ABSTRACT

Title of Dissertation: NON-EQUILIBRIUM THERMODYNAMICS
OF CYTOSKELETAL SELF-ORGANIZATION

Carlos Floyd
Doctor of Philosophy, 2021

Dissertation Directed by: Professor Garegin A. Papoian
Department of Chemistry and Biochemistry

The actin-based cytoskeleton is a polymer network that plays an essential role in cell biology. By self-organizing into various local architectures, the cytoskeleton performs physiological functions that allow the cell to physically interact with its environment. It is also an example of biological active matter, consuming chemical free energy at a local scale to produce directed motion and do mechanical work. While it is well-known that cytoskeletal free energy transduction occurs, it has been a challenge to say anything quantitative about this far-from-equilibrium process due to the difficulty of making the necessary experimental measurements. This lack of methodology to quantify cytoskeletal energetics significantly hinders our understanding of the self-organization process underlying the cytoskeleton's physiological functionality. To address this research gap, we develop in this thesis an explicit computational method to quantify chemical and mechanical free energy changes during simulated cytoskeletal self-organization using the software package MEDYAN (Mechanochemical Dynamics of Active Networks). We then apply this tool in several studies to advance our understanding of the self-organization process and its thermodynamic characteristics. For instance, we analyze the thermodynamic efficiency of mechanical stress generation and the network's time-dependent dissipation rates under a range of network conditions. We also investigate the recent experimentally discovered phenomenon of cytoskeletal avalanches, which we identify in simulation as anomalous mechanical dissipation events. Our analysis clarifies the phenomenology and underlying mechanism of these avalanche events, which we propose may play an important role in cellular information processing. The *in silico* method developed in this thesis provides a new perspective on cytoskeletal self-organization and may be extended to investigate other biological active matter systems.

NON-EQUILIBRIUM THERMODYNAMICS OF CYTOSKELETAL SELF-ORGANIZATION

by

Carlos Floyd

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Garegin A. Papoian, Co-Chair/Co-Advisor
Professor Christopher Jarzynski, Co-Chair/Co-Advisor
Professor Michelle Girvan
Professor Sergei Sukharev
Professor Arpita Upadhyaya

Acknowledgments

This work would not have been possible without the people whom I've known over the past six years. First, I would like to express my sincere gratitude toward my advisors, Garyk Papoian and Chris Jarzynski. I have benefited from many helpful discussions with you on scientific topics, developing research ideas, navigating career decisions, and life in general. I'm very fortunate to have spent this time in graduate school working with such great scientists and mentors, and I hope there will be plenty of chances to interact again down the road. My committee of Michelle Girvan, Sergei Sukharev, and Arpita Upadhyaya also deserve my sincere thanks. Your work in guiding this thesis has been of great help, and I appreciate the opportunities I've had along the way to learn from each of you.

I would like to thank the members of the Papoian and Jarzynski labs, who have been colleagues and friends throughout my time in graduate school. The members of the Papoian lab, including Mary Pitman, Hao Wu, Haiqing Zhao, and the MEDYAN developers Aravind Chandrasekaran, James Komianos, Haoran Ni, and Qin Ni have made it a pleasure to come to lab everyday. I am especially indebted to the MEDYAN developers for our close working relationship - tracking down segfaults would have been a much more annoying without your help. I am also grateful to the members and associates of the Jarzynski lab, including Andrew Smith, Wade Hodson, Debankur Bhattacharyya, Guilherme De Sousa, Long Him Cheung, Joshua Chiel, and Kanupriya Sinha. Discussing aspects of non-equilibrium and quantum physics with all of you has been a very rewarding part of my graduate school experience. To the members of both labs, thanks for making the day-to-day work of research more enjoyable.

I would also like to thank the University of Maryland organizations I have been a member of, including the Biophysics program, the Institute for Physical Science and Technology, the Physics of Living Systems network, and the Computation and Mathematics of Biological Networks (COMBINE) program. The faculty and students in these organizations have greatly enriched my time in graduate school. It's been eye-opening to engage with the broad interdisciplinary research being done at this and other universities through seminars and discussions with the members of these organizations. I am particularly grateful for the training and financial support from the COMBINE program. By exposing me to ideas in network science and machine learning and by providing clear advice for professional development, the COMBINE program has had a large positive impact on my time here.

I would like to acknowledge our co-workers at the University of Oxford, Radek Erban and Ravinda Gunaratne. While it isn't presented in this thesis, I have learned a lot from our work together. I have also benefited from a productive collaboration with Herbert Levine developing the work on cytoskeletal avalanches. I wish to thank my undergraduate advisor, Rosa Tamara Branca, as well, for guiding me through my first research experience and helping with my graduate school applications.

Finally, I wouldn't have been able to reach this milestone in my life without the support of my family and friends. My parents, Jim and Katie Floyd, my brother and sister-in-law, Cam and Kim Floyd, and my girlfriend, Michelle Swanson (and our cat Bowie), are a constant source of strength for which I am deeply grateful. To my friends in Maryland, North Carolina, and elsewhere, I thank you as well.

Contents

List of Figures

List of Tables

List of Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
AUC	Area under the curve
BC	Brooks-Carlsson
CCDF	Complementary cumulative distribution function
CT	Constant tip
F-actin	Polymerized actin
FPR	False positive rate
G-actin	Unpolymerized actin
GTP	Guanosine triphosphate
in silico	"in silicon"
in vitro	"in the glass"
in vivo	"in the living"
MEDYAN	Mechanochemical dynamics of active networks
ML	Machine learning
MTP	Main treadmilling pathway
NMIIA	Non-muscle myosin II A
ODE	Ordinary differential equation
PCA	Principal component analysis
PCM	Parallel cluster model
PDF	Probability density function
Pi	Inorganic phosphate
QSS	Quasi-steady state
QSSA	Quasi-steady state approximation
RDME	Reaction-diffusion master equation
ROC	Receiver operating characteristic
SOC	Self-organized criticality
STORM	Stochastic optical reconstruction microscopy
TPR	True positive rate

Chapter 1

Introduction

The fundamental unit of life on earth is the cell, and inside every cell is a structural protein network called the *cytoskeleton* [1]. Like the musculoskeletal system in our bodies, the cytoskeleton supports the body of the cell and is involved in dynamic physiological processes that require it to exert coherent mechanical forces [2–4]. This allows the cell to physically negotiate with its surroundings. These dynamic physiological processes, which we give examples of below, result from the cytoskeleton consuming chemical fuel to emergently and collectively self-organize into specific local architectures tailored to the physiological task. Whereas some hard-shelled organisms get by with no internal skeleton, there are no "cellular invertebrates" that lack a cytoskeleton; it is present in some form in the cells of all three domains of life (eukaryotes, bacteria, and archaea) [5–7]. Understanding in detail how cytoskeletal systems self-organize to perform their physiological functions is therefore an important piece of understanding life in general.

Much of the underlying chemistry and physics relevant to cytoskeletal networks have already been mapped out, as experimental investigations have uncovered a wealth of knowledge about the structural and mechanical properties of cytoskeletal proteins as well as the chemical reactions in which those proteins participate [8–19]. As with other complex systems, though, knowing about the components of cytoskeletal networks does not fully explain their emergent behaviors, and it is a challenge to synthesize this experimental knowledge to address questions about the collective dynamics of many interacting cytoskeletal components [20–22]. An important set of such questions pertains to the *thermodynamics* of cytoskeletal self-organization, i.e. the changes in the network's chemical and mechanical energy and the exchange of heat with its surroundings [23, 24]. Thermodynamics is a pillar of classical physics and a natural framework through which to analyze the engine-like transduction of chemical fuel into mechanical work that enables cytoskeletal self-organization and its physiological functionality [25–27]. While experiments can probe the energetics of processes involving individual cytoskeletal components, it is nearly impossible to experimentally access the energetics of entire cytoskeletal networks [28–31].[1] As a result, very little is known about system-wide free energy transduction in cytoskeletal networks, a vital research gap that hinders our theoretical understanding of cytoskeletal self-organization.

As with other complex systems, a powerful methodological approach to understanding the emergent behaviors of cytoskeletal networks is *computer simulation* [37, 38]. Mesoscopic simulations allow one to reproduce the complicated collective behaviors observed in living cytoskeletal networks while also giving access to ample amounts of information generally not available through experimental measurements [39–42]. Although the computational modeling of cytoskeletal dynamics is becoming a mature and well-developed technique, it has not yet been effectively used to monitor cytoskeletal thermodynamics at the network level. Doing so requires combining in a consistent way the biophysical models and protocols used in simulation with a thermodynamic treatment of mesoscopic non-equilibrium systems.

In this thesis, we develop such a method to computationally quantify the thermodynamics of cytoskeletal networks using the simulation platform MEDYAN (Mechanochemical Dynamics of Active Networks), and apply it to investigate open problems in the theory of cytoskeletal self-organization. In the introduction, we first describe in greater detail the cytoskeleton, its interpretation as an active matter system, and options for

---

[1] There is an interesting partial exception to this limitation of experimental methods. Recent theoretical work has illustrated that the presence of non-equilibrium driving in complex systems can be detected as non-zero probability currents in a low-dimensional projection of the system's phase space [32, 33]. This work was extended to not just detect but actually quantify the approximate entropy production of cytoskeletal networks [34, 35]. This method underestimates the total entropy production by only analyzing currents in the low-dimensional projection, but certain interesting trends can still be observed using this approach [36]. We compare our explicit computational method to quantify entropy production with this indirect experimental method in Chapter 4.

modeling it computationally. We then we give an overview of the thermodynamics of mesoscopic systems and how it may help to explain active matter self-organization, and at the end we outline the subsequent chapters of the thesis.

1.1 The cytoskeleton as biological active matter

1.1.1 Functions and components of the cytoskeleton

The cytoskeleton plays a central role in the life of a cell. It is an interconnected network of polymers and accessory proteins which extends throughout the cell, adopting varied local architectures in different regions (Figures 1.1 and 1.2).[2] As the cell's "skeleton" it structurally supports the cell membrane like a scaffold, and it further participates in numerous processes which generally result in the production of useful mechanical work through the consumption of an out-of-equilibrium supply of chemical fuel, typically adenosine triphosphate (ATP) or guanosine triphosphate (GTP) [2, 43, 44]. Specific cellular processes which rely on cytoskeletal activity include: endocytosis, in which extracellular material is brought inside the cell via the cytoskeleton's coordinated mechanical manipulations of the cell membrane [45–47]; cell migration, in which the leading edge of the cell is pushed forward through active polymerization of actin fibers while the rear of the cell is pulled through the contraction of actin stress fibers (Figure 1.2) [48, 49]; cytokinesis, in which the cytoskeleton cinches the cell along the equatorial plane during cell division [50, 51]. These examples begin to illustrate the variety of function and structure exhibited by the cytoskeleton, a variety which is perhaps nowhere more evident than in neurons [52]. By selectively activating accessory proteins which interact with the polymer network in various ways, the cell can finely control the behavior of the cytoskeletal machinery. These accessory proteins fall into different classes based on their function [43, 53, 54]. These classes include: capping proteins which inhibit polymerization of filaments; cross-linking proteins which bind nearby filaments together; severing proteins

---

[2] Although a cytoskeleton is found in the cells of all domains of life, we will focus throughout this thesis on eukaryotes. Thus the particular protein components we mention are specific to the eukaryotic cytoskeleton.

which cut filaments into smaller pieces; sequestering proteins which remove monomers from the available pool and slow down polymerization; nucleators which promote the creation of new filaments. From *in vivo* experiments it is becoming clear which accessory proteins are implicated in which physiological processes, although in many cases the detailed mechanisms are not well understood [55–62].

Even though living cells utilize an array of accessory proteins to control cytoskeletal behavior, a rich variety of interesting dynamical behaviors can be observed using just a subset of these many components. For instance, reconstituted systems of only cytoskeletal filaments and molecular motors are sufficient to produce several qualitatively different dynamic behaviors, including orientationally coherent flocking, rotating vortices, traveling plane waves, moving nematic defects, and large-scale contractility [67–69]. In this thesis we primarily focus on a parsimonious subset of cytoskeletal proteins constituting a system called an *actomyosin network* (Figure 1.3.A). We will be concerned with three proteins: actin filaments, which are $\sim 8\ nm$ in diameter semi-flexible polar polymers of varying length that act as mechanical scaffold (Figure 1.3.B) [8, 55]; non-muscle myosin II A (NMIIA) minifilaments, which are $\sim\ 200\ nm$ long aggregates of tens of myosin heads that bind to actin filaments and walk in a directed manner toward the filament plus ends (Figure 1.3.C) [19, 70]; α-actinin, a cross-linking protein $\sim 35\ nm$ long that stably binds to adjacent filaments to mechanically connect them (Figure 1.3.D) [71, 72]. We will elaborate on additional aspects of these molecules as they become relevant throughout the thesis.

1.1.2 Active matter

The field of *active matter* reaches across many scales and concerns the collective behavior of interacting agents that can individually convert energy into directed motion [73]. Active matter research is an exciting and recently established field that mixes condensed matter physics, non-equilibrium statistical mechanics, chemistry, and other disciplines to understand and even control the dynamics of driven collectives [74–78]. Active matter systems include swarming flocks of birds [79–81], groups of robots programmed with

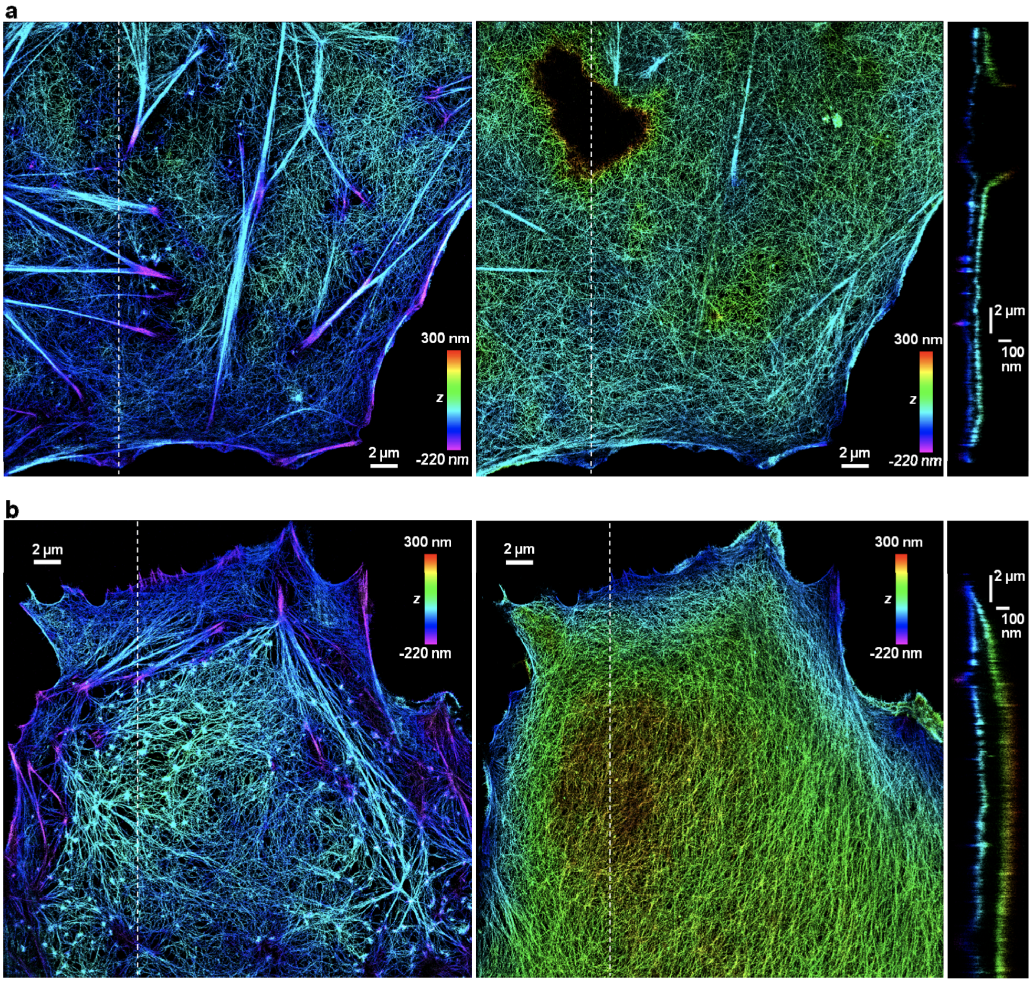

FIGURE 1.1: Images using stochastic optical reconstruction microscopy (STORM) of two layers of the actin cytoskeleton in COS-7 fibroblast cells. Using this technique, individual actin filaments can be resolved in three-dimensional space, showing a remarkable heterogeneous network structure. Actin filaments are colored according to their vertical position in the cell. The left two panels show the ventral part of two cells (which are shown on the top and bottom rows), and the central two panels show the dorsal parts. The rightmost panels show a slice along the dotted lines gray lines indicating more clearly the vertical actin positions. Part of the cell imaged on the top was a bulged structure that was excluded from the imaging range used. Image reprinted from Ref. 63 with permission from *Springer Nature.*

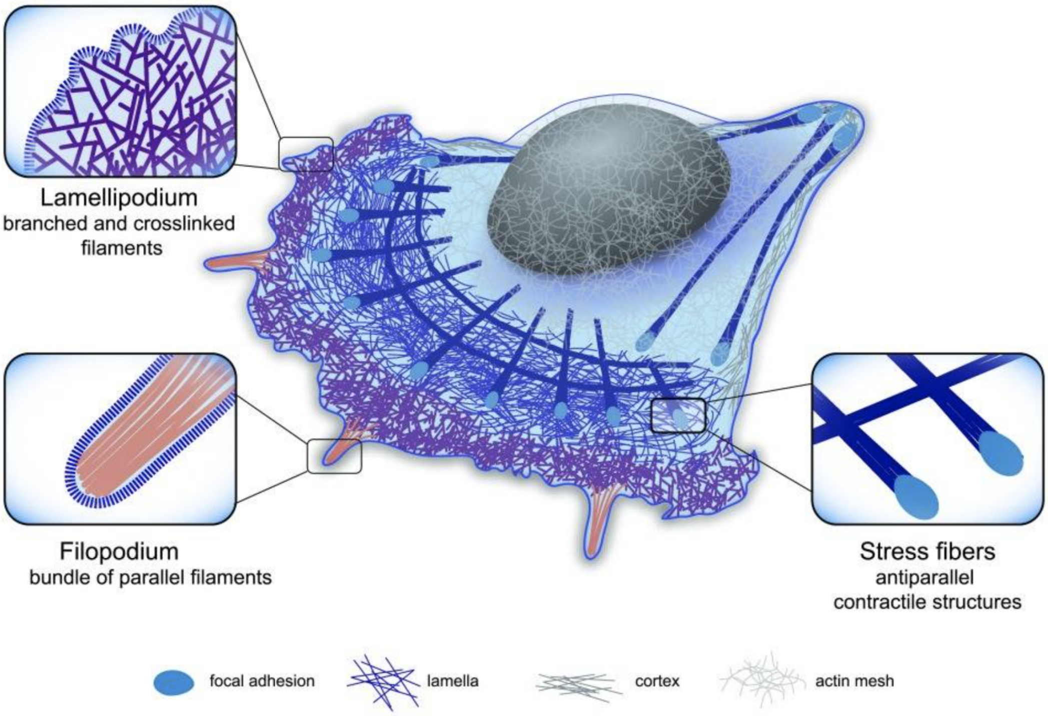

Figure 1.2: A schematic illustration of the various local architectures adopted by the cytoskeleton during cell migration. The cortical cytoskeleton, shown in gray, sits underneath and supports the cell membrane. Stable bundles of actin filaments form stress fibers, shown in dark blue, which produce a strong contractile force that pulls on focal adhesions attached to the underlying substrate, moving the cell along its polarity axis. Filopodial structures, shown in orange, are also formed from bundles of filaments and protrude from the edge of a cell to detect changes in the surrounding chemical environment, guiding cell migration. The lamellipodium, shown in purple, is a highly branched network which protrudes on the cell's edge through directed polymerization to push it forward. This combined process of pushing at the front and pulling up the rear requires a coordination of cytoskeletal self-organization by chemical signaling that differentially activates associated proteins in different regions of the cell. Image reprinted from Ref. 64 with permission from *Faculty of 1000*.

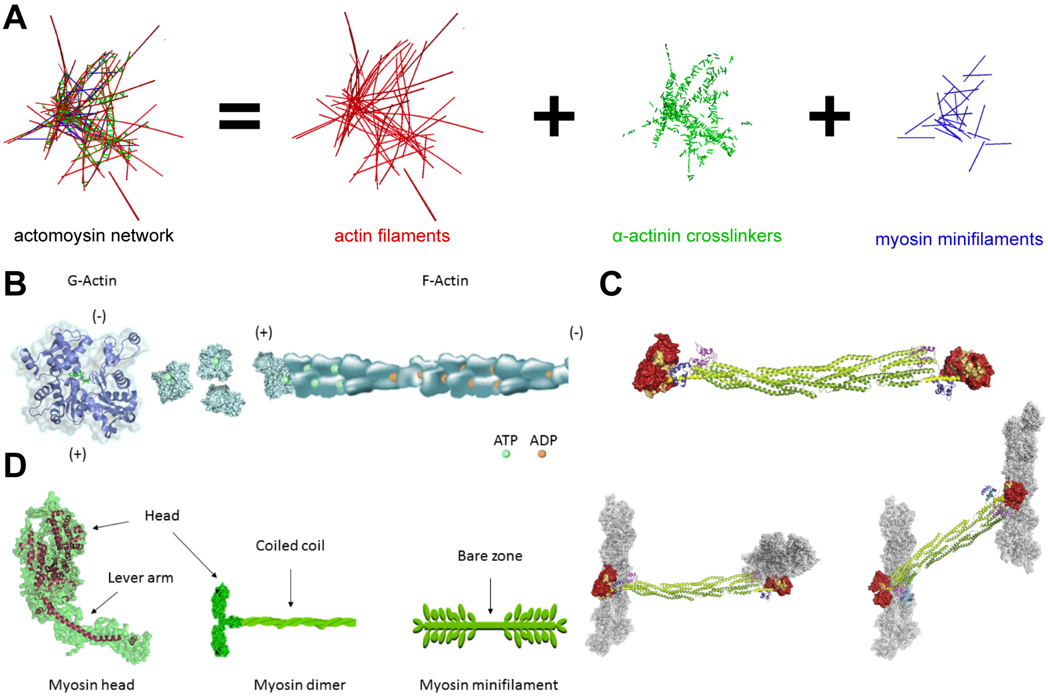

Figure 1.3: Composite image showing the primary components of actomyosin networks considered in this thesis. A) A snapshot from a MEDYAN simulation is shown illustrating how the different components colocalize in space in a self-organized actomyosin network. The color labeling of each protein is indicated by the text color underneath the images. B) An image of the polar, double helical actin filament is shown along with the secondary protein structure of a G-actin monomer. The binding location of adenosine nucleotides is shown as a small green molecule located in the monomer's minus-end cleft and as colored spheres in the filament image. Signaling molecules such as profilin can bind to unpolymerized monomers and control their polymerization kinetics. Image reprinted from Ref. 65 with permission from *IOP Publishing*. C) The secondary structure of a bare α-actinin molecule is shown along with two snapshots of its configuration bound to pairs of actin filaments. Signaling molecules can bind to the ends of the dumbbell-like molecule, shown in red, which activate its actin-binding functionality. Image reprinted from Ref. 66 with permission from *Elsevier*. D) Cartoons of the non-muscle myosin II minifilament and dimers are shown along with the secondary structure of an individual myosin head. The primary features of the myosin molecule including its head, rotating lever arm, and coiled coil tail region are indicated. Near the lever arm are regulatory regions which allow for selective activation through various cell signaling pathways. Tens of myosin heads aggregate to form a myosin minifilament, which has enhanced processivity compared to individual heads. Image reprinted from Ref. 65 with permission from *IOP Publishing*.

simple rules [82, 83], bacterial colonies [84], and actomyosin networks, the subject of our work [85, 86]. We highlight the fact that actomyosin networks are an active matter system because it helps to generalize the work presented in this thesis beyond the specifics of the cytoskeleton; a similar set of questions to those addressed in this thesis could be asked about a broad class of other systems as well.

Like other biological active matter systems, actomyosin networks couple the consumption of a chemical potential energy source to drive certain processes "uphill" away from their equilibrium [26]. This free energy transduction process can produce conformational displacements against a tension, so that mechanical work is done as a result of a spontaneous chemical reaction [87]. Free energy transduction occurs via two molecular mechanisms in actomyosin networks: the power-stroke cycle of myosin motors [88–90] and actin filament polymerization [91, 92].

The power-stroke cycle of myosin is a sequence of chemical states whose net effect is the movement of the myosin head one binding site forward on the actin filament and the hydrolysis of one solvated ATP molecule [70, 90, 93]. Thus, the thermodynamic driving force of an ATP molecule's chemical potential is coupled through the power-stroke cycle to the generation of tension along the myosin minifilament. The force production by myosin walking is transmitted throughout the cytoskeleton partly through the binding of passive (i.e. not ATP-consuming) cross-linking elements, e.g. α-actinin. Recent work has shown that cross-linkers are in fact necessary to produce robust force generation and contractility in actomyosin networks [94–97].

Actin polymerization involves the hydrolysis of ATP and subsequent release of inorganic phosphate by polymerized monomers in the actin filament, converting ATP-bound polymerized monomers to ADP-bound monomers [14]. ATP-bound monomers tend to accumulate near the filament plus ends while ADP-bound monomers tend to accumulate near and depolymerize from the minus ends [98]. The chemical reaction network describing this process has a steady state in which monomers polymerize at the plus end at the same rate that they depolymerize from the minus end, causing the filament to move forward without its length changing, a situation termed "treadmilling" [99]. The thermodynamic driving force of ATP hydrolysis in this active polymerization process can be coupled via a ratcheting mechanism to push against a resistive load, such as the cell membrane during lamellipodial protrusions in migrating cells [100–102]. In this way, the free energy of ATP is transduced through actin polymerization into cellularly useful mechanical work.

Another important (though not ATP consuming) facet of actomyosin network dynamics are the force-sensitive kinetic reaction rates controlling cross-linker and myosin filament unbinding as well as myosin filament walking. At high tension, cross-linkers will unbind more quickly ("slip-bond") whereas motors will unbind and walk less quickly ("catch-bond" and "stalling") [72, 87, 103]. Additionally, the polymerization rate of an actin filament decreases as the load against which it polymerizes is increased [101]. Intuitively, these force-sensitive reactions can be understood as having binding energy landscapes whose shapes change when a force is applied to the molecule, altering the kinetic rates and either stabilizing or destabilizing the bound state. The mathematical forms for these nonlinear dependencies on force are provided in Appendix A. These force-sensitive reactions control the actomyosin network connectivity, which in turn determines the ability of the network to globally distribute stress [104]. Thus, the force-sensitive feedback introduces nonlinear coupling between the stress sustained by an actomyosin network and the network's ability to reorganize in response to that stress.

1.1.3 Computational modeling of the cytoskeleton

The complexity of cytoskeletal networks makes it difficult to study their emergent dynamics using only mathematical analysis, although certain phenomena such as nonlinear rheological properties and glassy dynamics have been partly rationalized using mathematical models [105–109]. Luckily, however, these emergent properties can be studied in great detail using computer simulations. A key consideration in choosing a simulation method for studying cytoskeletal dynamics is identifying the characteristic time and length scales of the behaviors one is interested in. While highly resolved atomistic simulations can distinguish, for example, the mechanism by which tensile forces alter an

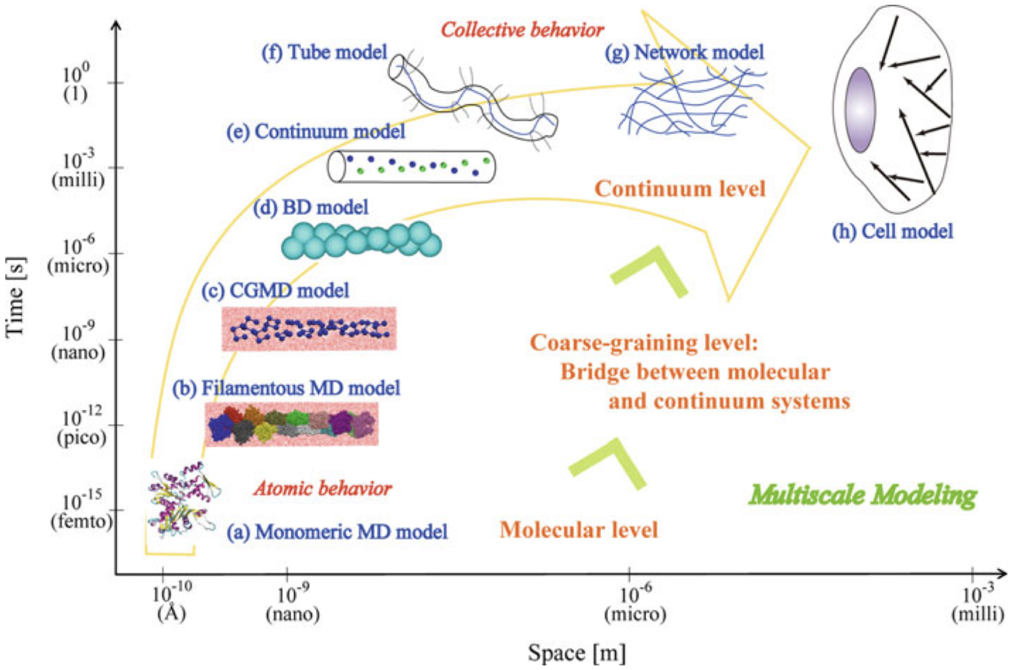

Figure 1.4: The computational modeling the actin cytoskeleton can be done at several length and time scales, and the choice of a model depends on the types of behavior one is interested to study. All-atom models of actin monomers and short filament segments can be coarse-grained, assigning one bead to each domain of the actin monomer. Another coarse-graining step can be taken to assign one bead to each monomer. Further steps eventually yield a model for many interacting filaments, allowing collective network dynamics to be studied. Using the MEDYAN model one can simulate lengths of several micrometers and times of thousands of seconds. Image reprinted from Ref. 112 with permission from *Springer Nature*.

actin monomer's catalytic activity, such simulations could not describe many interacting filaments on a time scale where interesting collective behaviors are known to occur [110, 111]. A hierarchy of simulation methods at different scales is therefore needed to run the gamut of multiscale cytoskeletal dynamics, as illustrated in Figure 1.4. In this thesis we are interested in the emergent dynamics of many cytoskeletal proteins occurring at length scales of up to hundreds of micrometers and time scales of thousands of seconds, the scales at which physiological cytoskeletal processes occur. We therefore require a *coarse-grained* simulation platform for studying semi-flexible polymer networks.

Several such platforms for studying network-level dynamics are available, including AFiNeS, CytoSim, the model of Kim and coworkers, and MEDYAN [39–42]. Each of these allows simulating fascinating emergent cytoskeletal phenomena while striving to preserve realistic microscopic physics. These models use similar spatial resolutions to represent an actin filament, assigning $\sim 10 - 100$ actin monomers to a single discrete element. Additionally, the mechanical energy of the filaments is modeled similarly in these platforms, accounting for the elastic stretching of and bending between the discrete elements. While the mechanical potentials are comparable, the ways that these platforms treat the chemical reaction dynamics are quite different. The chemistry model used in MEDYAN is the most detailed, accounting for the spatial distribution and local reaction propensities of all chemical species in the system rather than assuming a globally well-mixed reservoir. To study the thermodynamics of active cytoskeletal networks we will need to track the chemical energetics at high resolution, for which MEDYAN is the best suited.[3] Indeed, we will find in Chapter 4 that a large contribution to the system's free energy comes the diffusive flux of solvated species, which could not be accounted for in other platforms. In Appendix A, we provide an overview of the MEDYAN simulation platform for reference.

1.2 Non-equilibrium thermodynamics of active matter

1.2.1 Overview of non-equilibrium theories

As with our choice of a computational modeling approach, our choice of a thermodynamic description for cytoskeletal self-organization should reflect the appropriate time

---

[3] One technical drawback of using MEDYAN to study thermodynamics is that the mechanical dynamics do not obey the principle of microscopic reversibility, which roughly says that if the probability of a transition in one direction ($A \to B$) is non-zero then the probability of the reverse transition ($B \to A$) should also be non-zero [113, 114]. In MEDYAN, the chemical dynamics are stochastic, but for a given chemical state and network configuration the new configuration is obtained deterministically through mechanical equilibration. In other words, the mechanical dynamics occur at zero temperature, and reverse transition probabilities are zero when their forward counterparts are non-zero. This approach has greater computational efficiency than, for example, numerically integrating a stiff Langevin equation at small time steps, and it could in principle be straightforwardly amended by including thermal fluctuations in the network's configuration [40]. The methods for tracking thermodynamic quantities presented in this thesis should after some minor modifications still apply for such an amended model, and we would expect only small quantitative corrections to the results presented here. This is because the active fluctuations from ATP-consuming processes which MEDYAN accounts for should dominate over the thermal fluctuations which MEDYAN neglects. This issue is nonetheless an avenue for future work.

and length scales, leveraging where possible the advantages of coarse-graining. Several non-equilibrium theories useful at different scales have been developed, and we first review these theories before discussing which is best suited for tracking cytoskeletal thermodynamics using the MEDYAN platform.

Ignoring the quantum level (whose thermodynamic description encompasses its own growing field of research; see Refs. 115–117), the most fine-grained, microscopic description of a non-equilibrium system coupled to a reservoir tracks the probability density over phase space of all of the system and reservoir's classical degrees of freedom, which obey Hamiltonian dynamics [118–120]. The Liouville equation describes the probability density's time evolution and requires knowing the explicit form for the Hamiltonian. Thermodynamic quantities, such as the work done by the varying an external control parameter and the system and reservoir ensemble entropies, can be defined as functions of their degrees of freedom and their occupied phase space volume (although some technical details, such as whether the system is ergodic and mixing, need to be considered; see Ref. 121).

Assuming weak coupling between the system and reservoir can facilitate calculations within the Liouville framework, but for many systems of interest this interaction and the many degrees of freedom needed to track for the reservoir hopelessly complicate this microscopic approach. This difficulty motivates a coarse-graining step in which the deterministic Hamiltonian dynamics are replaced by stochastic dynamics reflecting the statistical properties of the system-reservoir interaction. In this step the interaction with the reservoir is renormalized into an effective diffusion constant that enters into the either the Fokker-Planck equation describing the time evolution of the system's probability density over a continuous state space, or the Langevin equation describing a single stochastic system trajectory [122, 123]. One could also use a jump process to model the stochastic evolution of the system's microstates interacting with the implicit reservoir. The reservoir's dynamics are assumed to be sufficiently fast for it to instantaneously be at equilibrium for a given system microstate, allowing the system's dynamics to be assumed Markovian. For a single fluctuating trajectory through the system's microstates, the work done on the system by an external driving protocol or non-conservative force as well as the heat exchanged with the implicit reservoir can obtained using functions of the system's degrees of freedom [124, 125]. The heat exchanged with a constant temperature reservoir can be unambiguously associated with the change in the reservoir entropy, but the change in the entropy of the system itself is less clearly defined in the Langevin description since the entropy is typically interpreted as a property of an ensemble, not one trajectory. This issue was resolved through a consistent definition of a fluctuating system entropy which is equal to the usual ensemble entropy when averaged over the time-dependent probability density obtained from the Fokker-Planck solution [126, 127]. These definitions for the thermodynamic quantities of a Markovian system interacting with an implicit reservoir are the basis for the field of "stochastic thermodynamics" [128].

Some physical processes can be well-approximated as transitions through a discrete set of locally equilibrated "mesostates" rather than through the system's full set of microstates. The transitions are still stochastic, with rates that can be derived from the underlying Langevin or jump dynamics [36, 129]. The energy landscapes for which this approximation is realistic can be thought of as "hilly," characterized by local minima separated by relatively high barriers. For such systems, the microstate dynamics can be replaced by a jump process through the discrete mesostate space in another coarse-graining step [130, 131]. The hilly landscapes allowing for the partitioning of state space into a discrete set of local minima typically also allow for a time scale separation between the long time spent occupying a given minimum and the short time spent transitioning over the barriers between minima [129]. If the relaxation timescale within the minimum is also short, then the system loses memory of its previous location quickly enough to assume that the jump dynamics are Markovian. Additionally, this time scale separation can allow one to assume that when the system occupies a given minimum the system explores a locally equilibrated ensemble of configurations before transitioning out, so that a well-defined free energy may be assigned to the mesostate [132]. This notion of a locally equilibrated ensemble of system microstates associated with a given energy minimum is a key part of our thermodynamic description of the cytoskeleton, a system whose dynamics are driven by slow chemical reactions and fast solvent relaxation with a hilly, time scale separated energy surface as just described [109, 133]. We elaborate on this time scale separation in Chapter 3. The state space used in MEDYAN is partly discrete (to describe the copy numbers and locations of diffusing chemical species) and partly continuous (to describe the positions of the network's mechanical elements), and can be viewed as mesoscopic in scale. Thus, our computational thermodynamic description of the cytoskeleton will be at this mesoscopic level, in which the system moves stochastically through a series of locally equilibrated states to which well-defined thermodynamic free energies may be assigned [134].

There is a further set of coarse-graining steps which may be taken to give a macroscopic description of the system. If the system is large enough that fluctuations about the mean behavior are negligible, then one can replace the stochastic dynamics with another deterministic dynamics [25, 135]. Unlike the Hamiltonian dynamics, however, which gives the deterministic evolution of the microscopic degrees of freedom, these give the deterministic evolution of macroscopic thermodynamic forces and fluxes such as chemical concentrations, chemical potential, temperature, and heat. Much research has been dedicated to the macroscopic regime in which the thermodynamic driving forces are small, allowing one to obtain linear transport equations such as Fick's law of diffusion and Fourier's law of heat flow, as well as Onsager's reciprocity relations describing how different types of transport coefficients are interdependent [25].

Finally we mention that, in general, a system may be driven out of thermodynamic equilibrium by two different methods [136]. In the first, the dynamics of the system obey the principle of detailed balance but the system is externally driven through the time-dependent variation of some externally manipulated control parameter (such as the position of a piston wall). By definition of the principle of detailed balance, if the control parameter came to rest then the system would eventually settle to equilibrium with no net probability currents in its phase space. In the second method, there is no externally manipulated control parameter to drive the system, but there is an externally maintained non-zero thermodynamic force on the system (such as out-of-equilibrium chemostatted concentrations of chemical reactants and products), so that the dynamics of the system do not obey detailed balance. This produces a steady state with non-zero net probability currents in phase space. The second method applies to actomyosin networks, which are driven out-of-equilibrium due to a large chemical potential difference between ATP and its hydrolysis products. Throughout this thesis this chemical potential will be fixed for a given simulated network due to an assumption of chemostatted concentrations of ATP, ADP, and Pi.

1.2.2 Using non-equilibrium thermodynamics to explain active matter self-organization

As described above, a defining feature of active matter systems is the consumption and dissipation of free energy by the system's individual constituents, such as myosin motors and actin monomers. Because they are driven, active matter systems can be analyzed using non-equilibrium thermodynamics. Having described the different scales at which the thermodynamics of a non-equilibrium system can be analyzed, we next review several proposed principles that offer ways to explain the self-organization of a non-equilibrium system in terms of its thermodynamic quantities.

In the 1960's, Prigogine proposed that non-equilibrium conditions can allow for the spontaneous emergence of ordered (i.e. seemingly low entropy) spatiotemporal patterns, which he called "dissipative structures." The organization of these structures facilitates the eventual relaxation to equilibrium, but they can persist at steady state if the non-equilibrium boundary conditions are held fixed [23, 137]. In Prigogine's description, dissipative structures represent a broken symmetry of the system. They emerge at the point when certain boundary condition parameters pass a threshold that causes the equilibrium organization of the system to become unstable with respect to thermal fluctuations. A classic example is the nonlinear, auto-catalytic chemical reaction network called the "Brusselator" [138]. As the chemostatted concentrations of the input reactants are increased, a Hopf bifurcation occurs and the steady state concentrations of the dynamical species approach a limit cycle of oscillating concentrations, which is understood a dissipative structure that breaks time translation symmetry. We can view biological organisms and their material parts, such as actomyosin networks, as complex dissipative structures operating within the boundary conditions of their environment. This is an interesting interpretation that places living organisms in the same category as hurricanes and circulating convective fluid patterns called Rayleigh-Bérnard cells, but it is does not have much predictive power and we mention it mainly out of completeness [139, 140].

Among the most provocative developments in macroscopic non-equilibrium thermodynamics are the various extremal principles characterizing non-equilibrium steady states. These principles claim that the manifested steady states of a non-equilibrium system are those for which the entropy production rate is optimized. In these arguments the entropy production rate is formulated as a functional of time-varying yet deterministic thermodynamic forces and fluxes, and whether it is minimized or maximized depends on the boundary conditions and how far away from equilibrium (in the linear or nonlinear regime) the system is. Examples of such principles have been put forward by Helmholtz, Onsager, Prigogine, Gyamarti, and Ziegler [135, 141–144]. For instance, Prigogine's minimum entropy production principle says that the steady state of a non-equilibrium system in the linear response regime corresponds to a minimization of the entropy production over all choices of the unconstrained thermodynamic forces and fluxes. While seeming to offer a powerful unifying approach to understanding non-equilibrium systems, the applicability and accuracy of these principles are heavily debated in the literature, with some researchers pointing to counter-examples as evidence that no such principle could apply generally to all systems [145–149]. Due to these complications it is not clear *a priori* whether we should expect actomyosin networks to obey an extremal principle. Rather than delve into the technical details of these theories, we will take an empirical approach to assess whether an extremal principle applies in our case.

A recent result in non-equilibrium statistical mechanics allows one to express the irreversibility of transitions between the macrostates of fluctuating systems (defined as arbitrary groups of microstates) in terms of the entropy production accompanying the set of microstate trajectories which connect the two macrostates. In principle this allows one to make predictions about the system's likely macrostate transitions in terms of its dissipation. The basis for these results is the Crooks fluctuation theorem, which applies to a system whose dynamics obey the principle of detailed balanced but are non-autonomous due to the variation of some work parameter by some external agent [114, 150, 151]. In 2013, England derived an extension of this principle by averaging Crooks' relation over all paths connecting some beginning and final microstates, and then averaging the resulting relation over all microstates belonging to some beginning and ending macrostates. The implication of this result, which England has termed "dissipative adaptation", is that the irreversibility of macrostate transitions (rather than just the microstate trajectories used in Crooks' relation) increases with the dissipated work accompanying typical system trajectories [152–155]. The fact that high work absorption and dissipation should be associated with high irreversibility is expected, since the work absorption facilitates the traversal of high kinetic barriers, and if the absorbed work is dissipated then returning over the barrier is unlikely, which makes the barrier traversal more irreversible. Actomyosin networks and their architectures are most conveniently described in terms of macrostates, i.e. according to some coarse property such as its radius of gyration, rather than microstates, i.e. the specific arrangement of each polymer in the network. This result of England's is then at a coarse enough level for analyzing transitions between different types of actomyosin architectures whose probabilities of emerging we would like to understand. However, the result of England's is intractable to exactly test for non-"toy model" systems, as it requires knowing both the forward and reverse probabilities of a macrostate transition. The reverse transitions will be exceedingly rare and difficult to measure the probability of. This result also makes no claims about the steady state behavior of a system, instead claiming that the transient dynamics have an irreversibility related to the dissipated work. Due to these issues, it is not clear whether the dissipative adaptation principle is useful to understand cytoskeletal self-organization.

A slightly more operative principle, called "low rattling," was proposed by Chvykov and England in 2018 [83, 156]. It does not actually make any reference to the thermodynamic quantities of a system (such as its entropy production), but focuses instead on the system's kinetic rates to make predictions about the typical steady sate behavior. The intuition behind this principle is that a system will tend to end up in a place it that it is slow to leave from, i.e. in a state where the kinetic rates leading the system away are small. While non-equilibrium systems can be contrived which do not obey this principle (because for instance the states with small exit rates also have small entrance rates), it has been shown that a "typical" system obeys the principle, which can thus be interpreted as a tendency rather than a definite law [156]. In actomyosin networks, the force-sensitive reaction rates (slip and catch-bonds and motor stalling) set the stage for kinetics that depend on the system's configuration, and the low rattling principle may well apply in these systems.

Finally, we consider the non-equilibrium dynamical phenomenon of avalanches, which can appear as heavy tails in the distributions of dissipation event sizes. Avalanche-like processes occur in many systems throughout the natural sciences when there is a combination of slow driving on the system and fast, configuration-dependent dissipation [157]. For certain systems (and evidently in the earth) the event sizes tend to power-law distributions for large systems, a phenomenon called self-organized criticality (SOC) [158–162]. The name comes from the fact that power-law distributions also characterize the fluctuations of order parameters at the critical points of equilibrium statistical systems, i.e. when the system's tuning parameters are at particular critical values [163]. In the non-equilibrium case, though, no parameter fine-tuning is needed to bring about the critical point. As long as the system is slowly driven, rapidly dissipates energy to the environment, and conservatively transfers energy internally, then the critical point is a dynamical attractor towards which the system spontaneously self-organizes [164]. SOC can be observed in many disparate systems and may apply to cytoskeletal networks if the conditions of cytoskeletal driving and dissipation are allowing [165–167]. Even if not technically SOC, cytoskeletal networks may exhibit avalanche-like dynamics in which a cascading dissipation of stored energy results from a comparatively slower energy accumulation process, leading to heavy-tailed (if not power-law) distributions of event sizes. However, the connection between cooperative, avalanche-like non-equilibrium dynamics and cytoskeletal self-organization remains to be established. There is strong experimental motivation for understanding this connection, as *in vivo* studies have discovered heavy-tailed cytoskeletal displacement statistics, and these may explain the unpredictability of measured cellular traction forces [168–170].

1.3 Outline of chapters

Our aim in this thesis is to develop a computational method for quantifying the non-equilibrium thermodynamics of the cytoskeleton and apply it to study open questions in the theory of cytoskeletal self-organization. Some of these open questions will be about determining which of the above principles for non-equilibrium systems apply to cytoskeletal networks. The chapters cover topics of both method development and application and are presented in a logically dependent order, such that earlier chapters should help to understand later chapters. We next outline these chapters and how they relate to each other.

In Chapter 2 we analyze the mean-field chemical dynamics of actin filament polymerization, aging, and depolymerization. This chemical system is active due to the hydrolysis of ATP by the actin subunits and has a non-equilibrium steady state with non-zero net polymerization and depolymerization rates. We present analytical work in which the fast and slow time scales of the chemical dynamics are identified and their separation leveraged to simplify the model. This allows new insights into the transient and steady state behavior of the system. The model of the filament (de)polymerization chemistry in this chapter is later used in more complete simulations of cytoskeletal self-organization based on the MEDYAN platform.

In Chapter 3 we examine how to monitor the Gibbs free energy of a system of solvated chemical species whose dynamics are modeled using the reaction diffusion master equation (RDME). In this approach the copy numbers of the chemical species are discrete quantities and can be comparable in size to the stoichiometric coefficients of the reactions in which the species participate. We find that in this regime certain textbook expressions for the Gibbs free energy change of a reaction are highly inaccurate, requiring us to derive new expressions from first principles. The derivations and formulas presented in this chapter are crucial for tracking cytoskeletal thermodynamics in MEDYAN, which uses the RDME to propagate the chemical dynamics.

Chapter 4 is the central methodological component of this thesis. In this chapter the results of Chapters 2 and 3 as well as additional new modeling considerations are synthesized within the MEDYAN platform to develop a method for tracking the thermodynamics of cytoskeletal self-organization. The method complies with the built-in MEDYAN simulation protocol and accounts for chemical energy consumption as well as the accumulation and dissipation of mechanical energy. We discuss how the method, which could be used for other biological active matter systems, is parameterized specifically for actomyosin networks. The method is then applied in this chapter to study the thermodynamics of actin filament treadmilling, which we compare to the mean-field model predictions of Chapter 2, as well as the efficiency of the transduction of chemical to mechanical energy by motors, which depends on accessory protein concentrations. We generally find dissipation rates to reach a minimum at steady state, in apparent agreement with the minimum entropy production principle mentioned above.

In Chapter 5 we explore several interrelated aspects of cytoskeletal structure and dissipation: the effect of the system size, the role of force-sensitive motor stalling, and the description of cytoskeletal architectures using a graph representation. Using community-detection routines on the graph representation allows for a systematic way to identify higher-order topological features (i.e. filament bundles) in cytoskeletal networks. The system size scaling of the graph metrics and dissipation rate are investigated. We also show that motor stalling can significantly decrease the dissipation rate during the course of cytoskeletal self-organization, which we suggest agrees with the low rattling principle discussed above.

Chapter 6 is dedicated to exploring anomalously large events in cytoskeletal displacements, which have been termed “cytoquakes” by experimenters. Understanding these cooperative mechanical relaxation events and their underlying cause is an open problem, which may have significant biological importance due to a plausible role that cytoquakes might play in enhancing cellular information processing. Our free energy tracking method is applied in this chapter together with several additional new analysis techniques to characterize cytoquake phenomenology and its causes. One new technique is normal mode decomposition of the network, whose trends across different experimental conditions we tangentially explore. Our main results indicate that cytoquakes can indeed be explained as an avalanche-like process of slow energy accumulation punctuated by occasional large, cascading energy dissipation events, although the system is not technically SOC.

Finally, in Chapter 7 we summarize the work presented in this thesis and suggest future research directions. Appendix A presents a self-contained description of the MEDYAN simulation platform for the reader’s reference, and Appendices B-E have supporting information for Chapters 2, 3, 4, and 6.

Chapter 2

Low-dimensional manifold of actin polymerization dynamics

This chapter and Appendix B are adapted from: Floyd, C., Jarzynski, C., & Papoian, G. (2017). Low-dimensional manifold of actin polymerization dynamics. New Journal of Physics, 19(12), 125012.

2.1 Introduction

Actin filaments are an integral part of the cytoskeleton of eukaryotes and are involved in functions such as controlling cell shape, cell motility, organelle redistribution, and mechanical coupling with the cellular environment. These filaments are formed of globular subunits which polymerize in a non-equilibrium process that *in vivo* is modulated by an array of accessory proteins. They are helical and polar, with distinct plus ("barbed") and minus ("pointed") ends at which subunits have different rates of association and dissocation [1]. Hydrolysis of cellular ATP leads to filament "treadmilling", in which there are equal and opposite rates of polymerization at the barbed and pointed ends of the filaments, which drives the polymerization process away from equilibrium and allows actin networks to be responsive to cellular signals on relatively fast timescales [171–173]. Each actin subunit molecule is bound to a nucleotide, which can be in one of several hydrolysis states: adenosine triphophate (ATP), adenosine diphosphate (ADP), or an intermediate state ADP-Pi, in which ADP is still bound to a hydrolyzed inorganic phosphate molecule. Release of inorganic phosphate by ADP-Pi converts it to ADP [174]. The hydrolysis state of the bound nucleotide has dramatic effects on the kinetic polymerization and depolymerization rate constants of the globular subunit [11]. In addition, these hydrolysis states affect the binding affinity of accessory proteins as well as structural properties such as filament persistence length [175, 176]. Thus it is of interest to be able to predict the hydrolysis state of the nucleotide bound to each actin subunit in a filament, or at least the fraction of actin subunits bound to nucleotides in a certain hydrolysis state.

Over several decades a variety of models describing actin polymerization dynamics have been put forward, and these models have evolved alongside the growth of experimental knowledge about the nature of actin. Some early models tracked the number of filaments with a certain degree of polymerization under different assumptions about the filament polarity, geometry, and the cooperativity of polymerization, among other factors [14, 16, 27, 177–179]. Polymerization and depolymerization rate constants for ATP and ADP-bound actin were measured for the first time in 1986 [11]. A subset of more recent models have investigated aspects such as the effects of accessory proteins on actin polymerization via tracking the time-varying concentration of actin subunits distinguished by their polymerization state and by the hydrolysis state of the nucleotide they are bound to. A variable is assigned to the concentration of each species and complexes between certain species, and equations of motion in terms of mean-field kinetic rate constants are written for each. In this context, “mean-field” refers to the assumption that the the solution of actin subunits and accessory proteins is homogeneous and obeys mass-action kinetics. The resulting coupled ordinary differential equations (ODEs) are solved numerically, and the effects of varying parameters such as reservoir ATP/ADP disequilibrium, total filament concentration, fraction of capped plus ends, free actin concentration, and profilin concentration are investigated [180, 181]. One point of contention is whether transitions between hydrolysis states of polymerized subunits occurs in a random fashion, in which hydrolysis states of a subunit’s neighbors do not affect the hydrolysis rate of that subunit, or in a vectorial fashion, in which an ATP-bound subunit will only hydrolyze ATP if its neighbor towards the minus end is ADP-Pi bound, leading to a contiguous ATP-bound cap at the plus end. Recent models suggest that the truth is in the middle, such that coupling exists in ATP cleavage rates between neighboring polymerized subunits, but not such that the process is truly vectorial [92, 182]. Most mean-field models make the assumption of random ATP hydrolysis for simplicity.

An important disadvantage of such mean-field models aimed at resolving the roles of accessory proteins is that their level of detail inhibits analytical solutions to the time courses and, hence, obscures deeper insights into dynamical behaviors of these systems. While this approach has successfully allowed modelers to, for example, rule out certain mechanisms of profilin's action on critical concentrations [98], one might ask what is the simplest such model that reproduces time courses from more detailed models. This is the aim of the present chapter. The model reduction here is based on a 2009 model by Brooks and Carlsson (BC) [183], which presents a system of differential equations that admits only numerical solution but does not include extra detail by accounting for accessory proteins. It is useful for predicting the process of polymerization when a pool of subunits are added to an initial concentration of seed filaments, and is sufficiently simple to be incorporated into larger-scale cellular models without too much additional computational cost.

In this chapter, we report on two successive reduction schemes of the 11-dimensional BC model: a quasi-steady-state approximation that leverages fast dynamics of the filament tips, leading to a 5-dimensional system of ODEs, and a subsequent linearization approximation. The latter equations admits an analytical solution whose implications we investigate, revealing interesting features of actin polymerization process projected on the slow dynamical manifold. Our analytical model reduction approaches show excellent agreement with the results obtained from stochastic simulations of the full BC model and also when compared with diffusion mapping analyses of stochastic trajectories.

2.2 Methods

2.2.1 Brooks-Carlsson model

The BC model of actin polymerization is an 11-dimensional system of ODE's tracking the concentrations of non-tip actin subunits in different states as well as the concentrations of filament tip subunits in different states [183]. It is assumed that the number concentration of filaments N remains constant, implying an absence of filament nucleation, splitting, or joining. Additionally, the total concentration of actin subunits M is assumed to remain constant, such that actin subunits are not created or destroyed in any reaction. Since there are typically many actin subunits belonging to a given actin filament, we have $N \ll M$. All species are assumed to be well-mixed and in large enough quantities to be treated effectively via a mean-field description. In other words, the size of the stochastic fluctuations is negligible compared to the concentrations of the species. Unpolymerized (globular) actin subunits are referred to as G-actin, while polymerized (filamentous) actin subunits are referred to as F-actin. Actin filaments are helical, but they are more easily modeled as linear chains, which is a realistic approximation if one assumes that the reaction propensities of a given F-actin subunit are determined only by the state of the nucleotide bound by that subunit and not by the subunit's neighbors. Such a chain is displayed in Figure 2.1, along with some of the reactions allowing interconversion between subunit types. The variables representing the concentrations of these actin species are superscripted by the hydrolysis state of the bound nucleotide (for what follows we refer more simply to a subunit being in a certain hydrolysis state as opposed to the nucleotide attached to a subunit as being in that state). The hydrolysis states are ATP, ADP-Pi, and ADP, denoted T, Pi, and D, respectively. The tip subunits are denoted T and are further subscripted according to which tip they are on. Thus, for example, the concentration of tip subunits at the plus end bound to ADP-Pi is denoted T_+^{Pi}. Because inorganic phosphate rapidly dissociates from G-actin, G^{Pi} is taken to be 0 and is not tracked. With the 3 hydrolysis states of each of the 2 tips, the 3 states of the F-actin and the 2 states of G-actin, the number of tracked variables is 11.

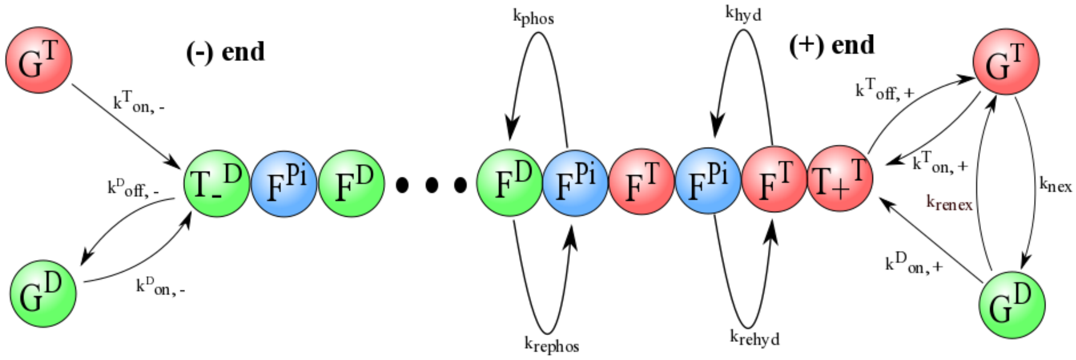

FIGURE 2.1: A linear actin filament, with subunits colored according to hydrolysis state. Random, as opposed to vectorial, hydrolysis is assumed here. Some of the reactions are slightly misleading as drawn: for example a hydrolysis reaction converting F^T to F^{Pi} would happen at a single location in the filament, i.e. the subunit would not change into its neighbor as shown here. Also, the polymerization of G^T onto the minus end would convert T_-^D into T_-^T, and a similar statement applies to G^D polymerizing to the plus end. The depolymerization of $T_\pm^{Pi}$ is not shown.

The different subunit types interconvert through chemical reactions. These reactions can be classified as polymerization/depolymerization reactions which change G actin to F actin and vice versa, or as reactions in which the hydrolysis state of the subunit changes. The evolution of the concentrations of a given tip subunit hydrolysis states in principle depends on the hydrolysis state of the subunit adjacent to the tip, which itself depends on the hydrolysis state of the next subunit in the filament, and so on. Every subunit in the filament then requires explicit tracking, causing the dimensionality of the model to be roughly equal to the degree of polymerization of a filament, which typically contains hundreds of subunits. The major accomplishment of the BC model is to truncate this set of recursion equations by assuming that the hydrolysis state of the subunit adjacent to the tip depends only on the hydrolysis state of the tip subunit. Using the results of stochastic simulations of a more complete model, they write empirical equations to capture these relationships, and in doing so they close off an 11-dimensional subset of the original hundreds of equations. They find close agreement between their truncated model and the full stochastic simulation over a wide and realistic range of parameters. The equations of motion for the 11 variables in the BC model can be written as a nonlinear system of ODE's:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \tag{2.1}$$

where $\mathbf{x}$ is a vector of the concentrations of the 11 species, and $\mathbf{f}$ is a nonlinear vector-valued function of $\mathbf{x}$. This system of equations must be solved numerically, but the results match well with simulations that accurately model experimental data [184]. The steady state vector $\mathbf{x}^{\mathrm{ss}}$ satisfying $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = \mathbf{0}$ is unique for given values of N and M and under the condition that all concentrations be real and non-negative, and it is attracting. We give more details of the BC model in Appendix B, where we list the 11 ODE's.

A separation of timescales exists between the dynamics of the non-tip subunit states and those of tip subunit states: the latter evolve much more rapidly than the former. Thus we partition the vector $\mathbf{x}$ into slow and fast variables: $\mathbf{x}_{\mathrm{s}} \equiv (G^T,\ G^D,\ F^T,\ F^{Pi},\ F^D)^{\intercal}$, $\mathbf{x}_{\mathrm{f}} \equiv (T_+^T,\ T_+^{Pi},\ T_+^D,\ T_-^T,\ T_-^{Pi},\ T_-^D)^{\intercal}$, where the subscripts "s" and "f" refer to "slow" and "fast". Terms of comparable magnitude appear in the equations of motion for both $\mathbf{x}_{\mathrm{f}}$ and $\mathbf{x}_{\mathrm{s}}$, but the elements in $\mathbf{x}_{\mathrm{f}}$ are typically much smaller than those in $\mathbf{x}_{\mathrm{s}}$ because $N \ll M$, and so $\mathbf{x}_{\mathrm{f}}$ reaches equilibrium more rapidly since the distance to equilibrium is not as large as it is for $\mathbf{x}_{\mathrm{s}}$. To see the separation timescales concretely, we refer the reader to the end of Section 2.3.1, where we show the spectrum of the Jacobian matrix of the BC model evaluated at steady state. With the new collective variables $\mathbf{x}_{\mathrm{s}}$ and $\mathbf{x}_{\mathrm{f}}$, Equation 2.1 can be usefully rewritten as follows:

$$\dot{\mathbf{x}}_{\mathrm{s}} = \mathbf{A}\mathbf{x}_{\mathrm{s}} + \mathbf{B}\mathbf{x}_{\mathrm{f}} \tag{2.2}$$

$$\dot{\mathbf{x}}_{\mathrm{f}} = \mathbf{h}(\mathbf{x}_{\mathrm{s}},\ \mathbf{x}_{\mathrm{f}}) \tag{2.3}$$

where $\mathbf{A}$ and $\mathbf{B}$ are matrices whose off-diagonal elements are combinations of kinetic rate constants and whose columns sum to zero due to conservation of actin, and $\mathbf{h}$ is a nonlinear vector-valued function containing terms that are up to cubic products of variables.

The BC model provides a computationally accessible model of the dynamics of actin polymerization, however we might ask for several other features in such a model. We ask that it: i) be low-dimensional, ii) capture the interesting and important timescales, and iii) be exactly solvable. To meet these goals, we make the decision to track only the vector $\mathbf{x}_\mathrm{s}$ explicitly. If we were to track the concentration of the tip subunit states, i.e. the elements in $\mathbf{x}_\mathrm{f}$, we would automatically increase the dimensionality of the model, and we will discuss reasons why we may assume that the tip subunits are evolving in such a way that keeping track of them explicitly is unnecessary. We make two approximations for how to treat $\mathbf{x}_\mathrm{f}$ in Equation 2.2, first utilizing the fact that a separation of timescale exists between the dynamics of $\mathbf{x}_\mathrm{f}$ and $\mathbf{x}_\mathrm{s}$, and then utilizing the fact that $|\mathbf{B}\mathbf{x}_\mathrm{f}| \ll |\mathbf{A}\mathbf{x}_\mathrm{s}|$ except when the system is near its steady state, at which point these terms have comparable magnitudes. This second approximation follows since $\mathbf{x}_\mathrm{f}$ contains terms up to $\mathcal{O}(N)$ and $\mathbf{x}_\mathrm{s}$ contains terms up to $\mathcal{O}(M)$.

It is also possible to demonstrate that a low-dimensional description of the slow dynamics is a valid approximation through the use of diffusion mapping on a stochastically generated data set based on the BC model. This analysis is described in the supporting information of Ref. 99.

2.2.2 Quasi-steady state approximation

The quasi-steady state approximation (QSSA) relies on a separation of timescales between fast and slow variables. This separation allows one to assume that the fast variables $\mathbf{x}_\mathrm{f}$ are always in equilibrium with respect to the slow variables $\mathbf{x}_\mathrm{s}$, and therefore that the values of the slow variables determine the values of the fast variables at any moment. We can imagine that $\mathbf{x}_\mathrm{f}$ is effectively being “dragged around” by the values of the elements in $\mathbf{x}_\mathrm{s}$. So, we can solve for functions $\mathbf{x}_\mathrm{f}^\mathrm{eq}(\mathbf{x}_\mathrm{s})$ relating the quasi-equilibrated fast variables in terms of the slow variables by imagining holding $\mathbf{x}_\mathrm{s}$ fixed and finding the equilibrium values of $\mathbf{x}_\mathrm{f}$. This amounts to the condition $\mathbf{h}(\mathbf{x}_\mathrm{f}^\mathrm{eq}\ ;\ \mathbf{x}_\mathrm{s}) = \mathbf{0}$. The functions $\mathbf{x}_\mathrm{f}^\mathrm{eq}(\mathbf{x}_\mathrm{s})$ are then substituted in the equations of motion for the slow variables giving the closed system of equations

$$\dot{\mathbf{x}}_\mathrm{s} = \mathbf{A}\mathbf{x}_\mathrm{s} + \mathbf{B}\mathbf{x}_\mathrm{f}^\mathrm{eq}(\mathbf{x}_\mathrm{s}). \tag{2.4}$$

This subsystem is lower-dimensional, though it is nonlinear since $\mathbf{x}_{\mathrm{f}}^{\mathrm{eq}}(\mathbf{x}_{\mathrm{s}})$ is nonlinear, and it describes the evolution of the system on the slow timescales.

In the BC model, the condition $\mathbf{h}(\mathbf{x}_{\mathrm{f}}^{\mathrm{eq}}\ ;\ \mathbf{x}_{\mathrm{s}}) = \mathbf{0}$ implies the following algebraic systems of equations (see Appendix B):

$$\begin{aligned} \frac{dT_{\pm}^{T}}{dt} &= 0 \\ \frac{dT_{\pm}^{Pi}}{dt} &= 0 \\ \frac{dT_{\pm}^{D}}{dt} &= 0. \end{aligned} \tag{2.5}$$

Only four of these six equations are linearly independent due to the conservation of number of plus and minus end filament tips, so we use the following supplementary equations to find a solution of the combined systems of equations:

$$T_{\pm}^{T} + T_{\pm}^{Pi} + T_{\pm}^{D} = N. \tag{2.6}$$

The system of Equations 2.5, 2.6 can be solved numerically resulting in tabulated functions of the forms $T_{\pm}^{T}(G^{T},\ G^{D})$, $T_{\pm}^{Pi}(G^{T},\ G^{D})$, and $T_{\pm}^{D}(G^{T},\ G^{D})$. These functions do not depend on F^{T}, F^{Pi}, and F^{D} because these variables do not enter into the the function $\mathbf{h}$. Subsequently, the nonlinear (slow) system described by Equation 2.4 is numerically integrated. Figure 2.2 displays a comparison of the QSSA approximation to the original 11-dimensional BC model.

This approximation succeeds in reducing the dimensionality of the model to 5. We note, however, that the dynamics of this model lie on a 4-dimensional submanifold of the full 5-dimensional manifold due to the fact that $\mathbf{A}$ and $\mathbf{B}$ are both singular, corresponding to conservation of actin. This model also captures the interesting timescales describing polymerization events and dynamics of the hydrolysis states of F and G actin, while it neglects the fast dynamics corresponding to such events taking place at the tips. However the model is still nonlinear and analytically unsolvable, so we propose an alternative model reduction scheme.

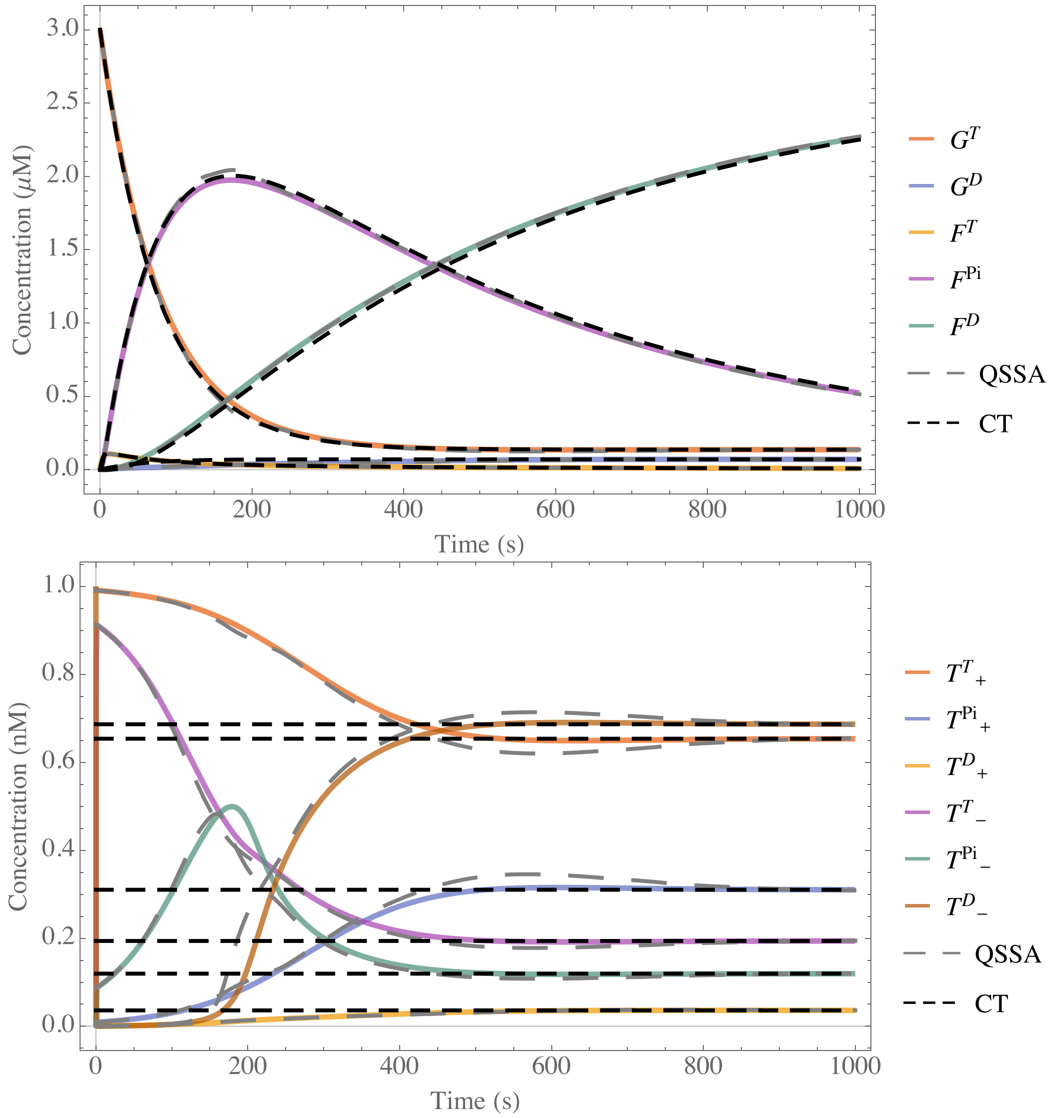

Figure 2.2: The numerical integration of the BC model is displayed as solid colors, the numerical integration of the QSSA model (Section 2.2) as a long dashed gray line, and the exact solution of the CT model (Section 2.3) as a short dashed black line. The time courses of the concentrations of the various species in $\mathbf{x}_\mathrm{s}$ (top) and $\mathbf{x}_\mathrm{f}$ (bottom) are shown. The CT and QSSA models have the same steady state behavior as the BC model, and the QSSA model approximately retains the dependence of the tip subunit states on the slow non-tip subunit states.

2.2.3 Constant tip approximation

In addition to having sufficiently fast dynamics as to be effectively described as in quasi-equilibrium with respect to $\mathbf{x}_\mathrm{s}$, the vector $\mathbf{x}_\mathrm{f}$ is also small in magnitude compared to $\mathbf{x}_\mathrm{s}$, and this fact can be utilized to drastically simplify the QSSA model. Although the tip dynamics are fast, one can profitably assume that the concentration of filament tip subunit states, that is, the elements in $\mathbf{x}_\mathrm{f}$, are constant in time. We refer to this assumption as the constant tip (CT) approximation. Certainly this assumption is not realistic since the tips have fast dynamics, but the effect of this error on the equations of motion of the other actin species is small in the regime where the number concentration of filaments is much smaller than the total amount of actin in the system, i.e. when $N \ll M$, or when the tip subunit states quickly attain their limiting values. This assumption allows the model to be reduced to a 5-dimensional linear system of equations that can be solved analytically. The procedure is to replace the dynamical variables T_j^i by constants $N\Gamma_j^i$, chosen such that the steady state of the CT model coincides with the steady state of the BC model. We define Γ_j^i as the fraction of the filament tips at the j (plus or minus) end that are in the i (T, Pi, D) hydrolysis state. These constants determine the rate of depolymerization reactions, and replacing the variables T_j^i with them causes the term $\mathbf{B}\mathbf{x}_\mathrm{f}$ to be a constant source and sink term in Equation 2.2, which as a result becomes linear.

The steady states of the BC and CT models will be the same if

$$\begin{aligned} N\Gamma_\pm^T &= \lim_{t\to\infty} T_\pm^T(t) \\ N\Gamma_\pm^{Pi} &= \lim_{t\to\infty} T_\pm^{Pi}(t) \\ N\Gamma_\pm^D &= \lim_{t\to\infty} T_\pm^D(t). \end{aligned} \tag{2.7}$$

In other words, the constant tip subunit state fractions in the CT model should be chosen as the steady state values of the tip subunit state fractions in the BC model. As a result, only the approach to steady state will be different between the two models. These limiting values can be found by numerically solving the algebraic system of equations

$$\begin{aligned}
&\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = 0\\
&T_{\pm}^{T} + T_{\pm}^{Pi} + T_{\pm}^{D} = N\\
&G^{T} + G^{D} + F^{T} + F^{Pi} + F^{D} = M.
\end{aligned} \tag{2.8}$$

where $\mathbf{x}^{\mathrm{ss}}$ is the 11-dimensional steady state vector of concentrations in the BC model, and taking the real non-negative solution. Equation 2.8 determines the values of Γ_j^i which will be unique for a given N and M.

We define

$$\mathbf{b} \equiv \mathbf{B}\mathbf{x}_{\mathrm{f}}^{\mathrm{ss}} \tag{2.9}$$

where $\mathbf{x}_{\mathrm{f}}^{\mathrm{ss}}$ contains the constants $N\Gamma_j^i$ instead of the variables T_j^i. The equation of motion for $\mathbf{x}_{\mathrm{s}}$ is

$$\dot{\mathbf{x}}_{\mathrm{s}} = \mathbf{A}\mathbf{x}_{\mathrm{s}} + \mathbf{b} \tag{2.10}$$

where

$$\mathbf{A} = \begin{pmatrix} -a-b & c & 0 & 0 & 0\\ b & -c-d & 0 & 0 & 0\\ a & 0 & -e & f & 0\\ 0 & 0 & e & -f-g & h\\ 0 & d & 0 & g & -h \end{pmatrix} \tag{2.11}$$

$$\mathbf{b} = \begin{pmatrix} i\\ j+k\\ -i\\ -k\\ -j \end{pmatrix} \tag{2.12}$$

with

$$\begin{aligned}
a &= N(k^T_{\text{on, }+} + k^T_{\text{off, }+}) & g &= k_{\text{phos}} \\
b &= k_{\text{nex}} & h &= k_{\text{rephos}} \\
c &= k_{\text{renex}} & i &= N(\Gamma^T_+ k^T_{\text{off, }+} + \Gamma^T_- k^T_{\text{off, -}}) \\
d &= N(k^D_{\text{on, }+} + k^D_{\text{off, }+}) & j &= N(\Gamma^D_+ k^D_{\text{off, }+} + \Gamma^D_- k^D_{\text{off, -}}) \\
e &= k_{\text{hyd}} & k &= N(\Gamma^{Pi}_+ k^{Pi}_{\text{off, }+} + \Gamma^{Pi}_- k^{Pi}_{\text{off, -}}) \\
f &= k_{\text{rehyd}} & &
\end{aligned} \quad . \tag{2.13}$$

Equation 2.10 is a linear system of differential equations which can be solved exactly. One point of difficulty in solving this system is that the columns of $\mathbf{A}$ sum to zero due to conservation of actin, causing $\mathbf{A}$ to be singular. In chemical reaction network theory, one often has such systems with linear conservation laws. If the system is linear, with only first order or pseudo-first order reactions, then such a singular system can be solved cleanly by a method using the Drazin inverse $\mathbf{A}^{\mathcal{D}}$ of the matrix $\mathbf{A}$. We believe that this method has certain advantages over other approaches to solving singular systems of differential equations. If $\mathbf{A}$ has Jordan decomposition

$$\mathbf{A} = \mathbf{V} \begin{pmatrix} J_1 & 0 \\ 0 & J_0 \end{pmatrix} \mathbf{V}^{-1} \tag{2.14}$$

where J_1 and J_0 correspond to the non-zero and zero eigenvalues respectively, then

$$\mathbf{A}^{\mathcal{D}} = \mathbf{V} \begin{pmatrix} J_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{V}^{-1}. \tag{2.15}$$

In the supporting information of Ref. 99 the details of solving Equation 2.10 as well as the advantages of this method are discussed [185]. The solution is

$$\mathbf{x}_{\text{s}} = \Big(-\mathbf{A}^{\mathcal{D}} + e^{\mathbf{A}t}\mathbf{G} \Big)\mathbf{b}. \tag{2.16}$$

The matrix

$$\mathbf{G} = \frac{1}{|\mathbf{b}|^2}\mathbf{x}_{\text{s}}(0)\mathbf{b}^{\intercal} + \mathbf{A}^{\mathcal{D}} \tag{2.17}$$

encodes information about the initial conditions $\mathbf{x}_{\mathrm{s}}(0)$.

2.3 Results

2.3.1 Eigenvalues of A

Perhaps the most important benefit of an exactly solvable model is the ability to formally determine the eigenvalues governing the linear dynamics. These eigenvalues characterize the relaxation times of the components of a perturbation from equilibrium, in the basis of the eigenvectors. The expressions for the eigenvalues thus shed light on the timescales that describe the different chemical processes. In our modeling we have included the reverses of kinetically dominant forward reactions, and we have set the rate constants of these reactions to be equal to something on the order of the corresponding forward reaction rate constants multiplied by a small parameter ϵ. Thus $b = \epsilon b^*$, where $b^* \sim c$, $f = \epsilon f^*$, where $f^* \sim e$, and $h = \epsilon h^*$, where $h^* \sim g$. By writing reverse rate constants this way, we can Taylor expand the expressions for the eigenvalues around the point $\epsilon = 0$ to simplify the result. Doing so to first order in ϵ, we have

$$\lambda_1 = 0 \tag{2.18}$$

$$\begin{aligned}\lambda_2 &= \frac{1}{2}\Big(- a - \epsilon b^* - c - d + \sqrt{(a - d + \epsilon b^* - c)^2 + 4\epsilon b^* c}\Big) \\ &\approx -c - d + \frac{b^* c}{a - c - d}\epsilon \end{aligned}\tag{2.19}$$

$$\begin{aligned}\lambda_3 &= \frac{1}{2}\Big(- a - \epsilon b^* - c - d - \sqrt{(a - d + \epsilon b^* - c)^2 + 4\epsilon b^* c}\Big) \\ &\approx -a - \frac{b^*(a - d)}{a - c - d}\epsilon \end{aligned}\tag{2.20}$$

$$\begin{aligned}\lambda_4 &= \frac{1}{2}\Big(- e - \epsilon f^* - g - \epsilon h^* + \sqrt{(e - g + \epsilon f^* - \epsilon h^*)^2 + 4\epsilon f^* g}\Big) \\ &\approx -g - \left(h^* - \frac{g f^*}{e - g}\right)\epsilon \end{aligned}\tag{2.21}$$

$$\begin{aligned}\lambda_5 &= \frac{1}{2}\Big(- e - \epsilon f^* - g - \epsilon h^* - \sqrt{(e - g + \epsilon f^* - \epsilon h^*)^2 + 4\epsilon f^* g}\Big) \\ &\approx -e - \frac{e f^*}{e - g}\epsilon. \end{aligned}\tag{2.22}$$

The fact that the nonzero eigenvalues are negative implies the stability of the steady state. These eigenvalues have no dependence on M, the total concentration of actin subunits, but they do depend on N, the number concentration of filaments, since this term appears in the expressions for a and d. For the parameterization used here and with $N = 1\ nM$, the eigenvalues can be ordered by magnitude as follows:

$$|\lambda_5| > |\lambda_3| > |\lambda_2| > |\lambda_4| > |\lambda_1|. \tag{2.23}$$

Equations 2.18-2.22 are in terms of reaction rate constants and can be interpreted as representing certain collective subprocesses in the chemical system corresponding to combinations of those reactions. The ordering indicates the comparative rates of those subprocesses. We interpret the zeroth order terms of the eigenvalues as follows:

- $\lambda_1 = 0$ because actin subunits are conserved in this system, causing $\mathbf{A}$ to be singular. Equivalently, one can say that the dynamics unfold on a 4-dimensional submanifold of the 5-dimensional manifold, and that this submanifold is determined by M.

- λ_2 represents the combination of the forward nucleotide exchange reaction ($G^D \rightarrow G^T$) and the polymerization of G^D. Both of these reactions convert G^D into another species, so this eigenvalue represents the subprocess of G^D depletion.

- λ_3 represents the polymerization of G^T.

- λ_4 represents the release of phosphate by F^{Pi} to form F^D.

- λ_5 represents the hydrolysis of ATP converting F^T to F^{Pi}.

Whether the magnitudes of these eigenvalues are increased or decreased due to the presence of reversible reactions (i.e. when $\epsilon > 0$) depends on the parameterization, since the sign of the first order terms depend on the comparative sizes of certain parameters.

In the full BC 11-dimensional model, one can numerically evaluate the Jacobian matrix of $\mathbf{f}(\mathbf{x})$ at steady state:

$$\mathbf{J}^* \equiv \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^{\mathrm{ss}}}. \tag{2.24}$$

We find that, for the same parameterization, the smallest four non-zero eigenvalues of $\mathbf{J}^*$ are exactly equal to the non-zero eigenvalues of $\mathbf{A}$. This implies that the CT model has captured the slowest dynamics of the BC model by ignoring the processes involving the tip subunits. The main benefit of the CT approximation is that these slow linear dynamics can now be easily analyzed. These dynamics provide information about the polymerization process, the nucleotide composition of the filaments, and nucleotide exchange of globular actin. For most purposes, these aspects are of primary interest, and the processes at the tips are of lesser importance.

To quantitatively judge the separation of timescales of the BC model, we list the nonzero numerically evaluated eigenvalues of $\mathbf{J}^*$ when $N = 1$ nM and $\epsilon = 0$, in decreasing magnitude: $(-5.49,\ -1.41,\ -0.833,\ -0.393,\ -0.300,\ -0.0130,\ -0.0129,\ -0.002)$. For comparison, we do the same with the eigenvalues of $\mathbf{A}$: $(-0.300,\ -0.0130,\ -0.0129,\ -0.002)$. While there is no large spectral gap, there is certainly a wide range of timescales, and it would suffice to even keep the smallest 3 non-zero eigenvalues. We later show how, by combining certain species, the dimensionality can be reduced to 3.

2.3.2 Steady state concentrations

One might expect that if we increase the amount of actin subunits in the system by a different choice of initial conditions, the concentrations of the different species at steady state would change. An interesting feature of this system is that this is true only for some species, and which species it is true for depends on whether or not we have included reversible reactions (if $\epsilon > 0$). Additionally, this can be shown to be true in both the CT model and the BC model. We demonstrate it first in the CT model.

We find the steady state vector of concentrations $\mathbf{x}_{\mathrm{s}}^{\mathrm{ss}}$ by taking the limit of Equation 2.16 as $t \to \infty$:

$$\mathbf{x}_{\mathrm{s}}^{\mathrm{ss}} \equiv \lim_{t\to\infty} \mathbf{x}_{\mathrm{s}}(t) = -\mathbf{A}^{\mathcal{D}}\mathbf{b} + \mathbf{CGb} \tag{2.25}$$

where

$$\mathbf{C} \equiv \lim_{t\to\infty} e^{\mathbf{A}t}. \tag{2.26}$$

We eigendecompose $\mathbf{A}$ as $\mathbf{UDV}^{\dagger}$ and use it in the expression for $\mathbf{C}$:

$$\mathbf{C} = \mathbf{U} \lim_{t\to\infty} e^{\mathbf{D}t}\mathbf{V}^{\dagger}. \tag{2.27}$$

With the exception of λ_1, which is zero, the eigenvalues of $\mathbf{A}$ are negative, so we have

$$\begin{aligned} \mathbf{C} &= \mathbf{U}\ \mathrm{diag}(1,0,0,0,0)\ \mathbf{V}^{\dagger} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{fh}{eg} & 0 & 0 & 0 & 0 \\ \frac{h}{g} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{V}^{\dagger} \end{aligned} \tag{2.28}$$

where $(0, 0, \frac{fh}{eg}, \frac{h}{g}, 1)^{\intercal}$ is the eigenvector corresponding to λ_1. Thus the top two rows of $\mathbf{C}$ are zero, and the top four rows would be zero if we exclude reversible reactions. This is also true of the term $\mathbf{CGb}$. Now, the matrix $\mathbf{G}$ is the only place where the initial conditions appear in Equation 2.25. So if the top two rows $\mathbf{CGb}$ are zero, then the first two elements of $\mathbf{x}_{\mathrm{s}}^{\mathrm{ss}}$ cannot have any dependence on initial conditions. In other words, G^D and G^T always reach the same concentrations at steady state no matter what the initial concentrations of any of the species are. If we have no reversible reactions, the same is also true for the species F^T and F^{Pi}.

Consider the following thought experiment, assuming for simplicity $\epsilon = 0$. We start with some initial amount M of actin in any form and let the system come to steady state. We then add an amount ΔM more actin to the system, in any form, and wait until the system has reached steady state again. We would find that the only difference between the two steady states is that the concentration of F^D had increased by ΔM. If $\epsilon > 0$, then we would find that the steady state values of F^T, and F^{Pi}, F^D had all increased, and the sum of these changes would be ΔM.

This lack of dependence of the steady state concentrations of some species on M is not an artifact of the CT approximation; it is also the case in the BC model. It can not be shown to be true by taking the limit $t \to \infty$ as is the case here, but it can instead be shown by observing that a subset of the system of algebraic equations $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = 0$ is closed, and that a solution for the subset could be obtained without specifying M. This implies that the steady state concentrations of the species represented by that solution have no dependence on M. We give the details of this argument in Appendix B.

2.4 Discussion

We have argued that the dynamics of the polymerization of actin subunits into filaments can effectively be subdivided as follows: fast nonlinear dynamics govern the states of the filament tip subunits, and slow linear dynamics govern the change in polymerization and hydrolysis states of non-tip subunits. One cannot completely separate the tip subunit dynamics from the non-tip subunit dynamics because they are coupled; the tip subunit states depend on the concentrations of G^T and G^D through polymerization reactions, and the non-tip subunit states depend on the tip subunit states via depolymerization reactions. However, because of the typical size of N compared to M, the non-tip subunit states depend only comparatively weakly on the tip subunit states during most of a typical trajectory. We have shown two ways to approximate this coupling to achieve a significant reduction in dimensionality of the system. First, in the QSSA model, it is assumed that, on the slow timescale, the number of tips in a certain hydrolysis state depends only on G^T and G^D. This allows one to write a closed system of equations of motion of the non-tip subunits, describing evolution of the entire system on the slow 5-dimensional submanifold of the full 11-dimensional space. This model is physically realistic and quite accurate because it preserves the dependence of the tip subunit states on the concentration of the non-tip subunits, however the resulting equations of motion are analytically intractable.

In the CT model, we make the seemingly unrealistic assumption that the tip subunits have no dependence on the non-tip subunits and in fact do not evolve at all, but remain fixed for all times at their steady state values. In this sense we turn the tip subunit concentrations from variables into constants, and the equations of motion for the non-tip subunits become 5-dimensional and linear with the depolymerization terms involving the tip subunits entering as a non-homogenous term $\mathbf{b}$. By choosing the fixed values of the tip subunits as the resting values, we ensure that the steady states of the two models will be the same. The CT assumption is valid because of the weak dependence of the non-tip subunit states on the tip subunit states. In other words, $\mathbf{b}$ is comparatively small, and the discrepancy between trajectories of the CT and BC model due to $\mathbf{b}$ not containing realistic values for all times is not pronounced. In exchange for the cost of this error, there is an important benefit, which is the ability to write symbolic expressions for the eigenvalues of the matrix $\mathbf{A}$. We note that these eigenvalues agree with the numerically calculated smallest nonzero eigenvalues of $\mathbf{J}^*$ of the full 11-dimensional model, indicating that indeed the linear non-tip subunit dynamics are the slowest of all of the processes in the BC model. In addition, qualitative results about the nature of the dependence of the steady state concentrations on the initial conditions agree for the full and reduced model. Dimensionality reduction is also possible by diffusion map analysis of a simulated trajectory of the BC model (supporting information of Ref. 99)). The results of this analysis indicate that fewer than 6 or 7 dimensions suffice to faithfully reproduce the polymerization dynamics.

Eigenvalue analysis of $\mathbf{A}$ allows one to understand the timescales that govern the linear non-tip subunit dynamics as well as how these timescales depend on the parameters. These timescales approximately represent the following processes: depletion of G^D via polymerization and conversion to G^T, polymerization of G^T, hydrolysis of ATP converting F^T to F^{Pi}, and phosphate release converting F^{Pi} to F^D. As might be expected, the timescales involving polymerization depend on N, and their magnitude compared to that of the other timescales may change significantly. We have treated the presence of the reverses of some nearly irreversible reactions essentially as perturbations by regarding the rate constants of backward reactions as equal to the rate constant of the corresponding forward reaction multiplied by a small parameter ϵ. As shown above, the inclusion of these reactions introduce small corrections to the eigenvalues. These timescales allow one to understand a typical trajectory of the system. Such a trajectory can be visualized in three dimensions by combining multiple species into a single species and choosing to not to visualize a variable whose value is determined by the other three due to conservation of actin. In Figure 2.3, we combine F^T and F^{Pi} together, since they have similar structural properties in the filament, and we do not visualize F^D. Thus we further reduce dimensionality to 4 (of which only 3 dimensions are independent) by neglecting the timescale corresponding to ATP hydrolysis, i.e. λ_5. In the trajectory depicted, G^T and G^D are quickly made small via polymerization and nucleotide exchange reactions. The polymerization of G^T causes F^{T+Pi} to increase, and when G^T has become small, F^{T+Pi} converts to F^D via the slow process of phosphate release, and at the end, nearly all of the actin is in the form F^D.

Conclusion

The main results of this chapter are the elucidation of the degree to which not explicitly accounting for tip subunit state dynamics during actin polymerization is a passable assumption, and the resulting insight into the hierarchy of processes involved in the slow linear dynamics of the non-tip subunit states. The CT model is overly simple but useful to understand basic features of the polymerization process. In other more detailed models, actin related proteins are incorporated either by introducing new variables representing the proteins and the protein-subunit complexes, or by including new parameters that multiply certain terms in the equations of motion [98, 180, 181]. These adaptations could be straightforwardly included in the modeling done here. Additionally, the effect

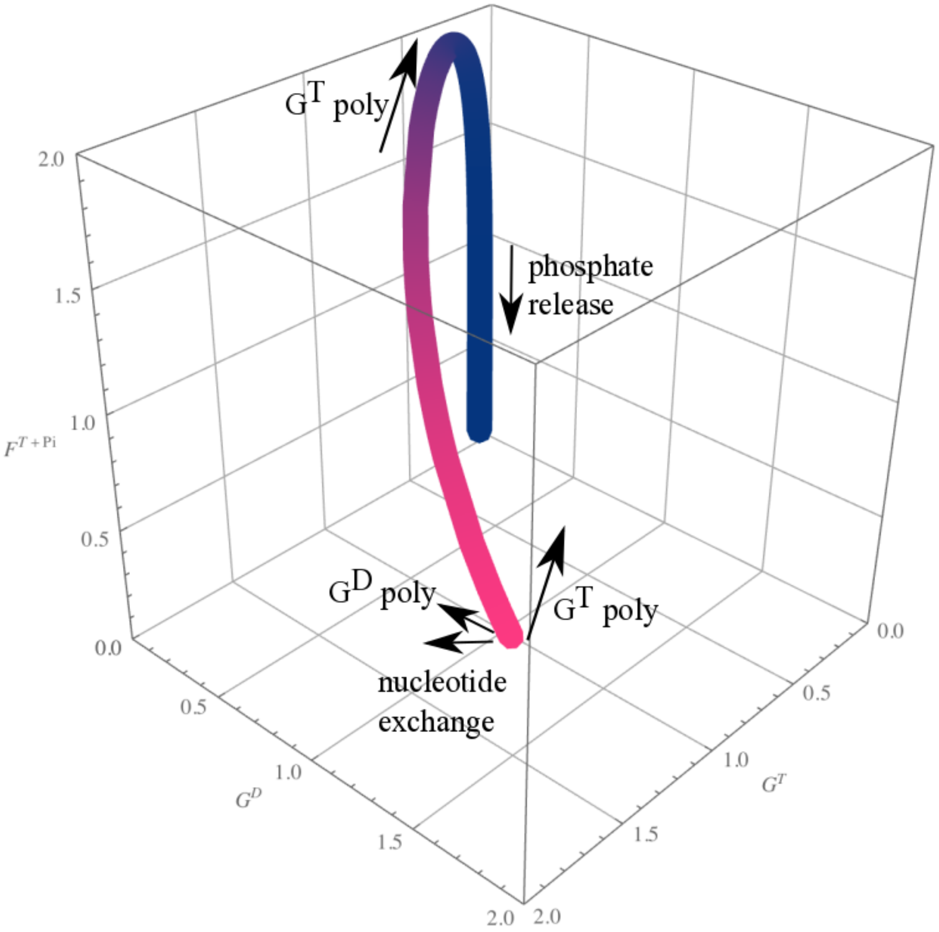

FIGURE 2.3: Visualization of a 3,000 s trajectory of the CT model beginning from 1 μM of G^T and 1 μM of G^D, with $N = 1$ nM and with F^D not visualized. The curve is colored according to time, with pink representing early times. All units are μM. Vectors are drawn, not to scale, and labeled to represent the direction that certain reactions pull the trajectory and at which point in the trajectory those pulls become dominant.

of different concentrations of solvated ATP, ADP-Pi, and ADP could be investigated by changing the values of the pseudo-first order reaction rate constants, or even by regarding those reactions as second order and tracking the concentrations of the nucleotide species as separate variables. These modifications run counter to the goal here of model reduction, however. Different choices in modeling are of course suited to different purposes, and the choices made here address a desire to have a simple mental picture of an otherwise obscured and complex process.

Chapter 3

Gibbs free energy change of a discrete chemical reaction event

This chapter and Appendix C are adapted from: Floyd, C., Papoian, G. A., & Jarzynski, C. (2020). Gibbs free energy change of a discrete chemical reaction event. The Journal of chemical physics, 152(8), 084116.

3.1 Introduction

In recent years the coarse-grained computational modeling of intracellular environments has enjoyed significant advances. An important paradigm shared by many such models is to treat the evolution of reacting chemical species' copy numbers and spatial distributions by simulating a reaction-diffusion master equation (RDME) [186]. In this approach, the system volume is divided into compartments, each with local values of the copy numbers and chemical potentials of the different chemical species (Figure 3.1). The RDME is a differential equation describing the evolution of the probability $P(\mathbf{N}, t)$ of observing the vector of copy numbers $\mathbf{N} = \{N_{i,A}\}_{i \in L, A \in \Omega}$ of chemical species i in compartment A at time t, where L is the set of solute species and Ω is the set of compartments in the system. The RDME can be written schematically as

$$\frac{dP\left(\mathbf{N}, t\right)}{dt} = \left(\hat{M} + \hat{D}\right) P\left(\mathbf{N}, t\right) \tag{3.1}$$

where $\hat{M}$ and $\hat{D}$ represent operators describing chemical reactions and inter-compartment diffusion, respectively [187]. An in-depth description of the RDME approach can be found in Refs. 188, 189, where the forms of the operators are discussed. Rather than directly solving Equation 3.1, one often simulates trajectories of the vector $\mathbf{N}$ obeying the stochastic dynamics encoded in the RDME, using a variant of the Gillespie algorithm [190]. The sizes of the compartments are commonly determined by the Kuramoto lengths, the mean free diffusional path for a species before it participates in a chemical reaction. Within each compartment, the spatial distributions of the reacting species are assumed to be homogeneous, allowing the use of mass-action kinetics with the compartment's local values of the species' concentrations to describe the stochastic chemical reaction propensities. Molecules can additionally jump between adjacent compartments in "diffusion events" (whose propensities also depend on the compartments' local concentrations of species) to give rise to concentration gradients on the scale of the compartment length. This modeling approach is appropriate when the Kuramoto length is small compared to the system size (i.e. the assumption of homogeneity over the system volume fails), yet large compared to the intermolecular distance scale [186]. Examples of simulation platforms employing such an approach include Virtual Cell [191, 192], lattice microbes [193, 194], and MEDYAN [42, 195].

One important aspect of simulating non-equilibrium biological systems is the computation of thermodynamic forces that drive the observed flux on the network of chemical reactions [26, 196, 197]. Determining these forces can allow for the quantification of entropy production in chemically reactive systems [198, 199]. In several research groups, measuring entropy production in biological active matter has been a recent goal [34, 152]. For instance, in Chapter 4 we quantify the entropy production rates of self-organizing non-equilibrium actomyosin networks in MEDYAN using the expressions derived here as a first step [195]. The ability to measure dissipation in active matter systems will allow to test the applicability of different physical organizing principles relating the production of entropy to the likelihood of observing certain trajectories [200, 201]. For isothermal, isobaric, chemically reactive solutions, which include most mesoscopic biological systems,

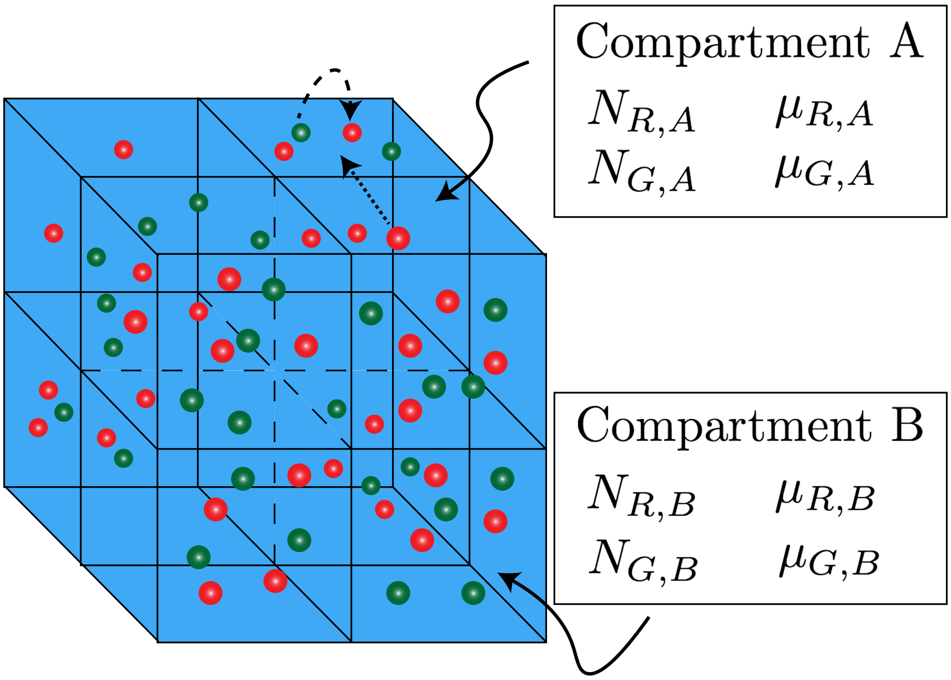

FIGURE 3.1: An example of a cubic compartment grid used in the simulation of a RDME. Each compartment, labeled with letters $A, B, ...$, has local values of the quantities N_R, N_G, referring to the copy numbers of the red and green molecules in the compartment, and of μ_R, μ_G, referring to the chemical potentials of those molecules. Molecules can react with each other within compartments (long dashed arrow), as well as hop between adjacent compartments, representing diffusion (short dashed arrow).

measuring the total entropy production amounts to determining the change in Gibbs free energy that accompanies chemical reactions and diffusion down concentration gradients [23, 25, 202]. A ubiquitous textbook expression for the change in Gibbs free energy G accompanying a chemical reaction is

$$\Delta G = k_B T \log K_{eq} + k_B T \log Q \tag{3.2}$$

where K_{eq} is the equilibrium constant, Q is the reaction quotient (we give definitions of these quantities below), k_B is Boltzmann's constant, and T is the temperature [1, 44]. At equilibrium, $Q = K_{eq}^{-1}$ and as a result $\Delta G = 0$. However, in this chapter, we argue that Equation 3.2 is a biased approximation to the exact value of ΔG accompanying a chemical reaction which holds when the copy numbers of the reacting molecules are large, such as on the order of Avogadro's number. When the system is small, such as when copy numbers are on the order of 100 as is often the case in RDME simulations of intracellular environments, then thermodynamic expressions such as Equation 3.2 require corrections [203, 204].

As a simple motivating example, consider a mixture of an even number of two chemical species, red and green, which inter-convert at equal rates. At equilibrium, the copy numbers of these molecules will be equal, and $Q = K_{eq}^{-1} = 1$. Now if a reaction were to occur at equilibrium to produce an additional red molecule and one fewer green molecule, then we would expect that the Gibbs free energy of the system had increased, since we have left equilibrium where the free energy attains its minimum. However the prediction of Equation 3.2 is that $\Delta G = 0$ for this reaction. The assumption whose violation leads to Equation 3.2 being incorrect is that the copy number of chemical species is a continuous quantity. When these variables are considered as discrete, then a different expression for ΔG must be used to give correct behavior.

Similarly for diffusion between adjacent compartments, a common expression for the change in Gibbs free energy accompanying the jump of a molecule i from compartment

A with copy number $N_{i,A}$, to compartment B where its copy number is $N_{i,B}$, is

$$\Delta G = k_B T \log \frac{N_{i,B}}{N_{i,A}}. \tag{3.3}$$

Imagine however we have a situation where $N_{i,A} = N_{i,B}$, and a molecule jumps from compartment A to B. The Gibbs free energy should have increased since we have departed from the highest entropy distribution of the molecules over the two compartments, however Equation 3.3 will predict that $\Delta G = 0$.

In this chapter we derive exact expressions for the change in Gibbs free energy accompanying chemical reactions within compartments and diffusion events between compartments, and we further show how these expressions relate to the familiar textbook formulas Equations 3.2 and 3.3 through a series of approximations. We also discuss the assumptions involved in defining the Gibbs free energy of a grid of homogeneously mixed compartments which can exchange energy and particles, such as that used in a simulation of a RDME. Finally we present numerical simulations using MEDYAN to demonstrate the need to use these more exact expressions for ΔG in order to obtain sensible results when copy numbers are treated as discrete variables. Only these more exact expressions will give correct, unbiased behavior when measuring entropy production in mesoscopic *in silico* studies of biological non-equilibrium systems that rely on the RDME formalism.

3.2 Methods

3.2.1 ΔG of reactions

Here we make successive approximations to the formula for ΔG accompanying a chemical reaction, and our notation reflects the level of approximation in which certain quantities are being used: when appropriate, we subscript quantities with a parenthesized number, i.e. $\Delta G_{(0)}$, where increasing numbers represent more approximate versions. The symbol $\widetilde{}$ will indicate that the quantities of chemical species are being represented by mole fractions χ_i, rather than concentrations C_i. In this section we treat the case that our system comprises a single closed compartment of a homogeneous dilute solution in which a chemical reaction has occurred and derive an expression for the change in the Gibbs free energy. In this system the number of solvent molecules is fixed and the solute molecules participate in chemical reactions, causing their copy numbers to change. In the next sections we consider a system with multiple weakly interacting compartments, each of which represents a homogeneous solution with local copy numbers of solvent and solutes and between which both solvent and solutes can diffuse. The nearly exact result Equation 3.21 obtained in this section will also apply to those systems, as we argue below.

Before restricting to the case of a single closed compartment, we establish notation for properties of the chemical species in a compartment grid. The chemical potential for species i in compartment A can be expressed either as depending upon the mole fraction, $\chi_{i,A}$, or upon the concentration, $C_{i,A}$, of that species in the compartment:

$$\mu_{i,A} = \widetilde{\mu_i^0}(T,p) + k_B T \log \chi_{i,A} = \mu_i^0(T,p) + k_B T \log C_{i,A} \tag{3.4}$$

where k_B is Boltzmann's constant, $\widetilde{\mu_i^0}(T,p)$ is the standard state chemical potential at temperature T and pressure p when working with the dimensionless $\chi_{i,A}$, and $\mu_i^0(T,p)$ is the same when working with $C_{i,A}$. $C_{i,A}$ and $\chi_{i,A}$ both play the role of the "composition variable" leading to these different, yet equivalent expressions for the chemical potential [205]. We make the distinction between dependence upon copy number and dependence on concentration in order to establish parameters that can be used in simulation, which commonly works with copy numbers, based on those given in the literature, which typically use units of concentration. Here we make the assumption of an ideal-dilute solution and neglect the coefficient of activity of the solute species [206]. The mole fraction can be written

$$\chi_{i,A} = \frac{N_{i,A}}{N_A} = \frac{N_{i,A}}{\sum_{j \in M} N_{j,A}} \tag{3.5}$$

where $N_{i,A}$ is the copy number of species i in compartment A, N_A is the total copy number of molecules in compartment A, and M is the set of all species including the solvent in the system. Similarly, the concentration can be written

$$C_{i,A} = \frac{N_{i,A}}{\Theta_A} \tag{3.6}$$

where

$$\Theta_A = N_{\mathrm{Av}} V_A \tag{3.7}$$

is a conversion factor, N_{Av} is Avogadro's number, and V_A is the compartment volume (we assume constant pressure and that the fluctuations in volume are negligible for these liquid systems, allowing the use of Gibbs free energy). Using Equations 3.4, 3.5, and 3.6, we can write[1]

$$\widetilde{\mu_i^0}(T,p) = \mu_i^0(T,p) + k_B T \log \frac{N_A}{\Theta_A}. \tag{3.8}$$

Using standard arguments concerning thermodynamic extensivity, it is possible to establish that the Gibbs free energy of the solution in compartment A can be written as a weighted sum over the chemical potentials of the species:

$$G_A(\mathbf{N}_A) = \sum_{i \in M} N_{i,A} \mu_{i,A}(\chi_{i,A}) \tag{3.9}$$

where $\mathbf{N}_A = \{N_{i,A}\}_{i \in M}$ is the vector of species copy numbers $N_{i,A}$, and where we have explicitly indicated the dependency of $\mu_{i,A}$ upon mole fraction $\chi_{i,A}$ via Equation 3.5 [202, 207]. One may be concerned that Equation 3.9 fails to apply when the copy numbers of solutes are small. As explained below, we make the assumption that boundary effects are still negligible, which allows us to treat G as a first order homogenous function of the number of copies of the system [204]. This is the necessary property to establish Equation 3.9, so the small copy numbers of solutes does not render this approach invalid. We rely on Equations 3.4 and 3.9 to derive changes in Gibbs free energy accompanying chemical reactions and inter-compartment diffusion.

---

[1] The argument of the logarithm, $\frac{N_A}{\Theta_A}$, has dimensions of concentration, however it is understood that such quantities are taken with reference to some standard concentration, typically 1 M.

Consider a reaction of the general form

$$\nu_1 X_1 + \nu_2 X_2 + \ldots \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \upsilon_1 Y_1 + \upsilon_2 Y_2 + \ldots \tag{3.10}$$

where X_i represent reactants, Y_j represent products, ν_i and υ_j are stoichiometric coefficients, and the rate of reaction is k_+ to the right and k_- to the left. We have dropped the subscript A indicating the compartment in which the reaction takes place, and now restrict to the case that our system is a single compartment. When this reaction has occurred once to the right, the copy numbers of reactants have changed $N_i \to N_i - \nu_i$, and those of the products have changed $N_j \to N_j + \upsilon_j$. We calculate the change in Gibbs free energy accompanying this reaction by considering it as resulting from these finite, discrete changes in copy numbers [204], not from infinitesimal changes. Using Equations 3.4, 3.5, and 3.9, the Gibbs free energy before the reaction has occurred can be written as

$$\begin{aligned} G^{\text{initial}} = &\sum_{i \in R} N_i \left(\widetilde{\mu_i^0} + k_B T \log \frac{N_i}{N} \right) + \sum_{j \in P} N_j \left(\widetilde{\mu_j^0} + k_B T \log \frac{N_j}{N} \right) \\ &+ N_s \left(\widetilde{\mu_s^{0*}} + k_B T \log \frac{N_s}{N} \right) \end{aligned} \tag{3.11}$$

where R is the set of reactants, P is the set of products, the subscript s refers to the solvent, and where we have dropped the dependence of the standard state chemical potential on T and p. $\widetilde{\mu_i^0}$ and $\widetilde{\mu_j^0}$ describe the chemical potential at a reference concentration of the solute in the solvent (also referred to as the solute standard state), whereas $\widetilde{\mu_s^{0*}}$ describes the chemical potential at a reference state of pure solvent (also referred to as the solvent standard state) [206]. We assume here for simplicity and without loss of generality that there are no solute species which have not participated in the reaction (i.e. spectator solute species). These species would also contribute terms to Equation 3.11 but, when we subtract the initial from the final Gibbs free energy, their inclusion would not give a different result. The final Gibbs free energy is

$$G^{\text{final}} = \sum_{i \in R} (N_i - \nu_i) \left(\widetilde{\mu_i^0} + k_B T \log \frac{N_i - \nu_i}{N + \sigma} \right) + \sum_{j \in P} (N_j + \upsilon_j) \left(\widetilde{\mu_j^0} + k_B T \log \frac{N_j + \upsilon_j}{N + \sigma} \right) \\ + N_s \left(\widetilde{\mu_s^{0*}} + k_B T \log \frac{N_s}{N + \sigma} \right) \tag{3.12}$$

where

$$\sigma = \sum_{j \in P} \upsilon_j - \sum_{i \in R} \nu_i \tag{3.13}$$

is what we refer to as the "stoichiometric difference", or the amount by which the total species copy number N has changed. As described in Refs. 205, 208, when $\sigma \neq 0$ it is important to account for the solvent species in the calculation of free energy differences. This is because the free energy of the solvent, which is the most numerous species in the reaction volume, will not be the same after the reaction has taken place since its mole fraction will change as N changes to $N + \sigma$. Neglecting the solvent species when $\sigma \neq 0$ leads to expressions for ΔG that are off by an amount $\sigma k_B T$ [205]. Whereas the authors of Refs. 205, 208 describe the appearance of this erroneous term while formulating the Gibbs free energy as a function of a continuous degree of advancement of reaction $d\xi = -dN_i/\nu_i = dN_j/\upsilon_j$, here we treat the extent of reaction as a discrete quantity. In the limit that $\nu_i/N_i \to 0$, $\upsilon_j/N_j \to 0$ for each reactant and product, the discrete case passes into the continuum case, however under the assumption of small copy numbers we do not satisfy this limit. In Appendix C we discuss further differences between the continuum treatment and the discrete treatment, as well as the relation to the Gibbs-Duhem Equation.

After some algebra (using the fact that $\sum_{i \in R} N_i + \sum_{j \in P} N_j + N_s = N$), we can write the change in Gibbs free energy as

$$\Delta G_{(0)} = G^{\text{final}} - G^{\text{initial}} \\ = \widetilde{\Delta G^0} + k_B T \log \prod_{i \in R} \frac{(N_i - \nu_i)^{N_i - \nu_i}}{N_i^{N_i}} \prod_{j \in P} \frac{(N_j + \upsilon_j)^{N_j + \upsilon_j}}{N_j^{N_j}} \frac{N^N}{(N + \sigma)^{N + \sigma}} \tag{3.14}$$

where

$$\widetilde{\Delta G^0} = \sum_{j\in P} \upsilon_j \widetilde{\mu_j^0} - \sum_{j\in R} \nu_i \widetilde{\mu_i^0}. \tag{3.15}$$

Equation 3.14 is exact, but we would like to avoid specifying N in simulation since the solvent is typically not modeled explicitly, as we elaborate on in the next section. We would also like to determine $\widetilde{\Delta G^0}$ from literature values of ΔG^0. To these ends we first rewrite Equation 3.14 as

$$\Delta G_{(0)} = \widetilde{\Delta G^0} + k_B T \log \widetilde{Q}_{(0)} + k_B T \log \frac{N^N}{(N+\sigma)^{N+\sigma}} \tag{3.16}$$

where

$$\widetilde{Q}_{(0)} = \prod_{i\in R} \frac{(N_i - \nu_i)^{N_i-\nu_i}}{N_i^{N_i}} \prod_{j\in P} \frac{(N_j + \upsilon_j)^{N_j+\upsilon_j}}{N_j^{N_j}}. \tag{3.17}$$

From Equations 3.8 and 3.15 we can write

$$\widetilde{\Delta G^0} = \Delta G^0 + \sigma k_B T \log \frac{N}{\Theta} \tag{3.18}$$

where

$$\Delta G^0 = \sum_{j\in P} \upsilon_j \mu_j^0 - \sum_{j\in R} \nu_i \mu_i^0. \tag{3.19}$$

Inserting this to Equation 3.16, we have

$$\begin{aligned}\Delta G_{(0)} &= \Delta G^0 + k_B T \log \left(\frac{N}{\Theta}\right)^{\sigma} \frac{N^N}{(N+\sigma)^{N+\sigma}} + k_B T \log \widetilde{Q}_{(0)} \\ &= \Delta G^0 - \sigma k_B T \log \Theta + k_B T \log \left(\frac{N}{N+\sigma}\right)^{N+\sigma} + k_B T \log \widetilde{Q}_{(0)}.\end{aligned} \tag{3.20}$$

We now make the approximation that $N \gg \sigma$, which is certainly reasonable for most realistic parameterizations of the compartment grid (in the example of a 0.125 μm^3 compartment filled with water, $N \sim 10^9$ while $\sigma \sim 1$). With this we can write the third term in Equation 3.20 as $-\sigma k_B T$,[2] giving

$$\Delta G_{(1)} = \Delta G^0 - \sigma k_B T \log \Theta - \sigma k_B T + k_B T \log \widetilde{Q}_{(1)}, \tag{3.21}$$

where $\widetilde{Q}_{(1)} = \widetilde{Q}_{(0)}$ (we update the subscript to indicate the use of this quantity in a more approximate version of the formula for ΔG). We recommend using Equation 3.21 in simulation because it allows to incorporate literature values for ΔG^0 and avoids specification of N. To understand the term $-\sigma k_B T$ in Equation 3.21 and to understand the relationship between Equations 3.21 and 3.16 and the textbook expression for ΔG, we make further approximations leveraging the large sizes of the solute copy numbers compared to their stoichiometric coefficients. First, we rewrite $\widetilde{Q}_{(1)}$ as

$$\widetilde{Q}_{(1)} = \prod_{i \in R} \left(1 - \frac{\nu_i}{N_i}\right)^{N_i} (N_i - \nu_i)^{-\nu_i} \prod_{j \in P} \left(1 + \frac{\upsilon_j}{N_j}\right)^{N_j} (N_j + \upsilon_j)^{\upsilon_j}. \tag{3.22}$$

Assuming $N_i \gg \nu_i$ and $N_j \gg \upsilon_j$, and using the limit

$$\lim_{x \to \infty} \left(1 \pm \frac{y}{x}\right)^x = e^{\pm y} \tag{3.23}$$

we obtain

$$k_B T \log \widetilde{Q}_{(1)} \approx \sigma k_B T + k_B T \log \prod_{i \in R} (N_i - \nu_i)^{-\nu_i} \prod_{j \in P} (N_j + \upsilon_j)^{\upsilon_j}. \tag{3.24}$$

Inserting this into Equation 3.21, canceling the term $\sigma k_B T$, gives

$$\Delta G_{(2)} = \Delta G^0 - \sigma k_B T \log \Theta + k_B T \log \widetilde{Q}_{(2)}, \tag{3.25}$$

where

$$\widetilde{Q}_{(2)} = \prod_{i \in R} (N_i - \nu_i)^{-\nu_i} \prod_{j \in P} (N_j + \upsilon_j)^{\upsilon_j}. \tag{3.26}$$

---

[2] We have $-\log\left(\frac{N+\sigma}{N}\right)^{(N+\sigma)} = -\log\left(1 + \frac{\sigma}{N}\right)^{(N+\sigma)} \approx -\sigma$ for large N.

Finally, introducing

$$\widetilde{Q}_{(3)} = \prod_{i \in R} N_i^{-\nu_i} \prod_{j \in P} N_j^{\upsilon_j} \tag{3.27}$$

and discarding terms in $\Delta G_{(2)}$ that scale like $\frac{\nu_i}{N_i}$ or $\frac{\upsilon_j}{N_j}$, we arrive at

$$\Delta G_{(3)} = \Delta G^0 - \sigma k_B T \log \Theta + k_B T \log \widetilde{Q}_{(3)} = \Delta G^0 + k_B T \log Q \tag{3.28}$$

where

$$Q = \prod_{i \in R} C_i^{-\nu_i} \prod_{j \in P} C_j^{\upsilon_j}. \tag{3.29}$$

We see that Equation 3.28 is obtained from Equation 3.25 upon making the approximations $N_i - \nu_i \approx N_i$ and $N_j + \upsilon_j \approx N_j$. This final result in Equation 3.28 is a standard textbook expression for the change in Gibbs free energy [1, 44].

We thus have four expressions for ΔG of chemical reactions:

- Equation 3.14 for $\Delta G_{(0)}$ is exact, however it requires specifying N.
- Equation 3.21 for $\Delta G_{(1)}$ uses the approximation $N \gg \sigma$. We recommend the use of this expression because it is the most exact expression for which we need not specify N, and since it is written in terms of ΔG^0, for which literature values can be found.
- Equation 3.25 for $\Delta G_{(2)}$ uses the approximations $N_i \gg \nu_i$ and $N_j \gg \upsilon_j$.
- Equation 3.28 for $\Delta G_{(3)}$ uses the approximations $N_i \gg \nu_i$ and $N_j \gg \upsilon_j$ again.

In Appendix C, we provide expressions for the accuracy of these approximations. Note that, without specifying the copy number of solvents N_s, we can only approximately compute changes in Gibbs free energy, and not the instantaneous Gibbs free energy of the system. Typically only the changes are of interest.

The definition of Gibbs free energy states that

$$\Delta G = \Delta H - T \Delta S. \tag{3.30}$$

Comparing this expression to Equation 3.28, we identify

$$\Delta G^0 = \Delta H - T\Delta S^0, \tag{3.31}$$

and

$$-T\Delta S_{\text{mix}} = k_B T \log Q, \tag{3.32}$$

where ΔH is the enthalpy of reaction, and ΔS^0 and ΔS_{mix} represent the changes in entropy due to the molecular conformations and the translational motions, respectively:

$$\Delta S = \Delta S^0 + \Delta S_{\text{mix}}. \tag{3.33}$$

It is instructive to realize that the discrepancies between the various expressions for ΔG are due to how the ΔS_{mix} term in Equation 3.33 is written.

When employing any of the above expressions that involve the logarithms of products of copy numbers that are raised to the power of other copy numbers, we recommend splitting the logarithm of products into a sum of logarithms as well as using $\log x^y = y \log x$ in order to prevent overflow resulting from computing very large numbers. The results of this section do not assume a system consisting of multiple, weakly-interacting compartments, which we discuss next.

3.2.2 Thermodynamics of a reaction-diffusion compartment grid

Now that we have treated the scenario of a reaction event occurring within a single compartment, we want to generalize to the case of a grid of compartments. For this we develop an argument based on timescales which will allow us to track the copy numbers of the reactive solutes while ignoring those of the inert solvent. This is a necessary modeling feature due to the computational infeasibility of tracking the solvent copy numbers in each compartment. After describing the thermodynamic framework for a grid of compartments, in the next section we derive the change in G resulting from diffusion of solutes between adjacent compartments.

In simulating a RDME, one commonly treats diffusion between adjacent compartments and chemical reactions within compartments using an augmented set of all species and reactions in the system that treats species as distinct if they belong to separate compartments. Thus if there are $|L|$ reacting species and $|\Omega|$ compartments, where L and Ω are the sets of solute species and compartments respectively, then in the augmented set there are $|L||\Omega|$ species tracked. The solvent species is not tracked explicitly, as we discuss below. The number of reactions in the augmented system, including r chemical reactions per compartment and roughly $z|L|$ diffusion events per compartments (where z is the assumed constant number of neighbors of each compartment, ignoring boundary compartments), is $|\Omega|(r+z|L|)$. This augmented set of species and reactions is then simulated using the Gillespie algorithm, in which the reaction propensities are appropriately scaled according to the compartment volumes [189].

Crucial to the justification of this strategy to simulate a RDME is the assumption that within each compartment, the reacting species can be considered homogeneously distributed, so that one may use mass-action kinetics to determine the propensities. This assumption amounts to the condition that the timescale describing diffusion within compartments, τ_D, is much less than the timescale of chemical reactions, τ_C:

$$\tau_D \ll \tau_C. \tag{3.34}$$

This comparison should be done for each diffusing and reacting species, and the timescale of the fastest reaction (taken as the inverse of the propensity) for each species should be used. If the condition holds, then the process of intra-compartmental diffusion will homogenize the solution faster than chemical reactions occur, so the assumption of mass-action kinetics holds. Let the dimension of the space be d, the length of the (here assumed cubical) compartments be h, and the diffusion constant of a species be D. Then

$$\tau_D \approx \frac{h^2}{2dD}. \tag{3.35}$$

The Kuramoto length is given by

$$l_K = \sqrt{2dD\tau_C}, \tag{3.36}$$

so one can see that the condition $\tau_D \ll \tau_C$ is equivalent to the condition $l_K \gg h$, and thus one can enforce this condition by choosing a smaller compartment size h. For fixed total volume, there is a trade-off between h and $|\Omega|$, which determines the size of the augmented system, and therefore the computational efficiency. The timescale of intra-compartment diffusion τ_D is approximately equal to the timescale of inter-compartment diffusion (which can be given as the inverse of $k_D = \frac{D}{h^2}$), so a third way of describing this condition is that the frequency of jumps between adjacent compartments is much greater than the frequency of chemical reactions inside the compartments, which can be checked empirically in simulation [186, 189]. We note that in the literature these expressions may differ up to a constant coefficient depending on the reference.

In order to approximately describe the thermodynamics of this system, we distinguish between the inert solvent and the dilute, chemically reactive solutes. We assume here that the system is impermeable to the flow of either kind of species to the exterior. Within the system, all species are permeable (though local diffusion constants may be incorporated for each species [186]). We treat the solutes explicitly (at the level of their compartment concentrations), whereas we model the solvent implicitly through an appropriate limit as in the steps leading Equation 3.21 above. To this end we assume the existence of a laboratory timescale τ_l whose purpose is to define the temporal resolution of our measurements, such that any changes occurring on a timescale longer than τ_l will be measured. We assume that τ_l is much longer than the timescale describing the local compartment fluctuations in the solvent copy numbers τ_s, yet shorter than but on the order of the timescale describing the diffusion of the solutes τ_D. On this timescale τ_l, then in the time between the chemical reactions and diffusion events involving the solutes, the system is quasi-equilibrated with respect to fluctuations involving the fast processes of solute and solvent homogenization, and thus the system may be assigned well-defined values of Gibbs free energy [132]. This hierarchy of timescales can be written as

$$\tau_s \ll \tau_l \lesssim \tau_D \ll \tau_C. \tag{3.37}$$

Typical ratios of the diffusion constants for solute to solvent are in the range of 1/10 to 1/100, placing τ_D/τ_s in the range of 10 to 100 [209].

Thus on the timescale τ_l there is enough temporal resolution to track the diffusion and chemical reactions of the solutes, while allowing averaging over the fluctuations of the solvent. To describe activity occurring over the large grid of compartments for extended systems, we introduce new timescales τ_x^g, where x refers to any of the timescales defined above. If we hold the compartment size h and the chemical concentrations fixed and add more compartments to our system, the rates of solute diffusion events and chemical reactions occurring anywhere in the system will scale as $|\Omega|$, the number of compartments, and thus the timescales needed to describe them scale as $|\Omega|^{-1}$, i.e. $\tau_l^g \sim \tau_l/|\Omega|$. One might be concerned that τ_l^g will be less than τ_s for large systems, ostensibly violating our requirement that we be able to average over the solvent fluctuations to define quasi-equilibrated states. However, the timescale of solvent fluctuations across the whole grid, τ_s^g, will also scale inversely with $|\Omega|$, so the condition for being able to average over solvent fluctuations expressed in Equation 3.37 does not ultimately depend on the number of compartments. In other words, for systems with many compartments, as long as Equation 3.37 holds for one compartment, we can be sure that our laboratory timescale that describes the whole grid, τ_l^g, will be short enough to describe activity of the solutes while long enough to allow averaging over the fluctuations of the solvent occurring locally in each compartment.

We assume that the exterior of the system acts as a reservoir for the thermodynamic variables p and T. The volume V of the system also remains constant, however under the assumption of the solvent being an incompressible liquid, the change in quantity pV is approximately zero, and for each reaction the change in Gibbs free energy is equal to that of the Helmholtz free energy. Thus it is inconsequential whether we consider

p or V to be reservoir variables, and we choose p in order to speak of the Gibbs free energy of the system. We can write the Gibbs free energy of the system as $G(\mathbf{N}, p, T)$, where $\mathbf{N} = \{\{N_{i,A}\}_{i\in M}\}_{A\in\Omega}$ represents the set of copy numbers of all solute and solvent species in each compartment in the grid. We further assume the compartments to be only weakly-interacting; that is, they can exchange energy and particles, but the interaction of the two subsystems does not contribute a term to the Gibbs free energy of the system. This is equivalent to assuming that the Gibbs free energy of the compartments G_A are linearly additive:

$$G = \sum_{A\in\Omega} G_A \tag{3.38}$$

without any terms of the form G_{AB}. To justify this, we first note that the interaction free energy between two adjacent compartments will primarily be due to the interaction of the solvent at the interface. This interfacial free energy will, even for mesoscopically sized compartments, be small compared to the free energy due to the bulk of the compartment. Further, our main interest will be in changes in the Gibbs free energy of the system due to solute diffusion between compartments and chemical reactions, neither of which will significantly affect the interfacial free energy, so all terms of the form G_{AB} will drop out of the expression for ΔG.

In our approach of averaging over fluctuations in the solvent amounts and taking the limit that this average is large compared to the copy numbers of the solutes, we are choosing to neglect the changes in Gibbs free energy of the system owing to the solvent fluctuations. We justify this with an argument that these fluctuations are small compared to those resulting from the activity of the solute molecules; it also because it arises out of necessity due to the computational intractability of tracking the solvent fluctuations. On the laboratory time scale τ_l, each compartment has an average copy number of solvent molecules, $\overline{N_{s,A}}$. The fluctuations in this quantity will have a standard deviation on the order of $\overline{N_{s,A}}^{1/2}$ [132]. As indicated previously (to arrive at Equation 3.21), we avoid specifying $\overline{N_{s,A}}$ by taking the limit that it is much larger than the number of solute species, and thus fluctuations in this quantity don't matter when computing nearly exact changes in the Gibbs free energy accompanying reactions involving the solutes. We need to establish that the change in Gibbs free energy of a compartment due to a fluctuation in the solvent copy number on the order of $\overline{N_{s,A}}^{1/2}$ is small compared to the change in Gibbs free energy accompanying chemical reactions and inter-compartment solute diffusion events, and thus that the fluctuations in this quantity are not outweighing the changes that we are measuring. In Appendix C we show that if the concentrations of solutes are very different in the two compartments, then this Gibbs free energy change is on the order of $\epsilon\overline{N_{s,A}}^{1/2}k_BT$, where ϵ is the ratio of solutes to solvent. We show that this quantity is typically much less in magnitude than the change in Gibbs free energy accompanying a solute diffusion event. If the concentrations are nearly equal, then the Gibbs free energy is on the order of ϵk_BT, and is thus negligibly small. We make the modeling choice to ignore these changes in the Gibbs free energy resulting from solvent fluctuations because they tend to be small and they do not represent the processes we are interested in which involve the activity of the solute molecules.

To summarize, we track the changes in the Gibbs free energy of the system by computing a value of ΔG whenever a chemical reaction occurs within a compartment or a solute diffusion event occurs between adjacent compartments. By the linear additivity of the compartments' free energies, any change in the free energy of a single compartment is equal to the change in free energy of the whole system (i.e. $\Delta G_A = \Delta G$). We assume that the activity of the solvent only contributes small, fluctuating, unbiased changes to the free energy which we ignore; we also assume that the amount of solvent is so large that one can neglect the fluctuations in this quantity when computing the changes in free energy for processes involving the solute copy numbers.

3.2.3 ΔG of diffusion

To describe the change in Gibbs free energy accompanying a solute diffusion event between neighboring compartments we use a similar approach to the one used above for chemical reactions. The key difference here is that as opposed to multiple species being involved in a reaction taking place in a single compartment, we now have a single species involved in a diffusion event taking place between two compartments.

Consider species i diffusing from compartment A, where its initial copy number is $N_{i,A}$, to compartment B, where its intial copy number is $N_{i,B}$. The total copy numbers of molecules in the compartments A and B are N_A and N_B. Assume there is just one spectator species constituting the solvent, labeled s with copy numbers $N_{s,A}$ and $N_{s,B}$ (we remove our uncertainty in the exact values of these numbers by taking a limit later). As a result of the diffusion event, we have the following changes in these quantities: $N_{i,A} \rightarrow N_{i,A}-1$, $N_{i,B} \rightarrow N_{i,B}+1$, $N_{s,A} \rightarrow N_{s,A}$, $N_{s,B} \rightarrow N_{s,B}$, $N_A \rightarrow N_A - 1$, and $N_B \rightarrow N_B + 1$. The initial Gibbs free energy is

$$\begin{aligned} G^{\text{initial}} = {} & N_{i,A}\left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,A}}{N_A}\right) + N_{i,B}\left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,B}}{N_B}\right) \\ & + N_{s,A}\left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,A}}{N_A}\right) + N_{s,B}\left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,B}}{N_B}\right), \end{aligned} \tag{3.39}$$

and the final Gibbs free energy is

$$\begin{aligned} G^{\text{final}} = {} & (N_{i,A}-1)\left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,A}-1}{N_A-1}\right) + (N_{i,B}+1)\left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,B}+1}{N_B+1}\right) \\ & + N_{s,A}\left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,A}}{N_A-1}\right) + N_{s,B}\left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,B}}{N_B+1}\right). \end{aligned} \tag{3.40}$$

The difference of these two expressions leads to an exact formula, similar to Equation 3.16, which we omit here. Analogously to the approximation $N \gg \sigma$ made in the context of chemical reactions, here we assume that $N_A,\ N_B \gg 1$, which amounts to setting $N_A - 1 \approx N_A$ and $N_B + 1 \approx N_B$ in Equations 3.39 and 3.40. Using this approximation, we can express the change in Gibbs free energy as

$$\Delta G = k_B T \log \frac{(N_{i,A}-1)^{(N_{i,A}-1)}}{N_{i,A}^{N_{i,A}}} \frac{(N_{i,B}+1)^{(N_{i,B}+1)}}{N_{i,B}^{N_{i,B}}} \tag{3.41}$$

We see that we, analogously to making the approximation $N_i \gg \nu_i$ above, if we here take $N_{i,A}, N_{i,B} \gg 1$, then Equation 3.41 reduces to the common expression

$$\Delta G = k_B T \log \frac{N_{i,B}}{N_{i,A}}. \tag{3.42}$$

The right hand side of Equation 3.42 will always be less than that of Equation 3.41, which, although the difference is typically very slight, can lead to systematically biased calculations, as we show next.

3.3 Results

Here we perform stochastic simulations to illustrate the effects of approximating ΔG for reactions and diffusion when copy numbers are considered small and discrete. We use MEDYAN, a simulation platform designed to study active networks at high resolution, which is equipped with a RDME simulation engine of the type described above [42]. With this, we report on two different simulation set-ups, to test the discrepancy between the nearly exact and approximate forms of ΔG corresponding to reactions and to diffusion. We compare the nearly exact Equations 3.21 and 3.41 for reactions and diffusion, respectively, with their approximate counterparts, Equations 3.28 and 3.42. We observe that only the nearly exact expressions result in sensible behavior, i.e. that the rate of change of Gibbs free energy on average is 0 k_BT/s when the system is at equilibrium.

3.3.1 ΔG of reactions

To test the effect of approximation for ΔG of reactions, we consider the simple reaction scheme

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} C \tag{3.43}$$

where the rate constants to the right and left are $k_+ = 0.05\ \mu M^{-1}s^{-1}$ and $k_- = 0.01$ s^{-1} respectively, giving an equilibrium constant of $K_{eq} = k_-/k_+ = 0.2\ \mu M$. We perform stochastic simulations with the Next Reaction Method in MEDYAN [210], using a single compartment of size 0.125 μm^3 (i.e. in this example there is no diffusion). To employ Equation 3.21 in simulation, we first compute $\Delta G^0 - \sigma k_B T \log \Theta - \sigma k_B T = 3.71\ k_B T$ for the forward reaction and $-3.71\ k_B T$ for the reverse reaction, and then when each reaction fires during simulation, the quantity $\widetilde{Q}_{(1)}$ is computed from the instantaneous values of the copy numbers to determine $\Delta G_{(1)}$. A similar approach is taken to employ Equation 3.28. We begin with $N_A = 100$, $N_B = N_C = 50$, and repeat a simulation of 100 s duration 3,000 times to obtain averages of the trajectory of rates $\partial_t \Delta G(t)$ resulting from the forward and reverse reactions (the notation $\partial_t \Delta G(t)$ implies that the measured rates vary smoothly as a function of time, however this quantity is calculated as the total change in Gibbs free energy resulting from discretely timed chemical events during 1 second-long windows). Figure 3.2 displays the results of these simulations. Note how, while the two trajectories bear close similarity, the equilibrium value of $\partial_t \Delta G(t)$ centers around 0 $k_B T/s$ for the nearly exact formulation Equation 3.21, yet erroneously centers around $\sim -0.08\ k_B T/s$ for the approximate version, Equation 3.28.

3.3.2 ΔG of diffusion

To test the effect of approximating ΔG for diffusion in a compartment-based reaction-diffusion scheme, we next employed MEDYAN to simulate diffusion of 1,000 molecules with diffusion constant 20 $\mu m^2 s^{-1}$ in a $2 \times 2 \times 2$ grid of compartments, each a cube with volume 0.125 μm^3. The initial distribution of molecules is uniformly random over the compartment grid, and thus the system begins near equilibrium and is then allowed to stochastically evolve for 100 s, i.e. the molecules hop randomly between adjacent compartments. For each diffusion event, the value of ΔG is determined using the nearly exact Equation 3.28 for one run, and in another run the approximate Equation 3.42 is used. The difference in the trajectories of $\partial_t \Delta G$ for these simulations is stark, as displayed in Figure 3.2. While the trajectory centers around 0 $k_B T/s$ for the nearly exact formulation of ΔG, it erroneously centers around $\sim -1,930\ k_B T/s$ for the approximate version. Diffusion events are very frequent in this system, occurring around 240,000 times per second (this number can be calculated from the parameters of the system and

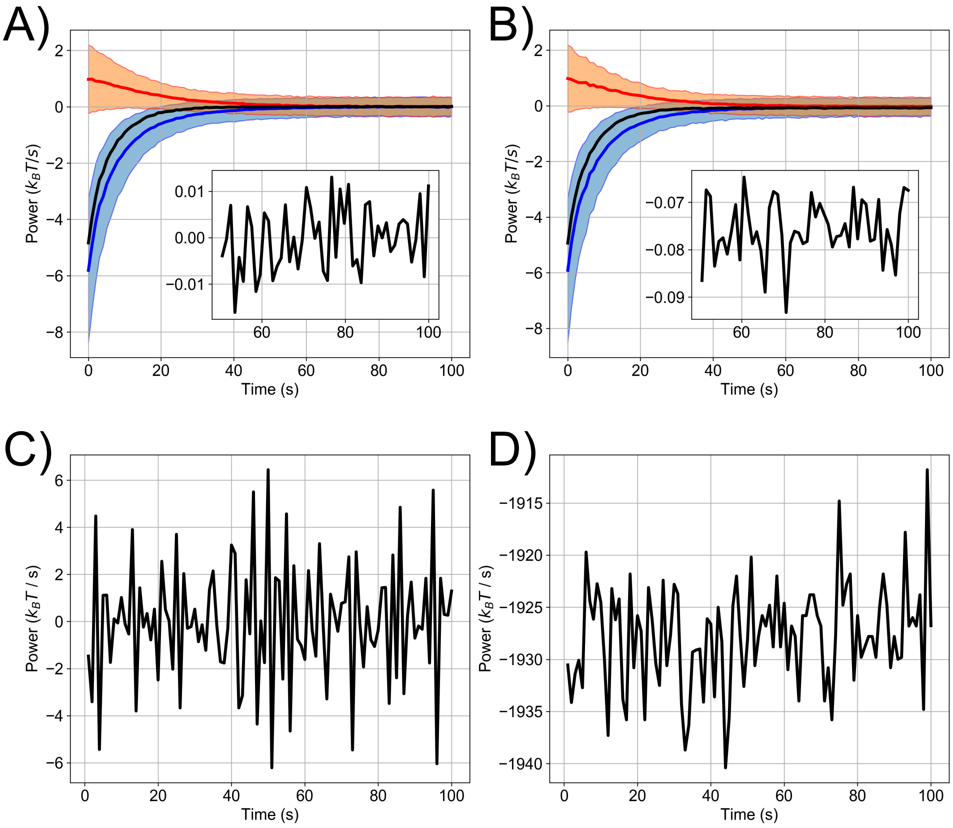

FIGURE 3.2: Numerical results illustrating the difference between nearly exact and approximate formulations of ΔG for reactions and diffusion. A) Averages over 3,000 simulations of the chemical scheme $A+B \rightleftharpoons C$, using the nearly exact Equation 3.21. The blue curve represents the trajectory of $\partial_t \Delta G$ resulting from the forward reaction, the red curve represents that from the reverse reaction, and the black curve represents their sum. Shaded regions represent the standard deviation over the 3,000 repeated trials. The inset displays a blow-up of the black curve once the system has gotten close to equilibrium. B) The same as just described, but using the approximate Equation 3.28. C) A single trajectory of the diffusion of 1,000 molecules over a 1 μm^3 cubic grid of 8 compartments, beginning from a random initial spatial distribution. Values of ΔG are calculated using Equation 3.41. D) The same as just described, however values of ΔG are calculated using Equation 3.42.

is also observed during simulations). Thus, since Equation 3.42 is always less than the nearly exact quantity, even by a small amount on the order 0.05 $k_B T$ for this system, this systematic bias is amplified by the frequency of diffusion events to produce significant differences from expected behavior, necessitating the use of a more exact formula for ΔG. Lastly, we performed a simulation involving both reactions and diffusion across multiple compartments. We found again that only when the nearly exact formulas were used did the rate of change of Gibbs free energy center on 0 $k_B T/s$ at equilibrium. It is not additionally illuminating to show the data, so we do not display it here.

3.4 Discussion

We have argued that when the copy numbers of the reactants and products are treated as small, discrete quantities, then certain approximations leading to the textbook formulas for ΔG of reaction and diffusion, Equations 3.2 and 3.3, break down and lead to biased results. We emphasize that this is true only when the copy numbers and reaction occurrences are treated as discrete; when they are treated as continuous, one should use the textbook formulas. This can be shown by considering a continuous version of the chemical system described by Equation 3.43. The time evolution of the concentrations of the chemical species is obtained by solving a system of ordinary differential equations that employ mass-action kinetics, and from this solution, the rates $\partial_t \Delta G(t)$ resulting from the forward and reverse reactions are computed using the instantaneous values of the species' copy numbers which enter into Equations 3.21 and 3.28. Figure 3.3 displays the results of these calculations. Here, the total rate of $\partial_t \Delta G(t)$ only approaches 0 $k_B T\ /s$, as it must at equilibrium, when Equation 3.28 is used.

When the copy numbers of the chemical species are treated as continuous, there is no notion of a single occurrence of a reaction; instead the evolution of the system is parameterized by the continuous variable ξ which quantifies the extent of advancement of the reaction [211]. In this framework, which is adopted in classical thermodynamics, the copy number of any species never jumps instantaneously from N_i to $N_i + \nu_i$, and

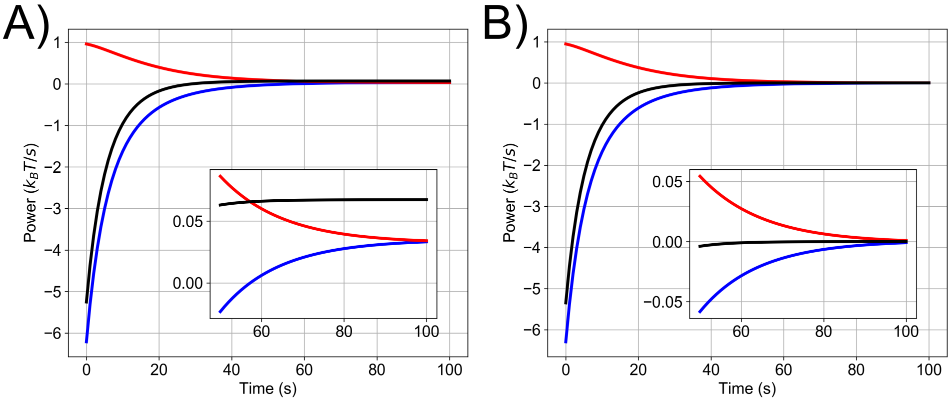

FIGURE 3.3: Analytical results illustrating that, when copy numbers are treated as continuous variables, the textbook formula for ΔG of reaction should be used. A) Calculation of the rates $\partial_t \Delta G(t)$ using Equation 3.21 resulting from the forward (blue curve) and reverse (red curve) reactions, as well as their sum (black curve), as described in the main text. The inset shows a blow-up of these curves as the system approaches equilibrium. B) The same as just described, but using Equation 3.28.

thus the premise of the derivation presented above leading to Equation 3.16 falls apart. This explains why the seemingly more accurate Equation 3.21 is wrong when applied to chemical dynamics that are described using continuous variables to represent the copy number of species. When the chemical dynamics are modeled this way, the textbook expression Equation 3.28 for the change in Gibbs free energy is valid. Another way to understand the difference in these expressions is to view Equation 3.28 as giving the instantaneous slope of G with respect to the degree of advancement of the reaction when these quantities are viewed as continuous variables [205, 208], whereas Equation 3.16 effectively integrates the continuously varying quantity G through a single discrete reaction occurrence.

If the copy numbers of chemical species are not large compared to the stoichiometric coefficients (i.e. N_i is not much greater than ν_i), it is best to describe the chemical dynamics by treating the copy numbers as discrete variables participating in stochastically timed chemical reactions. This is the philosophy adopted by several recent models of intracellular environments, where copy numbers of molecules of interest are sometimes quite small. In these cases, adoption of the more exact expression for ΔG can not only improve precision, but can ensure correct behavior. If concentration gradients are expected to be a strong source of entropy production (e.g. diffusion of monomeric actin along the lengths of filopodia, see Refs. 212–214), then using the nearly exact formula presented here can ensure that the resulting measurements of dissipation are not strongly biased.

Chapter 4

Quantifying dissipation in actomyosin networks

This chapter and Appendix D are adapted from: Floyd, C., Papoian, G. A., & Jarzynski, C. (2019). Quantifying dissipation in actomyosin networks. Interface focus, 9(3), 20180078.

4.1 Introduction

The actin-based cytoskeleton is a dynamic supramolecular structure that, by sustaining and releasing mechanical stress in response to various physiological cues, mediates the exertion of force by cells both on their environments and within their bodies [2, 215]. These cytoskeletal structures are composed of long (on the order of 1 μm *in vivo* [216]) actin polymers which are interconnected by various cross-linkers, as well as by myosin motor filaments, resulting in a three-dimensional network-like organization referred to as an "actomyosin network" [217, 218].[1] Part of the intricacy of actomyosin network dynamics is due to the mechanosensitive kinetic reaction rates controlling cross-linker and myosin filament unbinding as well as myosin filament walking: at high tension, cross-linkers will unbind more quickly (slip-bond) whereas motors will unbind and walk less quickly (catch-bond and stalling). [72, 87, 103]. These reactions control the actomyosin

---

[1] In our terminology, we will distinguish between "cross-linking proteins," which will include both active (e.g. myosin filaments) and passive (e.g. α-actinin and fascin) proteins that bind to adjacent actin filaments, and "cross-linkers," which refer exclusively to passive cross-linking proteins.

network connectivity, which in turn determines the ability of the network to globally distribute stress [104]. Thus the mechanosensitive feedback introduces nonlinear coupling between the stress sustained by an actomyosin network and the network's ability to reorganize in response to that stress. In order to be responsive to physiological cues, the dynamics of these systems occur far from thermodynamic equilibrium; the hydrolysis of an out-of-equilibrium concentration of ATP molecules fuels a) the stress-generating activity of the myosin motor filaments, and b) filament treadmilling [88–92]. Filament treadmilling is a steady state situation in which the polymerization at the plus end of the filament is compensated by the depolymerization at the minus end, resulting in the filament moving forward without its length changing. As a result of these these local free energy-consuming processes, actomyosin networks constitute an interesting and biologically important example of soft active matter. Active matter is composed of agents that individually transduce free energy from some external source, in this case the chemical potential energy of many ATP molecules [26, 73, 86]. Dissipation in these systems results when the free energy consumed ΔG is greater than the quantity of work W done by the system on its environment, with the remainder $\Delta G - W$ serving to increase the total entropy.

The viewpoint of actomyosin networks as active matter systems has been fruitfully adopted in recent theoretical and experimental studies, yet a lack of ability to quantify the rates of free energy transduction by these systems has hindered development of some of these lines of study. The emergence of distinctive dynamical states (for instance pulsing actin waves or vortices) during the self-organization of actomyosin systems has been documented in several *in vitro* experiments [67, 219, 220]. These emergent patterns depend sensitively on the concentrations of myosin filaments and cross-linkers: myosin filament concentration controls the level of active stress generation, and cross-linker concentration controls the degree of mechanical coupling of actin filaments, which has been described using the language of percolation theory [104, 221]. While these emergent dynamic patterns have been characterized in detail, a general mechanism explaining why these patterns emerge under given conditions has not yet been proven. It might be expected, given that these systems operate away from thermodynamic equilibrium, that the quantity of free energy dissipated during a system's evolution is optimized, similar to the principle of minimum entropy production in the near-equilibrium theory of irreversible thermodynamics [198]. However, this minimum entropy production principle breaks down in the far-from-equilibrium, nonlinear-response regime, where many active matter systems including actomyosin networks operate [201]. It has recently been proposed that another optimization principle applies arbitrarily far from equilibrium. This idea, referred to as dissipative adaptation, suggests that, in general, a coarse-grained trajectory of some non-equilibrium system will be more likely than all alternative trajectories if the amount of free energy absorbed and dissipated along that trajectory is maximal [152, 155, 200]. This organizing principle has been borne out in model systems [153, 154], yet has also been shown to have certain counter examples [222]; it remains actively debated. In the case of actomyosin systems, it has not yet been tested because of the difficulty in measuring dissipated free energy using most experimental approaches. In this chapter we take the first steps toward such a test, by developing a simulation methodology allowing the quantification of dissipated free energy during the self-organization of actomyosin networks.

In addition to being of interest in the field of active matter systems, dissipation in actomyosin networks has also been an important factor in recent experimental studies of cell mechanics. Rheological properties of actomyosin networks largely determine rheological properties of the whole cell, and it has been discovered that the dissipative component of the cell's viscoelastic response to mechanical oscillations (called the loss modulus) is partly attributable to the ATPase activity of myosin motor filaments in actomyosin cytoskeleton [223–225]. In the context of cell mechanics, myosin filaments have several roles: they produce mechanical stress by pulling on actin filaments, they dissipate mechanical stress by disassembling actin filaments and higher order stress-sustaining filament structures, and they consume chemical energy through the hydrolysis of ATP molecules. It has been proposed that a lack of detailed understanding of the effects of these processes, and of dissipation in actomyosin systems more generally, underlies inconsistent, widely variable traction force microscopy measurements of cell migration [170]. Progress along this line is hindered by an absence of methods to study dissipation in actomyosin networks directly and at sufficiently high spatio-temporal resolution.

To address these needs, we introduce a computational approach to measure dissipation during simulations of actomyosin network self-organization using the simulation platform MEDYAN [42, 226]. MEDYAN simulations marry stochastic reaction-diffusion chemistry algorithms with detailed mechanical models of actin filaments, cross-linkers, myosin motor filaments, and other associated proteins, and it also accounts for mechanosensitive reaction rates. This combination of simulation features makes this software uniquely capable of probing the complexity of actomyosin network dynamics. For instance, past studies utilizing MEDYAN have investigated the dependence of network collapse on myosin filament and cross-linker concentrations, as well as the origin of local contractility in actomyosin networks [42, 94]. We refer the reader to the paper describing MEDYAN for a detailed discussion of the various aspects of the simulation platform [42], while here we describe an extension of that platform that allows for calculation of the energetics of the chemical and mechanical events occurring during simulation. We utilize these new capabilities to characterize the dissipation resulting from filament treadmilling, for which we further introduce a mean-field model, as well as from myosin filament walking. We study both the time-dependence and the distributions of dissipation rates as concentrations of cross-linkers and myosin filaments are varied, observing that transduction of chemical energy to stored mechanical energy is more efficient at denser network organizations. For these simulations, we first explore systems with “plain” myosin filaments and cross-linkers that are not mechanosensitive, in order to simplify the overall dynamics. We then introduce their mechanochemical coupling to understand its effect on the observed trends. We end by discussing how this new methodology can provide a valuable technique to advance the studies of actomyosin networks mentioned above.

4.2 Methods

4.2.1 Measuring dissipation in MEDYAN

We first give a brief overview of the MEDYAN simulation platform. MEDYAN employs a stochastic chemical evolution algorithm in conjunction with mechanical representations of polymers and cross-linking proteins to simulate the dynamics of networks with active components, including but not limited to actomyosin networks. The simulation space comprises a grid of reaction-diffusion compartments, inside which chemical species (e.g. unpolymerized subunits or cross-linking proteins) are assumed to be homogeneously distributed without specified locations, and which participate in reactions (e.g. (de)polymerization or (un)binding) according to mass-action kinetics; the species can additionally jump between compartments in diffusion events. When an unpolymerized subunit polymerizes to or nucleates a filament, it becomes part of the mechanical subsystem, gaining location coordinates in the simulation volume and becoming subject to mechanical potentials depending on its interaction with other mechanical elements. Through chemical reactions such as myosin filament binding and walking, the mechanical energy of the system changes, and the new net forces are then periodically relaxed in a mechanical equilibration phase, using conjugate gradient minimization. We fill in salient details of the above overview as they become relevant below. A user provides input data including system size and simulation length, mechanical parameters (e.g. stretching and bending constants and excluded volume cutoff distances), size of the polymer subunits, energy minimization algorithm parameters, chemical simulation algorithm parameters, choices for the modeling of force-sensitive reaction rates, initial conditions of the filaments (either specified or randomly generated), a list of reacting species and their associated parameters, a list of reactions involving those species and their associated parameters, and a list of desired output information. The output of a simulation is a set of trajectory files containing information at each time point, which can include positions of the mechanical network elements, tensions on the elements, and copy numbers of the chemical species, among other things. In Appendix D we discuss parameterization of the simulations analyzed in this chapter. MEDYAN is extensible in that it is possible to implement new types of outputs, depending on the experimental needs; in this chapter we describe a novel output that reports the changes in the Gibbs free energy of the system.

As a MEDYAN simulation progresses, the Gibbs free energy of the system continually changes due to occurrences of chemical reactions and structural rearrangements of the polymer network. These processes are driven by an out-of-equilibrium concentration of ATP which fuels filament treadmilling and myosin filament walking. Dissipation measurement in MEDYAN works by calculating running totals of the chemical and mechanical energy changes. The running totals can then be converted into instantaneous rates by taking the numerical derivative at each time point using the forward difference quotient. The algorithm for tracking these energy changes is compatible with the following sequence of consecutive procedures that make up one iterative cycle of a MEDYAN simulation [42]:

1. Evolve system with stochastic chemical simulation for time t_{min}.
2. Calculate the changes in the mechanical energy resulting from the reactions in Step 1.
3. Mechanically equilibrate the network based on the new stresses calculated in Step 2.
4. Update the reaction rates of force-sensitive reactions based on the new forces.

Dissipation tracking is done by calculating for each of these four steps a change in free energy, and then adding these free energy changes to determine the total change in free energy resulting from each cycle, $\Delta G_{\text{dissipated}}$. Since, at least in this study, the actomyosin network is not mechanically coupled to any work reservoir external to the simulation volume (i.e. $W = 0$), $\Delta G_{\text{dissipated}}$ is indeed dissipated energy [227]. This methodology could be straightforwardly extended to account for work exchanged with an external system in future studies, however. Step 4 in the MEDYAN simulation cycle does not result in a change in free energy, as explained in Appendix D. The notation

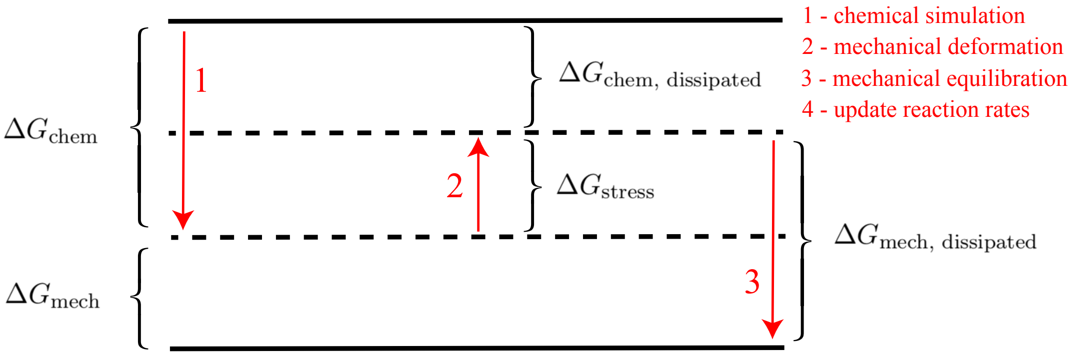

FIGURE 4.1: Energy level diagram indicating which changes in free energy are tabulated during the 4 procedures constituting one cycle in MEDYAN simulation. Step 4, mechanochemical update of reaction rates, does not result in a free energy change. Dotted lines represent energy levels that are intermediate during the iterative cycle, and solid lines represent the energies at the beginning and end of one cycle.

chosen for these free energy changes, as well as the direction in which the energy is changing during each procedure, is illustrated in Figure 4.1.

From this picture, we have the following relations:

$$\Delta G_{\text{dissipated}} = \Delta G_{\text{chem}} + \Delta G_{\text{mech}} \tag{4.1}$$

$$\Delta G_{\text{dissipated}} = \Delta G_{\text{chem, dissipated}} + \Delta G_{\text{mech, dissipated}} \tag{4.2}$$

$$\Delta G_{\text{chem, dissipated}} = \Delta G_{\text{chem}} + \Delta G_{\text{stress}} \tag{4.3}$$

$$\Delta G_{\text{mech, dissipated}} = \Delta G_{\text{mech}} - \Delta G_{\text{stress}} \tag{4.4}$$

For the depicted relative position of energy levels, the sign convention is such that all values of ΔG except for ΔG_{stress} will be negative (indicated by the arrow's direction), since G refers to the free energy of the system, not of its environment, and will therefore be tend to be negative as the system moves down the free energy landscape. The usual intuition that the total dissipation is positive can be stated

$$\Delta G^{+}_{\text{dissipated}} = -\Delta G_{\text{dissipated}} > 0, \tag{4.5}$$

where the superscript “+” indicates the positive change in the total entropy.

Equation 4.3 says that, given some change in the system’s chemical potential energy, ΔG_{chem}, resulting from reactions occurring during Step 1, a portion of that energy is used to deform the polymer network (e.g. via myosin filament pulling on actin filaments). This increases the mechanical energy in the network by an amount ΔG_{stress}. Only the portion of ΔG_{chem} which has not gone into ΔG_{stress} has been dissipated as heat. In Step 3 the network is mechanically equilibrated, resulting in relaxation of net forces (though not of all stresses) and updating of the network elements’ positions. We refer to the decrease in mechanical energy resulting from this relaxation as $\Delta G_{\text{mech, dissipated}}$.

The calculation of ΔG_{stress} and $\Delta G_{\text{mech, dissipated}}$ is based on a set of mechanical potentials describing interactions between elements of the actomyosin network. Polymers are modeled as a sequence of thin, unbendable, yet extensible cylinders that are joined at their ends by beads whose positions define the polymer’s configuration. The structural resolution of MEDYAN is at the level of the cylinders, which in this study are 27 nm long and have effective diameters of approximately 5 nm, however the diameter is not a parameter of the simulation, being instead effectively determined by the strength of the excluded volume interaction between cylinders. Cross-linking proteins (e.g. α-actinin and myosin filaments) are modeled as Hookean springs connecting these cylinders by attaching to discrete binding sites. Included among the mechanical potentials are various modes of filament deformation, excluded volume interactions, and stretching of cross-linkers and myosin filaments. Mechanical equilibration is accomplished by constrained minimization of the mechanical energy with respect to the positions of the network elements. A full description of the mechanical potentials and equilibration protocols is given in Ref. 42. Determining ΔG_{stress} and $\Delta G_{\text{mech, dissipated}}$ requires evaluating the instantaneous total mechanical energy of the system at certain points during the iterative simulation cycle and taking the difference of those values.

The calculation of ΔG_{chem} and $\Delta G_{\text{chem, dissipated}}$ is accomplished by incrementing a running total of the chemical energy G_{chem} whenever a reaction stochastically occurs during Step 1, and finding the accumulated change at the end of the protocol. Chemical stochastic simulation in MEDYAN uses a Gillespie-like reaction-diffusion algorithm over a grid of compartments which constitutes the simulation volume. Diffusing species are assumed to be homogeneous (i.e. obey mass-action kinetics) inside the compartments, and can jump between the compartments leading to concentration gradients at the scale of the compartment length (taken to be roughly the Kuramoto length of diffusing G-actin, following [226]). The evolving polymer network is overlaid on this compartment grid, with each piece of a polymer reacting with diffusing species according to the concentrations in its local compartment. Again, we refer the reader to [42] for a more detailed description of the chemical dynamics. For the present purpose of measuring dissipation, we introduced into this simulation protocol a precise formula for the change in Gibbs free energy corresponding to the occurrence of various reactions as a function of the instantaneous compartment concentrations. This formula was derived in Chapter 3. The set of chemical reactions used to describe actomyosin networks in this study is based on a previous model of actin polymerization dynamics that explicitly treats hydrolysis states of the nucleotide bound to each actin subunit [99, 183]. This level of detail allows to quantify the dissipation resulting from ATP hydrolysis during filament treadmilling. To increase computational efficiency, we neglect the dynamics of nucleotide hydrolysis states of the tips of the filaments. This has been shown in previous work to be a valid approximation to the full dynamics which includes the states of the filament tips [99]. We refer to the resulting set of reactions describing actin polymerization dynamics as the Constant Tip (CT) model. However unlike in the original CT model, here we explicitly include G-actin bound to ADP-Pi as a reacting species, for completeness and since the extra computational strain of doing so is small. The actin subunit species tracked in this model are distinguished by their polymerization state and by the hydrolysis states of the nucleotide to which they are bound. We notate a species as G or F, to represent globular (un-polymerized) or filamentous (polymerized) actin respectively, superscripted by T, Pi, or D to represent that it is bound to ATP, ADP-Pi, or ADP, respectively; thus for instance filamentous actin bound to ADP-Pi is notated F^{Pi}. We also include reactions describing cross-linker (un)binding and myosin filament (un)binding and walking.

We exclude filament nucleation, severing, destruction, and annealing reactions, thus the number of filaments is constant throughout the simulation trajectories. In Appendix D we describe how we compute the change in Gibbs free energy for each reaction in this set, as well as how we parameterize the simulations.

Lastly, we developed a mean-field model to describe just the dissipation resulting from reactions in the CT model, i.e. excluding cross-linkers and myosin filaments. We describe the model and present its results in Appendix D. We find, among other things, that the steady state dissipation rate from filament treadmilling counter-intuitively does not depend on the total amount of actin, but only on the number of filaments.

4.3 Results

4.3.1 Total dissipation rates of disordered networks do not increase

We studied dissipation rates accompanying the process of myosin-driven network self-organization. We first excluded in these simulations the force-sensitivity of the reaction rates describing cross-linker unbinding and myosin filament unbinding and walking. This allowed us to understand a simplified version of the dynamics (we later discuss the the effect of including mechanochemical feedback). We analyzed the trajectories of the quantities $\Delta G_{\text{dissipated}}$, $\Delta G_{\text{chem, dissipated}}$, $\Delta G_{\text{mech, dissipated}}$, ΔG_{chem}, and ΔG_{mech} over a set of simulations with identical initial concentrations but different random filament distributions. In Figure 4.2, we display these trajectories averaged over 10 runs. We used a simulation volume of 1 μm^3, divided into 8 compartments, and initial conditions of equal amounts (10 μM each) of G^T and G^D actin in a 0.08 μM pool of seed filaments containing F^T, as well as 0.2 μM myosin and 1.0 μM cross-linkers. Similar trajectories for other conditions are shown in the supporting information of Ref. 195. Points lying outside 3 median absolute deviations (MAD) have been excluded for this visualization [228].[2]

---

[2] For certain statistical aggregations in this chapter we use the MAD since it is the estimate of scale most robust against outliers, having a breakdown point of 50%. However, we use the unscaled version of the MAD and, as a result, do not claim that this statistic is a consistent estimator of the standard deviation [229]. The distributions of most of the quantities of interest are too pathological

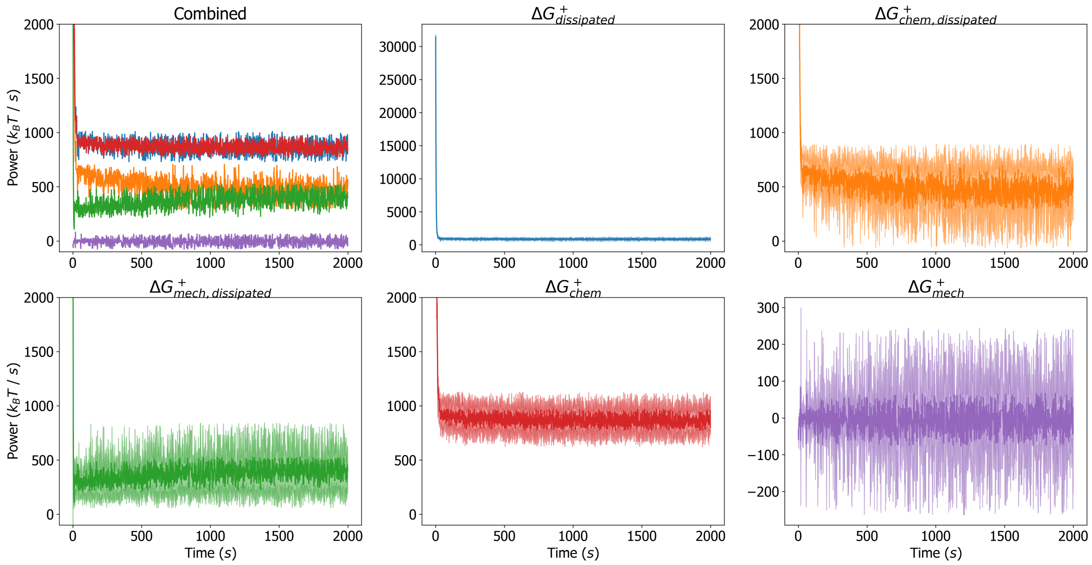

Figure 4.2: *Top Left*: Combined trajectories of 5 quantities tracked during a MEDYAN simulation, averaged over 10 separate runs. The color coding is indicated by the remaining panels. Note the close overlap between $\Delta G^+_{\text{dissipated}}$ and ΔG^+_{chem}. In the remaining panels, the individual trajectories are visualized with their standard deviations at each time point over the 10 runs visualized as lighter curves above and below the main curve. In the plots for $\Delta G^+_{\text{dissipated}}$ and ΔG^+_{mech}, the full range is visualized, however this range is cropped in the other plots to aid visibility.

For each tracked quantity, there is an initial transient phase (lasting a few tens of seconds) followed by fluctuations around a roughly steady value; we do not observe in this set of simulations any slow approach to a significantly different total dissipation rate, which might correspond to large reorganizations of the network. However for high concentrations of cross-linkers ($C_{CL} = 5.0\ \mu M$) and myosin ($C_M = 0.4\ \mu M$), we observe that the contribution of $\Delta G^+_{\text{mech, dissipated}}$ to $\Delta G^+_{\text{dissipated}}$ tends to increase relative to that of $\Delta G^+_{\text{chem, dissipated}}$ (see the supporting information of Ref. 195). It is known that network percolation and collapse occurs only under certain conditions of myosin and cross-linker concentrations [42, 221], and under these conditions which result in collapse we observe an increase in $\Delta G^+_{\text{mech, dissipated}}$ relative to $\Delta G^+_{\text{chem, dissipated}}$ (Figure 4.8).

---

to allow for a straightforward choice of scaling factor, so we simply us the unscaled version, defined as $\text{MAD} = \text{median}_i|x_i - \text{median}_j x_j|$, where x_i are elements in the data set.

This indicates that more mechanical stress is being created by myosin filament walking as the network collapses and becomes more densely cross-linked. The transient phase corresponds to the initial polymerization of the seed filaments followed by the initial mechanical coupling of filaments by cross-linkers and myosin filaments. Following the transient phase, the networks in this study are generally disordered (Figure 4.3). The dissipation rate corresponding to the initial polymerization of the seed filaments is much larger than the chemical dissipation resulting from myosin filament activity. However, following the transient phase, after which the treadmilling dissipation has reached its steady state (Appendix D), the contribution from myosin filament activity outweighs the dissipation resulting from filament treadmilling. Tracking the instantaneous rate of change in $\Delta G^{+}_{\text{chem}}$ resulting from each reaction separately, we observe myosin filament walking to contribute the majority to $\Delta G^{+}_{\text{chem}}$ after the transient phase, however the amount depends on C_{CL} and C_M. The integrated contributions of each reaction to $\Delta G^{+}_{\text{chem}}$ for different conditions are shown in Figure 4.4 and the supporting information of Ref. 195. Interestingly, diffusion contributes an appreciable fraction to $\Delta G^{+}_{\text{chem}}$; plus and minus ends of actin filaments tend to localize together (Figure 4.3) and through treadmilling deplete the concentrations of G^T and G^D in certain reaction compartments relative to others, leading to significant diffusion gradients on the scale of the compartment length. This suggests that the establishment of concentration gradients is an important driving force in actomyosin self-organization, as has been noted in other work [212, 230]. In the initial transient phase, the mechanical energy changes appreciably as the filaments grow and are initially coupled to each other. Following this, however, the rate of $\Delta G^{+}_{\text{mech}}$ is on average near zero. This indicates that, despite the process of mechanical stress generation through myosin filament activity and treadmilling, the resulting stress is dissipated through fast relaxation such that, on a slower timescale, the mechanical energy of the system does not change in a significant, persistent way.

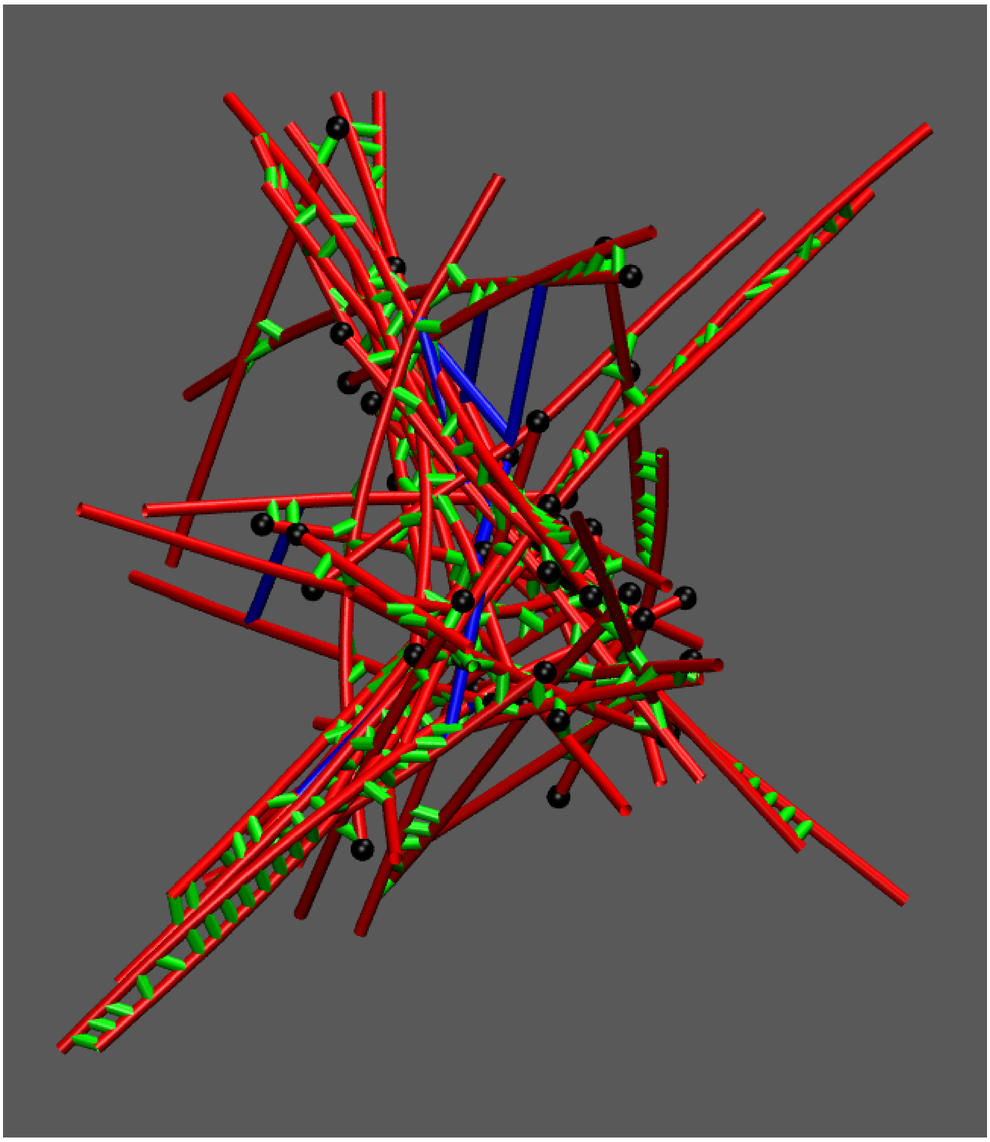

Figure 4.3: Snapshot of a percolated actomyosin network in MEDYAN under conditions $C_{CL} = 5.0\ \mu M$ and $C_M = 0.4\ \mu M$. Actin filaments are drawn in red, cross-linkers are drawn in green, myosin filaments are drawn in blue, and the plus ends of filaments are drawn as black spheres.

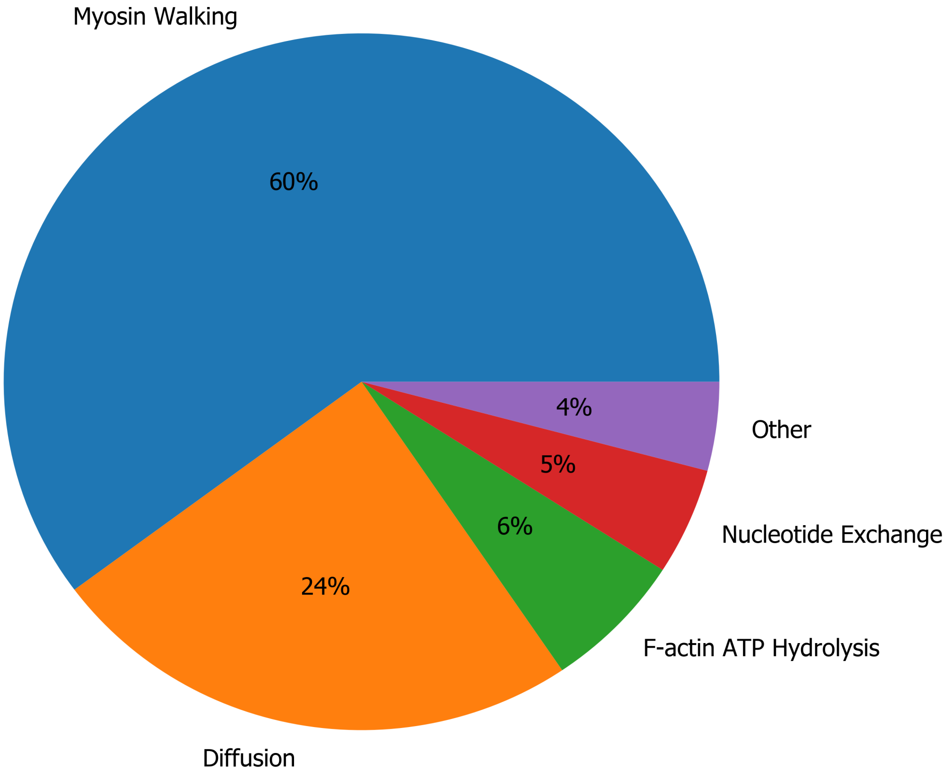

FIGURE 4.4: Integrated contributions of each reaction in MEDYAN to the total ΔG_{chem} along a simulation trajectory with $C_{CL} = 1.0\ \mu M$ and $C_M = 0.2\ \mu M$.

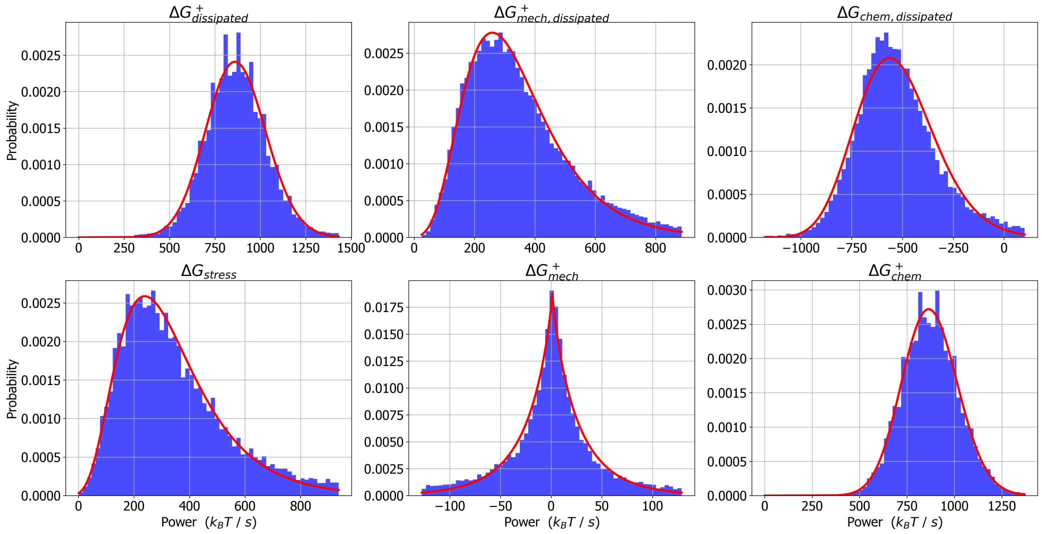

FIGURE 4.5: Histograms and fitted probability distribution functions for 6 tracked quantities. For each histogram, the full trajectory for each of 10 runs is combined into a single data set. A log-normal distribution was used to fit the histograms of $\Delta G^{+}_{\text{dissipated}}$ and $\Delta G^{+}_{\text{chem}}$, a generalized normal distribution was used to fit $\Delta G^{+}_{\text{mech}}$, and the rest were fit with gamma distributions. All distributions are fit using the SciPy package to determine shape, scale, and location parameters [231]. Quantities were made positive or negative in order to produce the best fits.

Aggregating each of the 2000 s of the 10 trajectories into a collective data set for each condition of C_M and C_{CL}, we next analyzed the distribution of the instantaneous rates of $\Delta G^{+}_{\text{dissipated}}$, $\Delta G^{+}_{\text{mech, dissipated}}$, $\Delta G^{+}_{\text{chem, dissipated}}$, $\Delta G^{+}_{\text{stress}}$, $\Delta G^{+}_{\text{mech}}$, and $\Delta G^{+}_{\text{chem}}$. In Figure 4.5 we plot histograms of these 6 quantities for the conditions $C_{CL} = 1.0\ \mu M$, $C_{CL} = 0.2\ \mu M$. In the supporting information of Ref. 195 the same plots for other conditions are displayed.

Each distribution contained heavy tails which, for the purpose of fitting, we suppressed by excluding data lying outside 5 MAD's of the median. No distributions were sufficient to cleanly fit the full set of data, so we focus here on the center of the distribution and include a qualitative discussion of the heavy tails below. With the exception of $\Delta G^{+}_{\text{mech}}$ which was fit with a generalized normal distribution, each distribution exhibited significant skew and could be fit reasonably well with a log-normal or a gamma distribution. At higher concentrations of myosin and cross-linkers, the histograms were less cleanly fit by any standard distributions (see the supporting information of Ref. 195).

Log-normal distributions are fairly ubiquitous across different fields and systems [232]. For instance, it has been shown that the distributions of concentrations of species in a chemical reaction network are log-normal [233]. Gamma distributions are similarly common and often difficult to discriminate from log-normal distributions [234]. A speculative explanation for the gamma distribution of ΔG_{stress} is as follows: ΔG_{stress} can be viewed as resulting from a number of myosin filament steps that, according to the central limit theorem, is approximately normally distributed given a sufficiently long time between simulation snapshots t_{snap} (these stepping events are not truly i.i.d., but to a first approximation we may assume they are). The main effect of each of these steps is to increase the harmonic stretching potential on the myosin filaments as well as on the actin filament cylinders by a roughly fixed stepping distance. Thus the increase in mechanical energy, ΔG_{stress}, is approximately a quadratic function of the normally distributed number of myosin filament steps per t_{snap}. As shown in Figure 4.6, the resulting distribution of ΔG_{stress} is well-fit by a gamma distribution and bears qualitative similarity to the histogram of ΔG_{stress} in Figure 4.5. In Figure 4.5, ΔG_{stress} and $\Delta G^{+}_{\text{mech, dissipated}}$ have similar distributions, indicating that, for each MEDYAN cycle, almost all the stress accumulated following Step 1 is then immediately relaxed. The remainder goes into $\Delta G^{+}_{\text{mech}}$, whose distribution is centered on zero with little skew. Furthermore, the distribution of $\Delta G^{+}_{\text{mech, dissipated}}$ has particularly heavy tails, indicating infrequent yet large relaxation events. This tendency has some precedent in avalanche-prone systems, whose hallmarks are self-organized criticality, intermittency, and scale-invariance in their distribution of avalanche event sizes [161, 235–237]. While we do not observe true scale-invariance in this set of simulations (i.e. power-laws cannot fit these distributions cleanly), it will be interesting to continue to explore actomyosin networks dynamics in the framework of self-organized criticality [166].

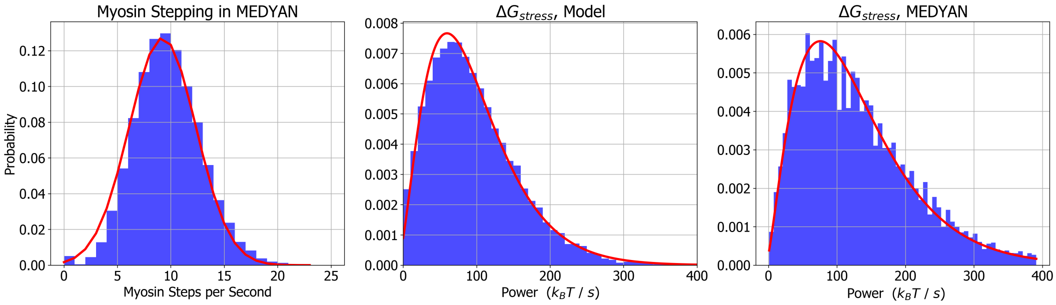

FIGURE 4.6: A simple model to rationalize the distribution for the ΔG_{stress} rate at for the condition of $C_{CL} = 0.1\ \mu M$, $C_M = 0.1\ \mu M$. *Left*: Histogram of the number of myosin filament steps per second measured over a set of 10 trajectories of 2000 s each. The fitted curve shows the normal approximation to this distribution. *Middle*: The fitted normal distribution approximating the number of myosin filament steps per second ($\mu = 9.22$, $\sigma = 3.13$) is sampled from 20,000 times. Each sample of the number of myosin steps per second is squared and multiplied by a constant $\alpha \sim 1$, giving a sample of ΔG_{stress} for a one second window. This procedure models the increase in mechanical stress by a spring that increases its length beyond equilibrium by a normally distributed number of fixed-length steps each second. The resulting histogram is well fit by a gamma distribution. *Right*: For comparison, the measured histogram of ΔG_{stress} is shown. This drastically oversimplified model of stress generation in actomyosin networks thus captures qualitative features describing the shape of this distribution.

4.3.2 More compact networks are more efficient

We simultaneously varied concentrations of myosin filaments C_M, and cross-linkers, C_{CL}, which is known to produce a range of network architectures [42, 238]. Using cross-linker concentrations of 0.1, 1.0, and 5.0 μM, and myosin concentrations of 0.1, 0.2, and 0.4 μM, we studied the effects on $\Delta G^{+}_{\text{dissipated}}$, $\Delta G^{+}_{\text{chem, dissipated}}$, and $\Delta G^{+}_{\text{mech, dissipated}}$. The concentrations of actin subunits and filaments are the same as described above. For each of the 9 conditions, we ran 10 simulations of 2000 s. In Figure 4.7, we display the median values of $\Delta G^{+}_{\text{chem, dissipated}}$ and $\Delta G^{+}_{\text{mech, dissipated}}$ along each trajectory and over each repeated trajectory for these conditions. The median is used here because of its insensitivity to outliers such as are found in the heavy-tailed distributions of these quantities [229].

We find that as C_M is increased, the median rates of $\Delta G^{+}_{\text{dissipated}}$, $\Delta G^{+}_{\text{chem, dissipated}}$, and $\Delta G^{+}_{\text{mech, dissipated}}$ all tend to increase. Further, the value of of $\Delta G^{+}_{\text{mech, dissipated}}$ relative to $\Delta G^{+}_{\text{chem, dissipated}}$ increases. As C_{CL} is increased with C_M fixed, then $\Delta G^{+}_{\text{chem, dissipated}}$ tends to decrease. Increased concentrations of myosin filaments obviously have a strong effect on the dissipation simply because they are the chief active agents in the system once treadmilling has reached its steady state. We can define a measure of efficiency for the present purpose as:

$$\eta = \frac{\Delta G_{\text{stress}}}{\Delta G^{+}_{\text{chem}}} = \frac{\Delta G^{+}_{\text{mech, dissipated}} - \Delta G^{+}_{\text{mech}}}{\Delta G^{+}_{\text{dissipated}} - \Delta G^{+}_{\text{mech}}} \approx \frac{\Delta G^{+}_{\text{mech, dissipated}}}{\Delta G^{+}_{\text{dissipated}}} \tag{4.6}$$

where the approximation follows since, as illustrated in Figure 4.2, the average of $\Delta G^{+}_{\text{mech}}$ is close to zero. Thus it is evident that, perhaps surprisingly, as C_M increases, η increases: the more motors are in our system, the more efficiently can chemical energy be converted into mechanical stresses. Further, as C_{CL} increases, η tends to increase because $\Delta G_{\text{chem, dissipated}}$ decreases relative to $\Delta G_{\text{mech, dissipated}}$. At higher levels of cross-linking,

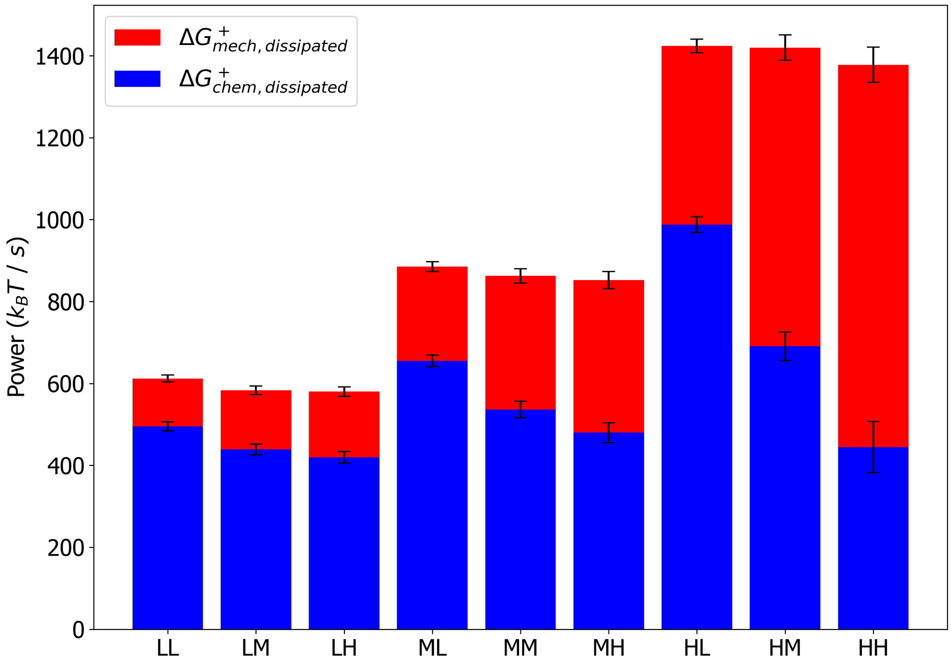

FIGURE 4.7: Bar plot representing the contributions of $\Delta G^+_{\text{chem, dissipated}}$ and $\Delta G^+_{\text{mech, dissipated}}$ to the total, $\Delta G^+_{\text{dissipated}}$. The letters in the abscissa labels designate "low," "medium," and "high." The first letter represents the concentration of myosin, C_M: "L" $= 0.1$ μM, "M" $= 0.2$ μM, "H" $= 0.4$ μM, and the second letter represents the concentration of crosslinkers, C_{CL}: "L" $= 0.1$ μM, "M" $= 1.0$ μM, "H" $= 5.0$ μM. The median of each quantity is taken over 10 runs of 2000 s, and error bars represent 1 MAD.

which introduce mechanical constraints on the filaments, the walking of myosin motor filaments will produce more stress compared to at low levels of cross-linking, when filament sliding can result from myosin filament walking, producing less stress.

We repeated the above experiments with the inclusion of mechanochemical feedback on the reaction rates controlling cross-linker unbinding and myosin filament unbinding and walking (see the supporting information of Ref. 195). We observed, somewhat surprisingly, no qualitative differences compared to the results described for the case of no feedback, however there were quantitative differences in the stress production and radius of gyration for certain conditions. These quantitative differences result from the fact that, for the concentrations C_M and C_{CL} used above, we did not observe significant collapse of the actomyosin network when mechanochemical feedback was included due to the stalling and catching of the myosin filaments. As a result the networks were less densely cross-linked, and the efficiency was lower. The total dissipation rates are largely unaffected by the inclusion of mechanochemical feedback; it was primarily the degree to which chemical energy had been converted to stress that was different for certain conditions.

It is worthwhile to mention how these results compare with *in vitro* studies of dissipation in actomyosin systems. While the computational approach described in this chapter allows uniquely highly-resolved and direct measurement of free energy changes, other experimental methods have produced qualitatively similar results to those obtained here. Rheological experiments have determined that a single cell's response to compression is similar in nature to a muscle's response to increasing load, suggesting that the actomyosin network underlies the cell's mechanical responsiveness [224]. Further, this responsiveness is modulated by blebbistatin, a myosin ATPase inhibitor, highlighting myosin's role in negotiating how the network rearranges in response to the sustained stress. Using the metric of mechanical dissipation to measure the degree of structural rearrangements and release of stress, we confirm that these processes indeed sensitively

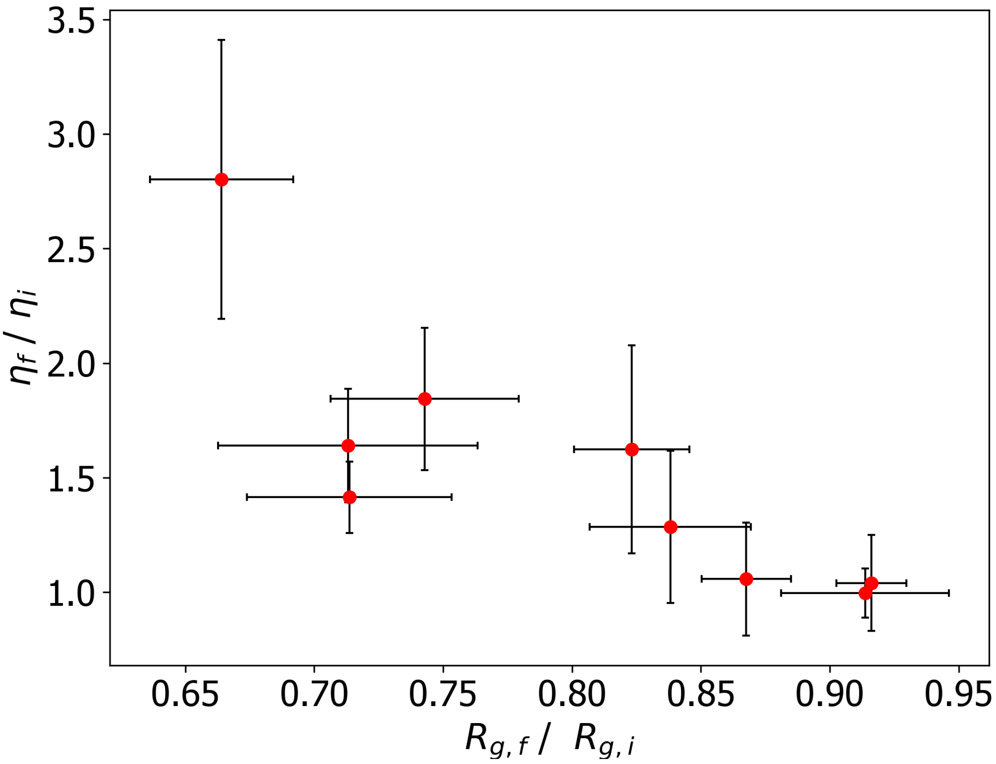

Figure 4.8: Networks which tend to collapse also tend to increase their ability to transduce chemical energy into mechanical stress. The change in the radius of gyration averaged over the last 100 s of a 2000 s simulation, $R_{g,f}$ divided by the radius of gyration averaged over the initial 100 s, $R_{g,i}$ measures the degree of network collapse over the trajectory. The same ratio is formed for the efficiency η, by taking the median over the same 100 s windows. These quantities are measured for each of the 9 conditions, and are averaged over the 10 repetitions of each. The error bars indicate the standard deviations over these sets of 10 measurements. The radius of gyration is defined as $R_g = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - r_{GC})^2}$, where r_{GC} is the geometric center of the actin network and r_i is the position of the i^{th} bead in the network. The efficiency is defined as $\eta = \frac{\Delta G_{\text{stress}}}{\Delta G^{+}_{\text{chem}}}$.

depend on myosin activity, which we control here through its concentration. Additional rheological studies have probed the mechanical dissipation of actin cortices more directly, using the loss modulus as a readout. These have also indicated that inhibiting myosin activity reduces the mechanical dissipation of the system, causing it to be behave more elastically [223]. Lastly, we mention a recent study that quantified mechanical dissipation of actin filaments using a novel experimental method [34]. By measuring the flow through a low-dimensional phase space defined by the amplitudes of the filament bending modes [32], they determine the entropy production of fluctuating actin filaments in different phases of contractility. They relate the entropy production to the degree of transverse bending of the filaments, as opposed to sarcomeric filament sliding, caused by myosin filament walking. Similarly, we here relate the hindrance of filament sliding due to cross-linker density to increased mechanical dissipation rates. Quantitative comparisons of these two approaches to mechanical dissipation measurement will be an interesting future direction. We note finally that a unique capability of quantifying dissipation using MEDYAN is the ability to simultaneously measure the energetics of chemical reactions in addition to the changes in the mechanical energy, which is not currently available using *in vitro* methods.

4.4 Discussion

We have introduced a methodology for tracking the energetics of chemical and mechanical events during a MEDYAN simulation, allowing us to probe the properties of actomyosin networks as dissipative active matter systems. The distinction between dissipation's mechanical and chemical origins is natural in the context of MEDYAN's iterative simulation procedure which carries out chemical stochastic simulation, mechanical deformation, and mechanical relaxation at separate times. As explained in Ref 42, this procedure exploits a separation of timescales between the characteristic mechanical relaxation times of actomyosin networks [239] compared to typical waiting time between reactions that introduce mechanical stresses, such as myosin filament walking [88] or filament growth [12]. Ultimately the source of all dissipation is the chemical potential of ATP molecules driving treadmilling and myosin filament activity. This is reflected by the near equality of $\Delta G^{+}_{\text{chem}}$ and $\Delta G^{+}_{\text{dissipated}}$ in Figure 4.2. On a fast timescale (that of the characteristic mechanical relaxation time), however, the free energy of chemical reactions cause small force deformations of the network which are then quickly and almost fully relaxed. Thus, for a myosin filament stepping event, only a portion of the chemical free energy $\Delta G^{+}_{\text{chem, dissipated}}$ is immediately dissipated as heat, with the rest going into temporarily increased mechanical energy of the actomyosin network, ΔG_{stress}. The fast relaxation of this new mechanical stress constitutes what we refer to as mechanical dissipation, $\Delta G^{+}_{\text{mech, dissipated}}$, and the small residual stress after this relaxation has balanced all net forces acting on the system results in a change of the mechanical energy of the system on a slow timescale, $\Delta G^{+}_{\text{mech}}$.

One interpretive framework which is useful to understand the flow of free energy in actomyosin networks is illustrated in Figure 4.9. We can think of the different forms of free energy storage, including chemical potentials, mechanical stress, concentration gradients (which could also be considered as arising from chemical potential differences across compartments), and dissipated energy, as nodes on a directed graph, where edges represent transduction of energy from one form of storage to another. The weights of these edges represent the amount of free energy flowing through them. In this picture, we can describe the process of network percolation, which occurs at increasing concentrations of cross-linkers and myosin filaments, as widening the edge flowing from chemical potential into mechanical stress, while thinning the edge from chemical potential directly to dissipation. The edge weights corresponding to the establishment of concentration gradients and the resulting diffusive dissipation will not be affected dramatically by the onset of percolation except to the extent that percolated networks might lead to the formation of more bundles, and therefore more significant concentration gradients. At steady state, we have a stationary current on the graph, fueled by the chemical potential of the assumed limitless supply of ATP. Of course, at this level of description, we

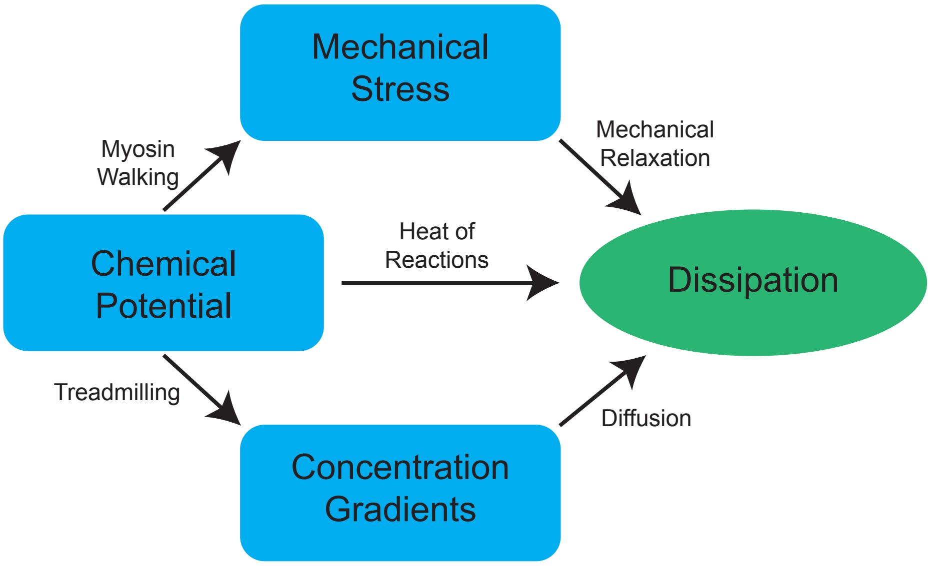

FIGURE 4.9: A simple schematic illustrating the flow of free energy in actomyosin network systems. Blue compartments represent forms of free energy storage, arrows represent the transfer of free energy from one form to another, and the green compartment indicates dissipation as ultimate destination of all free energy flows in this non-equilibrium system. Arrows are labeled with the mechanisms by which these free energy transduction from one form to another are achieved. In this depiction, the sizes of compartments and the widths of the arrows indicating the magnitudes of the represented quantities are not to scale.

have coarse-grained away the details of the specific chemical reaction networks and mechanical potentials which constitute the system, however by doing so we gain a clearer understanding of how percolation of the network alters the flows of free energy.

The new capabilities of MEDYAN allowing detailed energetics computations should provide a way to address outstanding issues in the the fields of active matter systems and of cell mechanics. In this study we observe dissipation rates to stay at a low fluctuating steady state following a high initial transient phase, apparently in contrast with the predictions of the dissipative adaptation hypothesis. However we also do not observe slow reorganization of the actomyosin network into different higher order structures such as bundles (for the concentrations of components used here we observe disordered networks only), and as a result we can not rule out that such reorganizations correspond to marked changes in the dissipation rates. A dedicated test of the hypothesis of dissipative adaptation is in principle possible using this methodology. For instance, a simulated gliding assay, which has been shown *in vitro* to lead to diverse dynamical patterns [67], may indicate that these emergent patterns correspond to an optimization of the chemical dissipation from myosin filament walking, as argued in Ref. 152. In the context of cell migration, studies investigating the mechanically dissipative activity of myosin filaments are also feasible if one incorporates in simulation an external substrate against which the actomyosin network pulls. In this type of study, it should be straightforward to determine how the amount of stress which is sustained against the substrate is altered by myosin activity given the energetics calculations in the methodology described here. It is also feasible to do simulated measurements of the dynamic shear modulus by compressing the actomyosin network at different frequencies. It should then be possible to directly observe the degree to which the dissipation of elastically stored energy is attributable to myosin walking. In fact, this last research question has already been investigated to some extent using an alternate computational model to the one presented here [240]. We hope some of these potential future experiments will shed light on outstanding questions in the studies of actomyosin networks.

Chapter 5

Graph representations, motor stalling, and volume scaling effects

5.1 Introduction

In Chapter 4, we studied dissipation in actomyosin networks across a range of cross-linker and motor concentrations which had a strong effect on the efficiency of chemical to mechanical energy transduction. An important parameter whose impact we have not yet discussed in much detail is the sensitivity of myosin motor stalling, which controls the coupling between the motors' ATP consumption rate and the tension sustained by the motors [87, 90]. Additionally, we have not yet discussed the role of the system size in determining dissipation rates [25]. For larger systems, it is useful to introduce a more detailed way to describe the structure of an actomyosin network than by, for example, the radius of gyration. One approach to describing structure is to develop a graph representation of the actomyosin network, allowing the use of metrics from the network science literature to describe cytoskeletal structure [241]. With these considerations in mind, the purpose of this chapter is three-fold: i) to describe a method for constructing graph objects from actomyosin networks; ii) to show how motor stalling in percolated networks leads to an appreciable decrease in the energy consumption rate; iii) to discuss how the graph objects and dissipation rates depend on the system volume. These issues are interrelated, as a key determinant of whether significant motor stalling occurs is the frustration of filament sliding caused by bound cross-linkers which is naturally characterized using metrics defined on the graph representation. The system boundary can also affect the graph metrics and dissipation rates, causing these quantities to scale inhomogeneously with the system volume.

5.2 Methods

5.2.1 Graph mapping

Recent decades have seen an explosion in the development of network science, partly motivated by an increasing awareness of the prevalence of graph objects occurring throughout the natural sciences [241–243]. These include the Internet, social networks, gene regulatory networks, and many others [244–246]. Network science offers a way to summarize key features of these different complex systems in a unified way, by abstracting away the specific nature of the interaction between components to focus instead on how the presence of these interactions (called edges) is distributed among the components (called nodes). To use network science to describe actomyosin networks, we first need a protocol which takes the actomyosin network's configuration (i.e. positions and binding information of the filaments and associated proteins) and produces a graph object $\mathcal{G}(\mathcal{E}, \mathcal{V})$ (i.e. a collection of nodes $\mathcal{V}$ and edges $\mathcal{E}$). Any such protocol can be viewed as an onto but not one-to-one (and hence not invertible) mapping, as we are inevitably discarding some information from the full configuration in producing the graph representation. This implies that the mapping is not unique, and indeed we can construct several different protocols which capture different types of information about the full configuration in the graph representation, as discussed below. In the mapping we use here, the nodes $\mathcal{V} = \{v_i\}_{i=1}^{N_f}$ are the actin filaments (where N_f is the number of filaments), and they are connected by un-directed weighted edges $\mathcal{E} = \{e_{i,j}\}_{i,j=1}^{N_f}$ whose weights $E_{i,j} = E_{j,i}$ represent the number of passive cross-linkers (e.g. α-actinin molecules) bound to the pair of filaments i and j. By representing the pattern of cross-linker binding, this mapping captures the topological structure of the actomyosin network and contains information on the degree of filament frustration caused by mechanical constraints in the form of bound cross-linkers [104].

A ubiquitous higher-order structural feature of actomyosin networks is a bundle of several nearby and highly inter-connected (cross-linked) filaments [2, 247]. One could even interpret these bundles as being the elementary objects in a coarse-grained representation of the original actomyosin network. To express this idea in the language of network science, we can identify bundles as communities (i.e. highly inter-connected groups of nodes) in the graph representation [248, 249]. We can systematically determine these bundles using community detection methods such as the Girvan-Newman, Clauset-Newman-Moore, or Louvain algorithms [248, 250, 251]. These methods produce a partitioning of the nodes $\mathcal{C}_K = \{\{v_i\}_{i\in\mathcal{I}_k}\}_{k=1}^K$, where the K communities are defined as the sets of node indices $\mathcal{I}_k$. This is done in a hierarchical fashion over the range of K values from 1 to N_f, producing a dendrogram of partitions. Each $\mathcal{C}_K$ is chosen to approximately maximize over all choices of K partitions the modularity

$$\mathcal{Q}_K = \frac{1}{2m}\sum_{k=1}^{K}\sum_{i,j\in\mathcal{I}_k}\left(E_{i,j} - \frac{k_i k_j}{2m}\right) \tag{5.1}$$

where m is the sum of all edge weights in the network, and $k_i = \sum_j E_{i,j}$ is the total edge weight belonging to node i [249, 251]. The modularity reflects the amount of edge weight contained within the communities relative to the total edge weight, including edges between communities, in the graph. The value K^* is then chosen as that which gives the maximum modularity over all partitions the dendrogram, i.e. $K^* = \arg\max_K\{Q_K\}_{K=1}^{N_f}$. To implement community detection in our graphs representing actomyosin networks we use the Louvain algorithm, a fast method which performs greedy optimization to locally maximize Q_K over the K partitions.

With a partition $\mathcal{C}_{K^*}$ it becomes possible to create a coarse-grained network $\widetilde{\mathcal{G}}(\widetilde{\mathcal{E}}, \widetilde{\mathcal{V}})$ by collapsing the identified communities $\mathcal{I}_k$ into nodes $\widetilde{v}_k$, which now have node weights

$$\widetilde{V}_K = |\mathcal{I}_k| \tag{5.2}$$

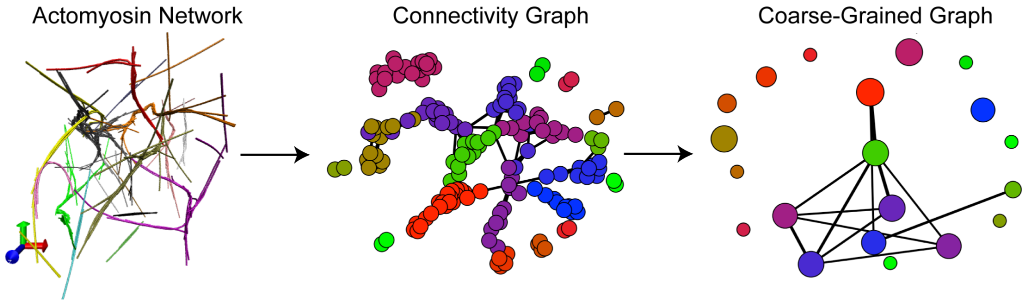

FIGURE 5.1: Schematic illustration of the pipeline mapping an actomyosin network to a graph $\mathcal{G}$, on which pruning, community detection, and coarse-graining is performed to produce a new network $\widetilde{\mathcal{G}}$. The color of the nodes in $\mathcal{G}$ indicates the node's community, and this color is presereved in the drawing of $\widetilde{\mathcal{G}}$. The colors of the filaments in the actomyosin networks also indicate the identified communities (bundles) of the filament, but these colors do not correspond to those in the drawings of $\mathcal{G}$ and $\widetilde{\mathcal{G}}$. To aid visibility of the bundles, a smaller actomyosin network is shown on the left than the one used to produce $\mathcal{G}$ and $\widetilde{\mathcal{G}}$.

given by the number of filaments in the bundle. The new edges $\widetilde{e}_{k,l}$ have weights $\widetilde{E}_{k,l}$ which sum over those between the members of the corresponding communities:

$$\widetilde{E}_{k,l} = \sum_{i \in \mathcal{I}_k} \sum_{j \in \mathcal{I}_l} E_{i,j}. \tag{5.3}$$

We can also optionally prune the networks $\mathcal{G}$ by removing edges with a weight below some threshold E_T, which we usually take to be 5. This excludes connections which correspond to very weakly cross-linked filaments and which aren't mechanically coupled in the way we consider bundled filaments to be. The pipeline summarizing the graph mapping methods described above is schematically illustrated in Figure 5.1.

5.2.2 Volume scaling study

To study the effect of volume size on the structural properties and dissipation rates of actomyosin networks, we held the protein concentrations fixed and varied the side length of a cubic system in increments of 0.5 μm from 0.5 μm to 5 μm. This corresponds to volumes ranging from 0.125 μm^3 to 125 μm^3, covering four orders of magnitude (Figure

5.2). We also tested two conditions of myosin stalling force F_{stall} (i.e. the force at which the motor heads stop walking; see Appendix A) and two concentrations of cross-linkers C_{CL}. The concentrations used are 12 μM for actin monomers, 0.02 μM for actin filaments, 0.015 μM for myosin minifilaments, and either 0.25 μM (low) or 5 μM (high) for α-actinin. With these concentrations of actin monomers and filaments the mean filament length is ~ 1.6 μm. The two stall forces tested are 15 pN (easily stalled) and 300 pN (hard to stall). We performed 7 trials for each of the 40 conditions of size, cross-linker concentration, and stall force. Larger systems take longer to run and not all systems completed the full 1000 s of requested simulation time, but interesting scaling behaviors can still be observed.

To aid computational efficiency, we increased the cylinder size used to discretize the filament backbones from 27 nm, used for system side lengths up to 1.5 μm, to 108 nm, used for system side lengths 2.0 μm and greater. This switch occurs when the side length exceeds the mean filament length. Because of this, the detailed mechanical properties of the networks are not comparable across the system sizes, and we instead focus on topological and chemical scaling behaviors. In our investigations of the anomalous mechanical energy fluctuations described in the Chapter 6 we perform a similar scaling study over a smaller range of sizes, and in that study we do control for the cylinder size. We also note for the remainder of the thesis the model of filament treadmilling is simplified and does not distinguish between different nucleotide-bound actin subunits. This aids in computational efficiency, and the mean-field model of treadmilling dissipation described in earlier chapters already provides an essentially complete understanding. The (de)polymerization rate constants of this simplified model, along with the other simulation parameters used in this chapter, are displayed in Table E.1.

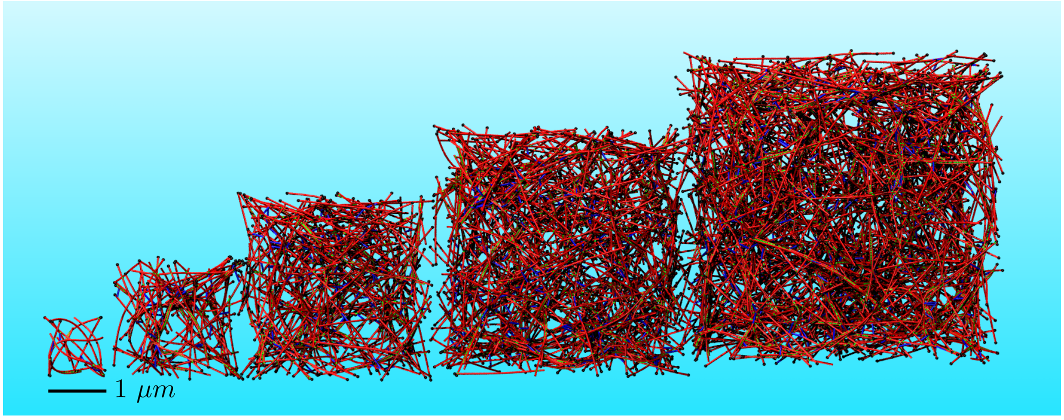

Figure 5.2: Visualization of 5 actomyosin networks with cubic boundaries whose side lengths range from 1 to 5 μm in increments of 1 μm. 5 additional network sizes than visualized here with non-integer side lengths were also used in the volume scaling study.

5.3 Results

5.3.1 Self-organization produces percolated, coalesced networks

There are many available metrics in the network science literature which can be used to describe various properties of graphs. Here we report on just a few such metrics which display interesting scaling behavior with the system volume, however the mapping described above opens the door for a much broader array of network features to be used to characterize actomyosin network structure.

We first show in Figure 5.3 the degree of network percolation, as quantified by the proportion γ of total filaments belonging to the largest connected component of the unpruned graph $\mathcal{G}$. The trajectory of $\gamma(t)$ is a sigmoidal curve across the range of conditions tested, however $\gamma(t)$ only saturates at 1, indicating full network percolation, when the concentration of cross-linkers is high. Furthermore, for low concentration of cross-linkers $\gamma(t)$ robustly saturates around 0.9 only for system volumes above 3.375 μm^3. Below this volume the fluctuations appear to more significantly affect the trajectory which does not reach a definite steady state value. We elaborate on the scaling of the fluctuations as a function of system size below, when we discuss fluctuations of the dissipation rate. We note that this separation in behavior at 3.375 μm^3 could be due to the change in the cylinder length, but this it not likely to be the case since for other conditions and other reported trends we do not see a similar separation at this volume. Finally, we observe that there is minimal dependence of $\gamma(t)$ on the stall-force of myosin.

Next we show that the number of identified communities normalized by the system size, or the bundle density ρ_b, appears to reach a limiting value for large system sizes at steady state (Figure 5.4). We refer to the aggregation of filaments into a small set of bundles as coalescence of the network. As an intensive quantity, the density of bundles should not depend on the system size as long as the system is large enough that boundary effects become negligible. It should depend on protein concentrations, and we observe the limiting density to be greater for the low cross-linker concentration than for the high cross-linker concentration. This is expected since the filaments are less able to coalesce into bundles at low cross-linker concentrations. The trajectories of $\rho_b(t)$ for large systems are non-monotonic, with an initial increase to a maximum ρ_b^* during the initial polymerization of seed filaments and first cross-linking events based on random filament proximity. The trajectories $\rho_b(t)$ agree for all volumes up to this point, after which they separate. Larger systems show a decrease after this due to the coalescing of bundles together into a smaller number of communities. We conjecture that for the smallest system size this coalescence does not occur (and $\rho_b(t)$ continues to increase) because too much of the actin mass is at the simulation boundary, where it cannot interact with enough other filaments to form bundles. We also observe that for high concentration of cross-linkers this coalescence process occurs at a faster rate compared to low concentration. Finally, we also observe minimal dependence of these trends on the stall force of myosin.

We next discuss the mean node connectivity of the coarse-grained graph $\widetilde{\mathcal{G}}$, which has an interesting but slightly ambiguous scaling behavior (Figure 5.5). Given a pair of nodes v_i and v_j in a graph $\mathcal{G}$, the connectivity between these nodes is defined as

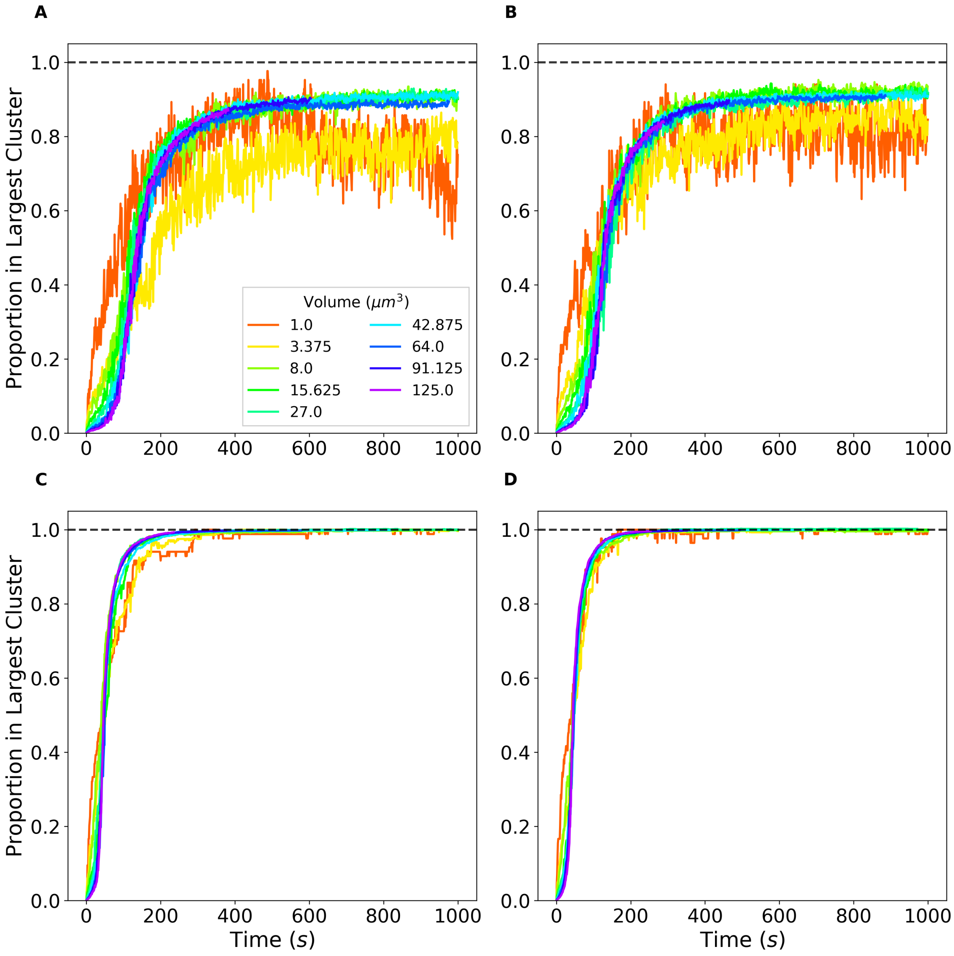

Figure 5.3: Plots of the proportion $\gamma(t)$ of total filaments belonging to the largest connected component of the graph $\mathcal{G}(t)$. A) Averages of trajectories over 7 trials of each system size are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 0.25\ \mu M$. The uncertainty in these trajectories is not visualized to aid visibility. The smallest volume $0.125\ \mu m^3$ is not visualized as its fluctuations are so large that it impedes visibility of the other trajectories. B) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 0.25\ \mu M$. C) Averages of trajectories for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. C) Averages of trajectories are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. D) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 5\ \mu M$.

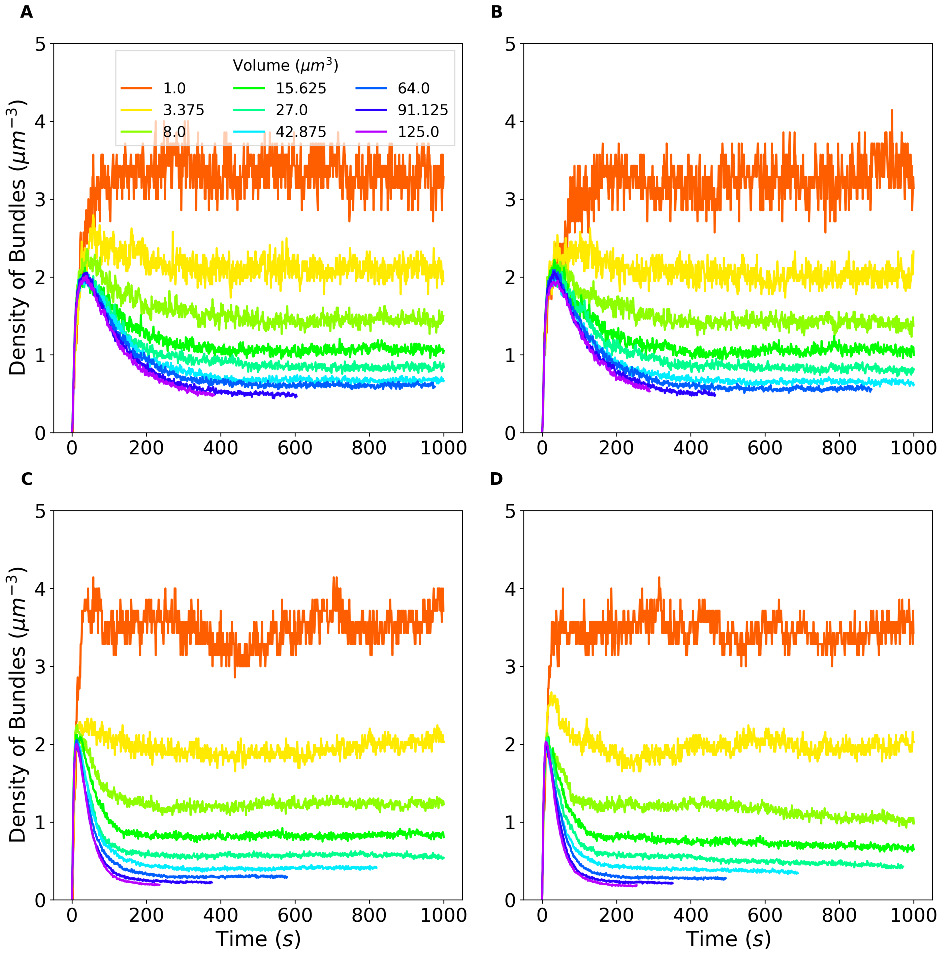

FIGURE 5.4: Plots of the density of bundles $\rho_b(t)$. A) Averages of trajectories over 7 trials of each system size are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 0.25\ \mu M$ (c.f. Figure 5.3). B) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 0.25\ \mu M$. C) Averages of trajectories for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. C) Averages of trajectories are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. D) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 5\ \mu M$.

the minimum number of edges that is necessary to delete in order to remove all paths connecting the pair, not accounting for edge weights. The mean node connectivity ϕ is defined as the average of the node connectivity over all pairs of nodes in $\mathcal{G}$. In the context of actomyosin networks, ϕ can be interpreted as the number of force chains connecting the typical pair of filaments or bundles. It is an expensive metric to calculate for large graphs, but for the coarse-grained graphs $\widetilde{\mathcal{G}}$ it is easy to determine and informs on the number of redundant paths at the level of filament bundles. $\phi(t)$ is generally observed to monotonically increase and then plateau for all conditions and sizes. For an un-percolated network, we expect that as the network size increases ϕ obtains a limiting value related to the typical size of the connected graph components. This appears to be borne out in the data for the low cross-linker concentration conditions. For a percolated network, we expect that ϕ continues to increase as some function of the network size, since additional nodes added to the percolated network create additional paths contributing to ϕ. This expected behavior is observed up at least up to intermediate system sizes, after which it is unclear whether the trajectories $\phi(t)$ begin to collapse on one another or continue to increase with system size. The limiting value of $\phi(t)$ for percolated networks thus appears to have a non-trivial dependence on the system size, but our investigations did not find a definite empirical form for this scaling that holds for all sizes. This issue could benefit from additional future research. Finally, we note again that there is minimal observed dependence of these trends on the stall force of myosin motors.

5.3.2 Dissipation rates are extensive and depend on motor stalling

In previous chapters we introduced how the dissipation rate $\dot{D} \equiv \partial_t \Delta G_{\text{dissipated}}(t)$ is dynamically calculated in simulation. Here we discuss how these rates depend on the system size and the experimental conditions of stall force and cross-linker concentration.

We first show that the dissipation rate behaves like a first-order homogeneous function of the system size, even for small systems. This means that $\dot{D}(\lambda V) = \lambda \dot{D}(V)$ for any positive scaling parameter λ, i.e. $\dot{D} \propto V$. The linear dependence on system size is

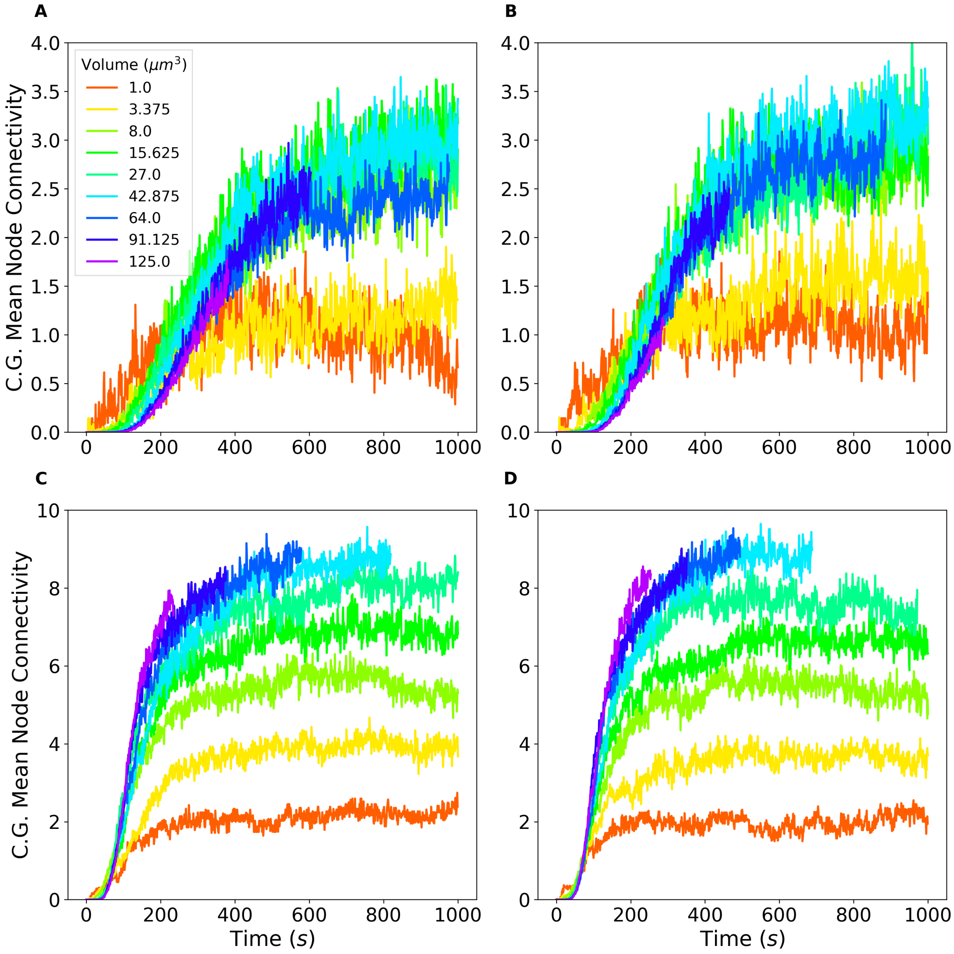

FIGURE 5.5: Plots of the mean node connectivity $\phi(t)$ for the coarse-grained graph $\widetilde{\mathcal{G}}$. A) Averages of trajectories over 7 trials of each system size are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 0.25\ \mu M$ (c.f. Figure 5.3). B) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 0.25\ \mu M$. C) Averages of trajectories for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. C) Averages of trajectories are shown for $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$. D) Averages of trajectories are shown for $F_{\text{stall}} = 300\ pN$ and $C_{CL} = 5\ \mu M$.

illustrated in Figure 5.6 for data collected across all system sizes and averaged over the time interval from 200 to 250 s and across the 7 trials. The standard deviation $\sigma_{\dot{D}}$ pooled over the trials is displayed in the inset of Figure 5.6. We find that $\sigma_{\dot{D}}$ scales with the square root of the system size, i.e. $\sigma_{\dot{D}} \propto V^{1/2}$. This behavior is familiar from the theory of equilibrium statistical mechanics. For example, the relative fluctuations σ_E/N of the system energy E for an equilibrium canonical ensemble scales with the number of particles N like $N^{-1/2}$, so that $\sigma_E \propto N^{1/2}$ [132, 163]. It is interesting to find that the same scaling is observed for the non-equilibrium thermodynamic quantity $\dot{D}$.

$\dot{D}$ is an extensive quantity, and we could consider a corresponding density $\dot{d} \equiv \dot{D}/V$. For this intensive quantity, we expect the fluctuations to decrease with the system size, since $\sigma_{\dot{d}} = \sigma_{\dot{D}}/V \propto V^{-1/2}$. For intensive quantities such as the energy consumption rate density (Figure 5.8), bundle density (5.4), and the proportion of filaments in the connected cluster (Figure 5.1), we indeed find the $V^{-1/2}$ scaling of the fluctuations to be approximately obeyed (data not shown).

It is evident from Figure 5.6 that the condition corresponding to $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 0.25\ \mu M$ has a lower mean dissipation rate than the other conditions. We next show that this results from the stalling of the myosin motors in percolated networks. In Figure 5.7 we display the trajectories of chemical energy consumption by myosin motors for the 27 μm^3 system, averaged over the 7 trials. The chemical energy consumption by motors directly depends on their walking rate, which is slowed for motors with low stall force under large tensions. Clearly, the low stall force causes the energy consumption rate to decrease as the network self-organizes for this condition. Interestingly, the low stall force alone is not sufficient to produce a decrease in the energy consumption rate, as filament frustration from bound a high concentration of cross-linkers is also needed to cause the typical tension experienced by motors to be large enough to cause stalling. That the energy consumption rate only appears to decrease implies that the long-time limit behavior corresponds to motors that tend to be maximally stalled for the given

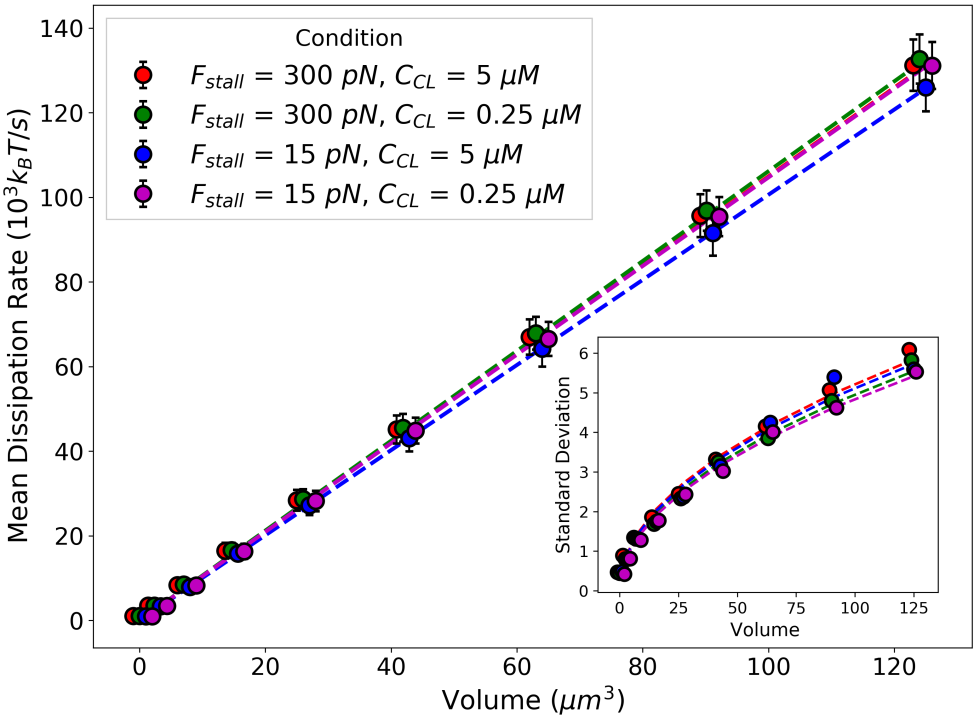

FIGURE 5.6: The dissipation rate $\dot{D}$ is shown as a function of the system size and across the four experimental conditions, averaged as described in the main text. The error bars show the standard deviation. A small horizontal offset is added to the scatter plot data to aid visibility. Functions of the form aV are fit to determine a for each condition and shown as dotted lines. The inset shows the standard deviations as a function of the system size for each condition. Functions of the form $aV^{1/2}$ are fit to determine a for each condition and shown as dotted lines.

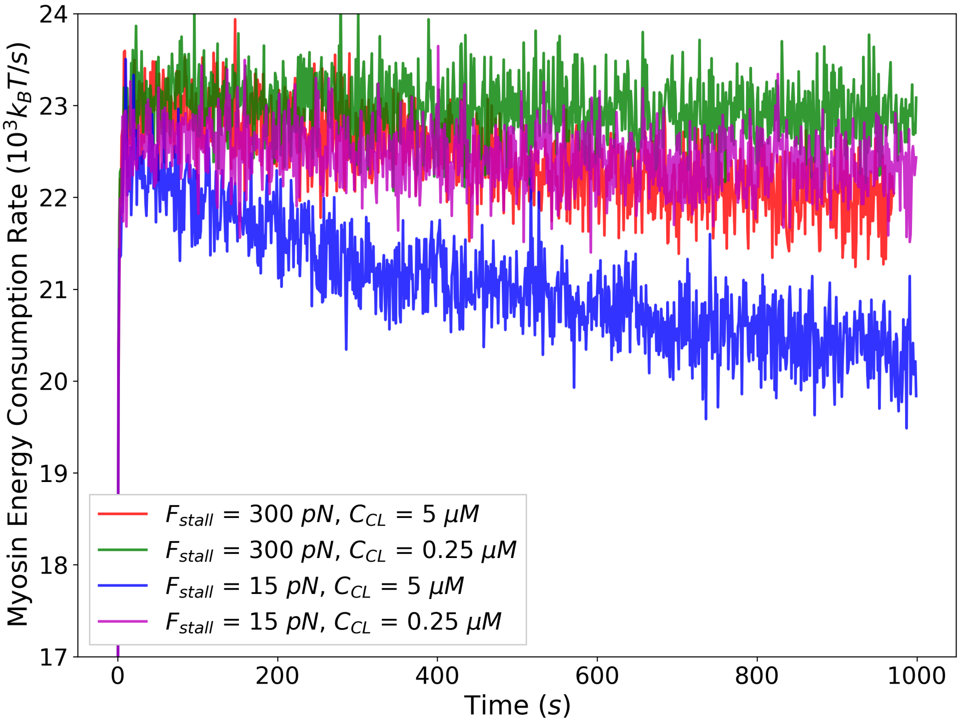

FIGURE 5.7: The chemical energy consumption rate by myosin motors is shown for the $27\mu m^3$ system and for the four experimental conditions, averaged over the 7 trials. The uncertainty in these trajectories is not substantial for this system size and is omitted to aid visibility.

conditions. Finally, we show in Figure 5.8 how this trend depends on the system size by plotting the chemical energy consumption rate normalized by the system volume for the condition mentioned above. All trajectories of this rate collapse onto a single curve, indicating that the system volume plays no appreciable role in the stalling-induced decrease in the energy consumption rate.

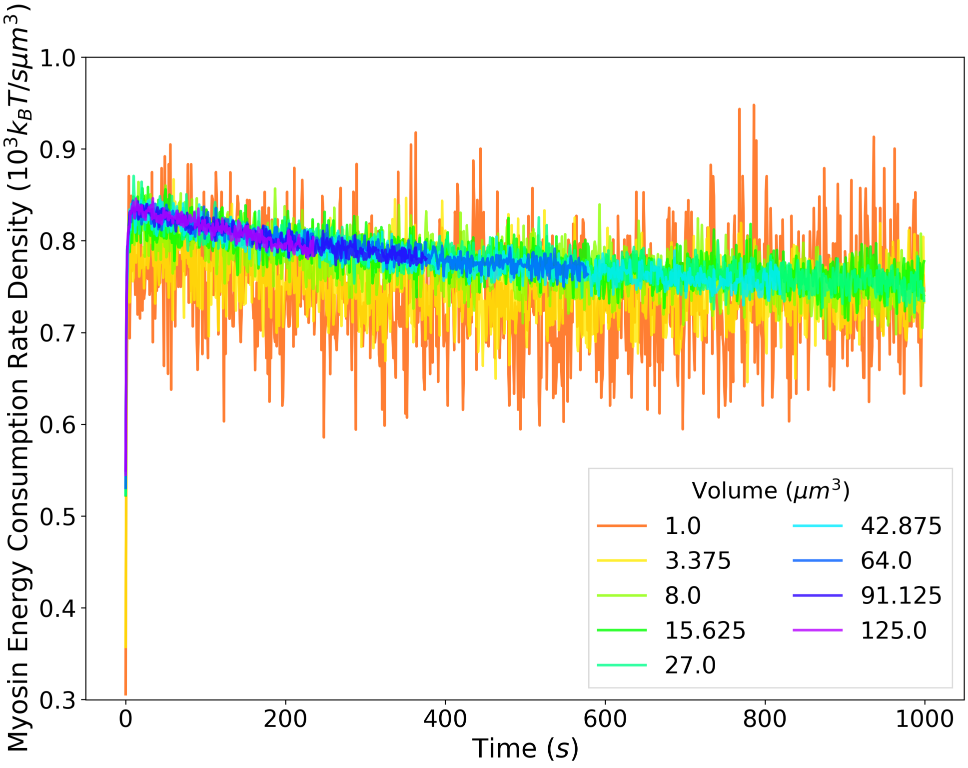

FIGURE 5.8: The density of the chemical energy consumption rate by myosin motors is shown for the condition $F_{\text{stall}} = 15\ pN$ and $C_{CL} = 5\ \mu M$ as system size is varied, averaged over the 7 trials (c.f. Figure 5.3).

5.4 Discussion

The graph mapping protocol we have introduced can be placed in the context of other possible mappings given the configuration of an actomyosin network. We briefly list some other mappings:

1. *Nodes*: Filament plus or minus ends. *Edges*: Binary, representing whether nodes are less than a distance cutoff from each other other.
2. *Nodes*: Filaments. *Edges*: Weighted, representing the total tension magnitudes of cross-linkers or motors connecting the nodes.
3. *Nodes*: Myosin motors. *Edges*: Binary, representing whether nodes are less than a distance cutoff from each other other.
4. *Nodes*: Filaments. *Edges*: Binary, representing whether the minimal distance along the filament lengths is less than a distance cutoff [252].

The choice of a mapping reflects which property of the full configuration one wishes to study. For instance mapping 1) allows studying the co-localization of filament end points during self-organization, whereas mapping 2) allows to study more quantitatively the force chains spreading throughout the network [104, 195]. The mapping discussed used this chapter is useful to study the frustration of filaments due to the mechanical constraints from bound cross-linkers. Thus, graph representations offer a flexible approach to describe abstract features of cytoskeletal networks and may be of use more broadly in the study of soft-matter systems. One important consideration is the amount of information available when constructing a graph representation. For instance, experimental methods using fluorescence-based imaging will not be able to easily determine the binding information of all cross-linking molecules. Such methods could determine the proximity of various elements, though, allowing the use of mappings 1), 3), and 4).

Our volume scaling study was inspired by the powerful analysis tool of finite-size scaling, which has been used to study the near-critical behavior of lattice-based statistical systems [253, 254]. However, it is important to note that a 125 μm^3 cubic box of actomyosin is of interest in more of a theoretical sense rather than an applied biological one. In real cells, most of the actomyosin mass is concentrated at the cell cortex, a shell of thickness $0.1 - 1$ μm that sits underneath the cell membrane [255]. If we assume that a typical cell is a spherical cow with a diameter of 20 μm, then the volume of actomyosin is in the range $124 - 1135$ μm^3, with an outer and inner surface area of 1256 μm^2 and 1231 μm^2 respectively. [209]. The surface area-to-volume ratio for the cortex is therefore $\sim 2 - 20$ μm^{-1}, whereas for the cubic system it is 1.2 μm^{-1}. In our volume scaling study we considered the extrapolation to an infinite bulk system, for instance to argue that density of bundles approaches a fixed value (Figure 5.4). We argued that for small to intermediately sized systems the boundary plays an appreciable role in preventing the coalescence of bundles. We can therefore expect the boundary to play an important role in the structure of cell cortices and conjecture that cortical actomyosin is less coalesced than an equal amount of bulk actomyosin.

The observed dependence of the chemical energy consumption rate on motor stalling can be placed in the context of several developments in the study of active matter systems. First, a recent study has highlighted how the force-sensitive catch-bond behavior (in which unbinding is slower at high tension) of myosin motors in breaking the detailed balance of motor binding reactions [256]. This was then argued to give rise to network-level effects, such that highly force-sensitive motors produced a jammed, solid-like phase with small actin flow compared to the liquid-like phase produced by less force-sensitive or force-insensitive motors. This tendency of force-sensitivity to lead to a actomyosin state characterized by diminished motor activity and associated actin movement is also reflected in our results showing that myosin energy consumption steadily decreases for highly sensitive motor stalling. In contrast, however, we do not find that motor stalling sensitivity (which was not included in the study referenced above) is a strong determinant for the topology of cross-linker binding, as evidenced by the minimal change in the trends of the graph representation metrics. We note that in our study the catch-bond force-dependence is held fixed at an experimentally validated value and only the stalling sensitivity is varied. Second, a recent theory of active matter assemblies makes predictions for the typical steady state behavior purely in terms of kinetic factors [83, 156]. This “low rattling” theory formalizes the basic intuition that a system will tend to stay for a long time in a place that it is slow to leave from. Thus the typical non-equilibrium steady state configuration of a system should be the state that has the smallest associated kinetic factors leading the system out of it. The monotonic increase in cross-linker-induced filament frustration and concomitant motor stalling implies a decrease in the motor walking kinetics. Thus the eventual actomyosin state is one which the system will be the slowest to leave, at least via motor-induced transitions. This seemingly agrees with the low-rattling idea. Future work could further elucidate to what degree actomyosin dynamics can be understood in the context of this new theory.

Chapter 6

Understanding cytoskeletal avalanches using mechanical stability analysis

6.1 Introduction

The actin-based cytoskeleton is an active biopolymer network that plays a central role in cellular physiology, providing the cell with a means to control its shape and produce mechanical forces during processes such as migration and cytokinesis [2, 3, 257, 258]. These cellular-level forces arise from the collective non-equilibrium activity of molecular motors interacting with the actin filament scaffold, enabling dynamic, driven-dissipative cytoskeletal remodeling [106, 259, 260]. Recent experimental efforts have uncovered a remarkable phenomenon exhibited by cytoskeletal networks *in vivo*: these networks undergo large, sudden structural rearrangements significantly more frequently than predicted by a Gaussian distribution [168, 169]. Heavy-tailed distributions of event sizes are well-known in seismology, where the Gutenberg-Richter law describes the power-law relationship between the energy released by an earthquake and such an earthquake's frequency [162, 261]. Due to this analogy the term "cytoquake" which we adopt here, has been coined by experimenters to describe large cytoskeletal remodeling events. In previous work we have reported the first *in silico* observations of this phenomenon, appearing as heavy tails in the distributions of mechanical energy released by cytoskeletal networks [195]. These findings suggest that avalanche-like processes may play a fundamental role in cytoskeletal dynamics.

The physics underlying cytoquakes is not well understood, as current explanations based on experimental data are mostly speculative and rely on qualitative comparisons to systems amenable to computational study which similarly exhibit non-exponential relaxation, such as jammed granular packings and spin glasses [168, 169, 262, 263]. In addition, in previous studies little emphasis has been given to the possible biological roles played by cytoquakes. We propose one such role, that, as suggested by the fluctuation-dissipation theorem, these large mechanical fluctuations are concomitant with a large susceptibility, allowing the cytoskeleton to be highly sensitive to physiological cues arriving via various cell signaling pathways [264]. Dynamic instability is already an acknowledged feature of certain cytoskeletal components such as microtubules and filopodia [265]. A similar design principle may also apply to larger cytoskeletal structures to allow fast remodeling. For instance, avalanche-like dynamics may serve a useful purpose in the lamellipodia of migrating cells, which probe local chemical gradients and must quickly collapse protrusions in unsuccessful search directions [3]. However, to investigate such possible biological roles we first need a more detailed account of cytoquake physics, which is the subject of this chapter.

6.2 Methods

We study a subsystem of the full cytoskeleton called an actomyosin network. This consists of semi-flexible actin filaments and associated proteins, including active molecular motors (e.g. minifilaments of non-muscle myosin IIA) and passive cross-linkers (e.g. α-actinin). The actin filaments hydrolyze ATP molecules in a directed polymerization process which reaches a steady state called treadmilling. The myosin minifilaments ($\sim$200 nm in length) transiently bind to pairs of actin filaments and also hydrolyze ATP as fuel to walk along the filaments, generating motion and mechanical stresses. These active process drive the network away from equilibrium. The cross-linkers ($\sim$35 nm) bind more stably to nearby filaments, serving to transmit the force produced by motors and to both store and through unbinding dissipate the resulting energy, heating the environment [4, 71, 88, 93, 94, 170, 266]. Dissipation of stored mechanical energy also occurs as filaments relax out of strained configurations, in a manner which can depend on constraints from bound cross-linkers and motors. Additionally, the rates of motor walking and unbinding as well as of cross-linker unbinding depend exponentially on the forces sustained by these molecules, giving rise to nonlinear coupling between the mechanical state of the network and its chemical propensities [87, 267]. These means by which the ability of the network to mechanically relax depends on its current state set the stage for avalanche-like dynamics. An actomyosin network as represented in simulation is visualized in Figure 6.1.

To computationally study cytoskeletal networks at high spatio-temporal resolution, we use the simulation platform MEDYAN [42, 195, 247, 268, 269]. MEDYAN simulations combine stochastic chemical dynamics with a mechanical representation of filaments and associated proteins. Simulations proceed iteratively in a sequence of four steps: 1) stochastic chemical simulation for a time δt (here 0.05 s), 2) computation of the resulting new forces, 3) quasi-equilibration via minimization of the mechanical energy, and 4) updating of force-sensitive reaction rates. Extensions to the MEDYAN platform presented in Chapter 4 allow calculation of the change in the system's Gibbs free energy during each of these steps [195, 270], originally applied to study the thermodynamic efficiency of myosin motors in converting chemical free energy to mechanical energy under various conditions of cross-linker and motor concentration. We employ this methodology here and focus on the statistics of the system's mechanical energy U as it self-organizes.

We performed MEDYAN simulations of small cytoskeletal networks consisting of 50 actin filaments in 1 μm^3 cubic boxes with varying concentrations of α-actinin cross-linkers ($[\alpha]$) and of NMIIA minifilaments ($[M]$). MEDYAN simulations combine stochastic chemical dynamics with a mechanical representation of filaments and associated proteins. Simulations proceed iteratively in a sequence of four steps: 1) stochastic chemical

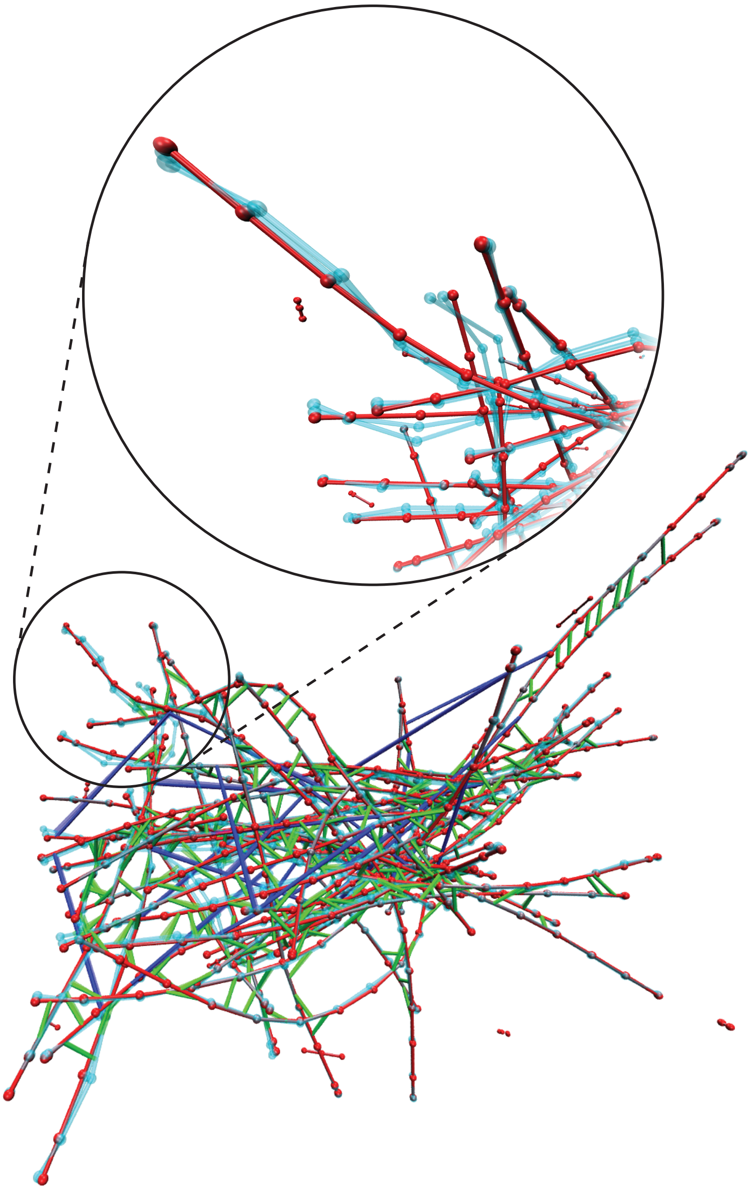

Figure 6.1: A snapshot from a MEDYAN trajectory of an actomyosin network in a 1 μm^3 box for the condition $C_{3,3}$. Actin filaments are shown in red, α-actinin is shown in green, and myosin motors are shown in blue. Beads representing the joined points (i.e. hinges) of thin cylinders (here 54 nm long) are visualized as red spheres. The cyan filaments represent motion of the network corresponding to a soft, delocalized vibrational mode determined from Hessian analysis, as described below. In the inset we zoom in on part of the network and exclude associated proteins to show greater detail of this vibrational motion.

simulation for a time δt (here 0.05 s), 2) computation of the resulting new forces, 3) equilibration via minimization of the mechanical energy, and 4) updating of force-sensitive reaction rates such the as slip-bonds of cross-linkers, catch-bonds of motors, and motor stalling. The boundaries of the box exert an exponentially repulsive force against the filaments with a short screening length of 2.7 nm. Five concentrations of α-actinin (ranging from 0.17 to 5.48 μM) and five concentrations of myosin minifilaments (ranging from 0.003 to 0.08 μM) were used with a constant G-actin monomer concentration of 13.3 μM, in the regime of physiological concentrations [209]. This led to a steady state filament length distribution with mean 0.48 μm and standard deviation 0.26 μm. We label these conditions $C_{i,j}$, where $i = 1, ..., 5$ represents the rank of the cross-linker concentration and $j = 1, ..., 5$ represents the rank of the myosin motor concentration. The length of the simulation cycle δt was chosen as 0.05 s, although we explore dependence on this parameter below (Figure 6.3). Five runs of each condition $C_{i,j}$ were simulated. We omit here other associated proteins, such as the branching agent Arp2/3, finding that our minimal system is sufficient to produce heavy-tailed distributions of event sizes, although it has recently been discovered that branching acts to enhance avalanche-like processes [271].

6.3 Results

6.3.1 Asymmetric heavy-tailed distributions and correlations of ΔU

We first characterize the observed occurrence of avalanche-like dynamics in these simulations. The simulations begin with short seed filaments that quickly polymerize (tens of seconds) to their steady state lengths. Following this, the slower process (hundreds of seconds) of dominantly myosin-driven self-organization occurs which for most conditions results in geometric contraction to a percolated network (see Movie 1) [94, 272]. The mechanical energy $U(t)$ fluctuates near a quasi-steady state (QSS) value, which we analyze as a stochastic process. In Figure 6.2.A we display the trajectory of $U(t)$ for condition $C_{3,3}$ (with α-actinin concentration $[\alpha] = 2.82$ μM, and motor concentration $[M] = 0.04$

μM. We tracked the net changes of the mechanical energy $\Delta U(t) = U(t + \delta t) - U(t)$ resulting from each complete cycle of simulation steps 1) - 4).[1] For the purpose of analyzing the observed asymmetric heavy tails in the distribution of ΔU, we treat the negative increments ΔU_- (energy release) and positive increments ΔU_+ (energy accumulation) as samples from separate distributions with semi-infinite domains. The complementary cumulative distribution functions (CCDFs or "tail distributions", the probability $P(X \geq x)$ of observing a value of the random variable X above a threshold x, as a function of x) of the observed samples collected from all five runs at QSS are illustrated in Figure 6.2.B. Both distributions display striking heavy tails relative to a fitted half-normal distribution. The CCDFs are better fit by stretched exponential (Weibull) functions of the form $P(X \geq x) = e^{-((x/\lambda)^k)}$ [273]. We justify this choice of distribution using Weibull plots, as discussed in Appendix E. We find $k = 0.60 \pm 0.06$ for $|\Delta U_-|$ and $k = 0.83 \pm 0.07$ for ΔU_+ with uncertainty taken over the five runs, indicating shallower tails for energy release compared to energy accumulation. We also measured the non-Gaussian parameter $\eta = \frac{\langle x^4 \rangle}{3\langle x^2 \rangle^2} - 1$, where $\langle x^m \rangle$ is the m^{th} moment about zero; for a half-normal distribution $\eta = 0$, and $\eta > 0$ quantifies heavy-tailedness. We find $\eta = 11.37 \pm 5.37$ for $|\Delta U_-|$ and $\eta = 1.96 \pm 0.58$ for ΔU_+. This, along with the shallower tails of the fitted stretched exponential functions, indicates greater deviation from Gaussianity for energy release compared to energy accumulation. These results support the picture that typically energy accumulates comparatively slowly and is released via large avalanche-like events.

The heavy-tailed distributions of $|\Delta U_-|$, the magnitudes of the negative energy increments which are the chief subject of this chapter, may have strong dependence on certain key parameters governing the mechanical equilibration protocol. To ensure that these distributions are not artifacts of simulation we investigate whether changing the parameters F_T and δt alters the qualitative properties of the distributions. In Figure 6.3 we compare these distributions using 3 runs for each parameter choice. Only weak dependence on F_T is observed. We find strong dependence on δt, however for each parameter

---

[1] In previous chapters the symbol ΔG_{mech} was used for ΔU, the net change of mechanical energy around the simulation cycle. ΔU will be used throughout this chapter to keep the notation simpler. We also use here the bracket notation for concentrations (e.g. $[\alpha]$) which in previous chapters were notated as C_{CL}. It is hoped that the notation is clear from context and will not cause confusion.

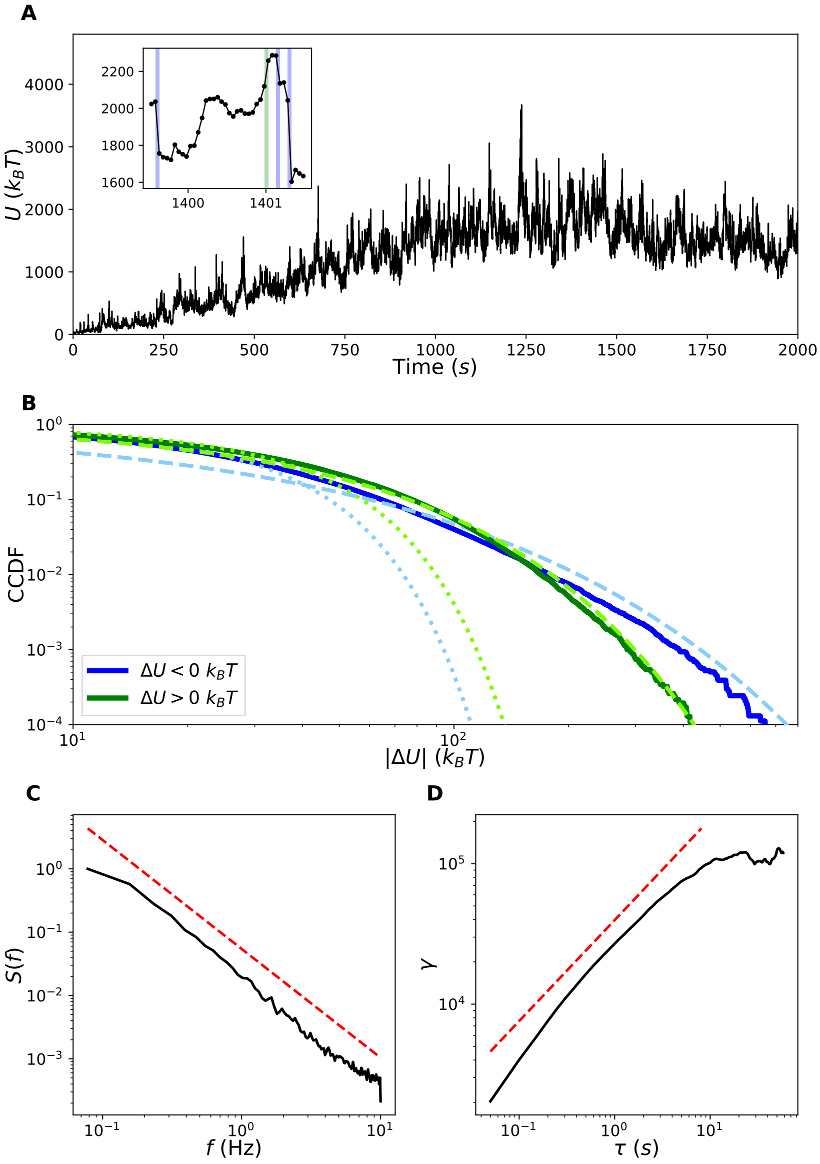

FIGURE 6.2: Statistics of ΔU. **A**: Trajectory of the network's mechanical energy $U(t)$ for condition $C_{3,3}$. Inset: A blow-up of the trajectory to show instances of rare events ($|\Delta U| > 100$ k_BT) of energy release (blue) and accumulation (green). **B**: CCDFs of $|\Delta U_-|$ (blue) and ΔU_+ (green) collected from five runs when the system is at QSS after 1000 s. Dotted lines in lighter colors represent fits to the data of a half-normal CCDF, and dashed lines represent fits of stretched exponentials. **C**: The normalized power spectral density of $U(t)$ for a single run at QSS from which the spectral exponent $\beta = 1.72$ is determined by fitting a power-law, shown offset in red. **D**: The semivariogram obeys the scaling relationship $\gamma \sim \tau^{2H_a}$ over the scaling range.

choice heavy tails exist and thus we may conclude that the cytoquake phenomenon is not an artifact despite their frequency and magnitude having dependence on δt. While a smaller choice for F_T and δt should correspond more closely to reality, we find that for the smallest of the tested values for these parameters the simulations did not complete in the allotted computer wall time of 2 weeks. Thus our choices for these parameters used in this chapter are chosen to be small while still allowing us to run full 2,000 s simulations.

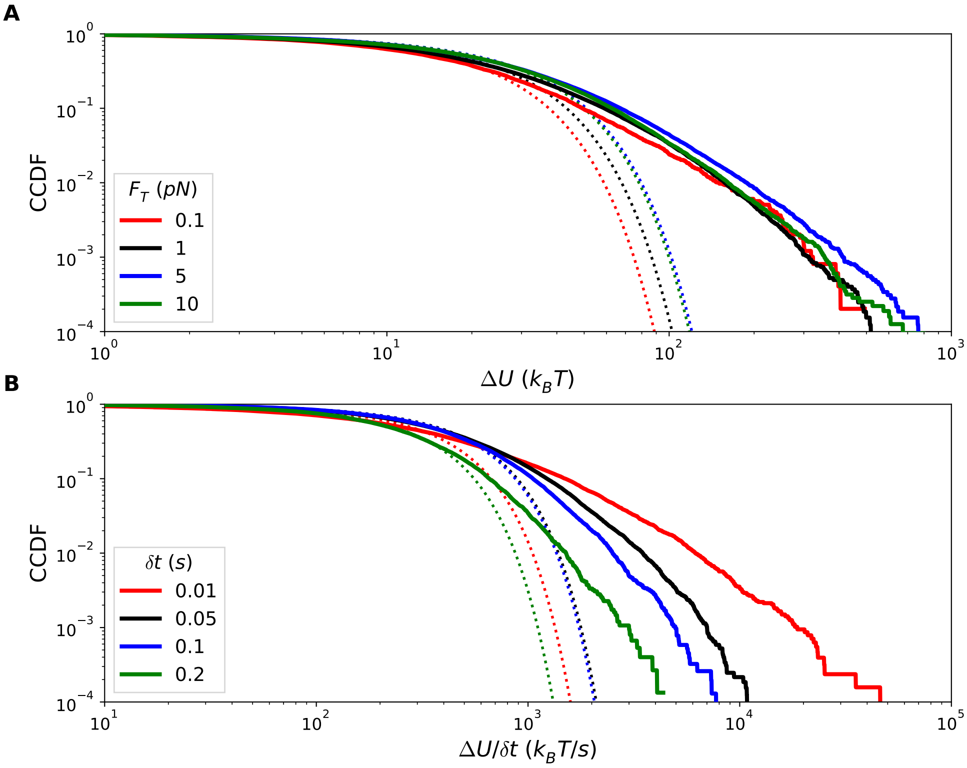

Figure 6.3: **A**: Complementary cumulative distribution functions of the negative increments $|\Delta U_-|$ at QSS for various choices of the force tolerance parameter F_T plotted against fitted half-normal CCDFs. For these runs condition $C_{3,3}$ is used with $\delta t = 0.05\ s$. **A**: Complementary cumulative distribution functions of the negative increments $|\Delta U_-|$ at QSS for various choices of the time between minimization, δt. The energy increments are normalized by δt for more direct comparison between these curves. For these runs condition $C_{3,3}$ is used with $F_T = 2\ pN$.

We next analyze the temporal correlations of $U(t)$ at QSS. A self-affine stochastic time series $G(t)$, for which $G(t)$ and $|\zeta|^{H_a} G(t/\zeta)$ have the same statistics for any scaling parameter ζ, has a power spectral density $S(f)$ exhibiting a power-law dependence on frequency f: $S(f) \propto f^{-\beta}$, where the spectral exponent β is the persistence strength, related to the color of the signal [274, 275]. We find $\beta = 1.72 \pm 0.02$ for $U(t)$ as shown in Figure 6.2.C, painting $U(t)$ as a pinkish brown signal; thus $U(t)$ is non-stationary and has temporally anti-correlated increments ΔU. Self-affine time series further obey the theoretical relationship $\beta = 2H_a + 1$ when $1 \leq \beta \leq 3$, where H_a is the Hausdorff exponent determined from the scaling of the semivariogram $\gamma(\tau) = \frac{1}{2}\overline{(G(t+\tau) - G(t))^2} \sim \tau^{2H_a}$, and where the bar represents temporal averaging [159, 276]. We find that this relationship is satisfied by $U(t)$, as shown in Figure 6.2.D, yielding $H_a = 0.36 \pm 0.01$ and confirming that $U(t)$ is self-affine. Such non-Markovian and self-affine time series and spatial patterns commonly arise in various complex geophysical processes (e.g. the temporal variation of river bed elevation), further supporting the connection between the cytoskeleton and earth systems [277, 278].

6.3.2 Distinguishing features of cytoquake events

We find that cytoquakes are correlated with several changes in the state of the network. In Figure 6.4 we show that rare large events of energy accumulation correspond to a greater than usual number of myosin motor steps whereas rare large events of energy release correspond to greater than usual total displacement of the actin filaments and a slightly greater number of linker unbinding events. These large total displacements do not come from highly localized motions, and instead depend on many filaments each being displaced more than usual (Figure 6.5). This agrees with the notion of cytoquakes as collective structural rearrangements of the network.

We also observe cytoquakes to induce a spatial homogenization of the tension sustained by the network during large events of energy release, as quantified by changes in the Shannon entropy of the spatial tension distribution $H(t) = -\sum_{ijk} P_{ijk}(t) \ln P_{ijk}(t)$. The tension distribution P_{ijk} is constructed by discretizing the simulation volume of $1\ \mu m^3$ into a grid of 10^3 voxels indexed by i, j, k, and computing the proportion of the total network tension belonging to the mechanical elements (filament cylinders, crosslinkers, and motors) inside each voxel; additional details can be found in Appendix E.

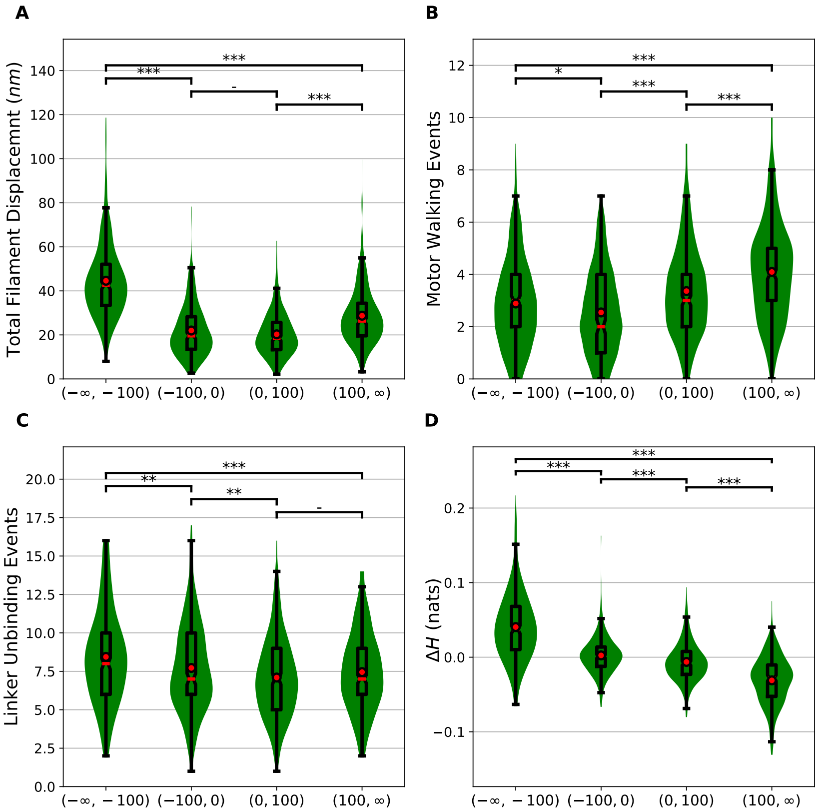

FIGURE 6.4: **A**: Differences in the total filament displacement between simulation cycles for which ΔU is less than $-100\ k_BT$, cycles for which $\Delta U \in (-100\ k_BT,\ 0\ k_BT)$, cycles for which $\Delta U \in (0\ k_BT,\ 100\ k_BT)$, and finally cycles for which ΔU is greater than $100\ k_BT$. To compare these distributions, the two-sided p-value of the Wilcoxon rank-sum test between pairs of cycle types is reported as being either not significant: - ($p \geq 0.05$), significant at level 1: * ($p < 0.05$), at level 2: ** ($p < 0.01$), or at level 3: *** ($p < 0.001$). The total filament displacement is computed as the sum over all filaments of the distance between a filament at time t and that filament at time $t + \delta t$. The calculation of distance between filaments is described in Figure E.3. Since there many more simulation cycles for the categories $\Delta U \in (-100\ k_BT, 0\ k_BT)$ and $\Delta U \in (0\ k_BT, 100\ k_BT)$, we took a random sub-sample (~ 300 each) of all events for these categories to be roughly equal to the number of events for which $\Delta U < -100\ k_BT$ and for which $\Delta U > 100\ k_BT$. In these combination violin and box-and-whisker plots, the red circle represents the mean, the red bar represents the median, and the notches in the box represent the 95% confidence interval of the median. **B**: Differences in the number of motor walking events between the different cycle types as just described. **C**: Differences in the number of α-actinin unbinding events between the different cycle types. **D**: Differences in the Shannon entropy of the spatial tension distribution of network tension.

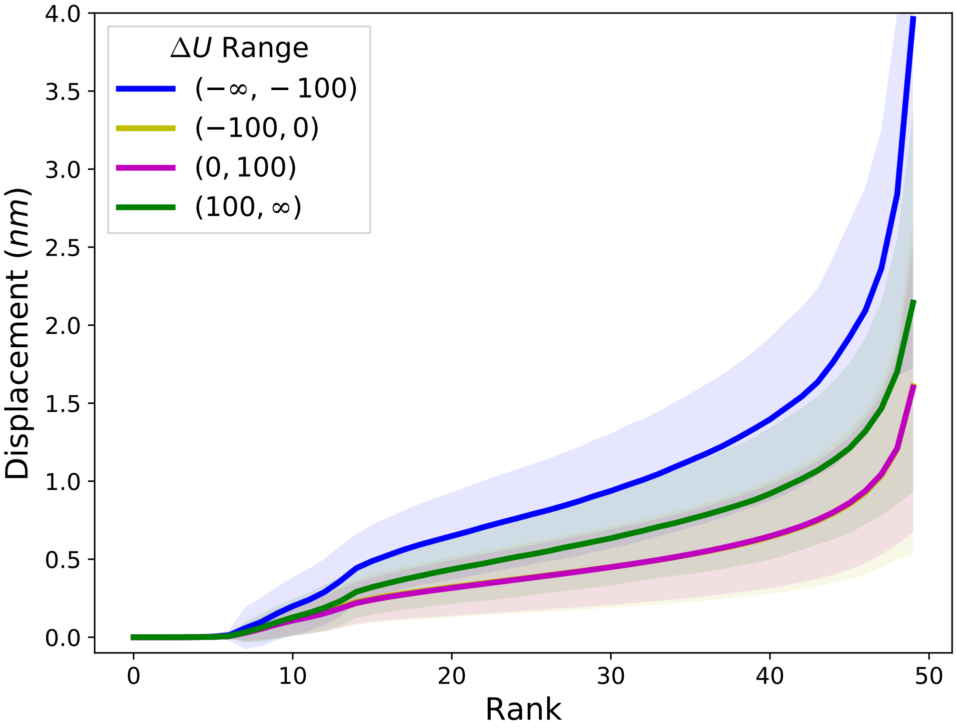

FIGURE 6.5: Rank-size distribution of the displacements experienced by each of the 50 filaments during simulation cycles when ΔU is in different ranges, in units of k_BT. For each cycle, the filaments are ranked according to their displacement and these ranks are plotted against the corresponding displacement. The average and standard deviation of these rank-displacement curves are taken over each cycle in a given category. The curves for the categories $\Delta U \in (-100\ k_BT,\ 0\ k_BT)$ and $\Delta U \in (0\ k_BT,\ 100\ k_BT)$ are almost exactly coincident. This data is collected from one run of condition $C_{3,3}$ at QSS.

6.3.3 Testing different concentrations and system sizes

We next discuss how these results generalize to different concentrations of associated proteins and different system sizes, and then connect our results to existing experiments. Five concentrations of α-actinin (ranging from 0.17 to 5.48 μM) and five concentrations of myosin minifilaments (ranging from 0.003 to 0.08 μM) were tested with a constant G-actin monomer concentration of 13.3 μM, in the regime of physiological concentrations [209]. At the lowest concentrations of cross-linkers and motors, the network did not contract, representing a very different actomyosin phase which we omit comparisons to. For all of the conditions producing contracting networks, we found that asymmetric heavy-tailed distributions of ΔU persist, with large values of the non-Gaussian parameter for $|\Delta U_-|$ ($\eta \sim 5 - 20$) and ΔU_+ ($\eta \sim 2 - 5$), although η for negative increments was observed to decrease with the motor concentration (Figure 6.6). We conclude that the avalanche-like energy fluctuations discussed above are not highly sensitive to associated protein concentrations. These fluctuations may depend on the parameters of the force-sensitive reaction rates (which are taken here to correspond to experimental values), but we leave this interesting question for future work.

To check the robustness of anomalous energy fluctuations in cytoskeletal networks, we measured the non-Gaussian parameters η for $|\Delta U_-|$ and ΔU_+ at a range of concentrations of cross-linkers and myosin motors. For the lowest concentrations of these proteins, the networks did not contract (conditions $C_{1,j}$ and $C_{i,1}$, $i, j = 1 \ldots 5$), and we omit them from the comparison. The results are displayed in Figure 6.6. Despite a trend toward smaller η for $|\Delta U_-|$ with increasing motor concentration, we may conclude that the asymmetric heavy tails of ΔU persist across these conditions.

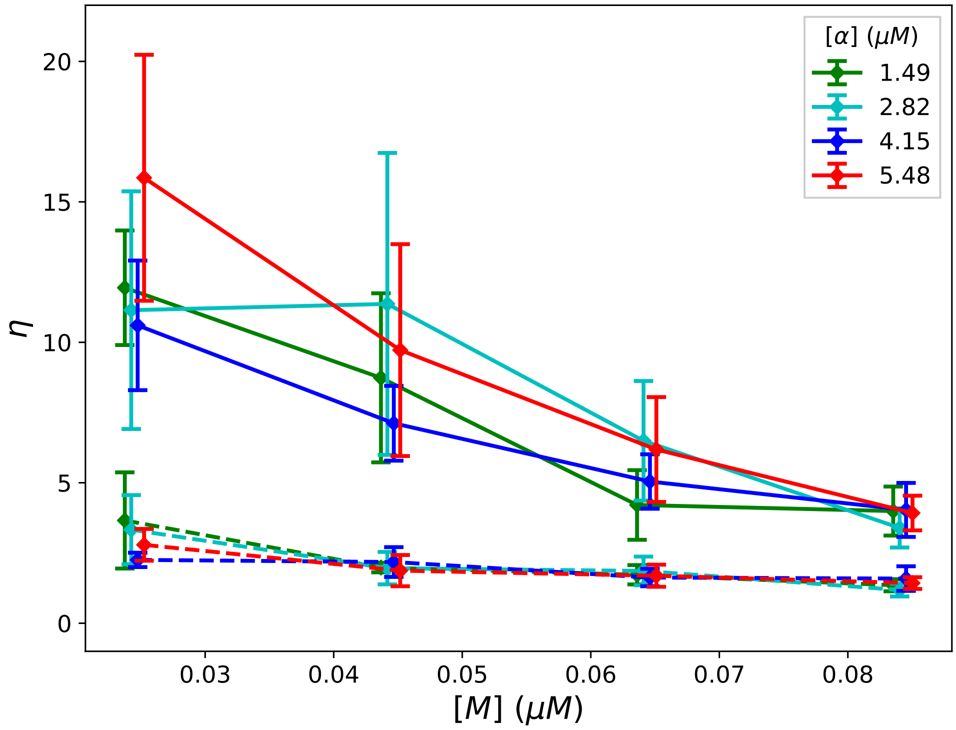

FIGURE 6.6: Plots of the non-Gaussian parameter η for the distributions of $|\Delta U_-|$ (solid lines) and of ΔU_+ (dashed lines) at QSS for various concentrations of myosin motor ($[M]$) and α-actinin cross-linkers ($[\alpha]$). The mean and standard deviation is shown over five runs of each condition. A small horizontal offset is added to the points to ease visibility.

We performed a finite-size scaling study by holding the concentrations of condition $C_{3,3}$ fixed (with α-actinin concentration $[\alpha] = 2.82\ \mu M$ and motor concentration $[M] = 0.04\ \mu M$) and varying the system volume V. Holding the concentrations of all chemical species for condition $C_{3,3}$ fixed, we performed simulations using cubic volumes with side lengths ranging from 0.5 to 2.5 μm. Larger systems reach QSS at later times, and our simulations of larger systems did not reach QSS in the allotted computational time. As a result, we collected samples of $|\Delta U_-|$ for these systems on the approach to QSS, from 300 to 800 s, once the networks had all nearly fully percolated (i.e. nearly all filaments belonged to a single component connected by cross-linkers), trusting that the relevant scaling behavior could still be observed. Stretched exponential functions approximately fit the distributions of ΔU_+ and $|\Delta U_-|$ for all system sizes (see Figure 6.7.A for the fits of $|\Delta U_-|$). Larger systems displayed steeper tails as indicated by the observed power-law decay of η for $|\Delta U_-|$ and ΔU_+ (Figure 6.7.B), although interestingly η for $|\Delta U_-|$ is larger than that for ΔU_+ by a constant factor of roughly 3 for all systems sizes. The steeper tails are also evidenced by the slow growth of the Kohlrausch exponents k with V (Figure 6.7.C). Thus, the distributions of energy release and accumulation across the entire network become narrower and more Gaussian for large systems. This, in contrast to driven-dissipative systems that exhibit self-organized criticality (SOC), suggests the existence of some intrinsic and finite scale for avalanche-like releases of energy in cytoskeletal networks. By summing over many local energy fluctuations of this finite scale, the distribution of the fluctuations in the total energy U becomes increasingly Gaussian for large systems owing to the central limit theorem. This intrinsic scale may be partly determined by the non-conservative transfer (dissipation) of mechanical energy as it spreads through the network during avalanches [4, 164].

6.3.4 Connection to experiments

Existing experimental definitions of cytoquakes define them as large local displacements of the cytoskeleton probed using transmembrane attached microbeads or flexible micropost arrays, rather than as large changes in the cytoskeleton's total energy U as done here [168, 169]. To compare our results to experiments, we make the corresponding local measurements of the displacements of individual filaments, finding that the resulting distributions are heavy-tailed with values of the non-Gaussian parameter for most filaments in the range $\eta \sim 1 - 5$, in semi-quantitative agreement with *in vivo* measurements (Figure 6.8) [169]. We also find our estimate of typical actin filament displacement speeds ($\sim 10\ nm/s$) to be consistent in order of magnitude with separate *in vitro* measurements of contractile networks [69]. In addition, we find that the heavy tails of these local measurements persist even for large system sizes, whereas global measurements such as the total energy or total filament displacement become increasingly Gaussian for large systems. This distinction between local and global measurements may be important in interpreting future studies of anomalous statistics in cytoskeletal self-organization. We

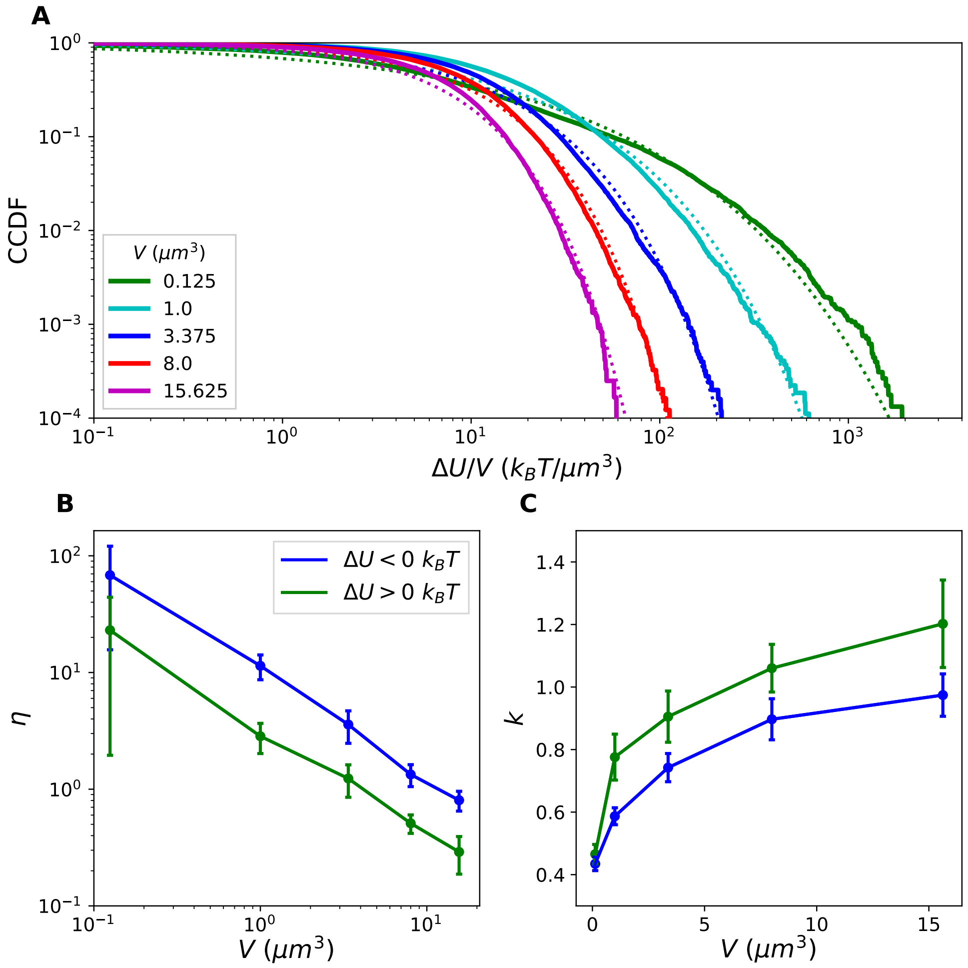

Figure 6.7: **A**: CCDFs of $|\Delta U_-|$ normalized by the system volume V collected for 5 runs for increasing system sizes plotted against the fitted stretched exponential functions. **B**: The non-Gaussian parameter η for $|\Delta U_-|$ and ΔU_+ are plotted with uncertainty taken over the different runs. **C**: The Kohlrausch exponents k for $|\Delta U_-|$ and ΔU_+.

finally mention in connection to experiments that it has recently been argued that more detailed understanding of mechanical dissipation by cytoskeletal networks should help to more precisely control traction-based measurements of cellular force production [170].

Previous *in vivo* studies have framed cytoquakes as large local displacements of the cytoskeleton, which were measured by tracking the motion of apically attached microbeads as well as the deflection of micropost arrays under the basal surface [168, 169]. Here, we have leveraged the information accessible from simulation to frame cytoquakes as large releases of stored mechanical energy, which we show typically correspond to large global displacements. This energy-centered point of view deepens the analogy between cytoquakes and earthquakes, which are thought to result from an avalanche-like process of energy accumulation and release in the earth's crust that manifests as occasional large displacements [159, 276]. Our measurements of displacements in Figures 6.4 Figure 6.5 are global in the sense of being summed for each simulation cycle over all filament in the network, capturing more modes of network buckling than the local displacement measurements obtained experimentally, e.g. using attached microbeads. However, we can also make a similar local measurement to compare to experiments. Rather than summing over all filaments, we track each filament individually and measure the set $\{\eta_f\}_{f=1}^{N_f}$, where $N_f = 50$ is the number of filaments, of the non-Gaussian parameter corresponding to each filament f's distribution of displacements from 300 to 800 s. Displacements are calculated according to Figure E.3. The displacements experienced by a filament should depend on the filament's location in the network, but we observe for nearly all filaments a heavy-tailed distribution characterized by $\eta_f > 0$, as shown in Figure 6.8. In addition, we find the distribution of η_f itself to be heavy-tailed, which appears to agree with experiments using micropost arrays [169]. Finally, we observe similar values of η_f across simulation volumes V.

These local measurements, obtained by tracking each filament individually, can be compared to global measurements, obtained by summing over every filament to get the total displacement. The distribution of total displacements is closer to Gaussian, characterized by $\eta \approx 0$ for most volumes tested (Figure 6.8). As with the increasing Gaussianity of ΔU for large systems, this can be attributed to the central limit theorem since many filaments were summed over to determine the total displacement. We therefore conclude that in large systems, metrics can be heavy-tailed when measured locally but Gaussian when measured globally.

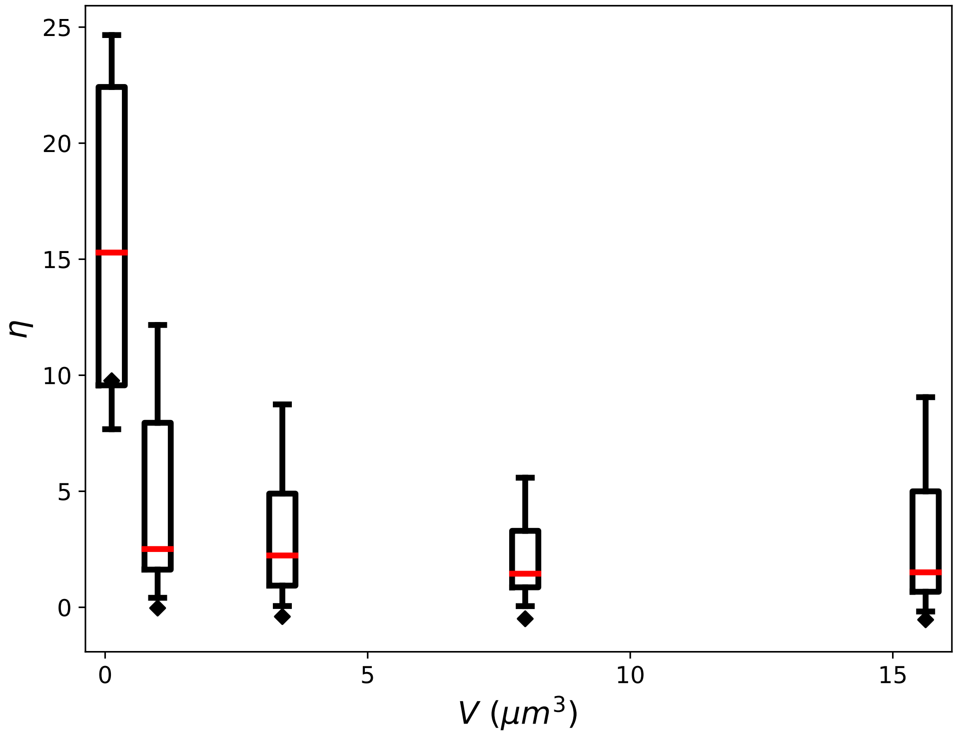

FIGURE 6.8: Plots of the distributions of non-Gaussian parameter η_f for the distributions of displacements of individual filaments, for different simulation volumes V. For a single run of each V, the displacement of each filament from time point to time point in the system is tracked and, from these time series, η_f is determined for each filament f. The box and whisker plots summarize the distribution of η_f for all filaments in the system, with the median shown as the red bar, the box extending from the first to the third quartiles, and the whiskers extending across the range of η, omitting outliers. The diamonds indicate the value of η obtained when, instead of tracking each filament's displacement individually, the total summed displacement of all filaments from time point to time point is tracked. These global measurements of displacement are more Gaussian (with $\eta \approx 0$) than the corresponding local measurements obtained from tracking filaments individually.

6.3.5 Normal mode decomposition probes network's mechanical state

Having described the statistics of the increments ΔU, we next aim to connect the occurrence of cytoquakes, defined as large values of $|\Delta U_{-}|$, to the cytoskeletal network's mechanical stability. To this end we implemented a method to compute the Hessian matrix $\boldsymbol{\mathcal{H}}$ of the mechanical energy function U. The eigen-decomposition of $\boldsymbol{\mathcal{H}}$ is $\boldsymbol{\Lambda} = \{\lambda_k\}_{k=1}^{3N}$, where $3N$ is the number of mechanical degrees of freedom in the system, which comprises N "beads" that are used to discretize the actin filaments. $\boldsymbol{\Lambda}$ is related to the mechanical stability of the cytoskeletal network: the eigenvectors $\mathbf{v}_k$ are the normal vibrational modes of the network, and the eigenvalues λ_k indicate the stiffness ($|\lambda_k|$) and stability ($\mathrm{sgn}(\lambda_k)$) of the corresponding mode. An example vibrational mode is illustrated in Figure 6.1. We draw inspiration for studying $\boldsymbol{\Lambda}$ in the current context from several sources: in single-molecule molecular dynamics studies, the saddle-points of U (i.e. points in the landscape with some imaginary frequencies) are associated with transition states [280, 281]; studies of polymer networks show that internal stresses produce non-floppy vibrational modes even below the isostatic threshold [282]; in simulations of glass-forming liquids, the instantaneous normal mode spectrum allows inference about proximity to the glass transition and determination of incipient plastic deformation regions [283–285]; in deep learning models for predicting earthquake aftershock distributions, it was found that certain metrics also related to stability (e.g. the von-Mises criterion) are informative model inputs [286, 287]. We describe some interesting observed trends on metrics defined on $\boldsymbol{\Lambda}$ below.

In MEDYAN, semi-flexible filaments are represented as a connected sequence of thin cylinders whose joined endpoints (i.e. hinges) are called beads. The set of potentials defining the mechanical energy of the filaments and associated proteins is outlined below. The mechanical energy U is a function of these beads' positions, and elements of the Hessian matrix are defined as

$$\mathcal{H}_{i\mu,j\nu} = \frac{\partial^2 U}{\partial x_{i\mu} \partial x_{j\nu}} = -\frac{\partial F_{i\mu}}{\partial x_{j\nu}} = -\frac{\partial F_{j\nu}}{\partial x_{i\mu}}, \tag{6.1}$$

where $x_{i\mu}$ is the μ^{th} Cartesian component of the position of the i^{th} bead. We have $\mu = x, y, z$ and $i = 1, ..., N$ where N is the number of beads in the network, so $\boldsymbol{\mathcal{H}}$ is a square symmetric $3N$-dimensional matrix. The number of beads $N(t)$ will change as filaments (de)polymerize; in these simulations, at QSS a single filament of length 0.5 μm comprises $\sim$ 10 cylinders (11 beads), each $\sim$ 50 nm in length. After each mechanical minimization, $\boldsymbol{\mathcal{H}}(t)$ is constructed by numerically computing the derivatives on the right of Equation 6.1. The derivative $\frac{\partial F_{i\mu}}{\partial x_{j\nu}}$ is found using a second-order central difference approximation by moving the j^{th} bead in the $\pm\nu$ directions by a small amount and determining the changes in the force component $F_{i\mu}$ [288]. Due to issues of numerical accuracy, we do not assume the symmetry of the matrix $\boldsymbol{\mathcal{H}}$, but instead directly compute each component $\mathcal{H}_{i\mu,j\nu}$ and then symmetrize the result: $\frac{1}{2}(\boldsymbol{\mathcal{H}}^{\intercal} + \boldsymbol{\mathcal{H}}) \rightarrow \boldsymbol{\mathcal{H}}$.

We next describe some interesting observed trends of metrics defined on $\boldsymbol{\Lambda}$. We distinguish between unstable, stable, soft, and stiff modes: for unstable modes $\lambda_k < 0$, for stable modes $\lambda_k \geq 0$, for soft modes $0 \leq \lambda_k < \lambda_T$, and for stiff modes $\lambda_k \geq \lambda_T$, where we define the threshold $\lambda_T = 40$ pN/nm to discriminate between the twin peaks in the density of states (Figure 6.11.B). The set $\{\lambda_k\}_{k=1}^{3N}$ is visualized with these modes labeled in Figure 6.11.A for a QSS time point of condition $C_{3,3}$. A very small number of unstable modes persist after each minimization cycle, later iterations stopping once the maximum force on any bead is below a threshold F_T (here 1 pN). Thus the minimized configurations are in fact saddle-points of U; this is expected as it is known from the theory of minimizing loss functions that the ratio of saddle-points to true local minima increases exponentially with the dimensionality of the domain [289]. We expect that in the space of all possible network topologies (i.e. patterns of cross-linkers and motors binding to filaments), the energy landscape will be rugged, leading to the well-appreciated glassy dynamics of non-equilibrium cross-linked networks [109, 133]. For a fixed topology, however, which is the result of the chemical reactions occurring during step 1) of the iterative simulation cycle, the energy landscape should be smooth (i.e. non-rugged) with respect to the beads' positions, with a single nearby local minimum being sought during mechanical minimization in step 3). The residual unstable modes are therefore thought to be an unimportant artifact of thresholded stopping in the conjugate-gradient minimization routine, and not representative of some physical feature of cytoskeletal networks. The observed quantitative dependence of the number of residual unstable modes on F_T supports this conclusion and is illustrated in Figure 6.9.

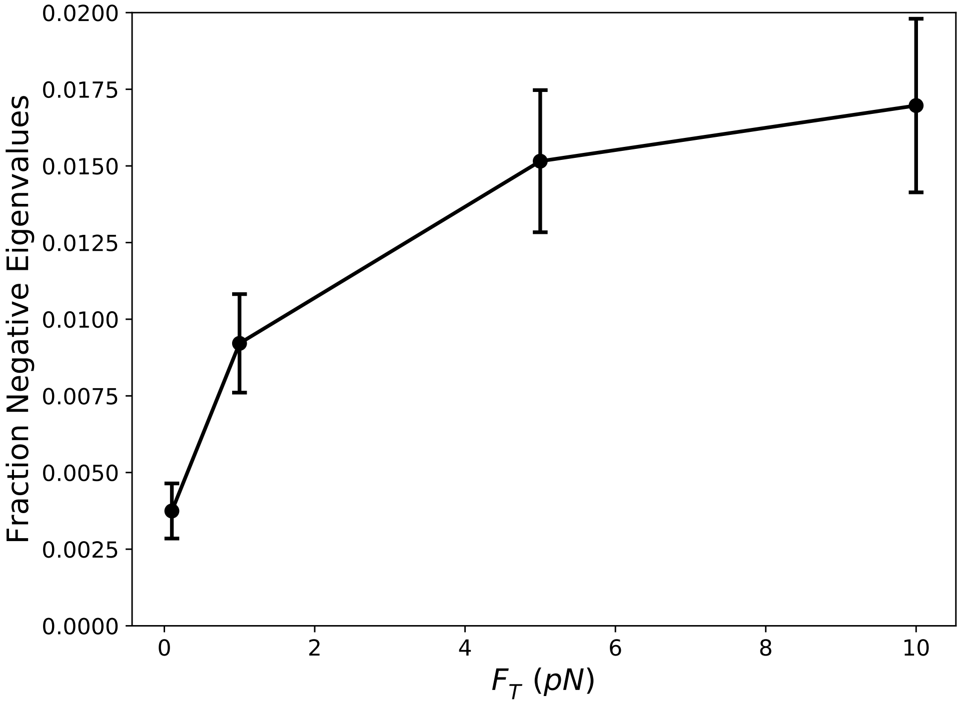

Figure 6.9: Scatter plot showing the fraction of negative eigenvalues remaining after mechanical minimization when different choices of the parameter F_T are used. The data is collected from QSS for 3 runs of $C_{3,3}$, with the standard deviation taken over time and over the runs.

We quantify the number of degrees of freedom involved in a given normalized eigenvector $\mathbf{v}_k$ using the inverse participation ratio [283]:

$$r_k = \left(\sum_{i=1}^{N} \sum_{\mu=1}^{3} (v_{k,i\mu})^4 \right)^{-1} . \tag{6.2}$$

If the eigenmode involves only one degree of freedom, then one component of $\mathbf{v}_k$ will be one and the rest will be zero, and $r_k = 1$. On the other hand, if the eigenmode is evenly spread over all $3N$ degrees of freedom, then each component $v_{k,i\mu} = (3N)^{-1/2}$, and $r_k = 3N$. In Figure 6.11.B we plot r_k for the unstable, soft, and stiff modes along with the density of states, showing that the soft modes involve many degrees of freedom while the stiff and unstable modes are comparatively localized.

We find that the mean value $\langle r_k \rangle$ varies non-monotonically with myosin motor concentration $[M]$ and α-actinin concentration $[\alpha]$ (Figure 6.11.C). To understand this trend we implemented a mapping from the cytoskeletal network into a graph and measured its mean node connectivity, a purely topological measure of network percolation. We construct a graph to represent the cross-linker binding topology of cytoskeletal networks. Nodes in the graph correspond to actin filaments, and weighted edges (which may be thresholded and converted to binary edges in an unweighted graph) correspond to the number of cross-linkers connecting the pair of filaments. The mean node connectivity is defined as the average over all pairs of nodes in the unweighted graph of the number of edges necessary to remove in order to disconnect them, thus quantifying the typical number of force chains between filaments, or equivalently the extent of network percolation [104, 290]. Revealingly, the mean node connectivity correlates closely with $\langle r_k \rangle$ for the stable modes across the various conditions $C_{i,j}$ (Figure 6.11.D). We also find the number of connected components of $\boldsymbol{\mathcal{H}}$ and of the graph's adjacency matrix to match for most time points, supporting this connection between network topology and stable mode delocalization. Intermediate concentrations of myosin motors enhance the network percolation, but as $[M]$ continues to increase the motors act to disconnect cross-linked network structures causing the mean node connectivity and $\langle r_k \rangle$ to decrease.

We observe that as a network contracts and becomes percolated during the process of myosin-driven self-organization, the stable modes steadily delocalize ($\langle r_k \rangle$ increases) and stiffen (the geometric mean $\langle \lambda_k \rangle_g$ increases), as shown in Figures 6.11.E and Figures 6.11.F. During this process we also witness a qualitative change in the level spacing statistics of the very soft and delocalized modes ($\lambda_k < 10\ pN/nm$, $r_k > 100$) from a Poisson to a Wigner-Dyson distribution (Figure 6.10). This indicates that in the percolated state these vibrational modes interact and exhibit level repulsion, similar to soft particles near the jamming transition [169, 262, 291]. Future studies may reveal further similarities between these systems and other marginally stable solids [108, 133].

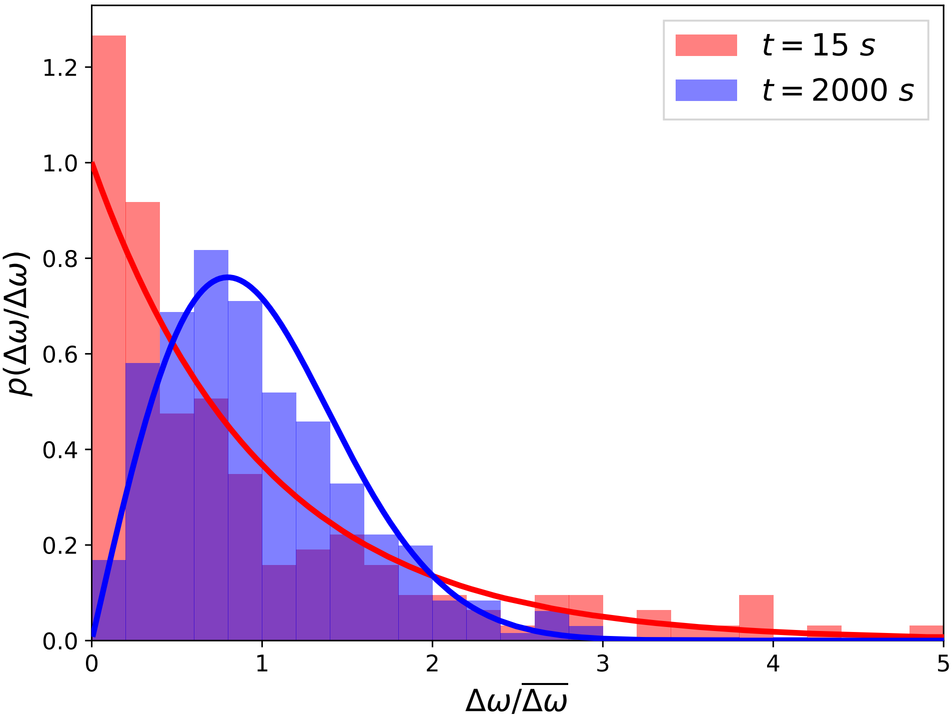

FIGURE 6.10: Histograms of the level spacings $\Delta\omega = \omega_{k+1} - \omega_k$, where $\omega_k = \sqrt{\lambda_k}$ are ordered so that ω_k increases as k increases, normalized by their average $\overline{\Delta\omega}$ for the very soft ($\lambda_k < 10\ pN/nm$) and delocalized ($r_k > 100$) vibrational modes at different times of a run of condition $C_{3,3}$. The Poisson distribution $p(\Delta\omega/\overline{\Delta\omega}) = e^{-\Delta\omega/\overline{\Delta\omega}}$ and the Wigner-Dyson distribution $p(\Delta\omega/\overline{\Delta\omega}) = \frac{\pi}{2}(\Delta\omega/\overline{\Delta\omega})e^{-\frac{\pi}{4}(\Delta\omega/\overline{\Delta\omega})^2}$ are plotted as red and blue solid lines. This transition in distributions signifies that in the percolated network at 2000 s the frequencies of these modes are no longer randomly spaced and begin to interact, exhibiting level repulsion for small $\Delta\omega/\overline{\Delta\omega}$.

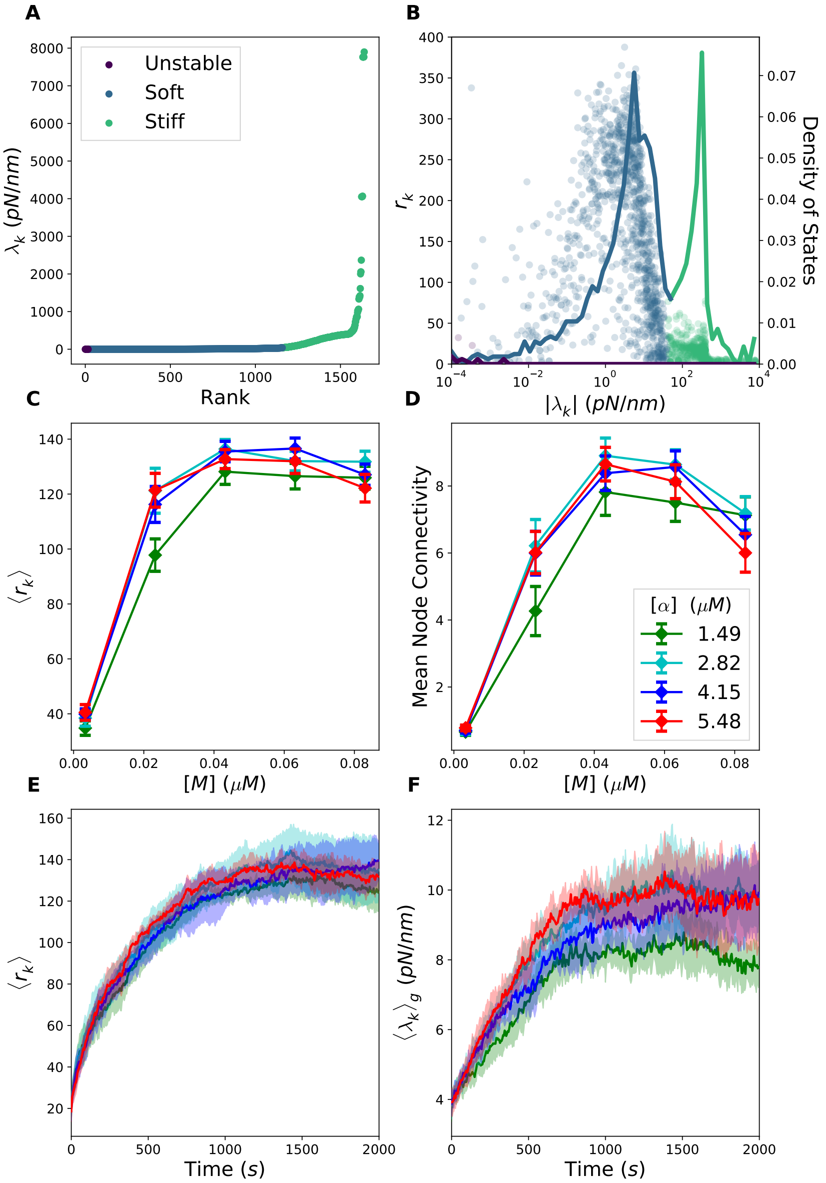

FIGURE 6.11: Metrics defined on Hessian eigen-decomposition. **A**: Ordered eigenspectrum $\{\lambda_k\}_{k=1}^{3N}$ at a QSS time point for condition $C_{3,3}$. **B**: Scatter plot of the pairs $|\lambda_k|, r_k$ (circles) plotted against the density of states (solid lines), i.e. the proportion of eigenvalues between λ and $\lambda + d\lambda$. **C**: The mean value at QSS of $\langle r_k \rangle$ for the stable modes for various conditions $C_{i,j}$. The conditions $C_{1,j}$ with low linker concentrations are not visualized as these networks did not percolate and obscure visualization for the remaining conditions. The mean is taken over the last 500 s and over different runs. **D**: The mean value of the mean node connectivity for various conditions. **E**: Trajectories of $\langle r_k \rangle$ of the stable modes as the network self-organizes for the conditions $C_{2,3}$, $C_{3,3}$, $C_{4,3}$, and $C_{5,3}$, with the mean and standard deviation taken over the different runs. **F**: Similar trajectories of $\langle \lambda_k \rangle_g$ of the stable modes.

6.3.6 Cytoquake forecasting

Can the eigen-decomposition of the Hessian matrix be used to forecast cytoquake occurrence? Intuition suggests that, in analogy with the connection between imaginary frequencies (i.e. unstable modes) and molecular transition states, the vibrational modes of the cytoskeletal network may contain information that a large structural rearrangement is poised to occur [280, 281]. To test this idea, and without detailed *a priori* knowledge about which features in $\mathbf{\Lambda}$ would be informative, we implemented a machine learning model using the eigen-decomposition as the input and outputting the predicted probability of observing a large event of energy release ($\Delta U < -100\ k_BT$) occurring within the next 0.15 s. As detailed in Appendix E, we found that, indeed, the Hessian eigenspectrum $\mathbf{\Lambda}$ contains sufficient information to forecast cytoquake occurrence with significant accuracy compared to a random model. We first reduced the dimensionality of $\mathbf{\Lambda}(t)$ using principal component analysis, finding that 30 dimensions sufficed to explain $> 95\%$ of the variance across time points, and then used the reduced input in a three layer feed-forward neural network. We validated our model using receiver operating characteristic curves, achieving an area under the curve (AUC) of 0.70 when using data from five runs of condition $C_{3,3}$. This improvement in prediction performance over a random model (which would have an AUC of 0.5) implies that mechanical instability, as encoded in the Hessian eigenspectrum, precedes the occurrence of cytoquakes.

6.3.7 Cytoquakes have enhanced displacement along soft modes

To further study the connection between cytoquakes and mechanical stability, we measured the projections of the network's displacements onto the vibrational normal modes $\{\mathbf{v}_k\}_{k=1}^{3N}$. Network displacements $\mathbf{d}(t)$ were found by tracking the movement of each of the $N(t)$ beads during simulation cycles. As a working approximation, beads that depolymerized during a cycle were assigned a displacement of 0, and beads that newly polymerized were not assigned elements in $\mathbf{d}(t)$. The $3N$-dimensional displacement vectors $\mathbf{d}$ were then normalized to have unit length. We define $d_k = \mathbf{d} \cdot \mathbf{v}_k$ as the projections of $\mathbf{d}$ onto the eigenmodes $\mathbf{v}_k$, which obey $\sum_k d_k^2 = 1$ owing to the normalization of $\mathbf{d}$ and

$\mathbf{v}_k$. Thus the quantity d_k^2 is the weight of the displacement $\mathbf{d}$ along the k^{th} eigenmode. With this we define the effective stiffness $\lambda_P = \sum_k d_k^2 \lambda_k$ as the displacement-weighted average of the eigenvalues. In Figure 6.12 we display a scatter plot of the pairs $\Delta U(t)$, $\lambda_P(t)$ measured during QSS for a run of condition $C_{3,3}$, along with a kernel density estimate of their joint probability density function (PDF). We distinguish between soft ($0 \leq \lambda_k < \lambda_T$) and stiff ($\lambda_k \geq \lambda_T$) eigenmodes, where $\lambda_T = 40\ pN/nm$ separates the twin peaks in the density of states (Figure 6.11). The structure of the joint PDF is markedly asymmetric about $\Delta U = 0$ and shows that λ_P during cytoquake events is almost always soft, whereas for all other simulation cycles λ_P could be soft or stiff with similar probabilities. We also consider $n_k = \frac{d_k^2}{r_k}$ as the weight of the displacement along eigenmode k per degree of freedom involved in the eigenmode, where r_k is the inverse participation ratio [283]. We define n_{soft} and n_{stiff} as the mean of n_k over the soft and stiff subsets. Values of $n_{\text{soft}}/n_{\text{stiff}}$ for different simulation cycle types are displayed in the inset of Figure 6.12, showing that $n_{\text{soft}} < n_{\text{stiff}}$ typically only during cytoquakes. Based on this analysis, we conclude that during the large collective rearrangements corresponding to cytoquakes, cytoskeletal networks exhibit enhanced displacement along the soft vibrational modes. We qualify these results by observing that, since cytoquakes involve particularly large network displacements, it may be inappropriate to interpret them using the local harmonic approximation to U implicit in Hessian analysis [285]. In addition, changes in network topology from linker and motor (un)binding cannot be captured using normal mode decomposition of instantaneous network configurations. The eigenspectrum $\mathbf{\Lambda}(t)$ still informs on the stability of the energy minimized configuration before a cytoquake, but caution should be used in interpreting the cytoquake motion from t to $t + \delta t$ as decomposing cleanly into non-interacting motions along the normal modes $\mathbf{v}_k$. We leave a detailed analysis of the anharmonicity of cytoquake deformations to future work.

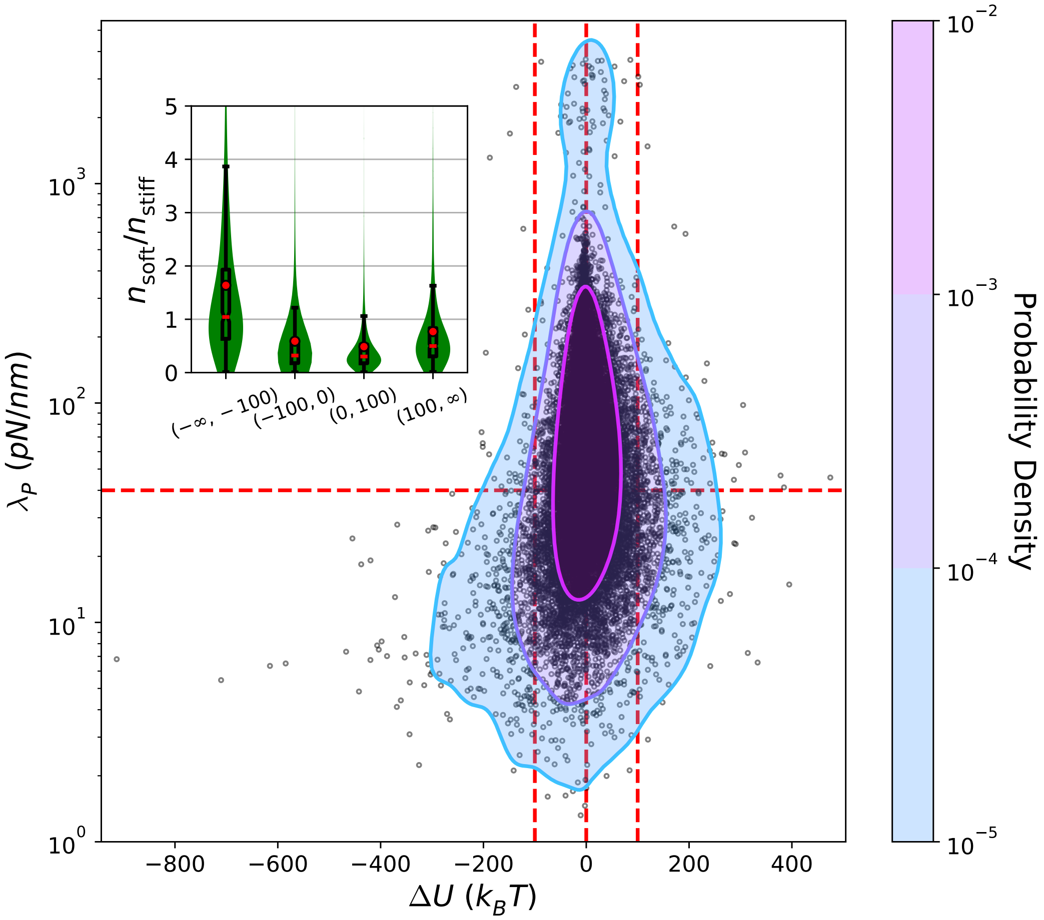

FIGURE 6.12: Scatter plot of the pairs ΔU, λ_P measured during QSS for a run of condition $C_{3,3}$. From these points, a Gaussian kernel density estimate of the joint PDF (treating λ_P on a log-scale) is constructed and shown as a contour plot. Red guidelines demarcate regions of interest. Inset: Combination violin and box-and-whisker plots showing the ratio $n_{\mathrm{soft}}/n_{\mathrm{stiff}}$ for different categories of simulation cycles, c.f. Figure 6.4. The inset is not blocking any of the scatter plot data.

6.4 Discussion

We have shown evidence supporting the following picture of active cytoskeletal network self-organization: cytoskeletal networks explore a rugged mechanical energy landscape in a stochastic process characterized by occasional, sudden jumps out of metastable configurations [109, 133]. These jumps entail non-Gaussian dissipation of mechanical energy and are accomplished by an avalanche-like process of spreading destabilization, resulting in a collective structural rearrangement and a homogenization of tension. These collective motions have large projections along the soft, delocalized vibrational modes, and furthermore, properties of these modes can be used to predict when such relaxation events will occur. Future research could help elucidate the role of force-sensitive reactions in controlling collective relaxation as well as the importance of anharmonicites in the energy landscape of cytoskeletal networks.

Chapter 7

Summary and future directions

We have presented in this thesis a computational method for quantifying the energetics of non-equilibrium cytoskeletal processes and applied it to gain new insights into the nature of cytoskeletal self-organization. As described in Chapter 1, this self-organization under different concentrations of accessory proteins is responsible for the diverse physiological functionality of the cytoskeleton. Thus, the new insights into cytoskeletal self-organization have biological implications. To conclude this thesis, we next summarize some of our key results and their biological implications and suggest future research directions.

We observed a generic tendency of cytoskeletal networks under fixed chemical conditions to decrease their dissipation rates over time, in both the active processes of actin subunit hydrolysis (Chapters 2 and 4) and motor walking (Chapters 4 and 5). Cytoskeletal conditions can be finely controlled by the cell by activating accessory proteins to varying degrees, but for a given concentration of activated proteins the total dissipation rate reaches a minimum in the steady state. This observation can be used as a predictive principle for steady state organizations of cytoskeletal structures. The minimization of dissipation from ATP hydrolysis by actin subunits appears consistent with the minimum entropy production principle of Prigogine applying to chemically reactive systems which, however, was only proved by Prigogine in the near-equilibrium regime [135]. In our simulations of contractile networks, polymerizing filaments only minimally interact with a boundary, so that the Brownian ratchet-like decrease in polymerization rates due to forces on the filament tips is not a major effect [101]. Polymerization against the leading edges of cells in lamellipodia may be strongly affected by this force-sensitivity, however, and it would be interesting to see how our results extend to that scenario. Presumably the net decrease in polymerization rates would only drive the average steady state dissipation rates lower, but recent work has shown that the collective interactions of branched filaments with a flexible boundary can give rise to scale-free fluctuations in the membrane shape, which could manifest as non-trivial distributions of dissipation rates [166, 292]. The decrease in dissipation from motor walking results from motors stalling and catching in networks with increasing typical tensions, such that the steady state is the most kinetically arrested. This agrees with the low rattling principle proposed by Chvykov and England [83, 156]. It would be interesting to study the how the typical time scale of attaining minimized dissipation rates compares with the time scales on which cells modulate the activation of accessory proteins. The transient behavior of the network likely plays an important role in cellular processes like chemotaxis or electrotaxis, where the cell is actively responding to sometimes rapidly changing external cues [293, 294]. It is also possible to experimentally manipulate myosin activation using light, which can allow one to, for example, control the dynamics of nematic defects in actomyosin gels [69, 77]. Computationally studying transient dissipation dynamics under time varying protein activation would thus be a promising future research direction.

An additional insight gained from the studies presented in this thesis pertains to the nature of mechanical relaxation by cytoskeletal networks. It was argued in Chapter 6 that mechanical relaxation can be highly collective, with large events involving many filaments, rather than local, with events involving just a few filaments. The cooperativity of mechanical relaxation can lead to a wide range of relaxation event sizes and may be a root cause of the observed variability in traction force measurements of migrating cells [170]. As mentioned in Chapter 6, it is plausible that this cooperativity allows for enhanced mechanical responsiveness to a changing cellular environment, which could be a significant physiological advantage. Several open questions remain about these cytoquake dynamics, though. One question regards the characteristic time and length scales of cytoquakes. Our study was at a fixed time resolution of 20 Hz and measured energy fluctuations summed over the whole 1 μm^3 network, so the cytoquakes (or "cytoquake-like events") we observed were limited to these scales. In a larger system, though, a cytoquake may have an extended lifetime as it cascades throughout the network, opening the door to a broader range of spatiotemporal scales than analyzed here. Current work underway in Reich's group is investigating this question *in vivo*, and this work could be complemented using MEDYAN simulations of large networks that spatially resolve the release of mechanical energy [169]. Another question is about the degree to which cooperative mechanical relaxation depends on force-sensitive reactions. Cooperativity should result from just the structural properties of a network with insensitive reaction rates, where the ability of the network to relax depends on its configuration and the mutual mechanical constraints filaments impose on each other. However, we expect cooperativity to be amplified when the reactions controlling network connectivity sensitively depend on the network forces. The degree to which avalanche-like dynamics can be controlled through the parameters of force-sensitive reaction remains to be investigated.

One possible application of the dissipation tracking method presented here would be to compare it with various indirect methods for quantifying entropy production based solely on observations of the system's trajectory or even snapshots of its structure [32, 34, 35, 130, 295, 296]. A benefit of these indirect methods is that they can be applied to experimentally accessible low-resolution representations of the system trajectory, although as a result it underestimates the true entropy production [36]. One could compare these calculations of entropy production in cytoskeletal networks with our explicit computational approach to assess the accuracy of the indirect estimates in real biological systems.

An additional principle potentially useful for understanding cytoskeletal self-organization via its non-equilibrium thermodynamics is the set of thermodynamic uncertainty relations [297]. Under various assumptions on the system's state space and dynamics, these relations provide bounds on the uncertainty in an arbitrarily defined current of a non-equilibrium system in terms of the system's dissipation rate, such that a larger dissipation rate implies a smaller uncertainty in the current. Originally established within a limited range of applicability, they have since been quite broadly generalized and appear to hold in some form for many systems of interest [130, 298–301]. These results have been applied to determine a bound on the efficiency of a discrete-state molecular motor moving on a linear track, and it is reasonable to expect that they also have an implication for higher-order cytoskeletal self-organization (in which many such motors collectively cause changes in network morphology) [302]. For instance, the network contraction rate may have an uncertainty that depends on the dissipation of consumed ATP energy by molecular motors. This could be relevant to cells with low concentrations of ATP, leading not just to slower physiological processes but a greater uncertainty in such processes. Testing the applicability of the thermodynamic uncertainty relations for cytoskeletal self-organization thus remains a promising future research direction using our computational method.

A theoretical question which would be of interest beyond cytoskeletal networks is about how entropy production varies as a function of length scale for active matter systems, where energy is injected at small scales. By comparison one can consider a turbulent fluid driven by out-of-equilibrium boundary conditions. This system has a so-called Kolmogorov cascade of energy flow, in which energy is transferred down from large scale coherent fluid motions, or eddies, to progressively smaller eddies until it is finally dissipated due to viscous interactions at the smallest length scale, lost in the incoherent degrees of freedom of molecular vibrations and disorganized motion [303, 304]. In contrast, energy in active matter systems is injected not at the global level (e.g. through boundary conditions) but at the local level (e.g. by molecular motors). Does a similar energy cascade picture apply in this case? Recent work on coarse-graining non-equilibrium chemical reaction networks suggests that a cascade-like picture does apply, such that dissipation occurs maximally at small scales where energy is injected, and dissipation decreases when measured using coarser reaction network representations [305]. It would be interesting to investigate how the spatial scale at which one is looking affects the measured dissipation in cytoskeletal networks. This topic may be relevant to the graph coarse-graining procedure for describing bundled cytoskeletal architectures discussed in Chapter 5.

A final future research direction that we would like to mention, and which is currently underway, is to drastically improve the realism of the mechanical model used in MEDYAN and to apply the improved model to study collective chiral symmetry breaking phenomena in cytoskeletal systems. Experiments *in vitro* and *in vivo* have revealed the existence of remarkable rotating dynamical phases of active networks [67, 306, 307]. Emerging at micrometer length scales, these collective swirling motions display a sensitive dependence on experimental conditions such as cross-linker concentration, which can even reverse the direction of the rotation. Their emergence may be explainable in terms of the dissipation minimization and low rattling principles mentioned earlier, and hence it is of interest to study such phenomena computationally using our free energy tracking method. To model such phenomena using MEDYAN, however, filament torques must be accounted for. We are addressing this by developing a new, finite-width filament model based on the Cosserat theory of elastic rods [308, 309]. This model will allow to more closely study how mechanical energy is stored in cytoskeletal networks (i.e. in shearing, stretching, bending, or twisting deformations), as well as to investigate novel rotational phenomena of cytoskeletal systems. Preliminary applications of the model have for instance shown that motorized filaments undergo axial rotations in a stochastic process similar to biased diffusion. This "active rotational diffusion" may be closely connected to dissipation by myosin motors and remains to be further investigated.

In addition to the future projects mentioned above, which will likely require some appreciable additional modeling effort, there are also some relatively straightforward foreseeable applications. Several accessory proteins known to strongly modulate cytoskeletal architecture, such as capping and branching molecules, have not been considered throughout this thesis [268, 271]. Studying entropy production in systems including these additional agents may allow to test the generality of the observed dissipation minimization, and they could additionally affect the cooperative mechanical relaxation leading to cytoquakes. Certain common cytoskeletal structures, such as stable cross-linked stress fibers and the cortical actomyosin shell, could also be reproduced in simulation and the concomitant dissipation signatures of these structures investigated [247]. Our computational method to track cytoskeletal thermodynamics is available for public use in the currently released version of MEDYAN, and we hope that continuation of the work presented in this thesis may provide new insights into cytoskeletal self-organization and its physiological role in the life of a cell.

Appendix A

Description of MEDYAN simulation platform

A detailed description and justification of the MEDYAN model can be found in Ref. 42, and additional extensions and applications of MEDYAN to study the dynamics of actomyosin networks are described in Refs. 94, 195, 247, 268–270. Here we briefly outline the relevant aspects of MEDYAN to facilitate understanding the results in this thesis, and direct the reader to the above references for a more thorough introduction.

A.1 Simulation protocol

A MEDYAN simulation proceeds by iterating a cycle of four steps which propagate the chemical and mechanical dynamics forward while maintaining a tight coupling between the two. The steps are as follows:

1. Evolve system using stochastic chemical simulation for a time δt.
2. Compute the changes in the mechanical energy resulting from the reactions that occurred in step 1).
3. Mechanically equilibrate the network in response to the new stresses from step 2).
4. Update the reaction rates of force-sensitive reactions based on the new tensions from step 3).

This protocol reflects a separation of timescales between the slow chemical dynamics and the fast mechanical response, such that the mechanical subsystem is assumed to always remain near equilibrium and to adiabatically follow the chemical changes in the network. As argued in Ref. 42, supported using experimental evidence from Refs. 12, 88, 239, this timescale separation holds for typical cytoskeletal networks which experience localized force deformations with fast relaxation times compared to the typical waiting time between myosin motor walking steps and filament growth-induced deformations.

A.2 Chemistry

In MEDYAN, diffusing chemical species are represented with discrete copy numbers belonging to several compartments, which form a regular grid comprising the simulation volume. The compartment size is chosen so that it may be assumed that inside the compartments the diffusing species are well-mixed, allowing the use of mass-action kinetics to determine their instantaneous propensities to participate in chemical reactions within compartments and diffusion events between adjacent compartments. The minimum Kuramoto length (i.e. the mean free diffusional path length of a reactive species before it participates in a chemical reaction) among the species sets this compartment size to ensure that the well-mixed assumption holds [310]. The diffusing chemical species may participate in local chemical reactions according to the copy numbers of the reactants in its compartment, or else they may jump to an adjacent compartment in a diffusion event with a propensity determined by its copy number in the original compartment [189]. The algorithm for stochastically choosing which event (including local reactions or jumps between compartments) will occur next is the Next Reaction Method, an accelerated variant of the Gillespie algorithm [189, 190]. These are Monte Carlo methods which randomly select both the time to any next event and which event will occur at that time in accordance with each event's instantaneous propensity.

The user specifies the different chemical species and the reactions that they participate in. Several types of reactions are possible. Regular reactions involve only diffusing species (e.g. the conversion of ADP-bound to ATP-bound G-actin monomer). Polymerization reactions result in the subtraction of a diffusing monomer from the local compartment and its conversion into a filament species, and depolymerization reactions do the opposite. Filaments in MEDYAN's have definite spatial coordinates, rather than the compartment-level description of the diffusing species' positions. This network of spatially resolved filaments is overlaid on the compartment grid, so that sections of filaments are able to react with diffusing species according to their local copy numbers. In addition, filaments have mechanical properties which will be discussed in the next section. A filament may react with a diffusing species such as a cross-linker (e.g. α-actinin), branching (e.g. Arp2/3), or molecular motor (e.g. NMIIA). Binding reactions involve a discrete set of binding sites along the filament, and they stochastically occur as chemical reaction events according to the number of those binding sites and the local copy number of diffusing binding molecules. A bound molecular motor may participate in a walking reaction, which causes it to move one of its ends to an adjacent binding site, stretching the motor and generating forces. Other reactions not used in this thesis but encompassed by MEDYAN include filament nucleation, filament destruction, filament severing, and filament branching reactions.

A.3 Mechanics

The mechanical energy U of networks in MEDYAN is a function of the positions of the filament beads and the lengths of the molecules bound to the filaments. There are also potentials describing a branched filament's energy which are not included in this thesis. Filament beads mark the joined end points (i.e. hinges) of the cylinders comprising the filament. Individual cylinders can stretch but not bend, but a bending energy term is included for pairs of adjacent cylinders. The energy term for the stretching of cylinders is

$$U_{\text{str}} = \frac{1}{2} K_{\text{fil,str}} (l - l_0)^2, \tag{A.1}$$

where $l = ||\mathbf{r}_{i+1} - \mathbf{r}_i||$ is the length of the cylinder whose beads are at positions $\mathbf{r}_{i+1}$ and $\mathbf{r}_i$, l_0 is the cylinder's equilibrium length, and K_{str} is the spring constant of this harmonic potential. The energy term for the bending of adjacent cylinders is

$$U_{\mathrm{bend}} = \epsilon_{\mathrm{bend}} \left(1 - \cos(\theta_{i,i+1})\right), \tag{A.2}$$

where ϵ_{bend} parameterizes the strength of the interaction and $\theta_{i,i+1}$ is the angle between the cylinders. Molecules bound to pairs of filaments (e.g. α-actinin and NMIIA) of stretched length l_{bound} have a harmonic stretching energy term:

$$U_{\mathrm{bound,str}} = \frac{1}{2} K_{\mathrm{bound,str}} (l_{\mathrm{bound}} - l^0_{\mathrm{bound}})^2, \tag{A.3}$$

where the subscript "bound" indicates that the variables and parameters are specific to the bound molecule. An excluded volume interaction is included to prevent cylinders from overlapping. The analytical formula for this interaction is complicated but can be expressed as a double integral over the two lengths of the participating cylinders i and j:

$$U_{\mathrm{vol},ij} = K_{\mathrm{vol}} \int_0^1 \int_0^1 \frac{ds dt}{||\mathbf{r}_i(s) - \mathbf{r}_j(t)||^4}, \tag{A.4}$$

where $\mathbf{r}_i(s) = \mathbf{r}_i + s(\mathbf{r}_{i+1} - \mathbf{r}_i)$ is the position along the i cylinder, which is parameterized by a variable s running from 0 to 1 along the cylinder's length. These positions $\mathbf{r}_i(s)$ are also therefore functions of the cylinders' bead positions, $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$. See Ref. 311 for a detailed derivation of the excluded volume potential. Finally, an exponentially decaying boundary repulsion term prevents the filaments from poking outside the simulation volume:

$$U_{\mathrm{boundary}} = \epsilon_{\mathrm{boundary}} e^{-d_i/\lambda}, \tag{A.5}$$

where $\epsilon_{\mathrm{boundary}}$ parameterizes the interaction strength, d_i is the distance from the boundary to the nearest endpoint of cylinder i, and λ parameterizes the interaction screening length.

At the end of each chemical evolution cycle, the positions of the bound molecules

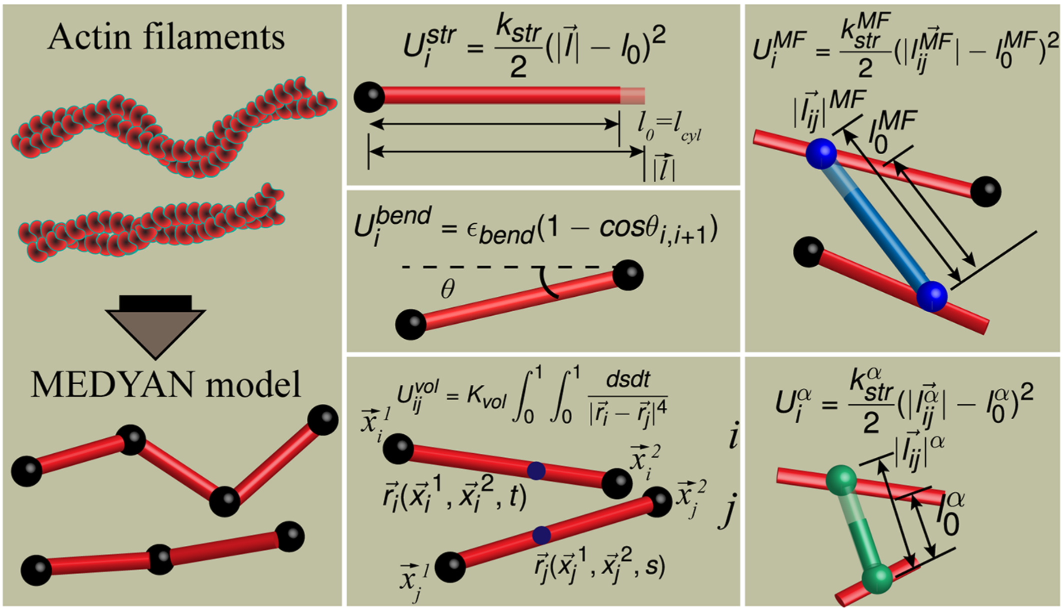

Figure A.1: Depiction of the mechanical potenetials used in MEDYAN. The double helical structure of filaments is coarse-grained into a sequence of cylinders (shown in red) whose connected hinges are called beads (shown as black spheres). Cylinders expereince a stretching potential (Equation A.3), a bending poential (Equation A.2), and an excluded volume repulsion (Equation A.4). Myosin minifilaments (shown in blue) and α-actinin cross-linkers experience a stretching potential (Equation A.3). Note that the mathematical notation differs slightly from what it used in the main text. Image reprinted from Ref. 247 with permission from *PLoS*.

and the filament beads will have changed due to the chemical reactions which occurred, displacing the system from near-equilibrium. The positions of the filament beads are then updated in a mechanical equilibration cycle by minimizing the total mechanical energy function U. This is accomplished using the conjugate-gradient minimization algorithm [312]. The minimization procedure ends when the maximum net force remaining in the network is below a user-specified force tolerance F_T, as result of which the system returns to near mechanical equilibrium.

A.4 Mechanochemical coupling

An important facet of the dynamics of actomyosin networks is that the chemical reaction rates of the associated proteins depend on the forces they sustain: at high tension the myosin minifilaments will walk and unbind more slowly (stalling and catch-bond behavior) whereas the passive cross-linkers are modeled as unbinding more quickly under tension (slip-bond behavior) [87, 267]. These force-sensitive behaviors thus play the role of nonlinearly coupling the mechanical state of the actomyosin network to its stochastic chemical dynamics.

The myosin motors used in MEDYAN are modeled after non-muscle myosin IIA (NMIIA), which exists in the cell as a minifilament consisting of tens of individual myosin heads. The chemical dynamics of the myosin minifilaments are based on the Parallel Cluster Model of Erdmann et al. [90, 93]. In this model, a myosin minifilament contains a number N_{total} of individual myosin heads and has a binding rate to the actin filament pair equal to

$$k_{\text{fil,bind}} = N_{\text{total}} k_{\text{head,bind}}, \tag{A.6}$$

where $k_{\text{head, bind}}$ is the individual myosin head binding rate. In MEDYAN, N_{total} is uniformly randomly selected between a minimum and maximum number of heads each time a minifilament binds. The bound myosin minifilament has a number of bound heads N^0_{bound} under zero tension equal to the duty ratio ρ times the total number of heads:

$$N^0_{\text{bound}} = \rho N_{\text{total}}. \tag{A.7}$$

The duty ratio is determined by the individual head unbinding rate:

$$\rho = \frac{k_{\text{head,bind}}}{k^0_{\text{head,unbind}} + k_{\text{head,bind}}}, \tag{A.8}$$

where $k^0_{\text{head,unbind}}$ is the head unbinding rate under zero tension. Under tension F_{ext} the bound myosin minifilament has altered walking and unbinding rates as well as an altered number of bound heads. The number of bound heads under tension is given by

$$N_{\text{bound}}(F_{\text{ext}}) = \min\left\{N_{\text{total}},\ N^0_{\text{bound}} + \beta\frac{F_{\text{ext}}}{N_{\text{total}}}\right\}, \tag{A.9}$$

where the parameter $\beta = 2.0$ is chosen to fit experimental data. The myosin minifilament walking rate under zero tension is

$$k^0_{\text{fil,walk}} = s\frac{1-\rho}{\rho}k_{\text{head,bind}}, \tag{A.10}$$

where s is called the stepping fraction, defined as the ratio of the user-specified real distance between binding sites on the filament d_{step} to the distance between binding sites on the computational cylinder representing the filament segment d_{total}: $s = \frac{d_{\text{step}}}{d_{\text{total}}}$. Equation A.10 is based on the PCM and is explained Refs. [42, 93]. Under tension, the myosin minifilament walking rate is altered according to a formula of the Hill type:

$$k_{\text{fil,walk}} = \max\left\{0.0,\ k^0_{\text{fil,walk}}\frac{F_{\text{stall}} - F_{\text{ext}}}{F_{\text{stall}} + F_{\text{ext}}/\alpha}\right\}, \tag{A.11}$$

where the stall force F_{stall} is the maximum tension a minifilament can sustain before it stops walking, and where $\alpha = 0.2$ is another parameter chosen to fit to experimental data. The myosin minifilament will unbind from the pair of actin filaments under zero tension with a rate

$$k^0_{\text{fil,unbind}} = \frac{k_{\text{head,bind}}N_{\text{total}}}{\exp\left(\log\left(\frac{k^0_{\text{head,unbind}}+k_{\text{head,bind}}}{k^0_{\text{head,unbind}}}\right)N_{\text{total}}\right) - 1}. \tag{A.12}$$

This non-obvious expression is the inverse of the mean residence time of the minifilament as determined using the PCM. Under tension, the myosin minifilament unbinding is modeled with Kramers-type catch-bond behavior:

$$k_{\text{fil,unbind}}(F_{\text{ext}}) = k^0_{\text{fil,unbind}}\max\left\{0.1,\ \exp\left(\frac{-F_{\text{ext}}}{N_{\text{bound}}(F_{\text{ext}})F_{0,\text{head}}}\right)\right\}, \tag{A.13}$$

where $F_{0,\text{head}}$ is the characteristic force a single myosin head catch-bond, and the minimum unbinding factor 0.1 is a parameter to chosen to ensure the possibility to unbind under arbitrarily large tension. We assume for myosin minifilaments that the stretching constant is given by

$$K_{\mathrm{bound,str}} = K_{\mathrm{head,str}} N_{\mathrm{bound}}, \tag{A.14}$$

where $K_{\mathrm{head,str}}$ is the stretching constant for an individual head; this equation assumes the bound heads share the load in parallel.

The unbinding of passive cross-linkers (e.g. α-actinin) are modeled as Kramers-type slip-bond:

$$k_{\mathrm{linker,unbind}}(F_{\mathrm{ext}}) = k^0_{\mathrm{linker,unbind}} \exp\left(\frac{F_{\mathrm{ext}}}{F_{0,\mathrm{linker}}}\right), \tag{A.15}$$

where $F_{0,\mathrm{linker}}$ is the characteristic force of the cross-linker slip-bond.

Finally, the actin filament will polymerize with a rate that exponentially decreases with the component of the sustained force along the polymerizing tip, F_{ext}. This dependence is based on the Brownian ratchet model of Peskin et al. [101]:

$$k_{\mathrm{poly}}(F_{\mathrm{ext}}) = k^0_{\mathrm{poly}} \exp\left(-\frac{F_{\mathrm{ext}}}{F_{0,\mathrm{poly}}}\right), \tag{A.16}$$

where $F_{0,\mathrm{poly}}$ is the characteristic force of the Brownian ratchet model, and k^0_{poly} is the zero-force polymerization rate.

Any of the above characteristic forces F_0 may be converted to a corresponding characteristic distance x_0 via

$$F_0 = k_B T / x_0, \tag{A.17}$$

where $k_B T$ is the thermal energy, casting expressions of the form F_{ext}/F_0 to the form $F_{\mathrm{ext}} x_0 / k_B T$.

Appendix B

Supporting information for Chapter 2

B.1 Details of BC model

The BC model consists of the following set of 11 coupled ODE's:

$$\begin{aligned}\frac{dG^T}{dt} &= k^T_{\text{off, +}} T^T_+ + k^T_{\text{off, -}} T^T_- + k_{\text{nex}} G^D - k_{\text{renex}} G^T \\ &\quad - G^T N(k^T_{\text{on, +}} + k^T_{\text{on, -}})\end{aligned} \tag{B.1}$$

$$\begin{aligned}\frac{dG^D}{dt} &= k^D_{\text{off, +}} T^D_+ + k^D_{\text{off, -}} T^D_- + k^{Pi}_{\text{off, +}} T^{Pi}_+ + k^{Pi}_{\text{off, -}} T^{Pi}_- + k_{\text{renex}} G^T \\ &\quad - k_{\text{nex}} G^D - G^D N(k^D_{\text{on, +}} + k^D_{\text{on, -}})\end{aligned} \tag{B.2}$$

$$\begin{aligned}\frac{dF^T}{dt} &= G^T N(k^T_{\text{on, +}} + k^T_{\text{on, -}}) - k^T_{\text{off, +}} T^T_+ - k^T_{\text{off, -}} T^T_- - k_{\text{hyd}} F^T \\ &\quad + k_{\text{rehyd}} F^{Pi}\end{aligned} \tag{B.3}$$

$$\begin{aligned}\frac{dF^{Pi}}{dt} &= k_{\text{hyd}} F^T - k_{\text{rehyd}} F^{Pi} - k^{Pi}_{\text{off, +}} T^{Pi}_+ - k^{Pi}_{\text{off, -}} T^{Pi}_- - k_{\text{phos}} F^{Pi} \\ &\quad + k_{\text{rephos}} F^D\end{aligned} \tag{B.4}$$

$$\begin{aligned}\frac{dF^D}{dt} &= G^D N(k^D_{\text{on, +}} + k^D_{\text{on, -}}) + k_{\text{phos}} F^{Pi} - k_{\text{rephos}} F^D - k^D_{\text{off, +}} T^D_+ \\ &\quad - k^D_{\text{off, -}} T^D_-\end{aligned} \tag{B.5}$$

$$\begin{aligned}\frac{dT_{\pm}^{T}}{dt} &= k_{\text{on},\,\pm}^{T} G^{T}(T_{\pm}^{Pi} + T_{\pm}^{D}) + (k_{\text{off},\,\pm}^{Pi} T_{\pm}^{Pi} + k_{\text{off},\,\pm}^{D} T_{\pm}^{D})\eta_{\pm}^{T} - k_{\text{hyd}} T_{\pm}^{T} \\ &\quad + k_{\text{rehyd}} T_{\pm}^{Pi} - k_{\text{off},\,\pm}^{T} T_{\pm}^{T}(\eta_{\pm}^{Pi} + \eta_{\pm}^{D}) - k_{\text{on},\,\pm}^{D} G^{D} T_{\pm}^{T} \end{aligned} \tag{B.6}$$

$$\begin{aligned}\frac{dT_{\pm}^{Pi}}{dt} &= k_{\text{hyd}} T_{\pm}^{T} - k_{\text{rehyd}} T_{\pm}^{Pi} + (k_{\text{off},\,\pm}^{T} T_{\pm}^{T} + k_{\text{off},\,\pm}^{D} T_{\pm}^{D})\eta_{\pm}^{Pi} - k_{\text{phos}} T_{\pm}^{Pi} \\ &\quad + k_{\text{rephos}} T_{\pm}^{D} - k_{\text{off},\,\pm}^{Pi} T_{\pm}^{Pi}(\eta_{\pm}^{T} + \eta_{\pm}^{D}) - T_{\pm}^{Pi}(k_{\text{on},\,\pm}^{T} G^{T} + k_{\text{on},\,\pm}^{D} G^{D}) \end{aligned} \tag{B.7}$$

$$\begin{aligned}\frac{dT_{\pm}^{D}}{dt} &= k_{\text{on},\,\pm}^{D} G^{D}(T_{\pm}^{T} + T_{\pm}^{Pi}) + (k_{\text{off},\,\pm}^{T} T_{\pm}^{T} + k_{\text{off},\,\pm}^{Pi} T_{\pm}^{Pi})\eta_{\pm}^{D} + k_{\text{phos}} T_{\pm}^{Pi} \\ &\quad - k_{\text{rephos}} T_{\pm}^{D} - k_{\text{off},\,\pm}^{D} T_{\pm}^{D}(\eta_{\pm}^{T} + \eta_{\pm}^{Pi}) - k_{\text{on},\,\pm}^{T} G^{T} T_{\pm}^{D} \end{aligned} \tag{B.8}$$

where

$$\eta_{\pm}^{T} = \frac{T_{\pm}^{T}}{N}\left(1 - \frac{T_{\pm}^{Pi}}{N}\right) \tag{B.9}$$

$$\eta_{\pm}^{D} = \frac{T_{\pm}^{D}}{N} \tag{B.10}$$

$$\eta_{\pm}^{Pi} = 1 - \eta_{\pm}^{T} - \eta_{\pm}^{D}. \tag{B.11}$$

In Table B.1 we list the meaning and values of the rate constants used in our implementation of the BC model.

Label	**Reaction**	**Value**
$k^{T}_{\text{on, +}}$	Polymerization of G^{T} to barbed end	11.6 μM^{-1} s^{-1}
$k^{T}_{\text{on, -}}$	Polymerization of G^{T} to pointed end	1.3 μM^{-1} s^{-1}
$k^{T}_{\text{off, +}}$	Depolymerization of F^{T} from barbed end	1.4 s^{-1}
$k^{T}_{\text{off, -}}$	Depolymerization of F^{T} from pointed end	0.8 s^{-1}
$k^{D}_{\text{on, +}}$	Polymerization of G^{D} to barbed end	2.9 μM^{-1} s^{-1}
$k^{D}_{\text{on, -}}$	Polymerization of G^{D} to pointed end	0.13 μM^{-1} s^{-1}
$k^{D}_{\text{off, +}}$	Depolymerization of F^{D} from barbed end	5.4 s^{-1}
$k^{D}_{\text{off, -}}$	Depolymerization of F^{D} from pointed end	0.25 s^{-1}
$k^{Pi}_{\text{off, +}}$	Depolymerization of F^{Pi} from barbed end	1.4 s^{-1}
$k^{Pi}_{\text{off, -}}$	Depolymerization of F^{Pi} from pointed end	0.8 s^{-1}
k^{hyd}	ATP hydrolysis converting F^{T} to F^{Pi}	0.3 s^{-1}
k^{rehyd}	ATP condensation converting F^{Pi} to F^{T}	ϵ 0.3 s^{-1}
k^{nex}	Nucleotide exchange converting G^{D} to G^{T}	0.01 s^{-1}
k^{renex}	Nucleotide exchange converting G^{T} to G^{D}	ϵ 0.01 s^{-1}
k^{phos}	Inorganic phosphate release converting F^{Pi} to F^{D}	0.002 s^{-1}
k^{phos}	Inorganic phosphate capture converting F^{D} to F^{Pi}	ϵ 0.002 s^{-1}

Table B.1: Rate constants in the BC model. The prefix “re” indicates the reverse of a kinetically dominant forward reaction (i.e. nearly irreversible reactions). The value of rate constants for these reverse reactions is taken to be equal to the corresponding forward reaction rate multiplied by a small parameter ϵ. We typically take $\epsilon = 0.01$. Some of these reactions, such as the nucleotide exchange reaction, are second order reactions. For example the proper rate of reaction for conversion of G^{D} to G^{T} is $k^{*}_{\text{nex}}[G^{D}][ATP]$. We treat such cases as pseudo-first order reactions by assuming that the concentration of the species which we don't track is constant and that its concentration is contained in the rate constant used. Thus $k_{\text{nex}} = k^{*}_{\text{nex}}[ATP]$ in our model. This assumption of constant concentrations of free ATP, ADP, and Pi is reasonable in cellular environments. All values are taken from Refs. 12, 183.

We note that the original equations in the BC model did not include reversible reactions as shown here. This amounts to setting k_{renex}, k_{rehyd}, and k_{rephos} to 0 Equations B.1-B.8. The interpretation of η^{i}_{j} is the probability that the subunit adjacent to the j tip is in the i hydrolysis state. The equations of motion of these variables in principle depend on the hydrolysis state of the subunit next to them toward the center of the filament, and Equations B.1-B.8 represent the truncation of this set of recursion equations. This is accomplished by assuming that η^{i}_{j} depends only on the tip subunit hydrolysis state through Equations B.6-B.8. These equations were arrived at by inspecting the time

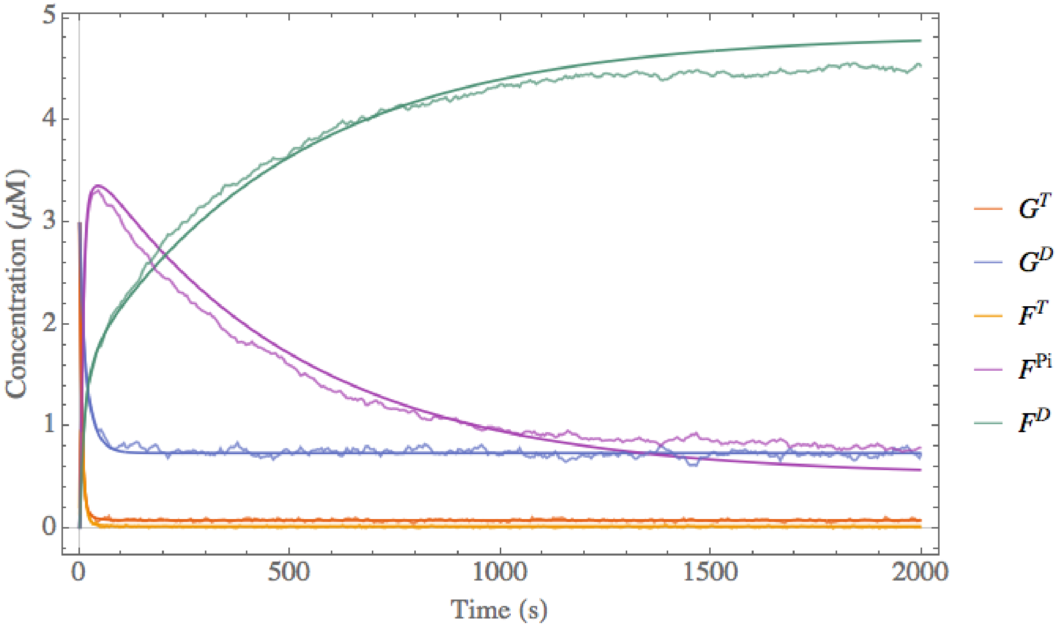

Figure B.1: Time course of the concentrations of the various species following the addition of 3 μM of G^T and 3 μM of G^D actin to a bath with a number concentration $N = 0.017$ μM of seed filaments. The mean-field BC (smooth curve) model accurately predicts the shape of the time-courses resulting from the stochastic simulation (noisy curve).

course of the η_j^i and the T_j^i variables and discerning the equations which approximately related the two. The system of Equations B.1-B.8 does not admit and analytical solution but can be numerically integrated with appropriate initial conditions specified [183].

We check the accuracy of the truncation assumption of the BC model by comparing a predicted time course to the results of a simulation using the software package MEDYAN. MEDYAN was developed to perform coarse-grained simulations of active networks and combines stochastic chemical algorithms with detailed mechanics as well as coupling between reaction rates of force-sensitive chemical reactions and the mechanical state of the species involved [42]. In Figures B.1 and B.2 we show the simulated time courses as well as the mean-field prediction of the BC model.

The steady state vector $\mathbf{x}^{\mathrm{ss}}$ satisfying $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = \mathbf{0}$ can be found by numerically solving the root of the right hand sides of Equations B.1-B.8. Equations B.1-B.5 sum to zero, reflecting the conservation of total actin M encoded in these reactions. Also each of the two sets of Equations B.6-B.8 sum to zero, reflecting the conservation of plus end

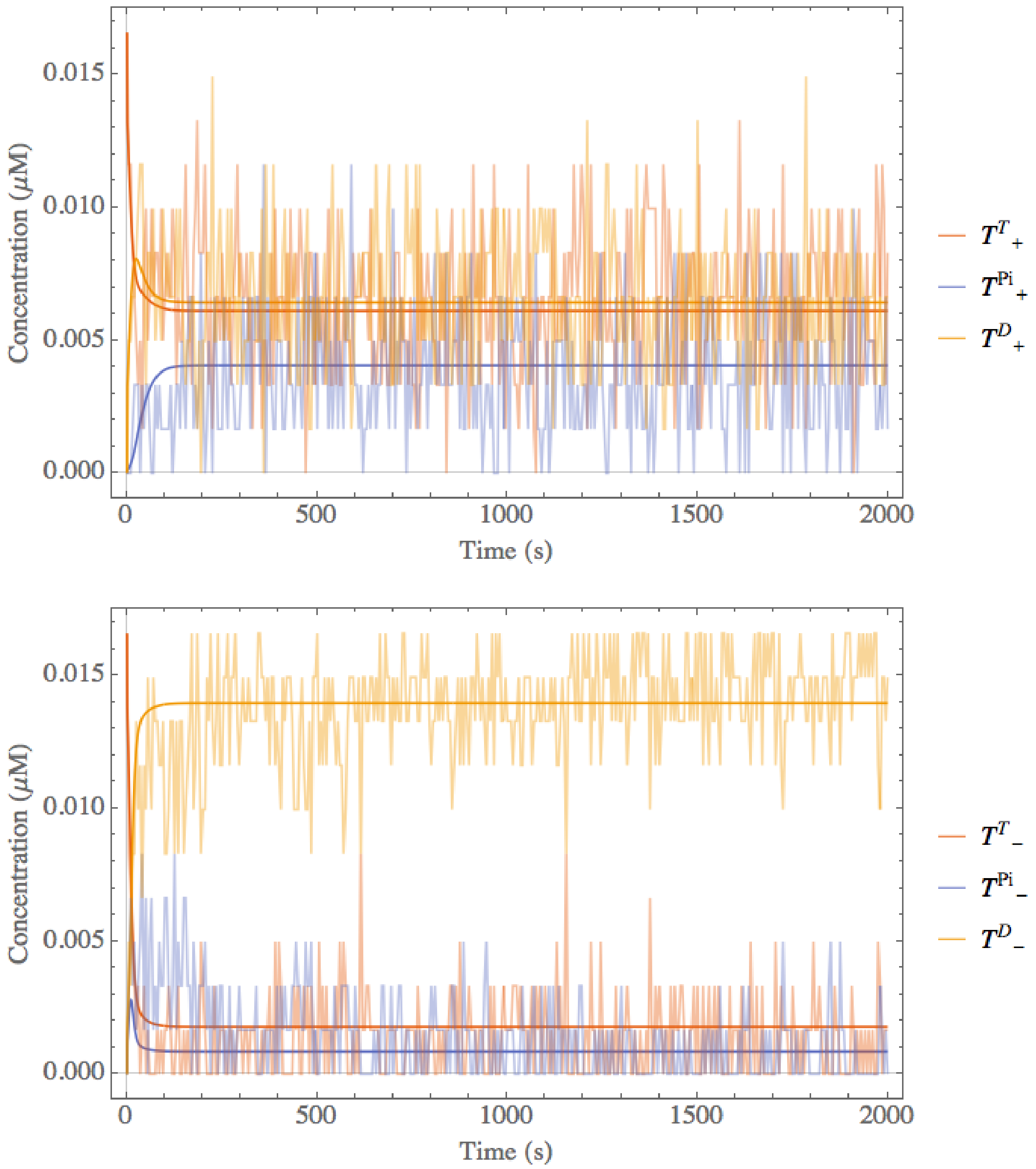

Figure B.2: Time course of the concentrations of the various tip species at the plus (A) and minus (B) ends, following the addition of 3 μM of G^T and 3 μM of G^D actin to a bath with a number concentration $N = 0.017$ μM of seed filaments. The noisiness of the stochastic trajectories results from the small copy number of filaments. Note how quickly the these concentrations attain their steady state values with this parameterization. This implies that not tracking these variables explicitly, as in the QSSA and CT models is valid for most of the trajectory.

tips and minus end tips. Thus the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ represented by Equations B.1-B.8 is linearly dependent, and no solution to exists $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = 0$ unless we specify additional equations. These additional equations are

$$G^T + G^D + F^T + F^{Pi} + F^D = M \tag{B.12}$$

$$T^T_\pm + T^{Pi}_\pm + T^D_\pm = N_\pm. \tag{B.13}$$

For unbranched filaments considered here, the number of plus end tips is equal to the number of minus end tips: $N_+ = N_- = N$. Solving $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = 0$ gives one solution for which all variables are non-negative, so there is a unique realistic steady state solution for a given set of parameters M and N. The eigenvalues of the BC Jacobian evaluated at the equilibrium point $\mathbf{J}^* = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\big|_{\mathbf{x}=\mathbf{x}^{\mathrm{ss}}}$ indicate the stability of the steady state. 3 of the 11 eigenvalues are zero, corresponding to the 3 linear conservation laws in our system. This implies that the dynamics of the BC model lie on an 8-dimensional submanifold of the full 11-dimensional variable space, and this submanifold is determined by the parameters M and N. The remaining eigenvalues are negative, indicating that the unique non-negative vector $\mathbf{x}^{\mathrm{ss}}$ is attracting and stable. We note that this equilibrium point of the dynamics actually corresponds to a non-equilibrium steady state of the chemical system, since this state corresponds to actin treadmilling, fueled by ATP hydrolysis. This chemical driving is manifested in the values of certain pseudo-first order reaction rate constants such as k_{nex}.

B.2 Steady state concentrations in the BC model

We show here that the independence of the steady state concentrations of some species on M is also present in the BC model. We do this by consideration of the system of algebraic equations $\mathbf{f}(\mathbf{x}^{\mathrm{ss}}) = \mathbf{0}$. The Jacobian matrix of system at the steady state (see Equation 2.24) is visualized in Figure B.3. Now that we can grasp which species are coupled to which other species, we observe that no species depend on F^T, F^{Pi}, or F^D, except those species themselves. Therefore we could ignore the equations of motions of these variables

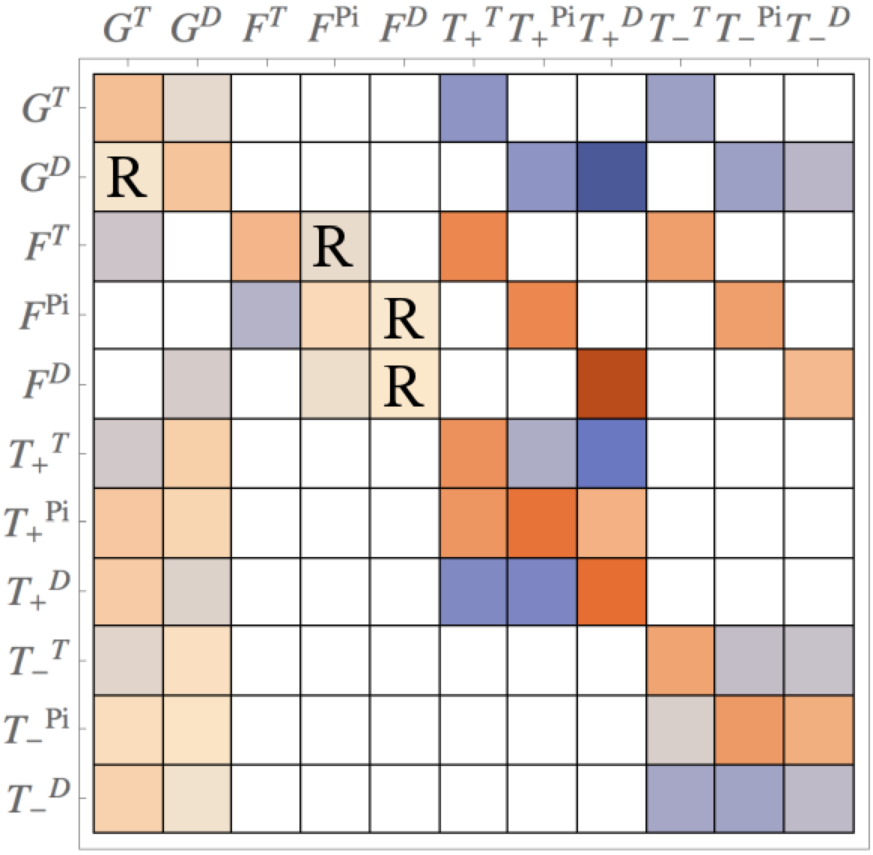

Figure B.3: Visualization of $\mathbf{J}^*$ to understand the couplings in the BC model. If $\frac{\partial x_i}{\partial x_j} \neq 0$, then J^*_{ij} will be nonzero. If this coupling causes x_i to increase, the entry in the matrix is colored blue, and if it causes x_i to decrease, the entry is red. The saturation of the color indicates the magnitude of the coupling. An “R” is placed in the plot if that element is nonzero only due to the inclusion of slow reversible reactions, i.e. when $\epsilon > 0$. The corresponding plot for the CT model is identical to the top left 5 by 5 matrix in this plot.

and the resulting subsystem of equations would be closed. Now that subsystem will be linearly dependent, but we can supplement it with the following equations to make it independent:

$$T_{\pm}^{T} + T_{\pm}^{Pi} + T_{\pm}^{D} = N. \tag{B.14}$$

The subsystem of equations could now be solved, giving the steady state concentrations of all species except for F^T, F^{Pi}, and F^D. M, the total amount of actin, does not enter into any of the equations of the subsystem, and so the solution of that system does not depend on M. Therefore, the steady state concentration of each species except F^T, F^{Pi}, and F^D does not depend on the initial conditions, however they do depend on the parameter N.

Now we consider the case where $\epsilon = 0$. Referring to Figure B.3, we see that this would mean that no species depends on F^D. By reasoning similar to the above, this implies that the steady state concentration of each species except F^D could be determined without specifying M. Thus, the steady state concentration of only F^D depends on the initial conditions in this case.

Appendix C

Supporting information for Chapter 3

C.1 Discrete variables and the relation to the Gibbs-Duhem equation

The Gibbs-Duhem Equation from classical thermodynamics states

$$\sum_{i \in M} N_i d\mu_i = -SdT + Vdp = 0 \tag{C.1}$$

where M represents all chemical species in the system including the solvent, and the last equality holds at constant temperature and pressure [25]. Applying the product rule[1] to the expression for the Gibbs free energy, $G = \sum_{i \in M} N_i \mu_i$, we get

$$dG = \sum_{i \in M} \mu_i dN_i + \sum_{i \in M} N_i d\mu_i = \sum_{i \in M} \mu_i dN_i. \tag{C.2}$$

Evaluating $\sum_{i \in M} \mu_i dN_i$ using stoichiometric coefficients in place of dN_i can be shown to lead to the expression $\Delta G_{(3)}$.

If the copy numbers are small, we are better served using the discrete difference operator Δ and not the differential difference operator d. The product rule for the

---

[1] The product rule can be derived as $d(fg) = (f + df)(g + dg) - fg = fdg + gdf + dfdg = fdg + gdf$ since the term $dfdg$ is assumed to be negligible. We give this reminder to emphasize the presence of the cross term, which we do not neglect in the discrete case.

discrete difference operator applied to G gives

$$\Delta G = \sum_{i\in M} \mu_i \Delta N_i + \sum_{i\in M} N_i \Delta\mu_i + \sum_{i\in M} \Delta\mu_i \Delta N_i. \tag{C.3}$$

Here we do not neglect the cross-term as we do in the differential product rule, Equation C.2. We evaluate this expression term by term:

$$\sum_{i\in M} \mu_i \Delta N_i = \Delta G_{(3)} \tag{C.4}$$

$$\sum_{i\in M} N_i \Delta\mu_i = k_B T \log \prod_{i\in R} \left(\frac{N_i - \nu_i}{N_i}\right)^{N_i} \prod_{j\in P} \left(\frac{N_i + \upsilon_j}{N_j}\right)^{N_j} \left(\frac{N}{N+\sigma}\right)^{N} \tag{C.5}$$

$$\sum_{i\in M} \Delta N_i \Delta\mu_i = k_B T \log \prod_{i\in R} \left(\frac{N_i - \nu_i}{N_i}\right)^{-\nu_i} \prod_{j\in P} \left(\frac{N_i + \upsilon_j}{N_j}\right)^{\upsilon_j} \left(\frac{N}{N+\sigma}\right)^{\sigma} \tag{C.6}$$

Comparing these to the expressions for the accuracy of the various approximations given in Appendix B, we have

$$\sum_{i\in M} N_i \Delta\mu_i + \sum_{i\in M} \Delta N_i \Delta\mu_i = \Delta G_{(0)} - \Delta G_{(3)} \tag{C.7}$$

and thus, combining all the terms in Equation C.3, we recover our exact expression $\Delta G_{(0)}$, and this approach can be seen as an alternate derivation of the main result. To summarize, for small, discrete copy numbers, we obtain corrections to the expression for the change in Gibbs free energy accompanying chemical reactions that are not captured by the constraints imposed by the Gibbs-Duhem Equation, which assumes that the copy numbers are continuous quantities.

C.2 Accuracy of the approximations

We calculate the accuracy of the approximations $\Delta G_{(1)}, \Delta G_{(2)}$, and $\Delta G_{(3)}$ by taking the difference of these quantities with $\Delta G_{(0)}$. We have

$$(\Delta G_{(1)} - \Delta G_{(0)})/k_B T = \log \left(\frac{N+\sigma}{N} \right)^{N+\sigma} - \sigma \tag{C.8}$$

$$(\Delta G_{(2)} - \Delta G_{(0)})/k_B T = \log \left(\frac{N+\sigma}{N} \right)^{N+\sigma} + \log \prod_{i \in R} \left(\frac{N_i}{N_i - \nu_i} \right)^{N_i} \prod_{j \in P} \left(\frac{N_j}{N_j + \upsilon_j} \right)^{N_j} \tag{C.9}$$

$$(\Delta G_{(3)} - \Delta G_{(0)})/k_B T = \log \left(\frac{N+\sigma}{N} \right)^{N+\sigma} + \log \prod_{i \in R} \left(\frac{N_i}{N_i - \nu_i} \right)^{N_i - \nu_i} \prod_{j \in P} \left(\frac{N_j}{N_j + \upsilon_j} \right)^{N_j + \upsilon_j} \tag{C.10}$$

The accuracy of $\Delta G_{(1)}$ depends only on N for a given reaction, whereas the remaining approximations depend also on the values of N_i and N_j. These observations reflect the fact that, in order to arrive at the expression for $\Delta G_{(1)}$, we leveraged the size of N compared to σ, and to arrive at the expressions for $\Delta G_{(2)}$ and $\Delta G_{(3)}$ we further successively leveraged the sizes of N_i, N_j compared to ν_i, ν_j.

C.3 ΔG of solvent fluctuations

Here we first calculate an approximate expression for the change in Gibbs free energy of the system accompanying a fluctuation of n solvent molecules from compartment A to compartment B. We have

$$\begin{aligned} G^{\text{initial}} &= \sum_{i \in L} N_{i,A} \left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,A}}{N_A} \right) + N_{s,A} \left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,A}}{N_A} \right) \\ &\quad + \sum_{i \in L} N_{i,B} \left(\widetilde{\mu_i^0} + k_B T \log \frac{N_{i,B}}{N_B} \right) \\ &\quad + N_{s,B} \left(\widetilde{\mu_s^{0,*}} + k_B T \log \frac{N_{s,B}}{N_B} \right), \end{aligned} \tag{C.11}$$

where L is the set of solute species. After the transfer of n solvent molecules from A to B we have

$$\begin{aligned} G^{\text{final}} &= \sum_{i\in L} N_{i,A}\left(\widetilde{\mu_i^0} + k_B T\log\frac{N_{i,A}}{N_A - n}\right) + N_{s,A}\left(\widetilde{\mu_s^{0,*}} + k_B T\log\frac{N_{s,A}-n}{N_A - n}\right) \\ &\quad + \sum_{i\in L} N_{i,B}\left(\widetilde{\mu_i^0} + k_B T\log\frac{N_{i,B}}{N_B + n}\right) \\ &\quad + N_{s,B}\left(\widetilde{\mu_s^{0,*}} + k_B T\log\frac{N_{s,B}+n}{N_B + n}\right). \end{aligned} \tag{C.12}$$

Taking the difference and simplifying, we arrive at the exact expression

$$\Delta G = k_B T \log \frac{N_A^{N_A}}{(N_A - n)^{N_A - n}} \frac{(N_{s,A} - n)^{N_{s,A}-n}}{N_{s,A}^{N_{s,A}}} \frac{N_B^{N_B}}{(N_A + n)^{N_A + n}} \frac{(N_{s,B} + n)^{N_{s,B}+n}}{N_{s,B}^{N_{s,B}}}. \tag{C.13}$$

To understand the magnitude of ΔG for typical values of n, $N_{s,A}$, and $N_{s,B}$ compared to N_A and N_B, we first make the assumption that the two compartments initially have the same number of solvent molecules, i.e. that $N_{s,A} = N_{s,B} \equiv N_s$. Next we assume that the solvent molecules dominate the proportion of total molecules, allowing us to write $N_A \approx N_B \equiv N$. These approximations will hold in the limit that the number of solute molecules is much less than the number of solvent molecules for each compartment. We next introduce the small parameters

$$\epsilon_A = \frac{\sum_{i\in L} N_{i,A}}{N}, \tag{C.14}$$

and

$$\epsilon_B = \frac{\sum_{i\in L} N_{i,B}}{N}, \tag{C.15}$$

which capture the dilutions of the two compartments, and

$$\xi = \frac{n}{N}, \tag{C.16}$$

which represents the relative size of the fluctuation. For $N = 10^9$, we typically have $\xi \sim 10^{-4.5}$ (since $n \sim N^{1/2}$) and $\epsilon \sim 10^{-6}$. Substituting these parameters in Equation

C.13, we have

$$\begin{aligned}\Delta G &= k_BT\log\left(\frac{N^{2N}}{(N(1-\xi))^{N(1-\xi)}(N(1+\xi))^{N(1+\xi)}}\times\right.\\ &\qquad\left.\frac{(N(1-\epsilon_A-\xi))^{N(1-\epsilon_A-\xi)}(N(1-\epsilon_B+\xi))^{N(1-\epsilon_B+\xi)}}{(N(1-\epsilon_A))^{N(1-\epsilon_A)}(N(1-\epsilon_B))^{N(1-\epsilon_B)}}\right). \qquad \text{(C.17)}\end{aligned}$$

We expand this expression to first order in ϵ_A and ϵ_B, and then to second order in ξ. The result is

$$\begin{aligned}\Delta G &\approx (\epsilon_A-\epsilon_B)N\xi k_BT+\frac{1}{2}(\epsilon_A+\epsilon_B)N\xi^2k_BT\\ &= (\epsilon_A-\epsilon_B)nk_BT+\frac{1}{2}(\epsilon_A+\epsilon_B)\frac{n^2}{N}k_BT. \qquad \text{(C.18)}\end{aligned}$$

The observed numerical agreement between Equations C.13 and Equations C.18 is close for realistic values of the parameters: for $N_A = 10^9$, $N_B = 5\times10^8$, $\epsilon_A = 10^{-6}$, $\epsilon_B = 3\times10^{-6}$, and $n = N_A^{1/2}$, the prediction of Equation C.13 is $-0.0632401\ k_BT$, and the prediction of Equation C.18 is $-0.0632436\ k_BT$. The first term in Equation C.18 will dominate if the solute dilutions in the two compartments are very different from each other. In this case, we may compare the size of this change in Gibbs free energy to that accompanying the diffusion of a solute from compartment B to compartment A. This latter change in Gibbs free energy will be approximately $k_BT\log\frac{\epsilon_A}{\epsilon_B}$. If we now set $\epsilon_A = a\epsilon_B$, where a is of order 1 (typically it will fall in the range $[1/10, 10]$), then ΔG for the solvent fluctuation will be $k_BT(a-1)\epsilon_Bn$ and ΔG for the solute diffusion will be $k_BT\log a$. The product ϵ_Bn will typically on the order of $\sim 10^{-1.5}$, so one can see that for typical values of the parameters the change in Gibbs free energy from a solute diffusion event will be significantly greater in magnitude than that from a solvent fluctuation. If the dilutions are very similar, then $\epsilon_A-\epsilon_B\approx 0$, and the second term in Equation C.18 dominates. This term is on the order of $(\epsilon_A+\epsilon_B)k_BT$ since $\frac{n^2}{N}\sim 1$. These changes in Gibbs free energy will typically be much smaller than those accompanying a chemical reaction or inter-compartment diffusion of the solute. Thus we may neglect the activity of the solvent in tracking the Gibbs free energy of the system.

Appendix D

Supporting information for Chapter 4

D.1 Parameterization

Parameterization of the CT model for the purpose of tracking free energy changes during simulation trajectories consists of choosing values of the rate constants (kinetic parameters) and of ΔG^0 (thermodynamic parameters) for all reactions in the model. Wherever possible, values from the literature are used. Experimental measurements have determined rate constants for every reaction, however for some reactions the value of ΔG^0 hasn't been reliably measured, to the best of our knowledge. Below, we describe a technique to solve for these unknown values.

For reversible reactions, where the forward and reverse rate constants k_+ and k_- are known, such as (de)polymerization and (un)binding of cross-linkers, ΔG^0 can be found from

$$\Delta G^0 = k_B T \log K_{\mathrm{eq}} = k_B T \log \frac{k_-}{k_+} \tag{D.1}$$

Literature values for the equilibrium constants or of ΔG^0 are used for irreversible reactions, for which k_- is often too small to determine from direct measurement. Irreversible reactions in this system include all reactions except for (de)polymerization reactions. For reactions for which literature values of ΔG^0 are unavailable, it is possible to solve for ΔG^0 values based on a self-consistency condition [98, 180]: the sum of the ΔG^0 values around a closed loop of reactions in which the number of molecules has not experienced a net change must be zero since the free energy is a state function (equivalently, by Equation D.1, the product of equilibrium constants around any such loop must be equal to one). Writing several such closed loops of reactions leads to a system of equations that can be solved for the unknown variables. Not all possible loops result in independent equations, but we were able to determine the values of two unknown parameters using the loops illustrated in Figure D.1.

The loops in Figure D.1 imply the following independent system of equations:

$$K^T K_{\text{hyd}} K_{\text{rel}} K_{\text{nex}} = K^{Pi} K_{\text{ATP}} \tag{D.2}$$

$$K^T K_{\text{hyd}} K_{\text{phos}} K_{\text{nex}} = K^D K_{\text{ATP}} \tag{D.3}$$

where K_{hyd} represents the hydrolysis of ATP by F^T, K_{phos} represents the release of Pi by F^{Pi}, K_{nex} represents nucleotide exchange converting G^D to G^T, K_{ATP} represents the hydrolysis of ATP in solution producing ADP and Pi, and K_{rel} represents the release of phosphate by G^{Pi}.

Values from the literature [12, 89, 91] can be used to determine 6 of the 8 variables in Equations D.2 and D.3, which thus represent two equations in two unknowns: K_{rel} and K_{nex}. The resulting parameters are listed in Table D.1.

Note that it is possible to draw loops such as those in Figure D.1 that would imply that the equilibrium constants for polymerization and depolymerization of, for example G^T, should be the same at the plus and minus ends of the filaments. This condition is not borne out by the experimental values of these equilibrium constants, and this discrepancy is a recognized outstanding problem [12]. Here, we use the literature values for these equilibrium constants and employ the reaction loop method only to determine the parameters K_{rel} and K_{nex}.

For reversible binding of myosin filaments, results from the PCM are used to describe binding and unbinding rates, and therefore K_{eq} [90, 93]. The filament binding rate is

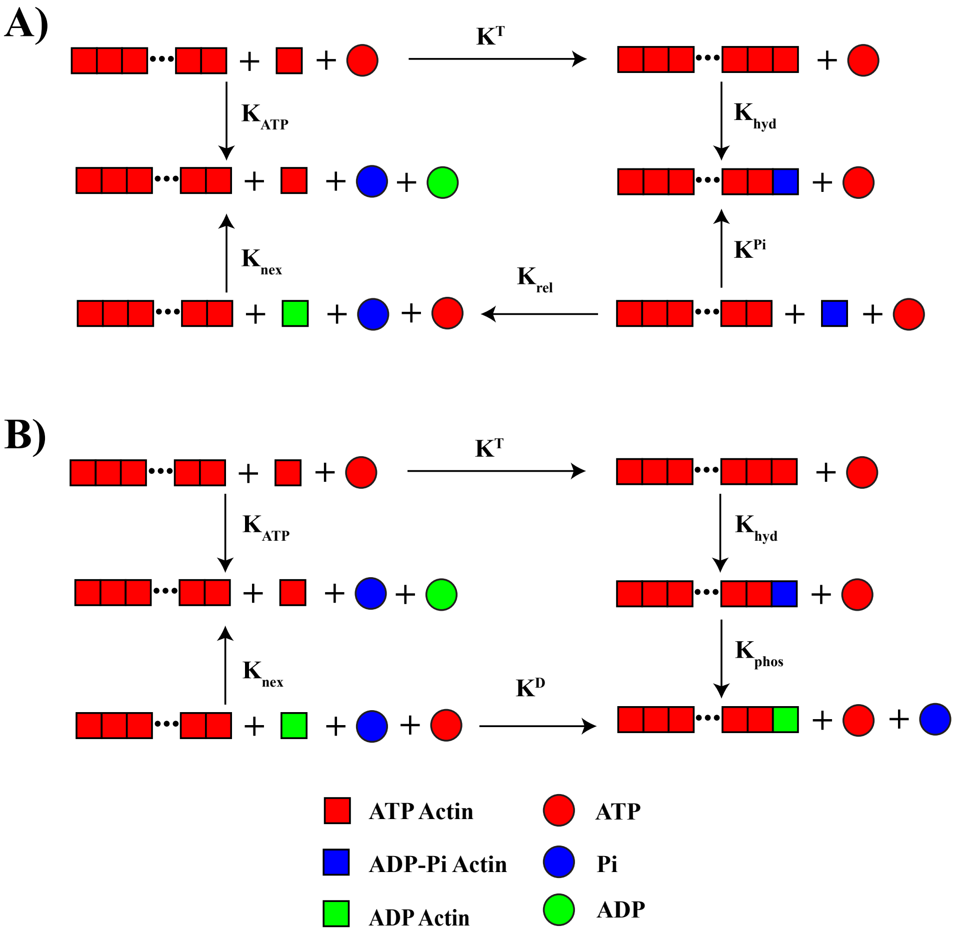

Figure D.1: Diagrams representing sequences of reactions, with the involved species drawn between each reaction, resulting in independent relations between the equilibrium constants. The meaning of these equilibrium constants is provided in the main text. Loops are assumed to proceed in the clockwise direction, and arrows that point opposite to this direction indicate that that reaction is occurring “backwards”, i.e. from products to reactants. Polymers are shown as connected chains of subunits, while the “+” sign indicates different solvated species.

given as

$$k_{\text{fil, bind}} = N_t k_{\text{head, bind}}, \tag{D.4}$$

and the unbinding rate is

$$k_{\text{fil,unbind}} = N_t k_{\text{head, bind}} \left[\left(\frac{k_{\text{head, bind}} + k_{\text{head, unbind}}}{k_{\text{head, unbind}}} \right)^{N_t} - 1 \right]^{-1}, \tag{D.5}$$

where $k_{\text{head, bind}}$ and $k_{\text{head, unbind}}$ describe the binding kinetics of a single myosin head, and N_t is the number of heads in the filament (c.f. Appendix A) [93]. The resulting expression for ΔG^0 is

$$\Delta G^0 = kT \frac{k_{\text{fil, unbind}}}{k_{\text{fil, bind}}} = -kT \log \left[\left(\frac{k_{\text{head, bind}} + k_{\text{head, unbind}}}{k_{\text{head, unbind}}} \right)^{N_t} - 1 \right]. \tag{D.6}$$

We assume that we have chemostatted concentrations of ATP, ADP, and Pi, which we account for implicitly via the effect of these concentrations on the kinetic and thermodynamic parameters of certain reactions. Thus these species are not explicitly tracked. The concentrations of these species affect the change in Gibbs free energy associated with the following reactions:

- myosin filament walking
- nucleotide exchange ($G^D \to G^T$)
- phosphate release by F and G-actin ($F^{Pi} \to F^D$, $G^{Pi} \to G^D$)

The effect of the concentrations of ATP, ADP, and Pi for the these reactions is to simply change the reaction quotient Q which changes ΔG via $\Delta G = \Delta G^0 + k_B T \log Q$. For instance the nucleotide exchange reaction can be written explicitly as

$$ATP + G^D \to ADP + G^T \tag{D.7}$$

and the change in Gibbs free energy is

$$\Delta G = \Delta G^0 + k_B T \log \left(\frac{C_{ADP} C_{G^T}}{C_{ATP} C_{G^D}} \right). \tag{D.8}$$

To treat the concentrations of ATP and ADP implicitly, we rewrite the reaction as

$$G^D \to G^T \tag{D.9}$$

for which the change in Gibbs free energy is

$$\Delta G = \Delta G^{0\prime} + k_B T \log\left(\frac{C_{G^T}}{C_{G^D}}\right), \tag{D.10}$$

where

$$\Delta G^{0\prime} = \Delta G^0 + k_B T \log\left(\frac{C_{ADP}}{C_{ATP}}\right). \tag{D.11}$$

A similar approach is taken for the other reactions mentioned above.

The concentration of just ATP (since neither ADP or Pi appear implicitly as reactants in any of the reactions of the CT model) affects the kinetics of the following reactions:

- myosin filament walking
- nucleotide exchange

To understand the effect of C_{ATP} on the myosin filament walking rate, we employ the five state cross-bridge model of a single myosin head described in Ref. 90. In that model unbinding of a head from the filament substrate occurs via two pathways: a slip path, with rate k_{35}, and a catch path, with an effective rate k_{345}. The catch path is a two-step reaction: the release of ADP with rate k_{34}, followed by the unbinding of the filament head and binding of ATP with rate $k_{45} = k_T C_{ATP}$. The effective rate constant of the approximate one-step representation of this reaction is

$$k_{345} = \frac{k_{34} k_T C_{ATP}}{k_{34} + k_T C_{ATP}}. \tag{D.12}$$

Unless C_{ATP} is very low, this reaction rate is limited by k_{34}. Because the head can unbind by the catch or slip pathway, the rate for unbinding is

$$k_{\text{head, unbind}} = k_{35} + \frac{k_{34} k_T C_{ATP}}{k_{34} + k_T C_{ATP}} \tag{D.13}$$

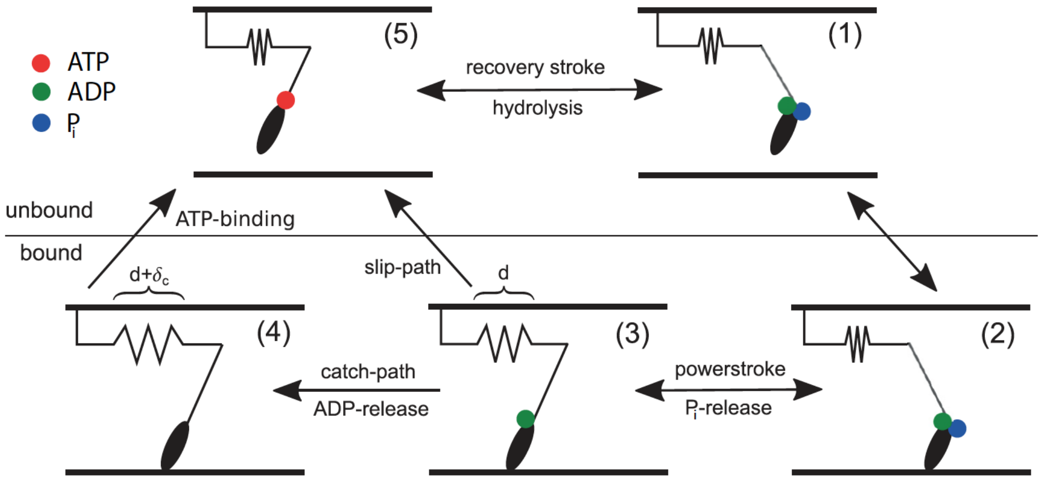

FIGURE D.2: A discrete model of the sequence of chemical reactions constituting the myosin head power-stroke cycle. In a minifilament, tens of such heads work in tandem to processively walk along the actin substrate. Image reprinted from Ref. [90] with permission from the *American Physical Society.*

The slip path should also be dependent on ATP concentration, since in state 5 the head is ATP-bound, however we follow the authors of [90] in neglecting this dependence since the slip path only becomes active under large load.

The rate of the nucleotide exchange reaction also depends on C_{ATP} and also occurs in two steps. The full reaction, with ATP and ADP explicitly included, is

$$G^D + ATP \rightarrow G^* + ATP + ADP \rightarrow G^T + ADP \tag{D.14}$$

where G^* represents actin with no bound nucleotide. Following [183], this reaction is approximated as a one-step irreversible reaction with ATP and ADP included explicitly, Equation D.9. We can write the rate k_{nex} of this approximate reaction as:

$$k_{\text{nex}} = \frac{k_{D\to *} k_{*\to T} C_{ATP}}{k_{D\to *} + k_{*\to T} C_{ATP}} \approx k_{D\to *} \tag{D.15}$$

where $k_{D\to *}$ is the dissociation rate of ADP, $k_{*\to T}$ is the second order rate constant of ATP association, and the approximation holds except at low concentrations of ATP when ADP dissociation is no longer the rate-limiting step. The values of $k_{*\to T}$ and $k_{D\to *}$ are presented and discussed in Refs. 13, 15, 313.

Reaction	**Rate Constant**	ΔG^0 ($k_B T$)
G^T poly at plus end	11.6 $(\mu M s)^{-1}$	-2.12 [a] [183]
G^T depoly at plus end	1.4 s^{-1}	2.12 [a] [183]
G^T poly at minus end	1.3 $(\mu M s)^{-1}$	-0.51 [a] [183]
G^T depoly at minus end	0.8 s^{-1}	0.51 [a] [183]
G^{Pi} poly at plus end	3.4 $(\mu M s)^{-1}$	-2.81 [a] [12]
G^{Pi} depoly at plus end	0.2 s^{-1}	2.81 [a] [12]
G^{Pi} poly at minus end	0.11 $(\mu M s)^{-1}$	-1.75 [a] [12]
G^{Pi} depoly at minus end	0.02 s^{-1}	1.75 [a] [12]
G^D poly at plus end	2.9 $(\mu M s)^{-1}$	0.59 [a] [183]
G^D depoly at plus end	5.4 s^{-1}	-0.59 [a] [183]
G^D poly at minus end	0.09 $(\mu M s)^{-1}$	1.03 [a] [183]
G^D depoly at minus end	0.25 s^{-1}	-1.03 [a] [183]
Pi release by F-actin	0.002 s^{-1}	-7.31 [b,e] [12, 183]
ATP hydrolysis by F-actin	0.3 s^{-1}	-10.0 [91, 183]
Pi release by G-actin	5 s^{-1}	-10.77 [c,e] -
Nucleotide exchange	0.01 s^{-1} [e]	-6.76 [c,e] [183]
Cross-linker binding	0.7 $(\mu M s)^{-1}$	-0.85 [a] [42]
Cross-linker unbinding	0.3 s^{-1}	0.85 [a] [42]
Myosin head binding	0.2 s^{-1} [d]	- [d] [42]
Myosin head unbinding	1.708 s^{-1} [d]	- [d] [42]
Myosin filament walking	- [f]	-14.5 [e] [209]

TABLE D.1: Kinetics and thermodynamic parameters describing reactions in the CT model as well as cross-linker and myosin filament (un)binding and myosin filament walking.

[a] - Values of ΔG^0 determined via Equation D.1.
[b] - ΔG^0 determined from K_{eq} given in Ref. 12.
[c] - ΔG^0 determined using constraints as described above.
[d] - Parameters describing the myosin filament obtained via Equations D.4, D.5, D.6.
[e] - Depends implicitly on C_{ATP}, C_{ADP}, and C_{Pi}; given value applies to standard state.

[f] - Calculated in simulation using results of PCM, see [42].

In all the studies in this thesis, implicit nucleotide concentrations are taken to be $C_{ATP} = 8\ mM$, $C_{ADP} = 7\ \mu M$, and $C_{Pi} = 1\ mM$, corresponding roughly to the amounts found in human muscle after exercise [209]. Other parameters of the system, including mechanical constants, diffusion rates, screening lengths, and boundary cutoffs are given below.

Parameter	**Description**	**Value**
	General Simulation Parameters	
k_BT	Thermal energy	$4.1\ pN \cdot nm$
L_{comp}	Cubic compartment side length	$500\ nm$
$N_x,\ N_y,\ N_z$	Number of compartments in each dimension	$2,\ 2,\ 2$
L_{cyl}	Filament cylinder equilibrium length	$27\ nm$
δt	Length of chemical evolution step	$0.05\ s$
F_T	Force tolerance of mechanical minimization	$1\ pN$
	Mechanical Parameters	
$K_{\text{fil,str}}$	Actin filament stretching constant	$100\ pN/nm$ [42]
ϵ_{bend}	Actin filament bending energy	$2690\ pN \cdot nm$ [42, 314]
K_{vol}	Cylinder excluded volume constant	$10^5\ pN/nm^4$ [42]
$K_{\text{head,str}}$	NMIIA head stretching constant	$2.5\ pN/nm$ [315]
$K_{\alpha\text{,str}}$	α-actinin stretching constant	$8\ pN/nm$ [316]
$\epsilon_{\text{boundary}}$	Boundary repulsion energy	$41\ pN \cdot nm$ [a]
λ	Boundary repulsion screening length	$2.7\ nm$ [b]
	Mechanochemical Parameters	
$N_{\text{NMIIA,bind}}$	Binding sites per cylinder for myosin motors	8 [c]
$N_{\alpha\text{,bind}}$	Binding sites per cylinder for α-actinin	4 [d]
d_{step}	NMIIA minifilament step size	$6.0\ nm$ [315]
$N_{\text{min}},\ N_{\text{max}}$	Range of number of NMIIA heads per minifilament	15, 25 [e] [317]
F_{stall}	Stall force of NMIIA minifilament	$50\ pN$ [f]
$F_{0,\text{head}}$	Characteristic force of NMIIA catch-bond	$12.6\ pN$ [93]
$F_{0,\alpha}$	Characteristic force of α-actinin slip-bond	$17.2\ pN$ [9]
$F_{0,\text{poly}}$	Characteristic force of actin Brownian ratchet	$1.5\ pN$ [10]
l_M	Equilibrium length of NMIIA minfilament	$175 - 225\ nm$ [42]
l_α	Equilibrium length of α-actinin	$30 - 40\ nm$ [42]
	Diffusion Constants	
$k_{\text{actin,diff}}$	Diffusion constant of actin monomer	$20\ \mu M s^{-1}$ [42]
$k_{\alpha\text{,diff}}$	Diffusion constant of α-actinin	$2\ \mu M s^{-1}$ [42, 226]
$k_{\text{motor,diff}}$	Diffusion constant of NMIIA minifilament	$0.2\ \mu M s^{-1}$ [42]

TABLE D.2: All other parameters used in the simulations reported in this chapter.

[a] - Chosen for the energy scale to be 10 k_BT.
[b] - Chosen as the the length of a G-actin monomer.
[c] - Chosen to allow the spacing between binding sites to be roughly equal to its physiological value near 6 nm [315].

[d] - Chosen to allow the spacing between binding sites to be roughly equal to its physiological value near 30 nm [17].

[e] - Chosen to given an average $N_{\text{total}} = 20$ in approximate agreement with literature values [317].

[f] - A wide range of values are found in the literature for the stall force of the minifilament. We take an order of magnitude estimate for this parameter based on the stall force of a single head (taken here as 5 pN, estimated as $d_{\text{step}} K_{\text{head,str}}$ [42]) times the number of bound heads in the minifilament (taken here as 10). This parameter choice is empirically valid as it yields observable network contraction.

Cylinder equilibrium lengths L_{cyl} are chosen as 27 nm with 4 binding sites per cylinder for myosin filaments and 1 binding site per cylinder for cross-linkers, giving approximately physiological values for stepping distances of myosin motor filaments and spacing along actin filaments of bound α-actinin. We note that the form of the mechanochemical models has been changed from those used in Ref. 42; the modeling used here is current as of MEDYAN v3.2, and we refer readers Appendix A for further details.

D.2 Mean-field model of treadmilling dissipation

To validate the methods for quantifying dissipation using MEDYAN against a simpler representation of actin filament treadmilling, we developed a mean-field description of the dissipation resulting from chemical reactions in the CT model (see Chapter 2). Mean-field models of the trajectory of the vector of concentrations of species, $\mathbf{C}(t)$, have been formulated previously as an 11-dimensional system of ODEs ([183]), and in the CT model as a 5-dimensional system of ODEs ([99]). These models describe the polymerization of a concentration N_{fil} of actin filaments in a pool of actin subunits of total concentration M. The subunit species tracked by these models are distinguished by their polymerization state and by the hydrolysis states of the nucleotide to which they are bound. The meaning of "mean-field" in this context is the assumption that reacting species are well-mixed over the entire system volume, therefore obeying mass-action kinetics and deterministic dynamics which can be represented by ODEs. For an instantaneous value of $\mathbf{C}$, we define a function $D_\Lambda(\mathbf{C})$ representing the instantaneous rate of dissipation due to a set of reactions Λ. Thus a solution of a mean-field $\mathbf{C}(t) = \mathbf{N}(t)/\Theta$, allows us to construct the trajectory of the dissipation rate, $D_\Lambda(\mathbf{C}(t))$. This function cannot capture the dissipation due to the activity of myosin filaments or cross-linkers, or from relaxation of mechanical stress, because these aspects are not included in the these mean-field models describing filament treadmilling. The benefit of such a mean-field model is that one can perform systematic variation of parameters with limited computational demands, and we use it here to study the effect of the parameters N_{fil} and M on the dissipation due to filament treadmilling.

The function $D_\Lambda(\mathbf{C}(t))$ can be written as a sum over the reactions $\lambda \in \Lambda$ of the instantaneous rate of change of the solution's Gibbs free energy due to that reaction:

$$D_\Lambda(\mathbf{C}(t)) = \sum_{\lambda \in \Lambda} \Delta G_\lambda(\mathbf{C}(t)) r_\lambda(\mathbf{C}(t)). \tag{D.16}$$

The expressions for $\Delta G_\lambda(\mathbf{C}(t))$ for different reactions are described above. The instantaneous rate of reaction λ, $r_\lambda(\mathbf{C}(t))$, is written as usual for mass-action kinetics as

$$r_\lambda(\mathbf{C}(t)) = \Theta k_\lambda \prod_{i \in R} C_i^{\nu_i} \tag{D.17}$$

where R is the set of reactant species for reaction λ, and where we have included the conversion factor Θ to convert $r_\lambda(\mathbf{C}(t))$ to units of s^{-1}.

To facilitate the study of dissipation due to filament treadmilling, we define a set of reactions Λ in the CT model which constitutes the dominant cycle a subunit undergoes in the treadmilling process: a) polymerization of G^T to the plus end, b) hydrolysis of ATP by F^T, c) release of Pi by F^{Pi}, d) depolymerization of G^D from the minus end, and e) nucleotide exchange converting G^D to G^T. We refer to this set of 5 reactions as the main treadmilling pathway (MTP). Alternative sequences of reactions whose net effect is similarly the conversion of one molecule of ATP to ADP and Pi are considered as of secondary importance and not included in this analysis.

This mean-field description of entropy production can be compared to the description of entropy production rates for chemically reactive systems that emerged from the

Brussels school of thermodynamics [318, 319]. The results of that school include the minimum entropy-production principle applicable in the linear regime [198], and the general evolution criterion applicable even in the the nonlinear regime [320]. Their formalism typically considers the total entropy production rate as a volume integral over the local entropy production rate density, which is itself written as a sum over fluxes multiplied by their corresponding thermodynamics forces defined at each point of the system. This sum is decomposed into terms representing diffusion and terms representing chemical reactions, and the terms representing the chemical reactions are written such that the fluxes reflect the net reaction rate at that point in the system. In contrast, our mean-field description neglects concentration gradients and resulting diffusion fluxes, as we assume a homogeneous distribution of the chemical species. Equation D.16 then represents only the chemical contribution to the entropy production, as a sum over the rates and affinities of the reactions in the system, implicitly integrating over the homogeneous system volume. We treat the forward and reverse direction for some chemical reaction as separate terms in Equation D.16, so these rates cannot be considered fluxes which would include the reverse rate as well. This allows for more general sets of reactions which might include effectively irreversible processes for which the reverse rate is negligible. However the set of reactions Λ could be chosen to include a reverse reaction for each forward reaction, with the result that these pairs of terms represent fluxes along the reversible reaction pathway. The parsimonious reaction set MTP is chosen not to fully describe the rate of entropy production in the system, but to allow easy visualization of the main contributions to the entropy production. We do not pursue the connection of our treatment to the formalism of the Brussels school further here, but we lastly note that our results are compatible with their minimum entropy production principle, as shown in Figure D.3.

We first verified that the mean-field model of MTP dissipation agreed with results from MEDYAN simulations, to illustrate consistency between these approaches. In Figure D.3 we display the close match between the trajectory of MTP dissipation over a 2000 s run from these two approaches. Note that, to allow direct comparison, only the changes of Gibbs free energy resulting from reactions in the MTP set are visualized for both approaches here, i.e. the contribution from diffusion and other non-MTP reactions in the MEDYAN simulation are not visualized. We also turned off in MEDYAN the force-sensitive decrease in polymerization rate when the filament tips push against the simulation hard-wall boundaries, since this effect is not captured in the mean-field modeling [101]. We used a simulation volume of 1 μm^3 and initial conditions of equal amounts (10 μM each) of G^T and G^D actin in a 0.08 μM pool of seed filaments containing F^T. The dissipation rate decreases nearly monotonically, attaining a minimal steady state value after tens of seconds. In Figure D.3 we also display the individual contributions to the sum in Equation D.16. Initial polymerization of G^T to the plus end constitutes the majority of the initial dissipation. As this polymerization process slows after about 1 second, the hydrolysis of ATP by the now relatively abundant F^T becomes the dominant contribution. As hydrolysis then slows after about 10 seconds, the total dissipation rate reaches a steady state value of roughly 80 $k_B T/s$. In Figure D.3 we also plot the mean-field prediction of the trajectory of the reacting species' concentrations.

We next simultaneously varied the total concentration of actin M and the concentration of actin filaments N_{fil}, and determined the total dissipation integrated along each trajectory as well as the steady state dissipation rate. As M was varied, we held the initial concentration of each species proportionally the same: 49 % G^T, 49 % G^D, and 2 % F^T. As shown in Figure D.4, the integrated dissipation over 2000 s was observed to increase monotonically with both M and N_{fil}. Quantitatively, the integrated dissipation depends on the choice of initial proportions, however we found that the shape of the dependence on N_{fil} and M is largely independent of initial proportions (data not shown). Total dissipation increases with M simply because more actin is available to hydrolyze ATP during the approach to steady state. For large amounts of actin, increasing N_{fil}, the number concentration of filaments, allows increased polymerization of G^T, which constitutes a large contribution to total dissipation during the early stages of the trajectory.

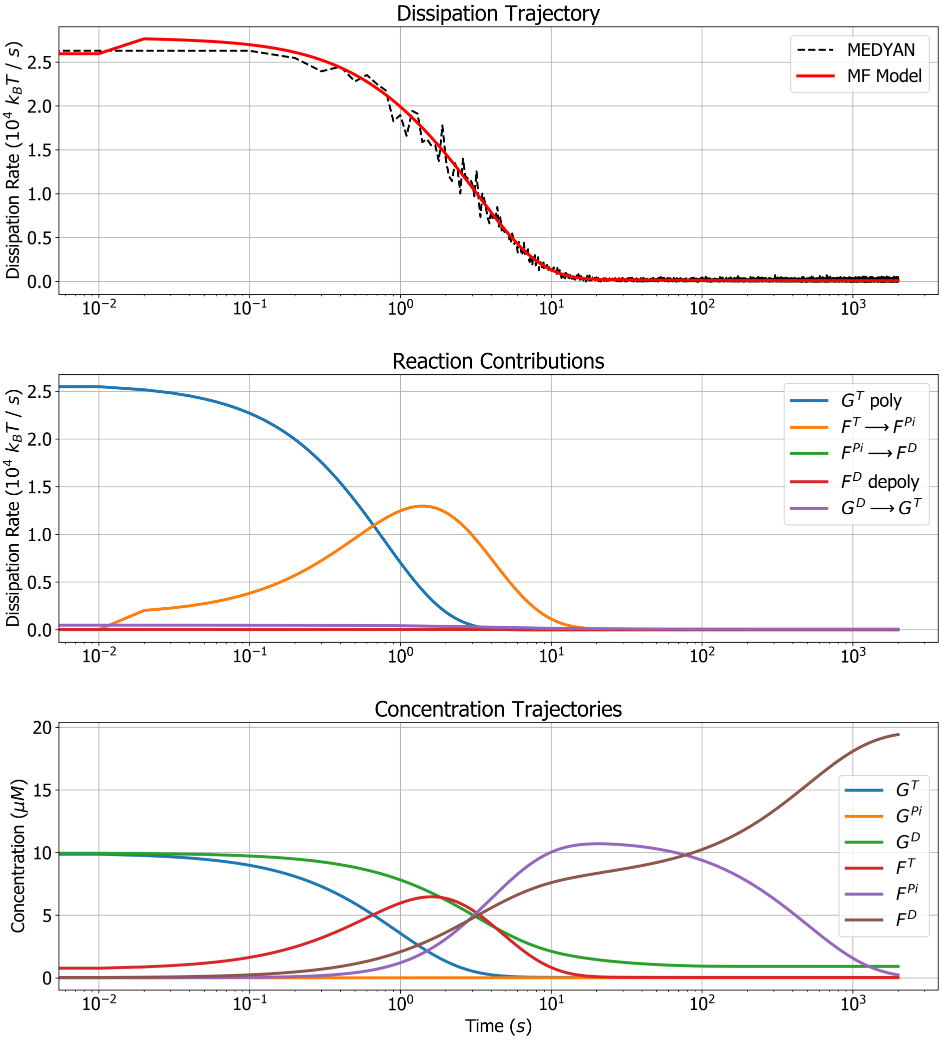

FIGURE D.3: Results of the mean-field modeling of MTP dissipation. *Top*: Comparison of $D_{\mathrm{MTP}}(\mathbf{C}(t))$ calculated using the mean-field model (with time increments of 0.01 s), with ΔG_{chem} rates measured during MEDYAN simulation (with time increments of 0.1 s). *Middle*: Plot of the contributions of each reaction in the MTP to the total dissipation. The items in the legend represent reactions in the MTP, which are described in the main text. *Bottom*: Plot of the trajectories of the reacting species concentrations. The notation for each species is described in the main text.

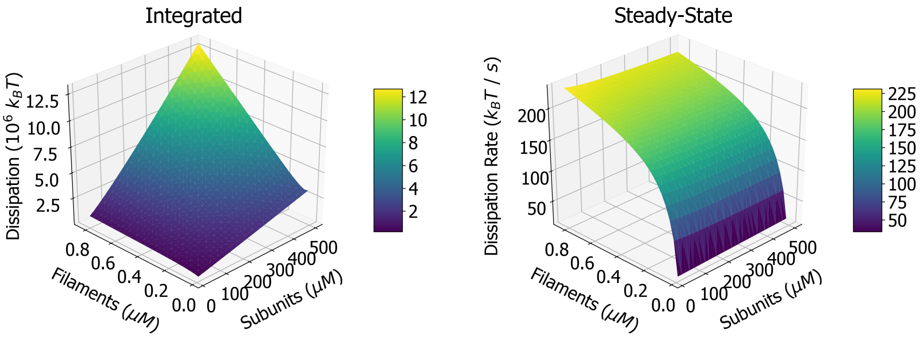

FIGURE D.4: *Left*: The total dissipation integrated over 2000 s trajectories as M and N_{fil} are varied. *Right*: The steady state dissipation rate over the same range of M and N_{fil}.

Increasing N_{fil} also shifts the steady state concentration of G^T downward (Figure D.5), implying that more G^T has been polymerized during the approach to steady state. A loose analogy can be made of a one lane road compared to a multi-lane highway during heavy traffic to describe this situation. As N_{fil} is increased with M fixed, the steady state dissipation rate increases concavely, as shown in Figure D.4. The steady state concentration of F^T also increases concavely (Figure D.5), representing higher rates of ATP hydrolysis. The contributions of each reaction to the total dissipation rate at steady state as N_{fil} is varied is illustrated in Figure D.6. The lack of dependence of the steady state dissipation rate on M can be explained by the fact that increasing M increases the steady state concentration of only F^D, not of any other species [99]. In other words, all extra actin accumulates in the form F^D as M is increased. This species is essentially inert, since the depolymerization rate of F^D is controlled by the concentration of filaments N_{fil}. Thus the steady state dissipation has no dependence on the total amount of actin. Furthermore, it has no dependence on the initial concentrations, since it is known that the steady state vector of concentrations does not depend on initial conditions [99].

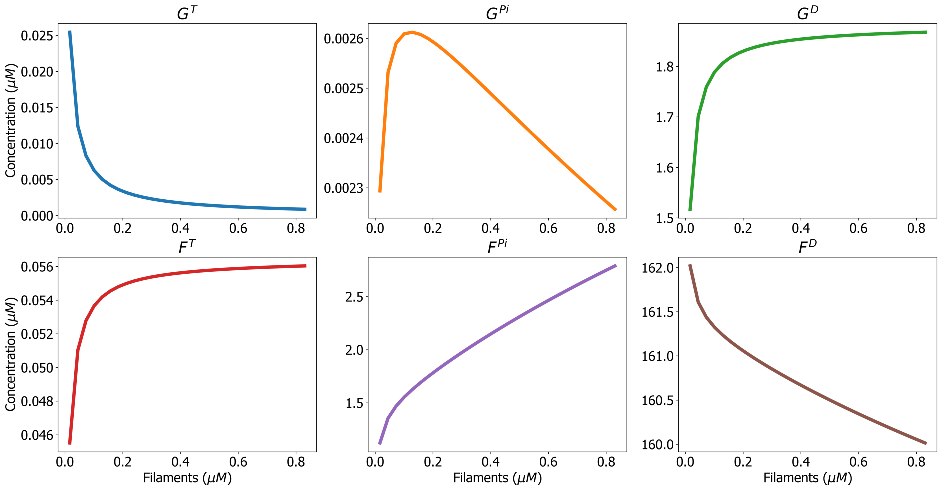

FIGURE D.5: The concentrations at steady state of the various actin subunit species as the concentration of filaments N_{fil} is varied. These curves have no dependence on initial conditions, except F^D which will increase linearly with the total concentration of actin subunits M; any additional actin subunits in the system will accumulate in the form F^D at steady state.

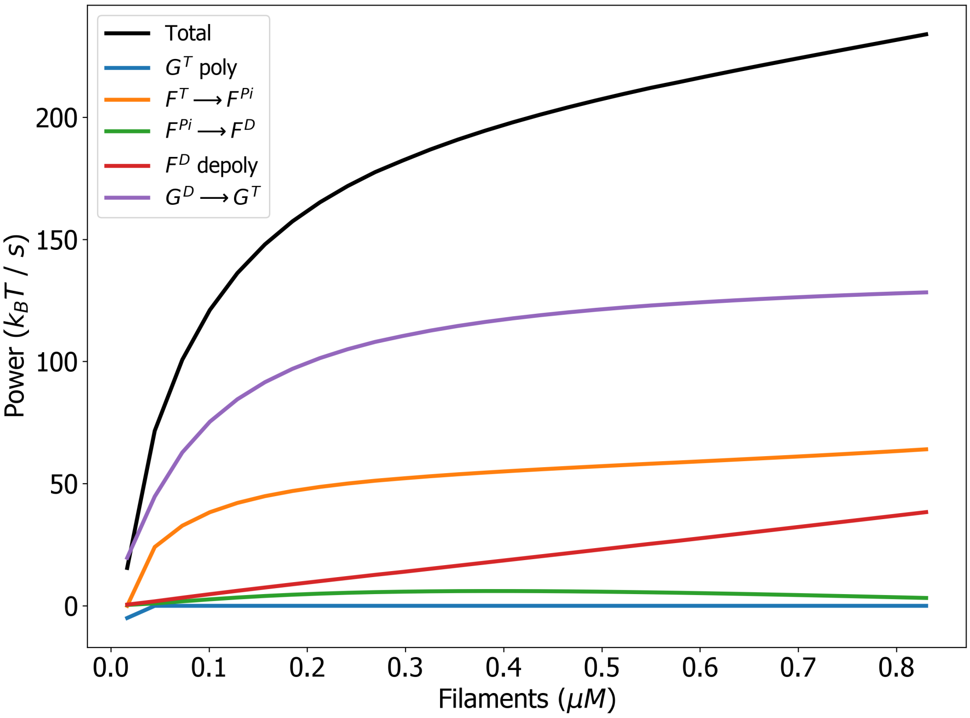

Figure D.6: Contributions of each reaction in the main treadmilling pathway to the total dissipation rate at steady-state, as the concentration of filaments N_{fil} is varied. These curves have no dependence on the initial concentrations of the different subunit species.

Appendix E

Supporting information for Chapter 6

E.1 Weibull plots

To assess whether stretched exponential functions describe the empirical distributions of $|\Delta U_-|$ and ΔU_+, we first define $Q(u) = \ln\left(-\ln\left(P(|\Delta U_-| \geq u)\right)\right)$. The degree to which the plots of $Q(u)$ against $\ln(u)$ appear to be linear serves as a check of the appropriateness of modeling $P(|\Delta U_-|)$ as a stretched exponential, or Weibull, distribution [273]. On the basis of these Figures E.1 and E.2 we conclude that the Weibull distribution is a satisfactory choice for all values of V. In the Chapter 6, the Weibull parameters k and λ were determined by fitting the stretched exponent $e^{-(x/\lambda)^k}$ to the observed CCDF $P(|\Delta U_-|)$ on a log-scale, that is, by fitting $-(x/\lambda)^k$ to $\ln\left(P(|\Delta U_-|)\right)$ using standard nonlinear fitting routines. Treating these functions on a log-scale ensured a better fit to the distribution tails which are of most interest in the present case.

E.2 Computing filament displacements

The area between the two filaments $\mathbf{x}$ and $\mathbf{y}$ is triangulated using the beads comprising the filaments ($\{\mathbf{x}_i\}_{i=0}^{n_{\mathbf{x}}-1}$ and $\{\mathbf{y}_j\}_{j=0}^{n_{\mathbf{y}}-1}$) as vertices, where $n_{\mathbf{x}}$ is the number of beads in $\mathbf{x}$ and similarly for $n_{\mathbf{y}}$. To compute the displacement of filament $\mathbf{x}$ during the time interval δt, we set $\mathbf{y}$ to the new configuration of $\mathbf{x}$ at the end of the interval. The triangles come

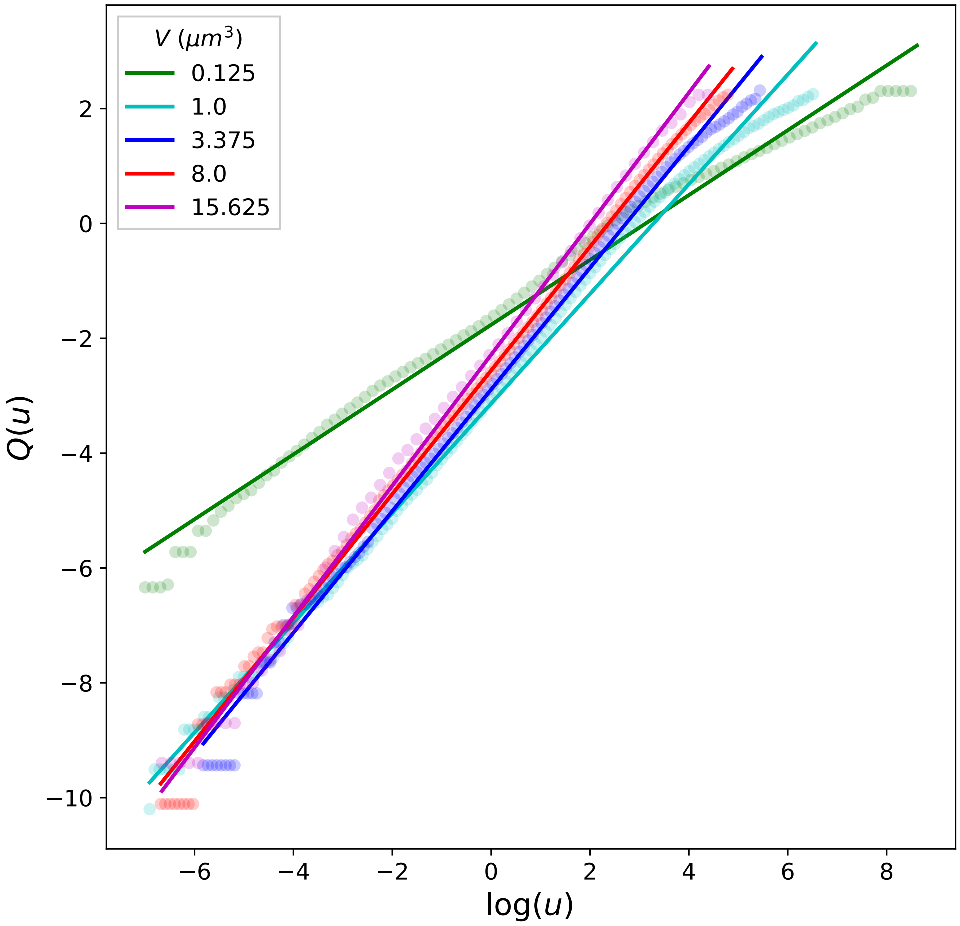

Figure E.1: Plots of the function $Q(u) = \ln(-\ln(P(|\Delta U_-| \geq u)))$ for different volumes V along with a fitted line, where $P(|\Delta U_-|)$ is the observed CCDF obtained from five runs of each volume.

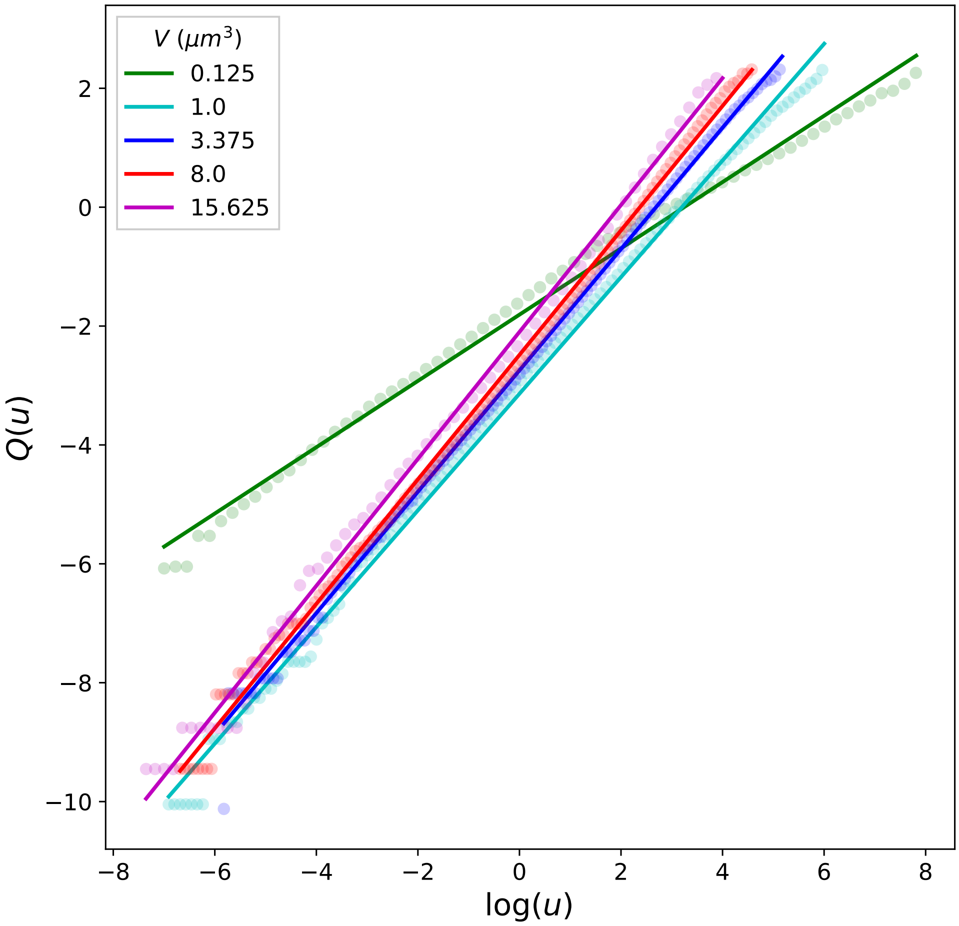

FIGURE E.2: Plots of the function $Q(u) = \ln(-\ln(P(\Delta U_+ \geq u)))$ for different volumes V along with a fitted line, where $P(\Delta U_+)$ is the observed CCDF obtained from five runs of each volume.

in pairs for most of the filament lengths, as shown using the dark and light colors of green of Figure E.3. If $n_{\mathbf{x}}$ and $n_{\mathbf{y}}$ are unequal (say $n_{\mathbf{x}} < n_{\mathbf{y}}$), extra triangles are added using the last bead in $\mathbf{x}$, $\mathbf{x}_{n_{\mathbf{x}}-1}$, as the only vertex in filament $\mathbf{x}$. The sum of these triangle areas A_{tot} is divided by the average of the two filament contour lengths $L_{\mathbf{x}}$ and $L_{\mathbf{y}}$ to give the measure of distance $d = \frac{2A_{\text{tot}}}{L_{\mathbf{x}}+L_{\mathbf{y}}}$.

E.3 Parameterization

The following table lists the parameters chosen for the simulations presented in this Chapters 5 and 6. However, in 5, the motor stall force is varied and does not correspond to values given below, and the cylinder discretization lengths are chosen differently. As described in that chapter, the cylinder lengths are either 27 nm (with $\epsilon_{\text{bend}} = 2690\ pN \cdot nm$) or 108 nm (with $\epsilon_{\text{bend}} = 672\ pN \cdot nm$). Additionally, in finite size scaling studies the number of components $N_x,\ N_y,\ N_z$ are varied.

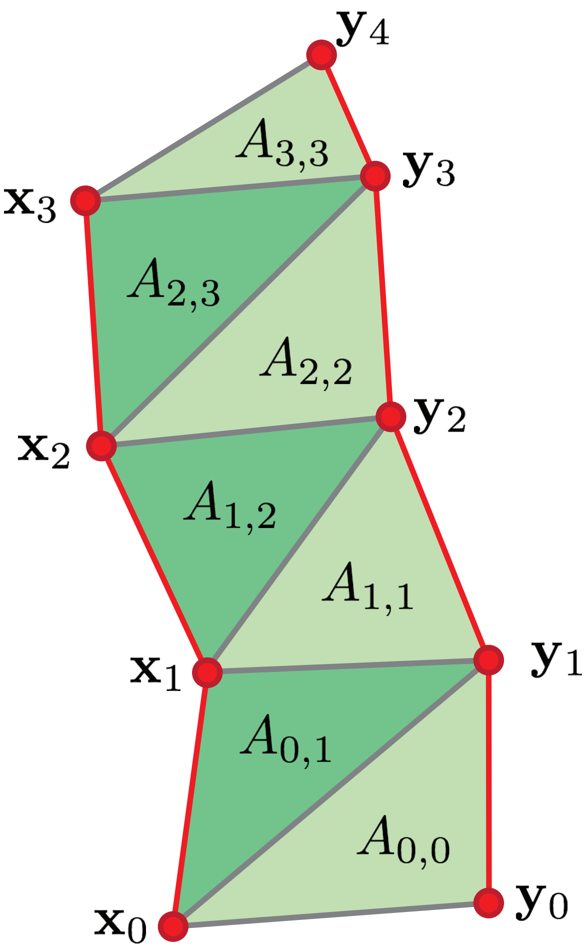

Figure E.3: Illustration of how the area between two filaments $\mathbf{x}$ and $\mathbf{y}$ is triangulated to allow calculation of the distance between them. The beads comprising the filaments are labeled $\mathbf{x}_i$, $\mathbf{y}_j$, and areas between triplets of beads are labeled $A_{i,j}$ where the lowest indices of the beads $\mathbf{x}_i$ and $\mathbf{y}_j$ in the triplet are used.

Parameter	Description	Value
	General Simulation Parameters	
k_BT	Thermal energy	$4.1\ pN \cdot nm$
L_{comp}	Cubic compartment side length	$500\ nm$
$N_x,\ N_y,\ N_z$	Number of compartments in each dimension	$2,\ 2,\ 2$
L_{cyl}	Filament cylinder equilibrium length	$54\ nm$
δt	Length of chemical evolution step	$0.05\ s$
F_T	Force tolerance of mechanical minimization	$1\ pN$
	Mechanical Parameters	
$K_{\text{fil,str}}$	Actin filament stretching constant	$100\ pN/nm$ [42]
ϵ_{bend}	Actin filament bending energy	$1344\ pN \cdot nm$ [42, 314]
K_{vol}	Cylinder excluded volume constant	$10^5\ pN/nm^4$ [42]
$K_{\text{head,str}}$	NMIIA head stretching constant	$2.5\ pN/nm$ [315]
$K_{\alpha,\text{str}}$	α-actinin stretching constant	$8\ pN/nm$ [316]
$\epsilon_{\text{boundary}}$	Boundary repulsion energy	$41\ pN \cdot nm$ [a]
λ	Boundary repulsion screening length	$2.7\ nm$ [b]
	Mechanochemical Parameters	
$N_{\text{NMIIA,bind}}$	Binding sites per cylinder for myosin motors	8 [c]
$N_{\alpha,\text{bind}}$	Binding sites per cylinder for α-actinin	4 [d]
d_{step}	NMIIA minifilament step size	$6.0\ nm$ [315]
$N_{\text{min}},\ N_{\text{max}}$	Range of number of NMIIA heads per minifilament	15, 25 [e] [317]
F_{stall}	Stall force of NMIIA minifilament	$100\ pN$ [f]
$F_{0,\text{head}}$	Characteristic force of NMIIA catch-bond	$12.6\ pN$ [93]
$F_{0,\alpha}$	Characteristic force of α-actinin slip-bond	$17.2\ pN$ [9]
$F_{0,\text{poly}}$	Characteristic force of actin Brownian ratchet	$1.5\ pN$ [10]
l_M	Equilibrium length of NMIIA minfilament	$175 - 225\ nm$ [42]
l_α	Equilibrium length of α-actinin	$30 - 40\ nm$ [42]
	Chemical Parameters	
$k_{\text{actin,diff}}$	Diffusion constant of actin monomer	$20\ \mu M s^{-1}$ [42]
$k_{\alpha,\text{diff}}$	Diffusion constant of α-actinin	$2\ \mu M s^{-1}$ [42, 226]
$k_{\text{motor,diff}}$	Diffusion constant of NMIIA minifilament	$0.2\ \mu M s^{-1}$ [42]
$k_{\text{actin,poly,+}}$	Actin plus-end polymerization	$11.6\ \mu M s^{-1}$ [12]
$k_{\text{actin,poly,-}}$	Actin minus-end polymerization	$1.3\ \mu M s^{-1}$ [12]
$k_{\text{actin,depoly,+}}$	Actin plus-end depolymerization	$1.4\ s^{-1}$ [12]
$k_{\text{actin,depoly,-}}$	Actin minus-end depolymerization	$0.8\ s^{-1}$ [12]
$k_{\text{head,bind}}$	NMIIA head binding	$0.2\ s^{-1}$ [88]
$k^0_{\text{head,unbind}}$	NMIIA head unbinding under zero tension	$1.7\ s^{-1}$ [42, 88]
$k_{\alpha,\text{bind}}$	α-actinin binding	$0.7\ \mu M s^{-1}$ [72]
$k^0_{\alpha,\text{unbind}}$	α-actinin unbinding under zero tension	$0.3\ s^{-1}$ [72]

TABLE E.1: All parameters used in the simulations reported in Chapter 6 and 7.

[a] - Chosen for the energy scale to be 10 k_BT.

[b] - Chosen as the the length of a G-actin monomer.

[c] - Chosen to allow the spacing between binding sites to be roughly equal to its physiological value near 6 nm [315].

[d] - Chosen to allow the spacing between binding sites to be roughly equal to its physiological value near 30 nm [17].

[e] - Chosen to given an average $N_{\text{total}} = 20$ in approximate agreement with literature values [317].

[f] - A wide range of values are found in the literature for the stall force of the minifilament. We take an order of magnitude estimate for this parameter based on the stall force of a single head (taken here as 10 pN, estimated as $d_{\text{step}} K_{\text{head,str}}$ [42]) times the number of bound heads in the minifilament (taken here as 10). This parameter choice is empirically valid as it yields observable network contraction.

E.4 Shannon entropy of tension distribution

The simulation volume of 1 μm^3 is discretized into 10^3 cubic voxels, each 0.1 μm in linear dimension. Let $i, j, k = 1, ..., 10$ index these voxels, which are an analysis tool and not related to the reaction-diffusion compartments used in MEDYAN. After each simulation cycle, the mechanical components of the cytoskeletal network (i.e. the filament cylinders, the myosin motors, and the passive cross-linkers) are each under some compressive or tensile force T_n, where n indexes the mechanical component. There are other mechanical potentials involving these components, but we focus here only on the tensions T_n. Each mechanical component has a center of mass $\mathbf{r}_n$, and we introduce the indicator function $\chi_{ijk}(\mathbf{r}_n)$ which is equal to 1 if $\mathbf{r}_n$ is inside voxel i, j, k and 0 otherwise. The total tension magnitude inside voxel i, j, k is

$$|T|_{ijk} = \sum_n |T_n| \chi_{ijk}(\mathbf{r}_n). \tag{E.1}$$

The discrete non-negative scalar field $|T|_{ijk}$ is converted to a probability distribution P_{ijk} by normalization:

$$P_{ijk} = \frac{|T|_{ijk}}{\sum_{ijk} |T|_{ijk}}. \tag{E.2}$$

Finally, we introduce the discrete Shannon entropy of this distribution at time t as

$$H(t) = -\sum_{ijk} P_{ijk}(t) \ln P_{ijk}(t). \tag{E.3}$$

The units of H are nats, and large values indicate a homogeneous spatial distribution of tension magnitudes throughout the network. Reported trends using this metric are essentially independent of the discretization length.

E.5 Machine learning implementation details

To forecast the occurrence of cytoquakes, we resorted to using a high-dimensional machine learning (ML) model (3 layer feed-forward neural network) after it was found that several simple features in the eigenspectrum which we believed might reflect mechanical stability (for instance the value of the smallest positive eigenvalue) did not by themselves significantly correlate with cytoquake occurrence. We pose the forecasting of cytoquakes as a binary classification problem. A trajectory $\Delta U(t) = U(t+\delta t) - U(t)$ at QSS (after 1,000 s) is converted to a binary sequence such that each t for which $\Delta U(t) \leq \Delta U_T$, as well as the $t_W = 0.15$ previous seconds (i.e. 3 previous time points) are classified as cytoquakes, and the rest are not. This t_W window is chosen to help overcome the stochasticity inherent in the chemical dynamics which, along with the instantaneous mechanical stability we are using as a predictor, controls cytoquake occurrence. We focus here on the five runs of conditions $C_{3,3}$. $\Delta U_T = -100\ k_B T$ is chosen to lie well in the tail of the distribution of $|\Delta U_-|$ for this condition and therefore distinguishes rare events, as shown in Figure 6.2. With these choices, $\sim 10\%$ of samples across all runs are labeled as events in the classification problem.

The predictors of the model capture information about the network's mechanical stability. The ordered sets of eigenvalues $\{\lambda_k\}_{k=1}^{3N}$ at each time t is padded by adding zero eigenvalues between the unstable ($\lambda_k < 0$) and stable ($\lambda_k \geq 0$) parts of the spectrum to maintain a fixed input dimension across all time points and runs. We then collect these eigenvalues into a tuple $\mathbf{M}(t)$ such that the first element of $\mathbf{M}(t)$ is the largest negative

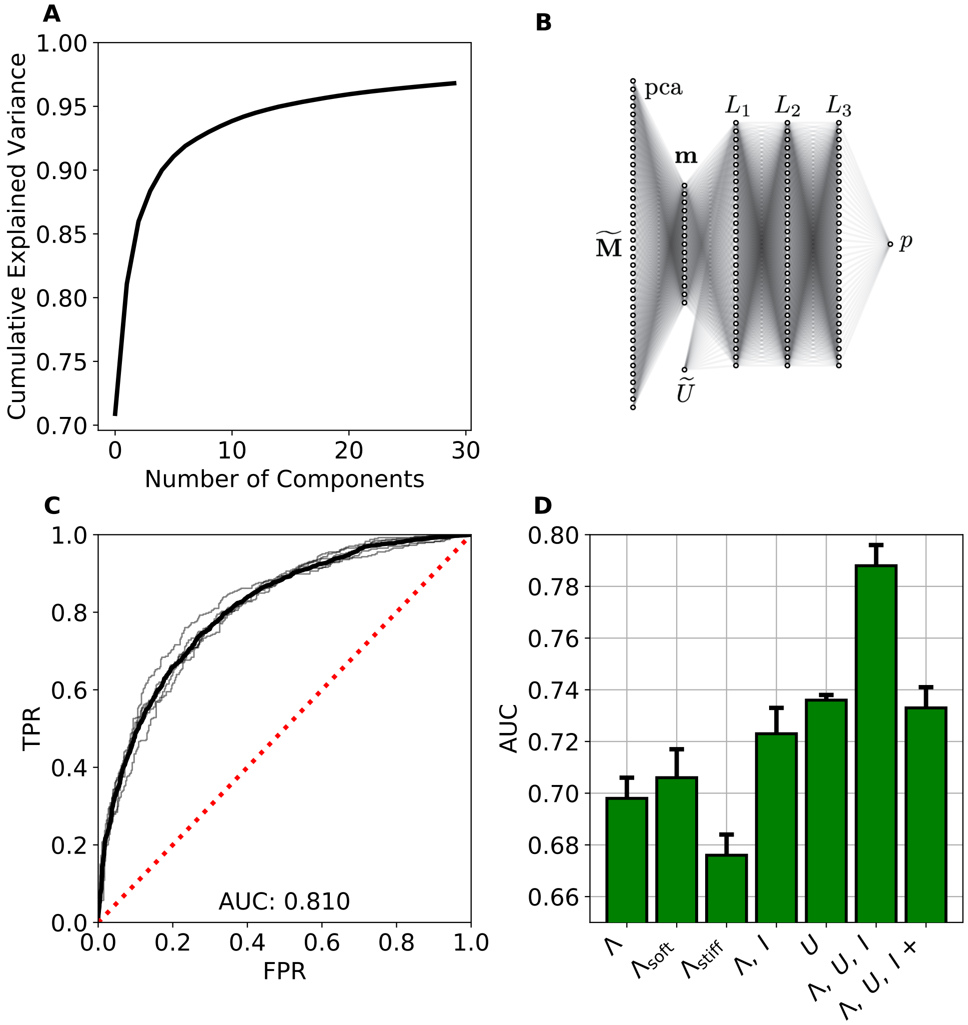

Figure E.4: **A**: Cumulative explained variance from PCA of the $\sim 1,600$ eigenvalues $\{\lambda_k(t)\}_{k=1}^{3N}$. **B**: Schematic depiction of the feed-forward neural network architecture. **C**: ROC curves for the model using only $\{\lambda_k\}_{k=1}^{3N}$ as input and trained on a single run of condition $C_{3,3}$, with five realizations of the stochastic batch network training and their average shown. The ROC curve of a random model is plotted as the red dotted line. **D**: Bar plot indicating the AUC of ROC curves using different combinations of inputs for the model trained on data collected from all runs of condition $C_{3,3}$. From left to right, the labels indicate that the model inputs are: $\{\lambda_k\}_{k=1}^{3N}$; $\{\lambda_k | 0 \leq \lambda_k < \lambda_T\}$; $\{\lambda_k | \lambda_T \leq \lambda_k\}$; $\{\lambda_k\}_{k=1}^{3N}$ and $\{r_k\}_{k=1}^{3N}$; U, using a logistic regression model; $\{\lambda_k\}_{k=1}^{3N}$, $\{r_k\}_{k=1}^{3N}$, and U; $\{\lambda_k\}_{k=1}^{3N}$, $\{r_k\}_{k=1}^{3N}$, and U with forecasting done for large positive increments $\Delta U > 100\ k_B T$. Error bars indicate uncertainty from five realizations of stochastic batch training.

λ_k at time t and the last element is the largest positive λ_k at time t. We optionally include the the inverse participation ratios $\{r_k\}_{k=1}^{3N}$ in this vector by first adding zeros in the places of the set $\{r_k\}_{k=1}^{3N}$ corresponding to where zeros were added in the set $\{\lambda_k\}_{k=1}^{3N}$, and then interleaving the λ_k and r_k in the now doubly sized tuple $\mathbf{M}(t)$, so that now for example the first two elements of $\mathbf{M}(t)$ correspond to the largest negative λ_k and the associated r_k at time t. The tuples $\mathbf{M}(t)$ are then linearly rescaled, so for each element $M_i(t)$ the average over all times of a run is 0 and the variance is 1. These rescaled tuples are labeled $\widetilde{\mathbf{M}}(t)$.

When only the λ_k are included then $\widetilde{\mathbf{M}}(t)$ has $\sim$ 1,600 dimensions, and with the r_k are also included it has $\sim$ 3,200 dimensions. To avoid overfitting the model, we first reduce the dimensionality of $\widetilde{\mathbf{M}}(t)$ via principal component analysis (PCA) using all QSS time points in a run. We choose 30 dimensions as the size of the reduced tuple $\mathbf{m}(t)$ because this allows for more than 95% of the variance of $\widetilde{\mathbf{M}}(t)$ to be explained when just the λ_k are included as shown in Figure E.4.A. Model performance appreciably decreases when fewer than 30 dimension are used and improves only marginally if more are used. A row of ones is added as a 31^{st} dimension to $\mathbf{m}(t)$ as a bias for the neural network. As an additional indicator of the network's mechanical stability we also consider its mechanical energy U at time t. $U(t)$ is linearly rescaled to give $\widetilde{U}(t)$ so that it has zero mean and unit variance. We then optionally augment with input tuple $\mathbf{m}(t)$ with the $\widetilde{U}(t)$ as a 32^{nd} dimension.

We can treat the data from all five runs of condition $C_{3,3}$ separately or combine all data together to train a larger model. Model performance is generally found to be better when trained on data from a single run, however by combining data from all runs we probe more general underlying trends that are not specific to the network organization of one run. When describing trends from varying model inputs, as in Figure E.4.D, we focus on results obtained by combining all runs due to their greater generality.

For a single run there are $\sim$ 20,000 samples, giving 100,000 samples when combining all runs. When combining runs, we first rescale and perform PCA on the predictors using only the data within a single run, and then concatenate the resulting $\mathbf{m}(t)$ with their associated labels into a larger data set. This way the relative variation of the predictors compared to their typical values for a particular organization of the actomyosin network is retained, and the typical values of particular network organizations themselves affect the model inputs to a lesser degree.

We used the Python modules scikit-learn and Keras with a Tensorflow back end to train a deep feed-forward neural network and a logistic regression model for the binary classification problem [321, 322]. The 31 or 32-dimensional (depending on if $\widetilde{U}(t)$ is included as a predictor) input tuple $\mathbf{m}$ is fed into three fully connected hidden layers L_i, $i = 1, 2, 3$, each with either 30 or 100 nodes depending on if the data consists of a single run (20,000 samples) or of all five runs (100,000 samples). Each node in the hidden layers uses a rectified linear unit activation function. The output of the network is two nodes using a softmax activation function whose values are p and $1 - p$, where p is the predicted probability of a cytoquake event at that time t. This architecture is schematically illustrated in Figure E.4.B. The network is trained for either 400 or 200 epochs using a categorical cross-entropy loss function with Adam optimization in stochastically chosen batches of either 1,000 or 10,000 samples, depending on the whether the single or multiple run data sets, respectively, are used. The cytoquake samples are given a higher weight ($\times 3$) than the non-cytoquake samples during training. A L2 penalty of 0.05 is used to curb overfitting. When using only $\widetilde{U}(t)$ as a predictor, a logistic regression model is fit using the same sample weights.

Of all the data samples, we use 2/3 to train the model with and validate the model on the remaining 1/3. We repeat these random training/testing set splits to gather statistics on model performance. The binary classification procedure involves the probability threshold p_T (such that $p > p_T$ means the model predicts a cytoquake). Model performance is measured by varying p_T from 0 to 1 and measuring the true positive rate (TPR, the proportion of actual cytoquakes correctly predicted as such) and false positive rate (FPR, the proportion of actual non-cytoquakes incorrectly predicted as cytoquakes) on the test data; the locus of these points forms the receiver operator characteristic (ROC) curve. A random model would have FPR = TPR, so an area under the curve (AUC) of the ROC curve greater than 0.5 indicates a good model, and a perfect model would have an AUC of 1. One can also consider precision-recall (PR) curves, which contain points in the space of model precision (the proportion of predicted cytoquakes which were actual cytoquakes) and recall (the same as TPR). A random model would have the same precision, equal to the proportion of actual cytoquakes in the testing data, for all values of recall as p_T is varied, giving an AUC equal to that proportion.

When the test data is unbalanced, i.e. when there are many more non-cytoquake events than cytoquake events, it has been shown that the AUC of the PR curve is a more faithful metric for model performance (since a model may score a high AUC of the ROC curve by overestimating that events are not cytoquakes) [323, 324]. To overcome this limitation of ROC curves, which we believe has a more intuitive interpretation that PR curves, we balance the testing data, keeping all cytoquake events and randomly keeping an equal number of non-cytoquake events. We confirmed that trends observed in the AUC of the ROC curves as the model is varied also hold when considering the AUC of PR curves on the full test set.

In Figure E.5 we show examples of these PR and ROC curves on the training and testing data for a model trained on a single run. The very high AUC of the PR and ROC curves evaluated on the training data indicates that the model has nearly perfected its prediction on those samples and may indicate overfitting, however this high performance generalizes nicely to the unseen testing data. Note that the AUC of the ROC evaluated on the testing data is significantly higher than shown in Figure E.4.D reflecting the generally higher performance of models trained on data from a single run compared to models trained on data from all runs.

Finally, as a sanity check, we confirmed that randomly shuffling the labels on the training set decreases performance on the training set and causes the performance on the test set to decrease to that of a random model, as shown in Figure E.6.

Applying the model using the Hessian eigenspectrum as the input, we obtained an AUC of 0.81 when using data from a single run of condition $C_{3,3}$ (Figure E.4.C) and of 0.70 when using data from five runs, i.e. from five different network realizations. In

Figure E.4.D, we display the effects of varying the machine learning model inputs on prediction performance, reflecting the degree to which cytoquake occurrence depends on the various inputs. We point out that these trends from varying the model inputs are not particularly strong, contributing only marginal changes (though greater the measured uncertainty) to the model performance. These differences are less than the difference resulting from combing all five runs in a data set rather than using one run. We report them here mainly out of completeness, rather than in support of some strong conclusion.

Uncertainty in AUC from five repetitions of stochastic batch training is roughly 0.01 for all reported values. Keeping only the eigenvalues of the soft modes does not harm performance (AUC 0.71), while keeping only the stiff modes does harm performance (AUC 0.68). Performance is not harmed (AUC 0.72) upon augmenting the input with the inverse participation ratios $\{r_k(t)\}_{k=1}^{3N}$. Interestingly, we found that a logistic regression model using only the mechanical energy $U(t)$ as an input feature performs well (AUC 0.74, with a smaller uncertainty around 0.002 for this simpler model), reminiscent of the debate concerning one neuron vs. deep learning models of earthquake aftershock prediction [286, 287]. This logistic regression model has learned an optimal cutoff for U that indicates instability and likely cytoquake occurrence. We may seemingly conclude that the machine learning model using the Hessian eigenspectrum as an input has merely learned what the mechanical energy is, however we find that by far the best performance results from combining $\{\lambda_k(t)\}_{k=1}^{3N}$, $\{r_k(t)\}_{k=1}^{3N}$, and $U(t)$ in the ML model, reaching an AUC of 0.79 when using data from all five runs. This suggests that the learned features of the Hessian eigenspectrum are not redundant given U, i.e. that their mutual information is low. Finally, we found that prediction of large positive increments ($\Delta U > 100\ k_BT$) is also possible, with an AUC of 0.74 when combining all inputs.

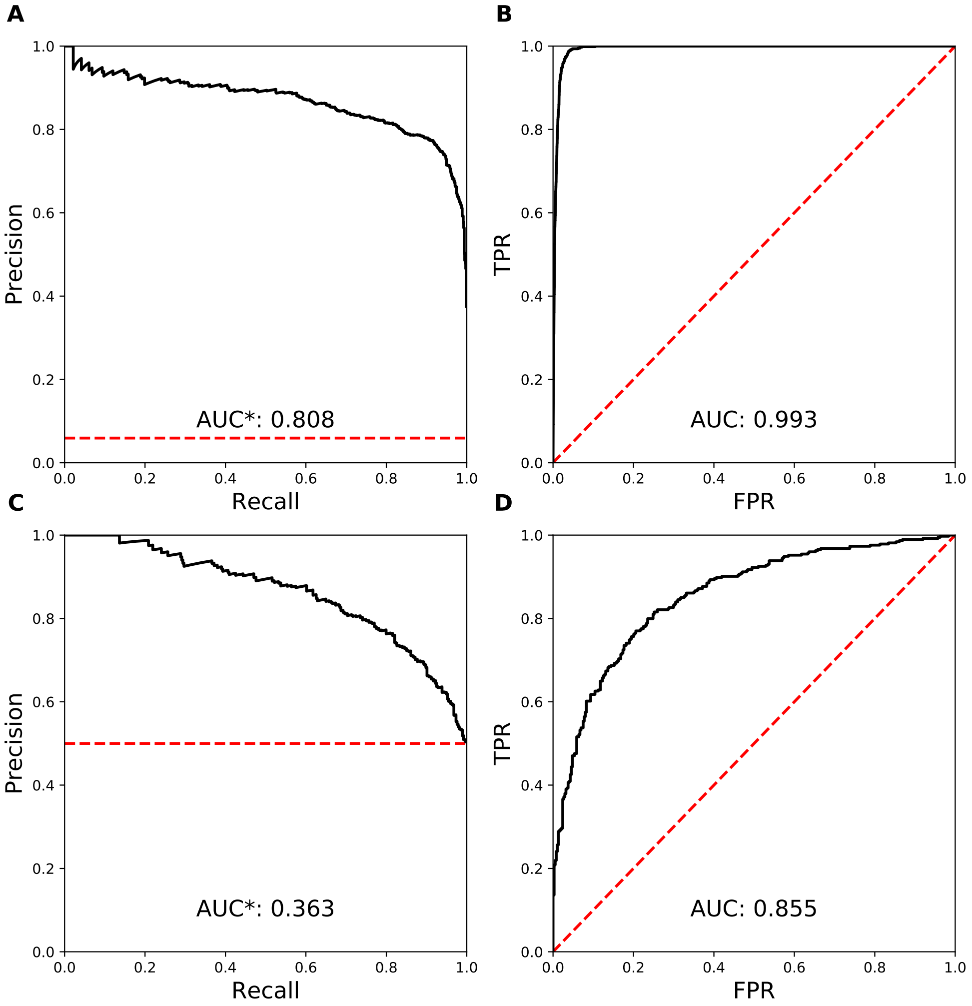

Figure E.5: **A**: PR curve evaluated for a model using $\{\lambda_k\}_{k=1}^{3N}$, $\{r_k\}_{k=1}^{3N}$, and U as inputs trained on data from a single run at QSS of condition $C_{3,3}$ and evaluated on the training data. The red line indicates the performance of a random model on the data set. The asterisk on the AUC indicates that the fraction of cytoquake samples in the data set (for this run ~ 0.06) has been subtracted from the actual AUC, to give the area between the black and red curves. **B**: ROC curve for the same model evaluated on the training data. **C**: PR curve for the same model evaluated on the balanced testing data. **D**: ROC curve for the same model evaluated on the balanced testing data.

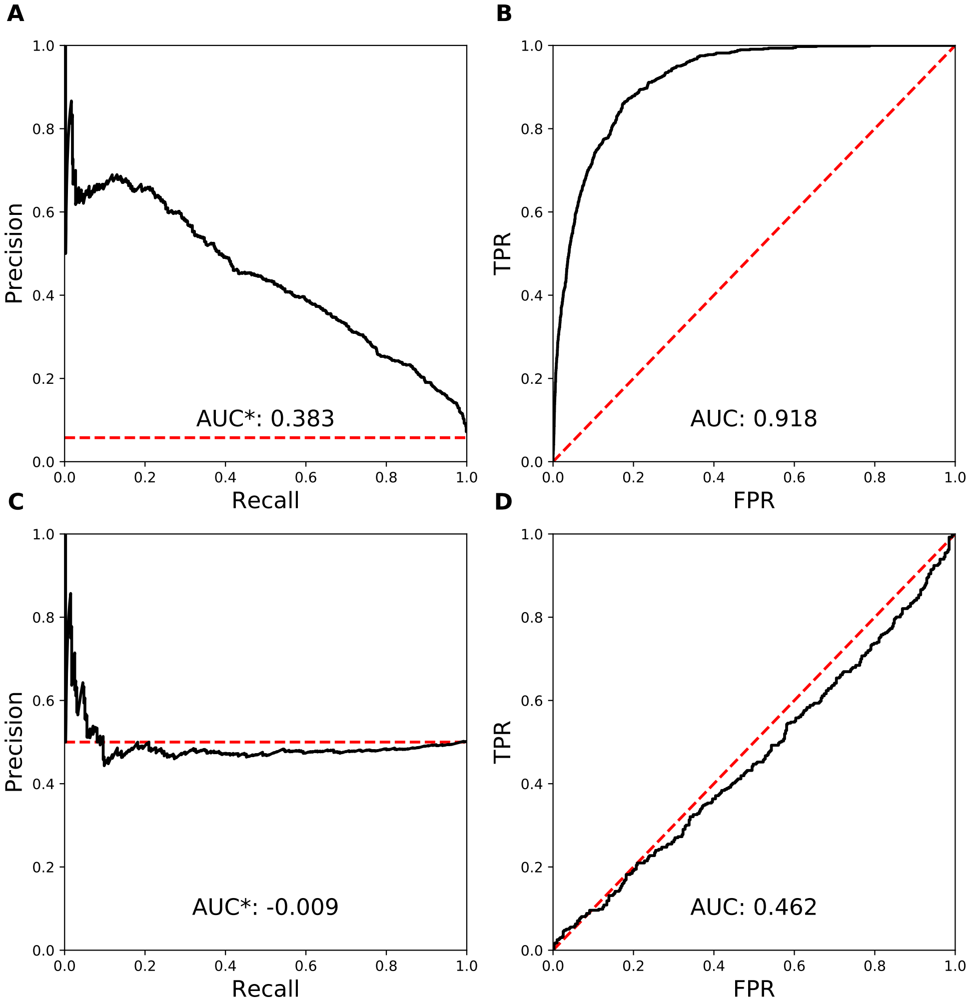

Figure E.6: **A**: PR curve evaluated for a model using $\{\lambda_k\}_{k=1}^{3N}$, $\{r_k\}_{k=1}^{3N}$, and U as inputs trained on data from a single run at QSS of condition $C_{3,3}$ and evaluated on the training data, when the training data labels have been randomly shuffled. The red line indicates the performance of a random model on the data set. **B**: ROC curve for the same model evaluated on the training data. **C**: PR curve for the same model evaluated on the balanced testing data. **D**: ROC curve for the same model evaluated on the balanced testing data.

Bibliography

[1] Harvey Lodish, Arnold Berk, Chris A Kaiser, Monty Krieger, Matthew P Scott, Anthony Bretscher, Hidde Ploegh, Paul Matsudaira, et al. *Molecular cell biology*. Macmillan, 2008.

[2] Daniel A Fletcher and R Dyche Mullins. Cell mechanics and the cytoskeleton. *Nature*, 463(7280):485, 2010.

[3] David Boal. *Mechanics of the Cell*. Cambridge University Press, 2012.

[4] Jonathon Howard. *Mechanics of motor proteins and the cytoskeleton*. Sinauer associates Sunderland, MA, 2001.

[5] Yu-Ling Shih and Lawrence Rothfield. The bacterial cytoskeleton. *Microbiology and Molecular Biology Reviews*, 70(3):729–754, 2006.

[6] Thijs JG Ettema, Ann-Christin Lindås, and Rolf Bernander. An actin-based cytoskeleton in archaea. *Molecular microbiology*, 80(4):1052–1061, 2011.

[7] Bill Wickstead and Keith Gull. The evolution of the cytoskeleton. *Journal of Cell Biology*, 194(4):513–525, 2011.

[8] Roberto Dominguez and Kenneth C Holmes. Actin structure and function. *Annual review of biophysics*, 40:169–186, 2011.

[9] Jorge M Ferrer, Hyungsuk Lee, Jiong Chen, Benjamin Pelz, Fumihiko Nakamura, Roger D Kamm, and Matthew J Lang. Measuring molecular rupture forces between single actin filaments and actin-binding proteins. *Proceedings of the National Academy of Sciences*, 105(27):9221–9226, 2008.

[10] Matthew J Footer, Jacob WJ Kerssemakers, Julie A Theriot, and Marileen Dogterom. Direct measurement of force generation by actin filament polymerization using an optical trap. *Proceedings of the National Academy of Sciences*, 104(7):2181–2186, 2007.

[11] Thomas D Pollard. Rate constants for the reactions of atp-and adp-actin with the ends of actin filaments. *The Journal of cell biology*, 103(6):2747–2754, 1986.

[12] Ikuko Fujiwara, Dimitrios Vavylonis, and Thomas D Pollard. Polymerization kinetics of adp-and adp-pi-actin determined by fluorescence microscopy. *Proceedings of the National Academy of Sciences*, 104(21):8827–8832, 2007.

[13] HJ Kinosian, LA Selden, JE Estes, and LC Gershman. Nucleotide binding to actin. cation dependence of nucleotide dissociation and exchange rates. *Journal of Biological Chemistry*, 268(12):8683–8691, 1993.

[14] Edward D Korn, Marie-France Carlier, and Dominique Pantaloni. Actin polymerization and atp hydrolysis. *Science*, 238(4827):638–645, 1987.

[15] Lynn A Selden, Henry J Kinosian, James E Estes, and Lewis C Gershman. Impact of profilin on actin-bound nucleotide exchange and actin polymerization dynamics. *Biochemistry*, 38(9):2769–2778, 1999.

[16] Albrecht Wegner. Head to tail polymerization of actin. *Journal of molecular biology*, 108(1):139–150, 1976.

[17] Rudolf K Meyer and Ueli Aebi. Bundling of actin filaments by alpha-actinin depends on its molecular length. *The Journal of cell biology*, 110(6):2013–2024, 1990.

[18] Emil Reisler, Craig Smith, and George Seegan. Myosin minifilaments. *Journal of molecular biology*, 143(1):129–145, 1980.

[19] Anne R Bresnick. Molecular mechanisms of nonmuscle myosin-ii regulation. *Current opinion in cell biology*, 11(1):26–33, 1999.

[20] Philip W Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[21] John H Holland. *Hidden order: How adaptation builds complexity.* Addison Wesley Longman Publishing Co., Inc., 1996.

[22] Stuart A Kauffman. Autocatalytic sets of proteins. *Journal of theoretical biology*, 119(1):1–24, 1986.

[23] Gregoire Nicolis. Self-organization in nonequilibrium systems. *Dissipative Structures to Order through Fluctuations*, pages 339–426, 1977.

[24] Dirk Helbing and Tamás Vicsek. Optimal self-organization. *New journal of physics*, 1(1):13, 1999.

[25] Herbert B Callen. Thermodynamics and an introduction to thermostatistics, 1998.

[26] Terrell L Hill. *Free energy transduction and biochemical cycle kinetics.* Courier Corporation, 2004.

[27] Terrell L Hill and Marc W Kirschner. Subunit treadmilling of microtubules or actin in the presence of cellular barriers: possible conversion of chemical free energy into mechanical work. *Proceedings of the National Academy of Sciences*, 79(2):490–494, 1982.

[28] Massimo Reconditi, Marco Linari, Leonardo Lucii, Alex Stewart, Yin-Biao Sun, Peter Boesecke, Theyencheri Narayanan, Robert F Fischetti, Tom Irving, Gabriella Piazzesi, et al. The myosin motor in muscle generates a smaller and slower working stroke at higher load. *Nature*, 428(6982):578–581, 2004.

[29] Haruo Sugi, Hiroyuki Iwamoto, Tsuyoshi Akimoto, and Hirohiko Kishi. High mechanical efficiency of the cross-bridge powerstroke in skeletal muscle. *Journal of experimental biology*, 206(7):1201–1206, 2003.

[30] Chris J Barclay, Rodger C Woledge, and Nancy A Curtin. Inferring crossbridge properties from skeletal muscle energetics. *Progress in biophysics and molecular biology*, 102(1):53–71, 2010.

[31] David W Maughan. Kinetics and energetics of the crossbridge cycle. *Heart failure reviews*, 10(3):175–185, 2005.

[32] Christopher Battle, Chase P Broedersz, Nikta Fakhri, Veikko F Geyer, Jonathon Howard, Christoph F Schmidt, and Fred C MacKintosh. Broken detailed balance at mesoscopic scales in active biological systems. *Science*, 352(6285):604–607, 2016.

[33] J Gladrow, N Fakhri, FC MacKintosh, CF Schmidt, and CP Broedersz. Broken detailed balance of filament dynamics in active networks. *Physical review letters*, 116(24):248301, 2016.

[34] Daniel S Seara, Vikrant Yadav, Ian Linsmeier, A Pasha Tabatabai, Patrick W Oakes, SM Ali Tabei, Shiladitya Banerjee, and Michael P Murrell. Entropy production rate is maximized in non-contractile actomyosin. *Nature communications*, 9(1):4948, 2018.

[35] Junang Li, Jordan M Horowitz, Todd R Gingrich, and Nikta Fakhri. Quantifying dissipation using fluctuating currents. *Nature communications*, 10(1):1–9, 2019.

[36] Massimiliano Esposito. Stochastic thermodynamics under coarse graining. *Physical Review E*, 85(4):041125, 2012.

[37] Daan Frenkel and Berend Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Elsevier, 2001.

[38] Herman Berendsen. *Simulating the physical world.* Cambridge University Press Cambridge, 2007.

[39] Simon L Freedman, Shiladitya Banerjee, Glen M Hocky, and Aaron R Dinner. A versatile framework for simulating the dynamic mechanical structure of cytoskeletal networks. *Biophysical journal*, 113(2):448–460, 2017.

[40] Francois Nedelec and Dietrich Foethke. Collective langevin dynamics of flexible cytoskeletal fibers. *New Journal of Physics*, 9(11):427, 2007.

[41] Taeyoon Kim, Wonmuk Hwang, Hyungsuk Lee, and Roger D Kamm. Computational analysis of viscoelastic properties of crosslinked actin networks. *PLoS Comput Biol*, 5(7):e1000439, 2009.

[42] Konstantin Popov, James Komianos, and Garegin A Papoian. Medyan: mechanochemical simulations of contraction and polarity alignment in actomyosin networks. *PLoS computational biology*, 12(4):e1004877, 2016.

[43] Chris A Kaiser, Monty Krieger, Harvey Lodish, and Arnold Berk. *Molecular cell biology.* WH Freeman, 2007.

[44] Rob Phillips, Julie Theriot, Jane Kondev, and Hernan Garcia. *Physical biology of the cell.* Garland Science, 2012.

[45] Jozef Šamaj, František Baluška, Boris Voigt, Markus Schlicht, Dieter Volkmann, and Diedrik Menzel. Endocytosis, actin cytoskeleton, and signaling. *Plant physiology*, 135(3):1150–1161, 2004.

[46] Britta Qualmann, Michael M Kessels, and Regis B Kelly. Molecular links between endocytosis and the actin cytoskeleton. *The Journal of cell biology*, 150(5):F111–F116, 2000.

[47] Elizabeth Smythe and Kathryn R Ayscough. Actin regulation in endocytosis. *J Cell Sci*, 119(22):4589–4598, 2006.

[48] Alan Rick Horwitz and J Thomas Parsons. Cell migration–movin'on. *Science*, 286(5442):1102–1103, 1999.

[49] Bernhard Wehrle-Haller and Beat A Imhof. Actin, microtubules and focal adhesion dynamics during cell migration. *The international journal of biochemistry & cell biology*, 35(1):39–50, 2003.

[50] Ulrike S Eggert, Timothy J Mitchison, and Christine M Field. Animal cytokinesis: from parts list to mechanisms. *Annu. Rev. Biochem.*, 75:543–566, 2006.

[51] Robert J Pelham Jr and Fred Chang. Actin dynamics in the contractile ring during cytokinesis in fission yeast. *Nature*, 419(6902):82, 2002.

[52] Roderick P Tas and Lukas C Kapitein. Exploring cytoskeletal diversity in neurons. *Science*, 361(6399):231–232, 2018.

[53] Thomas D Pollard and John A Cooper. Actin and actin-binding proteins. a critical evaluation of mechanisms and functions. *Annual review of biochemistry*, 55(1):987–1035, 1986.

[54] Steven J Winder and Kathryn R Ayscough. Actin-binding proteins. *Journal of cell science*, 118(4):651–654, 2005.

[55] Thomas D Pollard. Actin and actin-binding proteins. *Cold Spring Harbor perspectives in biology*, 8(8):a018226, 2016.

[56] Wenliang Lei, Omotola F Omotade, Kenneth R Myers, and James Q Zheng. Actin cytoskeleton in dendritic spine development and plasticity. *Current opinion in neurobiology*, 39:86–92, 2016.

[57] Ildikó Kristó, Izabella Bajusz, Csaba Bajusz, Péter Borkúti, and Péter Vilmos. Actin, actin-binding proteins, and actin-related proteins in the nucleus. *Histochemistry and cell biology*, 145(4):373–388, 2016.

[58] Alexandra Krol, Steven J Henle, and Lisa V Goodrich. Fat3 and ena/vasp proteins influence the emergence of asymmetric cell morphology in the developing retina. *Development*, 143(12):2172–2182, 2016.

[59] Jonathan D Winkelman, Colleen G Bilancia, Mark Peifer, and David R Kovar. Ena/vasp enabled is a highly processive actin polymerase tailored to self-assemble parallel-bundled f-actin networks with fascin. *Proceedings of the National Academy of Sciences*, 111(11):4121–4126, 2014.

[60] Shashank Shekhar, Mikael Kerleau, Sonja Kühn, Julien Pernier, Guillaume Romet-Lemonne, Antoine Jégou, and Marie-France Carlier. Formin and capping protein together embrace the actin filament in a ménage à trois. *Nature communications*, 6:8730, 2015.

[61] Kaustubh Wagh, Brittany A Wheatley, Maria K Traver, Imran Hussain, Brian C Schaefer, and Arpita Upadhyaya. Bcl10 is associated with actin dynamics at the t cell immune synapse. *Cellular Immunology*, 356:104161, 2020.

[62] Ivan Rey-Suarez, Brittany A Wheatley, Peter Koo, Anshuman Bhanja, Zhou Shu, Simon Mochrie, Wenxia Song, Hari Shroff, and Arpita Upadhyaya. Wasp family proteins regulate the mobility of the b cell receptor during signaling activation. *Nature communications*, 11(1):1–14, 2020.

[63] Ke Xu, Hazen P Babcock, and Xiaowei Zhuang. Dual-objective storm reveals three-dimensional filament organization in the actin cytoskeleton. *Nature methods*, 9(2):185–188, 2012.

[64] Gaëlle Letort, Hajer Ennomani, Laurène Gressin, Manuel Théry, and Laurent Blanchoin. Dynamic reorganization of the actin cytoskeleton. *F1000Research*, 4, 2015.

[65] Adar Sonn-Segev, Anne Bernheim-Groswasser, and Yael Roichman. Dynamics in steady state in vitro acto-myosin networks. *Journal of Physics: Condensed Matter*, 29(16):163002, 2017.

[66] Jr Ede A Ribeiro, Nikos Pinotsis, Andrea Ghisleni, Anita Salmazo, Petr V Konarev, Julius Kostan, Björn Sjöblom, Claudia Schreiner, Anton A Polyansky, Eirini A Gkougkoulia, et al. The structure and regulation of human muscle α-actinin. *Cell*, 159(6):1447–1460, 2014.

[67] Volker Schaller, Christoph Weber, Christine Semmrich, Erwin Frey, and Andreas R Bausch. Polar patterns of driven filaments. *Nature*, 467(7311):73, 2010.

[68] Tim Sanchez, Daniel TN Chen, Stephen J DeCamp, Michael Heymann, and Zvonimir Dogic. Spontaneous motion in hierarchically assembled active matter. *Nature*, 491(7424):431–434, 2012.

[69] Ian Linsmeier, Shiladitya Banerjee, Patrick W Oakes, Wonyeong Jung, Taeyoon Kim, and Michael P Murrell. Disordered actomyosin networks are sufficient to produce cooperative and telescopic contractility. *Nature communications*, 7(1):1–9, 2016.

[70] Miguel Vicente-Manzanares, Xuefei Ma, Robert S Adelstein, and Alan Rick Horwitz. Non-muscle myosin ii takes centre stage in cell adhesion and migration. *Nature reviews Molecular cell biology*, 10(11):778, 2009.

[71] Carol A Otey and Olli Carpen. α-actinin revisited: A fresh look at an old player. *Cell motility and the cytoskeleton*, 58(2):104–111, 2004.

[72] Daniel H Wachsstock, WH Schwartz, and Thomas D Pollard. Affinity of alpha-actinin for actin determines the structure and mechanical properties of actin filament gels. *Biophysical journal*, 65(1):205–214, 1993.

[73] Sriram Ramaswamy. The mechanics and statistics of active matter. *Annual Review of Condensed Matter Physics*, 1:323–345, 2010.

[74] Jonathan Colen, Ming Han, Rui Zhang, Steven A Redford, Linnea M Lemma, Link Morgan, Paul V Ruijgrok, Raymond Adkins, Zev Bryant, Zvonimir Dogic, et al. Machine learning active-nematic hydrodynamics. *Proceedings of the National Academy of Sciences*, 118(10), 2021.

[75] Sho C Takatori, Wen Yan, and John F Brady. Swim pressure: stress generation in active matter. *Physical review letters*, 113(2):028103, 2014.

[76] Trevor GrandPre, Katherine Klymko, Kranthi K Mandadapu, and David T Limmer. Entropy production fluctuations encode collective behavior in active matter. *Physical Review E*, 103(1):012613, 2021.

[77] Rui Zhang, Steven A Redford, Paul V Ruijgrok, Nitin Kumar, Ali Mozaffari, Sasha Zemsky, Aaron R Dinner, Vincenzo Vitelli, Zev Bryant, Margaret L Gardel, et al. Spatiotemporal control of liquid crystal structure and dynamics through activity patterning. *Nature Materials*, pages 1–8, 2021.

[78] Tyler D Ross, Heun Jin Lee, Zijie Qu, Rachel A Banks, Rob Phillips, and Matt Thomson. Controlling organization and forces in active matter through optically defined boundaries. *Nature*, 572(7768):224–229, 2019.

[79] Francesco Ginelli. The physics of the vicsek model. *The European Physical Journal Special Topics*, 225(11-12):2099–2117, 2016.

[80] Thierry Mora, Aleksandra M Walczak, Lorenzo Del Castello, Francesco Ginelli, Stefania Melillo, Leonardo Parisi, Massimiliano Viale, Andrea Cavagna, and Irene Giardina. Local equilibrium in bird flocks. *Nature physics*, 12(12):1153, 2016.

[81] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226, 1995.

[82] Christian Scholz, Michael Engel, and Thorsten Pöschel. Rotating robots move collectively and self-organize. *Nature communications*, 9(1):931, 2018.

[83] Pavel Chvykov, Thomas A Berrueta, Akash Vardhan, William Savoie, Alexander Samland, Todd D Murphey, Kurt Wiesenfeld, Daniel I Goldman, and Jeremy L England. Low rattling: A predictive principle for self-organization in active collectives. *Science*, 371(6524):90–95, 2021.

[84] D Dell'Arciprete, ML Blow, AT Brown, FDC Farrell, Juho S Lintuvuori, AF McVey, D Marenduzzo, and Wilson CK Poon. A growing bacterial colony in two dimensions as an active nematic. *Nature communications*, 9(1):4190, 2018.

[85] Jacques Prost, Frank Jülicher, and Jean-François Joanny. Active gel physics. *Nature physics*, 11(2):111–117, 2015.

[86] M Cristina Marchetti, Jean-François Joanny, Sriram Ramaswamy, Tanniemola B Liverpool, Jacques Prost, Madan Rao, and R Aditi Simha. Hydrodynamics of soft active matter. *Reviews of Modern Physics*, 85(3):1143, 2013.

[87] David Keller and Carlos Bustamante. The mechanochemistry of molecular motors. *Biophysical Journal*, 78(2):541–556, 2000.

[88] Mihály Kovács, Fei Wang, Aihua Hu, Yue Zhang, and James R Sellers. Functional divergence of human cytoplasmic myosin ii kinetic characterization of the non-muscle iia isoform. *Journal of Biological Chemistry*, 278(40):38132–38140, 2003.

[89] Xuejun C Zhang and Wei Feng. Thermodynamic aspects of atp hydrolysis of actomyosin complex. *Biophysics reports*, 2(5-6):87–94, 2016.

[90] Thorsten Erdmann, Kathrin Bartelheimer, and Ulrich S Schwarz. Sensitivity of small myosin ii ensembles from different isoforms to mechanical load and atp concentration. *Physical Review E*, 94(5):052403, 2016.

[91] Martin McCullagh, Marissa G Saunders, and Gregory A Voth. Unraveling the mystery of atp hydrolysis in actin filaments. *Journal of the American Chemical Society*, 136(37):13053–13058, 2014.

[92] Xin Li, Jan Kierfeld, and Reinhard Lipowsky. Actin polymerization and depolymerization coupled to cooperative hydrolysis. *Physical review letters*, 103(4):048102, 2009.

[93] Thorsten Erdmann, Philipp J Albert, and Ulrich S Schwarz. Stochastic dynamics of small ensembles of non-processive molecular motors: The parallel cluster model. *The Journal of chemical physics*, 139(17):11B604_1, 2013.

[94] James E Komianos and Garegin A Papoian. Stochastic ratcheting on a funneled energy landscape is necessary for highly efficient contractility of actomyosin force dipoles. *Physical Review X*, 8(2):021006, 2018.

[95] Hajer Ennomani, Gaëlle Letort, Christophe Guérin, Jean-Louis Martiel, Wenxiang Cao, François Nédélec, M Enrique, Manuel Théry, and Laurent Blanchoin. Architecture and connectivity govern actin network contractility. *Current Biology*, 26(5):616–626, 2016.

[96] Sihan Chen, Tomer Markovich, and Fred C MacKintosh. Motor-free contractility in active gels. *Physical Review Letters*, 125(20):208101, 2020.

[97] Wonyeong Jung, Michael P Murrell, and Taeyoon Kim. F-actin cross-linking enhances the stability of force generation in disordered actomyosin networks. *Computational particle mechanics*, 2(4):317–327, 2015.

[98] Elena G Yarmola, Dmitri A Dranishnikov, and Michael R Bubb. Effect of profilin on actin critical concentration: a theoretical analysis. *Biophysical journal*, 95(12):5544–5573, 2008.

[99] Carlos Floyd, Christopher Jarzynski, and Garegin Papoian. Low-dimensional manifold of actin polymerization dynamics. *New Journal of Physics*, 19(12):125012, 2017.

[100] Julie A Theriot. The polymerization motor. *Traffic*, 1(1):19–28, 2000.

[101] Charles S Peskin, Garrett M Odell, and George F Oster. Cellular motions and thermal fluctuations: the brownian ratchet. *Biophysical journal*, 65(1):316–324, 1993.

[102] Alexander Mogilner and George Oster. The physics of lamellipodial protrusion. *European biophysics journal*, 25(1):47–53, 1996.

[103] Philipp J Albert, Thorsten Erdmann, and Ulrich S Schwarz. Stochastic dynamics and mechanosensitivity of myosin ii minifilaments. *New Journal of Physics*, 16(9):093019, 2014.

[104] José Alvarado, Michael Sheinman, Abhinav Sharma, Fred C MacKintosh, and Gijsje H Koenderink. Force percolation of contractile active gels. *Soft matter*, 13(34):5624–5644, 2017.

[105] Cornelis Storm, Jennifer J Pastore, Fred C MacKintosh, Tom C Lubensky, and Paul A Janmey. Nonlinear elasticity in biological gels. *Nature*, 435(7039):191–194, 2005.

[106] Fred C MacKintosh and Alex J Levine. Nonequilibrium mechanics and dynamics of motor-activated gels. *Physical review letters*, 100(1):018104, 2008.

[107] Predrag Bursac, Guillaume Lenormand, Ben Fabry, Madavi Oliver, David A Weitz, Virgile Viasnoff, James P Butler, and Jeffrey J Fredberg. Cytoskeletal remodelling and slow dynamics in the living cell. *Nature materials*, 4(7):557–561, 2005.

[108] Shenshen Wang and Peter G Wolynes. Communication: Effective temperature and glassy dynamics of active matter, 2011.

[109] Shenshen Wang and Peter G Wolynes. Microscopic theory of the glassy dynamics of passive and active network materials. *The Journal of chemical physics*, 138(12):12A521, 2013.

[110] Vilmos Zsolnay, Harshwardhan H Katkar, Steven Z Chou, Thomas D Pollard, and Gregory A Voth. Structural basis for polarized elongation of actin filaments. *Proceedings of the National Academy of Sciences*, 117(48):30458–30464, 2020.

[111] Sriramvignesh Mani, Harshwardhan H Katkar, and Gregory A Voth. Compressive and tensile deformations alter atp hydrolysis and phosphate release rates in actin filaments. *Journal of Chemical Theory and Computation*, 2021.

[112] Hidetaka Yamaoka, Shinji Matsushita, Yoshitaka Shimada, and Taiji Adachi. Multiscale modeling and mechanics of filamentous actin cytoskeleton. *Biomechanics and modeling in mechanobiology*, 11(3):291–302, 2012.

[113] Richard C Tolman. The principle of microscopic reversibility. *Proceedings of the National Academy of Sciences of the United States of America*, 11(7):436, 1925.

[114] Gavin E Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *Journal of Statistical Physics*, 90(5-6):1481–1487, 1998.

[115] Sai Vinjanampathy and Janet Anders. Quantum thermodynamics. *Contemporary Physics*, 57(4):545–579, 2016.

[116] Sebastian Deffner and Steve Campbell. *Quantum Thermodynamics: An introduction to the thermodynamics of quantum information.* Morgan & Claypool Publishers, 2019.

[117] Andrew Smith, Yao Lu, Shuoming An, Xiang Zhang, Jing-Ning Zhang, Zongping Gong, HT Quan, Christopher Jarzynski, and Kihwan Kim. Verification of the quantum nonequilibrium work relation in the presence of decoherence. *New Journal of Physics*, 20(1):013008, 2018.

[118] J Willard Gibbs. *Elementary principles in statistical mechanics.* Courier Corporation, 2014.

[119] Christopher Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.*, 2(1):329–351, 2011.

[120] Christopher Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690, 1997.

[121] Edward Ott. *Chaos in dynamical systems.* Cambridge university press, 2002.

[122] Robert Zwanzig. *Nonequilibrium statistical mechanics.* Oxford University Press, 2001.

[123] Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.

[124] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[125] Ken Sekimoto. *Stochastic energetics*, volume 799. Springer, 2010.

[126] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Physical review letters*, 95(4):040602, 2005.

[127] Udo Seifert. Stochastic thermodynamics: principles and perspectives. *The European Physical Journal B*, 64(3):423–431, 2008.

[128] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on progress in physics*, 75(12):126001, 2012.

[129] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after kramers. *Reviews of modern physics*, 62(2):251, 1990.

[130] Todd R Gingrich, Grant M Rotskoff, and Jordan M Horowitz. Inferring dissipation from current fluctuations. *Journal of Physics A: Mathematical and Theoretical*, 50(18):184004, 2017.

[131] Crispin W Gardiner et al. *Handbook of stochastic methods*, volume 3. springer Berlin, 1985.

[132] Y Mishin. Thermodynamic theory of equilibrium fluctuations. *Annals of Physics*, 363:48–97, 2015.

[133] Tongye Shen and Peter G Wolynes. Stability and dynamics of crystals and glasses of motorized particles. *Proceedings of the National Academy of Sciences*, 101(23):8547–8550, 2004.

[134] JM Rubi. Mesoscopic thermodynamics. *Physica Scripta*, 2012(T151):014027, 2012.

[135] Ilya Prigogine, Ilya Prigogine, Chimiste Physicien, Ilya Prigogine, Chemist Physicist, and Ilya Prigogine. *Introduction to thermodynamics of irreversible processes*, volume 7. Interscience Publishers New York, 1961.

[136] Vladimir Y Chernyak, Michael Chertkov, and Christopher Jarzynski. Path-integral analysis of fluctuation theorems for general langevin processes. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(08):P08001, 2006.

[137] Ilya Prigogine and Isabelle Stengers. *Order out of chaos: Man's new dialogue with nature*. Verso Books, 2018.

[138] René Lefever, Grégoire Nicolis, and Pierre Borckmans. The brusselator: it does oscillate all the same. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases*, 84(4):1013–1023, 1988.

[139] Marja Bister and Kerry A Emanuel. Dissipative heating and hurricane intensity. *Meteorology and Atmospheric Physics*, 65(3):233–240, 1998.

[140] Ernst L Koschmieder. *Bénard cells and Taylor vortices*. Cambridge University Press, 1993.

[141] H von Helmholtz. Zur theorie der stationären ströme in reibenden flüssigkeiten. *Wiss. Abh*, 1:223–230, 1868.

[142] Lars Onsager. Reciprocal relations in irreversible processes. i. *Physical review*, 37(4):405, 1931.

[143] Istvan Gyarmati et al. *Non-equilibrium thermodynamics*. Springer, 1970.

[144] Hans Ziegler. A possible generalization of onsager's theory. In *Irreversible Aspects of Continuum Mechanics and Transfer of Physical Characteristics in Moving Fluids*, pages 411–424. Springer, 1968.

[145] Umberto Lucia. Maximum or minimum entropy generation for open systems? *Physica A: Statistical Mechanics and its Applications*, 391(12):3392–3398, 2012.

[146] A Di Vita. Maximum or minimum entropy production? how to select a necessary criterion of stability for a dissipative fluid or plasma. *Physical Review E*, 81(4):041137, 2010.

[147] Edwin T Jaynes. The minimum entropy production principle. *Annual Review of Physical Chemistry*, 31(1):579–601, 1980.

[148] Walter T Grandy Jr. *Entropy and the time evolution of macroscopic systems*, volume 10. Oxford University Press on Demand, 2008.

[149] Leonid M Martyushev and Vladimir D Seleznev. Maximum entropy production principle in physics, chemistry and biology. *Physics reports*, 426(1):1–45, 2006.

[150] Gavin E Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721, 1999.

[151] Gavin E Crooks. Path-ensemble averages in systems driven far from equilibrium. *Physical review E*, 61(3):2361, 2000.

[152] Jeremy L England. Dissipative adaptation in driven self-assembly. *Nature nanotechnology*, 10(11):919, 2015.

[153] Jordan M Horowitz and Jeremy L England. Spontaneous fine-tuning to environment in many-species chemical reaction networks. *Proceedings of the National Academy of Sciences*, 114(29):7565–7570, 2017.

[154] Tal Kachman, Jeremy A Owen, and Jeremy L England. Self-organized resonance during search of a diverse chemical space. *Physical review letters*, 119(3):038001, 2017.

[155] Nikolay Perunov, Robert A Marsland, and Jeremy L England. Statistical physics of adaptation. *Physical Review X*, 6(2):021036, 2016.

[156] Pavel Chvykov and Jeremy England. Least-rattling feedback from strong time-scale separation. *Physical Review E*, 97(3):032115, 2018.

[157] James P Sethna, Karin A Dahmen, and Christopher R Myers. Crackling noise. *Nature*, 410(6825):242–250, 2001.

[158] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, 59(4):381, 1987.

[159] Stefan Hergarten. *Self organized criticality in earth systems*, volume 2. Springer, 2002.

[160] Donald L Turcotte. Fractals in geology and geophysics. *Pure and applied Geophysics*, 131(1-2):171–196, 1989.

[161] Per Bak and Chao Tang. Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research: Solid Earth*, 94(B11):15635–15637, 1989.

[162] Per Bak, Kim Christensen, Leon Danon, and Tim Scanlon. Unified scaling law for earthquakes. *Physical Review Letters*, 88(17):178501, 2002.

[163] David Chandler. Introduction to modern statistical. *Mechanics. Oxford University Press, Oxford, UK*, 40, 1987.

[164] Chon-Kit Pun, Sakib Matin, W Klein, and Harvey Gould. Prediction in a driven-dissipative system displaying a continuous phase transition using machine learning. *Physical Review E*, 101(2):022102, 2020.

[165] Lucilla de Arcangelis, Carla Perrone-Capano, and Hans J Herrmann. Self-organized criticality model for brain plasticity. *Physical review letters*, 96(2):028107, 2006.

[166] Luca Cardamone, Alessandro Laio, Vincent Torre, Rajesh Shahapure, and Antonio DeSimone. Cytoskeletal actin networks in motile cells are critically self-organized systems synchronized by mechanical interactions. *Proceedings of the National Academy of Sciences*, 108(34):13978–13983, 2011.

[167] Josué X De Carvalho and Carmen PC Prado. Self-organized criticality in the olami-feder-christensen model. *Physical review letters*, 84(17):4006, 2000.

[168] Adriano Mesquita Alencar, Mariana Sacrini Ayres Ferraz, Chan Young Park, Emil Millet, Xavier Trepat, Jeffrey J Fredberg, and James P Butler. Non-equilibrium cytoquake dynamics in cytoskeletal remodeling and stabilization. *Soft matter*, 12(41):8506–8511, 2016.

[169] Yu Shi, Christopher L Porter, John C Crocker, and Daniel H Reich. Dissecting fat-tailed fluctuations in the cytoskeleton with active micropost arrays. *Proceedings of the National Academy of Sciences*, page 201900963, 2019.

[170] Laetitia Kurzawa, Benoit Vianay, Fabrice Senger, Timothée Vignaud, Laurent Blanchoin, and Manuel Théry. Dissipation of contractile forces: the missing piece in cell mechanics. *Molecular biology of the cell*, 28(14):1825–1832, 2017.

[171] Julie A Theriot and Timothy J Mitchison. Actin microfilament dynamics in locomoting cells. *Nature*, 352(6331):126, 1991.

[172] Anja Schmidt and Michael N Hall. Signaling to the actin cytoskeleton. *Annual review of cell and developmental biology*, 14(1):305–338, 1998.

[173] Viola Vogel and Michael Sheetz. Local force and geometry sensing regulate cell functions. *Nature reviews Molecular cell biology*, 7(4):265–275, 2006.

[174] Marie-France Carlier, Valérie Laurent, Jérôme Santolini, Ronald Melki, Dominique Didry, Gui-Xian Xia, Yan Hong, Nam-Hai Chua, and Dominique Pantaloni. Actin depolymerizing factor (adf/cofilin) enhances the rate of filament turnover: implication in actin-based motility. *The Journal of cell biology*, 136(6):1307–1322, 1997.

[175] Ernesto Andrianantoandro and Thomas D Pollard. Mechanism of actin filament turnover by severing and nucleation at different concentrations of adf/cofilin. *Molecular cell*, 24(1):13–23, 2006.

[176] Herve Isambert, Pascal Venier, Anthony C Maggs, Abdelatif Fattoum, Ridha Kassab, Dominique Pantaloni, and Marie-France Carlier. Flexibility of actin filaments derived from thermal fluctuations. effect of bound nucleotide, phalloidin, and muscle regulatory proteins. *Journal of Biological Chemistry*, 270(19):11437–11444, 1995.

[177] Albrecht Wegner and Juergen Engel. Kinetics of the cooperative association of actin to actin filament. *Biophysical chemistry*, 3(3):215–225, 1975.

[178] Roger Cooke. Role of the bound nucleotide in the polymerization of actin. *Biochemistry*, 14(14):3250–3256, 1975.

[179] Dominique Pantaloni, Terrell L Hill, Marie-France Carlier, and Edward D Korn. A model for actin polymerization and the kinetic effects of atp hydrolysis. *Proceedings of the National Academy of Sciences*, 82(21):7207–7211, 1985.

[180] Paul A Dufort and Charles J Lumsden. How profilin/barbed-end synergy controls actin polymerization: A kinetic model of the atp hydrolysis circuit. *Cytoskeleton*, 35(4):309–330, 1996.

[181] M Bindschadler, EA Osborn, CF Dewey, and JL McGrath. A mechanistic model of the actin cycle. *Biophysical journal*, 86(5):2720–2739, 2004.

[182] Evgeny B Stukalin and Anatoly B Kolomeisky. Atp hydrolysis stimulates large length fluctuations in single actin filaments. *Biophysical journal*, 90(8):2673–2685, 2006.

[183] FJ Brooks and AE Carlsson. Nonequilibrium actin polymerization treated by a truncated rate-equation method. *Physical Review E*, 79(3):031914, 2009.

[184] FJ Brooks and AE Carlsson. Actin polymerization overshoots and atp hydrolysis as assayed by pyrene fluorescence. *Biophysical journal*, 95(3):1050–1062, 2008.

[185] Stephen L Campbell, Carl D Meyer, Jr, and Nicholas J Rose. Applications of the drazin inverse to linear systems of differential equations with singular constant coefficients. *SIAM Journal on Applied Mathematics*, 31(3):411–425, 1976.

[186] Ramon Grima and Santiago Schnell. Modelling reaction kinetics inside cells. *Essays in biochemistry*, 45:41–56, 2008.

[187] Natsuki Tanaka and Garegin A Papoian. Reverse-engineering of biochemical reaction networks from spatio-temporal correlations of fluorescence fluctuations. *Journal of theoretical biology*, 264(2):490–500, 2010.

[188] Florence Baras and M Malek Mansour. Reaction-diffusion master equation: A comparison with microscopic simulations. *Physical Review E*, 54(6):6139, 1996.

[189] David Bernstein. Simulating mesoscopic reaction-diffusion systems using the gillespie algorithm. *Physical Review E*, 71(4):041103, 2005.

[190] Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977.

[191] James Schaff, Charles C Fink, Boris Slepchenko, John H Carson, and Leslie M Loew. A general computational framework for modeling cellular structure and function. *Biophysical journal*, 73(3):1135–1146, 1997.

[192] Leslie M Loew and James C Schaff. The virtual cell: a software environment for computational cell biology. *TRENDS in Biotechnology*, 19(10):401–406, 2001.

[193] Elijah Roberts, John E Stone, and Zaida Luthey-Schulten. Lattice microbes: High-performance stochastic simulation method for the reaction-diffusion master equation. *Journal of computational chemistry*, 34(3):245–255, 2013.

[194] Michael J Hallock and Zaida Luthey-Schulten. Improving reaction kernel performance in lattice microbes: particle-wise propensities and run-time generated code. In *2016 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 428–434. IEEE, 2016.

[195] Carlos Floyd, Garegin A Papoian, and Christopher Jarzynski. Quantifying dissipation in actomyosin networks. *Interface focus*, 9(3):20180078, 2019.

[196] Tyler M Earnest, John A Cole, and Zaida Luthey-Schulten. Simulating biological processes: stochastic physics from whole cells to colonies. *Reports on Progress in Physics*, 81(5):052601, 2018.

[197] Daniel A Beard and Hong Qian. Relationship between thermodynamic driving force and one-way fluxes in reversible processes. *PloS one*, 2(1):e144, 2007.

[198] Ilya Prigogine. Introduction to thermodynamics of irreversible processes. *New York: Interscience, 1967, 3rd ed.*, 1967.

[199] Riccardo Rao and Massimiliano Esposito. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Physical Review X*, 6(4):041064, 2016.

[200] Jeremy L England. Statistical physics of self-replication. *The Journal of chemical physics*, 139(12):09B623_1, 2013.

[201] Ilya Prigogine and Grégoire Nicolis. Biological order, structure and instabilities. *Quarterly Reviews of Biophysics*, 4(2-3):107–148, 1971.

[202] Phil Attard. *Thermodynamics and statistical mechanics*. CUP Archive, 2002.

[203] Terrell L Hill. Perspective: nanothermodynamics, 2001.

[204] Terrell L Hill. *Thermodynamics of small systems*. Courier Corporation, 1994.

[205] J de Heer. The free energy charge accompanying a chemical reaction and the gibbs-duhem equation. *J. Chem. Educ.*, 63(11):950, 1986.

[206] Thomas Engel, Philip Reid, et al. *Physical chemistry*. Pearson, 2006.

[207] Daniel M Anderson, James D Benson, and Anthony J Kearsley. Foundations of modeling in cryobiology—i: Concentration, gibbs energy, and chemical potential relationships. *Cryobiology*, 69(3):349–360, 2014.

[208] Gordon M Barrow. Free energy and equilibrium: the basis of g0=-rt in k for reactions in solution. *J. Chem. Educ.*, 60(8):648, 1983.

[209] Ron Milo and Rob Phillips. *Cell biology by the numbers.* Garland Science, 2015.

[210] Michael A Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889, 2000.

[211] Javier Borge. Reviewing some crucial concepts of gibbs energy in chemical equilibrium using a computer-assisted, guided-problem-solving approach. *Journal of Chemical Education*, 92(2):296–304, 2014.

[212] Radek Erban, Mark B Flegg, and Garegin A Papoian. Multiscale stochastic reaction–diffusion modeling: application to actin dynamics in filopodia. *Bulletin of mathematical biology*, 76(4):799–818, 2014.

[213] Ulrich Dobramysl, Garegin A Papoian, and Radek Erban. Steric effects induce geometric remodeling of actin bundles in filopodia. *Biophysical journal*, 110(9):2066–2075, 2016.

[214] Yueheng Lan and Garegin A Papoian. The stochastic dynamics of filopodial growth. *Biophysical journal*, 94(10):3839–3852, 2008.

[215] Isabelle Delon and Nicholas H Brown. Integrins and the actin cytoskeleton. *Current opinion in cell biology*, 19(1):43–50, 2007.

[216] S Burlacu, PA Janmey, and J Borejdo. Distribution of actin filament lengths measured by fluorescence microscopy. *American Journal of Physiology-Cell Physiology*, 262(3):C569–C577, 1992.

[217] Marco Fritzsche, Christoph Erlenkämper, Emad Moeendarbary, Guillaume Charras, and Karsten Kruse. Actin kinetics shapes cortical network structure and mechanics. *Science advances*, 2(4):e1501337, 2016.

[218] Jonathan Stricker, Tobias Falzone, and Margaret L Gardel. Mechanics of the f-actin cytoskeleton. *Journal of biomechanics*, 43(1):9–14, 2010.

[219] Tzer Han Tan, Maya Malik Garbi, Enas Abu-Shah, Junang Li, Abhinav Sharma, Fred C MacKintosh, Kinneret Keren, Christoph F Schmidt, and Nikta Fakhri. Self-organization of stress patterns drives state transitions in actin cortices. *arXiv preprint arXiv:1603.07600*, 2016.

[220] Simone Köhler, Volker Schaller, and Andreas R Bausch. Structure formation in active networks. *Nature materials*, 10(6):462, 2011.

[221] José Alvarado, Michael Sheinman, Abhinav Sharma, Fred C MacKintosh, and Gijsje H Koenderink. Molecular motors robustly drive active gels to a critically connected state. *Nature Physics*, 9(9):591, 2013.

[222] Marco Baiesi and Christian Maes. Life efficiency does not always increase with the dissipation rate. *Journal of Physics Communications*, 2(4):045017, 2018.

[223] Martial Balland, Alain Richert, and François Gallet. The dissipative contribution of myosin ii in the cytoskeleton dynamics of myoblasts. *European Biophysics Journal*, 34(3):255–261, 2005.

[224] Demosthene Mitrossilis, Jonathan Fouchard, Axel Guiroy, Nicolas Desprat, Nicolas Rodriguez, Ben Fabry, and Atef Asnacios. Single-cell response to stiffness exhibits muscle-like behavior. *Proceedings of the National Academy of Sciences*, 106(43):18243–18248, 2009.

[225] Brenton D Hoffman, Gladys Massiera, Kathleen M Van Citters, and John C Crocker. The consensus mechanics of cultured mammalian cells. *Proceedings of the National Academy of Sciences*, 103(27):10259–10264, 2006.

[226] Longhua Hu and Garegin A Papoian. Mechano-chemical feedbacks regulate actin mesh growth in lamellipodial protrusions. *Biophysical journal*, 98(8):1375–1384, 2010.

[227] Giulio Ragazzon and Leonard J Prins. Energy consumption in chemical fuel-driven self-assembly. *Nature nanotechnology*, page 1, 2018.

[228] Christophe Leys, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4):764–766, 2013.

[229] Peter J Huber. *Robust statistics.* Springer, 2011.

[230] Longhua Hu and Garegin A Papoian. Molecular transport modulates the adaptive response of branched actin networks to an external force. *The Journal of Physical Chemistry B*, 117(42):13388–13396, 2013.

[231] Eric Jones, Travis Oliphant, and Pearu Peterson. {SciPy}: open source scientific tools for {Python}. 2014.

[232] Eckhard Limpert, Werner A Stahel, and Markus Abbt. Log-normal distributions across the sciences: Keys and clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability—normal or log-normal: That is the question. *AIBS Bulletin*, 51(5):341–352, 2001.

[233] Chikara Furusawa, Takao Suzuki, Akiko Kashiwagi, Tetsuya Yomo, and Kunihiko Kaneko. Ubiquity of log-normal distributions in intra-cellular reaction dynamics. *Biophysics*, 1:25–31, 2005.

[234] Debasis Kundu and Anubhav Manglick. Discriminating between the log-normal and gamma distributions. *Journal of the Applied Statistical Sciences*, 14:175–187, 2005.

[235] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. *Physical review A*, 38(1):364, 1988.

[236] A Sornette and D Sornette. Self-organized criticality and earthquakes. *EPL (Europhysics Letters)*, 9(3):197, 1989.

[237] M Bottiglieri and C Godano. On-off intermittency in earthquake occurrence. *Physical Review E*, 75(2):026101, 2007.

[238] Simon L Freedman, Glen M Hocky, Shiladitya Banerjee, and Aaron R Dinner. Nonequilibrium phase diagrams for actomyosin networks. *Soft matter*, 14(37):7740–7747, 2018.

[239] Tobias T Falzone, Savanna Blair, and Rae M Robertson-Anderson. Entangled f-actin displays a unique crossover to microscale nonlinearity dominated by entanglement segment dynamics. *Soft matter*, 11(22):4418–4423, 2015.

[240] William M McFadden, Patrick M McCall, Margaret L Gardel, and Edwin M Munro. Filament turnover tunes both force generation and dissipation to control long-range flows in a model actomyosin cortex. *PLoS computational biology*, 13(12):e1005811, 2017.

[241] Mark Newman. *Networks: an introduction.* Oxford university press, 2010.

[242] Albert-László Barabási. The network takeover. *Nature Physics*, 8(1):14–16, 2012.

[243] Albert-László Barabási. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.

[244] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Diameter of the world-wide web. *nature*, 401(6749):130–131, 1999.

[245] Evelien Otte and Ronald Rousseau. Social network analysis: a powerful strategy, also for the information sciences. *Journal of information Science*, 28(6):441–453, 2002.

[246] Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature reviews Molecular cell biology*, 9(10):770–780, 2008.

[247] Aravind Chandrasekaran, Arpita Upadhyaya, and Garegin A Papoian. Remarkable structural transformations of actin bundles are driven by their initial polarity, motor activity, crosslinking, and filament treadmilling. *PLoS computational biology*, 15(7):e1007156, 2019.

[248] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002.

[249] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.

[250] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004.

[251] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[252] Yossi Eliaz, Francois Nedelec, Greg Morrison, Herbert Levine, and Margaret S Cheung. Insights from graph theory on the morphologies of actomyosin networks with multilinkers. *Physical Review E*, 102(6):062420, 2020.

[253] John Cardy. *Finite-size scaling.* Elsevier, 2012.

[254] Vladimir Privman. *Finite size scaling and numerical simulation of statistical systems.* World Scientific, 1990.

[255] Andrew G Clark, Kai Dierkes, and Ewa K Paluch. Monitoring actin cortex thickness in live cells. *Biophysical journal*, 105(3):570–580, 2013.

[256] Alan Pasha Tabatabai, Daniel S Seara, Joseph Tibbs, Vikrant Yadav, Ian Linsmeier, and Michael P Murrell. Detailed balance broken by catch bond kinetics enables mechanical-adaptation in active materials. *Advanced Functional Materials*, page 2006745, 2020.

[257] Alexander Mogilner and George Oster. Cell motility driven by actin polymerization. *Biophysical journal*, 71(6):3030–3045, 1996.

[258] Alex Mogilner. On the edge: modeling protrusion. *Current opinion in cell biology*, 18(1):32–39, 2006.

[259] R TyleráMcLaughlin et al. Collective dynamics of processive cytoskeletal motors. *Soft matter*, 12(1):14–21, 2016.

[260] Toshihiro Toyota, David A Head, Christoph F Schmidt, and Daisuke Mizuno. Non-gaussian athermal fluctuations in active gels. *Soft Matter*, 7(7):3234–3239, 2011.

[261] Beno Gutenberg and Charles Richter. *Seismicity of the earth and associated phenomena.* Princeton University Press, 1949.

[262] Martin van Hecke. Jamming of soft particles: geometry, mechanics, scaling and isostaticity. *Journal of Physics: Condensed Matter*, 22(3):033101, 2009.

[263] Jean-Philippe Bouchaud. Weak ergodicity breaking and aging in disordered systems. *Journal de Physique I*, 2(9):1705–1713, 1992.

[264] Pavel I Zhuravlev and Garegin A Papoian. Molecular noise of capping protein binding induces macroscopic instability in filopodial dynamics. *Proceedings of the National Academy of Sciences*, 106(28):11570–11575, 2009.

[265] Tim Mitchison and Marc Kirschner. Dynamic instability of microtubule growth. *nature*, 312(5991):237–242, 1984.

[266] O Lieleg, Mireille Maria Anna Elisabeth Claessens, Y Luan, and AR Bausch. Transient binding and dissipation in cross-linked actin networks. *Physical review letters*, 101(10):108101, 2008.

[267] Yuriy V Pereverzev, Oleg V Prezhdo, Manu Forero, Evgeni V Sokurenko, and Wendy E Thomas. The two-pathway model for the catch-slip transition in biological adhesion. *Biophysical journal*, 89(3):1446–1454, 2005.

[268] Qin Ni and Garegin A Papoian. Turnover versus treadmilling in actin network assembly and remodeling. *Cytoskeleton*, 2019.

[269] Xiaona Li, Qin Ni, Xiuxiu He, Jun Kong, Soon-Mi Lim, Garegin A Papoian, Jerome P Trzeciakowski, Andreea Trache, and Yi Jiang. Tensile force-induced cytoskeletal remodeling: Mechanics before chemistry. *PLOS Computational Biology*, 16(6):e1007693, 2020.

[270] Carlos Floyd, Garegin A Papoian, and Christopher Jarzynski. Gibbs free energy change of a discrete chemical reaction event. *The Journal of Chemical Physics*, 152(8):084116, 2020.

[271] James Liman, Carlos Bueno, Yossi Eliaz, Nicholas P Schafer, M Neal Waxham, Peter G Wolynes, Herbert Levine, and Margaret S Cheung. The role of the arp2/3 complex in shaping the dynamics and structures of branched actomyosin networks. *Proceedings of the National Academy of Sciences*, 2020.

[272] Shenshen Wang and Peter G Wolynes. Active contractility in actomyosin networks. *Proceedings of the National Academy of Sciences*, 109(17):6446–6451, 2012.

[273] DN Prabhakar Murthy, Min Xie, and Renyan Jiang. *Weibull models*, volume 505. John Wiley & Sons, 2004.

[274] Bruce D Malamud and Donald L Turcotte. Self-affine time series: I. generation and analyses. *Adv. Geophys*, 40:1–90, 1999.

[275] Jon D Pelletier and Donald L Turcotte. Self-affine time series: Ii. applications and models. In *Advances in Geophysics*, volume 40, pages 91–166. Elsevier, 1999.

[276] Donald L Turcotte. *Fractals and chaos in geology and geophysics.* Cambridge university press, 1997.

[277] Annette Witt and Bruce D Malamud. Quantification of long-range persistence in geophysical time series: conventional and benchmark-based improvement techniques. *Surveys in Geophysics*, 34(5):541–651, 2013.

[278] ZC Williams, JD Pelletier, and Thomas Meixner. Self-affine fractal spatial and temporal variability of the san pedro river, southern arizona. *Journal of Geophysical Research: Earth Surface*, 124(6):1540–1558, 2019.

[279] Henrik Jeldtoft Jensen. *Self-organized criticality: emergent complex behavior in physical and biological systems*, volume 10. Cambridge university press, 1998.

[280] Tamar Schlick. *Molecular modeling and simulation: an interdisciplinary guide: an interdisciplinary guide*, volume 21. Springer Science & Business Media, 2010.

[281] Andrew R Leach and Andrew R Leach. *Molecular modelling: principles and applications.* Pearson education, 2001.

[282] EM Huisman and Thomas C Lubensky. Internal stresses, normal modes, and nonaffinity in three-dimensional biopolymer networks. *Physical review letters*, 106(8):088301, 2011.

[283] Minhaeng Cho, Graham R Fleming, Shinji Saito, Iwao Ohmine, and Richard M Stratt. Instantaneous normal mode analysis of liquid water. *The Journal of chemical physics*, 100(9):6672–6683, 1994.

[284] Scott D Bembenek and Brian B Laird. Instantaneous normal modes and the glass transition. *Physical review letters*, 74(6):936, 1995.

[285] David Richard, Geert Kapteijns, Julia A Giannini, M Lisa Manning, and Edan Lerner. Simple and broadly applicable definition of shear transformation zones. *Physical Review Letters*, 126(1):015501, 2021.

[286] Phoebe MR DeVries, Fernanda Viégas, Martin Wattenberg, and Brendan J Meade. Deep learning of aftershock patterns following large earthquakes. *Nature*, 560(7720):632–634, 2018.

[287] Arnaud Mignan and Marco Broccardo. One neuron versus deep learning in aftershock prediction. *Nature*, 574(7776):E1–E3, 2019.

[288] Bernard R Brooks, Dušanka Janežič, and Martin Karplus. Harmonic analysis of large systems. i. methodology. *Journal of computational chemistry*, 16(12):1522–1542, 1995.

[289] Yann N Dauphin, Razvan Pascanu, Caglar Gulcehre, Kyunghyun Cho, Surya Ganguli, and Yoshua Bengio. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In *Advances in neural information processing systems*, pages 2933–2941, 2014.

[290] Mark Newman. *Networks*. Oxford university press, 2018.

[291] Zorana Zeravcic, Wim van Saarloos, and David R Nelson. Localization behavior of vibrational modes in granular packings. *EPL (Europhysics Letters)*, 83(4):44001, 2008.

[292] Mayte Bonilla-Quintana, Florentin Wörgötter, Elisa D'Este, Christian Tetzlaff, and Michael Fauth. Reproducing asymmetrical spine shape fluctuations in a model of actin dynamics predicts self-organized criticality. *Scientific reports*, 11(1):1–17, 2021.

[293] Evelyn F Keller and Lee A Segel. Model for chemotaxis. *Journal of theoretical biology*, 30(2):225–234, 1971.

[294] Guangping Tai, Brian Reid, Lin Cao, and Min Zhao. Electrotaxis and wound healing: experimental methods to study electric fields as a directional signal for cell migration. In *Chemotaxis*, pages 77–97. Springer, 2009.

[295] Alireza Seif, Mohammad Hafezi, and Christopher Jarzynski. Machine learning the thermodynamic arrow of time. *Nature Physics*, 17(1):105–113, 2021.

[296] Laura Tociu, Gregory Rassolov, Etienne Fodor, and Suriyanarayanan Vaikuntanathan. Inferring dissipation from static structure in active matter. *arXiv preprint arXiv:2012.10441*, 2020.

[297] Jordan M Horowitz and Todd R Gingrich. Thermodynamic uncertainty relations constrain non-equilibrium fluctuations. *Nature Physics*, 16(1):15–20, 2020.

[298] Andre C Barato and Udo Seifert. Thermodynamic uncertainty relation for biomolecular processes. *Physical review letters*, 114(15):158101, 2015.

[299] Todd R Gingrich, Jordan M Horowitz, Nikolay Perunov, and Jeremy L England. Dissipation bounds all steady-state current fluctuations. *Physical review letters*, 116(12):120601, 2016.

[300] Jordan M Horowitz and Todd R Gingrich. Proof of the finite-time thermodynamic uncertainty relation for steady-state currents. *Physical Review E*, 96(2):020103, 2017.

[301] Timur Koyuk and Udo Seifert. Thermodynamic uncertainty relation for time-dependent driving. *Physical Review Letters*, 125(26):260604, 2020.

[302] Patrick Pietzonka, Andre C Barato, and Udo Seifert. Universal bound on the efficiency of molecular motors. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(12):124004, 2016.

[303] Gregory Falkovich. *Fluid mechanics: A short course for physicists*. Cambridge University Press, 2011.

[304] Stephen B Pope. Turbulent flows, 2001.

[305] Qiwei Yu, Dongliang Zhang, and Yuhai Tu. Inverse power law scaling of energy dissipation rate in nonequilibrium reaction networks. *Physical Review Letters*, 126(8):080601, 2021.

[306] M Fritzsche, D Li, H Colin-York, VT Chang, E Moeendarbary, JH Felce, E Sezgin, G Charras, E Betzig, and C Eggeling. Self-organizing actin patterns shape membrane architecture but not cell mechanics. *Nature communications*, 8(1):1–14, 2017.

[307] Yee Han Tee, Tom Shemesh, Visalatchi Thiagarajan, Rizal Fajar Hariadi, Karen L Anderson, Christopher Page, Niels Volkmann, Dorit Hanein, Sivaraj Sivaramakrishnan, Michael M Kozlov, et al. Cellular chirality arising from the self-organization of the actin cytoskeleton. *Nature cell biology*, 17(4):445–457, 2015.

[308] Oliver M O'Reilly. *Modeling nonlinear problems in the mechanics of strings and rods*. Springer, 2017.

[309] Mattia Gazzola, LH Dudte, AG McCormick, and L Mahadevan. Forward and inverse problems in the mechanics of soft filaments. *Royal Society open science*, 5(6):171628, 2018.

[310] Olaf Wolkenhauer, Peter Wellstead, Kwang-Hyun Cho, Ramon Grima, and Santiago Schnell. Modelling reaction kinetics inside cells. *Essays in biochemistry*, 45:41–56, 2008.

[311] Carlos Floyd, Aravind Chandresekaran, Haoran Ni, Qin Ni, and Garegin A Papoian. Segmental lennard-jones interactions for semi-flexible polymer networks. *arXiv preprint arXiv:2102.11446*, 2021.

[312] Jorge Nocedal and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[313] Ewa Nowak and Roger S Goody. Kinetics of adenosine 5'-triphosphate and adenosine 5'-diphosphate interaction with g-actin. *Biochemistry*, 27(23):8613–8617, 1988.

[314] A Ott, M Magnasco, A Simon, and A Libchaber. Measurement of the persistence length of polymerized actin using fluorescence microscopy. *Physical Review E*, 48(3):R1642, 1993.

[315] Andrej Vilfan and Thomas Duke. Instabilities in the transient response of muscle. *Biophysical Journal*, 85(2):818–827, 2003.

[316] BA DiDonna and Alex J Levine. Unfolding cross-linkers as rheology regulators in f-actin networks. *Physical Review E*, 75(4):041909, 2007.

[317] Neil Billington, Aibing Wang, Jian Mao, Robert S Adelstein, and James R Sellers. Characterization of three full-length human nonmuscle myosin ii paralogs. *Journal of Biological Chemistry*, 288(46):33398–33410, 2013.

[318] G Nicolis and I Prigogine. I (1977) self-organization in nonequilibrium systems. *From Dissipative structures to Order through Fluctuations, Mir, Moscow, Russia.*

[319] Dilip Kondepudi and Ilya Prigogine. *Modern thermodynamics: from heat engines to dissipative structures*. John Wiley & Sons, 2014.

[320] P Glansdorff and I Prigogine. Sur les propriétés différentielles de la production d'entropie. *Physica*, 20(7-12):773–780, 1954.

[321] François Chollet et al. Keras. `https://keras.io`, 2015.

[322] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[323] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240, 2006.

[324] Takaya Saito and Marc Rehmsmeier. The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PloS one*, 10(3), 2015.