A method for detecting nonequilibrium dynamics in active matter

by

Garrett Watson

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of
Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2018

© Garrett Watson, MMXVIII. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.
A method for detecting nonequilibrium dynamics in active matter

by

Garrett Watson

Submitted to the Department of Physics on May 12, 2018, in partial fulfillment of the requirements for the degree of Bachelor of Science

Abstract

Active force generation is an important class of out-of-equilibrium activity in cells. These forces play a crucial role in vital processes such as tissue folding, cell division and intracellular transport. It is important to determine the extent of such nonequilibrium activity during cellular processes to understand cell function. Here we present a framework for measuring nonequilibrium activity in biological active matter using time reversal asymmetry based on the Kullbeck-Leibler Divergence (KLD), also known as relative entropy. We estimate the KLD from a stationary time series using a $k$-nearest neighbors estimator, comparing the time-forwards process to the time-reversed process. Using time series data of probe particles embedded in the actin cortex, we establish a lower bound for the entropy production of cortical activity. Our results demonstrate a reliable way to measure the breaking of detailed balance in mesoscopic systems.

Thesis Supervisor: Nikta Fakhri
Title: Associate Professor, Department of Physics
Acknowledgments

The author would like to acknowledge the support and guidance of his advisor, Professor Nikta Fakhri, as well as that of postdocs Jordan Horowitz and Todd Gingrich as well as graduate student Tzer Han Tan who collected all of the data.
Contents

1 Introduction 13

2 Stochastic Thermodynamics 15
   2.1 Review of Thermodynamics .................................. 16
   2.2 Advances in Thermodynamics ................................. 20
   2.3 Entropy and Irreversibility .................................. 24
      2.3.1 Jarzynski Relation and Crooks Fluctuation Theorem .... 24
      2.3.2 KPB Theorem ............................................. 26

3 Approximating the Kullback-Leibler Divergence 29
   3.1 k-Nearest Neighbor Estimator ................................ 29
      3.1.1 k-Nearest Neighbors Estimate for Shannon Entropy .... 30
      3.1.2 Convergence to true value of relative entropy .......... 33
   3.2 Applying the KLD to trajectories ............................ 33
      3.2.1 Addressing correlations in the data ..................... 36

4 Calculating Irreversibility in Active Systems 39
   4.1 Experimental System ........................................... 39
      4.1.1 Perturbations ............................................. 41
      4.1.2 Velocity as our Observable ............................... 43
   4.2 Analysis and Results .......................................... 44
      4.2.1 Convergence Plots ....................................... 44
      4.2.2 Choice of k-value ........................................ 45

7
4.2.3 Effect of lag on KLD rate calculation ........................................... 47
4.2.4 Rank ordering of perturbations .................................................. 50

5 Concluding Remarks ................................................................. 51

5.0.1 Finding meaning in the numbers ............................................. 52
5.0.2 Further developments ............................................................ 52
List of Figures

2-1 (a) An ideal gas system at two different states, $U(\lambda_0)$ with large volume and low pressure and $U(\lambda_r)$ with small volume and high pressure. (b) A graphical depiction of the two states (black dots) and two paths through the $(p, V)$ plane connecting them. Along the orange (B) path, as the gas is compressed, the pressure increases linearly; along the blue (A) path the pressure changes in a more complicated way. The work $W$ and heat $Q$ exchanged with the system are different along each path while their sum is the same for both paths. .......................... 18

3-1 Determination of $\rho$ for $k = 1$. The blue diamonds are samples $x_i$ of $p$. In this example $m = 2$. ................................................................. 31

3-2 In addition to the samples $x_i$ of $p$ (blue diamonds), we now have $y_i$ samples of $q$ (orange circles). $\rho$ and $\sigma$ are depicted for $k,l = 1$. Both $\rho_i$ and $\sigma_i$ are defined from the same sample $x_i$ from $p$. .............. 32

3-3 Example time series of a particle moving along a stochastic trajectory. The left side is the time-forwards trajectory while the rightmost plot is the time-reversed version. The full trajectory is split into smaller segments used in the KLD calculation. Example segments of length $m$ are highlighted in orange for the time-forwards case, and their corresponding reverse paths in green for the time-reversed image. ....... 35

4-1 (a) Images of the actin filament network and endogenous probe particles, respectively. (b) A schematic depicting myosin pulling on the actin filaments[4]. ......................................................... 40
4-2 (a) An example MSD curve for our trajectories. (b) The probability distribution of the MSD for different cellular perturbations.

4-3 (a) The displacement of probe particles over time. Since all trajectories start at zero, we do not have a steady state distribution. (b) Particle velocities are linear and representative of a steady state.

4-4 (a) Convergence of 12 oocytes (2 MRLC [red], 4 WT [blue] and 4 NaN3 [yellow]) with number of samples. (b) The same oocytes but is a plot of $d_4 = D_4 - D_3$. Both plots are shown at $k = 7$ with a lag of 80s.

4-5 Two oocytes are shown from each of the following perturbations: MRLC (red), wild type / no perturbation (blue) and NaN3 (yellow). The KLD rate $d_m$ is plotted versus trajectory length $m$ for $k = 1$ and a lag of (a) 5s and (b) 20s.

4-6 Wild type (unperturbed) oocyte at a lag of 20s. $d_4$ estimate calculated at a range of $k$ values and plotted to show convergence.

4-7 (a) Data from a high activity (myosin overexpression) sample with a lag of 20s. (b) How $k$ may act as coarse graining in phase space. In the kNN method we assume that the density is constant within the ball of radius $\rho$. We see that $\rho_2 > \rho_1$ where $\rho_2$ is the distance to the second nearest neighbor and $\rho_1$ is the distance to the closest neighbor. Thus we are approximating a larger volume as constant.

4-8 (a) A representative example $d_m$ varying with $m$ (x-axis) and lag (dt, y-axis). (b) 6 oocytes (2 MRLC, 2 WT and 2 NaN3) are plotted to compare how $m$ and lag (dt) affect $d_m$ values.

4-9 $D_8$ and $d_8$ for six oocytes at $k = 1$. Lag values plotted are dt = 2s, 5s, 6.5s, 8s, 10s, 20s and 60s.

4-10 (a) $d_m$ as a function of effective time $m$*lag. (b) has a $\sqrt{\text{lag}}$ scaling on both vertical and horizontal axis. $k = 1$, $m$ values plotted are 2 through 8. Lag values plotted are dt = 2s, 5s, 6.5s, 8s, 10s, 20s and 40s.
List of Tables
Chapter 1

Introduction

Nonequilibrium activity describes many of the processes in nature. A marker of nonequilibrium activity is irreversibility. Irreversibility means that the process does not (or is very unlikely to) run in reverse. Irreversible dynamics implies broken detailed balance in the system. If we could detect or measure broken detailed balance, then we would be making a measurement on the irreversibility of the system – a quantifying of the arrow of time. Quantifying irreversibility is useful in active systems: it can determine if a system is in thermal equilibrium or being driven by an active mechanism[8].

What we outline in this work is a method to apply this concept to an experimental system, and compare the energetics of different systems by creating a hierarchy of irreversibility. In the first section, we give an overview of classical thermodynamics and transition into the developments of stochastic thermodynamics and fluctuation theorems. Stochastic thermodynamics and fluctuation theorems are the foundation for studying the energetics of nonequilibrium activity at the mesoscale. They are the theories that allow us to extend our understanding of thermodynamics outside of an equilibrium context. This culminates in the work of Kawai, et al. [5] who drew a connection between time irreversibility and entropy production. This development is to use the Kullback-Leibler divergence to distinguish between time-forwards processes and time-reversed processes quantitatively.

Having established the value of the Kullback-Leibler divergence, we then explain
how to calculate this quantity. We do so by using a $k$-nearest neighbors estimator. That is, we calculate the distance from a point in phase space to nearby points to estimate the probability density. While this estimator is well known, it is derived for systems of a random variable. We provide a framework for applying the estimator on a time series trajectory and interpreting those results. The value in using a $k$-nearest neighbors estimator is that it is more robust to rare events than a histogram method and eliminates the need to coarse grain in physical space.

Lastly, we apply our nearest neighbors estimator of the Kullback-Leibler divergence on an in vivo actin cortex. We demonstrate the convergence of our algorithm and how parameter choices impact the estimated divergence. Our results indicate that a rank ordering of systems can be established based on the irreversibility of their driving processes. Additionally, our results suggest that the KL divergence of a process has temporal information about the energetics embedded within.

The power of our methodology is that the nearest neighbors estimator for KL divergence assumes no physical information about the series, taking solely a trajectory as an input. The quantity itself is purely information-theoretic and can work for any time series. The surprising result is that this leads to an understanding of the time irreversibility and entropy production of the system that created the time series. There are many processes where there is no sense of the energy consumption or that are embedded in larger systems where we have an idea of the energy consumption of the larger system but not the process of interest. The methodology outlined here allows us to gain insight and derive a lower bound for the energy consumption for those processes through a non-invasive recording of the process.
Chapter 2

Stochastic Thermodynamics

Newton's laws of motion allow us to predict the trajectory of individual particles very easily and accurately. They even allow us to model how different bodies interact and give us a complete picture of our system. However, these methods are computationally arduous. If we want to describe the motion of gas in a balloon, for example, we would need to follow every gas molecule which would require on the order of $N \sim 10^{23}$ calculations. Because of this, we cannot realistically simulate real world systems on the molecular level with Newton’s laws.

The solution to this dilemma is to instead focus on statistical and probabilistic methods, averaging over the properties of several molecules[2]. There are two toolsets that allow us to take this approach. The first is that of thermodynamics, which focuses on the flow of heat and energy through a system that allow us to make powerful statements with little specific information about the system. The second is statistical mechanics, which connects the microscopic descriptions of each particle to a system’s thermodynamic properties.

Thermodynamics, in a technical sense, is a phenomenological description of equilibrium properties of macroscopic systems [3]. The laws of thermodynamics are derived from empirical observations. It applies to systems whose global properties do not change appreciably over the time of interest, referred to as a system in "equilibrium." These properties are called "state variables" and fully characterize a thermodynamic system.
Statistical mechanics is a probabilistic approach to understanding the equilibrium properties of macroscopic systems, with large numbers of degrees of freedom[3]. Instead of being empirically motivated, as in thermodynamics, these properties are derived from the microscopic description of the system.

Classical thermodynamics, however, is not sufficient for describing the natural world. While it allows us to describe macroscopic systems, it only rigorously applies to those systems in equilibrium, a small fraction of what we encounter daily. Statistical mechanics works best on a system with a large number of degrees of freedom $N$, as fluctuations from the average occur on order $1/\sqrt{N}$. If we wish to apply these tools at small scales, for example to study systems at the mesoscale, these fluctuations become more important to the overall dynamics[8]. So called "fluctuation theorems" have been derived in recent decades that allow us to extend the concept of thermodynamics to the nonequilibrium regime at small scales. The work presented here is method for harnessing these tools in nonequilibrium active matter systems.

2.1 Review of Thermodynamics

Thermodynamics is a set of tools for studying a system through properties that only emerge when there are a large number of degrees of freedom. Such properties, for example states of matter or pressure of a gas, play a large part in understanding the macroscopic world. The toolset of thermodynamics lets us predict and explain natural phenomena of everyday systems. Here we'll demonstrate the primary features of interest using the canonical example of gas trapped in a box Fig. 2-1.

Given a system in thermodynamic equilibrium, let us define internal energy, $U$, to be the sum of all internal contributions to the energy of the system. This includes the kinetic energy of individual molecules in a gas or the potential energy stored in atomic bonds. It is important to note that the internal energy of a system is defined in isolation, thus there is no contribution from an overall kinetic energy of the system or potential due to an external field (e.g. gravity)[2]. In our example of a gas trapped in a box, it is natural to define the internal energy $U$ in terms of measurable properties
such as pressure $p$ which captures the motion of the gas particles and its volume $V$ which conveys the size of our system.

Now consider a transition of the system from an initial internal energy $U_i(p_{low}, V_{high})$ to a different one $U_f(p_{high}, V_{low})$. We induce this transition of the system by moving a piston to compress the gas. This transition is characterized by a parameter $\lambda(t)$ for $t \in [0, \tau]$ which traces out a path in the $(p, V)$ plane that the system (all of the gas molecules) follows as the piston is moved (Fig. 2-1b). To move the piston and cause a transition from one equilibrium state to another, we need to add energy to the gas. Energy can be added through mechanically moving the piston, doing work $W$ on the system or by exchanging heat $Q$ with the gas, letting changes in pressure move the piston without us needing to do work to move it. We are allowed to add work $W$ or $Q$ heat however we like, and different combinations of $W$ and $Q$ trace out different paths in the $(p, V)$ plane (Fig. 2-1b). However, since we always start at $U_i$ and end at $U_f$, state variables that define the system, the total contributions from work $W$ and heat $Q$ must be the same regardless of the path $\lambda(t)$ taken. Thus we arrive at the first law of thermodynamics.

$$\Delta U = Q + W$$  \hspace{1cm} (2.1)

It is important to note that there are many ways to do work $W$ on and add heat $Q$ to a system that yield the same change in internal energy $U$. The values of $W$ and $Q$ are dependent on the path $\lambda(t)$ taken from $U_i \rightarrow U_f$ in Fig. 2-1b. For example, we can compress the gas while at the same time extracting or adding heat to keep the pressure of the gas proportional to its volume (path B). Alternatively, we might extract a lot of heat initially to drop the pressure so moving the piston takes less work, but in moving the piston the pressure of the gas then rises once more. If we are to write the first law in a differential form, we need to convey that $W$ and $Q$ are path dependent so we express it as the following:

$$dU = \delta Q + \delta W$$ \hspace{1cm} (2.2)
Consider a cycle \( C \) from \( U_i \rightarrow U_f \rightarrow U_i \), where we end up at the same state that we start in. It is clear that \( \Delta U = \int_C dU = 0 \). However, it is not true that \( Q = \int_C \delta Q \) is also zero, since we can take a different path backwards. We simply need to add or extract work from the system so that \( W = -Q \neq 0 \) and the first law is still satisfied. However, we can define a state function related to \( Q \) by including a system property called temperature \( T \). Let us call this function entropy, and it is defined as

\[
S = \int_C \frac{\delta Q}{T} \tag{2.3}
\]

Since entropy is path independent, it is a property of our macroscopic system that we can use to compare different states \( U \) and will become an important property in our discussions here.

In the interpretation of the first law thus far, we assume that at every step of the process the system is in equilibrium. This means that at any point along our trajectory \( \lambda(t) \), we can define \( U \) and other state functions that only make sense at
equilibrium. In this way we are free to retrace our steps if necessary. If the process is done quickly, however, then the system won’t have time to equilibrate and our path is ill-defined. The points $U(\lambda)$ along our trajectory $\lambda(t)$ do not exist where the system didn’t settle into equilibrium. Because of this, we call these processes irreversible. We cannot run the process backwards by adding the same amount of heat that the system lost or extracting the work that was put into the system[3]. However, the entropy change $\Delta S$ is the same in both cases, as entropy is a state function. In fact, the entropy change $\Delta S$ is always greater than the amount of heat added irreversibly.

$$\Delta S \geq \frac{Q}{T}$$  \hspace{1cm} (2.4)

where equality occurs if the process is done reversibly. This result is known as the second law of thermodynamics. One result of this is that if a system is driven irreversibly from one state $U_i$ to another $U_f$, then there’s a discrepancy between the heat added to the system and the total change in entropy of the system. In this process we have

$$\Delta S = \frac{Q}{T} + S_{\text{prod}}$$  \hspace{1cm} (2.5)

where $S_{\text{prod}}$ is the entropy produced during the transition $U_i \rightarrow U_f$. It is clear from Eq. 2.4 that $S_{\text{prod}} \geq 0$ for a process where the internal energy increases and no heat has been added to the system. We may interpret each term in this Eq. 2.5 as follows: $\Delta S$ is the change in entropy of the system, $Q/T$ is the entropy we add to the system and $S_{\text{prod}}$ is entropy in the universe generated to account for that difference[8].

In addition to internal energy $U$ and entropy $S$, it will be helpful to define a third state function $F = U - TS$ which is the free energy of the system. Let us go back to the first law Eq. 2.1 to see where this is useful by combining it with the second law Eq. 2.4.
\[ W + Q - \Delta U = 0 \]
\[ W + T\Delta S - \Delta U \geq 0 \]
\[ W - (\Delta U - T\Delta S) \geq 0 \]
\[ W - \Delta F \geq 0 \] (2.6)

Where equality holds for a reversible process, following from Eq 2.4. The free energy difference can be interpreted as the amount of work to bring a system from state \( U_i \) to \( U_f \) in a reversible manner. If the process is done irreversibly, there is more work done on the system than the free energy difference \( \Delta F \) between the states. The excess work \( W \) needed to drive a transition is referred to as work dissipated[8]

\[ W_{\text{diss}} = W - \Delta F \] (2.7)

We see from Eq. 2.6 that \( W_{\text{diss}} \geq 0 \). If our system undergoes a closed cycle, then the free energy difference \( \Delta F \) between the initial and final states is 0 and \( W_{\text{diss}} \) is the difference between the work applied in the first half of the cycle to that extracted in the second half of the cycle. The way we arrived at \( W_{\text{diss}} \) is very similar to how we defined \( S_{\text{prod}} \), and indeed there is a relation between the two that follows directly from their definitions and the first law Eq. 2.1:

\[ W_{\text{diss}} = TS_{\text{prod}} \] (2.8)

### 2.2 Advances in Thermodynamics

Thermodynamics is a discipline constructed to describe processes in macroscopic systems from an energetic perspective, using descriptions of the system such as entropy production and work dissipation[7]. However, thermodynamic principles apply at the microscopic scale as well. The caveat is that there exist statistical deviations
from these variables. Thermodynamic variables are defined as an ensemble average, where each the constituents of the ensemble have properties associated with the variable given by a Gaussian distribution. The deviations from the norm are negligible classically but in small systems have a measurable effect. We are building up to a description of stochastic systems. Stochastic systems are characterized by dynamics that are not entirely deterministic but instead probabilistic elements of randomness.

Think again of our gas molecules in a box. When two molecules collide, there is a transfer of energy between them that will be of order $k_B T$. We can think of this process as creating the differences given by the distribution of molecular energies. In this case, the differences between molecular energies are on the order of $k_B T$. If we consider the full energy in our system, with $N$ molecules in the box, we see that the total energy is of the order $N k_B T$. This implies that the magnitude of deviations in the system are roughly $1/N$. For large systems, $1/N \sim 0$, whereas in small systems of size $n \ll N$, $1/n$ deviations can have an impact on the dynamics of the system.

When we start considering systems away from equilibrium, we would like ways of relating equilibrium properties such as free energy $F$, entropy $S$ or other state functions with nonequilibrium processes. Linear response theory is the first part of this processes. Linear response theory relates the transport properties of materials through equilibrium functions[9]. This applies only to small deviations from equilibrium, however. Exact results to expand the range of validity of thermodynamic statements beyond linear response into far-from-equilibrium regimes consider distribution functions of thermodynamic quantities such as exchanged heat, applied work, or entropy production. These results are particularly relevant for small systems, where fluctuations need to be taken into account.

Some systems of interest where these considerations are necessary include colloidal particles, biopolymers and molecular motors. We will focus on this class of systems for concreteness. All of these are embedded in aqueous solutions, so we also take care to describe the state of the medium in addition to that of the system[9]. One of the possible nonequilibrium processes one can observe in these systems is called a nonequilibrium steady state (NESS). A NESS is driven out of equilibrium by a
time independent force that creates a current in the phase space of the system. It is considered to be at steady state if the probability for the system to be in a given state is constant in time.

The dynamics that we are interested in are equivalently described by the Langevin equation, path integrals and the Fokker-Planck equation[9]. We include in our descriptions a systematic force $F(x, \lambda)$ that is the sum of an internal conservative force $V(x, \lambda)$ and an applied force $f(x, \lambda)$. The Langevin equation takes the form below

$$\dot{x} = \mu F(x, \lambda) + \xi = \mu(-\partial_x V(x, \lambda) + f(x, \lambda)) + \xi$$

(2.9)

where $\xi$ is uncorrelated Gaussian white noise. Langevin dynamics generate trajectories $x(t)$ with diffusion constant $D = T\mu$ and associated action

$$\mathcal{A}([x(\tau), \lambda(\tau)]) \equiv \int_0^t d\tau[(\dot{x} - \mu F)^2/4D + \mu \partial_x F/2]$$

(2.10)

Additionally, it will be useful to write the Fokker-Planck equation which we will solve to find the probability $p(x, \tau)$ of the particle to be at at position $x$ at time $\tau$.

$$\partial_\tau p(x, \tau) = -\partial_x(\mu F(x, \lambda)p(x, \tau) - D\partial_x p(x, \tau))$$

(2.11)

The driving in our system is once again given by parameter $\lambda(t)$ and we allow our particle to feel a potential $V(x, \lambda(t))$ and we may choose to apply an external force $f$. Where $f = 0$ we have equilibrium, and $f \neq 0$ is a process that has some entropy production. In the notation of this section, we use lowercase letters $(w, s, q)$ to represent particle/trajectory properties. One of the purposes of developing a stochastic theory of thermodynamics is to be able to develop the concepts of work $w$ and heat $q$ along a particle trajectory instead of an ensemble of particles at equilibrium. The differential work applied to a particle takes the form below[9]

$$\delta w = (\partial V/\partial \lambda)d\lambda + f dx$$

(2.12)

where the first term is a result of changing the potential by parameter $\lambda(t)$, and
the latter the external force acting on the particle. We can subsequently define the heat lost by the particle to the system as

\[ \delta q = \delta w - dV = Fdx \quad (2.13) \]

We integrate these path-dependent quantities over time as follows:

\[ w[x(\tau)] = \int_0^t [(\partial V/\partial \lambda)\dot{\lambda} + f\dot{x}]d\tau \quad (2.14) \]
\[ q[x(\tau)] = \int_0^t F\dot{x} \quad (2.15) \]

In stochastic systems, state variables become functionals evaluated along a trajectory[9]. We can alternatively describe the heat \( q[x(\tau)] \) in terms of the action defined in Eq. 2.10

\[ q[x(\tau)] = T\frac{\mathcal{A}[x(\tau), \lambda(\tau)]}{\mathcal{A}[x(t-\tau), \lambda(t-\tau)]} = T \log \frac{p[x(\tau), \lambda(\tau)]}{p(\tilde{x}(\tau), \tilde{\lambda}(\tau))} \quad (2.16) \]

\[ p[x(\tau), \lambda(\tau)] = \mathcal{N} e^{-\mathcal{A}[x(\tau), \lambda(\tau)]} \quad (2.17) \]

Where \( p[x(\tau), \lambda(\tau)] \) is the joint probability density of following path \( x(\tau) \) given the process \( \lambda(\tau) \) normalized by a factor \( \mathcal{N} \). This defines heat \( q[x(\tau)] \) as a ratio involving the probabilities to follow path \( x(\tau) \) and the probability of following the time-reversed path \( \tilde{x}(\tau) \equiv x(t-\tau) \) under the reverse process \( \tilde{\lambda}(\tau) \equiv \lambda(t-\tau) \) which hints at the relation between entropy production and time reversal[9].

The last thermodynamic functional to be defined is entropy \( s \), using \( p(x(\tau), \tau) \) from solving the Fokker-Planck equation.

\[ s(\tau) \equiv -\log p(x(\tau), \tau) \quad (2.18) \]

Since the stochastic entropy \( s(\tau) \) depends on the solution to the Fokker-Planck
equation, \( s(\tau) \) depends on the full state of the ensemble, not just the starting point of the trajectory.

This definition of \( s(\tau) \) matches with the ensemble entropy \( S \) upon averaging.

\[
\langle s \rangle = \int \mathcal{D}[x(\tau)] p(x(\tau), \tau) \log p(x(\tau), \tau) = S
\]

(2.19)

### 2.3 Entropy and Irreversibility

#### 2.3.1 Jarzynski Relation and Crooks Fluctuation Theorem

Let us consider a system in an initial equilibrium state \( U(\lambda_0) \), that is driven arbitrarily far-from-equilibrium through a time-dependent control parameter \( \lambda(t) \). The Jarzynski relation is then[9]:

\[
\langle e^{-W/k_B T} \rangle = e^{-\Delta F / k_B T}
\]

(2.20)

where \( \Delta F = F(\lambda_t) - F(\lambda_0) \) is the free energy difference between the initial state and the equilibrium state corresponding to \( \lambda_t \). With this relation, the free energy difference (an equilibrium property of a system) is defined in terms of the average of nonequilibrium, driven processes. This is a powerful implication that equilibrium properties can be determined from nonequilibrium measurements. Explicitly this can be done by taking an average of the work done over multiple realizations of the process and rewriting Eq. 2.20 as

\[
\Delta F = -k_B T \ln \langle e^{W/k_B T} \rangle
\]

(2.21)

Since the exponential is a convex function, Jensen’s inequality can be applied and we arrive at a statement the second law[8] \( \langle W \rangle \geq \Delta F \). We see in the Jarzynski relation an encoding of fundamental thermodynamic laws within the context of nonequilibrium processes.

Closely related to the Jarzynski relation is Crooks fluctuation theorem. In Crooks theorem, we begin as we do in the Jarzynski relation. Consider a system driven out of equilibrium through a control parameter \( \lambda(t) \) as \( t \) is varied from 0 to \( \tau \). Let us
denote $\lambda(t)$ as the forward process, from which we also consider the time-reversed process given by $\tilde{\lambda}(t) = \lambda(\tau - t)$ where the system is driven in the opposite manner as to the forward process. That is to say that $\lambda(0) = \tilde{\lambda}(\tau)$ and $\lambda(\tau) = \tilde{\lambda}(0)$ where the work spent along $\lambda$ and $\tilde{\lambda}$ are equal in magnitude and opposite in sign. Since the driving requires work to be done or extracted from the system, we can define a probability density function for work applied or spent by the system in the forward process $p(W)$ and the reverse process $\tilde{p}(-W)$. We will consider the initial state for both the forward $\lambda$ and reverse $\tilde{\lambda}$ process to be at equilibrium. Explicitly, $U(\lambda(0))$ and $U(\tilde{\lambda}(0))$ is well defined but is not required for $U(\lambda(\tau))$ or $U(\tilde{\lambda}(\tau))$. In this case there is a well defined free energy difference $\Delta F$ between the states. Crooks showed that we can then relate the two probability density functions $p(W)$ and $\tilde{p}(-W)$ by the following relation

$$\frac{p(W)}{\tilde{p}(-W)} = e^{(W - \Delta F)/k_BT} \quad (2.22)$$

The Crooks fluctuation theorem is initially derived for systems that satisfy microscopically reversible dynamics. Given a trajectory $x(t)$, we say that the system dynamics are microscopically reversible if the probability to travel that path $P[x(t)]$ and the reverse path $\tilde{P}[\tilde{x}(t)]$ satisfy the following condition

$$\frac{P[x(t)]}{\tilde{P}[\tilde{x}(t)]} = e^{-Q[x(t)]/k_BT} \quad (2.23)$$

Both the Jarzynski relation and Crooks theorem are very insightful to nonequilibrium thermodynamics by relating nonequilibrium measurements to equilibrium properties. These equalities are often a stepping stone for developments in stochastic thermodynamics and as such their fundamental arguments are implicit in much of the discussion here. The idea behind describing the Jarzynski relation (Eq. 2.20) and Crooks fluctuation theorem (Eq. 2.22) is to set up the KPB theorem, described in the next section, which underpins the methodology we use here.
2.3.2 KPB Theorem

Kawai, Parrando and Van den Broeck (2007) related the average work dissipated in a system $\langle W_{\text{diss}} \rangle$ with the Kullback-Leibler divergence (KLD) between the forward and time reversed phase space probability densities. The KLD is a measure of the distinguishability between two probability distributions. For two probability distributions $p$ and $q$ for a random variable $X$, the KLD between $p(x)$ and $q(x)$ where $x$ is a given realization of $X$ is

$$D[p(x)||q(x)] = \int dx p(x) \log \frac{p(x)}{q(x)}$$ (2.24)

Starting with Crooks fluctuation theorem Eq. 2.22, we seek to calculate $\langle W_{\text{diss}} \rangle$ as defined in Eq. 2.7.

$$w - \Delta F = k_B T \log \frac{\bar{p}(w)}{\bar{p}(-w)}$$

$$\langle w - \Delta F \rangle = k_B T \int dx p(w) \log \frac{p(w)}{\bar{p}(-w)}$$

$$\langle w_{\text{diss}} \rangle = k_B T D[p(w)||\bar{p}(-w)]$$ (2.25)

Where we note that when taking the average $\langle w_{\text{diss}} \rangle$ we are calculating $D[p(w)||p(-w)]$. Since work $w$ is path dependent, we can consider the above equation to be a path integral over all possible trajectories $x(t)$. In this case, $p(w)$ is replaced with Eq. 2.26 and $\bar{p}(-w)$ is similarly replaced with Eq. 2.27

$$\mathcal{P}[x(t)] = \int D[x(t)] p(w) \delta(w[x(t)] - w)$$ (2.26)

$$\tilde{\mathcal{P}}[\tilde{x}(t)] = \int D[\tilde{x}(t)] \tilde{p}(-w) \delta(w[\tilde{x}(t)] + w)$$ (2.27)

This argument holds assuming all the information of $p(w)$ is contained in $\mathcal{P}[x(t)]$. Even in the case where we have coarse grained the trajectory, so our path probability $\mathcal{P}[x(t)]$ contains only partial information of $p(w)$ this formulation yields useful results.
Consider the case where our path is described by two independent variables, so we have $\mathcal{P}[x(t), y(t)]$ and the time reversed $\tilde{\mathcal{P}}[\tilde{x}(t), \tilde{y}(t)]$. Then we have the following result

$$D[\mathcal{P}[x, y]||\tilde{\mathcal{P}}[\tilde{x}, \tilde{y}]] = \int \int \mathcal{D}[x] \mathcal{D}[y] \mathcal{P}[x, y] \log \frac{\mathcal{P}[x, y]}{\tilde{\mathcal{P}}[\tilde{x}, \tilde{y}]}$$

$$= \int \mathcal{D}[x] \mathcal{P}[x] \log \frac{\mathcal{P}[x]}{\tilde{\mathcal{P}}[\tilde{x}]} + \int \mathcal{D}[y] \mathcal{P}[y|x] \int \mathcal{D}[y] \mathcal{P}[y] \log \frac{\mathcal{P}[y|x]}{\tilde{\mathcal{P}}[\tilde{y}]}$$

$$= D[\mathcal{P}[x]||\tilde{\mathcal{P}}[\tilde{x}]] + D[\mathcal{P}[y|x]||\tilde{\mathcal{P}}[\tilde{y}]] \quad (2.28)$$

$$\geq D[\mathcal{P}[x]||\tilde{\mathcal{P}}[\tilde{x}]] \quad (2.29)$$

The inequality in the last line follows from the KLD always being greater than or equal to 0. We find equality in the event that the relation between $x$ and $y$ is some one-to-one function, where the second integral on the second line vanishes. Eq. 2.28 is known as the chain rule for the KLD, which is an essential property for being able to apply this to a stochastic process and extract a lower bound to the entropy production. Therefore, we arrive at the following conclusion which we refer to as the KPB Theorem

$$\langle W_{diss} \rangle \geq k_B T D[\mathcal{P}[x(t)]||\mathcal{P}[x(\tau - t)]] \quad (2.30)$$

$$\langle S_{prod} \rangle \geq k_B D[\mathcal{P}[x(t)]||\mathcal{P}[x(\tau - t)]] \quad (2.31)$$
Chapter 3

Approximating the Kullback-Leibler Divergence

3.1 k-Nearest Neighbor Estimator

Suppose $P$ and $Q$ are probability distributions on some metric space. One way of comparing these two distributions is by defining a distance between them, known as a divergence. The distance between two probability distribution should tell how distinguishable $P$ and $Q$ are from each other. Let $P$ be absolutely continuous with $Q$, then the divergence between $P$ and $Q$ is called their relative entropy, also known as their Kullback-Leibler divergence (KLD)[10]

$$D[P||Q] = \int dP \log \frac{P}{Q}$$

(3.1)

If the densities of $P$ and $Q$ exist, denoted by $p(x)$ and $q(x)$ respectively, where $x$ is in $m$-dimensional space $\mathbb{R}^m$ then we can write the divergence as

$$D[p||q] = \int p(x) \log \frac{p(x)}{q(x)} d^m x$$

(3.2)

where the convention $0 \log 0/0 \equiv 0$ and $p \log p/0 \equiv \infty$ is chosen. Our goal is to derive an unbiased estimator for $D[p||q]$ given samples drawn from $p$ and $q$. To do so, we want to find density estimators $\hat{p}$ and $\hat{q}$ for $p$ and $q$ that will give a consistent
estimate for $D[p||q]$. We first note that the KLD is the difference between the Shannon entropy of $p$, $H(p)$, and cross entropy of $p$ and $q$, $H(p,q)$

$$D[p||q] = \int p(x) \log p(x) - \int p(x) \log q(x)$$

$$= -H(p) + H(q,p)$$

We now follow a heuristic derivation of a k-Nearest Neighbors (kNN) estimator for Shannon entropy $H(p)$ from Krasov, Stogbauer and Grassberge (2004)[6]. We then extend their argument to derive a kNN estimator for the cross-entropy $H(q,p)$. Combining these we arrive at a k-Nearest Neighbors estimator for Kullback-Leibler divergence (kNN-KLD).

### 3.1.1 k-Nearest Neighbors Estimate for Shannon Entropy

Consider $(x_1, x_2, \ldots, x_N)$ to be $N$ independent and identically distributed (i.i.d) samples drawn from from probability density $p$. We can see from Eq. 3.3 that Shannon entropy can be thought of as an average of $\log p(x)$ (with the inclusion of a minus sign). If we can obtain an unbiased estimator of $\log p(x)$, then we would have an estimator for $H(p)$. Using hats to denote the estimator, we are looking for the following equation

$$\hat{H}(X) = -\frac{1}{N} \sum_{i=1}^{N} \log \hat{p}(x)$$

In order to obtain the estimate $\hat{\log p(x)}$, we consider the probability distribution $M_k(\rho)$ for the distance between $x_i$ and its $k$th nearest neighbor (see Fig. 3-2). The probability $M_k(\rho) d\rho$ is the likelihood that there is one point within distance $d \in [\rho, \rho + d\rho]$ from $x_i$, that there are $k - 1$ points at $d < \rho$ and $N - k - 1$ points where the distance from $x_i$ is $d > \rho$[6]. Let us denote $\mu_i(\rho)$ as the mass of a ball with radius $\rho$ centered at $x_i$. Then $\mu_i(\rho) = \int_{||\xi - x_i|| < \rho} d\xi p(\xi)$ and we can write $M_k(\rho) d\rho$ as the following
Figure 3-1: Determination of \( \rho \) for \( k = 1 \). The blue diamonds are samples \( x_i \) of \( p \). In this example \( m = 2 \).

\[
M_k(\rho) d\rho = \frac{(N-1)!}{1!(k-1)!(N-k-1)!} \frac{d\mu_i(\rho)}{d\rho} d\rho \times \mu_i^{k-1} \times (1 - \mu_i)^{N-k-1} \tag{3.5}
\]

\[
M_k(\rho) = k \binom{N-1}{k} \frac{d\mu_i(\rho)}{d\rho} \mu_i^{k-1}(1 - \mu_i)^{N-k-1} \tag{3.6}
\]

We now wish to find \( \log p(x) \) using \( \mu_i(\rho) \). Let us assume that \( p(x) \) is constant within the ball of radius \( \rho \). This implies \( \mu_i(\rho) \approx c_m \rho^m p(x_i) \), where \( x_i \in \mathbb{R}^m \) and \( c_m \) is the volume of the \( m \)-dim unit ball[6]. We then have that the expectation \( \mathbb{E} \) of \( \log p(x) \) is

\[
\mathbb{E}[\log p(x)] \approx -m \mathbb{E}[\log \rho] - \log c_m + \mathbb{E}[\log \mu_i(\rho)] \tag{3.7}
\]

To calculate \( \mathbb{E}[\log \mu_i(\rho)] \), we note the following

\[
\mathbb{E}[\log \mu_i(\rho)] = \int_0^\infty d\rho M_k(\rho) \log \mu_i(\rho)
= k \binom{N-1}{k} \int_0^1 dp p^{k-1}(1 - p)^{N-k-1} \log p
= \psi(k) - \psi(N) \tag{3.8}
\]
Figure 3-2: In addition to the samples $x_i$ of $p$ (blue diamonds), we now have $y_i$ samples of $q$ (orange circles). $\rho$ and $\sigma$ are depicted for $k, l = 1$. Both $\rho_i$ and $\sigma_i$ are defined from the same sample $x_i$ from $p$.

where in the last line, $\psi(z)$ is the digamma function. If we combine Eq. 3.8 with Eq. 3.7 and 3.4, we arrive at the an $k$-Nearest Neighbors estimator for Shannon entropy

$$
\hat{H}(p) = -\psi(k) + \psi(N) + \log c_m + \frac{m}{N} \sum_{i=1}^{N} \log \rho_i
$$

(3.9)

Here $\rho_i$ is the distance from $x_i$ to it’s $k^{th}$ nearest neighbor, and we are summing over all samples. We also assume $k$ is kept fixed for every point in the calculation. It is possible to pick a new $k_i$ for each point $x_i$, in which case $\psi(k)$ would be replaced with $\sum_{i=1}^{N} \psi(k_i)/N$.

We will now extend this argument for the cross entropy $H(p, q)$. For the cross entropy, we want an estimator for $\log q(x)$ with respect to the probability density $p(x)$. To do this, we consider $M_i(\sigma)$ which is the distance $\sigma$ from $x_i$ to its $l^{th}$ nearest neighbor in $q$ (the $l^{th}$ closest point $y_i$ in Fig. 3-2). The derivation then follows exactly as in the case for Shannon entropy giving us

$$
H(p, q) = -\psi(l) + \psi(N) + \log c_m + \frac{m}{N} \sum_{i=1}^{N} \log \sigma_i
$$

(3.10)

If we plug in Eq. 3.9 and Eq. 3.10 into Eq. 3.3 we arrive at the following estimator
for KL divergence, where we allow $k$ and $l$ to vary per sample and denote $\rho_{k,i}$ and $\sigma_{l,i}$ for the distances from $x_i$ to the $k^{th}$ neighbor in $p$ and $l^{th}$ neighbor in $q$, respectively.

$$\hat{D}_{KL}[p||q] = \frac{m}{N} \sum_{i=1}^{N} \left[ \log \frac{\sigma_{l,i}}{\rho_{k,i}} + \psi(k_i) - \psi(l_i) \right]$$ \hspace{1cm} (3.11)

When we apply this algorithm, we fix $k = l$ for all samples, which then reduces the calculation to

$$\hat{D}_{KL}[p||q] = \frac{m}{N} \sum_{i=1}^{N} \log \frac{\sigma_i}{\rho_i}$$ \hspace{1cm} (3.12)

### 3.1.2 Convergence to true value of relative entropy

Wang, Kukami and Verdu[10] showed that the $k$-NN estimator for the KLD has vanishing bias and variance as sample size increases as long as the true KLD $D[p||q]$ is bounded.

The choice of $k$ is in principle independent to the convergence; however, Kraskov, et al. suggests that $k$ has an effect on the size of systematic and statistical errors, with smaller $k$ being susceptible to larger statistical errors but reduced systematic errors[6]. In particular, small $k$ has much smaller systematic errors than other estimators for the KLD. Larger choices of $k$ reduces statistical errors at the cost of higher systematic error.

### 3.2 Applying the KLD to trajectories

As we saw in the previous section, Kawai, Parrando and van den Broeck demonstrated the following relationship

$$\langle S_{\text{prod}} \rangle \geq k_B D[\mathcal{P}[x(t)]||\tilde{\mathcal{P}}[\tilde{x}(t)]]$$ \hspace{1cm} (3.13)

Where $\mathcal{P}[x(t)]$ and $\tilde{\mathcal{P}}[\tilde{x}(t)]$ are the path probabilities of a particle undergoing stochastic dynamics to take the trajectory $x(t)$ and the time-reversed trajectory $\tilde{x}(t)$, respectively. $\langle S_{\text{prod}} \rangle$ is the average entropy produced by the process driving the par-
article. If we take a system in a nonequilibrium steady state (NESS), a property of interest is the entropy production rate. The rate of entropy production provides insight into how energy consumption affects these dynamics and organization. To determine the entropy production rate from Eq. 3.13, we consider the limit of an infinitely long trajectory, $x(t)$ where $t$ varies from 0 to $\tau$[8]

$$
\langle \dot{S}_{\text{prod}} \rangle = k_B \lim_{\tau \to \infty} \frac{1}{\tau} D[\mathcal{P}[x(t)]||\tilde{\mathcal{P}}[\tilde{x}(t)]]
$$

In practice (simulation or experiment), we don’t observe the full trajectory of the system. Our observables $x$ are sampled at a finite frequency, producing a discrete time series $x^m_1 \equiv (x_1, x_2, \ldots, x_m)$. When we discretize our trajectory, the relationship still holds if all the information regarding the entropy production of the system is encoded in the discretized time series. If the time series does not completely describe the entropy production, then we are losing information about the system in the coarse graining process. In this case, the KLD calculation outputs a strictly lower value than with the full information of the system. Because of this, the KLD is a lower bound to the total entropy production, where equality occurs when the trajectory $x^m_1$ contains the full information of the entropy production of the system.

Given a stochastic process $X$, consider $x^m_1$ to be a realization of the process observed at fixed time intervals $dt$. Then $p(x^m_1)$ is the probability of observing the ordered sequence $(x_1, x_2, \ldots, x_m)$ in a realization of $X$. To measure the irreversibility, we want to compare this probability $p(x^m_1)$ to that of observing the time-reversed sequence $x^1_m \equiv (x_m, x_{m-1}, \ldots, x_1), p(x^1_m)$. When we consider trajectories in this manner, the KLD $D[p(x^m_1)||p(x^1_m)]$ becomes a measure of the entropy production as a function of the number of symbols $m$. We can relate the entropy production per symbol to the entropy production rate using the time interval $dt$ of sampling to convert from a per symbol rate to a per time rate. Using the symbol notation, we define the $m$-th order KLD for a process $X$ as the KLD calculated for $m$ realizations of $X$ of the probability of observing the forward sequence $p(x^m_1)$ against the probability of observing the reverse sequence $p(x^1_m)[8]$. 

34
As an example, let's consider Fig. 3-3. On the left is a realization of process $X$, $X_\tau$. The right plot is the time-reversed sequence $X_\tau$. Our goal is to determine the likelihood of taking the path on the left versus the path on the right in a given realization of $X$. To do this from a single instance is not possible; however, we can split the full path into shorter segments $x_1^m$ and $x_1^1$ (highlighted in orange and green, respectively) where $m < \tau$. It is then possible to ask what the probability of observing $x_1^m$ is in the trajectory $X_\tau$ by either counting how frequently the sequence occurs or using a density estimator as outlined in §3.1. We can then take the time-reversed segment and ask the same question, what is the likelihood of finding the sequence $x_1^m$ (highlighted in green) in the full trajectory $X_\tau$.

The above description contextualizes the use of $D_m$ as a proxy for the KLD of the full trajectory. The longer sequences $x_1^m$ that we consider in the example, the more information we have and the larger our estimate of $D_m$ becomes. However, we expect that the increase of $D_m$ eventually saturates as $m$ gets larger due to $x_1^m$ saturating in terms of its information content about the entropy production of the process. This establishes the hierarchy of $D_m$ from which we can calculate an estimator for the system entropy production and extrapolate to the true entropy production of the system.
As stated earlier, we are interested in determining the entropy production rate for the process $X$, so we define a KLD rate as the growth rate of $D_m$ with the number of symbols

$$\frac{\langle S_{\text{prod}} \rangle}{k_B} \geq d^X = \lim_{m \to \infty} \frac{D_m^X}{m}$$  \hspace{1cm} (3.16)

It will be useful to also define an $m$-th order KLD rate $d_m = D_m - D_{m-1}$. Taking $\lim_{m \to \infty} d_m$ converges to $d^X$ faster than Eq. 3.16. We will consider $d_m$ and its limit to be our estimator for the entropy production.

### 3.2.1 Addressing correlations in the data

The $k$-Nearest Neighbors estimator assumes that the observables $x_i$ are independently distributed. Given that we are interested in particle trajectories, and our observables are consecutive measurements of a particle undergoing stochastic dynamics, the observables $x_i$ are necessarily correlated. This causes an uneven distribution of $p(x)$ in phase space. The dimensions of the observable $x$ will not be orthogonal, given that $x_i$ is related to $x_{i-1}$ by a stochastic equation. The uneven distribution makes the method unreliable as the phase space is anisotropic, thus it is more likely for the nearest neighbor to appear in a specific direction. Not having an isotropic phase space makes it difficult for approximating the time-reversed probability distributions and will cause a bias towards larger entropy production estimates.

To counteract this, we apply a whitening filter to the data which maps the covariance matrix of our observables to the identity matrix. Whitening transformations are linear and therefore reversible, so no information about the system is lost in the process (meaning we don’t worsen our lower bound). This rounds out the phase space and creates a relatively uniform distribution.

The variable-$k$ algorithm is one way to address this issue as well. However, when applying the whitening filter in addition to variable-$k$ yields worse results than fixed-$k$ on the whitened data.

For the input into our $k$-NN algorithm, we also ask that the data be stationary
(otherwise there are more correlations). This is implicit in the NESS assumption, however care needs to be taken to insure this validity. For example, the observable \( x \) could express drift along the full length of the time series, which would bias the calculation towards higher estimates. To account for this, it is necessary to remove the trend to find a realistic probability distribution for the time-reverse path probabilities.
Chapter 4

Calculating Irreversibility in Active Systems

4.1 Experimental System

Active matter systems are ones where the elements making up the system consume energy and form complex structures, thus displaying nonequilibrium dynamics. Active matter systems can be found in many areas of biology and physics. They are an opportune place to study stochastic thermodynamics. We apply our method to a living biological system, which is necessarily out of equilibrium. As a model system we choose the actomyosin cortex of a starfish oocyte (Fig. 4-1a). The actomyosin cortex is made up of an actin filament network and myosin motors. The cortex is responsible for determining cell shape and structure. The actin filament network is a mesh of polymers that are continuously being polymerized and depolymerized and pulled on by myosin motor proteins. Myosin motors consume ATP, acting as a force dipole, pulling actin filaments and driving the system out of equilibrium (Fig. 4-1b). ATP, or adenosine triphosphate, is the form of energy storage most dominant in cells. When myosin consume ATP, they are removing a phosphate group (turning it into ADP, adenosine diphosphate). The removal of the phosphate group releases energy that the myosin motor can use. The concentration of ATP (high energy state) versus ADP (low energy state) creates a free energy difference $\Delta F$ which drive the
nonequilibrium activity. The fluctuations that occur as myosin motors pull on the actin network creates the entropy that we aim measure. We record the fluctuations by endogenous probes that are embedded in the actin network. We then record each particles' trajectory which eventually becomes the input into our nearest neighbors KLD estimator.

Given that for our system the cortex (to first order) can be considered as isotropic and in a steady state over the duration of measurement, we assume ergodicity. With the assumption of ergodicity we increase the statistics of a sample by considering each probe as part of an ensemble of trajectories, instead of working with a single probe trajectory. This allows us to better approximate the trajectory probabilities.

We use a starfish oocyte as a model cell because before maturation it stays in
a nonequilibrium steady state (NESS) for long time scales (several hours). This is optimal because it allows us to prepare the cells in different environmental conditions as an experimental test of our algorithm. Ultimately, we seek to understand how cells utilize activity to organize and maintain their functional structures.

4.1.1 Perturbations

One of the advantages of using an actomyosin gel is that we have control over the amount of mechanical fluctuations within the network. Within a batch of oocytes, we injected a portion of them with one of the perturbations, while leaving some unperturbed as a control (wild type cells, or WT).

We can drastically reduce the presence of ATP within the cell using the drug sodium azide (Na$_3$N). This removes the energy source for myosin motors, resulting in significantly less fluctuations and a halting of nonequilibrium activity in the cell.

Another drug we can add to our system is blebbistatin, which inhibits myosin activity. Cells with the blebbistatin perturbation have their myosin rendered inoperable. As in the sodium azide case, we expect this perturbation to result in much lower to no nonequilibrium activity.

A perturbation we can add to increase myosin activity is nocodazole. Nocodazole depolymerizes microtubules in the network. As a result, a protein is released that creates a signal which enhances myosin activity.

Another perturbation we can apply to our cell is over expression of the myosin motor protein, known as MRLC. By injecting a molecule (RNA) encoding the protein, we can utilize the internal cell machinery to produce more myosin motors. A higher concentration of myosin motors implies a lot more activity within the cortex as it increases the number of force dipoles within the system.

We can independently check that the added drugs have affected the cell through a mean square displacement (MSD) calculation (Fig. 4-2). MSD is calculated by
MSD $\equiv \langle (x - x_0)^2 \rangle = \frac{1}{N} \sum_{n=1}^{N} (x_n(t) - x_n(0))^2$  \hspace{1cm} (4.1)

Where we've taken the average over $N$ trajectories of the same oocyte. The MSD generally follows a power law as a function of $t$, so the curve is traditionally plotted in log-log space, where the slope of the curve determines the exponent, which gives a sense of how far particles travel in the medium. The MSD curve is influenced by both the material particles of the medium the particle is in and temporal characteristics of the forces driving the particle[1]. The MSD curve for our probes in the actomyosin cortex, displayed in Fig. 4-2a, have two distinct linear regions. The reason for this is due to myosin activity occurring on the time scale of 5-8s. Before this timescale, the correlation of a particle's position is low as it is undergoing thermal motion. At timescales longer than myosin activity, the force of the myosin causes the particle to move further on average, causing the MSD exponent to grow.

We collect this information for each particle trajectory in a given sample, which is then averaged to get a rough distribution of the MSD exponent for each perturbation (Fig. 4-2b). The differences that are seen between the MSD distributions at long timescales is reflective of the perturbations imposed. For the lower activity pertur-
bations, we see the MSD exponent for timescales longer than myosin (α in the Fig. 4-2) activity is shifted left, whereas the high activity perturbations are shifted right and have a much tighter distribution.

### 4.1.2 Velocity as our Observable

When working with our experimental data, our tracking software records the position of each probe particle on each frame. This builds an ensemble of trajectories at a steady state distribution for calculating the KLD. However, the software imposes a grid on the sample so to create our ensemble we look at displacement instead of absolute position. When we do this, we no longer start with a steady state distribution since we force every particle to start at the origin.

To get around this, we look at particle velocities (Fig. 4-3). The velocities have the property of being stationary, which is necessary for an accurate relative entropy estimator using the k-nearest neighbors method. Physically, this also is intuitive, because we are using the particles as probes for driving forces in the actin cortex and the particle velocity is the dominant term in the overdamped Langevin equation (Eq. 2.9). Velocity is a linear function of position, so we lose no information of our system and maintain a similar bound to what we would measure from a stationary distribution of positions.
This has the advantage of allowing us to probe timescales in our system. Because velocity is calculated as the difference between position measurements, we can control how far apart in time the position measurements are that we take the difference between. Explicitly, \( x_i \) and \( x_{i+1} \) are recorded 0.5s apart, but we can also take the difference of \( x_i \) and \( x_{i+10} \) which is 5s apart. We use the term lag to mean the amount of time between position measurements used in the velocity calculation

\[
v_i = x_{i+\text{lag}} - x_i
\]  

(4.2)

When we increase the lag, we are time averaging the velocity over a longer period. By time averaging over a long period, signal to noise is reduced and the entropy per symbol in the velocity string is larger. However, at these longer time averages, we are coarse graining out more of the dynamics and worsen our lower bound of entropy production. Using lower lags allows us to get a better estimate of the true entropy production rate of the system, but it there is larger error due to higher levels of noise.

4.2 Analysis and Results

4.2.1 Convergence Plots

For our discussion here to be relevant, it is necessary to know that we have converged values for our calculations and that each sample is similarly converged. Fig. 4-4 shows the convergence of the 4th-order KLD value \( D_4 \) and KLD rate \( d_4 \) for 9 oocytes. These values are plotted against "number of samples" where each sample is one sequence \( x_1^m \). Since we are looking at \( m = 4 \), then the number of samples are the number of quadruples taken from the ensemble of trajectories in velocity. Due to experimental limitations, we cannot ensure a consistent number of samples recorded from each run, which is reflected in the different horizontal axis values.

It's clear, unfortunately, that the current data does not appear converged. Lower estimates of dissipation are converging slightly better than higher estimates. The convergence for \( d_4 \) is slightly worse than that for \( D_4 \) because \( d_4 \) depends on both \( D_3 \)
and $D_4$ to be converged, before itself has converged. This is due to how we calculate $d_m$ as the difference between $D_m$ and $D_{m-1}$. We can also compare how the different $m$ values impact convergence. We see that lower $m$ values converge better than higher values.

In addition to ensuring the values of $d_m$ are individually converged, we want to check that our estimate for entropy production rate $\lim_{m \to \infty} d_m$ is converges. This convergence is heavily dependent on how much we time average the in velocity calculation (Fig. 4-5). For estimates where velocity was time averaged over a long period (where we expect a worse lower bound) the values of $d_m$ converge or start to converge around $m = 6$ (Fig. 4-5a). When time averaging over shorter timescales (where we expect a tighter lower bound), we fail to see convergence up to $m = 8$ (Fig. 4-5b).

### 4.2.2 Choice of k-value

The value of $k$ in the $k$-Nearest Neighbors method is, principally, a statistical choice. As discussed in the previous section, the $k$-value can be chosen based on data size or varied in a variable-$k$ method. Processing the data on multiple $k$ values, we notice that the choice of $k$ does not affect the convergence of $d_m$ with number of samples. Across a range of $k$ values, as we increase the number of samples used in
Figure 4-5: Two oocytes are shown from each of the following perturbations: MRLC (red), wild type / no perturbation (blue) and NaN₃ (yellow). The KLD rate $d_m$ is plotted versus trajectory length $m$ for $k = 1$ and a lag of (a) 5s and (b) 20s.

In the calculation, all samples converge similarly, but to different values (Fig. 4-6). It's possible that if we increase our number of samples by a couple of orders of magnitude that they converge to the same value, doing so slowly. However, in the samples with large estimated dissipation rates, at current sampling levels they do not appear to converge together (even for low estimated dissipation rates, it appears there will be a finite difference between $k$ values).

We observe that higher values of $k$ reduce our estimate of entropy production rates (Fig 4-7a). This is notable because past work determined that the converged value was $1.5 \pm 0.5$.

Figure 4-6: Wild type (unperturbed) oocyte at a lag of 20s. $d_4$ estimate calculated at a range of $k$ values and plotted to show convergence.
4.2.3 Effect of lag on KLD rate calculation

In Fig 4-8, we show a 3D surface plot to visualize how KLD rates $d_m$ vary with both $m$ and lag values increasing. Fig. 4-8a is an example plot, where we can see as both $m$ and lag increases, our measurements of $d_m$ grow, i.e. for fixed $m$ increasing lag can yield larger $d_m$ values and for fixed lags larger $m$ values give larger $d_m$, as we expect. Fig. 4-8b shows a set of six oocytes in the same plot format as Fig. 4-8a. The angle has been adjusted to better see if the ordering is affected by these parameters. Visually, we can tell that changing lag and $m$ don’t disrupt the ordering, with the differences between each sample growing with increasing lag and $m$-value.
Figure 4-8: (a) A representative example $d_m$ varying with $m$ (x-axis) and lag (dt, y-axis). (b) 6 oocytes (2 MRLC, 2 WT and 2 NaN₃) are plotted to compare how $m$ and lag (dt) affect $d_m$ values.

Clearly there is some dependence on the lag we choose for $d_m$, so it is a valid question to ask whether we can account for this fact and convert to an entropy production per time (which would be consistent across all lags) from the entropy production per symbol $m$ plotted in Fig. 4-8. One method worth considering is to try and convert the per symbol rate to a per time rate by multiplying by the symbol frequency, $1/\text{lag}$. This will normalize the $d_m$ values with respect to the amount of time averaging in their calculation. Doing so results in Fig. 4-9, where we’ve selected $m = 8$ for demonstration. First we notice that increasing the lag from 2s to 10s creates a sharp rise in the estimated KLD value and KLD rate. However, when we go beyond that, the estimates start to drop. The decrease implies that our estimate of $d_m$ decreases as we coarse grain in time. It’s likely that at large lags, we are averaging over several cycles in phase space, masking the dynamics. The sharp rise, on the other hand, is harder to explain. It’s possible that there’s a peak timescale to observe dissipation. At low time scales, we do not fully capture the dynamics of our system with a sequence of 8 symbols, but doubling our lag we do capture those dynamics. A priori we know from our experimental system that myosin activity is in the range of 5-8s, so we might hope this is indicative of thermal activity below 5-8s and after surpassing that timescale we see the full dynamics where our estimate subsequently decays after larger coarse graining.

The next attempt to capture the time scaling in the lag discussion is to consider
the effective time being probed in the KLD. If we think about the \( m \)-th order KLD, \( D_m \), we are looking at the entropy of a trajectory that occurs over a time of \( t = 0 \) to \( t = \tau_{eff} = m \times \text{lag} \). This comparison is plotted in Fig. 4-10a. We see that increasing lag moves the \( d_m \) curves to the right. Higher lag values appear to converge to larger \( d_m \) rates. This is intuitive as each symbol has more total entropy production. One might expect the convergences of each curve to align, but this is not the case. We see that the largest few lags are starting or have converged at these time scales, whereas the lowest few lags have not yet done so. Importantly, rank ordering has still been maintained in these samples. One outcome of this type of plot is that if we look closely at the differences between various choices for the lag, we see that for a given effective time, the curves corresponding to lower lag values estimate higher dissipation. This is in line with our intuition that higher lag values have lower information content and are therefore a worse lower bound. The advantage of using higher lag values, though, is that we can probe higher effective time scales. This also suggests that if we consider \( d_m \) to be a function of both \( m \) and lag, higher \( m \) values contribute more to the calculation than higher lag values if we fix \( \tau_{eff} \).

Another way we might account for the observed differences is by finding a scale factor that collapses these curves together. Empirically, we can achieve this by plotting \( d_m / \sqrt{\text{lag}} \) versus \( m \sqrt{\text{lag}} \), shown in Fig. 4-10b. While this correction brings the curves together, it is not exactly precise. Under this scaling, \( d_m \) is larger for larger
Figure 4-10: (a) $d_m$ as a function of effective time $m\cdot\text{lag}$. (b) has a $\sqrt{\text{lag}}$ scaling on both vertical and horizontal axis. $k = 1$, $m$ values plotted are 2 through 8. Lag values plotted are $dt = 2s, 5s, 6.5s, 8s, 10s, 20s$ and 40s.

lag values instead of larger $m$ values as in the previous case. However, the values that each curve is converging to may be larger for the lower $m$ values, as the only curve that looks converged is the largest lag = 40s. In light of these difficulties, it is unclear if there is any physical information in the quantity $d_m$ except in the limit of $m \to \infty$ from the theory.

4.2.4 Rank ordering of perturbations

Across all choices of $k$ and lags, at high enough $m$ values we see a clear separation from the three perturbations presented here. The two oocytes with myosin overexpression (MRLC, in red) consistently are the highest. The two ATP depleted (NaN3, yellow)) oocytes have estimated values near zero, lower than the other samples. The untreated oocytes (blue) are estimated to have an entropy production rate in between the other two treatments. Importantly, the difference in $d_m$ between two samples of the same perturbation are much smaller than the differences between samples of a different perturbation (in most parameter regimes). This suggests that we have constructed a method that qualitatively represents the rank ordering of entropy production rates among similar systems.
Chapter 5

Concluding Remarks

In this thesis we have outlined a method for determining the entropy production rate of a process by quantifying its time irreversibility. In particular, we have studied the relationship between dissipation and irreversibility within the context of stochastic thermodynamics for systems that are in a nonequilibrium steady state (NESS).

After reviewing classical thermodynamics, we explored recent developments in extending the classical theory to the nonequilibrium regime for systems where fluctuations from the mean impact the dynamics. These developments led to the results of Kawai, et al[5] connecting the work dissipated from a process to the Kullback-Leibler divergence (KLD) between the probability to observe a trajectory of the process and the probability to observe its time-reversed trajectory. We then described a process for dividing a trajectory that is a realization of a stochastic process into subsequences and using the space created by those subsequences for estimating the KLD. The KLD can be estimated in this space using a $k$-nearest neighbors approach as described by Wang, et al[10].

We demonstrated this algorithm in an in vivo actin cortex of a starfish oocyte. Perturbing the oocyte with drugs that affect the nonequilibrium driving in the cell, we were able to obtain a rank-ordering of the irreversibility in the oocytes corresponding to the understood effects of the drugs. By calculating the KLD using particle velocity we observed a time dependence in the value and convergence of the KLD rate, which lower bounds entropy production. Our method provides a framework for
comparing the dissipation or time irreversibility of a process across similar processes in a noninvasive manner.

5.0.1 Finding meaning in the numbers

The next steps of this work is to draw conclusions from the quantitative results. We currently have entropy production rates in units of $k_B$ per symbol, but there's little sense of error and relation to the true entropy production rate of the process. Our values for $d_m$ with sampling are not currently converged, so more experiments are needed to achieve high enough sampling for a converged value. The value of our entropy production rate estimator, $d_m$ for large $m$, is only converged at large lag values, so higher $m$ values are needed to be probed at low lags to get a better picture of the entropy production rate and time scaling in the system.

To get an understanding of the error we will use statistical bootstrapping to determine an error bound for the values we obtain. This will determine if the estimated values of the entropy production rate are statistically significant. This will also help to verify that the lower bound is monotonic with respect to irreversibility, so that the qualitative picture is accurate. Using a general Langevin model of our system, we can also simulate trajectories of known dissipation rates and compare with the results of our method to get an understanding of how tight our lower bound is that we calculate. If we could experimentally determine the entropy production rate of our system, then comparing our estimate with the true value would be very helpful in understanding the information content of the trajectories and accuracy of the calculation.

5.0.2 Further developments

Our method discussed in this thesis is a black box with respect to the dynamics, due to the information theoretic nature of our estimator. As such, future projects would be to apply this framework to different active matter systems. The first part of this proposal would be to observe qualitative differences between different environmental perturbations of the same process. Additionally, one could probe this method
by calculating the KLD through different constructed phase spaces in addition to the particle velocity we use in this work. This would speak to the generality of this method. The second part of the proposal would be to compare the results between unrelated processes. Assuming the estimator is truly monotonic with respect to entropy production, and that the tightness of the lower bound is independent of the process generating the trajectory, then this method allows us to compare the dissipation of unrelated processes. It would also be beneficial to apply this method to nonequilibrium systems that are not in steady state. If this is shown to be useful in non-NESS systems, it would allow new experimental insights into stochastic thermodynamics.

By having a noninvasive measurement of dissipation in living systems, this provides us the opportunity to study the responses of cells under different environments under a thermodynamic perspective. This insight will allow us to study the efficiency of energy use and provide a better understanding of the effects of different perturbations within the system.
Bibliography


