

**MATHEMATICAL MODELS OF BRANCHING
ACTIN NETWORKS: RESULTS AND METHODS.**

by

Daniel B. Smith

BS, University of Minnesota, 2006

MA, University of Pittsburgh, 2009

Submitted to the Graduate Faculty of
the Department of Mathematics in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH
DEPARTMENT OF MATHEMATICS

This was presented

by

Daniel B. Smith

It was defended on

August 13, 2012

and approved by

Jon Rubin, Department of Mathematics

David Swigon, Department of Mathematics

Bard Ermentrout, Department of Mathematics

Michael Grabe, Department of Biological Sciences

Jian Liu, NHLBI, NIH

Committee Chair: Jon Rubin, Department of Mathematics

Committee co-chair: David Swigon, Department of Mathematics

Copyright © by Daniel B. Smith
2012

MATHEMATICAL MODELS OF BRANCHING ACTIN NETWORKS: RESULTS AND METHODS.

Daniel B. Smith, PhD

University of Pittsburgh, 2012

Branching actin networks are made up of polymeric actin filaments and play a principle role in cell motility and other cellular processes. Inside the lamellipodium, the thin extension at the leading edge of a motile cell, there is a dense actin network composed of branched filaments. That network organizes into regular patterns near the membrane and serves as an engine moving the membrane forward. There are good models explaining how an individual actin filament is able to generate force against a load, but it is not well understood how filament networks collectively generate force. Multiple patterns have been observed in the force-velocity relationships of actin networks. The first part of this dissertation uses a agent-based stochastic to attempt to explain those patterns. We find that the rate of filament turnover can determine the nature of the force-velocity relationship.

Electron micrographs of actin networks have shown surprisingly regular patterns in the angle of filaments relative to the membrane normal. Several continuum models have been proposed to explain this regularity. In the second part of this work, the limiting behavior of those models are characterized. It has been hypothesized, with numerical evidence, that the models select for some small number of optimal orientation patterns. The results in the second section imply that both orientation models uniquely select for an optimal orientation pattern. Also, a fitness function for each orientation pattern is derived.

A number of properties of actin filaments have been studied by using atomistic models of actin monomers and filaments. Calculating those properties amounts a difficult sampling problem as the properties are statistical in nature. In order to calculate statistical properties

of these models, the conformational space needs to be effectively sampled. Current computing capabilities are unable to do so directly, so some form of enhanced sampling algorithm is needed. However, there is no standard way to compare existing methods nor test new methods. The last part of this dissertation proposes a model that would allow for standardized testing of a large class of enhanced sampling methods.

TABLE OF CONTENTS

1.0 INTRODUCTION	1
2.0 SIMULATIONS OF BRANCHING ACTIN NETWORKS SUGGEST LOAD-SHARING MECHANISM DETERMINES FORCE-VELOCITY RELATIONSHIP.	5
2.1 Background	5
2.2 Defining the Model	12
2.2.1 Methods	12
2.3 Simulation Results	18
2.4 Discussion	31
3.0 ACTIN ORIENTATION DENSITY EQUATIONS SELECT FOR A UNIQUE, STABLE EQUILIBRIUM.	37
3.1 Introduction	37
3.1.1 Definitions and Assumptions	41
3.1.2 Survey of Results	42
3.2 Proof of First-Order Theorem	45
3.3 Proof of Second Theorem	56
3.4 Proof of Zeroth-Order Propositions	57
3.5 Simulations	59
3.5.1 Methods	60
3.5.2 Results	61
3.6 Perturbation Expansion	64
3.7 Calculating Equilibrium Distributions	70

3.8	Progress on Zeroth-Order Stability	71
3.9	Generalization to Measures	76
3.10	Discussion	80
4.0	MDMS: MOLECULAR DYNAMICS META-SIMULATOR FOR EVALUATING EXCHANGE TYPE SAMPLING METHODS.	82
4.1	Introduction	82
4.2	System Definition and Basic Properties	85
4.2.1	Assumptions	85
4.2.2	Transition Rates	86
4.2.3	Equilibrium Distribution	88
4.2.4	Limits	90
4.2.5	Energies	91
4.2.6	Exchanges	93
4.2.7	Generalizing to Free Energies	93
4.2.8	Heat Capacity and Entropy	95
4.3	Example Models	98
4.3.1	Diffusion-Limited Cases	98
4.3.2	Non-Diffusion-Limited Cases	99
4.3.3	Combining Models	100
4.4	Simulations	100
4.4.1	Simulation Details	100
4.4.2	Test Problem	101
4.4.3	Results	101
4.5	Discussion	108
5.0	DISCUSSION	110
	APPENDIX. PSEDUO-CODE FOR THE AGENT-BASED MODEL	115
	Bibliography	120

LIST OF TABLES

2.1	Table of actin network model parameters.	13
4.1	Table of MDMS energy values.	101

LIST OF FIGURES

2.1	Actin motility diagram	7
2.2	Example convex and concave curves	9
2.3	Pseudo-code for event calculation.	14
2.4	Fast response of actin network.	19
2.5	Varying the capping rate changes the shape of the force-velocity curve.	20
2.6	Actin network velocities rapidly converge.	21
2.7	Cases where actin network velocity doesn't converge.	22
2.8	Correlation of contacting filaments with velocity	22
2.9	Diagram showing how the average lifetime/length of the filaments could influence the number of contactin filaments.	24
2.10	Plots of simulations testing for hysteresis in the force-velocity relationship.	26
2.11	Plot of $f_{\frac{1}{2}}$ for a variety of parameters.	27
2.12	Equilibrium velocity reduction in response to a small force.	28
2.13	Plots regarding actin network orientation.	29
2.14	Angular density plots for various branching conditions	30
2.15	Spatial density of filaments	32
2.16	Force-velocity curves for restricted branching.	33
3.1	Filament angle diagram	39
3.2	Equilibrium orientation distributions	62
3.3	The observed L^1 distances from equilibrium as a function of time	63
4.1	Theoretical and observed MDMS energy distributions.	102
4.2	MDMS model without exchanges converges.	103

4.3	Simple replica exchange simulation converges significantly faster than an MD simulation.	105
4.4	Replica exchange simulations with varying exchange frequencies.	106
4.5	Replica exchange simulations with two different sets of temperature intervals.	107
A1	A flow chart diagramming the basic simulation structure.	116

1.0 INTRODUCTION

Actin is a globular protein involved in many cellular processes ranging from regulating gene transcription to acting as a motor in cell motility [27]. It is one of the most conserved proteins in all eukaryotic cells. Actin is a protein that is seen both as isolated, globular proteins (g-actin) and as a polymeric filament (f-actin) where many actin monomers are bound together to form a thin filament [100, 126]. Given the broad function of actin, there has been much study of actin's function both in the laboratory and through mathematical modeling. The results presented here contribute both to the mathematical modeling of actin and to the development of methods for future modeling.

Actin filaments form highly branched networks near the leading edge of motile cells [80]. While actin monomers will spontaneously polymerize in physiological conditions, inside these branched networks, new filaments are generated by branching off of existing filaments [100]. New filaments are nucleated by the actin related proteins 2 and 3 complex (Arp2/3). To maintain a consistent supply of actin monomers, actin filaments are eventually severed and depolymerized. Filament density is regulated by capping protein binding to the filament tips, ceasing polymerization [16]. Combined with filaments growing by the addition of new monomers, these processes create a dynamic network that serves as the engine in certain types of cell motility [100, 104]. When growing filaments reach the membrane, the filaments continue to grow, albeit at a slower pace, which pushes the membrane forward. A more detailed explanation of this process is included in chapter 2.

Mathematical modeling of actin has flourished in recent years [32–34, 37, 38, 74, 97]. Models of actin have been constructed across huge ranges of time and spatial scales [16, 97]. At one extreme, atomistic models of actin have studied the structure and binding properties of actin monomers and filaments. At the other extreme, branching actin networks have been

modeled as a gel with fluid dynamics techniques [74, 87, 97]. The first papers modeling actin dynamics were perturbation calculations modeling how an individual actin filament could generate force against an object [77, 93]. Growing from that work, a number of agent-based models have been developed to study how a branching actin network could generate force [114, 117, 137]. Models have also been constructed studying the geometric arrangement of the branched network [68, 115, 137].

Chapter 2 reports a set of simulations investigating how a branching actin network could collectively generate force. A number of experiments have been performed probing the velocity of a branching actin network growing against a given resisting force [10, 11, 46, 52, 69, 73, 81, 89, 101, 135, 139]. Those measurements have been used to generate force-velocity curves. Different experiments have shown varying shapes for these curves. Several models have been developed to attempt to explain this disparity [117, 137]. The model developed in chapter 2 builds upon the work by Weichsel and Schwarz [137] investigating the observed force-velocity curves. We find that the two force-velocity curves can be explained by the rate of filament turnover that is regulated by the ratio of capping and branching to the filament growth rate. The total number of filaments growing at the leading edge determines the velocity of the network and could explain other experimental effects.

Electron micrographs of branched actin networks have shown surprisingly regular patterns of filament orientation [68, 124, 131, 132, 138]. There is some amount of disagreement as to the exact nature of the pattern [68, 138], but organized patterns have generally been reported. In spite of these observations, there is not a generally accepted mechanism even for the simple question of why the growing end of actin filaments point towards the membrane. There have been a number of modeling studies to examine what produces the observed regular patterns [68, 115, 137]. However, there has not been rigorous study to show that the models are well-posed. The models amount to solutions of an ODE on an infinite-dimensional Banach space, so proving that solutions exist is relatively straightforward. However, the nature of the solutions is non-obvious. For example, it is important to know if the models uniquely select for some number of orientation patterns. Chapter 3 proves that one model uniquely selects for a single orientation pattern and provides strong evidence that another model does also. Those results justify some of the numerical calculations. It also shows that

the orientation pattern fulfills an optimization problem giving credence to the notion that the models are selecting for an optimal orientation pattern.

Molecular dynamics (MD) simulations of actin monomers have been used to study how the internal dynamics of individual monomers affects filament formation and behavior [19, 23, 143]. MD simulations are used to study proteins at an atomic level. Those simulations are based on empirical energy functions fit to some experimental or quantum-level calculations. Energy functions used include all atom models [143] and coarse-grained models [19, 20] where groups of atoms are modeled in unison. These models can be used to calculate physical properties of proteins including energies [20] or physical properties such as filament stiffness [45]. In order to calculate statistical properties such as energies, the protein conformational landscape needs to be effectively sampled. However, given current computing power, many systems of biological interest suffer the ‘curse of dimensionality’ and cannot be efficiently sampled directly. Enhanced sampling methods are algorithms designed to increase the rate of sampling given a certain number of computer cycles or a certain amount of wall clock time [134, 146]. Development of new methods remains a focus of extensive research currently.

Some examples of MD calculations done on actin include simulations to estimate the behavior of actin monomers in different binding states with ATP or ADP [19, 23, 143]. However, the system size was generally too large for any statistical quantities such as energy differences between states to be calculated without significant coarse-graining such as in [20]. To estimate such quantities for large systems, new enhanced sampling methods need to be developed. In order for calculations to be done on systems the size of the actin simulations above, new computational methods will need to be developed.

In contrast to the previous two chapters, chapter 4 is focused on method development for studying atomistic models. While enhanced sampling methods have been used extensively, there does not exist any way to systematically compare methods [146]. Molecular Dynamics Meta-Simulator (MDMS) models are proposed in chapter 4 to begin to bridge that gap for at least one class of enhanced sampling methods. Replica exchange methods are a type of enhanced sampling method that entails running parallel simulations with different conditions and exchanging states. One common example is running simulations at different temperatures, utilizing the faster sampling of the replicas run at higher temperatures [30,

[127](#), [146](#)]. MDMS abstracts away the dynamics of the system in between exchanges to allow for fast testing of replica exchange methods. We are able to exactly derive the equilibrium distribution and thermodynamic properties such as entropy and heat capacity for MDMS models. While we do not provide an analytical formulation, MDMS allows for an exact characterization of the the benefit of replica exchange in terms of the eigenvalues of the matrix associated with the Markov chain. Finally, we run a few elementary tests to demonstrate the usefulness of MDMS simulations. We show that a recently proposed temperature selection scheme [[110](#)] does provide better results than the previously accepted geometric selection method.

2.0 SIMULATIONS OF BRANCHING ACTIN NETWORKS SUGGEST LOAD-SHARING MECHANISM DETERMINES FORCE-VELOCITY RELATIONSHIP.

2.1 BACKGROUND

Actin is a globular protein that forms thin filaments inside cells [100]. Actin is termed to be g-actin when in globular form and f-actin when in filament form [126]. Actin filaments are also polarized with a barbed end and a pointed end. The actin filaments grow primarily at their barbed end [100]. Individual actin filaments join together to form densely branched networks [80]. Those branching actin networks are important to many cellular processes. Dendritic actin filament networks can exert considerable force against a load, which is important in cellular processes ranging from cell motility [98] to endocytosis [28]. Understanding the basic process by which an actin network is able to exert force against a load is a fundamental part of understanding a number of cellular processes [99].

Actin-based motility can be thought of as involving three processes: protrusion, contraction and adhesion [41, 74]. Networks of actin filaments push the cell membrane forward in the form of lamellipodia or filopodia. Focal adhesions connect the actin network to the external surface the cell is moving along. New focal adhesions form underneath the extended lamellipodial protrusion. Finally, the cell membrane retracts to the newly established focal adhesions, which are close to the leading edge of the lamellipodium. As the older, more rearward focal adhesions detach, the cell moves forward as it is centered around the new focal adhesions.

Lamellipodia are broad, flat membrane protrusions that are filled with a dense, laminar actin network. The actin network inside the lamellipodium is highly branched [80, 123]

protrusion that is broad and largely flat; the thickness of the lamellipodium ranges from 0.1 to 0.2 μm top to bottom [125]. The lamellipodium is the leading edge of a migrating cell, and the protrusion of the lamellipodium is the considered the first step in the protrusion/contraction model of cell motility [76]. Filopodia are small, spiked protrusions of the membrane at the front edge of a lamellipodium generated by bundles of parallel actin filaments. Filopodia allow cells to sense the surrounding environment [72] and serve as the leading point of cell migration.

Focal adhesions serve as the mechanical connection between a branching actin network inside of a cell and the surrounding environment. Focal adhesions are made up of a variety of proteins, including transmembrane proteins, that can bind to the extracellular matrix (ECM) surrounding the cell [41]. As the lamellipodium moves forward, new focal adhesions are generated inside the advancing lamellipodium. The lamellipodium later contracts, and the focal adhesions limit how far back the membrane can retract. The focal adhesions thereby fix the location of the cell. As focal adhesions age, their connection to the ECM weakens [41] causing older, rearward adhesions to release. This process occurs relatively slowly, on the order of 10-20 minutes. When those rearward adhesions release, the cell stabilizes over the more forward adhesions. The dynamics of this process can be seen in the diagram in figure 2.1 A.

The biochemistry of an actin network growing *in vivo* involves many proteins and a number of pathways [98]. Surprisingly, however, only a few proteins are required to generate a branching actin network capable of generating force [65]. Loosely speaking, the processes required are filament growth, new filament branching, filament capping which stops growth, and filament severing/pointed-end depolymerization to maintain the supply of g-actin. These processes are discussed in more detail below. This simplified biochemistry opens the door to detailed modeling to try to understand the behavior of branching actin networks.

The organization of actin filaments inside the lamellipodium of a motile cell is a branched dendritic array [80, 100]. New filaments branch off of existing filaments at a characteristic angle of $\sim 70^\circ \pm 7^\circ$ [80]. Actin related proteins 2 and 3 create a complex (the Arp2/3 complex) that serves as the branching point for new filaments and generates the characteristic branching angle. WiskottAldrich Syndrome protein (WASp) and related proteins activate

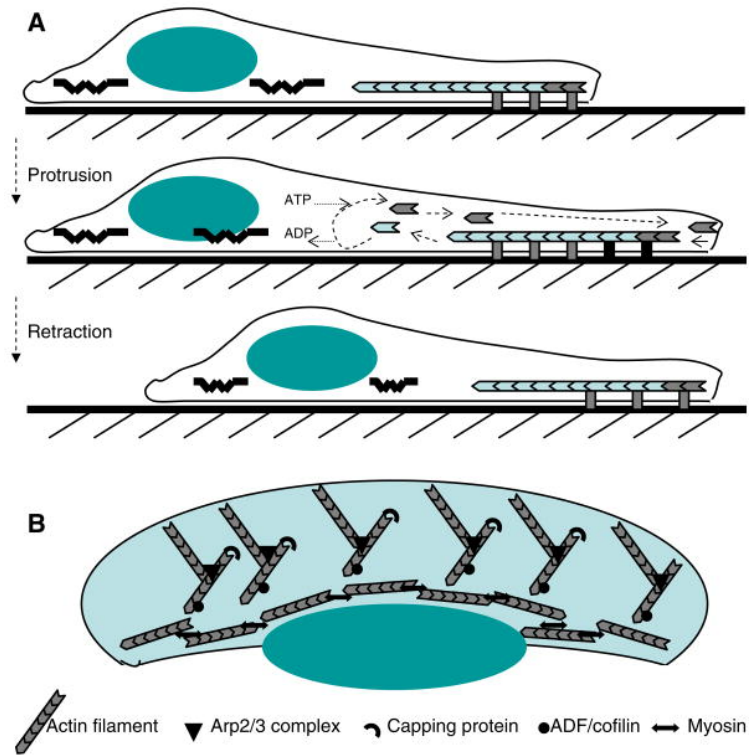


Figure 2.1: This diagram shows some of the basic processes involved in actin-based cell motility. A) shows the dynamics of actin-based lamellipodia protrusion followed by retraction. B) shows the biochemical processes involved in the growth of actin networks. Reprinted from Mogilner 2009 [76] Copyright ©2009 Springer-Verlag. Reprinted with permission. All rights reserved.

the Arp2/3 complex at the membrane to generate new filaments. Filaments grow through the addition of new actin monomers to the barbed end of uncapped filaments. Capping proteins stop the growth of filaments by binding to the filament barbed end preventing new monomers from binding [112]. Finally, the actin network moves the membrane forward by continuing to elongate even after the filament tip has reached the membrane [100]. The three relevant processes are shown in figure 2.1 B.

One way to characterize an actin network’s ability to function as a motor is to measure the network velocity when pushing against a number of different forces. The results of such an experiment are termed the force-velocity relationship or curve. These experiments have been both *in vivo* [52, 73, 101] and *in vitro* using beads [10, 11, 81, 139], lipid vesicles [46, 135] and mechanical levers [69, 89]. Different experiments have led to somewhat contradictory results. There have been two basic shapes for the force-velocity curve observed. One type of curve is a simple convex curve likely indicating a passive response to increasing forces. The other curve features no significant reduction in velocity in response to small forces, a force-independent range, and a rapid reduction in velocity after some critical force. For the purposes of this document, the second type of curve is termed ‘concave’. It is important to note that we are using the terms convex and concave in a qualitative sense and do not mean that the curves fulfill the mathematical definitions. An example of each type of curve can be seen in figure 2.2.

Both types of force-velocity curves have been observed both *in vitro* and *in vivo*. A particularly stark pair of similar experiments with dissimilar results can be seen in [69] and [89]. The first study measured the velocity of an actin network growing against a flexible microfiber [69] which resisted the network with a constant force, and the resulting force-velocity relationship was convex. In another experiment, the actin network grew against the cantilever of an atomic force microscope where the force progressively increased. In that second experiment, the network showed a concave force-velocity relationship [89]. The simulation results presented here indicate that the rate of network adaptation could explain these two results.

The force-velocity relationship of a branching actin network is determined by a number of factors. Those factors include the ability of an individual filament to support a load

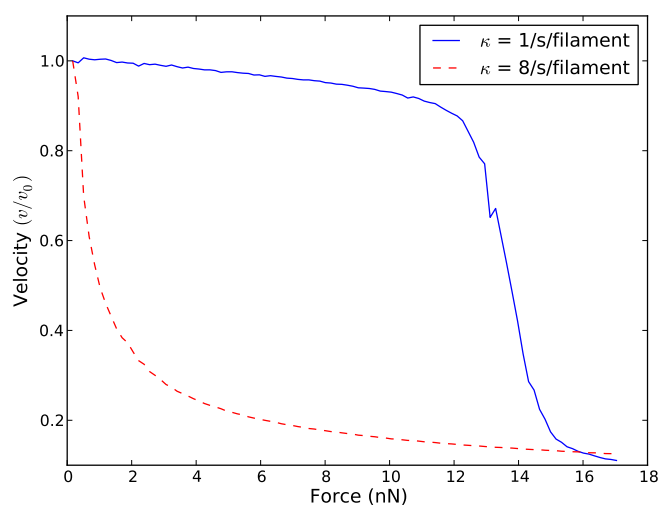


Figure 2.2: Our simulations were able to reproduce both types of force velocity curves. The red dashed line is a good example of a convex force-velocity relationship, and the solid blue line is an example of a concave curve.

[10, 77, 78, 93, 119], the mechanism that allows multiple filaments to distribute the load [114], the number of actively growing filaments at the membrane detached from the membrane [35, 36, 78, 145], and the structure of the actin network itself. Our simulations were designed to try to isolate the impact the actin network geometry and internal structure generated by the simplest actin biochemistry has upon the force-velocity relationship.

There have been a number of models of actin dynamics in the past few decades [76]. The models particularly relevant to this chapter relate to lamellipodial protrusion and the type of experiments listed above. The two primary components in these models are the structure of the actin network and the ability of actin filaments to generate force, moving the membrane forward. The structure of an actin network can be thought of as comprising three factors: the average filament length, the distribution of filament orientations relative to the leading edge, and the spatial distribution of filament tips. Filaments are believed to generate force through rectified thermal motion termed the ‘Brownian ratchet’ [77, 78, 93].

One of the first models of actin filament force generation relied on thermodynamic motion of the load allowing new actin monomers to bind to an existing actin filament [93]. A filament pressing against a load subject to Brownian motion would only be physically in contact with the load for some fraction of the time. At other times, there would be a gap in between the load and the filament tip. When that gap became sufficiently large, an additional actin monomer could bind to the filament tip. That additional monomer binding would move the average position of the filament tip, and therefore the load, forward. When the on-rate for actin monomers binding is sufficiently larger than the unbinding rate, the average growth rate will be positive, and the filament growth will move the leading edge forward. Peskin et al. [93] derived an expression from this growth rate:

$$v(F) = v_0 \exp \left[-\frac{F\delta}{k_B T} \right] \quad (2.1)$$

where F is the force resisting filament growth, v_0 is the free growth velocity, δ is the length of an individual actin monomer, k_B is Boltzmann’s constant, and T is the temperature.

Peskin et al.’s model predicted that the velocity of a growing actin network pushing against a load should be proportional to the size of the load fluctuations, and thus inversely proportional to load size. However, Goldberg and Theriot [48] found that bacteria of vastly

different sizes moved at approximately the same speed. To account for this fact, Mogilner and Oster [77] developed a model based on gaps between the filament and the load arising from filament fluctuations. They argued that filament fluctuations were sufficiently faster than load fluctuations that the two could be treated separately. Singular perturbation theory allowed them to calculate an expression for a number of limiting cases. The cases most relevant to physiological conditions involved short filaments, relative to the filament persistence length, and large forces irrespective of filament length. In those limits, the expression derived by Mogilner and Oster [77] is the same as the expression derived by Peskin et al. [93]. In cells, filaments range in average length from ~ 100 nm to a few microns [96, 132] much shorter than the ~ 17 μm persistence length of actin filaments [88], justifying the short filament assumption.

While the Brownian ratchet describes the behavior of one filament well, a unified model of the behavior of a network of filaments has yet to be accepted. One of the first questions needing to be addressed to develop such a model is how to model the kinetic dynamics of branching and capping. The capping rate has generally been modeled as a first-order rate, i.e. the capping rate is proportional to the number of freely growing filaments. This implicitly assumes that the concentration of capping protein is close enough to constant as to not influence the capping rate. There is more controversy regarding the branching process. The seminal paper on the topic was written by Carlsson [15]. In that paper, Carlsson considered the implications of having a first-order or zeroth-order branching rate. He assumed that each filament was able to exert an equal amount of force and used the Brownian ratchet mechanism to determine the velocity. From that, a zeroth-order branching rate was predicted to lead to a convex force-velocity curve, and the first-order branching was predicted to have a flat force-velocity curve. The first-order branching was able to explain the force-independent range in the force-velocity curve, but some other mechanism would be required to explain the critical force and eventual stalling of the network. Similarly, the zeroth-order branching explained the convex curve, but some other mechanism would be needed to explain the concave curve.

Several papers have added additional detail to the Carlsson model for studying the force-velocity relationship. Schaus and Borisy [114] added spatial detail and studied the performance of a fixed number of filaments when varying the way that the resisting force

was distributed across the filaments. They found that the performance of the network at any given force was vastly changed by the force sharing, but that the qualitative shape of the force-velocity curve remained unchanged. Building on similar work by [61], Schreiber et al. [117] combined first-order branching with a finite spatial restriction. They found that excluded volume effects from filament interactions limited filament density and led to the concave force-velocity relationship. However, they were unable to reproduce the convex force-velocity curve also observed.

Building on the work from the Borisy group [114, 115], Weichsel and Schwarz [137] suggested that the pattern of filament orientations relative to the membrane could explain both force-velocity curves. In the previous work studying the orientation pattern [4, 68, 115], only one stable orientation pattern was seen. That orientation pattern was characterized by two density peaks centered around -35° and 35° . However, in their work, Weichsel and Schwarz found two stable patterns that had different force-velocity profiles. One pattern was the previously observed $-35^\circ/35^\circ$, and the other had peaks at -70° , 0° and 70° . Their model, however, did not include any explicit force generation dynamics by actin filaments. The work presented here was originally conceived to extend their work to include specific force generation by individual actin filaments. A more detailed discussion of actin orientation patterns is included in chapter 3.

The model presented here incorporates features from a number of previous models including [77, 114, 137]. The model suggests that the two force-velocity curves may be explained by the rate at which actin networks recruit new filaments to the leading edge. We also find results relating to the network orientation pattern that explicitly contradict [115] and [137].

2.2 DEFINING THE MODEL

2.2.1 Methods

Actin Model

Each filament was modeled as a point in a two-dimensional plane representing the barbed

Table 2.1: Table of actin network model parameters.

δ	2.7 nm	Length of an actin subunit. [1]
$k_B T$	4.6 pN·nm	Absolute temperature.
λ	200 κ	Default branching rate.
κ	1-8/s	Default capping rate. [112]
v_{free}	100 δ /s	Default velocity. [1]
N_f	200	Average number of free filament barbed ends. $= \lambda/\kappa$.
Θ_{br}	70°	Mean branching angle. [80]
σ_{br}	5°	Branching standard deviation

end of an individual actin filament. The plane was bounded by a hard leading edge in the principle direction of motion (X) and a periodic boundary in the perpendicular (Y) direction. Each filament had three properties: an X-coordinate, a Y-coordinate and an angle of growth θ which was relative to the X-direction. As the filament grew, the X and Y coordinates would change in time, e.g. $\dot{x}_j = v_j \cos(\theta_j)$ for filament j growing at an angle θ_j , while the angle (θ) for each filament did not change. The filaments were limited to having angles in $-90^\circ < \theta < 90^\circ$ as filaments growing against the principle direction of growth (X) are not seen in experiment [74] and would quickly grow too far away from the leading edge to contribute to the network velocity. The default parameters for the model are listed in Table 2.1.

The filaments grew against the force exerted by the leading edge (at point $X=0$). A filament was treated as being in contact with the boundary if the tip of the filament was within one subunit length (δ) from the leading edge—i.e. if adding an additional actin subunit would induce an energy penalty from moving the leading edge. Identifying the filaments that were in contact with the leading edge was performed at the beginning of each time step.

Within the simulations, new filaments were generated by branching from existing filaments. The filaments branched at a constant rate, calculated using zeroth-order Poisson

```

1  Nbranch = -1;                                // Number of branching events counter
2  R = ran1;                                     // Uniform [0,1) random number
3  r = 0;
4  while (r < R)
5      Nbranch++;
6  r += exp(-lambda*dt)*(lambda*dt)**Nbranch/factorial(Nbranch);

```

Figure 2.3: This pseudo-code shows how the Poisson statistics for branching and capping events were calculated.

statistics. The pseudo-code in figure 2.3 shows how the number of branching events was calculated. The particular filament that the new filament branched from was chosen by drawing a random index from the set of filaments not contacting the leading edge. A constant rate of branching new filaments essentially assumes that the rate limiting factor is the concentration of Arp 2/3 [15, 100, 113, 137]. The initial tip of a new branch was placed at a point along the initial filament randomly selected from a uniform distribution extending a distance of 5δ backwards from the tip of the initial filament. The difference between the angle of the branching filament (θ_b) and the angle of the initial filament (θ_i) was drawn from a normal distribution (\mathcal{N}) with mean 70° and variance $\sigma^2 = 25^\circ$, i.e. $|\theta_b - \theta_i| \sim \mathcal{N}(70, 25)$, and the filaments branched in both directions ($\theta_b > \theta_i$ or $\theta_b < \theta_i$) with equal probability [80]. Gaussian random numbers were drawn using the gasdev algorithm in Press et al. [102]. The filaments were capped as a first-order reaction, with first-order Poisson statistics, meaning the capping rate was dependent upon the number of active filaments not in contact with the leading edge. For the capping process, the term 'lambda' was replaced by 'kappa*Nf' in the above pseudo-code where 'kappa' was the capping rate and 'Nf' was the number of filaments not in contact with the leading edge. Once a filament was capped, it was no longer able to grow, branch or exert force upon the leading edge [112]. The first-order statistics were calculated based on the total number of active, uncapped filaments not in contact with the leading edge. At each time step, the number of filaments added and capped during the time were calculated using a uniform [0,1) random number using the ran1 algorithm in Press et

al. [102]. Filaments in contact with the leading edge were neither branched nor capped the same as was done in previous models [14, 114, 115].

In our model, filaments grew at a deterministic, constant rate in the direction defined by θ_j , based on the assumption that the rate of actin subunits adding to the filament is much faster than the rate of the network reorganization. The principle goal of the model was to study the role of the geometric structure of the network and modeled only average growth dynamics. Similar to *in vitro* conditions, the actin monomer concentration was modeled as saturated leading to a constant growth rate of $100 \delta/s$ [1]. The velocity of each individual filament in contact with the leading edge was reduced by the Boltzmann factor derived from the Brownian-ratchet mechanism [77, 78, 93]. The force was shared across all filaments using a modified version of the optimal force sharing from Schaus and Borisy [114]. Each time a filament would add an additional actin subunit, the energy penalty associated with the new actin monomer moving the leading edge forward would be proportional to $F_{tot}\delta \cos(\theta)$, where F_{tot} is the total force applied by the leading edge and θ is the angle between the filament and the normal direction to the leading edge. To model ideal conditions, the force was shared proportionally between all filaments. Thus, the sum of the forces resisting each individual filament was set to add up to F_{tot} . Dividing each force by $\sum_j \cos(\theta_j)$ gives the correct total force (recall filaments are required to have orientation strictly inside $(-90^\circ, 90^\circ)$). The velocity term for each filament i in contact with the leading edge was therefore:

$$v_i = v_0 \exp \left(\frac{-F_{tot}\delta \cos(\theta_i)}{k_B T \sum_j \cos(\theta_j)} \right) \quad (2.2)$$

where the sum j is over all of the filaments in contact with the leading edge and F_{tot} is the total force applied to the system. All other filaments grew at their equilibrium velocity.

At each time step, calculations were done in the following order. First, the location of the leading edge was calculated, which was defined as the largest X coordinate of the active filaments. The filaments which were then within δ of the leading edge were marked as in contact with the leading edge and were no longer able to be capped or branch new filaments. Next, the number of branching and capping events between times t and $t + dt$ was calculated by comparing a uniform random number in $[0,1)$ and the cumulative distribution function

for the associated Poisson distribution (see pseudo-code above). Individual filaments were chosen randomly from the population of filaments not in contact with the leading edge to be capped. Following that, filaments were again randomly chosen to serve as the branching point for new filaments. Then, the normalizing constant was calculated by summing over the filaments in contact with the boundary to get the average force felt by each filament. Finally, the positions of the filaments are advanced by $v_i dt$ where v_i is the velocity of filament i .

All simulations started with independent, identically distributed initial conditions. The simulations started with 200 filaments uniformly distributed in the y direction and uniformly distributed within a box of length 20δ in the X direction, perpendicular to the leading edge, with the constraint that half of the filaments were in the first 10δ and half in the second 10δ . The initial filament orientations were randomly drawn from a uniform distribution on $(-90^\circ, 90^\circ)$. The time step, dt , used in each simulation was 10^{-2} s. The code for the simulations was written in C++ and run on Ubuntu and CentOS servers.

Simulation Details

One thousand simulations were run with independent, identically distributed initial conditions as above. The simulations were averaged at each time step to generate Figure 2.4(A). Figure 2.4(B) was generated by taking the minimum velocity after force was applied. Parameter values were chosen to be $\kappa = 1$ and $\lambda = 200$ for the short time-scale simulations. Those parameter were chosen to emphasize that only convex force-velocity curves were observed on short timescales because they provided the most extreme concave case in our long time-scale simulations.

Each data point from Figure 2.5 was the result of averaging 10 equilibrium simulations with constant force starting from the initial conditions described above. Simulations were run for a total of 10,000 s with a dt of 10^{-2} s and data was sampled every 0.2 s. Reported data was sampled from the second half of the simulation to minimize the influence of initial conditions. The equilibration can be seen in Figure 2.6.

The hysteresis simulations were performed with the same initial conditions. Simulations were run for a total of 7,500 s. The first 2,500 s were run at an initial (low) force f_0 to equilibrate the system to the velocity observed in Figure 2.5. At time 2,500 s, the force was

increased to f_1 . The system was allowed to equilibrate again to a lower velocity. Finally, at time 5,000 s the force was reduced back to f_0 where the velocity rapidly rose before converging back to equilibrium.

Filaments were made to 'stick' to the leading edge by reducing the amount of force felt by a filament close to δ away from the leading edge. While we do not know the exact nature of such an interaction between the filament tips and the load surface, actin tethering has been theoretically proposed [78] and has some experimental evidence [13, 21, 46, 56, 130, 135]. The new expression force felt by each filament was $F_j = F_j^* q(x)$ where F_j^* is the force used in (2.2) and $q(x)$ is a cubic polynomial of the distance between the leading edge and the filament time, x such that:

$$q(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \gamma \\ -2 \left(\frac{\delta-x}{\delta-\gamma} \right)^3 + 3 \left(\frac{\delta-x}{\delta-\gamma} \right)^2 & \text{if } \gamma \leq x \leq \delta \\ 0 & \text{if } x > \delta \end{cases} \quad (2.3)$$

where $\gamma = 0.9\delta$ for our simulations. The effect of (2.3) was to increase the velocity of filaments between $\gamma\delta$ and δ away from the leading edge. Those filaments then grew faster than filaments with the same orientation closer to the leading edge. That prevented filaments from growing slow enough relative to the velocity of the leading edge to become greater than δ away from the leading edge.

Curve Characterization

The implication of Figure 2.5(B) is that force-velocity curves can be characterized by the length of the force-insensitive region, where the velocity reduction is small. Optimally, the location of the rapid velocity reduction, and by proxy the length of the plateau, could be found by finding a spike in the second derivative and the associated inflection point. However, numerical estimates of the second derivative were quite noisy. As an approximation, we characterized the force-velocity curves by their $f_{1/2}$, the force at which the relative velocity was reduced to $\frac{1}{2}$ relative to the low-force velocity (at 850 pN). Twenty constant force simulations were performed for each pair of branching and capping rates ranging from 850 pN to 17 nN. $f_{1/2}$ was estimated by linearly interpolating between the two successive forces where the velocities surrounded $\frac{1}{2}$. When the velocity at 17 nN was greater than one half,

we used the largest force calculated (~ 17 nN) for the purposes of figure 2.11. Likewise, for the purely convex curves with $f_{1/2} < 4$ nN the value 4 nN was used in the graph for clarity.

Angular Density Plots

The density plot in Figures 2.13 and 2.14 were made using a Gaussian kernel density estimator. The bandwidth parameter was fit using the algorithm outlined in [7] because it generally provides better fits for multi-modal data than other bandwidth estimators including plug-in estimators. The algorithm was implemented into the Numpy extension to Python [2, 54, 85, 86]. The number of repeated samples, i.e. angles observed more than once due to filaments persisting across samples or lack of precision in saved data, led to the jagged appearance of the density curves. Other estimators, e.g. histograms or more naive bandwidth selection methods, also led to jagged curves.

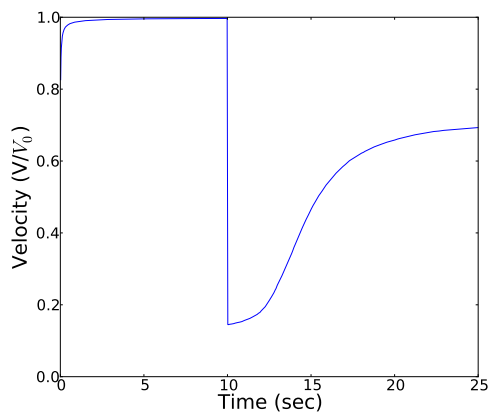
2.3 SIMULATION RESULTS

The simulations were focused on how the collective properties of a branching actin network influence the ability of the network to exert force onto a flat, uniform load barrier. That was done in part for simplicity and stemming from the result in [114] that showed no significant influence of membrane shape upon network performance, in the sense of the efficiency with which the network moved the leading edge.

The first set of simulations tested the temporal response of the velocity of branching actin network against a fixed load force. Figure 2.4(A) shows that upon loading force, the velocity drops almost instantly, which then recovers in a longer timescale (\sim minutes) reaching a value lower than before the application of a load. New filaments reaching the leading edge takes time, and that defines the principal timescale of adaptation. Taking the velocity at the bottom of the initial response to force gives a force-velocity curve for the fast timescale response of the network. As Figure 2.4(B) shows, the force-velocity curve is convex.

Running the simulations for an extended amount of time allows for the study of the equilibrium force velocity relationship. Figure 2.5(A) shows that the model reproduces both convex and concave-type force-velocity curves. The only difference between the sets

A)



B)

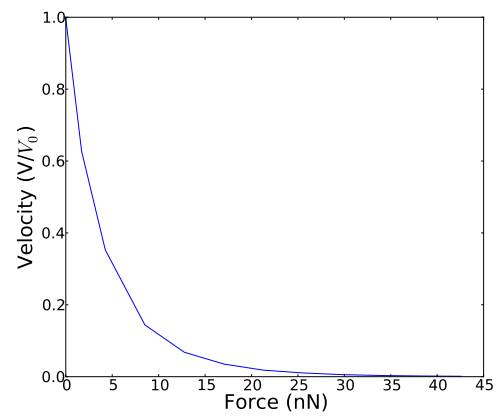
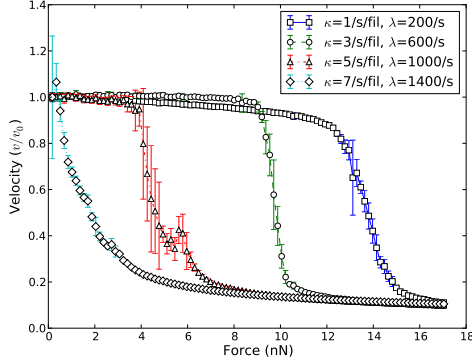


Figure 2.4: The network velocity rapidly drops in response to force application. A) shows a characteristic time trace of the velocity response to force applied at 100 seconds, and B) shows the convex force-velocity curve generated by the initial response of the network to force.

A)



B)

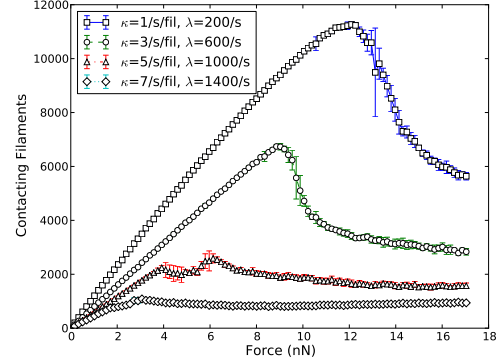


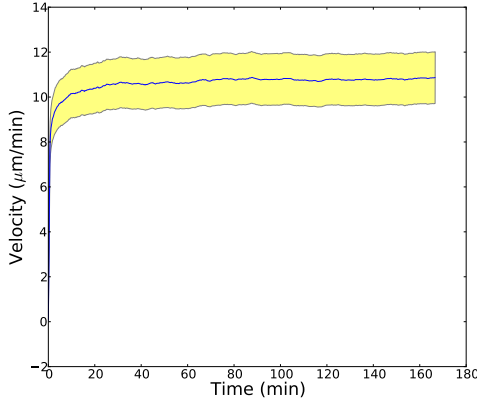
Figure 2.5: Varying the capping rate (κ) over an order of magnitude changes the shape of the force-velocity curve. A) shows the continuous deformation of the force-velocity curve for a few capping rates. B) shows the relationship between force and the equilibrium number of contacts for the same capping rates as in B. The error bars represent the standard deviation estimated from 10 simulations.

of simulations is the absolute value of the capping and branching coefficients. Their ratio, and therefore the average number of filaments, was fixed. We have hypothesized that the network is able to reinforce itself by bringing more filaments to the leading edge.

The hypothesis that the network reinforces itself by filaments growing to the leading edge would suggest that a network with short filament lifetimes, being less likely to grow to the leading edge, would stall at lower forces. Increasing the capping and branching rates simultaneously causes the filaments to grow for shorter periods of time even though the total number of filaments is kept constant. In fact, Figure 2.5(A) shows that by changing the capping rate the force-velocity curve continuously deforms from concave to convex. Fewer filaments reinforce the leading edge leading to lower and lower stall forces. The level of reinforcement can be seen by the number of filaments in contact with the leading edge. Figure 2.5(B) shows that more filaments are recruited to the leading edge when the capping rate is low for the same level of force.

For low-force/low-capping and branching rate cases, the simulations did appear to be

A)



B)

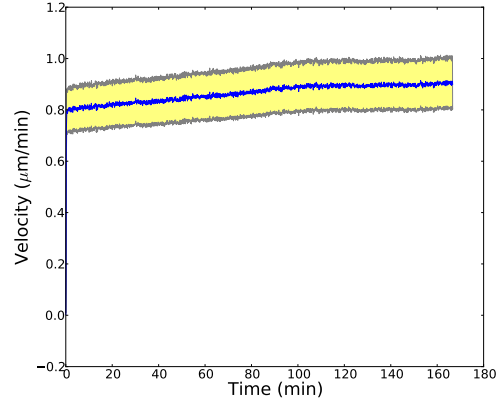


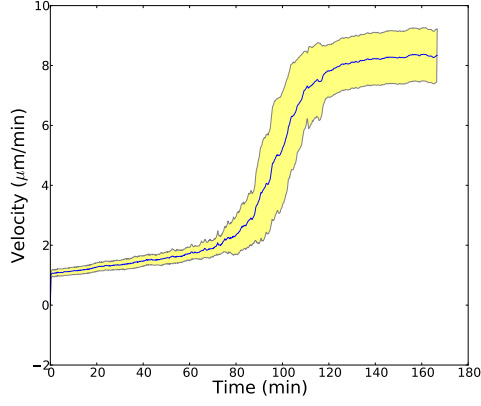
Figure 2.6: For small and large forces, the network velocity rapidly converged. A) $\kappa = 1/\text{s/filament}$ and $F = 2.55 \text{ nN}$ and B) $\kappa = 3/\text{s/filament}$ and $F = 16.2 \text{ nN}$

sampling from an equilibrium distribution. For the simulations where the force-velocity relationship was loosely flat (i.e. $\frac{\partial V}{\partial F} \sim 0$), the velocity converged to an equilibrium quite rapidly, as can be seen in Figure 2.6. The curves shown are the average of 10 simulations with the highlighted region representing the estimated \pm standard deviation.

However, where the force-velocity curve was sharp ($\frac{\partial V}{\partial F} \ll 0$), the velocity converged only at non-physiological time-scales, if at all. This can be seen in Figure 2.7. Because filaments at the leading edge are protected from branching, it is reasonable to hypothesize that filaments will eventually accumulate there to facilitate a loosely flat force-velocity curve. The number of filaments at the leading edge controls the velocity, as can be seen in Figure 2.8, which is how filament accumulation could determine the force-velocity relationship. However, with increasing force or capping and branching rates, the amount of time for that to happen takes arbitrarily long. The adaptation time could explain both the convex and the concave force-velocity curves, because running an experiment for less time than it takes the network to adapt would look identical to the network never adapting.

We were able to estimate the stall force per filament from Figure 2.5(A). We defined the network to be stalled when an increase in force of 170 pN led to a decrease in velocity of

A)



B)

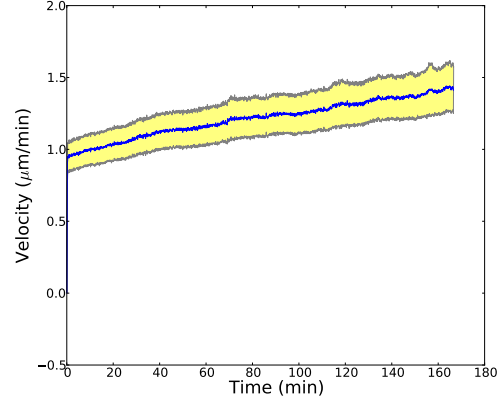
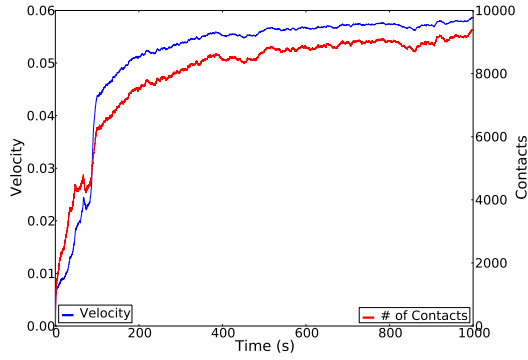


Figure 2.7: For intermediate forces, the actin network velocity does not appear to converge. $\kappa = 3/\text{s}/\text{filament}$ for both and A) $F = 9.4$ and B) $F = 11.1$ nN

A)



B)

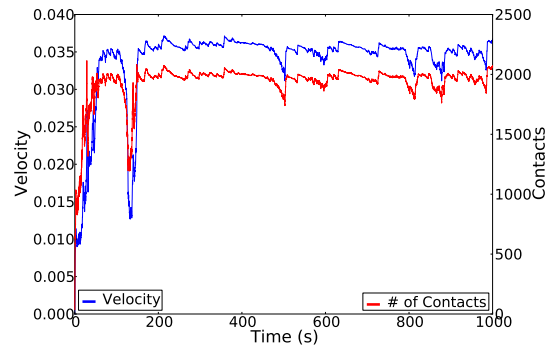


Figure 2.8: Network velocity and the number of contacting filaments were very strongly correlated. A) $\kappa=1$, $\lambda=20$ and $F=8.52$ nN. B) $\kappa=5$, $\lambda = 100$, and $F=3.41$ nN.

less than 1%. Our model could not truly stall due to the lack of depolymerization meaning (2.2) could never truly be zero, but we considered further increases in force failing to further slow the network as indicative of stalling. The stall forces per filament estimated ranged between 0.9 pN and 2.4 pN for $1 \leq \kappa \leq 5/\text{s/filament}$ with a mean of 1.3 pN, which is in close agreement with the reported value of 1.7 ± 0.8 [52]. The reported stall force per filament provides evidence that actin networks use close to optimal force sharing.

An actin network can be visualized as a population of filaments contacting the boundary and another population of filaments trailing the leading edge in reserve. The network reinforces the filaments at the leading edge when trailing filaments grow to reach the leading edge. That remodeling response simply depends on the rate at which trailing filaments are able to catch up to the leading edge. Figure 2.9 is an explanatory diagram showing a hypothetical branching pattern. The first filament is in contact with the boundary and is unable to branch or be capped. The second filament is further back and serves as a source for new filaments. The expected lifetime/length of the filament primarily determines the likelihood that the third filament reaches the boundary. Such a scenario repeated throughout the network would explain the differing rates of adaptation. If the average lifetime of filaments is too short, the network will not sufficiently increase the number of filaments on the leading edge leading to a drastic reduction in overall network velocity.

Figure 2.5(B) shows that the number of filaments in contact with the leading edge increases with increasing force before peaking in all cases. However, the cases with lower capping rates generate a larger increase before peaking and stalling. As can be seen in Figures 2.4, 2.6 and 2.7, the network velocity is equilibrating at multiple time scales. When the longest time-scale is allowed to converge, there is a linear increase in density as the network adapts to larger and larger forces. However, our model indicates that that longest time scale gets exponentially long with increasing force or increased capping/branching rates. When the longest time-scale becomes longer than your simulation or experiment, the network doesn't fully adapt, and you see a convex force-velocity curve.

The accumulation of filaments at the boundary could also explain some of the hysteresis observed in actin networks. As seen in Figure 2.5, the number of filaments in contact with the boundary increases with increasing force. Subjecting the network to a large force and

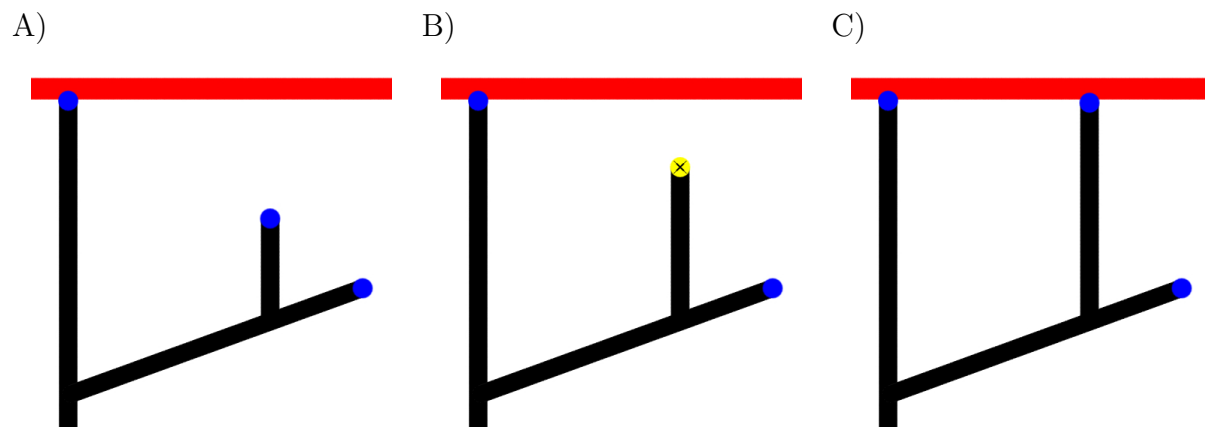


Figure 2.9: This diagram shows how the average lifetime/length of the filaments could influence the number of filaments in contact with the leading edge. Filaments are represented by black lines; the red line is the leading edge; the blue circles are actively growing barbed ends; and the yellow circle represents a barbed end that has been capped. A) A hypothetical scenario involving one filament barbed contacting the leading edge with a second filament growing behind the leading edge. B) When the capping rate is high, the reserve filament is capped (yellow circle) before it reaches the leading edge. C) When the capping rate is low, sufficiently long filaments can grow and contact the leading edge, increasing the leading edge velocity.

subsequently releasing that force should leave excess filaments in contact with the leading edge. Figure 2.10(A) shows the velocity of a simulation where the network pushed against low force for the first third of the run, followed by high force in the middle third, and finally the original low force. That is comparable to the experiment in Parekh et al. [89]. The velocity initially shoots back up in response to the reduced force, but it rapidly decays back to the initial equilibrium force similar to [15]. The reduction in the load force speeds up the growth rate of all the contacting filaments at the leading edge. Due to the angle dependence of the load sharing for individual filaments, the speed-up of growth rates is heterogeneous across the contacting filaments. As a result, some filaments grow faster while staying in contact with the load, and slower growing filaments slide off the leading edge and are capped. Ultimately, the number of contacting filaments relaxes back to the velocity corresponding to the original force and completely loses its memory of the previous loading force. Figure 2.10(B) shows that sustained hysteresis can be realized in the model, if we incorporate a factor that causes the actin filaments to stick to the leading edge. While the exact nature of such an interaction between the filament tips and the load surface is unknown, actin tethering to the load has been theoretically proposed [78] and has some experimental evidence [13, 21, 46, 56, 130, 135].

These results suggest that the actin network remodels itself by changing the number of filaments in contact with the leading edge. That remodeling in turn determines the shape of the force velocity relationship. In particular, it determines the length of the concave portion of the curve. To obtain a more systematic, qualitative understanding of the rate dependence of the actin force-velocity relationship, we performed a coarse parameter search over the capping and branching rates. We used the force at which the velocity drops to 50% of the small load force to characterize the shape of the force velocity curve. Figure 2.11 shows the estimated $f_{1/2}$ values at each parameter value that constitutes the principle prediction of the model: larger capping rates lead to less concave force-velocity curves, and larger branching rates lead to more concave curves.

Another result is that the model qualitatively reproduces the large velocity reduction in response to small forces observed in experiment [101]. Figure 2.12 shows the equilibrium velocity v/v_{free} in response to a 170 pN force. The only way for the network to maintain

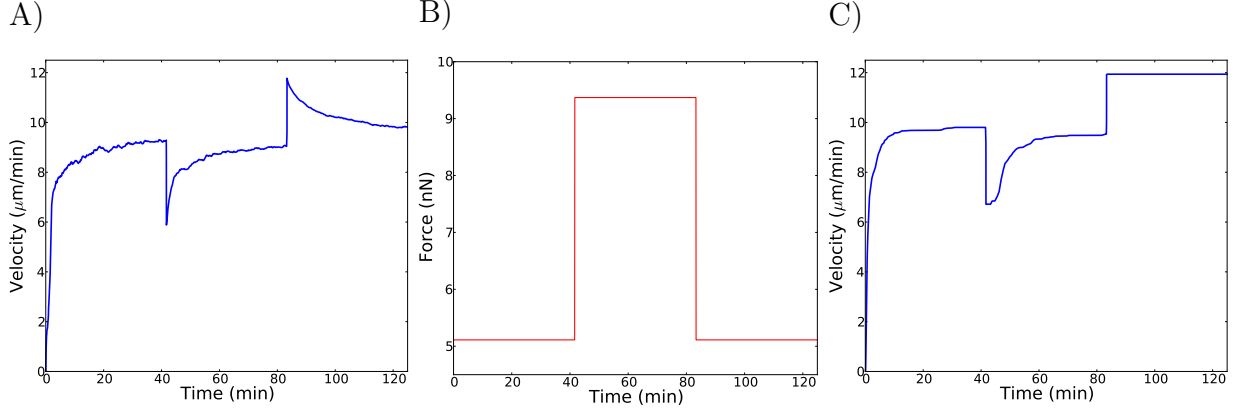


Figure 2.10: Plots of simulations testing for hysteresis in the force-velocity relationship. A) shows a typical simulation result with a transient hysteresis effect. B) shows the time dependent force used in both A and C. C) shows a typical simulation when filaments reaching the leading edge stuck to the leading edge.

velocity with increasing force is to increase the number of filaments at the leading edge. However, when the leading edge is moving close to the speed of a freely growing filament, filaments trailing the leading edge are unable to catch up. Thus, leading edge must be sufficiently slowed down before the force-velocity curve can plateau. Thus, the force-independent velocities must be significantly slower than v_{free} .

The time-scale argument appears to explain all of the results presented here, but it is worthwhile to compare other results from our simulations to those in [137]. The first question is whether or not our results reproduce the orientation patterns seen by Weichsel and Schwartz. That result attempts to reproduce the regular pattern of angles between actin filaments and the membrane as seen in electron micrographs [124]. They see two dominant orientation patterns. At low and high forces, they see a $-70^\circ/0^\circ/70^\circ$ pattern. However, at intermediate forces, they see a dominant $-35^\circ/35^\circ$. The only orientation pattern we were able to observe was the $-70^\circ/0^\circ/70^\circ$ pattern seen in Figure 2.13. In fact, the system rapidly converges to that orientation pattern. To see that, define the two numbers n_0 and n_{35} as the number of filaments with orientation in $(-5^\circ, 5^\circ)$ and $(30^\circ, 40^\circ)$, we can define the orientation

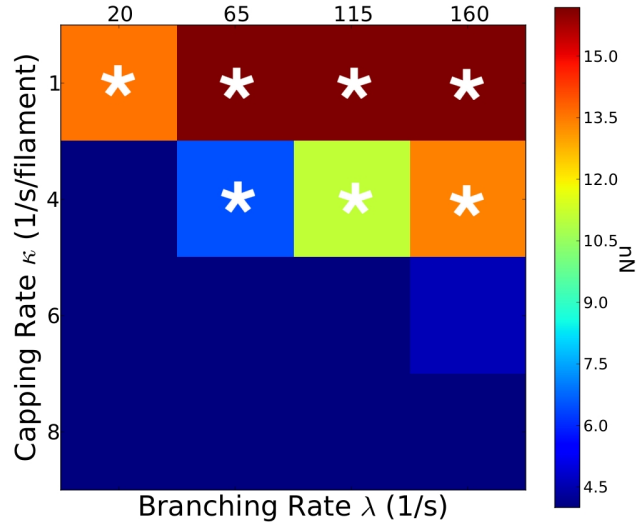


Figure 2.11: Plot of $f_{\frac{1}{2}}$ for a variety of parameters. The color represents the estimated $f_{\frac{1}{2}}$ for a variety of branching and capping rates. The cases with concave force-velocity curves are labeled with white asterisks. Decreasing the capping rate and increasing the branching rate serve to generate more filaments which shifts $f_{\frac{1}{2}}$ rightward, meaning a more concave curve.

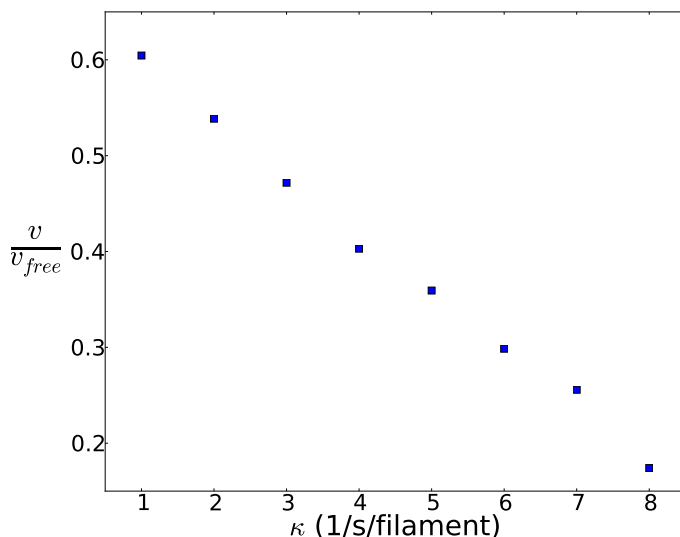


Figure 2.12: Equilibrium velocity reduction in response to a small (~ 170 pN) force by capping rate.

pattern parameter, Dist, from [137]:

$$\text{Dist} = \frac{n0 - n35}{n0 + n35}$$

That parameter is close to -1 for orientation patterns close to $-35^\circ/35^\circ$ and close to +1 for orientation patterns close to $-70^\circ/0^\circ/70^\circ$. Plotting that parameter versus time as in Figure 2.13 shows rapid convergence to $-70^\circ/0^\circ/70^\circ$.

Next, we tried to reproduce the orientation patterns Weichsel and Schwartz observed by modifying our model. In their model, new filaments were only able to branch from existing filaments whose tips were within 2δ of the leading edge. By adding that restriction to our model, we were able to reproduce their results as can be seen in Figure 2.14. At very low and very high forces, the $-70^\circ/0^\circ/70^\circ$ pattern dominates, but the $-35^\circ/35^\circ$ pattern is seen at intermediate forces. The exact shape of the distributions is different because of our differing boundary value specifications. However, this effect goes away by increasing the branching range even to 4δ .

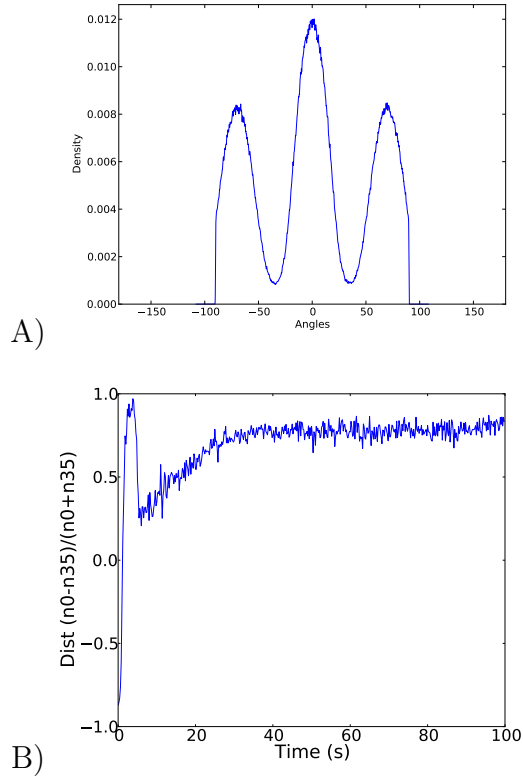
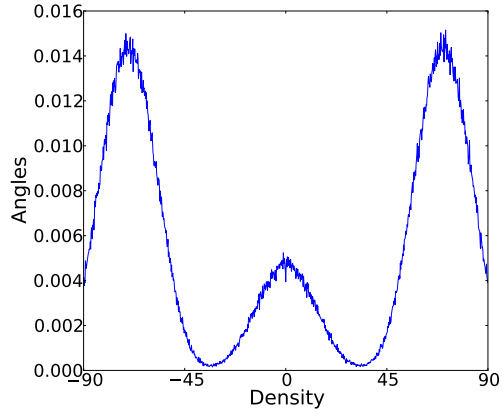
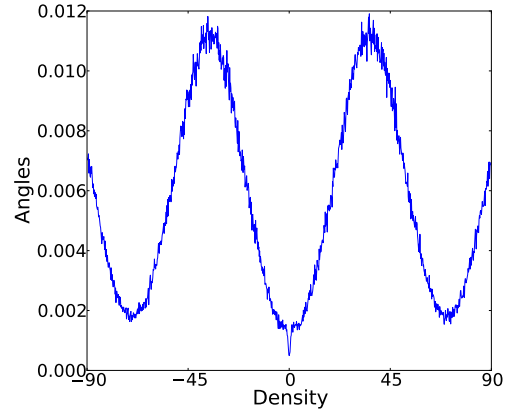


Figure 2.13: The $-70^\circ/0^\circ/70^\circ$ pattern was dominant in our simulations. A) shows a representative orientation pattern observed in our standard simulations. B) uses the orientation pattern parameter described in the text to show rapid convergence to $-70^\circ/0^\circ/70^\circ$.

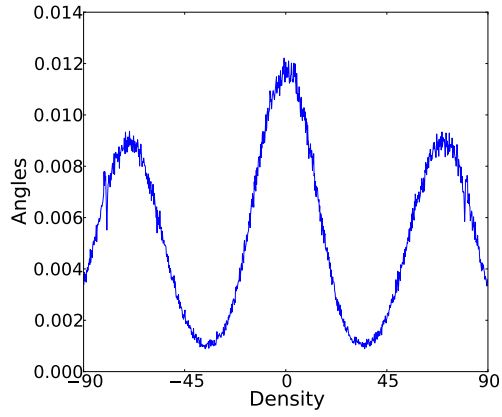
A)



B)



C)



D)

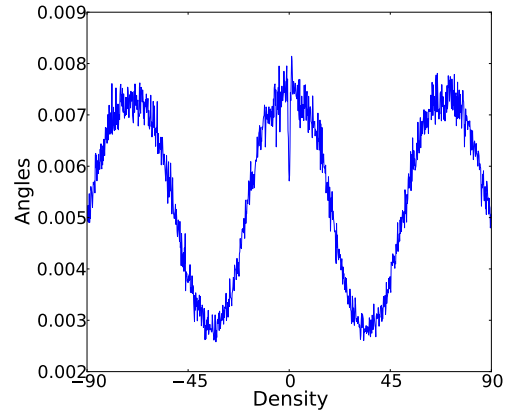


Figure 2.14: The angular density varied significantly when the branching zone was restricted. A-C) Angular density plots for branching restricted to 2δ , $\kappa = 3$, $\lambda = 600$, and A) $F = 0.34$ nN, B) $F = 6.81$ nN, C) $F = 13.63$ nN. D) Angular density for branching restricted to 4δ , $\kappa = 3$, $\lambda = 600$ and $F = 6.81$ nN.

That change in orientation pattern between limiting branching to 2δ and 4δ is only a function of filaments that are very close to the leading edge. Figure 2.15 shows a high density of filaments very close to the leading edge and a slow decrease in density behind that for moderate forces. For higher forces, almost all of the density is immediately behind the leading edge. A similar effect was seen in [114]. After about 2δ , the change in filament density is very gradual, which would explain the lack of effect of branching restriction on the orientation pattern.

Given that the size of the branching range changes the network orientation pattern, it is worthwhile to test whether or not that parameter change would affect the force-velocity relationship. Figure 2.16 shows no qualitative difference for any of the branching ranges we tried. This is not surprising given that more than about 2δ away from the leading edge the filament density does not change very much.

2.4 DISCUSSION

The model outlined here suggests a simple mechanism could explain multiple actin network force experiments. The question of how to explain the two distinct force-velocity curves observed experimentally has been a focus of a number of theoretical studies recently [61, 117, 137]. Our results indicate that the time-scale of network reorganization could explain the two observed force-velocity curves. When the network is given sufficient time between force application and velocity measurement, the force-velocity relationship is relatively flat. However, the initial response of the network to force application is always convex. The accumulation of filaments at the leading edge observed in our simulations also could explain the hysteresis effects seen in [89]. Finally, our results on filament orientation cast doubt on the results in [115, 137] explaining the dominance of the $-35^\circ/35^\circ$. Their result only holds when new filament branching is restricted to a small, likely unphysical, zone near the leading edge. Unfortunately, our results here do not yet provide an alternate explanation.

The number of active actin filaments at the boundary determining the force-velocity relationship has previously been studied [14, 15, 89]. In particular, Carlsson [15] suggested

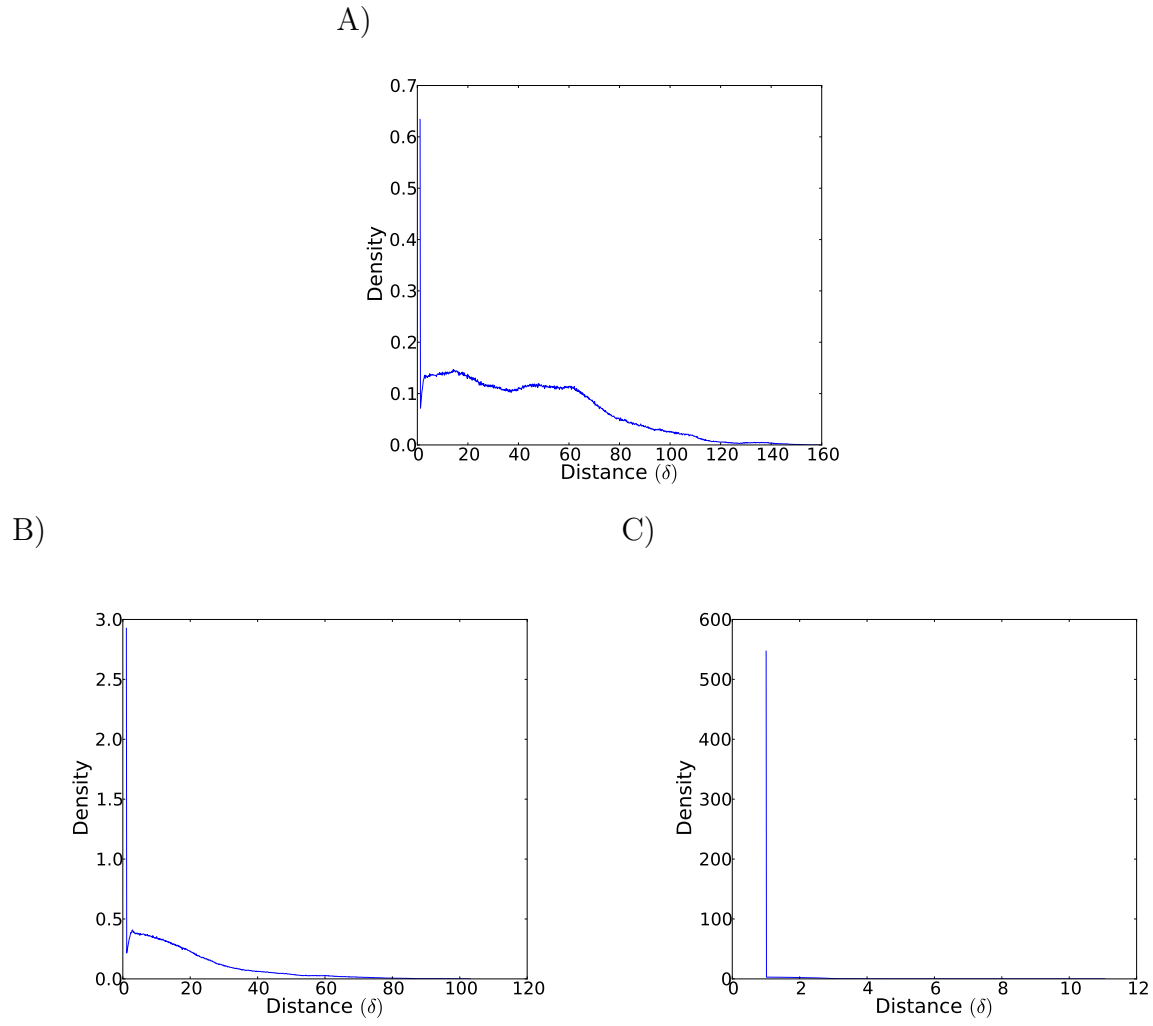


Figure 2.15: Filament density accumulated near the leading edge with a slow decline behind that for $\kappa=3$ and $\lambda=600$. A) $F = 3.42$ nN. B) $F = 8.52$ nN. C) $F = 11.93$ nN.

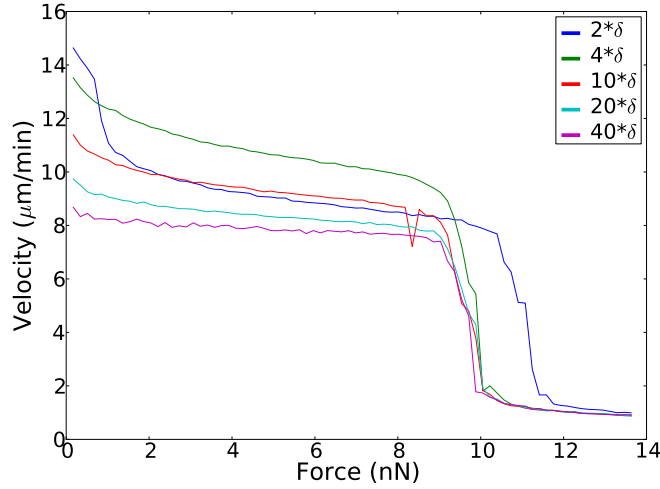


Figure 2.16: The force-velocity curves when new filament branching is restricted to a few different ranges are qualitatively similar to the unrestricted case for $\kappa = 3$ and $\lambda = 600$. The curves are reported in absolute velocities.

that actin networks with auto-catalytic branching kinetics would have an unbounded accumulation of filaments at the boundary leading to completely force independent velocities. However, some external feedback would be necessary to limit the total density to produce the inevitable reduction in velocity at high forces that we saw in Figure 2.12. His model also does not predict the reduction of velocity in response to small forces. Although branching events are auto-catalytic as demonstrated in experiments [49], there will only be a finite supply of Arp2/3 in the cell leading to a saturation of the branching rate. Our model explains both force-velocity curves with zeroth-order branching, i.e. a non-auto-catalytic model. We believe that our results suggest that the zeroth-order statistics is the more reasonable model. In particular, zeroth-order branching could be explained by Arp2/3 availability being the rate-limiting step.

Much like our model, Carlsson's model also predicts a transient hysteresis effect where additional filaments at the leading edge causes a spike in velocity when the resisting force is reduced. However, experiments have observed a sustained hysteresis effect [89]. Adding

a term which causes filaments to stick to the membrane, like we did in Figure 2.10, would allow either model to explain the hysteresis effect equivalently. Thus, the hysteresis results cannot be used to differentiate our model from his.

Our model does not preclude any other negative feedback mechanism that limits the density of actin filaments. It is likely that increased filament density would lead to excluded-volume effects at large forces [117], which would limit the filament density at the leading edge. However, any external negative feedback mechanism would only limit the length of the force-insensitive region of the force-velocity curve. It is important to note that, due to the exponential term in equation (2.2), even a relatively small change in filament density would lead to a large change in velocity. A doubling of the number of filaments, N , would lead to an approximately $\exp\left[\frac{F\delta}{2Nk_B T}\right]$ -fold increase in the velocity.

Figure 2.5 should be qualitatively reproducible in experiment and could serve as an excellent test of several hypotheses for branching-actin networks. The branching and capping rates could be manipulated by changing the concentration of the relevant proteins as in Cameron et al. [12]. That experiment could be performed both with constant force [69] or with a constantly increasing force [89]. Changing the concentrations of Arp 2/3 and capping protein simultaneously would isolate the average lifetime of filaments and be the most direct test of our predictions. Showing that the stall force of an actin network is dependent upon the Arp 2/3 and capping protein concentrations would be strong evidence for both the hypothesis that actin density controls the force-velocity relationship and an Arp 2/3 concentration-dependent branching rate.

The majority of the simulations were performed with constant force. However, this model is relevant to both constant and non-constant force because it requires no equilibrium assumptions. The model showed a relaxation time before the network reached an equilibrium velocity. Changing the force slower than this relaxation time would allow the network to continuously adapt to the increased forces and show strong hysteresis effects as in [89]. Increasing the force substantially faster than the relaxation time would not allow the network to restructure itself leading to results similar to the instantaneous application of force as in figure 2.4. An experimentally observed value for this relaxation time could be found, and further experiments manipulating the force more slowly than the observed relaxation time

should show these hysteresis effects.

The sharp bifurcation in behavior occurring when the branching range increases between 2δ and 4δ stems from the accumulation of filaments at the leading edge as seen in figure 2.15. It is important to note that no physical basis has been given for the choice of 2δ . That parameter choice originates in Schaus and Borisy [115]. In that model, branching was restricted to a small zone, and capping was restricted to outside that zone. There is evidence that new filament branching occurs close to the leading edge [4, 100], where Arp2/3 is more likely to be activated, but there is not any experimental evidence giving a precise estimate of how close to the leading edge branching occurs. Similarly, interactions with the leading edge could occlude filament capping. However, there is no reason to believe that the two areas are mutually exclusive. Without that assumption, Schaus and Borisy [115] would have likely have seen different results.

Our model implements a load-sharing mechanism where the contacting filaments collectively share the load across the leading edge similar to [114]. That is, filament growth is only resisted by a fraction of the total load force. In the context of our model, our results are likely not dependent upon the exact load-sharing mechanism. However, some amount of load-sharing is valid as long as the hypotheses for the Brownian ratchet mechanism hold. The Brownian ratchet mechanism assumes that the thermal fluctuations between the filament tip and the load surface are significantly faster than the addition of new actin monomers [77, 78]. For a filament to grow, fluctuations must be large enough for a new monomer to fit in the gap between the tip and the load. Smaller and or slower fluctuations reduce the efficiency of the mechanism. Experiments have demonstrated that reducing thermal fluctuations by lowering the temperature strongly hinders the efficiency of filament growth [119]. That provides evidence for the Brownian ratchet mechanism. The separation of time scales underlying the mechanism also implies that the load force felt by each contacting filament is averaged over many fluctuations. Those fluctuations are also likely influenced by the density of filaments resisting the membrane. Consequently, only the partial load force shared across filaments would dictate the network growth rate.

The simple physical model shown here gives insight into the behavior of actin networks growing against a load. In particular, the network velocity dependence upon the number

of filaments pushing against the leading edge provides a simple mechanical mechanism to explain a number of experimental effects. Further investigation into actin network properties, both physical and biochemical, that determine how many active filaments a network is able to recruit to the leading edge will further the understanding of actin network growth.

3.0 ACTIN ORIENTATION DENSITY EQUATIONS SELECT FOR A UNIQUE, STABLE EQUILIBRIUM.

3.1 INTRODUCTION

As mentioned in previous chapters, actin filaments form densely branched networks in a motile cell's lamellipodium [80]. Any individual filament can be characterized by the angle between its orientation, and the normal direction of the membrane. One obvious question is whether or not these angles form any regular pattern. While the question has not been extensively studied, there is some experimental evidence that the networks indeed organize into regular patterns relative to the cell membrane [68, 138]. A few models have been proposed to explain the existence of such patterns [68, 115, 137]. While these models have been numerically solved, there has been no rigorous work proving the existence, uniqueness or stability of these solutions. This chapter presents a few results that characterize the solutions to two equations modeling the angular density of branching actin networks.

All of the models proposed to explain the orientation distribution have used a continuum approximation. There is some question as to whether or not ignoring the stochastic fluctuations of actin networks is justified [104]. However, none of the models make specific predictions about the kinetics of network organization, and there is currently no evidence that correlations between filaments lead to changes in the equilibrium orientation pattern. For the rest of this chapter, we will assume the approximation is justified and focus on long-time equilibrium behavior.

Some of first few models to study orientation patterns in actin and similar networks studied the existence and persistence of peaks in the orientation pattern [43, 75, 79]. The analysis was based on Fourier series and small perturbations which greatly limited their

generality. Their analysis led to the qualitative result that peaked orientation patterns are likely to be observed. Similar methods have been used on models of orientation and space [8, 31]. Stability analysis has also been done on similar models, termed “ring models” in the neuroscience literature [6, 140].

The first model we consider here was proposed by Maly and Borisy [68]. Their insight was that if filaments were capped at different rates based on the filament orientation, filament branching and capping could generate stable orientation patterns. The model they proposed takes into account branching and capping explicitly and filament growth implicitly. New filaments branch off of existing filaments at a characteristic angle $\sim 70^\circ$ with some variance around that. We can write out the branching kernel as a probability of a mother filament with angle θ_M having a branched daughter filament with angle θ_D :

$$\mathcal{B}(\phi) = P(\theta_D = \theta_M - \phi) \quad (3.1)$$

An explanatory diagram can be seen in figure 3.1. Adding up the contribution of all filaments with density $u(\phi)$ gives the total branching rate at angle θ :

$$\text{BR}(\theta) \propto \int \mathcal{B}(\theta - \phi)u(\phi) d\phi \quad (3.2)$$

The limits for the above equation were left off intentionally as different models used different limits. The Maly and Borisy model only considered filaments growing faster than the leading edge, i.e. filaments with orientation $|\theta| \leq \theta_{\text{crit}} = \arccos \left[\frac{v}{v_{\text{max}}} \right]$ where $\frac{v}{v_{\text{max}}}$ is the velocity of the leading edge relative to the maximum velocity of filament growth. When combined with the total branching rate λ , the Maly and Borisy branching function was:

$$\text{BR}(\theta) = \lambda \int_{-\theta_{\text{crit}}}^{\theta_{\text{crit}}} \mathcal{B}(\theta - \phi)u(\phi) d\phi \quad (3.3)$$

They also proposed that the capping rate was proportional to the amount of time the filament would be not in contact with the leading edge, much like [78], but used the capping function $\frac{1}{\cos(\theta)}$. Combining the two terms gives the full equation:

$$\dot{u}(\theta, t) = \lambda \int_{-\theta_{\text{crit}}}^{\theta_{\text{crit}}} \mathcal{B}(\theta - \phi)u(\phi) d\phi - \frac{u(\theta, t)}{\cos(\theta)} \quad (3.4)$$

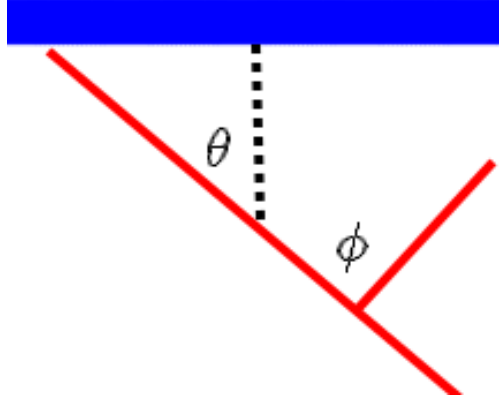


Figure 3.1: Diagram for filament branching. The blue line is the leading edge, and the dotted line is the inward normal to the leading edge. The network grows in the direction of the outward normal to the leading edge. The two red lines are a mother and daughter filament respectively. ϕ is the angle between the mother filament and the daughter filament.

where \dot{u} indicates the time derivative. The capping rate they chose, $\frac{1}{\cos(\theta)}$, was derived from an argument about how much time a filament would spend physically in contact with the leading edge. The equation is defined on $(-\theta_{\text{crit}}, \theta_{\text{crit}}) \times \mathbb{R}^+$ with absorbing boundary conditions.

Maly and Borisy performed two analyses on (3.4). The first analysis was to approximate solutions of (3.4) by solving the equation for two points in orientation space. Solutions to the two-point approximation supported the argument that the equation selected for a unique orientation ‘type’ that grew exponentially at the fastest rate. The second analysis was to use numerical quadrature [5] to approximate the eigenfunctions of the right-hand side of (3.4). However, the existence and uniqueness of the eigenfunction solutions were never rigorously shown. They explained their results by using an evolutionary selection metaphor. In this chapter, I show that a version of (3.4) with stricter hypotheses on the capping rate uniquely selects for a most ‘fit’ orientation pattern with a fitness function defined on the unit ball of orientation functions.

A very similar model was proposed by Weichsel and Schwartz [137] to explain both the orientation patterns and the force-velocity curve (see the previous chapter). There were two primary differences between their model and the Maly and Borisy model. First, orientations were defined on the entire circle, S^1 . The effect of filaments growing slower than the leading edge was incorporated into the capping rate. The second difference was to solve the problem that solutions to (3.4) appear to exponentially diverge was to normalize the branching rate so that the total branching rate was equal to λ . The Weichsel and Schwartz model was:

$$\dot{u}(\theta, t) = \frac{\lambda}{\int_{S^1} u(\phi, t) d\phi} \int_{S^1} \mathcal{B}(\theta - \phi) u(\phi, t) d\phi - \kappa(\theta) u(\theta, t) \quad (3.5)$$

where the capping rate is proportional to the difference between the velocity of the leading edge and a filament with a given orientation:

$$\kappa(\theta) = k + c(v_{LE} - v_0 \cos(\theta))^+ \quad (3.6)$$

where v_{LE} is the velocity of the leading edge, v_0 is the rate of filament growth, and c is a constant.

Weichsel and Schwartz performed the same two analyses as in the Maly and Borisy paper. They found that, for certain parameters, there were two stable equilibria in the two-point approximation to (3.5). Numerical simulations also seemed to support that hypothesis. Finally, they used numerical techniques to calculate the equilibrium distributions. The calculated equilibrium distributions seemed to agree with the discrete approximation results, but no bistable results were reported. The results in this chapter explicitly contradict the bistable hypothesis but show some evidence for the stability of a unique, positive equilibrium.

The techniques here also would go a long way towards solving two problems suggested in [18]. In that paper, analysis of a type of nonlocal diffusion equation proceeded using the Fourier transform. For fairly obvious reasons, that technique was not available for the analysis presented here. At the end of that paper, they presented several open problems for related equations that were not solvable by Fourier analysis. In particular, corollaries of theorem 1 demonstrates the asymptotic behavior for kernels with non unit integral. The techniques presented in this chapter should be generalizable to unbounded domains and provide satisfactory answers to those questions.

3.1.1 Definitions and Assumptions

As mentioned above, this chapter is studying the properties of two related nonlocal diffusion equations describing branching actin networks. The equations are a continuum approximation in the angular density. We will describe the two equations as first-order and zeroth-order relating to the order of the branching rate. The first-order branching is such that the branching rate $\text{BR} \propto \int u(\theta) d\theta$ and zeroth-order so that $\text{BR} \propto 1$.

The first-order branching equation is:

$$\dot{u}(\theta, t) = \lambda(\mathcal{B} \star u)(\theta, t) - \kappa(\theta)u(\theta, t) \quad (3.7)$$

where λ is the branching rate, \mathcal{B} is the branching kernel, and $\kappa(\theta)$ is the angularly dependent capping function. We define the linear operator, A_λ , to be the function of $u(\theta)$ on the right hand side. We are principally interested in positive, integrable solutions to the Cauchy problem with initial condition $u_0 \in L^1(S^1)$. The equation is formally defined on $u(\theta, t) \in L^1(S^1) \times \mathbb{R}^+$ where S^1 is the circle. It is straightforward to show that solutions with positive initial data remain positive. An interesting question going forward would be to extend the analysis here to a more general setting, such as the space of countably-additive measures. In actuality, we are interested in angular densities with finite mass, so the requirement that the density be absolutely continuous with the Lebesgue measure is a simplification for ease of analysis.

Due to the fact that first-order branching usually leads to pathological asymptotics, the zeroth-order branching equation was proposed. It was hypothesized that a more realistic model involves a zeroth-order (constant) branching rate. The resulting equation would be the nonlinear nonlocal diffusion equation:

$$\begin{aligned} \dot{u}(\theta, t) &= \frac{\lambda_0}{U(t)}(\mathcal{B} \star u)(\theta, t) - \kappa(\theta)u(\theta, t) \\ U(t) &= \int_{S^1} u(\omega, t) d\omega \end{aligned} \quad (3.8)$$

where λ_0 is the total branching rate, $U(t)$ is the total density of filaments at time t . For future reference, define the non-linear operator G to be the function of $u(\theta)$ on the right hand side of the above equation. The branching rate is zeroth-order in the sense that:

$$\int \text{BR} = \int \frac{\lambda_0}{\int u(\phi) d\phi} (\mathcal{B} \star u)(\theta) d\theta = \lambda_0 \int \mathcal{B}(\theta) d\theta \frac{\int u(\theta) d\theta}{\int u(\phi) d\phi} = \lambda_0 \quad (3.9)$$

The non-linearity from the normalizing term introduces greater difficulty in analyzing the asymptotic behavior. As such, this chapter contains a conjectured result with numerical evidence backing up the conjecture.

Assumptions on \mathcal{B} and κ :

1. \mathcal{B} is real, positive, symmetric C^2 function with $\|\mathcal{B}\|_1 = 1$.
2. κ is a real, strictly positive, symmetric C^2 function

The assumptions are likely stronger than necessary, but generalizing the problem is a question for further study. They also do not exactly capture the dynamics for either paper. The first paper, by Maly and Borisy [68], would require an infinite capping rate. However, as that is likely unphysical, the equations at hand should be sufficient. For Weichsel and Schwarz [137], the capping rate they used is continuous but not differentiable. The primary role of the C^2 hypothesis is to ensure compactness, and weaker hypotheses should be quite feasible. However, the smoothness hypothesis is merely technical and should have no effect on the interpretation of the results presented here.

3.1.2 Survey of Results

The most important result of this chapter is a Perron-Frobenius type theorem that characterizes the dynamics of the first-order branching equation. Without loss of generality, we will assume $\lambda = 1$ and refer to A_1 as A during the proof of theorem 1. The exact result is:

Theorem 1. *A has a principal, largest eigenvalue μ_0 with a positive, integrable, bounded eigenfunction $\phi(\theta)$. Moreover, μ_0 is an isolated element of the spectrum $\sigma(A)$.*

That result is proven with the following steps:

1. A is bounded and self-adjoint over L^2 and therefore has no residual spectrum over L^2 .

2. The operator $(\mathcal{B} \star u)(\theta)$ is compact.
3. All elements of the spectrum outside of $[-\sup \kappa, -\inf \kappa]$ are either eigenvalues or in the residual spectrum for all L^p spaces with $1 \leq p \leq \infty$.
4. The spectrum of A over L^2 outside of $[-\sup \kappa, -\inf \kappa]$ has no elements in the residual spectrum and therefore is equal to the spectrum of A over L^1 .
5. The spectrum of A over L^1 is real.
6. $\sup_{\|u\|_2=1} \langle Au, u \rangle$ is either an eigenvalue or an element of the continuous spectrum.
7. The largest element of the spectrum is greater than $-\inf \kappa$.
8. Eigenvalues do not accumulate outside of $[-\sup \kappa, -\inf \kappa]$. That implies the principal eigenvalue is isolated.
9. The principal eigenvalue is simple and has a positive, bounded eigenfunction.

The second first-order result shows that solutions converge to zero or diverge to infinity for almost all parameters:

Theorem 2. *For a given \mathcal{B} and κ , there exists precisely one branching rate λ such that (3.7) has a non-trivial equilibrium.*

Theorem 2 is shown through the following steps:

1. Show that μ_0 has the same sign as the leading eigenvalue of the associated operator $\lambda \frac{(\mathcal{B} \star u)(\theta)}{\kappa(\theta)} - 1$ with the following two steps:
 - a. Show connection between the inner product spaces associated with A and $\frac{A}{\kappa(\theta)}$.
 - b. Use the supremum to show the leading eigenvalues must have the same sign.
2. Simple algebra shows the equation $\lambda \mu^* = 1$ is only solvable for one λ since $\mu^* \neq 0$ where μ^* is the principal eigenvalue of $\frac{(\mathcal{B} \star u)(\theta)}{\kappa(\theta)}$.

Similar equations have been shown to be dynamics along an energy surface $E(u)$, i.e. $\dot{u} = -DE(u)$ where D is the Frechét derivative. In fact, we can write (3.7) as the L^2 gradient along the energy surface:

$$E(u) = \frac{1}{4} \iint_{S^1 S^1} \mathcal{B}(\theta - \omega) (u(\theta) - u(\omega))^2 d\theta d\omega + \frac{1}{2} \int_{S^1} (\kappa - 1) u(\theta)^2 d\theta \quad (3.10)$$

A straightforward calculation shows that $E(u) = -\langle Au, u \rangle / 2$. Heuristically, we can interpret (3.10) by the first term penalizing non-uniformity in solutions and the second term penalizing

density accumulation where the capping rate is high. By theorem 1, the function that minimizes the energy, E , is the principal eigenfunction which dominates the dynamics of (3.7). That fits in nicely with the fitness concepts espoused in [68, 115, 137]. We can define the fitness of an orientation as $F = -E$. That leads to the obvious corollary:

Corollary 1. *The dynamics of (3.7) are dominated by the most fit orientation pattern as defined by $\sup_{\|u\|_2=1} F = -\inf_{\|u\|_2=1} E$ from (3.10). Moreover, the sign of the principal eigenvalue is the same as the sign of the fitness for that eigenfunction.*

Proof. By the proof of theorem 1, the principal eigenvalue solves the maximization:

$$\sup_{\|u\|_2=1} \langle Au, u \rangle. \text{ Straightforward calculation gives } F(u) = \langle Au, u \rangle. \quad \square$$

The principal result for zeroth-order branching is considerably weaker:

Theorem 3. *There exists at most countable equilibria to the equation (3.8). Only one equilibrium is positive. That equilibrium uniquely solves a related maximization problem.*

Since solutions with positive initial data remain positive, there is only one possible equilibrium for positive initial data.

Existence and uniqueness of solutions to (3.8) can be shown by adapting the Picard-Lindelöf theorem to the Banach space L^1 . Global existence for all $t \geq 0$ can be shown by observing that solutions with positive, integrable initial data stay in the closed set:

$$\mathcal{U} := \{u \in L^1 : u(\theta) \geq 0 \text{ and } 0 < c \leq \|u\|_1 \leq c' < \infty\}$$

$G(u)$ is uniformly Lipschitz on \mathcal{U} . It is obvious that solutions with positive initial data remain positive. Simply observe for any $u(\theta, t) = 0$ where $u(\theta, t) \geq 0$:

$$\dot{u}(\theta, t) = (\mathcal{B} \star u)(\theta, t) - \kappa(\theta)u(\theta, t) = (\mathcal{B} \star u)(\theta, t) \geq 0 \quad (3.11)$$

It is also straightforward to show that there exists c and c' for the definition of \mathcal{U} . First, observe that

$$\frac{\partial}{\partial t} \int u(\theta, t) d\theta = \lambda_0 - \int \kappa(\theta)u(\theta, t) d\theta \quad (3.12)$$

The mean value theorem gives:

$$(\inf \kappa) \int u(\theta, t) d\theta \leq \int \kappa(\theta)u(\theta, t) d\theta \leq (\sup \kappa) \int u(\theta, t) d\theta \quad (3.13)$$

We can then write out explicit expressions for c and c' :

$$c = \min \left\{ \int u(\theta, 0) d\theta, \frac{\lambda_0}{\sup \kappa} \right\} \quad c' = \max \left\{ \int u(\theta, 0) d\theta, \frac{\lambda_0}{\inf \kappa} \right\} \quad (3.14)$$

All that remains is to show that G is uniformly Lipschitz on \mathcal{U} . The uniform lipschitz result comes from the straightforward result that there exists C such that $\|DG_u\|_{op} < C$ for all $u \in \mathcal{U}$. It is easy to see that a coarse estimate for C is:

$$\|DG_u\|_{op} \leq \frac{2}{c} + \|\kappa\|_{\infty}$$

Finally, simulation results and a pertrubation expansion are presented in support of the following conjecture:

Conjecture 1. *The unique, positive equilibrium of (3.8) is locally-stable on the set \mathcal{K} :*

$$\mathcal{K} := \{u \in L^1(S^1) | u(\theta) \geq 0 \wedge \int u(\theta) d\theta > 0\}$$

3.2 PROOF OF FIRST-ORDER THEOREM

The proof of theorem 1 will be broken down into the steps outlined above and proven in pieces. First, we will show that the operator is well-defined on all of the relevant L^p spaces. That immediately leads to the result that A has no residual spectrum on L^2 .

Step 1 *A is bounded and self-adjoint over L^2 and therefore has no residual spectrum over L^2 .*

Proof. Since $\nu(S^1)$ is finite, where ν is the Lebesgue measure, we have the series of inclusions $L^q \subset L^p$ where $1 \leq p < q \leq \infty$. This can be seen by Hölder's inequality:

$$\int_{S^1} |f(\theta)|^p d\theta \leq \left(\int_{S^1} |f(\theta)|^q d\theta \right)^{\frac{p}{q}} \left(\int_{S^1} 1 d\theta \right)^{\frac{q}{q-p}} \leq \|f\|_q^p (2\pi)^{\frac{q}{q-p}} \quad (3.15)$$

by choosing Hölder conjugates $\frac{q}{p}$ and $1 - \frac{p}{q}$. The above inequality implies

$\|f\|_p \leq \|f\|_q (2\pi)^{\frac{q}{qp-p^2}}$. Using Young's inequality we see that $\mathcal{B} \star u \in L^\infty$ when $u \in L^p$ for any $1 \leq p \leq \infty$ since:

$$\|\mathcal{B} \star u\|_\infty \leq \|\mathcal{B}\|_{(1-\frac{1}{p})} \|u\|_p \leq 2\pi \|\mathcal{B}\|_\infty \|u\|_p < \infty \quad (3.16)$$

where $\|\mathcal{B}\|_\infty < \infty$ by continuity. Similarly, Hölder's inequality shows that $\kappa u \in L^p$ for $u \in L^p$:

$$\int_{S^1} |\kappa(\theta)u(\theta)|^p d\theta \leq \|\kappa^p\|_\infty \|u^p\|_1 = \|\kappa^p\|_\infty \|u\|_p^p < \infty \quad (3.17)$$

where $\|\kappa^p\|_\infty < \infty$ again by continuity. Combining two inequalities shows that A is a bounded linear operator from $L^p \rightarrow L^p$ for all $1 \leq p \leq \infty$:

$$\|Au\|_p \leq \|\mathcal{B} \star u\|_p + \|\kappa u\|_p \leq 2\pi \|\mathcal{B}\|_\infty \|u\|_p + \|\kappa^p\|_\infty^{1/p} \|u\|_p \quad (3.18)$$

Since \mathcal{B} is real and symmetric, direct calculation shows:

$$\begin{aligned} \langle Au, u \rangle &= \iint_{S^1 S^1} \mathcal{B}(\theta - \omega) u(\omega) d\omega u(\theta) - \kappa(\theta) u(\theta)^2 d\theta \\ &= \iint_{S^1 S^1} \mathcal{B}(\omega - \theta) u(\theta) d\theta u(\omega) - \kappa(\omega) u(\omega)^2 d\omega = \langle u, Au \rangle \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ is the usual real L^2 inner product. The equalities hold by Fubini's theorem and the inclusion $L^2 \subset L^1$. \square

We are primarily interested in eigenvalues, elements of the point spectrum, so we will use the compactness of the convolution operator to ensure the existence of eigenvalues.

Step 2

The operator $(\mathcal{B} \star u)(\theta)$ is compact.

Proof. When we are considering the operator $\mathcal{B} \star \cdot$ over L^2 , we can simply observe that $\mathcal{B}(\theta - \omega)$ is a Hilbert-Schmidt kernel and that implies that the convolution is compact. However, proving compactness over L^1 is slightly more difficult. We will use the Arzelà-Ascoli theorem to show that $\mathcal{B} \star \cdot$ maps bounded sequences to sequences with a convergent subsequence.

Take a sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ where $\|f_n\|_1 \leq 1$. Using Young's inequality as above, we obtain a uniform bound on $\|\mathcal{B} \star f_n\|_\infty$:

$$\|\mathcal{B} \star f_n\|_\infty \leq \|\mathcal{B}\|_\infty \|f_n\|_1 \leq \|\mathcal{B}\|_\infty \quad (3.19)$$

Since $\mathcal{B} \in C^2$, we know that $\|\mathcal{B}'\|_\infty < \infty$. We can again apply Young's inequality to show that the derivative of $\mathcal{B} \star f_n$ is uniformly bounded:

$$\|(\mathcal{B} \star f_n)'\|_\infty = \|\mathcal{B}' \star f_n\|_\infty \leq \|\mathcal{B}'\|_\infty \|f_n\|_1 \leq \|\mathcal{B}'\|_\infty \quad (3.20)$$

The $(\mathcal{B} \star f_n)'$ being uniformly bounded implies that $\{\mathcal{B} \star f_n\}$ is uniformly Lipschitz. That means that Arzelà-Ascoli holds and $\mathcal{B} \star \cdot$ is compact. \square

We will now use the compactness of $\mathcal{B} \star \cdot$ to show that A has no continuous spectrum outside of the range of $-\kappa$. For the sake of notation, define $\sigma_1(A)$ to be the spectrum of A over L^1 and $\sigma_2(A)$ the spectrum of A over L^2 .

Step 3

All elements of the spectrum outside of $[-\sup \kappa, -\inf \kappa]$ are either eigenvalues or in the residual spectrum for all L^p spaces with $1 \leq p \leq \infty$.

Proof. This result holds equally for all the L^p spaces. I will prove the result for L^1 . The argument holds by simply replacing the metric $\|\cdot\|_1$ with $\|\cdot\|_p$. The eigenfunctions over L^1 are bounded, so are in all of the L^p spaces. Fix a number $\mu \in \sigma_1(A)$ with $\mu \notin [-\sup \kappa, -\inf \kappa]$ and in either the continuous spectrum or the point spectrum. Since μ is not in the residual spectrum, we have a sequence $\{u_n\}_{n \in \mathbb{N}} \subset L^1$ with $\|u_n\|_1$ such that:

$$\lim_{n \rightarrow \infty} \|(A - \mu I)u_n\|_1 = 0$$

By invoking the compactness of $\mathcal{B} \star \cdot$ from the previous step, there exists a subsequence u_{n_k} such that:

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} \|(\mathbf{A} - \mu \mathbf{I})u_{n_k}\|_1 = \lim_{k \rightarrow \infty} \|(\mathcal{B} \star u_{n_k})(\theta) - (\kappa(\theta) + \mu)u_{n_k}(\theta)\|_1 \\ &= \lim_{k \rightarrow \infty} \|v(\theta) - (\kappa(\theta) + \mu)u_{n_k}(\theta)\|_1 \end{aligned}$$

where v is the limit of $\mathcal{B} \star u_{n_k}$. By hypothesis, $\kappa(\theta) + \mu \neq 0$, so $\frac{1}{\kappa(\theta) + \mu}$ is bounded. We now have that:

$$\lim_{k \rightarrow \infty} u_{n_k}(\theta) = \frac{v(\theta)}{\kappa(\theta) + \mu} = w(\theta)$$

almost everywhere. By the fact that $\mathcal{B} \star \cdot$ is closed, $(\mathcal{B} \star w)(\theta) = v(\theta)$. Applying the above identities shows that $\mathbf{A}w = \mu w$. Finally, observe that $w \in L^\infty$ since:

$$w = \frac{\mathcal{B} \star w}{\kappa + \mu} \tag{3.21}$$

$\mathcal{B} \star w$ is bounded by Young's inequality and $\frac{1}{\kappa(\theta) + \mu}$ is bounded by hypothesis. \square

We are interested in solutions to (3.7) over finite densities, so we are studying solutions on L^1 . However, it is easier to study the spectrum of \mathbf{A} over L^2 where we can use Hilbert space geometry. To do so, we will show that the two spectrums coincide outside the range of $-\kappa$. We have already shown that the continuous and point spectra coincide outside of $-\kappa$ because all such points are associated with an L^∞ eigenfunction. The next step will finish proving the correspondence.

First, a quick lemma:

Lemma 1. *For any given sequence $f_n \subset L^q(S^1)$, $\lim_{n \rightarrow \infty} \|f_n\|_q = 0 \implies \lim_{n \rightarrow \infty} \|f_n\|_p = 0$ for all $1 \leq p < q \leq \infty$.*

Proof. Applying Hölder's inequality as above:

$$\lim_{n \rightarrow \infty} \|f_n\|_p \leq \lim_{n \rightarrow \infty} (2\pi \|f_n\|_q)^{1/p} = 0 \tag{3.22}$$

\square

Step 4

The spectrum of A over L^1 outside of $[-\sup \kappa, -\inf \kappa]$ has no elements in the residual spectrum and therefore is equal to the spectrum of A over L^2 .

Proof. To show this, we will consider the spectrum in three parts, the point spectrum, the continuous spectrum, and the residual spectrum. Any eigenvalue of A on L^2 is an eigenvalue on L^1 by the inclusion $L^2 \subset L^1$. The reverse inclusion comes from the fact that all of the eigenvalues over L^1 outside of $[-\sup \kappa, -\inf \kappa]$ are in $L^\infty \supset L^2$. We now have that the two point spectrums are equal. The result in step 3 implies that there is no elements of the continuous spectrum outside of $[-\sup \kappa, -\inf \kappa]$, which implies the continuous spectrums are equal. All that remains is to show that A has no residual spectrum on either L^2 or L^1 .

The natural embedding of L^q into L^p where $1 \leq p < q \leq \infty$ is a dense embedding. Continuous functions are dense in L^1 as can be seen by approximating simple functions by continuous functions. Since continuous functions are in L^∞ , that implies L^∞ is dense in L^1 . By the inclusion $L^q \subset L^1$, continuous functions are dense in L^q for all $1 < q \leq \infty$.

The last step remains to show that A has no residual spectrum over neither L^2 nor L^1 . By self-adjointness, A has no residual spectrum over L^2 . The fact that A has no residual spectrum over L^1 follows immediately from the density of the embedding L^2 in L^1 . We know that $A - \mu I$ has dense range in L^2 for all $\mu \in \mathbb{C}$ whenever μ is not an eigenvalue. Assume μ is not an eigenvalue, the dense embedding and $A - \mu I$ having dense range in L^2 implies that $A - \mu I \Big|_{L^2}$ is dense in L^1 . That implies that $A - \mu I \Big|_{L^1}$ is dense in L^1 and that μ is not in the residual spectrum. \square

To further characterize the spectrum, we will show the spectrum is real, which allows us to put an order on the elements of the spectrum.

Step 5

$\sigma_1(A)$ is real.

Proof. We can decompose the spectrum $\sigma_1(A)$ into two sets: $\sigma_1(A) \cap [-\sup \kappa, -\inf \kappa]$ and $\sigma_1(A) \setminus [-\sup \kappa, -\inf \kappa]$. The first set is real by definition. Step 4 implies that:

$$\sigma_1(A) \setminus [-\sup \kappa, -\inf \kappa] = \sigma_2(A) \setminus [-\sup \kappa, -\inf \kappa] \subset \mathbb{R}$$

The equality above follows from the fact that A is self-adjoint over L^2 . □

Step 6

$\sup_{\|u\|_2=1} \langle Au, u \rangle$ is either an eigenvalue or an element of the continuous spectrum.

Proof. By the previous step, we have that both $\sigma_1(A)$ and $\sigma_2(A)$ are real, so the supremum is well defined. Assume $\mu = \sup_{\|u\|_2=1} \langle Au, u \rangle$. It then suffices to show that $A - \mu I$ is not invertible.

By the definition of μ , $\mu I - A$ is a positive operator. Take a sequence $u_n \subset L^2$, $\|u_n\|_2 = 1$ and $\lim_{n \rightarrow \infty} \langle Au_n, u_n \rangle = \mu$. We can apply the generalized Cauchy-Schwartz inequality for positive operators to observe:

$$\begin{aligned} \lim_{n \rightarrow \infty} \|Au_n - \mu u_n\|_2^4 &= \lim_{n \rightarrow \infty} |\langle (\mu I - A)u_n, (\mu I - A)u_n \rangle|^2 \\ &\leq \lim_{n \rightarrow \infty} \langle (\mu I - A)u_n, u_n \rangle \langle (\mu I - A)^2 u_n, (\mu I - A)u_n \rangle = 0 \end{aligned}$$

since $\lim_{n \rightarrow \infty} \langle (\mu I - A)u_n, u_n \rangle = 0$ by hypothesis. Therefore, $\mu \in \sigma_2(A)$ and $\max \sigma_2(A) \geq \sup_{\|u\|_2=1} \langle Au, u \rangle$.

We can actually show the stronger result that $\max \sigma_2(A) = \sup_{\|u\|_2=1} \langle Au, u \rangle$. To show that, assume $\mu = \max \sigma_2(A)$. That implies there exists another sequence as above such that $\lim_{n \rightarrow \infty} \|Au_n - \mu u_n\|_2 = 0$. That implies:

$$\lim_{n \rightarrow \infty} \langle Au_n - \mu u_n, u_n \rangle \leq \lim_{n \rightarrow \infty} \|Au_n - \mu u_n\|_2 = 0$$

by the usual Cauchy-Schwartz inequality. □

To show the existence of at least one eigenvalue, we will show that $\max \sigma_2 > -\inf \kappa$. That implies that there is at least one eigenvalue, μ , with $\mu > -\inf \kappa$.

Step 7

The largest element of the spectrum is greater than $-\inf \kappa$.

Proof. From the above step, it is sufficient to show that there exists $u \in L^2$ such that $\langle A + (\inf \kappa)u, u \rangle > 0$. By continuity of κ and compactness of the circle, $\kappa(\theta) - \inf \kappa = 0$ for at least one θ . For the sake of notation, define $g(\theta) = \kappa(\theta) - \inf \kappa$.

Observe that for any function of the form $u(\theta) = c + f(\theta) \geq 0$ with $c, f(\theta) \geq 0$ where $\int_{S^1} u(\theta) = 1$:

$$\begin{aligned} \langle \mathcal{B} \star u, u \rangle &= \iint_{S^1 S^1} \mathcal{B}(\theta - \omega)(c + f(\omega)) \, d\omega (c + f(\theta)) \, d\theta \\ &\geq c \iint_{S^1 S^1} \mathcal{B}(\theta - \omega)(c + f(\theta)) \, d\omega \, d\theta = c \end{aligned}$$

since $f(\theta) \geq 0$ and $\int \mathcal{B} = 1$. Without loss of generality, assume $g(0) = 0$. Also, observe that $g'(0) = 0$ since $g(0)$ is a local minimum of a smooth function. Finally, observe that g'' is bounded since κ is C^2 . Putting those three facts together gives the relation:

$$\begin{aligned} g(\theta) &= \int_0^\theta g'(\omega) \, d\omega + g(0) \\ &= \int_0^\theta \int_0^\omega g''(\phi) + g'(0) \, d\phi \\ &\leq Q\theta^2 \end{aligned}$$

where $Q = \sup \kappa''(\theta)$. Define f_ε as:

$$f_\varepsilon = \frac{\mathbf{1}_{[-\varepsilon, \varepsilon]}}{2\varepsilon}$$

where $\mathbf{1}$ is the usual indicator function. Note that $\int f = 1$. We now have the two relations:

$$\begin{aligned} \int_{S^1} f_\varepsilon(\theta)g(\theta) \, d\theta &= \int_{-\varepsilon}^\varepsilon f_\varepsilon(\theta)g(\theta) \, d\theta \leq \frac{2}{\varepsilon} \int_0^\varepsilon Q\theta^2 \, d\theta = Q\frac{2\varepsilon^2}{3} \\ \int_{S^1} f_\varepsilon(\theta)^2 g(\theta) \, d\theta &= \int_{-\varepsilon}^\varepsilon f_\varepsilon(\theta)^2 g(\theta) \, d\theta \leq \frac{2}{\varepsilon^2} \int_0^\varepsilon Q\theta^2 \, d\theta = Q\frac{2\varepsilon}{3} \end{aligned}$$

Combining those relations gives:

$$\begin{aligned}\langle g(c + f_\varepsilon), (c + f_\varepsilon) \rangle &= \int_{S^1} g(\theta)(c + f_\varepsilon(\theta))^2 d\theta \\ &= \int_{S^1} g(\theta) \left(c^2 + 2cf_\varepsilon(\theta) + f_\varepsilon(\theta)^2 \right) d\theta \leq Rc^2 + cQ \frac{4\varepsilon^2}{3} + Q \frac{2\varepsilon}{3}\end{aligned}$$

where $R = \int g(\theta) d\theta$.

First, assume $R \geq 2\pi$ and $Q \geq 1$. Fix $c = \frac{1}{2R}$, $\varepsilon = \frac{1}{4RQ}$ and $c' = 1 - 2\pi c < 1$. Fix $u = c + c'f_\varepsilon$. Putting all of the above together gives:

$$\begin{aligned}\langle Au + (\inf \kappa)u, u \rangle &= \langle \mathcal{B} \star (c + c'f_\varepsilon), c + c'f_\varepsilon \rangle - \langle g(c + c'f_\varepsilon), c + c'f_\varepsilon \rangle \\ &\geq c - Rc^2 - cc'Q \frac{4\varepsilon^2}{3} - c'^2Q \frac{2\varepsilon}{3} \\ &\geq \frac{1}{2R} - \frac{R}{4R^2} - \frac{Q}{R} \frac{2\varepsilon^2}{3} - Q \frac{2\varepsilon}{3} \\ &\geq \frac{1}{4R} - \frac{Q}{3R} \frac{1}{8R^2Q^2} - \frac{Q}{6RQ} \\ &\geq \frac{1}{4R} - \frac{1}{24R} - \frac{1}{6R} = \frac{1}{24R} > 0\end{aligned}$$

If $R < 2\pi$, then set $c = \frac{1}{2\pi}$ and $c' = 0$. If $R \geq 2\pi$ and $Q < 1$, set c and c' as above and $\varepsilon = \frac{1}{4R}$. That shows we have constructed such a u and $\max \sigma_2(A) > -\inf \kappa$. A proof with a weaker smoothness condition is outlined in section 3.9. \square

Step 8

Eigenvalues do not accumulate outside of $[-\sup \kappa, -\inf \kappa]$. That implies that the principal eigenvalue is isolated.

Proof. By the relation $\sigma_1(A) \setminus [-\sup \kappa, -\inf \kappa] = \sigma_2(A) \setminus [-\sup \kappa, -\inf \kappa]$ proven in Step 4, it is sufficient to show the result for $\sigma_2(A)$.

Assume there exists a sequence of elements of the spectrum $\{\mu_j\}_{j \in \mathbb{N}} \subset \sigma_1(A)$ such that $\mu_j \rightarrow \mu$. Without loss of generality, assume $\mu_j < \mu_{j+1} < \mu$ and $\mu_j > -\inf \kappa$ for all j . That implies that all of the μ_j 's are eigenvalues. Choose a set of associated eigenfunctions

$\{u_j\}_{j \in \mathbb{N}} \subset L^\infty$ such that $\|u_j\|_2 = 1$. By the self-adjointness of A , we have that $\langle u_j, u_k \rangle = \delta_{jk}$ and $\|u_j - u_k\|_2 = \sqrt{2}(1 - \delta_{jk})$.

$$\|u_j - u_k\|_2^2 = \|u_j\|_2^2 - 2\langle u_j, u_k \rangle + \|u_k\|_2^2 = 2(1 - \delta_{jk}) \quad (3.23)$$

By the compactness of $\mathcal{B} \star \cdot$, choose a convergent subsequence $\{u_{j_k}\}_{k \in \mathbb{N}}$. We now arrive at the relations:

$$\begin{aligned} 0 &= \lim_{k \rightarrow \infty} \|\mathcal{B} \star u_{j_k} - \mathcal{B} \star u_{j_{k+1}}\|_2^2 = \lim_{k \rightarrow \infty} \|(\kappa + \mu_{j_k})u_{j_k} - (\kappa + \mu_{j_{k+1}})u_{j_{k+1}}\|_2^2 \\ &= \lim_{k \rightarrow \infty} \left(\|(\kappa + \mu)(u_{j_k} - u_{j_{k+1}})\|_2^2 + (\mu - \mu_{j_k})^2 \|u_{j_k}\|_2^2 + (\mu - \mu_{j_{k+1}})^2 \|u_{j_{k+1}}\|_2^2 \right) \\ &= \lim_{k \rightarrow \infty} \|(\kappa + \mu)(u_{j_k} - u_{j_{k+1}})\|_2^2 \\ &\geq \lim_{k \rightarrow \infty} (\inf \kappa + \mu)^2 \|u_{j_k} - u_{j_{k+1}}\|_2^2 \\ &= \lim_{k \rightarrow \infty} (\inf \kappa + \mu)^2 \left(\|u_{j_k}\|_2^2 - 2\Re(\langle u_{j_k}, u_{j_{k+1}} \rangle) + \|u_{j_{k+1}}\|_2^2 \right) = 2(\inf \kappa + \mu)^2 > 0 \end{aligned}$$

where $\Re(z)$ is the real part of z . The above is a contradiction since $\inf \kappa + \mu > 0$ by hypothesis. For the case where $\mu < -\sup \kappa$, simply replace $\inf \kappa + \mu$ with $\mu + \sup \kappa < 0$. \square

Step 9

The principal eigenvalue is simple and has a positive, bounded eigenfunction.

Proof. Observe that we have the related eigenvalue problem:

$$\frac{\mathcal{B} \star u}{\kappa + \mu_0} = \gamma u$$

where μ_0 is the principal eigenvalue of A . Assume there exists an eigenpair γ, u such that $\gamma > 1$ and γ is the largest eigenvalue.

$$\frac{\mathcal{B} \star u}{\kappa + \mu_0} = \gamma u \quad \Longleftrightarrow$$

$$\mathcal{B} \star u = (\kappa + \mu_0)\gamma u \quad \Longleftrightarrow$$

$$\mathcal{B} \star u - (\kappa u + \mu_0)u = (\gamma - 1)(\kappa + \mu_0)u \quad \Longleftrightarrow$$

$$Au - \mu_0 u = (\gamma - 1)(\kappa + \mu_0)u$$

which would give:

$$\langle Au - \mu_0 u, u \rangle = (\gamma - 1) \langle (\kappa + \mu_0)u, u \rangle > 0$$

The last inequality holds since $\kappa(\theta) + \mu_0 > 0$ for all θ . That is in contradiction to the fact that μ_0 is the principal eigenvalue of A which implies $\sup_{\|u\|_2=1} \langle Au, u \rangle = \mu_0$. It is obvious that μ_0 is a simple eigenvalue with positive eigenfunction if and only if $\gamma = 1$ is a simple eigenvalue with positive eigenfunction. We can use the Krein-Rutman theorem to complete the proof [29, 60]. For the sake of notation, define the operator A' :

$$A'u = \frac{\mathcal{B} \star u}{\kappa} \quad (3.24)$$

Inspection of the above equation shows that any eigenfunction u must be continuous because of the convolution. Thus, we will consider A' as an operator on the space $C(S^1)$, the space of continuous functions on the circle with supremum norm. A' is compact on this space as can be seen by following the argument in step 2. The set of positive functions in $C(S^1)$ forms a cone K such that $C(S^1) = K + (-K)$ and K is solid, it has non-empty interior. $A'K \subset K$ by the fact that the convolution of two positive functions multiplied by a positive function is itself positive. A' is therefore a compact, positive operator, which implies that there exists a largest positive eigenvalue μ with $\mu = \rho(A')$, the spectral radius. There is also a positive eigenfunction, $u \geq 0$, associated with μ .

The last thing that remains to be shown is that the eigenvalue is simple. The Krein-Rutman theorem implies that it is sufficient to show that $(A')^n$ is strongly positive on continuous functions for some n . That is shown in Lemma 2. \square

Lemma 2. *For any given kernel \mathcal{B} , there exists n such that $(A')^n$ is strongly positive, i.e.:*

$$u \geq 0 \implies (A')^n u > 0$$

whenever u is continuous and not uniformly zero.

Proof. We are only concerned with whether or not $\left(A'\right)^n$ is positive and not on the specific value of $\left(A'\right)^n$, so it is sufficient to show the result for B^n where $Bu = \mathcal{B} \star u$. Define Σ to be the σ -algebra associated with the Lebesgue measure on S^1 . We can define the set mapping $T : \Sigma \rightarrow \Sigma$ as:

$$T\Omega = \text{supp}\{\mathcal{B} \star \mathbf{1}_\Omega\} \quad (3.25)$$

where $\Omega \in \Sigma$. Choose some open interval $(y - \delta, y + \delta) \subset \text{supp}\{\mathcal{B}\}$ for $y \neq 0$, $y \in \mathbb{R} \setminus \mathbb{Q}$. Assume Ω contains an open interval $(x - \varepsilon, x + \varepsilon)$. Observe that for every $z \in (x + y - \delta, x + y + \delta)$:

$$\mathcal{B} \star \mathbf{1}_\Omega(z) = \int \mathcal{B}(z - s) \mathbf{1}_\Omega(s) ds \geq \int_{x-\varepsilon}^{x+\varepsilon} \mathcal{B}(z - s) ds = \int_{y-\varepsilon}^{y+\varepsilon} \mathcal{B}\left((z - x + y) - s'\right) ds' > 0 \quad (3.26)$$

The above argument also holds for $z \in (x - y - \delta, x - y + \delta)$. Notice that while the existence of the interval was used in the above calculation, there is no explicit dependence on ε beyond that $\varepsilon > 0$. Iterating T , we can observe that:

$$T^{2n}\Omega \subset \cup_{0 \leq j \leq n} \left((x + 2jy - \delta, x + 2jy + \delta) \cup (x - 2jy - \delta, x - 2jy + \delta) \right) \quad (3.27)$$

By the fact that $x + 2y$ is an irrational rotation, $\{x + 2jy\}_{j \in \mathbb{N}}$ is dense in S^1 and the sets $\{(x + 2jy - \delta, x + 2jy + \delta)\}_{j \in \mathbb{N}}$ form an open covering of S^1 . The compactness of S^1 implies that there exists $n \in \mathbb{N}$ such that $S^1 \subset \cup_{1 \leq j \leq n} (x + 2jy - \delta, x + 2jy + \delta)$. Rotational symmetry in S^1 implies that n has no dependence on x . Fix some function $u \in C(S^1)$ with $u \geq 0$ and u not uniformly zero. We can set $\Omega = \text{supp}\{u\}$ and observe Ω contains an open interval containing some x' . The above discussion implies:

$$T^{2n}\Omega \supset \cup_{1 \leq j \leq n} (x' + 2jy - \delta, x' + 2jy + \delta) \supset S^1 \quad (3.28)$$

The above set relation implies $B^{2n}u > 0$. □

3.3 PROOF OF SECOND THEOREM

This section shows that there is only one λ such that (3.7) has a stable equilibrium. As such, we will recall that A_λ is defined as:

$$A_\lambda u = \lambda \mathcal{B} \star u - \kappa u$$

Define the related operator and inner product spaces:

$$A' u = \frac{\mathcal{B} \star u}{\kappa} \quad \langle f, g \rangle_\kappa = \int_{S^1} f(\theta) g(\theta) \kappa(\theta) d\theta$$

We can now prove the result.

Proof. A' is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\kappa$:

$$\begin{aligned} \langle A' f, g \rangle_\kappa &= \iint_{S^1 S^1} \frac{\mathcal{B}(\theta - \omega)}{\kappa(\theta)} f(\omega) d\omega g(\theta) \kappa(\theta) d\theta \\ &= \iint_{S^1 S^1} \mathcal{B}(\omega - \theta) f(\omega) g(\theta) d\omega d\theta \\ &= \iint_{S^1 S^1} \frac{\mathcal{B}(\omega - \theta)}{\kappa(\omega)} g(\theta) d\theta f(\omega) \kappa(\omega) d\omega = \langle f, A' g \rangle_\kappa \end{aligned}$$

Also, we have the relation between A_λ and A' :

$$\begin{aligned} \langle A_\lambda u, u \rangle &= \iint_{S^1 S^1} \lambda \mathcal{B}(\theta - \omega) u(\omega) d\omega u(\theta) - \kappa u(\theta)^2 d\theta \\ &= \iint_{S^1 S^1} \lambda \frac{\mathcal{B}(\theta - \omega)}{\kappa(\theta)} u(\omega) d\omega u(\theta) \kappa(\theta) d\theta - \int_{S^1} u(\theta)^2 \kappa(\theta) d\theta \\ &= \lambda \langle A' u, u \rangle_\kappa - \langle u, u \rangle_\kappa \end{aligned}$$

From the result in step 9 of theorem 1, we have that A' has a simple principal eigenvalue, $\mu'_0 > 0$. We know that μ_0 and μ'_0 can be defined by the following:

$$\mu_0 = \sup_{\|u\|_2 \neq 0} \frac{\langle A u, u \rangle}{\langle u, u \rangle} \quad \text{and} \quad \mu'_0 = \sup_{\langle u, u \rangle_\kappa \neq 0} \frac{\langle A' u, u \rangle_\kappa}{\langle u, u \rangle_\kappa} \quad (3.29)$$

There is a relationship between the sign of μ_0 and μ'_0 :

$$\begin{aligned}
\text{sign}[\lambda\mu'_0 - 1] &= \text{sign} \left[\sup_{\langle u, u \rangle_\kappa \neq 0} \frac{\lambda \langle A'u, u \rangle_\kappa - \langle u, u \rangle_\kappa}{\langle u, u \rangle_\kappa} \right] \\
&= \sup_{\langle u, u \rangle_\kappa \neq 0} \left[\text{sign} \frac{\lambda \langle A'u, u \rangle_\kappa - \langle u, u \rangle_\kappa}{\langle u, u \rangle_\kappa} \right] \\
&= \sup_{\langle u, u \rangle_\kappa \neq 0} \left[\text{sign} \frac{\lambda \langle A'u, u \rangle_\kappa - \langle u, u \rangle_\kappa}{\langle u, u \rangle} \right] \\
&= \sup_{\|u\|_2 \neq 0} \left[\text{sign} \frac{\lambda \langle A'u, u \rangle_\kappa - \langle u, u \rangle_\kappa}{\langle u, u \rangle} \right] \\
&= \sup_{\|u\|_2 \neq 0} \left[\text{sign} \frac{\langle A_\lambda u, u \rangle}{\langle u, u \rangle} \right] \\
&= \text{sign} \left[\sup_{\|u\|_2 \neq 0} \frac{\langle A_\lambda u, u \rangle}{\langle u, u \rangle} \right] = \text{sign}[\mu_0] \tag{3.30}
\end{aligned}$$

The argument holds since $\frac{\langle u, u \rangle_\kappa}{\langle u, u \rangle} > 0$ by hypothesis of the supremum and does not change the sign of the argument. We know that (3.7) has a non-trivial equilibrium if and only if A_λ has a zero eigenvalue. That equilibrium is stable if and only if all of the other elements of the spectrum have negative real part. That is the case if and only if the zero eigenvalue is the largest eigenvalue, i.e. $\mu_0 = 0$. The calculation above therefore implies there exists only one λ where (3.7) has a stable equilibrium since there is only on λ such that $\mu_0 = \lambda\mu'_0 - 1 = 0$. \square

3.4 PROOF OF ZERO-ORDER PROPOSITIONS

The following proposition provides a formula for equilibria and shows there are at most countable equilibria. Similar to the proof of theorem 1, we will assume without loss of generality that $\lambda_0 = 1$ to simplify notation.

Proposition 1. *A function $u \in L^1$ is an equilibrium of equation (3.8) if and only if it is a solution to the eigenvalue problem:*

$$\frac{\mathcal{B} \star u}{\kappa(\theta)} = \mu u(\theta) \tag{3.31}$$

where $\mu \neq 0$ and $\int_{S^1} u(\omega) d\omega \neq 0$.

Proof. Assume you have an equilibrium $u \in L^1$ with $\int u(\omega) d\omega \neq 0$, i.e. $G(u) = 0$. We know that:

$$\frac{(\mathcal{B} \star u)(\theta)}{\int_{S^1} u(\omega) d\omega} - \kappa(\theta)u(\theta) = 0$$

Simple algebra gives:

$$\frac{(\mathcal{B} \star u)(\theta)}{\kappa(\theta)} = \int_{S^1} u(\omega) d\omega u(\theta)$$

That implies u is an eigenfunction with eigenvalue $\int u(\omega) d\omega \neq 0$. For the other direction, assume that we have:

$$\frac{\mathcal{B} \star u}{\kappa} = \mu u$$

and the listed hypotheses above. Simple algebra again:

$$\mathcal{B} \star u - \mu \kappa u = 0$$

Assume $\int_{S^1} u(\omega) d\omega = 1$. This is justified as long as $\int_{S^1} u(\omega) d\omega \neq 0$. The existence of at least one positive eigenfunction with non-zero integral is ensured by the Krein-Rutman theorem as in section 3.7. $v = \mu u$ is now an equilibrium to (3.8). \square

The above proposition allows us to study the set of equilibria of (3.8) by studying the spectra of the operator A' via the eigenvalue problem (3.31). The following proposition about A' implies that there is only one non-negative equilibrium to (3.8).

Proposition 2. *Only one eigenfunction of A' is non-negative.*

Proof. From previous results, we know the principal eigenvalue of A' is simple and the associated eigenfunction, u_0 , is non-negative. We want to show the stronger result that $u_0 > 0$. The argument follows the same format as in lemma 2. Assume $u_0(x) = 0$. By constructing the set mapping T as before, that implies that u_0 is zero on the set $(x - y - \delta, x - y + \delta)$. Iterating that argument shows that u_0 is zero on $\{(x - jy - \delta, x - jy + \delta)\}_{j \in \mathbb{N}} \supset S^1$. Since u_0 is not uniformly zero, we have reached a contradiction and can infer that $u_0 > 0$.

By the self-adjointness of A' , we have the all eigenfunctions $u_k \neq u_0$ are orthogonal to u_0 :

$$0 = \langle u_0, u_k \rangle_\kappa = \int_{S^1} u_0(\theta) u_k(\theta) \kappa(\theta) d\theta$$

However, we know that $\kappa u_0 > 0$. That implies the above can only be true if $u_k \equiv 0$ almost everywhere or u_k is negative on some set with non-zero measure. \square

Given the above results, we can observe that (3.8) must either converge to a unique, stable equilibrium or not converge to an equilibrium at all. We also know that the unique equilibrium to (3.8) is the principal eigenvalue of A' . We also know that there cannot be any bistable solutions to (3.8) since the equilibrium is unique. This directly contradicts [137].

3.5 SIMULATIONS

Due to the difficulty in showing stability for the zeroth-order branching equation, (3.8), simulations were run to support conjecture 1. To demonstrate stability, two branching kernels, two capping functions and four initial conditions were selected to feature some amount of generality.

The two branching kernels were:

$$\mathcal{B}_1(\theta) = \frac{\phi(\theta + \frac{\pi}{2}) + \phi(\theta - \frac{\pi}{2})}{2} \quad \text{and} \quad \mathcal{B}_2(\theta) \propto \begin{cases} 1 - \theta^6 + 3\theta^4 - 3\theta^2 & \text{if } |\theta| < 1 \\ 0 & \text{if } |\theta| \geq 1 \end{cases} \quad (3.32)$$

where \mathcal{B}_2 was normalized to have integral one and $\phi(\theta)$ was a von Mises distribution:

$$\phi(\theta) = \frac{\exp[\sigma^{-2} \cos(\theta)]}{2\pi I_0(\sigma^{-2})} \quad (3.33)$$

where $\sigma = \frac{7\pi}{180}$. \mathcal{B}_1 was based on the branching kernels used in [68] and [137]. Those two papers used truncated Gaussian distributions centered around $\pm 70^\circ$. Here, the von Mises distribution was used to avoid truncating the Gaussian or using the more complicated, formally correct wrapped Gaussian distribution. Also, the offset of $\pm \frac{\pi}{2}$ was used to simplify the radians conversion. Finally, the constant σ was chosen to be in line with previous numerical studies [114, 115, 137]. Both branching kernels were C^2 and symmetric.

The two capping functions used were:

$$\kappa_1(\theta) = 1 - \frac{1}{2} \cos(\theta) \quad \text{and} \quad \kappa_2(\theta) = 1 + \frac{3}{4} \cos(4\theta^2) \quad (3.34)$$

The first capping function, κ_1 was based upon [137], and the second was chosen to have multiple minima and maxima and non-uniform oscillations.

Finally, the four initial conditions chosen were:

$$\begin{aligned} u_1(\theta) &= 1 & u_2(\theta) &= \frac{1}{\left|\theta - \frac{\pi}{3}\right|^{1/2}} \\ u_3(\theta) &= \mathbf{1}_{\left(\frac{7\pi}{8}, \pi\right)} + \mathbf{1}_{\left(-\frac{3\pi}{4}, -\frac{2\pi}{3}\right)} & u_4(\theta) &= \begin{cases} -\frac{\theta}{\pi} & \text{if } \theta < 0 \\ 1 - \frac{\theta}{\pi} & \text{if } \theta \geq 0 \end{cases} \end{aligned} \quad (3.35)$$

They were chosen to include a mix of symmetric, non-symmetric, smooth and non-smooth functions. Also, u_2 was chosen so that $u_2 \in (L^1 \setminus L^2)$.

3.5.1 Methods

All simulations were run using the Numpy [2, 54, 85, 86] extension to Python. The equations of motion were integrated using a simple Euler method. The simulations were run for 100 time units with a time step of 0.01 time units. The circle was discretized using 2^{11} equally spaced points from $-\pi$ to π (the power of 2 was used to speed up the fast Fourier transform). All integrals were taken using the trapezoidal method included in Numpy. The convolution was performed by taking the real fast Fourier transform of the branching kernel \mathcal{B} and the density $u(\theta, t)$, multiplying, and taking the inverse real fast Fourier transform. The built-in Numpy fast convolution method was not used because that method pads the two convolved functions with extra zeros to prevent circular convolution, but the equations used here explicitly call for the circular convolution. The convolution was normalized by dividing \mathcal{B} by the integral of the convolution of \mathcal{B} with the constant function $\mathcal{B} \star \frac{1}{2\pi}$. Finally, the total branching rate was normalized by integrating the density at the previous time step, i.e.:

$$u(\theta, t + dt) = dt \left(\frac{(\mathcal{B} \star u)(\theta, t)}{\int u(\theta, t) d\theta} - \kappa(\theta)u(\theta, t) \right) \quad (3.36)$$

Using the integral from the previous time step and not a more sophisticated prediction-correction methods is justified by the following inequality:

$$\begin{aligned} |u(\theta, t + dt) - u(\theta, t)| &\leq dt \left(\int (\mathcal{B} \star u)(\theta, t) d\theta + \int \kappa(\theta) u(\theta, t) d\theta \right) \\ &\leq \left(1 + \sup_{\theta \in S^1} \kappa(\theta) \right) dt \|u(\theta, t)\|_1 \leq 3 dt \|u(\theta, t)\|_1 \end{aligned} \quad (3.37)$$

Since simulations remain bounded, the bound above can be made uniform.

The equilibrium was calculated a priori by iterating the equilibrium operator A' . Explicitly, a sequence of functions was generated by:

$$v_{n+1}(\theta) = \frac{1}{\int v_n(\theta) d\theta} A' v_n(\theta) \quad (3.38)$$

where the discretization and convolution were performed exactly as above and $v_0 = \frac{1}{2\pi}$. The theoretical justification for using this method is outlined in Section 3.7. For three combinations of \mathcal{B} and κ , $\|v_{10^4} - v_{2 \times 10^4}\|_1$ was less than numerical precision. For the combination of \mathcal{B}_1 and κ_2 , v_{10^4} was not sufficiently converged, so v_{10^6} was used. That decision was based on the condition that $\|v_{10^6} - v_{2 \times 10^6}\|_1$ was less than numerical precision. That level of precision was used to ensure that the convergence could be seen even when $\|u(\theta, t) - v_n\|_1 \leq 10^{-6} - 10^{-10}$. Even 10^6 iterations of A' only took several minutes on a standard Linux desktop system concurrently running other programs, a number that could be reduced with further optimization.

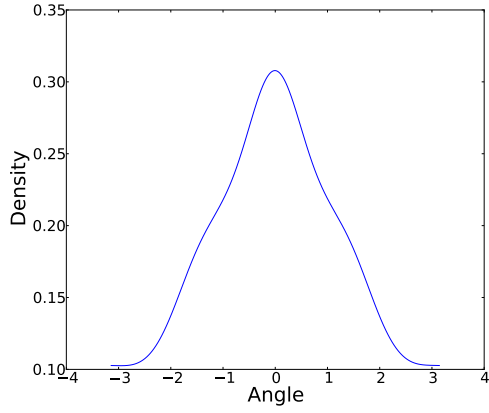
At each time-step, the L^1 distance between the state of the system $u(\theta, t)$ and the equilibrium was calculated. That quantity is plotted as a function of time in Figure 3.3.

3.5.2 Results

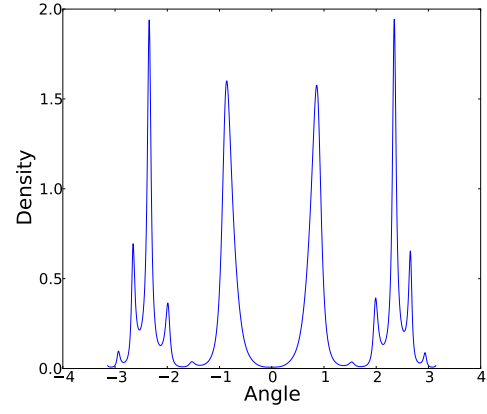
The first calculations run were to estimate the equilibrium distribution using the method in section 3.7. The method was iterated until the $\|\cdot\|_1$ difference between successive iterations was less than double precision. The equilibrium distribution appeared to have a qualitatively stronger dependence on κ than on \mathcal{B} as can be seen in Figure 3.2.

All of the simulations run converged (asymptotically) exponentially to the equilibrium. The equilibrium was calculated by iterating A' as outlined in the following section. Figure

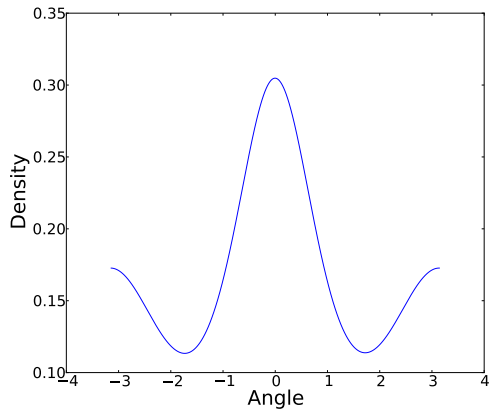
A)



B)



C)



D)

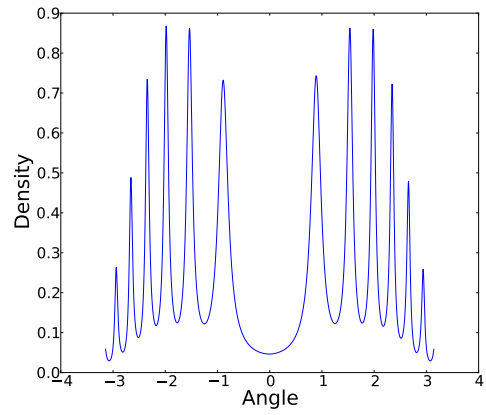


Figure 3.2: The plots here show the equilibrium distributions calculated using the method in section 3.7 for the four systems. A) \mathcal{B}_1 and κ_1 , B) \mathcal{B}_1 and κ_2 , C) \mathcal{B}_2 and κ_1 , and D) \mathcal{B}_2 and κ_2 .

3.3 shows the L^1 distance between the simulation result and the calculated equilibrium on a log scale. The log scale was used to make the graphs legible and to show the exponential convergence.

All of the initial conditions appear to asymptotically converge at the same rate. That gives evidence that a stronger result that conjecture 1 may be possible. The simulations provide evidence for exponential stability with some constant determined solely by \mathcal{B} and κ .

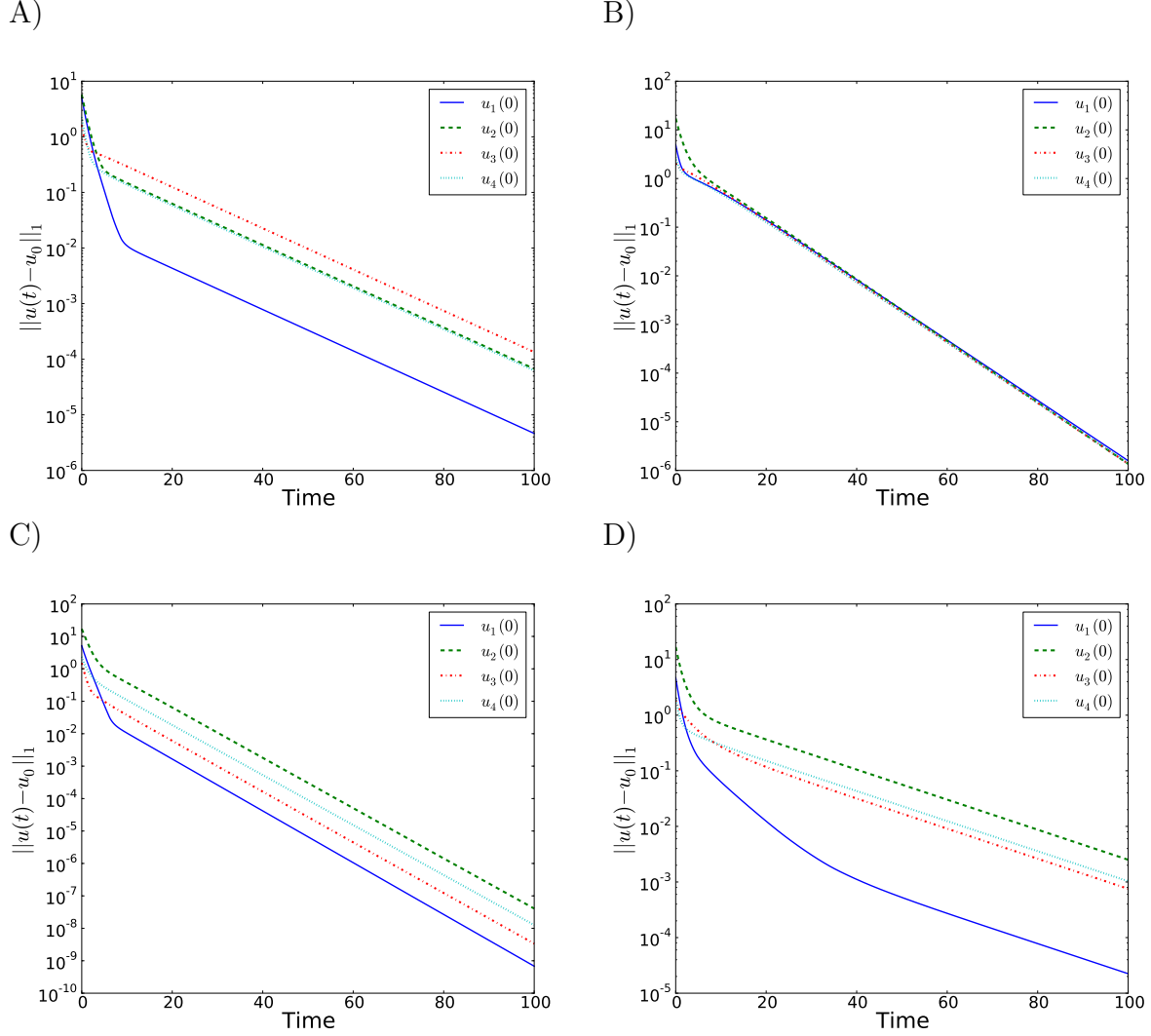


Figure 3.3: The observed L^1 distances from equilibrium as a function of time are plotted here on a log scale. The four lines on each plot each correspond to the solution starting with a different initial condition. A) \mathcal{B}_1 and κ_1 . B) \mathcal{B}_1 and κ_2 . C) \mathcal{B}_2 and κ_1 . D) \mathcal{B}_2 and κ_2 .

3.6 PERTURBATION EXPANSION

Another form of evidence for the stability of the zeroth-order branching equation is to show the stability of a first-order perturbation expansion. For this section, we will again consider functions $u \in L^1$. However, we will again appeal to Hilbert space techniques when necessary.

Assume that the capping function can be written out as:

$$\kappa(\theta) = c + \varepsilon\phi(\theta) \quad (3.39)$$

where ϕ is smooth, has integral zero, and reasonably small so that $\varepsilon\phi$ is close to zero. The equation of motion is thus:

$$\dot{u}(\theta, t) = \frac{(\mathcal{B} \star u)(\theta, t)}{\int u(\omega, t) d\omega} - (c + \varepsilon\phi(\theta))u(\omega, t) \quad (3.40)$$

For the Cauchy problem with $u(\theta, 0) = u^*(\theta)$, we will look for solutions of the form:

$$u(\theta, t) = \sum_{j=0}^{\infty} \varepsilon^j u_j(\theta, t) \quad (3.41)$$

with the initial conditions $u_0(\theta, 0) = u^*$ and $u_j(\theta, 0) = 0$ for all $j \geq 1$. Showing that both u_0 and u_1 converge to the equilibrium defined in equation (3.31) provides some evidence for the stability of the full equation.

First, we will calculate the first couple of terms of the equilibrium using equation (3.31):

$$\frac{(\mathcal{B} \star w)(\theta)}{1 + \varepsilon\phi(\theta)} = \mu w(\theta) \quad (3.42)$$

where we have the two expansions:

$$w(\theta) = \sum_{j=0}^{\infty} \varepsilon^j w_j(\theta) \quad \mu = \sum_{j=0}^{\infty} \varepsilon^j \mu_j \quad (3.43)$$

Expanding out the terms in (3.42) gives:

$$\begin{aligned} \left(\sum_{j=0}^{\infty} \frac{(-\varepsilon\phi)^j}{c^{j+1}} \right) \left(\sum_{j=0}^{\infty} \varepsilon^j \mathcal{B} \star w_j \right) &= \left(\sum_{j=0}^{\infty} \varepsilon^j \mu_j \right) \left(\sum_{j=0}^{\infty} \varepsilon^j w_j \right) \\ \sum_{j=0}^{\infty} \varepsilon^j \sum_{k=0}^j \frac{(-\phi)^k}{c^{k+1}} \mathcal{B} \star w_{j-k} &= \sum_{j=0}^{\infty} \varepsilon^j \sum_{k=0}^j \mu_j w_{j-k} \end{aligned} \quad (3.44)$$

The first term (ε^0) is simply the eigenproblem for the unperturbed problem:

$$\mathcal{B} \star w_0 = c\mu_0 w_0 \quad (3.45)$$

The only positive solution of the above is $\mu_0 = \frac{1}{c}$ and w_0 constant. Since we know the equilibrium has the same integral as the eigenvalue, we know that $w_0 = \frac{1}{c2\pi}$. The ε^1 term gives:

$$\frac{\mathcal{B} \star w_1}{c} - \frac{\phi \mathcal{B} \star w_0}{c^2} = \mu_0 w_1 + \mu_1 w_0 \quad (3.46)$$

Filling in the known quantities and rearranging gives:

$$\mathcal{B} \star w_1 - w_1 = \frac{\phi}{c} + \frac{\mu_1}{2\pi} \quad (3.47)$$

The left hand side is a self-adjoint Fredholm operator with nullspace spanned by the constant function, w_0 . By the Fredholm alternative, we know that (3.47) is solvable if and only if the right hand side is orthogonal to w_0 , i.e. $\phi(\theta) + \frac{\mu_1}{2\pi}$ has integral zero. To have integral zero, we know that $\mu_1 = -\frac{2\pi}{c} \int \phi(\omega) d\omega = 0$.

We can write out w_1 in terms of the Neumann series. By hypothesis, ϕ is bounded, and therefore $\phi \in L^2$. We know that $\mathcal{B} \star$ has operator norm strictly less than one on the space of functions orthogonal to w_0 : $\{w_0\}^\perp \subsetneq L^2$, which implies the Neumann series converges in L^2 norm:

$$w_1 = \sum_{j=0}^{\infty} \mathcal{B}^j \phi \quad (3.48)$$

where $Bu = \mathcal{B} \star u$ and $B^{j+1}u = \mathcal{B} \star B^j u$. The above expansion implies that $\int w_1 = 0$. We know that $\int Bu = \int u$ which gives:

$$\int \sum_{j=0}^n (\mathcal{B}^j \phi)(\theta) d\theta = 0 \quad (3.49)$$

Since norm convergence implies weak convergence, we can finish the proof by observing:

$$\lim_{n \rightarrow \infty} \int w_1 - \int \sum_{j=0}^n \mathcal{B}^j \phi = \lim_{n \rightarrow \infty} \left\langle w_1(\theta) - \sum_{j=0}^n \mathcal{B}^j \phi, 1 \right\rangle = 0 \quad (3.50)$$

We can now consider the dynamics. We can write out (3.40) in terms of our power series:

$$\sum_{j=0}^{\infty} \varepsilon^j \dot{u}_j(\theta, t) = \frac{1}{\sum_{j=0}^{\infty} \varepsilon^j \int u_j(\omega, t) d\omega} \sum_{j=0}^{\infty} \varepsilon^j (\mathcal{B} \star u_j)(\theta, t) - (c + \varepsilon \phi(\theta)) \sum_{j=0}^{\infty} \varepsilon^j u_j(\theta, t) \quad (3.51)$$

A formal treatment of the integrand could be considered, but without confidence of convergence, that seems unnecessary. We are only calculating u_0 and u_1 , so we will ignore terms of $o(\varepsilon^2)$ or higher in the integrand:

$$\sum_{j=0}^{\infty} \varepsilon^j \dot{u}_j(\theta, t) = \left(\sum_{j=0}^{\infty} \frac{(-\varepsilon \int u_1(\omega, t) d\omega)^j}{(\int u_0(\omega, t) d\omega)^{j+1}} \right) \left(\sum_{j=0}^{\infty} \varepsilon^j (\mathcal{B} \star u_j)(\theta, t) \right) - (c + \varepsilon \phi(\theta)) \sum_{j=0}^{\infty} \varepsilon^j u_j(\theta, t) \quad (3.52)$$

The above equation allows us to solve for the first two terms of the perturbation expansion.

The equation for the first term $u_0(\theta, t)$ from the ε^0 expansion is:

$$\dot{u}_0(\theta, t) = \frac{(\mathcal{B} \star u_0)(\theta, t)}{\int u_0(\omega, t) d\omega} - u_0(\theta, t) \quad (3.53)$$

We can explicitly solve for the time-dependent total density by observing:

$$\frac{\partial}{\partial t} \int u(\omega, t) d\omega = 1 - \int u(\omega, t) d\omega \quad (3.54)$$

Solving the above gives:

$$\int u(\omega, t) d\omega = 1 + A \exp[-t] \quad (3.55)$$

where $A = \int u^*(\omega) d\omega$. Substituting in to (3.53) gives:

$$\dot{u}_0(\theta, t) = \frac{(\mathcal{B} \star u_0)(\theta, t)}{1 + A \exp[-t]} - u_0(\theta, t) \quad (3.56)$$

The right hand side of the above clearly depends upon the denominator $1 + A \exp[-t]$ continuously in almost any operator topology. Since we are primarily concerned with the asymptotic dynamics, it is sufficient to show that the asymptotic equation converges:

$$\dot{u}_0 = \mathcal{B} \star u_0 - u_0 \quad (3.57)$$

From the results in Theorem 1, we know that the above equation converges to a multiple of the principal eigenfunction w_0 . Equation (3.55) implies that the total density is equal to $\frac{1}{c}$ and $u_0(\theta, t) \rightarrow w_0(\theta)$ in norm as was required.

The second term of the expansion (ε^1) is slightly more complicated:

$$\dot{u}_1(\theta, t) = \frac{(\mathcal{B} \star u_1)(\theta, t)}{\int u_0(\omega, t) d\omega} - \frac{\int u_1(\omega, t) d\omega}{(\int u_0(\omega, t) d\omega)^2} (\mathcal{B} \star u_0)(\theta, t) - u_1(\theta, t) - \phi(\theta)u_0(\theta, t) \quad (3.58)$$

Substituting in known quantities gives:

$$\dot{u}_1 = \frac{\mathcal{B} \star u_1}{1 + A \exp[-t]} - \frac{\int u_1}{(1 + A \exp[-t])^2} \mathcal{B} \star u_0 - u_1 - \phi u_0 \quad (3.59)$$

We can now solve for the integral of u_1 :

$$\begin{aligned} \frac{\partial}{\partial t} \int u_1 &= \frac{\int \mathcal{B} \star u_1}{1 + A \exp[-t]} - \frac{\int u_1}{1 + A \exp[-t]} - \int u_1 - \int \phi u_0 \\ &= - \int u_1 - \int \phi u_0 \end{aligned} \quad (3.60)$$

That gives the explicit solution:

$$\int u_1(\omega, t) d\omega = - \exp[-t] \int_0^t \exp[s] \int \phi(\omega) u_0(\omega, t) d\omega ds \quad (3.61)$$

Weak convergence of $u_0 \rightarrow w_0$ is sufficient to show that $\int u_1 \rightarrow - \int \phi$. Substituting the asymptotic forms into (3.58) similar to the first term gives:

$$\dot{u}_1 = \mathcal{B} \star u_1 - u_1 + \int \phi - \frac{\phi}{2\pi} \quad (3.62)$$

asymptotically. Finally, to show convergence, we can write u_1 as:

$$u_1(\theta, t) = w_1(\theta) + \epsilon(\theta, t) \quad (3.63)$$

where $\epsilon(\theta, 0) = -w_1(\theta)$. Putting that form into (3.62) gives:

$$\dot{\epsilon} = \mathcal{B} \star \epsilon - \epsilon \quad (3.64)$$

Inspection of (3.47) shows that $w_1(\theta)$ is bounded and is therefore in L^2 . Since $\mathcal{B} \star u - u$ has the Fourier eigenpairs $\{(\gamma_n, \frac{1}{\sqrt{2\pi}} \exp[-in\theta])\}_{n \in \mathbb{N}}$ as an orthonormal basis, we can write the L^2 norm of ε as:

$$\|\varepsilon(\theta, t)\|_2^2 = \sum_{n \in \mathbb{N}} \exp[\gamma_n t] |\hat{w}_1(n)|^2 \quad (3.65)$$

where $\hat{w}_1(n)$ is the n 'th Fourier component of $w_1(\theta)$. We know that $\gamma_n < 0$ for all $n > 0$ ¹ and $\sup_{n \geq 1} \gamma_n < 0$. Combining those facts with the fact that $\hat{w}_1(0) = 0$ because $\int w_1 = 0$ gives:

$$\|\varepsilon(\theta, t)\|_2^2 = \sum_{n \geq 1} \exp[\gamma_n t] |\hat{w}_1(n)|^2 \leq \exp \left[\left(\sup_{n \geq 1} \gamma_n \right) t \right] \sum_{n \geq 1} |\hat{w}_1(n)|^2 \quad (3.66)$$

That implies that $u_1(\theta, t) \rightarrow w_1(\theta)$ asymptotically exponentially.

We have now shown that the first two components of the perturbation expansion of (3.40) converge to an equilibrium. Calculating further terms would not provide any additional insight as the estimate from equation 3.44 gives a worse estimate after the first order. Recall from the proof of proposition 1 that the integral of the equilibrium w has to be equal to the eigenvalue μ . However, it is easy to observe that the argument showing that $\int w_1 = 0$ holds for all w_j , which implies that $\int \sum_{j=0}^k \varepsilon^j w_j = \frac{1}{c}$ for all k . However, $\sum_{j=0}^k \varepsilon^j \mu_j \neq \mu_0 = \frac{1}{c}$ for all $k \geq 2$. We can show that by calculating the second term in the perturbation expansion.

Using equation 3.44, we can gather the ε^2 terms:

$$\frac{1}{c} \mathcal{B} \star w_2 - \frac{\phi}{c^2} \mathcal{B} \star w_1 + \frac{\phi^2}{c^3} \mathcal{B} \star w_0 = \mu_0 w_2 + \mu_1 w_1 + \mu_2 w_0 \quad (3.67)$$

Observe that $\mathcal{B} \star w_1 = w_1 + \frac{\phi}{c}$, and filling in other known quantities gives the relation:

$$\begin{aligned} \frac{1}{c} (\mathcal{B} \star w_2 - w_2) &= \frac{\phi}{c^2} \mathcal{B} \star w_1 - \frac{\phi^2}{c^3} + \mu_2 w_0 \\ \mathcal{B} \star w_2 - w_2 &= \frac{\phi^2}{c^2} - \frac{\phi^2}{c^2} + \frac{\phi w_1}{c} + \frac{\mu_2}{2\pi} \\ \mathcal{B} \star w_2 - w_2 &= \frac{\phi w_1}{c} + \frac{\mu_2}{2\pi} \end{aligned} \quad (3.68)$$

It is obvious that $\mu_2 = 0 \iff \int \phi w_1 = 0$. However, we can observe that:

$$\begin{aligned} \int \frac{\phi w_1}{c} &= \int \phi (\mathcal{B} - \mathcal{I})^{-1} \frac{\phi}{c} = \left\langle \phi, (\mathcal{B} - \mathcal{I})^{-1} \frac{\phi}{c} \right\rangle \\ &= \left\langle (\mathcal{B} - \mathcal{I}) w_1, (\mathcal{B} - \mathcal{I})(\mathcal{B} - \mathcal{I})^{-1} w_1 \right\rangle = \left\langle (\mathcal{B} - \mathcal{I}) w_1, w_1 \right\rangle \end{aligned} \quad (3.69)$$

¹the γ_n 's are real since \mathcal{B} is symmetric.

by observing that $(B - I)w_1 = \frac{\phi}{c}$. We can use the Cauchy-Schwarz inequality to show:

$$\langle Bw_1, w_1 \rangle - \langle w_1, w_1 \rangle \leq \|Bw_1\|_2 \|w_1\|_2 - \|w_1\|_2^2 < 0 \quad (3.70)$$

The last inequality comes from the fact that $\|Bf\|_2 \leq \|B\| \|f\|_2 = \|f\|_2$ and that equality holds if and only if $f \equiv C$.

An explicit example can give insight into the non-convergence discussed in the previous paragraph. Assume $\mathcal{B} = \frac{1}{2\pi}$ and $\kappa = 1 + \varepsilon\phi$ where $|\varepsilon\phi| < 1$. It is easy to see that the principle solution to 3.42 is $w = \frac{1}{1+\varepsilon\phi}$ and $\mu = \int \frac{1}{1+\varepsilon\phi}$. Moreover, we have the exact expansion in ε :

$$w = \sum_{j=0}^{\infty} \varepsilon^j w_j = \sum_{j=0}^{\infty} \varepsilon^2 (-\phi)^j \quad (3.71)$$

By our hypothesis $|\varepsilon\phi| < 1$, we know the above converges. It is easy to see that:

$$(B - I)^{-1} \phi = - \sum_{j=0}^{\infty} B^j \phi = -\phi \quad (3.72)$$

which is the exact solution for the first term of the asymptotic expansion. However, the second term gives an incorrect solution:

$$(B - I)^{-1} \left(\phi^2 - \int \phi^2 \right) = \sum_{j=0}^{\infty} B \left(\int \phi^2 - \phi^2 \right) = \int \phi^2(\omega) d\omega - \phi^2(\theta) \neq \phi^2(\theta) \quad (3.73)$$

Further terms would show the same difficulty.

3.7 CALCULATING EQUILIBRIUM DISTRIBUTIONS

The results in this chapter justify the use of a naive eigenvalue calculation algorithm. Calculating eigenvalues of integral equations is a non-trivial problem. Investigation into open questions regarding the generality of orientation patterns across branching and capping patterns, such as in [103], may require calculating the equilibrium solution to equations like the ones analyzed here. Moreover, in the previous section, equilibrium distributions were calculated a priori to show that simulations converged. The method below has proven to be very efficient for the work in this chapter.

We will consider calculating the leading eigenvalue of the equilibrium operator for zeroth-order branching, A' . As A' is self-adjoint and compact, we can represent its range as the sum of eigenfunctions. We can explicitly calculate the n -th iterate of A' in terms of its (orthonormal) eigenfunctions:

$$\left(A'\right)^n v = \left(A'\right)^n \left(\sum_j c_j u_j\right) = \sum_j \mu_j^n c_j u_j \quad (3.74)$$

where $c_j = \langle v, u_j \rangle_\kappa$. We know that μ_0 is equal to the spectral radius from the Krein-Rutman theorem as in step 9 in the proof of theorem 1. That step also gives that the eigenvalue is simple. Finally, since A' is a compact operator, we know that there must be a spectral gap, i.e. $\mu_0 - |\mu_j| > c > 0$ for some c and all $j \neq 0$.

All that remains necessary to show that the above iteration converges to the positive equilibrium is to show that $c_j = \langle v, u_0 \rangle_\kappa \neq 0$. If v equals the constant function, that condition is fulfilled. However, a stronger result is possible. By a result in [24], we know that $u_0(\theta) > 0$. The continuity of u_0 gives that $\inf u_0 > 0$. Thus, we have the inequality:

$$\langle v, u_0 \rangle_\kappa \geq \inf_{\theta \in S^1} (u_0(\theta) \kappa(\theta)) \int_{S^1} v(\theta) d\theta > 0 \quad (3.75)$$

which implies that the iterative procedure will converge.

3.8 PROGRESS ON ZERO-ORDER STABILITY

The current work I have been doing towards proving that the unique equilibrium to the zeroth-order branching equation is stable relies on showing that the spectrum of the linearization around u_0 has strictly negative real part.

The first step is to prove the Fréchet derivative is:

$$DG_u(v) = \frac{(\mathcal{B} \star v)(\theta)}{\int_{S^1} u(\omega) d\omega} - \kappa(\theta)v(\theta) - \int_{S^1} v(\omega) d\omega \frac{(\mathcal{B} \star u)(\theta)}{\left(\int_{S^1} u(\omega) d\omega\right)^2} \quad (3.76)$$

for $\int u \neq 0$.

Proof. Observe that for sufficiently small $\varepsilon(\theta) \in L^1$, we have the series:

$$\frac{1}{\int_{S^1} u(\omega) + \varepsilon(\omega) d\omega} = \sum_{j=0}^{\infty} \frac{\left(-\int_{S^1} \varepsilon(\omega) d\omega\right)^j}{\left(\int_{S^1} u(\omega) d\omega\right)^{j+1}}$$

To prove the derivative, we show the following:

$$\lim_{\|\varepsilon\|_1 \rightarrow 0} \frac{\|G(u + \varepsilon) - G(u) - DG_u(\varepsilon)\|_1}{\|\varepsilon\|_1} = 0$$

Expanding out the first term gives:

$$G(u + \varepsilon) = \sum_{j=0}^{\infty} \frac{\left(-\int_{S^1} \varepsilon(\omega) d\omega\right)^j}{\left(\int_{S^1} u(\omega) d\omega\right)^{j+1}} \left((\mathcal{B} \star u)(\theta) + (\mathcal{B} \star \varepsilon)(\theta) \right) - \kappa(\theta)(u(\theta) + \varepsilon(\theta))$$

Subtracting out $G(u)$ gives:

$$\begin{aligned} G(u + \varepsilon) - G(u) &= \sum_{j=1}^{\infty} \frac{\left(-\int_{S^1} \varepsilon(\omega) d\omega\right)^j}{\left(\int_{S^1} u(\omega) d\omega\right)^{j+1}} \left((\mathcal{B} \star u)(\theta) + (\mathcal{B} \star \varepsilon)(\theta) \right) \\ &\quad + \frac{(\mathcal{B} \star \varepsilon)(\theta)}{\int_{S^1} u(\omega) d\omega} - \kappa(\theta)\varepsilon(\theta) \end{aligned}$$

Subtracting out $DG_u(\varepsilon)$ using our definition from above:

$$\begin{aligned} G(u + \varepsilon) - G(u) - DG_u(\varepsilon) &= \sum_{j=2}^{\infty} \frac{\left(-\int_{S^1} \varepsilon(\omega) d\omega\right)^j}{\left(\int_{S^1} u(\omega) d\omega\right)^{j+1}} \left((\mathcal{B} \star u)(\theta) + (\mathcal{B} \star \varepsilon)(\theta) \right) \\ &\quad - \frac{\int_{S^1} \varepsilon(\omega) d\omega}{\left(\int_{S^1} u(\omega) d\omega\right)^2} (\mathcal{B} \star \varepsilon)(\theta) \end{aligned}$$

By hypothesis on \mathcal{B} , we have the inequality:

$$\|\mathcal{B} \star u\|_1 \leq \|u\|_1$$

Formally writing out the norm in terms of the series gives:

$$\begin{aligned} \|G(u + \varepsilon) - G(u) - DG_u(\varepsilon)\|_1 &\leq \sum_{j=2}^{\infty} \frac{\|\varepsilon\|_1^j}{\|u\|_1^{j+1}} (\|u\|_1 + \|\varepsilon\|_1) + \frac{\|\varepsilon\|_1^2}{\|u\|_1^2} \\ &= \left(\frac{\|\varepsilon\|_1^3}{\|u\|_1^3} + \frac{\|\varepsilon\|_1^2}{\|u\|_1^3} \right) \sum_{j=0}^{\infty} \left(\frac{\|\varepsilon\|_1}{\|u\|_1} \right)^j + \frac{\|\varepsilon\|_1^2}{\|u\|_1^2} \\ &= \left(\frac{\|\varepsilon\|_1^3}{\|u\|_1^3} + \frac{\|\varepsilon\|_1^2}{\|u\|_1^3} \right) \frac{1}{1 - \frac{\|\varepsilon\|_1}{\|u\|_1}} + \frac{\|\varepsilon\|_1^2}{\|u\|_1^2} \end{aligned}$$

where the equalities hold for sufficiently small $\|\varepsilon\|_1$. We can now evaluate the limit by observing:

$$\lim_{\|\varepsilon\|_1 \rightarrow 0} \frac{\|G(u + \varepsilon) - G(u) - DG_u(\varepsilon)\|}{\|\varepsilon\|_1} \leq \lim_{\|\varepsilon\|_1 \rightarrow 0} \left(\frac{\|\varepsilon\|_1^2}{\|u\|_1^3} + \frac{\|\varepsilon\|_1}{\|u\|_1^3} \right) + \frac{\|\varepsilon\|_1}{\|u\|_1^2} = 0$$

which finishes the proof. □

Since DG_{u_0} is bounded, the spectrum $\sigma(DG_{u_0})$ is closed and bounded, and we only need to prove that all γ with $\Re \gamma \geq 0$ is in the resolvent set. An identical argument to step 3 in the proof of theorem 1 shows that any elements of the spectrum outside of $[-\sup \kappa, -\inf \kappa]$ are eigenvalues. I have shown that the following regions do not contain any eigenvalues. The proofs rely on the following lemma:

Lemma 3. *μ such that $\mu \neq 0$ and $\Re \mu \geq 0$ is in the spectrum if and only if v defined as:*

$$v = \left(DG_{u_0} - \mu I \right)^{-1} \kappa u_0$$

has integral one.

First, a quick simplification of the expression for DG_{u_0} . For the case DG_{u_0} , without loss of generality assume $\int u_0 = 1$, we have the following expression:

$$\begin{aligned} DG_u(v) &= \frac{(\mathcal{B} \star v)(\theta)}{\int_{S^1} u(\omega) d\omega} - \kappa(\theta)v(\theta) - \int_{S^1} v(\omega) d\omega \frac{(\mathcal{B} \star u)(\theta)}{(\int_{S^1} u(\omega) d\omega)^2} \\ &= (\mathcal{B} \star v)(\theta) - \kappa(\theta)v(\theta) - \left(\int_{S^1} v(\omega) d\omega \right) \kappa(\theta)u_0(\theta) = Av - \langle v, 1 \rangle \kappa u_0 \end{aligned} \quad (3.77)$$

We can now prove the lemma:

Proof. First, observe that any eigenfunctions with eigenvalue as in the statement of the lemma must have non-zero integral. Assume the contrary. From (3.77), we have:

$$Av = \mu v \quad (3.78)$$

However, from theorem 1, we know the spectrum of A is real and has largest eigenvalue 0, in contradiction to the assumption. We can therefore assume any eigenfunction has non-zero integral.

Fix μ as in the statement of the lemma. Assume that $v = (A - \mu I)^{-1} \kappa u_0$ has integral one. That implies:

$$Av - \mu v = \kappa u_0 = \int v(\omega) d\omega \kappa u_0 \quad (3.79)$$

That implies that μ, v is an eigenpair.

Assume that μv is an eigenpair and $\int (A - \mu I)^{-1} \kappa u_0 \neq 1$. Define the constant C as:

$$C = \int_{S^1} (A - \mu I)(\kappa u_0)(\omega) d\omega \quad (3.80)$$

Since v is an eigenfunction, we know:

$$v = (A - \mu I)^{-1} \left(\int v(\omega) d\omega \right) \kappa u_0 \quad (3.81)$$

Integrating both sides gives:

$$\int v(\omega) d\omega = \left(\int v(\omega) d\omega \right) \int (A - \mu I)^{-1} (\kappa u_0)(\theta) d\theta = C \int v(\omega) d\omega \quad (3.82)$$

The above is solvable if and only if $C = 1$ or $\int v(\omega) d\omega = 0$. Since we have already excluded the latter case, we can conclude $C = 1$. \square

The last calculation needed before proving the results is to split apart the real and imaginary parts of any eigenfunctions v . Direct calculation shows that we have the two equations:

$$\begin{aligned} A\Re v - \Re\mu\Re v + \Im\mu\Im v &= \kappa u_0 \\ A\Im v - \Re\mu\Im v - \Im\mu\Re v &= 0 \end{aligned} \tag{3.83}$$

Whenever $\Re\mu \neq 0$, we can solve the second equation since $A - \Re\mu I$ is invertible:

$$\Im v = \Im\mu \left(A - \Re\mu I \right)^{-1} \Re v \tag{3.84}$$

Writing $w = \Re v$, $a = \Re\mu$ and $b = \Im\mu$, we have the following real equation for w :

$$Aw - aw + b^2 \left(A - aI \right)^{-1} w = \kappa u_0 \tag{3.85}$$

μ **with** $\Im\mu = 0$

Proof. When $b = \Im\mu = 0$, we can directly solve (3.85) for w :

$$\begin{aligned} \mathcal{B} \star w - \kappa w - aw &= \kappa u_0 \\ \frac{\mathcal{B} \star w}{\kappa + a} - w &= \frac{\kappa}{\kappa + a} u_0 \\ w &= - \left(I - \frac{\mathcal{B} \star w}{\kappa + a} \right)^{-1} \frac{\kappa u_0}{\kappa + a} \\ w &= - \sum_{j \in \mathbb{N}} B_a^j \frac{\kappa u_0}{\kappa + a} \end{aligned} \tag{3.86}$$

where $B_a v = \frac{\mathcal{B} \star v}{\kappa + a}$. The last equality can be derived from the fact that the spectral radius $\rho(B_a) < 1$. Since B_a maps positive functions to positive functions, we can conclude that $w \leq 0$. By lemma 3, we can conclude μ is not an eigenvalue. \square

μ **such that** $(\Im\mu)\rho((A - \Re\mu I)^{-1}) \leq 1$

Proof. The above case can be expanded further. We can rewrite (3.85) as:

$$w - b^2(A - aI)^{-2}w = (A - aI)^{-1}\kappa u_0 \quad (3.87)$$

Whenever $b < \rho((A - aI)^{-1})$, we can explicitly write out the solution for w :

$$w = \sum_{j \in \mathbb{N}} (A - aI)^{-2j-1} \kappa u_0 \quad (3.88)$$

Since $(A - aI)^{-1}$ maps positive functions to negative functions, odd powers of the operator will do the same. Therefore $w \leq 0$ and μ is not an eigenvalue. This result can be extended to the case $b = \rho((A - aI)^{-1})$ by observing that the inverse operator is compact and κu_0 is not an eigenvalue. \square

μ **with** $\Re \mu = 0$

Proof. First, assume $\mu = 0$. That would imply that:

$$\frac{\mathcal{B} \star v}{\kappa} - v = \left(\int v(\omega) d\omega \right) u_0 \quad (3.89)$$

The nullspace of the right hand side is spanned by u_0 . That implies $\int v \neq 0$. However, the Fredholm alternative states that the above is only solvable if the right hand side is perpendicular to u_0 . As u_0 is not perpendicular to u_0 , the equation is not solvable.

Next, assume $\Im \mu \neq 0$. From (3.84):

$$\frac{\mathcal{B} \star \Im v}{\kappa} - \Im v = \Im \mu \Re v \quad (3.90)$$

The compact term above, A' is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\kappa$. The Fredholm alternative implies that the above is solvable if and only if $w = \Re v$ is perpendicular to u_0 , i.e. $\langle w, \kappa u_0 \rangle = 0$. That implies:

$$\langle Aw, w \rangle + b^2 \langle A^{-1}w, w \rangle = \langle \kappa u_0, w \rangle = 0 \quad (3.91)$$

However, A and A^{-1} are negative semi-definite and $\langle Aw, w \rangle = 0$ if and only if $w = u_0$, in contradiction to the hypothesis that $\langle \kappa u_0, w \rangle = 0$. \square

$|\Im \mu| > C$

Proof. If we have $\left\| \left(A - aI + (A - aI)^{-1} \right)^{-1} \right\|_1 < 1$ where the norm is the operator norm associated with L^1 , we know that there cannot be any eigenvalues since:

$$\int w(\omega) d\omega \leq \|w\|_1 \leq \left\| \left(A - aI + (A - aI)^{-1} \right)^{-1} \right\|_1 \|\kappa u_0\|_1 < 1 \quad (3.92)$$

since $\|\kappa u_0\|_1 = 1$. We can use the spectrum of $A - aI$ to approximate the operator norm. We know that if $\gamma \in \sigma(A - aI)$, then $\gamma + \frac{b^2}{\gamma} < 0$ is in the spectrum of $A - aI + (A - aI)^{-1}$. By differentiating, we can easily show that $\gamma + \frac{b^2}{\gamma} \leq -2b$ for all $\gamma < 0$. That implies that the spectral radius of $\rho \left(\left(A - aI + (A - aI)^{-1} \right)^{-1} \right) < \frac{1}{2b}$. We can show that in the Banach algebra of operators on L^1 , $B(L^1)$, that $A - aI + (A - aI)^{-1}$ is Hermitian in the sense that:

$$\left\| \exp \left[-ic \left(A - aI + (A - aI)^{-1} \right)^{-1} \right] \right\| = 1 \quad (3.93)$$

for all real c . If that Banach algebra is a C^* algebra, we then have that $\left\| \left(A - aI + (A - aI)^{-1} \right)^{-1} \right\| = \rho \left(\left(A - aI + (A - aI)^{-1} \right)^{-1} \right)$ by Sinclair's theorem [120], which gives us $C = \frac{1}{2}$. If that is not the case, we simply have that $\left\| \left(A - aI + (A - aI)^{-1} \right)^{-1} \right\| \leq \frac{\pi}{2} \rho \left(\left(A - aI + (A - aI)^{-1} \right)^{-1} \right)$ by [90]. \square

3.9 GENERALIZATION TO MEASURES

Defining equations (3.7) and (3.8) on the space L^1 was done solely for mathematical convenience. However, there is no scientific justification for why the densities must be absolutely continuous with respect to the Lebesgue measure. In particular, it is interesting to consider situations such as having initial conditions $u(\theta, 0) = \delta(\theta)$ where δ is the Dirac delta. For many diffusion processes, we could observe that solutions are regularizing and restrict ourselves to analysis to the smoothness of the regularization. However, for nonlocal diffusion equations such as the one considered here, solutions are not regularizing, and analysis must be done on whatever space the initial conditions are in. Also, the hypotheses on \mathcal{B} and κ were largely arbitrary. This section outlines a proof of proposition similar to theorem 1 on a less regular space and discusses the role of the hypotheses on \mathcal{B} and κ .

The space under consideration here is the space of finitely-additive, finite measures. There is a trade-off in using that more abstract space. You lose the machinery of working with functions but gain nicer topological properties. Also, solutions will obviously need to be defined in the weak sense. Define the space $ba(\Sigma)$ as the space of countably additive measures with respect to the Borel σ -algebra, Σ , with the usual topology on the circle. Define the space $B(\Sigma)$ to be the dual space of $ba(\Sigma)$. $B(\Sigma)$ is the space of Σ' -measurable functions with finite $\|\cdot\|_{\text{sup}}$ -norm where:

$$\|f\|_{\text{sup}} = \sup_{\theta \in S^1} |f(\theta)| \quad (3.94)$$

Note that that $B(\Sigma)$ and L^∞ are very different spaces. Up to this point, all of the discussion in this chapter has treated the L^p directly as function spaces, ignoring the fact that the spaces are quotient spaces modular functions with zero integral. That quotient means that some measures, such as the δ measure cannot be defined on the space L^∞ . ba is a Banach space with the total variation of the measure as the metric. ba is also reflexive, which means it is weakly compact. That fact opens new tools.

Working with the space ba also allows us to significantly weaken the hypotheses on \mathcal{B} and κ . For this section: $\mathcal{B} \in C^1$, $\int \mathcal{B} = 1$, κ is Hölder continuous, i.e. $\kappa \in C^{0,\alpha}$ where $\alpha \geq 1$ and $\mathcal{B}, \kappa \geq 0$. We can define A exactly as above:

$$Au = \mathcal{B} \star u - \kappa u \quad (3.95)$$

where $u \in ca$, \star is interpreted as a convolution of measures, and equality is interpreted in the weak sense. The weaker hypotheses do not lead to a slightly weaker conclusion:

Conjecture 2. *A , defined on $ba(\Sigma)$, has a principal, largest eigenvalue μ_0 with a positive, finite eigenmeasure ϕ . The eigenmeasure ϕ is absolutely continuous with respect to the Lebesgue measure and has bounded Radon-Nikodym derivative. Moreover, μ_0 is an isolated element of the spectrum $\sigma(A)$.*

Boundedness of the Radon-Nikodym derivative $\frac{d\phi}{d\nu}$ can be recovered by adding the conditions that κ is Hölder continuous with $\alpha \geq 1$ and $\kappa > 0$. A counter-example showing that $\kappa > 0$ and Hölder continuity with $\alpha < 1$ is not sufficient for the result to hold is given towards the end of the section. Most of the proof of theorem 1 still holds. Below is a sketch of how the proof may work.

Proof Sketch:

Direct calculation readily shows that A is a bounded linear operator on ba . The first steps are to characterize the structure of the spectrum. The weaker hypotheses still give that A is from L^2 to L^2 , so we can draw the same conclusions regarding the spectrum $\sigma_{ba}(A)$ for the non-residual spectrum, including that it is real. Also, $\mathcal{B} \star \cdot$ is still compact. However, showing that requires slightly more care. The range of $\mathcal{B} \star ba$ is isometric to a subset of L^∞ via the Radon-Nikodym derivative, but the range must be considered as a set of measures. The compactness of $\mathcal{B} \star$ allows us to characterize the non-residual spectrum outside of $[-\sup \kappa, -\inf \kappa]$ in the same way as before. The final step is to prove that A has no residual spectrum. Since L^2 is not necessarily dense in ba , another technique must be used. A heuristic argument can be made by decomposing any measure, u , in ba into a part continuous to the Lebesgue measure u_c and a part singular to the Lebesgue measure u_s . Since $\mathcal{B} \star v$ is always absolutely continuous with respect to the Lebesgue measure, we can solve the two equations for any $\mu \in \mathbb{C}$:

$$-(\kappa + \mu)v_s = u_s \quad \text{and} \quad \mathcal{B} \star v_c = u_c - \mathcal{B} \star v_s \quad (3.96)$$

The first equation can be solved by division when $\mu \notin [-\sup \kappa, -\inf \kappa]$, and the second is solvable, or at least approximately solvable, by the same arguments as in the original proof since $u_c - \mathcal{B} \star v_s$ is effectively a L^∞ function. More care would need to be taken for $\mu \in [-\sup \kappa, -\inf \kappa]$, if necessary.

The next step is to show the existence of the principal eigenvalue. In the original proof, κ being C^2 was required for the existence of an eigenvalue outside of the range $[-\sup \kappa, -\inf \kappa]$. A different argument can be made to relax that requirement simply to C^1 . Simple division shows that $\mu \in \sigma(A)$ if and only if 1 is in the spectrum of the operator: $A'_\mu = \frac{\mathcal{B} \star}{\kappa + \mu}$ whenever $\frac{1}{\kappa + \mu} \in B(\Sigma)$. A'_μ for $\mu > -\inf \kappa$ always has a positive eigenvalue equal to its spectral

radius by Krein-Rutman. Moreover, the spectral radius is equal to the operator norm with respect to the $\langle \cdot, \cdot \rangle_{\kappa+\mu}$ inner product. We can write the function $\rho(\mu)$ as the spectral radius of A'_μ . The fact that Gelfand's formula uniformly converges as a function of μ gives that ρ is continuous. We know that μ is an eigenvalue of A whenever $\rho(\mu) = 1$. The principal eigenvalue is thus the largest value μ where $\rho(\mu) \geq 1$. Therefore, it is sufficient to show that $\rho(\mu) \geq 1$ for at least one $\mu > -\inf \kappa$. However, by the definition of $\rho(\mu)$, we have:

$$\rho(\mu) \geq \left\langle A'_\mu \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}} \right\rangle = \frac{1}{2\pi} \frac{1}{\kappa + \mu} \quad (3.97)$$

However, the last quantity goes to infinity as $\mu \rightarrow -\inf \kappa$ since $\kappa \in C^{0,\alpha}$. We now show the existence of a principal eigenvalue with a positive eigenmeasure.

The final step is to show that the eigenvalue is simple. De Masi et al. showed that the Markov chain generated by A' was ergodic, which implied that the associated eigenfunction must be simple. Without having worked through the details, it would seem likely that the same argument would hold for this case as smoothness does not appear to be used in that proof.

Other Hypotheses

As mentioned before, κ being Hölder continuous is a necessary hypothesis for the existence of an eigenvalue. Take the family of systems where $\mathcal{B} = \frac{1}{2\pi}$ and $\kappa = c|\theta|^\alpha + c'$ where $c, c' > 0$ and $0 < \alpha < 1$. Obviously, κ is α -Hölder continuous. Assume we have an eigenpair $\mu, v(\theta)$. Dividing through wherever $v(\theta) \neq 0$, we have:

$$v = \frac{(\mathcal{B} \star v)(\theta)}{c|\theta|^\alpha + c + \mu} = \frac{\int v(\omega) d\omega}{c|\theta|^\alpha + c' + \mu} \quad (3.98)$$

That means $\int v(\omega) d\omega = 1$ and $v = \frac{1}{c|\theta|^\alpha + c' + \mu}$. However, if we set:

$$c = \frac{1}{2} \int_{S^1} |\theta|^{-\alpha} d\theta \quad (3.99)$$

the above equations are not solvable for any pair c, μ since $\int v(\omega) d\omega \leq \frac{1}{2}$. That implies that there is not any eigenvalues for A .

In [24], compactness of A' was used to prove the existence of a principal eigenvalue. When we relax the hypotheses on \mathcal{B} , we cannot directly use the results from Krein-Rutman nor [24]

because A' may not be compact. However, for the space $ba(\Sigma)$, we have weak compactness from the space. Schauder-Tykhonov can then be used to give the existence of a positive eigenvalue. If simplicity of the eigenvalue and the spectral gap can also be recovered without compactness, the far weaker hypotheses $\mathcal{B} \in B(\Sigma)$, $\kappa \in C^{0,\alpha}$ $\alpha \geq 1$, $\mathcal{B} \geq 0$, and $\kappa > 0$ can be used.

3.10 DISCUSSION

The results presented here form a framework to study continuum models of actin filament orientation density and formally justify some of the numerical calculations reported in the literature. For the first-order branching equation, (3.7), theorem 1 completely characterizes the dynamics. Solutions converge exponentially such that $u(\theta, t) \sim A \exp[\mu_0 t] u_0(\theta)$. While solutions to the zeroth-order branching equation, (3.8), have not been completely characterized, the results in section 3.4 do greatly limit the possible behaviors. The simulations in section 3.5 provide strong evidence that solutions converge to a unique equilibrium. The results seen in section 3.8 provide a strong start to proving that the unique equilibrium to the zeroth order branching equation is at least locally stable.

The results from Maly and Borisy [68] hold up surprisingly well. The eigenvalues and eigenfunctions they calculated make sense if and only if the results in theorems 1 and 2 hold for their eigenproblem. It is not obvious that the steady state, when it exists, would be from a simple eigenvalue. Without both results, the eigenfunctions calculated would not necessarily be unique. Finally, a modified version of their suggested concept of orientation pattern fitness turns out to be mathematically justified. Maly and Borisy attempted to define a fitness value for each particular orientation θ . Casual inspection of (3.7) shows that such a function is impossible as the entire density contributes to the dynamics at any given orientation. However, there does exist a fitness function on the total orientation pattern as seen in corollary 1.

The results relevant to Weichsel and Schwarz [137] present a more mixed picture. The Weichsel and Schwarz model can be thought of as adding detail to the time-dependent

dynamics to the Maly and Borisy model. In both papers, a steady-state assumption was used to calculate the equilibrium. The conjecture 1 being proven would demonstrate that the dynamics in the Weichsel and Schwarz model justify the steady state assumption. However, the bistable hypothesis put forth by Weichsel and Schwarz has been proven wrong. While the hypothesis may have some physical merit, as is the thrust of Weichsel et al. [138], the conclusion is not mathematically justified.

One important physical question building upon this work is to study the multiplicity of orientation patterns within potential models of branching and capping dynamics. This has been done for a discrete approximation in [103]. It is easy to show that, given the stronger hypotheses in this chapter, zero must be a local minimum or maximum. Under further hypotheses based on experimental results, it is an open question on how common one orientation pattern is over the other. Given that the current work on the topic, including the aforementioned papers and Atilgan et al. [4], only report results for one particular choice of \mathcal{B} and κ , it is interesting to investigate how solutions are distributed given some reasonably physical set of possible \mathcal{B} and κ .

Beyond the conjecture, another line of future mathematical work stems from generalizing the results here for both equations from L^1 to weak solutions over some space of finite measures as in section 3.9. While many nonlocal diffusion equations feature unbounded diffusion kernels, the equations here can be studied on the space of finite, countably-additive measures without appealing to Schwarz spaces or some other space of generalized functions which admit some notion of smoothness. As seen in the proof sketch for conjecture 2, very little of the analysis for theorem 1 uses results which cannot be generalized to the larger Banach space. Similar results for zeroth-order branching would be similarly interesting. There also remains questions regarding the smoothness required for \mathcal{B} and κ . As seen in the counter-example given in section 3.9, κ must at least be Hölder continuous with $\alpha \geq 1$. Physical considerations justify the hypothesis that κ must be continuous, but the stronger Hölder hypothesis appears to be a novel constraint. On the other hand, it appears that much weaker conditions on \mathcal{B} may be required. Those generalizations are relevant to consider cases such as in [137] where κ is Lipschitz continuous but not C^1 .

4.0 MDMS: MOLECULAR DYNAMICS META-SIMULATOR FOR EVALUATING EXCHANGE TYPE SAMPLING METHODS.

4.1 INTRODUCTION

There have been a number of computer simulation studies of actin and related proteins at the atomic level [19, 20, 22, 26, 71, 118]. These have led to some interesting experimental predictions [50]. However, these simulations involve large molecules, actin monomer simulations include $\sim 80\text{k}$ atoms and actin filament studies include $\sim 600\text{k}$ atoms [126], where it is difficult to calculate free energies and other statistical quantities, . For that reason, enhanced conformational sampling algorithms are needed. Such techniques have already been used for at least one study of actin protein folding [94]. Sampling from a protein’s conformational space remains one of the hardest and most important problems in computational biophysics. In order to calculate important properties of a protein, such as free energies or rigidity, thorough sampling of likely conformations is required. For small proteins, standard Langevin molecular dynamics (MD) or Markov chain Monte Carlo methods are sufficient [116]. However, for larger proteins, some form of enhanced sampling method becomes necessary [134].

Many conformational sampling methods for macromolecules have been developed in recent years [55, 64, 134]. With so many algorithms being developed, it is important to have standardized tests to compare the methods’ ability to search multiple local minima and preserve the desired ensemble. One popular type of enhanced sampling method is called parallel tempering [51] or replica exchange (RE) [127]. RE employs parallel Markov chain [44, 133] simulations where, neighboring simulation states are periodically exchanged. There currently does not exist a standardized model to compare sampling algorithms, [146] nor

does one exist for just exchange based RE methods. To start to fill that gap, we are proposing a class of models called *Molecular Dynamics Meta-Simulators* (MDMS) for evaluating RE methods.

RE methods use parallel simulations (replicas) run with different conditions to utilize the increased rate of transitions between local minima in the conformational phase space by modifying parameters. The most common form of RE is temperature RE. Temperature RE methods use parallel replicas run at different temperatures to utilize the increased rate of transitions between local minima in the conformational phase space at higher temperatures. The lowest temperature replica is usually the temperature of interest, while the higher temperatures are chosen to balance the faster transition rates at high temperatures and the reduced probability that two replicas far apart in temperature will exchange states [110, 127].

The effectiveness of RE methods are determined by multiple factors including the relative sampling rate of the added replicas and how much the probability distribution differs between replicas. One consideration for using RE is whether the increase in transitions between states observed in the simulation(s) run with the replica(s) of interest is great enough to outweigh the cost in additional computing hardware from the added replicas. However, RE has the advantage of being easily parallelizable, which can limit the added computing time needed compared to a more direct approach. In cases where a sufficient number of parallel nodes are available, RE can save wall clock time even if it is more expensive in terms of total CPU cycles. RE has been successfully used to enhance conformational sampling and calculate temperature dependent properties for biological systems [40, 91, 95, 111, 144]. However, for large systems with slow transition kinetics, standard temperature RE simulations remain computationally expensive. Numerous methods have been proposed to improve the effectiveness of RE simulations. [39, 62, 63, 66, 67, 83, 84, 106, 107, 128, 129]. MDMS can serve to compare and evaluate these new methods.

Even within temperature RE, there are a number of open questions about how best to use RE. Two related questions that have spurred much inquiry are what the optimal exchange frequency is [3, 121, 122] and what temperatures to choose to achieve that [92, 110]. Other studies have tried to understand RE from a more theoretical vantage point [108, 141, 142], but the mapping between those theoretical studies and atomistic models is very vague. MDMS

can serve as a useful tool in further research on all of these questions.

Any test problem must be computationally cheap enough to allow for rapid testing of multiple algorithms, have a tractable metric to compare sampling efficacy, yet retain enough physical detail to be relevant to practical simulations. Short peptides, like poly-Alanine chains, are computationally efficient to simulate as test models, but their simple topologies may mask complications. New methods developed on such peptides may not be transferable to larger, biologically relevant systems. A model which allows for complicated kinetics while remaining computationally cheap would allow for much faster testing of RE methods. While we present a simple example in this chapter, MDMS is flexible enough to readily admit complicated systems with many local minima, complicated topologies and many degrees of freedom. Additionally, the cost of adding additional complexity to an MDMS model is very small, especially compared to simulating larger molecules. In fact, MDMS can even be tuned to approximate a particular system of interest for an effective comparison of methods before running large-scale simulations. It can also be run on a standard desktop machine allowing for a fast comparison of many algorithms.

In order to be efficient, MDMS abstracts away the dynamics of a simulation between successive exchanges. At any given exchange attempt, the only important factors are a model’s state and the energy of that state. MDMS represents the state of the model by discretizing conformational space into n local minima. Then, a locally-harmonic energy approximation to generate a continuous distribution of energies when the system is in state i . The total energy of the system is the sum of the energy of the local minimum, E_i , and the energy of a system of harmonic oscillators. The resulting energy distribution is independent of the choice of specific harmonic energy function and is functionally defined by the number of degrees of freedom and one parameter, the determinant of the harmonic energy matrix.

This model can be classified as an extension to the two-state kinetic network model introduced by Zheng et. al. [141, 142]. The current model allows building a system with an arbitrary number, n , states with explicitly defined free energy landscape and barriers between minima. Since all energy values and barriers are user defined, a realistic landscape with different intermediate states and folding pathways can be easily created and simulated. Defining only two minima with a single barrier will simplify the kinetics to those described

by Zheng et al. [141, 142]. A more flexible energy landscape is essential to be able to test advanced RE methods such as reservoir RE [84, 107] where multiple free energy minima are required to construct a realistic reservoir. We do note, however, that our presentation here only includes linear activation entropy functions restricting each transition to only Arrhenius or anti-Arrhenius behavior. The model is easily extendable to non-linear free energy functions to recover the folding behavior in [141].

The model we propose is a new tool that will greatly decrease the amount of time between new RE method idea and practical code. Using this model, a developer can test several new RE methods in the space of an afternoon, compare new methods to existing algorithms, and make informed decisions for which methods should be coded into a full molecular dynamics software such as CHARMM [9] and AMBER [17]. This model will also allow for further analytical investigation of current and new RE methods.

4.2 SYSTEM DEFINITION AND BASIC PROPERTIES

Traditional replica exchange simulations will run dynamics for some number of steps in between a set of exchange attempts where the parallel replicas probabilistically exchange states. The goal in developing this model is to abstract away the dynamics between exchanges allowing for faster simulations. When testing the efficacy of replica exchange methods, nothing is to be gained from knowing the state of the system in between exchange attempts, so we constructed a system where we can exactly solve for the transition probabilities between states for the amount of simulated time between exchange attempts. MDMS dynamics are based on continuous-time discrete-state Markov chain where we solve for the discrete-time δ -skeleton.

4.2.1 Assumptions

MDMS uses a discretized approximation of an energy surface. Each discrete state represents a local minimum of the energy function and the energy function around the local minimum is

approximated by a system of harmonic oscillators. For any smooth energy function, Taylor’s theorem implies that such an approximation can be made for small displacements around the minimum. The individual states are then separated by finite energy barriers. The locally-harmonic approximation is used to allow for a simple discretization of phase space without concern to optimal partitioning (e.g. [136]). MDMS assumes essentially instant equilibration within each energy well. With this approach, MDMS is not able to study systems where the exchange frequencies are faster than the rate of convergence within an energy well. However, later in this chapter, an implementation which allows for slow convergence within the harmonic dimensions is discussed.

Transition state theory traditionally uses the free energy difference between the stable state and the energy barrier to determine the transition rates. The free energy difference is dependent both upon the enthalpy difference and the entropy difference. It is useful to note here that enthalpy is a thermodynamic quantity that cannot be defined for an individual conformation like the situation in MD simulations as the pressure of an individual conformation is not formally defined. In terms of practical MD simulations, this can be thought of as including an implicit solvation term in the potential energy or an expanded energy function which includes both the molecule of interest and the solvating environment (e.g. explicit water).

For a simple introduction, we will treat energies strictly as enthalpies. Entropic effects can be integrated by varying the number of states i with equal energies and small or zero energy barriers separating them. This first approximation leads to a large variety of behaviors as can be seen by the models outlined in Section 4.3. Later in the section, MDMS energies will be generalized to free energies.

4.2.2 Transition Rates

Transition state theory (TST) uses Arrhenius equation based rates to determine transition rates between states. The Arrhenius equation determines that rate from the size of the energy barrier between the two states. MDMS uses TST-like energy barriers that can be

directly generalized to any number of states by requiring that for all states i, j :

$$E_i + \Delta E_{ij} = E_j + \Delta E_{ji} \quad (4.1)$$

where E_i is the energy of state i and ΔE_{ij} is the energy barrier in between states i and j . The rate of transitions $i \rightarrow j$ from the Arrhenius equation is then:

$$k_{ij} = A_{ij} \exp \left[-\frac{\Delta E_{ij}}{k_B T} \right] \quad (4.2)$$

Symmetric prefactors, i.e. $A_{ij} = A_{ji}$, are necessary for a Boltzmann equilibrium. Pairs of states that are inaccessible from each other can be defined by setting either $A_{ij} = 0$ or $\Delta E_{ij} = \infty$.

Care must be taken to ensure that all pairs of states are connected through a finite number of accessible transitions to guarantee ergodicity. A model system with more complicated, non-symmetric, dynamics (e.g. $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$) could be constructed, but it would not necessarily be straightforward to ensure a Boltzmann-type temperature dependent equilibrium. Also, note that there is no requirement for $\Delta E_{ij} > 0$, meaning anti-Arrhenius behavior could be included.

From the TST rates, we construct a continuous-time discrete-state Markov system with the rates k_{ij} and $k_{ii} = -\sum_{j \neq i} k_{ij}$. Define $\mathbf{K} = (k_{ij})$. Defining $\mathbf{P}(\mathbf{x})$ as the row vector of state probabilities, the master equation for this system is thus:

$$\frac{d}{dt} \mathbf{P}(x=i) = \sum_j k_{ji} \mathbf{P}(x=j) \quad (4.3)$$

Equivalently:

$$\frac{d}{dt} \mathbf{P}(\mathbf{x}) = \mathbf{P}(\mathbf{x}) \mathbf{K} \quad (4.4)$$

Since we are interested in the system state at discrete time intervals, δt , we can model the system by the transition probabilities from time t to time $t + \delta t$. Define $\mathbf{M}_{\delta t} = \exp[\delta t \mathbf{K}]$. From (4.4), the probability vector of our system at time t given a probability vector $\mathbf{P}(x(s))$ at time s with $t > s$ is:

$$\mathbf{P}(x(t)) = \mathbf{P}(x(s)) \mathbf{M}_{t-s} \quad (4.5)$$

That solution gives the necessary tool to calculate the transition probabilities exactly. For any $t > 0$:

$$P(x(t + \delta t)) = P(x(t))M_{\delta t} \quad (4.6)$$

It is important to note that the above equation has transition matrix independent of t . When running simulations, we will use that fact by calculating $M_{\delta t}$ before starting the run. For each simulation step, however, we are not interested in general probability vectors. For every time t during the simulation, we know exactly which state the system is in. Thus, $P(x(t) = j) = \delta_{ij}$ where i is the current state of the system, and we can use the i 'th standard basis vector as our initial condition giving the convenient relation:

$$P(x(t + \delta t) = j | x(t) = i) = (\mathbf{e}_i M_{\delta t})_j = \sum_k \delta_{ik} (M_{\delta t})_{kj} = (M_{\delta t})_{ij} \quad (4.7)$$

That means that all of the relevant properties of our system can then be derived from the matrices K and $M_{\delta t}$.

4.2.3 Equilibrium Distribution

First, we want to show that the system has a reasonable, temperature-dependent equilibrium distribution. In fact, we want a Boltzmann distribution as a unique equilibrium, and for solutions to (4.4) to be ergodic. Specifically, we want:

$$\lim_{t \rightarrow \infty} \mathbf{p} M_t = \mathbf{b} \quad (4.8)$$

where \mathbf{b} is our Boltzmann distribution and \mathbf{p} is a probability vector: $p_j \geq 0$ and $\sum_j p_j = 1$. Basic linear algebra tells us that if K has zero as a simple eigenvalue with eigenvector \mathbf{b} and all other eigenvalues are negative, we have the following:

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbf{p} M_t &= \lim_{t \rightarrow \infty} \mathbf{p} U \text{diag}\{1, \exp[\mu_1 t], \dots, \exp[\mu_{n-1} t]\} U^* \\ &= \mathbf{p} U \text{diag}\{1, 0, \dots, 0\} U^* \\ &= \mathbf{p} Q_{\mathbf{b}} = c \mathbf{b} \end{aligned} \quad (4.9)$$

where U is a unitary matrix, $\text{diag}\{\dots\}$ is a diagonal matrix, $Q_{\mathbf{b}}$ is a projection onto \mathbf{b} , and for some constant c . We can show that c equals 1 by observing:

$$\frac{d}{dt} \sum_j (\mathbf{p}M_t)_j = \sum_j (\mathbf{p}K)_j = \sum_j \sum_i k_{ji} p_i = \sum_j \sum_{i \neq j} k_{ji} p_i - \sum_j \sum_{i \neq j} k_{ji} p_i = 0 \quad (4.10)$$

by the definition of k_{jj} . Since $\sum_j b_j = \sum_j p_j = 1$ by construction, c must equal 1.

Since the rows of K sum to zero, we know K has a zero eigenvalue. We want to show that a Boltzmann-type distribution is associated with that eigenvalue. Define the Boltzmann distribution as:

$$\mathbf{b}(x = i) = \frac{1}{\mathcal{Z}(T)} \exp \left[-\frac{E_i}{k_B T} \right] \quad (4.11)$$

$\mathcal{Z}(T)$ is the partition function:

$$\mathcal{Z}(T) = \sum_j \exp \left[-\frac{E_j}{k_B T} \right] \quad (4.12)$$

Showing that \mathbf{b} is in the kernel of K is straightforward:

$$\begin{aligned} (\mathbf{b}K)_i &= \sum_j b_j k_{ji} = - \sum_{j \neq i} b_i k_{ij} + \sum_{j \neq i} b_j k_{ji} \\ &= - \sum_{j \neq i} A_{ij} \exp \left[-\frac{\Delta E_{ij} + E_i}{k_B T} \right] + \sum_{j \neq i} A_{ij} \exp \left[-\frac{\Delta E_{ji} + E_j}{k_B T} \right] = 0 \end{aligned} \quad (4.13)$$

where the last equality comes because of the TST relation in (4.1).

It remains to show that the eigenvalue is simple, and that all of the other eigenvalues have negative real part. Uniqueness is assured as long as the graph of the transition rates is irreducible [42]. In other words, for every pair of states, (i, k) , there must be a path of transitions between i and k with finitely many states and finite energy barriers. An application of the Perron-Frobenius theorem gives shows all the other eigenvalues have negative real part [42], finishing the proof of (4.8).

The TST relation also ensures the detailed balance condition:

$$b_j k_{ji} = b_i k_{ij} \quad (4.14)$$

for all i and j .

4.2.4 Limits

To be physically reasonable, MDMS should display accurate limiting behavior. For example, the probability of the system jumping from state i to another state j as a function of the temperature should go to zero as the temperature goes to zero. This follows directly from the following:

$$\lim_{T \rightarrow 0} P(x(t + \delta t) = i | x(t) = i) = \lim_{T \rightarrow 0} \exp[k_{ii}\delta t] = \exp\left[\delta t \lim_{T \rightarrow 0} k_{ii}\right] = 1 \quad (4.15)$$

where k_{ii} is defined as the negative sum of the Arrhenius rates in (4.2). Indeed, the matrix M_s converges to the identity matrix as $T \rightarrow 0$ in operator norm for all $s \geq 0$.

The high temperature limit does not behave as nicely. From the Arrhenius rates, it is obvious that the transition rate is bounded above. Increasing A_{ij} is indistinguishable from increasing δt . It is important to choose $A_{ij}\delta t$ sufficiently large that higher temperature replicas will see frequent state changes. It is easy to see that $\lim_{T \rightarrow \infty} k_{ij} = A_{ij}$ for $i \neq j$ and $\lim_{T \rightarrow \infty} k_{ii} = -\sum_{j \neq i} A_{ij}$. The probability of a transition occurring during the window $(t, t + \delta t)$ is:

$$\lim_{T \rightarrow \infty} P(x(t + \delta t) \neq i | x(t) = i) = \lim_{T \rightarrow \infty} 1 - \exp[\delta t k_{ii}] = 1 - \exp\left[-\delta t \sum_{j \neq i} A_{ij}\right] \quad (4.16)$$

In practice, RE methods can only be used on systems with bounded temperature, so the lack of a physical infinite temperature limit behavior is not detrimental to the applicability of MDMS.

We are also able to take the high and low temperature limits for the equilibrium distribution. The high temperature limit is simply $\mathbf{b}(x = i) = \frac{1}{n}$. This can be seen via:

$$\lim_{T \rightarrow \infty} P(x = i) = \lim_{T \rightarrow \infty} \frac{\exp\left[-\frac{E_i}{k_B T}\right]}{\sum_j \exp\left[-\frac{E_j}{k_B T}\right]} = \frac{1}{\sum_j 1} = \frac{1}{n} \quad (4.17)$$

An ‘infinite temperature’ replica could be generated by choosing states at each time point from the uniform distribution of states. Similarly, the low temperature limit is $\mathbf{b}(x = i) = \delta_{i0}$

where E_0 is the lowest energy state and δ_{ij} is the Kronecker delta. Without loss of generality, assume E_0 is a unique energy minimum. The limit can be seen by:

$$\lim_{T \rightarrow 0} P(x = i) = \lim_{T \rightarrow 0} \frac{\exp\left[-\frac{E_i}{k_B T}\right]}{\sum_j \exp\left[-\frac{E_j}{k_B T}\right]} = \lim_{T \rightarrow 0} \frac{\exp\left[\frac{E_0 - E_i}{k_B T}\right]}{1 + \sum_{j>0} \exp\left[\frac{E_0 - E_j}{k_B T}\right]} = \frac{0}{1} \quad (4.18)$$

for $i > 0$ and:

$$\lim_{T \rightarrow 0} P(x = 0) = \lim_{T \rightarrow 0} \frac{\exp\left[-\frac{E_0}{k_B T}\right]}{\sum_j \exp\left[-\frac{E_j}{k_B T}\right]} = \lim_{T \rightarrow 0} \frac{1}{1 + \sum_{j>0} \exp\left[\frac{E_0 - E_j}{k_B T}\right]} = 1 \quad (4.19)$$

If there are k states with energy E_0 , the last limit simply becomes $\frac{1}{k}$.

4.2.5 Energies

In order for exchanges between replicas to occur, there has to be sufficient energy overlap between the replicas. For systems with many states, with a relatively dense energy distribution, energy overlap would not be a problem. However, for systems with large energy gaps or small systems with few states, some continuous energy distribution needs to be included. Approximating a system near a local minimum by harmonic oscillators is a standard approach. To each state i , we assign a $m \times m$ positive-definite matrix H_i that gives the harmonic energy function:

$$V_i(\mathbf{y}) = \mathbf{y}^T H_i \mathbf{y} \quad (4.20)$$

where m is the number of harmonic oscillators/dimension of the harmonic oscillator space and is chosen to get a reasonable spread of energies.

For this section, we will mildly abuse notation. The formally correct way to write out probabilities is:

$$P(\mathbf{x} \in \mathcal{A}) = \int_{\mathcal{A}} d\mu(\mathbf{x}) \quad (4.21)$$

All of the probability measures we are concerned with here are absolutely continuous with respect to the Lebesgue measure, so we will write:

$$P(\mathbf{x} = \mathbf{y} \, d\mathbf{y}) = \pi(\mathbf{y}) \, d\mathbf{y} \quad (4.22)$$

for our probability density $\pi(\mathbf{y})$.

We can exactly calculate the distribution of energies for a harmonic oscillator with positive-definite energy matrix H_i over \mathbb{R}^m . The probability distribution is a Gaussian distribution with pdf:

$$P(\mathbf{x} = \mathbf{y} \, d\mathbf{y}) = \frac{|H_i|^{1/2}}{(\pi k_B T)^{m/2}} \exp \left[-\frac{\mathbf{y}^T H_i \mathbf{y}}{k_B T} \right] d\mathbf{y} \quad (4.23)$$

For any energy level $E \geq 0$, we can define the surface S_E :

$$S_E := \mathbf{y} : \mathbf{y}^T H_i \mathbf{y} = E \quad (4.24)$$

That gives the expression for the probability of an energy level E :

$$P(E = F \, dF) = \frac{|H_i|^{1/2}}{(\pi k_B T)^{m/2}} \int_{S_F} \exp \left[-\frac{\mathbf{y}^T H_i \mathbf{y}}{k_B T} \right] d\mathbf{y} \quad (4.25)$$

There is no dF term in the above expression since the integral is the m -dimensional measure of a $(m-1)$ -dimensional surface, which can be thought of as the infinitesimal. Since H is positive-definite, we can define the matrix $H^{1/2}$ and perform the change of variables $\mathbf{x} = H_i^{1/2} \mathbf{y}$ which gives:

$$P(E = F \, dF) = \frac{1}{(\pi k_B T)^{m/2}} \int_{B_{\sqrt{F}}} \exp \left[-\frac{\mathbf{x}^T \mathbf{x}}{k_B T} \right] d\mathbf{x} \quad (4.26)$$

where $B_{\sqrt{F}}$ is the Euclidean ball with radius \sqrt{F} . We have already shown that energy density does not on the choice of H_i and can now proceed in a fashion identical to the usual derivation of the χ^2 distribution:

$$P(E = F \, dF) = \frac{\exp \left[-\frac{F}{k_B T} \right]}{(\pi k_B T)^{m/2}} \int_{B_{\sqrt{F}}} d\mathbf{x} = \frac{\exp \left[-\frac{F}{k_B T} \right]}{(\pi k_B T)^{m/2}} \frac{F^{(m-1)/2} 2\pi^{m/2}}{\Gamma(m/2)} d\sqrt{F} \quad (4.27)$$

where the second equality is obtained by the surface area of the $(m-1)$ -sphere times the coordinate normal to the sphere and $d\sqrt{F}$ is the infinitesimal volume element of the sphere of radius \sqrt{F} . Observe that by defining $F = \frac{k_B T}{2} G$ and substituting $d\sqrt{F} = \frac{dF}{2\sqrt{F}}$ we have:

$$\begin{aligned} \frac{F^{(m/2-1)} \exp \left[-\frac{F}{k_B T} \right]}{(k_B T)^{m/2} \Gamma(m/2)} dF &= \frac{\left(\frac{2}{k_B T} G \right)^{(m/2-1)} \exp \left[-\frac{\frac{2}{k_B T} G}{k_B T} \right]}{(k_B T)^{m/2} \Gamma(m/2)} \frac{2}{k_B T} dG \\ &= \frac{G^{(m/2-1)} \exp \left[-\frac{G}{2} \right]}{2^{m/2} \Gamma(m/2)} dG \end{aligned} \quad (4.28)$$

The last term is simply the pdf for the χ^2 distribution implying $E \sim \frac{k_B T}{2} \chi^2$. That means that we can simply add m unitary Gaussian random numbers and multiply by $\frac{k_B T}{2}$ to get the harmonic energies. For $m > 50$, a Gaussian approximation can be made with mean $\frac{mk_B T}{2}$ and variance $\frac{m}{2}(k_B T)^2$.

4.2.6 Exchanges

MDMS is flexible enough to admit a number of exchange criteria. For our test simulations, see below, we used the usual exchange probabilities [44] ($P_x(T \rightarrow T')$):

$$P_x(T \rightarrow T') = \min \left\{ 1, \exp \left[\left(\frac{1}{k_B T} - \frac{1}{k_B T'} \right) (E - E') \right] \right\} \quad (4.29)$$

where the energies E , E' are calculated by adding the harmonic energy to the energy E_i of the current state of the system. The exchange criterion can be interpreted as comparing the probabilities on the extended state space on a state space $\Omega \subset \mathbb{R}^m \times \mathbb{N}$, which includes the spatial component to the harmonic oscillator dimension. Exact exchange frequencies can be calculated to tune the temperature gaps before running simulations.

4.2.7 Generalizing to Free Energies

MDMS is flexible enough to admit free energies or other temperature-dependent energies. In order to maintain the relation (4.1), we will define the three temperature dependent quantities $E_i(T)$, $E_j(T)$ and $E_{ij}(T) = E_{ji}(T)$, where $E_{ij}(T)$ is the energy of the barrier. That gives $\Delta E_{ij}(T) = E_{ij}(T) - E_i(T)$. While any type of temperature dependent energy function can be used, we will only consider a free energy based function.

The above can be interpreted as a system with free energies on a state space $\Omega \subset \mathbb{R}^m \times \mathbb{N}$. The Helmholtz free energy of each state i is simply $E_i(T) = E_i - k_B T \ln(\mathcal{Z})$. Both the exchanges defined in (4.29) and the equilibrium in (4.11) hold for the marginal distribution on \mathbb{N} . For the energy function given in (4.20), the Helmholtz free energy of the state i , $E_i(T)$

can be explicitly calculated:

$$\begin{aligned} E_i(T) &= \hat{E}_i - k_B T \ln \left(\int_{\mathbb{R}^m} \exp \left[-\frac{V_i(\mathbf{y})}{k_B T} \right] d\mathbf{y} \right) \\ &= \hat{E}_i - \frac{k_B T}{2} \left(m(\ln k_B T + \ln \pi) - \ln |\mathbf{H}_i| \right) \end{aligned} \quad (4.30)$$

where $|\mathbf{H}_i|$ is the determinant of the harmonic energy matrix for state i and \hat{E} is the energy minimum of the state. For the sake of relative energies, the $\ln \pi$ and $m \ln k_B T$ terms can be dropped. The energy barrier term can be treated similarly:

$$E_{ij}(T) = \hat{E}_{ij} + \frac{k_B T}{2} \ln |\mathbf{A}_{ij}| \quad (4.31)$$

The term $\ln |\mathbf{A}_{ij}|$ simply behaves like a prefactor:

$$\begin{aligned} k_{ij} &= \exp \left[-\frac{E_{ij}(T) - E_i(T)}{k_B T} \right] \\ &= \exp \left[-\frac{\hat{E}_{ij}(T) - \hat{E}_i(T)}{k_B T} + \frac{\ln |\mathbf{A}_{ij}| - \ln |\mathbf{H}_i|}{2} \right] \\ &= \left(\frac{|\mathbf{A}_{ij}|}{|\mathbf{H}_i|} \right)^{\frac{1}{2}} \exp \left[\frac{\hat{E}_{ij} - \hat{E}_i}{k_B T} \right] \end{aligned} \quad (4.32)$$

The prefactors are no longer symmetric since the equilibrium distribution is a Boltzmann distribution of the free energies of the states i and not the enthalpies. When the $|\mathbf{H}_i|$'s are identical, the equations become identical to the enthalpic case listed above.

The full Boltzmann distribution can be written out:

$$P(y = (\mathbf{x}, i)) = \frac{1}{\mathcal{Z}(T)} \exp \left[-\frac{\hat{E}_i + V_i(\mathbf{x})}{k_B T} \right] \quad (4.33)$$

A straightforward calculation yields:

$$\begin{aligned} P(y = i) &= \frac{1}{\mathcal{Z}(T)} \int_{\mathbb{R}^m} \exp \left[-\frac{\hat{E}_i + V_i(\mathbf{x})}{k_B T} \right] d\mathbf{x} \\ &= \frac{\exp \left[-\frac{\hat{E}_i}{k_B T} \right]}{\mathcal{Z}(T)} \exp \left[\frac{-k_B T \ln \left(\int_{\mathbb{R}^m} \exp \left[-\frac{V_i(\mathbf{y})}{k_B T} \right] d\mathbf{y} \right)}{-k_B T} \right] \\ &= \frac{1}{\mathcal{Z}(T)} \exp \left[-\frac{E_i(T)}{k_B T} \right] \end{aligned} \quad (4.34)$$

That is exactly the distribution generated by the dynamics described above. The infinite temperature limit is no longer uniform:

$$\lim_{T \rightarrow \infty} \frac{P(y=i)}{P(y=j)} = \lim_{T \rightarrow \infty} \frac{\exp \left[-\frac{E_i(T)}{k_B T} \right]}{\exp \left[-\frac{E_j(T)}{k_B T} \right]} = \left(\frac{|H_i|}{|H_j|} \right)^{\frac{1}{2}} \quad (4.35)$$

An infinite temperature replica can be generated from the above relation. The two limits also show how melting behavior can be generated where the lowest enthalpy state dominates at low temperatures and a higher enthalpy state dominates at higher temperatures (data not shown).

4.2.8 Heat Capacity and Entropy

The optimal distribution of replica temperatures has been suggested to be tied to the heat capacity or entropy of the system [58, 110]. Another advantage of our model is that these quantities can be explicitly calculated. While the expressions are not particularly intuitive, it is easy to plot the two quantities and estimate the necessary values.

First, we will consider the enthalpic picture. It is easy to write down the expected energy of the system:

$$\langle E \rangle = \frac{1}{\mathcal{Z}(T)} \sum_j E_j \exp \left[-\frac{E_j}{k_B T} \right] + \sum_j \frac{k_B T}{2} = \sum_j E_j b_j + m \frac{k_B T}{2} \quad (4.36)$$

where b_j are the entries of \mathbf{b} the equilibrium Boltzmann distribution, and the $m \frac{k_B T}{2}$ comes from the harmonic oscillators. Since we are concerned with exchange probabilities, we need to include the harmonic oscillator contributions to the entropy and heat capacity. The square of the energy $\langle E^2 \rangle$ can also be easily calculated. This is done for the free energy case in (4.42). The enthalpic case simply drops the \hat{E} notation. From that, the expected energy variance of the system is similarly easy to write out, and we can write out the heat capacity in terms of the energy variance:

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = \frac{1}{k_B T^2} \left(\sum_j E_j^2 b_j - \left(\sum_j E_j b_j \right)^2 \right) + \frac{k_B}{2} \quad (4.37)$$

The discrete energy variance remains bounded, so the heat capacity of the system decays asymptotically like the inverse square of the temperature to $\frac{k_B}{2}$. A straightforward calculation shows that the heat capacity exponentially decays to $\frac{k_B}{2}$ as the temperature goes to zero. From the low temperature limit, we know that $\langle E \rangle \rightarrow E_0$, so we simply need to calculate the limit of $\langle E^2 \rangle$:

$$\begin{aligned}
\lim_{T \rightarrow 0} \langle E^2 \rangle &= \lim_{T \rightarrow 0} \left(\frac{\sum_j E_j \exp \left[-\frac{E_j}{k_B T} \right]}{\mathcal{Z}(T)} \right)^2 \\
&= \lim_{T \rightarrow 0} \left(\frac{\sum_j E_j \exp \left[\frac{E_0 - E_j}{k_B T} \right]}{\sum_j \exp \left[\frac{E_0 - E_j}{k_B T} \right]} \right)^2 \\
&= \lim_{T \rightarrow 0} \left(\frac{E_0 + \sum_{j>0} E_j \exp \left[\frac{E_0 - E_j}{k_B T} \right]}{1 + \sum_{j>0} \exp \left[\frac{E_0 - E_j}{k_B T} \right]} \right)^2 = E_0^2
\end{aligned} \tag{4.38}$$

since $E_0 - E_j < 0$ for all $j > 0$ where E_0 is the lowest energy state. Putting those two together gives:

$$\lim_{T \rightarrow 0} C(T) = \lim_{T \rightarrow 0} \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = 0$$

If more than one state has the same energy as E_0 , the above calculation can be repeated with addition of the appropriate integers. The zero and infinite temperature limits and the fact that the heat capacity is positive implies that there is a maximum to the heat capacity curve.

Similar to the heat capacity, the entropy in the enthalpic picture is a relatively simple expression. Calculating directly from the Boltzmann distribution **(b)**:

$$\begin{aligned}
S &= -k_B \sum_j b_j \ln[b_j] + k_B \frac{m}{2} \ln[\pi e k_B T |H|] \\
&= -k_B \sum_j \frac{\exp \left[-\frac{E_j}{k_B T} \right]}{\mathcal{Z}(T)} \ln \left[\frac{\exp \left[-\frac{E_j}{k_B T} \right]}{\mathcal{Z}(T)} \right] + k_B \frac{m}{2} \ln[\pi e k_B T |H|] \\
&= k_B \sum_j \frac{\exp \left[-\frac{E_j}{k_B T} \right]}{\mathcal{Z}(T)} \left(\frac{E_j}{k_B T} + \ln[\mathcal{Z}(T)] \right) + k_B \frac{m}{2} \ln[\pi e k_B T |H|] \\
&= \frac{\langle E \rangle}{T} + k_B \ln[\mathcal{Z}(T)] + k_B \frac{m}{2} \ln[\pi e k_B T |H|]
\end{aligned} \tag{4.39}$$

where the \ln term comes from the harmonic oscillators. A more detailed calculation is shown below.

The calculations for the free energy case are slightly more difficult. One consideration is that we want to use the truncated form of $E_i(T)$. We will again use the energy fluctuations to calculate the heat capacity. From our implementation, we would like to see an expression of the form:

$$\langle E \rangle = \sum_j b_j \langle E_j \rangle = \sum_j b_j \left(\hat{E}_j + m \frac{k_B T}{2} \right) \quad (4.40)$$

That above form uses $E_j(T)$ only in the Boltzmann probabilities b_j , meaning we can use the abbreviated form as desired. Direct calculation gives exactly that result:

$$\begin{aligned} \langle E \rangle &= \frac{1}{\mathcal{Z}(T)} \sum_j \int_{\mathbb{R}^m} \left(\hat{E}_j + V_j(\mathbf{x}) \right) \exp \left[\frac{-\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right] d\mathbf{x} \\ &= \frac{1}{\mathcal{Z}(T)} \sum_j \left(\hat{E}_j \int_{\mathbb{R}^m} \exp \left[-\frac{\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right] d\mathbf{x} \right. \\ &\quad \left. + \int_{\mathbb{R}^m} \exp \left[-\frac{\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right] d\mathbf{x} \frac{\int_{\mathbb{R}^m} V_j(\mathbf{x}) \exp \left[-\frac{V_j(\mathbf{x})}{k_B T} \right] d\mathbf{x}}{\int_{\mathbb{R}^m} \exp \left[-\frac{V_j(\mathbf{x})}{k_B T} \right] d\mathbf{x}} \right) \\ &= \sum_j b_j \left(\hat{E}_j + \langle V_j(\mathbf{x}) \rangle \right) = \sum_j b_j \langle E_j \rangle \end{aligned} \quad (4.41)$$

A similar calculation shows:

$$\langle E^2 \rangle = \sum_j b_j \langle E_j^2 \rangle = \sum_j b_j \left(\hat{E}_j^2 + \hat{E}_j m \frac{k_B T}{2} + 3 \left(m \frac{k_B T}{2} \right)^2 \right) \quad (4.42)$$

We now have a form for our heat capacity:

$$\begin{aligned} C &= \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} = \frac{\sum_j \hat{E}_j^2 b_j - \left(\sum_j \hat{E}_j b_j + m \frac{k_B T}{2} \right)^2}{k_B T^2} + \frac{\sum_j \hat{E}_j b_j}{2T} + m \frac{3k_B}{4} \\ &= \frac{\sum_j \hat{E}_j^2 b_j - \left(\sum_j \hat{E}_j b_j \right)^2}{k_B T^2} + m \frac{k_B}{2} \end{aligned} \quad (4.43)$$

The heat capacity goes to $\frac{k_B}{2}$ at zero and very high temperatures with the same peak as for the enthalpic system with the energies $\{\hat{E}_j\}$.

The entropy of the free-energy case looks very much like the expression for the enthalpic case:

$$\begin{aligned}
S &= -k_B \sum_j \int_{\mathbb{R}^m} \frac{\exp \left[-\frac{\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right]}{\mathcal{Z}(T)} \ln \left[\frac{\exp \left[-\frac{\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right]}{\mathcal{Z}(T)} \right] d\mathbf{x} \\
&= \sum_j \int_{\mathbb{R}^m} \left(\hat{E}_j + V_j(\mathbf{x}) \right) \frac{\exp \left[-\frac{\hat{E}_j + V_j(\mathbf{x})}{k_B T} \right]}{\mathcal{Z}(T)} d\mathbf{x} + k_B \ln [\mathcal{Z}(T)] \\
&= \frac{\langle E \rangle}{T} + \ln \left[\sum_j \exp \left[\frac{E_j(T)}{k_B T} \right] \right]
\end{aligned} \tag{4.44}$$

Note that from equations (4.34) and (4.40), the above entropy can be calculated without explicitly evaluating any of the integrals associated with the harmonic oscillator dimensions.

4.3 EXAMPLE MODELS

In this supplement, we outline a small 'model zoo' of potential models that could be used for evaluating methods. This list is very far from complete, but we believe that these particular models demonstrate the flexibility of our general model and can serve as a few effective tests for situations that occur in protein systems.

4.3.1 Diffusion-Limited Cases

These cases consider situations where there are many states separated by small—or even zero—energy barriers where convergence rates are limited by diffusion rates, in our case the Arrhenius factors A_{ij} .

2-state, 1-barrier

This model has two energy levels with diffusion-limited transitions within each energy level and a larger barrier in between the two energy levels. Labeling the two energy levels E_0 and E_1 with large energy barriers ΔE_ℓ and small energy barriers ΔE_s gives a linear state picture:

$$E_0 \xrightleftharpoons{\Delta E_s} E_0 \xrightleftharpoons{\Delta E_s} \dots \xrightleftharpoons{\Delta E_s} E_0 \xrightleftharpoons{\Delta E_\ell} E_1 \xrightleftharpoons{\Delta E_s} E_1 \xrightleftharpoons{\Delta E_s} \dots \xrightleftharpoons{\Delta E_s} E_1$$

The two energy levels also include a significant entropy difference by having more states with one energy level than the other energy level.

2-pathway

A second diffusion-limited model features two diffusion-limited pathways between two end states. If you have end-states E_0 and E_1 , this system looks like:

$$E_0 \leftrightarrow \left\{ \begin{array}{cccccc} \xleftrightarrow{\Delta E_s} & E_{j_1} & \xleftrightarrow{\Delta E_s} & \dots & \xleftrightarrow{\Delta E_s} & E_{j_n} & \xleftrightarrow{\Delta E_s} \\ \xleftrightarrow{\Delta E_s} & E_{k_1} & \xleftrightarrow{\Delta E_s} & \dots & \xleftrightarrow{\Delta E_s} & E_{k_m} & \xleftrightarrow{\Delta E_s} \end{array} \right\} \leftrightarrow E_1$$

where E_{j_i} and E_{k_i} are distinct pathways. Note that n and m may not be equal. It is also topologically equivalent to a circle:

$$E_0 \xleftrightarrow{\Delta E_s} E_{j_1} \xleftrightarrow{\Delta E_s} \dots \xleftrightarrow{\Delta E_s} E_{j_n} \xleftrightarrow{\Delta E_s} E_1 \xleftrightarrow{\Delta E_s} E_{k_1} \xleftrightarrow{\Delta E_s} \dots \xleftrightarrow{\Delta E_s} E_{k_m} \xleftrightarrow{\Delta E_s} E_0$$

This model is very easily expanded to an arbitrary number n -pathways.

4.3.2 Non-Diffusion-Limited Cases

These models have large barriers in the system that can become the rate-limiting factor.

2-state, n -barrier

To create n barriers between two states, we can explicitly define the energy of the intermediate state to generate an arbitrary number of barriers. If you are interested in pathway k between states E_0 and E_1 , the energy levels can be explicitly written as $E_0 + \Delta E_{0k} = E_k$ and $E_1 + \Delta E_{1k} = E_k$. That gives the usual TST relation: $E_0 + \Delta E_{0k} = E_1 + \Delta E_{1k}$. A 3-barrier system would look like:

$$E_0 \leftrightarrow \left\{ \begin{array}{ccc} \xleftrightarrow{\Delta E_{0j_1}} & E_{j_1} & \xleftrightarrow{\Delta E_{j_1 1}} \\ \xleftrightarrow{\Delta E_{0j_2}} & E_{j_2} & \xleftrightarrow{\Delta E_{j_2 1}} \\ \xleftrightarrow{\Delta E_{0j_3}} & E_{j_3} & \xleftrightarrow{\Delta E_{j_3 1}} \end{array} \right\} \leftrightarrow E_1$$

2-state, n -pathway

There is no reason all the pathways between two states need to be involved in a single barrier. Two states could be separate by two (or more) pathways, one of which is diffusion-limited and the other features a large barrier. As an example, consider a system with two

states $E_0 = 1 k_B T_0$ and $E_1 = 2 k_B T_0$. One barrier state has energy $E_2 = 10 k_B T_0$, a very large barrier. The other pathway, comprised of states $E_{j_k} = 0.05 * k k_B T_0$. The network would look like:

$$E_0 \leftrightarrow \left\{ \begin{array}{ccccccc} & \xleftrightarrow{\Delta E_{03}} & & E_3 & & \xleftrightarrow{\Delta E_{31}} & \\ \xleftrightarrow{\Delta E_s} & E_{j_1} & \xleftrightarrow{\Delta E_s} & \dots & \xleftrightarrow{\Delta E_s} & E_{j_{19}} & \xleftrightarrow{\Delta E_s} \end{array} \right\} \leftrightarrow E_1$$

The diffusion limited path uses small barriers, ΔE_s . At low temperatures, the system would move almost entirely through the diffusion-limited pathway. However, for sufficiently high temperatures, most transitions would occur through the large barrier because of the lower entropy cost.

4.3.3 Combining Models

All of the models listed here are only concerned with the transitions between two primary states. These models can easily be combined by chaining them together linearly or by a more complicated matrix of connections.

4.4 SIMULATIONS

4.4.1 Simulation Details

All of the simulations were run using the Numpy package in Python [2, 54, 85, 86]. Before the start of each simulation, the transition matrices $M_{\delta t}$ were calculated for all of the temperatures. At each exchange attempt, first, the states of the system were updated according to the transition matrices $M_{\delta t}$ by comparing to a uniform random number. Then, the energies were calculated using Gaussian random numbers. Finally, neighboring replica exchange attempts were performed. All random number generators were based upon the Mersenne twister algorithm [70].

In order to demonstrate the effectiveness of MDMS, we devised a simple test case that allowed us to test the effectiveness of using parallel tempering with two temperature distributions versus standard constant temperature dynamics. For this test case, we will only

use the enthalpic case.

4.4.2 Test Problem

Our test case is a three-state model with large barriers between the first state and the other two states. The energy values can be seen in Table 4.1. The two large barriers lead to very slow convergence times. The correlation time of the model at $k_B T_0 = 1$ is relatively long (~ 6.5 ns with $A_{ij} = 1/\text{ps}$), and many samples would be required to get good statistics. However, increasing the temperature to $2.828 k_B T$ gives a correlation time of ~ 15 ps.

We ran four replicas with temperatures at $k_B T = 1, 1.414, 2$ and 2.828 , respectively. Those temperatures follow the geometric recommendation in [82]. In order to get exchanges at a reasonable frequency, we set $m = 10$, all of the energy functions V_i equal. The probability density functions are shown in Figure 4.1-A.

Another advantage of MDMS is that energy density graphs such as Figure 4.1-A can be easily made analytically, and the exchange frequency at equilibrium can be explicitly calculated. When we ran the corresponding RE simulations, we obtain virtually identical energy distributions at each temperature (Figure 4.1-B).

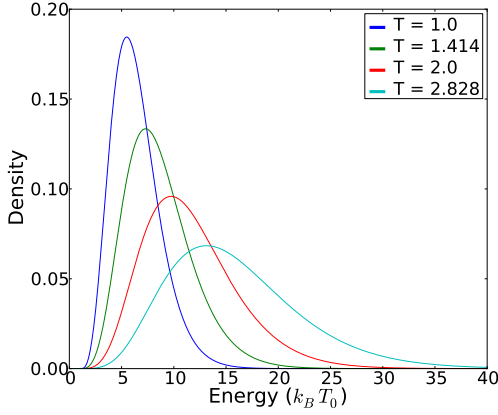
4.4.3 Results

The first test we performed was to see if MDMS can simulate the transitions between local minima effectively. We ran 10,000 independent “MD” simulations, which represent molecular dynamics simulations with no exchanges, for 2000 ps starting from random initial states at

Table 4.1: Table of energy values for each state and barriers between them.

$E_1 = 1 k_B T_0$	$\Delta E_{12} = 12 k_B T_0$
$E_2 = 2 k_B T_0$	$\Delta E_{13} = 10 k_B T_0$
$E_3 = 3 k_B T_0$	$\Delta E_{23} = 2 k_B T_0$
where $k_B T_0 = 1$	

A)



B)

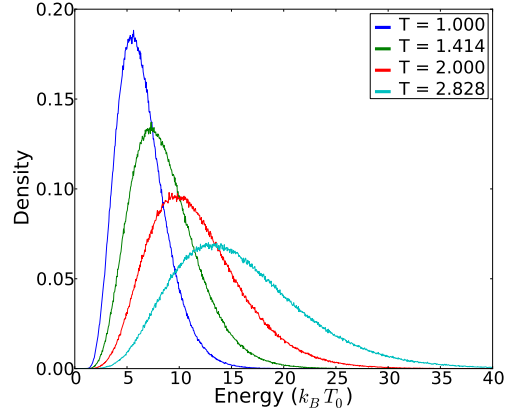


Figure 4.1: The numerically calculated energy distributions almost exactly match the analytically calculated distribution. A) Theoretical energy distributions at each temperature. B) Observed temperature distributions for each replica.

a temperature of $k_B T = 1$. To ensure rapid convergence, we used energy barriers smaller than those listed in Table 4.1. The reduced barrier heights are listed in the figure legend for Figure 4.2. We then compared the average population of the lowest energy state (with energy $1 k_B T_0$) to the correct equilibrium distribution. Three sets of calculations were performed varying the sampling window δt , and no dependence on δt was observed. This can allow for more challenging systems with even higher barriers between states requiring longer simulation time.

To generate representative energies for each replica, we chose the harmonic dimension parameter $m = 10$ because it gave a reasonable spread without overwhelming the energy gaps. Larger m would simply require smaller gaps between the temperatures of successive replicas as is seen in practice.

Next we ran replica exchange simulations using four replicas with temperatures $k_B T = 1, 1.414, 2$ and 2.828 respectively. As with the straight MD runs, we ran 10,000 parallel RE simulations using the standard replica exchange rule in (4.29), independent simulations started with random initial conditions so that each state was uniformly populated. The

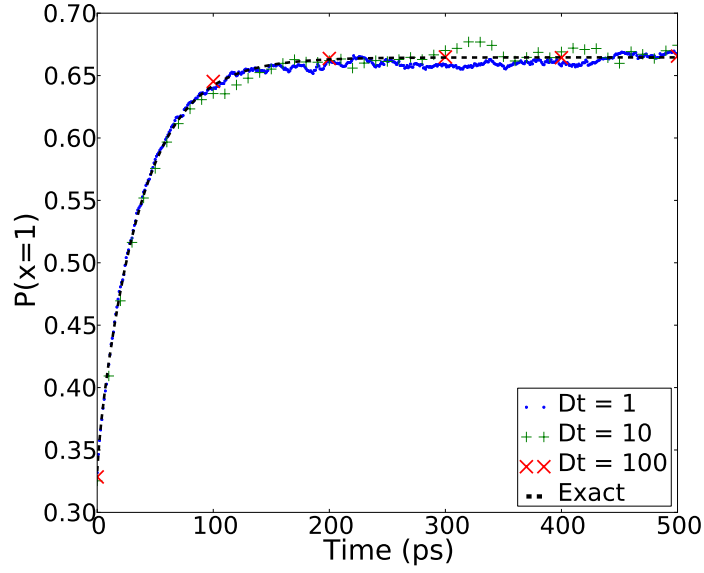


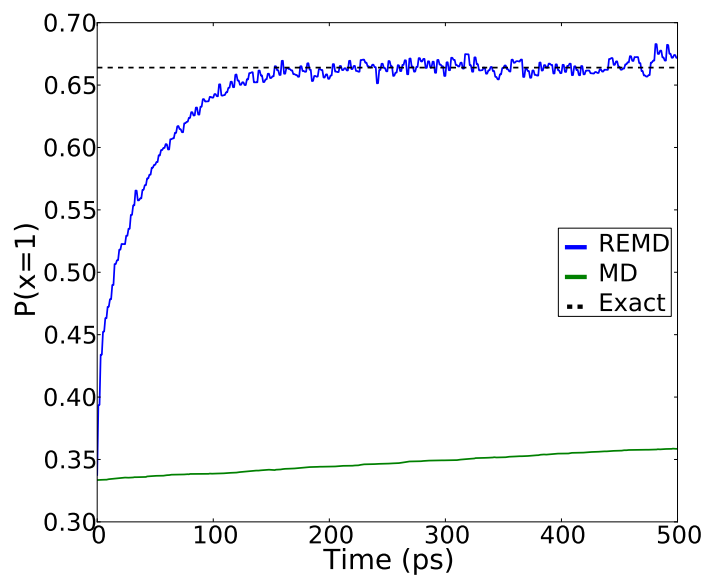
Figure 4.2: The standard MDMS MD model converges to the Boltzmann equilibrium. Also, the convergence behavior of the system is independent of the sampling window δt . The y -axis is the proportion of simulations where the lowest temperature replica is in the 1st state at time t . The dashed line is the exact solution for uniform initial conditions. For the sake of fast convergence, the energy barriers used were $E_{12} = 6$, $E_{13} = 5$ and $E_{23} = 2$.

population of the lowest energy state was compared to the exact equilibrium populations (Figure 4.3). Figure 4.3 shows the convergence of the population of the lowest energy state to the analytical result. The simulation energies are shown in Table 4.1. The standard MD simulations would converge to equilibrium after about 30 ns, while the replica exchange simulations reach their equilibrium values much faster (within 150 exchange attempts) and maintain that population throughout the simulation. The difference in rate of convergence is orders of magnitude.

To investigate the benefits of frequent exchange attempts, we ran four simulations where we varied the interval between exchange attempts δt . As seen in Figure 4.4, going from $\delta t = 10$ ps to $\delta t = 1$ ps shows a pronounced increase in the rate of convergence, in agreement with previous studies [108, 121, 122]. However, smaller δt show little additional benefit. There are obvious constraints to the sampling rate, not least of which being the rate of sampling by the highest temperature replica. MDMS is currently limited by the lack of an equilibration time within each well, but this result provides evidence for using frequent exchanges to increase the rate of sampling. Using more frequent exchange rates does have to be compared to the computational cost of doing more frequent exchange attempts.

The last test we ran was to explicitly test the recommended temperature intervals suggested in [110]. We assumed that the entropy contribution of the harmonic oscillators was small and only used the entropy from the discrete states for our calculation. As the total entropy is bounded, we were only able to select three temperatures with a constant entropy increase: $k_B T = 1, 1.414$, and 3.4 . For comparison, the simulations with geometric temperature differences were run with $k_B T = 1, 1.414$ and 2 . Figure 4.5 shows significantly faster convergence using the constant entropy increase temperature selection from [110]. These calculations further show the power of MDMS in that the total time spent calculating entropies and running simulations was less than three hours.

A)



B)

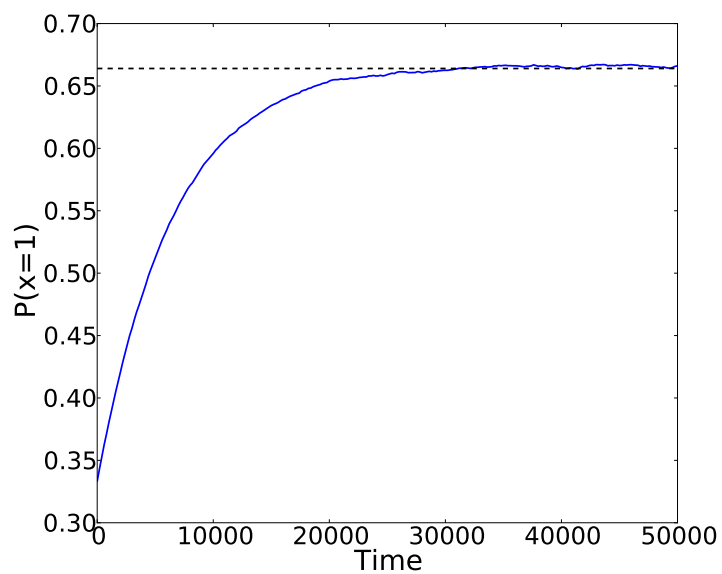


Figure 4.3: A simple replica exchange simulation converges significantly faster than an MD simulation with the parameters listed in the table. A) shows that the MD simulation is close to constant on the time-scale of REMD convergence. B) shows that the MD simulation does converge but at a time-scale approximately 100 times longer.

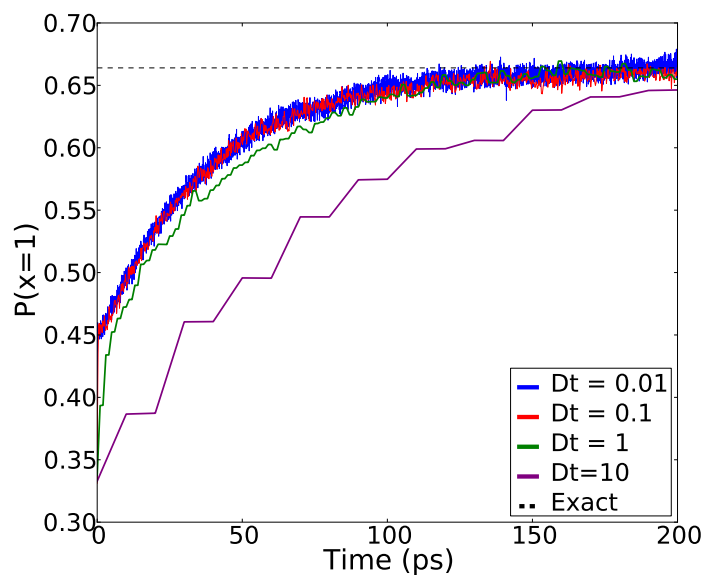


Figure 4.4: Replica exchange simulations were run with varying exchange frequencies. Convergence rates are increased with faster exchanges, but there is a limit to the rate of convergence leading to diminishing returns.

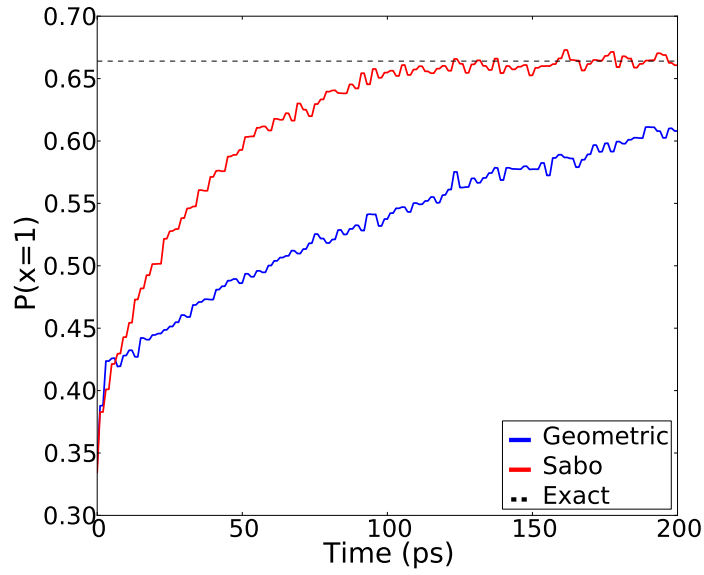


Figure 4.5: Replica exchange simulations were run with two different sets of temperature intervals. The first set of simulations were run according to the geometric progression of $k_B T = 1, 1.414$ and 2 . The second set were run with a temperature distribution based on [110] with $k_B T = 1, 1.414$ and 3.4 .

4.5 DISCUSSION

The simulation results we have presented here are very simple and do not break new ground, but MDMS is very general and can be easily expanded to investigate a number of exchange based methods and potentially even more general sampling methods. The model zoo in the supplement gives a small selection of ideas that could be used in developing a systematic test set for new methods. Our theoretical model may become a test-bed for future exchange based sampling methods where many parameters can be tested and evaluated simultaneously.

In the results of this chapter, only two temperature selection schemes have been considered. However, there has been considerable work on this topic [3, 25, 57, 92, 105, 110]. In previous attempts to quantify the effects of temperature selection such as [3], there was not an obvious definition of optimal. For MDMS, the obvious metric is the real part of the slowest eigenvalue. While analytical results might be difficult to obtain, the analysis in [3] could be repeated numerically for a few fixed models. In fact, a program could be written that generates random models and automatically calculates convergence rates based on whatever temperature selection method you choose.

As one example of a more general exchange method, Hamiltonian replica exchange methods could be implemented by simply modifying the barriers in the rate matrix and using that to define the kinetics of each replica. The reduction in energy barriers could be estimated directly from the particular changes to the energy function or empirically with short simulations. For some types of Hamiltonians, such as changing the dihedral potentials such as in [59], the change in barrier height is easy to calculate. However, even in a less straightforward case, such as changing Lennard-Jones potentials like in [53], potential of mean force calculations could give an estimate for the change in barrier height.

One possible improvement would be to implement some mechanism that would allow for non-instant equilibration in the harmonic oscillator dimensions. Before the simulation, a family of energy distributions based on the convergence in time would be calculated. Then, at each exchange attempt the time of the last transition would be calculated, and the energy would be drawn from the relevant distribution. The simplest way to do that would be to implement specific dynamics in between exchanges, e.g. using a method such as Gillespie's

algorithm [47]. It also should be possible to explicitly calculate the distribution of state changes within the interval $(t, t + \delta t)$ subject to the relevant boundary constraints. That distribution would allow for calculation of a last transition time at each exchange attempt again negating the need for explicit dynamics.

RE methods developed using MDMS will expand the size of systems for which we can calculate statistical quantities. That means that more detailed atomistic actin-related models can be used to calculate physical quantities. Whether it be parameterizing coarse-grained models [19, 20], calculating binding and conformational change free energies [118], or estimating physical quantities of actin filaments using atomistic models [26, 71], enhanced computational sampling methods will greatly aid the study actin networks going forward.

In conclusion, we have demonstrated a new model that allows for rapid testing of replica exchange methods. We have developed a tool that can quickly generate many replica exchange simulations on a pre-defined system with exactly known statistical properties. Results calculated with new methods can be compared to known quantities and guide method development. We intend to use this model to test the applicability of reservoir replica exchange methods [84, 107] and investigate their strengths and weaknesses in a systematic manner and propose enhancements to overcome possible deficiencies and improve their efficiency.

5.0 DISCUSSION

The results presented here both serve to improve our understanding of branching actin networks and justify some related numerical methods. Three different approaches are used in studying actin networks from different perspectives. Chapter 2 used agent-based stochastic simulations to probe how a branching actin network could collectively generate force and experimentally observed force-velocity relationships. While chapter 3 is not necessarily methodological, it is significantly more theoretical. The nature of solutions to a continuum model of filament orientation was studied. The continuum model was shown to uniquely select for an ‘optimal’ orientation pattern with a well-defined fitness function. However, the results do provide new methods and perspective towards future work studying actin network orientation patterns. In contrast, chapter 4 is purely methodological. Understanding actin monomer binding properties and basic, physical properties of actin filaments, such as stiffness, is very important for the study of branching actin networks. However, calculating such things is close to or beyond the limit of current computing power. Even with improved computational power, the curse of dimensionality will still persist. To improve our ability to develop new methods for this problem, Molecular Dynamics Meta-Simulators (MDMS) is proposed as a way to test new sampling methods.

Multiple patterns have been observed in the experimental force-velocity relationships of branching actin networks. However, there is no accepted mechanism for how an actin network can collectively generate force against a load [74]. We developed a stochastic simulation that modeled each filament tip and the branching, capping and filament growth dynamics explicitly. That model was able to qualitatively reproduce both general force-velocity relationships experimentally observed [69, 89]. Our interpretation of the model is that the temporal dynamics of filament turnover, i.e. new filaments branching and old filaments

becoming capped, determines the nature of the force-velocity relationship. The network appears to converge to an equilibrium number of filaments at the leading edge over multiple time-scales. If an experiment is conducted on a fast time-scale, the network will not recruit enough filaments to the leading edge to sustain the network velocity. However, if the force is ramped up slowly or the experiment measures the velocity over a long enough time window, the slower time-scales in the network will converge sustaining the velocity in spite of increasing velocity.

Our simulation results also explicitly contradicted previous results regarding the geometry of the network. We found that filaments tended to organize in a fashion where most of the filaments were perpendicular to the leading edge. However, previously reported simulations where filaments tended to have a $\sim 35^\circ$ degree angle relative to the normal direction of the leading edge [115, 137]. We were able to reproduce their results for a very narrow parameter range, which we argue is not likely physically accurate. Most experimental work has shown the filaments organize with symmetric peaks in the orientation density somewhere between 25° and 45° relative to the normal direction of the membrane, consistent with previous simulations. The fact that their results are consistent with experiment only for a narrow parameter range suggests some other mechanism may be involved. Though, recent experiments have cast doubt on the actual orientation pattern [138].

There remains much to be learned by using stochastic models of branching actin networks. Our model did not include the dynamics of filopodia, bundles of actin filaments present at the leading edge of a cell, which are believed to play a significant role in cell motility [72]. Motile cells also display wave-like behavior with the lamellipodium expanding and contracting [16]. Both of these phenomena remain areas where further stochastic modeling would likely be beneficial. Actin networks have also been successfully modeled as a visco-elastic gel [74, 97]. However, as of yet, the visco-elastic properties of actin have not been successfully connected to the basic physical processes of actin networks.

There is also a need to develop new methods for studying such stochastic models, including for modeling actin. One particularly active area of current research involves developing new methods for simulating point-processes. The canonical example for this is the Gillespie algorithm [47]. However, models such as the one in chapter 2 have a significant spatial com-

ponent complicating the simulation details. Future modeling work will also likely require spatial resolution [104]. The simulations presented here were very time consuming and run on a network with significant computational resources. However, a less naive computational method would likely significantly reduce the computational resources required. Reproducing and extending our work would greatly benefit from such a method.

As observed in electron micrographs, branching actin networks appear to have a regular geometry relative to the leading edge [68, 138]. The angle between the filaments and the membrane appears to be quite regular. The exact mechanism regulating these patterns is not currently known. Moreover, it is not even known what pathway or signal is used that causes almost all actin filaments to point towards the membrane. One hypothesized mechanism is that the capping rate of filaments is somehow orientation dependent [68, 137]. Continuum models have been developed to test this hypothesis. Those models were nonlocal diffusion equations, so the nature of the solutions is non-obvious. The analysis in chapter 3 was done to place the numerically calculated solutions in context. We were able to completely characterize the solution of one type of model and greatly restrict the possible dynamics of a second model. For the model with filaments branching with a first-order rate constant, we found that the model uniquely selected for an optimal orientation pattern in the sense that solutions $v(\theta, t) \sim \exp[\mu t]u_0(\theta)$. When filament branching is assumed to occur at a constant, or zeroth-order, rate, we were not able to completely characterize solutions. However, we were able to show that there exists a unique positive equilibrium. The stability of that equilibrium remains in doubt. Thus, both models select for a unique orientation pattern in some sense. We also found that that orientation pattern is the unique extremum of a related optimization problem.

Several mathematical problems immediately follow from that work. Obviously, a proof of conjecture 1 would complete the analysis of equation (3.8). There is also an obvious generalization of theorem 1. The densities of interest are not necessarily absolutely continuous with respect to the Lebesgue measure. Both (3.7) and (3.8) can be interpreted as being defined on the space of positive, finite, countably-additive measures, which is a Banach space when equipped with the total variation norm. Solutions are then considered in the weak sense with respect to the dual space. The same result should still hold over that more general

space.

The results in chapter 3 prove the existence of an optimal orientation pattern but do not characterize that solution. The two orientation patterns generally described can be characterized by whether or not 0° is a local minimum or a local maximum in the density. It is easy to show that the symmetry hypotheses imply that 0° must be one of the two. It remains an open question as to what conditions are necessary to give one pattern over the other. Quint and Schwarz [103] used a discretization to attempt to characterize how general the two patterns were but did not come to any strong conclusions. Further analysis could be performed on the continuous equations to come to a stronger characterization. The numerical method presented in section 3.7 could be used to efficiently calculate a variety of equilibria. The numerical results presented here, however, were done with stability in mind and include non-physical branching and capping functions. Physical considerations could greatly limit the possible candidates for the capping and branching functions, κ and \mathcal{B} . For example, it has previously been assumed that $\kappa(\theta)$ reaches a minimum at 0° . For those candidates, the two orientation patterns previously reported, as in [68, 137], can be differentiated by whether 0° is a local minimum or maximum of the density function u . That it has to be one or the other can be seen, for the first-order case, by:

$$\begin{aligned} u'(0) &= \frac{\partial}{\partial \theta} \left. \frac{\int_{S^1} \mathcal{B}(\theta - \omega) u(\omega) d\omega}{\kappa(\theta) + \mu} \right|_0 \\ &= \frac{(\kappa(0) + \mu) \int_{S^1} \mathcal{B}'(-\omega) u(\omega) d\omega - \kappa'(0) \int_{S^1} \mathcal{B}(-\omega) u(\omega) d\omega}{(\kappa(0) + \mu)^2} = 0 \end{aligned} \quad (5.1)$$

and the fact that u is symmetric. The hypothesis that $\kappa(0)$ is a minimum gives that $\kappa'(0) = 0$. The hypothesis that \mathcal{B} is symmetric implies \mathcal{B}' is odd. When combined with u being even, that implies that the integral $\int \mathcal{B}' u = 0$. Using the smoothness assumption from the proof of theorem 1, we can characterize most solutions by the sign of the second derivative:

$$u''(0) = \frac{(\kappa(0) + \mu) \int_{S^1} \mathcal{B}''(-\omega) u(\omega) d\omega - \kappa''(0) \int_{S^1} \mathcal{B}(-\omega) u(\omega) d\omega}{(\kappa(0) + \mu)^2}$$

Analyzing the above equation could lead to understanding the shape dependence on κ and \mathcal{B} .

The development of new enhanced sampling algorithms has been a very active area of research for a number of years [55, 64, 134, 146]. One class of enhanced sampling methods, called replica exchange (RE) methods, use parallel simulations run in parallel while periodically exchanging states [44, 133]. A number of new RE methods have been developed [39, 62, 63, 66, 67, 83, 84, 106, 107, 128, 129], but there does not exist a way to effectively compare methods and make informed decisions on which method to use [146]. We proposed MDMS as a start to fill this gap, at least for RE methods. The most time consuming step of testing RE methods is running the dynamics between exchange attempts. MDMS abstracts those dynamics, which allows us to rapidly test exchange criteria. It drastically reduces the amount of time it takes to initially test a new RE method.

Current and future work using MDMS is focused on developing new replica exchange methods. In particular, MDMS is currently being used to test reservoir replica exchange methods [62, 84, 107]. Those methods rely on exchanging states with a reservoir of conformations that correspond to local minima of the energy function that are previously calculated using some brute-force sampling method. MDMS allows for energy functions and state spaces that can be complex while still being completely understood. For example, we know all of the local minima for an MDMS model and can test the effect of have imperfect sampling in the reservoir. We can also specify exact properties of the reservoir relative to properties of the system without any expensive calculations. There is also the potential to expand MDMS beyond temperature RE and related methods. MDMS should also be able to model Hamiltonian replica exchange methods such as [39, 53]. The transition rates could be modified using an estimate of the change in energy barriers from using the modified Hamiltonian. MDMS also could allow for analytical calculation of replica exchange efficiencies similar to [108, 109]. Unlike previous attempts such as [3], MDMS comes with a pre-defined metric for how to define optimal sampling.

Further work may show that MDMS is applicable even beyond the methods outlined here. In particular, we would like to eventually extend MDMS to be able to test non-RE sampling methods to begin to allow for rigorous comparison of the numerous enhanced sampling algorithms available.

APPENDIX

PSEDUO-CODE FOR THE AGENT-BASED MODEL

The following is a pseudo-code representation of the algorithm used in chapter 2. As such, all of the memory handling, writing to file, array resizing, etc. are omitted. All calculations are done in angles, and the trigonometric functions are assumed to be consistent with that. `ran1` generates uniform random numbers in $[0,1)$, and `gasdev` generates samples from the standard normal distribution with mean 0 and variance 1. Figure A1 outlines the basic structure of the simulation.

```
1 // Variable definitions
2 system[i][j] // 2-dim array of filament states
3
4 system[i][0] // Filament angle
5 system[i][1] // x-position
6 system[i][2] // Distance from leading edge
7 system[i][3] // Force applied to filament
8
9 contact[i] // Boolean array for filament
   contacting
10 // leading edge
11
12 T // Total number of time steps
13 dt // Time step length
14
15 nfibers // Current number of filaments
16 lambda // Branching constant
17 Cl=exp(-lambda) // Branching Poisson constant
18 kappa // Capping constant
19 seed // Random seed
20 nfront // Number of filaments contacting
21 // leading edge
22
23 vf // Free filament velocity
```

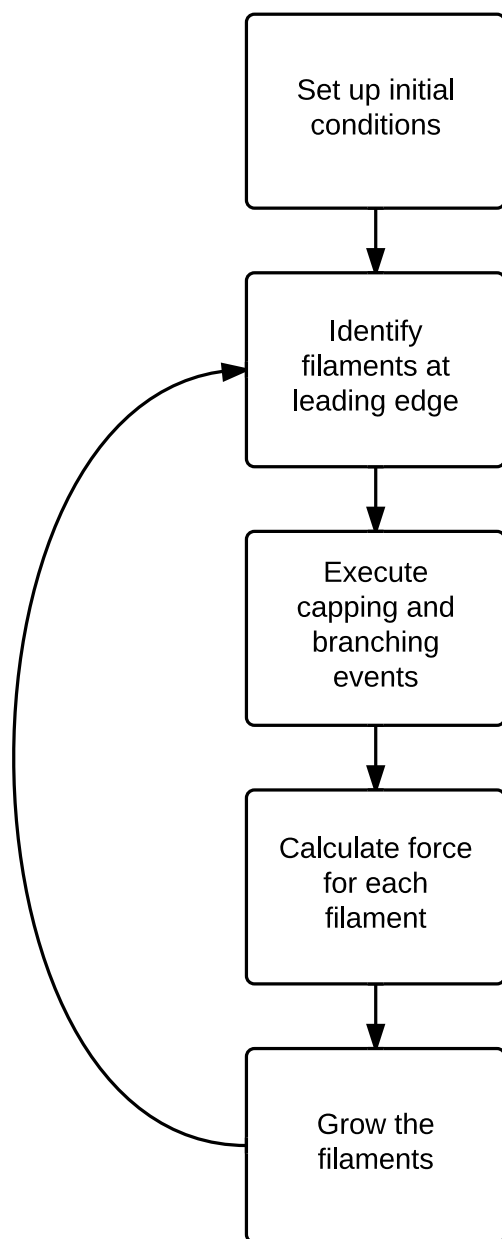


Figure A1: A flow chart diagramming the basic simulation structure.

```

24 delta                                // Length of individual actin subunit
25 ftot                                // Total force
26
27 // Randomly assign uniform initial conditions
28
29 for (iN < nfibers) {
30     system[iN][0] = -90 + 180*ran1(seed)
31     system[iN][1] = ran1(seed)
32 }
33
34 // Iterate for T steps
35
36 for (it < T) {
37
38     // Identify position of furthest filament
39     // This defines the position of the leading edge
40
41     Front = 0                                // Reset leading edge position
42     for (iN < nfibers) {
43         TestFront = Front - system[iN][1]
44         if (TestFront < 0) {
45             Front = system[iN][1]
46         }
47     }
48
49     // Test to see if filament is within one actin subunit of leading edge
50
51     nfront = 0                                // Reset number of contacting filaments
52     system[iN][2] = Front - system[iN][1]
53     if (system[iN][2] < delta) {
54         contact[iN] = 1                                // Mark filament as contacting l.e.
55         nfront++
56     } else {
57         contact[iN] = 0
58     }
59
60     // Calculate number of branching and capping events
61
62
63     Ck = exp(-kappa*(nfibers-nfront)*dt);
64
65     p = 0                                // Capping Poisson constant
66     K = -1                                // Poisson dummy variable
67     P = ran1(seed)                        // Reset number of capping events
68     while (p < P) {
69         K++
70         p += Ck*pow(kappa*(nfibers-nfront)*dt, K)/factorial(K)
71     }
72
73     p = 0
74     L = -1                                // Reset number of branching events
75     P = ran1(seed)
76     while (p < P) {
77         L++

```

```

78     p+=Cl*pow(lambda*dt,L)/factorial(L)
79 }
80
81 // Implement capping and branching events
82 iK=0
83 while (iK < K) {
84     iM = choose(contact, nfibers, nfront, seed)
85                                     // Uniform random selection from
86                                     // filaments not contacting l.e.
87     delete system[iM]
88     system[iM] = system[nfibers-1] // Reassign filament nfibers-1 to iM
89     contact[iM] = contact[nfibers-1]
90     nfibers —
91     iK++
92 }
93
94 iL=0;
95 while (iL < L) {
96     contact[nfibers] = 0 // New filament not contacting l.e.
97     iM = choose(contact, nfibers, nfront, seed)
98     p = ran1(seed)
99     system[nfibers][1] = system[iM][1] — cos(system[iM][0])*2*delta*p
100                                     // Place new filament backwards along
101                                     // existing filament
102     P = 5*gasdev(seed) + 70 // Choose absolute branching angle
103     if (p < 0.5) // Choose branching direction
104         angle = system[iM][0] + P
105     else
106         angle = system[iM][0] — P
107     if (abs(angle)>90) { // If filament not pointing forwards
108         iL++ // reject branching event, add to
109         continue // counter and continue
110     }
111     system[nfibers][0] = angle
112     nfibers++
113     iL++
114 }
115
116 // Calculate forces
117
118 alpha = 0 // Reset force-sharing factor
119 for (iN < nfibers) {
120     if (contact[iN])
121         alpha += cos(system[iN][0])
122 }
123
124 for (iN < nfibers) {
125     if (contact[iN])
126         system[iN][3] = ftot*cos(system[iN][0])
127 }
128
129 // Advance the position of the filament tips
130
131 for (iN < nfibers) {

```



```

132     if (contact[iN])
133         sysetm[iN][1] += dt*vf*delta*cos(system[iN][0])
134     else
135         sysetm[iN][1] += dt*vf*delta*cos(sysetm[iN][0])*
136             exp(-system[iN][3]*delta)*cos(system[iN][0])
137 }

```

Bibliography

- [1] V. C. Abraham, V. Krishnamurthi, D. L. Taylor, and F. Lanni. The actin-based nanomachine at the leading edge of migrating cells. *Biophys J*, 77(3):1721–1732, 1999.
- [2] D. Ascher, P. F. Dubois, K. Hinsen, J. Hugunin, and T. Oliphant. *Numerical Python*. Lawrence Livermore National Laboratory, Livermore, CA, ucrl-ma-128569 edition, 1999.
- [3] Y. F. Atchadé, G. O. Roberts, and J. S. Rosenthal. Towards optimal scaling of metropolis-coupled markov chain monte carlo. *Statistics and Computing*, 21(4):555–568, 2011.
- [4] E. Atilgan, D. Wirtz, and S. X. Sun. Morphology of the lamellipodium and organization of actin filaments at the leading edge of crawling cells. *Biophys J*, 89(5):3589–3602, 2005.
- [5] C. T. H. Baker. *The Numerical Treatment of Integral Equations*. Clarendon Press, Oxford, 1977.
- [6] R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*, 92(9):3844–3848, 1995.
- [7] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 2010.
- [8] P. Bressloff. Euclidean shift-twist symmetry in population models of self-aligning objects. *SIAM Journal on Applied Mathematics*, 64(5):1668–1690, 2004.
- [9] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. Charmm: The biomolecular simulation program. *J. Comput. Chem.*, 30(10):1545–1614, 2009.

- [10] L. A. Cameron, M. J. Footer, A. van Oudenaarden, and J. A. Theriot. Motility of acta protein-coated microspheres driven by actin polymerization. *Proc Natl Acad Sci U S A*, 96(9):4908–4913, 1999.
- [11] L. A. Cameron, T. M. Svitkina, D. Vignjevic, J. A. Theriot, and G. G. Borisy. Dendritic organization of actin comet tails. *Curr Biol*, 11(2):130–135, 2001.
- [12] L. A. Cameron, J. R. Robbins, M. J. Footer, and J. A. Theriot. Biophysical parameters influence actin-based movement, trajectory, and initiation in a cell-free system. *Mol Biol Cell*, 15(5):2312–2323, 2004.
- [13] M.-F. Carlier and D. Pantaloni. Control of actin assembly dynamics in cell motility. *J Biol Chem*, 282(32):23005–23009, 2007.
- [14] A. E. Carlsson. Growth of branched actin networks against obstacles. *Biophys J*, 81(4):1907–1923, 2001.
- [15] A. E. Carlsson. Growth velocities of branched actin networks. *Biophys J*, 84(5):2907–2918, 2003.
- [16] A. E. Carlsson. Actin dynamics: From nanoscale to microscale. *Annual Review of Biophysics*, 39(1):91–110, 2010.
- [17] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. The amber biomolecular simulation programs. *J. Comput. Chem.*, 26(16):1668–1688, 2005.
- [18] E. Chasseigne, M. Chaves, and J. D. Rossi. Asymptotic behavior for nonlocal diffusion equations. *Journal de Mathematiques Pures et Appliques*, 86(3):271 – 291, 2006.
- [19] J. W. Chu and G. A. Voth. Allostery of actin filaments: molecular dynamics simulations and coarse-grained analysis. *Proc Natl Acad Sci U S A*, 102(37):13111–13116, 2005.
- [20] J.-W. Chu and G. A. Voth. Coarse-grained free energy functions for studying protein conformational changes: A double-well network model. *Cell*, 93(11):3860–3871, 2007.
- [21] C. Co, D. T. Wong, S. Gierke, V. Chang, and J. Taunton. Mechanism of actin network attachment to moving membranes: barbed end capture by n-wasp wh2 domains. *Cell*, 128(5):901–913, 2007.
- [22] P. Dalhaimer and T. D. Pollard. Molecular dynamics simulations of arp2/3 complex activation. *Biophys J*, 99(8):2568–2576+, 2010.
- [23] P. Dalhaimer, T. D. Pollard, and B. J. Nolen. Nucleotide-mediated conformational changes of monomeric actin and arp3 studied by molecular dynamics simulations. *Journal of Molecular Biology*, 376(1):166 – 183, 2008.

- [24] A. De Masi, E. Olivieri, and E. Presutti. Spectral properties of integral operators in problems of interface dynamics and metastability. *Markov Processes and Related Fields*, 4(1):27–112, 1998.
- [25] R. Denschlag, M. Lingenheil, and P. Tavan. Optimal temperature ladders in replica exchange simulations. *Chemical Physics Letters*, 473(1-3):193 – 195, 2009.
- [26] M. A. Deriu, T. C. Bidone, F. Mastrangelo, G. D. Benedetto, M. Soncini, F. M. Montevocchi, and U. Morbiducci. Biomechanics of actin filaments: A computational multi-level study. *Journal of Biomechanics*, 44(4):630 – 636, 2011.
- [27] R. Dominguez and K. C. Holmes. Actin structure and function. *Annual Review of Biophysics*, 40(1):169–186, 2011.
- [28] D. G. Drubin, M. Kaksonen, C. Toret, and Y. Sun. Cytoskeletal networks and pathways involved in endocytosis. *Novartis Found Symp*, 269:35–42; discussion 43–6, 223–30, 2005.
- [29] Y. Du. *Order structure and topological methods in nonlinear partial differential equations. Vol. 1*, volume 2 of *Series in Partial Differential Equations and Applications*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2006. Maximum principles and applications.
- [30] D. J. Earl and M. W. Deem. Parallel tempering: theory, applications, and new perspectives. *Phys Chem Chem Phys*, 7(23):3910–3916, Dec 2005.
- [31] L. Edelstein-Keshet and B. Ermentrout. Models for branching networks in two dimensions. *SIAM Journal on Applied Mathematics*, 49(4):1136–1157, 1989.
- [32] L. Edelstein-Keshet and G. B. Ermentrout. Models for the length distributions of actin filaments: I. simple polymerization and fragmentation. *Bull Math Biol*, 60(3):449–475, 1998.
- [33] L. Edelstein-Keshet and G. B. Ermentrout. Models for spatial polymerization dynamics of rod-like polymers. *J Math Biol*, 40(1):64–96, 2000.
- [34] L. Edelstein-Keshet and G. B. Ermentrout. A model for actin-filament length distribution in a lamellipod. *J Math Biol*, 43(4):325–355, 2001.
- [35] M. Enculescu, A. Gholami, and M. Falcke. Dynamic regimes and bifurcations in a model of actin-based motility. *Phys Rev E Stat Nonlin Soft Matter Phys*, 78(3 Pt 1):031915, 2008.
- [36] M. Enculescu, M. Sabouri-Ghomi, G. Danuser, and M. Falcke. Modeling of protrusion phenotypes driven by the actin-membrane interaction. *Biophys J*, 98(8):1571–1581, 2010.

- [37] G. B. Ermentrout and L. Edelstein-Keshet. Models for the length distributions of actin filaments: II. polymerization and fragmentation by gelsolin acting together. *Bull Math Biol*, 60(3):477–503, 1998.
- [38] G. B. Ermentrout and J. B. McLeod. Existence and uniqueness of travelling waves for a neural network. *Proceedings of the Royal Society of Edinburgh, Section: A Mathematics*, 123(03):461–478, 1993.
- [39] H. Fukunishi, O. Watanabe, and S. Takada. On the hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *J. Chem. Phys.*, 116(20):9058–9067, 2002.
- [40] A. E. Garcia and K. Y. Sanbonmatsu. Exploring the energy landscape of a beta hairpin in explicit solvent. *Proteins: Struct., Funct., Bioinf.*, 42(3):345–354, 2001.
- [41] M. L. Gardel, I. C. Schneider, Y. Aratyn-Schaus, and C. M. Waterman. Mechanical integration of actin and adhesion dynamics in cell migration. *Annu Rev Cell Dev Biol*, 26:315–333, 2010.
- [42] C. Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics)*. Springer, 3rd edition, 2004. ISBN 3540208828.
- [43] E. Geigant, K. Ladizhansky, and A. Mogilner. An integrodifferential model for orientational distributions of f-actin in cells. *SIAM Journal on Applied Mathematics*, 59(3): pp. 787–809, 1998.
- [44] C. J. Geyer. Markov chain Monte Carlo maximum likelihood. In *Computing Science and Statistics: Proc. 23rd Symp. Interface*, pages 156–163, 1991.
- [45] H. Ghodsi and M. Kazemi. Elastic properties of actin assemblies in different states of nucleotide binding. *Cellular and Molecular Bioengineering*, 5:1–13, 2012. 10.1007/s12195-011-0181-z.
- [46] P. A. Giardini, D. A. Fletcher, and J. A. Theriot. Compression forces generated by actin comet tails on lipid vesicles. *Proc Natl Acad Sci U S A*, 100(11):6493–6498, 2003.
- [47] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [48] M. B. Goldberg and J. A. Theriot. Shigella flexneri surface protein icsa is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A*, 92(14):6572–6576, 1995.
- [49] E. D. Goley and M. D. Welch. The arp2/3 complex: an actin nucleator comes of age. *Nat Rev Mol Cell Biol*, 7(10):713–726, 2006.

- [50] G. W. Greene, T. H. Anderson, H. Zeng, B. Zappone, and J. N. Israelachvili. Force amplification response of actin filaments under confined compression. *Proceedings of the National Academy of Sciences*, 2009.
- [51] U. H. E. Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.*, 281(1-3):140 – 150, 1997.
- [52] F. Heinemann, H. Doschke, and M. Radmacher. Keratocyte lamellipodial protrusion is characterized by a concave force-velocity relation. *Biophys J*, 100(6):1420–1427, 2011.
- [53] J. Hritz and C. Oostenbrink. Hamiltonian replica exchange molecular dynamics using soft-core interactions. *J Chem Phys*, 128(14):144121, 2008.
- [54] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001–.
- [55] K. Klenin, B. Strodel, D. J. Wales, and W. Wenzel. Modelling proteins: Conformational sampling and reconstruction of folding kinetics. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics*, 1814(8):977–1000, 2011.
- [56] S. A. Koestler, S. Auinger, M. Vinzenz, K. Rottner, and J. V. Small. Differentially oriented populations of actin filaments generated in lamellipodia collaborate in pushing and pausing at the cell front. *Nat Cell Biol*, 10(3):306–313, 2008.
- [57] D. A. Kofke. On the acceptance probability of replica-exchange Monte Carlo trials. *The Journal of Chemical Physics*, 117(15):6911–6914, 2002.
- [58] A. Kone and D. A. Kofke. Selection of temperature intervals for parallel-tempering simulations. *J Chem Phys*, 122(20):206101, 2005.
- [59] G. König and S. Boresch. Non-boltzmann sampling and bennett’s acceptance ratio method: how to profit from bending the rules. *J Comput Chem*, 32(6):1082–1090, 2011.
- [60] M. G. Kreĭn and M. A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Uspehi Matem. Nauk (N. S.)*, 3(1(23)):3–95, 1948.
- [61] K.-C. Lee and A. J. Liu. Force-velocity relation for actin-polymerization-driven motility from brownian dynamics simulations. *Biophys J*, 97(5):1295–1304, 2009.
- [62] H. Li, G. Li, B. A. Berg, and W. Yang. Finite reservoir replica exchange to enhance canonical sampling in rugged energy surfaces. *J. Chem. Phys.*, 125(14):144902, 2006.
- [63] P. Liu, B. Kim, R. A. Friesner, and B. J. Berne. Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proc. Natl. Acad. Sci. U. S. A.*, 102(39):13749–13754, 2005.

- [64] A. Liwo, C. Czaplewski, S. Ołdziej, and H. A. Scheraga. Computational techniques for efficient conformational sampling of proteins. *Current Opinion in Structural Biology*, 18(2):134–139, 2008.
- [65] T. P. Loisel, R. Boujemaa, D. Pantaloni, and M. F. Carlier. Reconstitution of actin-based motility of listeria and shigella using pure proteins. *Nature*, 401(6753):613–616, 1999.
- [66] E. Lyman and D. M. Zuckerman. Resolution exchange simulation with incremental coarsening. *J. Chem. Theory Comput.*, 2(3):656–666, 2006.
- [67] E. Lyman and D. M. Zuckerman. Ensemble-based convergence analysis of biomolecular trajectories. *Biophys. J.*, 91(1):164 – 172, 2006.
- [68] I. V. Maly and G. G. Borisy. Self-organization of a propulsive actin network as an evolutionary process. *Proc Natl Acad Sci U S A*, 98(20):11324–11329, 2001.
- [69] Y. Marcy, J. Prost, M.-F. Carlier, and C. Sykes. Forces generated during actin-based propulsion: A direct measurement by micromanipulation. *Proceedings of the National Academy of Sciences of the United States of America*, 101(16):5992–5997, 2004.
- [70] M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, 1998.
- [71] S. Matsushita, Y. Inoue, M. Hojo, M. Sokabe, and T. Adachi. Effect of tensile force on the mechanical behavior of actin filaments. *Journal of Biomechanics*, 44(9):1776 – 1781, 2011.
- [72] P. K. Mattila and P. Lappalainen. Filopodia: molecular architecture and cellular functions. *Nat Rev Mol Cell Biol*, 9(6):446–454, 2008.
- [73] J. L. McGrath, N. J. Eungdamrong, C. I. Fisher, F. Peng, L. Mahadevan, T. J. Mitchison, and S. C. Kuo. The force-velocity relationship for the actin-based motility of listeria monocytogenes. *Curr Biol*, 13(4):329–332, 2003.
- [74] A. Mogilner. Mathematics of cell motility: have we got its number? *J Math Biol*, 58(1-2):105–134, 2009.
- [75] A. Mogilner and L. Edelstein-Keshet. Selecting a common direction. *Journal of Mathematical Biology*, 33:619–660, 1995.
- [76] A. Mogilner and K. Keren. The shape of motile cells. *Curr Biol*, 19(17):R762–R771, 2009.
- [77] A. Mogilner and G. Oster. Cell motility driven by actin polymerization. *Biophys J*, 71(6):3030–3045, 1996.

- [78] A. Mogilner and G. Oster. Force generation by actin polymerization ii: the elastic ratchet and tethered filaments. *Biophys J*, 84(3):1591–1605, 2003.
- [79] A. Mogilner, L. Edelstein-Keshet, and G. B. Ermentrout. Selecting a common direction. ii. peak-like solutions representing total alignment of cell clusters. *J Math Biol*, 34(8): 811–842, 1996.
- [80] R. D. Mullins, J. A. Heuser, and T. D. Pollard. The interaction of arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc Natl Acad Sci U S A*, 95(11):6181–6186, 1998.
- [81] V. Noireaux, R. M. Golsteyn, E. Friederich, J. Prost, C. Antony, D. Louvard, and C. Sykes. Growing an actin gel on spherical surfaces. *Biophys J*, 78(3):1643–1654, 2000.
- [82] Y. Okamoto, M. Fukugita, T. Nakazawa, and H. Kawai. Alpha-helix folding by monte carlo simulated annealing in isolated c-peptide of ribonuclease a. *Protein Eng*, 4(6): 639–647, 1991.
- [83] A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak, and C. Simmerling. Improved efficiency of replica exchange simulations through use of a hybrid explicit/implicit solvation model. *J. Chem. Theory Comput.*, 2(2):420–433, 2006.
- [84] A. Okur, D. Roe, G. Cui, V. Hornak, and C. Simmerling. Improving convergence of replica-exchange simulations through coupling to a high-temperature structure reservoir. *J. Chem. Theory Comput.*, 3(2):557–568, 2007.
- [85] T. E. Oliphant. *Guide to NumPy*. Provo, UT, 2006.
- [86] T. E. Oliphant. Python for scientific computing. *Computing in Science & Engineering*, 9(3):10–20, 2007.
- [87] G. F. Oster and A. S. Perelson. Cell spreading and motility: A model lamellipod. *Journal of Mathematical Biology*, 21:383–388, 1985.
- [88] Ott, Magnasco, Simon, and Libchaber. Measurement of the persistence length of polymerized actin using fluorescence microscopy. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 48(3):R1642–R1645, 1993.
- [89] S. H. Parekh, O. Chaudhuri, J. A. Theriot, and D. A. Fletcher. Loading history determines the velocity of actin-network growth. *Nat Cell Biol*, 7(12):1219–1223, 2005. in-vitro concave force-velocity curves shows a force-loading hysteresis.
- [90] J. R. Partington. The resolvent of a hermitian operator on a banach space. *Journal of the London Mathematical Society*, s2-27(3):507–512, 1983.

- [91] D. Paschek and A. E. Garcia. Reversible temperature and pressure denaturation of a protein fragment: A replica exchange molecular dynamics simulation study. *Phys. Rev. Lett.*, 93:238105, 2004.
- [92] A. Patriksson and D. van der Spoel. A temperature predictor for parallel tempering simulations. *Phys. Chem. Chem. Phys.*, 10:2073–2077, 2008.
- [93] C. S. Peskin, G. M. Odell, and G. F. Oster. Cellular motions and thermal fluctuations: the brownian ratchet. *Biophys J*, 65(1):316–324, 1993.
- [94] J. Pfaendtner, D. Branduardi, M. Parrinello, T. D. Pollard, and G. A. Voth. Nucleotide-dependent conformational states of actin. *Proceedings of the National Academy of Sciences*, 106(31):12723–12728, 2009.
- [95] J. W. Pitera and W. Swope. Understanding folding and design: Replica-exchange simulations of “trp-cage” miniproteins. *Proc. Natl. Acad. Sci. U. S. A.*, 100(13):7587–7592, 2003.
- [96] J. L. Podolski and T. L. Steck. Length distribution of f-actin in dictyostelium discoideum. *J Biol Chem*, 265(3):1312–1318, 1990.
- [97] T. D. Pollard and J. Berro. Mathematical models and simulations of cellular processes based on actin filaments. *Journal of Biological Chemistry*, 284(9):5433–5437, 2009.
- [98] T. D. Pollard and G. G. Borisy. Cellular motility driven by assembly and disassembly of actin filaments. *Cell*, 112(4):453–465, 2003.
- [99] T. D. Pollard and J. A. Cooper. Actin, a central player in cell shape and movement. *Science*, 326(5957):1208–1212, 2009.
- [100] T. D. Pollard, L. Blanchoin, and R. D. Mullins. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu Rev Biophys Biomol Struct*, 29:545–576, 2000.
- [101] M. Prass, K. Jacobson, A. Mogilner, and M. Radmacher. Direct measurement of the lamellipodial protrusive force in a migrating cell. *J Cell Biol*, 174(6):767–772, 2006.
- [102] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C (2nd ed.): the art of scientific computing*. Cambridge University Press, New York, NY, USA, 1992. ISBN 0-521-43108-5.
- [103] D. Quint and J. Schwarz. Optimal orientation in branched cytoskeletal networks. *Journal of Mathematical Biology*, 63:735–755, 2011.
- [104] S. M. Rafelski and J. A. Theriot. Crawling toward a unified model of cell mobility: spatial and temporal regulation of actin dynamics. *Annu Rev Biochem*, 73:209–239, 2004.

- [105] N. Rathore, M. Chopra, and J. J. de Pablo. Optimal allocation of replicas in parallel tempering simulations. *J Chem Phys*, 122(2):024111+, 2005.
- [106] S. W. Rick. Replica exchange with dynamical scaling. *J. Chem. Phys.*, 126(5):054102, 2007.
- [107] A. Roitberg, A. Okur, and C. Simmerling. Coupling of replica exchange simulations to a non-boltzmann structure reservoir. *J. Phys. Chem. B*, 111(10):2415–2418, 2007.
- [108] E. Rosta and G. Hummer. Error and efficiency of replica exchange molecular dynamics simulations. *J Chem Phys*, 131(16):165102+, 2009.
- [109] E. Rosta and G. Hummer. Error and efficiency of simulated tempering simulations. *J Chem Phys*, 132(3):034102, Jan 2010.
- [110] D. Sabo, M. Meuwly, D. L. Freeman, and J. D. Doll. A constant entropy increase model for the selection of parallel tempering ensembles. *J Chem Phys*, 128(17):174109, 2008.
- [111] K. Sanbonmatsu and A. Garca. Structure of met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins: Struct., Funct., Bioinf.*, 46(2):225–234, 2002.
- [112] D. A. Schafer, P. B. Jennings, and J. A. Cooper. Dynamics of capping protein and actin assembly in vitro: uncapping barbed ends by polyphosphoinositides. *J Cell Biol*, 135(1):169–179, 1996.
- [113] S. Schaub, J.-J. Meister, and A. B. Verkhovsky. Analysis of actin filament network organization in lamellipodia by comparing experimental and simulated images. *J Cell Sci*, 120(Pt 8):1491–1500, 2007.
- [114] T. E. Schaus and G. G. Borisy. Performance of a population of independent filaments in lamellipodial protrusion. *Biophys J*, 95(3):1393–1411, 2008.
- [115] T. E. Schaus, E. W. Taylor, and G. G. Borisy. Self-organization of actin filament orientation in the dendritic-nucleation/array-treadmilling model. *Proc Natl Acad Sci U S A*, 104(17):7086–7091, 2007.
- [116] T. Schlick. *Molecular Modeling and Simulation*. Springer, 1 edition, 2002. ISBN 038795404X.
- [117] C. H. Schreiber, M. Stewart, and T. Duke. Simulation of cell motility that reproduces the force-velocity relationship. *Proc Natl Acad Sci U S A*, 107(20):9141–9146, 2010.
- [118] D. Sept and J. A. McCammon. Thermodynamics and kinetics of actin filament nucleation. *Biophysical Journal*, 81(2):667 – 674, 2001.

- [119] J. W. Shaevitz and D. A. Fletcher. Load fluctuations drive actin network growth. *Proc Natl Acad Sci U S A*, 104(40):15688–15692, 2007.
- [120] A. M. Sinclair. The norm of a hermitian element in a banach algebra. *Proceedings of the American Mathematical Society*, 28(2):pp. 446–450, 1971.
- [121] D. Sindhikara, Y. Meng, and A. E. Roitberg. Exchange frequency in replica exchange molecular dynamics. *J. Chem. Phys.*, 128(2):024103, 2008.
- [122] D. J. Sindhikara, D. J. Emerson, and A. E. Roitberg. Exchange often and properly in replica exchange molecular dynamics. *J. Chem. Theory Comput.*, 6(9):2804–2808, 2010.
- [123] J. V. Small. Lamellipodia architecture: actin filament turnover and the lateral flow of actin filaments during motility. *Semin Cell Biol*, 5(3):157–163, 1994.
- [124] J. V. Small, M. Herzog, and K. Anderson. Actin filament organization in the fish keratocyte lamellipodium. *J Cell Biol*, 129(5):1275–1286, 1995.
- [125] J. V. Small, T. Stradal, E. Vignal, and K. Rottner. The lamellipodium: where motility begins. *Trends Cell Biol*, 12(3):112–120, 2002.
- [126] T. Splettstoesser, K. C. Holmes, F. No, and J. C. Smith. Structural modeling and molecular dynamics simulation of the actin filament. *Proteins*, 79(7):2033–2043, 2011.
- [127] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314(1-2):141 – 151, 1999.
- [128] Y. Sugita and Y. Okamoto. Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape. *Chem. Phys. Lett.*, 329(3-4):261 – 270, 2000.
- [129] Y. Sugita, A. Kitao, and Y. Okamoto. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.*, 113(15):6042–6051, 2000.
- [130] T. Svitkina. N-wasp generates a buzz at membranes on the move. *Cell*, 128(5):828–830, 2007.
- [131] T. M. Svitkina and G. G. Borisy. Arp2/3 complex and actin depolymerizing factor/-cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *J Cell Biol*, 145(5):1009–1026, 1999.
- [132] T. M. Svitkina, A. B. Verkhovsky, K. M. McQuade, and G. G. Borisy. Analysis of the actin-myosin ii system in fish epidermal keratocytes: mechanism of cell body translocation. *J Cell Biol*, 139(2):397–415, 1997.
- [133] R. H. Swendsen and J.-S. Wang. Replica monte carlo simulation of spin-glasses. *Phys. Rev. Lett.*, 57:2607–2609, Nov 1986.

- [134] K. Tai. Conformational sampling for the impatient. *Biophys Chem*, 107(3):213–220, 2004.
- [135] A. Upadhyaya, J. R. Chabot, A. Andreeva, A. Samadani, and A. van Oudenaarden. Probing polymerization forces by using actin-propelled lipid vesicles. *Proc Natl Acad Sci U S A*, 100(8):4521–4526, 2003.
- [136] E. Vanden-Eijnden and F. A. Tal. Transition state theory: variational formulation, dynamical corrections, and error estimates. *J Chem Phys*, 123(18):184103, 2005.
- [137] J. Weichsel and U. S. Schwarz. Two competing orientation patterns explain experimentally observed anomalies in growing actin networks. *Proceedings of the National Academy of Sciences*, 107(14):6304–6309, 2010.
- [138] J. Weichsel, E. Urban, J. V. Small, and U. S. Schwarz. Reconstructing the orientation distribution of actin filaments in the lamellipodium of migrating keratocytes from electron microscopy tomography data. *Cytometry A*, 81(6):496–507, 2012.
- [139] S. Wiesner, E. Helfer, D. Didry, G. Ducouret, F. Lafuma, M.-F. Carlier, and D. Pantaloni. A biomimetic motility assay provides insight into the mechanism of actin-based motility. *J Cell Biol*, 160(3):387–398, 2003.
- [140] K. Zhang. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci*, 16(6):2112–2126, 1996.
- [141] W. Zheng, M. Andrec, E. Gallicchio, and R. M. Levy. Simulating replica exchange simulations of protein folding with a kinetic network model. *Proc. Natl. Acad. Sci. U. S. A.*, 104(39):15340–15345, 2007.
- [142] W. Zheng, M. Andrec, E. Gallicchio, and R. M. Levy. Simple continuous and discrete models for simulating replica exchange simulations of protein folding. *J Phys Chem B*, 112(19):6083–6093, 2008.
- [143] X. Zheng, K. Diraviyam, and D. Sept. Nucleotide effects on the structure and dynamics of actin. *Biophysical Journal*, 93(4):1277 – 1283, 2007.
- [144] R. Zhou, B. J. Berne, and R. Germain. The free energy landscape for beta hairpin folding in explicit water. *Proc. Natl. Acad. Sci. U. S. A.*, 98(26):14931–14936, 2001.
- [145] J. Zimmermann, M. Enculescu, and M. Falcke. Leading-edge-gel coupling in lamellipodium motion. *Phys Rev E Stat Nonlin Soft Matter Phys*, 82(5 Pt 1):051925, 2010.
- [146] D. M. Zuckerman. Equilibrium sampling in biomolecular simulations. *Annual Review of Biophysics*, 40(1):41–62, 2011.