in Eq. (A.2) for the Van der Pol system; here, the matrix is chosen at the discretion of the engineer. Sample matrices are given in Section 3.3. The system is solved, as shown in Eq. (3.4). The following list will illustrate the appropriate choice of inverse when solving this system of equations.

1. If A has fewer rows than columns, then the system has infinite exact solutions that can be found using the pseudo-inverse.

2. If A is square and full-rank, then the standard inverse can be used.

3. If A has more rows than columns, then no exact solution exists, but an approximate one can be found using the pseudo-inverse.

A.5.2 Multi-Cell System

With time-history data of the trajectories and the derivative data of the cells obtained from TrackMate, illustrated in Fig. 3.2, and the differentiation from the finite difference method, the multi-cellular system of equations can be assembled, shown in Eq. (A.23).

$$\begin{Bmatrix} | \\ \frac{\partial U}{\partial t} \\ | \end{Bmatrix} = \underbrace{\begin{bmatrix} | & | & | & | & | & | & | \\ U & U^2 & U^3 & \frac{\partial U}{\partial x} & \frac{\partial U}{\partial x}U & (\frac{\partial U}{\partial x})^2 & \frac{\partial^2 U}{\partial x^2} \\ | & | & | & | & | & | & | \end{bmatrix}}_{A} \begin{Bmatrix} | \\ \xi \\ | \end{Bmatrix} \tag{A.23}$$

The structure given in Eq. (A.23) is made possible by reshaping the matrices into vectors that are stacked on top each other, as is done in Eq. (A.24), for example, where $u \in \mathbb{R}^{m-2 \times n-2}$ and $U \in \mathbb{R}^{(m-2)*(n-2)\times 1}$.

$$U = \begin{Bmatrix} u(2, 2:n-1)^T \\ u(3, 2:n-1)^T \\ \vdots \\ u(m-1, 2:n-1)^T \end{Bmatrix} \tag{A.24}$$

A.6 Dynamic Equations

A.6.1 Single-Cell System

Several different structures for the A-matrix are explored here to study the effects of the accuracy of the recovered dynamics. The data from the trajectory illustrated in Fig. 3.2 is used for this study.

Structure of A - Version 1

Using a matrix with one term, shown in Eq. (A.25), the dynamics are recovered as shown in Eq. (A.25).

$$A = \begin{bmatrix} | \\ x \\ | \end{bmatrix} \rightarrow \begin{Bmatrix} \dot{x} \\ \dot{y} \end{Bmatrix} = 10^{-5} \begin{bmatrix} 1.11621 & 0 \\ 0.850172 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{A.25}$$

Structure of A - Version 2

Using a matrix with two terms, shown in Eq. (A.26), the dynamics are recovered as shown in Eq. (A.26).

$$A = \begin{bmatrix} | & | \\ x & y \\ | & | \end{bmatrix} \rightarrow \begin{Bmatrix} \dot{x} \\ \dot{y} \end{Bmatrix} = 10^{-5} \begin{bmatrix} 0.108295 & 4.31462 \\ 1.76421 & -3.91278 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{A.26}$$

Structure of A - Version 3

The matrix containing four different terms is shown in Eq. (A.27).

$$A = \begin{bmatrix} | & | & | \\ x & y & xy \\ | & | & | \end{bmatrix} \tag{A.27}$$

The dynamics are recovered as shown in Eq. (A.28).

$$\begin{Bmatrix} \dot{x} \\ \dot{y} \end{Bmatrix} = 10^{-3} \begin{bmatrix} -0.0138 & 0.2372 & -0.0003 \\ 0.0355 & -0.2727 & 0.0004 \end{bmatrix} \begin{bmatrix} x \\ y \\ xy \end{bmatrix} \tag{A.28}$$

Structure of A - Version 4

The matrix containing four different terms is shown in Eq. (A.29).

$$A = \begin{bmatrix} | & | & | & | & | \\ 1 & O & O^2 & O^3 & t \\ | & | & | & | & | \end{bmatrix} \tag{A.29}$$

The dynamics are recovered as shown in Eq. (A.30).

$$\begin{Bmatrix} \dot{x} \\ \dot{y} \end{Bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \begin{Bmatrix} 1 \end{Bmatrix} + \begin{bmatrix} -2.757 & 10.448 \\ 4.096 & -15.965 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$+10^{-2} \begin{bmatrix} -9.930 & 1.974 & 9.303 \\ 14.865 & -2.906 & -2.906 \end{bmatrix} \begin{Bmatrix} x^2 \\ xy \\ y^2 \end{Bmatrix}$$

$$+10^{-5} \begin{bmatrix} -3.523 & -38.243 & 22.035 & 20.041 \\ 5.155 & 56.770 & -32.645 & -30.144 \end{bmatrix} \begin{Bmatrix} x^3 \\ x^2y^1 \\ xy^2 \\ x^3 \end{Bmatrix}$$

$$+10^{-7} \begin{bmatrix} 18.842 \\ -7.257 \end{bmatrix} \begin{Bmatrix} t \end{Bmatrix} \tag{A.30}$$

The plots of the original TrackMate trajectory data, in addition to the plots from the solution of the dynamics subject to the boundary conditions in Table A.1 obtained from the DMD, are shown in

Fig. (A.2), where for the graphs on the right-hand side, the blue and red curves represent the x and y-positions of the cell, respectively. It's apparent that the trajectory from the DMD-based dynamics approaches the trajectory of the cell obtained by TrackMate as the number of approximating terms in the dynamics increases. This gives further confidence that the PDE regression solution will give reliable results.

A.6.2 Multi-Cell System

$$(\delta_t u_i)^j = \frac{-3u_i^j + 4u_i^{j+1} - u_i^{j+2}}{2\Delta t} \approx \frac{\partial u_i^j}{\partial t}\Big|_{x_i,t_j} \tag{A.31}$$

$$(\delta_t u_i)^j = \frac{-11u_i^j + 18u_i^{j+1} - 9u_i^{j+2} + 2u_i^{j+3}}{6\Delta t} \approx \frac{\partial u_i^j}{\partial t}\Big|_{x_i,t_j} \tag{A.32}$$

A.7 Optimization

A.7.1 ODE Optimization

The path constraint, $\mathbf{C}(\mathbf{x}(t), t)$, is assumed to be zero since there will be no obstacles. The kinetic energy of the cell is defined in Eq. (A.33), where m_i represents the mass of the i^{th} cell, I_i represents the inertia tensor of the i^{th} cell, $\dot{\mathbf{x}}_i$ represents the velocity of the i^{th} cell, and ω_i represents the angular velocity of the i^{th} cell.

$$T_i = \frac{1}{2}\dot{\mathbf{x}}_i^T m_i \dot{\mathbf{x}}_i + \frac{1}{2}\omega_i^T I_i \omega_i \tag{A.33}$$

The relationship between force and the potential is defined in Eq. (A.34), where F_i represents the traction force of the i^{th} cell, V_i represents the potential field of the i^{th} cell, and u_i represents

the displacement of the i^{th} cell as defined in Eq. (A.10).

$$F_i = -\frac{\partial V_i}{\partial u_i} \tag{A.34}$$

Values of the traction forces during cell migration range from 0 nN to 120 nN [26]. The traction force will be assumed to linearly decrease from this maximum to the minimum value during its migration. The potential field is modeled in Eq. (A.35), where the c_1, c_2, and c_3 terms arise as constants of integration considering just one cell, $i = 1$.

$$V_i(u_i) = \sum_{k=1}^{3} c_k u_i^{k-1} \tag{A.35}$$

For the case of the cell's motion that is governed by the ODE, Table A.1 gives the values that are used in the simulation. The choice of 1, T, and T-V for the L$(\cdot)$ function is to find a minimum time, energy, and action, respectively, for the cell migrational problem. The mass of the system is given as m=2.29 ng [65]. The velocity bounds in Eq. (A.36) are obtained from experimentation.

$$\text{-0.0033 } \frac{\mu m}{s} \leq \sqrt{\dot{x}_i^2 + \dot{y}_i^2} \leq \text{0.0033 } \frac{\mu m}{s} \tag{A.36}$$

The optimized single-cell system is shown in Fig. A.3. The x and y directional motion of the cells are fixed subject to the discovered dynamics; the vertical motion is free. The endpoints are fixed in time and space. The results of these simulations from the hybrid synthesis presented serve as good predictors of cell movement that are at the interface of the scratch. However, it is still desirable to be able to predict the movement of the cells that are located on the scratch and wall boundaries.

A.7.2 Optimization and Prediction of Multiple Cell Trajectories

With the framework developed, the trajectories of multiple cells can be obtained. The obstacles could be represented by other cells or fibers in the extracellular matrix.

Table A.1: The values of the states and time at the boundary.

State	$t_0 = 0s$	$t_f =$ Unknown
x	354.060 μ m	438.904 μ m
$\dot{x}$	0 μ m/s	Known from Eq. (A.25,A.26,A.27,A.29)
y	53.906 μ m	121.759 μ m
$\dot{y}$	0 μ m/s	free

3D cell model predicts the stress fiber localization and thickness as a material resource available to cells

A more sophisticated and realistic geometry of the cell can be considered by constructing a model in three dimensions with a goal to predict the stress fiber formation in 3D space. A problem like this can be solved with MATLAB's fmincon as was done for the simplified cell model. However, due to geometric complexity, the structural optimization is performed in Ansys Topology Optimization. The optimization results from Ansys is shown in Fig. A.4. The optimization results compare closely to the optimizations that were shown in Fig. 4.3. This serves as a verification that the optimization results are converging on the appropriate solution.

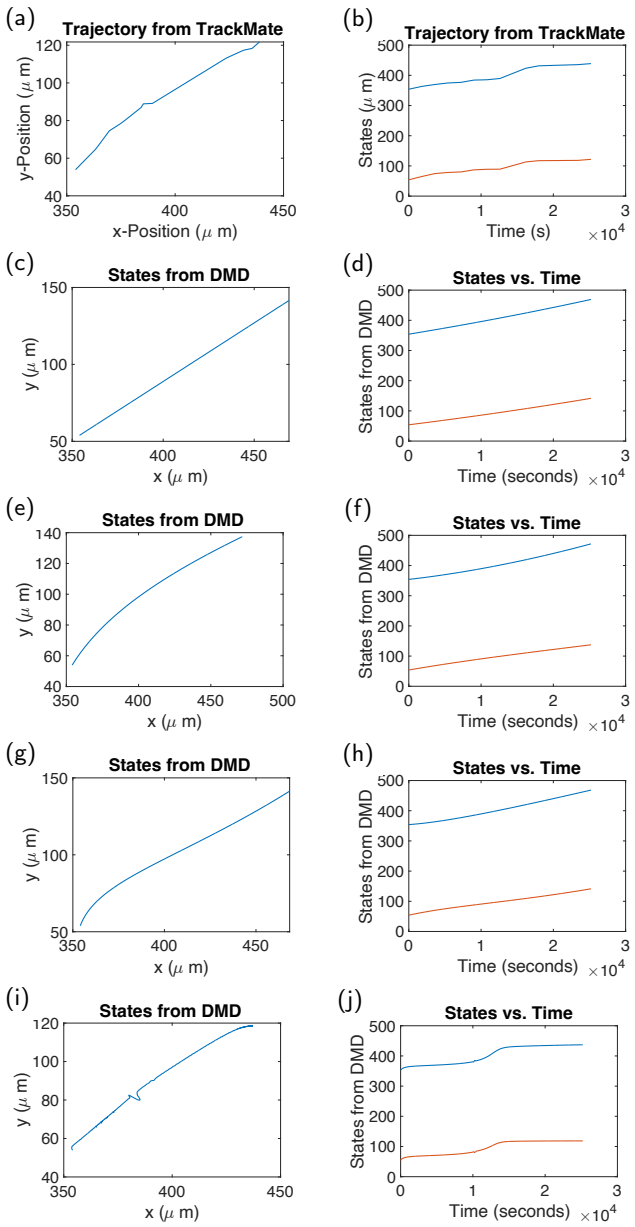


Figure A.2: (a) The original y(x) and (b) y(t), shown in red, and x(t), shown in blue, experimental data from TrackMate is shown. The y(x) trajectories of Eqs. (A.25), (A.26), (A.28), and (A.30) are shown in (c), (e), (g), and (i), respectively. The x(t), in blue, and y(t), in red, trajectories of Eqs. (A.25), (A.26), (A.28), and (A.30) are shown in (d), (f), (h), and (j), respectively.

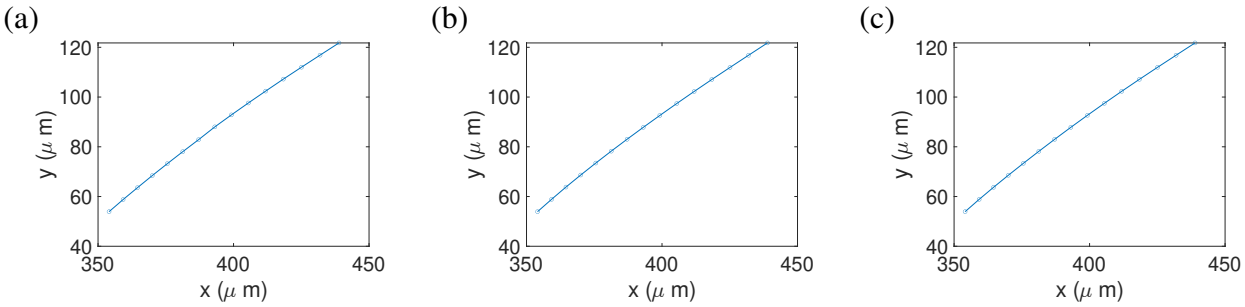


Figure A.3: Graphs of the hybrid DMD-optimization trajectories for the single cell, with a Lagrangian cost of (a) time. (b) energy. (c) action.

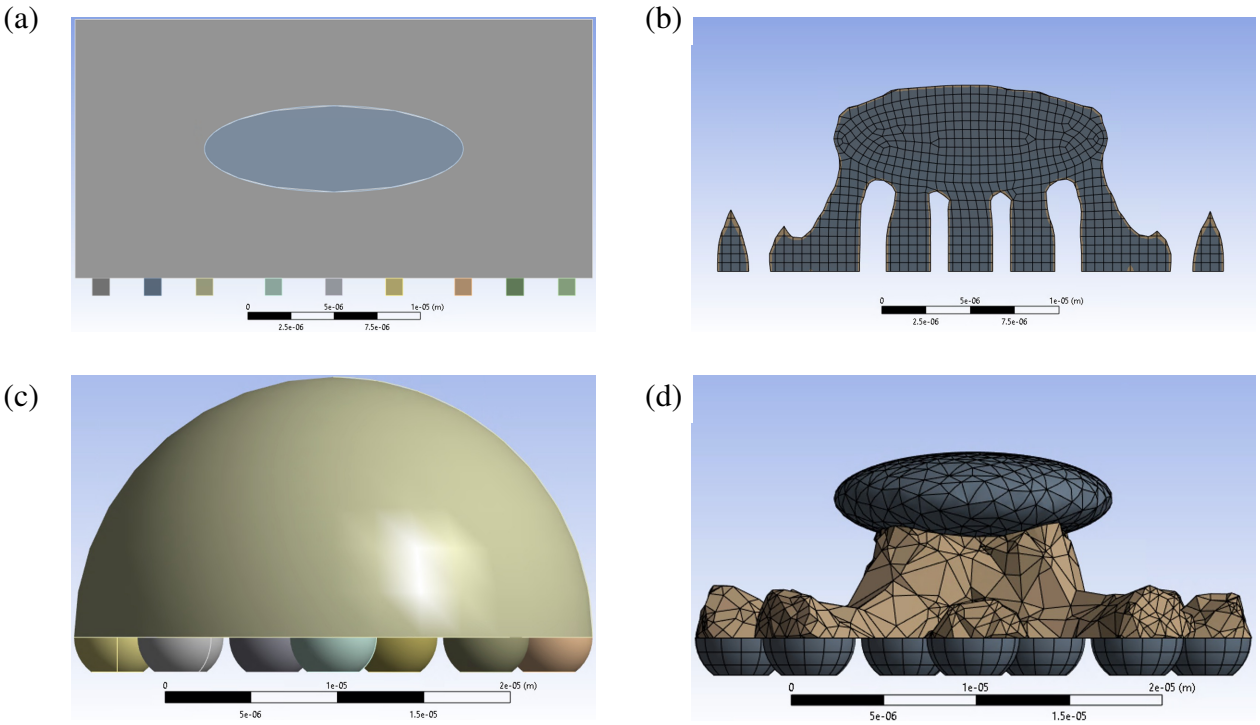


Figure A.4: Structural optimizations of the cells in ANSYS are shown. (a) The 2D model of the cell before the optimization. (b) The minimum mass optimization results for the 2D case. (c) The 3D model of the cell before the optimization. (d) The minimum mass optimization results for the 3D case.