Partial Differential Equation Models of Collective Migration During Wound Healing

by

John Nardini

B.S., North Carolina State University, 2013

M.S., University of Colorado, 2016

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of Doctor of Philosophy

Department of Applied Mathematics

2018

This thesis entitled:
Partial Differential Equation Models of Collective Migration During Wound Healing
written by John Nardini
has been approved for the Department of Applied Mathematics

_____
Prof. David M. Bortz

_____
Prof. Vanja Dukic

_____
Prof. Zachary Kilpatrick

_____
Prof. Xuedong Liu

_____
Prof. Fatemeh Pourahmadian

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Nardini, John (Ph.D., Applied Mathematics)

Partial Differential Equation Models of Collective Migration During Wound Healing

Thesis directed by Prof. David M. Bortz

This dissertation is concerned with the derivation, analysis, and parameter inference of mathematical models of the collective migration of epithelial cells. During the wound healing process, epidermal keratinocytes collectively migrate from the wound edge into the wound area as a means to re-establish the outermost layer of skin. This migration into the wound is stimulated by the presence of epidermal growth factor. Accordingly, this dissertation focuses on the migratory response of epidermal keratinocytes in response to this growth factor. Such studies will suggest suitable clinical treatments to consider for chronic wounds and invasive cancers.

We begin with a study into the role of cell-cell adhesions on keratinocyte migration during wound healing. We use an inverse problem methodology in combination with model validation to show that cells use these connections to promote migration by pulling on their follower cells as they migrate into the wound. We next derive a biochemically-structured version of Fisher's Equation that provides a framework to study how patterns of biochemical activation influence migration into the wound. We prove the existence of a self-similar traveling wave solution. In considering a more complicated scenario where cell migration depends on biochemical activity levels, we show numerically that the threshold parameter where all cells in the population become activated yields the simulations that migrate farthest into the wound. Lastly, we consider the role of numerical error on an inverse problem methodology. The numerical approximation of a cost function is dominated by either numerical or experimental error in computations, which leads to different rates of convergence as numerical resolution increases. We use residual analysis to derive an autocorrelative statistical model for cases where numerical error is the main source of error for first order schemes. This autocorrelative statistical model can correct confidence interval computation for these methods and hence improve uncertainty quantification.

Dedication

In memory of Josephine A. Cosco.

We miss and love you, Nonnie.

Acknowledgments

Only my name goes on the title page of this dissertation, but I simply wouldn't be here without the help of so many people over the past 27 years. I am extremely grateful to my advisor, Dr. David Bortz for helping me develop as an applied mathematician and person for the past five years. Your constant belief in me has brought many opportunities that I never would have thought possible.

I am also thankful for my committee members, Drs. Vanja Dukic, Zachary Kilpatrick, Xuedong Liu, and Fatemeh Pourahmadian. These individuals have given me personal guidance all throughout my graduate education by helping me prepare for preliminary exams, letting me attend lab meetings, teaching me in class, reading over manuscripts and research drafts, and so much more.

Many other people have also made this dissertation possible. Members of the Math Biology group at CU have improved my research by asking good questions whenever I gave lab meeting, and never judged me for asking weird (or obvious) questions. Dr. Douglas Chapnick performed the experiments discussed and analyzed in Chapter 2 and provided helpful feedback in preparing my first manuscript. I am also grateful to the IQ Biology Program at the BioFrontiers Institute at CU for funding support but also the opportunity to pursue interdisciplinary research that fit my interests. I would particularly like to thank Dr. Thomas Cech, Dr. Jana Watson-Capps, Kim Kelly, Kim Little, Janice Jones, Amber McDonnell, Dr. Kristin Powell, and especially Dr. Andrea Stith. I am also grateful to the Applied Mathematics staff, including Emily O' Connor, Ian Cunningham, Desiree Holtz, Mary Fentress, and Catriona Allcock for always making my life as easy as possible.

I have also been blessed with a wonderfully supportive family. I can't thank my mom enough for volunteering at the local school and as a foster parent when I was young, and letting me help in these endeavors. Through this, I learned the importance of education and helping others at a very young age. I'll never understand how my dad had the perseverance to spend his Saturdays taking me to hockey practice at 6 AM, basketball practice at 10 AM, indoor soccer games at 6 PM, and the library somewhere in between all of this. Almost more surprising was my parents' willingness to let me pursue mathematics at NC State when we didn't know much about careers in mathematics or North Carolina. Thank you both for trusting me and for being so accepting at every step along the way. My older brother Stephen was my first best friend, and my oldest brother Nick continues to set the bar higher and higher for me. I would also like to thank my second mother, Laurie Maillette,

Contents

List of Tables

List of Figures

Chapter 1

Introduction

A population of cells is considered to undergo *collective migration* when neighboring cells move coherently instead of autonomously. The mechanical and chemical signals underlying individual cell migration both *in vitro* and *in vivo* are very well-studied, yet how these processes translate to the collective migration of a population of cells is poorly understood [Ridley et al., 2003, Friedl, 2004]. Collective migration occurs in a wide variety of biological processes, including tissue repair [Maini et al., 2004] and tumorigenesis [Anderson, 2005], so a better understanding of it is fundamental for improving how we treat different types of wounds and prevent the spread of invasive cancers [Friedl and Gilmour, 2009, Pastar et al., 2014].

This thesis will focus on the collective migration of epithelial cells, which comprise the tissue that lines organs of the body, although collective migration also arises in many other cell types [Stuelten et al., 2018]. An illustrative example of collective migration occurs during the re-epithelialization phase of wound healing, in which keratinocytes from the epidermis migrate into the wound area to re-establish the outermost layer of skin [Clark and Henson, 1995]. The epidermal layer of skin is a key component of the immune system, as it provides the main barrier that prevents pathogens from entering the body, so re-epithelialization is a critical part of the wound healing process [Pastar et al., 2014]. Failure of this process often lead to wounds that are prone to infection, which can cause disastrous results that include death or amputation [Tsang et al., 2003].

The full wound healing process is generally split into three overlapping stages: inflammation, re-epithelialization, and tissue remodeling [Clark and Henson, 1995]. The inflammatory phase of wound healing consists of the skin's initial response to injury, in which monocytes and neutrophils enter

the wound area. These cells will remove pathogens from the wound area, form fibrin clots to seal disrupted blood vessels, and remove any remaining cell debris [Werner and Grose, 2003]. When re-epithelialization is ready to begin, these inflammatory cells release growth factors such as epidermal growth factor (EGF) and tumor necrosis factor (TNF) to attract epidermal keratinocytes into the wound area. These cytokines and growth factors cause keratinocytes to differentiate from immobile epithelial cells to highly migratory cells that collectively migrate into the wound [Pastar et al., 2014]. These migratory keratinocytes show marked differences in their cytoskeleton and receptor compositions due to activity along the mitogen-activated protein kinase (MAPK) pathways, which are known to cause cells to increase motility in certain contexts [Seger and Krebs, 1995]. In healthy tissue, these activated keratinocytes will migrate over the denuded area until they meet other cells in the wound area. Once the epidermis has been re-established, the re-epithelialization phase of wound healing is complete, and the tissue remodeling phase begins. Basal keratinocytes are again stimulated by various growth factors (including EGF, heparin-binding (HB)-EGF, and transforming growth factor (TGF)-α) but now differentiate into highly proliferative cells as a means to supply cells to fill in the wound [Pastar et al., 2014].

From above, we see that epidermal keratinocytes play many crucial roles during the wound healing process. They are able to do so by properly responding to different chemical and physical cues and differentiating into the appropriate phenotypes. For example, these epithelial cells are firmly connected to their neighbors through physical connections called adherens junctions or cell-cell adhesions. Such connections between cells hold them together in tissue to provide the barrier function of the epidermis. During re-epithelialization, cells lose some of these cell-cell adhesions to allow their entry into the wound site. But the amount of regulation is intricate, as the cells do maintain some connections to their neighbors during this process [Clark and Henson, 1995, Perez-Moreno et al., 2003]. Recent studies have shown these physical connections to be necessary for some cellular processes to occur, such as migration during morphogenesis, but how cells use these connections during wound healing is not well understood [Perez-Moreno et al., 2003].

Various growth factors such as EGF are crucial for regulating the re-epithelialization process by causing keratinocytes to differentiate via activation of the MAPK signaling pathways. The importance of these pathways is best exhibited during chronic wounds, which are defined as wounds that never properly heal. This is often due to the failure of re-epithelialization, which may occur

because of the failure of EGF signaling in cells [Brem et al., 2007, Pastar et al., 2014]. Some proposed therapies for chronic wounds have included treating the wounds topically with EGF, but these methods have proved ineffective in clinical trials [Tsang et al., 2003]. This therapy is likely ineffective because of a lack of keratinocyte sensitivity to EGF in chronic wounds [Falanga et al., 1992], suggesting that effective therapies must focus on the bioavailability of keratinocytes to EGF and other signaling chemicals. HB-EGF and insulin-like growth factor (IGF) are known to act synergistically *in vivo* to promote keratinocyte proliferation, suggesting that treatment with multiple growth factors may be an effective strategy to promote healthy re-epithelialization in chronic wounds [Marikovsky et al., 1996]. We thus observe that understanding how cells respond to these growth factors, as well as the effects of these growth factors on migration, will greatly enhance future therapies for chronic wounds.

The most popular experimental protocol used to study collective migration currently is the two-dimensional scratch assay. Such a protocol allows for the study of wound healing events without using test animals [Monsuur et al., 2016]. In these experiments, a layer of cells is grown to confluence in a well plate, and then a significant portion of the cells is removed (by pipette scratching, removal of a boundary, etc.). Time-lapse microscopy is then used to track the remaining population as it migrates into the denuded area over time [Justus et al., 2014]. These experiments have proved vital in studying cells from many layers of the skin, including dermal fibroblasts, endothelial cells that line blood vessels, epidermal keratinocytes, and the interactions between dermal fibroblasts and epidermal keratinocytes [Monsuur et al., 2016]. Scratch assays have also proved successful in aiding our understanding of cancer, as a previous study investigated the rates of diffusion and proliferation of fibroblasts and melanoma cells that were co-cultured [Haridas et al., 2017]. An alternative experimental protocol to study cell migration during wound healing and cancer invasion includes the human skin equivalent, in which a three dimensional matrix is grown from primary human keratinocytes [McGovern et al., 2013]. This protocol is expensive, however, which has prevented its wider use.

Recent experiments have focused on the influence of the MAPK signaling cascades on wound healing. Experimental wounding assays of madine darby canine kidney cells (MDCKs) in [Matsubayashi et al., 2004] yielded a transient pulse of extracellular signal-regulared kinase (ERK) 1/2 (a specific MAPK protein) activity in the cell sheet that only lasted for a few minutes. This single

pulse of activity was followed by a slow wave of activity that propagated from the wound margin to submarginal cells over the course of several hours. The second wave was determined to be crucial for regulating MDCK sheet migration. The authors of [Matsubayashi et al., 2004] proposed that these fast and slow waves of ERK 1/2 activity could be caused by the production of reactive oxygen species (ROS) and EGF, respectively. Similar experiments with fibroblasts also demonstrated this first transient wave of ERK 1/2 activity but not the following slow wave or resulting collective sheet migration, demonstrating that the response to the wound is cell-type specific. The authors of [Chapnick and Liu, 2014] found that human keratinocyte (HaCaT) cells exhibit ERK 1/2 activity primarily at the wound margin during similar experimental wound healing assays with a high density in response to treatment with transforming growth factor-β (TGF-β). While these patterns of biochemical activity are known to be crucial in regulating cell migration in various contexts, a quantitative understanding of biochemical activation and its influence on cell population phenotype is currently lacking.

Partial differential equation (PDE) models are now a vital tool in interpreting the mechanisms underlying experimental data from scratch assay experiments [Thackham et al., 2008b]. For example, previous mathematical models have been used to determine the rates of diffusion and proliferation of mesothelial cells [Maini et al., 2004], the effect of cell density on the diffusive rates of prostate cancer cells [Jin et al., 2016], and how wounds of different shapes and sizes influence healing rates [Arciero et al., 2013, 2011]. While PDE models have proved successful in describing collective migration, most mathematical models are very simple and neglect the plethora of biological complexity underlying how a group of cells can coherently migrate as one unit. This lack of complexity has limited the ability of these models to interpret more complicated aspects of collective migration, such as the effects of physical connections between cells or patterns of biochemical activity.

The goal of this dissertation is to develop and analyze mathematical models that describe some of the relevant physical and chemical cues underlying the collective migration during epidermal wound healing more realistically than previous models. Focus will also be given on how numerical error in approximating a PDE solution propagates in a deterministic inverse problem methodology that is commonly used to interpret experimental data. This will include the derivation of nonlinear diffusion equations to interpret the influence of cell-cell adhesions on keratinocyte migration in response to EGF treatment. We will also develop a framework to study biochemically-structured

transport equations, in which different patterns of biochemical activation influence transport into the wound. These equations admit self-similar traveling wave solutions when the parameters are independent of biochemical activation. For scenarios where the transport parameters depend on biochemical activation, we derive a time-dependent equation that governs how the cell population migrates in time and space in response to temporal biochemical patterns. This time-dependent equation can be simulated numerically very quick and easily. We then study how error in numerically computing PDE models influences a common inverse problem methodology used to estimate unknown parameters. In Section 1.1, we discuss some important previous models of cell migration. In Section 1.2, we briefly discuss the results from each chapter in this dissertation.

1.1   Previous Mathematical Models of Wound Healing

Mathematical models of wound healing have arisen in many different forms over the past 25 years. One of the most common types of equations that arise are reaction-diffusion equations, which are written as:

$$\frac{\partial u}{\partial t} = D\nabla^2 u + R(u),$$

where $u \in \mathbb{R}^n$ denote a vector of unknown quantities under consideration, D denotes a diagonal matrix representing the diffusive spread of the quantities, and $R : \mathbb{R}^n \to \mathbb{R}^n$ is an operator representing how the different quantities interact [Fife, 1979]. Such equations arise in a variety of different contexts, notably excitable media in neuroscience [Keener and Sneyd, 2009], species migration in ecology [Hastings et al., 2005], and biology [Murray, 2002]. They admit a variety of interesting mathematical behaviors and patterns, most notably pattern formation and traveling waves solutions.

A traveling wave solution to a PDE is a solution that maintains a fixed profile over time and moves in one direction at a constant speed. If x is one-dimensional, then a traveling wave solution to a PDE is written as $U(z) = u(x, t)$ where $z = x - ct$ and $U(z)$ denotes the traveling wave profile [Murray, 2002]. Here c denotes the speed of the traveling wave solution. Notably, this change of variable transforms a PDE into an ordinary differential equation (ODE) that defines the solution's profile. The experimental data observed in wound healing experiments often moves into the wound area at a constant speed, which has prompted the significant study of traveling wave solutions to

PDE models in the wound healing literature [Ai, 2008, Dale et al., 1994, Horstmann and Stevens, 2004, Landman et al., 2005, Li and Wang, 2011].

One of the earliest mathematical studies into epidermal wound healing is presented in [Sherratt and Murray, 1990], in which different stimuli for the enhanced rate of epidermal mitosis during wound healing was considered. The stimulus considered includes the presence of some biochemical activator or inhibitor of mitosis. The final model is written as

$$u_t = D\nabla^2 u + s(c) \cdot u \cdot (2 - (u/K)) - ku \tag{1.1}$$
$$c_t = D_c\nabla^2 c + f(u) - \lambda c.$$

where subscripts denote differentiation, u denotes a cell density in time and space, c denotes a chemical concentration in time and space, the parameters D and D_c denote the rates of diffusion of cells of chemical respectively, K is the unwounded cell density, and k and λ denote natural death and decay terms for the cells and chemical, respectively. The function $s(c)$ denotes chemical control of mitosis, which the authors consider to be increasing or decreasing with c to represent an activator or inhibitor of mitosis, respectively. The term $f(u)$ denotes how cells produce the chemical. In the two cases where the chemical either enhanced or inhibited the mitotic activity of cells, the authors were able to fit the models to *in vivo* experimental data on radial wound closure from [van den Brenk, 1956]. This suggests that biochemical regulation of mitosis is a crucial aspect of epidermal wound healing.

Equation (1.1) was used to study corneal epithelial wound healing in [Dale et al., 1994]. This study considered $c(x, t)$ to denote the concentration of EGF, so this represents a scenario where the chemical is an activator of mitosis. The authors also analyzed the speed and existence of traveling wave solutions to the equation by deriving a complicated dispersion relation that can be solved numerically. By doing so, the authors showed that their model had traveling wave solutions at a speed that matched experimental data. Because traveling wave solutions to (1.1) exist when the rate of diffusion is independent of EGF concentration, the authors also concluded that the effect of EGF on cell migration is negligible in the corneal layer of the epidermis.

Fisher's Equation has also been widely used in the wound healing literature due to its simple

nature [Cai et al., 2007, Jin et al., 2016, Maini et al., 2004]. Fisher's Equation is written as

$$u_t = Du_{xx} + \lambda u \left(1 - \frac{u}{K} \right) \tag{1.2}$$

where u denotes a density of cells, D represents their diffusivity, λ represents the rate of proliferation, and K denotes the carrying capacity. Note that Equation (1.2) is a simplification of (1.1). R.A. Fisher introduced this equation in 1937 to model the advance of an advantageous gene in a population [Fisher, 1937] and showed that Equation (1.2) has a positive and monotonic profile for $|c| \geq 2\sqrt{D\lambda/K}$, which is biologically relevant when $u(t,x)$ denotes a population of cells. Kolmogoroff proved in 1937 that any solution to (1.2) with a compactly-supported initial condition (such as a wound) will converge to this solution [Kolmogoroff et al., 1937]. Fisher's Equation was used to interpret wound healing assays of mesothelial cells in [Maini et al., 2004]; since then, many variations of this equation have been considered in the math biology literature [Cheeseman et al., 2014, Curtis and Bortz, 2012, Gourley, 2000, Hammond and Bortz, 2011, Johnston et al., 2015, Yahyaoui et al., 2017].

Another important class of mathematical models from the wound healing literature include chemotaxis equations, in which biological organisms move in response to a chemical gradient. Chemotaxis equations where first considered in [Keller and Segel, 1971] to mimic the chemotactic migration of bacteria. This equation is given by

$$b_t = (\mu(s)b_x)_x - (b\chi(s)s_x)_x$$
$$s_t = -k(s)b + Ds_{xx},$$

where b denotes the concentration of bacteria, s denotes the chemoattractant chemical, μ is the rate of bacterial diffusion, χ is the rate of chemotaxis, k is the rate of consumption of chemoattractant, and D is the rate of chemoattractant diffusion. In this paper, criteria for the speed and existence of traveling wave solutions were investigated. Since then, chemotaxis equations have been widely used to model the movement of cells in response to chemicals, such as EGF [Ai et al., 2015, Hillen and Painter, 2009, Landman et al., 2008, Simpson et al., 2006]. These equations again often lead to traveling wave solutions and the hyperbolic nature of these equations can lead to solutions with

discontinuous profiles (as opposed to the parabolic Equation(1.1)) [Landman et al., 2008, Newgreen et al., 2003].

It should lastly be noted that the continuous PDE models considered above are approximations to the true wound healing process, as any continuous model must assume that cells in the population are infinitely small, which is of course unrealistic. Due to increases in computing ability in recent years, agent-based models of wound healing are becoming a more and more popular method to study experiments [Johnston et al., 2012, 2014, 2016, Simpson et al., 2011]. In these models, a finite number of agents are used to represent the cells in a population. The migration of the whole population is then governed by sets of rules for how each agent in the population can move and produce more agents. For cells of small diameter, these rules governing cell migration can often be approximated with continuum models. For example, in [Simpson et al., 2014], simple and realistic rules are given that can be approximated by Fisher's Equation.

1.2   Dissertation Overview

When epithelial cells are migrating during tissue repair, they maintain physical cell-cell adhesions to the nearest neighbors. As these cell-cell adhesions connect cells to the actin cytoskeletons of their neighbors, these physical connections influence population-wide migration [Friedl and Gilmour, 2009, Clark and Henson, 1995]. For instance, the loss of these connections between cells that gives tumorous epithelial cells their invasive phenotype and ability to develop into mesenchymal cells during the epithelial-to-mesenchymal transition (EMT), a crucial process in tumorigenesis. This rise in invasive capability allows the cancerous mesenchymal cells to invade local and distant areas of the body [Hollier et al., 2009]. The observation that loss of physical connections gives rise to invasiveness in cancer biology has caused several mathematical models of wound healing to assume that these cell-cell adhesions also hinder healthy cell migration during wound healing [Anguige and Schmeiser, 2009, Johnston et al., 2012, Mi et al., 2007]. In Chapter 2, we will investigate this suspect hypothesis by deriving two nonlinear diffusion equation models with different assumptions on the main role of cell-cell adhesions during keratinocyte migration into the wound. The first model assumes that cell-cell adhesions *hinder* migration into the wound, while the second model assumes that cell-cell adhesions *promote* migration into the wound. We will fit both of these models

to experimental data on the leading edge propagation of keratinocyte and observe that the latter model can better fit experimental data. This observation prompted us to perform the same scratch assay in keratinocyte populations with decreased levels of cell-cell adhesion expression. We show that these new experiments are unable to maintain healthy rates of migration into the wound. Leader and follower cells are known to exist in migrating cell populations, and leader cells are often proposed to act by pulling follower cells into the wound with them [Chapnick and Liu, 2014]. To the best of our knowledge, this is the first use of a mathematical model to demonstrate and quantify this behavior.

While PDE models have proved crucial in interpreting wound healing experiments, they have not yet been used to quantify the effects of biochemical activation on cell migration and proliferation. In Chapter 3, we aim to bridge this gap by deriving and analyzing biochemically-structured reaction-diffusion equations. This chapter will begin with an introduction to continuous stage-structured equations and demonstrate how to solve them analytically using the method of characteristics. We will then discuss the relevance of a biochemically-structured trait in the context of wound healing and derive a biochemically-structured Fisher's Equation. All terms in the equation will be described using terminology from [de Roos, 1996]. We will then show the existence of unique self-similar traveling wave solutions to a simple version of this equation. For a more realistic equation, we will derive a time-dependent version of Fisher's Equation (in time and space) to account for how various patterns of biochemical activation influence the cell population's migration into the wound. We will use numerical simulations to demonstrate the effects of different patterns of biochemical activation on cell migration into the wound. These simulations are aided by some analytical criteria to deduce for which parameter values we will observe the cell population become activated. From these simulations, we observe that a healthy combination of cell diffusion and proliferation is needed for the cell population to migrate farthest into the wound, which occurs at the threshold value at which all cells in the population become activated.

The PDE models used to interpret wound healing assays are often too complicated to solve with a closed-form solution. Instead, numerical solutions to these models are typically used to fit their approximations to the equation solution to experimental data. In Chapter 4, we take a numerical analysis approach to investigate how numerical error influences an inverse problem methodology that is commonly used in the sciences to estimate the parameters from a system under consideration.

This is done on a simple hyperbolic advection equation, where numerical difficulties for simulating discontinuous solutions are well documented [Thackham et al., 2008a]. We will show how numerical solvers of different orders cause the order of a numerical cost function to differ in various scenarios. We then demonstrate how residual analysis is altered when numerical error is the dominant form of error, and we derive an autocorrelative statistical model to correct these residual terms for first order methods. We also compute confidence intervals to demonstrate how sensitive these are to the numerical solvers and precision. We demonstrate that the autocorrelative statistical model is often able to correct confidence interval computation.

We finish in Chapter 5 by discussing the relevance of this work to future clinical trials for chronic wounds and discussing future work.

Chapter 2

Modeling Keratinocyte Wound Healing Dynamics: Cell-cell Adhesion Promotes Sustained Collective migration[1]

Collective cellular migration is a critical component of many biological processes, including embryo development [McLennan et al., 2015], tissue repair [Maini et al., 2004], and tumorigenesis [Anderson, 2005]. A group of cells is considered to be migrating collectively when neighboring cells adopt similar migration directions. In many cases of collective migration, neighboring cells are physically linked through some form of cell-cell junction [Friedl and Gilmour, 2009]. The exact role of these cell-cell junctions has not been identified in this process, nor is it clear how temporal regulation of these junctions may influence the migration behavior of a collectively migrating group of cells. To date, cell-cell adhesion is believed to act as a component for cell-cell coupling during epithelial migration [Ilina and Friedl, 2009], but has also been shown experimentally to affect cell migration both positively [Geisbrecht and Montell, 2002, Hazan et al., 2000] and negatively [Friedl and Gilmour, 2009, Hazan et al., 2000] in different situations. However, the majority of previous mathematical models of cell migration assume that cell-cell adhesion affects cell migration negatively [Anderson, 2005, Anguige and Schmeiser, 2009, Johnston et al., 2012]. In this study, we investigate the role of cell-cell adhesion during *in vitro* keratinocyte wound healing.

During the re-epithelialization phase of wound healing in mammalian skin, the migration of both fibroblasts and keratinocytes is required to reestablish the physically disrupted barrier between the organism and the surrounding environment [Clark and Henson, 1995]. In this process, collective

---

[1]This chapter has been published in Nardini et al. [2016].
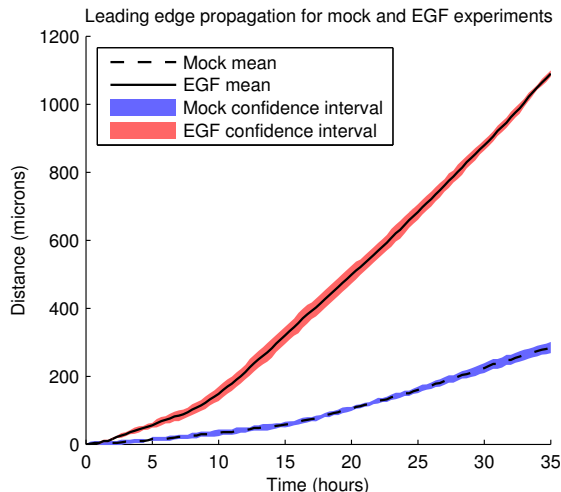
Figure 2.1: Mean leading edge propagation of mock (black dashed line in blue strip) and EGF (black line in red strip) experiments between 0 and 35 hours. Our leading edge computation finds where the normalized cell sheet profile reaches a certain value and is discussed in Section 2.1.3. The value used in this figure is 0.3. The total height of the colored strips correspond to two standard deviations of the leading edge data over time.

migration of layers of keratinocytes that are physically linked through adherens junctions allows for the completion of wound healing and reformation of the epidermis [Ilina and Friedl, 2009]. Adherens junctions are composed of E-cadherin, α- and β- catenin, and are bound to the actin cytoskeleton through the adaptor protein vinculin, which binds α-catenin to filamentous actin [Friedl and Gilmour, 2009, Haley and Gullick, 2008, Juliano, 2002]. Thus, adherens junctions serve as a bond between both the plasma membranes and actin cytoskeletons of adjacent cells.

Our experimental system of artificially constructed sheets of spontaneously immortalized human keratinocyte (HaCaT) cells (described previously in [Chapnick and Liu, 2014]) exhibits many similarities to *in vivo* keratinocyte behavior. These sheets of physically connected cells migrate into a wound area in response to epidermal growth factor (EGF) treatment in the same way that multiple layers of keratinocyte layers migrate during *in vivo* wound healing. Whereas fibroblasts secrete EGFR ligands *in vivo* [Werner et al., 2007], addition of exogenous EGF allows us to stimulate this keratinocyte migration into the wound. The stimulatory effect of EGF treatment on cell migration in this system is demonstrated in Figure 2.1, where we have displayed the leading edge propagation of untreated (denoted as mock) and EGF-treated keratinocyte sheets. The EGF-treated cell sheets

migrate more than three times as far as the mock cell sheets after 35 hours.

A plethora of recent quantitative studies have analyzed *in vitro* wounding assays to investigate aspects of collective cell migration for various cell types [Arciero et al., 2011, 2013, Cai et al., 2007, Landman et al., 2007, Posta and Chou, 2010, Poujade et al., 2007, Simpson et al., 2014]. For example, the continuum model developed in [Arciero et al., 2011] investigated how wound area, shape, and aspect ratio influence gap wound healing as a means to improve predicted wound healing times in intestinal enterocyte cells [Arciero et al., 2013]. In this current study, we investigate the role of cell-cell adhesion on collective cell migration during *in vitro* wound healing assays of keratinocyte cell sheets. To do so, we develop and compare two competing mathematical models to describe how cell sheets migrate into the wound. Both models are nonlinear diffusion equations based on assumptions of how cell-cell adhesion influences the space filling response of cells to a wound. Our first model (Model H) assumes cell-cell adhesion *hinders* migration into the wound through a drag force, while the other model (Model P) assumes cell-cell adhesion *promotes* this migration with a pulling force. We simulate both models with time-dependent rates of cell-cell adhesion to accurately fit the leading edge propagation of experiments from our model system. Model P is more robust than Model H to changes in the definition of the leading edge, so we determine it to be an appropriate model of keratinocyte migration during wound healing. We also show that it can reliably predict leading edge propagation from our experimental system.

Performing the same experimental protocol on cell sheets with decreased α-catenin expression demonstrates that cell sheets with weakened cell-cell junctions initially enter the wound area rapidly but do not display collective migration and are unable to maintain migration into the wound. Intact adherens junctions thus allow the cell sheet to sustain the wound healing response for long periods of time through collective migration. The agreement between sheet migration behavior in wild-type sheets and model P simulations, in addition to the inability of cell sheets with decreased α-catenin expression to maintain migration, leads us to conclude that cell pulling, which is mediated by cell-cell adhesion, promotes sustained collective migration during wound healing.

In Section 2.1, we present our two model derivations based on different assumptions for the role of cell-cell adhesion on cell migration. In Section 2.2, we demonstrate how both models can fit experimental leading edge propagation data, but that Model P is robust to changes in the leading edge definition. We then use keratinocyte sheets with decreased α-catenin expression to demonstrate

that cell-cell adhesion is needed to maintain a sustained wound healing response. We discuss our conclusions in Section 2.3 and discuss the implications of these results as well as plans for future work in Section 2.4.

2.1 Model Development

To investigate the role of cell-cell adhesion on wound healing, we develop two contrasting models in Section 2.1.1. Model H assumes that cell drag *hinders* migration, while Model P assumes that cell pull *promotes* migration. In Section 2.1.2, we discuss the data collection procedure. In Section 2.1.3, we discuss our parameter estimation method.

2.1.1 Nonlinear Diffusion Models

Let $u(t, x)$ denote the normalized cellular sheet density at location x and time t. We first discretize our solution domain with a uniform grid for both time and space, *i.e.*, we set $x_i = x_{\min} + i\Delta x, i = 0, ..., M - 1$ and $t_j = j\Delta t$, $j = 0, ..., N - 1$, where M and N denote the number of spatial and time points used, respectively. We simplify notation by writing the cell density at a fixed location over time as $u_i = u(t, x_i)$.

Similar to [Anguige and Schmeiser, 2009], we first define discrete transition rates for our two models and then derive their continuum limits. The transition rate, denoted τ_i^+, is the rate at which u_i (the cell density at location x_i) will move forward to the location x_{i+1}. Both models incorporate two terms in their transition rates that denote space-filling migration into the wound and the effect of cell-cell adhesion on this migration. For Model H, in which cell-cell adhesion hinders migration, we denote the transition rate $\tau_{i,H}^+$ as:

$$\tau_{i,H}^+ = \frac{D(1 - u_{i+1})(1 - \alpha u_{i-1})}{\Delta x^2} \tag{2.1}$$

where D denotes the rate of cell diffusion and the first term $(1 - u_{i+1})$ represents space-filling migration of the cell sheet (*i.e.*, the cell density will migrate towards lower density areas). The parameter α denotes the rate of cell-cell adhesion, and thus the second term $(1 - \alpha u_{i-1})$ in Equation (2.1) represents the neighboring density in the direction opposite migration decreasing the transition rate. In this scenario, strong physical connections to other cells will *hinder* cell migration, as the added weight of other cells provides a drag force that impedes migration into the wound. The

transition rates $\tau_{i-1,H}^+, \tau_{i,H}^-$, and $\tau_{i+1,H}^-$ are all defined analogously in Table 1 in the appendix (τ_i^- denotes the transition rate of u_i moving from position x_i to x_{i-1}).

We similarly define $\tau_{i,P}^+$ for model P, in which cell-cell adhesion promotes migration, as:

$$\tau_{i,P}^+ = \frac{D(1 - u_{i+1})(1 + \alpha u_{i+1})}{\Delta x^2} \tag{2.2}$$

where the first term again denotes space-filling migration, and the second term denotes cells located in the direction of migration increasing the transition rate. In this scenario, strong physical connections to other cells will *promote* migration as cells are pulled forward with the cells migrating in front of them. The transition rates $\tau_{i-1,P}^+, \tau_{i,P}^-$, and $\tau_{i+1,P}^-$ are also given in Table 1 in the appendix.

Given these transition probabilities, we derive the continuum limit by setting the time derivative of u_i equal to the flux into and out of x_i by:

$$\frac{du_i}{dt} = \tau_{i-1}^+ u_{i-1} + \tau_{i+1}^- u_{i+1} - (\tau_i^+ + \tau_i^-)u_i \tag{2.3}$$

and taking the limit as $\Delta x \to 0^+$. We follow the steps of Anguige and Schmeiser [2009] and derive the continuum limit for Model H as

$$u_t = ((D + 3\gamma(u - 2/3)^2 - 4/3\gamma)u_x)_x \tag{2.4}$$

where $\gamma = D\alpha$ represents our new rate of cell-cell adhesion. We note that backwards diffusion may occur with Equation (2.4) if $4\gamma > 3D$ and address this further in Section 2.1.3. In the appendix, we derive the continuum limit for Model P to be

$$u_t = ((D + \gamma u^2)u_x)_x \tag{2.5}$$

where again $\gamma = D\alpha$. This diffusion term is always positive, so there is no potential for backwards diffusion to occur.

Adherens junctions change dynamically in response to a cell's environmental cues, such as the presence of EGF and wound induction [Friedl and Gilmour, 2009]. We will investigate time-dependent rates of cell-cell adhesion (denoted as $\Gamma(t)$) in response to EGF treatment, as previous

studies have suggested using time-dependent parameters for modeling cell migration in response to EGF treatment [Johnston et al., 2015]. Because the cell sheets transition from stationary to highly motile after wound induction and EGF treatment, we consider decreasing $\Gamma(t)$ terms for Model H, as a decrease in cell drag should allow the cell sheet to increase its rate of migration. Similarly, we consider increasing $\Gamma(t)$ terms for Model P. We write our final mathematical models as:

$$u_t = ((D + 3\Gamma(t)(u - 2/3)^2 - 4/3\Gamma(t))u_x)_x \qquad \text{(Model H)} \qquad (2.6)$$

and

$$u_t = ((D + \Gamma(t)u^2)u_x)_x, \qquad \text{(Model P)} \qquad (2.7)$$

and we will present the terms considered for $\Gamma(t)$ in Section 2.2.1. We note that cell proliferation is negligible in our experimental system, as demonstrated in the appendix, so terms for proliferation have not been included in our model. We briefly review the two models, along with their terms and assumptions in Table 2.1. A list of all parameters used throughout this study is also given in Table 2.2.

Model features	H		P	
Cell-cell adhesion	Hinders migration		Promotes migration	
Treatment	mock	EGF	mock	EGF
$\Gamma(t)$	γ_1	$\gamma_1 + \gamma_2(1 - t/t_{final})$	γ_1	$\gamma_1 + \gamma_2 t/t_{final}$
Equation	(2.4)	(2.6)	(2.5)	(2.7)

Table 2.1: Summary of equations and assumptions relating to Models H and P.

Parameter	Description (units)
D	Baseline rate of diffusion $\left(\text{microns}^2/\text{hr}\right)$
γ_1	Baseline cell-cell adhesion rate between adjacent cells $\left(\text{microns}^2/\text{hr}\right)$
γ_2	Response of the rate of cell-cell adhesion to EGF treatment $\left(\text{microns}^2/\text{hr}^2\right)$

Table 2.2: List of parameters for the two models. Note that the fit values for different simulations are given in Tables 2.3 and 2.5.

If w is the site of the wound, we use an indicator function for our initial condition that is set equal to 1 for $x < w$ and 0 for $x > w$. This initial condition represents a high density behind the wound and no initial cell density in the wound area. Note that in experimental videos, we only

observe a 3.24 mm × 3.24 mm field of view, whereas the experimental domain is actually 7 mm long and 5 mm wide. We accordingly use zero Neumann boundary conditions 4.86 mm behind and in front of the field of view to simulate no flux conditions at the walls of the well plate.

2.1.2 Description of experiments

In vitro wound healing experiments, as previously described in Chapnick and Liu [2014], begin with the growth of a confluent monolayer of cells to a density of about 4,000 cells/mm^2 in a well plate. After 24 hours, half of the cell population on one side of the well plate is manually removed with a pipette tip. EGF-treated populations are globally treated with a concentration of 100 nM of EGF. An ImageXpress MicroXL high throughput wide-field fluorescence microscope was used to image the cell populations every 20 minutes for 48 hours. We performed both mock and EGF experiments in triplicate[2]. In Section 2.2.3, we perform the same protocol in triplicate on a cell line with decreased α-catenin expression using RNAi.

2.1.3 Parameter Estimation

In order to fit our model simulations to data on wound healing, we will use an inverse problem procedure to estimate the parameter vector $\vec{q} = [D, \gamma_1, \gamma_2]^T$ for both models for mock and EGF data. We do so by comparing the locations of the leading edge for both the experimental data and model simulations, as has been done in previous quantitative wound healing studies [Johnston et al., 2014, 2015, Maini et al., 2004, Treloar et al., 2014]. In the literature, there are many different definitions of the leading edge [Chapnick and Liu, 2014, Johnston et al., 2014, Maini et al., 2004]. Similar to [Maini et al., 2004], we denote the leading edge as where the sheet profile reaches a certain value. In Figure 2.2, we have included some typical video snapshots along with their resulting data profile and leading edge designation. To calculate the data profile, we sum over the vertical axis for each image matrix and then normalize by the largest density in the population over each time step. We then calculate the leading edge by finding the location where the leading edge density is equal to some value, β, and denote its location over time as $\ell^{\beta}_{data}(t)$. In Figure 2.2, we display this calculation for $\beta = 0.3$ but will consider $\beta = 0.1, 0.2, 0.3$ throughout this study. Note that the

---

[2]Such an experiment in response to EGF treatment is given in the left frame of Video 1 in the supplementary material.
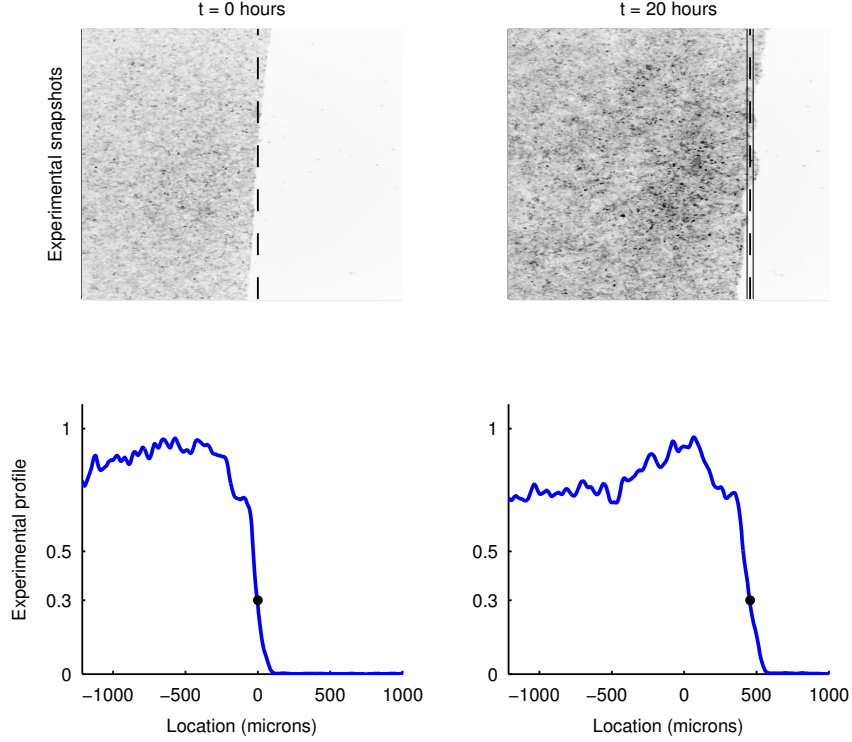
Figure 2.2: Experimental snap shots (top row) of the experiment at $t = 0$ and 20 hours and their resulting data profiles (bottom row), whose calculation is described in the text. The dashed lines against the snapshots denote the calculated leading edge location, and the thin lines denote one sample standard deviation. Note that there is no standard deviation at $t = 0$ because the leading edge for all experiments starts at 0. The dots depicted against the data profiles in the bottom row denote the calculated value for $\ell_{data}^{0.3}(t)$.

cell sheets move right as a coherent whole even through the leading edge is not completely vertical.

Given a model simulation $u(t, x)$, we calculate the leading edge for a relative density of β in a similar manner by approximating the location $\tilde{x}$ such that $u(t, \tilde{x}) = \beta$. We denote the leading edge of a model simulation resulting from a vector of parameter values $\vec{q}$ by $\ell_{model}^{\beta}(t, \vec{q})$.

As a means to estimate the true $\vec{q}$ using our models, we will implement an inverse problem in which we minimize a cost function between the model and data leading edge locations. In this work, we will use the cost function given by an ordinary least squares estimate:

$$J(\vec{q}) = \sum_{j=1}^{N} |\ell_{model}^{\beta}(t_j, \vec{q}) - \ell_{data}^{\beta}(t_j)|^2,$$

where $\{t_j\}_{j=1}^{N}$ are the time points sampled and N is the number of time points considered. Note

18
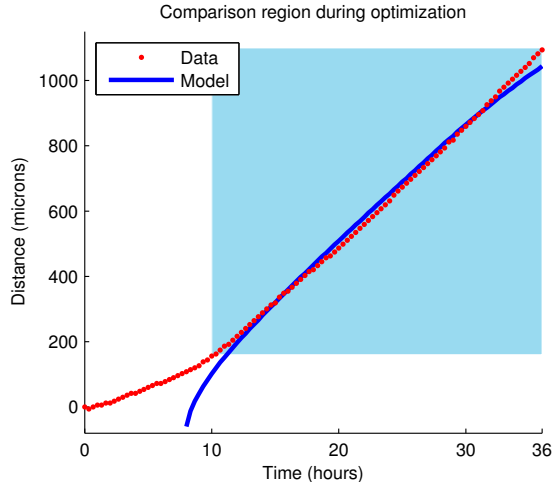
Figure 2.3: Example of fitting the leading edge propagation of a model simulation (blue curve) to experimental leading edge propagation data (red dots) over time. Model P has been used here for illustration. The blue shaded region denotes where we compare the model to the data.

that the experiment is performed for 48 hours, but we only compare the model to data between 10-35 hours for EGF experiments because there is an initial lag period before cell migration, and some leading edges run off the field of view after 35 hours. For mock experiments, we compare the models and data after $t = 20$ hours, because there is a longer initial lag period and cells do not run off the field of view. For example, in Figure 2.3 the blue rectangle depicts when we compare the model and data for the cost function. To find the $\hat{q}$ estimate that minimizes the cost function, we use the Nelder-Mead algorithm as implemented in MATLAB's fminsearch command. For Model H simulations, we use a barrier function to restrict $4D > 3(\gamma_1 + \gamma_2)$ to ensure that backwards diffusion does not occur (as mentioned in Section 2.1.1).

2.2 Results

We now investigate the performances of Models H and P in describing keratinocyte migration during wound healing. We determine the form of the time-dependent rate of cell-cell adhesion, $\Gamma(t)$, for mock and EGF experiments in Section 2.2.1 and then fit $\ell^{\beta}_{model}(t, \vec{q})$ to $\ell^{\beta}_{data}(t)$ for EGF experiments for $\beta = 0.1, 0.2, 0.3$ with Models H and P to investigate how the models respond to changes in β. Model P is more robust than Model H to changes in the leading edge definition, so we declare it a more appropriate model. We demonstrate the predictive capability of Model P in Section 2.2.2.
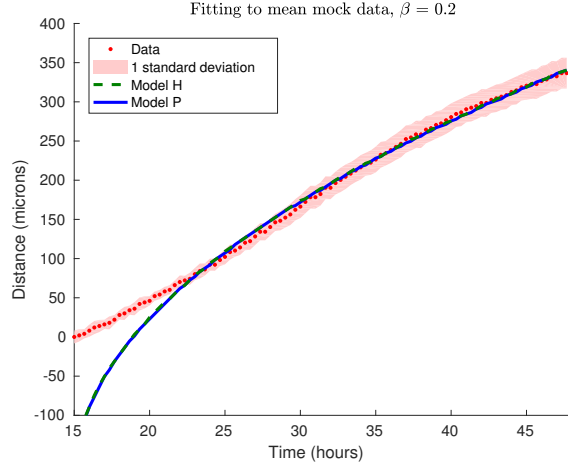
Figure 2.4: Fitting the model leading edge propagation to mean experimental leading edge propagation data. The red region denotes a 95% confidence interval for the experimental data, the green dashed line denotes the leading edge propagation of Model H, and the solid blue line denotes the leading edge propagation of Model P. A constant rate of cell-cell adhesion is used in both model simulations shown.

β	Model H parameters		Model P parameters	
	D ($\mu\mathrm{m}^2/hr$)	γ_1 ($\mu\mathrm{m}^2/hr$)	D ($\mu\mathrm{m}^2/hr$)	γ_1 ($\mu\mathrm{m}^2/hr$)
0.2	32300 ± 600	21300 ± 400	10.8 ± 0.2	20000 ± 3300

Table 2.3: Mock $\hat{q} = [\hat{D}, \hat{\gamma}_1]^T$ estimate table when fitting $\ell^{0.2}_{model}(t, \vec{q})$ to $\ell^{0.2}_{data}(t)$ for Models H and P. Estimates are calculated by fitting the models to the mean leading edge distance traveled and the standard deviation is calculated from the standard deviation of the three estimates used to fit the individual data sets.

Inspired by the results of Section 2.2.1, we perform the same experimental assay (as described in Section 2.1.2) on cell sheets with decreased α-catenin expression in Section 2.2.3 and observe that these sheets are unable to mediate a sustained wound healing response without intact adherens junctions.

2.2.1   Parameter estimation

Models H and P can both accurately fit $\ell^{0.2}_{data}(t)$ for mock data with a constant rate of cell-cell adhesion ($\Gamma(t) \equiv \gamma_1$), as depicted in Figure 2.4. We used the parameter estimation method outlined in Section 2.1.3 to obtain the estimate values $\hat{q} = [\hat{D}, \hat{\gamma}_1]^T$ for these simulations. The resulting estimate values for both models are depicted in Table 2.3. Note that the two models use very different simulations to achieve these similar leading edge locations: Model H uses large $\hat{D}$ and $\hat{\gamma}_1$ estimates while Model P uses a large $\hat{\gamma}_1$ estimate and a very small $\hat{D}$ estimate.

Because adherens junctions are dynamically changing in response to their environment, we consider four simple terms for the rate of cell-cell adhesion over time in response to EGF treatment: constant, linear, quadratic, and square root. The exact forms for each are given in Table 2.4. Note that the $\Gamma(t)$ terms we consider for model H decrease with time whereas the $\Gamma(t)$ terms for model P increase with time, as these match the assumptions discussed in Section 2.1.1. Best-fit simulations for Models H and P to $\ell_{data}^{0.2}(t)$ using these four different $\Gamma(t)$ terms are depicted in Figure 2.5. As opposed to mock data, both models are unable to match the leading edge propagation data with a constant rate of cell-cell adhesion, as they overestimate the experimental leading edge location from $t = 10 - 25$ hours, and then underestimate the leading edge location afterwards. The linear term for $\Gamma(t)$ appears to be a suitable candidate, as both model fits stay within the 95% confidence interval of the data for the majority of the experiment. This linear cell-cell adhesion term may be indicative of cells constantly increasing or decreasing their surface display of adherens junctions over time in response to EGF treatment. One of the two models fails to fit the experimental data well with the quadratic or square root terms, so we decide to further investigate the two models with a linear rate of cell-cell adhesion throughout the rest of this study.

$\Gamma(t)$	Model H	Model P
constant	γ_1	γ_1
linear	$\gamma_1 + \gamma_2(1 - t/t_{final})$	$\gamma_1 + \gamma_2 t/t_{final}$
quadratic	$\gamma_1 + \gamma_2(1 - t/t_{final})^2$	$\gamma_1 + \gamma_2(t/t_{final})^2$
square root	$\gamma_1 + \gamma_2(1 - t/t_{final})^{1/2}$	$\gamma_1 + \gamma_2(t/t_{final})^{1/2}$

Table 2.4: Different terms considered for the time-dependent rate of cell-cell adhesion, $\Gamma(t)$, when fitting Models H and P to experimental EGF data. The parameter γ_1 denotes a baseline rate of cell-cell adhesion, where γ_2 denotes the response of the rate of cell-cell adhesion to EGF treatment. Models H and P are fit to experimental EGF data using all four terms in Figure 2.5.

In Figure 2.6, we depict best-fit $\ell_{model}^{\beta}(t, \hat{q})$ simulations for Models H and P with a linear term for $\Gamma(t)$ to mean leading edge propagation data, $\ell_{data}^{\beta}(t)$, for $\beta = 0.1, 0.2$, and 0.3. While both models can fit $\ell_{data}^{0.2}(t)$ very well, we observe that model H overestimates $\ell_{data}^{0.1}(t)$ from $t = 15$ to 25 hours and fails to accurately match $\ell_{data}^{0.3}(t)$ after $t = 20$ hours. Model P provides an excellent match to $\ell_{data}^{0.3}(t)$, but underestimates $\ell_{data}^{0.1}(t)$ after $t = 30$ hours. We present the best-fit parameter estimates $\hat{q} = [\hat{D}, \hat{\gamma}_1, \hat{\gamma}_2]^T$ when fitting to $\ell_{data}^{\beta}(t)$ with a linear $\Gamma(t)$ term in Table 2.5 for $\beta = 0.1, 0.2, 0.3$. We observe a wide range of $\hat{q}$ estimates needed for Model H to fit $\ell_{data}^{0.1}(t)$, $\ell_{data}^{0.2}(t)$, and $\ell_{data}^{0.3}(t)$ in comparison to Model P. For Model H, $\hat{D}$ varies 347%, $\hat{\gamma}_1$ varies 160%, and $\hat{\gamma}_2$ varies 882%. For
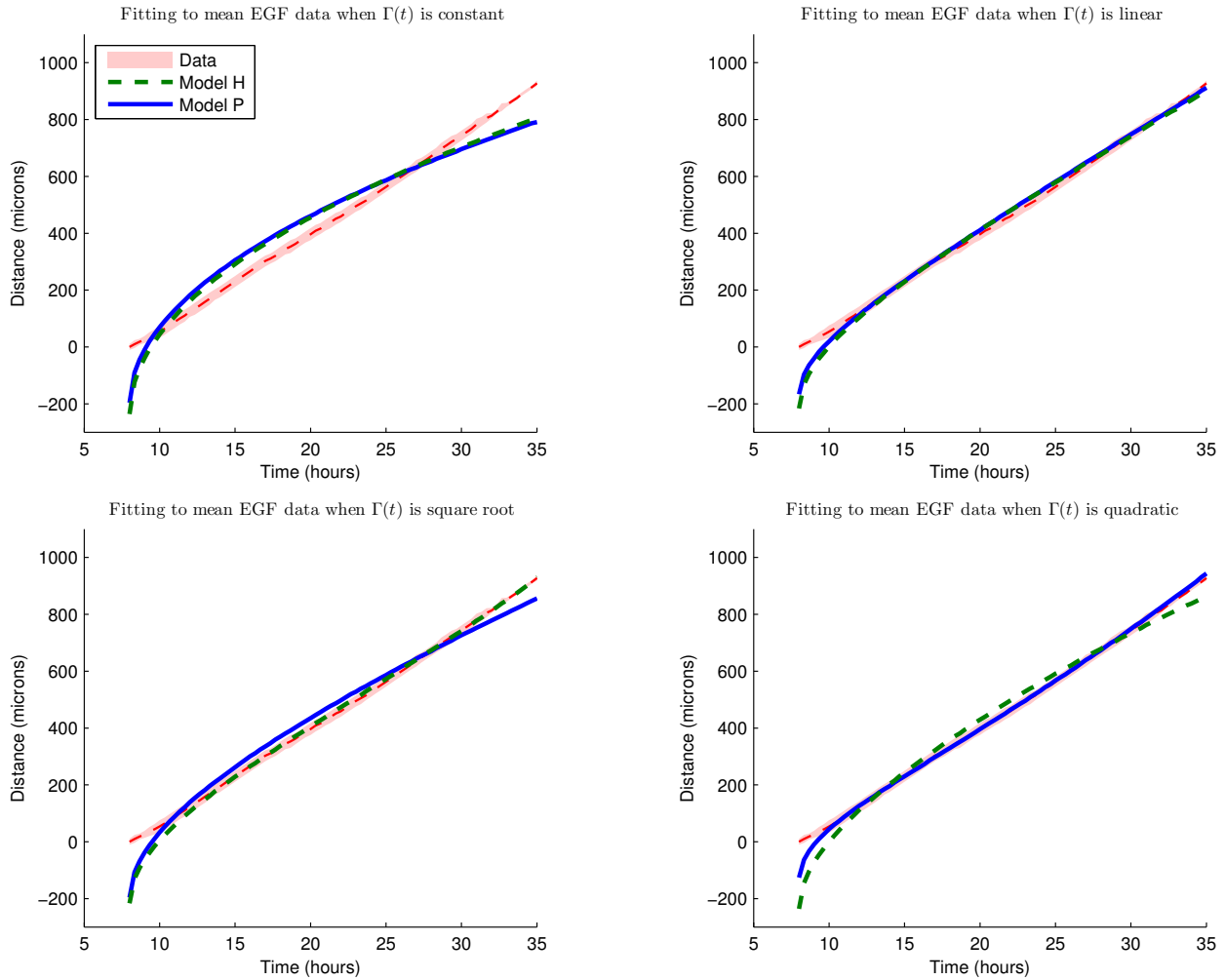
Figure 2.5: Determining a suitable form for the time-dependent rate of cell-cell adhesion, $\Gamma(t)$, by fitting $\ell^{0.2}_{model}(t, \vec{q})$ to $\ell^{0.2}_{data}(t)$. We plot the best-fit simulations, $\ell^{0.2}_{model}(t, \hat{q})$, for Models H and P with different rates of cell-cell adhesion in each frame. The four rates of cell-cell adhesion considered are constant (top left), linear (top right), square root (bottom left), and quadratic (bottom right). Each term is given in Table 2.4. The red region in each figure denotes a 95% confidence interval for the experimental data, the green dashed line denotes the leading edge propagation of Model H, and the solid blue line denotes the leading edge propagation of Model P.
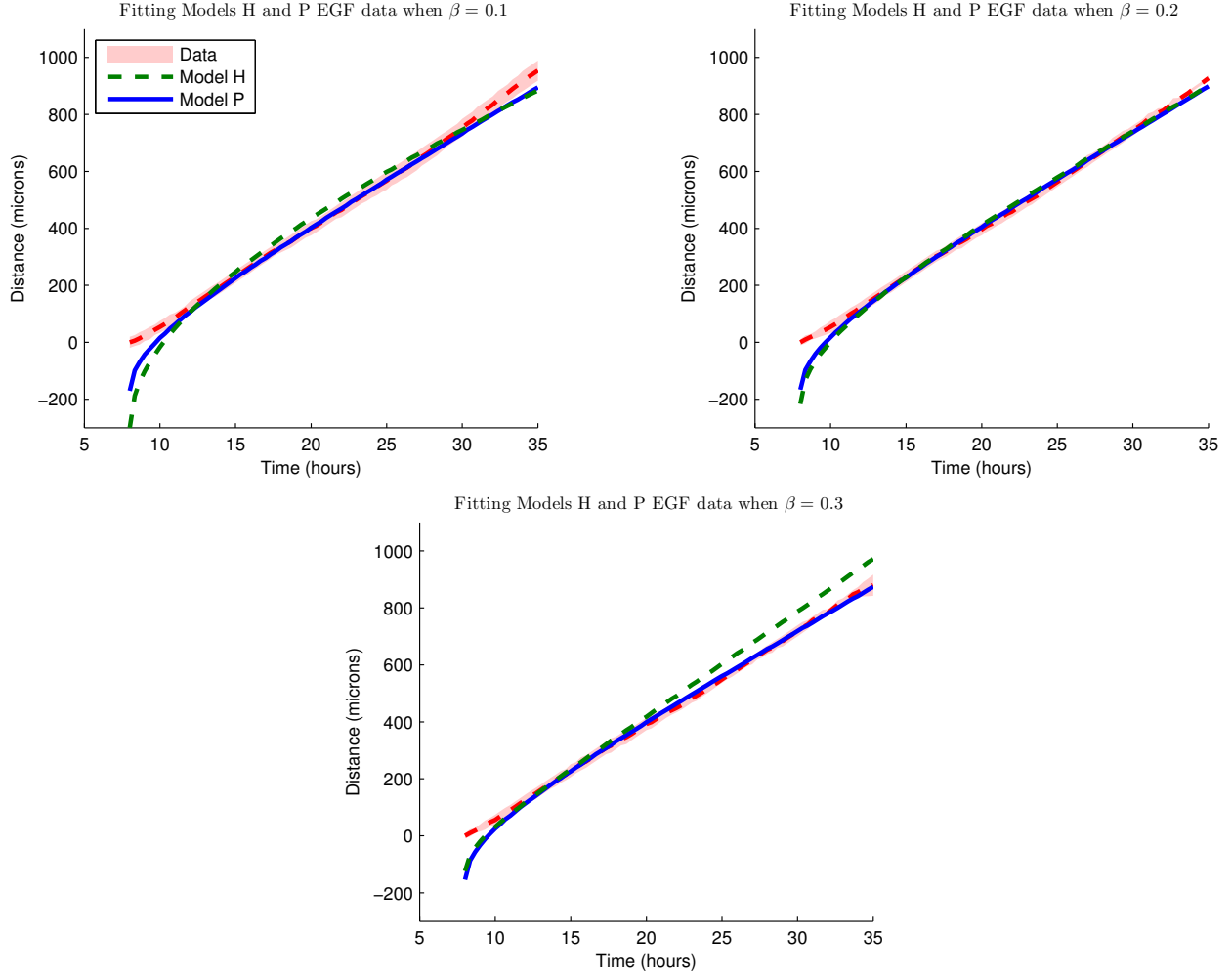
Figure 2.6: Best fits of $\ell^{\beta}_{model}(t, \vec{q})$ to $\ell^{\beta}_{data}(t)$ for $\beta = 0.1$ (top left), 0.2 (top right), and 0.3 (bottom left) when using a linear term for $\Gamma(t)$ for Model H (green dashed line) and Model P (solid blue line). The red shaded region denotes a 95% confidence region for the experimental data.

Model P, $\hat{D}$ varies 146%, $\hat{\gamma}_1$ varies 165%, and $\hat{\gamma}_2$ varies 30%. Model H estimates thus appear more sensitive to changes in the leading edge definition than Model P estimates. We also note that both models use different types of simulations to fit the experimental data, as Model H yields large estimate values for $\hat{D}$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$ whereas Model P uses small estimate values for $\hat{D}$ and $\hat{\gamma}_1$ and large estimate values for $\hat{\gamma}_2$. These differences in the parameter estimates prompted us to investigate the model profiles that led to these simulations.

We depict the three model profiles that best fit $\ell^{\beta}_{model}(t, \vec{q})$ to $\ell^{\beta}_{data}(t)$ for $\beta = 0.1, 0.2, 0.3$ for Models H and P against experimental snapshots over time in Figures 2.7 and 2.8, respectively. Model H exhibits rapid expansion into the wound, which is likely due to its high $\hat{D}$ estimate values.

23

β	Model H			Model P		
	$\hat{D}$	$\hat{\gamma_1}$	$\hat{\gamma_2}$	$\hat{D}$	$\hat{\gamma_1}$	$\hat{\gamma_2}$
0.1	32400 ± 1100	13100 ± 200	10300 ± 700	6.5 ± 13	40 ± 300	152600 ± 22500
0.2	73100 ± 2700	20500 ± 600	32900 ± 1500	16 ± 4	50 ± 1500	169900 ± 9400
0.3	144800 ± 9800	7900 ± 700	100700 ± 1100	13 ± 2	20 ± 400	198900 ± 12600
% change	347%	160%	882%	146%	165%	30%

Table 2.5: EGF $\hat{q} = [\hat{D}, \hat{\gamma_1}, \hat{\gamma_2}]^T$ estimate table when fitting $\ell^{\beta}_{model}(t, \vec{q})$ to $\ell^{\beta}_{data}(t)$ for Models H and P for $\beta = 0.1, 0.2, 0.3$. Estimates are calculated by fitting the models to the mean leading edge distance traveled and the standard deviation is calculated from the sample standard deviation of the three estimates used to fit the individual data sets.

This rapid expansion causes Model H to predict a high cell density in areas that are empty in the experimental snapshot when fitting to $\ell^{0.3}_{data}(t)$ (such as at $x = 1000$ microns when $t = 25$ hours) and to predict a low cell density in areas that appear confluent when fitting to $\ell^{0.1}_{data}(t)$ (such as at $x = 0$ when $t = 25$ hours). Contrast Model H's rapid expansion into the wound with the limited local expansion of Model P's profile in Figure 2.8. Model P maintains a sharp front that propagates into the wound as a coherent unit over time. The sharp front allows Model P to qualitatively predict high and low density areas of the cell sheet well, and its profile simulations do not change significantly when fitting between $\ell^{0.1}_{data}(t), \ell^{0.2}_{data}(t)$, or $\ell^{0.3}_{data}(t)$.

Due to Model P's robustness to different leading edge definitions and its ability to qualitatively match experimental cell sheets, we suggest that Model P is a suitable model for keratinocyte migration during wound healing.

2.2.2 Fitting and predicting of individual data sets

We perform a type of cross validation as a means to demonstrate the predictive capacity of Model P. We do so by fitting $\ell^{\beta}_{model}(t, \vec{q})$ from Model P simulations to $\ell^{\beta}_{data}(t)$ for each data replicate and use these individual best-fit simulations to predict the other data replicate. The three leading edge replicates for $\ell^{0.2}_{data}(t)$ are depicted in the top left frame of Figure 2.9. Because the data sets begin their wound healing response at different times, we predict a data set from a fit simulation by aligning their leading edge locations at $t = 15$ hours (after all data sets have started migrating).

In the top right frame of Figure 2.9, we show the best-fit Model P simulation to $\ell^{0.2}_{data}(t)$ for data set 1, and in the the bottom rows, we demonstrate how this simulation can accurately predict $\ell^{0.2}_{data}(t)$ for data sets 2 and 3. We observe similar results after fitting to the other data sets and after
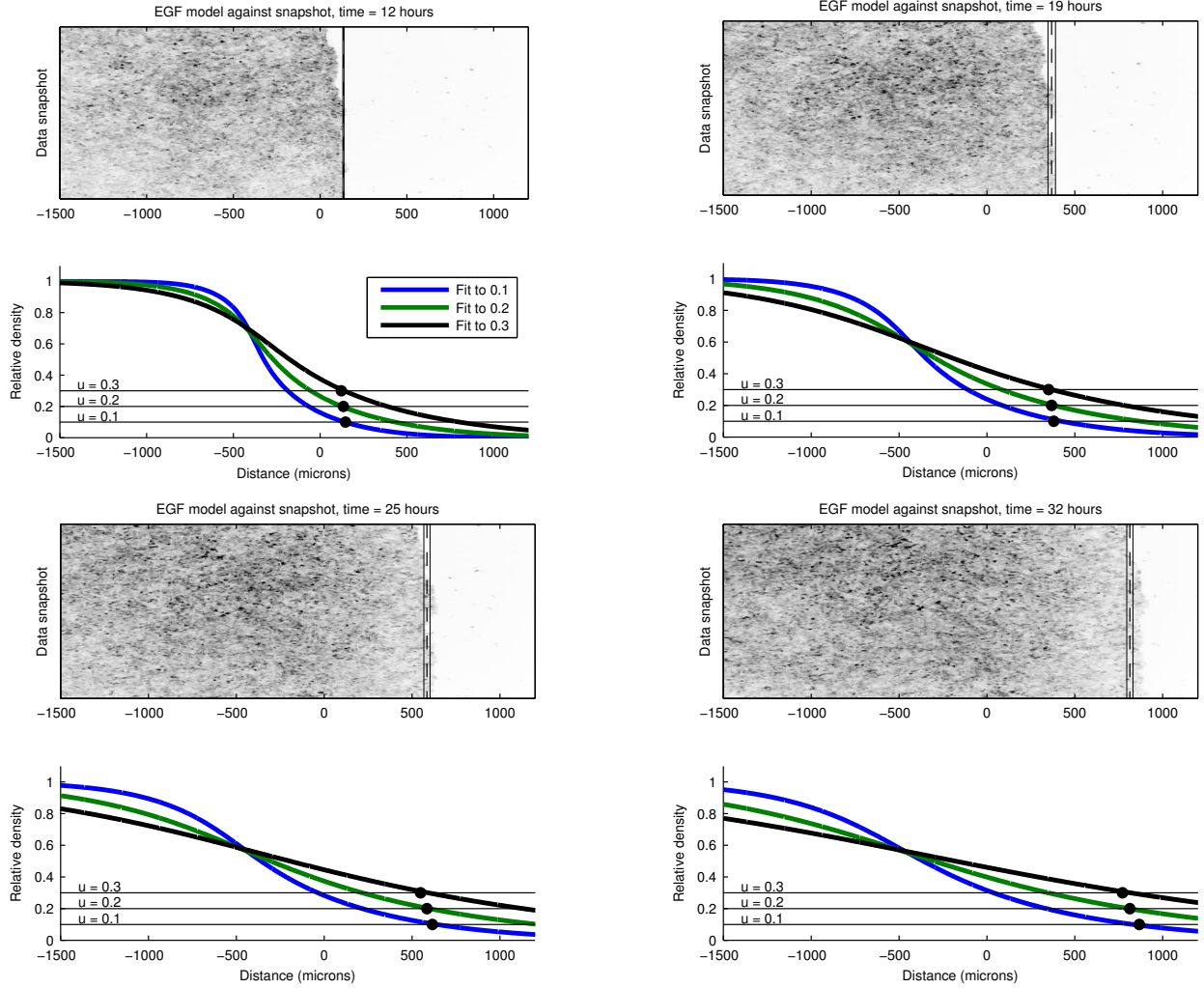
Figure 2.7: Model H simulations against experimental snapshots. The bottom frame depicts Profile plots of Model H that have been fit to $\ell_{data}^{\beta}(t)$ for data set 3 when $\beta = 0.1$ (blue), 0.2 (green), 0.3 (black). The calculated value for $\ell_{data}^{0.2}(t)$ is depicted against the experimental snap shot with a dashed black line and the solid thin lines denote one standard deviation of the data. Against the model profile, we've depicted the lines $u = 0.1, 0.2, 0.3$ for easy identification of the model leading edge location. The black dots denote the experimental leading edge locations.
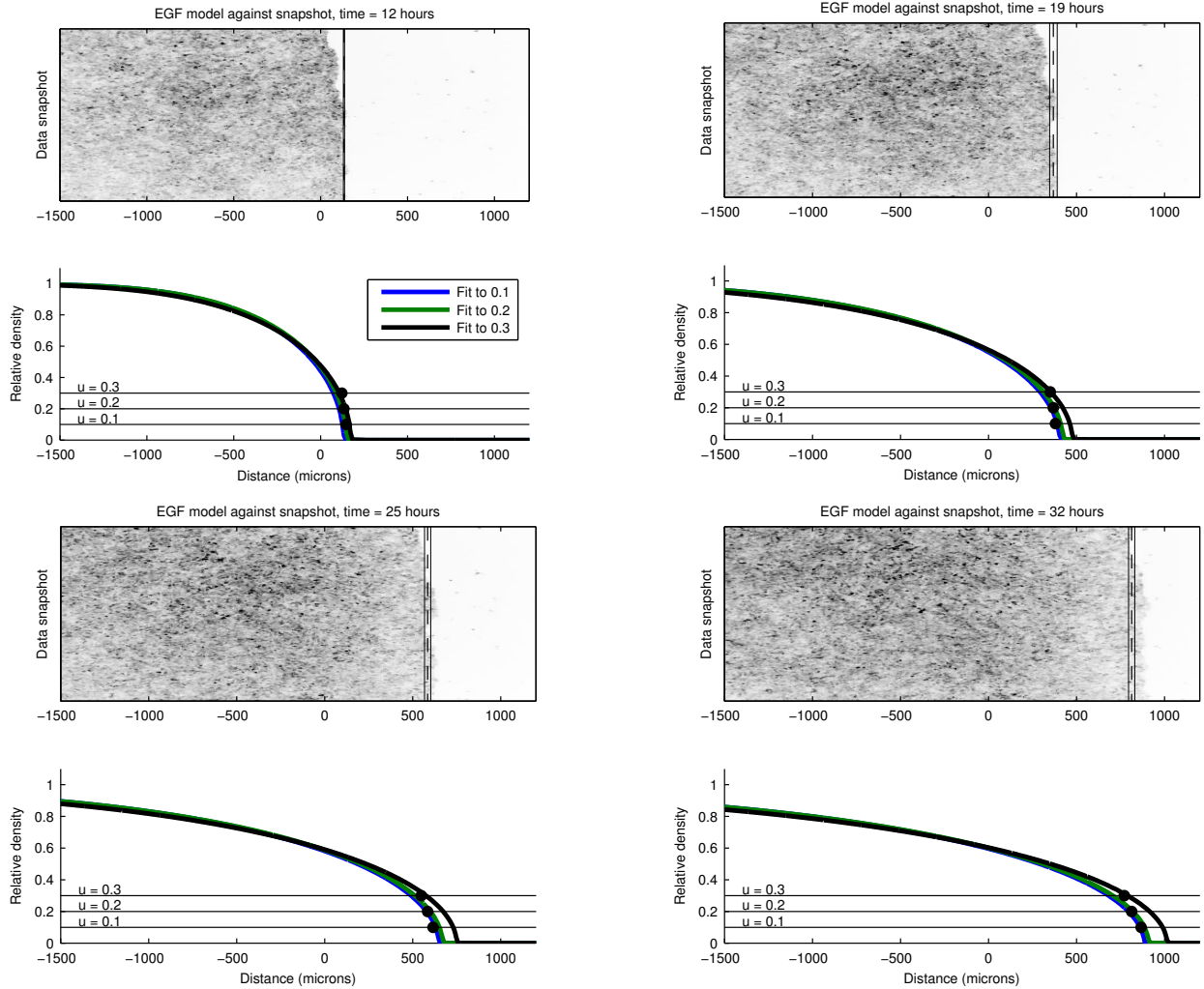
Figure 2.8: Model P simulations against experimental snapshots. The bottom frames depicts Profile plots of Model P that have been fit to $\ell_{data}^{\beta}(t)$ for data set 3 when $\beta = 0.1$ (blue), 0.2 (green), 0.3 (black). The calculated value for $\ell_{data}^{0.2}(t)$ is depicted against the experimental snap shot with a dashed black line and the solid thin lines denote one standard deviation of the data. Against the model profile, we've depicted the lines $u = 0.1, 0.2, 0.3$ for easy identification of the model leading edge location. The black dots denote the experimental leading edge locations.
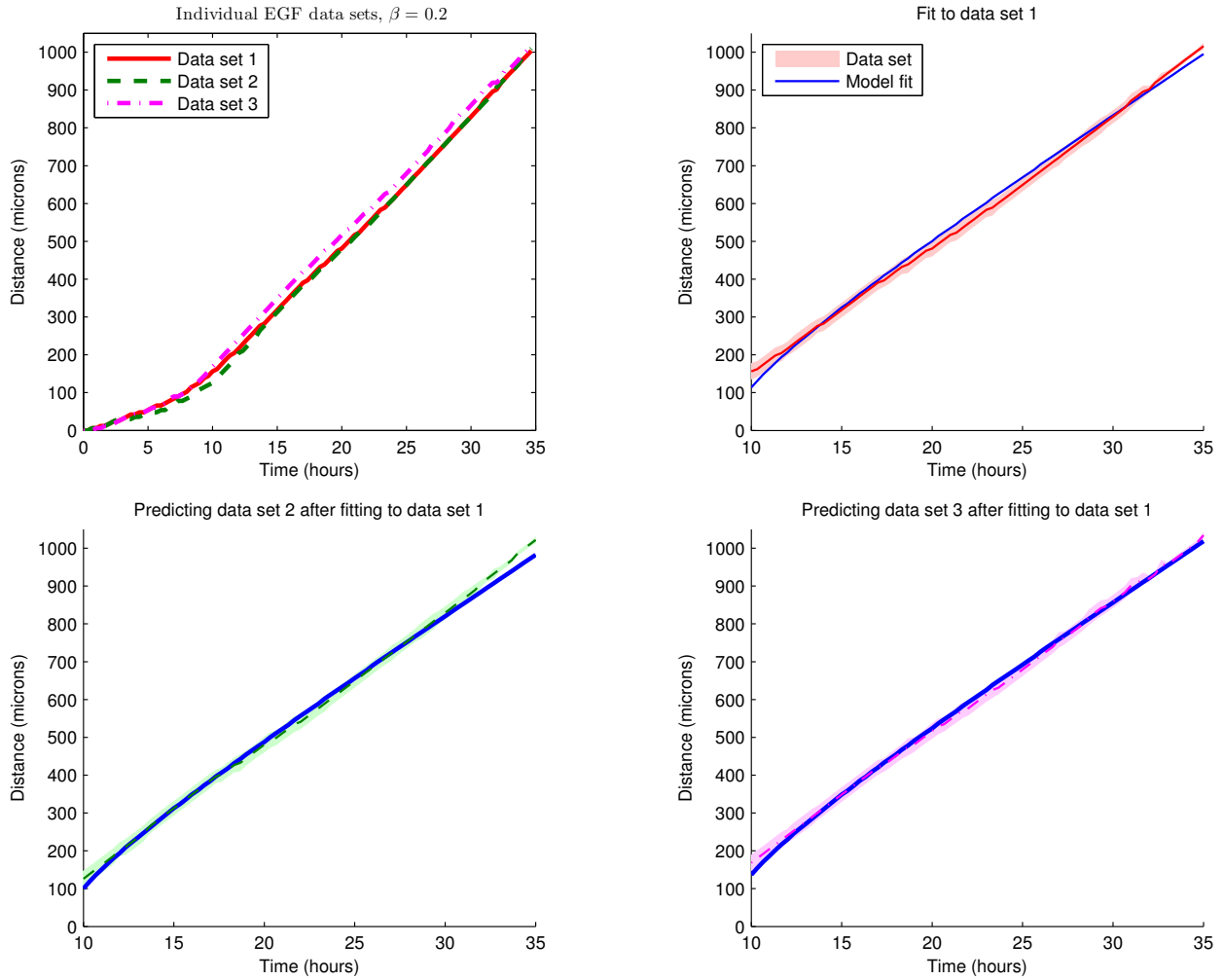
Figure 2.9: Using Model P fits to individual data sets to predict independent data sets. In the top left frame, we depicted experimental data for three replicates of wild type cell sheets in response to EGF treatment. In the top right frame, we fit Model P to data set 1. In two bottom frames, we use this Model P fit to predict the other two data sets by aligning the model simulation and data at $t = 15$ hours. Leading edge computation here was done for $\beta = 0.2$. Examples for $\beta = 0.1$ and 0.3 are given in the appendix.

fitting to and then predicting $\ell_{data}^{0.1}(t)$ and $\ell_{data}^{0.3}(t)$. Examples of fits and predictions to $\ell_{data}^{0.1}(t)$ and $\ell_{data}^{0.3}(t)$ are given in the appendix. We thus see that fit Model P simulations can reliably predict independent data sets given their leading edge location at $t = 15$ hours.

2.2.3 Decreasing cell-cell adhesion levels with RNA interference

We now investigate how wound healing is affected in a cell sheet with reduced cell-cell adhesion expression by performing RNA interference for α-catenin, a critical cell-cell adhesion protein [Friedl and Gilmour, 2009, Haley and Gullick, 2008]. We denote these experiments as *sh α-catenin cell sheets*[3] and perform the same experimental protocol described in Section 2.1.2 in triplicate, in which a scratch wound is created with a pipette tip, and the sheet is treated with EGF to stimulate migration. We depict one sh α-catenin sheet against a wild type sheet in Video 1 of the supplementary material and observe that the leading edge of both experiments move at a similar rate. As another measure of wound healing, however, we consider an "area of cell invasion," and define a wound healing ratio by:

$$\text{wound healing ratio} = \frac{\#\text{ cells in area of cell invasion}}{\text{confluent cell count}}. \tag{2.8}$$

The area of cell invasion is defined by the rectangle whose height spans from the initial leading edge location to 1200 μm into the wound and whose width is defined by the field of view. We count the number of cells that enter this area of invasion over time using imageJ's FindMaxima function. To calculate the confluent cell count, we again use the FindMaxima function in a rectangle of the same size in the back of the sheet and use the largest cell count found in this region over the course of the experiment. This value does not change significantly between experiments. For an analogous wound healing ratio for model simulations, we calculate $\int_0^{1200} u(t, x)dx$ over time using Simpson's rule, where $x \in [0, 1200]$ denotes a location in the area of invasion, and normalize this calculation by its final time point. This calculation measures the total cell mass in the area of invasion over time.

The experimental wound healing ratios over time are depicted for wild type and sh α-catenin cell sheets in Figure 2.10 with a 95% confidence interval of the wound closure ratio depicted in the blue

---

[3]The sh denotes treatment with short hairpin (sh) RNA used for RNA interference.
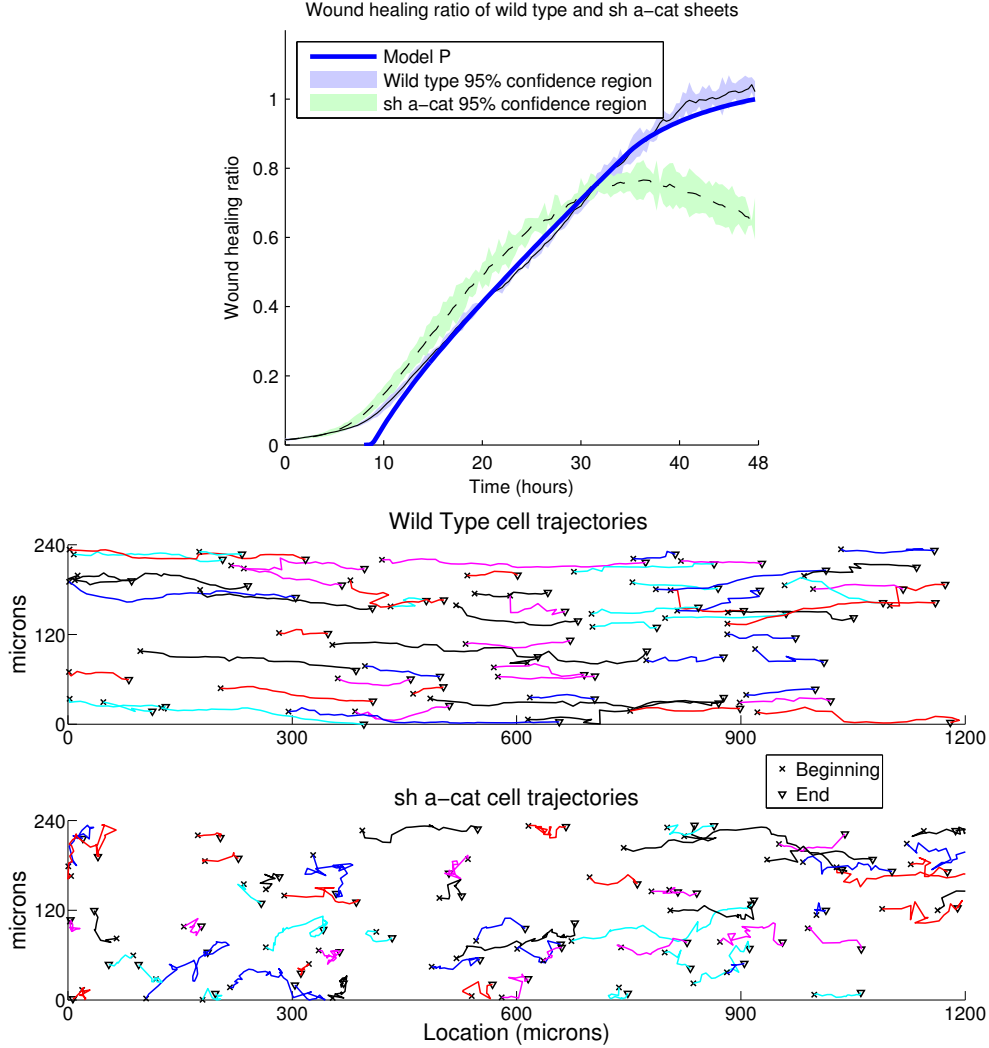
Figure 2.10: Comparing wild type and sh α-catenin cell sheet wound healing. Top : We compare experimental wound closure ratios in the area of invasion (shaded regions) for wild type and sh α-catenin cell sheets using Equation 2.8. The mean experimental ratio values are given in the black and dashed lines for wild type and sh α-catenin sheets, respectively. The blue solid line denotes the wound healing ratio for the best-fit Model P simulation to $\ell_{data}^{0.2}(t)$. Bottom: We use the Pathfinder software to display the individual cell trajectories of randomly picked cells in from experimental replicates after $t = 35$ hours. For each trajectory, an x indicates the cell's initial location, and a triangle indicates the cell's final location.

and green shaded regions, respectively. The experimental sh α-catenin cell sheets initially migrate vigorously into the wound but are unable to maintain this response. The constantly-increasing wild type cell sheet wound healing ratio passes the sh α-catenin sheet wound healing ratio around $t = 30$ hours. We also depict the model wound healing ratio in the solid blue line in Figure 2.10 from the best-fit model P simulation to $\ell_{data}^{0.2}(t)$, which recapitulates the constantly-increasing experimental wound healing ratio well. Note that both wild type ratios have an elbow around $t = 35$ hours, which we expect to see because cells at the leading edge begin moving past the area of invasion.

In Video 1 of the supplementary material, The sh α-catenin sheets appear to stop migrating into the wound around $t = 35$ hours, as demonstrated in the bottom frame of Figure 2.10, where we used the Pathfinder software to track individual cell trajectories after $t = 35$ hours in a thin slice of the area of invasion [Chapnick et al., 2013].The sh α-catenin sheets show only modest migration towards the wound, especially for cells located farther back in the cell sheet (between $x = 0$ and 300 microns). This lack of forward migration causes the wound healing ratio to decrease after $t = 35$ hours, as cells near the leading edge leave the area of invasion, but few cells in the back of the sheet migrate into the area of invasion. Compare this to the wild type sheets, which continue to migrate towards the wound for the entire 48 hours, even as cells move past the area of invasion. We thus observe that cell sheets with decreased α-catenin expression initially migrate into the wound efficiently but are unable to sustain a proper wound healing response.

2.3   Conclusions

In this study, we have developed two mathematical models for the migration of keratinocyte cell sheets during wound healing in response to EGF treatment. These two models are used to investigate the role of cell-cell adhesion on *in vitro* keratinocyte wound healing assays. Model H assumes that cell-cell adhesion *hinders* migration during wound healing due to a drag force, whereas Model P assumes that cell-cell adhesion *promotes* migration with a pulling force. Neither model is able to match experiment EGF-treated data with a constant rate of cell-cell adhesion, but adjusting the rate of cell-cell adhesion to be linear with time allows both models to fit the data well. In its best-fit simulations, Model H exhibits rapid initial expansion into the wound, whereas Model P has a sharp front that propagates into the wound. Model P maintains consistent $\hat{q} = [\hat{D}, \hat{\gamma}_1, \hat{\gamma}_2]^T$

estimates and profile shapes for different leading edge definitions, while Model H varies drastically in both of these aspects. We accordingly propose that Model P appropriately describes keratinocyte migration during wound healing assays, suggesting that keratinocytes use cell-cell adhesion to pull other cells forward during wound healing. The sharp profile front of Model P suggest that cell junctions may also hold cells together to limit drastic changes in density throughout the cell sheet.

We perform the same experimental protocol with HaCaT cell sheets with decreased α-catenin expression. These sheets display initially high migration into the wound, but this efficient migration response can not be maintained without intact adherens junctions. The experimental wild type cell sheets exhibit constant collective migration into the wound over the entire course of the experiment. We thus conclude that cell-cell adhesion acts to hold cells in close proximity to allow migrating cells to pull their neighboring cells into the wound during *in vitro* keratinocyte wound healing assays.

2.4   Future Work

The positive effect of EGF on wound repair is well documented [Clark and Henson, 1995, Dale et al., 1994, Martin, 1997], however, topical EGF treatment has had only moderate effects in clinical wound repair trials [Berlanga et al., 1998, Tsang et al., 2003]. Effective clinical treatment of chronic wounds (such as foot ulcers in diabetic patients) with EGF will require a thorough understanding of its role during wound healing [Tsang et al., 2003]. The knowledge that cell-cell adhesions promote migration through cell pulling may guide future investigations in determining methods to increase the duration and speed of cell migration during wound repair. Such endeavors will aid in developing treatments to shorten the time needed for re-epithelialization to occur or cause chronic wounds to heal.

The simple nature of our first investigation into the effects of EGF treatment on wound healing in keratinocytes leaves ample opportunities for future studies. We note that all experiments considered in this study have the same high cell density (4000 cells/mm^2) and EGF concentration treatment (10 nM). Recent studies have used various forms of Fisher's equation (initially used to model the propagation of an advantageous gene [Fisher, 1937]) to analyze the effects of cell density and EGF concentration on invasion of the prostate cancer PC-3 cell line [Jin et al., 2016, Johnston et al., 2015]. In [Jin et al., 2016], the authors found that a linear cell diffusion term in response to cell density was suitable for less sensitive parameter estimation than a constant diffusion term. Our assumptions on cell-cell adhesion for Model P led to a quadratic cell diffusion term that increases with cell density

and yields consistent parameter estimates based on different leading edge definitions.

While this study provides insight into re-epithalialization, future work should also investigate other phases of wound healing such as inflammation and tissue remodeling. Inflammation is the body's innate response to injury while tissue remodeling occurs mostly after re-epithelialization to restore dermal integrity [Clark and Henson, 1995]. Tissue remodeling involves the interaction of several different cell types, such as macrophages and fibroblasts, and various cytokines and chemokines [Clark and Henson, 1995]. Such studies would necessitate further work into the complicated interactions of fibroblasts with the extracellular matrix (ECM), as fibroblasts synthesize and deposit ECM during migration, but the ECM also affects fibroblast migration and ECM production [Werner et al., 2007]. There is a vast literature on continuum models of haptotaxis, chemotaxis, and mechanotaxis to inform and guide such studies [Olsen et al., 1997, Othmer and Hillen, 2002, Thackham et al., 2008a].

From the accuracy of Model P, and the time-dependent rate of cell-cell adhesion needed to fit experimental data, we propose that cell-cell adhesion strength increases over time during keratinocyte wound healing. Future studies may also investigate the biochemical mechanisms that influence this time-dependent cell-cell adhesion strength. For example, the mitogen-activated protein kinase (MAPK) signaling cascade is downstream of EGFR and known to stimulate cell migration [Chapnick and Liu, 2014, Huang et al., 2004]. In [Posta and Chou, 2010], the propagation of EGF ligand, reactive oxygen species, and EGF's intracellular protease (which is directly linked to activation of the MAPK signaling cascade) was modeled in madine darby canine kidney cells, though cell migration was neglected. Some recent computational studies have modeled molecular signaling processes during inflammation with kinetic and agent based models and found macrophage flux to be a key regulator of the inflammatory response [Nagaraja et al., 2014, 2015, Ziraldo et al., 2015]. The cytokines tumor necrosis factor-α and transforming growth factor-β (both of whose receptors are upstream of MAPK) were determined to be reliable predictors of the the development of chronic inflammation. Incorporating our cell migration model with this signal propagation model and observations on inflammation may help elucidate the interaction between the MAPK signaling cascade and cell migration.

Chapter 3

Investigation of a Structured Fisher's Equation with Applications in Biochemistry[1]

Traveling wave solutions to partial differential equations (PDEs) are often used to study the collective migration of a population of cells during wound healing [Cai et al., 2007, Denman et al., 2006, Landman et al., 2005, 2007], tumorigenesis [Kuang et al., 2015], and angiogenesis [Pettet et al., 1996, Sherratt and Chaplain, 2001]. R.A. Fisher introduced what is now referred to as Fisher's Equation in 1937 to model the advance of an advantageous gene in a population [Fisher, 1937]. Since then, it has been used extensively in math biology literature to model the migration of a two-dimensional monolayer of cells during experimental wound healing assays [Cai et al., 2007, Jin et al., 2016, Maini et al., 2004].

Fisher's Equation is written as

$$u_t = Du_{xx} + \lambda u \left(1 - \frac{u}{K}\right) \tag{3.1}$$

with subscripts denoting differentiation with respect to that variable and $u = u(t, x)$ representing a population of cells over time t at spatial location x. The first term on the right hand side of (3.1) represents diffusion in space with rate of diffusion, D, and the second term represents logistic growth of the population with proliferation rate, λ, and carrying capacity, K. As shown in [Murray,

---

[1]This chapter has been accepted for publication in the SIAM Journal on Applied Mathematics.

2002, § 11.2], (3.1) admits traveling wave solutions of the form

$$u(t, x) = U(z), \ z = x - ct$$

where c denotes the speed of the traveling wave solution and $U(z)$ denotes the traveling wave profile. Traveling wave solutions to (3.1) thus maintain a constant profile, $U(z)$, over time that moves leftward if $c < 0$ or rightward if $c > 0$ with speed $|c|$. It is also shown that (3.1) has a positive and monotonic profile for $|c| \geq 2\sqrt{D\lambda}$, which is biologically relevant when $u(t, x)$ denotes a population of cells. Kolmogoroff proved in 1937 that any solution to (3.1) with a compactly-supported initial condition will converge to a traveling wave solution with wavespeed $c = 2\sqrt{D\lambda}$ [Kolmogoroff et al., 1937]. See [Murray, 1977, § 5.4] for a proof of this. There is also a wide literature on studies into extensions of Fisher's Equation, such as Fisher's Equation coupled with chemotaxis Ai et al. [2015], Landman et al. [2005], time-dependent rates of proliferation and diffusion [Hammond and Bortz, 2011], and space-dependent rates of diffusion [Curtis and Bortz, 2012].

Continuously-structured population models, or PDE models with independent variables to distinguish individuals by some continuously-varying properties, were first investigated via age-structured models in the early 20th century [McKendrick, 1927, Sharpe and Lotka, 1911]. The 1970s saw a revival in structured population modeling after the introduction of methods to investigate nonlinear structured population models [Gurtin and Maccamy, 1974], which led to our current understanding of semigroup theory for linear and nonlinear operators on Banach spaces [Webb, 2008]. Recent math biology studies have used structured equations in various contexts. Several studies have proved the existence of traveling wave solutions to structured population models [Ducrot, 2011, Ducrot and Magal, 2009, Gourley et al., 2007, So et al., 2001]. The authors of [Domschke et al., 2017] developed a spatio-temporal-structured framework to describe the molecular binding process for cell membranes and connected it with cell population dynamics to investigate several examples of cancer invasion. Another study used an independent variable representing subcellular β-catenin concentration to investigate how signaling mutations can cause intestinal crypts to invade healthy neighboring crypts [Murray et al., 2010]. An age-structured equation has also been used to describe demographic data for *Daphnia magna*, a species of water flea [Rutter et al., 2017a].

Recent biological research has focused on the influence of biochemical signaling pathways on the

migration of a population of cells during wound healing. Particular emphasis has been placed on the mitogen-activated protein kinase (MAPK) signaling cascade, which elicits interesting patterns of activation and migration in response to different types of cytokines and growth factors in various cell lines [Chapnick and Liu, 2014, Matsubayashi et al., 2004]. For example, experimental wounding assays of madine darby canine kidney cells (MDCKs) in [Matsubayashi et al., 2004] yielded a transient pulse of ERK 1/2 (a specific MAPK protein) activity in the cell sheet that only lasted for a few minutes. This pulse of activity was followed by a slow wave of activity that propagated from the wound margin to submarginal cells over the course of several hours. The second wave was determined to be crucial for regulating MDCK sheet migration. The authors of [Matsubayashi et al., 2004] proposed that these fast and slow waves of ERK 1/2 activity could be caused by the production of reactive oxygen species (ROS) and epidermal growth factor (EGF), respectively. Similar experiments with fibroblasts also demonstrated this first transient wave of ERK 1/2 activity, but not the following slow wave. The authors of [Chapnick and Liu, 2014] found that human keratinocyte (HaCaT) cells exhibit ERK 1/2 activity primarily at the wound margin during similar experimental wound healing assays with a high density in response to treatment with transforming growth factor-β (TGF-β).

In this study, we detail an approach to investigate a continuously-structured extension of Fisher's Equation that is motivated by the above experimental observations. Previous structured population models (with the exception of [Domschke et al., 2017]) have been restricted to traits that primarily increase over time, such as age or size, but our analysis allows for both activation and deactivation along the biochemical activity dimension.

In Section 3.1, we develop our structured population model and devote Section 3.2 to a review of relevant material from size-structured population models. We demonstrate the existence of self-similar traveling wave solutions to the model in Section 3.3. In Section 3.4, we study a more realistic version of our model where migration and proliferation of the population depend on MAPK activity levels. We derive two criteria for the cell population to activate and use numerical simulations to demonstrate that the cell population is most sensitive to parameter changes when only part of the cell population is activated. Some simulations also migrate furthest at the threshold between part of the population activating and all of the population activating. We make final conclusions and

discuss future work in Section 3.5.

3.1 Model Development

We model a cell population during migrating into a wound, $u(t, x, m)$, by

$$u : [0, \infty) \times \mathbb{R}^n \times [m_0, m_1] \to \mathbb{R}$$

where t denotes time, x denotes spatial location, and m denotes activation along a biochemical signaling pathway with minimum and maximum levels m_0 and m_1, respectively. As a first pass, we assume that any cells of the same MAPK activity level will activate identically over time in the same environment. This assumption allows us to model the activation distribution of the population over time deterministically by considering how cells of all possible MAPK activity levels activate and deactivate over time. We note that biochemical signaling is an inherently heterogeneous process, so our approach would benefit from a further investigation with stochastic differential equations. While cell migration is realistically a three-dimensional process, two-dimensional scratch assays are currently the most common experiments used to study cell migration during wound healing or cancer invasion [Chapnick and Liu, 2014, Haridas et al., 2017, Johnston et al., 2016]. For ease of analysis and numerical computation, we will use $n = 1$ here for one-dimensional migration. Note that one-dimensional equations have been used frequently to interpret two-dimensional scratch assays [Cai et al., 2007, Jin et al., 2016, Mi et al., 2007, Nardini et al., 2016], and the results here extend to two or three dimensions [Hastings et al., 2005].

As discussed in [de Roos, 1996], crucial aspects of a structured population model include the individual state, the environmental state, external forcing factors, and feedback functions. The *individual state* is a dimension used to distinguish between individuals of a population and is typically based on physiological properties such as age or size. As activation of biochemical signaling pathways influences cell migration through diffusive and proliferative properties of cells, we incorporate the biochemical activity dimension, m, as an individual state for our model.

The *environmental state* of a population is the external factors that influence individual behavior. Recall that external cytokines and growth factors, such as ROS, TGF-β, and EGF, influence activation of the MAPK signaling cascade and promote migration during wound healing. The cell

population will not directly affect the level of external growth factor in this work, so an *external forcing factor* will be used to represent treatment with these chemicals here. The external chemical concentration at time t will be denoted by $s(t)$, and the activation response of cells to this chemical will be given by the function $f(s)$.

A *feedback function* included in our work will be the inhibition of individual cell proliferation in response to a confluent density. As proliferation is hindered by contact inhibition, we introduce a new variable,

$$w(t, x) := \int_{m_0}^{m_1} u(t, x, m) dm \tag{3.2}$$

to represent the population of cells at time t and spatial location x. Proliferation of the population will accordingly vanish as $w(t, x)$ approaches the carrying capacity, K.

Our model, which we term as a *structured Fisher's Equation*, is given by the PDE model

$$
\begin{aligned}
u_t + \underbrace{(f(s(t))g(m)u)_m}_{activation} &= \underbrace{D(m)u_{xx}}_{diffusion} + \underbrace{\lambda(m)u\left(1 - \frac{w(t,x)}{K}\right)}_{population\ growth} \tag{3.3} \\
w(t, x) &= \int_{m_0}^{m_1} u(t, x, m) dm \\
u(t = 0, x, m) &= \phi_1(m)\phi_2(x) \\
u(t, x, m = m_1) &= 0 \\
w(t, -\infty) = K \qquad & w(t, x = +\infty) = 0
\end{aligned}
$$

The function $g(m) \in C^1([m_0, m_1])$ denotes the rate of biochemical activation in the population, $s(t) \in L^\infty(\mathbb{R}^+)$ denotes the external chemical concentration in the population, $f(s) \in L^1_{loc}(0, \infty)$ denotes the activation response of cells to the level of signaling factor present, $D(m)$ and $\lambda(m)$ denote biochemically-dependent rates of cell diffusion and proliferation, and $\phi_1(m)\phi_2(x)$ denotes the initial condition of u. We assume a separable initial condition for simplicity. The spatial boundary conditions specify that the cell density has a confluent density at $x = -\infty$ and an empty wound space at $x = +\infty$. We use a no flux boundary condition at $m = m_1$ so that cells cannot pass this boundary. In the remainder of this study, we will write $f(s(t))$ as $f(t)$ for simplicity, though we note that this function will differ between cell lines that respond differently to the same chemical

during wound healing[2].

The solution space of (3.3), $\mathcal{D}$, is defined with inspiration from [Webb, 2008] and Volpert et al. [1994, § 1.1]. If we let Z denote the space of bounded and twice continuously differentiable functions on $\mathbb{R}$, then we define

$$\mathcal{D} := \left\{ u(t,x,m) \,\Big|\, \int_{m_0}^{m_1} u(t,x,m)dm \in Z \right\},$$

i.e., $u(t,x,m) \in \mathcal{D}$ if $w(t,x) \in Z$ for all $t > 0$. We note that $\int_{m_0}^{m_1} \phi(x,m)dm$ need only be bounded and piecewise continuous with a finite number of discontinuities [Volpert et al., 1994]. If $\phi(x,m)$ is not sufficiently smooth in m, we obtain generalized solutions of (3.3) [Webb, 2008].

In Section 3.3, we will investigate (3.3) with constant rates of diffusion and proliferation (i.e., $D(m) = D, \lambda(m) = \lambda$) and $f(t) = 1$. By substituting

$$
\begin{aligned}
u^* &= \frac{m_1 - m_0}{K} u, & t^* &= \lambda t, f^*(t^*) = f(t^*/\lambda) \\
x^* &= x\sqrt{\lambda/D}, & m^* &= (m - m_0)/(m_1 - m_0), \\
g^*(m^*) &= \frac{g(m^*(m_1 - m_0) + m_0)}{\lambda \cdot (m_1 - m_0)}, \\
w^* &= \int_0^1 u^*(t^*, x^*, m^*)dm^* &&\text{(3.4)}
\end{aligned}
$$

and dropping asterisks for simplicity, (3.3) can be non-dimensionalized to

$$
u_t + \underbrace{(f(t)g(m)u)_m}_{activation} = \underbrace{u_{xx}}_{diffusion} + \underbrace{u\left(1 - \int_0^1 u(t,x,m)dm\right)}_{population\ growth} \tag{3.5}
$$

$$
\begin{aligned}
w(t,x) &= \int_0^1 u(t,x,m)dm \\
u(t=0,x,m) &= \phi_1(m)\phi_2(x) \\
u(t,x,m=1) &= 0 \\
w(t,x=-\infty) = 1 & \qquad w(t,x=+\infty) = 0.
\end{aligned}
$$

In Section 3.4, we will consider the full model (3.3) when the rates of cellular diffusion and proliferation are piece-wise constant functions of m and numerically investigate how different functions for $f(t)$ lead to increased and decreased levels of population migration. These rates of diffusion and

---

[2]Note that an extension for modeling the dynamics governing $s(t)$ will be considered in a future study.

proliferation are given by

$$D(m) := \begin{cases} D_1 & m \in M_{inact} \\ D_2 & m \in M_{act} \end{cases}, \; \lambda(m) := \begin{cases} \lambda_1 & m \in M_{inact} \\ \lambda_2 & m \in M_{act} \end{cases} \tag{3.6}$$

for $M_{inact} := [m_0, M_{crit})$ and $M_{act} := [m_{crit}, m_1]$. We can perform the same normalization as (3.4) but now use

$$x^* = x\sqrt{\lambda_1/D_1}, \;\; t^* = \lambda_1 t$$
$$f^*(t^*) = f(t^*/\lambda_1)$$

in which case (3.3) becomes

$$\begin{aligned} u_t + (f(t)g(m)u)_m &= \tilde{D}(m)u_{xx} + \tilde{\lambda}(m)u\left(1 - \frac{w}{K}\right) \\ w &= \int_{m_0}^{m_1} u(t,x,m)dm \\ u(t=0,x,m) &= \phi_1(m)\phi_2(x) \\ u(t,x,m=m_1) &= 0 \\ w(t,x=+\infty) = 0 \qquad & w(t,x=-\infty) = 1. \end{aligned} \tag{3.7}$$

with

$$\tilde{D}(m) := \begin{cases} 1 & m \in M_{inact} \\ \bar{D} & m \in M_{act} \end{cases}, \; \tilde{\lambda}(m) := \begin{cases} 1 & m \in M_{inact} \\ \bar{\lambda} & m \in M_{act}. \end{cases} \tag{3.8}$$

3.2 Background Material from Size-Structured Population Modeling

Before investigating the existence of traveling-wave solutions to (3.5), it is useful to review some key topics used to solve size-structured population models, as discussed in [Webb, 2008]. These topics will be useful in analyzing (3.3) in later sections. A reader who is familiar with using the method of characteristics to solve size-structured population models may briefly skim over this section to pick up on the notation used throughout our study.

As an example, we consider the size-structured model given by

$$u_t + (g(y)u)_y = Au \tag{3.9}$$

$$u(t = 0, y) = \phi(y)$$

where $u = u(t, y) : [0, \infty) \times [y_0, y_1] \to \mathbb{R}$ denotes the size distribution over y of a population at time t, y_0 and y_1 denote the minimum and maximum population sizes respectively, and $g(y) \in C^1(([y, y_1])$ denotes the physical growth rate[3] of individuals of size y. In this section, we will work in the Banach space $\mathbb{X} = L^1(([y, y_1] \to \mathbb{R})$ and assume $A \in \mathcal{B}(\mathbb{X})$, the space of bounded and linear operators on $\mathbb{X}$. The method of characteristics will facilitate solving (3.9).

The goal of this section is to discuss techniques that can be used to solve (3.9) analytically, and provide an example problem to provide insight. In Section 3.2.1, we discuss the functions $\sigma(y; \underline{y})$ and $\sigma^{-1}(t; \underline{y})$, which will prove useful in solving (3.9). In Section 3.2.2, we discuss how to solve (3.9) along the characteristic curve $y = \sigma^{-1}(t; \underline{y})$. In Section 3.2.3, we discuss how to obtain the final solution form.

3.2.1 The characteristic curves

For a fixed size $\underline{y} \in [y_0, y_1]$, the function

$$\sigma(y; \underline{y}) := \int_{\underline{y}}^{y} \frac{1}{g(y')} dy' \tag{3.10}$$

provides *the time it takes for an individual to grow from the fixed size $\underline{y}$ to arbitrary size y*. If $g(y)$ is positive and uniformly continuous on $[y_0, y_1]$, then $\sigma(y; \underline{y})$ is invertible [Webb, 2008]. We denote its inverse function, $\sigma^{-1}(t; \underline{y})$, as the *size curve*, which represents *the size of an individual over time that starts at size $\underline{y}$ at time $t = 0$*. For instance, if an individual has size $\underline{y}$ at $t = 0$, then that individual will have size $\sigma^{-1}(t_1; \underline{y})$ at time $t = t_1$. Some helpful properties of the growth curve are that $\sigma^{-1}(0; \underline{y}) = \underline{y}$ and

$$\frac{d}{dt} \sigma^{-1}(t; \underline{y}) = g(\sigma^{-1}(t; \underline{y})). \tag{3.11}$$

---

[3]Note that in this section, $g(y)$ denotes a growth rate with respect to size, y, whereas throughout the rest of our study, $g(m)$ denotes an activation rate with respect to biochemical activity, m.

See Section A.6 in the appendix for the derivation of (3.11).

As an example problem, consider $g(y) = (1 - y)/2, y_0 = 0, y_1 = 1$. Then we can compute

$$\sigma(y; \underline{y}) = 2 \log \left(\frac{1 - \underline{y}}{1 - y} \right)$$

$$\sigma^{-1}(t; \underline{y}) = 1 - (1 - \underline{y})e^{-t/2}.$$

We depict plots of $\sigma(y; \underline{y})$ in Figure 3.1(a) and $\sigma^{-1}(t; \underline{y})$ in Figure 3.1(b). Note that $\sigma(0.4; \underline{y} = 0) \approx 1$, so it takes an individual of size zero about 1 time unit to reach a size of 0.4. Conversely, $\sigma^{-1}(1; \underline{y} = 0) \approx 0.4$, so an individual that begins with a size of zero will be close to size 0.4 after 1 time unit. These computations are confirmed in Figure 3.1.

3.2.2 Solving along characteristic curves

In order to solve (3.9) with the method of characteristics, we set $y = \sigma^{-1}(t; \underline{y})$ to define the variable $v(t; \underline{y})$ as

$$v(t; \underline{y}) := u(t, y = \sigma^{-1}(t; \underline{y})). \tag{3.12}$$

As shown in Section A.7 of the appendix, substitution of (3.12) into (3.9) yields the characteristic equation

$$v_t = -g'(\sigma^{-1}(t; \underline{y}))v + Av, \tag{3.13}$$

where primes denote differentiation with respect to y. This characteristic equation has size $\underline{y}$ at time $t = 0$ and can be solved explicitly as[4]

$$v(t; \underline{y}) = \frac{g(\underline{y})}{g(\sigma^{-1}(t; \underline{y}))} e^{At} \phi(\underline{y}). \tag{3.14}$$

Returning to our example from Section 3.2.1, if we set $A = 0$, then we can compute the solution

---

[4]To derive this, use separation of variables and with the help of (3.11) note that $\int_0^t g'(\sigma^{-1}(\tau; \underline{y}))d\tau = \ln[g(\sigma^{-1}(t; \underline{y}))/g(\underline{y})]$.
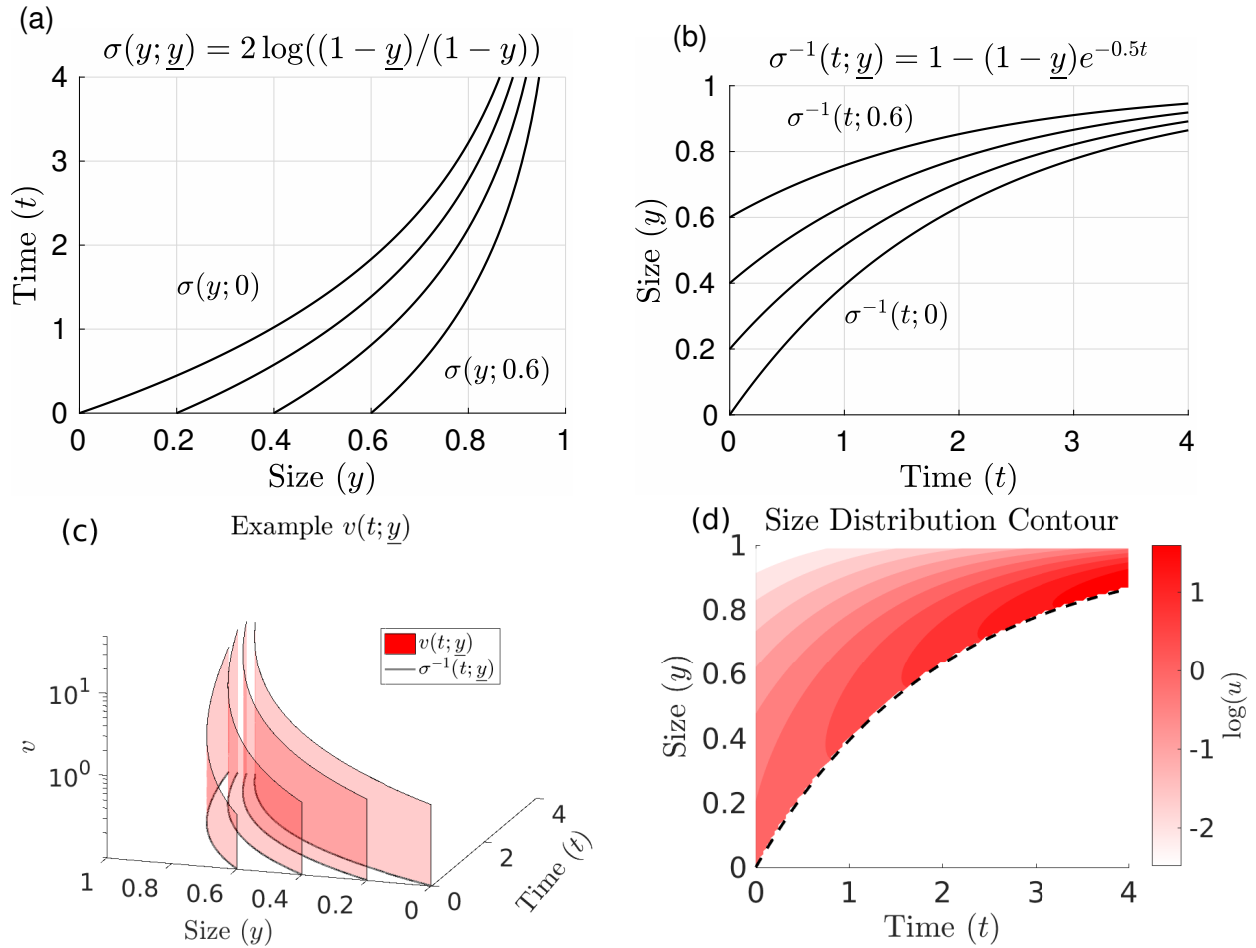
Figure 3.1: Solving (3.9) analytically for $g(y) = (1 - y)/2$, $A = 0$, and $\phi(y) = e^{-(y-0.1)^2/.16}$. (a) Plots of $t = \sigma(y; \underline{y})$ for $\underline{y} = 0, 0.2, 0.4$, and 0.6. (b) Plots of $y = \sigma^{-1}(t; \underline{y})$ for $\underline{y} = 0, 0.2, 0.4$, and 0.6. (c) Plots of $v(t; \underline{y})$ for $\underline{y} = 0, 0.2, 0.4, 0.6$. The curves in the (t, y)-plane denote $y = \sigma^{-1}(t; \underline{y})$. (d) Log-scale plot of $u(t, y)$. The dashed line denotes the minimum level of support over time.

to (3.9) along $y = \sigma^{-1}(t; \underline{y})$ as

$$v(t; \underline{y}) = \frac{(1 - \underline{y})/2}{(1 - \underline{y})e^{-t/2}/2}\phi(\underline{y})$$
$$= e^{t/2}\phi(\underline{y}).$$

Some three-dimensional plots of $v(t; \underline{y})$ are depicted in Figure 3.1(c). We depict $\sigma^{-1}(t; \underline{y})$ in the (t, y)-plane below these curves to emphasize that we are solving (3.9) along the curves $y = \sigma^{-1}(t; \underline{y})$.

3.2.3   Final Solution Form

As (3.14) provides the solution to (3.9) along the arbitrary characteristic curve with initial size $\underline{y}$, we use it to solve the whole equation with the substitution $\underline{y} = \sigma^{-1}(-t, y)$, in which we find

$$u(t, y) = \begin{cases} \frac{g(\sigma^{-1}(-t,y))}{g(y)}e^{At}\phi(\sigma^{-1}(-t, y)) & \sigma^{-1}(t; y_0) \leq y \leq y_1 \\ 0 & y_0 \leq y < \sigma^{-1}(t; y_0). \end{cases} \tag{3.15}$$

If $\phi(y) \notin C^1(y_0, y_1)$, then (3.15) is viewed as a generalized solution to (3.9). Note that a piecewise form is needed for (3.15) because we do not begin with any individuals below the minimum size, y_0, and thus the minimum possible size at time t is given by $\sigma^{-1}(t; y_0)$. If the population is assumed to give birth to individuals of size y_0 over time, then the appropriate renewal equation representing population birth would replace the zero term in the piecewise function (see [Banks and Tran, 2009, § 9.5] for an example in size-structured populations and [Gourley et al., 2007] for an example in age-structured populations).

Completing our example, we can compute the final solution to (3.9) when $g(y) = (1 - y)/2$ as

$$u(t, y) = \begin{cases} e^{t/2}\phi(1 - (1 - y)e^{t/2}) & 1 - e^{-t/2} \leq y \leq 1 \\ 0 & y < 1 - e^{-t/2} \end{cases}. \tag{3.16}$$

A contour of this solution is depicted in Figure 3.1(d). Note that the dashed line denotes $y = 1 - e^{-t}$, so $u(t, y) = 0$ for all points below this curve.

43

3.3 Existence of Traveling Wave Solutions to the Structured Fisher's Equation and Other Transport Equations

In this section, we aim to prove the existence of unique self-similar traveling wave solutions to (3.5), as summarized in Theorem 1. We do so with techniques described in the previous section.

Note that in this section, $g(m)$ is a function of biochemical activity level, and the characteristic curves, which will be given by

$$h(t; \underline{m}) := \sigma^{-1}(F(t); \underline{m}) \tag{3.17}$$

for $F(t) := \int_0^t f(t')dt'$, will compute *the activity level of an individual over time that starts at level $\underline{m}$ at time $t = 0$*. We will thus now refer to these as the *activation curve*s.

We now state the main theorem:

Theorem 1. *Equation (3.5) admits unique self-similar traveling wave solutions of the form*

$$u(t, x, m) = \begin{cases} \frac{g(\sigma^{-1}(-F(t); m))}{g(m)} W(z)\phi_1(\sigma^{-1}(-F(t); m)), & h(t; 0) \leq m < 1 \\ 0 & otherwise \end{cases}, \tag{3.18}$$

for $z = x - ct$, $c \geq 2$, $\phi(m) := \phi(x = -\infty, \underline{m})$, and $W(z)$ denoting the traveling wave profile to the classical Fisher's Equation.

In Section 3.2.3, we plot a numerical solution to (3.5) to demonstrate the equation's behavior and how it relates to (3.18). We prove Theorem 1 in Section 3.3.2. We then comment on how this theorem can be extended to general reaction-diffusion and chemotaxis equations in Section 3.3.3.
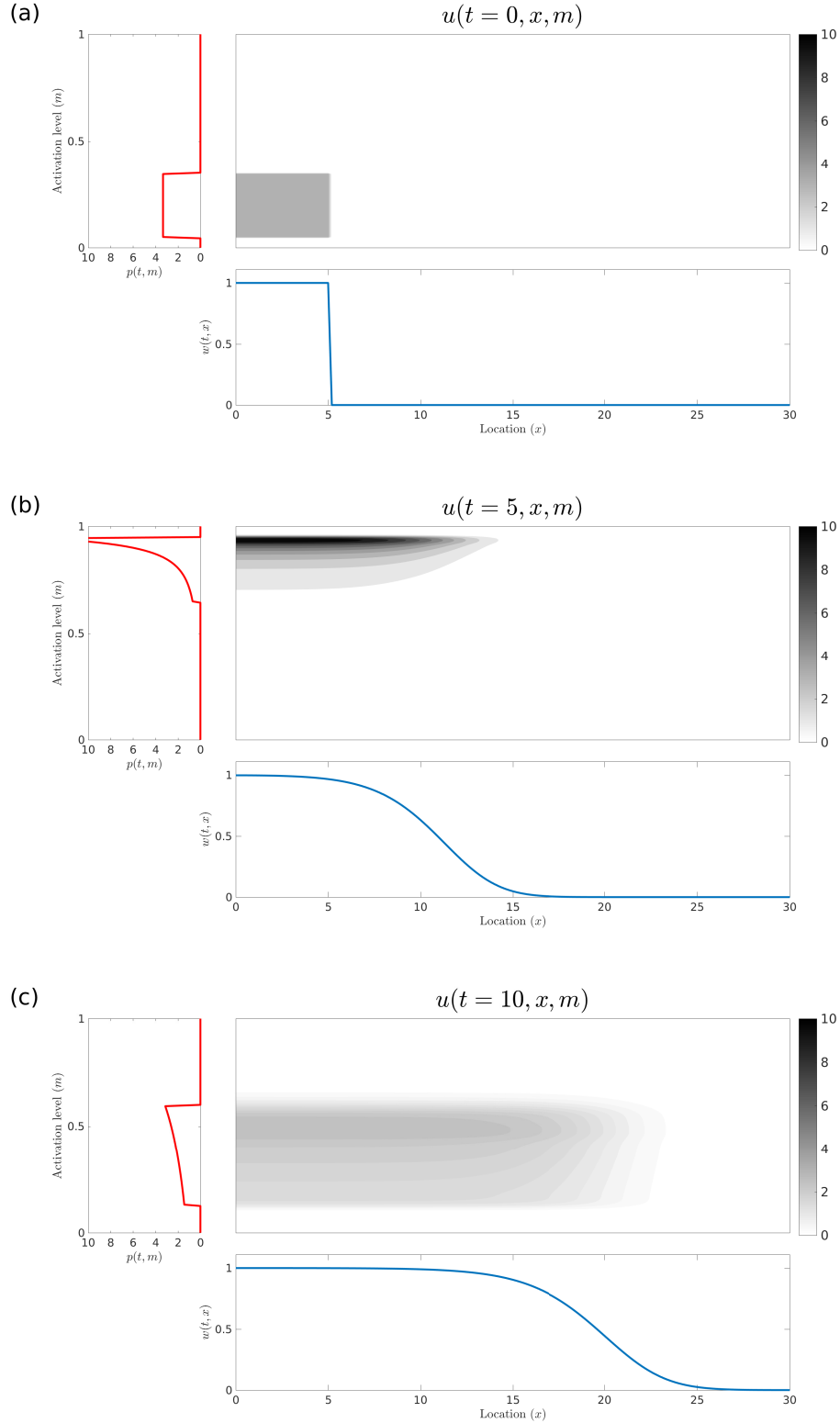
Figure 3.2: Solutions to (3.5) with $f(t) = \beta e^{\gamma t} - 1, g(m) = \alpha m(1 - m)$ at $t = 0, 5,$ and 10 for $\alpha = 1, \beta = 3, \gamma = -1/4$. In each frame, the grayscale contour depicts the numerical simulation of $u(t, x, m)$ along, the red profile is a plot of $p(t, m)$ to denote the activation profile at time t, and the blue profile is the result of numerical integration of $u(t, x, m)$ along the m dimension to depict $w(t, x)$.

3.3.1 Numerical Simulation of (3.5)

In Figure 3.2, we depict a numerical simulation of (3.5) with $f(t) = \beta e^{\gamma t}, g(m) = \alpha m(1-m)$. Note that these terms will be explained in Section 3.4.4. For numerical implementation, we use a standard central difference scheme along the x-dimension, an upwind scheme with flux limiters (similar to those described in [Thackham et al., 2008a]) along the m dimension, and a Crank-Nicholson scheme to integrate along time. We discretized each independent variable of a uniform grid with step sizes of $\Delta x = 0.2, \Delta m = 0.00625, \Delta t = 10^{-3}$.

In each frame of Figure 3.2, the grayscale contour denotes the numerical simulation of $u(t, x, m)$ along x and m for $t = 0, 5,$ or 10. The blue profile under the contour is the result of numerical integration of $u(t, x, m)$ along the m dimension to depict $w(t, x)$. To the left of the contour in red, we plot the activation profile along m, $p(t, m)$, which has the exact form

$$p(t, m) = \begin{cases} \frac{g(\sigma^{-1}(-F(t), m))}{g(m)} \phi_1(\sigma^{-1}(-F(t), m)) & h(t; m_0) \leq m \leq m_1 \\ 0 & m_0 \leq m < h(t; m_0). \end{cases} \tag{3.19}$$

Note that (3.19) is similar to (3.18) but is missing the $W(z)$ term.

In Figure 3.2, we observe that the cell population initially activates because $f(t) > 0$, so the cell population moves up the $m-$dimension, as observed at $t = 5$. $f(t)$ soon becomes negative, in which case the cell population will begin to deactivate and move down the m-dimension, which is observed at $t = 10$. If we continued plotting $u(t, x, m)$ at later time points, we would observe the whole contour approaching $m = 0$ as the population continues to deactivate. Meanwhile, along the spatial domain we see the cell population moving to the right at a constant speed, as we would expect from the $W(z)$ term from (3.18). Note that there is numerical error in this simulation due to numerical diffusion along the m dimension, which is unavoidable but can be decreased with finer values of Δm during computation.

3.3.2 Proof of Theorem 1

Proof. After taking the time derivative of $w(t, x)$, which was defined in (3.2), we can rewrite (3.5) as a system of two coupled PDEs[5]:

$$
\begin{aligned}
u_t + (f(t)g(m)u)_m &= u_{xx} + u(1 - w) \\
w_t &= w_{xx} + w(1 - w).
\end{aligned}
\tag{3.20}
$$

We next set up the characteristic equation for u by setting $m = h(t; \underline{m})$ for a fixed value of $\underline{m}$, which gives

$$
v(t, x; \underline{m}) := u(t, x, m = h(t; \underline{m})).
\tag{3.21}
$$

Substituting (3.21) into (3.20) provides our characteristic equation

$$
\begin{aligned}
v_t &= v_{xx} + v[1 - w - f(t)g'(h(t; \underline{m}))] \\
w_t &= w_{xx} + w(1 - w),
\end{aligned}
\tag{3.22}
$$

a time-dependent system of two coupled PDEs in time and space. Note that the bottom equation for (3.22) is Fisher's Equation, which has positive monotonic traveling wave solutions for any speed $c \geq 2$ (see [Murray, 2002, § 11.2]).

We next aim to derive traveling wave solutions to (3.22). From our knowledge of size-structured population models from Section 3.2, we intuit the ansatz of a self-similar traveling wave solution, which we write as

$$
\begin{aligned}
v(t, x; \underline{m}) &= \frac{g(m)}{h(t; \underline{m})} V(z), \ z = x - ct \\
w(t, x) &= W(z).
\end{aligned}
\tag{3.23}
$$

In this ansatz, $V(z)$ will define a traveling wave profile for v and $\frac{g(m)}{h(t; \underline{m})}$ will provide the height of

---

[5]Note that either $g(m_0) = 0$ or $u(t, m = m_0, x) = 0$ for $t > 0$, so that the activation term drops out when integrating over m for w.

the function over time. With the aid of the chain rule, we observe that:

$$
\begin{aligned}
v_t(t, x; \underline{m}) &= -\frac{f(t)g'(h(t; \underline{m}))g(\underline{m})}{g(h(t; \underline{m}))}V - c\frac{g(\underline{m})}{g(h(t; \underline{m}))}V_z \\
v_{xx}(t, x; \underline{m}) &= \frac{g(\underline{m})}{h(t; \underline{m})}V_{zz},
\end{aligned}
$$

where subscripts denote differentiation with respect to t, x, or z and primes denote differentiation with respect to m. Substituting (3.23) into (3.22) reduces it to the autonomous system

$$
\begin{aligned}
-cV_z &= V_{zz} + V(1 - W) \\
-cW_z &= W_{zz} + W(1 - W).
\end{aligned} \tag{3.24}
$$

The existence of a traveling wave solution can now be proven by showing the existence of bounded solutions to (3.24) with the properties

$$
(V, V_z, W, W_z)(-\infty) = (\phi_1(\underline{m}), 0, 1, 0), \ (V, V_z, W, W_z)(\infty) = (0, 0, 0, 0). \tag{3.25}
$$

We note that the equations for W are independent of V and are the system of differential equations that yield the traveling wave profile for Fisher's equation [Murray, 2002]. We thus have that $W(-\infty) = 1, W_z(-\infty) = W(\infty) = W_z(\infty) = 0$ are already satisfied for $c \geq 2$. For a fixed value of c, call this solution W_c. We then observe by substitution that $\phi_1(\underline{m})W_c$ satisfies the equations for V. We now have that (3.18) is a solution to (3.5).

We now show that $(\phi_1(\underline{m})W_c(z), W_c(z))^T$ is the unique solution (for c fixed) of (3.24) and (3.25) as follows. Let V, W be solutions to (3.24). We can then multiply the first equation of (3.24) by W and the second by V and subtract the two solutions to find

$$
-cV_zW + cVW_z = V_{zz}W - VW_{zz} = (V_zW - VW_z)_z.
$$

The above gives that

$$
V_zW - VW_z = ke^{-cz}
$$

and boundedness of $V, V_z, W,$ and W_z ensures that $k = 0$. Because $W > 0$, we find that

$$\left(\frac{V}{W}\right)_z = 0,$$

which suggests that

$$V = \phi_1(\underline{m})W$$

is the unique solution to the upper portion of (3.24). $\qquad\square$

3.3.3 Note on the generality of results

We have focused on Fisher's Equation in this work due to its applicability to the math biology literature and cell migration in particular. The methods provided in Section 3.3.2 for the existence of a unique self-similar traveling wave solution are also applicable for a wide variety of structured transport equations, including reaction-diffusion and chemotaxis equations. For example, consider the nondimensionalized structured reaction-diffusion equation given by

$$u_t + (f(t)g(m)u)_m = u_{xx} + uR(w) \tag{3.26}$$

$$w_t = w_{xx} + wR(w). \tag{3.27}$$

We have the following corollary:

Corollary 1.1. *Suppose $R(w) \in C^1(\mathbb{R})$ has a single value, $w_l > 0$, such that $R(w_l) = 0$ or two values, $0 < w_l^1 < w_l^2$, such that $R(w_l^1) = R(w_l^2) = 0$ and $d/dw(wr(w)) > 0$ at $w = w_l^1$. Then (3.26) will have the unique self-similar traveling wave solution given by*

$$u(t,x,m) = \begin{cases} \frac{g(\sigma^{-1}(-F(t),m))}{g(m)}W_c(z)\phi(\sigma^{-1}(-F(t),m)), & h(t;0) \leq m < 1 \\ 0 & \textit{otherwise} \end{cases} \tag{3.28}$$

where $W_c(z)$ is a traveling wave solution with speed c for (3.27).

Proof. Note that the given conditions for $R(w)$ are sufficient for the existence of a traveling wave solution for $w(t,x)$, which we will write as $W_c(z)$ [Fife, 1979, § 4.4]. The speed of this solution, c, will depend on the form of $R(w)$ and possibly the initial condition for $w(t,x)$.

We may then assume the characteristic equation $v(t, x; \underline{m}) := u(t, x, m = h(t, \underline{m}))$ takes the form

$$v(t, x; \underline{m}) = \frac{g(m)}{h(t, \underline{m})} V(z)$$

and $w(t, x) = W(z)$. This yields the autonomous system of ordinary differential equations

$$-cV_z = V_{zz} + VR(W)$$

$$-cW_z = W_{zz} + WR(W),$$

from which we observe that $V = \phi_1(\underline{m})W_c$ and $W = W_c$ will provide a solution to this system. The proof of uniqueness of this solution is identical to the proof of uniqueness of the solution for (3.24). This proves the existence of unique self-similar traveling wave solutions to structured reaction diffusion equation given by (3.26).

$\square$

Note that the form of the reaction-diffusion equation given in (3.27) is general and well studied. If R is given by $R(w) = a(w - 1)(\alpha - w)$ for constants $a > 0$ and $\alpha \in (0, 1)$, then (3.27) represents the bistable equation. A traveling wave solution then exists for a unique speed, c, which we will write as $W_c(z)$ [Keener and Sneyd, 2009]. Corollary 1 now guarantees that (3.26) will have the unique self-similar traveling wave solution given by (3.28).

The same method could also be applied to a structured chemotaxis equation given by

$$u_t + (f(t)g(m)u)_m = -(us_x)_x + \alpha u(1 - w) \tag{3.29}$$

$$w_t = -(ws_x)_x + \alpha w(1 - w) \tag{3.30}$$

$$s_t = bs(1 - s) - ws.$$

Here, $u = u(t, x, m)$ now denotes a cell density simultaneously activating along the biochemical pathway, m, and migrating up the gradient of some chemoattractant $s = s(t, x)$, and $w(t, x) = \int u(t, x, m)dm$ denotes the cell density in time and space. Conditions for the existence of smooth- and discontinuous-fronted traveling wave solutions for (3.30), denoted as $W_c(z)$, are provided in [Newgreen et al., 2003]. From the same proof methods detailed above, our results now also guarantee

the existence of a unique self-similar traveling wave solution to (3.29) when $W_c(z)$ has infinite support.

3.4 Structured Fisher's Equation with MAPK-dependent Phenotype

We now study a version of Fisher's Equation where cellular migration and proliferation depend on biochemical activity, m, as described by (3.7). Various cell lines have reduced rates of proliferation and increased migration in response to MAPK activation [Chapnick and Liu, 2014, Clark and Henson, 1995, Matsubayashi et al., 2004], so we let m denote activity along the MAPK signaling cascade in this section. We consider a model with two subpopulations: one with a high rate of diffusion in response to MAPK activation and the other with a high rate of proliferation when MAPK levels are low. MAPK activation will depend on an external forcing factor to represent the presence of an extracellular signaling chemical, such as ROS, TGF-β, or EGF. While the method of characteristics is not applicable to spatial activation patterning here due to the parabolic nature of (3.3) in space, we can investigate temporal patterns of activation and deactivation. *We will exhibit simple scenarios that give rise to three ubiquitous patterns of biochemical activity: 1.) a sustained wave of activation, 2.) a single pulse of activation, and 3.) periodic pulses of activation.*

Before describing these examples, we first introduce some tools to facilitate our study of (3.7). We will detail some activation criteria in Section 3.4.2, and discuss numerical issues and the derivation of a time-dependent averaged Fisher's Equation in Section 3.4.3 before illustrating the different activation patterns and their effects on migration in Section 3.4.4.

3.4.1 Model Description

Recall that the biochemically-dependent structured Fisher's Equation is given by (3.7). Given some $m_{crit} \in (0, 1)$, we define two subsets of $[0, 1]$ as $M_{inact} := [0, m_{crit}), M_{act} := [m_{crit}, 1]$, and the rates of diffusion and proliferation by (3.8) for $\bar{D} > 1$ and $\bar{\lambda} < 1$ Hence for $m \in M_{inact}$, the population is termed as *inactive* and primarily proliferates whereas for $m \in M_{act}$, the population is termed as *active* and primarily diffuses.

We let $supp(\phi_1(m)) = [\underline{m}_{\min}, \underline{m}_{\max}]$ for $\underline{m}_{\max} < m_{crit}$ and assume that $\int_0^1 \phi_1(m)dm = 1$

51

so $\phi_1(m)$ represents a probability density function for the initial distribution of cells in m. We accordingly denote

$$\Phi_1(m) := \begin{cases} 0 & m \leq 0 \\ \int_{m_0}^{m} \phi_1(m')dm' & 0 < m \leq 1 \\ 1 & 1 < m \end{cases}$$

as the cumulative distribution function for $\phi_1(m)$.

3.4.2 Activation Criteria

To facilitate future numerical simulations, we will derive a condition for a cell population with an initial condition contained in M_{inact} to enter M_{act}. Note that to do so, we can ask when the characteristic curve for the maximum level of initial support will reach the threshold value, m_{crit}. This implies that the solution will enter the active population if $h(t; \underline{m}_{\max}) > m_{crit}$, which is analogous to

$$F(t) > \sigma(m_{crit}; \underline{m}_{\max}) \tag{3.31}$$

for some values of t. By standard calculus arguments, (3.31) will occur if

$$F(t_{\max}) > \sigma(m_{crit}; \underline{m}_{\max}) \tag{3.32}$$

where the maximum for $F(t)$ occurs at $t = t_{\max}$. We denote (3.32) as the *activation criterion* for (3.7). By the same argument, for the entire population to enter M_{act} at some point, then we can derive the *entire activation criterion* as

$$F(t_{\max}) > \sigma(m_{crit}; \underline{m}_{\min}). \tag{3.33}$$

3.4.3 Numerical Simulation Issues and Derivation of an Averaged time-dependent Fisher's Equation

We depict the $u = 1$ isocline for a numerical simulation of (3.7) in Figure 3.3 with $g(m) = \alpha m(1-m)$, $f(t) = \beta \sin(\gamma t)$, $\alpha = 1/2, \beta = 1$, and $\gamma = 1.615$. These terms will be explained in Example 3 below. The numerical scheme is the same as that described in Section 3.3.1, except $\Delta m = 0.025$ here. From
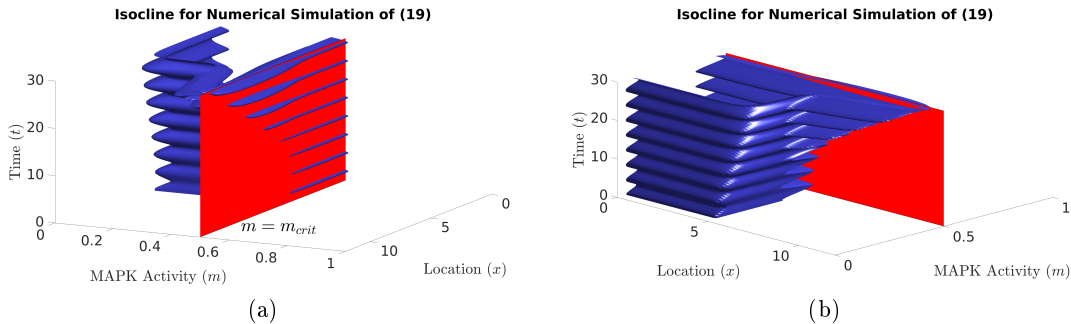
Figure 3.3: Two views of the isocline for $u = 1$ from a numerical simulation of (3.7) with $g(m) = \alpha m(1 - m)$ and $f(t) = \beta \sin(\gamma t)$ for $\alpha = 0.5, \beta = 1, \gamma = 1.615, \bar{D} = 100, \bar{\lambda} = 1/100$, and an initial condition of $\phi_1(m) = {}^{10}/3 I_{[.05,0.35]}(m)$ and $\phi_2(x) = I_{[x \leq 5]}(x)$. The numerical scheme is discussed in Section 3.4.3 and the step sizes used are $\Delta m = {}^{1}/80, \Delta x = {}^{1}/5, \Delta t = 10^{-3}$. From (3.32), the simulation should not cross the $m = m_{crit}$ plane, which is given by the red plane. We see in frame (a) that the simulation does cross the $m = m_{crit}$ plane due to numerical diffusion, which causes the high rate of diffusion along x seen in frame (b).

(3.32), we see that this simulation should not enter the active population with an initial condition of $\phi_1(m) = {}^{10}/3 I_{[.05,0.35]}(m)$, where $I_M(m)$ denotes an indicator function with support for $m \in M$. In Figure 3.3, however, we observe that the numerical simulation does enter the active population, which causes a significant portion of the population to incorrectly diffuse into the wound at a high rate.

Numerical simulations of advection-driven processes have been described as an "embarrassingly difficult" task, and one such problem is the presence of numerical diffusion [Leonard, 1991, Thackham et al., 2008a]. Numerical diffusion along the m-dimension is hard to avoid and here causes a portion of the cell population to enter the active population in situations where the it should not. Numerical diffusion can be reduced with a finer grid, but this can lead to excessively long computation times. With the aid of the activation curves given by (3.17), however, we can track progression of cells in the m-dimension analytically and avoid the problems caused by numerical diffusion completely.

To avoid the problems caused by numerical diffusion, we derive a time-dependent Fisher's Equation for $w(t, x)$ in the appendix that represents the average behavior along m with time-dependent

53

diffusion and proliferation terms. This equation is given by

$$w_t = D(t)w_{xx} + \lambda(t)w(1-w), \qquad (3.34)$$

$$w(t=0,x) = \phi_2(x)$$

$$w(t, x = -\infty) = 1 \qquad w(t, x = \infty) = 0$$

where

$$D(t) = \bar{D} + (1-\bar{D})\Phi_1(\psi(t))$$

$$\lambda(t) = \bar{\lambda} + (1-\bar{\lambda})\Phi_1(\psi(t))$$

for $\psi(t) = \sigma^{-1}(-F(t); m_{crit})$ denoting the threshold value in $\underline{m}$ that separates the active and inactive regions.

3.4.4 Three biologically-motivated examples

We next consider three examples of (3.7) that pertain to common patterns of biochemical activity during wound healing. We will use numerical simulations of (3.34) to investigate how different patterns of activation and deactivation over time affect the averaged cell population profile. We will also investigate how the profile changes when crossing the activation and entire activation thresholds derived in (3.32) and (3.33). In each example, we fix $m_{crit} = 0.5, \bar{D} = 100, \bar{\lambda} = 1/100,$ $\phi_1(m) = {}^{10}/{3}I_{[.05,0.25]}(m), \phi_2(x) = I_{(-\infty,5]}(x)$ and $g(m) = \alpha m(1-m),$ and use a different terms for $f(t)$ to mimic different biological situations. The choice for $g(m)$ ensures that the distribution along m stays between $m = 0$ and $m = 1$. The method of lines is used to compute (3.34), with a standard central difference discretization scheme in space and MATLAB's ode45 command is used for time integration.

Example 1: Single Sustained MAPK activation wave: $f(t) = 1$

In this example, we consider a case where we observe the entire cell population approach a level of $m = 1$ over time. Such a scenario may represent the sustained wave of ERK 1/2 activity observed in MDCK cells from [Matsubayashi et al., 2004]. We use $f(t) = 1$ to observe this behavior. We
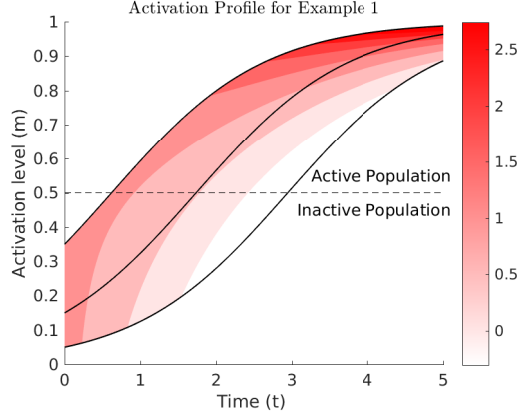
Figure 3.4: The analytical solution for the activation profile, $p(t,m)$, for Example 1 for $\alpha = 0.5$, and $\phi_1(m) = I_{(0.05,0.35)}(m)$. The solid black curves denote $h(t;\underline{m})$ for $\underline{m} = 0.05, 0.15$, and 0.35 and the dashed line denotes $m = m_{crit}$. Note that a log scale is used along p for visual ease.

provide the functions for $\sigma(m;\underline{m}), h(t;\underline{m})$, and $\psi(t)$ that result from this choice of terms in the appendix.

In Figure 3.4, we plot (3.19) to depict an example activation profile, $p(t,m)$, over time to show the activation behavior of the population. As expected, we observe all of $p(t,m)$ converging to $m = 1$. We include some specific plots of the activation curves, $h(t;\underline{m})$, for this example. Note that the density changes along these curves by the height function $\frac{g(\sigma^{-1}(-F(t),m))}{g(m)}$, which is equivalent to the height function of the self-similar traveling wave ansatz made in (3.23).

In Figure 3.5(a), we depict a numerical simulation of $w(t,x)$ over time using (3.34). The Regions denoted as "P" and "D" denote when the population is primarily proliferating ($\Phi_1(\psi(t)) > 1/2$) or diffusing ($\Phi_1(\psi(t)) \leq 1/2$) over time. The profile maintains a high cell density but limited migration into the wound during the proliferative phase and then migrates into the wound quickly during the diffusive phase but can not maintain a high cell density throughout the population. In Figure 3.5(b), we investigate how the numerical simulation profile of $w(t = 40, x)$ changes as α varies from $\alpha = 0$ to $\alpha = 0.2$. In the region denoted "No activation", the entire population is still in the inactive population at $t = 40$ and thus does not progress far into the wound or change with α. The population is split between the active and inactive populations at $t = 40$ for simulations in the region denoted "Activation." The profiles here are very sensitive to increasing values of α, as they migrate further into the wound while maintaining a high density near $x = 0$. Simulations from the region denoted as "Entire Activation" are entirely in the active population by $t = 40$. As α increases, these

Figure 3.5: Numerical simulations of the averaged time-dependent Fisher's equation for Example 1. In (a), we depict a simulation of $w(t,x)$ over time for $\alpha = 0.05$. The letters "P" and "D" denote when the population is primarily proliferating or diffusing, respectively. In (b), we depict how the profile for $w(t=40,x)$ changes for various values of α. The descriptions "No Activation", "Activation", and "Entire Activation" denote values of α for which the population is entirely in the inactive population, split between the active and inactive populations, or entirely in the active population at $t = 40$, respectively.

simulations do not migrate much further into the wound but do have decreasing densities at $x = 0$. These results suggest that a combination of proliferation and diffusion must be used to maximize population migration while maintaining a high cellular density behind the population front. The optimal combination for a far-moving front with a confluent density appears to occur at the entire activation threshold.

Example 2: Single pulse of MAPK activation: $f(t) = \beta e^{\gamma t} - 1$

We now detail an example that exhibits a pulse of activation in the m dimension, which may represent the transient wave of ERK 1/2 activation observed in MDCK cells in [Matsubayashi et al., 2004]. The authors of [Posta and Chou, 2010] proposed that this wave may be caused by the rapid production of ROS in response to the wound, followed by its quick decay or consumption by cells. We now let $f(t) = \beta e^{\gamma t} - 1$. This forcing function arises if ROS is present but decaying exponentially over time and modeled by $s(t) = \beta e^{\gamma t}, \beta > 0, \gamma < 0$ and cells activate linearly in response to the presence of ROS but have a baseline level of deactivation, which may be given by $f(s) = s - 1$. We provide the functions for $\sigma(m; \underline{m}), h(t; \underline{m})$, and $\psi(t)$ that result from this choice of terms in the appendix, as well as the activation criteria.

In Figure 3.6, we use (3.19) to depict the activation profile, $p(t,m)$, over time to show the

Figure 3.6: The analytical solution for the activation profile, $p(t, m)$, for Example 2 for $\alpha = 0.5, \beta = 3, \gamma = -1/4$ and $\phi_1(m) = I_{(0.05, 0.35)}(m)$. The solid black curves denote $h(t; \underline{m})$ for $\underline{m} = 0.05, 0.15$, and 0.35 and the dashed line denotes $m = m_{crit}$. Note that a log scale is used along p for visual ease.

activation behavior of the population. We also include some specific plots of the activation curves, $h(t; \underline{m})$, which show a pulse of MAPK activity in the population that starts decreasing around $t = 5$. Note that $h(t; 0.35)$ crosses the $m = m_{crit}$ line but $h(t; 0.05)$ does not, so (3.32) is satisfied for this parameter set (the population becomes activated) but (3.33) is not (the entire population does not become activated).

In Figure 3.7(a), we depict a numerical simulation of (3.34) for this example. The population quickly transitions to a diffusing stage due to the pulse of MAPK activation and shows the smaller densities (u approximately less than 0.2) migrating into the wound rapidly while the density behind the population front drops. As the pulse of MAPK activation ends and the population transitions back to a proliferating phenotype, the populations restores a high density behind the cell front and begins to develop a traveling wave profile, as suggested by the parallel contour lines. In Figure 3.7(b), we investigate how the numerical solution profile for $w(t = 30, x)$ changes as β varies from $\beta = 2$ to $\beta = 9$ while keeping all other parameters fixed. We observe that the profile is the same for all inactive simulations as (3.32) is not satisfied. As β increases past the activation threshold, the profile shows increased rates of migration into the wound. After passing the entire activation threshold (3.33), the profile continues to migrate further as β increases, but appears less sensitive to β. This increased migration is likely due to the population spending more time in the active population for larger values of β. Note that for all simulations shown, the pulse of MAPK activation has finished by $t = 30$.
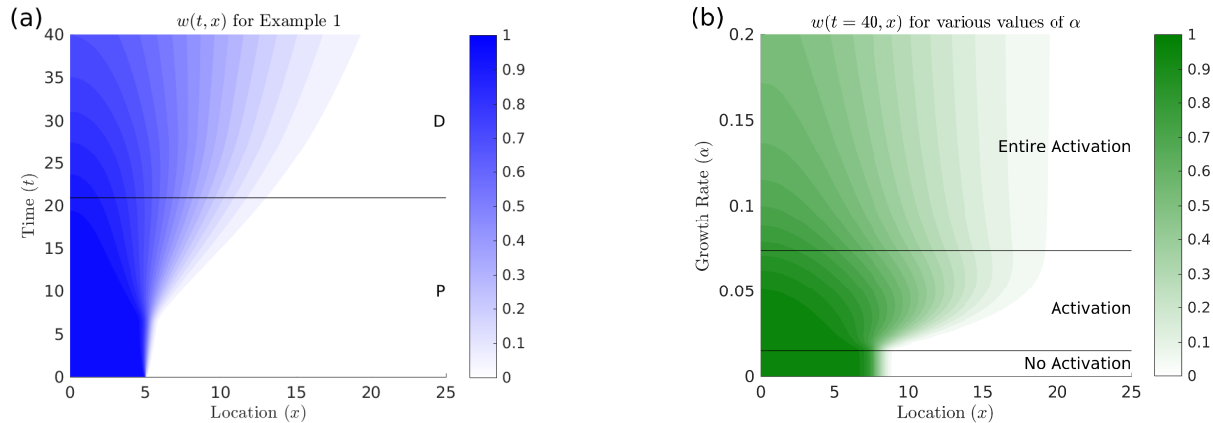
57

Figure 3.7: Numerical simulations of the averaged time-dependent Fisher's equation for Example 2. In (a), we depict a simulation of $w(t,x)$ over time for $\alpha = 1, \beta = 8, \gamma = -1$. Regions denoted with a "P" or "D" denote when the population is primarily proliferating or diffusing, respectively. In (b), we depict how the profile for $w(t = 30, x)$ changes for various values of β. The descriptions "No activation", "Activation", and "Entire Activation" denote values of β for which the population is entirely in the inactive population, split between the active and inactive populations, or entirely in the active population at $t = t_{\max}$.

Example 3: Periodic pulses of MAPK activation: $f(t) = \beta \sin(2\pi t/\gamma)$

As a last example, we exhibit a scenario with periodic waves of activity. Such behavior was observed in some of the experiments performed in [Zi et al., 2011], in which cell cultures of the HaCaT cell line were periodically treated with TGF-β to investigate how periodic treatment with TGF-β affects activation of the SMAD pathway (the canonical pathway for TGF-β, which also influences cell proliferation and migration). We let $f(t) = \beta \sin(2\pi t/\gamma), \beta, \gamma > 0$, which occurs if the concentration of TGF-β over time is given by $s(t) = 1 + \sin(2\pi t/\gamma)$, and cells activate linearly in response to s and have a baseline rate of deactivation, given by $f(s) = \beta(s - 1)$. Note that $f(t)$ has frequency γ. We provide the functions for $\sigma(m; \underline{m}), h(t; \underline{m})$, and $\psi(t)$ that result from this choice of terms in the appendix, as well as the activation criteria.

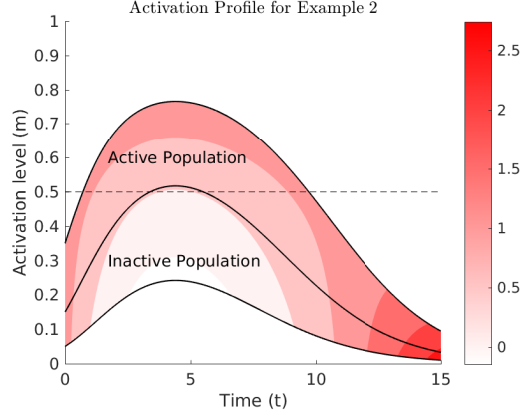In Figure 3.8, we use (3.19) to depict the activation profile, $p(t, m)$, over time to show the activation behavior of the population. We also include some specific plots of the activation curves $h(t; \underline{m})$, which demonstrate periodic waves of activation along m. Note that $h(t; 0.05)$ crosses the $m = m_{crit}$ line, so (3.33) is satisfied, and the entire population becomes activated at some points during the simulation.

In Figure 3.9(a), we depict a numerical simulation of (3.34) for this example. The population
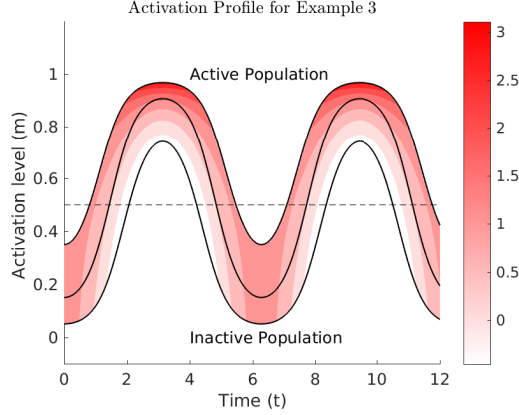
Figure 3.8: The analytical solution for the activation profile, $p(t, m)$, for Example 3 for $\alpha = 1/2, \beta = 4, \gamma = 1$ and $\phi_1(m) = I_{(0.05, 0.35)}(m)$. The solid black curves denote $h(t; \underline{m})$ for $\underline{m} = 0.05, 0.15$, and 0.35 and the dashed line denotes $m = m_{crit}$. Note that a log scale is used along p for visual ease.

phenotype has a period of 4π, and we see that the lower densities migrate into the wound most during the diffusive stages, whereas all densities appear to migrate into the wound at similar speeds during the proliferative stages. In Figure 3.9(b), we investigate how the profile for $w(t = 40, x)$ changes as γ varies between $\gamma = 0$ and $\gamma = 37$ while keeping all other parameters fixed. All "No Activation" profiles appear the same because (3.32) is not satisfied. As γ increases past this threshold, more of the population becomes activated during the simulation, culminating in a maximum propagation of the population at the entire activation threshold given by (3.33). As γ increases above this threshold, the population tends to migrate less far, although the population does migrate far for γ near 33. In the "Entire Activation" Region, the population appears to spend too much time in the active population and diffuses excessively with limited proliferation. These simulations lead to shallow profiles that do not migrate far into the wound.

3.5   Discussion and Future work

We investigate a structured Fisher's Equation that incorporates an added dimension for biochemical activity that influences population migration and proliferation. The method of characteristics proves to be a useful way to track the progression along the population activity dimension over time. We prove the existence of a unique self-similar traveling wave solution to the equation when diffusion and proliferation do not depend on MAPK activity. The height function of the self-similar traveling wave ansatz along characteristic curves is demonstrated in Figures 3.4, 3.6, and 3.8. These methods
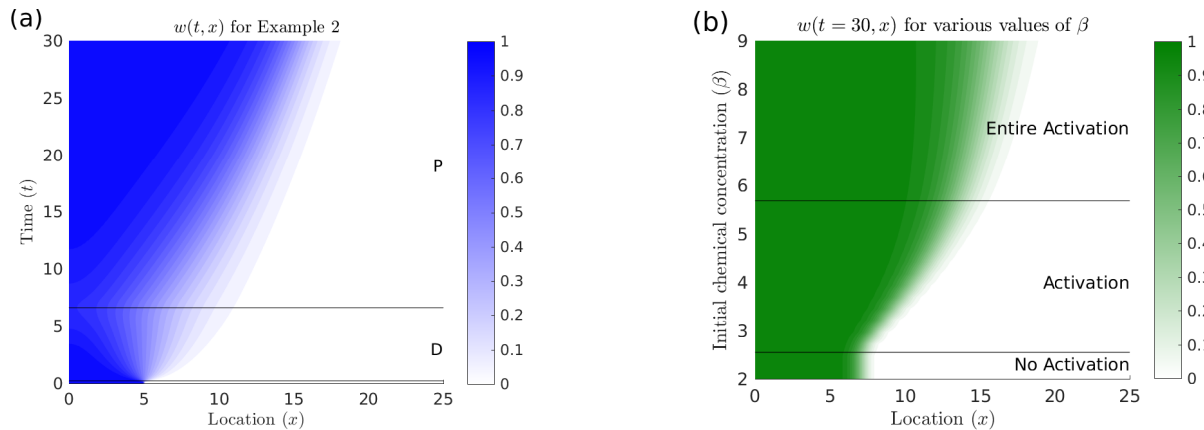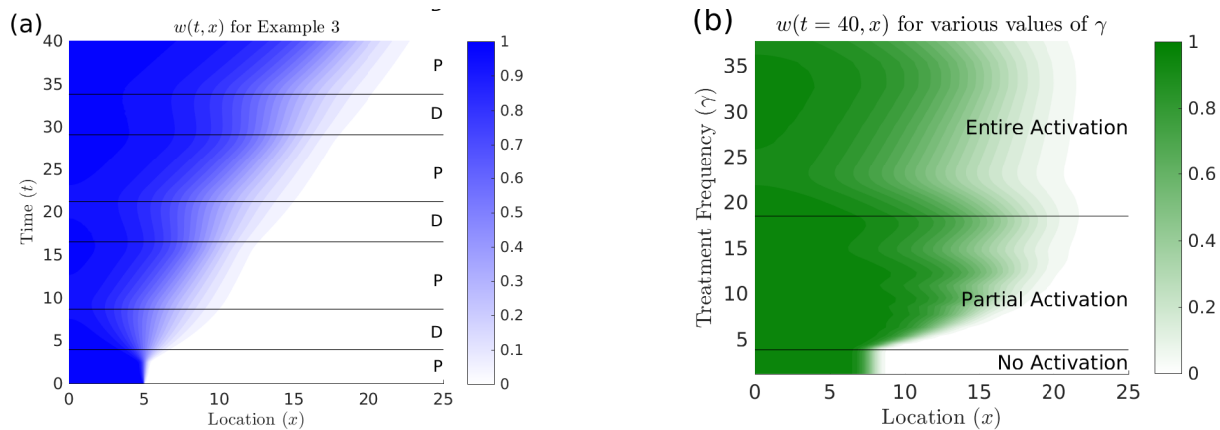
Figure 3.9: Numerical simulations of the averaged time-dependent Fisher's equation for Example 3. In (a), we depict a simulation of $w(t, x)$ over time for $\alpha = 0.5, \beta = 1$, and $\gamma = 4\pi$. Regions denoted with a "P" or "D" denote when the population is primarily proliferating or diffusing, respectively. In (b), we depict $w(t = 40, x)$ for various values of γ. The descriptions "No activation", "Activation", and "Entire Activation" denote values of γ for which the population is entirely in the inactive population, split between the active and inactive populations, or entirely in the active population at $t = t_{\max}$.

are also applicable to structured reaction-diffusion and chemotaxis equations.

Activation of the MAPK signaling cascade is known to influence collective migration during woung healing through cellular migration and proliferation properties. For this reason, we also consider a structured PDE model in which the rates of cellular diffusion and proliferation depend on the levels of MAPK activation in the population. We derive two activation criteria for the model to establish conditions under which the population will become activated during simulations. As numerical simulations of the structured equation are prone to error via numerical diffusion, we derive a time-dependent PDE equation in time and space to represent the average population behavior along the biochemical activity dimension. Using this time-dependent equation, we exhibit three simple examples that demonstrate biologically relevant activation patterns and their effects on population migration: a sustained wave of activity, a pulse of activity, and periodic pulses of activity. We find that the population tends to migrate farthest while maintaining a high cell density at the entire activation threshold value, given by (3.33), for the sustained wave and periodic pulse patterns of activation. The single pulse case continues migrating further into the wound after passing the entire activation threshold but appears less sensitive after doing so.

A natural next step for this analysis is to use a structured population model of this sort in combination with biological data to thoroughly investigate the effects of MAPK activation and

60

deactivation on cell migration and proliferation during wound healing. Previous mathematical models have focused on either collective migration during wound healing assays in response to EGF treatment (while neglecting the MAPK signaling cascade) [Johnston et al., 2015, Nardini et al., 2016] or MAPK propagation during wound healing assays (while neglecting cell migration) [Posta and Chou, 2010]. To the best of our knowledge, no mathematical models have been able to reliably couple signal propagation and its effect on cell migration during wound healing. The examples detailed in this work intentionally used the simplest terms possible as a means to focus on the underlying mathematical aspects. With a separate in-depth study into the biochemistry underlying the MAPK signaling cascade and its relation with various cytokines or growth factors, more complicated and biologically relevant terms for $g(m), f(s)$, and $s(t)$ can be determined to help elucidate the effects of MAPK activation on cell migration during wound healing.

As various pathways become activated and cross-talk during wound healing to influence migration [Guo and Wang, 2009], future studies could also investigate a population structured along multiple signaling pathways, $u(t, x, \vec{m})$ for the vector $\vec{m} = (m_1, m_2, \ldots, m_n)^T$. Because the cell population also produces cytokines and growth factors for paracrine and autocrine signaling during wound healing, these models would also benefit from variables representing ROS, TGF-β, EGF, etc.

While the main motivation for this study is epidermal wound healing, there are potential applications in other areas of biology. Fisher's equation has also been used to study population dynamics in ecology and epidemiology [Ai and Huang, 2005, Hastings et al., 2005, Shigesada and Kawasaki, 1997]. Our framework could be extended to a case where an environmental effect, such as seasonal forcing, impacts species migration or susceptibility of individuals to disease. The results presented here may thus aid in a plethora of mathematical biology studies.

Chapter 4

The Influence of Numerical Error on an Inverse Problem Methodology[1]

4.1 Introduction

Partial Differential Equations (PDEs) are frequently used to study various scientific systems. Due to their complicated nature, deriving analytical solutions is frequently difficult or impossible, so scientists must frequently use numerical methods to approximate their solutions. Numerical error is inevitable in these approximations, but any useful numerical method will have a predetermined error structure. How numerical approximation error influences various aspects of scientific studies, such as parameter estimation routines and confidence interval computation, is an interesting and poorly-understood question [Banks and Fitzpatrick, 1990, Xue et al., 2010].

In [Ackleh and Thibodeaux, 2008], the authors consider an advection-driven model of erythyropoiesis, an important step in red blood cell development. The authors show that an inverse problem is asymptotically well-posed as the numerical step size used for grid computation, h, approaches zero if an upwind numerical method is used to approximate the equation.. In practice, we cannot let h approach zero but must instead choose a finite value for h to estimate the parameters. Furthermore, advection equations like those used in [Ackleh and Thibodeaux, 2008] are known to cause a multitude of numerical issues [Leonard, 1991]. For instance, numerical diffusion or unstable oscillations occur for many numerical schemes, especially when the true solution is discontinuous [Randall J. Leveque, 1992, Thackham et al., 2008a]. Upwind methods for advection-dominated equations are popular because they can avoid spurious oscillations by satisfying certain monotonic-

---

[1]This chapter is being prepared for a publication.

ity constraints, such as the Courant-Friedrichs-Lewy (CFL) Condition [LeVeque, 2007]. Upwind methods cause their own trouble, however, as they can lead to large amounts of numerical diffusion near points of discontinuity, which is disadvantageous for studying systems with discontinuous solutions.

In this study, we will use a simple advection equation to demonstrate how various numerical finite difference methods can be used to approximate an analytic solution and estimate the true parameter estimate values. We begin in Section 4.2 by introducing some preliminary information, including the equation under consideration and its analytic solution, the inverse problem methodology, and the numerical methods used in this study. In Section 4.3, we introduce a numerical cost function that is used in the inverse problem computations and discuss its asymptotic properties in terms of both data and numerical precision. In Section 4.4, we discuss our results in using various numerical methods on different-precision numerical grids to estimate parameters from various data sets. We use residual analysis in Section 4.5 to infer how an autocorrelation statistical method can be used to quantify numerical error in the inverse problem for first order numerical approximations and how it improves confidence interval computation. Based on these results, we provide some suggestions for practitioners in Section 4.6 to ensure the results of their inverse problem routines are as reliable as possible. We make concluding remarks and discuss future work in Section 4.7.

4.2 Mathematical Preliminaries

In this section, we outline some necessary information regarding our inverse problem methodology. In particular, we discuss the advection equation and choice of parameterization in Section 2.1, how we generate artificial data for this study in Section 2.2, the numerical scheme used are discussed in Section 2.3, and we conclude with our numerical cost function used to compare numerical simulations to data in Section 2.4.

4.2.1   Model Equation

We will consider the advection equation given by

$$u_t + (g(x|\theta)u)_x = 0, \ u = u(t,x) \tag{4.1}$$

$$u(t = 0, x) = \phi(x)$$

$$x \in X, t \in T, \theta \in Q$$

where $g(x|\theta)$ is a spatially-dependent advection rate with parameter vector θ from the parameter space Q, $\phi(x)$ is the initial condition, and $u(t, x|\theta)$ denotes the quantity of interest at time t from the time domain, T, and spatial location x from the spatial domain, X. Throughout this chapter, we let $X = [0, 1]$ and $T = [0, 10]$. We will suppress the dependence of $g(x)$ and $u(t, x)$ on θ when this dependence is unnecessary throughout this document.

The method of characteristics can be used to show the analytical solution to Equation (4.1) of

$$u_0(t, x) = \begin{cases} \dfrac{g(\sigma^{-1}(-t, x))}{g(x)} \phi(\sigma^{-1}(-t, x)) & \sigma^{-1}(t, 0) \leq x \leq 1 \\ 0 & \text{o.w.} \end{cases} \tag{4.2}$$

where $\sigma^{-1}(t, x)$ is the characteristic curve that satisfies

$$\frac{\partial}{\partial t}\sigma^{-1}(t, x) = g\left(\sigma^{-1}(t, x)\right), \ \sigma^{-1}(t = 0, x) = x.$$

See [Webb, 2008] for more information and [Nardini and Bortz, 2018] for an illuminating application. If $g(x) > 0$ and $g'(x) > 0$, then Equation (4.1) could be used to mimic a cell population that is chemotaxing over time, in which cells at the wound edge move faster and faster [Keller and Segel, 1971]. For this reason, we choose

$$g(x) = \alpha \sqrt[\beta]{x}, \quad \alpha, \beta > 0$$

for $x > 0$ and for $(\alpha, \beta) \in Q = \mathbb{R}^2$. The choice of $g(x)$ above yields the characteristic curves $\sigma^{-1}(t, x) = \left[\alpha(1 - 1/\beta)t + x^{1-1/\beta}\right]^{\beta/(\beta-1)}$.

We will consider two initial conditions to test their influence on numerical convergence and the

inverse problem. We will denote the first initial condition as Gaussian, which is given by

$$\phi_1(x) = e^{-(x-0.2/\sqrt{.005})^2} \tag{4.3}$$

and the second initial condition as a front, which is given by

$$\phi_2(x) = \begin{cases} 5 & x \leq 0.2 \\ 0 & otherwise. \end{cases} \tag{4.4}$$

We will focus on the main results for $\phi(x) = \phi_2(x)$ in the main body of this document, but most of the corresponding results when $\phi(x) = \phi_1(x)$ will be given throughout the appendix. We will make note of how the results change between these two initial conditions.

4.2.2 Artificial Data Generation

We will consider various artificial data sets to from Equation (4.1). These data sets will be generated by adding Gaussian noise to the analytical solution

$$Y_{i,j} = u_0(t_i, x_j | \theta_0) + \mathcal{E}_{i,j}, \ \ \mathcal{E}_{i,j} \sim \mathcal{N}(0, \eta^2), \ \ i = 1, .., M \ \ j = 1, .., N. \tag{4.5}$$

using the uniform grids:
$$t_i = \left\{ \frac{10(i-1)}{M} \right\}_{i=0}^{M}, \ x_j = \left\{ \frac{j-1}{N} \right\}_{j=0}^{N}$$

for $\theta_0 \in Q$. Note that this is a very simple form of data generation, and others could have been used, including multiplicative noise (see Banks and Tran [2009, § 3.2]). Changing the way we generate this model would also require changing the cost function that we use in Equation 4.6. We will generate several data sets for both initial conditions given by Equations (4.3) and (4.4). To investigate the influence of data error and number of data points on the inverse problem, we will choose various values of N and η to generate these different data sets. M will be fixed at 6 for simplicity for all cases. We choose $\theta_0 = (0.3, 0.4)^T$ for data sets where $\phi(x) = \phi_1(x)$ and $\theta_0 = (0.3, 0.5)^T$ for data sets where $\phi(x) = \phi_2(x)$. An example data sets is depicted for $\eta = 0.1, N = 11$ in Figure 4.1.
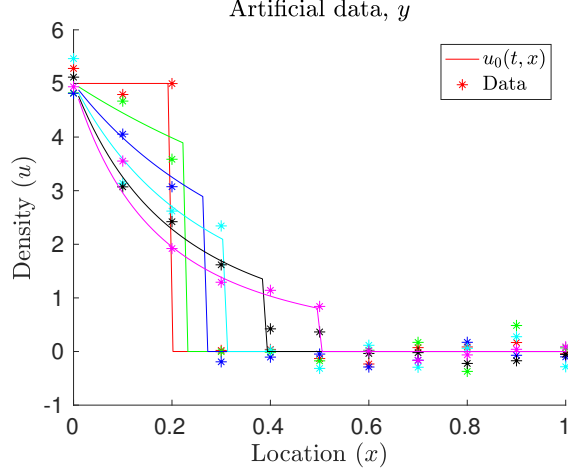
Figure 4.1: Artificial data from Equations (4.1) and (4.5) for $\eta = 0.1, N = 10$ $\phi(x) = \phi_2(x)$. Dots denote the artificial data and the solid lines denote the analytical solution given by Equation (4.2). Red plots denote $t = 0$, green plots denote $t = 2$, blue plots denote $t = 4$, cyan plots denote $t = 6$, black plots denote $t = 8$, and magenta plots denote $t = 10$.

4.2.3   Numerical Methods

We will consider four commonly-used finite difference schemes to simulate Equation (4.1). These four schemes are the upwind, Lax-Wendroff, Beam-Warming methods, and the upwind method with flux limiters. The first three methods are discussed and presented in the popular monograph by Leveque [LeVeque, 2007], and the final method is discussed in [Randall J. Leveque, 1992, §16.2]. We will denote $u(t, x|h, \theta) : T \times X \times \mathbb{R} \times Q \to \mathbb{R}^{MN}$ as the numerical approximation to equation (4.1) that has been computed with a spatial step size of h for parameter set θ and then interpolated to the grid used for data comparison.[2] We will also denote $u_0(t, x|\theta) : [T, X] \to \mathbb{R}^{MN}$ as the analytical solution evaluated at the grid points chosen for data comparison throughout this document.

A common practice in numerical analysis is to compute the order of convergence of a numerical method. Guided by [LeVeque, 2007], we define the error for a numerical scheme as

$$E(h) = \|u(t_i, x_j|h, \theta) - u_0(t_i, x_j|\theta)\|_1,$$

where $\| \cdot \|_1$ denotes the 1-norm in $\mathbb{R}^{MN}$. We then estimate the order of the scheme by fitting a line to the natural log of the error, $\ln(E(\vec{h}))$ against $\ln(\vec{h})$, where $\vec{h}$ denotes the vector of step sizes

---

[2] Note that the interpolation order is chosen larger than the numerical error so that this step does not influence convergence rates.

h used to compute each $u(x, t|\theta, h)$. The slope of these lines gives the order of convergence, and we denote this *numerical order of convergence* as p throughout this document.

The upwind method is first order accurate, meaning that $\|u(t, x|h, \theta) - \tilde{u}_0(t, x|\theta)\|_1 = \mathcal{O}(h^1)$ when $u_0(t, x)$ is continuous and $\|\cdot\|_1$ denotes the 1-norm for vectors in $\mathbb{R}^{MN}$. The Lax-Wendroff and beam warming method are second order accurate so that $\|u(t, x|h, \theta) - u_0(t, x|\theta)\|_1 = \mathcal{O}(h^2)$ with these methods. The upwind and beam warming methods use one-sided derivative approximations from the direction where information is coming from, which generally makes them more accurate than their respective same-order method when computing advection equations.

While these schemes are often referred to as first- or second-order accurate, this is only true when the analytical solution, $u_0(t, x)$, is continuous with respect to x. The order for these schemes can be computed when $u_0(t, x)$ is discontinuous using the theory of modified equations, which is described in [Randall J. Leveque, 1992, § 11], and the relevant results are presented in this and the following paragraph. The upwind methods is order $1/2$, and the Lax-Wendroff and Beam-Warming methods are order $2/3$ when $u_0(t, x)$ is discontinuous. The modified equation for a numerical scheme is derived by adding the local truncation error from the numerical method to the original PDE model. So when a numerical method is first (second) order accurate for the PDE model under consideration, it will be second (third) order accurate for the modified equation.

The modified equations for the first order upwind and Lax-Friedrichs methods can be computed as the advection-diffusion equation

$$u_t + (g(x)u)_x = D(h, x)u_{xx}, \ D(h, x) = \mathcal{O}(h)$$

so we can expect error from these methods to be diffusive. The modified equations for the second order Lax-Wendroff and Beam-Warming methods can be computed as the advection-dispersion equation

$$u_t + (g(x)u)_x = \mu(h, x)u_{xxx}, \ \mu(h, x) = \mathcal{O}(h^2)$$

so we can expect error from these methods to be dispersive as modes with different wave numbers propagate forward at different speeds.

Flux limiters are another popular method to simulate advection equations with sharp-fronts

[Sweby, 1984, Thackham et al., 2008a]. When flux limiters are used, the spatial gradients near each computational point are estimated before computation. Flux limiters with an upwind scheme make computations approximately second-order accurate near smooth spatial points and first order accurate near points of discontinuity. These scheme thus prevents spurious oscillations near the discontinuity, and instead allows for some numerical diffusion in this region. In this chapter, we will use the Van-Leer flux limiter [Randall J. Leveque, 1992].

In Table 4.2, we depict the calculated values of p for $\phi(x) = \phi_2(x)$ and $\theta = \theta_0$, and see that our calculated numerical rate of convergence for the upwind scheme is consistent with the theory (close to $1/2$), but the order for the Lax-Wendroff and Beam-Warming schemes are smaller and larger than expected, respectively (p is calculated as 0.4737 for Lax-Wendroff and 0.7876 for Beam-Warming, when theory suggests these both should be $2/3$). The upwind scheme with flux limiters has a calculated numerical order of convergence of 0.9570. We show in Table 2 in the appendix that the calculated rates of numerical order of convergence for $\phi(x) = \phi_1(x)$ are consistent with theory for continuous solutions. We also calculate that the upwind scheme with flux limiters has an order of convergence of 0.9183 in this case.

4.2.4 Inverse problem

For a given data set $\{y_{i,j}\}_{i=1,..,M}^{j=1,..,N}$, and step size, h, we estimate the parameter set, $\hat{\theta}_{OLS}^{M,N}(h) \in Q$, that minimizes the cost functional for M spatial and N temporal data points given by

$$J^{M,N}(h,\theta) := \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (y_{i,j} - u(t_i, x_j|h,\theta))^2. \tag{4.6}$$

We refer the reader interested in more details about inverse problems to [Banks and Fitzpatrick, 1990, Banks and Tran, 2009, Banks et al., 2014]. The proofs regarding the consistency of $\hat{\theta}_{OLS}^{M,N}(h)$ as an estimator for $\theta_0 = (\alpha_0, \beta_0)$, the true parameter values, require one to work on a compact parameter space, $Q_{ad} \subset Q$, so we choose $Q_{ad} = [0,10]^2$ in this chapter. For each h, we also use temporal step size $k = \lambda h, \lambda = 1/2$ to ensure numerical stability. We use MATLAB's built-in command fmincon to perform numerical optimization. We will perform this optimization on every data set for values of h given by $h = \frac{1}{10}\frac{1}{2^i}$, $i = 0, 1, .., 6$ to obtain $\hat{\theta}_{OLS}^{M,N}(h)$. We will also perform the inverse problem for multiple data sets with multiple numbers of data points and data

error levels. For $\phi(x) = \phi_1(x)$, we consider data sets for $N = 11, 31$, or 51 and $\eta = 0, 10^{-4}, 5 \times 10^{-4}, 10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}$, or 2×10^{-1}. For $\phi(x) = \phi_2(x)$, we consider data sets for $N = 11, 31$, or 51 and $\eta = 0, 10^{-1}, 1.5 \times 10^{-1}, 2 \times 10^{-1}, 3 \times 10^{-1}, 5 \times 10^{-1}$, or 1.

4.3 Inverse Problem Theory

In this Section, we will discuss asymptotic theories of the inverse problem as data increases ($M, N \to \infty$) and as numerical resolution increase ($h \to 0$). In Section 4.3.1, we discuss some of the theory regarding $\hat{\theta}_{OLS}^{M,N}(h)$, and in Section 4.3.2, we discuss the asymptotic theory regarding $J_{OLS}^{M,N}(h, \theta)$.

We introduce some useful notational conveniences here. Throughout this section, $u_0(t_i, x_j)$ denotes the analytical solution at the point (t_i, x_j),

$\vec{u}_0(t, x) = [u_0(t_1, x_1), ..., u_0(t_1, x_N), u_0(t_2, x_1), ..., u_0(t_2, x_{1N}), ..., u_0(t_M, x_N)]^T$ denotes the $MN \times 1$ vector of the analytical solution at all data points $(t_i, x_j)_{i=1,..,M}^{j=1,...,N}$. Similarly, $u(t_i, x_j | h)$ denotes a numerical approximation at the point (t_i, x_j) and $\vec{u}(t, x | h)$ denotes the $MN \times 1$ vector of the numerical approximation at the data points. The terms $u_0(t, x)$ and $u(t, x | h)$ denote the function defined on $T \times X$. The vector $\nabla \vec{u}_0(t, x)$ is the $MN \times 2$ matrix of the gradient of the analytical solution with respect to θ (also known as the sensitivity equations). The sensitivity equation at a point is denoted as $\nabla u_0(t_i, x_j)$. The same notation is used for the numerical approximations, where $\nabla \vec{u}(t, x | h)$ is an $MN \times 2$ matrix, and $\nabla u(t_i, x_j | h)$ denotes the sensitivity equation at one point. The vector $\vec{\epsilon}$ denotes the $MN \times 1$ vector of realizations of Gaussian error terms, and $\epsilon_{i,j}$ denotes one realization at location (t_i, x_j). The vector $\vec{y}$ denotes the $MN \times 1$ vector of data points.

4.3.1 Theory on $\hat{\theta}_{OLS}^{M,N}(h)$

Before discussing the convergence of the numerical cost function, $J^{M,N}(h, \theta_0)$, it is useful to discuss the asymptotic properties of the OLS parameter estimator, $\hat{\theta}_{OLS}^{M,N}(h)$. In [Banks and Fitzpatrick, 1990], it is shown that this is a consistent estimator, meaning that as $M, N \to \infty$ and $h \to 0$, then $\hat{\theta}_{OLS}^{M,N}(h) \to \theta_0$ almost surely. It is also useful to discuss the known properties of the distribution of the estimator as M, N get asymptotically large and h asymptotically small. These asymptotic properties are widely discussed in the absence of numerical error and are derived in [Seber and Wild, 1988, § 12]. The most important result being that as $M, N \to \infty$, this estimator becomes

asymptotically normally distributed as

$$\hat{\theta}_{OLS}^{M,N} \sim \mathcal{N}(\theta_0, \eta^2 C), \ C = (\nabla \vec{u}_0(t,x|\theta_0)^T \nabla \vec{u}_0(t,x|\theta_0))^{-1}. \tag{4.7}$$

Note that the normality of the error terms, $\epsilon_{i,j}$, is not needed for this asymptotic result, so long as their mean is zero, their variance is finite, and $u_0(t,x|\theta)$ is sufficiently smooth with respect to θ.

In Section A.12 in the appendix, we show some of the main steps to show Equation (4.7), as well as similar steps to show that

$$\hat{\theta}_{OLS}^{M,N}(h) \sim \mathcal{N}(\theta_0, \eta^2 C_h), \ |C_h - C| = \mathcal{O}(h^p)$$

where we thus see that the entries of the covariance matrix of $\hat{\theta}_{OLS}^{M,N}(h)$ are order p convergent to the entries of the true covariance matrix.[3] Note that C_h is the inverse of a positive-definite matrix, so C_h is also positive definite. We will denote its Cholesky decomposition as $C_h = R_h^T R_h$.

4.3.2 Convergence of $J^{M,N}(h,\theta)$

The cost function from Equation (4.6) is commonly used in inverse problem methodologies [Banks and Tran, 2009]. This cost functional was investigated in-depth in [Banks and Fitzpatrick, 1990], where it was shown as a consistent estimator as $h \to 0$ and $M, N \to \infty$ under some reasonable assumptions, meaning that $\hat{\theta}_{OLS}^{M,N}(h) \to \theta_0$ as $M, N \to \infty$ and $h \to 0$. In this section, we discuss the asymptotic properties of this function as $h \to 0$ and $M, N \to \infty$ to elucidate the analysis from future sections.

---

[3]It is important to note that this *does not* imply that C_h is close in the norm-sense to the inverse of a matrix close to the matrix $\nabla \vec{u}_0(t,x|\theta_0)^T \nabla \vec{u}_0(t,x|\theta_0)$ [Higham, 1996].

	Asymptotic Properties ($h \to 0$)	Asymptotic Properties ($M, N \to \infty$)
A	-	Converges in distribution to $\mathcal{N}(\eta^2, \frac{2\eta^2}{MN})$
B	Random variable with mean η^2/MN variance unknown	Converges to $J^*(\theta)$
C	Converges to 0 with order $\mathcal{O}(h^p)$	Converges to $\int (u_0(t, x\|\theta) - u(t, x\|h, \theta))^2 d\nu(t) d\chi(x),$
D	Random variable with mean 0 variance will converge with order $\mathcal{O}(h^{2p})$	Converges to 0 by LLN
E	Converges to 0 with order $\mathcal{O}(h^p)$	converges to 0 by LLN
F	Random variable with mean 0 variance converges to $\eta^2 \nabla \vec{u}_0(\theta_0)^T \left(\nabla \vec{u}_0(\theta_0)^T \nabla \vec{u}_0(\theta_0) \right)^{-1} \nabla \vec{u}_0(\theta_0)$ with order $\mathcal{O}(h^{2p})$	Converges to 0 by LLN

Table 4.1: Asymptotic limits for the six terms comprising the numerical cost function given by equation (4.6). A "-" indicates that this term is independent of the limiting variable.

Observe that (omitting indices and independent variables for simplicity)

$$
\begin{aligned}
J^{M,N}(h, \theta) &= \frac{1}{MN} \sum [u_0(\theta_0) + \mathcal{E} - u(h, \theta) + u_0(\theta) - u_0(\theta)]^2 \qquad (4.8) \\
&= \frac{1}{MN} \sum \mathcal{E}^2 + \frac{1}{MN} \sum [u_0(\theta_0) - u_0(\theta)]^2 + \frac{1}{MN} \sum [u_0(\theta) - u(h, \theta)]^2 \\
&\quad + \frac{2}{MN} \sum \mathcal{E}(u_0(\theta_0) - u_0(\theta)) + \frac{2}{MN} \sum \mathcal{E}(u_0(\theta) - u(h, \theta)) \\
&\quad + \frac{2}{MN} \sum [(u_0(\theta) - u(h, \theta))(u_0(\theta_0) - u_0(\theta))] \\
&=: A + B + C + D + E + F
\end{aligned}
$$

We thus observe that the numerical cost function can be broken down into six separate terms. We will now discuss how we can expect these six terms to act asymptotically (as $h \to 0$ and/or $M, N \to \infty$). The asymptotic limits for the terms A through F are summarized in Table 4.1. The explanations for these limits are provided in the appendix in Section A.13.

4.4 Inverse Problem Results

4.4.1 Behavior of Numerical Cost Function

In Figure 4.2, we depict log-log plots of $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N})$ for various values of h for an initial condition of $\phi(x) = \phi_2(x)$. Here, we observe that the cost function appears to converge for larger values of

h before flattening as h becomes smaller and $J^{M,N}(h, \hat{\theta}_{OLS}^{MN}(h))$ approaches η^2. This suggests that numerical error dominates over experimental error from the data until $J^{M,N}(h, \hat{\theta}_{OLS}^{MN}(h))$ reaches η^2. Once $J^{M,N}(h, \hat{\theta}_{OLS}^{MN}(h))$ approaches η^2, it appears that term A from Equation (4.8) is the dominant term, so now experimental error is the dominant term. We depict the log-log plots of $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N})$ for all data sets considered in the appendix in Figure 5 for $\phi(x) = \phi_1(x)$ and in Figure 7 for $\phi(x) = \phi_2(x)$. These figures support the observation that $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N})$ is dominated by numerical error when it converges and experimental error when it flattens.

In Figures 4.2, we observe that $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ converges faster for the second order Lax-Wendroff and Beam-Warming methods than the first order upwind methods. The numerical method used can thus significantly influences the convergence rate of the numerical cost function. We also observe that the upwind scheme with flux limiters can significantly improve computation of $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$, where this can often outperform the two second order schemes. This is interesting, because this is technically a first order scheme, but not surprising because flux limiters have been proposed previously as suitable methods to simulate discontinuous solutions [Thackham et al., 2008a].

In Figure 4.3 we depict plots of $\|\hat{\theta}_{OLS}^{M,N}(h) - \theta_0\|_2$ against h for $\phi(x) = \phi_2(x)$. The Upwind auto method will be discussed in Section 4.5.1. Interestingly, $\hat{\theta}_{OLS}^{M,N}(h)$ can significantly improve, even when $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ appears flat. For example, when $\eta^2 = 0.04, N = 30$ we see that $J^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ does not change with h for the four methods, but $\hat{\theta}_{OLS}^{M,N}(h)$ in general improves for this data set as h decreases for all of these methods. We also observe that it is hard to predict which scheme will estimate θ_0 best. For example, the Beam-Warming scheme tends to do the best overall for all methods considered, but it is outperformed by the upwind scheme with flux limiters when $N = 11$ and $\eta^2 = 0$ or by the upwind scheme when $\eta^2 = 1$ and $N = 30$. We depict plots of $\|\hat{\theta}_{OLS}^{M,N}(h) - \theta_0\|_2$ for all data sets considered in the appendix in Figure 6 for $\phi(x) = \phi_1(x)$ and in Figure 8 for $\phi(x) = \phi_2(x)$. These figures show that the Beam-Warming and Lax-Wendroff schemes typically do best for $\phi(x) = \phi_1(x)$ and that confirm that the best method is hard to predict for $\phi(x) = \phi_2(x)$.

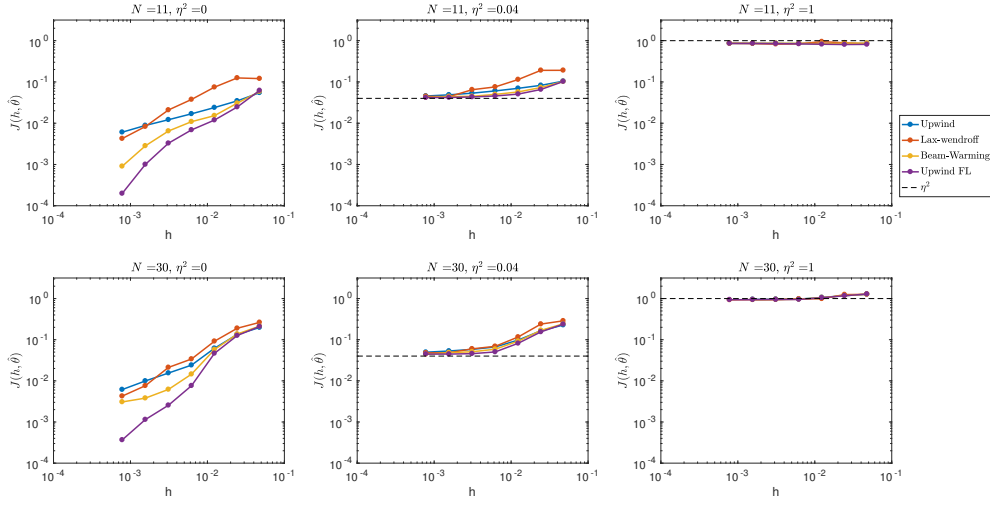Figure 4.2: Plots of $J(h, \hat{\theta})$ for the four schemes considered for various values of h with $\phi(x) = \phi_2(x)$. We depict the results for $N = 11$ or 30 and $\eta = 0, 0.04$, or 1.
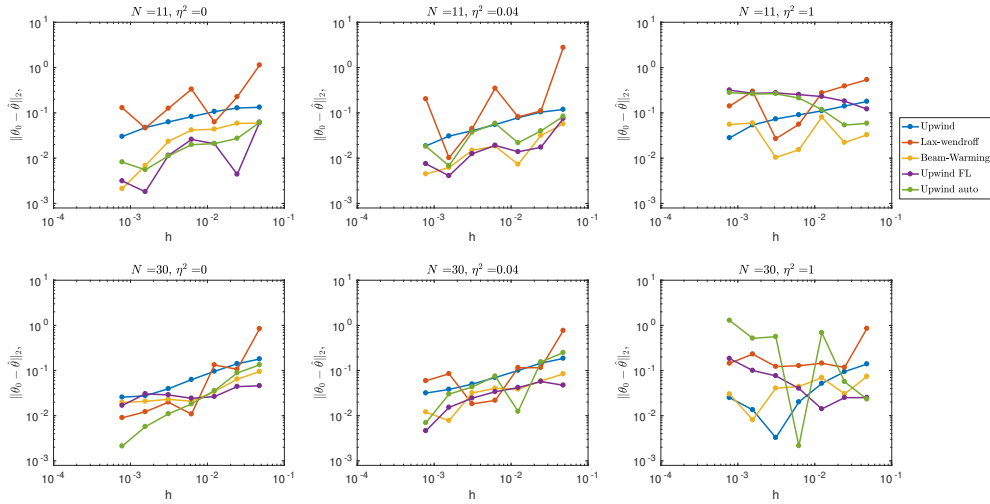

Figure 4.3: Plots of $\|\theta_0 - \hat{\theta}_{OLS}^{M,N}(h)\|_2$ for the four schemes considered for various values of h with $\phi(x) = \phi_2(x)$. We depict the results for $N = 11$ or 30 and $\eta = 0, 0.04$, or 1.

Numerical Method	p	p_J			
		N \ η	0	2×10^{-1}	1
Upwind	0.5839	11	0.517	0.208	-0.002
		30	0.612	0.226	0.040
Lax-Wendroff	0.4737	11	0.966	0.490	-0.011
		30	0.878	0.387	0.062
Beam-Warming	.7876	11	0.785	0.367	0.000
		30	0.987	0.441	0.040
Upwind FL	.9570	11	1.285	0.409	-0.020
		30	1.338	0.505	0.037

Table 4.2: Order estimates for the numerical cost function when $\phi(x) = \phi_2(x)$ and the various schemes are used to estimate θ_0.

4.4.2 Order Estimates

In the previous section, we observed that the numerical cost function appeared to converge at different rates depending on the scheme being used. In this section, we estimate the order of the numerical cost function by fitting the best-fit line between $\log\left(J^{M,N}(\vec{h}, \hat{\theta}_{OLS}^{M,N}(h))\right)$ and $\log(\vec{h})$. The slope of these lines will provide the order of the numerical cost function, and we will denote this *statistical order of convergence* as p_J throughout this document. Note that we use values of h where $\log\left(J^{M,N}(h, \hat{\theta})\right)$ has not yet flattened out when computing the order (for example, for $\phi(x) = \phi_2(x), N = 30, \eta^2 = 0.04$, we use the four coarsest points to compute the order for the upwind scheme with flux limiters). We present the results for $\phi(x) = \phi_2(x)$ in Table 4.2. We observe that p_J is about the same as p for the upwind and Beam-Warming schemes and double the value of p_J for the Lax-Wendroff Scheme when $\eta^2 = 0$ and $\phi(x) = \phi_2(x)$. There is no apparent pattern between p_J and p for the upwind scheme with flux limiters. In the appendix, we depict the values for p_J for all data sets considered for $\phi(x) = \phi_1(x)$ in Table 2 and for $\phi(x) = \phi_2(x)$ in Table 3. For the continuous solutions when $\phi(x) = \phi_1(x)$, we observe that p_J is often double the value of p. The order tends to decrease as η increases for both continuous and discontinuous solutions, eventually reaching zero when experimental error dominates numerical error.

In Figures 4.4 and 4.5, we depict the components A-F against $J_{OLS}^{M,N}(h, \hat{\theta}_{OLs}^{M,N}(h))$ for some data set when $\phi(x) = \phi_2(x)$ and the upwind or Lax-Wendroff schemes are used for computation. Here we observe why p_J/p differs for these schemes. For the upwind scheme, that $J_{OLS}^{M,N}(h, \hat{\theta}_{OLs}^{M,N}(h))$ appears to follow the order $2p$ term C for smaller values of h until $J_{OLS}^{M,N}(h, \hat{\theta}_{OLs}^{M,N}(h))$ approaches the value

Figure 4.4: Plots of $J_{OLS}^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ with components A-F when the upwind method is used for various data sets for $\phi(x) = \phi_2(x)$.

of A and then flattens. For larger values of h, the order p terms B and F are all significant, which likely causes $p_J \approx p$. For the Lax-Wendroff scheme, the order $2p$ term C tends to dominate the numerical cost function as h decreases until it flattens out at the A terms. Terms C and F appear insignificant to the cost function with this scheme, as they are often an order of magnitude below $J_{OLS}^{M,N}(h, \hat{\theta}_{OLs}^{M,N}(h))$, which causes $p_j \approx 2p$.

Surprisingly, all terms B, C, and F appear to converge at order $2p$ for $\phi(x) = \phi_1(x)$, which is why $p_J \approx 2p$. For the Beam-Warming method with $\phi(x) = \phi_2(x)$, we observe that the order p terms C and F are significant, which likely makes $p_J \approx p$ for this scheme.

4.4.3 Numerical Simulation Profiles

In Figure 4.6, we depict plots of $u(t, x | h, \hat{\theta}_{OLS}^{M,N}(h))$ against numerical data for the four schemes considered when $\phi(x) = \phi_2(x)$. These plots confirm the results discussed in Section 4.2.3 regarding modified equations: the first order schemes are diffusive, which prevents them from accurately matching the data, and the second order methods are dispersive. The Lax-Wendroff method is excessively dispersive, as it displays many small oscillations but still fits the general trend of the data. These spurious oscillations likely explain the somewhat unpredictable $\hat{\theta}_{OLS}^{M,N}(h)$ estimates for this method from Figure 8; these oscillations likely cause problems for the scheme to accurately match the data near θ_0, and the oscillations that occur at farther away points may happen to better
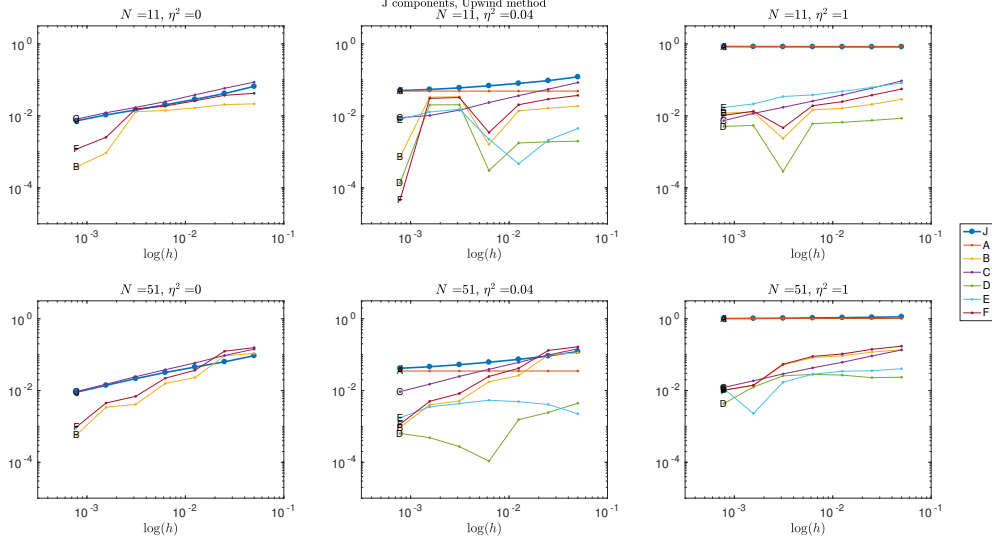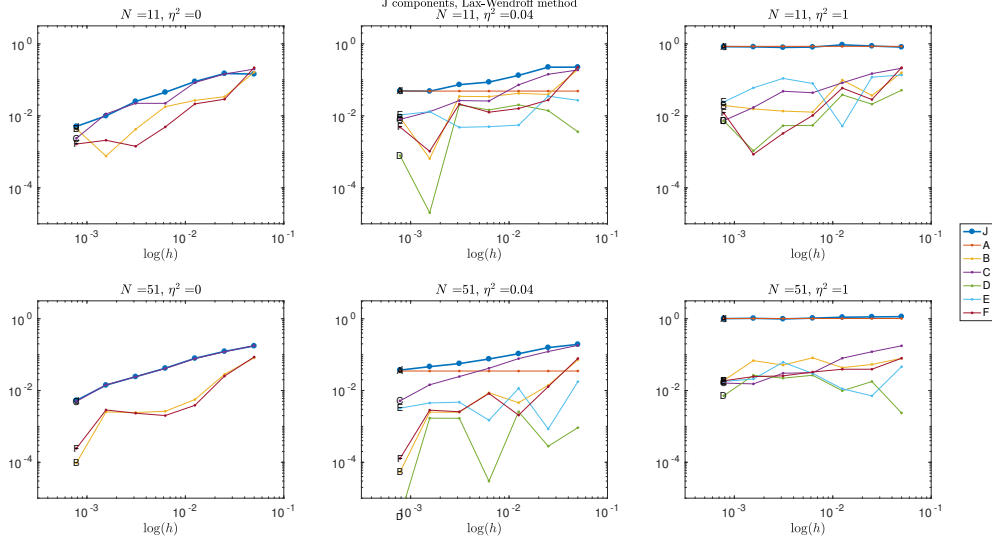
Figure 4.5: Plots of $J_{OLS}^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ with components A-F when the Lax-Wendroff method is used for various data sets for $\phi(x) = \phi_2(x)$.

match the data points. In general, as $h \to 0$, a subsequence of $\hat{\theta}_{OLS}^{M,N}(h)$ appears to converge to θ_0 with the Lax-Wendroff method. The Beam-Warming method gives the most accurate profile simulation of these four, but it still becomes negative just after the front. The upwind method with flux limiters provides the most realistic profile, as it maintains a sharp front, but also maintains a nonnegative profile.

4.5 Residual Analysis and Confidence Intervals

4.5.1 Residual Analysis

Residuals can provide insight into the underlying statistical nature of the data and model under consideration [Banks et al., 2012]. Here we attempt to use residual analysis to ascertain how numerical error propagates in the inverse problem. We define the residual at time t_i and spatial location x_j as

$$r_{i,j} = u(t_i, x_j | h, \hat{\theta}) - y_{ij}.$$

By minimizing the OLS cost function in Equation (4.6), we are implicitly assuming that each residual value is independent all other residuals. We depict the residuals for the upwind method, along with the model simulation, in Figure 4.7 for $h = 1/(10 \cdot 2^3)$ for $\eta = 0.1$. We see in this figure
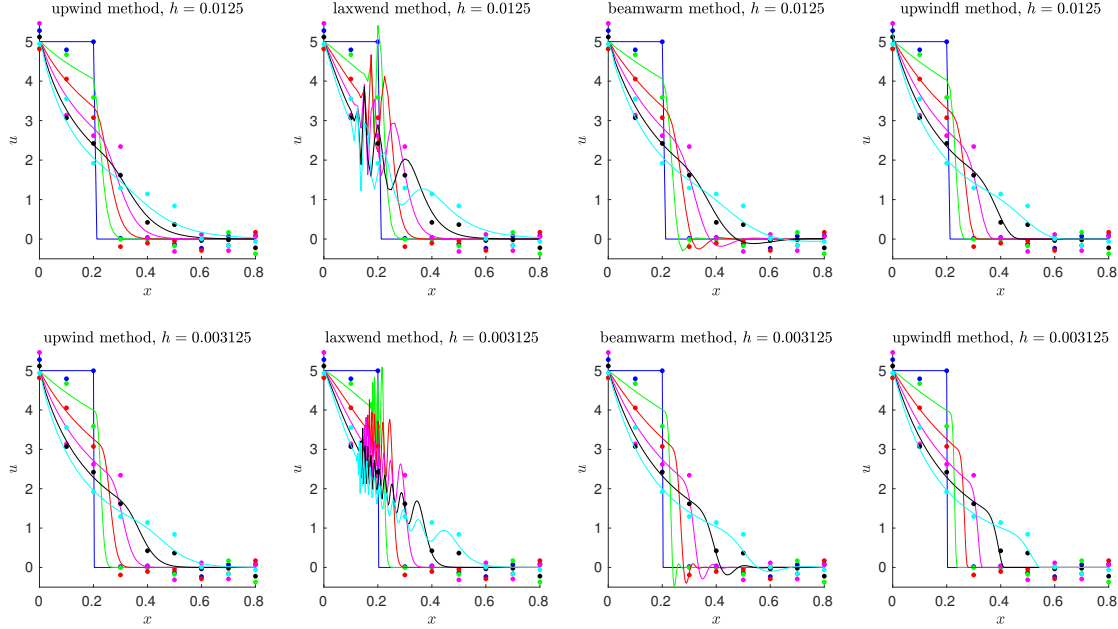
Figure 4.6: Numerical solution profiles (solid lines) plotted against artificial data (dots) for the five schemes considered when $\phi(x) = \phi_2(x)$ for two different step sizes.

that the residuals are clearly not independent of the neighboring points. The residuals are highest near the front and appear very correlated with the neighboring residual values. This type of error is expected when computing the discontinuous solution with a first order method, as we know from Section 4.2.3 that the main source of error will be diffusive. The correlation between neighboring data points indicates that an autocorrelative statistical model may be suitable in capturing the numerical error underlying the inverse problem when using a first order method to solve Equation (4.1).

To quantify the level of autocorrelated error that arises from numerical diffusion, we incorporate the first order autocorrelation assumption from [Seber and Wild, 1988, § 6], in which we assume that the residuals at the fixed time t_i satisfy

$$r_{i1} = \epsilon_1$$

$$r_{i2} = \gamma_i r_{i1} + \epsilon_2$$

$$r_{i3} = \gamma_i r_{i2} + \epsilon_3$$

$$\vdots$$

$$r_{iN} = \gamma_i r_{iN-1} + \epsilon_N \tag{4.9}$$

77
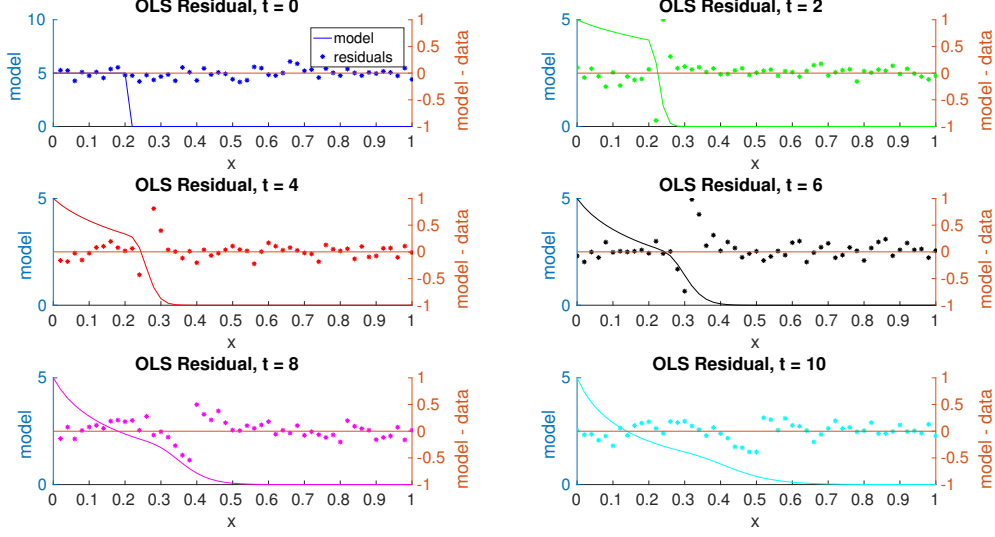
Figure 4.7: Plots of $r_{i,j}$ (dots) against simulations of $u(x,t|h,\hat{\theta})$ for the upwind method with $h = 0.00625$ with $\eta = 0.1$.

for $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \eta^2)$. The correlation between two spatial points may thus be written as

$$\text{corr}(r_{ij}, r_{ik}) = \gamma_i^{|j-k|}$$

where γ_i is the autocorrelation constant at each time point, t_i. This assumption gives that

$$\text{Var}(\epsilon) = \eta^2 V,$$

where $V_{jk} = \gamma_i^{|j-k|}$. We can then compute the inverse of V as

$$V^{-1} = \frac{1}{1 - \gamma_i^2} \begin{bmatrix} 1 & -\gamma_i & 0 & \dots & 0 \\ -\gamma_i & 1 + \gamma_i^2 & -\gamma_i & & 0 \\ 0 & -\gamma_i & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 + \gamma_i^2 & -\gamma_i \\ 0 & 0 & \dots & -\gamma_i & 1 \end{bmatrix}$$

$$= R^T R$$

for

$$R = \frac{1}{\sqrt{1-\gamma_i^2}} \begin{bmatrix} \sqrt{1-\gamma_i^2} & 0 & 0 & \dots & 0 \\ -\gamma_i & 1 & 0 & & 0 \\ 0 & -\gamma_i & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & 0 & \dots & -\gamma_i & 1 \end{bmatrix}.$$

By combining Equations (4.5) and (4.9), we thus see that

$$R(y_{i,j} - u(t_i, x_j | h, \theta_0)) \overset{i.i.d.}{\sim} \eta \mathcal{N}(0, 1) \tag{4.10}$$

when $u(t, x | h, \theta)$ is used to approximate $u_0(t, x | \theta)$ and adds numerical diffusion error in the process.

Note that we perform this autocorrelation statistical model for every time point t_i considered in the inverse problem. From Figure 4.7, we also notice that the residuals at a fixed time point appear to have two separate autocorrelated models occurring: one before the point of discontinuity and one occurring after the point of discontinuity. For this reason, we split the set of residuals at time t_i into two regions: residuals that occur before the point of discontinuity and points that occur after the point of discontinuity and then perform two separate autocorrelation models on both of them.

To estimate θ_0 and quantify numerical error with an autocorrelation model, we perform the following two-stage estimation routine for a data set with a given step size, h (taken from [Seber and Wild, 1988, § 6.2.3]):

1. Fit the model by finding the estimator, $\hat{\theta}_{OLS}^{M,N}(h)$, that minimizes Equation (4.6).

2. Compute the corresponding OLS residuals, r_{ij}, and estimate γ_i with the formula[a]

$$\gamma_i = \frac{\sum_{j=1}^{n-1} r_{ij} r_{ij+1}}{\sum_{j=2}^{n-1} r_{ij}^2}.$$

3. Now fit the model by find the estimator, $\hat{\theta}_{auto}^{M,N}(h)$, that minimizes

$$J_{auto}^{M,N}(h, \theta) = \frac{1}{MN} \vec{r}^T V^{-1} \vec{r}.$$

---

[a] *Recall that in practice, we do this estimation for each i and for points before, and after the discontinuity, so a more realistic labeling may be γ_i^{before} and γ_i^{after}.*

We performed this autocorrelation optimization method for the upwind method and depict
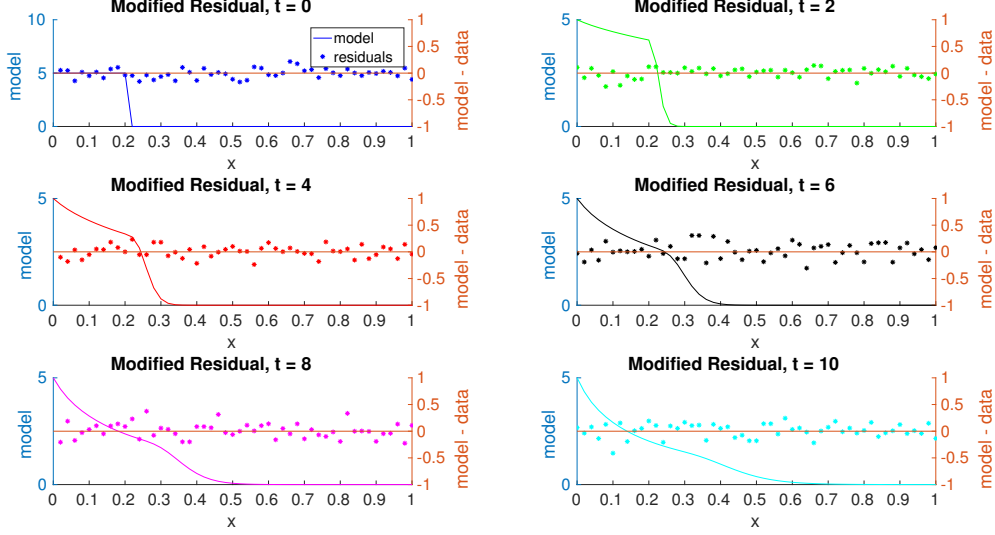
Figure 4.8: Plots of modified residuals (dots) for the upwind method with $h = 0.00625$ with $\eta = 0.1$.

the resulting autocorrelative residuals, $R\vec{r}$, in Figure 4.8. Here we see that the modified residuals appear more independent and identically distributed than the OLS residuals, suggesting that the autocorrelation method is capable of accurately quantifying how numerical error propagates in the inverse problem. We only showed the results for one data set here, but others show similar results.

The goal of using the autocorrelative statistical model is not only to improve residuals but to improve estimation of θ_0. In Figure 8 , we depict $\|\hat{\theta}^{M,N}_{auto}(h) - \theta_0\|_2$ for all data sets considered using the upwind scheme. Here we see that $\hat{\theta}^{M,N}_{auto}(h)$ is significantly improved over $\hat{\theta}^{M,N}_{OLS}(h)$ for $\eta^2 \leq 0.0225$ and larger values of h. The improvement is marginal (and sometimes worse) for larger error values with the exception of $\eta^2 = 0.25$ when $N = 30$. This observation suggests that when numerical error overrides experimental error when $\eta^2 \leq 0.0225$, so that the autocorrelation scheme can facilitate estimation of θ_0. For larger values of η, it appears that experimental error dominates numerical error, so estimation cannot be improved with the autocorrelation estimation routine.

4.5.2 Confidence Interval Computation

If for some matrix, Q, we assume that the residuals satisfy $(Q\vec{r})_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \eta^2)$, and let the parameter estimator $\hat{\theta}^{M,N}(h)$ be defined as

$$\hat{\theta}^{M,N}(h) = \arg\min_\theta \vec{r}^T Q^T Q \vec{r},$$

then asymptotically as $M, N \to \infty$,

$$\hat{\theta}^{M,N}(h) \sim \mathcal{N}_{k_\theta}(\theta_0, H_0^{M,N}) \approx \mathcal{N}_{k_\theta}\left(\theta_0, \eta^2\left[(Q\nabla\vec{u}_0(\theta_0))^T (Q\nabla\vec{u}_0(\theta_0))\right]^{-1}\right),$$

where k_θ denotes the number of parameters being estimated. See [Banks and Tran, 2009, Banks et al., 2014, Burnham et al., 2002] and [Seber and Wild, 1988] for more details. Observe that $Q = I$ when minimizing the OLS cost function and $Q = R$ when minimizing the autocorrelation cost function described in the previous section. From this, we can show that the $(1 - a)100\%$ confidence interval for the kth component of θ_0 is given by the interval

$$\hat{\theta}_k^{M,N} \pm SE_k(\hat{\theta}^{M,N})t_{1-1-a/2}^{MN-k_\theta} \tag{4.11}$$

where

$$SE_k(\hat{\theta}) = \sqrt{\hat{\eta}^2 \hat{H}_{kk}(\hat{\theta})},$$

$$\hat{H}(\hat{\theta}) = \left[(Q\nabla\vec{u}_0(\theta_0))^T (Q\nabla\vec{u}_0(\theta_0))\right]^{-1}, \hat{\eta}^2 = \frac{1}{MN - k_\theta}r^T Q^T Q r$$

and $t_{1-a/2}^n$ is the value such that $P(T \geq t_{1-a/2}^n) = a/2$ for T sampled from the student's t distribution with n degrees of freedom.

In Figure 4.9, we depict 95% OLS confidence intervals when $\phi(x) = \phi_2(x)$ for some data sets. We observe that with more data points, the confidence interval width gets smaller and we have a harder time enclosing θ_0. We depict the 95% OLS confidence intervals for the Beam-Warming method in Figure 4.10, and observe that the confidence intervals can capture θ_0 well for $N = 11$, but do struggle for $N = 30$. Note that these confidence intervals are close to θ_0, yet their widths are very small. All schemes appear to get very close to θ_0 as $h \to 0$, but they do struggle to capture it in the confidence intervals for all data sets. This is especially true for $N = 51$.

The flux limiters can improve the upwind scheme in almost all data sets where $N = 10$. This scheme can even outperform the second order Beam-Warming method for larger values of h.

Figure 4.11 depicts the calculated 95% confidence intervals for $\theta = (\alpha, \beta)^T$ when $\phi(x) = \phi_2(x)$ when using the autocorrelative statistical model from Section 4.5.1 with the upwind method. Almost all computed intervals for contain θ_0 regardless of h, whereas the upwind method struggled to have
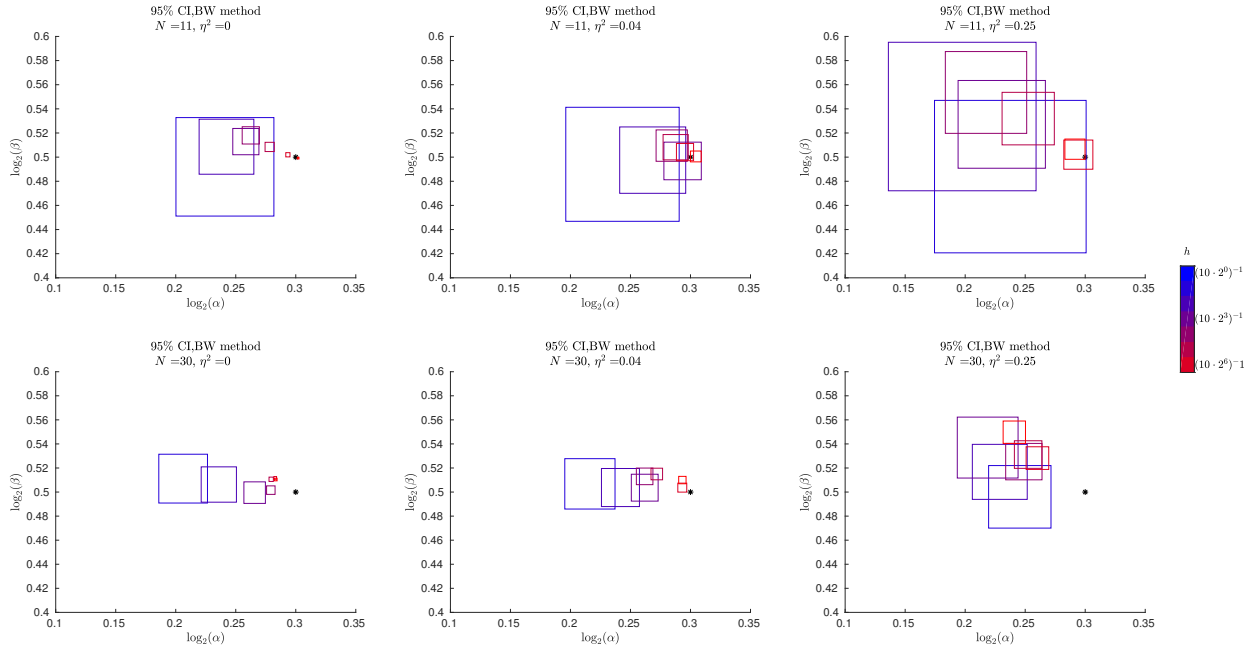
Figure 4.9: 95% autocorrelative confidence intervals for $\theta = (\alpha, \beta)^T$ using Equation (4.11) with an upwind scheme and $\phi(x)$ is a discontinuous front. The asterisk denotes θ_0.
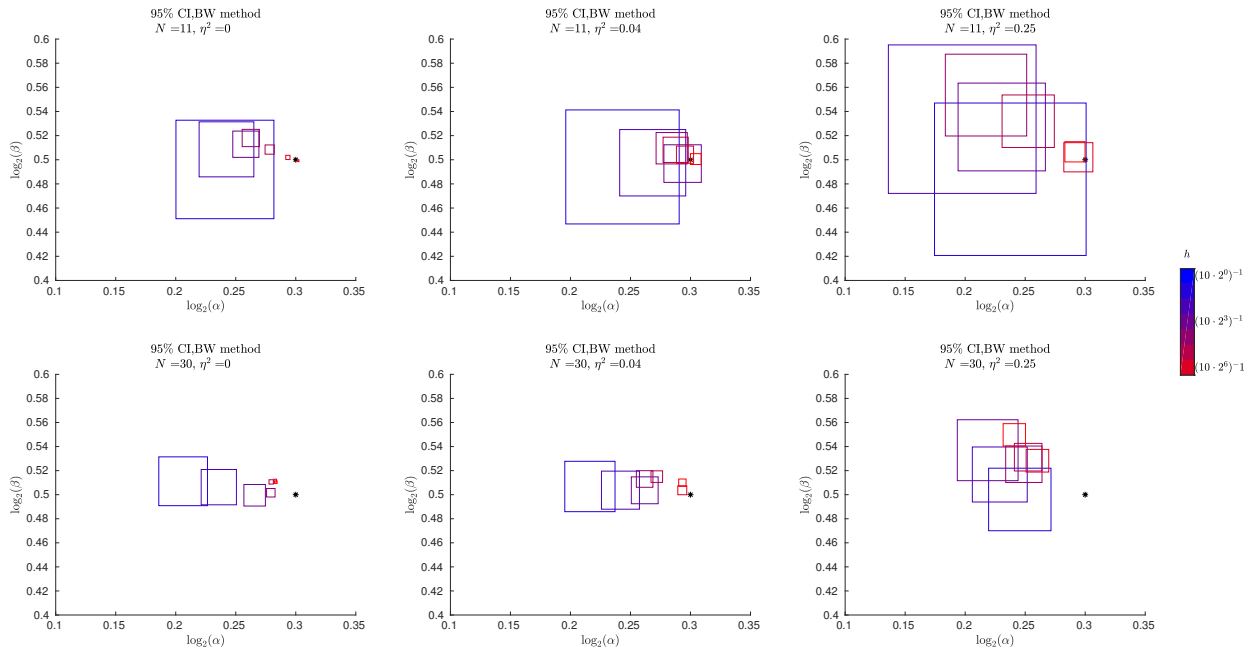
Figure 4.10: 95% confidence intervals for $\theta = (\alpha, \beta)^T$ using Equation (4.11) with a Beam-Warming scheme and $\phi(x)$ is a discontinuous front. The asterisk denotes θ_0.
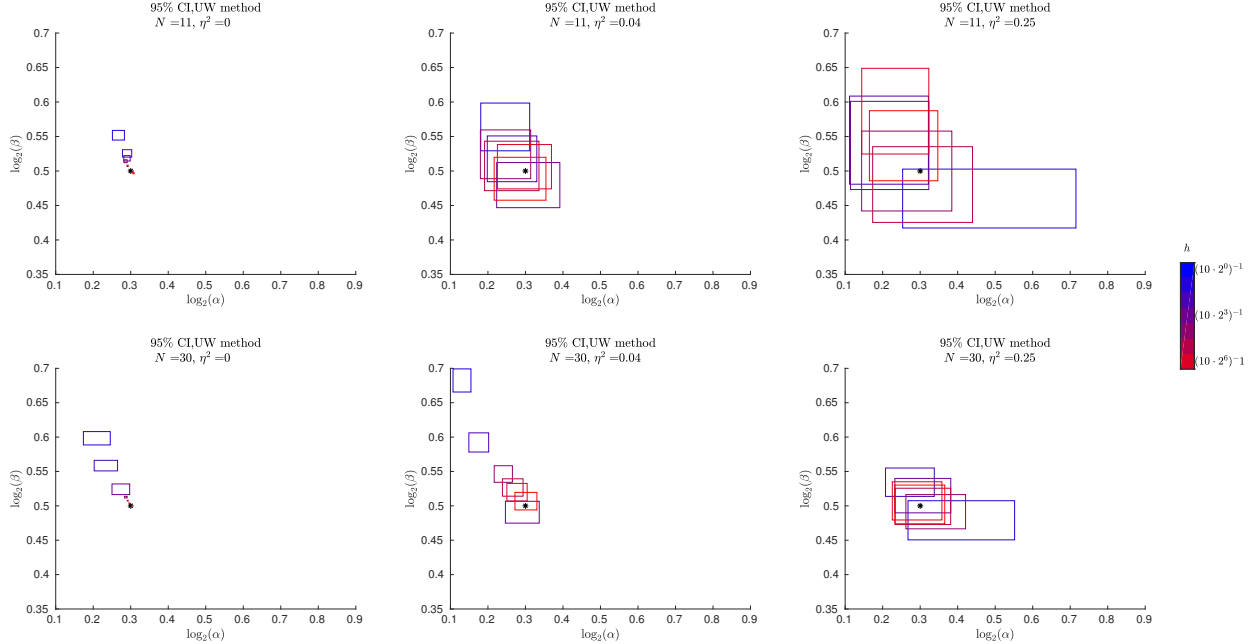
Figure 4.11: 95% autocorrelative confidence intervals for $\theta = (\alpha, \beta)^T$ using Equation (4.11) with an upwind scheme and $\phi(x)$ is a discontinuous front. The asterisk denotes θ_0.

the coarser grids contain θ_0 using the OLS cost function. In general, the method can significantly improve confidence interval computation for the upwind scheme. The method continues to struggle for $N = 51$, however.

4.6   Suggestions for practitioners

Based on the previous results, we suggest some strategies for practitioners to improve their inverse problem methodogies.

If a practitioner is concerned with minimizing $J_{OLs}^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ (and hence inferring the level of experimental error i in their calculation), he or she can ensure they have reached the minimum value by performing the inverse problem discussed here for multiple values of grid size. If the computed cost function appears to still be converging for smaller choice of h, then they can believe that numerical error is the dominant term in their computation. The practicioner can do better in finding the minimum value of $J_{OLs}^{M,N}(h, \hat{\theta}_{OLS}^{M,N}(h))$ by continuing the process with finer values of h, if possible. If the order of the numerical cost function appears to be zero (which can be confirmed by finding the best-fit line to $\ln\left(J_{OLs}^{M,N}(\vec{h}, \hat{\theta}_{OLS}^{M,N}(\vec{h}))\right)$ against $\ln(\vec{h})$), then the practitioner can be

confident that experimental error is the dominant term of the cost function. They thus have likely found a good estimate for η^2, assuming M, N are large.

There are a variety of different numerical schemes that a practitioner can use for their simulations. As we observed in Figure 4.3, these different numerical schemes can lead to better or worse parameter estimate values, which are important for properly understanding the system under consideration. To ensure the best possible parameter estimation, the practitioner should generate an artificial data set with the same number of data points and the estimate of the minimum value of the numerical cost function. With this data set, then use all numerical schemes being considered to determine which method can best estimate parameters. This one should be used to fit experimental data from the system. If an analytical solution is not available for data generation, then a very fine grid could be used to generate a simulation close to the true solution.

Lastly, we saw with the autocorrelation method with the upwind scheme could correct residual values and in turn, parameter estimation and confidence interval computation. Practitioners should actively plot their residuals for any inverse problem implementation and make any effort possible to modify their statistical model accordingly. Different types of statistical models are discussed in length in [Seber and Wild, 1988].

4.7   Discussion and Future Work

Numerical approximations for advection-dominated processes are a known challenge in the sciences [Leonard, 1991, Thackham et al., 2008a], and the effects of numerical error on an inverse problem have not been investigated thoroughly. In this document, we fit various numerical schemes with different precision numerical grids to artificial data with varying numbers of data points and error levels. We use a numerical cost function in a similar vein to that in [Banks and Fitzpatrick, 1990] to show that this cost function will converge at twice the order of the numerical scheme when $u_0(t, x)$ is continuous and at the same order of the numerical scheme when $u_0(t, x)$ is discontinuous for smaller noise levels. In general, the second order methods outperform the first order methods, as expected, but the first order methods can be significantly improved with the introduction of flux limiters. We also demonstrate an autocorrelative statistical model than be used to quantify the amount of numerical error introduced by first order methods and improve parameter estimation
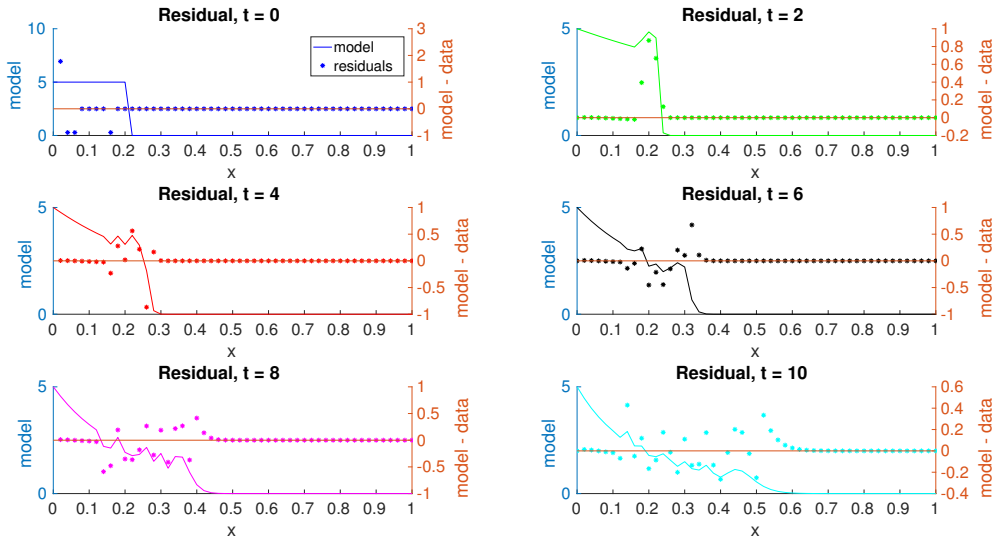
Figure 4.12: Plots of $r_{i,j}$ (dots) against simulations of $u(x, t | h, \hat{\theta})$ for the Lax-Wendroff method with $h = 0.0015625$ with $\eta = 0$.

results. Confidence interval computations suggest that one should be mindful of the scheme and precision used when solving an inverse problem, as different numerical methods perform differently for different noise levels and numbers of data points.

There are some aspects of this study that we have left for future work. In Figure 4.12, we depict the OLS residuals when fitting the Lax-Wendroff method to the artificial data when $\phi(x) = \phi_2(x)$. Recall that the modified equation for this second order method is dispersive, so the leading error terms are composed of high-frequency modes from the initial condition propagating at different speeds. This set of residuals shows patterns that would be much more difficult to quantify than those presented in Section 4.5.1 and warrants its own future work for a careful analysis into how their numerical error is influencing parameter estimation here.

Chapter 5

Conclusions

5.1 the Importance in Understanding Wound Healing Experiments

A thorough understanding of the physical and chemical cues underlying the collective migration of cells during wound healing experiments promises to inform how we treat several harmful diseases. Areas of particular relevance are the clinical treatment of nonhealing wounds and our understanding of cancer progression. In both of these processes, changes to the phenotype of healthy epithelial cells contributes to disease persistence and progression.

Chronic wounds are a significant burden to the US health care system, with wound care comprising an estimated 4% of total healthcare costs, and an estimated 6.5 million individuals with chronic wounds in the US in 2009 de Smet et al. [2017]. As chronic wounds can be caused by diabetes, the increasing prevalence of diabetes makes these number likely to rise [Thackham et al., 2008b]. There are several hallmarks of chronic wounds, such as hyperproliferation and the presence of nuclei in the upper epidermal layers of skin [Brem et al., 2007, Usui et al., 2008]. In healthy tissue, keratinocyte proliferation stops and nuclei disappear before reaching these upper layers. These alterations to the behavior of skin cells in the epidermis cause the re-epithelialization process of wound healing to fail, which leaves patients prone to harmful infections. There are several proposed treatments for helping heal chronic wounds, including topical treatment of HB-EGF [Tsang et al., 2003], providing supplemental oxygen to the wound area de Smet et al. [2017], and non-contact low frequency ultrasound therapy [Wiegand et al., 2017]. The failure of these methods to reliably cause chronic wounds to heal indicates we have a deficient understanding of the wound healing process and the

effects of these different therapies on cell behavior.

Cancer is most harmful when tumors are able to spread to new areas of the body [Kim and Choi, 2010]. This harmful spread of cells can result from different physical and chemical cues to either healthy or cancerous cells. For example, carcinomas gain their invasive phenotype through the epithelial-to-mesenchymal transition (EMT), in which healthy epithelial cells become cancerous, lose cell-to-cell contacts to their neighbors, and then gain mesenchymal phenotypes and invade new areas of the body [Friedl and Gilmour, 2009]. This spread of cells is the main cause of death in many different cancers, including melanoma and breast cancer [Garbe et al., 2011, Hollier et al., 2009]. Signaling through the EGF-receptor and the MAPK signaling cascade are also fundamental in stimulating EMT to proceed [Chapnick et al., 2011], and sustained activation of the MAPK signaling cascade likely causes the highly invasive nature of cancer cells [Kim and Choi, 2010]. Paracrine signaling is also fundamental to cancer biology, as diffusible cytokines and growth factors have been shown experimentally to give melanoma cells their invasive capability in skin Flach et al. [2011]. Understanding these physical and chemical cues that underly cancer progression will thus aid in how we can prevent its harmful spread.

Studying the chemical and physical cues underlying cell migration during healthy wound healing will provide insight into how abuse or disruption of these systems gives rise to cancer progression and chronic wounds. Mathematical models have proven a fundamental tool in recent years for studying cancer progression and how to enhance wound healing [Anderson, 2005, Arciero et al., 2013, Denman et al., 2006, Kuang et al., 2015, Rutter et al., 2017b].

5.2   Applications of this Dissertation

This dissertation focused on the derivation and analysis of mathematical models to better understand keratinocyte wound healing assays. Though we have focused on keratinocytes, the methods also apply to cells from other epithelial tissue. These improvements include models that incorporate the effects of cell-cell adhesions on the migration of cell populations over time, a framework to include the effects of biochemical patterns of activation on population phenotype, and a study into the effects of numerical error on an inverse problem methodology. These results all have the potential to aid our understanding of collective migration and suggest successful treatments for chronic wounds

and cancer progression.

In Chapter 2, we demonstrated the importance of cell-cell interactions on wound healing. If cell-cell adhesion levels are increased to promote cell migration in response to EGF treatment during healthy wound healing, then future chronic wound therapies may similarly target increases in cell-cell interactions to promote migration. The loss of cell-cell interactions has recently been implicated to cause hyperproliferation of keratinocytes, which is also a problem in chronic wounds [Bunker, 2017]. We thus suggest that targeting an increase in cell-cell interactions in chronic wounds may reduce hyperproliferation and stimulate population-wide migration to increase the success rate of re-epithelialization. But the amount of growth factor to supply to the chronic wounds is important, as too small of a concentration will have little effect on cell phenotype, whereas too much may cause excessive migration rates or other off-target effects. The introduction of biochemically-structured reaction-diffusion equations in Chapter 3 thus serves as a useful framework to quantify how different patterns of activation influence how a population migrates into the wound. For example, when cells are diffusing and proliferating in response to topical EGF treatment, Figure 3.9 suggests that cells migrate most at the entire activation threshold (given by Equation (3.33)). Such an observation could guide future treatment frequency schedules for the treatment of chronic wounds. Determining reliable $D(m)$, $\lambda(m)$, and $g(m)$ terms would need to be estimated from experiments before such a challenge is feasible, however. The results from Chapter 4 will ensure that we understand how numerical error in our computations influences the decisions we may make when attempting to estimate these terms.

The EMT process begins with the loss of cell-cell adhesions between epithelial cells in tissue [Friedl and Gilmour, 2009]. It is this loss of adhesions that gives these now mesenchymal cells their invasive phenotype, which suggests that Model H from Chapter 2 may be a reliable model for cancer progression. Notice in Figure 2.7 that Model H yields low cell densities very far away from the rest of the cell population (similar to the few cells that migrated very quickly into the wound experimentally in response to sh α-catenin in this same chapter). As excessive EGF signaling can cause the invasion of cancer cells, a biochemically-structured equation may be able to capture this phenomena. This equation could be used to study how a long pulse of MAPK activation allows a cell to detach from its tissue and invade a new area. Once the cell is in its new environment, it will need a new cue (or lack of regulation) to begin proliferating in an uncontrolled manner.

Understanding how biochemical patterns and the resulting rates of migration lead to this invasion of new areas may allow us to develop therapies to prevent the harmful spread of tumors.

The use of numerical mathematical models to better understand experiments requires understanding the effects of numerical error on model inference. This is why the results of Chapter 4 are important, as we developed a criteria to determine when numerical or experimental error are the dominant form of error in a numerically-computed cost function. This chapter will also provide researchers with the ability to understand how reliable their parameter estimate values are based off finite difference methods with different-sized grids. The autocorrelation scheme used to quantify numerical error also allows researchers a way to improve their parameter estimates and better quantify error in their computation if first order methods are needed for computation. Understanding the nature of the numerical cost function on coarse grids also provides researchers the ability to make some inference from more complicated models that cannot be simulated with very fine grids due to computational expenses.

There are still a variety of aspects of wound healing that are untouched by mathematical models. While a significant portion of this dissertation has focused on the response of a cell population to growth factor, this growth factor is supplied by another cell population *in vivo*. For example, macrophages often supply the EGF that promotes re-epithelialization [Pastar et al., 2014], and there is significant chemical interaction between keratinocytes and dermal fibroblasts during the wound healing process [Lu et al., 2013, Werner et al., 2007, Witte and Kao, 2005]. There is a significant literature on competition between species in the mathematical ecology literature [Shigesada and Kawasaki, 1997], but cooperation has received less focus. Future mathematical models should focus on the paracrine interactions between two or more cell populations to understand supply of chemical, as well as how it influences transitions from one stage of wound healing to another [Clark and Henson, 1995].

Part of the difficulty in causing chronic wounds to heal lies in their change in phenotype from healthy tissue. This suggests that mathematical models should also focus on the change in phenotype observed between healthy and diseased tissue if we are to understand how to treat chronic wounds. This must include the hyperproliferation, decreased migration, and decrease of bioavailability in response to growth factors that are observed in chronic wounds if we are use the resulting models to develop more effective treatments [Pastar et al., 2014].

Bibliography

Azmy Ackleh and Jeremy Thibodeaux. Parameter estimation in a structured erythropoiesis model. *Mathematical Biosciences and Engineering*, 5(4):601–616, October 2008. ISSN 1551-0018. doi: 10. 3934/mbe.2008.5.601. URL `http://www.aimsciences.org/journals/displayArticles.jsp?paperID=3701`.

Shangbing Ai. Traveling waves in the modelling of aerosolised skin grafts. *Physica D: Nonlinear Phenomena*, 237(21):2761–2766, November 2008. ISSN 0167-2789. doi: 10.1016/j.physd.2008.04. 006. URL `http://www.sciencedirect.com/science/article/pii/S0167278908001681`.

Shangbing Ai and Wenzhang Huang. Travelling waves for a reaction-diffusion system in population dynamics and epidemiology. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 135(4):663–675, August 2005. ISSN 1473-7124, 0308-2105. doi: 10.1017/S0308210500004054. URL `https://www.cambridge.org/core/journals/proceedings-of-the-royal-society-of-edinburgh-section-a-mathematics/article/travelling-waves-for-a-reactiondiffusion-system-in-population-dynamics-and-epidemiology/F382C5ECCAF88931E4A157A5CF807045`.

Shangbing Ai, Wenzhang Huant, and Zhi-an Wang. Reaction, Diffusion and chemotaxis in wave propagation. *Discrete and Continuous Dynamical System - B*, 20(1):1–21, January 2015. ISSN 1551-0018. doi: 10.3934/mbe.2015.12.717. URL `http://www.aimsciences.org/journals/displayArticlesnew.jsp?paperID=11013`.

A. R. A. Anderson. A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Mathematical Medicine and Biology*, 22(2):163–186, March 2005. ISSN 1477-8599,

1477-8602. doi: 10.1093/imammb/dqi005. URL `http://imammb.oupjournals.org/cgi/doi/10.1093/imammb/dqi005`.

K. Anguige and C. Schmeiser. A one-dimensional model of cell diffusion and aggregation, incorporating volume filling and cell-to-cell adhesion. *J. Math. Biol.*, 58(3):395, March 2009. ISSN 0303-6812, 1432-1416. doi: 10.1007/s00285-008-0197-8. URL `https://link.springer.com/article/10.1007/s00285-008-0197-8`.

Julia C. Arciero, Qi Mi, Maria F. Branca, David J. Hackam, and David Swigon. Continuum Model of Collective Cell Migration in Wound Healing and Colony Expansion. *Biophysical Journal*, 100 (3):535–543, February 2011. ISSN 00063495. doi: 10.1016/j.bpj.2010.11.083.

Julia C. Arciero, Qi Mi, Maria Branca, David Hackam, and David Swigon. Using a continuum model to predict closure time of gaps in intestinal epithelial cell layers: Gap closure time of epithelial cell layers. *Wound Repair and Regeneration*, 21(2):256–265, March 2013. ISSN 10671927. doi: 10.1111/j.1524-475X.2012.00865.x. URL `http://doi.wiley.com/10.1111/j.1524-475X.2012.00865.x`.

H. T. Banks and B. G. Fitzpatrick. Statistical methods for model comparison in parameter estimation problems for distributed systems. *J. Math. Biol.*, 28(5):501–527, September 1990. ISSN 0303-6812, 1432-1416. doi: 10.1007/BF00164161. URL `http://link.springer.com/article/10.1007/BF00164161`.

H. T. Banks, Zachary R. Kenz, and W. C. Thompson. An Extension of RSS-based Model Comparison Tests for Weighted Least Squares. Technical report, August 2012.

H. Thomas Banks and Hien T. Tran. *Mathematical and Experimental Modeling of Physical and Biological Processes*. CRC Press, Boca Raton, FL, 2009.

H.T. Banks, W. Clayton Thompson, and Shuhua Hu. *Modeling and Inverse Problems in the Presence of Uncertainty*. CRC Press, Boca Raton, 2014.

Jorge Berlanga, Jorge Lodos, Osvaldo Reyes, Juan F. Infante, Enrique Caballero, and Pedro Lopez-Saura. Epidermal Growth Factor Stimulated Re-epithelialization in Pigs: the Possible Role of Acute-Wound Proteases. *Biotecnol. Apl.*, 15:83–87, 1998.

Harold Brem, Olivera Stojadinovic, Robert F Diegelmann, Hyacinth Entero, Brian Lee, Irena Pastar, Michael Golinko, Harvey Rosenberg, and Marjana Tomic-Canic. Molecular Markers in Patients with Chronic Wounds to Guide Surgical Debridement. *Mol Med*, 13(1-2):30–39, 2007. ISSN 1076-1551. doi: 10.2119/2006-00054.Brem. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1869625/`.

Eric Bunker. Novel Insights into Tace Regulation of Autocrine and Paracrine Signaling in Epithelial Cells. *Chemistry & Biochemistry Graduate Theses & Dissertations*, January 2017. URL `https://scholar.colorado.edu/chem_gradetds/216`.

Kenneth P. Burnham, David Raymond Anderson, and Kenneth P. Burnham. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, New York, 2nd ed edition, 2002. ISBN 978-0-387-95364-9.

Anna Q. Cai, Kerry A. Landman, and Barry D. Hughes. Multi-scale modeling of a wound-healing cell migration assay. *Journal of Theoretical Biology*, 245(3):576–594, April 2007. ISSN 00225193. doi: 10.1016/j.jtbi.2006.10.024. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022519306004954`.

Douglas A. Chapnick and Xuedong Liu. Leader cell positioning drives wound-directed collective migration in TGF beta-stimulated epithelial sheets. *Mol. Biol. Cell*, 25(10):1586–1593, May 2014. ISSN 1059-1524, 1939-4586. doi: 10.1091/mbc.E14-01-0697. URL `http://www.molbiolcell.org/content/25/10/1586`.

Douglas A. Chapnick, Lisa Warner, Jennifer Bernet, Timsi Rao, and Xuedong Liu. Partners in crime: the TGF beta and MAPK pathways in cancer progression. *Cell & Bioscience*, 1:42, December 2011. ISSN 2045-3701. doi: 10.1186/2045-3701-1-42. URL `https://doi.org/10.1186/2045-3701-1-42`.

Douglas A. Chapnick, Jeremy Jacobsen, and Xuedong Liu. The Development of a Novel High Throughput Computational Tool for Studying Individual and Collective Cellular Migration. *PLOS ONE*, 8(12):e82444, December 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0082444. URL `http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0082444`.

Bevan L. Cheeseman, Donald F. Newgreen, and Kerry A. Landman. Spatial and temporal dynamics of cell generations within an invasion wave: A link to cell lineage tracing. *Journal of Theoretical Biology*, 363:344–356, December 2014. ISSN 0022-5193. doi: 10.1016/j.jtbi.2014.08.016. URL `http://www.sciencedirect.com/science/article/pii/S0022519314004706`.

R. A. F. Clark and P. Henson. *The Molecular and Cellular Biology of Wound Repair*. Plenum Press, New York, second edition, 1995.

Christopher W. Curtis and David M. Bortz. Propagation of fronts in the Fisher-Kolmogorov equation with spatially varying diffusion. *Physical Review E*, 86(6), December 2012. ISSN 1539-3755, 1550-2376. doi: 10.1103/PhysRevE.86.066108. URL `http://link.aps.org/doi/10.1103/PhysRevE.86.066108`.

Paul D. Dale, Philip K. Maini, and Jonathan A. Sherratt. Mathematical modeling of corneal epithelial wound healing. *Mathematical Biosciences*, 124(2):127–147, December 1994. ISSN 0025-5564. doi: 10.1016/0025-5564(94)90040-X. URL `http://www.sciencedirect.com/science/article/pii/002555649490040X`.

Andre M. de Roos. A gentle introduction to physiologically structured population models. In *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*, Population and Community Biology Series. 1996.

Gijs H.J. de Smet, Leonard F. Kroese, Anand G. Menon, Johannes Jeekel, Antoon W.J. van Pelt, Gert-Jan Kleinrensink, and Johan F. Lange. Oxygen therapies and their effects on wound healing. *Wound Rep and Reg*, pages n/a–n/a, 2017. ISSN 1524-475X. doi: 10.1111/wrr.12561. URL `http://onlinelibrary.wiley.com/doi/10.1111/wrr.12561/abstract`.

Paula K. Denman, D. L. Sean McElwain, and John Norbury. Analysis of Travelling Waves Associated with the Modelling of Aerosolised Skin Grafts. *Bull. Math. Biol.*, 69(2):495–523, June 2006. ISSN 0092-8240, 1522-9602. doi: 10.1007/s11538-006-9138-0. URL `http://link.springer.com/article/10.1007/s11538-006-9138-0`.

Pia Domschke, Dumitru Trucu, Alf Gerisch, and Mark A. J. Chaplain. Structured models of cell migration incorporating molecular binding processes. *J. Math. Biol.*, 75(6-7):1517–1561,

December 2017. ISSN 0303-6812, 1432-1416. doi: 10.1007/s00285-017-1120-y. URL `https://link-springer-com.colorado.idm.oclc.org/article/10.1007/s00285-017-1120-y`.

A. Ducrot and P. Magal. Travelling wave solutions for an infection-age structured model with diffusion. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 139(03):459–482, 2009. URL `http://journals.cambridge.org/abstract_S0308210507000455`.

Arnaud Ducrot. Travelling waves for a size and space structured model in population dynamics: Point to sustained oscillating solution connections. *Journal of Differential Equations*, 250(1): 410–449, January 2011. ISSN 0022-0396. doi: 10.1016/j.jde.2010.09.019. URL `http://www.sciencedirect.com/science/article/pii/S0022039610003566`.

Vincent Falanga, William H. Eaglstein, Brian Bucalo, Matthew H. Katz, Brian Harris, and Polly Carson. Topical Use of Human Recombinant Epidermal Growth Factor (h-EGF) in Venous Ulcers. *The Journal of Dermatologic Surgery and Oncology*, 18(7):604–606, July 1992. ISSN 1524-4725. doi: 10.1111/j.1524-4725.1992.tb03514.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1524-4725.1992.tb03514.x/abstract`.

Paul C. Fife. *Mathematical Aspects of Reacting and Diffusing Systems*. Number 28 in Lecture Notes in Biomathematics. Springer-Verlag, 1979.

Ronald A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7:353–369, 1937.

Edward H. Flach, Vito W. Rebecca, Meenhard Herlyn, Keiran S. M. Smalley, and Alexander R. A. Anderson. Fibroblasts Contribute to Melanoma Tumor Growth and Drug Resistance. *Mol. Pharmaceutics*, 8(6):2039–2049, December 2011. ISSN 1543-8384. doi: 10.1021/mp200421k. URL `https://doi.org/10.1021/mp200421k`.

Peter Friedl. Prespecification and plasticity: shifting mechanisms of cell migration. *Current Opinion in Cell Biology*, 16(1):14–23, February 2004. ISSN 0955-0674. doi: 10.1016/j.ceb.2003.11.001. URL `http://www.sciencedirect.com/science/article/pii/S0955067403001571`.

Peter Friedl and Darren Gilmour. Collective cell migration in morphogenesis, regeneration and

cancer. *Nat Rev Mol Cell Biol*, 10(7):445–457, July 2009. ISSN 1471-0072. doi: 10.1038/nrm2720. URL http://www.nature.com/nrm/journal/v10/n7/abs/nrm2720.html.

Claus Garbe, Thomas K. Eigentler, Ulrich Keilholz, Axel Hauschild, and John M. Kirkwood. Systematic Review of Medical Treatment in Melanoma: Current Status and Future Prospects. *The Oncologist*, 16(1):5–24, January 2011. ISSN 1083-7159, 1549-490X. doi: 10.1634/theoncologist. 2010-0190. URL http://theoncologist.alphamedpress.org/content/16/1/5.

Erika R. Geisbrecht and Denise J. Montell. Myosin VI is required for E-cadherin-mediated border cell migration. *Nat. Cell Biol.*, 4(8):616–20, August 2002. ISSN 1465-7392. doi: 10.1038/ncb830. URL http://www.ncbi.nlm.nih.gov/pubmed/12134162.

S. Gourley, R. Liu, and J. Wu. Some Vector Borne Diseases with Structured Host Populations: Extinction and Spatial Spread. *SIAM J. Appl. Math.*, 67(2):408–433, January 2007. ISSN 0036-1399. doi: 10.1137/050648717. URL http://epubs.siam.org/doi/abs/10.1137/050648717.

S. A. Gourley. Travelling front solutions of a nonlocal Fisher equation. *J Math Biol*, 41(3):272–284, September 2000. ISSN 0303-6812, 1432-1416. doi: 10.1007/s002850000047. URL http://link.springer.com/article/10.1007/s002850000047.

Xing Guo and Xiao-Fan Wang. Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Res*, 19(1):71–88, January 2009. ISSN 1001-0602. doi: 10.1038/cr.2008.302. URL http://www.nature.com/cr/journal/v19/n1/abs/cr2008302a.html.

Morton E. Gurtin and Richard C. Maccamy. Non-linear age-dependent population dynamics. *Arch. Rational Mech. Anal.*, 54(3):281–300, 1974. ISSN 0003-9527, 1432-0673. doi: 10.1007/BF00250793. URL http://link.springer.com/article/10.1007/BF00250793.

John D. Haley and William John Gullick, editors. *EGFR Signaling Networks in Cancer Therapy*. Humana Press, Totowa, NJ, 2008. ISBN 978-1-58829-948-2. doi: 10.1007/978-1-59745-356-1. URL http://www.springerlink.com/index/10.1007/978-1-59745-356-1.

Jason F. Hammond and David M. Bortz. Analytical solutions to Fisher's equation with time-variable coefficients. *Applied Mathematics and Computation*, 218(6):2497–2508, November 2011. ISSN

00963003. doi: 10.1016/j.amc.2011.03.163. URL `http://linkinghub.elsevier.com/retrieve/pii/S0096300311010137`.

Parvathi Haridas, Catherine J. Penington, Jacqui A. McGovern, D. L. Sean McElwain, and Matthew J. Simpson. Quantifying rates of cell migration and cell proliferation in co-culture barrier assays reveals how skin and melanoma cells interact during melanoma spreading and invasion. *J. Theor. Biol.*, 423:13–25, June 2017. ISSN 1095-8541. doi: 10.1016/j.jtbi.2017.04.017.

Alan Hastings, Kim Cuddington, Kendi F. Davies, Christopher J. Dugaw, Sarah Elmendorf, Amy Freestone, Susan Harrison, Matthew Holland, John Lambrinos, Urmila Malvadkar, Brett A. Melbourne, Kara Moore, Caz Taylor, and Diane Thomson. The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, 8(1):91–101, January 2005. ISSN 1461-0248. doi: 10.1111/j.1461-0248.2004.00687.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2004.00687.x/abstract`.

R.B. Hazan, G.R. Phillips, and R.F. Qiao. Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. *J. Cell Biol.*, 148(4):779–790, 2000. URL `http://jcb.rupress.org/content/148/4/779.abstract`.

Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1 edition, 1996.

T. Hillen and K. J. Painter. A users guide to PDE models for chemotaxis. *Journal of Mathematical Biology*, 58(1-2):183–217, January 2009. ISSN 0303-6812, 1432-1416. doi: 10.1007/s00285-008-0201-3. URL `http://link.springer.com/10.1007/s00285-008-0201-3`.

Brett G. Hollier, Kurt Evans, and Sendurai A. Mani. The Epithelial-to-Mesenchymal Transition and Cancer Stem Cells: A Coalition Against Cancer Therapies. *J Mammary Gland Biol Neoplasia*, 14(1):29–43, March 2009. ISSN 1083-3021, 1573-7039. doi: 10.1007/s10911-009-9110-3. URL `https://link-springer-com.colorado.idm.oclc.org/article/10.1007/s10911-009-9110-3`.

D. Horstmann and A. Stevens. A Constructive Approach to Traveling Waves in Chemotaxis. *J Nonlinear Sci*, 14(1):1–25, February 2004. ISSN 0938-8974, 1432-1467. doi: 10.1007/s00332-003-0548-y. URL `http://link.springer.com/article/10.1007/s00332-003-0548-y`.

Cai Huang, Ken Jacobson, and Michael D. Schaller. MAP kinases and cell migration. *J. Cell Sci.*, 117(Pt 20):4619–28, September 2004. ISSN 0021-9533. doi: 10.1242/jcs.01481. URL `http://www.ncbi.nlm.nih.gov/pubmed/15371522`.

Olga Ilina and Peter Friedl. Mechanisms of collective cell migration at a glance. *J. Cell Sci.*, 122(Pt 18):3203–8, September 2009. ISSN 1477-9137. doi: 10.1242/jcs.036525. URL `http://www.ncbi.nlm.nih.gov/pubmed/19726629`.

Wang Jin, Esha T. Shah, Catherine J. Penington, Scott W. McCue, Lisa K. Chopin, and Matthew J. Simpson. Reproducibility of scratch assays is affected by the initial degree of confluence: Experiments, modelling and model selection. *Journal of Theoretical Biology*, 390:136–145, February 2016. ISSN 00225193. doi: 10.1016/j.jtbi.2015.10.040. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022519315005676`.

Stuart T. Johnston, Matthew J. Simpson, and Ruth E. Baker. Mean-field descriptions of collective migration with strong adhesion. *Phys. Rev. E*, 85(5):051922, May 2012. doi: 10.1103/PhysRevE.85.051922. URL `https://link.aps.org/doi/10.1103/PhysRevE.85.051922`.

Stuart T. Johnston, Matthew J. Simpson, and D. L. Sean McElwain. How much information can be obtained from tracking the position of the leading edge in a scratch assay? *Journal of The Royal Society Interface*, 11(97):20140325, August 2014. ISSN 1742-5689, 1742-5662. doi: 10.1098/rsif.2014.0325. URL `http://rsif.royalsocietypublishing.org/content/11/97/20140325`.

Stuart T. Johnston, Esha T. Shah, Lisa K. Chopin, D. L. Sean McElwain, and Matthew J. Simpson. Estimating cell diffusivity and cell proliferation rate by interpreting IncuCyte ZOOM assay data using the Fisher-Kolmogorov model. *BMC Systems Biology*, 9(38), 2015. ISSN 1752-0509. doi: 10.1186/s12918-015-0182-y. URL `http://dx.doi.org/10.1186/s12918-015-0182-y`.

Stuart T. Johnston, Joshua V. Ross, Benjamin J. Binder, D. L. Sean McElwain, Parvathi Haridas, and Matthew J. Simpson. Quantifying the effect of experimental design choices for in vitro scratch assays. *Journal of Theoretical Biology*, 400:19–31, July 2016. ISSN 0022-5193. doi: 10.1016/j.jtbi.2016.04.012. URL `http://www.sciencedirect.com/science/article/pii/S0022519316300406`.

R. L. Juliano. Signal transduction by cell adhesion receptors and the cytoskeleton: functions of integrins, cadherins, selectins, and immunoglobulin-superfamily members. *Annu. Rev. Pharmacol. Toxicol.*, pages 283–323, 2002. doi: 0362-1642/02/0210-0283. URL `http://www.annualreviews.org/doi/abs/10.1146/annurev.pharmtox.42.090401.151133`.

Calvin R. Justus, Nancy Leffler, Maria Ruiz-Echevarria, and Li V. Yang. In vitro Cell Migration and Invasion Assays. *J Vis Exp*, (88), June 2014. ISSN 1940-087X. doi: 10.3791/51046. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4186330/`.

James P. Keener and James Sneyd. *Mathematical Physiology: I Cellular Physiology*, volume 8/I of *Interdisciplinary Applied Mathematics*. Springer, second edition, 2009.

Evelyn F. Keller and Lee A. Segel. Traveling bands of chemotactic bacteria: a theoretical analysis. *Journal of Theoretical Biology*, 30(2):235–248, 1971. URL `http://www.sciencedirect.com/science/article/pii/0022519371900518`.

Eun Kyung Kim and Eui-Ju Choi. Pathological roles of MAPK signaling pathways in human diseases. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1802(4):396–405, April 2010. ISSN 0925-4439. doi: 10.1016/j.bbadis.2009.12.009. URL `http://www.sciencedirect.com/science/article/pii/S0925443910000153`.

A. Kolmogoroff, I. Petrovsky, and N. Piscounoff. Etude de l'equation de la diffusion avec croissance de la quantite de matiere et son application a un probleme biologique. *Moscow Univ. Bull. Math*, 1:1–25, 1937.

Yang Kuang, Erica M. Rutter, and Tracy L. Stepien. A data-motivated density-dependent diffusion model of in vitro glioblastoma growth. *Mathematical Biosciences and Engineering*, 12 (6):1157–1172, August 2015. ISSN 1551-0018. doi: 10.3934/mbe.2015.12.1157. URL `http://www.aimsciences.org/journals/displayArticlesnew.jsp?paperID=11558`.

Alexander Kurganov and Eitan Tadmor. New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection-Diffusion Equations. *J. Comput. Phys.*, 160(1):241–282, May 2000. ISSN 00219991. doi: 10.1006/jcph.2000.6459. URL `http://linkinghub.elsevier.com/retrieve/pii/S0021999100964593`.

K. A. Landman, M. J. Simpson, J. L. Slater, and D. F. Newgreen. Diffusive and Chemotactic Cellular Migration: Smooth and Discontinuous Traveling Wave Solutions. *SIAM Journal on Applied Mathematics*, 65(4):1420–1442, 2005. ISSN 0036-1399. URL http://www.jstor.org.colorado.idm.oclc.org/stable/4096190.

Kerry A. Landman, Anna Q. Cai, and Barry D. Hughes. Travelling Waves of Attached and Detached Cells in a Wound-Healing Cell Migration Assay. *Bull. Math. Biol.*, 69(7):2119–2138, June 2007. ISSN 1522-9602. doi: 10.1007/s11538-007-9206-0. URL http://link.springer.com/article/10.1007/s11538-007-9206-0.

Kerry A. Landman, Matthew J. Simpson, and Graeme J. Pettet. Tactically-driven nonmonotone travelling waves. *Physica D: Nonlinear Phenomena*, 237(5):678–691, May 2008. ISSN 0167-2789. doi: 10.1016/j.physd.2007.10.003. URL http://www.sciencedirect.com/science/article/pii/S0167278907003806.

B. P. Leonard. The ULTIMATE conservative difference scheme applied to unsteady one-dimensional advection. *Computer Methods in Applied Mechanics and Engineering*, 88(1):17–74, June 1991. ISSN 0045-7825. doi: 10.1016/0045-7825(91)90232-U. URL http://www.sciencedirect.com/science/article/pii/004578259190232U.

Randall J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2007. ISBN 978-0-89871-629-0 0-89871-629-2.

Tong Li and Zhi-An Wang. Asymptotic nonlinear stability of traveling waves to conservation laws arising from chemotaxis. *Journal of Differential Equations*, 250(3):1310–1333, February 2011. ISSN 00220396. doi: 10.1016/j.jde.2010.09.020. URL http://linkinghub.elsevier.com/retrieve/pii/S0022039610003578.

Dongyuan Lu, Xiaofeng Liu, Yuxin Gao, Bo Huo, Yingyong Kang, Juan Chen, Shujin Sun, Li Chen, Xiangdong Luo, and Mian Long. Asymmetric Migration of Human Keratinocytes under Mechanical Stretch and Cocultured Fibroblasts in a Wound Repair Model. *PLOS ONE*, 8(9):e74563, September 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0074563. URL http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0074563.

Philip K. Maini, D.L. Sean McElwain, and David I. Leavesley. Travelling Waves in a Wound Healing Assay. *Applied Mathematics Letters*, 17:575–580, 2004. doi: 10.1016/j.aml.2003.01.011.

Moshe Marikovsky, Peter Vogt, Elof Eriksson, Jeffrey S. Rubin, William G. Taylor, Joachim Sasse, and Michael Klagsbrun. Wound Fluid-Derived Heparin-Binding EGF-Like Growth Factor (HB-EGF) Is Synergistic with Insulin-Like Growth Factor-I For Balb/MK Keratinocyte Proliferation. *Journal of Investigative Dermatology*, 106(4):616–621, April 1996. ISSN 0022-202X. doi: 10.1111/1523-1747.ep12345413. URL http://www.sciencedirect.com/science/article/pii/S0022202X15425035.

Paul Martin. Wound Healing–Aiming for Perfect Skin Regeneration. *Science (80-.)*, 276(5309):75–81, April 1997. ISSN 00368075. doi: 10.1126/science.276.5309.75. URL http://www.sciencemag.org/content/276/5309/75.shorthttp://www.sciencemag.org/cgi/doi/10.1126/science.276.5309.75.

Yutaka Matsubayashi, Miki Ebisuya, Sakiko Honjoh, and Eisuke Nishida. ERK Activation Propagates in Epithelial Cell Sheets and Regulates Their Migration during Wound Healing. *Current Biology*, 14(8):731–735, April 2004. ISSN 0960-9822. doi: 10.1016/j.cub.2004.03.060. URL http://www.sciencedirect.com/science/article/pii/S0960982204002532.

J.A. McGovern, J.R. Heinemann, L.J. Burke, R. Dawson, T.J. Parker, Z. Upton, J.D. Hooper, and K.J. Manton. Stratum basale keratinocyte expression of the cell-surface glycoprotein CDCP1 during epidermogenesis and its role in keratinocyte migration. *British Journal of Dermatology*, 168(3):496–503, March 2013. ISSN 1365-2133. doi: 10.1111/bjd.12119. URL http://onlinelibrary.wiley.com/doi/10.1111/bjd.12119/abstract.

A. G. McKendrick. *Applications of mathematics to medical problems*. 1927.

R. McLennan, L. J. Schumacher, J. A. Morrison, J. M. Teddy, D. A. Ridenour, A. C. Box, C. L. Semerad, H. Li, W. McDowell, D. Kay, P. K. Maini, R. E. Baker, and P. M. Kulesa. Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development*, 142(11):2014–2025, June 2015. ISSN 0950-1991, 1477-9129. doi: 10.1242/dev.117507. URL http://dev.biologists.org/cgi/doi/10.1242/dev.117507.

Qi Mi, David Swigon, Beatrice Riviere, Selma Cetin, Yoram Vodovotz, and David J. Hackam. One-Dimensional Elastic Continuum Model of Enterocyte Layer Migration. *Biophys J*, 93(11): 3745–3752, December 2007. ISSN 0006-3495. doi: 10.1529/biophysj.107.112326.

Hanneke N. Monsuur, Mireille A. Boink, Ester M. Weijers, Sanne Roffel, Melanie Breetveld, Amit Gefen, Lenie J. van den Broek, and Susan Gibbs. Methods to study differences in cell mobility during skin wound healing in vitro. *Journal of Biomechanics*, 49(8):1381–1387, May 2016. ISSN 0021-9290. doi: 10.1016/j.jbiomech.2016.01.040. URL `http://www.sciencedirect.com/science/article/pii/S002192901630121X`.

James D. Murray. *Lectures on nonlinear-differential equation models in biology*. Oxford University Press, 1977.

James D. Murray. *Mathematical Biology I. An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer New York, New York, NY, 3rd edition, 2002. ISBN 978-0-387-95223-9. URL `http://link.springer.com/10.1007/b98868`.

Philip J. Murray, Jun-Won Kang, Gary R. Mirams, Sung-Young Shin, Helen M. Byrne, Philip K. Maini, and Kwang-Hyun Cho. Modelling Spatially Regulated beta-Catenin Dynamics and Invasion in intestinal Crypts. *Biophysical Journal*, 99(3):716–725, August 2010. ISSN 0006-3495. doi: 10.1016/j.bpj.2010.05.016. URL `http://www.sciencedirect.com/science/article/pii/S0006349510006168`.

Sridevi Nagaraja, Anders Wallqvist, Jaques Reifman, and Alexander Y. Mitrophanov. Computational approach to characterize causative factors and molecular indicators of chronic wound inflammation. *J. Immunol.*, 192(4):1824–34, February 2014. ISSN 1550-6606. doi: 10.4049/jimmunol.1302481. URL `http://www.ncbi.nlm.nih.gov/pubmed/24453259`.

Sridevi Nagaraja, Jaques Reifman, and Alexander Y. Mitrophanov. Computational Identification of Mechanistic Factors That Determine the Timing and Intensity of the Inflammatory Response. *PLoS Comput. Biol.*, 11(12):e1004460, December 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004460. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4669096&tool=pmcentrez&rendertype=abstract`.

John T. Nardini and D. M. Bortz. Investigation of a Structured Fisher's Equation with Applications in Biochemistry. *SIAM Journal on Applied Mathematics*, (accepted), 2018.

John T. Nardini, Douglas A. Chapnick, Xuedong Liu, and David M. Bortz. Modeling keratinocyte wound healing: cell-cell adhesions promote sustained migration. *Journal of Theoretical Biology*, 400:103–117, July 2016. doi: 10.1016/j.jtbi.2016.04.015.

D. Newgreen, G. Pettet, and K. Landman. Chemotactic Cellular Migration: Smooth and Discontinuous Travelling Wave Solutions. *SIAM J. Appl. Math.*, 63(5):1666–1681, January 2003. ISSN 0036-1399. doi: 10.1137/S0036139902404694. URL http://epubs.siam.org/doi/abs/10.1137/S0036139902404694.

Luke Olsen, J. A. Sherratt, P. K. Maini, and F. Arnold. A mathematical model for the capillary endothelial cell-extracellular matrix interactions in wound-healing angiogenesis. *Math. Med. Biol.*, pages 261–281, 1997. URL http://imammb.oxfordjournals.org/content/14/4/261.abstract.

Hans G. Othmer and Thomas Hillen. The Diffusion Limit of Transport Equations II: Chemotaxis Equations. *SIAM J. Appl. Math.*, 62(4):1222–1250, 2002.

Irena Pastar, Olivera Stojadinovic, Natalie C. Yin, Horacio Ramirez, Aron G. Nusbaum, Andrew Sawaya, Shailee B. Patel, Laiqua Khalid, Rivkah R. Isseroff, and Marjana Tomic-Canic. Epithelialization in Wound Healing: A Comprehensive Review. *Adv Wound Care (New Rochelle)*, 3(7): 445–464, July 2014. ISSN 2162-1918. doi: 10.1089/wound.2013.0473. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4086220/.

Mirna Perez-Moreno, Colin Jamora, and Elaine Fuchs. Sticky Business: Orchestrating Cellular Signals at Adherens Junctions. *Cell*, 112(4):535–548, February 2003. ISSN 0092-8674. doi: 10.1016/S0092-8674(03)00108-9. URL http://www.sciencedirect.com/science/article/pii/S0092867403001089.

G. J. Pettet, H. M. Byrne, D. L. S. Mcelwain, and J. Norbury. A model of wound-healing angiogenesis in soft tissue. *Mathematical Biosciences*, 136(1):35–63, August 1996. ISSN 0025-5564. doi: 10.1016/0025-5564(96)00044-2. URL http://www.sciencedirect.com/science/article/pii/0025556496000442.

Filippo Posta and Tom Chou. A mathematical model of intercellular signaling during epithelial wound healing. *Journal of Theoretical Biology*, 266(1):70–78, September 2010. ISSN 00225193. doi: 10.1016/j.jtbi.2010.05.029. URL `http://linkinghub.elsevier.com/retrieve/pii/S0022519310002729`.

M. Poujade, E. Grasland-Mongrain, A. Hertzog, J. Jouanneau, P. Chavrier, B. Ladoux, A. Buguin, and P. Silberzan. Collective migration of an epithelial monolayer in response to a model wound. *Proc. Natl. Acad. Sci.*, 104:15988–15993, 2007.

Randall J. Leveque. *Conservation Laws.* Lectures in Mathematics. Birkhauser Verlag, 2 edition, 1992.

Anne J. Ridley, Martin A. Schwartz, Keith Burridge, Richard A. Firtel, Mark H. Ginsberg, Gary Borisy, J. Thomas Parsons, and Alan Rick Horwitz. Cell Migration: Integrating Signals from Front to Back. *Science*, 302(5651):1704–1709, December 2003. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.1092053. URL `http://science.sciencemag.org/content/302/5651/1704`.

Erica M. Rutter, H. T. Banks, Gerald A. LeBlanc, and Kevin B. Flores. Continuous Structured Population Models for Daphnia magna. *Bull Math Biol*, pages 1–22, September 2017a. ISSN 0092-8240, 1522-9602. doi: 10.1007/s11538-017-0344-8. URL `https://link.springer.com/article/10.1007/s11538-017-0344-8`.

Erica M. Rutter, Tracy L. Stepien, Barrett J. Anderies, Jonathan D. Plasencia, Eric C. Woolf, Adrienne C. Scheck, Gregory H. Turner, Qingwei Liu, David Frakes, Vikram Kodibagkar, Yang Kuang, Mark C. Preul, and Eric J. Kostelich. Mathematical Analysis of Glioma Growth in a Murine Model. *Sci Rep*, 7, May 2017b. ISSN 2045-2322. doi: 10.1038/s41598-017-02462-0. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5451439/`.

G. A. F. Seber and C. J. Wild. *Nonlinear Regression.* Wiley series in probability and statistics. Wiley, 1988.

R. Seger and E. G. Krebs. The MAPK signaling cascade. *FASEB J*, 9(9):726–735, June 1995. ISSN 0892-6638, 1530-6860. URL `http://www.fasebj.org/content/9/9/726`.

F. R. Sharpe and A. J. Lotka. A problem in Age-Distribution. *Philosphical Magazine*, 21(124): 435–438, 1911.

Jonathan A. Sherratt and Mark A. J. Chaplain. A new mathematical model for avascular tumour growth. *J Math Biol*, 43(4):291–312, 2001. ISSN 0303-6812, 1432-1416. doi: 10.1007/s002850100088. URL http://link.springer.com.colorado.idm.oclc.org/article/10.1007/s002850100088.

Jonathan A. Sherratt and J. D. Murray. Models of epidermal wound healing. *Proc. R. Soc. Lond. B*, 241(1300):29–36, July 1990. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.1990.0061. URL http://rspb.royalsocietypublishing.org/content/241/1300/29.

Nanako Shigesada and Kohkichi Kawasaki. *Biological Invasions: Theory and Practice*. Oxford Series in Ecology and Evolution. Oxford University Press, 1997.

Matthew J. Simpson, Kerry A. Landman, and Donald F. Newgreen. Chemotactic and diffusive migration on a nonuniformly growing domain: numerical algorithm development and applications. *Journal of Computational and Applied Mathematics*, 192(2):282–300, August 2006. ISSN 0377-0427. doi: 10.1016/j.cam.2005.05.003. URL http://www.sciencedirect.com/science/article/pii/S0377042705003298.

Matthew J. Simpson, Ruth E. Baker, and Scott W. McCue. Models of collective cell spreading with variable cell aspect ratio: A motivation for degenerate diffusion models. *Physical Review E*, 83(2), February 2011. ISSN 1539-3755, 1550-2376. doi: 10.1103/PhysRevE.83.021901. URL http://link.aps.org/doi/10.1103/PhysRevE.83.021901.

Matthew J. Simpson, Parvathi Haridas, and D. L. Sean McElwain. Do Pioneer Cells Exist? *PLoS ONE*, 9(1):e85488, January 2014. ISSN 1932-6203. doi: 10.1371/journal.pone.0085488. URL http://dx.plos.org/10.1371/journal.pone.0085488.

Joseph W.-H. So, Jianhong Wu, and Xingfu Zou. A reaction-diffusion model for a single species with age structure. I Travelling wavefronts on unbounded domains. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 457(2012):1841–1853, July 2001. ISSN 1364-5021, 1471-2946. doi: 10.1098/rspa.2001.0789.

Christina H. Stuelten, Carole A. Parent, and Denise J. Montell. Cell motility in cancer invasion and metastasis: insights from simple model organisms. *Nature Reviews Cancer*, March 2018. ISSN 1474-1768. doi: 10.1038/nrc.2018.15. URL `https://www.nature.com/articles/nrc.2018.15`.

P. Sweby. High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM J. Numer. Anal.*, 21(5):995–1011, October 1984. ISSN 0036-1429. doi: 10.1137/0721062. URL `http://epubs.siam.org/doi/abs/10.1137/0721062`.

Jennifer A. Thackham, D. L. Sean McElwain, and Ian W. Turner. Computational Approaches to Solving Equations Arising from Wound Healing. *Bull. Math. Biol.*, 71(1):211–246, December 2008a. ISSN 0092-8240, 1522-9602. doi: 10.1007/s11538-008-9360-z. URL `http://link.springer.com/article/10.1007/s11538-008-9360-z`.

Jennifer A. Thackham, D.L. Sean McElwain, and Robert J. Long. The use of hyperbaric oxygen therapy to treat chronic wounds: A review. *Wound Repair and Regeneration*, 16(3):321–330, May 2008b. ISSN 1524-475X. doi: 10.1111/j.1524-475X.2008.00372.x. URL `http://onlinelibrary.wiley.com/doi/10.1111/j.1524-475X.2008.00372.x/abstract`.

Katrina K. Treloar, Matthew J. Simpson, D. L. Sean McElwain, and Ruth E. Baker. Are in vitro estimates of cell diffusivity and cell proliferation rate sensitive to assay geometry? *J. Theor. Biol.*, 356:71–84, September 2014. ISSN 1095-8541. doi: 10.1016/j.jtbi.2014.04.026. URL `http://www.ncbi.nlm.nih.gov/pubmed/24787651`.

M. W. Tsang, W. K. R. Wong, C. S. Hung, K.-M. Lai, W. Tang, E. Y.N. Cheung, G. Kam, L. Leung, C. W. Chan, C. M. Chu, and E. K.H. Lam. Human Epidermal Growth Factor Enhances Healing of Diabetic Foot Ulcers. *Diabetes Care*, 26(6):1856–1861, June 2003. ISSN 0149-5992, 1935-5548. doi: 10.2337/diacare.26.6.1856. URL `http://care.diabetesjournals.org/cgi/doi/10.2337/diacare.26.6.1856`.

Marcia L. Usui, Jonathan N. Mansbridge, William G. Carter, Mayumi Fujita, and John E. Olerud. Keratinocyte Migration, Proliferation, and Differentiation in Chronic Ulcers From Patients With Diabetes and Normal Wounds. *J Histochem Cytochem.*, 56(7):687–696, July 2008. ISSN 0022-1554. doi: 10.1369/jhc.2008.951194. URL `https://doi.org/10.1369/jhc.2008.951194`.

H. a. S. van den Brenk. Studies in restorative growth processes in mammalian wound healing. *Br J Surg*, 43(181):525–550, March 1956. ISSN 1365-2168. doi: 10.1002/bjs.18004318115. URL `http://onlinelibrary.wiley.com/doi/10.1002/bjs.18004318115/abstract`.

Aizik Volpert, Vitaly Volpert, and Vladimir Volpert. *Traveling Wave Solutions of Parabolic Systems*, volume 140 of *Translations of Mathematical Monographs*. American Mathematical Society, 1994.

G. F. Webb. Population models structured by age, size, and spatial position. In *Structured Population Models in Biology and Epidemiology*, pages 1–49. Springer, 2008.

Sabine Werner and Richard Grose. Regulation of Wound Healing by Growth Factors and Cytokines. *Physiological Reviews*, 83(3):835–870, July 2003. ISSN 0031-9333, 1522-1210. URL `http://physrev.physiology.org/content/83/3/835`.

Sabine Werner, Thomas Krieg, and Hans Smola. Keratinocyte-Fibroblast Interactions in Wound Healing. *Journal of Investigative Dermatology*, 127(5):998–1008, May 2007. ISSN 0022-202X. doi: 10.1038/sj.jid.5700786. URL `http://www.sciencedirect.com/science/article/pii/S0022202X15333820`.

Cornelia Wiegand, Kyle Bittenger, Robert D. Galiano, Vickie R. Driver, and Gary W. Gibbons. Does noncontact low-frequency ultrasound therapy contribute to wound healing at the molecular level? *Wound Rep and Reg*, 25(5):871–882, September 2017. ISSN 1524-475X. doi: 10.1111/wrr.12595. URL `http://onlinelibrary.wiley.com/doi/10.1111/wrr.12595/abstract`.

Richard P. Witte and Weiyuan John Kao. Keratinocyte-fibroblast paracrine interaction: the effects of substrate and culture condition. *Biomaterials*, 26(17):3673–3682, June 2005. ISSN 0142-9612. doi: 10.1016/j.biomaterials.2004.09.054. URL `http://www.sciencedirect.com/science/article/pii/S0142961204008816`.

Hongqi Xue, Hongyu Miao, and Hulin Wu. Sieve Estimation of Constant and Time-Varying Coefficients in Nonlinear Ordinary Differential Equation Models by Considering Both Numerical Error and Measurement Error. *Ann Stat*, 38(4):2351–2387, January 2010. ISSN 0090-5364. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2995285/`.

Boutheina Yahyaoui, Mekki Ayadi, and Abderrahmane Habbal. Fisher-KPP with time dependent diffusion is able to model cell-sheet activated and inhibited wound closure. *Mathematical Biosciences*, 292(October):36–45, 2017. ISSN 0025-5564. doi: 10.1016/j.mbs.2017.07.009. URL http://www.sciencedirect.com/science/article/pii/S002555641730069X.

Zhike Zi, Zipei Feng, Douglas A. Chapnick, Markus Dahl, Difan Deng, Edda Klipp, Aristidis Moustakas, and Xuedong Liu. Quantitative analysis of transient and sustained transforming growth factor-beta signaling dynamics. *Molecular Systems Biology*, 7(1):492, January 2011. ISSN 1744-4292, 1744-4292. doi: 10.1038/msb.2011.22. URL http://msb.embopress.org/content/7/1/492.

Cordelia Ziraldo, Alexey Solovyev, Ana Allegretti, Shilpa Krishnan, M. Kristi Henzel, Gwendolyn a Sowa, David Brienza, Gary An, Qi Mi, and Yoram Vodovotz. A Computational, Tissue-Realistic Model of Pressure Ulcer Formation in Individuals with Spinal Cord Injury. *PLoS Comput. Biol.*, 11(6):e1004309, June 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004309. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4482429&tool=pmcentrez&rendertype=abstract.

Appendix

A.1   Cell sheets do not proliferate while migrating

In Figure 1, we display our quantitative results of cellular proliferation during sheet migration. Sheets of HaCaT cells were labeled with Edu between t=24 and t=36 hrs post wound formation. Proliferating cells that incorporate EdU into their DNA were stained using the Click-IT EdU fluorescence labeling kit (ThermoFisher, C10337) and labeled and total cells were quantified by cell counting in selected regions using ImageJ. Although EGF is a potent stimulator of sheet migration between t=24 and 36 hrs, there is little difference between the quantified proliferation of Mock (A) and EGF (B) stimulated sheets, which display an average percent proliferation per hour below 1% of the population. We do note that there is a spatial distribution of proliferation in both cases, where a low level of proliferation is concentrated towards the region in close proximity to the wound.

A.2   Derivation of Model P

For our second model, we assume that cell-cell adhesions *promote* migration as leader cells pull the cells behind them forward and denote it as Model P. In this scenario, the upregulation of cell-cell adhesionpromotes migration. We now denote the forward transition rate τ_i^+ of cell density u_i as

$$\tau_{i,P}^+ = \frac{D(1 - u_{i+1})(1 + \hat{\alpha}u_{i+1})}{\Delta x^2},\tag{1}$$

where the first term again represent space filling, but *the second term in Equation (1) represents our second assumption in which cells located in the direction of migration (denoted u_{i+1}) promote this space filling movement towards the wound by pulling on the cells behind them.* We note again that
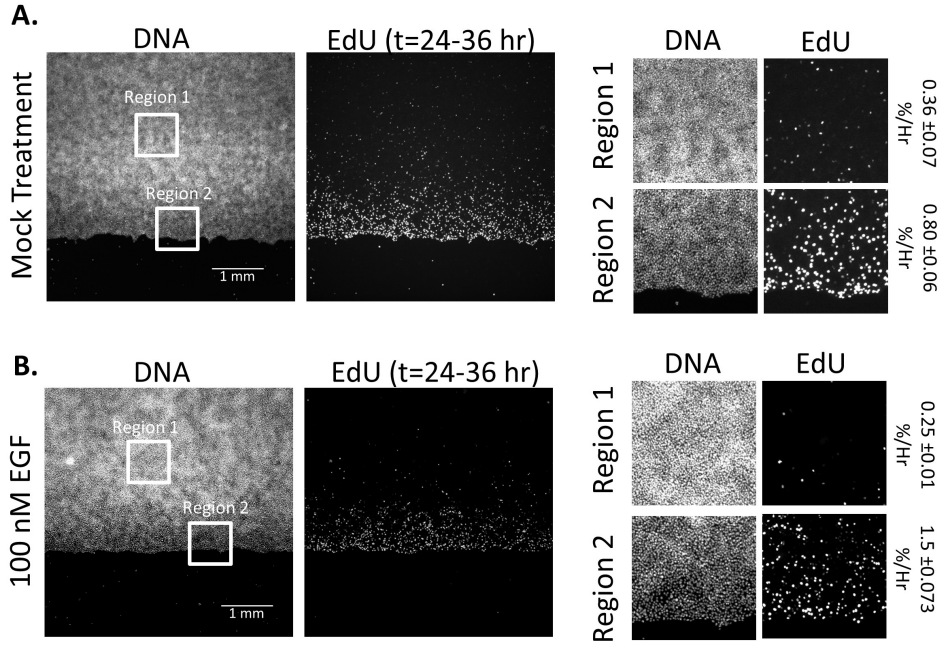
Figure 1: Quantitative results of cellular proliferation during sheet migration.

$\tau_{i-1,P}^{+}, \tau_{i,P}^{-}$, and $\tau_{i+1,P}^{-}$ are all defined analogously in Table 1. Substituting the transition probability from Equation (1) into Equation (2.3) yields

$$
\begin{aligned}
u_t &= \frac{1}{\Delta x^2}\left(D(1-u_i)(1+\hat{\alpha}u_i)u_{i-1} + D(1-u_i)(1+\hat{\alpha}u_i)u_{i+1} - D((1-u_{i+1})(1+\hat{\alpha}u_{i+1})+ \right. \\
&\quad \left. D(1-u_{i-1})(1+\hat{\alpha}u_{i-1}))u_i\right) \hspace{4cm} (2)\\
&= \frac{D}{\Delta x^2}\left(u_{i-1} + \hat{\alpha}u_{i-1}^2 u_i - \hat{\alpha}u_{i-1}u_i^2 + u_{i+1} - 2u_i - \hat{\alpha}u_i^2 u_{i+1} + \hat{\alpha}u_i u_{i+1}^2\right) \\
&= D\frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} + \alpha\frac{u_{i-1}^2 u_i - u_{i-1}u_i^2 - u_i^2 u_{i+1} + u_i u_{i+1}^2}{\Delta x^2}, \hspace{1.5cm} (3)
\end{aligned}
$$

where we can recognize the first term on the right hand side as the standard central difference approximation to the second derivative and we have set $\alpha = D\hat{\alpha}$. For the second term on the right hand side, we can simplify using Taylor series approximations as: $u_{i+1} \approx u_i + \Delta x \cdot u_i' + \Delta x^2/2 \cdot u_i''$, $u_{i-1} \approx u_i - \Delta x \cdot u_i' + \Delta x^2/2 \cdot u_i''$, where primes denote a spatial derivative term. Substituting

109

these terms into the second term on the right hand side of Equation (3) reveals

$$
\begin{aligned}
u_t &= \mathrm{D}\frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} + \alpha\frac{2\Delta x^2 u_i(u_i')^2 + u_i^2 u_i''\Delta x^2 + \mathcal{O}(\Delta x^4)}{\Delta x^2} \\
&= Du_i'' + \alpha(2u_i(u_i')^2 + u_i^2 u_i'') + \mathcal{O}(\Delta x^2) \\
&= Du_i'' + \alpha(u_i^2 u_i')' + \mathcal{O}(\Delta x^2)
\end{aligned}
$$

resulting in the dimensionless continuum limit to be

$$
u_t = ((D + \gamma u^2)u_x)_x, \quad \text{(Model P)} \tag{4}
$$

where D and γ again denote the rates of cell diffusion and cell-cell adhesion, respectively.

A.3 Fitting and Predicting individual data sets

In Figures 2 and 3, we depict the individual calculated experimental leading edges $\ell_{data}^{0.1}(t)$ and $\ell_{data}^{0.3}(t)$) for EGF experiments in the top left frames, along with the ability of both models to fit and predict this data using the parameter estimation method outlined in Section 2.2.2. In the top right of these figures, we demonstrate best-fit simulations of both models to one data set and in the the bottom rows, we demonstrate how this simulation can be used to accurately predict the leading edge locations of the other data sets. While we have only shown fits to one data set for both mock and EGF data, we see similar results after fitting to all three EGF data sets.

A.4 Numerical implementation

Note that for numerical implementation, we use a method of lines approach implemented with MATLAB's ode15s command for time integration. For spatial discretization, we use the second order scheme for convection-diffusion equations (without convection in our case) from Kurganov and Tadmor [2000] given by

$$
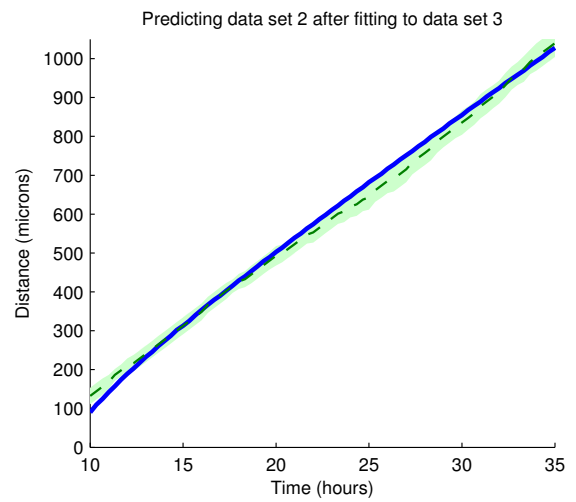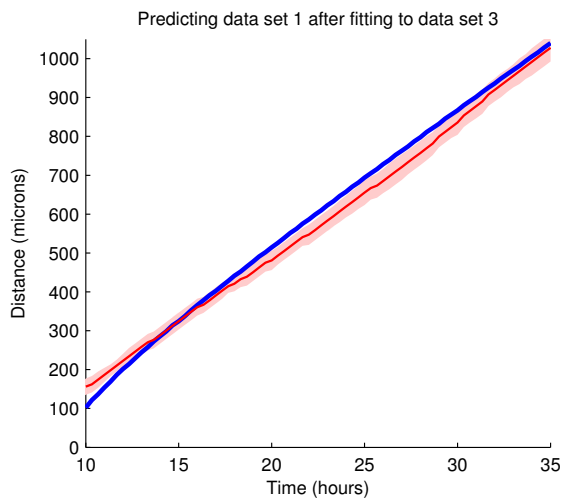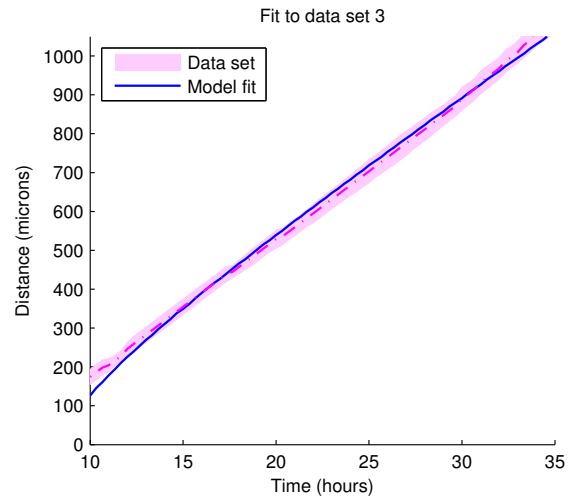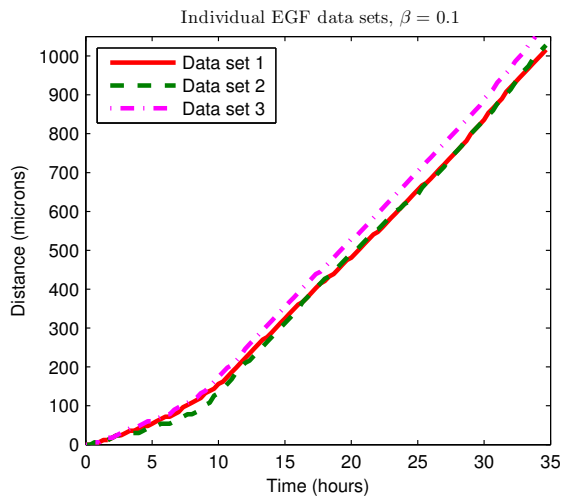\dot{u}_j(t) = \frac{P_{j+1/2}(t) - P_{j-1/2}(t)}{\Delta x}
$$

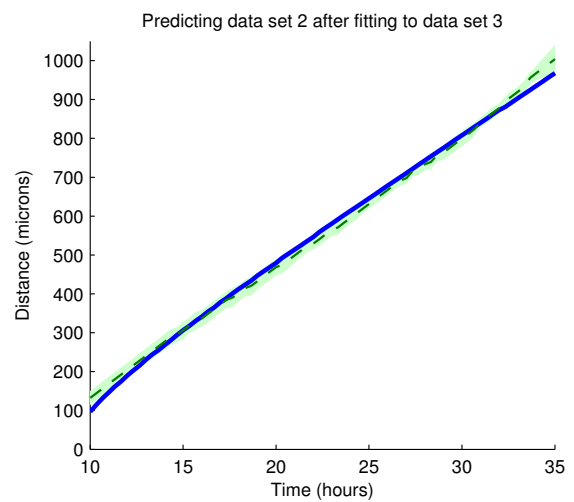Figure 2: Fits and predictions for a leading edge density of 0.1.
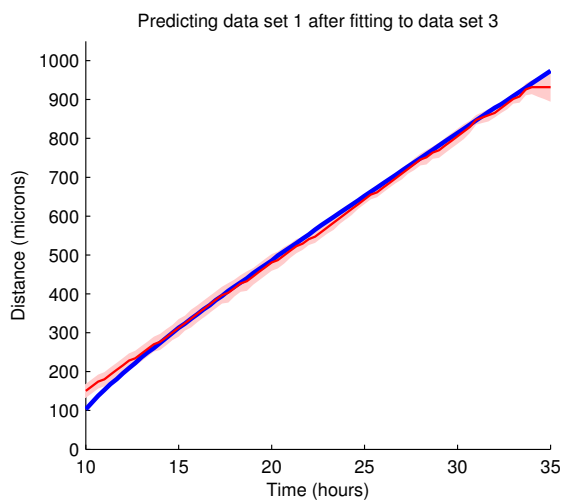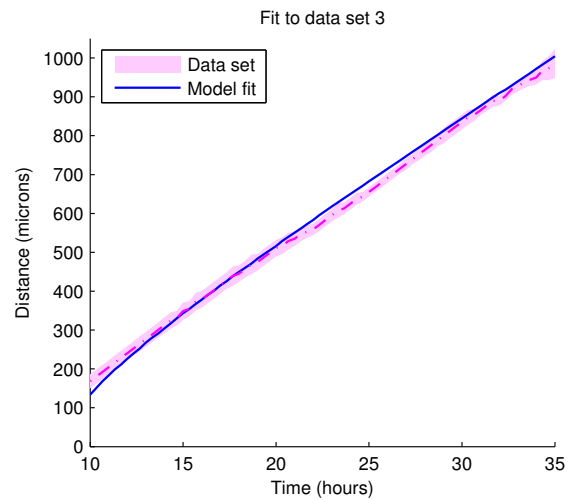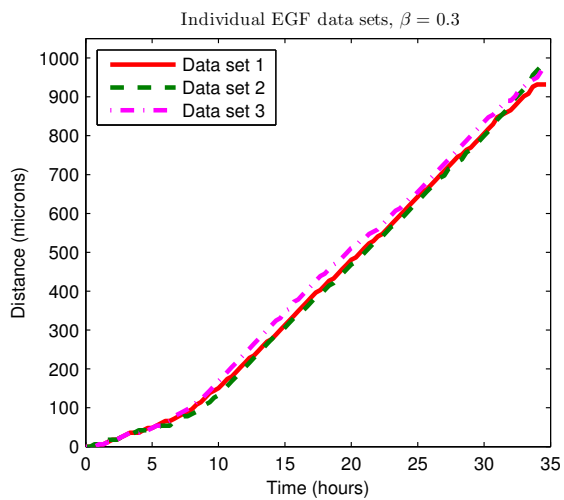
Figure 3: Fits and predictions for a leading edge density of 0.3.

where $P_{j+1/2}(t)$ is an approximation to the diffusive flux, given by

$$P_{j+1/2}(t) = \frac{1}{2}\left[Q\left(u_j(t), \frac{u_{j+1}(t) - u_j(t)}{\Delta x}\right) + Q\left(u_{j+1}(t), \frac{u_{j+1}(t) - u_j(t)}{\Delta x}\right)\right]$$

where $Q(u, u_x)$ denotes the cellular diffusion rate.

A.5 Table of transition probabilities

	Model H	Model P
τ_i^+	$D(1 - u_{i+1})(1 - \alpha u_{i-1})/\Delta x^2$	$D(1 - u_{i+1})(1 + \alpha u_{i+1})/\Delta x^2$
τ_i^-	$D(1 - u_{i-1})(1 - \alpha u_{i+1})/\Delta x^2$	$D(1 - u_{i-1})(1 + \alpha u_{i-1})/\Delta x^2$
τ_{i-1}^+	$D(1 - u_i)(1 - \alpha u_{i-2})/\Delta x^2$	$D(1 - u_i)(1 + \alpha u_i)/\Delta x^2$
τ_{i+1}^-	$D(1 - u_i)(1 - \alpha u_{i+2})/\Delta x^2$	$D(1 - u_i)(1 + \alpha u_i)/\Delta x^2$

Table 1: Table of transition probabilities, as discussed in Section 2.1.1 as a means to develop Models H and P. The transition rates for Model H denote cells in the direction opposite migration hindering migration through cell-cell adhesion, while the transition rates for Model P denote cells in the direction of migration promoting it through cell-cell adhesion.

A.6 Properties of the characteristic curves

If we assume that g is positive and uniformly continuous, then $\sigma^{-1}(t; \underline{y})$ exists and satisfies the following:

$$\frac{d}{dt}\sigma^{-1}(t; \underline{y}) = g(\sigma^{-1}(t; \underline{y})), \ \sigma^{-1}(0; \underline{y}) = \underline{y}. \tag{5}$$

To derive (5), see that

$$
\begin{aligned}
y(t) &= \sigma^{-1}(t; \underline{y}) \\
\Rightarrow \sigma(y(t); \underline{y}) &= t \\
\Rightarrow \frac{d}{dt}\left(\sigma(y(t); \underline{y})\right) = \frac{d}{dy}\sigma(y(t); s)\frac{dy}{dt} &= 1 \\
\Rightarrow \frac{1}{g(y(t))}\frac{dy}{dt} &= 1 \\
\Rightarrow \frac{dy}{dt} &= g(y(t)) \\
\Rightarrow \frac{d}{dt}\sigma^{-1}(t; \underline{y}) &= g(\sigma^{-1}(t; \underline{y})).
\end{aligned}
$$

and for the initial condition,

$$\sigma(\underline{y}, \underline{y}) \;=\; 0$$

$$\Rightarrow \sigma^{-1}(0, \underline{y}) \;=\; \underline{y}.$$

A.7 Derivation of (3.13)

In (3.12), we defined

$$v(t; \underline{y}) := u(t, y = \sigma^{-1}(t; \underline{y})).$$

Taking the derivative of $v(t; \underline{y})$ with respect to time, we find with the aid of the chain rule:

$$
\begin{aligned}
\frac{d}{dt}v(t;\underline{y}) \;&=\; \frac{\partial}{\partial t}u(t, y = \sigma^{-1}(t;\underline{y})) + \frac{\partial}{\partial y}u(t, y = \sigma^{-1}(t;\underline{y})) \cdot \frac{d}{dt}\sigma^{-1}(t;\underline{y}) \\
&=\; -\frac{\partial}{\partial y}\left[g\left(\sigma^{-1}(t;\underline{y})\right) u\left(t, y = \sigma^{-1}(t;\underline{y})\right) \right] \\
&\quad + Au\left(t, y = \sigma^{-1}(t;\underline{y})\right) + g\left(\sigma^{-1}(t;\underline{y})\right)\frac{\partial}{\partial y}u\left(t, y = \sigma^{-1}(t;\underline{y})\right) \\
&=\; -g'(\sigma^{-1}(t;\underline{y}))v(t;\underline{y}) + Av(t;\underline{y}).
\end{aligned}
$$

A.8 Derivation of the time-dependent Fisher's Equation

To investigate the averaged cell population behavior along m over time, we integrate (3.7) over m to find

$$
\begin{aligned}
w_t(t, x) \;&=\; \left(D_1 w_{xx} + \lambda_1 w(1 - w)\right) I_{[M_{inact}]}(m) \\
&\quad + \left(D_2 w_{xx} + \lambda_2 w(1 - w)\right) I_{[M_{act}]}(m).
\end{aligned}
\tag{6}
$$

An explicit form for (6) thus requires determining how much of the population is in the active and inactive populations over time. This is determined with the activation curves by calculating

$$h(t; \underline{m}) < m_{crit} \quad \Longleftrightarrow \quad F(t) < \sigma(m_{crit}; \underline{m})$$

$$\Longleftrightarrow \quad \underline{m} < \sigma^{-1}\left(-F(t); m_{crit}\right) =: \psi(t). \tag{7}$$

114

Thus, $\underline{m} = \sigma^{-1}(-F(t); m)$ maps the distribution along m at time t back to the initial distribution, $\phi_1(\underline{m})$, and $\psi(t)$ denotes the threshold value in $\underline{m}$ between the active and inactive populations over time. $\Phi_1(\psi(t))$ thus denotes the portion of the population in the inactive population, and $1 - \Phi_1(\psi(t))$ denotes the portion in the active population over time.

We thus derive a nonautonomous PDE for w, which we will term the *averaged nonautonomous Fisher's Equation*, as:

$$
\begin{aligned}
w_t &= D(t)w_{xx} + \lambda(t)w(1-w), \\
w(t=0, x) &= \phi_2(x) \\
w(t, x = -\infty) = 1 \qquad & w(t, x = \infty) = 0
\end{aligned}
$$

where

$$
\begin{aligned}
D(t) &= D_2 + (D_1 - D_2)\Phi_1(\psi(t)) \\
\lambda(t) &= \lambda_2 + (\lambda_1 - \lambda_2)\Phi_1(\psi(t)).
\end{aligned}
$$

A.9 Activation terms and criteria from Section 3.4

Recall that in each example, we set $m_{crit} = 0.5, D_1 = 0.01, D_2 = 1, \lambda_1 = .25, \lambda_2 = 0.0025, \phi_1(m) = {}^{10}\!/\!3 I_{[.05, 0.25]}(m), \phi_2(x) = I_{(-\infty, 5]}(x)$ and $g(m) = \alpha m(1-m)$.

A.9.1 Example 1

Using (3.10), (7), and $f(t) = 1$, we compute

$$
\begin{aligned}
\sigma(m; \underline{m}) &= \frac{1}{\alpha} \log\left(\frac{m}{1-m} \frac{1-\underline{m}}{\underline{m}} \right); \ \underline{m}, m \in (0, 1) \\
h(t; \underline{m}) &= \sigma^{-1}(t; \underline{m}) = \underline{m}\left((1 - \underline{m})e^{-\alpha t} + \underline{m} \right)^{-1} \\
\psi(t) &= \left(1 + e^{\alpha t} \right)^{-1}
\end{aligned}
$$

These functions demonstrate that the distribution along m is always activating along m but never reaches the $m = 1$ line, as $\sigma(m; \underline{m}) \to \infty$ as $m \to 1^-$ for any $\underline{m} \in (0, 1)$. The entire population

(excluding $\underline{m} = 0$) approaches $m = 1$ asymptotically, however, as $\lim_{t\to\infty} \sigma^{-1}(t; \underline{m}) = 1$.

A.9.2 Example 2

For $f(t) = \beta e^{\gamma t} - 1$, we compute

$$\sigma(m; \underline{m}) = \frac{1}{\alpha} \log\left(\frac{m}{1-m}\frac{1-\underline{m}}{\underline{m}}\right); \ \underline{m}, m \in (0, 1)$$

$$h(t; \underline{m}) = \underline{m}\left(\underline{m} + (1 - \underline{m})\exp\left[\alpha t - \frac{\alpha\beta}{\gamma}(\exp(\gamma t) - 1)\right]\right)^{-1}$$

$$\psi(t) = \left(1 + \exp\left[-\alpha t + \frac{\alpha\beta}{\gamma}(\exp(\gamma t) - 1)\right]\right)^{-1}.$$

Using (3.32), we determine our activation criterion for this example as

$$\frac{1 - \beta + \log\beta}{\gamma} > \frac{1}{\alpha} \log\left(\frac{m_{crit}}{1 - m_{crit}}\frac{1 - \underline{m}_{\max}}{\underline{m}_{\max}}\right).$$

If we fix $\gamma = -1, \alpha = 1, m_{crit} = 0.5, \underline{m}_{\max} = 0.35$, and $\underline{m}_{\min} = 0.05$, we find that the above inequality is satisfied for β approximately greater than 2.55. This may represent a scenario in which we know the decay rate of the ROS through γ, the activation rate of the MAPK signaling cascade through α, the MAPK activation distribution before ROS release with $\underline{m}_{\min}$ and $\underline{m}_{max}$, and the activation threshold with m_{crit}. The values of β denote the concentration of released ROS, which should be at least 2.55 to see the population activate. We similarly find that the entire population will activate at some time for $\beta > 5.68$.

A.9.3 Example 3

For $f(t) = \beta \sin(\gamma t)$, we compute

$$\sigma(m; \underline{m}) = \frac{1}{\alpha} \log\left(\frac{m}{1-m}\frac{1-\underline{m}}{\underline{m}}\right); \ \underline{m}, m \in (0, 1)$$

$$h(t; \underline{m}) = \underline{m}\left(\underline{m} + (1 - \underline{m})\exp\left[\frac{\alpha\beta}{\gamma}(\cos(\gamma t) - 1)\right]\right)^{-1}$$

$$\psi(t) = \left(1 + \exp\left[\frac{\alpha\beta}{\gamma}(1 - \cos(\gamma t))\right]\right)^{-1}.$$
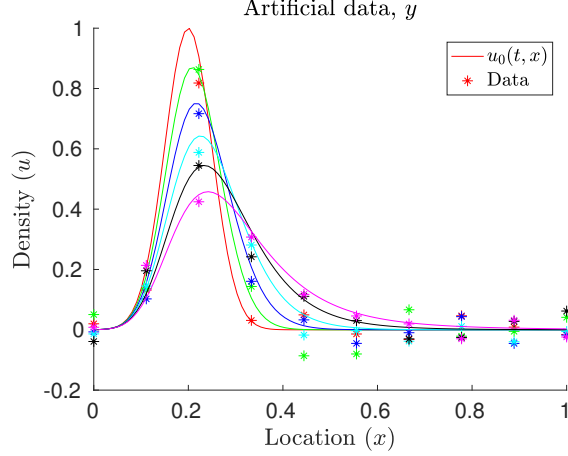
Figure 4: Artificial data from Equations (4.1) and (4.5) for $\eta = 0.1, N = 10$ and $\phi(x) = \phi_1(x)$. Dots denote the artificial data and the solid lines denote the analytical solution given by Equation (4.2). Red plots denote $t = 0$, green plots denote $t = 2$, blue plots denote $t = 4$, cyan plots denote $t = 6$, black plots denote $t = 8$, and magenta plots denote $t = 10$.

The activation criterion (3.32) can be solved as

$$\frac{2\beta}{\gamma} > \frac{1}{\alpha} \log \left(\frac{m_{crit}}{1 - m_{crit}} \frac{1 - \underline{m}_{\max}}{\underline{m}_{\max}} \right).$$

We thus calculate that if we fix $\beta = 1, \alpha = 1/2, \underline{m}_{max} = 0.35, \underline{m}_{\min} = 0.05$, and $m_{crit} = 0.5$, then the activation criterion (3.32) is satisfied for $\gamma < 1.615$ and the entire activation criterion (3.33) is satisfied for $\gamma < 0.34$. These estimates would tell us how frequently signaling factor treatment is needed to see different patterns of activation in the population.

A.10   Results for $\phi(x) = \phi_1(x)$

In Figure 4, we depict plots of artificially generated data when $\phi(x) = \phi_1(x)$. In Table 2, we depict the computed statistical orders of convergence for the numerical cost function when $\phi(x) = \phi_1(x)$ for all data sets considered. In Figure 5, we depict plots of the numerical cost function for all data sets and numerical schemes considered when $\phi(x) = \phi_1(x)$. In Figure 6, we depict plots of $\|\hat{\theta}_{OLS}^{M,N}(h) - \theta_0\|_2$ for all data sets and numerical schemes considered when $\phi(x) = \phi_1(x)$.

Method	p		η	p_J						
			N	0	10^{-4}	5×10^{-4}	10^{-3}	10^{-2}	5×10^{-2}	10^{-1}
UW	0.924		11	1.678	1.669	1.630	1.545	0.030	0.002	0.001
			31	1.882	1.869	1.881	1.670	0.173	0.005	0.002
			51	1.889	1.884	1.839	1.728	0.135	0.022	-0.003
LW	1.946		11	3.498	3.481	3.879	3.596	0.199	0.041	0.013
			31	3.768	3.746	3.900	3.580	0.225	0.037	0.012
			51	3.830	3.782	3.945	3.617	0.206	0.019	0.008
BW	1.838		11	2.760	2.762	2.533	2.510	0.141	0.007	0.000
			31	3.198	3.186	2.962	2.867	0.271	0.002	0.002
			51	3.235	3.227	2.960	2.903	0.242	0.030	-0.004
UW FL	0.918		11	2.201	2.208	2.043	1.888	0.145	0.017	0.002
			31	2.599	2.598	2.456	2.167	0.196	0.003	0.000
			51	2.663	2.659	2.485	2.141	0.165	0.020	-0.003

Table 2: Full list of results for order estimation when $\phi(x) = \phi_1(x)$. "UW" denotes the upwind method, "LW" denotes the Lax-Wendroff method, "BW" denotes the Beam-Warming method, and "UW FL" denotes the upwind method with flux limiters.

Method	p		η	p_J						
			N	0	10^{-1}	1.5×10^{-1}	2×10^{-1}	3×10^{-1}	5×10^{-1}	1
UW	0.584		11	0.517	0.357	0.237	0.208	0.128	0.078	-0.002
			31	0.612	0.393	0.309	0.226	0.156	0.089	0.040
			51	0.556	0.438	0.336	0.255	0.180	0.086	0.032
LW	0.474		11	0.966	0.826	0.573	0.490	0.263	0.176	-0.011
			31	0.878	0.601	0.498	0.387	0.257	0.222	0.062
			51	0.825	0.620	0.520	0.414	0.364	0.246	0.044
BW	.7878		11	0.785	0.516	0.441	0.367	0.195	0.173	0.000
			31	0.987	0.651	0.562	0.441	0.311	0.162	0.040
			51	0.896	0.706	0.611	0.491	0.327	0.216	0.025
UW FL	.957		11	1.285	0.756	0.512	0.409	0.275	0.145	-0.020
			31	1.338	0.824	0.642	0.505	0.295	0.155	0.037
			51	1.293	0.869	0.686	0.533	0.332	0.148	0.016

Table 3: Full list of results for order estimation when $\phi(x) = \phi_2(x)$. "UW" denotes the upwind method, "LW" denotes the Lax-Wendroff method, "BW" denotes the Beam-Warming method, and "UW FL" denotes the upwind method with flux limiters.
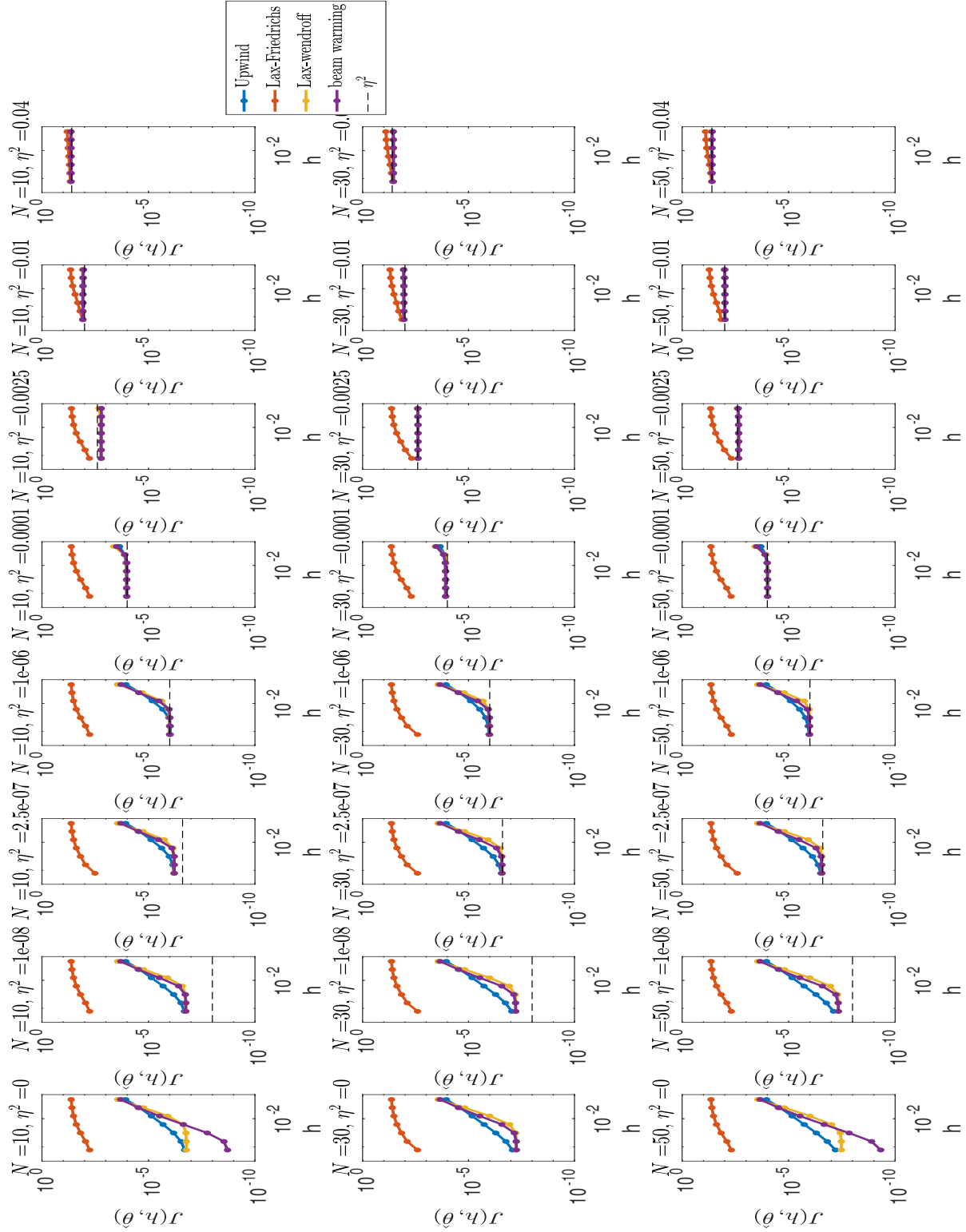
Figure 5: Plots of $J(h, \hat{\theta})$ for the four schemes considered for various values of h with $\phi(x)$ as a Gaussian function.
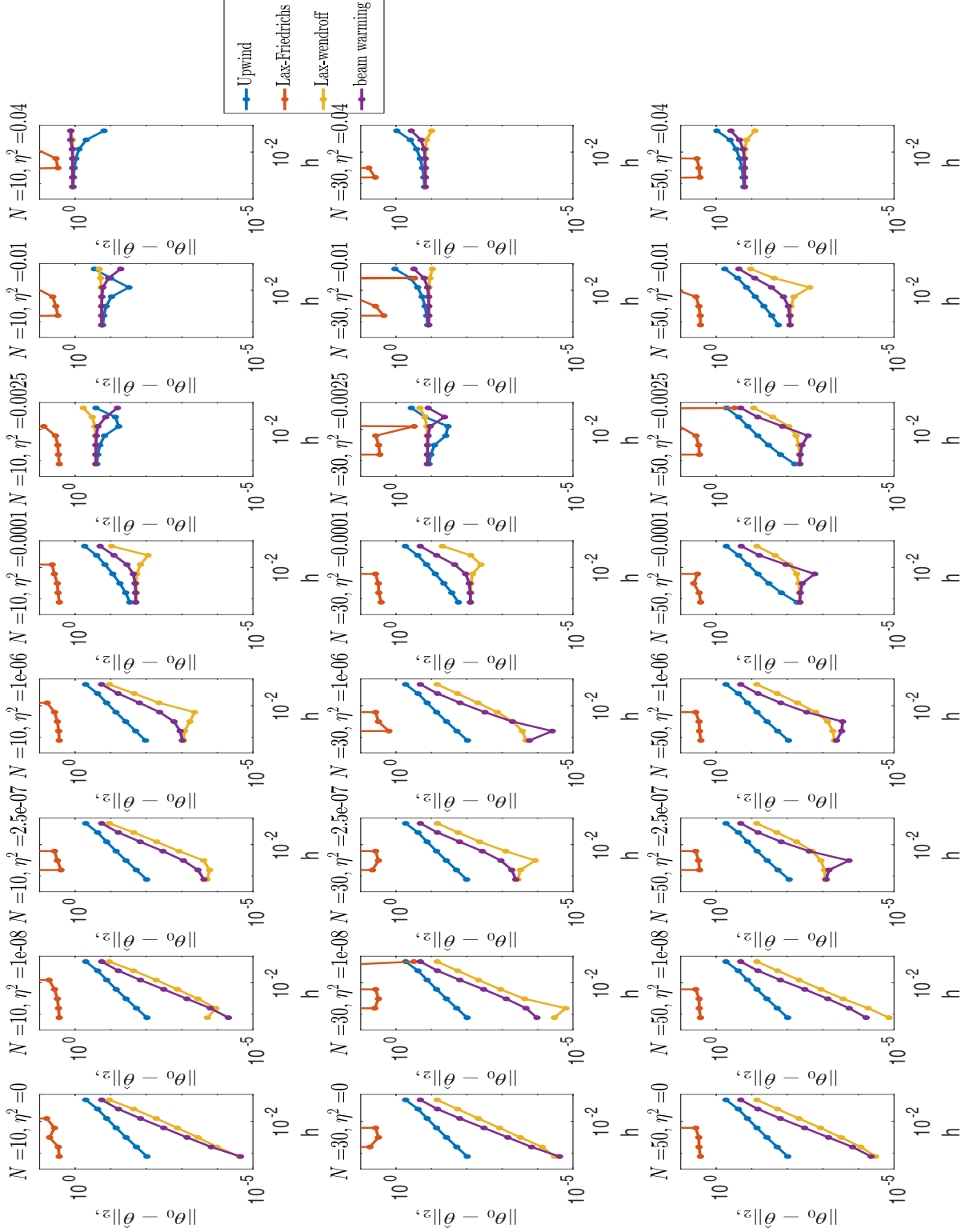
Figure 6: Plots of $\|\hat{\theta} - \theta_0\|_2$ for the four schemes considered for various values of h with $\phi(x)$ as a Gaussian function. .
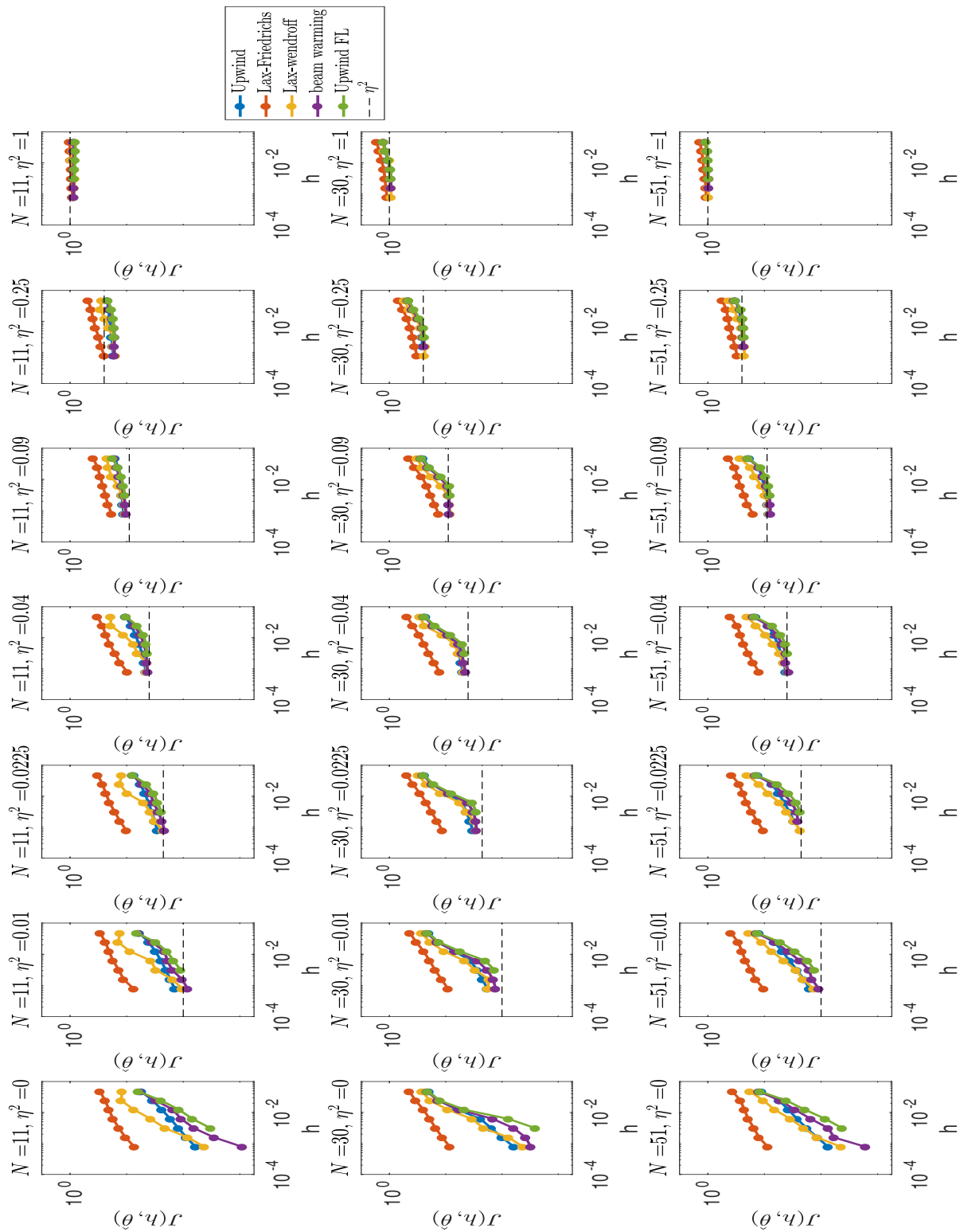
Figure 7: Plots of $J(h, \hat{\theta})$ for the four schemes considered for various values of h with $\phi(x)$ as a front. .
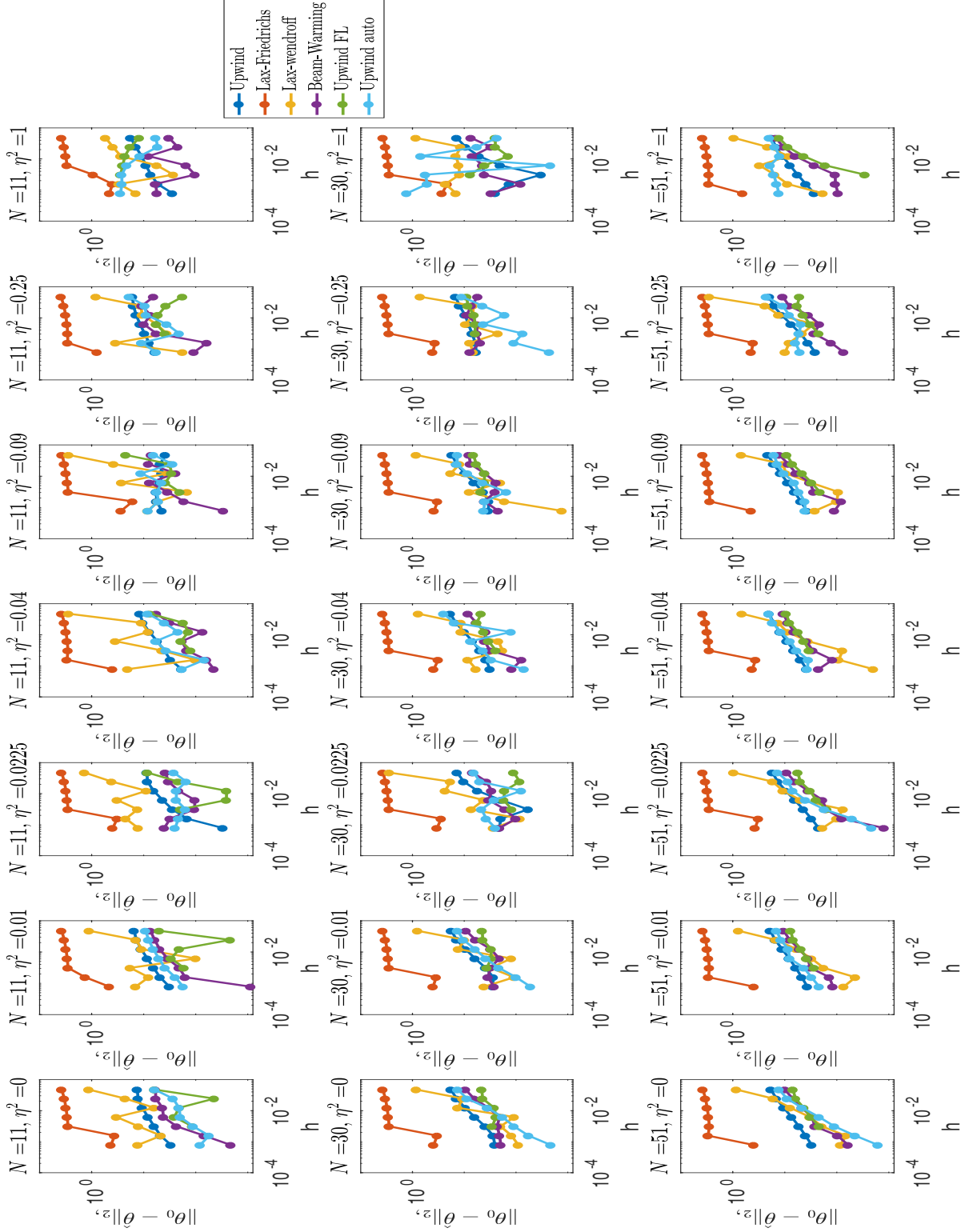
Figure 8: Plots of $\|\hat{\theta} - \theta_0\|_2$ for the four schemes considered for various values of h with $\phi(x)$ as a front. .

A.11 Extra Results for $\phi(x) = \phi_2(x)$

In Table 3, we depict the computed statistical orders of convergence for the numerical cost function when $\phi(x) = \phi_2(x)$ for all data sets considered. In Figure 7, we depict plots of the numerical cost function for all data sets and numerical schemes considered when $\phi(x) = \phi_2(x)$. In Figure 8, we depict plots of $\|\hat{\theta}_{OLS}^{M,N}(h) - \theta_0\|_2$ for all data sets and numerical schemes considered when $\phi(x) = \phi_2(x)$.

A.12 Theory on $\hat{\theta}_{OLS}^{M,N}(h)$

Consider

$$\hat{\theta}_{OLS}^{M,N}(h) = \arg \min_{\theta \in Q_{ad}} J^{M,N}(h, \theta).$$

In the standard theory, where numerical error is negligible, we have that as $M, N \to \infty$,

$$\hat{\theta}_{OLS}^{M,N} \approx \mathcal{N}(\theta_0, \sigma^2 C^{-1}), \ C = \nabla u(\theta_0)^T \nabla u(\theta_0). \tag{8}$$

This can be shown as follows from [Seber and Wild, 1988]:

near θ_0,

$$u(\theta) \approx u(\theta_0) + \nabla u(\theta_0)[\theta - \theta_0],$$

so

$$\begin{aligned} J(\theta) &= \|\vec{y} - \vec{u}(t, x|\theta)\|^2 \\ &\approx \|\vec{y} - \vec{u}(t, x|\theta_0) - \nabla \vec{u}(t, x|\theta_0)[\theta - \theta_0]\|^2 \\ &= \|\vec{z} - \nabla \vec{u}(\theta_0)\beta\|^2, \end{aligned} \tag{9}$$

a linear model in $\vec{z} = \vec{\epsilon}$ and $\beta = \theta - \theta_0$. Accordingly, we see that the value of β that will minimize the final part Equation (9) will be

$$\beta = [\nabla \vec{u}(t, x|\theta_0)^T \nabla \vec{u}(t, x|\theta_0)]^{-1} \nabla \vec{u}(t, x|\theta_0)\vec{z},$$

so in the asymptotic limit as $M, N \to \infty$, and $\hat{\theta}_{OLS}^{M,N}$ approaches θ_0,

$$\hat{\theta}_{OLS}^{M,N} - \theta_0 \approx [\nabla \vec{u}(t, x|\theta_0)^T \nabla \vec{u}(t, x|\theta_0)]^{-1} \nabla \vec{u}(t, x|\theta_0) \vec{\epsilon}, \tag{10}$$

a linear combination of normal random variables. so $\hat{\theta}_{OLS}^{M,N}$ is itself normally distributed. We see it has mean value θ_0.

Becase $\text{cov}(AX + a) = A\text{cov}(X)A^T$, we can also calculate that

$$\begin{aligned}
\text{cov}(\hat{\theta}) &= [\nabla u(\theta_0)^T \nabla u(\theta_0)]^{-1} \nabla u(\theta_0)^T \eta^2 [[\nabla u(\theta_0)^T \nabla u(\theta_0)]^{-1} \nabla u(\theta_0)^T]^T \\
&= \eta^2 \nabla u(\theta_0)^{-1} \nabla u(\theta_0)^{-T} \nabla u(\theta_0)^T \nabla u(\theta_0) \nabla u(\theta_0)^{-1} \nabla u(\theta_0)^{-T} \\
&= \eta^2 \nabla u(\theta_0)^{-1} \nabla u(\theta_0)^{-T} \\
&= \eta^2 [\nabla u(\theta_0)^T \nabla u(\theta_0)]^{-1}.
\end{aligned}$$

This proves Equation (8).

Now we consider the case where our numerical computation has a numerical order of convergence, p. In this case, when θ is near θ_0,

$$\begin{aligned}
\vec{u}(t, x|h, \theta) &\approx \vec{u}(t, x|h, \theta_0) + \nabla \vec{u}(t, x|h, \theta_0)[\theta - \theta_0] \\
&\approx \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p) + [\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p)][\theta - \theta_0].
\end{aligned}$$

So we see that our numerical cost function takes the form

$$\begin{aligned}
J(h, \theta) &= \|\vec{y} - u(t, x|h, \theta)\|^2 \\
&= \|\vec{y} - \vec{u}_0(t, x|\theta_0) - \mathcal{O}(h^p) - [\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p)][\theta - \theta_0]\|^2 \\
&\approx \|\vec{y} - \vec{u}_0(t, x|\theta_0) - [\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p)][\theta - \theta_0]\|^2 \\
&= \|\vec{z} - [\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p)]\beta\|^2
\end{aligned}$$

which has minimizer

$$\hat{\beta} = \left(X_h^T X_h\right)^{-1} X_h \vec{z} \tag{11}$$

where $X_h = \nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p)$. So, once $\hat{\theta}_{OLS}^{M,N}(h)$ is close to θ_0,

$$\hat{\theta}_{OLS}^{M,N}(h) \approx \mathcal{N}(\theta_0, \eta^2 (X_h^T X_h)^{-1}).$$

In general, we cannot say much about the inverse of $(X_h^T X_h)$ in regard to $\nabla u(\theta_0)^T \nabla u(\theta_0)$. However, note that $|X_h^T X_h| = |(\nabla u(\theta_0))^T + \mathcal{O}^T(h^p))(\nabla u(\theta_0)) + \mathcal{O}(h^p))| = |\nabla u(\theta_0)^T \nabla u(\theta_0) + \mathcal{O}(h^p)\nabla u(\theta_0)) + \mathcal{O}(h^{2p})| \approx |\nabla u(\theta_0)^T \nabla u(\theta_0) + \mathcal{O}(h^p)\nabla u(\theta_0))|$, meaning that we are perturbing each components of $\nabla u(\theta_0)^T \nabla u(\theta_0)$ with a relative error on the order of h^p. Then one can show

$$|(\nabla u(\theta_0)^T \nabla u(\theta_0))^{-1} - (X_h^T X_h)^{-1}| \leq \mathcal{O}(h^p)|(\nabla u(\theta_0)^T \nabla u(\theta_0))^{-1}||\nabla u(\theta_0)^T \nabla u(\theta_0)||(\nabla u(\theta_0)^T \nabla u(\theta_0))^{-1}|$$
$$+ \mathcal{O}(h^{2p}),$$

see [Higham, 1996, § 13.1] for details. Thus, the entries of the covariance matrix of $\hat{\theta}_{OLS}^{M,N}(h)$ should be within $\mathcal{O}(h^p)$ of the entries of the true covariance matrix for $\hat{\theta}_{OLS}^{M,N}$.

Also note that if $X_h^T X_h$ is invertible, then it is a positive definite matrix. In this case, $(X_h^T X_h)^{-1}$ is also positive definite.

A.13 Convergence of the terms of $J_{OLS}^{M,N}(h, \theta)$

Here we discuss the asymptotic properties of $J_{OLS}^{M,N}(h, \theta)$. We begin with the limits as $h \to 0$ in Section A.14 and as $M, N \to \infty$ in A.14.

For brevity, we denote $\hat{\theta}_{OLS}^{M,N}(h)$ as $\hat{\theta}$ for the rest of this section. Following the footsteps of this previous study, we present here some assumptions that must be made for the estimator to be consistent and to facilitate our analysis.

(A1) There exists finite measures χ and ν such that

$$\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} u_0(t_i, x_j|\theta) \to \int_X \int_T u_0(t, x|\theta) d\nu(t) d\chi(x)$$

and

$$\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} u(t_i, x_j|h, \theta) \to \int_X \int_T u(t, x|h, \theta) d\nu(t) d\chi(x)$$

125

for all $\theta \in Q_{ad}$ as $M, N \to \infty$, independent of h.

(A2) The functional

$$J^*(\theta) = \int_X \int_T (u_0(t, x|\theta_0) - u_0(t, x|\theta))^2 d\nu(t) d\chi(x)$$

has a unique minimizer in Q_{ad} at θ_0.

(A3) The numerical method used is order p accurate, so that

$$\sum_{i=1}^M \sum_{j=1}^N [u_0(t_i, x_j|\theta) - u(t_i, x_j|h, \theta)] \approx Kh^p$$

for some positive value, p, for all compact subsets of Q. Furthermore, for every h, $u(x, t|h, \theta)$: $Q_{ad} \to L^1(T \times X)$ is continuous.

Note that [Banks and Fitzpatrick, 1990] includes more assumptions than those listed here that are already satisfied by Equations (4.1) or (4.5) and that assumption (A3) is a slight modification of assumption (A5) in [Banks and Fitzpatrick, 1990]. Assumption (A3) will only be valid in practice when h is chosen small enough [LeVeque, 2007].

A.14 Limits as $h \to 0$

Term A is independent of h.

For Term B as $h \to 0$, and $\hat{\theta}$ approaches θ_0 (assuming that there are many data points used for this to occur). We can Taylor expand about θ_0 and find

$$B(\hat{\theta}) \approx \frac{1}{MN} \sum_{i,j=1}^{M,N} [\nabla u_0(t_i, x_j|\theta_0)[\hat{\theta} - \theta_0]]^2,$$

We thus observe that B has expectation

$$
\begin{aligned}
\mathbb{E}[B(\hat{\theta})] &= \frac{1}{MN} \sum_{i,j=1}^{M,N} \mathbb{E}\left(\left[\nabla u_0(t_i, x_j | \theta_0)[\hat{\theta} - \theta_0] \right]^2 \right) \\
&= \frac{1}{MN} \sum_{i,j=1}^{M,N} \left[\operatorname{var}\left(\nabla u_0(t_i, x_j | \theta_0)[\hat{\theta} - \theta_0] \right) + \mathbb{E}\left(\nabla u_0(t_i, x_j | \theta_0)[\hat{\theta} - \theta_0] \right)^2 \right] \\
&= \frac{1}{MN} \sum_{i,j=1}^{M,N} \nabla u_0(t_i, x_j | \theta_0)^T \operatorname{var}\left(\hat{\theta} - \theta_0 \right) \nabla u_0(t_i, x_j | \theta_0) \\
&= \frac{\eta^2}{MN} \sum_{i,j=1}^{M,N} \nabla u_0(t_i, x_j | \theta_0)^T C_h \nabla u_0(t_i, x_j | \theta_0) \\
&= \frac{\eta^2}{MN} \nabla \vec{u}_0(t, x | \theta_0)^T C_h \nabla \vec{u}_0(t, x | \theta_0)
\end{aligned}
$$

which as $h \to 0$ will approach

$$
\frac{\eta^2}{MN} \nabla \vec{u}_0(t, x | \theta_0)^T \left(\nabla \vec{u}_0(t, x | \theta_0)^T \nabla \vec{u}_0(t, x | \theta_0) \right)^{-1} \nabla \vec{u}_0(t, x | \theta_0)
$$

with order of convergence $\mathcal{O}(h^p)$. The variance of this random variable is difficult to compute explicitly.

The third term is given by

$$
C := \frac{1}{MN} \sum_{i,j=1}^{M,N} [u_0(t_i, x_j | \theta) - u(t_i, x_j | h, \theta)]^2,
$$

which may also be written in terms of the Euclidean-vector norm, from where we can then use equivalence of norms to show its order as $2p$ as $h \to 0$ using (A3)

$$
\begin{aligned}
C &= \frac{1}{MN} \| u_0(t_i, x_j | \theta) - u(t_i, x_j | h, \theta) \|_2^2 \\
&\leq \frac{K}{MN} \| u_0(t_i, x_j | \theta) - u(t_i, x_j | h, \theta) \|_1^2 \approx \frac{K_1}{MN} h^{2p}.
\end{aligned}
$$

We can also investigate the asymptotic properties as $h \to 0$ as follows. We begin with a Taylor

expansion about θ_0 :
$$D(\hat{\theta}) \approx \frac{2}{MN} \sum_{i,j=1}^{M,N} \epsilon_{i,j}(-\nabla u_0(t_i, x_j|\theta_0)(\hat{\theta} - \theta_0)).$$

We note that this is the sum of the product of two dependent random variables, so we use Equation (11) from the appendix and see

$$
\begin{aligned}
D(\hat{\theta}) &\approx \frac{-2}{MN} \sum_{i,j=1}^{M,N} \epsilon_{i,j}(\nabla u_0(t_i, x_j|\theta_0)C_h(\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p))^T \vec{\epsilon}. \\
&= \frac{-2}{MN} \vec{\epsilon}^T \nabla \vec{u}_0(t, x|\theta_0)^T C_h(\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p))^T \vec{\epsilon} \\
&= \frac{-2}{MN} tr\left(\nabla \vec{u}_0(t, x|\theta_0)^T C_h(\nabla \vec{u}_0(t, x|\theta_0) + \mathcal{O}(h^p))^T \vec{\epsilon}\vec{\epsilon}^T\right) \\
&= \frac{-2}{MN} \sum_{i,j=1}^{M,N} \nabla u_0(t_i, x_j|\theta_0)^T C_h(\nabla u_0(t_i, x_j|\theta_0) + \mathcal{O}(h^p))^T \epsilon_{i,j}^2
\end{aligned}
$$

where $tr(\cdot)$ denotes the trace function for a matrix. The exact distribution of D as $h \to 0$ is thus hard to ascertain because we are not summing identical random variables, however, we can deduce that the mean of D is zero, and the variance will be on the order of $\mathcal{O}(h^{2p})$. We thus conclude that this term should be negligible as $h \to 0$.

On the other hand, for M, N finite,

$$
\begin{aligned}
|E|^2 &\leq \frac{4}{M^2 N^2} \left(\sum_{i,j=1}^{M,N} |\epsilon_{i,j}|^2\right) \left(\sum_{i,j=1}^{M,N} |(u_0(x_i, t_j\theta) - u(x_i, t_j|h, \theta))|^2\right) \text{ as } h \to 0 \\
&\leq K \left(\sum_{i,j=1}^{M,N} |(u_0(x_i, t_j\theta) - u(x_i, t_j|h, \theta))|\right) \\
&\approx \mathcal{O}(h^p)
\end{aligned}
$$

where the first inequality is by the Cauchy-Schwarz Inequality, the second intequality is by the equivalence of finite-dimensional norms, and the final approximaiton is by assumption (A3).

The final term is written as

$$F := \frac{2}{MN} \sum_{i,j=1}^{M,N} \left[(u_0(x_i, t_j|\theta) - u(x_i, t_j|h, \theta))(u_0(x_i, t_j|\theta_0) - u_0(x_i, t_j|\theta))\right].$$

By Taylor expanding, we can approximate this at $\hat{\theta}$ as

$$F(\hat{\theta}) \approx \frac{2}{MN} \sum_{i,j=1}^{M,N} \left[(u_0(x_i, t_j|\theta_0) - u(x_i, t_j|h, \theta_0)) \left(\nabla u_0(x_i, t_j|\theta_0)[\hat{\theta} - \theta_0] \right) \right],$$

where the second term in the first part of the product has been neglected because it has zero expectation and a higher order variance than the remaining term. We next use the Cauchy-Schwartz Inequality to deduce

$$\begin{aligned}
\left| F(\hat{\theta}) \right| &\lesssim \frac{2}{MN} \|u_0(x_i, t_j|\theta_0) - u(x_i, t_j|h, \theta_0)\|_2 \left\| \nabla u_0(x_i, t_j|\theta_0)[\hat{\theta} - \theta_0] \right\|_2 \\
&\leq K \frac{2}{MN} \|u_0(x_i, t_j|\theta_0) - u(x_i, t_j|h, \theta_0)\|_1 \left\| \nabla u_0(x_i, t_j|\theta_0)[\hat{\theta} - \theta_0] \right\|_1 \\
&= \mathcal{O}(h^p) \left\| \nabla u_0(x_i, t_j|\theta_0)[\hat{\theta} - \theta_0] \right\|_1
\end{aligned}$$

where the term

$$\sum_{i,j=1}^{M,N} \nabla u_0(x_i, t_j|\theta_0)[\hat{\theta} - \theta_0]$$

has mean zero and variance

$$\eta^2 \nabla \vec{u}_0(t, x|\theta_0)^T C_h \nabla \vec{u}_0(t, x|\theta_0).$$

We accordingly suggest that Term F should have mean 0 and a variance that converges with order $\mathcal{O}(h^{2p})$ to $\eta^2 \nabla \vec{u}_0(t, x|\theta_0)^T \left(\nabla \vec{u}_0(t, x|\theta_0)^T \nabla \vec{u}_0(t, x|\theta_0) \right)^{-1} \nabla \vec{u}_0(t, x|\theta_0)$.

Limits as $M, N \to \infty$

Note that $\epsilon_{i,j}^2$ is distributed as η^2 times a degree-1 chi-squared random variable. We thus observe that A is distributed as η^2/MN times a degree-MN chi-squared random variable, which has mean η^2 and variance $2\eta^4/MN$. By the classical Central Limit Theorem,

$$\sqrt{MN}(A - \eta^2) \xrightarrow{D} \mathcal{N}(0, 2\eta^2)$$

as $M, N \to \infty$, where $\xrightarrow{D}$ denotes convergence in distribution.

Term B is the sum of the difference of the true solution squared when computed at θ_0 and θ . If we

assume that the sampling points $\{(x_i, t_j)\}_{i=1,..,M}^{j=1,..,N}$ are chosen properly to "cover" the spatio-temporal domain[1] then

$$B := \frac{1}{MN} \sum_{i,j=1}^{M,N} [u_0(t_i, x_j|\theta_0) - u_0(t_i, x_j|\theta)]^2 \to J^*(\theta).$$

as $M, N \to \infty$ by (A1).

The best we can say about C as $M, N \to \infty$ is that

$$C \to \int (u_0(t, x|\theta) - u(t, x|h, \theta))^2 d\nu(t) d\chi(x),$$

but there is not much more to say about this term.

The fourth term is written as

$$D := \frac{2}{MN} \sum_{i,j=1}^{M,N} \epsilon_{i,j}(u_0(t_i, x_j|\theta_0) - u_0(t_i, x_j|\theta)).$$

We can bound the second term in the series, so D will go to zero as $M, N \to \infty$ by the LLN.

The same logic can be used to show that for h fixed, the fifth term,

$$E := \frac{2}{MN} \sum_{i,j=1}^{M,N} \epsilon_{i,j}(u_0(t_i, x_j\theta) - u(x_i, t_j|h, \theta))$$

will go to zero as $M, N \to \infty$ by LLN, so long as the CFL condition is satisfied so that $u(x_i, t_j|h, \theta)$ is bounded.

For the sixth term, F, we can use the LLN to show that this term will converge to zero as $M, N \to \infty$.

---

[1] see [Banks and Fitzpatrick, 1990] for details.