Modeling and Analysis of Single-Molecule Experiments

by

James B. Witkoskie

B.S., Univ. of Delaware (2000)

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the
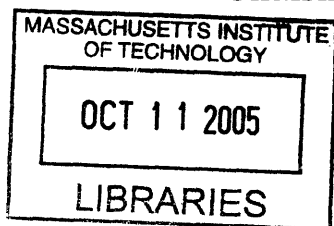
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Author ....
Department of Chemistry
August 3, 2005

Certified by........ ....
Jianshu Cao
Associate Professor of Chemistry
Thesis Supervisor

Accepted by ....
Robert W. Field
Chairman, Department Committee on Graduate Students
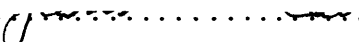
This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor Keith A. Nelson..
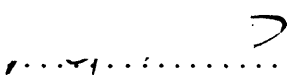
Chairman, Thesis Committee

Professor of Chemistry

Professor Jianshu Cao ....

Thesis Supervisor

Associate Professor of Chemistry

Professor Robert W. Field ....

Professor of Chemistry

2

Modeling and Analysis of Single-Molecule Experiments

by

James B. Witkoskie

Submitted to the Department of Chemistry
on August 3, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Single molecule experiments offer a unique window into the molecular world. This window allows us to distinguish the behaviors of individual molecules from the behavior of bulk by observing rare events and heterogeneity in the dynamics. This thesis discusses both models for single molecule experiments, including the stretching of DNA in hydrodynamic flows and the diffusion of tracer particles in heterogeneous environments, and methods to analyze single molecule data to allow determination of properties and models for single molecule experiments. These methods of analysis are based on combining information theory and Bayesian methods with physical insight and are applied to several experimental situations.

Thesis Supervisor: Jianshu Cao
Title: Associate Professor of Chemistry

Acknowledgments

Upon the completion of my PhD study in the Department of Chemistry, I am indebted to a number of people. First of all, I would like to express my sincerest thanks to my advisor, Prof. Jianshu Cao, who made the last five years educational and exciting. His insight was invaluable in the completion of this work. It has been a great experience to work with him.

I am truly grateful to Prof. Robert Silbey for helpful suggestions and useful discussions and guidance. Prof. Keith Nelson is greatly acknowledged for valuable comments while chairing the thesis committee. I also want to thank Prof. Robert Field for useful discussions and being part of my thesis committee, and Prof. Irwin Oppenheim for his insights.

During my graduate study at MIT, I have benefitted from fruitful discussions with Shilong Yang, Serhan Altunata, Steve Presse, Xiang Xia, Vas Lubchencko, Chiao-lun Cheng, Lorin Gutman, Yuan-Chung Cheng, and Jianlan Wu. It is my great pleasure to include collaborations with Shilong Yang as chapters of this thesis.

Last, but not the least, I would like to express my sincere gratitude to my wife, Kimberly Witkoskie, for her love and support over these years. She is the most important part of my life. I also thank my parents, James and Elaine Witkoskie, for their love and support. This thesis is dedicated to them.

Contents

3 Testing the properties of single molecules blinking sequences: detailed balance violations, renewal properties, and trajectory classifications   **99**

7 Continuous Time Random Walks: aging correlations and coupling to kinetics 251

List of Figures

15

17

18

24

List of Tables

Chapter 1

Introduction

Single molecule experiments have been around for almost 30 years. The first single molecule experiments involved ion channels [30, 211]. These experiments demonstrated discretized ion current levels when an electric potential is applied across membranes permeated with ion channels [30, 211]. The discretized levels were associated with the opening and closing of individual ion channels. Twenty years later, Orrit and Moerner were able to perform single molecule optical spectroscopy on single fluorophores at cryogenic temperatures [19]. Their results gave insight into spectral diffusion in glasses and the existence of two level systems [19]. Single molecule techniques were quickly extended to room temperature studies of more complex biological systems such as proteins and DNA [240, 246, 161]. Today, single molecule experiments are inundating the literature with over 700 papers published in 2004 alone [219].

The interest in single molecule experiments stems from the large amount of information that is accessible by single molecule experiments that is not available through bulk experiments. This information includes observations of rare events and spatial

and temporal heterogeneity. These phenomena are important in understanding the physics of the system. For example, the transition state of a reaction is often a rare event that dominates the kinetics of a reaction [94, 237]. The population of a transition state is so small that its signature is drowned out in bulk measurements and the transition state's existence must be inferred from observations of the reactant and product populations. However, single molecule experiments allow one to trace the evolution of a molecule from the reactant through the transition state to the product. This tracking is unique to single molecule experiments and is the largest advantage of single molecule experiments.

As another example of the advantages of single molecule experiments, consider an ensemble of molecules embedded in a matrix. In a bulk experiment, the measurements are averaged over the entire volume of the experiment. If molecules located in different parts of the matrix are subject to slightly different environments, one expects them to have slightly different attributes. However, the ensemble measurement only captures the average behavior [60]. We are unable to determine if the measurement is the result of the averaging of molecules with different attributes or if every molecule has the same attribute. This distinction can be determined by looking at each molecule separately. Using single molecule techniques, we can achieve this level of detail. Similarly, if the molecules are evolving as the result of a stationary process, the bulk experiment will only be sensitive to the average of this process. By contrast, single molecule techniques allow one to observe the different time evolution of each molecule. We should note that in neutron scattering techniques, one is able to observe time evolution on short time scales, but the single molecule techniques can be used

on much larger time scales because fewer particle are detectable and one does not lose track of the particles over longer periods of time [117].

All of these advantages suggest that single molecule techniques are superior to ensemble measurements, but there are some inherent difficulties with single molecule techniques. The techniques can only be applied to systems with certain attributes. For example, fluorescence experiments can only be successfully applied to systems that contain a good dye molecule. As most substances are not fluorescent, this greatly limits the application of fluorescent single molecule techniques [76]. This difficulty is a hard constraint on the types of systems that can be studied. Another difficulty that will be one focus of this thesis is the fact that the trajectories that we observe in single molecule experiments are a stochastic realization of an underlying mechanism of the system. From a collection of realizations, we need to infer the mechanism. For a bulk experiment, this is generally not a problem because the noise is small, due to the large number of molecules, and the averaged properties of interest are measured directly. If one could perform an infinite number of experiments on single molecules, one would have the same confidence in the attributes that can only be measured by single molecule experiments, but typical single molecule experiments take minutes to hours to perform. As a result of the time constraints, the amount of data is fairly limited and the properties measured have a fairly large amount of stochasticity, which is different than noise that can be removed by better instrumentation. The stochasticity requires one to use much more robust statistical methods to analyze data than is usually used within the physical sciences.

1.1 Overview

This thesis will examine several different types of single molecule experiments through modeling and data analysis. Chapter 2 discusses the use of more traditional experimental analysis and shows the difficulties of applying these methods to single molecule experiments. These difficulties mainly stem from the shot noise intrinsic to single molecule experiments and the time consuming nature of single molecule experiments that make the data insufficient for many of these analyses. Chapter 2 also demonstrates that Bayesian methods can compensate for the shot noise and data insufficiency to allow rigorous comparisons of data to models.

The problem with complete Bayesian analysis is its computationally intensive nature and the need to propose underlying models. These models quickly grow in complexity and full Bayesian analysis becomes cumbersome. In Chapter 3 we examine some properties of single molecule systems that can be determined without examining the entire single molecule data sequences or determining the exact underlying model. These properties include detailed balance violations, presence of correlations in single molecule *events* [35], experimental condition dependencies, non-ergodicity and the emergence of classes of single molecule behaviors, and identification of states. After developing a theoretical basis for the detailed balance violation signatures, simple Bayesian statistical methods are used to identify the detailed balance and these other signatures of single molecule systems.

The Bayesian methods are adapted to analyze a single protein photon counting experiment [240] in Chapter 4. Some simple one dimensional signatures are used

to construct viable models for the fluctuations in the photon lifetimes before a full Bayesian analysis is performed to test these models. The one dimensional analysis is non-parametric so one avoids imposing functional forms and artificial physics on the data. The analysis paints a simple physical picture based on diffusion in a harmonic well for the observed protein motions, which is validated by a more rigorous analysis.

The Bayesian approach is ideal for complex inhomogeneous systems like proteins, where the number of different types of interactions is large and first principles models are not possible. For more homogeneous systems, such as DNA, an approach based on first principles modeling is more appropriate. The preference for first principles modeling is strengthened when the number of different experiments testing the model is larger than the number of parameters so that the model is not simply a curve fitting procedure. To demonstrate first principles modeling, we examine the ability of the worm-like chain model of DNA [81] in both stretching and hydrodynamic flow experiments in Chapter 5. The stretching experiment is used to parameterize the model and the three different flow experiments–constant plug flow, elongational flow, shear flow–are used to test the model. The agreement between the four experiments for a model with a single adjustable parameter, the persistence length, is outstanding.

Although modeling single molecule behaviors is important in data analysis, single molecule experiments are often used as probes of complex environments. One of the most popular applications of single molecule methods to the examination of complex environments is single particle tracking in heterogeneous environments. In Chapters 6 and 7, we examine two models for single particle diffusion in random environments. These models are passive and the tracer particle does not influence the trapping

process. Instead the tracer particle is slaved to environmental fluctuations. The first model examines diffusion in a random potential (Chapter 6), while the second model examines trapping processes in diffusion (Chapter 7). The first model always maintains diffusivity, while the second model can show anomalous diffusion with a modification to the long time behavior.

Chapter 2

Single molecule kinetics:

theoretical analysis of indicators

and numerical Bayesian analysis

2.1 Introduction

Over the last few years, many scientists used single molecule experiments to reveal the nature of dynamic systems. The first elegant single molecule experiments by Moerner and Orrit explored low temperature glasses, where the spectral diffusion of a single chromophore probes the environment. Recent experiments extend the single molecule technique to probe complex biological systems at room temperature. The systems studied with single molecule techniques have become much more complex, which makes interpretation of data more difficult [148, 149]. Experiments by Chu, by Xie, and by other groups reveal the mechanisms of chemi-

cal reactions in biomolecular systems and the associated time scales of these reactions [138, 235, 240, 75, 110, 157, 94, 246, 33, 221, 237, 236]. Other experiments demonstrate single molecule spectroscopy's ability to distinguish between heterogeneous and homogeneous relaxation in glassy systems [106, 244, 60]. All of these analyses require the determination of characteristic times and pertinent configurations or states from the frequency and count of the photons emitted during the experiment. The switching between configurations and the photon statistics are stochastic processes that create large uncertainties in the data retrieved from experiments on these systems. Considering these uncertainties, analysis of these experiments requires the use of robust statistical methods.

The stochastic fluctuations in single molecule systems stimulated interest in the statistical mechanics community [216, 159, 11, 87, 21, 92, 241, 234, 35, 242, 243, 29, 14]. These single molecule experiments give exciting insight into microscopic systems including the role of system bath-interactions and fluctuations. For bulk experiments one is generally only able to measure "intensity" correlations in the system, but these single molecule experiments allowed theorists to proposed several new indicators of various dynamics in these systems, including the event-averaged quantities, "two-event echo", and number density, which we discuss in recent papers [35, 242, 243]. A more recent proposal suggests the use of a generating function to examine blinking sequences [29]. Most of these analyses depend on large amounts of data to remove uncertainties and accurately measure indicators. Since the typical experiment is on the order of seconds to minutes, these large data requirements may be experimentally impossible to obtain.

The chapter is an elaboration of previous work and a demonstration of the role of recently proposed indicators in the framework of single molecule kinetics [35, 242, 243, 29, 87, 21, 38, 16]. The results in this chapter are important in order for us to properly assess how to combine indicators to reveal the physics of the system. In this chapter we introduce a standard single molecule model using Poisson kinetics in section 2.2. We simulate data with a specific kinetic scheme in order to give a numerical example of the application of these indicators to single molecule experiments. Through both analytical work and applications to the simulated data, we demonstrate the information content of each indicator and discuss both strengths and possible difficulties related to these indicators. We show that the indicators contain similar information about connections between relaxation times in complementary forms, and this information may be easier to extract from one indicator or another. Understanding the relationships between the signatures of various indicators is important in unambiguously determining possible kinetic schemes.

Under common conditions the indicators can theoretically contain all the available information of the system, but extraction of this information may not be numerically feasible. We also discuss that non-uniqueness of the under-lying kinetics prevent determination of the exact kinetic scheme, but the degeneracy of the different schemes comes from linear transformations. In Chap. 2.7 we introduce a computational approach based on Bayesian statistics to analyze data from single molecule experiments, which can overcome the deficiencies in the indicators. This chapter concentrates on combining the insight from indicators with numerical methods to approach single molecule problems [224].

After discussing the standard four state Poisson model in Sec. 2.2, including the important role of initial conditions, we will examine the three common indicators, "intensity" correlation, "event" correlation and characteristic function, in Secs. 2.3, 2.4, and 2.5, respectively. Our analysis will include discussions of the information about relaxation times and connectivity contained in each indicator. For each indicator, we will also discuss situations where a limited number of moments of each indicator can theoretically contain all available information, but this information does not uniquely determine the kinetic schemes. The chapter also examine numerical examples of each indicator, which demonstrates several difficulties with the use of indicators.

2.2   System of interest

Our analysis will primarily focus on the blinking model that appears extensively in the literature [35, 242, 243, 29, 11, 21, 92, 87, 241, 234, 38, 16]. The model views the single molecule experiment as a stochastic process that switches between two different sets of states. A number of states, labeled "bright," emit photons in a laser field, but we cannot distinguish these states from each other. The other states, labeled "dark," do not emit photons and also cannot be distinguished from each other. The resulting signal resembles the telegraph signal demonstrated in Fig. 2-1, where we are able to see the molecule switch from a "bright" state to a "dark" state and vice versa, but we do not know which "bright" or "dark" state the system is in.

This simple model does not consider uncertainty in the state of the system, "bright" or "dark", and also neglects restrictions on temporal resolution. One can
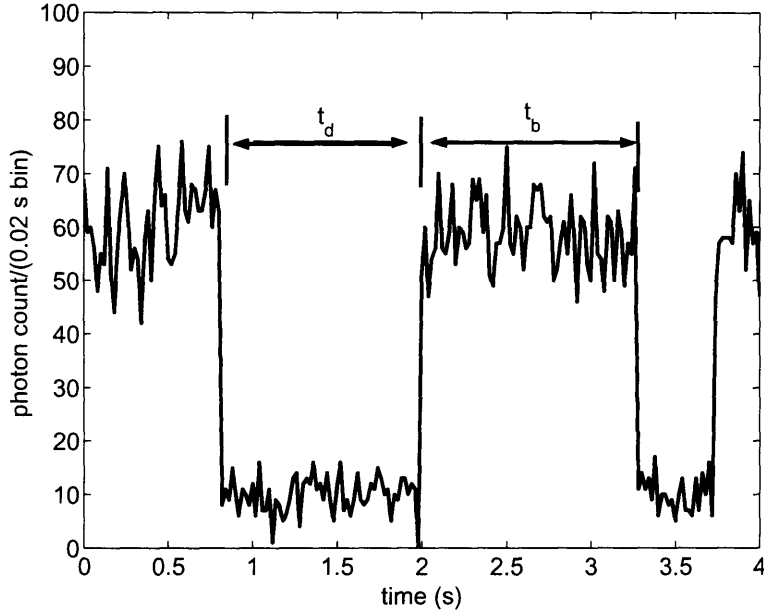
Figure 2-1: The telegraph signal with 20% Poisson shot noise for the standard single molecule experiment that we examine in this chapter. The molecule is either considered "bright" with high intensity of photon emission, or "dark" with low photon intensity. The "event" correlation function measures the duration of "bright" events and "dark" events, labeled t_b and t_d in the figure.

compensate for these simplifying assumptions, but for many systems the error introduced by these assumptions is minimal. The photon shot noise is instrumental in limiting both our ability to determine the state and the temporal resolution. In many of these experiments, the photon emission rate for the "bright" state is much larger than the "dark" state background photon emission rate. Our temporal resolution is limited by our ability to collect enough photons in a bin to determine the state of the system. If both the "bright" state photon emission is much larger than the blinking rates and the difference between the "bright" and "dark" state photon emission rates is also much larger than the blinking rates, the binning time can be chosen so that there is little ambiguity in the state or times of transitions. The unambiguous sharp transitions are apparent in the simulation in Fig. 2-1, where we show a single molecule

trajectory where the "bright" and "dark" state lifetimes are 1s (or other appropriate unit), the "bright" photon emission rate is 3000 photon/s and the "dark" emission rate is 500 photon/s. The bin size is 20 ms and the number of photons in each bin is given by a Poisson distribution. These relative rates are reasonable for many single molecule experiments. As can be seen from this trace, the state can be unambiguously assigned and the time resolution is adequate.

This simple kinetic model successfully explains correlations in the length of time spent in "bright" states in the experiments by Xie [138]. Mathematicians studied similar models related to ion channels [12, 93]. Our discussion primarily addresses systems that follow Poisson kinetics,

$$\dot{\rho} = -\mathbf{K}\rho, \qquad (2.2.1)$$

which is a type of Hidden Markov Chain (HMC). A HMC is a model where the kinetics of state transitions only depend on the current state, but we cannot directly observe the states. We need to infer the states from data, which do not specify the states uniquely or may only have a probabilistic dependence on the current state. Although most of our discussion will address discrete state Poisson kinetics, we can include the continuous limit to get a diffusion type of equation, like the fluctuating-bottle neck model whose equation is given in Eq. 2.5.23 (also see [35, 242, 243, 29, 11, 21, 20, 92, 87, 241, 234, 38, 16]).

The discrete state HMC can be used to interpret many experiments that fit multi-exponential distributions. These models explain dynamic heterogeneity that results in

long lived memory effects as we discussed in previous work [35, 242, 243]. Determination of this heterogeneity or the time scales of other underlying mechanisms in single molecule experiments requires detailed analysis of data acquired from experiments.

A Simple Model

As a demonstration of the philosophical approach to single molecule problems, we study a simple four state model, whose kinetic scheme is outlined in Fig. 2-2(a). The model has four states and the inter-conversion between the states is governed by Poisson kinetics. Two states are "bright" with labels "b1" and "b2," and two states are "dark" with labels "d1" and "d2." The equation for the probability density is given by the simple kinetic equation,

$$
\begin{bmatrix} \dot{\rho}_{b1} \\ \dot{\rho}_{b2} \\ \dot{\rho}_{d1} \\ \dot{\rho}_{d2} \end{bmatrix} = \begin{bmatrix} -(\gamma_{b1} + k_{b1}) & \gamma_{b2} & k_{d1} & 0 \\ \gamma_{b1} & -(\gamma_{b2} + k_{b2}) & 0 & k_{d2} \\ k_{b1} & 0 & -(\gamma_{d1} + k_{d1}) & \gamma_{d2} \\ 0 & k_{b2} & \gamma_{d1} & -(\gamma_{d2} + k_{d2}) \end{bmatrix} \begin{bmatrix} \rho_{b1} \\ \rho_{b2} \\ \rho_{d1} \\ \rho_{d2} \end{bmatrix}.
$$

$$(2.2.2)$$

For numerical calculations, $k_{b1} = 0.75s^{-1}$, $k_{d1} = 0.50s^{-1}$, $k_{b2} = 0.33s^{-1}$, $k_{d2} = 0.22s^{-1}$, and $\gamma_{b1} = \gamma_{b2} = \gamma_{d1} = \gamma_{d2} = 0.1s^{-1}$. Note that $b1$ is not connected to $d2$ and $b2$ is not connected to $d1$. We break the links between these states to avoid difficulties associated with non-uniqueness discussed in section 2.4. Except for certain ranges of parameters, the eight parameter model can reproduce the waiting time distribution for the four state model with more parameters due to the non-uniqueness [12]. The

41

numerical values of these kinetic constants are chosen to make the waiting time distributions not mono-exponential, which we determine from the indicators presented in Secs. 2.3, 2.4, and 2.5. However, the modulation rate between the states does not cause apparent time separation in any of the indicators.
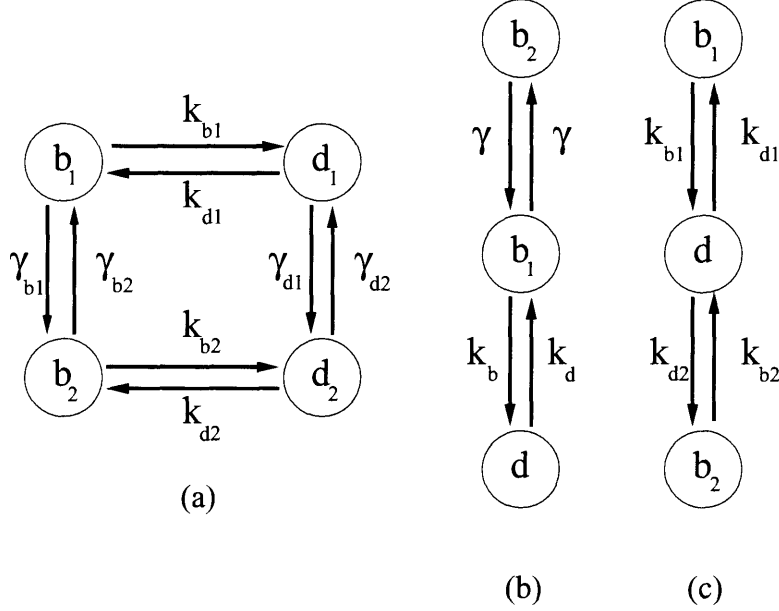
Figure 2-2: Possible kinetic schemes. (a) A diagrammatic depiction of the kinetic scheme that we analyze. (b) and (c) show two indistinguishable schemes. For example, the probability distribution for a sequence produced from the model in (b) with parameters $k_{b1} = 3s^{-1}$, $k_{d1} = 1s^{-1}$ and $\gamma = 2s^{-1}$ is indistinguishable from the model in (c) with $k_{b1} = 6s^{-1}$, $k_{d1} = \frac{2}{5}s^{-1}$, $k_{b2} = 1s^{-1}$, $k_{d2} = \frac{3}{5}s^{-1}$. These two sets of parameters correspond to cutting different connections in the model in (a), which one would expect to behave qualitatively differently, but this intuition is false.

Since many experiments fit the waiting time distributions to a bi-exponential, this four state model is a reasonable minimal model for many systems. For the particular constants we simulate, detailed balance holds, i.e. there is a vector ρ_{eq} such that $\dot{\rho}_{eq} = 0$ and $k_{ij}(\rho_{eq})_j = k_{ji}(\rho_{eq})_i$, but we do not need detailed balance or even Poisson kinetics to apply these approaches. In fact, detailed balance violations are generally difficult to determine [214]. The cholesterol oxidase experiment by Xie and the RNA hairpin experiment by Chu [138, 94, 246] monitor macromolecules that require a substrate, which the reaction depletes from the environment. The replenishment of the substrate is a transport process, which is not in equilibrium although the system reaches a steady state. When treated appropriately, the lack of detailed balance will

not affect any indicators discussed in this chapter or the numerical routine discussed in Sec. 2.7.

From the model system with the specified parameters, we generate a sequence of "bright" and "dark" states for 25 molecules with a duration of $300s$, which allows about 150 observed turn-over events per molecule and 7500 pieces of data [224]. These data sets are much smaller than the data sets collected in the experiments by Xie [138].

Matrix notation for Poisson kinetics

To discuss these HMC more generally we introduce a simple matrix notation for the kinetics of the system. We write the kinetic matrix as three contributions

$$
\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d = \begin{bmatrix} \mathbf{\Gamma}^{(++)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}^{(--)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{K}^{(+-)} \\ \mathbf{K}^{(-+)} & \mathbf{0} \end{bmatrix}
$$
$$
+ \begin{bmatrix} \mathbf{K}^{(++)} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^{(--)} \end{bmatrix}. \tag{2.2.3}
$$

We denote the "bright" state with a $+$ sign and the "dark" state with a $-$ sign. $\mathbf{K}^{(\pm\mp)}$ corresponds to transitions from a "dark" state to a "bright" state, or vice versa. $\mathbf{K}_d^{(\pm\pm)}$ is the decay caused by $\mathbf{K}^{(\pm\mp)}$. If probability is preserved then $(\mathbf{K}_d^{\pm\pm})_{jk} = -\sum_i \mathbf{K}_{ij}^{\mp\pm}\delta_{jk}$. The case of diagonal $\mathbf{K}^{(+-)}$ and $\mathbf{K}^{(-+)}$ corresponds to the modulated reaction models discussed extensively [215, 216, 159, 11, 87, 20, 21, 92, 241, 234, 35, 242, 243, 29]. The $\mathbf{\Gamma}^{(\pm\pm)}$ matrices correspond to unseen transitions between two

44

"bright" states or two "dark" states.

For the four state model, whose kinetic equation is outlined in Eq. 2.2.2, these matrices are

$$\mathbf{K}_d^{(++)} = -\mathbf{K}^{(-+)} = \begin{bmatrix} -k_{b1} & 0 \\ 0 & -k_{b2} \end{bmatrix},$$

$$\mathbf{K}_d^{(--)} = -\mathbf{K}^{(+-)} = \begin{bmatrix} -k_{d1} & 0 \\ 0 & -k_{d2} \end{bmatrix} \tag{2.2.4}$$

and

$$\mathbf{\Gamma}^{(++)} = \begin{bmatrix} -\gamma_{b1} & \gamma_{b2} \\ \gamma_{b1} & -\gamma_{b2} \end{bmatrix}, \qquad \mathbf{\Gamma}^{(--)} = \begin{bmatrix} -\gamma_{d1} & \gamma_{d2} \\ \gamma_{d1} & -\gamma_{d2} \end{bmatrix} \tag{2.2.5}$$

The waiting time distribution for the first visible transition into the "dark" state denoted by i given that we started in the "bright" state j at $t = 0$ has the Laplace transform $Q_{ij}^{(-+)}(s) = (\mathbf{K}^{(-+)} \left[\mathbf{1}s - \mathbf{K}^{(++)} - \mathbf{\Gamma}^{(++)} \right]^{-1})_{ij}$. Note that our notation is transposed relative to standard probability notation. A similar expression exists for the "dark" to "bright" transition, $Q_{ij}^{(+-)}(s) = (\mathbf{K}^{(+-)} \left[\mathbf{1}s - \mathbf{\Gamma}^{(--)} - \mathbf{K}^{(--)} \right]^{-1})_{ij}$. Below we will discuss a characteristic function, which requires us to define the combined matrix,

$$\mathbf{Q} = \begin{bmatrix} \mathbf{0} & \mathbf{Q}^{(+-)} \\ \mathbf{Q}^{(-+)} & \mathbf{0} \end{bmatrix} \tag{2.2.6}$$

From the matrices $\boldsymbol{\Gamma}$ and $\mathbf{K}$ or $\mathbf{Q}$, all observables of the system can be determined, but we want to determine the properties of the matrix from the observables, which is a non-trivial inversion problem.

Equilibrium Distribution

For various indicators the initial condition plays a pivotal role in determining the strength and clarity of various signatures. Since the system has Poisson kinetics, it has a well defined steady state distribution, which for convenience we will call the equilibrium distribution, ρ_{eq}, even if detailed balance does not hold. Generally, ρ_{eq} will not be the initial conditions of our calculations since we select certain configurations to start our measurement. For the "intensity" correlation discussed below, we start monitoring the system at an arbitrary time, but we only consider configurations that are in a specific state, "bright" or "dark" at that time. For this scenario, the initial condition is trivially, $\delta^{(\pm)}\rho_{eq}/|\delta^{(\pm)}\rho_{eq}|$, where

$$\delta^{(+)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \delta^{(-)} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{2.2.7}$$

and $|\delta^{(\pm)}\rho_{eq}| = \sum_i (\delta^{(\pm)}\rho_{eq})_i$. For the "event" correlation function, we start the observation times after an observable transition. If we start from an observed transition from the "dark" state to the "bright" state and average over all observed transitions, the initial condition is $\rho_{ic} = \mathbf{K}^{(+-)}\delta^{(-)}\rho_{eq}/|\mathbf{K}^{(+-)}\delta^{(-)}\rho_{eq}|$, which is the stationary flux introduced previously [35, 242, 243]. A similar expression exists for starting mea-

surements on a "bright" to "dark" transition. Generally, this initial condition is not as strongly influenced by long time correlations as the "intensity" correlation since long "bright" or "dark" periods do not have as large of a contribution to the initial condition as they do for the "intensity" correlation initial condition [212]. It is important to note that we must average over all observed transitions. This stationary flux initial condition is adequate if the duration of a "bright" or "dark" period in the molecular trajectory is shorter than the experimental measurement times so that many measurements can be made on the same time sequence. If this is not the case, the initial condition must be modified, accordingly.

2.3 Intensity Correlation

As discussed in section 2.2, the need to determine the time scales of the underlying conformational dynamics of these experiments motivated many authors, including our previous work, to propose various indicators. As a result, the recent literature contains extensive references to these binary "bright-dark" systems [35, 242, 243, 87, 21, 20]. The popular indicators measure correlations in the state of the system or analyze blinking between the "bright" and "dark" states. Below we will discuss three major approaches to interpreting single molecule experiments with binary–"bright-dark"–informations, which we refer to as "intensity" correlations, "event" correlations (see Sec.2.4) and characteristic functions (see Sec.2.5) [35, 242, 243, 29, 87, 21, 20].

The "intensity" correlation approach measures the probability of being in the "bright" or "dark" state at multiple times [35, 242, 243]. This correlation function

47

requires the existence of two sets of distinguishable states. Unlike some quantities discussed below, this correlation function cannot be applied to single photon statistics since single photons do not have a well defined duration. The time scales of the correlation functions determine the relaxation of the system to the equilibrium distribution from the initial state, but they do not contain the time scales of the sojourn times in the "bright" and "dark" manifolds. The one-time "bright-bright" intensity correlation function measures the joint probability of being in the "bright" state at $t = 0$ and at t. The general expression for a Poisson kinetic process is,

$$C_{\pm \ldots \pm}(t_1, \ldots, t_n) = 1^T \left\{ \prod_n \delta^{(\pm)} e^{-(\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d)(t_i - t_{i-1})} \right\} \frac{\delta^{(\pm)} \rho_{eq}}{|\delta^{(\pm)} \rho_{eq}|} \qquad (2.3.8)$$

where the matrices are time ordered. The 1^T is a unit row vector that denotes a sum over the remaining components and the resulting quantity is a scalar. The "bright-bright" correlation function measures the eigenvalues that correspond to the relaxation of the system from the non-equilibrium initial condition of being in the "bright" state at $t = 0$. The matrix that governs this relaxation is $\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d$.

If the eigenvalues of the matrix $\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d$ are unique, the two time correlation functions $C_{\pm \pm \pm}(t_1, t_2)$ contain all of the available information about the other correlation functions through a decomposition of the correlation function into a sum of multi-exponentials,

$$C_{\pm \ldots \pm}(t_1, \ldots, t_n) = \sum_{i_1, \ldots, i_n} a_{\pm i_1, \ldots, \pm i_n}^{\pm} \prod_{m=1}^{n} e^{\lambda_{i_m}(t_i - t_{i-1})}, \qquad (2.3.9)$$

48

where we can then relate the coefficients, $a^{\pm}_{\pm i_1,\ldots,\pm i_n}$ to each other. As a result, we can determine all of the available information about the system from accurate two time measurements [84]. For completeness, we outline the proof in appendix 2.10. We use the term all available information because different kinetic schemes can result in the same waiting time distributions and therefore produce the same sequence, although from a given kinetic matrix the waiting time distributions are uniquely determined [12]. We present an example of this non-uniqueness in section 2.4.

Although one can theoretically extract information on the duration of time spent in the "bright" or "dark" manifold by measuring an infinite number of multi-time correlations, it is not practically feasible. The approach also has a difficulty when one set of states is short-lived and the system is almost always "bright" or "dark" since the correlations stay near the baseline. One must simultaneously solve for all of the eigenvalues at once, whereas the "event" correlation methods that we will discuss in Sec. 2.4 below separates the "bright" and "dark" processes. The "intensity" correlation function does not contain any readily observable signatures of conformational dynamics[35, 242, 243]. Another difficulty with this approach is the need to accurately calculate the correlation function for a wide range of times and for multiple times, since the "intensity" correlation relaxation time scale will generally be larger than the duration of "events." The "intensity" correlation function does not probe individual elements of $\mathbf{\Gamma}$, $\mathbf{K}$, and $\mathbf{K}_d$, only the total matrix, so that determination of the kinetics is more difficult than the "event" correlation.

Numerical example

The deviations of the population from the equilibrium distribution by knowing that the system is in the "bright" or the "dark" state at a given time are smaller than knowing that the system made a transition from "bright" to "dark" or vice versa at that time. As a result, the signatures of memory can be weak. The weak signatures are made apparent in FIG 2-3. The figure compares the two time "bright-bright-bright" correlation function to the predictions for a single time correlation function, $C_{+++}(t_1, t_2) - C_{++}(t_1)C_{++}(t_2)$, for the model system discussed in Sec. 2.2. The correlation function was calculated with a sliding time window from 25 molecular trajectories. The maximum deviations in the correlation function are around 1% relative to the steady state value of the correlation function. This deviation is the same order as the deviations of the correlation function calculated from the data relative to the exact calculation, and because of correlations in the fluctuations and the high redundancy in the sliding window, the deviations appear to be systematic.

The deviations can be seen in Fig. 2-3(a), which compares the "bright-bright" correlation function, $C_{++}(t)$, determined from the data with the model prediction. The two time correlation function, $C_{+++}(t_1, t_2)$ shown in Fig. 2-3(b) does not have any obvious features that represent memory. Fig. 2-3(c) compares the values of C_{+++} for $t_1 = t_2$ with the square of the single time correlation function. The deviations between the two correlation functions coincide with deviations presented in Fig. 2-3(a), which indicates it is probably a data artifact. Fig. 2-3 (d) shows a contour plot of the deviations for $0 < t_1, t_2 < 10$. The maximum deviations are around 1% of the
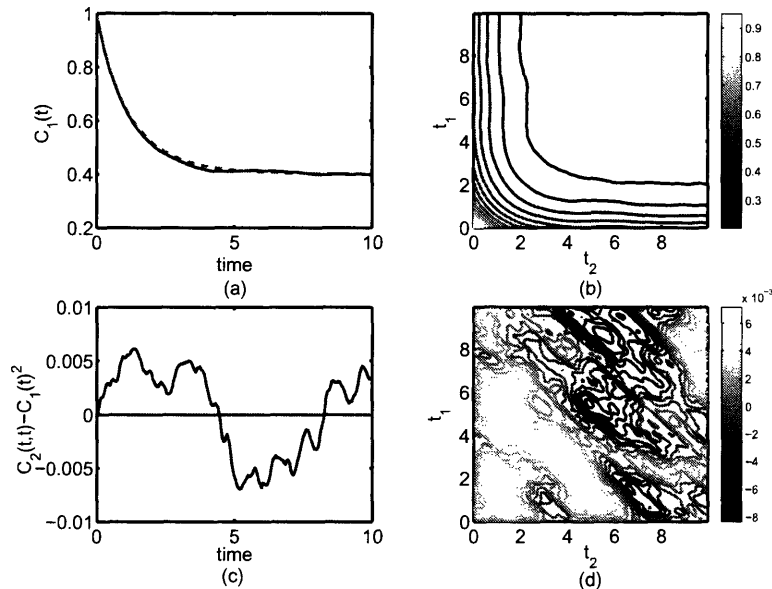
Figure 2-3: Memory in the correlation function generated from the sequences of 25 molecules with kinetic scheme discussed in Sec. 2.2. (a) shows the single time "bright-bright" correlation function, $C_{++}(t)$, measured from the simulation versus the exact solution for the model. As can be seen, the data is not sufficient for the determination of weak features. (b) shows a contour plot of the two time "bright-bright-bright" correlation function, $C_{+++}(t_1, t_2)$. (c) shows the deviation of $C_{+++}(t, t)$ from $C_{++}(t)^2$. The deviations coincide with deviations of the simulation data from the exact correlation function and cannot be considered an indication of memory. These deviations are less than 1% of the normalized correlation function and well within the range of noise. (d) shows a two dimensional plot of $C_{+++}(t_1, t_2) - C_{++}(t_1)C_{++}(t_2)$, which shows the scale of the noise in the system. Once again, the largest deviations are less than 1% of the maximum height of the correlation functions.

total magnitude of the function and may not indicate any memory effects considering the data's deviations from the exact correlation function. These difficulties make it almost impossible to determine the coefficients of the decomposition of the intensity correlations, a_{ij}, above, but suggest that other methods of determining the presence of memory are necessary. One should also note that the deviations do not give a clear signature of a specific kinetic motif since the deviations occur at short time which makes determination of specific features difficult.

2.4 "Event" correlation

The "event" correlation approach determines the statistics of "events." An "event" is the duration of a "bright" or "dark" period. In other words, the start of a "bright" event is the time when the molecule makes a transition from "dark" to "bright." The event ends when the molecule turns "dark" again. The definitions of *events* are demonstrated in Fig. 2-1. We can easily generalize these definitions to photon statistics, where the events are the arrival times of individual photons so the times of events are the times separating photons. The time scales extracted from the measurements pertain to the duration of events and not the overall relaxation measured by the "intensity" correlation approach. One determines the correlations in the length of duration of multiple events, such as two "bright" events separated by one "dark" event, or separated by two "dark" and one "bright" event, etc. The most important indicator of the "event" correlation family is a "bright" event versus an adjacent "dark" event. This quantity is the strongest, but also the most important for cases where we

can determine all available information. Previous work concentrated on two "bright" events separated by a single "dark" event, but these measurements are not as strong and cannot be simply inverted even when the eigenvalues are distinct [35, 242, 243]. The approach contains possible flags to infer non-Poisson or non-renewal behavior, but these flags also require a large amount of data to see them. One of these flags is the "two-event echo," which shows a rise in the two event "bright-bright" duration probability density function compared to the predictions of uncorrelated behavior for $t_1 \approx t_2 >> \tau$, where t_i is the duration time of an event and τ is a characteristic time of the system. The position and height of the echo is a measure of the memory of the system. [35, 242, 243].

The n-"event" probability density function, which we will call the n-"event" correlation function, directly specifies the probability of the path of the system. The correlation or any other quantity is a sum over all possible paths so the "event" correlation is in many ways the basis of all other measurements. The probability density for the 1st n transitions is

$$P(t_1, \ldots, t_n) = 1^T \left[\prod_{i=1}^{n} \mathbf{K} e^{(\mathbf{K}_d + \mathbf{\Gamma})(t_i - t_{i-1})} \right] \rho_{ic}^{(\pm)} \tag{2.4.10}$$

The initial condition $\rho_{ic}^{(\pm)}$ is chosen because the measurements start at a transition and one averages over all of the possible starting transitions. The superscript that appears in the initial condition is to specify the initially observed state. As has been pointed out by Verbeek and Orrit, the single "event" probability does not contain any information about the overall relaxation of the system. In order to extract

53

meaningful information about conformational kinetics, multiple time "event" densities are required, whereas the single time correlation function contains some information about conformational kinetics [35, 242, 243]. This expression is easily expressed in the Laplace domain as well. As discussed in Sec. 2.2, event densities can be sensitive to photon shot noise, which can cause uncertainty in the state of the molecule, and data binning, which can cause uncertainty in the time of transitions, if the photon emission rate is not high enough.

Similar to the two time "intensity" correlation function, when the matrices $\mathbf{\Gamma}^{(\pm\pm)} + \mathbf{K}_d^{\pm\pm}$ have distinct eigenvalues, the "two-event" correlation function contains all of the available information about the Poisson process. The derivation follows the results in appendix 2.10 with some redefinitions of various quantities and is also outlined by Fredkin and Rice [84]. It is important to note that the eigenvalues for the decay constants will be determined by the sub-matrices $\mathbf{K}_d^{(++)} + \mathbf{\Gamma}^{(++)}$ and $\mathbf{K}_d^{(--)} + \mathbf{\Gamma}^{(--)}$ instead of the complete matrix $\mathbf{K}_d + \mathbf{\Gamma} + \mathbf{K}$ and the coupling between the exponential decays for the two different times will be determined by the matrices $\mathbf{K}^{(+-)}$ and $\mathbf{K}^{(+-)}$ instead of $\delta^{(\pm)}$, so one can separate the contributions from $\mathbf{K}^{(\pm\mp)}$ and $\mathbf{\Gamma}^{(\pm\pm)} + \mathbf{K}_d^{\pm\pm}$.

The result implies that there is no additional information in higher densities, but the analysis of the "two-event" correlation function requires binning small amounts of data into a multidimensional array. Since the probability of falling into a specific bin is a Bernoulli variable, statistical fluctuations in the number of events that fall into a bin can affect determination of these relaxation times. One method of avoiding the binning difficulty is the calculation of the covariance of the times of "two-events." Although this calculation can indicate a correlation between adjacent events, one loses

much of the information contained in the density since the covariance is an averaged quantity.

"Event" correlations and non-uniqueness

The "n-event" density function can be rewritten in a revealing form

$$
\begin{aligned}
P(t_1, \ldots, t_n) &= 1^T \left[\prod_{i=1}^{n} \mathbf{K} e^{(\mathbf{K}_d + \mathbf{\Gamma})(t_i - t_{i-1})} \right] \rho_{ic} \\
&= q^T \left[\prod_{i=1}^{n} \mathbf{S}^{-1} \mathbf{K} \mathbf{S} e^{-\mathbf{\Lambda}(t_i - t_{i-1})} \right] p,
\end{aligned}
\tag{2.4.11}
$$

where $\mathbf{S}$ is the similarity transform that diagonalizes $\mathbf{K}_d + \mathbf{\Gamma}$ and $\mathbf{\Lambda}$ are the corresponding eigenvalues. The vectors, $p = \mathbf{S}^{-1} \rho_{ic}$ and $q^T = 1^T \mathbf{S}$, result from the similarity transform. We assume $\mathbf{\Lambda}$ is diagonal because having an algebraic multiplicity greater than one usually requires additional constraints. This form demonstrates a lack of uniqueness since any system with the same eigenvalues, the same matrix $\mathbf{S}^{-1} \mathbf{K} \mathbf{S}$, and the same vectors q^T and p give the same probability distribution for multiple time "events" [12]. In other words, since the density is restricted to probe along the vectors p and q^T complete information about the underlying matrix is lost even if we can determine all of the higher order correlations or "event" densities, but the different models are related through a linear transformation.

A simple example of two indistinguishable kinetic schemes is presented in Fig. 2-2(b) and (c). These are three state models with two bright and one dark state. We can map the three state model into a model with one "bright" state and one "dark" state with multi-exponential waiting-times for the duration of each state. Processes

with one "bright" state and one "dark" state with possibly non-mono-exponential waiting times are known as alternating renewal processes. If we set $k_{b1} = 3s^{-1}$, $k_{d1} = 1s^{-1}$, $k_{b2} = k_{d2} = 0s^{-1}$, $\gamma = 2s^{-1}$, the "bright" state waiting time distribution with dimensionless time is $\frac{2}{5}6e^{-6t} + \frac{3}{5}e^{-t}$. The "dark" state waiting time is a simple Poisson process, e^{-t}. The model in Fig. 2-2(c) can achieve the same waiting-time distribution with the parameters $k_{b1} = 6s^{-1}$, $k_{d1} = \frac{2}{5}s^{-1}$, $k_{b2} = 1s^{-1}$, $k_{d2} = \frac{3}{5}s^{-1}$. These models have different connections, but the ambiguity about the underlying kinetic scheme cannot be resolved.

Numerical example

The advantages of the "event" correlation over the other measurements mainly stems from its separation of the "bright" and "dark" events and the separation of $\mathbf{K}_d + \mathbf{\Gamma}$, which determines the eigenvalues, and $\mathbf{K}$ which determines the connectivity. The "event" correlation is the best indicator in many respects. As shown in Fig. 2-4(a) and (b), we are able to easily fit the waiting time distribution generated by 25 molecules to a bi-exponential form. This bi-exponential form gives us the number of states in both the "bright" and "dark" manifolds and also restricts the range of eigenvalues. A log-plot of the binned data is not obviously bi-exponential, but it is obviously not mono-exponential, which would justify performing a numerical fit to the bi-exponential, which is quite good.

The deviations of the two-event correlation versus the prediction from single waiting time distributions, $P(t_1, t_2) - P(t_1)P(t_2)$, are on the order of 2%, which is stronger
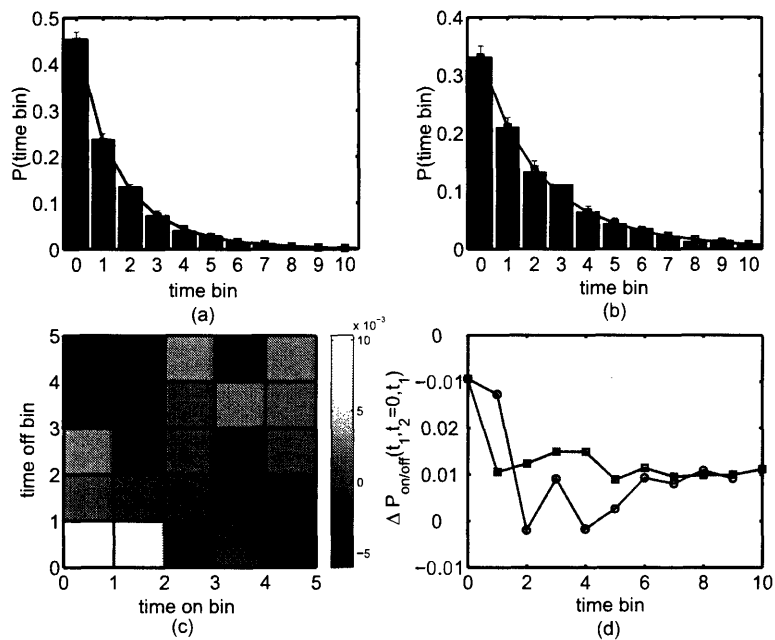
Figure 2-4: Event density generated from data of 25 molecules. (a) shows the "bright" waiting time distribution measured from the data for 25 molecules versus the predictions. (b) shows the "dark" waiting time distribution measured from 25 molecules versus predictions. (c) Shows the error of the two time "bright-dark" density versus the predictions of uncorrelated behavior, $\Delta P_{on/off}(t_1, t_2) = P_{on/off}(t_1, t_2) - P_{on}(t_1)P_{off}(t_2)$. (d) shows two slices of (c), $t_2 = 0$ ($\square$) and $t_2 = t_1$, ($\bigcirc$).

than the correlation function and the echo at later times indicates conformational dynamics. It is important to note that Fig. 2-4, presents the adjacent "bright-dark event" correlations. The "bright-bright event" correlation has weaker memory effects and cannot be used in the analysis for systems with distinct eigenvalues since one needs to know about the "dark" state to determine the connectivity. For the four state model, one only needs to fit the data to a bi-exponential, versus a tri-exponential and a constant for the correlation functions. The reduction in the number of fits also makes use of the "event" quantities more feasible.

2.5 Characteristic function

In this section we examine a recently proposed indicator, the characteristic function and the related moments that we previously referred to as "number densities" [35, 242, 243, 29, 87, 21, 20]. The characteristic function and related moments have an extensive history although suggestions of introducing them into the field of single molecules is fairly recent [243, 29, 87, 21, 20]. The "characteristic function" approach examines the probability of observing n transitions from "bright" to "dark" or vice versa in a period of time, $P(n, t)$ [29]. One classic indicator derived from the characteristic function is the Mandel Q parameter that measures deviations from a Poisson process,

$$\frac{\langle n^2(t) \rangle - \langle n(t) \rangle^2}{\langle n(t) \rangle} - 1, \tag{2.5.12}$$

where $\langle n^i(t) \rangle = \sum_n n^i P(n,t)$ is the expectation of the number of events. Wang and Wolynes and Silbey and co-workers also discuss using moments in several single molecule experiments [215, 14]. An attractive property of the "characteristic function" is the amount of information carried in quantities measured at a single time. The binning with respect to the number of observed events is more natural than binning time. Often one measures moments or "number densities," which resemble the use of correlations and avoids binning. These properties help us avoid problems associated with binning our data in multidimensional arrays.

If there is a visible transition into state j at $t = 0$, the probability of observing n more transitions by time t is given by

$$P(n, t|j, t = 0) = \sum_{i,k} \left[\int_t^\infty dt Q_{ik}(t) \right] * [\mathbf{Q}^{*n}(t)]_{k,j} \qquad (2.5.13)$$

In this expression, $\mathbf{Q}$ is the matrix defined in Eq. 2.2.6, $*$ is the convolution operator, and $*n$ denotes n convolutions of the same term. The $\mathbf{Q}(t)^{*n}$ term represents the probability of making n transitions before time t, and the $[\int_t^\infty dt Q_{ik}(t)]$ term represents the probability of not making the next transition. The expression in the Laplace domain is given by

$$P(n, s|j, t = 0) = 1^T \left[\frac{\mathbf{1} - \mathbf{Q}(s)}{s} \right] [\mathbf{Q}^n(s)]. \qquad (2.5.14)$$

The above equation is valid for any waiting time distributions, not just Poisson kinetics. The formalism can be easily applied to photons statistics, where a transition

corresponds to a detected photon. Since we do not know the original or final state we must introduce the initial distribution and sum over the initial states. As will be discussed below, the best initial condition starts from a switching event, switching from "bright" to "dark" or vice versa, since this initial condition allows simple analysis to determine if the process is the result of single waiting-time distributions. Supplemented with an initial condition $\rho_{ic}^{(\pm)}$, the probability of observing n transitions in a HMC with "bright" and "dark" states becomes

$$P(n, s|i.c.) = 1^T \left[\mathbf{1} s - \mathbf{\Gamma} - \mathbf{K}_d \right]^{-1} \left(\mathbf{K} \left[\mathbf{1} s - \mathbf{\Gamma} - \mathbf{K}_d \right]^{-1} \right)^n \rho_{ic} \qquad (2.5.15)$$

The characteristic function for the number of events, $n \to \zeta$, as a function of the Laplace variable $t \to s$ is

$$G(\zeta, s|ic) = \sum_n P(n, s|ic) e^{in\zeta} = 1^T \left[\mathbf{1} s - \mathbf{\Gamma} - \mathbf{K}_d - e^{i\zeta} \mathbf{K} \right]^{-1} \rho_{ic}. \qquad (2.5.16)$$

This characteristic function comes from the discrete Fourier Transform, which can be easily related to the generating function by Brown (a Z-transform) through analytic continuation [29]. For numerical applications the discrete Fourier transform has better software support and the general approach to inverting a Z-transform is by evaluating it on the unit circle, in which case it becomes the discrete Fourier transform.

Moments of the characteristic function

As with many functions, the Fourier or Laplace transforms generally do not give good insight into the probability function. Because of this difficulty, the most convenient and informative procedure for analyzing a characteristic function will come from the moments. If there is one state with arbitrary waiting time distribution, $Q(s)$, and we see the transitions from the state back into itself, an intermittent blink discussed in [35, 242], we have complete information by only knowing the expected number of blinks as a function of time. This scenario also applies to photon counting statistics, where the blinks are emitted photons. For photon statistics, the photon emission rates from the system will be lower than the "bright-dark" kinetics. As a result the background photon detection rate becomes important. For a Poisson kinetics with Poisson background, the kinetic rates in $\mathbf{K}_d$ need to be increased by the background rate and the emission from the background will correspond to a visible transition back into the same state. For a single state, the first moment is known as the renewal function and it is related to the waiting time distribution by the renewal theorem,

$$Q(s) = s\langle n(s)\rangle/(1 + s\langle n(s)\rangle), \qquad (2.5.17)$$

where $\langle n(s)\rangle$ is the Laplace transform of the expected number of events, $\langle n(t)\rangle = \sum_n nP(n,t)$, whose derivative is the "number density" that we discussed previously [35]. The relation is only valid if the process starts from a blink. If the process does not start from a blink, i.e. arbitrary start time, the expected number of blinks grows linearly in time and the relation is lost, $\langle n(t)\rangle = t/\tau$, where τ is the average time

61

between events, and any memory effects are contained in the higher moments, which are difficult to measure. The analysis can be applied to alternating renewal processes (one "bright" and one "dark" state with arbitrary waiting times) by measuring the expected number of transitions initially starting from a "bright" to "dark" transition and then from a "dark" to "bright" transition. Systems with a single waiting time for both "bright" and "dark" states have a simple memory. The system remembers how long it has been in a particular state, but does not know the path that it took to get there.

Deviations from simple memory

The characteristic function also contains information about the non-renewal nature of the process. For a process with a single waiting time distribution (renewal process),

$$\sum_n n^2 P(n, s) = \langle n^2(s) \rangle = \langle n(s) \rangle + 2s \langle n(s) \rangle^2, \tag{2.5.18}$$

and deviations from this relation imply more complex memory effects. Similar relationships can be derived for an alternating renewal process, supplemented with the appropriate initial conditions.

For the case of intermittent blinks discussed in [35, 242], the inverse Laplace of transform of Eq. 2.5.18 suggests a simple indicator

$$I(t) = \frac{\langle n^2(t) \rangle - \langle n(s) \rangle - 2 \int d\tau \langle n(t - \tau) \rangle \partial_t \langle n(\tau) \rangle}{\langle n(t) \rangle} \tag{2.5.19}$$

For long times $\langle n(t) \rangle \approx \mu t + b$, where μ is the average waiting time and b is an off set caused by the non-Poisson nature of the system. In the long time limit, the indicator function behaves as

$$I(t) \to \frac{\langle n^2(t) \rangle - \mu^2 t^2 - 4\mu b t - \mu t}{\mu t}. \tag{2.5.20}$$

Note that for a Poisson process, $b = 0$ and we recover the Mandel Q parameter.

Long lived correlations and the moments of the characteristic function

To give insight into the memory effects captured by these indicators, we consider a simple example with long lived bright states and intermittent dark states. The transition matrix $\mathbf{Q}$ is given by

$$\mathbf{Q}(s) = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \begin{bmatrix} \frac{1}{1+s} & 0 \\ 0 & \frac{2}{2+s} \end{bmatrix}, \tag{2.5.21}$$

and in the limit as $t \to \infty$ the indicator approaches $\lim_{t \to \infty} \langle n(t) \rangle I(t) = \frac{1}{18p}(\frac{1}{p} - 2)$. For $p = 0$, the indicator is infinite since the memory is infinite. For $p = \frac{1}{2}$ there is no memory of the previous transition and the indicator is zero, but the process is not Poisson. This is a simple renewal process (single waiting time distribution) with bi-exponential waiting time. The renewal indicator becomes negative for $p > \frac{1}{2}$ because the system prefers to flip-flop between states and a fast transition is followed

by a slow one, i. e. anti-correlated steps. The anti-correlation is a sign of a break-down of detailed balance resulting in circulation through the configurations. The important observation is that the indicator is based on first and second cumulants, which are easy to measure. Comparisons of higher-order cumulants of $P(n, t)$ are not as easily measured, and it is difficult to extract meaningful information from these higher cumulants.

In these long persistence situations, large deviations from a normal distribution for intermediate times are possible [29, 87, 21, 20]. For our example with $p \approx 0$ at intermediate times, one observes a superposition of two normal distributions. Although the "two-event" measurements contain some of this information, the information may not be as explicit if the two exponents are comparable $k_1 \approx k_2$ but p is still small.

Moments of a Poisson process and the extraction of kinetic schemes

From Eq. 2.5.16 for the Hidden Markov chain (HMC), the expected number of blinks (also know as renewals) can be written as

$$\langle n(s) \rangle = \frac{1}{s} 1^T \mathbf{K} \left[\mathbf{1}s - \mathbf{\Gamma} - \mathbf{K} - \mathbf{K}_d \right]^{-1} \rho_{ic}, \tag{2.5.22}$$

where $\left[\mathbf{1}s - \mathbf{\Gamma} - \mathbf{K} - \mathbf{K}_d \right]^{-1}$ is the matrix for the relaxation to the equilibrium distribution from the fluctuation that results in the initial transition at $t = 0$. These are the same decay constants measured by the "intensity" correlation functions, but the quantities are integrated because of the $\frac{1}{s}$ term, which often makes extraction of this

information difficult. Similar to discussions above, these eigenvalues are different than the "two-event" density eigenvalues, and the two measurements provide complementary information. One advantage of the characteristic function over the "intensity" correlation function is that the characteristic function directly probes the transition matrix, $\mathbf{K}$, which is also measured in the "event" density so the characteristic function can be viewed as a mixture of the "intensity" and "event" indicators.

For $Re(s) > 0$, the system has a convergent Taylor expansion in terms of the generating variables around $\zeta = 0$, so we only need to know the moments to determine the functional form. If the eigenvalues of the matrix $\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d$ are distinct we can use a similar procedure to those applied to the "intensity" correlation function to relate higher moments to the first few lower moments, so we can theoretically determine the entire generating function from the first few moments. This procedure is discussed in appendix 2.11. Unlike the "intensity" correlation function, the relations are highly non-linear and require higher order moments so this extraction is not practical. Introducing a multiple time moment expansion will prevent the non-linearities, but multiple times also removes the advantages of having a large amount of information contained in a two variable function.

Numerical examples

The above analysis demonstrates the difficultly in using the generating function to see explicit details of the system. Although the position and variance of the number of observed transitions give indications of memory effects, the underlying causes of these

65
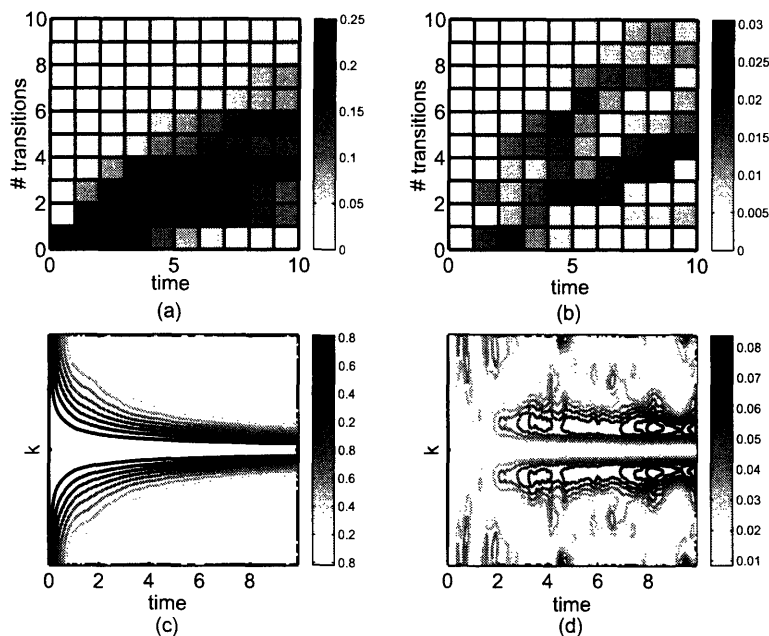
Figure 2-5: $P(n|t)$ and the absolute values of its generating function $G(k|t)$. (a) shows $P(n|t)$ determined from data of 25 molecules. (b) shows the errors between $P(n|t)$ and predictions for alternating renewal process. (c) shows the generating function, $G(k|t)$, determined from the data. (d) Shows the absolute errors between $G(k|t)$ and the predictions for alternating renewal processes.

memory effects are not obvious. The lack of specific features makes it difficult for the generating function to distinguish features in the data. The lack of other features can be seen in Fig. 2-5. In this figure, we examine the data generated for the model discussed in section 2.2. The initial condition is a "dark" to "bright" transition and every "dark" to "bright" transition in the trajectory is used as an initial condition. Fig. 2-5(a) shows a histogram of the number of renewals as a function of time, $P(n, t)$, determined from the data sequence.

The histogram is compared against the expected probability for an alternating renewal process in FIG 2-5(b). The alternating renewal waiting time distributions are determined from "event" correlations. The maximum error is 3% with typical errors around 0.5%. The characteristic function is calculated in Fig. 2-5(c). As can

66

be seen, most of the details of this distribution are hidden by noise in the data. The only notable features are the central peak, whose width we can measure through the second moment and the weak peaks at $k = \pm\pi$. The peaks at $k = \pm\pi$ are the result of the alternating renewal nature of the generating function. The "dark" events are longer lived than the "bright" events so the number of renewals is more likely to be odd than even if we start from a renewal into a "bright" state. Fig. 2-5(d) compares the generating function calculated from the data and the alternating renewal process predictions. The deviations between the alternating renewal prediction extracted from the data and the complete set of single molecule data are small less than $< 1\%$ relative to the maximum values of the function–unity–except in the vicinity of $k = 0, \pm\pi$. The error in this vicinity will grow in height but shrink in width. These peaks indicate a memory effect, but it is difficult to discern a motif associated with this memory effect.

Brown used the characteristic function approach to compare the signatures of the four state model with the fluctuating bottle-neck model [29]. The fluctuating bottle-neck model corresponds to a one dimensional diffusion process in a harmonic well with a reaction rate that depends quadratically on the coordinate.

$$\partial_t P_\pm(t) = D\nabla^2 P_\pm(t) + \nabla\left(kx P_\pm(t)\right) - \kappa_\pm x^2 P_\pm(t) + \kappa_\mp x^2 P_\mp. \qquad (2.5.23)$$

Consistent with previous notation $+$ represents the "bright" state and $-$ represents the "dark" state. He demonstrates that for perfect data the characteristic functions are different for the four state model and the fluctuating bottle-neck model, as long
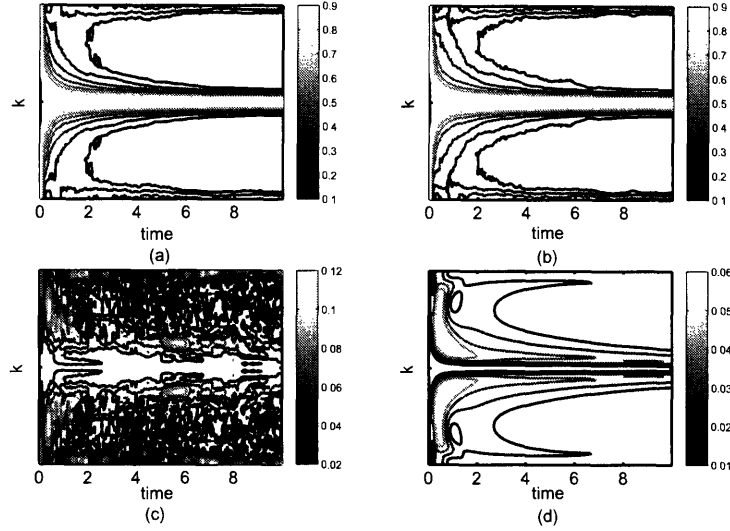
Figure 2-6: Comparison of the fluctuating bottle-neck model and the four state model. (a) The absolute value of $G(k|t)$ for the fluctuation bottle-neck model as determined from data of 25 molecules. (b) The absolute value of $G(k|t)$ for the four state model. (c) The absolute value of the difference in the two models. (d) The theoretical results of (c).

as the temporal fluctuations of the rate constant is slow in comparison with the rates of transitions. His example shows the exact generating functions with temporal fluctuations in the rate constants that are 25 times slower than the average rate of reaction. As demonstrated in Fig. 2-6, when these measurements are made on stochastic data with the more interesting scenario of the fluctuations in the rate constant being the same order as the average transition rates, the signature cannot be successfully deduced. The figure compares data generated from a diffusion process with $D = k = \kappa_+ = 1$ and $\kappa_{-1} = K_{eq}\kappa_+ = 2$, with a four state model with $\gamma_{b1} = \gamma_{d1} = 0.289706$, $\gamma_{b2} = \gamma_{d2} = 1.71029$, $k_{b1} = \frac{1}{2}k_{d1} = 0.417953$, $k_{b2} = \frac{1}{2}k_{d2} = 4.43615$. The constants for the four state model are chosen according to the procedures outlined in [29], which are close to maximizing the similarity of the two models.

The data includes the trajectories of 25 molecules run over 100 time units resulting

68

in an average of 150 turn-overs [224]. The initial condition starts from a "bright" to "dark" and the system is averaged over all possible initial transitions, which gives about 1250 total trajectories of 10 time units. Fig. 2-6(a) shows the absolute value of the fluctuating bottle-neck characteristic function determined from the data and Fig. 2-6(b) shows the absolute value of the four state model characteristic function determined from the data. Fig. 2-6(c) shows the absolute value of the difference between the two models calculated from the data and Fig. 2-6(d), is the ideal plot of Fig. 2-6(c). Even the ideal signal is only about 5% of the total signal, and the noise in the characteristic function is around 10%, which prevents any strong conclusions using generating function methods. The signal would be weaker if one did not specify the sequence starting on a transition at $t = 0$ [29]. In the long time limit the two models become indistinguishable since they have the same average rate of transitions and similar fluctuations in the number of transitions, but at short times there are possible signatures. It important to note that similar systematic deviations between characteristic functions generated from data sets from the same stochastic process also appear because of the averaging over the same data sequence for 25 molecules. As a result, the strongest possible conclusion is that the data might not be consistent with a four state model, but these measurements do not give a good quantitative measure.

2.6 Comparison of different indicators

This chapter presents a critical analysis of proposed indicators for single molecule experiments. All indicators have the potential to give qualitative insight into the dynamics of a single molecule system. The information contained in each indicator is similar, but it is convoluted differently in each measurement so that one indicator may have a relatively stronger signature than the others. An example is the long memory effects captured by the characteristic function.

The "intensity" correlation function is generally the weakest indicator. Measurements do not start from a transition so the deviations from equilibrium are small. Due to the initial condition, accurate measurements of the intensity correlation are sensitive to long "bright" or "dark" periods. The "intensity" correlation does not contain clear signatures of the dynamics of the system and only probes the total matrix, $\Gamma + \mathbf{K} + \mathbf{K}_d$, which does not give insight into individual contributions.

The "event" correlations are generally the most useful of the three indicators. They separate the contributions from $\mathbf{K}$ from $\Gamma + \mathbf{K}_d$ and the contributions from the "bright" and "dark" states. The "event" correlations also contain the two event echo, which is a signature of conformational dynamics, but the indicator requires data binning, instead of averaging, which causes a loss of temporal resolution.

The characteristic function and its moments are a hybrid between the "intensity" correlations and "event" correlations. The characteristic times are determined by the same matrix as the "intensity" correlation, $\Gamma + \mathbf{K} + \mathbf{K}_d$, but the coefficients allow exploration of $\mathbf{K}$ separately. The characteristic function allows averaging instead

of binning, but it does not allow separation of "bright" and "dark" states. The average number of transitions and its variance give insight into memory effects of the system, but there are no other salient features. Application of complete information for distinct eigenvalues becomes a highly non-linear problem for this indicator, unlike the other indicators.

In principle, Markovian systems with distinct eigenvalues only require two-time information to extract all available information about the process, but as shown in the simple example in Sec. 2.2 all available information does not give a unique kinetic scheme. The analysis can be extended to processes with a limited number of degeneracies, such as double degeneracies. These relations are theoretically interesting, but practical implementation is difficult if not impossible. The indicators can still give valuable information such as the time constant for relaxation or the time constants for the duration of an "event." A major cause of the difficulties with any indicator is the large data requirements and the lack of a unique solution. A more robust numerical approach that does not depend on the inversion of averaged data is required. Below we give a demonstration of combining the single molecule indicators with a Bayesian numerical approach to extract possible kinetic schemes.

2.7 Introduction to application of Bayesian statistics

As discussed above, a formidable theoretical obstacle facing single molecule analysis is the extraction of pertinent information about stochastic chemical processes from noisy data. Several authors have proposed simple indicators to attempt to reveal these processes [223]. Generally, these indicators give qualitative indications of deviations from simple Poisson or renewal processes (i.e. simple single waiting time processes), but quantitative extraction of the characteristics of these processes is difficult if not impossible. The major obstacles to using these indicators are the data binning and data averaging that they require, which removes much of the available information. The indicators may also be sensitive to instrument resolution and background counts, whose effects may be difficult to incorporate in the analysis.

Despite these difficulties, these indicators can give valuable qualitative insight into the behavior of the system and can be used to reduce the number of possible models that describe the system. After developing this insight and reducing the number of models, a more robust numerical routine that does not require the reduction of information through data binning or data averaging is required to move beyond a qualitative description of the system. One numerical routine that can perform a complete sequence analysis is Markov Chain Monte Carlo with Bayesian statistics [177]. Markov Chain Monte Carlo with Bayesian statistics starts from the use of Bayes formula, which compares the relative probability that two models, M and M', reproduce

a data set,

$$\frac{P(M|D)}{P(M'|D)} = \frac{P(D|M)P(M)}{P(D|M')P(M')}.$$
(2.7.24)

In this expression $P(M|D)$ is the posterior probability of the model given the data, $P(D|M)$ is the probability of the data given the model, and $P(M)$ is the a priori probability of the model, which is often assumed to be uniform or log-uniform or a conjugate prior [6, 86]. In other words, from our previous experience we start with a given distribution of possible models, $P(M)$. After examining the data through the Bayesian formula, we modify our guess about the possible model parameters, which results in the new distribution $P(M|D)$. This quantity tells us about our certainty in predicting the values of parameters.

As a demonstration of the philosophical approach to single molecule problems, we will apply the Bayesian method to the simple four state model that we examined above, Fig. 2-2. For numerical calculations we use the same constants as above, $k_{b1} = 0.75s^{-1}$, $k_{d1} = 0.50s^{-1}$, $K_{b1/b2} = 0.44$, $k_{b2} = k_{b1}K_{b1/b2} = 0.33s^{-1}$, $k_{d2} = k_{d1}K_{b1/b2} = 0.22s^{-1}$, and $\gamma_{b1} = \gamma_{b2} = \gamma_{d1} = \gamma_{d1} = \gamma = 0.1s^{-1}$. The four parameters we examine are k_{b1}, k_{d1}, $K_{b2/b1}$, and γ. The numerical values of these kinetic constants are chosen so that the waiting time distribution for both the "bright" and "dark" states are obviously not mono-exponential, which we determined from the indicators presented above, but the rate of modulation between the two states does not cause apparent time separation in any of the indicators. We also perform the analysis with all of the parameters free, but this information is difficult to present in a visual form so we concentrate on the four parameter fit.

From the model system with the specified parameters, we generate a sequence of "bright" and "dark" states for 25 molecules with a duration of $300s$, which allows about 150 observed turn-over events per molecule and 7500 pieces of data to determine the four parameters. The simulation details used to generate this data are presented in appendix 2.12.1. These data sets are much smaller than the data sets collected in the experiments by Xie [138]. By applying the "event" correlation indicator discussed in Sec. 2.4 to this data, one is able to deduce that the "bright" and "dark" decays can both be fit with bi-exponentials which suggests that they both contain two states. The "event" density also indicates a memory effect, which is confirmed by the characteristic function. This information allows one to reasonably suggest the four state model as a candidate to describe the system.

Given the kinetic scheme, we can simply calculate $P(D|M)$ through iterative matrix multiplication. Given the initial state of the system, "bright" or "dark", and the times of the transitions, $\{t_i\}$, the probability of the data given the model is

$$P(D = \{t_i\}|M) = 1^T \left[\prod_i \mathbf{K} e^{(\mathbf{K}_d + \mathbf{\Gamma})t_i} dt_i \right] \delta^{(\pm)} \rho_{eq}, \qquad (2.7.25)$$

where the matrix definitions follow those used previously [223, 35, 242, 243]. The $\delta^{(\pm)}$ is determined by the initial condition. Since we can calculate the relative probabilities, we can perform Monte Carlo on the probabilities to determine models that are consistent with the data. The exact method of calculating the probabilities and performing the Monte Carlo simulation are outlined in appendix 2.12.2.

The approach can also be extended to experiments with more complex data, such

74

as photon counting statistics and can incorporate data deficiencies or statistical fluctuations from sources other than the system such as instrument resolution and response [121, 225]. For most of the analysis we assume the switching between states is sufficiently fast so that these times are sharp variables relative to the duration of the "bright" and "dark" states, but generalizations can easily be made in this framework since we do not require the system to have Poisson kinetics [101]. From our proposed model, we determine the probability of the initial condition and the probability of the transitions at the recorded times. The probability of the sequence is the result of iterative matrix multiplication.

For many optimization applications, the Monte Carlo approach avoids difficulties associated with gradient based likelihood maximization. Gradient based maximum likelihood approaches successfully determined point estimates of the most likely set of parameters for single membrane ion channel experiments, but the calculation only determines the best fitting parameters and the curvature of the likelihood function, $P(D|M)$ at this point [172]. The Monte Carlo approach can show more detail in the probability distribution, such as multiple minima with similar probabilities [101]. The shot noise in the data creates large uncertainties and makes these maximum likelihood estimates inaccurate so a global estimation of the parameter distribution becomes important. Often, if the data is not sufficient or other difficulties arise that prevent the system from finding the most likely parameters, the probability density signals these difficulties.

2.8   Application of Bayesian Approach

We applied the Bayesian algorithm outlined in appendix 2.12.2 to the data set generated by the algorithm discussed in appendix. 2.12.1. Since the kinetic rates are positive quantities and we do not initially know the magnitude of these rates a natural initial (a priori) distribution for the kinetic rates is log-uniform, $P(M) \propto 1/M$ [6]. As a result the Monte Carlo jump sizes are proportional to the magnitude of the kinetic rates. This log-uniform distribution is not normalizable (improper), but the likelihood, $P(D|M)$, will give the necessary truncation to prevent any problems associated with this normalization. Other a priori distributions are possible based on the analysis with other indicators. From the multi-exponential fit determined by the "event" correlations, one may want to restrict the eigenvalues of the matrix to a small interval around the fitted parameters.

Fig. 2-7 shows 2-dimensional projections onto the principal axes of the posterior probability density for the parameters determined from a data set, and TABLE 2.1 contains the mean and covariance of these parameters. The predicted mean values for these constants are very accurate and the variances are fairly small, which shows that the data is sufficient to determine these parameters. We note that the maximum likelihood value is not necessarily in the center of the distribution, which shows that the distribution is not Gaussian. Because the eigenvalues and eigenvectors depend on the parameters through an inherently non-linear functional form, the asymmetry is not surprising. The largest uncertainty is in the constant $K_{b2/b1}$. This constant enters into the determination of the eigenvalues through multiplication with k_{b1} and

	Actual Values			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
	0.75	0.50	0.10	0.44
	Mean Values			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
	0.77	0.50	0.10	0.43
	Covariance $\times 10^3$			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
k_{b1}	0.86	0.33	0.42	-0.73
k_{d1}		0.39	0.27	-0.50
γ			0.89	-0.74
$K_{b2/b1}$				1.4

Table 2.1: Mean and Covariance for the Monte Carlo simulation presented in Fig. 2-7. Stability analysis determines normal modes with standard deviations of $5.1, 2.1, 1.8, 1.3 \times 10^{-2}$.

k_{d1} so the additional uncertainty comes from compensation for fluctuation in these other quantities.

The predicted maximum likelihood estimate is slightly off-set from the real parameters. Many sets of data examined in simulations converge to a roughly Gaussian distribution with a mean that is slightly off-set from the actual parameters although some simulations fail to converge and a few simulations converged to a set of parameters that are far from the actual set of parameters used to generate the data. Even for data sets with as few as 5 molecules, many simulations predict maximum likelihood estimates around the actual parameters, but more importantly, convergence to parameters far removed from the actual parameters are rare.

The most likely point estimates encountered during a Monte Carlo simulation for 500 different sets of data with 25 molecules and $300s$ trajectories are plotted in Fig. 2-8. Although the Bayesian philosophy concerns determining the entire probabil-
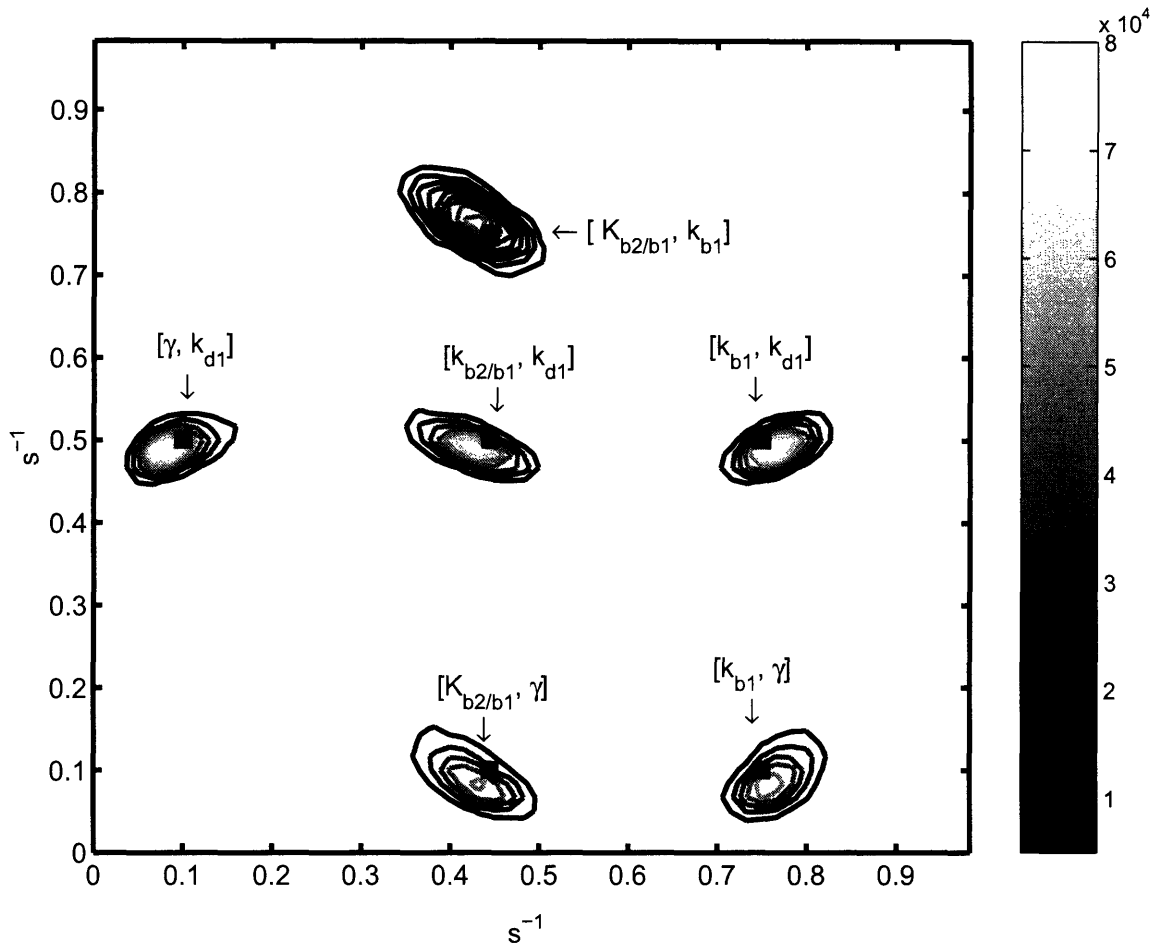
Figure 2-7: Contour plots of two dimensional projections onto the principal hyper-planes of the probability density generated by the Monte Carlo simulation that determines the four parameters for the model in Eq. 2.2.2. The gray-scale is measured in number of points for approximately $5 \cdot 10^6$ Monte Carlo Samples. The black squares mark the positions of the parameters that generated the data. Each peak is pointed to by a two component label for that peak, such as $[\gamma, k_{d1}]$. The first entry corresponds to the horizontal axis and the second entry corresponds to the vertical axis. $K_{b2/b1}$ has the largest uncertainty since it must adjust itself to fluctuations in k_{b1} and k_{d1}.

ity distribution, the point estimates give good insight into the reproducibility of the simulation. From this plot it is apparent that the predicted maximum likelihood parameters are distributed around the actual parameters, except for one outlying data point. The distribution of the maximum likelihood estimates resemble the distributions of the probability distribution of the parameters for a single set of data. The stochastic nature of the underlying dynamics causes these off-sets. This noise makes the use of simple maximum likelihood point estimates of the parameters statistically uncertain unless other analysis is performed.

To help determine the off-set caused by the noise in real single molecule experiments, we can break a large data set into several subsets and perform optimization of the parameters with these subsets and compare optimal parameters; this is known as cross-validation [99]. For this application, breaking data sets up by taking segments of single molecule sequences or by performing the analysis on different single molecule sequences has about the same effect on the cross-validation. In fact, future analysis concerns a data set that contains a single long sequence [225]. Mixing these subsets will improve our predictions of off-sets and allow us to understand the sensitivity of parameters to the intrinsic noise in these systems. Breaking the data up into smaller subsets will also allow the simulation to search larger regions of parameter space since the sensitivity of the Bayesian score scales linearly with the length of the sequence and number of molecules. In many ways, the length of the sequences corresponds to a fictitious inverse Boltzmann temperature.
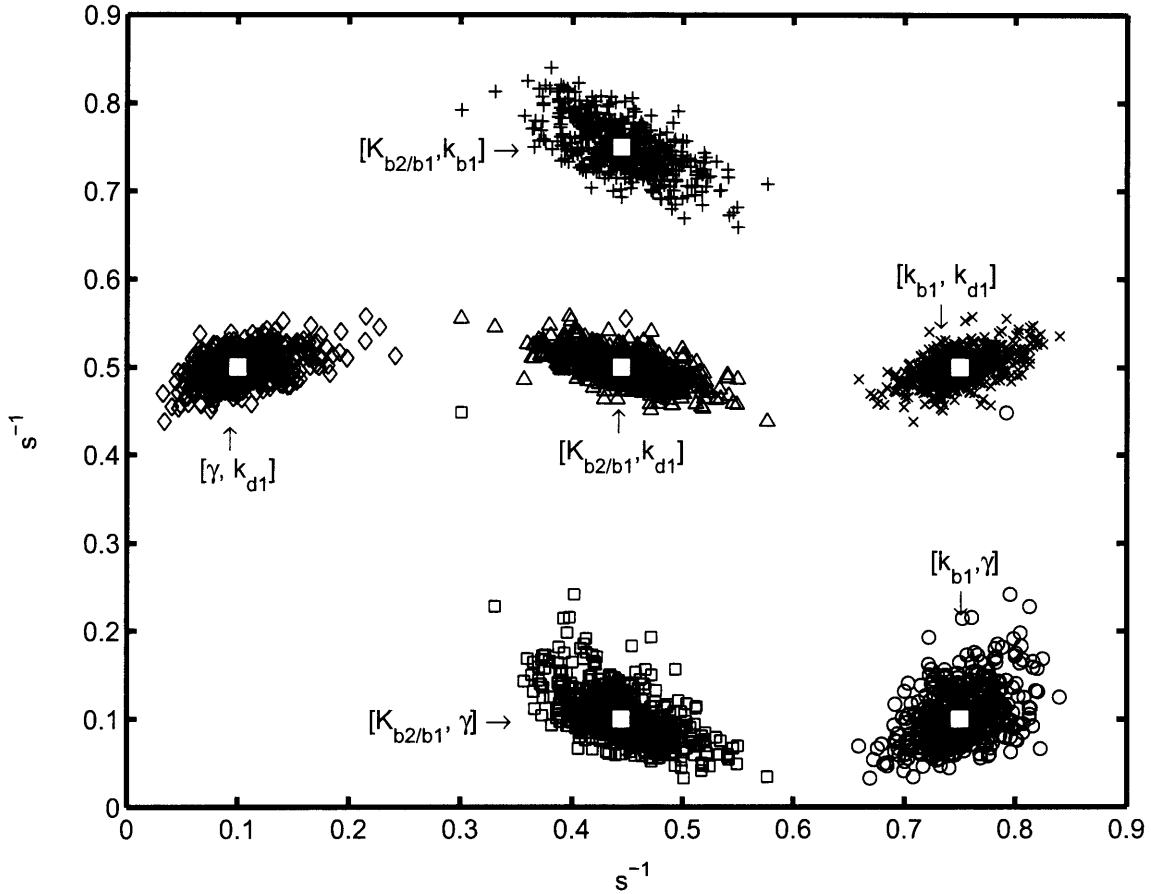
Figure 2-8: Two dimensional projections onto the principal hyper-plane of the vectors of the best fitting parameters encountered during a Monte Carlo simulation for 500 different data sets. In comparison, Fig. 2-7 is the total probability determined from a single data set, but this plot shows the peak position for several data sets. The optimal parameters are distributed around the white squares that label the parameters that generated the data sets. Different symbols are used for each projection to show the single outlier. Similar to Fig. 2-7, each cluster is pointed to by a label such as $[\gamma, k_{d1}]$. The first entry corresponds to the horizontal axis and the second entry corresponds to the vertical axis.

Variations of the simulation

The Monte Carlo simulation failed to converge to any value when the data sets are not consistent with the model. We tested this property by trying to fit the simple four parameter model to a sequence created from a stretched exponential waiting time distribution. Both the "bright" and "dark" waiting time distributions are given by $\frac{9}{8}e^{-(\frac{9}{4}t)^{1/2}}$, which has a characteristic time of 8/3, which is comparable to $\frac{k_{b1}^{-1}+k_{b2}^{-1}}{2}$ in Eq. 2.2.2. When the model tried to fit this data, it would either set many of its parameters to zero (large negative values for the log of the parameters) or wander through the parameter space without converging.

For situations where the difference between the proposed model and the actual kinetics is not as great as the stretched exponential, the manifestation of errors are more subtle. We demonstrate the subtlety by attempting to fit data generated from a five parameter model with the original four parameter model. The five parameter model is also a four state model with the same k_{b1},k_{d1} and $K_{b2/b1}$, but two different values for the γ's. For this demonstration, we set $\gamma_{b1} = \gamma_{d1} = 0.1s^{-1}$ and $\gamma_{b2} = \gamma_{d2} = 0.2s^{-1}$. Fig. 2-9 and TABLE 2.2 summarize the results of one of the best fits of this optimization. Although all of the parameters are shifted relative to the true parameters, the greatest uncertainty appears in the single γ since this quantity is not well defined in the model that generated the data.

As with all parameterizations, the data requirements scale with the number of parameters. For more complex models, more data may be needed. Since the amount of data serves the role of temperature in the sampling of a partition function, large

81
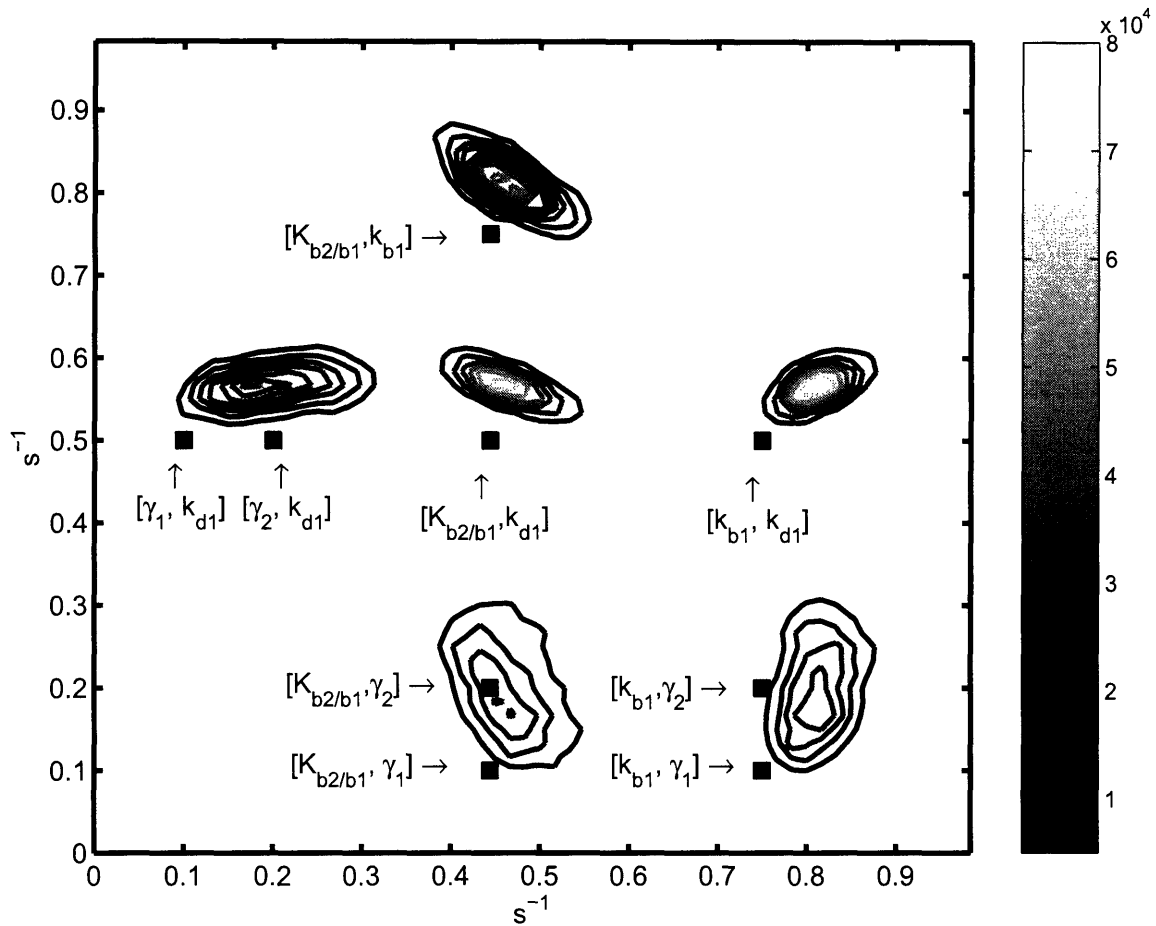
Figure 2-9: Contour plots of two dimensional projections of the probability density generated by the Monte Carlo simulation that determined the four parameter fit to the five parameter model discussed in section 2.8. The black squares mark the positions of the parameters that generated the data. The distribution of γ is much wider than the distribution presented in Fig. 2-7, which reflects the fact that this parameter is not defined for the model that generated the data. Similar to Fig. 2-7, each cluster is pointed to by a label such as $[\gamma, k_{d1}]$. The first entry corresponds to the horizontal axis and the second entry corresponds to the vertical axis.

	Actual Values			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
	0.75	0.50	0.10/0.20	0.44
	Mean Values			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
	0.82	0.57	0.22	0.47
	Covariance $\times 10^3$			
	k_{b1}	k_{d1}	γ	$K_{b2/b1}$
k_{b1}	0.88	0.33	0.52	-0.82
k_{d1}		0.42	0.35	-0.56
γ			4.0	-0.9
$K_{b2/b1}$				1.5

Table 2.2: Mean and Covariance for the Monte Carlo simulation presented in Fig. 2-9. Note that the data is generated from a model with two γ's, but the fit is performed with a single γ. Stability analysis determines normal modes with standard deviations of $6.7, 4.2, 1.8, 1.6 \times 10^{-2}$. Because the model with a single γ did not generate the data, there is a fairly large variance for γ relative to the other parameters.

amounts of data correspond to a very low temperature and can result in a failure to explore the relevant parameters space due to trapping in local minima. This trapping phenomenon is analogous to diffusion on a rough potential energy surface [53]. We found that the use of 250 or more molecules with a trajectory length of $300s$ leads to trapping in local minimum. This trapping shows that the surface is generally not monotonically decreasing to the global minimum. One should use standard approaches such as annealing to help the system find the global minimum [186]. One method of incorporating annealing is the addition and mixing of data during the simulation, i. e. exchanges data used in the optimization with unused data during the simulation. The mixing is important to preventing certain pieces of data from dominating the optimization and avoiding convergence to parameters that are consistent with all of the data.

	Actual Values							
	k_{b1}	k_{b2}	k_{d1}	k_{d2}	γ_{b1}	γ_{b2}	γ_{d1}	γ_{d2}
	0.75	0.33	0.50	0.22	0.10	0.10	0.10	0.10
	Mean Values							
	k_{b1}	k_{b2}	k_{d1}	k_{d2}	γ_{b1}	γ_{b2}	γ_{d1}	γ_{d2}
	0.75	0.32	0.50	0.19	0.12	0.09	0.11	0.14
	Covariance $\times 10^3$							
	k_{b1}	k_{b2}	k_{d1}	k_{d2}	γ_{b1}	γ_{b2}	γ_{d1}	γ_{d2}
k_{b1}	3.0	1.0	1.1	0.3	1.8	0.25	0.88	-1.4
k_{b2}		1.5	0.85	0.75	0.78	-0.80	0.37	0.83
k_{d1}			1.0	0.46	0.85	-0.22	6.7	-0.46
k_{d2}				0.87	0.61	-0.19	0.36	-0.92
γ_{b1}					2.7	1.2	0.11	-0.97
γ_{b2}						2.1	0.27	-0.01
γ_{d1}							1.6	0.55
γ_{d2}								2.1

Table 2.3: Mean and Covariance for the Monte Carlo simulation with eight parameters. Stability analysis determines the normal modes with standard deviations of $8.5, 5.8, 4.8, 3.4, 2.1, 1.7, 1.1, 0.9, \times 10^{-2}$.

The effects of excessive data, such as trapping in local minima, are not a consideration even if we perform the optimization of all 8 parameters with no constraints. Using 50 molecules, instead of 25, we generated data with the four parameter model, but let all eight parameters vary independently and then imposed $k_{b1} > k_{b2}$ at the end of each Monte Carlo iteration. Most of the simulations are able to locate a global minimum that is near the actual parameters. The results for one of these simulations is presented in TABLE 2.3. The largest standard deviation in a mode is $\sim 8.5\%$, which is a typical value for most simulations. For 25 molecules, the largest standard deviation of one of the normal modes was around $\sim 15 - 20\%$, which shows that the certainty in the parameters improve roughly as the expected $\frac{1}{\sqrt{n}}$ scaling as more data is added.

Four state model and the fluctuating bottle neck

The four state model is often compared with the fluctuating bottle-neck model (i.e. the diffusion controlled reaction) in the literature [29, 35, 242, 243, 92, 216]. The fluctuating bottle-neck model describes a one dimensional diffusion process in a harmonic well with a reaction rate that depends harmonically on the coordinate,

$$\partial_t P_\pm(t) = D\nabla^2 P_\pm(t) + \nabla\left(kx P_\pm(t)\right) - \kappa_\pm x^2 P_\pm(t) + \kappa_\mp x^2 P_\mp. \qquad (2.8.26)$$

We previously demonstrated that the difference between the characteristic function of the fluctuating bottle-neck model with $D = k = \kappa_+ = 1$ and $\kappa_{-1} = K_{eq}\kappa_+ = 2$, and a four state model with $\gamma_{b1} = \gamma_{d1} = 0.289706$, $\gamma_{b2} = \gamma_{d2} = 1.71029$, $k_{b1} = \frac{1}{2}k_{d1} = 0.417953$, $k_{b2} = \frac{1}{2}k_{d2} = 4.43615$ is small with the maximum deviation in the characteristic function of $\approx 5\%$, which is smaller than the approximate noise levels for the reasonable amount of data (See Sec. 2.5).

We consider the time traces of 25 molecules for 100 time units generated from both the fluctuating bottle-neck and the four state model. For the given model parameters, each molecule performs approximately ~ 150 turn-overs. From this data, we attempt to find a four state model that optimizes the fit to both sets of data. Applications of Bayesian statistics to a continuous diffusion model will be discussed in applications to single photon experiments [225]. TABLE 2.4 compares the means and covariances for a typical run (typical average value and variance). The first important observation is that the optimal fitting parameters to the fluctuating bottle-neck model are different from the parameters suggested by Brown's procedure [29].

	Actual Values For Four State				
	k_{b1}	k_{b2}	K_{eq}	γ_{b1}	γ_{b2}
	0.42	4.4	2.0	0.29	1.7
	Mean Values (Four State/Bottle-Neck)				
	k_{b1}	k_{b2}	K_{eq}	γ_{b1}	γ_{b2}
	0.44/0.41	4.6/4.1	2.0/2.0	0.26/0.43	1.7/2.2
	Covariance $\times 10^3$ (Four State/Bottle-Neck)				
	k_{b1}	k_{b2}	K_{eq}	γ_{b1}	γ_{b2}
k_{b1}	0.39/0.82	1.9/4.9	-0.67/-0.57	-0.20/-0.95	-0.18/-0.21
k_{b2}		43./51.	-8.4/-7.2	-1.0/-4.0	8.4/17
K_{eq}			6.4/6.6	$1.0 \cdot 10^{-2} / 2.0 \cdot 10^{-2}$	-0.81/-0.45
γ_{b1}				0.55/2.3	1.8/4.5
γ_{b2}					20./40.

Table 2.4: Mean and Covariance for the Monte Carlo simulations attempting to fit the four state model to data generated by a four state model and by the fluctuating bottle-neck model. Stability analysis determines the normal modes with standard deviations of $40., 21., 14., 8.6, 4.1 \times 10^{-2}$ for fitting the four state model to itself and standard deviations of $44., 23., 20., 8.1, 5.8 \times 10^{-2}$ for fitting the four state model to the fluctuating bottle-neck.

Maximizing statistical overlap of the sequences between two models is actually a non-trivial problem, and Brown's parameterization only matches correlation functions, which does not necessarily maximize the overlap of probability.

From the covariance matrices it is apparent that k_{b2} and K_{eq} have comparable variances for both data sets, but the variances of k_{b1}, γ_{b1}, and γ_{b2} are over twice as large for the fits to the fluctuating bottle-neck model. The larger uncertainty in the parameters can be used as a flag to suggest the exploration of other models, which can then be compared through the use of the Bayesian score. Similar to the choice of optimal parameters, one can use Bayesian statistics to choose from models with different physical features or complexity. Comparison of seemingly disjoint models has a rich history with several aspects including determination of when to increase

the number of parameters, so this step is omitted here for brevity, but several references address model comparison through Bayesian methods [47]. Many of these comparisons start to use the Bayesian score that we calculate to perform our Monte Carlo simulations.

2.9   Conclusion

Single molecule experiments offer an opportunity to gain significant insight into the physics of glasses, biomolecules, and other complex systems. The insight is limited by both the amount of collected data and the analysis performed on the data. Previously proposed indicators and other quantities used in the analysis of single molecule experiments contain useful information that give qualitative insight into the physics of the system, as demonstrated in previous references [35, 242, 243, 223]. The useful information includes various time constants for the relaxation of the system and connectivity between these relaxation times (i.e. memory effects), but the extraction of quantitative information from these indicators is difficult because the indicators require data binning and data averaging which removes large amounts of useful information.

To move beyond indicator analysis requires a numerical method that does not require ill-conditioned data inversion or averages out information contained in the data. Bayesian analysis with Monte Carlo optimization is one strong candidate. Implementing Bayesian analysis still requires the use of the previous indicators to determine constraints on possible models, such as the number of states or restric-

tions on eigenvalues, but the Bayesian approach gives quantitative estimates of the parameters and uncertainties in these parameters. By incorporating the constraints discovered through the indicators into the a priori distribution of possible models, one can use Monte Carlo with the Bayesian score as the Boltzmann energy to optimize the parameters. Applications of this approach to a simple four state model demonstrate its capabilities to reproduce the correct parameterization and give uncertainties in these parameters. The Bayesian approach also has the ability to distinguish different models as shown by the comparison of the fluctuating bottle-neck and four state models.

Many other scenarios exist and should be explored, such as the role of photon statistics and continuous distributions of states. The Bayesian approach performed well in our tests of some of these other scenarios, but additional analysis requires motivation from applications to real systems so we will not go into detail about other simulations with computer generated data. Preliminary results from the analysis of single photon emission events from fluorescence quenching experiment are very promising [240, 225]. Of all the existing approaches, the Bayesian approach is the most reliable and robust method of numerically analyzing single molecule data, and we encourage experimentalists to explore the application of this approach to their single molecule data. The Bayesian approach will become a standard method of single molecule analysis in the future.

2.10 Appendix: Proof of complete information content

It is important to demonstrate the completeness of the information so that we can compare results with other indicators. The derivation follows those in the references [84]. The n-time correlation function is written in the Laplace domain as,

$$1^T \left\{ \prod_n \delta^{(\pm)} \mathbf{S} \left[\mathbf{1} s_n + \Lambda \right]^{-1} \mathbf{S}^{-1} \right\} \frac{\delta^{(\pm)} \rho_{eq}}{|\delta^{(\pm)} \rho_{eq}|} = \sum_{i_1,\ldots,i_n} a^{\pm}_{\pm i_1,\ldots,\pm i_n} \prod_{m=1}^{n} \frac{1}{s_n + \lambda_{i_m}}. \quad (2.10.27)$$

The matrix $\mathbf{S}$ is a similarity transform of the matrix $\Gamma + \mathbf{K} + \mathbf{K}_d$, and Λ is the matrix of corresponding eigenvalues, λ_i. Since the eigenvalues are distinct this matrix is diagonal. The prefactor, $a^{\pm}_{\pm i_1,\ldots,\pm i_n}$, is determined by fitting the n-time correlation function to the functional form presented above. The initial condition appears in the superscript of the prefactor, $\pm$, and the $\pm i_m$ refers to the index of the eigenvalues and the state measured at the time t_m. Note that fitting the functional form can be done in the time domain. The functional form of the prefactor is

$$a^{\pm}_{\pm i_1,\ldots,\pm i_n} = 1^T \left\{ \prod_n \delta^{(\pm)} \mathbf{S} \delta_{i_n} \mathbf{S}^{-1} \right\} p_{\pm}. \quad (2.10.28)$$

In this expression, the matrices are time ordered and the matrix δ_{i_n} has all zeros except for the element, $\{i_n, i_n\}$, which is unity. The vector $p_{\pm}$ is $\delta^{(\pm)} \rho_{eq}/|\delta^{(\pm)} \rho_{eq}|$. The matrix $\mathbf{S} \delta_i \mathbf{S}^{-1}$ can be written as an outer product of two vectors, $\mathbf{S} \delta_i \mathbf{S}^{-1} = w_i v_i^T$ with the property that $w_i v_i^T w_i v_i^T = w_i v_i^T$ and $w_i v_i^T w_j v_j^T = 0$ for $j \neq i$. Starting from the trivial

identity $w_i v_i^T = \frac{a_{\pm i}^{\pm}}{a_{\pm i}^{\pm}} w_i v_i^T = (a_{\pm i}^{\pm})^{-1} w_i a_{\pm i}^{\pm} v_i^T$ with scalar $a_{\pm i}^{\pm} = 1^T \delta^{(\pm)} w_i v_i^T p_{\pm}$, which

follows from the definitions above, we get $w_i v_i^T = (a_{\pm i}^{\pm})^{-1} w_i 1^T \delta^{(\pm)} w_i v_i^T p_{\pm} v_i$. Since

$1^T \delta^{(\pm)} w_i$ and $v_i^T p_{\pm}$ are scalars, they commute resulting in the expression, $w_i v_i^T =$

$(a_{\pm i}^{\pm})^{-1} w_i v_i^T p_{\pm} 1^T \delta^{(\pm)} w_i v_i^T$. Substituting this expression for $\mathbf{S} \delta_{i_2} \mathbf{S}^{-1}$ in EQ. 2.10.28

gives a recursion relation for the values of

$$
\begin{aligned}
a_{\pm i_1, \ldots, \pm i_n}^{\pm} &= \\
1^T & \left\{ \prod_n \delta^{(\pm)} \mathbf{S} \delta_{i_n} \mathbf{S}^{-1} \right\} p_{\pm} = 1^T \left\{ \prod_n \delta^{(\pm)} w_{i_n} v_{i_n}^T \right\} p_{\pm} = \\
a_{\pm i_2}^{-1} & \left(1^T \left\{ \prod_{n=2} \delta^{(\pm)} w_{i_n} v_{i_n}^T \right\} p_{\pm} \right) \left(1^T \delta^{(\pm)} w_{i_2} v_{i_2}^T \delta^{(\pm)} w_{i_1} v_{i_1}^T p_{\pm} \right) \\
&= a_{\pm i_2, \ldots, \pm i_n}^{\pm} \frac{a_{\pm i_1, \pm i_2}^{\pm}}{a_{\pm i_2}^{\pm}},
\end{aligned}
\tag{2.10.29}
$$

which implies that determination of $a_{\pm i}^{\pm}$ and $a_{\pm i, \pm j}^{\pm}$ from fitting $C_{\pm\pm\pm}(t_1, t_2)$ to sums

of exponentials determines all of the higher order correlations. One can determine

the decay rates, λ_i, and the $a_{\pm i}^{\pm}$, from the single time correlation function, and use

the two time correlation function to fit the coefficients of the exponents, $a_{\pm i, \pm j}^{\pm}$ de-

termined from the one time correlation. If a limited number of degeneracies exist,

such as double degeneracy, a limited number of higher order moments can capture

this behavior. The result depends on the fact that eigenmodes never mix and does

not generalize to arbitrary waiting time distributions even if these distributions have

a single parameter.

As we discuss in previous work, the above analysis has an analogy to spectral

decomposition in quantum mechanics [35, 242, 243]. By writing,

$$1^T \left\{ \prod_n \delta^{(\pm)} \mathbf{S} \left[\mathbf{1} s_n + \mathbf{\Lambda} \right]^{-1} \mathbf{S}^{-1} \right\} p_\pm =$$

$$1^T \mathbf{S} \left\{ \prod_n \mathbf{S}^{-1} \delta^{(\pm)} \mathbf{S} \left[\mathbf{1} s_n + \mathbf{\Lambda} \right]^{-1} \right\} \mathbf{S}^{-1} p_\pm =$$

$$Tr \left(\mathbf{S}^{-1} p_\pm 1^T \mathbf{S} \left\{ \prod_n \mathbf{S}^{-1} \delta^{(\pm)} \mathbf{S} \left[\mathbf{1} s_n + \mathbf{\Lambda} \right]^{-1} \right\} \right), \qquad (2.10.30)$$

the n time "intensity" correlation maps into the spectral decomposition of the quantum time-correlation function with a density matrix $\rho \rightarrow \mathbf{S}^{-1} \delta^{(\pm)} p_\pm 1^T \mathbf{S}$, eigenfrequencies $i\omega_i \rightarrow -\lambda_i$, and transition matrices, $\mathbf{B} \rightarrow \mathbf{S}^{-1} \delta^{(\pm)} \mathbf{S}$.

2.11 Appendix: Application of information content to the generating function

For integer powers of m, the mth moment has the form,

$$\langle n^m(s) \rangle = \sum_{i=1}^n \frac{c_{m,i}}{s} 1^T \left[\mathbf{K} \left[\mathbf{1} s - \mathbf{\Gamma} - \mathbf{K} - \mathbf{K}_d \right]^{-1} \right]^i \rho_{ic}. \qquad (2.11.31)$$

In this expression, $c_{m,i}$ is a combinatorial factor, and we used $1^T \left[\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d \right] = 0$ to get the $\frac{1}{s}$ term. The terms with $i < m$ can be expressed as lower order moments. From these expressions it is apparent that each higher term contains new information in the form of the expression,

$$f_m(s) = \frac{1}{s} 1^T \left[\mathbf{K} \left[\mathbf{1} s - \mathbf{\Gamma} - \mathbf{K} - \mathbf{K}_d \right]^{-1} \right]^m \rho_{ic} \qquad (2.11.32)$$

91

We can use the same projection technique with projection operators of the form $\mathbf{S}\delta_i\mathbf{S}^{-1}$ to derive a recursive relationship between terms that contain m products of the matrices $\mathbf{K}$ and those with one and two $\mathbf{K}$ matrices. One should note that $\mathbf{S}$ is a similarity transform that diagonalizes the matrix $\mathbf{\Gamma} + \mathbf{K} + \mathbf{K}_d$, which is the same transformation used in the "intensity" correlation calculation, but the coefficients will be different. Using the projection operator we can write

$$f_m(s) = \frac{1}{s} \sum_{i_1,\dots,i_m} a_{i_1,\dots,i_m} \prod_{j=i_1,\dots,i_m} \frac{1}{s + \lambda_j}, \tag{2.11.33}$$

where the a's have a similar definition to the "intensity" correlation and "event" density, but they will be numerically different although the λ's will be the same as the "intensity" correlation function. The same recursion relation holds,

$$a_{i_1,\dots,i_m} = 1^T \left[\prod_{j=1}^{m} \mathbf{K}\mathbf{S}\delta_{i_m}\mathbf{S}^{-1} \right] \rho_{ic} = \frac{a_{i_1,i_2}}{a_{i_2}} a_{i_2,\dots,i_m}. \tag{2.11.34}$$

From the first moment we can find the a_i terms by fitting the transform to a sum of terms of the form $s^{-1}(s + \lambda_i)^{-1}$, but we cannot find the a_{i_1,i_2} terms from the simple second moments since we cannot distinguish ordering. It is important to note that the functional form does not give us any ordering of the eigenvalues, so we would fit $f_m(s)$ to

$$f_m(s) = \sum_{i_1 \geq \dots \geq i_m} b_{i_1,\dots,i_m} \prod_{j=i_1,\dots,i_m} \frac{1}{s + \lambda_j}, \tag{2.11.35}$$

with $b_{i_1,\dots,i_m} = \sum_{P\{i_1,\dots,i_m\}} a_{i_1,\dots,i_m}$, where $P\{\dots\}$ is the permutation operator. By

92

fitting the moments, $\langle n^m(s)\rangle$ or $f_m(s)$ as functions of the Laplace variables to the data, we can determine the $b_{i_1,\dots,i_m}$, and using the recursion relations, we write the $b_{i_1,\dots,i_m}$ in terms of products of a_{i_1,i_2} and $a_{i_1}^{-1}$. In principle, this process produces a set of equations that determines all of the a_{i_1,i_2} and the properties of the HMC, but the resulting equations are non-linear, which makes the existence of the solution difficult to ascertain. The higher order single time moments of the generating function can be used to extract the available information about these processes, but more of moments are needed than with time correlations. These moments also cannot be reliably extracted from experiments. The problem can be avoided by determining a two time renewal function, $\langle n(t_1)n(t_2)\rangle$, in which case the analysis becomes the same as Sec. 2.3 and only the first and second time moments are needed. Even using the two time renewal function requires one to either fit the function in Laplace space, or fit functions that contain convolutions, which results in additional polynomials in terms of time that must be fit. The result is theoretically interesting, but in practice, even for a model with two eigenvalues we need to accurately measure the Laplace transform of the 3rd moment and numerically solve a non-linear set of equations, which is not numerically feasible. The lack of certainty of the existence of a solution, as well as numerical difficulties in finding this solution, makes the use of the characteristic function difficult even when the eigenvalues are distinct.

2.12 Appendix: Simulation details

2.12.1 Generation of data

Four state model

The four state numerical example that we examine is depicted in Fig. 2-2 for the system are given by the kinetic matrix equation, Eq. 2.2.2. From the kinetic rates the steady state solution, ρ_{eq} can be determined. For each of the 25 molecular trajectories the initial state is randomly sampled from this steady state. After choosing the initial state, the time of a transition from this state to either the other connected states is drawn from an exponential distribution with characteristic time $(\gamma_s + k_s)^{-1}$, where $s = b1, b2, d1, d2$ denotes the current state. Once this time is chosen, the new state is chosen. The probability of making an unseen "bright-bright" or "dark-dark" transition is given by $\gamma_s / (\gamma_s + k_s)$, and one minus this quantity is the probability of making a visible transition. If a visible transition is observed, the transition time is recorded as part of the single molecule trajectory. After the new state is chosen, the simulation is continued until $300s$ has elapsed, but it is possible to incorporate photo-bleaching events. Although we assume that the transitions are sharp in this chapter, we can easily simulate systems with broader transition regions by adding the uncertainty in the transition time.

Fluctuating bottle-neck model

A similar simulation method is used to generate the data for the fluctuating bottle neck model. The equilibrium distribution is given by

$$(\rho_{eq})_{\pm} = \frac{k_{\mp}}{k_{\pm} + k_{\mp}} \frac{1}{\sqrt{2\pi k/D}} e^{-\frac{k}{2D}x^2},$$

(2.12.36)

This distribution can be easily sampled. Given the current position, x, we choose a small time step $\delta t \approx D \times 10^{-6}$. We calculate the probability that the system reacts $1 - e^{-k_{\pm}x^2\delta t}$. If the system reacts we record the time of the event. Then we propagate the system under normal Brownian motion in a harmonic oscillator to choose a new x value, so that our temporal resolution in $\approx 10^{-6}D$, which is much smaller than the kinetic rates $\sim 4.5D$

2.12.2   The Monte Carlo algorithm

Given the data generated in Sec. 2.12.1, we attempt to determine the relative likelihood of various possible parameters through a standard Metropolis Monte Carlo algorithm. More complex algorithms may be necessary, depending on the amount and complexity of the data, as well as the model being considered. The probability of a specific sequence is easily written in matrix notation as

$$P(D|M) = 1^T \left[\prod_i \mathbf{K} e^{(\mathbf{K}_d + \mathbf{\Gamma})t_i} dt_i \right] \delta_{\pm} \rho_{eq}$$

(2.12.37)

where the definitions of the matrices follow from above and the $\delta_\pm$ selects the initial condition [223]. The dt_i are the small widths of the time bins, which will generally be determined by instrument concerns as well as the photon emission rates from the "bright" and "dark" states. If dt_i is large relative to the kinetic constants, the appropriate binning needs to be considered. Note that the elements of $e^{(\mathbf{K}_d + \mathbf{\Gamma})t_i}$ must be evaluated through standard eigenmode analysis. Since we care about relative probabilities of different kinetic matrices, we can neglect the bin sizes, dt_i. For a long sequence the matrix products quickly go to zero. To prevent difficulties associated with zeroing out of the matrix products after each multiplication of $\mathbf{K}e^{(\mathbf{K}_d + \mathbf{\Gamma})t_i}$ we renormalize the resulting vector, $\rho_i = \eta_i \mathbf{K}e^{(\mathbf{K}_d + \mathbf{\Gamma})t_i}\rho_{i-1}$. The constant η_i is chosen so that $\sum_j (\rho_i)j = 1$. Here i denotes the number of matrix multiplications. The sum of the logs of all of the renormalization constants is the Bayesian score, $B = -\sum \log(\eta_i)$. Except for a constant correction of $\sum log(dt_i)$, the Bayesian score is the log of $P(D|M)$ which is necessary to apply Bayes formula. The Bayesian score plays the same role as the Boltzmann factor, $\beta E(\{t_i\})$ in statistical mechanics.

The Bayesian score allows us to perform Monte Carlo importance sampling to sample the posterior distribution of parameters that represent the data. We start the simulation at a random parameterization. At iteration j we have a set of parameters with Bayesian score B_j. We sample a new set of parameters, $j+1$, by multiplying the current parameters by e^r, where r is a uniform random number symmetric around the origin with step size Δ, i.e. $-\Delta < r < \Delta$. This step corresponds to a Markov Chain Monte Carlo importance sampling with an improper log-uniform a priori distribution, $P(M)dM \propto \frac{1}{M}dM$. This distribution is improper since it is not normalizable, but

as can be seen from the numerical examples in this chapter, the prior does not play a large role in the final distribution for these models. We choose this importance sampling since the kinetic constants must be positive and we cannot a priori set their magnitude so a log uniform distribution (Jeffrey's prior) is the most appropriate choice [6]. If we are imposing an ordering to preserve uniqueness, i.e. $k_{b1} > k_{b2}$, then after we choose the new values, we relabel the states to maintain this ordering. Once the new parameters, $j + 1$, are chosen we calculate the Bayesian score for these parameters. Following standard importance sampling in Monte Carlo if $B_{j+1} > B_j$ we accept the new set of parameters, else we conditionally accept the new parameters with probability $e^{B_{j+1} - B_j}$.

For a short sequence and a small number of states, we can readily evaluate the Bayesian score for every possible choice of parameters. For larger sets of states, more complicated models or data a more intelligent methods of sampling is required.

Chapter 3

Testing the properties of single molecules blinking sequences: detailed balance violations, renewal properties, and trajectory classifications

3.1 Introduction

Cellular processes involve highly connected biochemical networks that are operating at a steady state far away from equilibrium [151, 171, 78]. The steady state is maintained by constant influx of complex energy containing molecules (glucose) and efflux of

simpler molecules (H_2O and CO_2). Understanding the pathways that allow cells to use the energy influx to perform biological functions is the primary objective in the study of metabolism [151, 103]. The complexity of biochemical reaction networks in cells makes studies on the entire network difficult and motivates studying small segments such as single enzymatic reactions.

Enzymatic reactions may be the smallest component of cellular metabolism, but the enzyme itself is a fairly complex object with many degrees of freedom that can be affected by the energy influx of cellular metabolism. Although we generally associate simple Michaelis-Menten kinetics with enzymatic turnover [213], these additional degrees of freedom allow the enzyme to show complex behaviors with memory effects that can modify the molecule's properties in response to environmental or substrate fluctuations. These behaviors can be observed in single molecule experiments through fluorescent labeling [132]. The labels must be attached to many proteins of interest through chemical modification [98]. Avoiding changing the enzyme's reactivity through the chemical modification requires care, and the fluorescent probes will often be attached to the periphery of the protein so the coupling of the photophysics of the dye and the enzymatic reaction may not be obvious. This situation motivates assessment of the probe's coupling to the enzymatic reaction process. Since the enzyme is being pumped by the substrate reaction [95, 171, 78], a probe whose motions are coupled to the reaction are expected to violate detailed balance (DB) and determination of the nature of the DB violation in the probe's motion is the first step in understanding the coupling between the measured coordinate and the reaction process. The DB violation is a generic property of systems that can be studied without

determining a model for the exact kinetics or semi-Markov scheme. In this paper we discuss the properties of DB systems and different manifestations of DB violations. In a subsequent paper, we present robust statistical tests for these properties based on two dimensional histograms that do not depend on the exact model proposed [228].

Although generalizations are simple, we concentrate on the two manifold model, where the system exhibits hops between two intensities, labeled $+$ for *bright* and $-$ for *dark*, resulting in a blinking sequence. The duration of the $+$ or $-$ intensity has been referred to as *events* in literature [35, 243]. These hopping events are generally modeled by stochastic waiting time processes (semi-Markov) [79]. If the waiting times show correlations between adjacent transition times, additional hidden states are added. Including only a few states can greatly increase the complexity of the space of possible models so it is desirable to develop methods to get insights that do not require model optimization [30]. In the long time limit, the number of transitions from the $+$ manifold of states to $-$ manifold of states is necessarily equal to the number of transitions from $-$ to $+$, which results in a long time DB relation, $\langle K_{+-}(\rho_{eq})_- \rangle = \langle K_{-+}(\rho_{eq})_+ \rangle$, where ρ_{eq} is the equilibrium distribution, $K_{\pm\mp}$ is the flux operator, and $\langle \ldots \rangle$ is the average over the conformational coordinates (hidden states) [35, 243]. Monitoring each of the conformational substates may show more transitions from substate i to substate j than the reverse, resulting in a DB violation that does not appear on the mesoscopic scale of manifolds.

In this paper we examine the ability to detect and to characterize microscopic DB violations between the i and j substates in a system that appears to hold mesoscopic DB between the $+$ and $-$ manifolds. Although a substrate concentration dependence

is a direct indicator that the probed coordinate is coupled to the reaction, the concentration dependence does not always give insight into the underlying topology of the circulating loops leading to the DB violation. Using concentration dependences may also be difficult in some single molecule experiments. In ion channels the probe is the current from ions, which is also the substrate of the transport process. Changing the substrate concentration affects detection efficiency so the exact concentration dependence may not be easily determined. The substrate may also be a fairly ubiquitous compound, such as O_2, whose concentration is difficult to control, or one may desire testing for DB violations in the system to determine if the experiment is in equilibrium.

The classification of different mechanisms of DB violations without resorting to parametric models is the focus of this paper. By using the kinetic formalism discussed below, three signatures are derived, including a peak in the single waiting time distribution, a time reversibility violation, and a related lack of diagonal features in two dimensional joint waiting time distributions. The peak in the single waiting time distribution corresponds to a multiple step jumping process within a single manifold that indicates a flow through that manifold (See Fig. 3-1(b) for an example), while the time reversibility violation or a lack of a diagonal feature in the two dimensional joint waiting time distribution indicate a circulation loop that enters both manifolds twice (See Fig. 3-1(c)).
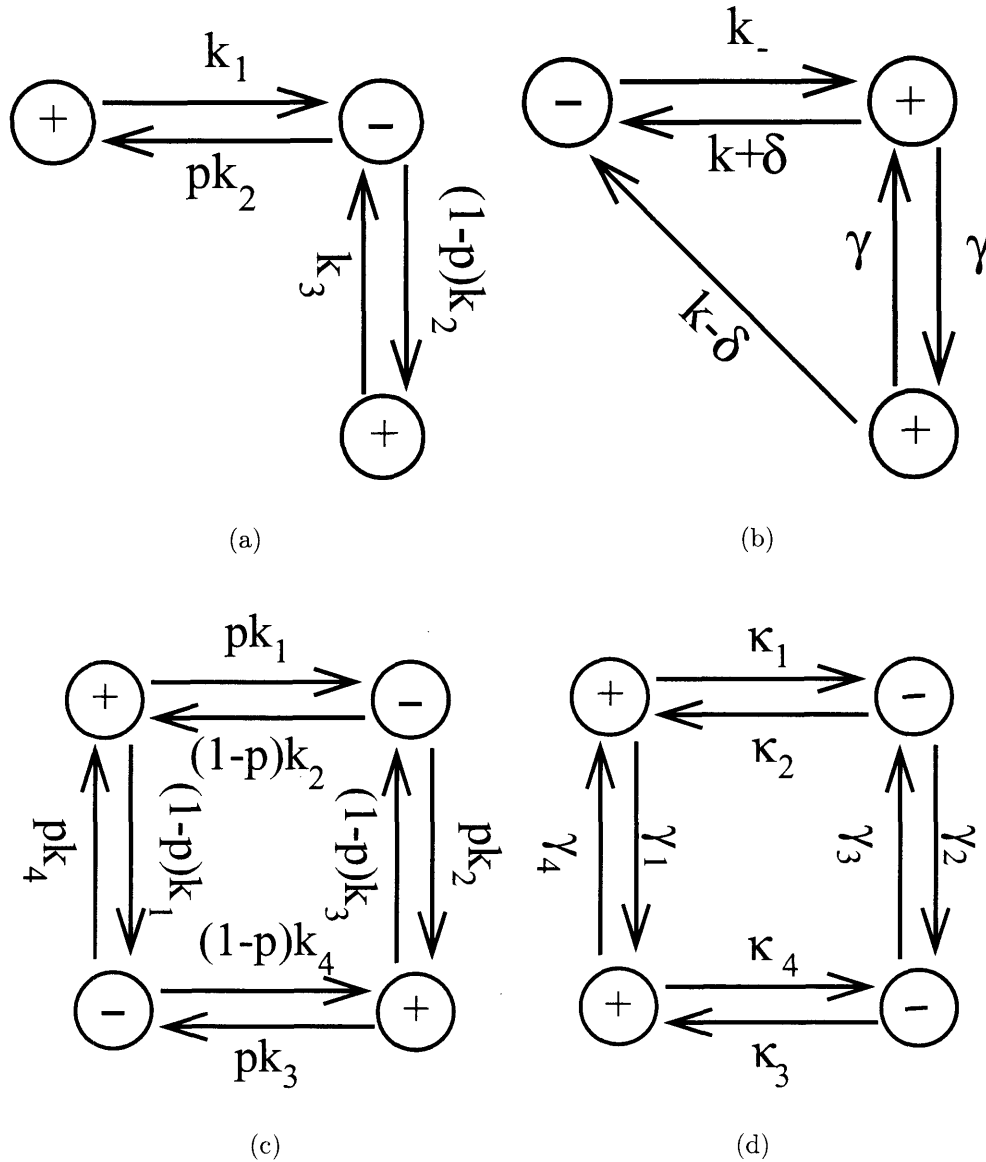
Figure 3-1: (a) A kinetic scheme that maps onto an alternating renewal process without violating DB. (b) A kinetic scheme that can be mapped onto (a) but violates DB. (c) A kinetic scheme that violates DB by alternating between waiting times. (d) The conformationally controlled kinetic scheme.

3.2 Kinetic Formalism

The statistical distribution of the times between *events* such as photon emission, switching between high fluorescent to low fluorescent state, and forming distinct contacts in a protein are generally the strongest indicators of the properties of single molecule systems [35, 243, 184, 138, 159]. If these *events* correspond to large scale rearrangements that cross high energy barriers, we expect the transitions follow a Poisson process. In order to discuss the features of a DB system (or DB violations), we need to model the system as a kinetic scheme [35, 243],

$$
\partial_t \begin{bmatrix} \rho_+ \\ \rho_- \end{bmatrix} = - \begin{bmatrix} K_{++} + \Gamma_{++} & -K_{+-} \\ -K_{-+} & K_{--} + \Gamma_{--} \end{bmatrix} \begin{bmatrix} \rho_+ \\ \rho_- \end{bmatrix}. \tag{3.2.1}
$$

Any semi-Markov process with a reasonable waiting time distribution is inverse Laplace transformable and can be expressed as a (possibly complex) kinetic scheme. To map this kinetic scheme into a semi-Markov process, one simply defines the waiting times for transitions from state j in the $\mp$ manifold to state i in the $\pm$ manifold as

$$
[G_{\mp \to \pm}(t)]_{ij} = \left[K_{\pm \mp} e^{-(K_{\mp \mp} + \Gamma_{\mp \mp})t} \right]_{ij} \tag{3.2.2}
$$

The determination of the exact waiting times from the data requires computationally expensive search algorithms for a complex topological space [30].

As an alternative, we determine features of DB violations in the one and two dimensional histograms of *events*. The histograms are determined from the probability

of n events [243],

$$P_{\pm \dots \pm}(t_1, \dots, t_n) = \sum \left[\prod_i K_{\pm \mp} exp-(K_{\mp \mp} + \Gamma_{\mp \mp})t_i \right] \frac{K_{\mp}(\rho_{eq})_{\mp}}{\sum K_{\mp}(\rho_{eq})_{\mp}} \qquad (3.2.3)$$

where the sums are carried over the final states. To get reduced probability matrices one integrates over the other *events*. If DB holds, the flux conservation law $(K_{-+})_{ij} [(\rho_{eq})_+]_j = (K_{+-})_{ji} [(\rho_{eq})_-]_i$ and $(\Gamma_{\pm\pm})_{ij} [(\rho_{eq})_\pm]_j = (\Gamma_{\pm\pm})_{ji} [(\rho_{eq})_\pm]_i$ implies that the kinetic matrix in Eq. 3.2.1 has a special form,

$$\partial_t \begin{bmatrix} \rho_+ \\ \rho_- \end{bmatrix} = - \begin{bmatrix} Q_{++} & -Q_{+-} \\ -Q_{+-}^T & Q_{--} \end{bmatrix} \begin{bmatrix} D_+ & 0 \\ 0 & D_- \end{bmatrix} \begin{bmatrix} \rho_+ \\ \rho_- \end{bmatrix} \qquad (3.2.4)$$

where $D_\pm = \delta_{ij} [(\rho_{eq})_\pm]_i^{-1}$ is the inverse of the equilibrium distribution, $Q_{\pm\pm}$ are symmetric matrices with positive eigenvalues (positive definite) that contain positive elements on the diagonals and negative elements off of the diagonal, and Q_{+-} is a matrix that contains only positive elements. The matrices satisfy the relationships, $\sum_i(Q_{++})_{ij} = \sum_i(Q_{+-})_{ji}$, and $\sum_i(Q_{--})_{ij} = \sum_i(Q_{+-})_{ij}$, $(Q_{++})_{ii} > \sum_j(Q_{+-})_{ij}$, $(Q_{--})_{ii} > \sum_j(Q_{+-})_{ji}$, and $\sum_{ij} Q_{++} = \sum_{ij} Q_{--} = \sum_{ij} Q_{\pm\mp}$.

Defining a symmetric operator with positive eigenvalues (positive definite), $H_{\pm\pm} = D_\pm^{\frac{1}{2}} Q_{\pm\pm} D_\pm^{\frac{1}{2}}$, allows us to rewrite $P_{+-\dots+-}$ for an even number of *events* starting from a + *event* as

$$P_{+-\dots+-}(t_1, \dots, t_{2n}) =$$

$$\sum \left[\prod_i Q_{+-} D_-^{\frac{1}{2}} e^{-H_{--}t_{2i}} D_-^{\frac{1}{2}} Q_{+-}^T D_+^{\frac{1}{2}} e^{-H_{++}t_{2i-1}} D_+^{\frac{1}{2}} \right] Q_{+-} \Big/ \sum_{ij} Q_{--}, \qquad (3.2.5)$$

105

where the matrix product over i is time ordered. Similar expressions exist for an odd number of events and for starting from the $-$ state. This representation will be the basis for much of the discussion below, but one may express $P_{\pm...}$ as a kernel of exponential components,

$$P_{\pm...\pm}(t_1,\ldots,t_n) = \int \prod \left[dk_i k_i e^{-k_i t_i}\right] P_{\pm...\pm}(k_1,\ldots k_n).\qquad(3.2.6)$$

The kernel, $P_{\pm...\pm}(k_1,\ldots k_n)$, need not be a probability distribution, but DB will impose conditions on this distribution (see Sec. 3.4). To simplify the presentation our focus is on $P(t_1,\ldots,t_n)$, but the results are independent of binning methods and can be easily applied to logarithmic binning, $P(lnt_1,\ldots,lnt_n)$. In fact, logarithmic binning is statistically robust for the determination of exponential decay times [196].

3.3   No sufficient condition for DB

Before continuing, we should note that all features of DB discussed below are necessary but not sufficient. If all states were distinguishable, DB violations are determined by examining the number of transitions between different states. When there is degeneracy in the states, a similarity transform can be performed that will not alter the statistics of *events*, but can change a DB obeying scheme into a DB violating scheme [223]. As a result, we can assess if DB is violated, but one can never establish if DB holds.

The statistical equivalence can be demonstrated by the schemes in Fig. 3-1(a)

and (b). Each model only has one $-$ state and the statistics of the system will be an alternating renewal process. If the $+$ and $-$ single waiting time distributions are identical between the two models, all statistics will also be identical. One model is a linear chain that must follow DB and has waiting time distributions,

$$P_- = k_2 e^{-k_2 t}; \qquad P_+ = pk_1 e^{-k_1 t} + (1-p)k_3 e^{-k_3 t}. \tag{3.3.7}$$

The other model has a one way loop that always violates DB and has distributions

$$P_- = k_- e^{-k_- t}; \qquad P_+ = q\kappa_- e^{-\kappa_- t} + (1-q)\kappa_+ e^{-\kappa_+ t}, \tag{3.3.8}$$

where $\kappa_\pm = k + \gamma \pm \sqrt{\delta^2 + \gamma^2}$, and

$$q = \frac{1}{2\sqrt{\delta^2 + \gamma^2}} \frac{\gamma(k - \delta) + (k + \delta)(\sqrt{\delta^2 + \gamma^2} - \delta)}{\gamma + \delta - \sqrt{\delta^2 + \gamma^2}} \tag{3.3.9}$$

For $\delta > 0$, we can choose k_1, k_2, and p such that $\kappa_- = k_1$, $\kappa_+ = k_3$ and $p = q$. The result is identical waiting time distributions and two indistinguishable models. From this example, it should be apparent that degeneracy eliminates the possibility of determining if DB holds, and one can only determine if DB is consistent with the statistics of the system.

3.4 Peaks in Single waiting time

The above arguments and examples show that there are no signatures of DB, but signatures of DB violations in the one and two dimensional histograms of *events* may still exist. To give some insight into possible manifestation of DB violations, we revisit the models in Fig. 3-1(a) and (b). The mapping from k_1, k_3, and p to γ, δ, and k is not possible for $\delta < 0$. The prefactor q is always positive, but it becomes greater than unity for $\delta < 0$, which implies that $1 - q$ is negative. The DB model in Fig. 3-1(a) cannot have a negative $(1 - p)$ value and cannot be mapped onto the kinetic scheme in Fig. 3-1(b). We want to examine if $\delta < 0$ creates a feature that corresponds to the DB violation that prevents the mapping between the two models.

Assuming that a DB violating feature increases with decreasing $\delta < 0$, we expect the feature to be the strongest when $\delta = -k$. The process continues to be a simple time reversible alternating renewal process, but the waiting time distribution for the + state becomes a two step process,

$$P_+(t) = \frac{\kappa_- \kappa_+}{\kappa_+ - \kappa_-} \left(e^{-\kappa_- t} - e^{-\kappa_+ t} \right) , \tag{3.4.10}$$

where $\kappa_\pm = k + \gamma \pm \sqrt{k^2 + \gamma^2}$ follows from above. The waiting time distribution is the convolution of two exponential processes and the function has a zero at $t = 0$ and a peak at $t = \frac{ln(\kappa_+) - ln(\kappa_-)}{\kappa_+ - \kappa_-}$. Intuitively, we expect this peak to correspond to a DB violation because the waiting time corresponds to a flow $(- \rightarrow +_1 \rightarrow +_2 \rightarrow - \rightarrow \ldots)$. In other words, the system prefers to hop into the + manifold through one state then

108

hop to a second + state before leaving the manifold, resulting in a circulation through the manifold.

To quantify this intuition, we note that by defining a vector w with coefficients

$$w_i = \sum_j \left(D^{\frac{1}{2}} Q_{+-} \right)_{ij} \Big/ \sqrt{\sum Q_{--}}, \tag{3.4.11}$$

we can write

$$P_+(t) = \sum K_{-+} e^{-(K_{++}+\Gamma_{++})t} \frac{K_{+-}(\rho_{eq})_-}{\sum K_{+-}(\rho_{eq})_-} = w^T e^{-H_{++}t} w \tag{3.4.12}$$

Differentiating with respect to t gives

$$\frac{dP_+(t)}{dt} = -w^T H_{++} e^{-H_{++}t} w. \tag{3.4.13}$$

Since H_{++} is positive definite, this expression must be strictly negative and a peak is not possible. Similarly, the second derivative must be strictly positive (convex). The lack of a peak in the single waiting time distribution imposes a restriction on the semi-Markov processes modeling equilibrium behavior.

A much stronger statement can be made about the single time waiting time distribution. Since H_{++} is a symmetric operator with positive eigenvalues, it has a complete eigenvalue decomposition and the waiting time distribution can be written

in the form of Eq. 3.2.6

$$P_+(t) = \sum_j a_j k_j e^{-k_j t} = \int dk k e^{-kt} P_+(k) \qquad (3.4.14)$$

where k_j are the eigenvalues of H_{++} and $a_j = k_j^{-1}(w \cdot V_j)^2$ is a weighted square of the dot product of the corresponding eigenvector onto w. The resulting kernel, $P_+(k) = \sum a_j \delta(k - k_j)$, is a probability density $(P_+(k) \geq 0, \int P_+(k) = 1)$ and the variance in the waiting time becomes

$$\langle \Delta t^2 \rangle = \langle k^{-1} \rangle^2 + 2\langle \Delta k^{-2} \rangle, \qquad (3.4.15)$$

where $\langle \Delta k^{-2} \rangle$ is the variance in k^{-1} that is calculated from $P_+(k)$. Eq. 3.4.15 demonstrates that the variance in the single waiting time distribution must be wider than the square of the mean waiting time for the DB schemes, $\langle t \rangle^2 = \langle k^{-1} \rangle^2$.

Coordinating parallel stochastic processes that have wide distributions of timescales would be difficult, so using energy to drive the reaction on a narrower timescale would be desirable. This driven reaction would result in a DB violation. As a simple example consider the kinetic scheme shown in Fig. 3-2(a), where a circular chain connects a block of $n +$ states with a block of $n -$ states that is driven clockwise with a transition rate of $k = k_0 n$ (mean waiting time independent of n). The system is a simple alternating renewal process with waiting time distributions of the form $P_\pm(t) = \frac{(k_0 n)^n t^{n-1}}{n!} exp(-k_0 n t)$, which is not allowed in DB schemes. Increasing n results in peaks that reduce the variance below the variance of an exponential process,

as shown in Fig. 3-2(b).

Assuming a small backwards reaction rate, k_b allows us to define a free energy loss per step, $\Delta G = k_b Tln(k/k_b)$. The total free energy dissipated in the cycle is $2n\Delta G$. A stronger peak corresponds to more steps, which implies a larger free energy loss. Achieving a narrow variance in the waiting time (i.e. be highly peaked) requires pumping by external energy sources. The narrow waiting time distributions also result in the peak in the power spectrum of the correlation function $C(t)$ explored by Qian [171, 78] and shown in Fig. 3-2(c). This power spectrum peak can be interpreted as a clocking mechanism in the enzyme.

3.5 Time reversibility

The peak in the waiting time distribution is the only one dimensional indicator of a DB violation. We expect many possible DB violating schemes to lack a peak in the waiting time, such as the scheme in Fig. 3-1(c), so it is necessary to look at higher dimensional measures for DB violations. Avoiding the difficulties of determining an underlying model restricts the applications of the analysis to the binning of data so two dimensional measures are the most useful for developing insight into the nature of DB violations. A previously studied two dimensional indicator of DB violations is time reversibility in the joint distribution of events [196]. The time reversibility can be understood by re-examining the expression for *n events* in Eq. 3.2.5. This
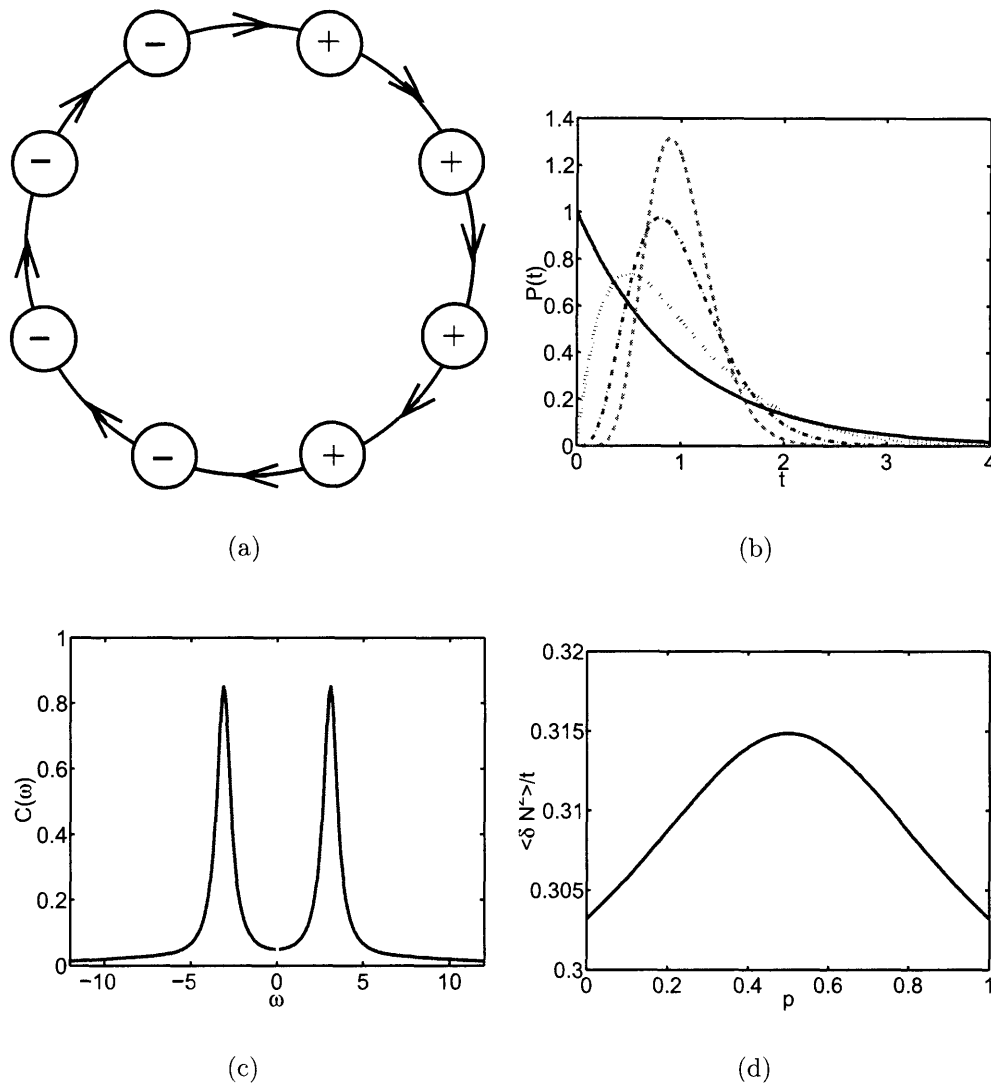
Figure 3-2: (a) A kinetic scheme with a narrow waiting time distribution. (b) The waiting time distribution, $\frac{k^n t^{n-1}}{n!}e^{-kt}$, for $n = 1$ (solid), 2 (dotted), 5 (dot-dash), and 10 (dashed) sequential processes with rates $k = k_0 n$ with $k_0 = 1$ (to ensure same mean). The peak is an indication of a DB violation. (c) The power-spectrum of the correlation function of the scheme shown in Fig. 3-2(a) for $n = 10$. (d) The long time rate of growth of the variance in the number of turnovers for the scheme in Fig. 3-1(c), with $k_1 = k_2 = k_4 = 1$ and $k_3 = 2$ as a function of p. The expected number of turnovers is $\frac{4}{7}t$ (see text) and the maximum variance occurs for $p = 1/2$ and corresponds to a renewal process obeying DB.

expression is a scalar so it is equal to its transpose, which is

$$P_{-+\ldots-+}(t_n, \ldots t_1) =$$

$$\sum \left[\prod_i Q^T_{+-} D^{\frac{1}{2}}_+ e^{-H_{++} t_{2i-1}} D^{\frac{1}{2}}_+ Q_{+-} D^{\frac{1}{2}}_- e^{-H_{--} t_{2i}} D^{\frac{1}{2}}_- \right] Q^T_{+-} \bigg/ \sum_{ij} Q_{--}. \quad (3.5.16)$$

Noting that $\sum_{ij} Q_{--} = \sum_{ij} Q_{++}$, $H_{\pm\pm}$ is symmetric, and the time ordering of the matrix product is reversed, we see that this expression is equivalent to observing the reverse of the sequence, time reversibility (i.e. $P_{\pm\ldots\mp}(t_1, \ldots, t_n) = P_{\mp\ldots\pm}(t_n, \ldots, t_1)$ and $P_{\pm\ldots\pm}(t_1, \ldots, t_n) = P_{\pm\ldots\pm}(t_n, \ldots, t_1)$). This time reversibility relation has been used to explore DB violations in ion channels [196]. As discussed elsewhere, one can perform statistically robust tests on 2-D contours to assess if time reversibility is violated [228].

The time reversal violation affects our inference about the waiting time of an *event* given that we have information about a different *event*. Suppose that we measured a + *event* and wanted to infer the properties of the − *events* immediately before and after this + *event*. If time reversal symmetry holds we would assume that the preceding and succeeding − *events* have identical statistics, but a time reversal symmetry violation implies that the statistics are different. The time of an *event* is a stochastic signature that can be used to infer which transition has occurred, $j \to i$, where we identify the transition through the index of the Green's function expression in Eq. 3.2.2. This definition defines the transition based on the initial substate entered in each manifold. An extreme example has $[G_{\pm\mp}(t)]_{ij} = \delta(t - t_{ij})$ for $t_{ij} \neq t_{k,l}$ so that the transition time uniquely identifies the transition. A time reversibility violation implies that the $j \to i$

transitions and the $i \to j$ transitions occur with different frequencies. This violation is the result of a circulation around a loop that enters the $+$ and $-$ manifolds at least twice, $(+ \to - \to + \to -)$.

A simple realization of this time reversibility violation has four states $+_1, +_2, -_1,$ and $-_2$ with each $+$ state only connected to $-$ states and vice versa (Fig. 3-1(c)). Topologies with no $+ \leftrightarrow +$ or $- \leftrightarrow -$ connections have been referred to as the Bauer-Kienker uncoupled (BKU) canonical form [30]. All models can be mapped into a canonical form with possibly unphysical rates (negative or complex), but the form reduces degeneracy in the number of identical schemes and the number of parameters, which simplifies model optimization. An arbitrary model will have $(N_+ + N_-)(N_+ + N_- - 1)$ kinetic rates connecting the $N_\pm$ substates in the $\pm$ manifolds, which should be compared with the $2N_+N_-$ rates of the BKU form. Studies in the literature have found that many ion-channel models have physical topologies that resemble this BKU form [30]. Another canonical form, whose topology has appeared in several physical models of ion channels, is the *manifest interconductance rank form*. This canonical form has several attributes, the most important property being that every $+$ state is connected to at most one $-$ state and vice versa. The conformationally controlled kinetic schemes, such as the four state model in Fig. 3-1(d) is an example of this canonical form [30].

Returning to the simple model in Fig. 3-1(c). Each state has a simple exponential waiting time distribution, $P_i(t) = k_i e^{-k_i t}$, and the conformational dynamics prefer to proceed in a circular sequence $+_1 \to -_1 \to +_2 \to -_2 \to +_1 \to \ldots$, with p representing the clockwise bias. The circular flipping process results in the detected DB violation

114

with the two *event* waiting times being

$$P_{+-}(t_1, t_2) = \frac{1}{2}p\left[k_2 e^{-k_2 t_2} k_1 e^{-k_1 t_1}\right] + \frac{1}{2}(1-p)\left[k_4 e^{-k_4 t_2} k_1 e^{-k_1 t_1}\right]$$
$$+ \frac{1}{2}p\left[k_4 e^{-k_4 t_2} k_3 e^{-k_3 t_1}\right] + \frac{1}{2}(1-p)\left[k_2 e^{-k_2 t_2} k_3 e^{-k_3 t_1}\right], \qquad (3.5.17)$$

and

$$P_{-+}(t_2, t_1) = \frac{1}{2}(1-p)\left[k_2 e^{-k_2 t_2} k_1 e^{-k_1 t_1}\right] + \frac{1}{2}p\left[k_4 e^{-k_4 t_2} k_1 e^{-k_1 t_1}\right]$$
$$+ \frac{1}{2}(1-p)\left[k_4 e^{-k_4 t_2} k_3 e^{-k_3 t_1}\right] + \frac{1}{2}p\left[k_2 e^{-k_2 t_2} k_3 e^{-k_3 t_1}\right], \qquad (3.5.18)$$

If $k_1 \ll k_3$, $k_2 \ll k_4$ and $p \ll 1$, we expect long + events will be followed by short − events, but long − events will be followed by long + events. If a short + *event* is observed, we infer that the state is probably the k_3 state and expect the preceding − *event* to be long and the succeeding − *event* to be short. As a result $P_{+-}(t_1, t_2)$ has a large contribution from the off diagonal, whereas $P_{-+}(t_2, t_1)$ does not. (See Fig. 3-3(a) and (b)).

3.6   A diagonal feature of DB systems

Degeneracy in some of the kinetic rates of Eq. 3.2.1 can lead to systems where the time reversibility violation and peaks in the waiting time distribution do not appear, despite the system violating DB. A simple example of this degeneracy is the existence of a path independent single rate determining step, such as transport limited reac-
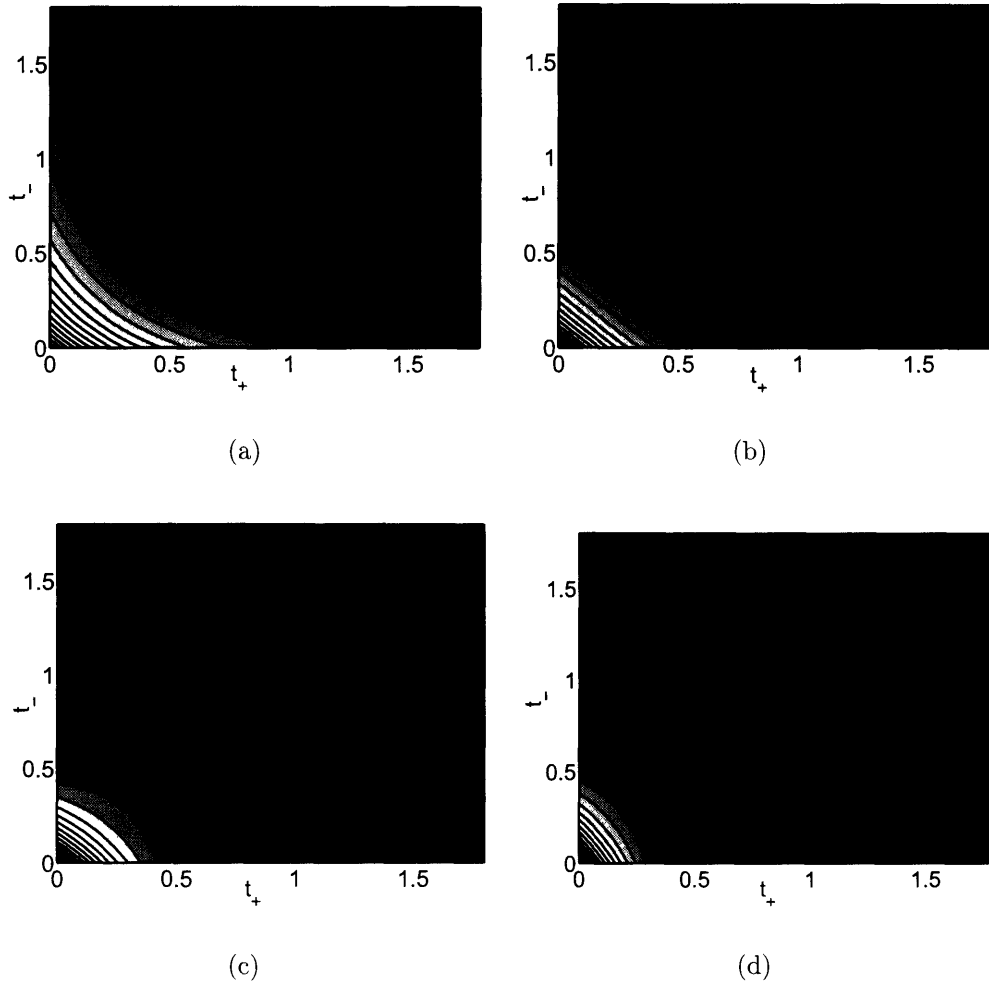
Figure 3-3: Comparison of $P_{+-}(t_+, t_-)$ in (a) with $P_{-+}(t_-, t_+)$ in (b) to show a time reversibility violation for the scheme in Fig. 3-1(c) with $4k_1 = 4k_2 = k_3 = k_4 = 4$, and $p = 0.1$. The maximum occurs at $t_1 = t_2 = 0$ and monotonically decays. (c) $P_{-+}(t_-, t_+) - P_{+-}(t_+, t_-)$ shows that the difference is not zero, which indicates a time reversibility violation. (d) $P_{-+}(t_-, t_+) - P_{+-}(t_+, t_-)$ with $4k_1 = 3k_2 = 2k_3 = k_4 = 4$ to demonstrate that the contours do not have to be symmetric.

tions. The scheme in Fig. 3-1(c) with $k_2 = k_4$ has this property. The $-$ waiting time

is mono-exponential so the $-$ waiting times cannot be used to detect DB violations,

and we must rely on examining the histogram of two $+$ events. For simplicity we

consider two $+$ events separated by a single $-$ event, $P_{++}(t_1, t_2)$, but generalizations

to greater separations is simple. If DB holds, this waiting time distribution can be

written as

$$P_{++}(t_1, t_2) = w^T e^{-H_{++}t_2} R_{++} e^{-H_{++}t_1} w \qquad (3.6.19)$$

where

$$R_{++} = D_+^{\frac{1}{2}} Q_{+-} Q_{--}^{-1} Q_{+-}^T D_+^{\frac{1}{2}} \qquad (3.6.20)$$

is a symmetric operator with positive eigenvalues (positive definite). From the triangle

inequality with the positive definite operator R_{++}, it is simple to show

$$[P_{++}(t_1, t_2)]^2 \leq P_{++}(t_1, t_1) P_{++}(t_2, t_2). \qquad (3.6.21)$$

As a result, the maximum of the probability distribution must appear on the diago-

nal, $t_1 = t_2$ and the function must be concave around the diagonal, i.e. for fixed t,

$P_{++}(t - \delta, t + \delta)$ must have a local maximum at $\delta = 0$. A possible DB violation P_{++} is

the contour in Fig. 3-4(a), where the wings in the waiting time distribution violate this

triangle inequality. The difference between the upper bound of the triangle inequal-

ity and the true distribution, $\Delta P_{++}(t_1, t_2) = P_{++}(t_1, t_2) - \sqrt{P_{++}(t_1, t_1) P_{++}(t_2, t_2)}$ is

plotted in Fig. 3-4(b). This kinetic scheme is time reversible and does not possess a
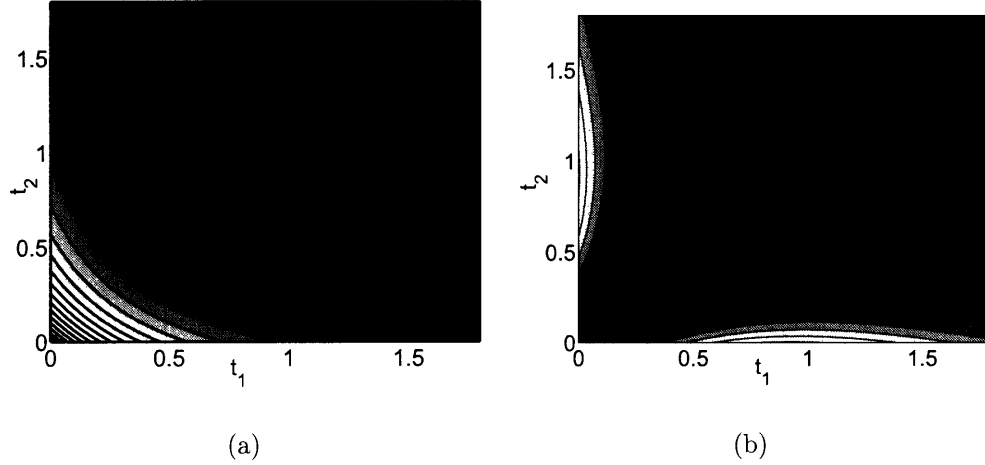
Figure 3-4: (a) $P_{++}(t_1, t_2)$ for the scheme in Fig. 3-1(c) with $k_1 = k_2 = k_4 = 1$, $k_3 = 4$, and $p = 0.05$. The large k_3 was chosen to accentuate the strong off diagonal features. (b) $\Delta P_{++}(t_1, t_2) = P_{++}(t_1, t_2) - \sqrt{P_{++}(t_1, t_1)P_{++}(t_2, t_2)}$ for the contour in (a). The positive peaks along the lines $t_1 = 0$ and $t_2 = 0$ and the zero along the line $t_1 = t_2$ are strong signatures of a DB violation.

peak in the single waiting time spectrum so the diagonal dominance test is the only one that can detect the violation. We should note that schemes that strongly violate DB have been used to create strong diagonal features so a strong diagonal feature does not infer even approximate consistency with DB [135].

Integrating Eq. 3.6.19 over a small interval, $\tilde{t}_i = t_i \pm \delta$ gives

$$\int d\tilde{t}_1 d\tilde{t}_2 P_{++}(\tilde{t}_1, \tilde{t}_2) = w^T H_{++}^{-1} \left(e^{-H_{++}(t_2-\delta)} - e^{-H_{++}(t_2+\delta)} \right)$$

$$R_{++} H_{++} \left(e^{-H_{++}(t_1-\delta)} - e^{-H_{++}(t_1+\delta)} \right) w \qquad (3.6.22)$$

Since H_{++} is symmetric and R_{++} is positive definite, the triangle inequality can also be applied to this expression. Histogramming data approximates this quantity so the relation can be applied to histograms. (If P_{ij} is a histogram for two + events, then $P_{ij}^2 \leq P_{ii} P_{jj}$) We discuss statistically testing these types of histograms in [228].

Similar to the DB violation, a violation of the triangle inequality follows from a circulation loop that enters both manifolds of states at least twice, but the circulation is being inferred from the ++ joint waiting time distribution since the − manifold is degenerate and does not give any information.

The standard renewal test $\delta P_{++}(t_1, t_2) = P_{++}(t_1, t_2) - P_+(t_1)P_+(t_2)$ corresponds to replacing R_{++} in the above expression with $(S_{++})_{ij} = (R_{++})_{ij} - \frac{\sum_{kl}(Q_{+-})_{ik}(Q_{+-})_{jl}}{\sum_{ij}Q_{--}}$ [35, 243]. This operator is symmetric and has positive eigenvalues except for a single zero eigenvalues (positive semi-definite operator) so the diagonal of the difference function must be positive and the maximum must occur on the diagonal. If there is a zero on the diagonal, as is the case for the two channel model considered in [35, 243], two lines of zero difference parallel to the $t_1 = 0$ and $t_2 = 0$ axes pass through that point.

The difference function also shows that

$$Corr(n) = \int dt_1 dt_2 t_1 t_2 \delta P_{++}(t_1, t_2) = \frac{\langle \delta t_1 \delta t_n \rangle}{\langle \delta t^2 \rangle} = \frac{w^T H_{++}^{-2} S_{++} H_{++}^{-2} w}{2 w^T H_{++}^{-3} w} \geq 0. \quad (3.6.23)$$

This expression implies that adjacent + events are positively correlated in DB obeying processes and results in a wider distribution of the number of turn-overs over a fixed period of time compared to the expectations from examining the single waiting time distributions, $P_+(t)$ [223]. A DB violation can make $\langle \delta t_1 \delta t_n \rangle \leq 0$ and narrow the distribution of the number of turnovers over a period of time.

A simple example of DB violations narrowing distributions in the number of turnovers and causing negative correlations in the duration of adjacent event is the scheme in Fig. 3-1(c) with $k_1 = k_2 = k_4 = 1$ and $k_2 = 2$. For any value of p, the

single waiting time distributions are $P_-(t) = e^{-t}$ and $P_+(t) = \frac{1}{2}[e^{-t} + 2e^{-2t}]$. In this case, $\langle t \rangle = \frac{3}{4}$ but $\langle \delta t_1 \delta t_2 \rangle = -\frac{(1-2p)^2}{8}$. In the long time limit, we expect the number of turnovers to grow as $\langle N(t) \rangle = \frac{4}{7}t$ since we need to include the $-$ waiting time. The variance in the number of turnovers for this model is $\langle (\delta N(t))^2 \rangle = \frac{4t}{343}\left[25 + \frac{1}{1-2p(1-p)}\right]$. As shown in Fig. 3-2(d), the variance in this distribution reaches a maximum at $p = \frac{1}{2}$, when the model is a renewal process obeying DB. As the DB violation increases $p \to 0, 1$, the variance monotonically decreases since slow events follow fast events.

3.7 Conclusion

The one and two dimensional features discussed above appear to be the exhaustive set of possible tests for detectable DB violations in the one and two dimensional histograms of *events* without resorting to determining the underlying model and examining the entire sequence. Testing for these features establishes necessary but not sufficient conditions for DB, since sufficient tests do not exist. A DB violating kinetic scheme may exhibit none, some, or all of these features, such as the three substate model in Fig. 3-1(b) that reduces to a time reversible alternating renewal process that lacks two dimensional features, or the four state model in Fig. 3-1(c) that lacks a peak in the single waiting time distribution, but shows two dimensional features.

The peak in the single waiting time distribution is caused by the system's tendency to make several transitions within either the $+$ or $-$ manifold before leaving the manifold and corresponds to a current through the manifold. Cyclically performing

many steps increases the DB violation, but allows the system to perform its function on a definite timescale. A time reversibility violation in comparing $P_{\pm\mp}$ to $P_{\mp\pm}$ or a lack of peaks on the diagonal of the $P_{\pm\pm}$ indicate a circulation loop that travels through both manifolds at least twice. These violations allow the system to alternate responses and can result in anti-correlated *event* times and narrowing the distribution of number of turnovers. These DB violations allow the protein to retain memory of past *events* and modify its response. We present statistical methods to determine these types of DB violations elsewhere [228].

3.8 Testing for renewal and detailed balance violations in single molecule blinking processes

In this section, we examine statistical tests of properties that appear in the one and two dimensional histograms of *events*. The tests capture important aspects of the system, including the non-Markovian/non-renewal nature of the system, violations of detailed balance, and similarity in behaviors of molecules in different experimental conditions [226]. The renewal behavior determines the existence of parallel paths, while the detailed balance violations give insight into circulation in the underlying topology, including the connectivities that lead to the circulation [226]. Concentration dependences indicate the role of cofactors in a single molecules function, such as metal ions in ribozyme folding and the energy transfer between the substrate and the macromolecule [171, 227].

The major difficulty with the two *event* waiting time distribution is the need to histogram the *events* (resulting in a histogram h_{ij}). In the infinite data and infinitesimal bin size limit, the traditional analysis discussed in the single molecule literature will suffice. However, these methods are sensitive to noise and cannot be easily implemented for finite sets of data. The implementation difficulties have several sources. One source is the binning methods, which are linear in many proposed applications, but should be logarithmic for examining exponentially distributed data [196]. In this paper, data will be presented logarithmically histogrammed event though the discussion will present linear probability distributions. Another issue is the scatter that is present in the histogram. The scatter in a histogram should be approximately Poisson distributed with the variance in the number of *events* in a bin being approximately equal to the number of *events* in the bin. As a result, the scatter will not be uniformly distributed throughout the histogram, which can cause misinterpretation in features that are measured from the differences between histograms. The nearly Poisson nature of the data scatter causes an additional problem for low count bins, where the deviations are not Gaussian. Avoiding these difficulties is the motivation behind introducing more rigorous statistical methods of assessing the relevance of apparent features. An important result is the establishment of the number of measurements necessary to elucidate the existence of features in the histograms properties and distinguish different models. These numbers should be used as a guideline for the types of models to compare to the data.

3.9 Testing models for histogrammed data

In order to avoid the combinatorially complex problem of finding the best model for the data, we construct histograms of *events* from the data and then test generic features of the histograms instead of the underling model that produced these features. The histograms, h_i or h_{ij}, include the single *events* (single sojourn time) and pairs of *events*, such as adjacent $+$ and $-$ sojourn times or two $+$ sojourn times separated by one $-$ sojourn. The objective is to test possible models for the histograms. The best fitting model will always be the histogram itself $P_{ij} = \frac{1}{N} h_{ij}$ with $N = \sum_{ij} h_{ij}$. The main issue is the establishment of other simpler models, $\tilde{P}_{ij}$ that are consistent with the data. For a model $\tilde{P}_{ij}$, the probability of getting the histogram h_{ij} is $P(h_{ij}|\tilde{P}_{ij}) = \prod_{ij} (\tilde{P}_{ij})^{h_{ij}}$. Taking the log of this probability gives

$$ln P(h_{ij}|\tilde{P}_{ij}) = \sum_{ij} h_{ij} ln(\tilde{P}_{ij}). \tag{3.9.24}$$

Since each *event* is a random variable, $\sum_{ij} h_{ij} ln(\tilde{P}_{ij})$ is the result of a sum of random variables that converges to a Gaussian distribution in the large N limit. As a result, the difference in the log probability of the histogram, $P_{ij} = \frac{1}{N} h_{ij}$, and $\tilde{P}_{ij}$, is a natural method of comparing the model to the data and assess if the model is adequate [133]. For the histogram h_{ij} the difference in the logarithms is

$$\sum_{ij} h_{ij}[ln(P_{ij}) - ln(\tilde{P}_{ij})] = \sum_{ij} h_{ij} ln(P_{ij}/\tilde{P}_{ij}). \tag{3.9.25}$$

Since $P_{ij} = \frac{1}{N}h_{ij}$, we are left with N times the Kullback-Liebler metric (KL) [133],

$$NI_{P|\tilde{P}} = N\sum_{ij}(dI_{P|\tilde{P}})_{ij} = N\sum_{ij}P_{ij}ln(P_{ij}/\tilde{P}_{ij}).\qquad(3.9.26)$$

The log-likelihoods should be compared against the expected variances,

$$N\delta I^2_{P|\tilde{P}} = N\left[\sum_{ij}P_{ij}ln(P_{ij})^2 - \left(\sum_{ij}P_{ij}ln(P_{ij})\right)^2 + \right.$$
$$\left.\sum_{ij}P_{ij}ln(\tilde{P}_{ij})^2 - \left(\sum_{ij}P_{ij}ln(\tilde{P}_{ij})\right)^2\right].\qquad(3.9.27)$$

If the KL metric is small compared to the variance estimate, $\tilde{P}$ is an adequate model for the data. It is simple to account for correlations in the data by modifying the variance. This testing method penalizes using the histogram, P_{ij}, itself by a factor that scales as $\sqrt{N}$ without regard to the number of parameters, whereas other methods penalize by factors of $ln(N)$ or unity with a parameter dependent prefactor [125]. The preferred method should depend on both the number of data points and number of parameters.

If $P_{ij} \approx \tilde{P}_{ij}$, then a Taylor expansion gives

$$I_{ij} \approx \sum_{ij}\frac{1}{2}\frac{(P_{ij} - \tilde{P}_{ij})^2}{P_{ij}} = \sum_{ij}\frac{1}{2}\frac{\Delta P^2_{ij}}{P_{ij}}\qquad(3.9.28)$$

since the linear term averages to zero. This result follows from the previous discussion about the approximately Poisson nature of the variations in the data. The variance in the number of *events* in each bin is equal to the number of *events* in a bin ($\sigma^2_{ij} = NP_{ij}$),

and this procedure reduces to least squares analysis in the large data limit [196]. The expression demonstrates that the relative instead of absolute values of deviations of $\tilde{P}_{ij}$ from P_{ij} is the important quantity. Large absolute deviations appear in regions with larger numbers of *events* and re-weighting the residual deviations is necessary.

3.10   Testing renewal behavior

As the simplest example of testing models, consider trying to determine if the histogram, h_{ij} corresponds to a simple renewal or alternating renewal process. Renewal or alternating renewal processes assume that the sojourn time of *events* are independent, which implies that the process does not exhibit multiple paths connecting the + and − manifolds [82]. To test the renewal property, the data, $P_{ij} = \frac{1}{N}h_{ij}$, must be compared with the best fitting model for independent *events*

$$\tilde{P}_{ij} = P_i P_j = \frac{1}{N^2} \left[\sum_j h_{ij} \right] \left[\sum_i h_{ij} \right]. \qquad (3.10.29)$$

The resulting Kullback-Liebler metric for comparing P and $\tilde{P}$ is $I_{P|\tilde{P}} = \sum_{ij} P_{ij} ln(P_{ij}/(P_i P_j))$, which is sometimes called the mutual information between the variable i and j [156]. The mutual information is always positive since P_{ij} is a better fit. To determine if the difference in fits is significant, we compare $N I_{P|\tilde{P}}$ with $N\delta I^2_{P|\tilde{P}}$. If zero falls within the 95% confidence interval (2 standard deviations), the *events* are considered independent and we adopt $\tilde{P}_{ij} = P_i P_j$.

We explore application of this test to the model in Fig. 3-5(a) as a function of p

and K for $k_1 = k_2 = K^{-1}k_3 = K^{-1}k_4 = 1$. In this case the two *event* function for a $+$ sojourn followed by a $-$ sojourn is,

$$P_{+-}(t_1, t_2) = \frac{1}{2}e^{-t_1}\left[pe^{-t_2} + (1-p)Ke^{-Kt_2}\right]$$
$$+\frac{1}{2}Ke^{-Kt_1}\left[pKe^{-Kt_2} + (1-p)e^{-t_2}\right] \tag{3.10.30}$$

The best fitting renewal process for P_{+-} has $p = \frac{1}{2}$. A comparison of the P_{+-} (in log time form) to the renewal prediction in Fig. 3-6(a) for $p = \frac{3}{4}$ and $K = 4$ shows that the true distribution is stretched along the diagonal compared to the renewal process.

In real experiments, the data is binned for the comparison, but the KL metric can be defined in the continuum limit, which maximizes the KL metric for comparing the true distribution P with the model $\tilde{P}$ [133]. Binning the data corresponds to coarse graining the distributions, which reduces the KL metric. The extreme example is the spacing being divided into a single bin, where all data points fall into the bin and the log likelihood of P and $\tilde{P}$ is zero. For any binning, the likelihood calculation determines the probability that the model histogram $\tilde{P}$ is an adequate representation of the data histogram, P. In the continuum limit, the Kullback-Liebler metric comes from an integration over a 2-D contour

$$\int dlnt_1 dlnt_2 dI_{P|\tilde{P}}(lnt_1, lnt_2) = \int dlnt_1 dlnt_2 P(lnt_1, lnt_2) ln(P(lnt_1, lnt_2)/\tilde{P}(lnt_1, lnt_2))$$
$$\tag{3.10.31}$$

that is shown in Fig. 3-6(b). We call $dI_{P|\tilde{P}}(lnt_1, lnt_2)$ the Kullback-Liebler (KL) difference function. To first order in ΔP the KL difference function resembles the

126
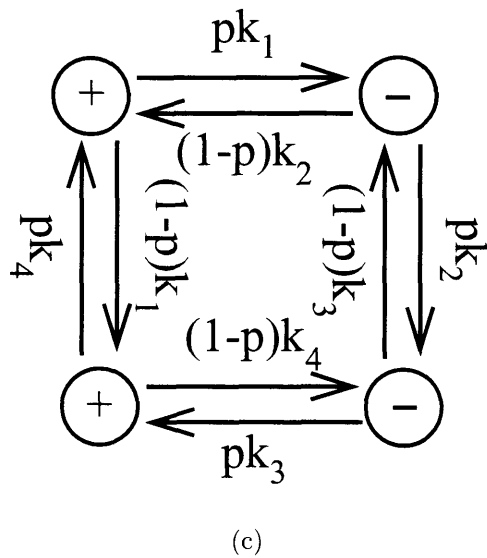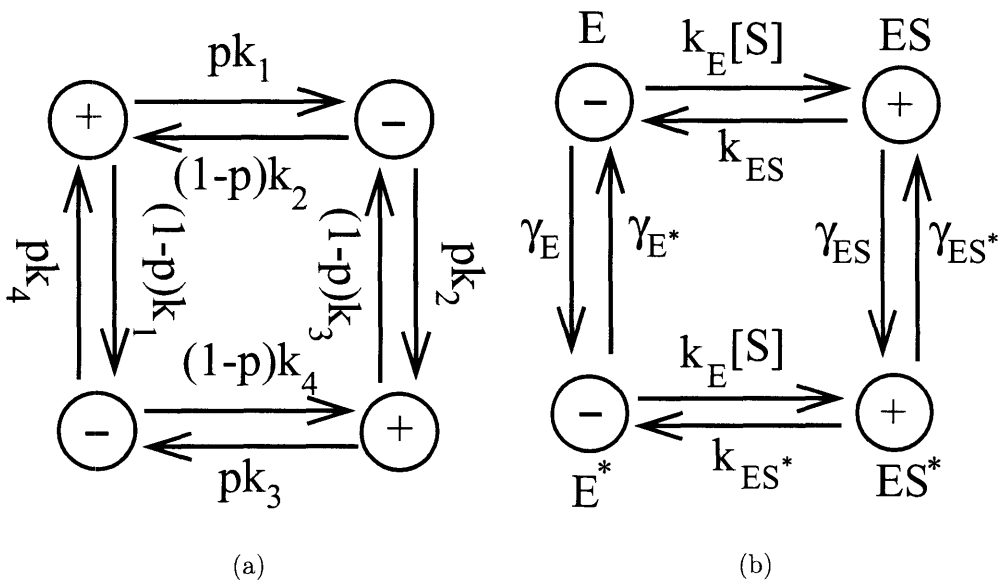
Figure 3-5: Kinetic schemes that violate detailed balance. (a) A scheme that has a circulation loop passing through both manifolds twice, which gives time reversibility and diagonal dominance violations. (b) A concentration dependent Michaelis-Menten scheme, where the substrate pumps the conformational coordinates of the system. (c) A kinetic scheme that has a circulation loop resulting in a peak in the single waiting time distributions.

(a)

(b)

(c)
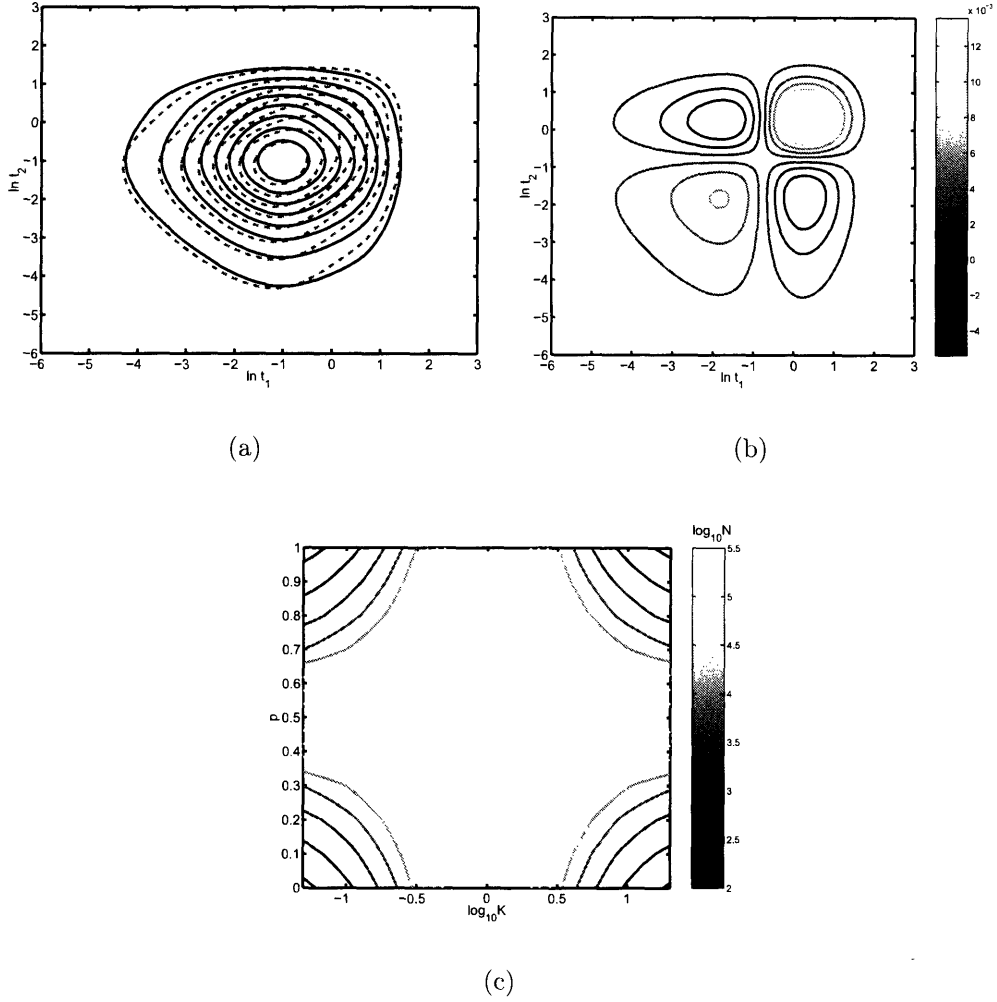
Figure 3-6: Testing the renewal hypothesis for the scheme in Fig. 3-5(a) with $p = 3/4$, and $2k_1 = k_2 = K^{-1}k_3 = K^{-1}k_4 = 1$ for $K = 4$. (a) The two *event* distribution, $P_{+-}(lnt_1, lnt_2)$ (dashed) is compared with the best fitting renewal process $\tilde{P}(lnt_1, lnt_2) = P_+(lnt_1)P_-(lnt_2)$ (solid). Note the log scales. (b) The KL difference, $dI_{P|\tilde{P}} = P_{+-}ln(\frac{P_{+-}}{P_+P_-})$ for comparing these two models. (c) The expected number of measurements required to distinguish a the data from the renewal model at the 95% confidence level. As the model gets closer to being a renewal process $p = 1/2$ or $K = 1$, more measurements are required to distinguish the two models.

traditional difference function, $\delta P(lnt_1, lnt_2) = P(lnt_1, lnt_2) - P(lnt_1)P(lnt_2)$, but this term does not contribute to test of $\tilde{P}$ [35, 242]. The expected minimum number of measurements necessary to distinguish the renewal process from the non-renewal process at a 95% confidence interval is plotted in Fig. 3-6(c).

The ability to detect a renewal violation depends on the magnitude of $p - \frac{1}{2}$ and $log(K)$. The degree of the renewal violation in the underlying scheme is captured by the magnitude of $p - \frac{1}{2}$, while the magnitude of $log(K)$ captures the ability to distinguish substates in the underlying manifold, which is necessary in order to detect the violation. The waiting time is a signature of the substate that was entered. If the waiting times are identical, $log(K) = 0$ the signatures cannot distinguish the states so the non-renewal nature defined by $p - \frac{1}{2}$ cannot be detected, while for $|log(K)|$ large, the two states are easily identifiable and the non-renewal nature is detectable. For systems that strongly violate the renewal property $p \approx 0, 1$, the number of measurements is reasonable (a couple thousand data points). Following our intuition, as the degree of the renewal violation decreases $p \rightarrow 1/2$, the number of measurements necessary to detect the violation increases and diverges for $p \approx 1/2$. Similar behavior is observed for the number of measurements necessary to distinguish different K values, since the system is a renewal process for $K = 1$.

The p parameter dominates the mixing of the process. For $p \approx 1/2$ the process mixes quickly and even a full sequence analysis cannot distinguish the renewal and non-renewal models. If the mixing is slow, but the kinetic rates are similar, $K \approx 1$, then a full sequence analysis can distinguish the data from a renewal model by detecting the weak but long lived correlations in the waiting times although the

proposed two dimensional test be weak. One may be able to use two dimensional analysis to overcome these difficulties by examining the probability distribution of sums of *events*, such as $P(t_1 + t_2, t_3 + t_4)$ or $P(t_1 + t_3, t_2 + t_4)$ (or the log equivalent). The first test would be sensitive to positive correlations, while the second test would be sensitive to negative correlations.

It is important to emphasize the difference between the mutual information and correlation analysis. Symmetry may make the first few correlations zero even if the measured quantities are correlated at higher moments, whereas the mutual information is only zero if the two quantities are independent. The properties that one measures with correlations also need to be characterized by a numerical value, but mutual information only requires binning of the data, which can be performed on data with qualitative labels or multidimensional data. An example is a traditional photon counting experiment, where the arrival time and fluorescence lifetimes of the photons are recorded [225]. In these experiments, the number of photons that arrive in a small time window and the average fluorescence lifetime in each bin may be recorded. To perform correlation analysis for the number and fluorescence lifetimes in two time windows separated by a fixed time t requires calculation of all possible correlations between the number of photons in different bins and the average fluorescence lifetime. For the mutual information, a two dimensional histogram of number of photons and average lifetime can be constructed and then the mutual information for all of these inputs results in a single number. If two variables are Gaussian distributed, the correlation function and mutual information are related since the correlation function defines the probability distribution.

3.11 Comparing experimental conditions

The simple test can be extended to the determination of the existence of concentration dependences in experiments [228]. Many single molecule experiments attempt to ascertain the mechanism of an enzyme's or ribozyme's reactivity by attaching a probe to the single molecule of interest by chemical modification. In this scenario, it becomes important to establish that the probe's motion is coupled to the reaction center of the single molecule by examining a substrate concentration dependence [237, 228].

We analyze the sensitivity in detecting changes for the model in Fig. 3-5(b), which is the reduced model that corresponds to Michaelis-Menten kinetics with extremely fast product release step, a diffusion limited substrate binding step, and a fluctuating kinetic rate for the enzyme-substrate to enzyme-product reaction. For simplicity we choose $K_E = K_{ES^*} = 1$, $K_{ES} = 2$, $\gamma_{ES^*} = \gamma_S = 0$, and $\gamma_{ES} = \gamma_{S^*} = 1/5$. The $-$ waiting time is a simple exponential process that depends linearly on the substrate concentration, $[S]$, i.e. $P_-(t, [S]) = [S]e^{-[S]t}$. The $+$ waiting time has a more complex substrate dependence,

$$P_+(t, [S]) = \frac{1}{66 + 30[S]}(121e^{-\frac{11}{5}t} + (11 + 30[S])e^{-t}). \qquad (3.11.32)$$

The first step in comparing two histograms $h_i^{(1)}$ and $h_i^{(2)}$ for different sets of experimental conditions is to construct the model for the two histograms being produced by the same underlying process, $\tilde{P}_i = 1/(N^{(1)} + N^{(2)})(h_i^{(1)} + h_i^{(2)})$. This model should be compared against the data to determine if a substrate dependence in the vari-

131

ous measurements exists. The test is very strong for the linearly dependent rates of the − waiting time. Even for concentration differences of 5%, the linear dependence in the waiting times can be detected with less than 500 measurements. Of greater interest is the ability to detect the more subtle dependences in the + waiting time distribution. Fig. 3-7(a) shows that even for a factor of 10 difference in the concentration, $[S] = 1, 10$, the waiting time distributions are similar, and the composite model $\tilde{P}_+(t) = \frac{1}{2}(P_+(t, [S] = 1) + P_+(t, [S] = 10)$ can be a very good fit. The difference measure $dI^{(i)}_{P|\tilde{P}} = P^{(i)}_+(lnt)ln(P^{(i)}_+(lnt)/\tilde{P}_+(lnt))$ is shown in Fig. 3-7(b), with $[S] = 1$ ($i = 1$ (solid)) and $[S] = 10$ ($i = 2$ (dashed)) The expected number of measurements needed to discriminate $\tilde{P}$ from the two true probability distributions at the 95% confidence interval is presented in Fig. 3-7(c).

The ability to distinguish the two waiting time distributions depends on the difference in the concentrations. This waiting time distribution is a weighted average of two exponentials with concentration independent decay constants, but concentration dependent weights. These changes in the weights saturate at high and low concentrations, which results in the plateaus in the ability to detect the concentration dependences at high and low concentrations. This comparison of two distributions can be used to test different single molecules in the same experimental conditions or segments of a single trajectory to determine if the experiment is ergodic. As will be discussed elsewhere, this idea can be extended to examining collections of single molecules in order to classify their behaviors [226].

(a)

(b)

(c)
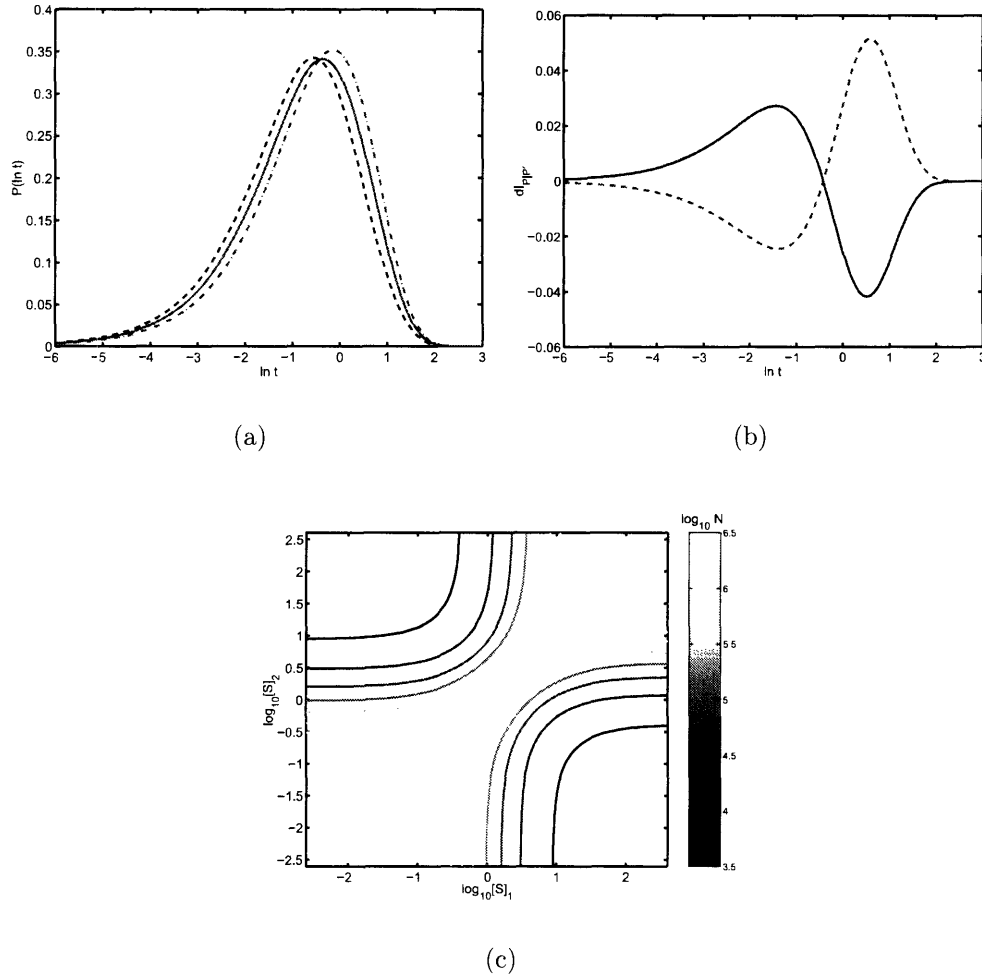
Figure 3-7: Ability to determine concentration dependent behaviors in $P_+(t_1)$ in the model depicted in Fig. 3-5(b) for $K_E = K_{ES^*} = 1$, $K_{ES} = 2$, $\gamma_{ES^*} = \gamma_S = 0$, and $\gamma_{ES} = \gamma_{S^*} = 1/5$. (a) $P_+(lnt)$ for $[S] = 1$ (dot-dashed) and $[S] = 10$ (dashed) are compared with with $\tilde{P} = \frac{1}{2}(P_+([S] = 1) + P_+([S] = 10)$ (solid). These probability distributions can be used to calculate $dI_{P|\tilde{P}}$, shown in (b) for $P_+([S] = 1)$ (solid) and $P_+([S] = 10)$ (dashed). (c) The expected number of measurements needed to discriminate behaviors between $[S]_1$ and $[S]_2$ at the 95% confidence level.

3.12 Time reversibility

The existence of detailed balance is an important property to establish for various protein systems since detailed balance violations imply that the conformational kinetics being probed are also pumped by an external source of energy, such as the substrate in an enzymatic turnover process or the ionic potential across a membrane that is often explored in ion channel experiments [228, 196]. The easiest test of the probed coordinate's motion violating detailed balance is a substrate concentration dependence, as discussed above. These concentration dependences are not always easy to detect [196], and do not always give insight into the topology of the system that leads to the detailed balance violation.

As discussed previously [228], there are several manifestations of detailed balance violations that can be seen in the 2-D *event* probability contours. These manifestation include a violation of time reversibility $P_{+-}(t_1, t_2) = P_{-+}(t_2, t_1)$, a violation of the triangle inequality of same *event* measurements, $P_{++}(t_1, t_2)^2 \leq P_{++}(t_1, t_1)P_{++}(t_2, t_2)$, and a peak in the single time waiting time distribution [228]. Most previous analysis concentrated on the time reversibility [196]. Here we will use one and two dimensional analysis to explore all three of these possible manifestations of detailed balance violation without resorting to examining the underlying model.

If detailed balance holds for a system, the system is time reversible and the statistics of the forward and backward process are identical [228],

$$P_{+-}(t_1, t_2) = P_{-+}(t_2, t_1). \tag{3.12.33}$$

A typical realization of a time reversibility violation occurs when there is a circulation loop in the underlying topology of the kinetic scheme that enters both manifolds at least twice. The simplest realization of this has only four substates $+_1, +_2, -_1$, and $-_2$ with the conformational dynamics preferring to proceed in a circular sequence $+_1 \rightarrow -_1 \rightarrow +_2 \rightarrow -_2 \rightarrow +_1 \rightarrow \ldots$. This situation is depicted in Fig. 3-5(a).

The time reversibility property is easily tested within the framework applied to test for a concentration dependence since time reversibility reduces to determining if two probability distributions are identical. Comparing P_{+-} and P_{-+} in the model in Fig. 3-5(a), with $k_1 = k_2 = K^{-1}k_3 = K^{-1}k4 = 1$ as a function of p and K results in Fig. 3-8. As shown in Fig. 3-8(a), for $p = \frac{3}{4}$ and $K = 2$, $P_{+-}(t_1, t_2)$ is elongated along the $t_1 = t_2$ line compared to $P_{-+}(t_2, t1)$ (remember logarithmic binning). The alternative hypothesis $P_{+-}(t_1, t_2) = P_{-+}(t_2, t_1) = \tilde{P}(t_1, t_2)$ with $\tilde{P} = \frac{1}{2}(P_{+-}(t_1, t_2) + P_{-+}(t_2, t_1))$ is similar to both distributions. The KL differences, $dI_{P|\tilde{P}}^{(+)} = P_{+-}ln(P_{+-}/\tilde{P})$ and $dI^{(-)} = P_{-+}ln(P_{-+}/\tilde{P})$, are plotted in Fig. 3-8(b) and (c), respectively. Since the P_{+-} distribution is elongated along the diagonal compared to P_{-+}, $dI^{(+)}$ is positive along the diagonal and negative on the off diagonal. The other KL difference, $dI^{(-)}$, shows the opposite behavior with a negative diagonal and positive off diagonal.

The necessary number of measurements to distinguish P_{+-} and P_{-+} from $\tilde{P}$ at the 95% confidence interval is plotted in Fig. 3-8(d). Similar to the renewal indicator, the ability to discriminate depends the magnitudes of $p - \frac{1}{2}$ and $log(K)$. The magnitude of $p - \frac{1}{2}$ is a measure of the detailed balance violation, while the magnitude of $log(K)$ is a measure of our ability to distinguish the two states. If two states are

(a)

(b)

(c)

(d)
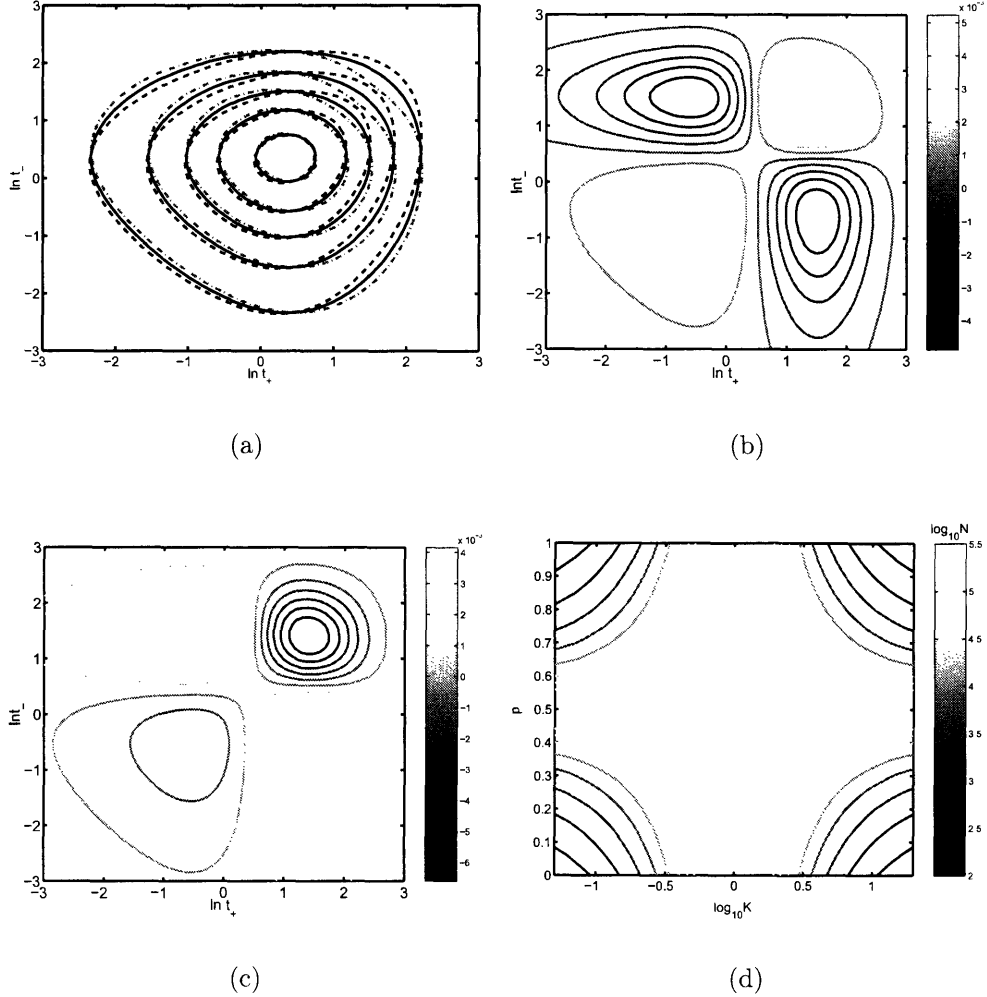
Figure 3-8: Determination of time reversibility in the model depicted in Fig. 3-5(a) for $k_1 = k_2 = 1$, $k_3 = k_4 = K = 2.7$, and $p = 3/4$. (a) $P_{+-}(lnt_1, lnt_2)$ (dash) and $P_{-+}(lnt_2, lnt_1)$(dot-dash) are compared against the time reversible model, $\tilde{P} = \frac{1}{2}(P_{+-} + P_{-+})$ (solid). (b) A contour of $dI^{(+)} = P_{+-}(lnt_1, lnt_2)ln\left(P_{+-}(lnt_1, lnt_2)/\tilde{P}(lnt_1, lnt_2)\right)$ (c) A contour of $dI^{(-)} = P_{-+}(lnt_2, lnt_1)ln\left(P_{-+}(lnt_2, lnt_1)/\tilde{P}(lnt_1, lnt_2)\right)$. (d) The expected number of measurements needed to discriminate P_{+-} and P_{-+} from $\tilde{P}$ at the 95% confidence level as a function of p and K.

136

distinguishable, ($|log(K)|$ large), it is easier to detect the detailed balance violation. If p is very different from $1/2$, ($p \rightarrow 0, 1$), but K is near unity, the time reversible and irreversible models may still be discriminated by either the complete sequence analysis or comparing sums of *events* in another 2D analysis as discussed in Sec. 3.10.

3.13 Diagonal Features

If the waiting time in one of the manifolds has a rate limiting step that is the same for all possible paths, such as substrate transport, we expect a degeneracy in the eigenspectrum of the waiting time distribution [228]. This degeneracy can make the sequence time reversible even if it violates detailed balance. In these cases, the distribution of two similar *events* such as two $+$ *events* separated by a $-$ *event* must be used to test for detailed balance violations. As an example, setting $k_2 = k_4$ for the model in Fig. 3-5(a) makes the $-$ waiting time a simple exponential, $P_-(t) = ke^{-kt}$, and the system time reversible, but there is a detailed balance violation and two adjacent $+$ *events* can show features of this violation.

One feature of a detailed balance obeying P_{++} distribution is diagonal dominance, where $P_{++}(t_1, t_2)^2 \leq \sqrt{P_{++}(t_1, t_1)P_{++}(t_2, t_2)}$ [228]. The violation of this diagonal dominance for the model in Fig. 3-5(a) with $k_1 = k_2 = k4 = 1$, $k_3 = K = 2.7$ and $p = 3/4$ is shown in Fig 3-9(a) and (b). Fig. 3-9(a) shows the isocontours of $\sqrt{P_{++}(t_1, t_1)P_{++}(t_2, t_2)}$ are narrower than those of the true distribution, so the distribution becomes greater than the theoretical detailed balance limit, resulting in the positive difference between $P_{++}(t_1, t_2)$ and $\sqrt{P_{++}(t_1, t_1)P_{++}(t_2, t_2)}$. In other

137

words,

$$\Delta P_{++} = P_{++}(t_1, t_2) - \sqrt{P_{++}(t_1, t1)P_{++}(t_2, t_2)} > 0 \qquad (3.13.34)$$

indicates a detailed balance violation.

Diagonal dominance holds for integration over t or lnt so

$$\left[\int_{t_1-\Delta/2}^{t_1+\Delta/2} dt' \int_{t_2-\Delta/2}^{t_2+\Delta/2} dt'' P_{++}(t', t'') \right]^2 >$$
$$\left[\int_{t_1-\Delta/2}^{t_1+\Delta/2} dt' \int_{t_1-\Delta/2}^{t_1+\Delta/2} dt'' P_{++}(t', t'') \right] \left[\int_{t_2-\Delta/2}^{t_2+\Delta/2} dt' \int_{t_2-\Delta/2}^{t_2+\Delta/2} dt'' P_{++}(t', t'') \right] \quad (3.13.35)$$

also indicates a detailed balance violation and the diagonal dominance test is also valid

for a histogram, $P_{ij} = a_{ij} P_i^{1/2} P_j^{1/2}$ [228]. Unlike previous tests, histogramming may

improve the KL measure of the detailed balance violation, by allowing comparison

of a narrow diagonal feature with broad off diagonal features, which can also violate

detailed balance and has appeared in some models that violate detailed balance [135].

Similar to the time reversibility test, a diagonal dominance violation occurs when

there is a circulation loop in the underlying topology of the kinetic scheme that

enters both manifolds at least twice, as depicted in Fig. 3-5(a).

To demonstrate a diagonal dominance test, we examine the model in Fig. 3-5(a).

The two large off diagonal peaks in the ΔP_{++} distribution in Fig. 3-9(a) indicate that

splitting the distribution into four quadrants along the diagonal, $t_1 = t_2$, should be

sufficient to test for diagonal dominance. The position of the split depends on p and

K. For the histogram with only four quadrants, diagonal dominance implies that

$\tilde{P}_{12}^2, \tilde{P}_{21}^2 \leq \tilde{P}_{11} \cdot \tilde{P}_{22}$. Assuming that we are testing detailed balance, time reversibility
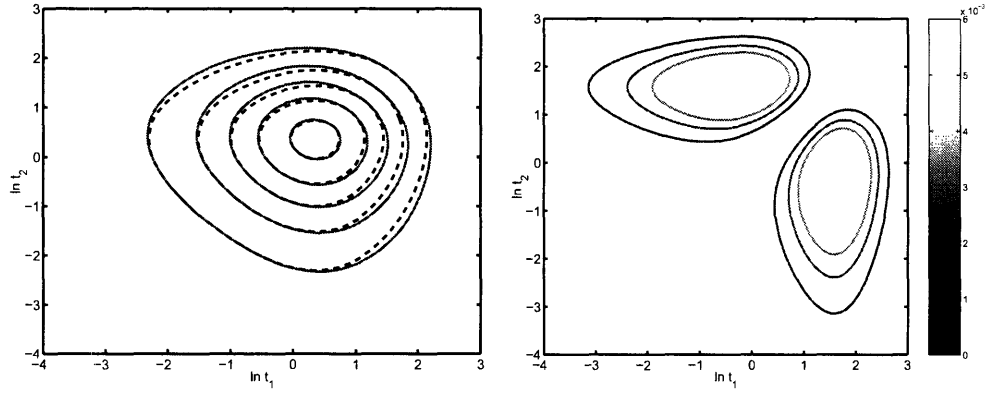
(a)

(b)
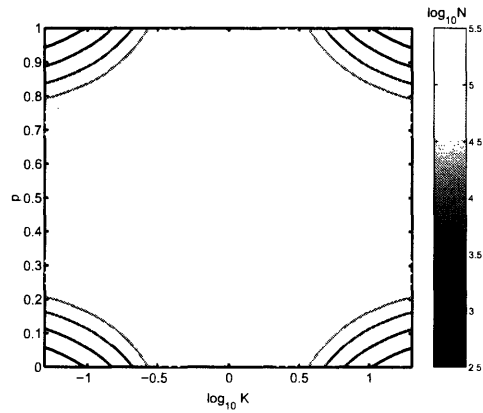
(c)

Figure 3-9: Ability to discriminate a diagonal dominance violation in the model depicted in Fig. 3-5(a) for $k_1 = k_2 = k_4 = 1$, $k_3 = K = 2.7$, and $p = 3/4$. (a) $P_{++}(lnt_1, lnt_2)$ (solid) is compared against $\sqrt{P_{++}(lnt_1, lnt_1)}$ $\sqrt{P_{++}(lnt_2, lnt_2)}$ (dashed). (b) $\Delta P_{++}(lnt_1, lnt_2) = P_{++}(lnt_1, lnt_2) - \sqrt{P_{++}(lnt_1, lnt_1)}$ $\sqrt{P_{++}(lnt_2, lnt_2)}$, which shows two positive off diagonal peaks indicating a detailed balance violation. c) The expected number of measurements needed to determine the existence of a diagonal dominance violation.

is also required and the optimal diagonally dominant time reversible distribution is given by $\tilde{P}_{12} = \tilde{P}_{21} = \frac{1}{2N}(h_{12} + h_{21})$ and $\tilde{P}_{ii} = \frac{1}{N}h_{ii}$, for $\frac{1}{4}(h_{12} + h_{21})^2 \leq h_{11}h_{22}$. If this inequality is not satisfied, we must modify the probability to,

$$\tilde{P}_{11} = \frac{(h_{11} + \frac{1}{2}(h_{12} + h_{21}))^2}{N^2}$$

$$\tilde{P}_{22} = \frac{(h_{22} + \frac{1}{2}(h_{12} + h_{21}))^2}{N^2}$$

$$\tilde{P}_{21} = \tilde{P}_{12} = \frac{(h_{11} + \frac{1}{2}(h_{12} + h_{21}))(h_{22} + \frac{1}{2}(h_{12} + h_{21}))}{N^2} \tag{3.13.36}$$

Following previous analysis we compare this model to the data, $P_{ij} = h_{ij}/N$, to determine the probability of a diagonal dominance violation. The number of measurements necessary to discriminate a diagonal dominance violation at the 95% confidence interval using the four quadrant test on the model in Fig. 3-5(a), is plotted in Fig. 3-9(c) as a function of p and K. The features are similar to those in the previous tests, with the discriminating power of the test depending on the magnitude of $p - \frac{1}{2}$ and $log(K)$ since these measure the detailed balance violation and distinguishability, respectively.

3.14 Single waiting time test

A one dimensional feature that indicates detailed balance violations is the existence of a peak in the single time waiting time distribution [228]. This detailed balance violation has a different origin than the previously discussed time reversibility violation and diagonal dominance violations that are usually associated with the circulation loop passing though the $+$ and $-$ manifolds at least twice. A peak in the single wait-

140

ing time distribution results from a flow within a single manifold so that the system has a tendency to enter the $+$ or $-$ manifolds through one substate and exit through another. This indicates a microscopic time reversibility violation between states in the same manifold even though the mesoscopic time reversibility may hold. If detailed balance holds, the single waiting time distribution can be expressed as a sum of exponentials $P(t) = \int dk P(k) k e^{-kt}$, where $P(k)$ is a proper probability density, $P(k) \geq 0$, $\int dk P(k) = 1$. A rigorous method of testing for a peak is to determine $P(k)$ from maximum entropy fits or another method and compare this probability distribution to the data [225].

As a simple example, we examine the $+$ waiting time distribution, $P_+(t)$, in the model depicted in Fig. 3-5(c) with $k_1 = k_2 = k_3 = K^{-1}k_4 = 1$. This waiting time distribution with $K = 2.7$ and $p = 1/10$ is compared against the best fitting detailed balance obeying distribution, $\tilde{P} = \int dk P(k) k e^{-kt}$, in Fig. 3-10(a). The detailed balance distribution is wider than the detailed balance violating scheme. Fig. 3-10(b) shows the KL difference, $dI_{P|P'}$. Similar to the previous tests, the ability to detect this detailed balance violation increases with increasing magnitude of $p - \frac{1}{2}$, but the ability to distinguish the peak varies inversely with magnitude of $log(K)$. Taking $p = 1$, the waiting time distribution is given by

$$P_+(lnt) = \frac{k_4}{k_4 - k_1} k_1 e^{-k_1 t} - \frac{k_1}{k_4 - k_1} k_4 e^{-k_4 t}, \qquad (3.14.37)$$

which as a zero at $t = 0$. If k_1 is much smaller than k_4 ($K \gg 1$), the waiting time of the system is nearly mono-exponential, $P_+(t) \approx k_1 e^{-k_1 t}$, with only a brief deviation

at short times, so it is difficult to detect the detailed balance violation. Similar results hold for $K \ll 1$, and it is only when $k_4 \approx k_1$ that the deviation from simple exponential behavior can be detected.

3.15   Summary of simple feature tests

As demonstrated above, one and two dimensional histograms can elucidate many properties of a system without solving the combinatorial complex problems of determining the exact underlying model. These methods allow detection of correlations in *events* through the renewal test, similarities in behaviors under different experimental conditions, and detailed balance violations. The renewal test determines if the transitions between the two manifolds correspond to multiple paths, and the experimental condition dependence indicates that the probe is coupled to the reaction coordinate. The detailed balance violations can result in time reversibility violations or a lack of diagonal dominance that indicate a circulation loop that goes through both manifolds as least twice, and peaks in the single waiting time distribution that indicate a multiple step circulation through a single manifold. Knowledge of these properties can give insight into the underlying topology of Markovian models without determining the specific model parameters.

The proposed tests not only allow determination of the existence of these properties, but also create rigorous bounds on the ability to discriminate models with one and two dimensional data. The number of measurements necessary to distinguish these model at the 95% confidence can not be reduced by introducing another mea-

(a)

(b)

(c)
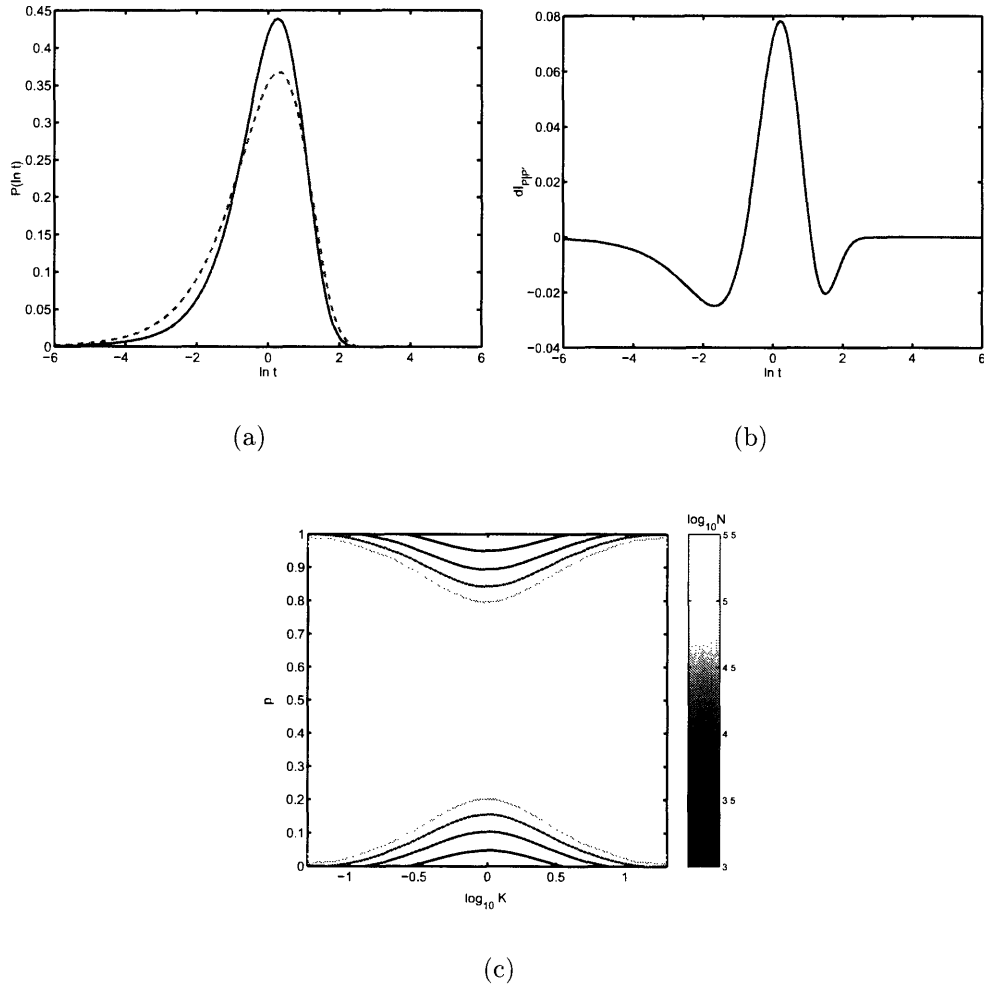
Figure 3-10: Ability to determine if the waiting-time distribution is not consistent with a detailed balance scheme. (a) The single + waiting time distribution, $P_+(lnt)$ in the model depicted in Fig. 3-5(c) for $k_1 = k_2 = k_3 = 1$, $k_4 = K = 2.7$, and $p = 1/10$ (solid) is compared to the best fit of a detailed balance obeying scheme, $\tilde{P}(t) = \int dk P(k) k e^{-kt}$, $P(k) > 0$ (dashed). (b) $dI_{P|\tilde{P}}$ for the distributions in (a). (c) The expected number of measurements needed to discern the detailed balance violation as a function of K and p. Unlike other tests, the ability to determine the existence of a peak in the waiting time distribution depends on $K \approx 1$.

sure on this data. The only major assumption that may need to be corrected is the independence of *events*, which may not be true if the histogram is constructed from parsing a long single trajectory, but this will only modify the variance estimate.

The analysis can be extended to higher dimensional binned data, but histogramming will not be practical. Instead, one needs to fit the data to flexible functional forms that continue to obey the restrictions on the properties, such as time reversibility [140]. These reduced information tests will never be as powerful as full sequence analysis, but even using flexible functional forms is orders of magnitude less computationally intensive than a full sequence analysis and does not require one to propose underlying topologies. These reduced information methods can also be expanded to classify different molecular trajectories, to test ergodicity, and to determine properties of the transition state ensemble [226]. The computational simplicity, along with the rigorous bounds in the ability to discriminate models, makes the information theoretic approach to reduced data representations an advantageous first step in performing single molecule analysis.

3.16 Classification of single molecule processes

The two major cited reasons for studying single molecule experiments over ensemble experiments is the ability to observe short lived transition states, whose signal is drowned out in an ensemble experiment, and to distinguish homogeneous and heterogeneous effects [158, 183, 77, 155]. Most of the successes in single molecule experiments have come from elucidating the short lived transition states, since their single

molecule signature is easily identified within many of these experiments [23, 247, 237]. A much more difficult task is assessing heterogeneity since this requires determining appropriate measures of the similarity between molecules.

Unambiguously determining the existence of heterogeneous behavior in single molecule trajectories is imperative in experiments that use single molecule probes to measure micro-environments, such as applications to transport processes in cells or glasses [95, 183, 77, 155]. A common approach to assessing heterogeneity has been mapping the measurements into a scalar, and then examining the distribution of this quantity [136, 134]. The selected scalar is arbitrarily chosen and little effort is dedicated to assessing if the scalar is a good indicator of heterogeneity. The two most common scalars are expectation values or parameters of a fit. If the expectation value is properly chosen, the scalar can discriminate quantitative differences, such as differences between mean waiting times, but they will not discriminate qualitative differences, such as distinguishing exponential distributions from normal distributions. Using higher order moments has been proposed to quantify differences in distributions, but higher order moments are sensitive to fluctuations in data and are not reliable [134].

Parameters determined by fits to specified functional forms can discriminate qualitative differences, but the ability to discriminate is restricted to the functional form of the fit that is often arbitrary. If the distributions show small peaks, but the functional form decays monotonically, such as an exponential or stretched exponential, the contribution of the peak features will be lost. The functional form may also lack a physical mechanism or (more dangerously) imply a physical mechanism that is not

present.

In this paper, we outline a statistical procedure based on the histogramming of single molecule *events* that will allow rigorous classification without calculating a simple scalar, such as a mean, or enforcing a fit to a specific functional form. The statistical distribution of *events*, such as photon emission, switching from high fluorescent to low fluorescent state, or forming distinct contacts in a protein, are generally the strongest indicators of the single molecule properties [223, 228, 35, 242]. Since most other analyzes can be performed on the bulk system, the statistics of *events* are the major distinction between single molecule and ensemble experiments. Histograms of *events* assign a fingerprint to each molecule that avoids the computational difficulties of performing full sequence analysis to determine underlying models for each molecule [30]. This classification scheme will allow the observation of heterogeneity in single molecule environments or determination of the effects of experimental parameters on the measured distributions. These approaches can be applied to ergodicity tests by examining several time windows in a single trajectory. Classification should be the first step in any single molecule experiment, since it gives insight into different types of behavior that may be hidden in the data sets and are not observable in bulk experiments.

3.17 Histogram classification of two molecules

The procedure may be applied to any characteristic histogram, including the joint sojourn time in blinking experiments and the joint distribution of photon arrival

146

times. For simplicity, we concentrate on the photon emission statistics that have become common in intra-molecular quenching experiments [240]. The *event* is the number of photons emitted over a small time interval. The data is coarse grained by histogramming the number of photons in each bin. For each molecule, j, this results in a histogram $h_i^{(j)}$. One could also histogram the average lifetime of the photons in each bin resulting in a two dimensional histogram, or histogram the joint distribution of the number of photons in two bins separated by a fixed distance.

From this histogram, we classify the different observed behaviors. First, suppose that the experiment was performed on two different molecules (or we examine two different sections of the same single molecule sequence to assess ergodicity) and the histograms for each molecule was constructed, $h_i^{(1)}$ and $h_i^{(2)}$. We want to determine if the molecules are similar or different (i.e. one or two classes). If the two molecules are behaving differently, the most likely model for each molecule is $P_i^{(j)} = \frac{1}{N^{(j)}} h_i^{(j)}$, with $j = 1, 2$. If the two molecules are produced by the same probability distribution, the most likely probability distribution is $\tilde{P}_i = \frac{1}{N^{(1)}+N^{(2)}}(h_i^{(1)} + h_i^{(2)})$.

For $\tilde{P}_i$, the probability of getting the histogram h_i is $P(h_i|\tilde{P}_i) = \prod_i (\tilde{P}_i)^{h_i}$. Taking the log of this probability gives

$$lnP(h_i|\tilde{P}_i) = \sum_i h_i ln(\tilde{P}_i). \tag{3.17.38}$$

Since each *event* is a random variable, $\sum_i h_i ln(\tilde{P}_i)$ is the result of a sum of random variables that converges to a Gaussian distribution in the large N limit. As a result, the difference in the log probability of the histogram given P_i and $\tilde{P}_i$ (log-likelihood)

147

is a natural method of comparing the restricted model $\tilde{P}_i$ to the data P_i. For the histograms $h_i^{(j)}$, the difference in the log probability is

$$\sum_{ij} h_i^{(j)}[ln(P_i^{(j)}) - ln(\tilde{P}_i)] = \sum_{ij} h_i^{(j)} ln(P_i^{(j)}/\tilde{P}_i) = \sum_{ij} N^{(j)} P_i^{(j)} ln(P_i^{(j)}/\tilde{P}_i) = NI_{P|P'}.$$

(3.17.39)

where $I_{P|P'}$ is the Kullback-Liebler metric (KL) [133],

The model $\tilde{P}_{ij}$ has restrictions, which generally implies less parameters, and should be adopted if the differences in the log-likelihoods is not statistically significant. The log-likelihoods should be compared against the estimated variance in this measure,

$$N\delta I_{P|\tilde{P}}^2 = \sum_j N^{(j)} \left[\sum_i P_i^{(j)} ln(P_i^{(j)})^2 - \left(\sum_i P_i^{(j)} ln(P_i^{(j)}) \right)^2 \right] +$$
$$\sum_j N^{(j)} \left[\sum_i P_i^{(j)} ln(\tilde{P}_i)^2 - \left(\sum_i P_i^{(j)} ln(\tilde{P}_i) \right)^2 \right].$$

(3.17.40)

It is simple to account for correlations in the data by modifying the variance. This testing method penalizes using the data histogram, $P_i^{(j)}$, by a factor that scales as $\sqrt{N}$ without regard to the number of parameters, whereas other methods penalize by factors of $ln(N)$ or unity with a parameter dependent prefactor [125]. The preferred method should depend on both the number of data points and number of parameters.

3.18 Histogram classification of M molecules

The results for comparing two sequences can be extended to several sequences. The objective is to use the fingerprint provided by the histogrammed data to classify a set

of single molecule trajectories into groups with similar traits. One can then examine the properties of these different groups to identify these features and gain insight into the properties of the system. Suppose that we have M molecular trajectories resulting in M histograms, $\{h_i^{(j)}\}$. The objective will be to make A probability distributions that capture the typical behaviors of the different molecules. The distributions are indexed by α and represent different classes of molecules. The probability of each class, $\mu^{(\alpha)}$ ($\mu^{(\alpha)} > 0$, $\sum_\alpha \mu^{(\alpha)} = 1$), and the representative histogram for each class, $P_i^{(\alpha)}$ ($P_i^{(\alpha)} > 0$, $\sum_i P_i^{(\alpha)} = 1$), are determined for fixed number of classes through the expectation maximization algorithm discussed in Sec. 3.18.1. Then the log-likelihoods and variances for various values of A are compared to determine the minimum number of classes, A, necessary to represent the molecular behaviors. The result of the calculation is the signature of the class determined by the representative histogram, $P_i^{(\alpha)}$, the percentage of occurrences of that class, μ^α, and the probability that a particular molecule is a member of that class.

3.18.1 Expectation maximization

Without going into detail about the expectation maximization algorithm [197], the best fitting class model are defined by $P^{(\alpha)}$ and $\mu^{(\alpha)}$ and can be calculated from an iterative algorithm. For fixed model histograms $P_i^{(\alpha)}$ and weights $\mu^{(\alpha)}$, the probability of each molecular histogram belonging to each class is calculated. Then $\mu^{(\alpha)}$ is the expected percentage of histograms that belong to each class, and $P_i^{(\alpha)}$ is the average of the molecular histograms weighted by the probability that the histogram belongs

to the class α. The resulting equations are

$$\mu_{t+1}^{(\alpha)} = \lambda_\mu^{-1} \sum_m P\left(\alpha \left| \left\{\mu_t^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}, h_i^{(m)}\right.\right) =$$

$$\lambda_\mu^{-1} \sum_m \frac{P\left(\alpha, h_i^{(m)} \left| \left\{\mu_t^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}\right.\right)}{P\left(h_i^{(m)} \left| \left\{\mu_t^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}\right.\right)} = \lambda_\mu^{-1} \sum_m \frac{P\left(\alpha, h_i^{(m)} \left| \left\{\mu_t^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}\right.\right)}{\sum_\alpha P\left(h_i^{(m)}, \alpha \left| \left\{\mu_t^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}\right.\right)}$$

$$= \lambda_\mu^{-1} \sum_m \frac{\mu_t^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)_t^{h_i^{(m)}}}{\sum_\alpha \mu_t^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)_t^{h_i^{(m)}}}$$

$$\left(P_i^{(\alpha)}\right)_{t+1} = \lambda_\alpha^{-1} \sum_m h_i^{(m)} P\left(\alpha \left| \left\{\mu_{t+1}^{(\alpha)}\right\}, \left\{\left(P_i^{(\alpha)}\right)_t\right\}, h_i^{(m)}\right.\right)$$

$$= \lambda_\alpha^{-1} \sum_m h_i^{(m)} \frac{\mu_{t+1}^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)_t^{h_i^{(m)}}}{\sum_\alpha \mu_{t+1}^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)_t^{h_i^{(m)}}}$$

$$(3.18.41)$$

where t denotes the iteration number. The proportionality constants, λ_μ and λ_α, are determined by normalization, $\sum_\alpha \mu^{(\alpha)} = \sum_i P_i^{(\alpha)} = 1$. One should note that local minima may exist so a thorough search is necessary.

Once the optimal parameters, $\mu_*^{(\alpha)}$ and $(P_i^{(\alpha)})_*$ are determined for a fixed A, the average log-likelihood and variance can be easily calculated,

$$\langle L_A \rangle = M^{-1} \sum_m ln(P(\{h_i^{(m)}\})|\{\mu_*^{(\alpha)}\}\{(P_i^{(\alpha)})_*\}) = M^{-1} \sum_m ln(\sum_\alpha \mu_*^{(\alpha)} \prod_i (P_i^{(\alpha)})_*^{h_i^{(m)}}).$$

$$(3.18.42)$$

and

$$\langle \delta L_A^2 \rangle = M^{-1} \sum_m \left[ln(P(\{h_i^{(m)}\})|\{\mu_*^{(\alpha)}\}\{(P_i^{(\alpha)})_*\})^2 - L_A \right]^2$$

$$= M^{-1} \sum_m \left[ln(\sum_\alpha \mu_*^{(\alpha)} \prod_i (P_i^{(\alpha)})_*^{h_i^{(m)}})^2 - L_A \right]^2 \qquad (3.18.43)$$

150

The log-likelihoods can then be compared at the 95% confidence intervals to determine the appropriate number of classes [228]. One can also calculate the probability of each molecule belonging to different classes,

$$P(\alpha|h_i^{(m)}) = \mu^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)^{h_i^{(m)}} \bigg/ \left(\sum_\alpha \mu^{(\alpha)} \prod_i \left(P_i^{(\alpha)}\right)^{h_i^{(m)}}\right). \qquad (3.18.44)$$

These probabilities can be used divide the single molecules into subensembles that allow further investigation into the causes and properties of these different classes.

3.19 An example: the telegraph model with photon counting

As a demonstration of the classification scheme, we consider a photon counting experiment. The monitored statistics correspond to the number of photons in each bin. The molecule switches between two states with a rate that is much smaller than the inverse bin size, $k \ll ms^{-1}$, so that we can approximate the transitions as occurring in between the $1ms$ bins with the probability of the transition $p \ll 1$. Given that the molecule is in state ω during bin n, the number of emitted photons is given by a Poisson distribution with parameter λ_ω, i.e. $P_i^{(\omega)} = \frac{\lambda_\omega^i}{i!} e^{-\lambda_\omega}$. For calculations below, we consider $\lambda_\omega = 1, 3$ for states $\omega = 1, 2$, respectively. These assumptions allow us to

151

write a Markov chain expression for the transitions between states 1 and 2,

$$
\begin{bmatrix} \rho_{n+1}^{(1)} \\ \\ \rho_{n+1}^{(2)} \end{bmatrix} = \begin{bmatrix} 1-p & p \\ \\ p & 1-p \end{bmatrix} \begin{bmatrix} \rho_n^{(1)} \\ \\ \rho_n^{(2)} \end{bmatrix}. \tag{3.19.45}
$$

We are concerned with sequences that are long enough to be able to develop histograms for each molecule, but short relative to the correlations in the data, $1 \ll N \leq \frac{1}{p}$. To study this scenario we consider, $p \to 0$ and $N \to \infty$, such that $pN \to 2t/\tau$, where τ corresponds to a fictitious relaxation time. Setting $\tau = 1$ makes the mixing time, t, and the number of time bins, N, the important parameters. Letting n_i denote the number of time bins that the system is in state $\omega = 1, 2$, and $z = n_1 - n_2$, the continuum approximate equation of motion for z as a function of t is the telegraph equation [154]

$$
\frac{\partial^2 \rho(z,t)}{\partial t^2} + \tau^{-1} \frac{\partial \rho(z,t)}{\partial t} = \frac{\partial^2 \rho(z,t)}{\partial z^2}, \tag{3.19.46}
$$

with initial conditions $\frac{\partial \rho(z,t)}{\partial t} = 0$ and $\rho(z,0) = \delta(x)$. Since the important quantity is the percentage of time spent in state 2, we want to solve the telegraph equation and substitute $x = \frac{1}{2} + \frac{z}{2N}$, resulting in

$$
\rho(x,t) = \frac{1}{2} e^{-t/(2\tau)} \left[\delta(x) + \delta(x-1) + \frac{t}{\tau} \left(I_0 \left(\frac{t}{\tau} \sqrt{x(1-x)} \right) + \frac{1}{2} \frac{I_1 \left(\frac{t}{\tau} \sqrt{x(1-x)} \right)}{\sqrt{x(1-x)}} \right) \right]
$$

$$
\tag{3.19.47}
$$

where $I_{0,1}$ are the modified Bessel functions. The x corresponds to the class of the molecule, which is determined by the number of time bins associated with distributions 1 versus 2. For this model, the classes are a continuous distribution. The δ

functions correspond to particles that are always in state 1 or 2, whereas the Bessel functions correspond to particles that have made at least one transition. For small mixing times, t, the δ functions dominate since the probability of a molecule switching distributions is small and the ballistic nature of the model is the major contribution. As a result we expect to see two classes of molecules, those that correspond to distribution $\omega = 1$ and $\omega = 2$. For large mixing times, the δ function contribution decays since every molecule will have switched states at least once, and the Bessel function contribution dominates. Due to the scaling, the distribution approaches a δ function at $x = 1/2$ and each molecule appears to be a member of the same class since ergodicity is achieved in the long time limit. The interesting behavior occurs at intermediate times where the optimal number of classes is not straight forward. Technically, an infinite spectrum of classes exists since the Bessel function contribution is spread out over the entire middle region for intermediate times, but these different classes should show similar behaviors and be captured by only a few classes.

The finite number of classes, $A = 1, 2, \ldots$, that approximate this continuous spectrum can be calculated numerically. Ignoring the stochastic nature of histograms, each class will have the functional form

$$P_i^{(\alpha)} = (1 - \beta_\alpha)P_i^{(\omega=1)} + \beta_\alpha P_i^{(\omega=2)}. \tag{3.19.48}$$

The different classes result from varying β_α. After finding the classes, we calculate the expected log-likelihood score, the expected variance in this score, and determine the number of single molecule trajectories necessary to distinguish models with differing

153

numbers of classes at the 95% confidence interval.

For modest amounts of data, one or two classes will always be the preferred model. Choosing a three class model over the two class model for any parameter t requires over 5000 histogrammed events per molecule for at least 1000 molecules, which is not reasonable. Although we know that a continuous distribution of classes exist, we cannot statistically verify the classes with reasonable amounts of data and it does not make sense to construct more complicated models that cannot be statistically verified. Between the one and two class models, the expected number of molecules necessary to distinguish the two class model at 95% confidence is plotted in Fig. 3-11(a) as a function of the number of turnovers per molecule, N, and the mixing time, t. Few molecules need to be examined for small mixing times, $t < 1$, and large number of samples per molecule, $N > 500$. For small t arguments, few molecules switched from distribution 1 to 2 or vice versa, so there are two distinct classes. For a large number of samples from each molecule (large N) it is easier to distinguish the classes.

The optimal classes for the two class models are not composed of purely distributions $\omega = 1, 2$. Fig. 3-11(b) shows the mixing coefficient for the two classes as a function of t for $N = 500$. The mixing occurs because some molecules have switched distributions. The probability of the two classes is also not symmetric, $\mu^{(\alpha=1)} \neq \mu^{(\alpha=2)} \neq \frac{1}{2}$. The class with a greater contribution from $P_i^{(\omega=2)}$ has a slightly higher weight, because wider distributions better account for variations in the data [241].

154
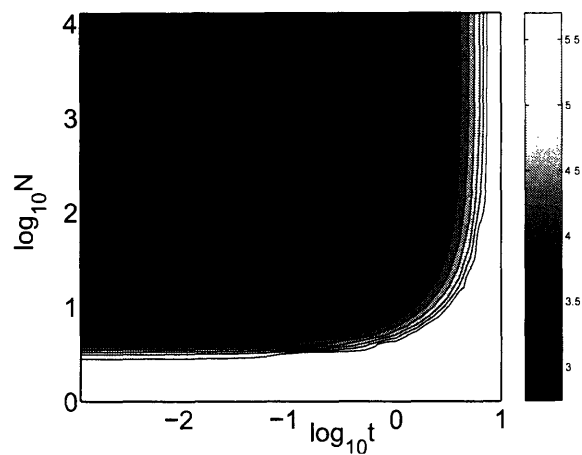
(a)
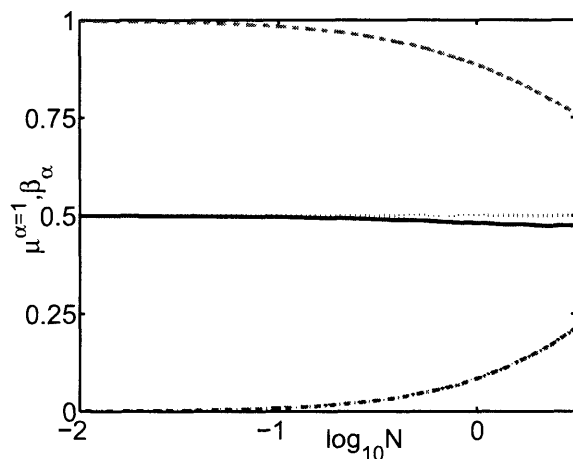
(b)

Figure 3-11: Ability to distinguish classification schemes with a single class from classification schemes with two classes for the random telegraph model discussed in the text. The important factors are the number of measurements per molecule, N, and the mixing time of the model, t. (a) The expected number of single molecule experiments necessary to determine the existence of two classes at the 95% confidence interval. (b) The β_α values of the distributions for $\alpha = 1$ (dot-dashed) and $\alpha = 2$ (dashed) as a function of t for $N = 500$. The solid line corresponds to the weight of the class with a higher contribution from $P_i^{(\omega=1)}$, $\mu^{(\alpha=1)}$, which is compared against $1/2$ (dotted line) to demonstrate the asymmetry.

155

3.20 Classification with physically restricted functional forms

The above model analysis used the histograms to classify the various molecule behaviors. This approach is very powerful at classifying molecules when the amount of data from each molecule is sufficient, even if the number of molecules is small. In some applications the amount of data from each molecule may not be sufficient for histogramming or one wants to make inference about the behavior of a small segment of a single molecule trajectory. A flexible functional form such as sums of Gaussians can replace histograms and compensate for a lack of data for histogramming [140]. The sum of Gaussians is flexible enough to avoid imparting artificial features and the parameters can be easily determined.

In several applications, the physics restrict the possible models so Gaussians may not be the most appropriate functional form. Examples include FRET and single photon counting processes, where the emission statistics follow a Poisson distribution, $P(n) = \frac{\lambda^n}{n!}e^{-\lambda}$ if the photon detection rate is much lower than the excitation rate so that the photons are approximately uncorrelated. Knowing the number of photons in a single bin allows one to construct a probability distribution for λ in each bin. This probability distribution can be used instead of a histogram to determine different classes and or states.

We concentrate on applications to FRET experiments, since application to photon counting is simpler. If we assume that the number of photons in each channel emitted during a single time bin comes from a Poisson processes with constants λ_d for the

donor and λ_a for the acceptor, then the probability of getting $\{m, n\}$ photons from the donor and acceptor respectively is given by

$$P(m, n | \lambda_d \lambda_a) = \frac{\lambda_d^m}{m!} e^{-\lambda_d} \frac{\lambda_a^n}{n!} e^{-\lambda_a}. \tag{3.20.49}$$

Often the intensity, $I = \lambda_d + \lambda_a$, is unstable due to drift in the alignment of the optics, but the relative intensities and resulting FRET value, $F = \lambda_a/(\lambda_a + \lambda_d)$, of the donor and acceptor are stable since detection efficiency is equally affected for both channels. As a result, we need to transform the variables $\lambda_{a,d}$ to the intensity, I and FRET value F. We assume that the intensity $I = \lambda_a + \lambda_d$ is a random variable that must be integrated out, while the FRET values are the characteristic of the class. From Bayes theorem, the probability of a specific FRET value for given number of photons, m and n, is

$$P(F | \{m, n\}) = \int dI \, P(F, I | \{m, n\}) = \int dI \frac{P(\{m, n\} | F, I) P(F, I)}{P(\{m, n\})}. \tag{3.20.50}$$

This expression is derived from assuming that m, n, I, and F are all random variables. For simplicity, we assume $P(F, I) = const$ since we are searching for the best fitting models and do not want the prior to influence the fit, but one may consider other priors [6]. Taking the log of Eq. 3.20.50 allows us to ignore the constant $P(F, I)$. Similarly, $P(\{m, n\})$ is a constant that can be neglected and the important contribution is the the likelihood of having $\{m, n\}$ photons in each channel for a fixed FRET

value,

$$P(F|\{m,n\}) = \frac{\Gamma(m+n+2)}{\Gamma(n+1)\Gamma(m+1)}(1-F)^m F^n, \qquad (3.20.51)$$

which is simply a β distribution. The maximum occurs at $F = \frac{n}{m+n}$, the generally reported FRET value. The probability distribution reflects the stochastic nature of the measurements. The variance in the probability of the FRET value decreasing with increasing $m+n$, since increasing $m+n$ is equivalent to increasing the number of experiments.

Since we have a probability distribution for various FRET values, given a certain number of photons from the donor or acceptor in each bin, we can determine the existence of classes of FRET values. As a simple example, consider a system that toggles between two states, $\pm$ with a fixed intensity, I and FRET values $F_\pm = \frac{1}{2} \pm \Delta F$ with equal probability. The probability distribution for $\{m, n\}$ photons is

$$P(m,n) = \frac{1}{2}\frac{(I(1-F_+))^m(IF_+)^n}{m!n!}e^{-I} + \frac{1}{2}\frac{(I(1-F_-))^m(IF_-)^n}{m!n!}e^{-I}, \qquad (3.20.52)$$

The most likely two class model will be the true model, and the best single FRET state model has $F = \frac{\langle n \rangle}{\langle n \rangle + \langle m \rangle} = \frac{1}{2}(F_+ + F_-) = \frac{1}{2}$, where the $\langle \ldots \rangle$ correspond to the expectation of n and m is over the distribution in Eq. 3.20.52.

The expected number of measurements needed to distinguish the true two state model from the one state model is calculated as a function of I and ΔF in Fig. 3-12. If ΔF is small the two states have similar statistics to the single state and it is difficult to distinguish models so it is not surprising that a large ΔF improves the ability

158
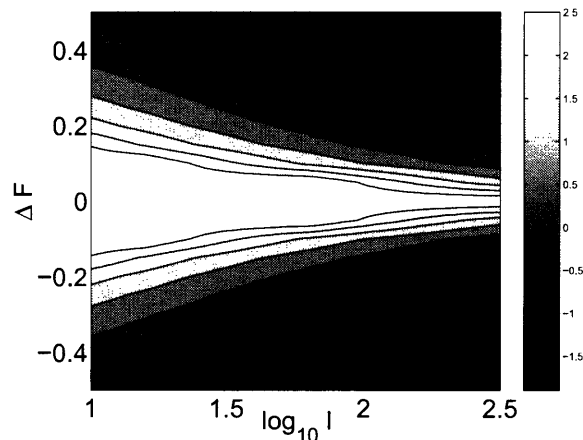
(a)

Figure 3-12: Ability to distinguish a classification scheme with a single class from the true model that contains two classes for the FRET model discussed in the text. The figure shows the number of measurements necessary to establish the existence of two classes as a function of intensity and differences in FRET values.

to distinguish two models. Increasing the total intensity, narrows the distribution of FRET values from each state (See Eq. 3.20.51), which reduces the overlap of the measured FRET values from the two states. The reduced overlap increases the ability to distinguish F_+ from F_- and infer the existence of two distinct classes (as seen in Fig. 3-12).

3.21   Conclusion

Testing the properties of single molecule systems requires both physical insight and rigorous statistical analysis. The physical insight into detailed balance violations lead to several possible features. Once these features and other features such as renewal character were identified, one is able to construct rigorous tests to determine if these features are truly present in the data or if the feature is a simple data artifact. These

159

tests are non-parametric since one does not have to propose a priori models and simply uses the histograms to represent the data in a coarse grained fashion.

The most powerful extension of these methods is to classification of single molecule behaviors. Classification schemes are a natural first step in performing single molecule analysis that allow the identification of different characteristic behaviors between molecules, the establishment of the existence/non-existence of ergodicity in single molecule trajectories, and the determination of states. Having insight into the variations in single molecule behavior is critical in additional analysis by allowing the assessment of the validity of average quantities. For example, the average number of photons in each bin is not very meaningful if the molecules form two distinct classes. These calculations also establish a theoretical limit in the ability to distinguish models and make assessments about the number of classes. Whenever a model is proposed, one should assess whether it is statistically feasible to distinguish this model from others. Using classification to identify molecular states, such as in the FRET example, is also the first step in examining the traditional multiple state kinetics. Rigorously identifying states makes examining transitions between states and inferring connectivities, which are the primary focus of many analyses, possible.

To minimize the information loss, one should examine histogramming several different quantities. In fact, the analysis can be simultaneously applied to several different histograms constructed from the same sequence. For example, if the data shows the system toggling between three states, a histogram of the waiting times for transitions between each of the states can be combined in a single classification procedure.

Classification procedures can flag interesting or problematic features in the data. Once classification procedures give initial insight into heterogeneity in the dynamics of the system, reductions in the possible models of the system are possible. After reducing these models, a more thorough analysis may be possible to find a model of the underlying dynamics.

Chapter 4

Analysis of the entire sequence of a single photon experiment on a flavin protein

4.1 Introduction

Single molecule methods are widely applied to the study of biomolecules [149, 246, 240, 75, 110, 221, 236]. The dynamics of biomolecules revealed by these single molecule techniques are complex with fluctuations on many time scales. Data collected from these experiments are inherently noisy with contributions from background photons, the photon statistics of the system (shot noise), and the stochastic nature of protein dynamics [38, 35, 242, 21, 14, 29]. These stochastic contributions to the data cause difficulties in interpreting single molecule data and necessitate the application of robust statistical methods.

This chapter uses a Bayesian/information theory framework to examine a possible model for a flavin protein (Fre) experiment by Xie and co-workers [240, 224, 4, 218]. The experiment collects single photons emitted from the system and shows photon correlations up to $100ms$ time scales. Maximum entropy analysis (MEM) shows long time multi-exponential relaxation [190, 198]. The MEM fit avoids using a pre-determined functional form so it does not introduce artificial physics through a parametric fit. Since many models may result in the same correlation functions, such as a harmonic diffusion model or a two state model with complex waiting times, the correlation function only contains a limited quantity of information relevant to understanding the mechanisms of biomolecules. Confirmation of the validity of the physical picture motivates the examination of the entire data sequence. By combining physical insight with statistical methods, this chapter shows that modeling the protein's motion as a collection of over-damped diffusive harmonic modes is consistent with the entire data sequence. Although diffusion in multi-dimensional harmonic potentials can be cast into a generalized Langevin equation with a smooth relaxation spectrum for the correlation function of the random force, the picture of a connected network of amino acids has physical appeal since it explains the Gaussian nature of the long lived correlations. This picture is a dynamic analogue of the elastic network models (ENM) used in determining the static root-mean-squared (rms) displacements of functional groups in proteins [7, 126, 68, 85].

This analysis is one of the first to utilize the entire data sequence of a single photon experiment on a single protein to determine a physical picture. Since the model fit incorporates all of the available information, one maximizes the ability to

test the legitimacy of models. The robustness of Bayesian analysis combined with the non-parametric MEM analysis give a complete description of the probed dynamics of Fre.

After discussing the experiment in Sec. 4.2, we perform simple preliminary statistical analysis in Sec. 4.3 to extract the limited information contained in correlation functions. This information is used to explore possible models in Sec. 4.4. We establish the ENM as an appropriate model for the system and discuss the reasons that an N state or trapping model is not a natural choice. In Sec. 4.6, we determine the adequacy of the harmonic diffusion model to fit all possible statistics through Bayesian analysis of the entire sequence.

4.2   Description of experiment

The Fre experiment examines a single flavin protein attached to a cover slip by exciting an electron in the flavin with a repetitive sequence of laser pulses. As shown in Fig 4-1(a), the excited electron can relax through the emission of a photon or through a two step electron transfer process between a nearby tyrosine, tyr^{35}, and the flavin molecule. The kinetic scheme associated with this system is shown in Fig. 4-1(b). The fluorescence rate is $k_f \approx 0.2ns^{-1}$. The first electron transfer rate is a dynamic quantity that fluctuates around $k_{ET}(t) \approx 1.0ns^{-1}$. The second electron transfer rate does not affect the ability to fluoresce and can be neglected. The experiment continually excites the flavin molecule with a pulse train separated by $13.2ns$, (Fig. 4-1(c)). As depicted in Fig. 4-1(c), the experiment detects the first photon and records

the arrival time of this photon, t_i, and the fluorescence lifetime, τ_i. The lifetime is the time difference between the excitation pulse time and the photon arrival time. Fig. 4-1(d) is a record of the lifetime versus chronological time for a short piece of the time sequence. The probability of detecting a photon is proportional to the fluorescence lifetime [240],

$$P_{photon}(t|\tau(t)) \propto \frac{\tau(t)}{t_f} = \frac{(\tau_f^{-1} + \tau_{ET}^{-1}(t))^{-1}}{\tau_f}. \tag{4.2.1}$$

Experiments reveal the exponential dependence of the electron transfer rate on the distance between the flavin and a specific tyrosine, tyr^{35}, $\tau_{ET}(t)/ns = e^{\beta r(t) - \beta r_0} = e^{R(t) - R_0}$, where R_0 accounts for the prefactor and $\beta \approx 1.4 \mathring{A}^{-1}$ is the empirically determined scaling coefficient [240]. Following Xie and co-workers, $\tau_{ET}^{-1} \gg \tau_f^{-1}$ implies

$$\tau(t)/ns \approx e^{(R(t) - R_0)}. \tag{4.2.2}$$

The objective of this chapter is the determination of the equations of motion for this coordinate $R(t)$.

4.3 Analysis with the maximum entropy method (MEM)

To gain insight into viable models for $R(t)$, we visualize the data through one dimensional measurements. The data is preprocessed to remove systematic errors, including monotonic intensity fluctuations in a peak corresponding to scattered photons from
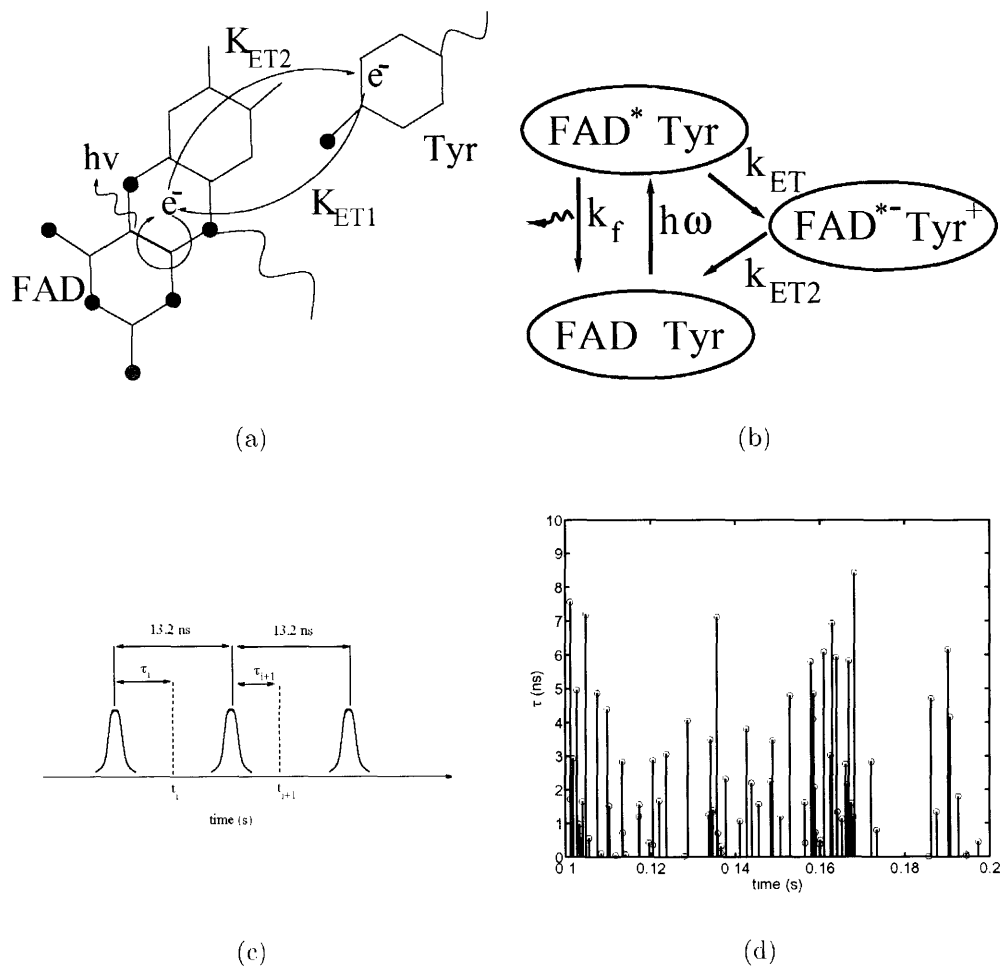
Figure 4-1: (a) The two competing mechanisms for relaxation of an excited electron to the ground state–photon emission and electron transfer. (b) The corresponding kinetic scheme. (c) Schematic of the pulse trail that defines the chronological time, t_i, and the fluorescence lifetime, τ_i. (d) Trace of the lifetime of a photon as a function of chronological time from the experiment.

the laser source and a drift in the zero time baseline for the lifetime. Then the photons are binned in $1ms$ time bins. After preprocessing, measurements with different segments of the sequence are consistent (stationary) and the background measurements show no correlations.

4.3.1   Static lifetime distribution

The photon statistics are complicated by the background counts contributing over half of the photons ($\approx 58\%$), and by the photon's lifetime, τ_i, being a random variable that depends stochastically on the lifetime of the system

$$P_s(\tau_i|\tau_{ET}(t)) \approx \tau_{ET}^{-1}(t)e^{-\tau_i/\tau_{ET}^{-1}(t)}, \tag{4.3.3}$$

which is convolved with the instrument response. These complications necessitate the examination of several averaged measurements to develop insight into possible models for this system.   The first averaged measurement is the static distribution of the fluorescence lifetimes. We perform this measurement for both the experiment and the background measurement to determine the typical lifetimes of photons emitted by the chromophore. The photon lifetime distributions for the experiment and background measurements are histogrammed in Fig. 4-2. Fig. 4-2 also shows maximum entropy fits (MEM) to sums of exponentials [190, 198]. MEM attempts to balance the ability to fit the data with the desire to have a featureless spectrum. This fitting scheme does not impose an arbitrary functional form on the data.

The distributions and MEM fits of the single molecule experiment and background

168
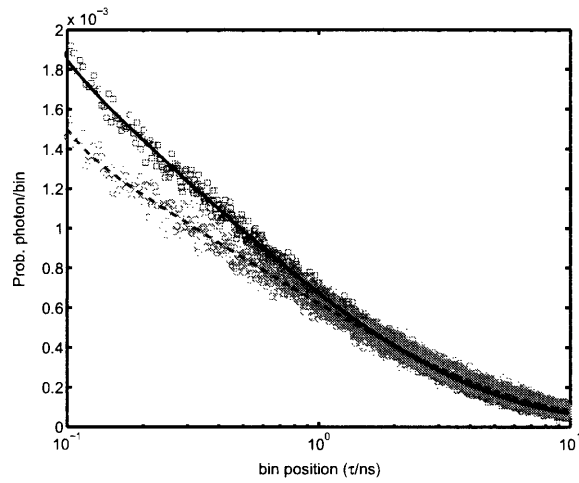
(a)
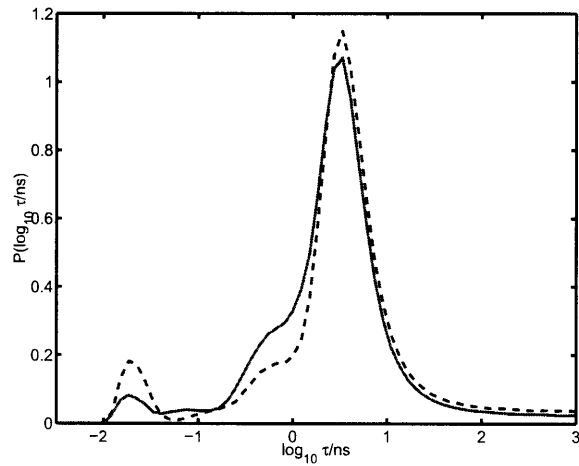
(b)

Figure 4-2: (a) The maximum entropy method (MEM) fit to the experimental (squares=data, black solid line=fit) and background photon lifetime measurements (circles=data, grey dashed line=fit). The two curves show similar long lifetime behavior, but differ in the intermediate times. The scatter in the data gives a good indication of the error bars. These error bars are not plotted for visual clarity. (b) The MEM spectrum for the fits. The fits differ in the amplitudes of the intermediate time shoulders, $0.1 \leq \tau/ns \leq 1.0$.

measurement differ quantitatively, but no strong features differentiate the two lifetime distributions (See Fig. 4-2). Both MEM fits show a small peak at short lifetimes, $\tau/ns \ll 1$ and a broad peak at $\tau/ns \approx 10$ caused by the instrument response with a shoulder at shorter lifetimes. The shoulder is larger for the single molecule experiment's data and indicates that many of the photons from the protein occur on the short time side of this peak, $0.1 < \tau/ns < 1.0$.

4.3.2   Intensity correlation function

To gain insight into the dynamics of the system, we examine the $1ms$ discretized trajectory to determine correlations between the number of photons in each bin. The objective of this chapter is to relate these temporal correlations in the intensity to the underlying dynamics, which are dominated by the fluorophore-quencher distance, $R(t)$ [240]. If the fluorophore-quencher distance, $R(t = j\Delta t) = R_j$ is constant over the $\Delta t = 1ms$ bin, the number of photons is Poisson with parameter $\lambda(j) = \lambda_b + \lambda_s(j)$,

$$P(n|R_j) = \lambda(j)^n/n!e^{-\lambda(j)}, \qquad (4.3.4)$$

where λ_b accounts for the background counts, and $\lambda_s(j) = A_0 e^{R_j}$ with prefactor A_0. Instrument considerations slightly modify these expressions.

For a Poisson process, the second moment for the number of photons in any two bins, i and j has the form

$$M_\lambda(j,k) = \langle n_j n_k \rangle = \langle \lambda(j)\lambda(k) \rangle + \delta_{jk} \langle \lambda \rangle \qquad (4.3.5)$$

170

with Kronecker delta δ_{jk} denoting an additional white noise term and $\langle \ldots \rangle$ is the expected value. After preprocessing to remove systematic instrument errors, the data is translationally invariant and allows time averaging. Subtracting the squared average,$\langle \lambda \rangle^2$ and $\delta_{jk}\langle \lambda \rangle$ gives the λ-correlation function,

$$C_\lambda(j) = \frac{M_\lambda(0,j) - \delta_{0,j}\langle \lambda \rangle - \langle \lambda \rangle^2}{M_\lambda(0,0) - \langle \lambda \rangle - \langle \lambda \rangle^2} \qquad (4.3.6)$$

plotted in Fig. 4-3(a).

To identify the time scales in this system, we perform a maximum entropy fit (MEM). The MEM fit avoids using parameterized models that can hide certain features in the data. The resulting spectrum reveals three time regimes (Fig. 4-3(c)) The fastest time scale corresponds to correlations that fall off within a few time bins less than 20 ms. This time scale appears to be broadly distributed since few data points contribute to the short time dynamics. A less broadly distributed second time scale decays around 50-100 ms, and a third narrowly distributed time scale decays around 400 ms. The narrow distribution at long times demonstrates an exponential decay of the correlation function at longer times. A fit to the Fractional Gaussian Noise model (FGN) to the first 3000 data points shows poor agreement, especially for the short time behavior ($t < 100ms$) [120] (see Fig. 4-3(a)). The $\chi^2 = N^{-1} \sum (x_i - \mu_i)^2 / 2\sigma_i^2 = 1.3$ for the FGN model, which shows that the fit is outside the 95% confidence interval, $\chi^2 = 1$, used to choose the MEM solution. The large number of data points in the tail force the FGN model to neglect the short time correlations. The parameters predict that the mean squared (MS) fluorophore-quencher

(a)

(b)

(c)
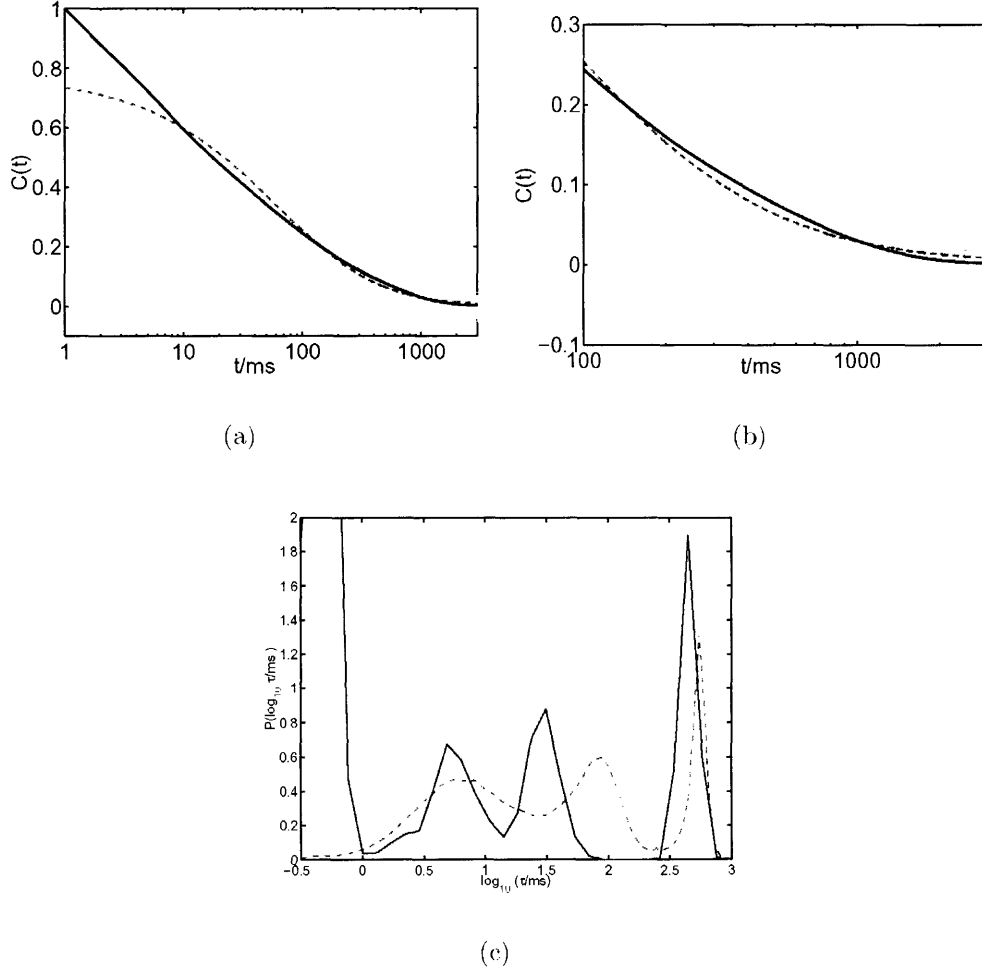
Figure 4-3: (a) Comparison of the MEM fit (solid line) to the FGN model (dashed line). The FGN model correlation function has the form, $C_\lambda = A(e^{C_{FGN}(t)}-1)$ with $C_{FGN} = B(\sum_n \frac{(t/\tau)^{\gamma n}}{\Gamma[n\gamma+1]})$. The FGN fit gives $\tau = 263ms$, $\gamma = 0.84$, $B = 2.32$ and $A = 0.084$. (b) The tail of (a). (c) Comparison of the MEM spectrum (dashed) to the Bayes spectrum averaged over the MC data(solid). Except for the fast fluctuations that cannot be captured by the simple correlation analysis, the Bayesian peaks overlap with the peaks in the MEM spectrum, showing that they are consistent. We normalize the spectrum to the $1ms$ bin contribution since the zero time correlation cannot be accurately measured.

distance fluctuation is approximately $1.2\mathring{A}^2$, which is significantly larger than experimental measurements [240, 104]. This large (MS) displacement reduces the amplitude of the long time power-law tail to less than 2% of the correlation function. As a result, exponential long time relaxation fits the data better, but FGN has reasonable long time agreement (See Fig. 4-3(b)). As discussed in Sec. 4.6, multiple diffusive harmonic modes are consistent with this MEM spectrum and give MS displacements of $0.32 \pm 0.021\mathring{A}^2$, which agree with experimental measurements. It is possible to achieve a better fit with a Fractional Gaussian noise model to the first few hundred data points (up to tenths of a second), but not the entire time range of interest. We emphasize that the MEM fit does not assume a functional form and favors a less structured relaxation spectrum, such as a power-law or stretched exponential, over a structured spectrum. As a result, one should have confidence that the data reflect these structures.

The millisecond motions like those captured by the MEM analysis have been observed in several fluorescence and NMR experiments and have been attributed to loop rearrangements, breathing motions in β sheets, rigid body motions of α-helices, and internally hindered rotations [126, 68, 144]. These low frequency millisecond motions often play pivotal roles in a protein?s function so the ability to resolve and model these motions is important [67]. Although the tyr[35]-flavin distance may not play a key role in functionality, other motions coupled to this displacement may.

The MEM analysis suggests that any physical model for the tyr[35]-flavin coordinate must reflect both the small scale fluctuations of $R(t)$ and the structured relaxation spectrum. It is also important to capture the non-Markovian fluctuations in the

intensity and fluorescence lifetimes (as demonstrated in Fig. 4-1(d)). The model must also account for the distribution being stationary after preprocessing the data since no aging effects are present. Armed with these insights we are now able to explore physically feasible models for this system in Sec. 4.4.

4.4   Slow motions in proteins: N state models, trapping models, and ENM

Many candidate models can reproduce the intensity correlation function and the lifetime distribution, so the other physical attributes discussed above also need to be considered in selecting a model. The physical basis of a model depends on the level of coarse graining required to capture the essential physics of a system. This point is illustrated by the hierarchical tier picture of protein energy landscapes/surfaces (PES) [83]. The potential energy landscape is high dimensional and complex with motions on many length and time scales. The motions on tier m are generally faster than motions on tier $n > m$, but slower than motions on tier $n < m$, and time separation arguments generally apply. If the motions that we are monitoring occur at tier m, we can homogeneously average over the degrees of freedom associated with the lower tiers $n > m$ and need to perform a quenched average over the higher tiers $n < m$. The quenched average would result in heterogeneity in the behavior of single molecules. As shown in Fig. 4-4, averaging over faster time scales results in a free energy potential instead of a detailed microscopic potential [68]. Considering that

174
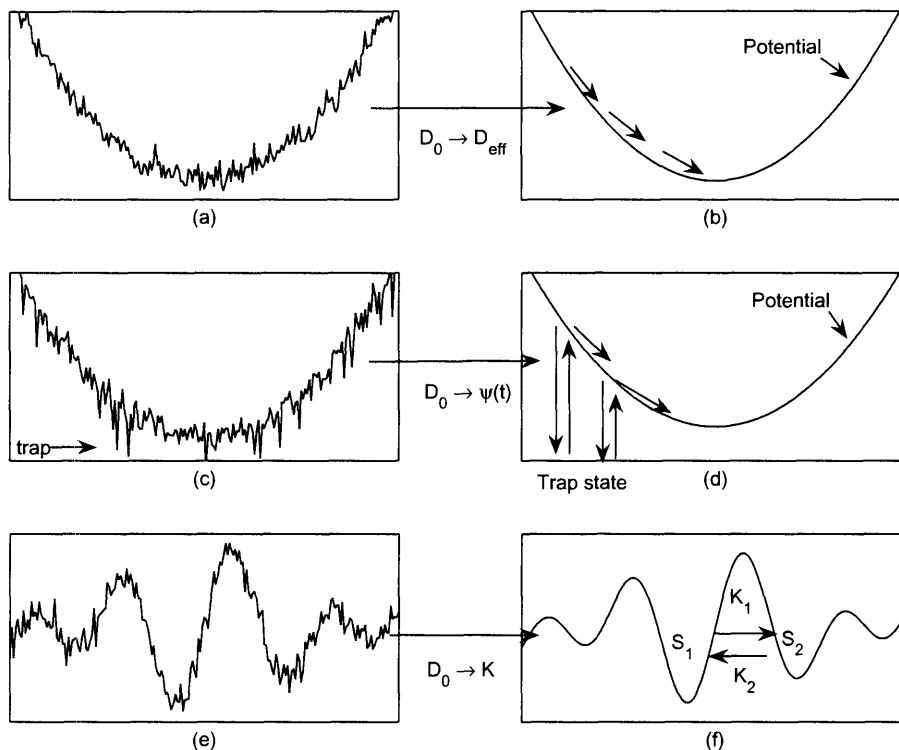
Figure 4-4: Caricature of the different potential energy surfaces (PES) after averaging over faster time scales. (a) A PES with a global curvature and Gaussian roughness. Homogeneously averaging over the smaller length scales results in normal diffusion with a renormalized diffusion constant, $D_{eff} = D_0 e^{-\langle(\beta V)^2(0)\rangle/d}$, as shown in (b). (c) A PES with an exponential distribution of traps, $P(E^\dagger) = \alpha e^{-\alpha E^\dagger}$, which results in deep spikes. Averaging over the smaller length scales replaces normal diffusion by a trapping process with a waiting time $\psi(t)$ as depicted in (d). (e) A more complicated PES with several local minima. Averaging over small length scales results in an N state model with Poisson kinetics depicted in (f).

the experimental time scales of the Fre experiment range from milliseconds to tenths

of seconds, these motions occur on the slowest time scales of the protein (no tiers

$n < m$), so we do not expect additional slower motions that must be heterogeneously

averaged over.

Three models that result from different coarse graining procedures include the N

state model, the trapping model, and the harmonic diffusion model. These models

originate from different topologies of the protein potential energy surface. As elabo-

175

rated below, the N state model results from the timescale of interest corresponding to motions over high energy barriers, while the trapping model corresponds to hopping over many smaller structures. For the harmonic diffusion model, the smaller scale structures result in a diffusion tensor.

4.4.1 The N state model

The N state model results from the tier of interest containing multiple minima separated by high barriers, (See Fig. 4-4(e) and (f)) Averaging over the faster degrees of freedom results in Kramer's barrier crossing kinetics,

$$\dot{P}_{R(t)=R_i} = -\mathbf{K}P_{R(t)=R_i}, \tag{4.4.7}$$

where $P_{R(t)=R_i}$ is the probability that the particle is in minima i with a corresponding tyr^{35}-flavin distance of R_i. One may also add fluctuating barrier heights that make $\mathbf{K}(t)$ time dependent [1]. The presence of three relaxation time scales suggests that a minimum of four states is necessary, but some of the kinetic rates would be slow, $\tau \sim 100ms$ and one would expect deviations from Gaussian behavior that are not seen in the data. Instead, the data analysis below suggests the additional non-Markovian fluctuations in $R(t)$ can be captured by a Gaussian model, although a model with both barrier crossing kinetics and intra-well relaxation may also be viable. These additional non-Markovian fluctuations can be captured by the inclusion of additional states. If enough states are included, the N state model can approximate any other model, but attempts to fit the data through complete sequence analysis with a reasonable

number of states, $N > 6$, did not achieve a desired fit to the data. The identity of these states is also ambiguous since the apparent fluctuations in the chromophore quencher distance is rather small, on the order of tenths of an $\mathring{A}$, compared to the larger scale motions that an N state model attempts to capture [240], although these motions may simply have a weak projection onto the probed coordinate.

4.4.2 The trapping model

Unlike the N state model, where the barriers that dominate the dynamics occur on the same tier as the motion of interest, the trapping model has important contributions from the smaller scale motions (See Fig. 4-4(c) and (d)). These small scale structures trap or hinder the motions of the coordinates of interest and the fluorescence lifetime becomes static for long periods of time. Fractional diffusion is an extreme example of this scenario, where the traps have energetic barriers that are exponentially distributed for large energy barriers, $P(E^\dagger) \sim \alpha e^{-\alpha E^\dagger}$ [150]. The exponential decay of the energy barrier distribution is the result of extreme value arguments with a strong emphasis on the functional form in the tail of the distribution. This formulation is hindered by the tails being slow to converge to the universal form [150]. The scenario results in a long time power-law decay and aging effects that are not seen in the data or the MEM fit [150]. We examined truncating the distribution of energy barriers, $P(E^\dagger)$, but this truncation imposes an interrupted aging effect that removes contributions from the short time trapping behavior and prevents the inclusion of a broad distribution of time scales in the stationary correlation function [227]. The trapping

models also depend on the system being large so that correlations in the trapping times can be ignored, which probably does not apply to finite sized proteins. As a result, the data do not support a trapping model.

4.4.3 Dynamic elastic network model (ENM)

If the barriers are not high within the tier of interest, the system demonstrates a diffusive behavior. Averaging over faster degrees of freedom results in a smooth convex free energy landscape (See Fig. 4-4(a) and (b)) [112, 7], and the slow motions of the protein can be approximated by diffusion of a collection of independent harmonic modes whose correlation function is a simple exponential $\langle R_\mu(t)R_\nu(0)\rangle = \delta_{\mu\nu}\frac{a_\mu^2}{b_\mu^2}e^{-t/\gamma_\mu}$. The motion of interest is a weighted sum of modes, $R(t) = \sum_\mu b_\mu R_\mu(t)$ and the process becomes Gaussian with correlation function $\langle R(t)R(0)\rangle = \sum_\mu a_\mu^2 e^{-t/\gamma_\mu}$. By defining $\rho(\gamma) = \sum_\mu a_\mu^2 \delta(\gamma - \gamma_\mu)$, we can define a relaxation spectrum and write

$$\langle R(t)R(0)\rangle = C_R(t) = \int d\gamma \rho(\gamma)e^{-t/\gamma}. \tag{4.4.8}$$

The probability distribution of $R(t)$ becomes a simple functional integral

$$P(\{R(t)\}) = \frac{1}{\sqrt{Det(2\pi C_R(t))}}e^{-\frac{1}{2}\int dt_1 dt_2 R(t_1)C_R^{-1}(t_1-t_2)R(t_2)}. \tag{4.4.9}$$

This model can be cast into an N state model with many states, but the kinetic rates have simple relationships with each other, which reduces the number of parameters. Unlike the trapping model, which puts strong emphasis on small scale fluctuations, the

178

harmonic diffusion model accounts for small scale motions through the more modest means of a renormalized diffusion tensor. The diffusion tensor for a Gaussian distribution of energy barriers can be shown to be isotropic with $D_{eff} = D_0 e^{-\langle \beta^2 V(0)^2 \rangle / d}$ [55], where β is the inverse temperature and $V(0)$ is the random potential. With this assumption for $R(t)$, the fluorescence intensity correlation function is approximately

$$C_\lambda(t) = A(e^{C_R(t)} - 1) \sim AC_R(t) + AC_R^2(t)/2! + \dots \qquad (4.4.10)$$

where A is a prefactor that accounts for A_0 and R_0 and the asymptotic relations result from the decay of the correlation function. This expression would be exact, except for experimental considerations such as instrument response and the finite detection window, but these deviations are small.

Large scale coarse grained normal modes have been used to study the dynamics and statistics of several systems [68, 10, 179]. The fast motions in proteins deviate from this model's simple harmonic motion [85], but the larger scale motions have been found to be independent of small scale anharmonicities due to laws of large numbers and motional narrowing effects [7, 187]. A Gaussian model is also the minimal information model consistent with the measured observables–the correlation function [170]. Even if deviations from harmonicity exist, the current data are not able to resolve characteristics of these deviations. As a result, the important features are characteristic relaxation times and the magnitude of the displacements of $R(t)$, which can be captured by a harmonic approximation.

The spectrum of the correlation function

The MEM analysis showed three time scales contributing to the system, which suggests that the major contribution to the long time relaxation of the system can be captured by a harmonic diffusion model with three relaxation time scales. Concurrent with our work, Kou and Xie suggested a harmonic diffusion model with a power-law distribution of relaxation times, which they capture with the Fractional Gaussian noise model [120]. This model is very interesting, but even ignoring the difficulty in fitting the dynamics, its applications to proteins needs to be justified.

Gaussian noise assumptions have two possible sources. An unlikely source for this protein system is the bath being much faster than the time scales of interest, which leads to multiple collisions and large number arguments. In this case, the structure of the bath does not matter, but long lived correlations in the random force cannot be introduced. The other possibility is that the bath has an intrinsically harmonic structure, as argued above for coarse grained descriptions of the protein. This coarse graining will add many time scales to the relaxation spectrum, and the major issue becomes the expected structure of the spectrum of relaxation times.

Power-law spectrum and scaling

A power-law spectrum has been suggested for this system [120]. This power-law may be the result of the protein showing a self similar structure. The simplest self similar structure is the Rouse polymer chain of N beads at positions x_n with local connectivities [189]. The beads undergo diffusion in a potential of the form $V_{Rouse} =$

$\sum_{i,a=x,y,z} \frac{k}{2}(a_i - a_{i+1})^2$. In the large monomer limit, $N \to \infty$, the eigenmodes $a_\omega(t) = \frac{1}{N}\sum e^{i\omega n} a_n(t)$ have a correlation function

$$C_\omega(t) = \langle a_{-\omega}(t)a_\omega(0) \rangle \approx \frac{D}{2k(1 - cos(\omega))}e^{-2k(1-cos(\omega))t}, \qquad (4.4.11)$$

where D is the diffusion constant, k is the spring constant of the Rouse chain, the friction coefficient $\zeta = 1$, and the end effects were ignored [231, 189]. The correlation function for $a_n - a_m$ is

$$C_{nm}(t) = \langle (a_n(t) - a_m(t))(a_n(0) - a_m(0)) \rangle =$$

$$\int \frac{d\omega}{2\pi} \frac{D}{k} \frac{1-cos(\omega(m-n))}{1-cos(\omega)} e^{-2k(1-cos(\omega))t} \sim \frac{D}{k}(m - n)^2 e^{-2kt} I_0(2kt) \qquad (4.4.12)$$

which decays as $t^{-\frac{1}{2}}$ (See Fig. 4-5) [189]. More generally, the lattice could have a self similar fractal structure, and the correlation would be expected to asymptotically decay as $t^{-d/2}$, where d is the fractal dimension. The power-law arises from the scale invariance of the system. By only having local connectivities, there is a translational invariance (if end effects are ignored). As a result, the relaxation times correspond to differing length scales resulting in the power-law. The power-law results will not be altered by adding bending rigidity or any other local interaction since the translational invariance will still hold in the large N limit.
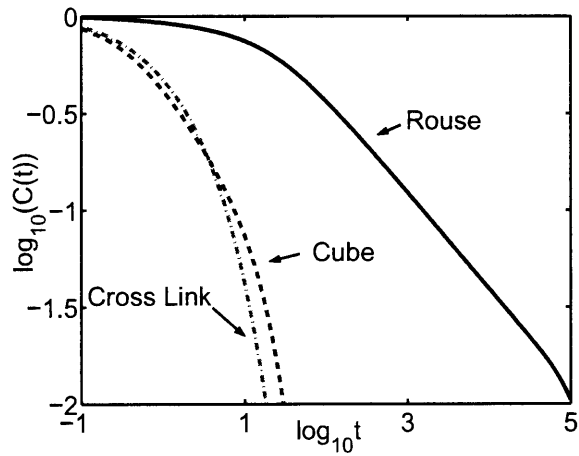
(a)
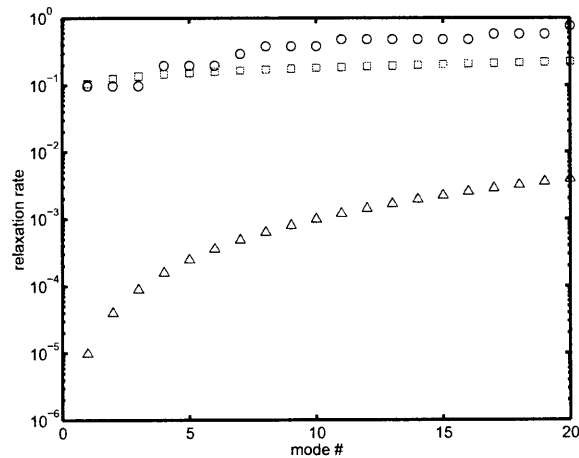
(b)

Figure 4-5: (a) The correlation function for $m = 493$ and $n = 507$ for a Rouse chain with 1000 beads, and the same correlation function averaged over random cross-linking of the Rouse chain polymer of 1000 beads with approximately 1 cross link per polymer. These correlation functions are compared with the correlation function for the opposite corners, $i = j = k = 1, 10$, of a $10 \times 10 \times 10$ cube. The cube also has 1000 beads, but it is a small object. The diffusivity, D and the force constants for all connected beads, k, are assumed to be unity $D, k = 1$. Except for the Rouse chain, the smallest relaxation rates $\lambda_1 \approx 0.1$, and the contribution of long time exponential relaxation is significant. (b) The lowest eigenvalues for the Rouse (triangle), cross-linked polymer (square), and the cube (circle).

Non-local contacts, finite sizes, and non-scaling behavior

Proteins have additional interactions that are non-local with respect to the position along the protein sequence so these self similar structures are probably not a good model for a protein. These non-local interactions destroy the scale invariance so a power-law spectrum cannot be universally applied to proteins. Although proteins may be large objects in terms of the one dimensional sequence, the three dimensional structure is much smaller than a crystal so edge effects destroy the scale invariance necessary for a power-law. A protein with one thousand amino-acids (Rouse beads) would only form a three dimensional structure around ten amino-acids in length (on the order of nanometers) [105].

The smallest eigenvalues for the relaxation spectrum of a 1000 bead Rouse chain show a near power-law behavior with very small eigenvalues, whereas the spectrum of an elastic body that is a 10 unit cube (also 1000 beads) shows a lower bound in the relaxation rates, (see Fig. 4-5). The cube undergoes the same diffusion process as the Rouse chain, but the potential for the cube is

$$
V_{cube} = \sum_{i,j,k=1...n;a=x,y,z} \frac{k}{2} \left[\left(a_{i,j,k} - a_{i+1,j,k} - a^0_{i,j,k,i+1,j,k} \right)^2 + \right.
$$
$$
\left. \left(a_{i,j,k} - a_{i,j+1,k} - a^0_{i,j,k,i,j+1,k} \right)^2 + \left(a_{i,j,k} - a_{i,j,k+1} - a^0_{i,j,k,i,j,k+1} \right)^2 \right], \quad (4.4.13)
$$

where $a^0_{i,j,k,i,j,k+1}$ denotes the equilibrium distance. In the large n limit, this expression is also exactly solvable with a t^{-1} power-law dependence, but the edge effects for $n = 10$ are quite significant. If force constant k and the diffusion constant D are assumed

to be unity (unit-less time and distance), the cube's spectrum (excluding rigid body motions) is bounded away from zero, $\lambda_1 \approx 0.1$, and long time exponential relaxation is expected. Even a less structured connectivity such as the average eigenspectrum of a randomly cross-linked Rouse chain (1000 beads) shows an eigenspectrum that is bounded away from zero (Fig. 4-5) [31]. In this model the potential is random with $V_{cross-link} = V_{Rouse} + \sum_{i>j;a=x,y,z} p_{ij} \frac{k}{2}(a_i - a_j)^2$, where $p_{ij} = 0,1$ with probabilities $(N-1)/N$ and $1/N$, respectively. This random linking results in a collapsed structure so finite size effects are expected [31]. Proteins have specific non-local connections to allow it to perform its function. As a results, one expects a greater variety of possible behaviors, which leads to a greater variety of possible relaxation spectra. As a result, it should not be surprising if the relaxation spectrum of proteins shows structures that are unique to the protein.

4.5   Implementation of Bayesian statistics to ENM

If the harmonic diffusion model is correct, the MEM analysis suggests that harmonic diffusion with three distinct time scales can capture the long time relaxation of the data. One can model the three time scales with three modes without introducing statistically significant errors. The remaining issue to address is verification of the consistency of the model with the data set as a whole. Additional coarse grained measures, including multiple time correlation functions, are too noisy to assess the model adequately, but the results are consistent with a harmonic diffusion model for the motions of $R(t)$. Concurrent with our work, Kou and Xie also demonstrated

this consistency by examining these averaged quantities so we will not go into detail about these tests of the Gaussian hypothesis [120]. To strengthen the legitimacy of harmonic diffusion a full sequence Bayesian analysis is necessary.

The Bayesian analysis is implemented by fixing the number of modes, M and performing a Monte Carlo (MC) simulation to sample the parameters of the model (the weights a_μ, the relaxation times, γ_μ, and auxiliary parameters such as A and R_0) that determine the statistics of the system. For given parameters, the probability of having n_j photons with lifetimes $\tau_{j1} \ldots \tau_{jn}$ in bin j for $R(t) = R_j$ is computed,

$$P(\{\tau_{j1} \ldots \tau_{jn}\}|R_j) = P(n|R_j) \prod_{m=1\ldots n} P(\tau_{jm}|R_j) \qquad (4.5.14)$$

where $P(n|R_j)$ is the probability of getting n photons as defined in Eq. 4.3.4 and $P(\tau_{jm}|R_j)$ is the probability of the lifetime of the photons given R_j,

$$P(\tau_{jm}|R_j) = \frac{\lambda_b}{\lambda_b + \lambda_s(j)} P_b(\tau_{jm}) + \frac{\lambda_s(j)}{\lambda_b + \lambda_s(j)} P_{sys}(\tau_{jm}|\tau_{ET} = e^{(R_j - R_0)}). \qquad (4.5.15)$$

Both sources of photons, the system, P_{sys}, and the background, P_b, are accounted for in this expression. The exact form of these probabilities is complicated by convolution with instrument response and other instrument considerations. Without these considerations P_{sys} is equal to P_s in Eq. 4.3.3. Given the probabilities of the photon emission events for all R_j, the mode positions $R_\mu(t = j\Delta t)$, are varied by randomly choosing one mode and statistically choosing its positions, $\{R_\mu(t = j\Delta t)\}$ keeping the other modes fixed through a forward-backward algorithm [211].

From this simulation we estimate the Bayesian score (log of the probability that the model produced the data) to determine the optimal parameters [211]. The score includes how well the sampled paths fit the data and the probability that the paths are produced by the diffusion model. The fit to the data is estimated from the log of Eq. 4.5.14 for the selected sequence $\{R_j\}$, and the fit to the model is estimated from the Fourier components of $R_\mu(t = j\Delta t)$, $\sum R_\mu(t = j\Delta t)e^{i\omega j}$. The Bayesian score was computed for different number of diffusive harmonic modes and compared to determine the appropriate number of modes.

4.6 Results

The simulation found that a fourth mode is necessary to account for fast fluctuations that are not consistent with the stochastic fluctuations (including the background) and would not be represented in the correlation function. The time constants and weights of the four oscillators are $a_\mu^2 = 0.595 \pm 0.023, 0.293 \pm 0.040, 0.292 \pm 0.028$, and 0.324 ± 0.041 for $\gamma_\mu = 0.42 \pm 0.10, 5.9 \pm 2.8, 28.0 \pm 8.2$ and $400. \pm 57$, respectively. The exponential components discovered by the Bayesian simulation fall into the time scales revealed by the maximum entropy fits, which shows that harmonic diffusion agrees with the basic features of the data and the Bayesian approach identifies the important time scales. The Bayesian spectrum is determined by averaging C_λ over the Monte Carlo simulation and is compared against the MEM simulation in Fig. 4-3(c).

For $\beta = 1.4\mathring{A}$, the mean square (MS) displacement of $r(t)$ is $\langle r^2(t) \rangle \approx 0.32 \pm 0.02\mathring{A}^2$, which is in agreement with other measurements [240, 104]. Crystal structure

data show that tyr^{35} has a MS displacement of $0.25\mathring{A}^2$ and that the isoalloxazine portion of FAD has a MS displacement of $0.10\mathring{A}^2$, so fluctuations around $0.35\mathring{A}^2$ are expected. For four modes, we predict that the average lifetime is around $\langle \tau \rangle \approx 0.310\pm0.011ns$. Since the number of photons emitted depends on $\tau(t)$, we must weigh the probability of $\tau(t)$ by the expected number of photons given $\tau(t)$ to determine the average lifetime of a photon emitted by the FAD, $\langle \tau_{photon} \rangle \approx 0.410 \pm 0.028ns$. As expected, this distribution suggests that the photons from the system occur on the short time shoulder peak of the MEM distribution in Fig. 4-2(b).

Additional modes, beyond four, slightly improve the fit to the data and the correlation function, but the improvement can not be justified statistically. For less modes, the paths selected by a two and three mode models have similar likelihoods to the paths of the four mode model, but the probability of these paths being produced by the harmonic model was much lower. In other words, following the variation in the data with less than four modes resulted in unlikely paths. The Fourier components of the sequence for the four mode model, $\sum R_\mu(t = j\Delta t)e^{i\omega j}$, are within the expected variances of the model so harmonic diffusion is consistent with the paths that fit the data. Similarly, the photon emission events are consistent with the model. The Bayesian scores and parameters for four oscillators show time translational invariance so the model is not over-fitting the data.

4.7 Conclusion and Discussion

This chapter examines a single photon experiment with a complex data set that is difficult to interpret from correlation analysis. Non-parametric fits by the MEM demonstrate a wide distribution of time scales with distinct structures in the relaxation spectrum that are neglected by phenomenological fits using smooth predetermined functional forms. The evidence for these structures (especially at long times) is strong so it is appropriate to discuss distributed lifetimes, but the existence of a stretched exponential, power-law, or other phenomenological functional form cannot be fully supported by the data. This experiment demonstrates the importance of introducing non-parametric methods into single molecule data analysis and the need for caution in interpreting model features such as power-law tails since models without these features may also be consistent with the data so it may be difficult to assign a physical meaning to the predicted power-law. From the non-parametric analysis of coarse grained measures, such as correlation functions, one can develop legitimate models to describe the behavior of the system. Although models should be consistent with the correlation analysis, correlation functions are only one or two dimensional information and cannot be the sole determiner of the properties of the system and generally fails to distinguish different models. The desire for a comprehensive test motivated the use of Bayesian methods in analyzing the entire data sequence. These tests are more time consuming than simple correlation analysis, but the conclusions are more reliable [120]. Through a complete sequence analysis on a single molecular trajectory, this chapter demonstrates that diffusive harmonic network with a few well

defined long time scales is a viable candidate for describing this system. These slow time modes may correspond to motions that influence protein structure and function.

Proteins are specific entities that perform specific tasks. The complexity of proteins may cause a broad distribution of time scales, but it is important to understand how the motions are associated with the specific tasks of the protein. Coarse grained diffusive normal modes incorporate the universality of large numbers by averaging over small scale fast fluctuations while maintaining features that are specific to the protein's structure and function. The fact that coarse grained models and simulations can capture these slower timescale motions while also being computationally tractable is a desirable feature that may allow comparison of simulation to experiment.

Chapter 5

First-principle path integral study of DNA under hydrodynamic flows

5.1 Introduction

The previous chapters mainly addressed the use of statistical techniques to make inferences about data without imposing artificial physics on the data. These methods are especially valuable for complex systems, such as proteins, where it is difficult to ascertain parameters and cross validate experiments. For more homogeneous systems, such as polymer systems, a first principles approach is possible since there are few parameters and these parameters can be independently verified. These systems also present opportunities to perform several different experiments and cross-validate results.

The mechanical properties of polymer systems are also important in many applications, including lubricants and plastics. The bulk visco-elastic properties of these

191

systems result from the microscopic deformation of the polymer chains when they are subject to external forces. This microscopic-macroscopic correspondence has generated interest in studying polymeric solutions at the microscopic level, including experiments at the single molecule level. Many single molecule experiments examine the behavior of single polymers, like DNA, subject to various stresses, including tensile stress and hydrodynamic flow [26, 199, 42, 9, 161, 163, 162, 173, 193, 192, 191]. One set of experiments by Chu's group visualizes fluorescently labeled DNA subject to the stresses discussed above. The experiment monitors the entire contour in real time and gives a complete picture of the polymer dynamics [161, 163, 162, 173, 193, 192, 191, 204, 97, 188, 8]. In this paper, we model DNA as a worm-like chain (WLC) with parameters previously determined in force-extension experiments and then use this model to examine the experimental results of Chu for the steady state configurations of DNA subject to constant plug, elongational, and shear flows [9, 161, 163, 162, 173, 193, 192, 191]. The ability to use parameters from one experiment to model different experiments confirms the validity of the WLC model for DNA.

This DNA system has also been the subject of many Brownian dynamics simulations [5, 69, 71, 70, 102, 108]. Predicting the properties of a complete contour requires a large number of long simulations performed on a large number of beads. With carefully chosen parameters, reasonable agreement between the experiments and these simulations exists, but these calculations are phenomenological and computationally intensive [69]. We propose a less computationally intensive first-principles path integral Monte Carlo algorithm based on equilibrium theory to study DNA subject

to various hydrodynamic flows. If the relaxation of the polymer is fast, the experimentally observed configurations correspond to a generalized equilibrium distribution, which has an associated potential. With a reasonable formulation of the potential, we evaluate time averaged ensemble quantities with Monte Carlo techniques. Although this approach will not describe dynamics, it is computationally more efficient than Brownian dynamics and allows the prediction of important time averaged quantities, like the mean extension of the polymer that we examine in this letter. Previously, Larson and Chu, as well as others, used similar Monte Carlo techniques on bead and spring models to describe time averaged properties of DNA polymers subject to constant plug flow [131].

DNA is a difficult polymer to model because it is semi-flexible with a large persistence length of 53 nm, compared to a typical length studied in experiments of 50 μm, and its contour length only extends under strong forces [69, 131]. Bead and spring simulations require a large number of beads to account for bending energy and constrained dynamics to maintain the contour length. We adopt a more natural model for DNA, the WLC model of Kratky and Porod [62]. The WLC replaces the Rouse-like bead and spring model with a continuous contour of fixed length and includes an energy associated with bending the polymer. The bending energy is experimentally measurable, which removes a fitting parameter. The first analytic treatment of the WLC model appear in a 1973 paper by Fixman and Kovac [81]. Later work by Marko and Siggia with improvements by Bouchiat demonstrate that the WLC model predictions agree extremely well with the force extension experiments on DNA by Bustamante's group [33]. The agreement suggests that the WLC captures the fun-

damental thermodynamics of DNA [142, 26, 193]. The WLC is also the basis for the force extension relations used in simulations by Doyle and in the analytic theory by Zimm [69, 248]. Doyle neglects hydrodynamic interactions while Zimm incorporates the interactions with a length dependent rescaling of the flow field determined by the Kirkwood approximation.

In the absence of an external field, the energy of the WLC is a simple contour integral, $\beta E = \int_0^L \frac{A}{2}|\partial_s \hat{\mathbf{t}}|^2 ds$, where L is the polymer's length [81]. The inverse temperature, β makes all quantities unitless and the tangent vector $\hat{\mathbf{t}}$ of the contour $\mathbf{R}(s)$ is normalized to fix the contour length, $|\hat{\mathbf{t}}| = 1$. An external potential, $\beta U(\mathbf{R}(s))$ modifies the energy resulting in a path integral partition function,

$$ Z = \int D\mathbf{R}(s) \exp\left(-\int_0^L \left\{\frac{A}{2}|\partial_s \hat{\mathbf{t}}|^2 + \beta U(\mathbf{R}(s))\right\} ds\right). \tag{5.1.1} $$

We can derive the form of $\beta U(\mathbf{R}(s))$ for many experiments. For example, in the experiments of Smith, Finzi and Bustamante, one end of the DNA strand is attached to a glass surface and magnetic tweezers stretch the other end of the DNA [193]. The external potential is $U(\mathbf{R}(s)) = -f|R_z(0) - R_z(L)| = -f|\int_0^L \hat{t}_z ds|$, where f is the force applied to the two ends and $\hat{t}_z$ is the component of the tangent vector in the direction of the force. The solution maps into the trajectory of a quantum rigid-rotor and has good agreement with experiment [81, 142, 26]. We parameterize the bending energy of the WLC model with the persistence length determined by these references, $A = 53nm$, to remove fitting parameters and validate the consistency of the WLC description of DNA in various experiments. We neglect the persistence

length's dependence on various experimental conditions, such as buffer concentrations and the dye used in imaging.

The potential is not easily defined for hydrodynamic flows because hydrodynamic flows are dynamic phenomena, but the flow still performs work on each monomer. If we ignore the intrachain hydrodynamic interactions of a polymer in a potential flow, like constant plug flow or elongational flow, the work is proportional to differences in the potential. For a free-draining polymer, we add the potential of all of the monomer units, which becomes an integral over the contour. A simple double integral over the contour can incorporate the two body potential, but this is not done here. Removal of the free draining approximation is more difficult, but previous studies show that hydrodynamic interactions lead to only modest corrections to many time averaged quantities [131, 108].

5.2 Constant Plug Flow

The constant plug flow potential is $V = -F \cdot z$, where F is the flow rate, and the polymer's potential is $-F\xi \int_0^L z(s)ds$, where ξ is the friction per unit length. Based on the findings of Larson et $al.$, we assume the friction does not vary with the flow rate or over the polymer, but we do not know the friction constant explicitly [131]. Most experiments report the viscosity of the polymer in the solution, whose relation to friction has a non-trivial geometric dependence. The friction is the only fitting parameter, but it is comparable in all simulations and simply scales the flow rates. The Kirkwood approximation in Zimm's calculation for constant plug flow replaces

the bare uniform force on each segment from the flow field with a dressed force that is also uniform and justifies a rescaled friction constant [248].

Perkins *et al.* performed the constant flow experiment with optical tweezers [163]. Fluid passes over the polymer creating a force along the entire polymer. Integration by parts gives an insightful formula for the potential, $-F \cdot \xi \int_0^L (L - s) \hat{t}_z ds$. The energy of the potential comes from a tension that scales linearly along the chain. The tension is the greatest at the end tethered to the bead because the whole chain pulls on this end, and it lessens further down the chain until it becomes zero at the free end. For a strong flow, the polymer is almost completely extended in the z-direction, and the components in the x and y directions are small perturbations of the linear configuration. The energy of the WLC is approximately

$$\beta E \approx \int_0^L \left[A/2 |\partial_s \hat{\mathbf{t}}_\perp|^2 + \frac{F\xi}{2} (L - s) |\hat{\mathbf{t}}_\perp|^2 \right] - \frac{F\xi}{2} L^2, \tag{5.2.2}$$

where $\hat{\mathbf{t}}_\perp$ is the transverse components of the tangent vector, the x and y directions, and the component in the z direction is approximately $t_z = 1 - \frac{1}{2} |\hat{\mathbf{t}}_\perp|^2$ [142]. The action corresponds to a quantum harmonic oscillator with a linearly ramped spring constant. The x and y components act independently and the partition function for each component is Gaussian,

$$\int \mathcal{D}(\hat{t}_\perp^{(x,y)}(s)) \exp \left(-\frac{A}{2} \hat{t}_\perp^{(x,y)}(s) \frac{d \hat{t}_\perp^{(x,y)}(s)}{ds} \bigg|_0^L \right.$$
$$\left. + \int_0^L \int_0^L ds ds' \delta(s - s') \left\{ \frac{A}{2} \hat{t}_\perp^{(x,y)}(s') \frac{d^2 \hat{t}_\perp^{(x,y)}(s)}{ds^2} - \hat{t}_\perp^{(x,y)}(s') \frac{F\xi}{2} (L - s) \hat{t}_\perp^{(x,y)}(s) \right\} \right).$$

Since the operator in the exponential is Hermitian, the Green's function for the average square of the transverse component, $u(s, s') = \langle (\hat{t}_\perp^{(x,y)})(s)(\hat{t}_\perp^{(x,y)})(s') \rangle$, is a weighted sum of eigenfunctions, $G(s, s') = \sum_n |u_n(s)\rangle \lambda_n^{-1} \langle u_n(s')|$, that satisfy the differential equation

$$-A\frac{d^2 u_n(s)}{ds^2} + F\xi(L - s)u_n(s) = \lambda_n u_n(s) \tag{5.2.3}$$

with the boundary conditions $u_n'(0) = u_n'(L) = 0$. The analytic solutions to the equation are sums of Bessel functions. The rms of the traverse displacement is $\sqrt{\sum_n \lambda_n^{-1} | \int_0^L u_n(s)ds|^2}$, which is shown in Fig. 5-1. Since $\langle t_z \rangle \approx 1 - \frac{1}{2}(\langle (\hat{t}_\perp^{(x)}(s))^2 \rangle + \langle (\hat{t}_\perp^{(y)}(s))^2 \rangle) = 1 - \sum \lambda_n^{-1} u_n^2(s)$, the average extension in the z-direction is approximately $L - \sum \lambda_n^{-1} \int_0^L ds u_n^2(s) = L - \sum \lambda_n^{-1}$. Since the eigenvalues change slowly with the flow rate in the high stretching regime, the width and extension also change slowly. The rms displacement displays a trumpet shape that is qualitatively similar to the shapes observed in experiments, simulations, and other theories [163, 131, 50, 28].

Without the large flow rate approximation, the action of the tangent vector corresponds to the imaginary time Shrödinger equation for a rigid-rotor in a time dependent potential,

$$\frac{\partial \Psi(s)}{\partial s} = \left[\frac{\hat{\mathbf{L}}^2}{2A} + F\xi(L - s)\cos(\theta) \right] \Psi(s). \tag{5.2.4}$$

In this equation, $\hat{\mathbf{L}}$ is the angular momentum operator, and $\cos(\theta)$ is the projection of the tangent vector onto the direction of the flow field. This equation resembles the

197
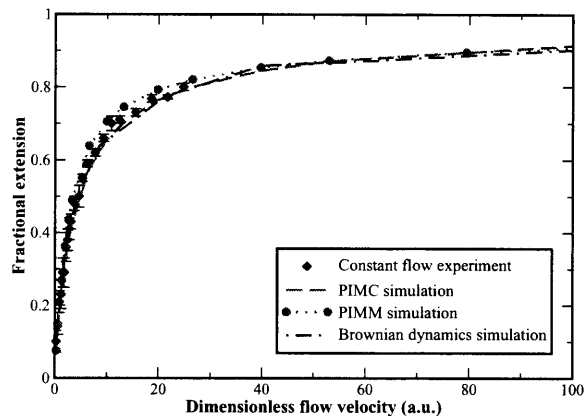
(a)

Figure 5-1: The root mean square displacement of the traverse component of the polymer in a strong constant plug flow. The displacement is plotted as a function of distance in the direction of the flow field. Note the resemblance to the trumpet shape observed in experiment and simulation.

constant force calculations with a simple time dependence $(L-s)$ [142]. This equation can be solved by using a spherical harmonic basis set and numerically propagating the resulting matrix [142, 26]. The only difficulty is the initial condition of the wave-function, for which we use results that are consistent with the high stretching calculation above, $t_z(0) = 1$ and $t_z(0) = t_y(0) = 0$.

Fig. 5-2 compares the predictions of mean fractional extension versus flow rate for the rotor Hamiltonian, the Monte Carlo simulation discussed below, the experimental results of Perkins et $al.$, and the Brownian dynamic simulations of Doyle [163, 69]. Since the experiment cannot determine the end-to-end distance, the fractional extension is the maximum distance between any points on the polymer contour compared against the contour length. The flow rate is in a dimensionless form, $Wi = \tau_F/\tau_p$, where τ_F is the characteristic time of the flow and τ_p is the longest relaxation rate of the free polymer, which is determined by equations in the references [69, 71, 70].
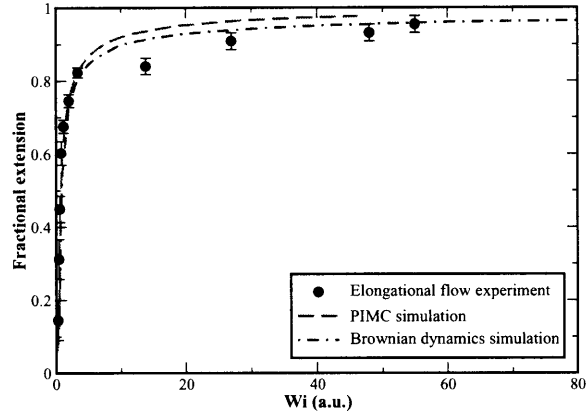
(a)

Figure 5-2: Comparison of the constant plug flow experiment in [163], the path-integral Monte Carlo simulation, the path-integral matrix multiplication method, and the Brownian dynamics simulation reported in [69]. Inset compares asymptotic behavior of the simulations for large flow rates.

For elongational and shear flows, the dimensionless form corresponds to Weissenberg numbers. Both simulations compare well with experiment. The mean extension initially increases rapidly with the flow rate. At about 60% of full extension, the rate of increase in the extension slows to an asymptotic approach to full extension in the large flow rate limit.

A slight discrepancy for moderate flow rates results from the initial condition, which is correct in the strong stretching limit and is not important for weak flows. Monte Carlo techniques correct the discrepancy, as would different initial conditions, which are not done here to avoid additional fitting parameters. Although we do not present these results since they appear elsewhere, the matrix multiplication method exactly reproduces the constant force results of Bouchiat, which agrees with the experiments of Bustamante, since no ambiguity about the initial conditions exists, and the rigid rotor equation is the same [26, 193, 33, 142].

199

5.3 Elongational Flow

The quantum rigid-rotor analogy extends to the elongational flow by changing the form of the external potential to $-F\xi \int_0^L ds(z^2(s) - 1/2(x^2(s) + y^2(s)))$, where F is the flow rate over length. The potential depends on the position instead of the tangent vector. Unless we know the starting position, which depends on the steady-state distribution, we do not know $\mathbf{R}(s)$ and cannot predict the s dependence of the field. To overcome the difficulties presented by this potential, we evaluate the path integral using Monte Carlo techniques.

As a calibration, we analyze the constant plug flow experiment with Monte Carlo and compare these results with our previous results. We discretized the polymer into 844 segments, two segments per persistence length for a DNA chain that is 22.4 μm in length. The discretization captures the rigidity of the polymer without incorporating phenomenological bending springs. The segments are fixed in length and only the angles are varied. We fix one end of the polymer at the origin and perform $12 \cdot 10^6$ Monte Carlo steps with the potential energy defined above. The Monte Carlo algorithm fits the experiment and Brownian dynamic simulation results better than the matrix multiplication method, as shown in Fig. 5-2. These results give us confidence in using this algorithm to evaluate more complicated flows.

Elongational flow corresponds to the experiments of Smith and Chu and of Perkins, Smith, and Chu [192, 162]. In these experiments, the DNA is freely flowing with the fluid. Since the forces caused by this flow are linear, we decompose the motion of the polymer into center of mass and relative motion of the polymer segments. The Monte

200
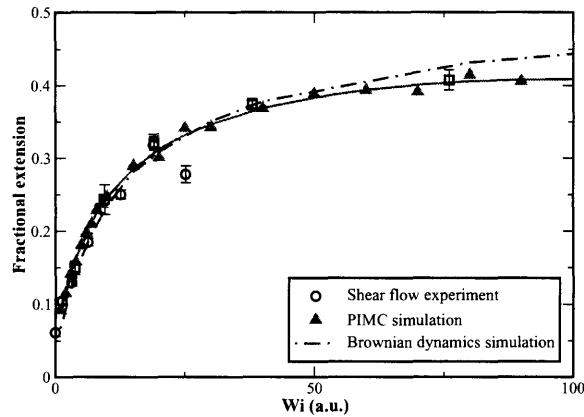
(a)

Figure 5-3: Comparison of the elongational flow experiment in [162], the path-integral Monte Carlo simulation, and the Brownian dynamics simulation reported in [69].

Carlo procedure is the same as for the constant plug flow, except that the energy is determined in the relative coordinate frame. Fig. 5-3 shows the fractional extension results compared against the experimental results of Chu and the Brownian dynamic simulations of Doyle [69].

Even for this more complicated flow, the fractional extensions predicted by both of the simulations agree extremely well with the experimental results. The fractional extension as a function of flow rate rises quickly to about 80% before a slow asymptotic approach to full extension. In the strong flow limit, the Monte Carlo simulation slightly overestimates the extension, as compared to the experiment and Brownian dynamics simulation, but all three results agree extremely well for weak and moderate flows. The agreement between the simulations and experiments for the constant plug and elongational flows demonstrates that the Monte Carlo technique successfully reproduces the results for potential flows and that the WLC model is a good description of DNA and possibly other semi-flexible biological polymers in potential flows.

5.4   Shear Flow

Encouraged by the success of the path integral Monte Carlo method on potential flows, we investigate the application of these methods to non-conservative flow fields like shear flow. The simple shear flow also has an analogous rigid-rotor Hamiltonian with an electrostatic potential $U \propto -x \cdot z$ and a non-conservative B-field, $\hat{\mathbf{r}} \times \hat{\mathbf{y}}$, where $\hat{\mathbf{r}}$ is the position vector and $\hat{\mathbf{y}}$ is the unit vector in the y direction. Similar to the elongational flow, we avoid the difficulty of the position dependence by evaluating the action with Monte Carlo.

The shear flow experiment that we analyze is similar to the elongational flow experiment [191, 69]. The DNA freely flows with the fluid, and we calculate the forces in the center of mass frame using the simple shear relations, $F_z = F\xi x$ and $F_x = F_y = 0$, where F is flow rate over length. Although shear flow is not a potential flow, the fluid still performs work on the system, which allows us to define a local energy change by integrating the force along the linear path connecting two configurations. The probability of a transition occurring is proportional to the energy difference in the local frame. Several authors used this approach to describe other polymer systems in shear flows [130, 202, 147, 238]. Since the potential changes as the polymer moves, detailed balance does not hold and the polymer rotates through space, but this simulation can be viewed as a Glauber dynamics [46]. The Monte Carlo algorithm for the shear flow follows the same steps as the constant plug flow and elongational flow with the potential defined locally. A trajectory dependence exists, which may require a larger number of simulation steps than the potential flows, but we still only

use $12 \cdot 10^6$ Monte Carlo steps.

The force extension relations for shear flow are plotted in Fig. 5-4. As with the potential flows, the path integral Monte Carlo method agrees extremely well with both experiment and the Brownian dynamics simulation. Because we only used $12 \cdot 10^6$ Monte Carlo steps, some scatter in the data exists. We added a trend line in Fig. 5-4 to help the eye follow the data. For weak shear rates there is a fast initial rise in the mean extension. After the initial rise, the data quickly asymptote to about a 40% extension. The small asymptotic value can be understood by examining the decomposition of the shear flow field into an elongational part and a rotational part. At an angle of about $\pi/4$ in the xz plane, the polymer gets stretched, but at $-\pi/4$ the polymer gets compressed [191, 69, 129]. The rotational part moves the polymer between these angles resulting in an averaging over these angles and a decreased total extension. This cycling from the extended to the compressed states has been observed in the experiment, the Brownian dynamics simulations, and our Monte Carlo simulations. The correspondence shows that the Glauber dynamics of a Monte Carlo simulation does capture some of the real dynamics of the system.

5.5   Summary and Conclusion

As demonstrated in this letter, the WLC is a good model for DNA and possibly other semi-flexible biopolymers. With a single fitting parameter, the friction constant, which linearly scales the flow rate, the solution to the path integral quantitatively agrees with experimental results for DNA subject to constant plug, elongational, and
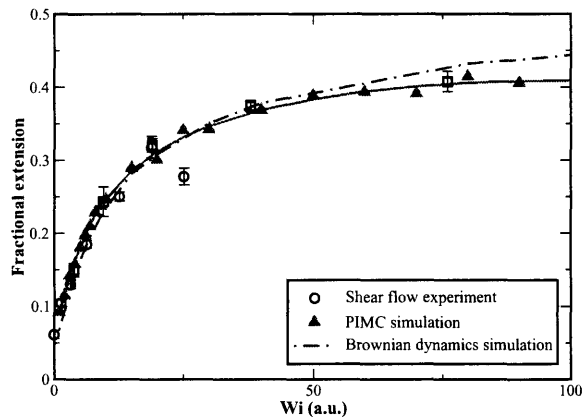
Figure 5-4: Comparison of the shear flow experiment in [191], the path-integral Monte Carlo simulation, and the Brownian dynamics simulation reported in [69]. To aid the eye, the solid curve follows the trend of the Monte Carlo data.

shear flows. The model is based on physical principles without phenomenological force extension relations. Although hydrodynamics are strictly dynamic phenomenon, time averaged quantities are quasi-equilibrium phenomenon in an effective potential. This description is possible because relaxation to the steady state distribution is fast and contributions from intra-chain hydrodynamics can often be neglected. The equilibrium partition function corresponds to an ensemble average, which demonstrates the correspondence between time-averages of single molecule trajectories and ensemble averages for ergodic systems. These techniques are computationally inexpensive since we do not have to run many trajectories to average over initial distributions. This path integral approach is applicable to other semi-flexible biopolymer systems.

204

Chapter 6

Brownian motion in dynamically disordered media

6.1 Introduction

Single particle tracking has become a popular method of probing local environments in complex systems, such as cells or glasses [220, 176]. These tracking experiments examine mesoscopic time and length scale motion of probe particles [61, 239, 220]. On these length scales, the random fluctuations of the solvent that influence the motion of the molecule cannot be modeled with the assumption that they are locally correlated in time and space since some large-scale motions of the solvent are on the time scale of the experiment. A result of these non-local correlations is deviation from Gaussian behavior on intermediate time scales which are measured in several experiments [239, 44, 61, 141].

Many stochastic problems can be modeled by the diffusion of Brownian particles

interacting with a random potential [24]. Although there is a macroscopic homogeneous environment, the mesoscopic heterogeneity caused by local environments often plays the central role of determining how the system behaves even on long time and length scales [55]. Several references have treated random media problems both analytically and numerically to find modifications to the diffusion constant to capture long-time behavior. An interesting issue that is not always examined is the deviations from the Gaussian behavior expected in the intermediate time scales of these random media problems, which can primarily be measured using single molecule experiments [73, 118, 217, 122, 165, 169, 24, 25, 55, 54, 51, 123, 124].

The heterogeneity on mesoscopic length and time scales has physical significance that is observable in experiments and simulations of many systems including systems near phase transitions like glasses and supercritical fluids.

Glasses

Several experiments in glasses demonstrate the effect of the heterogeneous fluctuating environments on molecules, including video microscopy, neutron scattering, NMR, and single-molecule tracking [220, 44, 100, 22, 15]. These problems are particularly interesting because of the many single molecule experiments performed on glassy systems. The systems exhibit regions of varying relaxation dynamics as well as collective behaviors, which lead to deviations from Gaussian behavior [209, 74, 44, 141, 63, 65, 66]. Kirkpatrick et al. attributes these varying dynamics to the constant formation and destruction of glassy clusters [115]. Experiments on colloidal systems by Weeks et al. and members of the Rice group show strong spatial-temporal correla-

tions in the system. Particles in these colloidal systems move collectively, resulting in long time correlations [113, 141, 209, 210]. Measurements of the rotational diffusion constants of colloidal spheres in glassy systems show similar stong spatial-temporal correlations [206, 111].

Supercritical fluids

Diffusion in supercritical fluids is another interesting example of dynamics in heterogeneous environments. The density of supecritical fluids has long length and time scale correlations, which lead to several anomalous experimental results [208, 27, 207, 45]. Although results differ, several experiments report dramatic changes in the diffusion constant of solutes in supercritical fluids [41, 180]. These density fluctuations have been observed in MD simulations and persist for long time scales [207, 91, 72].

These experiments and simulations show that there are intermediate length and time scale environmental fluctuations in several interesting systems, which motivates our study of a stochastic potential with intermediate scale correlations. Because two different stochastic processes, simple diffusion and the random potential, determine the movement of our Brownian particle, we expect significant deviations from the Gaussian form predicted by Einstein's equation. These deviations have been observed and quantified in several simulations and theories [141]. Several references showed that the simple assumption of fluctuating regions of two different diffusion constants, one for a cluster and another for free diffusion, results in significant changes

in the diffusion constant and deviations from Gaussian behavior on intermediate time scales [137, 249, 34]. Mode coupling theory also predicts significant deviations from the expected Gaussian behavior, especially for glasses [209, 210, 245, 89, 232]. Simulations by Donati et al. support these theoretical predictions of non-Gaussian behavior at intermediate times for glasses. Similar to the arguments of Kirkpatrick et al. they attribute these deviations to the formation of clusters [115, 66, 65]. Motivated by the observations of Donati and others we examine a simple model that exhibits similar clustering behavior.

6.1.1 The diffusion model

In order to address this Brownian motion problem analytically, we do not explicitly include the solvent. Instead, we develop a phenomenological model that captures the interactions of the solvent molecules with the Brownian particle without explicitly including them. In this paper, the microscopic time-scale fluctuations are still approximated as simple diffusion, but we extract the larger time scale motions and write them as a fluctuating potential with a time and spatial dependence. The approach follows the work of Deem and Chandler and has been discussed by Bouchaud and Georges but is generalized to allow a time dependent potential [55, 25]. This type of diffusion process is governed by the equation:

$$\frac{\partial G(\mathbf{x}, t)}{\partial t} = D_0 \nabla^2 G(\mathbf{x}, t) + \beta D_0 \nabla \cdot (G(\mathbf{x}, t) \nabla V(\mathbf{x}, t)), \qquad (6.1.1)$$

208

with

$$G(\mathbf{x}, 0) = \delta(\mathbf{x}),$$

where $G(\mathbf{x}, t)$ is the Green's function, D_0 is the diffusion constant modeling the short time scale interactions, β is the inverse temperature and $V(\mathbf{x}, t)$ is the random potential [55, 25, 54]. The difficulty in dealing with the random media problem is attempting to average $G(\mathbf{x}, t)$ over the disorder of $V(\mathbf{x}, t)$, even though we assume we know the moments of $V(\mathbf{x}, t)$ [25]. In this paper, we will make a Gaussian assumption for the random potential with

$$\langle V(\mathbf{x}, t) \rangle = 0$$

and

$$\langle V(\mathbf{x_1}, t_1) V(\mathbf{x_2}, t_2) \rangle = \chi(|\mathbf{x_1} - \mathbf{x_2}|, |t_1 - t_2|) \neq 0. \tag{6.1.2}$$

All other moments are either zero or can be expressed as a polynomial of χ. This assumption eases computation, but higher order cummulants can be incorporated using the methods below. Most of the equations in this paper are valid for arbitrary χ, but the explicit calculations used to generate the figures will correspond to 3 dimensions with χ defined by,

$$\chi(|\mathbf{x}|, |t|) = \frac{\alpha^{\frac{3}{2}} \chi_0}{(\alpha + \lambda|t|)^{\frac{3}{2}}} \exp\left[-\frac{|\mathbf{x}|^2}{4(\alpha + \lambda|t|)}\right]. \tag{6.1.3}$$

This potential-potential correlation is chosen because it has some of the features expected for a real potential of a fluid like system, but it is a model that also allows easier computation. Two of the important aspects of this paper are the various

techniques used to derive the Green's functions and the equivalence between them and the idea that the potential can mediate the clustering behavior of the system.

The constant $\chi_0 = \chi(|\mathbf{x}| = 0, t = 0)$, corresponds to the strength of the large scale interactions and is our perturbation expansion parameter. The choice of a Gaussian form allows us to omit a cutoff frequency that separates the small-scale motions that we approximate as simple diffusion and those that we treat as the stochastic potential. $\sqrt{\alpha}$ is the length scale of the heterogeneity of the system, which can be viewed as the size of wells in the potential energy surface. We refer to $\alpha^{\frac{d}{2}}$ as the size of the cluster to make contact with previous experiments and simulations [44, 63]. We incorporate decay of spatial correlation into our model by including the $\lambda|t|$ term. The $1/(\lambda|t|)$ dependence is chosen to give an exponential time dependence in Fourier space. This formulation corresponds to smaller scale fluctuations or clusters decaying faster than larger scale ones, a property prescribed by Trajus and Kivelson [203]. In analogy with several references, we refer to λ as the hopping rate, but we also emphasize that the power law dependence of our potential correlation function does not have a well-defined rate and, even for a strong disorder, the hopping fluctuations in the potential may not be associated with hopping of the Brownian particle since the particle has inertial mass [63, 88, 174].

6.1.2 Summary and outline

In the examination of the random media problem, a self-consistent equation for the one particle Green's function that is accurate to 2nd order in the disorder strength can

be determined by a perturbation expansion of equation 6.1.1 directly [165, 122, 118, 217]. But many papers calculated the Green's function and the diffusion constant by using the classical field formalism developed by Martin, Siggia, and Rose (MSR) [24, 55, 123, 51, 52]. To show the consistency of our perturbation method with the field theoretic methods of MSR, we solve for the Green's function using both methods and derive equations that are identical up to 2nd order in the disorder strength. This result should not be surprising since both methods attempt to describe the dynamics by a Dyson and Schwinger equation with self-energy. To make contact with previous work, we also solve the equations with an Edwards type of variational method that circumvents the Dyson and Schwinger equations [55, 24]. The variational method yields results that are similar to the our perturbation expansions as well as previous work but slightly more complex than perturbation. In the static limit the equations are identical to those derived by Deem and Chandler [55, 54, 164, 24, 51, 52].

The paper also contains a 1st order equation for the two-particle propagator determined by both the MSR and direct perturbation methods. Hydrodynamic and excluded volume interactions are not considered in this paper since the two particle propagator is intended to measure variation in different particle trajectories caused by the potential. As seen from equations 6.2.4 and 6.5.12 below these results are identicle to 1st order. As a final result, we use the MSR method to determine a 1st order renormalization group (RG) calculation of the diffusion constant, which compares well with our perturbation equations and reduces to previous RG calculations in the static limit [51, 55].

Our results are organized into six sections. Renormalized perturbation and nu-

merical results are presented in sections 6.2-6.4 followed by the conclusions in section 6.7. Readers interested in a more detailed discussion of MSR and RG are referred to sections 6.5 and 6.6.

6.2 Determination of the Green's function from direct perturbation

6.2.1 One-particle propagator

To begin our analysis, we perform a direct perturbation and re-summation of the Green's function equation to get a 2nd order self-consistent equation, which we compare with our field theoretic results. The renormalized perturbation expansion is analogous to the direct interaction approximation used to describe turbulent flows [54, 169, 73, 178, 119]. Fourier transforming the spatial variables ($\mathbf{x} \rightarrow \mathbf{k}$) gives the equation:

$$\frac{\partial}{\partial t}G(\mathbf{k}, t) = -D_0|\mathbf{k}|^2 G(\mathbf{k}, t)$$
$$- \frac{(\beta D_0)}{(2\pi)^d} \int d\mathbf{h}\{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h})(G(\mathbf{h}, t)V(\mathbf{k} - \mathbf{h}, t))\}, \qquad (6.2.1)$$

with

$$G(\mathbf{k}, 0) = 1.$$

The integration over $\mathbf{h}$ runs from $-\infty < |\mathbf{h}| < \infty$. We define $\partial_t + D_0|\mathbf{k}|^2$ as $G_0^{-1}(\mathbf{k}, t)$,

and $G_0(\mathbf{k}, t)$ as our usual fundamental solution for free diffusion,

$$G_0 = \exp\left(-D_0 t |\mathbf{k}|^2\right).$$

With this definition of $G_0(\mathbf{k}, t)$ the propagator for a specific realization of disorder is

$$G(\mathbf{k}, t) = G_0(\mathbf{k}, t)$$
$$- \frac{(\beta D_0)}{(2\pi)^d} \int d\mathbf{h} d\tau \{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h}) V(\mathbf{k} - \mathbf{h}, \tau) G_0(\mathbf{k}, t - \tau) G(\mathbf{h}, \tau)\}, \quad (6.2.2)$$

where τ has a range from 0 to t. Repetitively substituting the right-hand side of 6.2.2 for $G(\mathbf{k}, t)$ in the right-hand side of the equation 6.2.2 produces a perturbation expansion for $G(\mathbf{k}, t)$ in terms of $V(\mathbf{k}, t)$ and $G_0(\mathbf{k}, t)$. Since $G_0(\mathbf{k}, t)$ has no dependence on the random potential $V(\mathbf{k}, t)$, we are able to average over $V(\mathbf{k}, t)$ in this expression using Gaussian factorization. We re-sum terms so that the equation is accurate to various powers of the disorder strength, χ_0, and express these terms as self-consistent equations of $G(\mathbf{k}, t)$. The re-summation procedure corresponds to evaluating the self-energy in the Dyson expansion. The resulting 2nd order expression is

$$\langle G(\mathbf{k}, t) \rangle = G_0(\mathbf{k}, t)$$
$$- \frac{(\beta D_0)^2}{(2\pi)^d} \int d\mathbf{h} d\tau_1 d\tau_2 \{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h}) \mathbf{h} \cdot (\mathbf{k} - \mathbf{h}) \chi(|\mathbf{k} - \mathbf{h}|, |\tau_1 - \tau_2|)$$
$$\times G_0(\mathbf{k}, t - \tau_1) \langle G(\mathbf{h}, \tau_1 - \tau_2) \rangle \langle G(\mathbf{k}, \tau_2) \rangle \}$$
$$+ \frac{(\beta D_0)^4}{(2\pi)^{2d}} \int d\mathbf{h_1} d\mathbf{h_2} d\tau_1 d\tau_2 d\tau_3 d\tau_4 \{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h_1}) \mathbf{h_2} \cdot (\mathbf{k} - \mathbf{h_1})$$

$$\times \mathbf{h_1} \cdot (\mathbf{k} - \mathbf{h_1})(\mathbf{k} - \mathbf{h_1}) \cdot (\mathbf{k} - \mathbf{h_1} + \mathbf{h_2}) \chi(|\mathbf{k} - \mathbf{h_1}|, |\tau_1 - \tau_3|)$$

$$\times \chi(|\mathbf{h_1} - \mathbf{h_2}|, |\tau_2 - \tau_4|) G_0(\mathbf{k}, t - \tau_1) \langle G(\mathbf{h_1}, \tau_1 - \tau_2) \rangle \langle G(\mathbf{h_2}, \tau_2 - \tau_3) \rangle$$

$$\times \langle G(\mathbf{k} - \mathbf{h_1} + \mathbf{h_2}, \tau_3 - \tau_4) \rangle \langle G(\mathbf{k}, \tau_4) \rangle \}, \tag{6.2.3}$$

where the 2nd term is the first order correction and the 3rd term is the second order correction. The integrations in this expression are for $t > \tau_1 > \ldots > \tau_n > 0$ and $-\infty < |\mathbf{h_i}| < \infty$ and the $\langle \ldots \rangle$ represents the average over the disorder. The origin of the self-consistent equation can be easily demonstrated with a Feynman diagrammatic expansion and re-summation used in QED [80]. In these diagrams, we replace the propagator for photons with the propagator for χ and only use graphs that do not violate causality.

Although the sum of all graphs should converge to the solution, we are not guaranteed that the sum of any subsequence also converges. This lack of convergence may plague the 2nd order expression, since the graphs captured by this term are predominantly positive. The difficulty may also be purely numerical in nature due to approximations made in evaluating the 2nd order expression. In either case, a Páde approximation was used in the numerical calculations to aid in convergence of the solution. The specifics of these approximations are discussed in section 6.3.1.

6.2.2   Two particle propagator

Single particle motion is not sufficient to resolve the spatial-temporal correlations built into our model. The equations 6.2.3 show that the Green's function deviates from

the Gaussian form predicted by the free propagator, but the amount of the deviation is difficult to determine from examining the Green's function for the motions of a single particle. To overcome this difficulty we determine a perturbation expansion for a two-particle propagator. Section 6.5.2 shows the same calculation using the MSR perturbation method. This propagator measures the movement of two particles initially separated by a distance r. The correlation of these two particles determines the deviations from the behavior of two independent particles, which allows us to examine the spatial effects of the stochastic potential. As mentioned above, we are not considering any interactions between the particles, like excluded volume and hydrodynamic effects. Although numerical studies demonstrate the importance of both excluded volume and long-range hydrodynamic interactions between particles in many processes, including the glass transition, we are concerned with the roles of the heterogeneity of the solvent on the correlations between particles [205]. We are also not considering the effects of the particles on the solvent. The correlations we examine are strictly mediated through the potential and can be viewed as a demonstration of the deviations of different realizations of the path of the particle due to the potential if the particle was placed at different locations.

The perturbation calculation for the two-particle propagator starts from the expansion of the original Green's function equation. We do not average over the random potential and all terms are present. The propagator is centered at the origin and is denoted by $G^{(1)}(\mathbf{k_1}, t_1) = G(\mathbf{k_1}, t_1)$, where the superscript (1) denotes the particle label. We introduce a propagator for the 2nd particle that is displaced from the origin by a vector $\mathbf{r}$ and it is denoted $G^{(2)}(\mathbf{k_2}, t_2) = \exp[i\mathbf{k_2} \cdot \mathbf{r}] \cdot G(\mathbf{k_2}, t_2)$. Like the propa-

gator centered at the origin, we iteratively expand the equation and do not average the single-particle propagator over the random potential. We multiply the two propagators together and then average over the potential using Gaussian factorization. After resummation, the resulting first order self-consistent equation has a form that is similar to the equation for the single particle propagator.

$$
\begin{aligned}
G^{(1,2)}(\mathbf{k_1}, t_1 = t, \mathbf{k_2}, t_2 = t) &= \left\langle G^{(1)}(\mathbf{k_1}, t) G^{(2)}(\mathbf{k_2}, t) \right\rangle = \\
&\left\langle G^{(1)}(\mathbf{k_1}, t) \right\rangle \left\langle G^{(2)}(\mathbf{k_2}, t) \right\rangle - \frac{(\beta D_0)^2}{(2\pi)^d} \int d\mathbf{h} \int_0^t d\tau_1 \int_0^t d\tau_2 \\
&\times \left\{ \mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{h}) \mathbf{k_2} \cdot (\mathbf{k_1} - \mathbf{h}) \chi(|\mathbf{k_1} - \mathbf{h}|, |\tau_1 - \tau_2|) \right. \\
&\times \left. \left\langle G(\mathbf{k_1}, t - \tau_1) \right\rangle \left\langle G(\mathbf{k_2}, t - \tau_2) \right\rangle G^{(1,2)}(\mathbf{h}, \tau_1, \mathbf{k_1} + \mathbf{k_2} - \mathbf{h}, \tau_2) \right\}
\end{aligned} \tag{6.2.4}
$$

Two of the propagators are not labeled because they are single particle propagators with no multiplication by $\exp[i\mathbf{k} \cdot \mathbf{r}]$ and $\langle G(\mathbf{k}, t) \rangle$ corresponds to the single particle Green's function in Eq. 6.2.3. This equation can also be represented by a Feynman diagrammatic expansion and resummation [80].

6.3 Derivation of important statistical quantities from the Green's functions

The preceding section outlines the derivation of the self-consistent equations for various Green's functions, which correspond to different possible single molecule experiments. These experiments measure certain numerical quantities associated with

216

motions of the Brownian particles but they do not measure the Green's functions directly. In this section we outline methods of extracting some measurable quantities from the self-consistent equations derived above, including the diffusion constant and indicators of the deviation of the two-particle Green's function from the uncorrelated motion of two independent particles. We will refer to these quantities as non-Gaussian indicators, but the system is not necessarily Gaussian if these quantities are zero. Several non-Gaussian indicators are possible and we discuss a few that are easily determined from experiments [141, 66, 65, 74].

6.3.1 Diffusion constant

A simple Ansatz

As discussed in the introduction, macroscopic effects of mesoscopic disorder led to many studies that determine the diffusion constant from self-consistent equations similar to the ones derived above [55, 54, 164, 51, 52]. Most of these references address static disorder, but these methods can be generalized for the dynamic case. Our analysis parallels several of these approaches. The methods require numerical computation of the solution from a suitable basis set [55]. Before introducing these rigorous methods, we perform a simple analytic calculation based on a simple Ansatz that the single particle Green's function maintains a Gaussian form, but the diffusion constant is modified.

$$\langle G(\mathbf{k}, t) \rangle \approx \exp[-Dt|\mathbf{k}|^2] \tag{6.3.1}$$

For weak disorder, the Gaussian form is exact and the approximation should determine the correct initial slope of the diffusion constant. The diffusion constant can be expressed in terms of the Laplacian of the Green's function.

$$\lim_{t \to \infty} \left. \frac{-\nabla_{\mathbf{k}}^2 G(\mathbf{k}, t)}{2dt} \right|_{\mathbf{k}=0} = D \tag{6.3.2}$$

The Gaussian form allows analytic evaluation of the 1st order expression with the Gaussian form of the potential correlation functions, χ in Eq. 6.1.3, which results in an algebraic expression for the diffusion constant.

$$D = D_0 - \frac{D_0^2(D - \lambda)}{(D + \lambda)^2} \frac{\beta^2 \chi_0}{3} \tag{6.3.3}$$

The equation has no explicit dependence on the cluster size because it is the intrinsic scale of the problem which determines the magnitude of the other quantities. The equation exhibits the expected behavior from intuition. As the disorder strength, χ_0, increases, the diffusion constant will decrease, and as the hopping rate, λ, increases, the diffusion constant will increase, but this increase is dependent on the disorder strength. If $\lambda > D_0$ the equation suggests an increase in the diffusion constant. This result is not surprising in the weak disorder limit since the forces on the particle increase displacements, but there is no trapping because the potential rearranges quickly and is weak. The solution of this equation as a function of $\frac{\beta^2 \chi_0}{3}$ and λ is plotted in Fig. 6-1(a). The renormalization group (RG) result is also plotted as a function of $\frac{\beta^2 \chi_0}{3}$ and λ in Fig. 6-1(b). A derivation of the RG result and a more detailed comparison of these results are discussed in sections 6.6 and 6.4 respectivley.
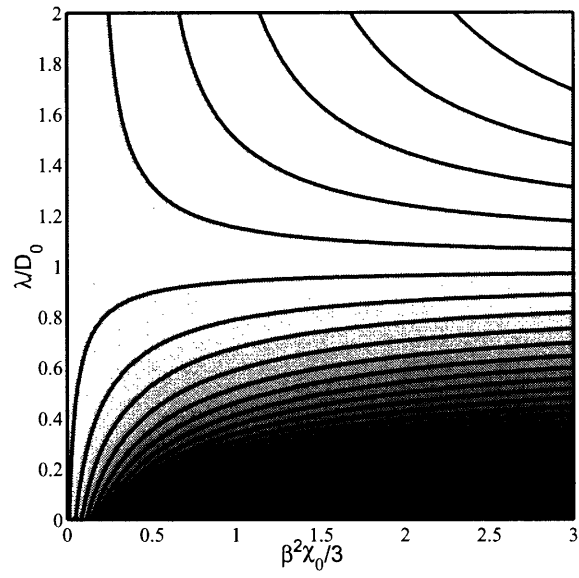
(a)
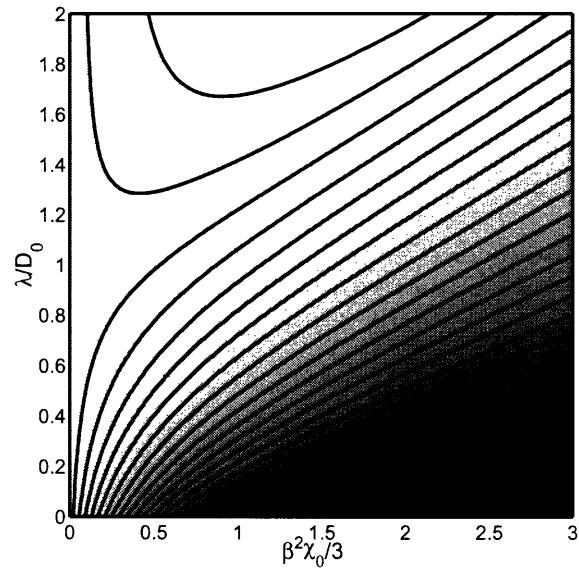
(b)

Figure 6-1: Contour plot of the diffusion constant as a function of λ and $\beta^2\chi_0/3$. (a) corresponds to the perturbation expansion of Eq. 6.3.3 and (b) is the renormalization group result.

Numerical determination of the diffusion constant

In order to go beyond the simple assumption of a Gaussian form, we need to implement a numerical approximation method for determining the Green's function and subsequently the diffusion constant. The numerical solutions are calculated in Laplace space using a basis set approach, which is similar to the approach of Deem and Chandler [55]. Approximate 1st and 2nd order solutions using a 16 function basis set were determined as functions of $\frac{\beta^2 \chi_0}{3}$ for $\lambda/D_0 = 0.0, 0.4, 1.2$. These results are presented in Fig. 6-2(a),(b), and (c) for $\lambda/D_0 = 0.0, 0.4, 1.2$, respectively. These figures also compare the RG calculation and the simple self-consistent result, Eq. 6.3.3.

The basis set equations are complicated non-linear integral equations that are difficult to solve analytically or numerically. Several numerical approximations used in evaluating these expressions introduce errors other than the choice of the basis set. To aid in convergence with these numerical approximations, a Páde approximation is introduced. This approximation has the correct 1st and 2nd order terms, but also includes additional higher order terms that may aid in convergence. This approximation combines our 1st and 2nd order correction terms in eq 6.2.3.

$$\frac{1^{st}\ order\ expression}{1 - \frac{2^{nd}\ order\ expression}{1^{st}\ order\ expression}} \tag{6.3.4}$$

The errors introduced in determining the numerical results are higher than second order in terms of disorder strength so the equation is still accurate at 2nd order. These approximations are not necessary for the $\lambda \to 0$ limit, and a comparison of the numerical solutions with and without the Páde approximation in this limit shows

220
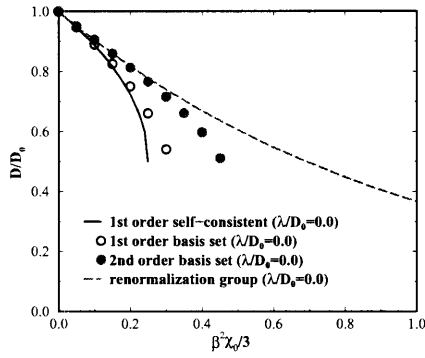
(a)
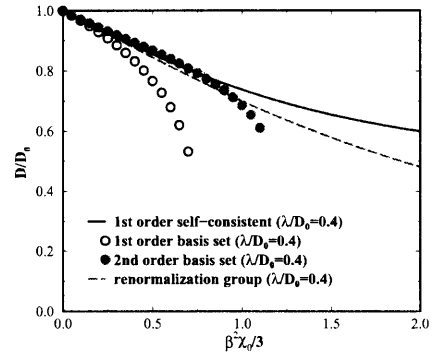
(b)
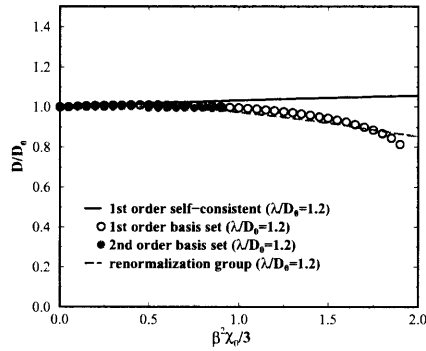
(c)

Figure 6-2: Comparison of the 1st order (empty circles) and 2nd order (filled circles) numerical solutions derived from the basis set approach outlined in section 6.3.1 and the results of Eq. 6.3.3 (solid line), and the renormalization group (RG) (dashed line) for (a) $\lambda/D_0 = 0.0$, (b) $\lambda/D_0 = 0.4$, and (c) $\lambda/D_0 = 1.2$.

good agreement at moderate disorder strengths. The Páde approximation prevents higher order terms contained in the 1st and 2nd renormalized expressions from strongly influencing the fit, which makes our calculation less sensitive to other approximations necessary to numerically calculate the diffusion constant. This is especially important since the terms captured by the 2nd order expression are predominately positive, which may prevent convergence since we are only summing some graphs. The extra higher order terms in the Gaussian reference calculation in equation 6.5.14 may also alleviate this sensitivity, as would including higher order terms in the perturbation expansion. Numerically, both the 2nd order Gaussian reference calculation and including higher order perturbation expansion terms would be difficult and is not performed here. A detailed discussion of the results for the calculations performed in this paper is presented in section 6.4.

6.3.2 Non-Gaussian indicators

In order to analyze the non-Gaussian nature of the model, we need to define quantities that measure the deviations from uncorrelated behavior. Experimentally determining deviations from Gaussian behavior is difficult with a single particle, so we analyze the model for two particles that is developed in section 6.2.2. The indicators should demonstrate collective motions like clustering because of structures in the system that prevent two particles from moving independently. Both of these effects are observed in experiments and simulations of real systems like glasses [141, 61, 63, 65, 174, 209].

Two correlations that are measured in several single molecule experiments are

222

the dot product of the displacement vectors of two particles initially separated by a distance r, $\delta\mathbf{R_1} \cdot \delta\mathbf{R_2}$, and the square of the displacements, $\delta\mathbf{R_1^2} \cdot \delta\mathbf{R_2^2}$ [43, 195]. The subscripts label the particle and the δ denotes the displacement from the initial position. The equation for the two-particle propagator is difficult to manipulate because it requires the solution for the one particle Green's function. To simplify the analysis, we perform an asymptotic expansion of the equations and only evaluate the non-renormalized terms. To normalize the function, we divide this quantity by $\beta^2 \chi_0$. The two correlation functions are

$$C_1(r,t) = \frac{\langle \delta\mathbf{R_1}(t) \cdot \delta\mathbf{R_2}(t)\rangle}{(\beta^2 \chi_0)\langle|\delta\mathbf{R}|^2\rangle} \approx \frac{\langle \delta\mathbf{R_1}(t) \cdot \delta\mathbf{R_2}(t)\rangle}{(\beta^2\chi_0)(2dD_0 t)} \tag{6.3.5}$$

$$C_2(r,t) = \frac{\langle|\delta\mathbf{R_1}(t)|^2|\delta\mathbf{R_2}(t)|^2\rangle - \langle|\delta\mathbf{R}|^2\rangle^2}{(\beta^2\chi_0)\langle|\delta\mathbf{R}|^2\rangle^2} \approx \frac{\langle|\delta\mathbf{R_1}(t)|^2|\delta\mathbf{R_2}(t)|^2\rangle - \langle|\delta\mathbf{R}|^2\rangle^2}{(\beta^2\chi_0)(2dD_0 t)^2}$$

$$\tag{6.3.6}$$

To first order these can be written in Fourier space as:

$$\frac{(\beta D_0)^2}{(\beta^2\chi_0)(2dD_0 t)(2\pi)^d} \nabla_{\mathbf{k_1}} \cdot \nabla_{\mathbf{k_2}} \left(\exp[-i\mathbf{k_2} \cdot \mathbf{r}] \int d\mathbf{h} \int_0^t d\tau_1 \int_0^t d\tau_2 \right.$$

$$\times \{\mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{h})\mathbf{k_2} \cdot (\mathbf{k_1} - \mathbf{h})\chi(|k_1 - h|, |\tau_1 - \tau_2|)G_0(\mathbf{k_1}, t - \tau_1)$$

$$\left. \times G_0(\mathbf{k_2}, t - \tau_2) \ G_0(\mathbf{h}, \tau_1) \exp[i(\mathbf{k_1} + \mathbf{k_2} - \mathbf{h}) \cdot \mathbf{r}] \ G_0(\mathbf{k_1} + \mathbf{k_2} - \mathbf{h}, \tau_2)\} \right)$$

$$\tag{6.3.7}$$

$$\frac{(\beta D_0)^2}{(\beta^2 \chi_0)(2dD_0 t)(2\pi)^d} \nabla^2_{\mathbf{k_1}} \nabla^2_{\mathbf{k_2}} \left(\exp[-i\mathbf{k'} \cdot \mathbf{r}] \int d\mathbf{h} \int_0^t d\tau_1 \int_0^t d\tau_2 \right.$$

$$\times \{ \mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{h}) \mathbf{k_2} \cdot (\mathbf{k_1} - \mathbf{h}) \chi(|\mathbf{k_1} - \mathbf{h}|, |\tau_1 - \tau_2|) G_0(\mathbf{k_1}, t - \tau_1)$$

$$\left. \times G_0(\mathbf{k_2}, t - \tau - 2) \; G_0(\mathbf{h}, \tau_1) \exp[i(\mathbf{k_1} + \mathbf{k_2} - \mathbf{h}) \cdot \mathbf{r}] G_0(\mathbf{k_1} + \mathbf{k_2} - \mathbf{h}, \tau_2) \} \right)$$

$$(6.3.8)$$

The 1st correlation C_1 contains both angular and radial information, while C_2 contains only radial information. These quantities are useful in micro-rheological experiments [43, 195]. Because the integral is Gaussian for our model potential, these expressions can be integrated analytically, resulting in complex expressions that are omitted. The two correlation functions behave similarly so we only present the features of C_2 in Fig. 6-3 (a) and (b). These figures are analyzed in section 6.4.

6.4 Analysis of results

In the previous section, sec 6.3, the diffusion constant and non-Gaussian indicators are defined and determined by renormalized perturbation for the random potential presented in eq 6.1.3. In this section we present a detailed discussion of the properties of the solutions to compare these results to previous models and demonstrate some of the unusual properties of this model.

6.4.1 Diffusion constant

Section 6.3 presents several approximations for determining the diffusion constant, and section 6.6 presents a renormalization group (RG) approximation of the diffusion
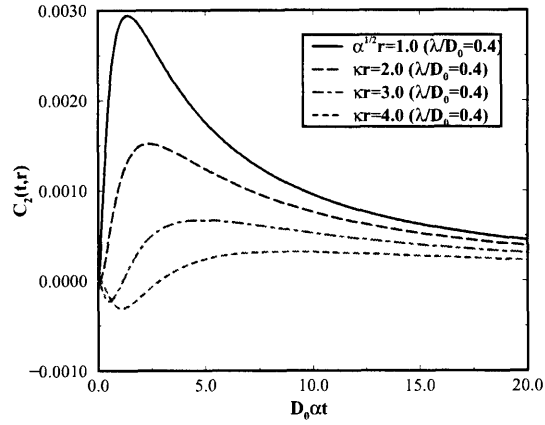
(a)
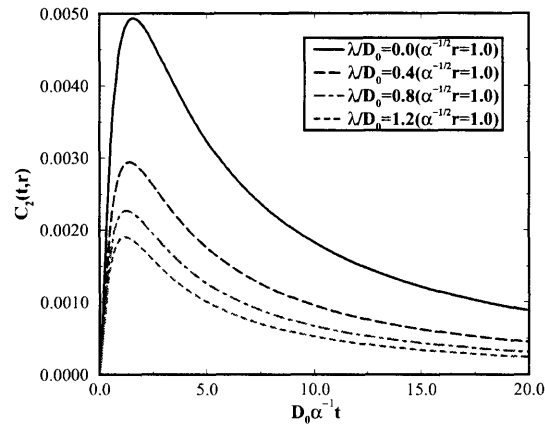
(b)

Figure 6-3: Time dependence of the non-Gaussian indicator C_2. (a) shows the behavior of C_2 as a function of time for several values of initial separation with $\lambda/D_0 = 0.4$. (b) shows the behavior of C_2 as a function of time for several values of λ with $r = \sqrt{\alpha}$.

constant. Due to the technical nature of the RG calculation, we compare the results of the perturbation and RG calculations in this section before discussing the details of the RG calculation in section 6.6. The analysis begins by defining $\sqrt{\alpha}$ as our length scale and $\frac{\alpha}{D_0}$ as our time scale. With the redefinition of our length scales, χ in Eq. 6.1.3 is uniquely determined by two dimensionless quantities $\frac{\beta^2 \chi_0}{3}$ and λ/D_0. All figures and discussions of the effects of the hopping rate, λ, and the disorder strength, χ_0, are in terms of these dimensionless quantities and the diffusion constant is also made dimensionless by dividing by D_0.

Figure 6-1 presents a contour plot of the solution that is generated from Eq. 6.3.3 and the RG calculation in section 6.6. Both plots appear similar for weak disorder strength because they capture the correct first order response of the diffusion constant. In the weak disorder regime, the diffusion constant is dominated by the disorder strength because the hopping rate's effect on the diffusion constant is proportional to the disorder strength. For moderate hopping rates, λ, the two figures agree up to stronger disorder strengths because the two equations yield similar results as long as D stays close to D_0.

For strong disorder strengths, the two equations show markedly different characteristics. For a small hopping rate, the perturbation solution, Eq. 6.3.3, predicts that the particle becomes trapped for moderate disorder. Because $\chi(0,0)$ is Gaussian, it can never be large and trapping should not occur [59]. The trapping predicted by perturbation expansions is common in disordered media problems because the effect of the disorder is over-emphasized [54]. For larger hopping rates, the perturbation result predicts that the diffusion constant approaches the hopping rate, which sug-

gests that the particle is trapped in a well that is moving at the hopping rate. As mentioned previously, a well-defined time constant for the hopping is not defined so this result is as unphysical as the trapping predicted for a small hopping rate.

Unlike the perturbation result, the RG calculation never predicts trapping. The solution remains above the solution determined by Dean and by Deem, which is above the lower bound determined by De Masi et al. [51, 55, 59]. In fact, in the $\lambda \to 0$ limit, the results of Dean and of Deem are recovered [51, 55]. For any value of the hopping rate, the diffusion constant approaches zero in the strong disorder limit so the contour lines of Fig. 6-1(b) never become parallel with the $\beta^2\chi_0/3$ axis. For $\lambda >> 0$, the diffusion constant approaches zero as a power-law, see Fig. 6-4. Except for small λ, the exponent is weakly dependent on λ and close to unity.

The numerical solutions are more computationally intensive than the simple self-consistent equation, Eq. 6.3.3, and the RG result so these equations are used to determine the diffusion constant for specific values of the hopping rate, $\lambda/D_0 = 0.0, 0.4, 1.2$. These results are compared against the results of Eq. 6.3.3 and the RG calculation in Fig. 6-2. All of the different techniques agree in the small disorder limit, but the numerical calculations predict trapping of the particle for finite disorder strength for all values of λ. The 2nd order numerical solution agrees with the RG solution for larger disorder strengths than the 1st order numerical solution. The better agreement between the RG and 2nd order solutions for diffusion problems in the static limit has been demonstrated in several papers, and numerical simulations in the static limit suggest that the RG calculation may be correct up to larger disorder strengths than would be expected from a first-order RG result [37, 164]. Because of these previous
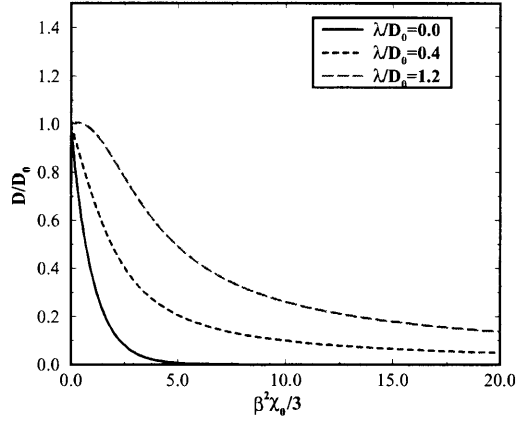
(a)
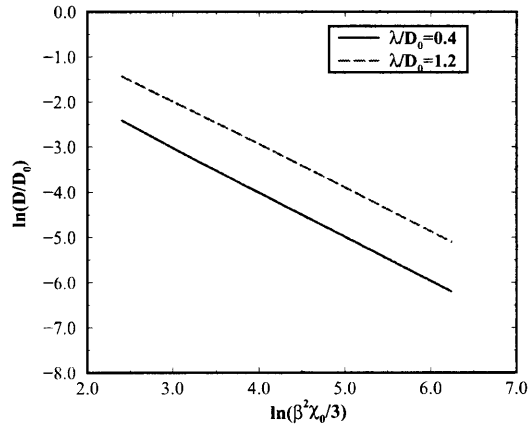
(b)

Figure 6-4: (a) shows the behavior of the diffusion constant predicted by the renormalization group calculation over a large range of disorder strengths for $\lambda/D_0 = 0.0, 0.4$, and 1.2. (b) is a log-log plot of the diffusion constant versus the disorder strength for large disorder for $\lambda/D_0 = 0.4$ and 1.2. The static limit is not plotted because it is an exponential. The straight line with nearly unity slope shows the power-law dependence of disorder for moderate values of λ

studies, the static disorder results are not surprising, and our RG calculation is a reasonable extension of these previous calculations.

The $\lambda/D_0 = 1.2$ case demonstrates interesting behavior because non-renormalized perturbation expansion predicts that the diffusion constant initially increases as discussed above. Because of the power-law time dependence of the potential correlations, one expects stronger disorder strength to eventually restrict the motion of the particle causing a decrease in the rate of diffusion. These intuitive arguments agree with the numerical calculations and the RG result and show that these results capture the physical aspects of the problem. Diffusion initially increases with increasing disorder strength but it eventually decreases in the strong disorder limit.

6.4.2 Non-Gaussian indicators

Determining the diffusion constant for disordered media is important in many industrial processes like chromatography, but single molecule experiments do not study macroscopic diffusion. The experiments study deviations from simple diffusion on intermediate time scales. Although more detailed calculations can be performed, the simple first order non-renormalized calculation demonstrates several of the interesting features of this model, including an apparent cluster-size and the effect of the dynamics of the random potential, λ (which can be shown from the C_1 calculation in Eq. 6.4.1 below). We defined two non-Gaussian indicators in section 6.3.2 in order to capture the angular and radial dependence of correlations involving two Brownian particles separated by an initial distance r.

The C_1 correlation function contains information that can be interpreted as clustering. Clustering of the two particles would correspond to the particles having a tendency to move together resulting in a decrease in the distance between them as compared to the motions of independent particles. As a result, $\langle |\mathbf{R_1} - \mathbf{R_2})|^2 \rangle$ would be less than the value predicted for independent particles. Because $\langle \mathbf{r} \cdot \delta\mathbf{R_1} \rangle = \langle \mathbf{r} \cdot \delta\mathbf{R_2} \rangle = 0$,

$$\left\langle |\mathbf{R_1} - \mathbf{R_2}|^2 \right\rangle = \left\langle |\delta\mathbf{R_1} - \mathbf{r} - \delta\mathbf{R_2}|^2 \right\rangle$$

$$= \left\langle |\delta\mathbf{R_1}|^2 \right\rangle + \left\langle |\delta\mathbf{R_2}|^2 \right\rangle + |\mathbf{r}|^2 - 2 \cdot \langle \delta\mathbf{R_1} \cdot \delta\mathbf{R_2} \rangle. \tag{6.4.1}$$

From this expression we see that a positive C_1 means that particles have a tendency to move closer together than predicted by independent motion and a negative C_1 corresponds to the opposite behavior.

The C_2 correlation function is a measure of correlations of the rates of diffusion for the two particles as compared to independent particles. We concentrate our analysis on C_2 because the two functions share similar characteristics with only minor quantitative differences which will be discussed below. Since there are only small differences, the graphs in Fig. 6-3(a) and (b) only correspond to the C_2 correlation function. We are examining the non-renormalized perturbation calculation and all quantities are proportional to the disorder strength. Similar to the diffusion constant, the natural length scale is $\sqrt{\alpha}$ and the natural time scale is α/D_0. Because of these relationships, α, D_0 and $\beta^2\chi_0/3$ are set to unity. The important parameters that have a qualitative effect on the non-Gaussian indicators are the initial separation of the two particles, r, and the relaxation dynamics of the potential, λ.

230

The effect of the initial separation is stronger than the dynamics of the potential since the time dependent factor is also dependent on the spatial factor, $\exp[-|\mathbf{r}|^2/(\alpha + \lambda|t|)]$. For short times, the temporal decay of the spatial correlations is not large and the correlations are completely dominated by the spatial separation. The strong dependence on r is apparent by examining Fig. 6-3 and Fig. 6-5. For small initial separations, the system shows strong positive correlations for both C_1 and C_2, resulting in a high positive peak. For larger values of r the strength of the correlations decrease and both C_1 and C_2 show anti-correlations in the short time behavior for some values of r. The C_1 correlation shows only a simple inversion for $r > \sqrt{6\alpha}$. As discussed above, negative correlations in C_1 correspond to a larger increase in the distance between the two particles than would be expected for independent particles. The C_2 correlation shows a more complex inverse response. For $1.9\sqrt{\alpha} > r > 4\sqrt{\alpha}$ the correlation function shows an initial inverse behavior, but some values of $r > 4\sqrt{\alpha}$ the correlation is initially positive for a short period of time before becoming negative. In the long time limit, all correlations become positive and decay as power-laws, $C_1 \propto t^{-1}$ and $C_2 \propto t^{-\frac{3}{2}}$. The power-law behavior comes from the normalization of the correlation functions, but it demonstrates that the correlations between the two particles remains significant even at large times.

Unlike the strong qualitative effect of the initial separation, the time dependence of χ only affects the quantitative features of the correlation functions. The qualitative shapes of the correlations do not change, but the heights and positions of the maxima change as λ changes. The time and height of the maxima of C_2 as a function of r for several values of λ are presented in Fig. 6-5, which demonstrates the qualitative
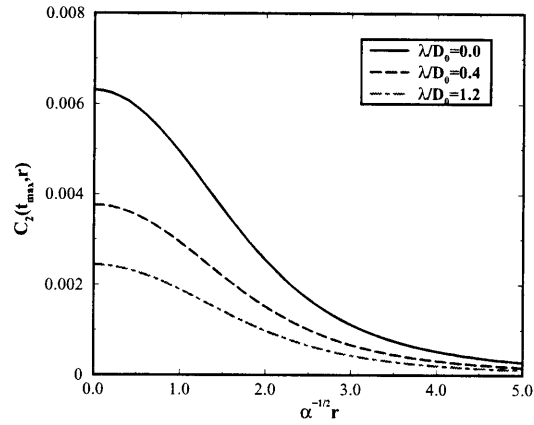
(a)
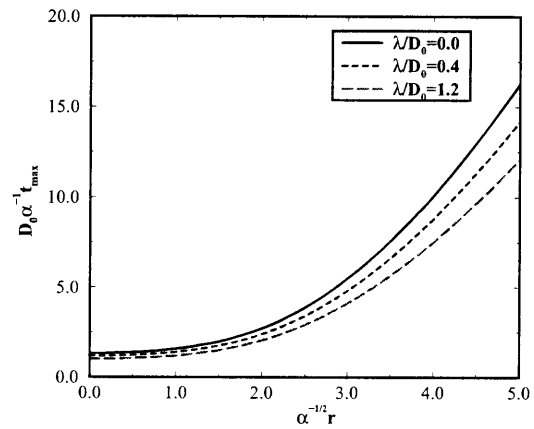
(b)

Figure 6-5: (a) shows the maximum value of C_2 as a function of initial separation, r, for several values of λ. (b)shows the time of the maximum. From the figures it is apparent that λ does not qualitatively affect the shape of these graphs.

effects of r and the quantitative effects of λ.

The strong positive correlations between particles separated by small distances suggest that the model does exhibit clustering behavior. The particles that are within $\sqrt{6\alpha}$ diffuse with similar speeds and have a tendency to move closer together than expected if they are uncorrelated. This behavior is similar to what is expected for particles in the "same cluster." If the particles are farther apart than the length of a cluster, they are in two "different clusters," which have a tendency to diffuse away from each other, leading to negative correlation in C_1.

6.5 MSR field theory

The perturbation expansion equations can be derived using the MSR formalism [143, 166, 109]. But most applications of MSR to diffusion problems have been centered around RG calculations and reference systems [25, 55, 51]. These efforts are stimulated by the quenched disorder problems where the Green's function instead of the generating function should be averaged over quenched disorder [25]. As stated in the introduction, section 6.1, the non-trivial terms for this particular problem are the same regardless of averaging over the generating function or the Green's function so we examine the perturbation expansion of this problem as well as a reference and RG calculation. This equivalence between the different averaging techniques for certain dynamic problems has been discussed elsewhere, and should not be surprising since the system will always be diffusive [49, 48, 24, 59, 36, 55]. For a more detailed discussion of MSR we refer readers to several references [143, 49, 48, 56, 57, 58, 166, 167,

168, 107, 109, 3, 24, 55]

6.5.1 The generating functional

To write the random media problem into the field theoretic formalism we note that the diffusion equation is also the equation for the change in the probability density at a point in space. With this idea we can write a replica generating function in a form that is similar to a Feynman path integral. Defining

$$G_0^{-1}(\hat{\mathbf{1}}, \mathbf{2}) = \delta(\mathbf{1} - \mathbf{2})(\partial_{t_2} + D_0|\mathbf{k_2}|^2) = \delta(\mathbf{1} - \mathbf{2})(-\partial_{t_1} + D_0|\mathbf{k_1}|^2), \tag{6.5.1}$$

and

$$\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4}) =$$

$$\frac{(\beta D_0)^2}{(2\pi)^d}\delta(t_1 - t_2)\delta(t_3 - t_4)\delta(\mathbf{k_1} - \mathbf{k_2} + \mathbf{k_3} - \mathbf{k_4})$$

$$\times \mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{k_2})\mathbf{k_3} \cdot (\mathbf{k_3} - \mathbf{k_4})\chi(|\mathbf{k_1} - \mathbf{k_2}|, |t_1 - t_3|)$$

the generating function is

$$Z^N[\xi_i, \hat{\xi}_i] = \int \mathcal{D}[\rho_i]\mathcal{D}[\hat{\rho}_i]$$

$$\exp\left[-G_0^{-1}(\hat{\mathbf{1}}, \mathbf{2})\hat{\rho}_i(\mathbf{1})\rho_i(\mathbf{2}) + \frac{1}{2}\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})\hat{\rho}_i(\mathbf{1})\rho_i(\mathbf{2})\hat{\rho}_j(\mathbf{3})\rho_j(\mathbf{4})\right.$$

$$\left. + \xi_i(\mathbf{1})\rho_i(\mathbf{1}) + \hat{\xi}_i(\mathbf{1})\hat{\rho}_i(\mathbf{1})\right] \tag{6.5.2}$$

234

where we integrate over repeated arguments and sum over indices i and j [109, 55]. The notation we use in equation 6.5.2 is consistent with Jensen's work. The variables $\mathbf{1} = [\mathbf{k_1}, t_1]$ and $\mathbf{2} = [\mathbf{k_2}, t_2]$ stand for all variable parameters, like space and time and the components of these variables are denoted with a subscript, 1 or 2. This generating function introduces a conjugate variable $\hat{\rho}(\mathbf{1})$, which acts as an infinitesimal density creation operator. Differentiation of $\ln Z$ with respect to $\xi_i(\mathbf{1})$ and $\hat{\xi}_i(\mathbf{1})$ determines the desired statistical quantities. For the perturbation expansion we take $N = 1$ since we are averaging over the generating function. In the reference calculation, where the Green's function is averaged, the $N \to 0$ limit is taken in the replica action, which eliminates several terms that are present in the $N = 1$ limit. It may appear that the two different limits cannot both be correct, but the additional terms in the perturbation expansion that are eliminated in the $N \to 0$ limit are non-causal and evaluate to zero as discussed above [49, 48]. This result gives us confidence that our MSR perturbation, where we do not take the $N \to 0$ limit, does not fail to properly average over the disorder [25]. The η is the Jacobian which depends on discretization and can be assumed to be constant.

The action in equations 6.5.2 is very similar to the actions in other references on diffusion in random media [55, 54, 37, 164, 25]. Our derivation is for the general case of dynamic disorder and our equations simplify to these previous results in the static limit. Because we consider dynamics, we must integrate over time or frequency, which can often be omitted in the static case [55]. Some slight differences also come from our definition of $\hat{\rho}$, which is equal to $-i\hat{\rho}$ in several references [55, 160, 123, 124]. The largest contrast comes from our action being defined in Fourier space so the signs

of some of the arguments, like **2** are reversed. Because of the domain of definition, the Fourier transform of $\hat{\rho}$ corresponds to density created at a specific wave-vector instead of a specific point in space.

The generating function gives us all of the desired statistical quantities by differentiating $\ln Z[\xi_i, \hat{\xi}_i]$ with respect to the generating variables ξ_i and $\hat{\xi}_i$. The replica trick that we use in our Gaussian reference calculation replaces the $\ln Z[\xi, \hat{\xi}]$ with $\lim_{N \to 0} \frac{Z^N - 1}{N}$, but we are still evaluating the same quantities. The Green's function is defined by

$$\langle\langle \rho(\mathbf{1})\hat{\rho}(\mathbf{2})\rangle\rangle = G_2(\mathbf{1}, \hat{\mathbf{2}}) = \left. \frac{\delta^2 \ln Z[\xi_i, \hat{\xi}_j]}{\delta \xi_i(\mathbf{1}) \delta \hat{\xi}_j(\mathbf{2})} \right|_{\xi_i = \hat{\xi}_j = 0}, \tag{6.5.3}$$

This quantity is the response function and it represents the creation of density at the wave vector $\mathbf{k_2}$ at time t_2 and subsequently measuring the density at the wave vector $\mathbf{k_1}$ at time t_1, which is similar to the definition of the Green's function defined in Eq. 6.1.1, except that the density creation is defined in Fourier space.

6.5.2 MSR perturbation theory

One particle propagator

The perturbation method follows the derivation of Jensen [109]. We explicitly construct Schwinger and Dyson equations from $Z[\xi, \hat{\xi}]$ instead of evaluating $\ln Z[\xi, \hat{\xi}]$. $N = 1$ and we are not introducing the unphysical replica trick. We also introduce the Legendre transform variable Γ,

$$\Gamma[G_1, \hat{G}_1] = \ln Z[\xi, \hat{\xi}] - \xi(\mathbf{1})G_1(\mathbf{1}) - \hat{\xi}(\mathbf{1})G_1(\hat{\mathbf{1}}) \tag{6.5.4}$$

236

where $\mathbf{i}$ and $\hat{\mathbf{i}}$ refer to $\rho(\mathbf{i})$ and $\hat{\rho}(\mathbf{i})$, respectively. When the variable is either $\rho(\mathbf{i})$ or $\hat{\rho}(\mathbf{i})$ it appears as $\tilde{\mathbf{i}}$. The Legendre transform formally closes the Dyson's equation. Note that the Green's functions generated by $Z[\xi, \hat{\xi}]$ are already averaged over the random potential and the $\langle \ldots \rangle$ is omitted. We assume that $\Gamma_2(\tilde{\mathbf{1}}, \tilde{\mathbf{2}}) = -G_2^{-1}(\tilde{\mathbf{1}}, \tilde{\mathbf{2}})$. This equality is not necessarily strict but it allows simple manipulation of the equations. The equations derived by MSR allow us to use perturbation theory to systematically expand and evaluate the self-energy, as demonstrated by Deker [56, 57, 58]. The resulting set of equations is:

$$G_2(\mathbf{1}', \hat{\mathbf{1}}'') = G_0(\mathbf{1}', \hat{\mathbf{1}}'')$$

$$+ G_0(\mathbf{1}', \hat{\mathbf{1}}) \gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, 4) \left[G_2(\mathbf{2}, \hat{\mathbf{5}}) G_2(\hat{\mathbf{3}}, 6) G_2(4, \hat{\mathbf{7}}) G_2(\hat{\mathbf{1}}'', 8) \Gamma_4(\hat{\mathbf{5}}, 6, \hat{\mathbf{7}}, 8) \right.$$

$$\left. + G_2(\mathbf{2}, \hat{\mathbf{3}}) G_2(4, \hat{\mathbf{1}}'') \right] \tag{6.5.5}$$

$$\Gamma_2(\mathbf{1}', \hat{\mathbf{1}}'') = -G_2^{-1}(\mathbf{1}', \hat{\mathbf{1}}'') =$$

$$-G_0^{-1}(\mathbf{1}', \hat{\mathbf{1}}'') + \gamma(\hat{\mathbf{1}}'', \mathbf{2}, \hat{\mathbf{3}}, 4) \left[G_2(\mathbf{2}, \hat{\mathbf{5}}) G_2(\hat{\mathbf{3}}, 6), G_2(4, \hat{\mathbf{7}}) \Gamma_4(\hat{\mathbf{5}}, 6, \hat{\mathbf{7}}, \mathbf{1}') \right.$$

$$+ \sum_{\mathbf{i} = \hat{\mathbf{5}}, 6, \hat{\mathbf{7}}} G_2(\mathbf{2}, \hat{\mathbf{5}}) G_2(\hat{\mathbf{3}}, 6) G_2(4, \hat{\mathbf{7}}) G_2(\tilde{\mathbf{9}}, \tilde{\mathbf{10}}) \Gamma_3(\tilde{\mathbf{9}}, \mathbf{j}(\neq \mathbf{i}), \mathbf{k}(\neq \mathbf{i}, \mathbf{j}))$$

$$\times \Gamma_3(\tilde{\mathbf{10}}, \mathbf{1}', \mathbf{i})$$

$$+ \sum_{\mathbf{i} = \mathbf{2}, \hat{\mathbf{3}}, 4} G_1(\mathbf{i}) G_2(\mathbf{j}(\neq \mathbf{i}), \tilde{\mathbf{5}}) G_2(\mathbf{k}(\neq \mathbf{i}, \mathbf{j}), \tilde{\mathbf{6}}) \Gamma_3(\tilde{\mathbf{5}}, \tilde{\mathbf{6}}, \mathbf{1}') \right]$$

$$+ \gamma(\hat{\mathbf{1}}'', \mathbf{1}', \hat{\mathbf{3}}, 4) \left[G_2(\hat{\mathbf{3}}, 4) + G_1(\hat{\mathbf{3}}) G_1(4) \right]$$

$$+ \gamma(\hat{\mathbf{1}}'', \mathbf{2}, \hat{\mathbf{3}}, \mathbf{1}') \left[G_2(\mathbf{2}, \hat{\mathbf{3}}) + G_1(\mathbf{2}) G_1(\hat{\mathbf{3}}) \right] \tag{6.5.6}$$

$$\Gamma_3(\mathbf{1}', \hat{\mathbf{1}}'', \tilde{\mathbf{1}}''') = \frac{\delta \Gamma_2(\mathbf{1}', \hat{\mathbf{1}}'')}{\delta G_1(\tilde{\mathbf{1}}''')} \qquad (6.5.7)$$

$$\Gamma_4(\mathbf{1}', \hat{\mathbf{1}}'', \mathbf{1}''', \hat{\mathbf{1}}^{iv}) = \frac{\delta^2 \Gamma_2(\mathbf{1}', \hat{\mathbf{1}}'')}{\delta G_1(\mathbf{1}'''), \delta G_1(\hat{\mathbf{1}}^{iv})} \qquad (6.5.8)$$

with $G_0(\mathbf{1}, \hat{\mathbf{2}})$ defined as

$$G_0(\mathbf{1}, \hat{\mathbf{2}}) = \theta(t_1 - t_2)\delta(\mathbf{k_1} - \mathbf{k_2}) \exp\left[-D_0 |t_1 - t_2||\mathbf{k_1} - \mathbf{k_2}|^2\right]$$

The MSR G_0 above differs from its analogue in the direct perturbation calculation in section 6.2, because it is a function on two different sets of coordinates. The MSR Green's function is a measure of the response of the system at the coordinate labeled $\mathbf{1}$ to the introduction of density at coordinate $\hat{\mathbf{2}}$ so the zeros of the system are still arbitrary. The step function, $\theta(t)$, enforces causality and the δ function enforces translations invariance. To recover the usual Green's function with the creation event centered at the origin in real space at time $t = 0$, we set $t_2 = 0$ and integrate over $\mathbf{k_2}$. The variables of the form $\tilde{\mathbf{i}}$ must be integrated over both arguments $\hat{\mathbf{i}}$ and $\mathbf{i}$, but many of the terms that they represent are zero. The result is similar to Deker's result for a cubic field at the 2nd order [56]. For the 2nd order expression, we evaluate Γ_4 to 1st order. Γ_4 has two terms to 1st order.

$$\Gamma_4(\mathbf{1}', \hat{\mathbf{1}}'', \mathbf{1}''', \hat{\mathbf{1}}^{iv}) = \gamma(\hat{\mathbf{1}}'', \mathbf{1}', \hat{\mathbf{1}}^{iv}, \mathbf{1}''') + \gamma(\hat{\mathbf{1}}'', \mathbf{1}''', \hat{\mathbf{1}}^{iv}, \mathbf{1}')$$

We substitute this expression for Γ_4 in Eq. 6.5.5 and G_2 becomes

$$G_2(\mathbf{1}', \hat{\mathbf{1}}'') = G_0(\mathbf{1}', \hat{\mathbf{1}}'')$$

$$+ G_0(\mathbf{1}', \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}'')$$

$$+ G_0(\mathbf{1}', \hat{\mathbf{1}})\gamma(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})G_2(\hat{\mathbf{3}}, \mathbf{6})G_2(\mathbf{4}, \hat{\mathbf{7}})G_2(\hat{\mathbf{1}}'', \mathbf{8})\gamma(\hat{\mathbf{5}}, \mathbf{6}, \hat{\mathbf{7}}, \mathbf{8})$$

$$+ G_0(\mathbf{1}', \hat{\mathbf{1}})\gamma(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})G_2(\hat{\mathbf{3}}, \mathbf{6})G_2(\mathbf{4}, \hat{\mathbf{7}})G_2(\hat{\mathbf{1}}'', \mathbf{8})\gamma(\hat{\mathbf{5}}, \mathbf{8}, \hat{\mathbf{7}}, \mathbf{6})$$

$$\text{(6.5.9)}$$

The last term is zero because it violates causality and it would be zero in the replica action because of the $N \to 0$ limit. Integrating over $\mathbf{k_1}''$ and setting $t_{1''} = 0$ results in the 2nd order expression, which is identical to equation 6.2.3.

$$G(\mathbf{k}, t) = G_0(\mathbf{k}, t)$$

$$- \frac{(\beta D_0)^2}{(2\pi)^d} \int d\mathbf{h} d\tau_1 d\tau_2 \{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h})\mathbf{h} \cdot (\mathbf{k} - \mathbf{h})\chi(|\mathbf{k} - \mathbf{h}|, |\tau_1 - \tau_2|)$$

$$\times G_0(\mathbf{k}, t - \tau_1)G(\mathbf{h}, \tau_1 - \tau_2)G(\mathbf{k}, \tau_2)\}$$

$$+ \frac{(\beta D_0)^4}{(2\pi)^{2d}} \int d\mathbf{h_1} d\mathbf{h_2} d\tau_1 d\tau_2 d\tau_3 d\tau_4 \{\mathbf{k} \cdot (\mathbf{k} - \mathbf{h_1})\mathbf{h_2} \cdot (\mathbf{k} - \mathbf{h_1})$$

$$\times \mathbf{h_1} \cdot (\mathbf{k} - \mathbf{h_1})(\mathbf{k} - \mathbf{h_1}) \cdot (\mathbf{k} - \mathbf{h_1} + \mathbf{h_2})\chi(|\mathbf{k} - \mathbf{h_1}|, |\tau_1 - \tau_3|)$$

$$\times \chi(|\mathbf{h_1} - \mathbf{h_2}|, |\tau_2 - \tau_4|)G_0(\mathbf{k}, t - \tau_1)G(\mathbf{h_1}, \tau_1 - \tau_2)G(\mathbf{h_2}, \tau_2 - \tau_3)$$

$$\times G(\mathbf{k} - \mathbf{h_1} + \mathbf{h_2}, \tau_3 - \tau_4)G(\mathbf{k}, \tau_4)\} \qquad \text{(6.5.10)}$$

The average of $G(\mathbf{k}, t)$ over disorder is implied. The perturbation expansion of

the 2nd order term has a simple graphical expression that follows the notation used by Deker. As can be seen from the Dyson series above Eqs. 6.5.5-6.5.8, the MSR perturbation method expresses the Green's function in terms of a self-energy term, which explains the equivalence between the directly renormalized perturbation and the MSR results.

Two particle propagator

The 1st order perturbation result with MSR is obtained from a simpler procedure than the direct perturbation result, but we do not determine a self-consistent equation. A self-consistent equation can be calculated with some complexity, but the definition of a two-particle propagator, $G^{(1,2)}$, is a natural result of the vertex functions. We start by setting $N = 2$ in the replica generating function and defining the two one-particle propagators, $G^{(1)}$ and $G^{(2)}$, where the superscripts have a similar meaning as section 6.2.2, but $G^{(2)}$ does not contain $\exp[i\mathbf{k} \cdot \mathbf{r}]$. We also define a two particle connected Green's function as

$$G_4(\mathbf{1}', \hat{\mathbf{2}}', \mathbf{3}', \hat{\mathbf{4}}') = \frac{\delta^4 Z[\xi_i, \hat{\xi}_i]}{\delta\xi_1(\mathbf{1}')\delta\hat{\xi}_1(\mathbf{2}')\delta\xi_2(\mathbf{3}')\delta\hat{\xi}_2(\mathbf{4}')}. \tag{6.5.11}$$

Similar to the one particle propagator, we must add the initial conditions by setting $t_{2'} = t_{4'} = 0$, multiplying by $\exp[i\mathbf{k_{4'}} \cdot \mathbf{r}]$ and integrating over $\mathbf{k_{4'}}$ and $\mathbf{k_{2'}}$. Finally, we remove the primes on the labels and change the label of $\mathbf{k_3}$ to $\mathbf{k_2}$ to recover the first correction to the two particle propagator which is captured in the expansion of the self-consistent equation, Eq. 6.2.4.

$$G^{(1,2)}(\mathbf{k_1}, t_1 = t, \mathbf{k_2}, t_2 = t) = G_2^{(1)}(\mathbf{k_1}, t)G_2^{(2)}(\mathbf{k_2}, t)\exp[i\mathbf{k_2}\cdot\mathbf{r}]$$

$$-\frac{(\beta D_0)^2}{(2\pi)^d}\int d\mathbf{h}\int_0^t d\tau_1\int_0^t d\tau_2\,\{\mathbf{k_1}\cdot(\mathbf{k_1}-\mathbf{h})\mathbf{k_2}\cdot(\mathbf{k_1}-\mathbf{h})$$

$$\times\chi(|\mathbf{k_1}-\mathbf{h}|,|\tau_1-\tau_2|)G^{(1)}(\mathbf{k_1}, t-\tau_1)G^{(2)}(\mathbf{k_2}, t-\tau_2)$$

$$\times\,G^{(1)}(\mathbf{h},\tau_1)G^{(2)}(\mathbf{k_1}+\mathbf{k_2}-\mathbf{h},\tau_2)\exp[i(\mathbf{k_1}+\mathbf{k_2}-\mathbf{h})\cdot\mathbf{r}]\} \qquad (6.5.12)$$

6.5.3 MSR with replica trick and Gaussian reference system

The MSR perturbation is an asymptotic expansion, which may not have good accuracy at large disorder strengths. To overcome the difficulties of asymptotic expansions, several references introduce a variational technique that attempts to minimize the errors. These variational methods are referred to in the literature as the Gaussian reference technique [18, 55, 54, 160]. To first order these techniques resemble the Edwards variational method. In the Gaussian reference technique the full action is fit with a new action that only contains a quadratic term, but the term is not necessarily Gaussian. The technique avoids the Dyson equation and the vertex renormalization used in section 6.5.2. With the time dependent generating function for $Z^N[0,0]$ derived in Eq. 6.5.2 the Gaussian reference technique follows the same procedure as several references except that it is necessary to integrate over the time or frequency variable as well as the spatial variables [25, 55, 54]. The results of this technique are time dependent analogues of the results in the references and reduce to

these previous expresssions in the static disorder limit. For 1st order we get the first
two terms derived in Eq. 6.5.9

$$G_2(\mathbf{1'}, \hat{\mathbf{1}}'') = G_0(\mathbf{1'}, \hat{\mathbf{1}}'') + G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}''), \qquad (6.5.13)$$

but the 2nd order expression is more complex

$$2G_0(\mathbf{1'}, \hat{\mathbf{1}}'') = 3G_2(\mathbf{1'}, \hat{\mathbf{1}}'') - G_2(\mathbf{1'}, \hat{\mathbf{1}})G_0^{-1}(\hat{\mathbf{1}}, \mathbf{2})G_2(\mathbf{2}, \hat{\mathbf{1}}'')$$

$$+ 4G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}'')$$

$$- 2G_2(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}'')$$

$$+ G_2(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})G_0^{-1}(\hat{\mathbf{5}}, \mathbf{6})G_2(\mathbf{6}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}'')$$

$$- G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{5}})\gamma(\hat{\mathbf{5}}, \mathbf{6}, \hat{\mathbf{7}}, \mathbf{8})G_2(\mathbf{6}, \hat{\mathbf{7}})G_2(\mathbf{8}, \hat{\mathbf{1}}'')$$

$$- G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})\gamma(\hat{\mathbf{5}}, \mathbf{6}, \hat{\mathbf{7}}, \mathbf{8})G_2(\mathbf{6}, \hat{\mathbf{7}})G_2(\mathbf{8}, \hat{\mathbf{3}})G_2(\mathbf{4}, \hat{\mathbf{1}}'')$$

$$- G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})G_2(\hat{\mathbf{3}}, \mathbf{6})G_2(\mathbf{4}, \hat{\mathbf{7}})G_2(\hat{\mathbf{1}}'', \mathbf{8})\gamma(\hat{\mathbf{5}}, \mathbf{6}, \hat{\mathbf{7}}, \mathbf{8})$$

$$+ G_0(\mathbf{1'}, \hat{\mathbf{1}})\gamma(\hat{\mathbf{1}}, \mathbf{2}, \hat{\mathbf{3}}, \mathbf{4})G_2(\mathbf{2}, \hat{\mathbf{5}})G_2(\hat{\mathbf{3}}, \mathbf{6})G_2(\mathbf{4}, \hat{\mathbf{7}})G_2(\hat{\mathbf{1}}'', \mathbf{8})\gamma(\hat{\mathbf{5}}, \mathbf{8}, \hat{\mathbf{7}}, \mathbf{6})$$

$$(6.5.14)$$

Expanding terms that contain G_0^{-1} or no G_0 demonstrates that this expression is the
2nd order perturbative term, Eq. 6.5.9, with additional third and higher order terms.
These additional terms may improve the fit but they generally make numerically
solving the self-consistent equations difficult by adding complexity to the equations.
Because of this difficulty, we do not solve the 2nd order reference system calculation

242

in this article.

The equality at 2nd order is not unexpected because there is only one 2nd order graph that is not contained in the graphical expansion of the 1st order self-consistent equation. The Gaussian reference method has many terms that appear redundant. At 2nd order the terms cancel but they do not necessarily cancel at higher order. As a result, this method over counts some graphs and subtracts graphs that should be added. The expansion is also much more complicated and does not allow systematic diagrammatic analysis. Determining additional graphs to include in the expansion is difficult so expanding beyond 2nd order requires one to start from the 3rd order variational expression and re-evaluate the higher order terms.

6.6   Renormalization Group Result

To further demonstrate the consistency of our results with other MSR methods, we perform a simple 1st order renormalization group calculation to determine the effective diffusion constant. The calculation parallels several other one-loop approaches [55, 37]. More sophisticated renormalization group calculations on the action in Eq. 6.5.2 have been used to determine the diffusion constant in static random media problems in several references as well [123, 124]. Our approach is not general, but it allows the incorporation of dynamic disorder and the spirit of these calculations can be implemented for other forms of disorder. This approach also recovers the general form suggested by Dean and by Deem in the static limit, $\lambda \to 0$ [51, 55].

The calculation begins by Fourier transforming the time variable, $t \to \omega$ so that

$\mathbf{1} = [\mathbf{k_1}, \omega_1]$ and $G_0^{-1} = i\omega + D|\mathbf{k}|^2$ in our generating function. Since the calculation is to first order we can take the number of replicas to be one. We introduce an artificial cutoff frequency in the spatial transform variable $\mathbf{k}$, denoted k_c, and define a momentum shell composed of frequencies that we eventually integrate over $(k_c/b \leq |\mathbf{k}| \leq k_c)$ with $b > 1$. To first order we can replace $\hat{\rho}(\mathbf{i})\rho(\mathbf{j})$ with $\mathbf{k_i}, \mathbf{k_j} > k_c/b$ with $\delta(\mathbf{i} - \mathbf{j})(i\omega_i + D|\mathbf{k_i}|^2)^{-1}$. With these substitutions the action in our functional is

$$-G_0^{-1}(\mathbf{1}, \mathbf{2})\hat{\rho}(\mathbf{1})\rho(\mathbf{2}) + \gamma(\mathbf{1}, \mathbf{1}', \mathbf{1}', \mathbf{2})G_0(\mathbf{1}', \mathbf{1}')\hat{\rho}(\mathbf{1})\rho(\mathbf{2})$$

$$+ \frac{1}{2}\gamma(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})\hat{\rho}(\mathbf{1})\rho(\mathbf{2})\hat{\rho}(\mathbf{3})\rho(\mathbf{4}) + const \qquad (6.6.1)$$

where the $\mathbf{1}'$ is integrated for $k_c/b \leq |\mathbf{k}'| \leq k_c$, the other $\mathbf{k}$ variables are integrated for $|\mathbf{k}| < k_c/b$, and all frequencies are integrated from $-\infty$ to ∞. The constant term comes from integrating the terms with $|\mathbf{k}| > k_c/b$ for all ρ and $\hat{\rho}$. This term only changes the normalization and will be omitted in further calculations.

Up to this point the random potential correlation function, χ has been general. Now we introduce the 3 dimensional χ in Eq. 6.1.3 and evaluate the integrals over ω' and $\mathbf{k}'$. Integrating over ω' is done in a straightforward manner, but the $\mathbf{k}'$ integral has some difficulty. Since the final form should resemble a free diffusion propagator, we will perform a Taylor expansion of the integral in terms of $\mathbf{k_2}$ up to 2nd order and assume the other terms are small. The new action is

$$-\int d\mathbf{1} d\mathbf{2} \Big\{ \delta(\mathbf{1} - \mathbf{2}) \left(i\omega_2 + \left(D - \frac{4}{3\sqrt{2\pi}}(2\alpha)^{3/2}\chi_0 \frac{(\beta D)}{D + \lambda} \right. \right.$$

$$\times \int dr \left\{ \left[2 - \alpha r^2 - \frac{\lambda}{D + \lambda} \right] r^2 \exp[-\alpha r^2] \right\} \Big) |\mathbf{k_2}|^2 \Big) \hat{\rho}(\mathbf{1})\rho(\mathbf{2}) \Big\}$$

244

$$+ \frac{1}{2}\left(\frac{\beta D}{(2\pi)^2}\right)^2 \int d1d2d3d4\,\delta(1 - 2 + 3 - 4)\left\{\mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{k_2})\mathbf{k_3} \cdot (\mathbf{k_3} - \mathbf{k_4})\right.$$

$$\times 2(2\pi)^{\frac{3}{2}}(2\alpha)^{\frac{3}{2}}\chi_0\lambda|\mathbf{k_1} - \mathbf{k_2}|^2 \frac{\exp[-\alpha|\mathbf{k_1} - \mathbf{k_2}|^2]}{\lambda^2|\mathbf{k_1} - \mathbf{k_2}|^4 + (\omega_1 - \omega_2)^2}$$

$$\left.\times \hat{\rho}(\mathbf{1})\rho(\mathbf{2})\hat{\rho}(\mathbf{3})\rho(\mathbf{4})\right\} \tag{6.6.2}$$

with $k_c/b < r = |\mathbf{k}| < k_c$ and $|\mathbf{k}| < k_c/b$. To get this result, a term of the form $(i\omega + (D + \lambda)r^2)^{-1}$ is replaced with $((D + \lambda)r^2)^{-1}$ since r is large and the major contribution from ω are for $\omega \approx 0$. Rescaling $\mathbf{k}$ by b, ω by b^2 and the field variables, ρ and $\hat{\rho}$, by $b^{-\frac{7}{2}}$ results in an action with the same form as our original and the proper limits of integration.

$$-\int d1d2\left\{\delta(\mathbf{1} - \mathbf{2})\left(i\omega_2 + \left(D - \frac{4}{3\sqrt{2\pi}}(2\alpha)^{3/2}\chi_0\frac{(\beta D)^2}{D + \lambda}\right.\right.\right.$$

$$\times \int dr\left\{[2 - \alpha r^2 - \frac{\lambda}{D + \lambda}]r^2 \exp[-\alpha r^2]\right\}\bigg)|\mathbf{k_2}|^2\bigg)\hat{\rho}(\mathbf{1})\rho(\mathbf{2})\bigg\}$$

$$+ \frac{1}{2b^3}\left(\frac{\beta D}{(2\pi)^2}\right)^2 \int d1d2d3d4\,\delta(\mathbf{1} - \mathbf{2} + \mathbf{3} - \mathbf{4})\left\{\mathbf{k_1} \cdot (\mathbf{k_1} - \mathbf{k_2})\mathbf{k_3} \cdot (\mathbf{k_3} - \mathbf{k_4})\right.$$

$$\times 2(2\pi)^{\frac{3}{2}}(2\alpha)^{\frac{3}{2}}\chi_0\lambda|\mathbf{k_1} - \mathbf{k_2}|^2 \frac{\exp[-\alpha\frac{\mu}{b^2}|\mathbf{k_1} - \mathbf{k_2}|^2]}{\lambda^2|\mathbf{k_1} - \mathbf{k_2}|^4 + (\omega_1 - \omega_2)^2}$$

$$\left.\times \hat{\rho}(\mathbf{1})\rho(\mathbf{2})\hat{\rho}(\mathbf{3})\rho(\mathbf{4})\right\} \tag{6.6.3}$$

μ is a new scale factor that modulates the $|\mathbf{k}|^2$ term in the exponential. Before the 1st iteration of the renormalization group, $\mu = 1$. From this equation we determine the relationship between the old parameters and the new parameters, denoted with

a prime.

$$\mu' \to \frac{\mu}{b^2}$$

$$\chi'_0 \to \frac{\chi_0}{b^3} \left(\frac{D}{D'}\right)^2$$

$$D' \to D - \frac{4}{3\sqrt{2\pi}} (2\alpha)^{3/2} \chi_0 \frac{(\beta D)^2}{D+\lambda} \int dr \left\{ \left[2 - \alpha\mu r^2 - \frac{\lambda}{D+\lambda}\right] r^2 \exp[-\alpha\mu r^2] \right\}$$

By rescaling χ_0 by $\left(\frac{D}{D'}\right)^2$ we replace the D^2 in the quartic term with $(D')^2$. Choosing $b = \exp[\Delta\zeta]$ with $\Delta\zeta \sim 0$ allows us to write approximate equations for the changes in these variables. These approximations will become accurate in the $\lim_{\Delta\zeta \to 0}$.

$$\Delta\mu \to -\mu 2\Delta\zeta$$

$$\Delta\chi_0 \to -\chi_0 3\Delta\zeta$$

$$+ \chi_0^2 \frac{8}{3\sqrt{2\pi}} (2\alpha)^{3/2} \frac{(\beta D)^2}{D+\lambda} \left\{ \left[2 - \alpha\mu k_c^2 - \frac{\lambda}{D+\lambda}\right] k_c^3 \exp[-\alpha\mu k_c^2] \right\} \Delta\zeta$$

$$\Delta D \to -\frac{4}{3\sqrt{2\pi}} (2\alpha)^{3/2} \chi_0 \frac{(\beta D)^2}{D+\lambda} \left\{ \left[2 - \alpha\mu k_c^2 - \frac{\lambda}{D+\lambda}\right] k_c^3 \exp[-\alpha\mu k_c^2] \right\} \Delta\zeta$$

The 2nd term in the flow expression for χ_0 can be made arbitrarily small by choosing a large enough cutoff frequency, k_c. This assumption fails when the μ becomes extremely small, but χ_0 can be arbitrarily small at the point of the failure, which allows us to neglect this term. This argument depends on the decay of χ_0 being faster than the decay of μ. From the first two equations, μ and χ_0 have a simple exponential form in terms of $\zeta = N\Delta\zeta$ where N is the number of iterations of the RG calculation.

246

$$\mu = \exp[-2\zeta]$$

$$\chi_0 = [\chi_0]_{\zeta=0} \exp[-3\zeta]$$

Substituting these expressions into the expression for D leads to the expression

$$
\begin{aligned}
\frac{dD}{d\zeta} = &-\frac{4}{3\sqrt{2\pi}}(2\alpha)^{3/2}[\chi_0]_{\zeta=0}\frac{(\beta D)^2}{D+\lambda} \\
&\times \left\{ \left[2 - \alpha\exp[-2\zeta]k_c^2 - \frac{\lambda}{D+\lambda} \right] k_c^3 \exp[-3\zeta]\exp[-\alpha\exp[-2\zeta]k_c^2] \right\}
\end{aligned}
$$

(6.6.4)

The new diffusion constant is the value of the solution to this equation at $\zeta = \infty$ with $D = D_0$ (the free diffusion value) at $\zeta = 0$. The equation simplifies further by introducing a new variable of integration $k = -\sqrt{\alpha}k_c\exp[-\zeta]$, and defining dimensionless quantities $D' = \frac{D}{D_0}$ and $\lambda' = \frac{\lambda}{D_0}$. In the limit as the cutoff frequency goes to infinity, the initial conditions are $D' = 1$ at $k = -\infty$, and the solution is D' at $k = 0$. With these substitutions the equation for D' is straightforward in form.

$$\frac{dD'}{dk} = -\sigma\frac{k^2 D'^2}{D'+\lambda'}\left(2 - k^2 - \frac{\lambda'}{D'+\lambda'} \right)\exp[-k^2]$$

(6.6.5)

with

$$\sigma = \frac{8}{3\sqrt{\pi}}(\beta^2[\chi_0]_{\zeta=0})$$

For $D \approx D_0$, the equation is just an integral and the result is the 1st order non-renormalized perturbation result

247

$$D = D_0 - D_0^2 \frac{(D_0 - \lambda)}{(D_0 + \lambda)^2} \frac{\beta^2 \chi_0}{3}. \tag{6.6.6}$$

This expression can also be derived by substituting $D = D_0$ in Eq. 6.3.3 and shows that the expression has the correct initial slope. The expression also recovers the static disorder limit ($\lambda \to 0$) reported in several references [55, 54, 51, 52, 164, 37],

$$D = D_0 \exp\left[-\frac{\beta^2 \chi_0}{3}\right]. \tag{6.6.7}$$

A discussion of the comparison between the RG results and the perturbation results was presented in section 6.4. As discussed in section 6.4.1, the RG calculation presented here gives the correct first order perturbation result and examination of the flow equation, Eq. 6.6.5, demonstrates that the solution does not go to zero for finite disorder strength. As a result, the particle avoids trapping. As mentioned in Sec. 6.4.1, the particle should exhibit this non-trapping behavior. The solution to this equation for the dynamic potential remains above $\exp[-\frac{\beta^2 \chi_0}{3}]$, the static disorder solution in Eq. 6.6.7, which is also above the lower bound predicted by De Masi et al. [55, 59].

6.7   Discussion of Results

In this paper, we have extended previous MSR results for diffusion to a random potential with both spatial and temporal correlations. We perform a Dyson expansion

248

to develop renormalized propagators for one- and two-particle systems. These propagators determine the characteristics of the system including the diffusion constant and non-Gaussian indicators. Most of the results are general for arbitrary dynamic potential correlation function χ or can be generalized by following the spirit of these calculations.

The field theoretic method developed by Martin, Siggia, and Rose is shown to be consistent with the direct Dyson expansion. A perturbation expansion using the MSR method yields the same single particle propagator to 2nd order and the same 1st order expression for the two-particle propagator. The field theoretic method of MSR can also be used to determine an Edwards type of variational fit of the propagator, which has the same 1st order expression as perturbation but a more complicated 2nd order expression. The variational approach is also consistent with previous static calculations [55]. The diffusion constant is also determined from a renormalization group calculation. These results are consistent with previous work in the static limit and give a reasonable generalization to dynamic disorder [51, 55].

The renormalized perturbation expansion used to determine the diffusion constant demonstrates the expected behavior of a perturbation expansion. The results match the non-renormalized expansion for small values of disorder strength but eventually they deviate from reasonable behavior and predict trapping. The dynamics of the potential correlation function, λ, increased the diffusion constant because any barriers to diffusion would eventually rearrange and allow the particle to move. Contrary to this intuitive result, the renormalized perturbation expansion still predicts trapping for finite disorder strength. A renormalization group calculation with the MSR for-

malism shows more reasonable behavior with the diffusion constant decreasing as an exponential with respect to the disorder strength in the static limit, which is consistent with previous calculations and simulations, and as a power-law for a system with non-zero λ [55].

The correlation function determined by the perturbation expansion for the two particle Green's functions exhibit collective behaviors that can be interpreted as clustering in this model. Particles that originate near each other have a tendency to diffuse with similar rates and move closer together. This behavior results in long lived correlations that are apparent even in the non-renormalized expressions. Since perturbation expansions have a tendency to over-emphasize the effects of the potential, we expect that the renormalized propagator will demonstrate even stronger correlations between the particles.

Although the model was chosen for computational convenience, the correlations exhibit behaviors that are similar to real systems like glasses and supercritical fluids. This study also demonstrates that the analytical and computational methods used in this paper can be applied to the diffusion of a solute in real systems with a potential-potential correlation function determined for these systems.

Chapter 7

Continuous Time Random Walks: aging correlations and coupling to kinetics

Several single molecule tracking experiments suggest the existence of dispersive transport mechanisms for the motions in glasses, cells and proteins [152, 181, 216, 2, 96, 128, 182]. The general feature of dispersive transport is complex non-exponential temporal relaxation, which can be phenomenologically fit with a long time power-law tail or a stretched exponential. Phenomenological dispersive transport models often incorporate the non-exponential relaxation into diffusion problems by replacing the exponential waiting-time distribution of traditional diffusion processes with a more complex waiting-time distribution through subordination, which results in a continuous time random walk (CTRW) or the related fractional Fokker-Planck equation

(FFPE) for sub-diffusion [79, 116, 194]. The origin of the complex waiting time has been extensively explored in simulations of glassy systems [64, 32] and is attributed to activation processes with randomly distributed energy barriers [233, 25, 153].

Many processes appear to obey power-law statistics over several decades before achieving an equilibrium state where the time correlation of the system becomes stationary [39, 225]. Examples include quantum dot blinking experiments that show a laser power dependent cutoff in the blinking statistics [39], and photon emission statistics in proteins, whose complex correlation function shows long time exponential relaxation [225]. Since aging would cause the system's behavior to change over time and additional processes that can truncate the aging in the system are almost always present, it is not surprising to see a cutoff in the power-law correlations observed in experiments [39, 225]. These truncations would have noticeable effects on biological processes where the complex transport dynamics are coupled to first order processes [200, 182, 127]. In these experiments, a kinetic rate of a process depends on a distance coordinate, which moves according to a CTRW or other dispersive transport model [77]. A simple example is an electron or proton transfer reaction where the rate of transfer depends on the distance between the donor and the acceptor [240]. If distance is modulated by a dispersive transport process, one expects non-exponential kinetics for the reaction.

The truncation of the aging processes and the inclusion of first order kinetics in dispersive transport models can be treated within the same framework by simply including a first order kinetic scheme that preserves probability to introduce a cutoff or removes probability to introduce reactions. In order to understand the applicability

of the anomalous diffusion models with first order kinetics to biological experiments, one must address issues associated with the coupling of these mechanisms. Several authors previously considered coupling of dispersive transport with a single δ function sink [185, 201, 200]. Here we consider more general scenarios and derive several possible limiting cases. To extend applications of FFPE to systems whose aging processes are interrupted by truncation or reaction requires us to develop a general description of trapping processes [139, 233, 25, 153].

7.1 General two-time Semi-Markov construction

A semi-Markov process is the most general description of trapping models with complex waiting times. In a semi-Markov process, the probability of making a transition from a state j to state i at a time t is given by the matrix element $Q_{ij}(t)dt$ with $Q_{ij}(t) \geq 0$ for $t > 0$ and $\sum_i \int dt Q_{ij}(t) \leq 1$ for all j [79]. If $\sum_i \int dt Q_{ij}(t) < 1$ the particle may stay at site i permanently, i.e. i is a trap. The single time propagator for this process in the Laplace domain, $t \to s$, is

$$\mathbf{G}_{ij}(s) = \frac{1 - \sum_k Q_{ki}(s)}{s} \left[\mathbf{I} - \mathbf{Q}(s)\right]_{ij}^{-1}. \tag{7.1.1}$$

In this expression $\mathbf{I}$ is the identity matrix and $\mathbf{Q}$ is the complete matrix of $Q_{ij}(s)$. The $\left[\mathbf{I} - \mathbf{Q}(s)\right]_{ij}^{-1}$ is the Laplace transform of the probability of jumping into the state i at time t given that we started in state j and $\frac{1 - \sum_k Q_{ki}(s)}{s}$ is the probability of not jumping out of the final state, i [79]. In this expression, the hopping process begins

at $t = 0$. In other words, the beginning of the hopping process is well defined, such as the temperature quenching time in glass forming experiments.

More generally, the measurement does not begin at the quenching time. In order to explore this scenario, we examine a two-time propagator in double Laplace space $t_1 \rightarrow s_1$ and $t_2 \rightarrow s_2$ [79],

$$
\mathbf{G}_{ijk}(s_1, s_2) = \frac{1 - \sum_l Q_{li}(s_2)}{s_2} \cdot \sum_m [\mathbf{I} - \mathbf{Q}(s_2)]_{im}^{-1} \frac{[Q_{mj}(s_1) - Q_{mj}(s_2)]}{s_2 - s_1} [\mathbf{I} - \mathbf{Q}(s_1)]_{jk}^{-1}
$$
$$
+ \delta_{ij} \frac{\left[\frac{1 - \sum_l Q_{lj}(s_1)}{s_1} - \frac{1 - \sum_l Q_{lj}(s_2)}{s_2} \right]}{s_2 - s_1} [\mathbf{I} - \mathbf{Q}(s_1)]_{jk}^{-1}
$$

$$(7.1.2)$$

The $\frac{1 - \sum_l Q_{li}(s)}{s}$ and $[\mathbf{I} - \mathbf{Q}(s)]^{-1}$ terms correspond to the same processes in Eq. 7.1.1, but $\frac{[Q_{mj}(s_1) - Q_{mj}(s_2)]}{s_2 - s_1}$ corresponds to the jump that spans both time intervals t_1 and t_2, and $\delta_{ij} \left[\frac{1 - \sum_l Q_{lj}(s_1)}{s_1} - \frac{1 - \sum_l Q_{lj}(s_2)}{s_2} \right] \frac{1}{s_2 - s_1}$ corresponds to failing to make a jump during the second time interval [79].

7.2 A microscopic trapping picture for $\mathbf{Q}(s)$

A possible origin of $\mathbf{Q}(s)$ is the trapping processes observed in glassy systems [64, 25, 153]. If the experiment monitors a few labeled tracer particles in a glassy matrix, the tracers spend most of their time trapped in cages formed by the matrix and can only move when a trap rearranges through an activation process. If the rearrangements are large enough, we would not expect correlations between the trapping times and/or the displacements [25, 153]. The trapping times would be site specific with no directional

254

dependence, $Q_{ij}(t) = W_{ij}\psi_j(t)$. The matrix $\mathbf{W}$ controls the directional aspects of the particle motions and can be approximated as $\mathbf{W} \approx e^{\mathbf{L}_{FP}\Delta t}$, where Δt is a small time step and $\mathbf{L}_{FP}$ can be a kinetic matrix or an operator, $\mathbf{L}_{FP}\rho(x, t) = \nabla \cdot [D(x)\nabla\rho(x, t)] + \nabla \cdot [\rho(x, t)\nabla V(x)]$, with diffusion constant $D(x)$ and potential $V(x)$.

Bouchaud [25, 153] introduced a waiting time that we denote by $\phi_j(t)$ through an activated process with possibly site specific random energy barriers, $\phi_j(t) = \int dE_j \phi(t|E_j) P(E_j)$ with $\phi(t|E_j) \approx \tau_{0j}^{-1} e^{-\beta E_j} \exp\left(\tau_{0j}^{-1} e^{-\beta E_j} t\right)$, and τ_{j0} is a weakly temperature dependent prefactor [25, 153]. If the energy barriers are exponentially distributed for each site $P(E_j) = \alpha_j e^{-\alpha_j E_j}$ the waiting time has an asymptotic power-law tail $\phi_j(t) \sim \tau_{0j}^{-1} \frac{\alpha_j}{\beta} \Gamma[1 + \frac{\alpha_j}{\beta}] [t/\tau_{0j}]^{-(1 + \frac{\alpha_j}{\beta})}$. The distribution of barrier heights, $P(E_j)$, is phenomenologically based on extreme value arguments that are not easily justified due to the sensitivity of the waiting time distribution to the non-universal tails of the barrier distribution [222].

The sensitivity to the non-universal tails suggests that the tail of the waiting time distribution should be modified. Wolynes introduced a cutoff in the free energy barriers of glasses because there are many competing processes resulting in a rearrangement with a finite sized barrier height always being possible [233]. We introduce a slightly different cutoff by adding a competing pathway of fixed energy barrier height, i.e. take the fastest times between an exponential process, $k_j e^{-k_j t}$ and $\phi_j(t)$. This approach is similar to the coupling scheme of first order and power-law dynamics discussed elsewhere [230]. Having only two such pathways is a simplification that follows if there are two possible classes of rearrangements to the system, ϕ corresponding to a broad distribution of energy barriers and ke^{-kt} corresponding to

a narrow distribution of deep energy barriers.

In the case of an exponential process competing with Bouchaud's waiting time distribution, our waiting time is $\psi_j(t) = \phi_j(t)e^{-k_i t} + k_j e^{-k_j t} \int_0^\infty \phi_j(t)$ [201]. If $\alpha_j = \alpha$, $\tau_{0j} = \tau_0$ and $k_j = k$ for all lattice sites, we can write an equation for the time evolution of the distribution of Bouchaud's energy barrier [25, 153],

$$\partial_t P(E) = -(k + \tau_0^{-1} e^{-\beta E}) P(E) + \alpha e^{-\alpha E} \int dE'(k + \tau_0^{-1} e^{-\beta E'}) P(E'). \qquad (7.2.3)$$

The high energy cutoff introduced by k creates a microscopic equilibrium for the distribution of energy barriers and the system achieves equilibrium in a finite amount of time, $t \approx k^{-1}$. The equilibrium distribution for the barrier heights, $\partial_t \rho_{eq} = 0$ is

$$\rho_{eq}(E) = \eta \frac{\alpha e^{-\alpha E}}{k + \tau_0^{-1} e^{-\beta E}}, \qquad (7.2.4)$$

where η is the normalization. On these long timescales, we expect the system to be diffusive. For $\beta > \alpha$ and $k = 0$ we recover the trapping result for a lack of equilibrium distribution [25]. The system is microscopically far away from equilibrium for time less than k^{-1} so the system is not in a linear regime for $t < k^{-1}$ and short time violations of the fluctuation dissipation theorem (FDT) should not be surprising. The traps arrest the motions on these timescales and the system cannot respond, but the FDT becomes valid for forces applied on timescales longer than the trapping time $(t > k^{-1})$.

In order to study the properties and effects of this waiting time distribution, we

perform a Laplace transform on the waiting time distribution [201]

$$\psi_j(s) = \phi_j(s + k_j) + k_j \frac{1 - \phi_j(s + k_j)}{s + k_j}. \tag{7.2.5}$$

In the long time limit, we approximate $\phi_j(s)$ with a Lévy distribution, $\phi_j(s) = e^{-(\tau_j s)^{\gamma_j}}$, where $\gamma_j = \frac{\alpha_j}{\beta}$ is the power-law exponent of the original waiting time distribution and τ_j is a constant that defines the tail amplitude. Eq. 7.2.5 allows us to write $\mathbf{Q}(s) = \mathbf{W}\psi(s)$ and

$$\mathbf{G}(s) = \frac{1 - \psi(s)}{s} \cdot [\mathbf{I} - \mathbf{W}\psi(s)]^{-1} \tag{7.2.6}$$

Substituting $\mathbf{W} = e^{\mathbf{L}_{FP}\Delta t}$ and taking the limit $\Delta t, \tau_j \to 0$ such that $\Delta t / \tau_j^{\gamma_j} \to D_j^{\gamma}$ lead to a Green's function for an anomalous diffusion process competing with a first order process,

$$\mathbf{G}(s) \approx \left[s - \mathbf{L}_{FP}\mathbf{D}^{\gamma}(s + \mathbf{k})^{1-\gamma} \right]^{-1} \tag{7.2.7}$$

where $\mathbf{D}^{\gamma}(s + \mathbf{k})^{1-\gamma} = \delta_{ij}D_j^{\gamma}(s + k_j)^{1-\gamma_j}$. If $\gamma_j = \gamma$ and $D_j^{\gamma} = D^{\gamma}$ are the same for all sites, the Green's function becomes $\mathbf{G}(s) \approx [s - D^{\gamma}\mathbf{L}_{FP}(s + \mathbf{k})^{1-\gamma}]^{-1}$ with corresponding Green's function equation,

$$s\mathbf{G} - \mathbf{I} = D^{\gamma}\mathbf{L}_{FP}(s + \mathbf{k})^{1-\gamma}\mathbf{G}. \tag{7.2.8}$$

Notice that $\mathbf{L}_{FP}$ is applied to the interrupted fractional derivative, $(s + \mathbf{k})^{1-\gamma}$ of the Green's function, so spatial inhomogeneities in the interruption rate, $\mathbf{k}$ result in a

change in the equilibrium population.

An extensively studied example of this effect on the equilibrium is the quantum dot, which has two states, *on* and *off*. The *on* time cutoff is experimentally verified to be shorter than the *off* time cutoff [39]. If $\gamma_{on} = \gamma_{off} = \gamma$ the quantum dot Green's function is

$$\mathbf{G}_{QD} = \begin{bmatrix} s + D_{on}^{\gamma}(s + k_{on})^{1-\gamma} & -D_{off}^{\gamma}(s + k_{off})^{1-\gamma} \\ -D_{on}^{\gamma}(s + k_{on})^{1-\gamma} & s + D_{off}^{\gamma}(s + k_{off})^{1-\gamma} \end{bmatrix}^{-1} \qquad (7.2.9)$$

and the equilibrium *on* population is $p_{on}(t \rightarrow \infty) = \frac{D_{off}^{\gamma} k_{off}^{1-\gamma}}{D_{off}^{\gamma} k_{off}^{1-\gamma} + D_{on}^{\gamma} k_{on}^{1-\gamma}}$. The $D_{on/off}^{\gamma}$ term plays the role of a determining the amplitude of the power-law tails versus the short time dynamics. When $\gamma = 1$ the intensity of the QD monotonically approaches equilibrium, but for $\gamma \neq 1$, the QD can show an initial rise to an intermediate plateau before monotonic decay to equilibrium as shown in Fig. 7-1 for $D_{on/off} = 1$, $\gamma = \frac{1}{2}$, $k_{on}^{-1} = 100$, and $k_{off}^{-1} = 10000$ [39]. The transient steady state is the result of the short time dynamics sufficiently scrambling the *on* and *off* probabilities before the long time asymptotics cause the decay to the equilibrium value, which has been observed in experiment [40].

Unlike the QD example, if $k_j = k$ is the same for all j, the interrupted fractional derivative commutes with $\mathbf{L}_{FP}$ and we recover the equilibrium distribution of $\mathbf{L}_{FP}$ as the equilibrium distribution of the interrupted process. For $k_j = k$, the Green's
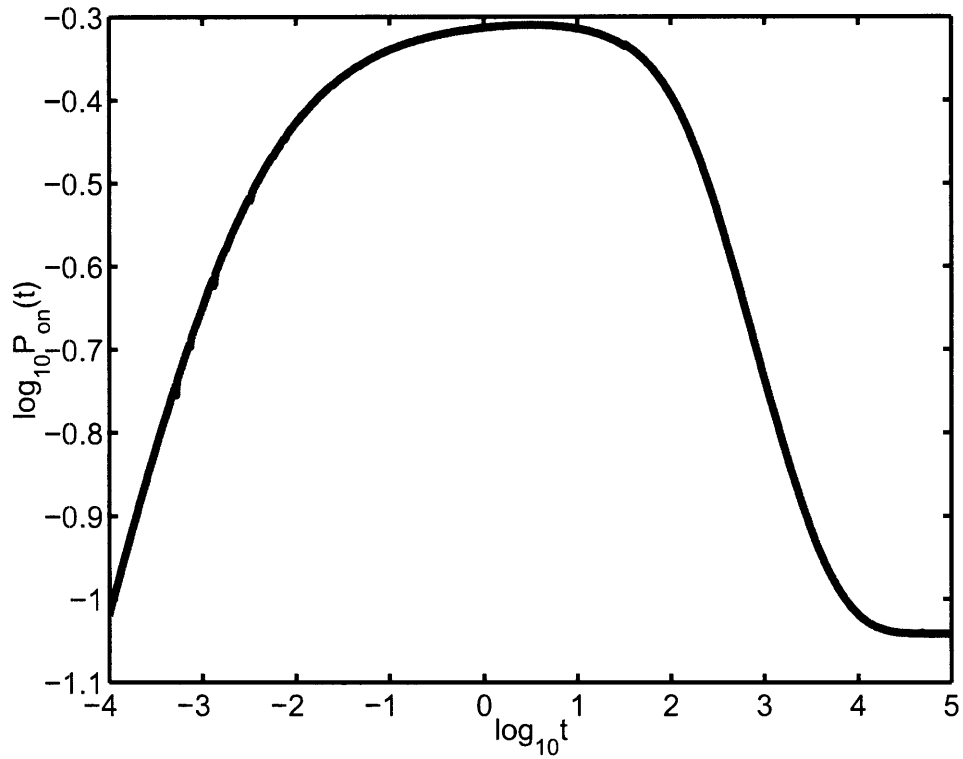
Figure 7-1: The *on* probability for a quantum dot (QD) modeled by the interrupted fractional diffusion propagator. The initial condition assumes the renewal process begins at $t = 0$ and the dot is *off*. Unlike a simple two state kinetic scheme that monotonically approaches equilibrium, the *on* probability shows a fast rise to an initial transient steady state before decaying to the long time equilibrium distribution.

function can also be re-expressed as

$$\mathbf{G}(s) = \int \frac{d\omega}{2\pi i} \frac{(s+k)^{\gamma-1}}{s(s+k)^{\gamma-1} + D^{\gamma}i\omega} \mathbf{G}_{FP}(-\omega) \tag{7.2.10}$$

where $\mathbf{G}_{FP}(\omega) = [-i\omega + \mathbf{L}_{FP}]^{-1}$ is the Green's function for the normal Fokker-Planck equation. This expression is the standard subordination result and can be rewritten in real time as [79]

$$\mathbf{G}(t) = \int_0^{\infty} dt' g(t', t) \mathbf{G}_{FP}(t') \tag{7.2.11}$$

where $g(t', t)$ is the inverse Laplace transform, $s \to t$, of

$$g(t', t) = \mathcal{L}^{-1} \left[(D^{\gamma})^{-1}(s+k)^{\gamma-1} e^{-s(s+k)^{\gamma-1}(D^{\gamma})^{-1}t'} \right]. \tag{7.2.12}$$

The interrupted fractional Green's function is simply an average of the normal diffusion Fokker-Planck Green's function over a stochastic distribution for t' determined by $g(t', t)$.

7.3 Interrupted aging and the two-time propagator

The single time propagator assumes that the trapping process begins at $t = 0$. More generally, the trapping process will begin at an earlier time than the observation time. Starting from Eq. 7.1.2 for the two-time semi-Markov propagator and following the

single time propagator derivation gives the two-time propagator

$$\mathbf{G}_{ijk}(s_1, s_2) = \frac{1 - \psi_i(s_2)}{s_2} \cdot \sum_m [\mathbf{I} - \mathbf{W}\boldsymbol{\psi}(s_2)]_{im}^{-1} \frac{\mathbf{W}_{mj}[\psi_j(s_1) - \psi_j(s_2)]}{s_2 - s_1} [\mathbf{I} - \mathbf{W}\boldsymbol{\psi}(s_1)]_{jk}^{-1}$$

$$+ \delta_{ij} \frac{\left[\frac{1 - \psi_j(s_1)}{s_1} - \frac{1 - \psi_j(s_2)}{s_2}\right]}{s_2 - s_1} [\mathbf{I} - \mathbf{W}\boldsymbol{\psi}(s_1)]_{jk}^{-1} ,$$

$$(7.3.13)$$

After performing the same limits as Eq. 7.2.7,

$$\mathbf{G}_{ijk}(s_1, s_2) \approx$$

$$+ [s_2 - \mathbf{L}_{FP}\mathbf{D}^\gamma(s_2 + \mathbf{k})^{1-\gamma}]_{ij}^{-1} \frac{(s_2 + k_j)^{\gamma_j} - (s_1 + k_j)^{\gamma_j}}{s_2 - s_1} (s + k_j)^{1-\gamma_j} [s_1 - \mathbf{L}_{FP}\mathbf{D}^\gamma(s_1 + \mathbf{k})^{1-\gamma}]_{jk}^{-1}$$

$$+ \delta_{ij} \frac{(s_1 + k_j)^{\gamma_j - 1} - (s_2 + k_j)^{\gamma_j - 1}}{s_2 - s_1} (s + k_j)^{1-\gamma_j} [s_1 - \mathbf{L}_{FP}\mathbf{D}^\gamma(s_1 + \mathbf{k})^{1-\gamma}]^{-1} \qquad (7.3.14)$$

The $[s - \mathbf{L}_{FP}\mathbf{D}^\gamma(s + \mathbf{k})^{1-\gamma}]$ terms are simply the single time interrupted fractional diffusion Green's functions. The $(s + \mathbf{k})^{1-\gamma}$ term is the interrupted version of the fractional derivative operator. The first s_1 versus s_2 difference term, $\frac{(s_2 + \mathbf{k})^\gamma - (s_1 + \mathbf{k})^\gamma}{s_2 - s_1}$ corresponds to the first jump made during the second time interval. Finally the third term, $\frac{(s_1 + \mathbf{k})^{\gamma - 1} - (s_2 + \mathbf{k})^{\gamma - 1}}{s_2 - s_1}$ corresponds to particles that fail to make a jump during the second time interval. After the first jump, the particle resumes normal interrupted fractional diffusion. We will discuss further details about these two terms after we introduce a memory kernel expression for these terms.

Similar to the derivation of the single time propagator, if $k_j = k$, $D_j^\gamma = D^\gamma$, and $\gamma_j = \gamma$ is independent of the lattice positions, the two-time propagator can be

rewritten as

$$G_2(s_1, s_2) = \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} c_1(\omega_1, \omega_2 | s_1, s_2) \tilde{G}_{FP}(-\omega_2) G_{FP}(-\omega_1)$$

$$+ \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} c_0(\omega_1, \omega_2 | s_1, s_2) \tilde{G}_{FP}(-\omega_2) G_{FP}(-\omega_1) \qquad (7.3.15)$$

where

$$c_1 = \frac{(s_2 + k)^{\gamma-1}}{s_2(s_2 + k)^{\gamma-1} + iD^\gamma \omega_2} \frac{(s_2 + k)^\gamma - (s_1 + k)^\gamma}{s_2 - s_1} \frac{1}{s_1(s_1 + k)^{\gamma-1} + iD^\gamma \omega_1} \qquad (7.3.16)$$

and

$$c_0 = \frac{(s_1 + k)^{\gamma-1} - (s_2 + k)^{\gamma-1}}{s_2 - s_1} \frac{1}{s_1(s_1 + k)^{\gamma-1} + iD^\gamma \omega_1}. \qquad (7.3.17)$$

The expression can be used to measure the response to a step potential applied at t_1, by choosing $\mathbf{G}_{FP} \neq \tilde{\mathbf{G}}_{FP}$. For the aging correlation function applied to the harmonic oscillator below, $\mathbf{G}_{FP} = \tilde{\mathbf{G}}_{FP}$. If we set $k = 0$, we recover two-time anomalous diffusion, which coincides with expressions previously derived for translationally invariant systems and harmonic oscillators [13, 17].

A few comments about Eq. 7.3.15 are necessary [90]. Consider the evolution during the 2nd time period, $(t_1, t_1 + t_2)$ for fixed t_1. The propagator corresponds to a random hopping process, where most hops are fast with the occasional long lived trap. The longer the particle hops, the more likely it will find a long lived cage. The expression for c_0 represents the probability that a particle is still in the cage that it found during the first time interval, which results in c_0 being independent of ω_2, while c_1 corresponds to all processes that made at least one hop during the 2nd time

262

interval. If t_1 is short ($s_1 \to \infty$), the c_0 term goes to zero because the particle did not find a long lived trap. If t_1 is long, but t_2 is short ($s_1 \to 0$, $s_2 \to \infty$) the c_0 term dominates since the particle found a long lived trap and has not hopped out of it during the 2nd time interval, which causes aging. Eventually, the particle will hop out of the trap found during the first time interval and resume its random walk so the c_0 term decays and the c_1 term dominates for $t_2 > k^{-1}$. The truncation of the c_0 term is the major contrast between interrupted and uninterrupted fractional diffusion.

7.4 Application to the Harmonic Oscillator

As a simple application of Eq. 7.3.15, consider fractional diffusion in a 1-D harmonic well with unit diffusion constant, viscosity, and force constant. The resulting Green's function for normal diffusion is

$$G_{FP}(x, x', t) = \tilde{G}_{FP}(x, x', t) = \frac{1}{\sqrt{2\pi(1 - e^{-2t})}} \exp\left[\frac{-(x - x'e^{-t})^2}{2(1 - e^{-2t})}\right] \qquad (7.4.18)$$

For simplicity, the initial condition is $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, and the system is always in macroscopic equilibrium, but not in microscopic equilibrium with respect to the distribution of activation barriers. We are concerned with the correlation function $C(t_2|t_1) = \langle x(t_1 + t_2)x(t_1)\rangle$. Setting $D^\gamma = 1$ (i.e. time and space are unit-less) and integrating over ω_i gives the aging correlation function

$$C(s_2|s_1) = \frac{(s_2+k)^{\gamma-1}}{s_2(s_2+k)^{\gamma-1}+1} \frac{(s_2+k)^\gamma - (s_1+k)^\gamma}{s_2 - s_1} \frac{1}{s_1(s_1+k)^{\gamma-1}}$$

$$+ \frac{(s_1+k)^{\gamma-1}-(s_2+k)^{\gamma-1}}{s_2-s_1} \frac{1}{s_1(s_1+k)^{\gamma-1}} \tag{7.4.19}$$

Setting $k = 0$ recovers the FFPE result [17]. This correlation function is rigorously unity at $t_2 = 0$. In the limit of no aging, $t_1 \to 0$, one recovers an interrupted version of the correlation function,

$$\lim_{s_1 \to \infty} s_1 C(s_2|s_1) = \frac{1}{s_2 + (k+s_2)^{1-\gamma}}. \tag{7.4.20}$$

Another important feature is the existence of a stationary correlation function in the limit as $t_1 \to \infty$,

$$\lim_{s_1 \to 0} s_1 C(s_2|s_1) = \frac{k^{1-\gamma}}{s_2} \left[k^{\gamma-1} - (k+s_2)^{\gamma-1} + \frac{(k+s_2)^{\gamma}\left(k^{\gamma}-(k+s_2)^{\gamma}\right)}{k+s_2+s_2(k+s_2)^{\gamma}} \right] \tag{7.4.21}$$

As a specific example we choose $\gamma = \frac{1}{2}$ and $k = 1/1000$. The solutions for several aging times $t_1 = 0, 1, 10, 100, \infty$ are plotted in Fig. 7-2. If the system has not aged, $t_1 = 0$, the correlation function demonstrates relaxation on many timescales before approaching exponential behavior at the interruption time $k^{-1} = 1000$. If the system is allowed to age for a short time, the correlation function remains approximately equal to unity for a short period of time before the onset of the distributed relaxation. The long time exponential behavior is not delayed so one observes a smaller region of multiple time-scale relaxation. In the long time limit, $t_1 \to \infty$, the multiple timescale relaxation is absent, and only approximate exponential relaxation is observed. The FDT become valid on these longer time scales and the traditional

264
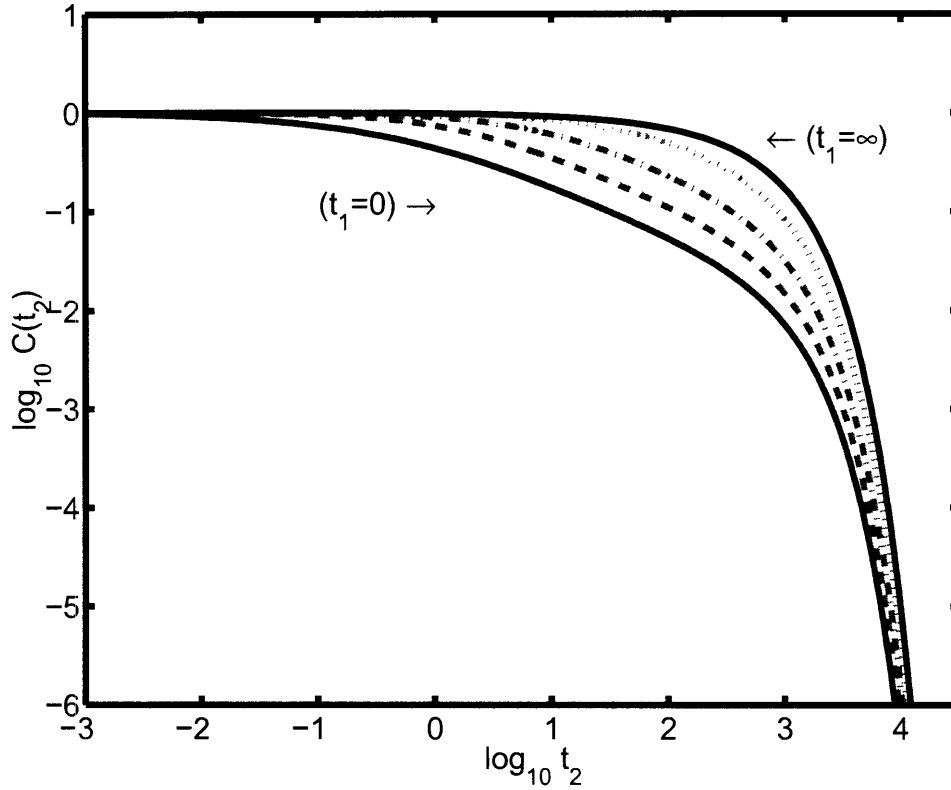
Figure 7-2: The correlation function for the interrupted fractional diffusion with $\gamma = \frac{1}{2}$, $k = 1/1000$ in a harmonic oscillator with induction times (from bottom to top), $t_1 = 0, 1, 10, 100, \infty$. The $t_1 = 0$ solution shows approximate power-law behavior for $1 < t_2 < 1000$, before switching over to exponential behavior in the long time limit. This is in contrast to the nearly exponential behavior (on a log scale) shown by the longer induction times. Aging removes many of the features that one tries to incorporate in the model through a complicated waiting time.

exponential relaxation associated with FDT returns.

7.5 Experimental measurements

The validity of the FDT at low frequencies and related aging effects must be explored

in any experimental system modeled by waiting time processes. If the time of the

trapping process cannot be determined, the aged correlation function in Eq. 7.4.21

should be used. This correlation function is not defined for $k = 0$. The longer the

first time interval, the longer the c_0 term dominates the correlation function and the particle does not move. The result is an essential system arrest for periods of time shorter than the characteristic relaxation time determined by the cutoff k^{-1}. This construction of interrupted fractional diffusion makes it apparent that certain quantities are not affected by aging, such as the ordering of events. If the particle is at position $x = 0$ at time, t_1 and we measure the probability of reaching position $x = 1$ before $x = -1$, this probability will not depend on t_1. These invariant measurements should be examined in experiments where interrupted fractional diffusion appears to fit the correlation function.

Interrupted aging processes are more realistic scenarios for modeling biological systems than the infinite aging of fractional diffusion. Although the stretched exponential approximation to the short time behavior of the non-aged FFPE and a long time power-law have been fit to the correlation functions of several different measurements on single room temperature biomolecules, the systems do not exhibit long time aging processes and appear to achieve equilibrium in a finite amount of time [225]. Similarly, the interruption of aging has been observed in quantum dot blinking statistics and may prove to be the rule rather than the exception [39].

7.6 Coupling first order kinetics to the fractional Fokker-Planck equation

The original applications of these dispersive transport models date back Scher and Montroll's description of the power-law arrival times of electrons in amorphous semiconductors [152, 181]. A more formal derivation of the CTRW in disordered media was provided by Klafter and Silbey [116]. More recently, the number of applications has grown to include the analysis of biological systems, such as CO binding in myoglobin and cellular reaction processes [200, 182], where the dispersive transport mechanism is coupled with first order kinetic processes [127]. In order to understand the applicability of the anomalous diffusion models with first order kinetics to biological experiments one must address issues associated with the coupling of these mechanisms. Several authors previously considered coupling of dispersive transport with a single δ function sink [185, 201, 200]. Here we consider more general scenarios and derive two possible limiting cases.

7.7 Naive approach

If one examines the fractional Fokker-Planck Equation [145], $\partial_t \mathbf{G} = {}_0 D_t^{1-\gamma} \mathbf{L}_{FP} \mathbf{G}$, one can naively write the first order kinetic equation as [201],

$$\partial_t \mathbf{G} = {}_0 \mathbf{D}_t^{1-\gamma} \mathbf{L}_{FP} \mathbf{G} - \mathbf{K} \mathbf{G}. \tag{7.7.22}$$

The Green's function is $\mathbf{G}$ and $\mathbf{K}$ is the first order process. The operator $\mathbf{L}_{FP}$ can be another kinetic operator or a time independent Fokker-Planck operator with potential V and diffusion constant D, $\mathbf{L}_{FP}\mathbf{G} = \nabla[D\nabla\mathbf{G}] + \nabla \cdot [\mathbf{G}\nabla\beta V]$. The fractional differential operator, $_0\mathcal{D}_t^{1-\gamma}$ is given by $_0\mathcal{D}_t^{1-\gamma}f(t) = D^\gamma \frac{1}{\Gamma[\gamma]}\frac{\partial}{\partial t}\int_0^t \frac{f(t')}{(t-t')^{1-\gamma}}$ with the constant D^γ having units of $[time]^{1-\gamma}$. In the Laplace domain $s \to t$, we can rewrite Eq. 7.7.22, $\mathbf{G}_0^{-1}\mathbf{G} = -\mathbf{K}\mathbf{G} + \mathbf{I}$. with $\mathbf{G}_0^{-1} = s - D^\gamma s^{1-\gamma}\mathbf{L}_{FP}$ and $\mathbf{I}$ being the identity operator. Iterating this solution gives

$$\mathbf{G} = \mathbf{G}_0\left[\mathbf{I} + \mathbf{K}\mathbf{G}_0\right]^{-1}, \tag{7.7.23}$$

which is equal to $\mathbf{G} = \left[\mathbf{I}s - s^{1-\gamma}D^\gamma\mathbf{L}_{FP} + \mathbf{K}\right]^{-1}$.

Suppose that $\mathbf{K}$ is a constant diagonal matrix $\mathbf{K} = \kappa\mathbf{I}$ and $\gamma \neq 1$ (i.e. depletion occurs at every point). The solution should be $\mathbf{G}_0(t)e^{-\kappa t}$, which becomes in the Laplace domain $\mathbf{G}_0(s + \kappa)$, but this result is not the solution of the Eq. 7.7.22, unless $\mathbf{L}_{FP} = 0$. The equation ignores the aging of the propagator each time the first order kinetic operator is applied. Perturbatively expanding the solution in Eq. 7.7.23 for a reactive sink in terms of the free fractional propagator gives

$$\mathbf{G} \approx \mathbf{G}_0(t, 0) - \int dt'\mathbf{G}_0(t, t')\kappa\mathbf{G}_0(t', 0) + \dots . \tag{7.7.24}$$

The first term corresponds to density that does not react, while the second term corresponds to a single reaction at some time defined by $\kappa\mathbf{G}_0(t, 0)$ that creates a negative density intended to cancel some of the density in $\mathbf{G}_0(t, 0)$ by being super-

imposed upon it. The fractional diffusion propagator is not a semi-group, $\mathbf{G}_0(t, 0) \neq \mathbf{G}_0(t, t')\mathbf{G}_0(t', 0)$, so the motion of the negative density at later times is different than the motion of the original density. As a result, the negative density does not remain super-imposed upon the original density, leading to the unphysical solution discussed above.

In order to construct a model for coupling first order and fractional processes, one must return to the original construction of semi-Markov processes and take appropriate limits. The construction will depend on the physical scenario considered. We discuss two possible scenarios below.

7.8   Passive first order kinetics

In the case of passive first order kinetics, the first order reaction process does not influence the anomalous process. An example is the fluctuating diffusion constant model [34]. The first order process is a diffusion process with time dependent diffusion constant, $G_{kin}(k, t) = e^{-\int \mathbf{K}(t)} = e^{\int D(t)\nabla^2}$, where kin stands for the kinetic operator $\mathbf{K}(t)$ which is $-\int D(t)\nabla^2$ for the fluctuating diffusion model. The diffusion constant changes between a set of discrete values, $D^{(j)}$, through a stochastic jumping process (semi-Markov) process [79]. In a semi-Markov process, the probability of making a transition from a state j to state i at a time t is given by the matrix element $Q_{ij}(t)dt$. The functions $Q_{ij}(t)$ must satisfy $Q_{ij}(t) \geq 0$ for $t > 0$ and $\sum_i \int dt Q_{ij}(t) \leq 1$ for all j. If $\sum_i \int dt Q_{ij}(t) < 1$ the particle may stay at site i permanently, i.e. i is a trap. For the rest of the derivation, we assume the sum is unity. The one time propagator

for the process governing the diffusion constant can be simply written in the Laplace domain $t \to s$ as [79]

$$G_{ij}^{(D)}(s) = \frac{1 - \sum_k Q_{ki}(s)}{s} [\mathbf{I} - \mathbf{Q}(s)]_{ij}^{-1} . \qquad (7.8.25)$$

In this expression $\mathbf{I}$ is the identity matrix and $\mathbf{Q}$ is the complete matrix of $Q_{ij}(s)$. The $[\mathbf{I} - \mathbf{Q}(s)]_{ij}^{-1}$ is the Laplace transform of the probability of jumping into the state i at time t given that we started in state j and $\frac{1 - \sum_k Q_{ki}(s)}{s}$ is the Laplace transform of the probability of not jumping out of the final state, i [79].

If the waiting time of the transitions is site specific but does not depend on the direction of the transition, $\mathbf{Q} = \mathbf{W}\boldsymbol{\psi}(t)$, where W_{ij} is the probability of the next transition being from site j to site i ($W_{ij} \geq 0$ and $\sum_i W_{ij} = 1$) and $\psi_{ij}(t) = \delta_{ij}\psi_j(t)$ is the waiting time for the transition out of site j ($\psi_j(t) \geq 1$, $\int dt \psi_j = 1$). The Green's function for the diffusion constant must be tensored with the Green's function for the kinetic process and the scalars $\psi_j(s)$ become matrices, $\mathbf{Q}_{ij} = \mathbf{W}_{ij}\psi_j(s + \mathbf{K}^{(j)})$, where $\mathbf{W}_{ij} = \mathbf{I}_{kin}W_{ij}$ and

$$\psi_j(s + \mathbf{K}^{(j)}) = \mathbf{S}^{(j)}\psi_j(s + \mathbf{\Lambda}^{(j)})(\mathbf{S}^{(j)})^{-1} \qquad (7.8.26)$$

with $\mathbf{S}^{(j)}$ diagonalizing the kinetic operator $\mathbf{K}^{(j)}$ with eigenvalues $\mathbf{\Lambda}^{(j)}$ and $\psi_j(s + \mathbf{\Lambda}^{(j)}) = \delta_{mn}\psi_j(s + \Lambda_m^{(j)})$ is a diagonal matrix. In the case of the diffusion process, $-\mathbf{K}^{(j)} = D^{(j)}\nabla^2$ and the eigenfunctions are the Fourier components, $x \to k$, with eigenvalues $D^{(j)}k^2$. The identity operator $\mathbf{I}_{kin}$ denotes a sub-matrix δ_{ij} with the

dimensions of the kinetic operator $\mathbf{K}^{(j)}$. One can allow the complex waiting time transitions to influence kinetic process by replacing $\mathbf{I}_{kin}$ with another Markov matrix. Similarly, the expression $\frac{1-\sum_k Q_{ki}(s)}{s}$ becomes $\frac{\mathbf{I}_{kin}-\psi_j(s+\mathbf{K}^{(j)})}{s+\mathbf{K}^{(j)}}$ with $\frac{1}{s+\mathbf{K}^{(j)}}$ representing a matrix inverse.

Often, the semi-Markov transitions have the form $W \approx e^{\mathbf{L}_{FP}\Delta t}$, where Δt is a small time that characterizes the displacement and $\mathbf{L}_{FP}$ can be a kinetic matrix or a Fokker-Planck operator as discussed above. Using the definitions above, Eq. 7.8.25 becomes

$$\mathbf{G}_{ij}(s) = \frac{\mathbf{I}_{kin} - \psi_i(s + \mathbf{K}^{(i)})}{s + \mathbf{K}^{(i)}} \left[\mathbf{I} - e^{\mathbf{L}_{FP}\Delta t}\psi(s + \mathbf{K})\right]^{-1}_{ij} \qquad (7.8.27)$$

In this expression, $\psi(s + \mathbf{K})$ is a block diagonal matrix,

$$\psi(s + \mathbf{K}) = \begin{bmatrix} \psi_j(s + \mathbf{K}^{(j)}) & \mathbf{0} \\ \mathbf{0} & \ddots \end{bmatrix}. \qquad (7.8.28)$$

and $\mathbf{L}_{FP}$ has the form

$$\mathbf{L}_{FP} = \begin{bmatrix} (L_{FP})_{11}\mathbf{I}_{kin} & \cdots & (L_{FP})_{1N}\mathbf{I}_{kin} \\ \vdots & \ddots & \vdots \\ (L_{FP})_{N1}\mathbf{I}_{kin} & \cdots & (L_{FP})_{NN}\mathbf{I}_{kin} \end{bmatrix}. \qquad (7.8.29)$$

The waiting time distribution is assumed to be a Lévy distribution $\psi_j(s) = e^{-\tau_j^\gamma s^{\gamma_j}}$, but the limiting results will be independent of the exact functional form. The limit

271

as $\tau_j^\gamma, \Delta t \ll \frac{1}{s+\lambda_j}$ with $\frac{\Delta t}{\tau_j^\gamma} \to const = D_j^\gamma$, gives

$$\mathbf{G}_{ij}(s) \approx \left[s + \mathbf{K} - \mathbf{L}_{FP}\mathbf{D}^\gamma \left[s + \mathbf{K}\right]^{1-\gamma} \right]^{-1}. \tag{7.8.30}$$

where

$$\mathbf{D}^\gamma = \begin{bmatrix} \mathbf{I}_{kin}D_j^\gamma & \mathbf{0} \\ \mathbf{0} & \ddots \end{bmatrix}, \quad \text{and} \quad \left[s+\mathbf{K}\right]^{1-\gamma} = \begin{bmatrix} \left[s+\mathbf{K}^{(j)}\right]^{1-\gamma_j} & \mathbf{0} \\ \mathbf{0} & \ddots \end{bmatrix}. \tag{7.8.31}$$

An important feature is the matrix ordering with $\mathbf{L}_{FP}$ operating on a matrix containing the first order reaction terms. Setting $D_j^\gamma = D^\gamma$, $\gamma_j = \gamma$, and $\mathbf{K}^{(j)} = 0$ recovers anomalous diffusion [145]. Similarly for $\gamma_j = 1$ and $D_j^\gamma = 1$, normal diffusion with reactive sinks is recovered. The intuitively expected fractional diffusion times an exponential decay, $\mathbf{G} = \mathbf{G}_0(t)e^{-\kappa t}$, results when $\mathbf{K}^{(j)} = \kappa \mathbf{I}_{kin}$, $\gamma_j = \gamma$, and $D_j^\gamma = D^\gamma$ for all j. .

For the fluctuating diffusion constant model with two states, $1, 2$, $(L_{FP})_{ij} = -2a\delta_{ij} + a$, $D_j^\gamma = 1$, and $\gamma_j = \gamma$, the Green's function in Fourier space becomes

$$\mathbf{G}_{ij}(k,s) = \left[\begin{bmatrix} s + D^{(1)}k^2 & 0 \\ 0 & s + D^{(2)}k^2 \end{bmatrix} \right.$$
$$\left. + \begin{bmatrix} a & -a \\ -a & a \end{bmatrix} \begin{bmatrix} (s+D^{(1)}k^2)^{1-\gamma} & 0 \\ 0 & (s+D^{(2)}k^2)^{1-\gamma} \end{bmatrix} \right]^{-1} \tag{7.8.32}$$

Note that we need to specify both the semi-Markov process, ij, and the kinetic

process, k. If the initial condition for the semi-Markov process is $\rho_1 = \rho_2 = \frac{1}{2}$ and $\delta(x)$ for the kinetic process, summing over the final state, $\sum_{i=1,2} \mathbf{G}(k, s)\rho$, allows us to calculate the variance, $\langle x^2(t) \rangle = (D^{(1)} + D^{(2)})t$, and the asymptotics of the non-Gaussian indicator $\frac{\langle x^4(t) \rangle - 3\langle x^2(t) \rangle^2}{\langle x^2(t) \rangle^2} \sim \frac{3(D^{(1)} - D^{(2)})^2(1-\gamma)}{(D^{(1)} + D^{(2)})^2)}$. In the long time limit, the system displays a non-zero non-Gaussian indicator for $\gamma \neq 1$ because the system is non-self averaging.

Another interesting case has $\mathbf{K} = v\nabla_x$ and $\mathbf{L}_{FP} = \nabla^2$ with $D_j^\gamma = 1$ (unit-less time and space) and $\gamma_j = \gamma$. This kinetic scheme corresponds to fractional diffusion in a glassy sample that we throw across the room. In Fourier space we get

$$\mathbf{G}(k, s) = \frac{1}{s - ivk - k^2(s - ivk)^{1-\gamma}} \tag{7.8.33}$$

As expected, the mean displacement is ballistic, $\langle x(t) \rangle = vt$, but the variance is anomalous, $\langle \delta x(t)^2 \rangle \propto \frac{1}{\Gamma[1+\gamma]} t^\gamma$. Throwing the sample across the room does not influence the anomalous diffusion occurring inside the sample. Similar behavior can be derived for fractional diffusion in systems with macroscopic flows that do not break up the traps, but convect the tracer particle and the trap together [146].

7.9 Interrupted Case

The reaction scheme can also correspond to the first order process competing with the more complex waiting time process, which is the construction applied the most extensively to models of CO myoglobin binding experiments [200]. For simplicity we

again assume that the semi-Markov process has site specific waiting time distribution, $\mathbf{Q} = \mathbf{W}\boldsymbol{\psi}(s)$ with $\sum_i W_{ij} = 1$ and $\int dt \psi_j(t) = 1$. We have a competing first order kinetic process, $P_{ij} k_j e^{-k_j t}$, with $P_{ij} \geq 0$ and $\sum_i P_{ij} \leq 1$. The semi-Markov matrix $\mathbf{Q}$ is modified to account for the competition between the kinetic process and the anomalous diffusion process $Q_{ij}(t) = P_{ij} k_j e^{-k_j t} \int_t^\infty \psi(t) + W_{ij} \psi_j(t) e^{-k_j t}$. If $P_{ij} = W_{ij}$, this expression is identical to an expression derived in the study of interrupted aging [229].

The Laplace transform of this expression is given by

$$Q_{ij} = P_{ij} k_j \frac{1 - \psi_j(s + k_j)}{s + k_j} + W_{ij} \psi_j(s + k_j). \tag{7.9.34}$$

and the probability of not jumping becomes $\frac{(1 - \psi_j(s + k_j))}{s + k_j}$. Using these substitutions in Eq. 7.8.25 gives the Green's function

$$\mathbf{G}_{ij}(s) = \left[\frac{1 - \psi_i(s + k_i)}{1s + k_i} \right] \left[\mathbf{I} - \mathbf{P} \left[\begin{matrix} k_j \frac{1 - \psi_j(s+k_j)}{s+k_j} & 0 \\ 0 & \ddots \end{matrix} \right] - \mathbf{W} \left[\begin{matrix} \psi_j(s + k_j) & 0 \\ 0 & \ddots \end{matrix} \right] \right]^{-1}_{ij}. \tag{7.9.35}$$

Substituting $\mathbf{W} = e^{\mathbf{L}_{FP}\Delta t}$ and $\psi_j(s) = e^{-\tau_j s^{\gamma_j}}$ and taking the same limiting procedure as above gives the Green's function,

$$\mathbf{G} \approx \left[s + \tilde{\mathbf{K}} - \mathbf{L}_{FP} \mathbf{D}^\gamma \left[\begin{matrix} (s + k_j)^{1 - \gamma_j} & 0 \\ 0 & \ddots \end{matrix} \right] \right]^{-1} \tag{7.9.36}$$

274

with

$$\tilde{\mathbf{K}} = [\mathbf{I} - \mathbf{P}] \begin{bmatrix} k_j & 0 \\ & \\ 0 & \ddots \end{bmatrix} \tag{7.9.37}$$

The $\tilde{\mathbf{K}}$ is equal to the first order kinetic matrix for normal diffusion. If $P_{ii} = 1$, $\tilde{\mathbf{K}} = 0$ and the interruption process results in restarting the system without depletion, so the equilibrium distribution is the same as the normal diffusion equilibrium distribution [229]. Note that $\tilde{\mathbf{K}}$ may be zero, but $\mathbf{L}_{FP}$ still acts on the kinetic matrix. Similar to the passive diffusion process, in the limits that $k_j = 0$, we recover fractional diffusion, and in the limit that $\gamma_j \to 1$ we recover normal diffusion. If the first order kinetics corresponds to a simple sink, $\tilde{\mathbf{K}} = \delta_{ij}k_j$, the interrupted and passive diffusion processes are identical and correspond to anomalous transport with site specific depletion rates, which is explored below.

7.10   Survival probability with a sink

An extensively studied problem is the survival probability of a CTRW in the presence of reactive sinks [185, 201]. The expression we derive does not depend on the reactive process being restricted to a small boundary and is intuitive without resorting to any limiting arguments. To simplify expressions, we assume $\gamma_j = \gamma$, $D_j^\gamma = D^\gamma$. A subset of the sites, j, have reactive sinks, which remove particles with a rate given by k_j. The other sites have no reaction, $k_{i \neq j} = 0$. So the Green's function has the form

$$
\mathbf{G} = \left[s + \begin{bmatrix} \mathbf{k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - D^\gamma \mathbf{L}_{FP} \begin{bmatrix} (\mathbf{I}s + \mathbf{k})^{1-\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}s^{1-\gamma} \end{bmatrix} \right]^{-1}
\tag{7.10.38}
$$

which can be rewritten as

$$
\mathbf{G} = \left[\mathbf{G}_0^{-1} + \begin{bmatrix} \mathbf{k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - D^\gamma \mathbf{L}_{FP} \begin{bmatrix} (s + \mathbf{k})^{1-\gamma} - s^{1-\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right]^{-1}
\tag{7.10.39}
$$

where $\mathbf{G}_0^{-1}$ is the inverse of the free fractional diffusion propagators in the absence of a sink. The survival probability is $S = \sum \mathbf{G}$. If $\mathbf{G}_0$ preserves probability, $\sum \mathbf{G}_0 = s^{-1}$ the survival probability can be written as

$$
S = \frac{1}{s} \sum \left[\mathbf{I} + \left(\mathbf{K} - D^\gamma \mathbf{L}_{FP} \begin{bmatrix} (\mathbf{I}s + \mathbf{k})^{1-\gamma} - \mathbf{I}s^{1-\gamma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) G_0 \right]^{-1}.
\tag{7.10.40}
$$

which can be simplified to

$$
S = \frac{1}{s} \left[1 - \sum \mathbf{k} \left[\mathbf{I} + (G_0)_{jj}\mathbf{k} + \left[s^{\gamma-1} - s^\gamma (G_0)_{jj} \right] \left[(s + \mathbf{k})^{1-\gamma} - s^{1-\gamma} \right] \right]^{-1} (G_0)_{ji} \right]
$$
$$
\tag{7.10.41}
$$

since $-D^\gamma \mathbf{L}_{FP} = s^{\gamma-1} \left[G_0^{-1} - s \right]$. By rewriting the solution in this form, we remove the $\mathbf{L}_{FP} \mathbf{D}^\gamma \left[s + \mathbf{K} \right]^{1-\gamma}$ term and can take a continuum approximation. For the simple diffusion limited recombination problem, $\mathbf{L}_{FP} = \nabla^2$, where we set $D = D^\gamma = 1$ (unitless space and time) and all the reactive sites have a rate of k and appear on a sphere

276

of radius σ. Symmetry allows us to write the survival probability as

$$S = \frac{1}{s} \left[1 - \frac{kG_0(\sigma, r_0, s)}{1 + kG_0(\sigma, \sigma, s) + [s^{\gamma-1} - s^\gamma G_0(\sigma, \sigma)]\left[(s+k)^{1-\gamma} - s^{1-\gamma}\right]} \right] \qquad (7.10.42)$$

with r_0 denoting the initial position. In three dimensions the Green's function is

$G_0(\sigma, r_0, s) = \frac{s^{\gamma-1}\sigma^2}{r_0(1+s^{\gamma/2})}e^{-(r_0-\sigma)s^{\gamma/2}}$ where angular integration has been performed [114]

(remember that s, k, σ, and r_0 are unit-less). The ultimate survival probability can

be calculated by taking the limit as $s \to 0$ of sS, which gives

$$\lim_{s \to 0} sS = 1 - \frac{1}{1 + k^{-\gamma}\sigma^{-1}} \frac{\sigma}{r_0} \qquad (7.10.43)$$

Note that the survival probability depends on $k^{-\gamma}$ and coincides with the expression

for $\gamma = 1$ [175]. This expression is much simpler than those presented previously [185,

201]. For $k > 1$ the survival probability decreases with increasing γ, and for $k < 1$ the

survival probability decreases with decreasing γ. For $\gamma = 0$, the survival probability

equals $\frac{1}{2}$. For $k = 1$, the survival probability is independent of γ, $S(t = \infty) =$

$1 - \frac{1}{1+\sigma^{-1}}$. The inverse σ dependence results from the apparent centrifugal force,

$F \propto 1/r$ that pushes the particle away from the reactive sites.

The result is slightly counter intuitive because we expect a smaller γ to correspond

to a slower process and to be more likely to react for all values of k. Smaller γ results

in a slower process in a mean waiting time sense, but the mean waiting time is not the

important contribution to the survival probability. The important contribution is the

probability of the anomalous translational process occurring before the rate process.

This probability is $e^{-\tau^\gamma k^\gamma}$ for the Lèvy distribution. For k large and $\tau = 1$ (time unit-less), a smaller γ increases this term and implies a greater survival probability during the encounter with the reactive sink. In fact, the fatter the power-law tail, $\gamma \ll 1$, the more likely there will be an extremely fast process, with only the occasional slow process to destroy the mean.

7.11 Conclusion

Several scenarios for the coupling of first order processes to anomalous transport are explored by constructing possible kinetic schemes with a limiting procedure performed on an appropriate semi-Markov process. The semi-Markov process is based on the trapping picture that has been explored in simulations of glassy systems. The limits are taken under the assumption that the kinetic process is a slow process, so that the non-universal features at the short time portion of the waiting time distributions can be ignored. The limiting procedure simplifies the mathematics and functional forms of the solutions into a form that allows one to explore the influences of topology and geometry without considering these non-universal features. Since an important aspect of proteins and cells is the geometry of internal cavities, this simplification is advantageous.

The first order processes? effects on equilibrium behavior, aging, and probability of reaction can be addressed within the same framework. The first order kinetic processes cannot be separated from the anomalous diffusion process, resulting in a term of the form $\mathbf{L}_{FP}\mathbf{D}^\gamma (s + \mathbf{K})^{1-\gamma}$, which can modify the equilibrium distribution

of the semi-Markov process. The operator form $\mathbf{L}_{FP}\mathbf{D}^{\gamma}(s+\mathbf{K})^{1-\gamma}$ also results in the aging of the semi-Markov system results in a modified correlation function. Most of the timescales that are present in the waiting time distribution are not apparent in the aged waiting time distribution. The $\mathbf{L}_{FP}\mathbf{D}^{\gamma}(s+\mathbf{K})^{1-\gamma}$ operator may also result in a non-trivial spatial limit if $\mathbf{K}$ becomes proportional to a spatial δ function. This spatial limit can be avoided in calculating the survival probability for a process with arbitrary sinks, and an explicit expression for an irreversible diffusion controlled reaction in three dimensions is derived. For the diffusion controlled reaction, the ultimate survival probability yields a simple expression that depends on $k^{-\gamma}$ and shows that slow diffusion can improve the survival probability if k is large. This result is not surprising because the probability of the power-law process being faster than an exponential process is different than the mean waiting time being a fast process.

Bibliography

[1] N. Agmon and J. J. Hopfield. Co binding to heme-proteins - a model for barrier height distributions and slow conformational-changes. *J. Chem. Phys.*, 79:2042, 1983.

[2] P. Allegrini, M. Buiatti, P. Grigolini, and B. J. West. Non-gaussian statistics of anomalous diffusion: The dna sequences of prokaryotes. *Phys. Rev. E*, 58:3640, 1998.

[3] H. C. Andersen. Functional and graphical methods for classical statistical dynamics. i. a formulation of the martin-siggia-rose method. *J. Math. Phys.*, 41:1979, 2000.

[4] M. Andrec, R. M. Levy, and D. S. Talaga. Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden markov models. *J. Phys. Chem. A*, 107:7454, 2003.

[5] N. C. Andrews and A. J. McHugh. Conformational and rheological dynamics of semiflexible macromolecules undergoing shear flow: A nonequilibrium brownian dynamics study. *J. Reol.*, 42:281, 1998.

[6] S. Assoudou and B. Essebbar. A bayesian model for markov chains via jeffrey's prior. *Commun. Stat-Theor M*, 32:2163, 2003.

[7] A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.*, 80:505, 2001.

[8] I. Auzanneau, C. Barreau, and L. Salome. Imaging by fluorescence videomicroscopy of individual single-stranded dna molecules in solution. *Comptes Rendus De L Academi Des Sciences Serie III- Sciences De La Vie-Life Sciences*, 316:459, 1993.

[9] H. P. Babcock, D. E. Smith, J. S. Hur, E. S. G. Shaqfeh, and S. Chu. Relating the microscopic and macroscopic response of a polymeric fluid in a shearing flow. *Phys. Rev. Lett.*, 85:2018, 2000.

[10] I. Bahar, A. Wallquist, D. G. Covell, and R. L. Jernigan. Correlation between native-state hydrogen exchange and cooperative residue fluctuations from a simple model. *Biochemsitry*, 37:1067, 1998.

[11] C. L. Bai, C. Wang, X. S. Xie, and P. G. Wolynes. Single molecule physics and chemistry. *Proc. Natl Acad. Sci*, 96:1175, 1999.

[12] F. G. Ball and J. A. Rice. Stochastic models for ion channels: introduction and bibliography. *Mathematical Biosciences*, 112:189, 1992. and references within.

[13] E. Barkai and Y. C. Cheng. Aging continuous time random walks. *J. Chem. Phys.*, 118:6167, 2003.

[14] E. Barkai, Y. Jung, and R. J. Silbey. Time-dependent fluctuations in single molecule spectroscopy: A generalized wiener-khintchine approach. *Phys. Rev. Lett.*, 87:207403, 2001.

[15] E. Barkai, R. J. Silbey, and G. Zumofen. Levy distribution of single molecule line shape cumulants in glasses. *Phys. Rev. Lett.*, 84:5339, 2000.

[16] V. Barsegov, V. Chernyak, and S. Mukamel. Multitime correlation functions for single molecule kinetics with fluctuating bottlenecks. *J. Chem. Phys.*, 116:4240, 2002.

[17] V. Barsegov and S. Mukamel. Multipoint fluorescence quenching-time statistics for single molecules with anomalous diffusion. *J. Phys. Chem. A*, 108:15, 2004.

[18] M. Barthelemy and H. Orland. Replica field theory for composite media. *Physica A*, 207:106, 1994.

[19] T. Basche, W. E. Moerner, and M. Orrit. Photon antibunching in fluorescence of a single dye molecule trapped in a solid. *Phys. Rev. Lett.*, 69:1516, 1992.

[20] A. M. Berezhkovskii, A. Szabo, and G. H. Weiss. Theory of single-moleucle fluorescence spectroscopy of two state systems. *J. Chem. Phys.*, 110:9145, 1999.

[21] A. M. Berezhkovskii, A. Szabo, and G. H. Weiss. Theory of the fluorescence of single molecules undergoing multistate conformational dynamics. *J. Phys. Chem. B*, 104:3776, 2000.

[22] A. M. Boiron, P. Tamarat, B. Lounis, R. Brown, and M. Orrit. Are the spectral trails of single molecules consistent with the standard two-level system of glasses at low temperatures. *Chem. Phys.*, 247:119, 1999.

[23] G. Bokinsky, D. Rueda, V. K. Misra, A. Gordus, M. M. Rhodes, H. P. Babcock, N. G. Walter, and X. Zhuang. Single-molecule transition-state analysis of rna folding. *Proc. Natl. Acad. Sci. USA*, 100:9302, 2003.

[24] J. P. Bouchaud, A. Comtet, A. Georges, and P. Ledoussal. Anomalous diffusion in random media of any dimensionality. *J. Physique*, 48:1445, 1987.

[25] J. P. Bouchaud and A. Georges. Anomalous diffusion in disordered media - statistical mechanisms, models and physical applications. *Phys. Rep.*, 195:127, 1990.

[26] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette. Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal*, 76:409, 1999.

[27] D. D. Brace, S. D. Gottke, H. Cang, and M. D. Fayer. Orientational dynamics of the glass forming liquid, dibutylphthalate: time domain experiments and comparison to mode coupling theory. *J. Chem. Phys.*, 116:1598, 2002.

[28] F. Brochard-Wyart. Polymer-chains under strong flows–stems and flowers. *Europhysics Letters*, 30:387, 1995.

[29] F. L. H. Brown. Single molecule kinetics with time dependent rates: a generating function approach. *Phys. Rev. Lett.*, 90:028302, 2003.

[30] W. J. Bruno, J. Yang, and J. E. Pearson. Using independent open-to-closed transitions to simplify aggregates markov models of ion channel gating kinetics. *Proc. Natl. Acad. Sci. USA*, 102:6326, 2005.

[31] J. D. Bryngelson and D. Thirumalai. Internal constraints induce localization in an isolated polymer molecule. *Phys. Rev. Lett.*, 76:542, 1996.

[32] S. Buchner and A. Heuer. Potential energy landscape of a model glass former: Thermodynamics, anharmonicities, and finite size effects. *Phys. Rev. E*, 60:6507, 1999.

[33] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith. Entropic elasticity of λ-phage dna. *Science*, 265:1599, 1994.

[34] J. Cao. Single molecule tracking of heterogeneous diffusion. *Phys. Rev. E*, 63:041101, 2001.

[35] J. S. Cao. Event averaged measurment of single molecule kinetics. *Chem. Phys. Lett.*, 327:38, 2000.

[36] A. K. Chakraborty, D. Bratko, and D. Chandler. Diffusion of ionic penetrants in charged disordered media. *J. Chem. Phys.*, 100:1528, 1994.

[37] L. G. Chen, M. Falcioni, and M. W. Deem. Disorder-induced time-dependent diffusion in zeolites. *J. Phys. Chem. B*, 104:6033, 2000.

[38] V. Chernyak, M. Schulz, and S. Mukamel. Stochastic-trajectories and nonpoisson kinetics in single-molecule spectroscopy. *J. Chem. Phys.*, 111:7416, 1999.

[39] I. H. Chung and M. G. Bawendi. Relationship between single quantum-dot intermittency and fluorescence intensity decays from collections of dots. *Phys. Rev. B.*, 70:165304, 2004.

[40] I. H. Chung, J. B. Witkoskie, J. S. Cao, and M. G. Bawendi. Description of the fluorescence intensity time trace of collections of cdse nanocrystal qds based on single qd fluorescence blinking statistics. *in preparation*, 2005.

[41] A. A. Clifford and S. E. Coleby. Diffusion of a solute in dilute solution in a supercritical fluid. *Procedings: Mathematical and Physical Sciences*, 433:63, 1991.

[42] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J. L. Viovy, D. Chatenay, and F. Caron. Dna: An extensible molecule. *Science*, 271:792, 1996.

[43] J. C. Crocker, M. T. Valentine, E. R. Weeks, T. Gisler, P. D. Kaplan, and A. G. Yodh D. A. Weitz. Two-point microrheology of inhomogeneous soft materials. *Phys. Rev. Lett.*, 85:888, 2000.

[44] B. X. Cui, B. H. Lin, and S. A. Rice. Dynamical heterogeneity in a dense quasi-two-dimensional colloidal liquid. *J. Chem. Phys.*, 114:9142, 2001.

[45] A. Cunsolo, G. Pratesi, R. Verbeni, D. Colognesi, C. Masciovecchio, G. Monaco, G. Ruocco, and F. Sette. Microscopic relaxation in supercritical and liquid neon. *J. Chem. Phys*, 114:2259, 2001.

[46] J. M. Nunes Da Silva and E. J. S. Jage. Glauber dynamics in the ising chain with random transition rates. *Phys. Lett. A*, 135:17, 1989.

[47] G. D'Agostini. Bayesian inference in processing experimental data: principles and basic applications. *Rep. Prog. Phys.*, 66:1383, 2003.

[48] C. De Dominicis. Dynamics as a substitute for replicas in systems with quenched random impurities. *Phys. Rev. B*, 18:4913, 1978.

[49] C. De Dominicis and L. Peliti. Field-theory renormalization and critical dynamics above tc: Helium, antiferromagnets, and liquid-gas systems. *Phys. Rev. B*, 18:353, 1978.

[50] P. G. De Gennes. Coil-stretch transition of dilute flexible polymers under untrahigh velocity gradients. *J. Chem. Phys.*, 60:5030, 1974.

[51] D. S. Dean, I. T. Drummond, and R. R. Horgan. Perturbation schemes for flow in random media. *J. Phys. A*, 27:5135, 1994.

[52] D. S. Dean, I. T. Drummond, and R. R. Horgan. Effective diffusivity in nonisotropic gradient flows. *J. Phys. A*, 28:6013, 1995.

[53] P. G. Debenedetti and F. H. Stillinger. Supercooled liquids and the glass transition. *Nature*, 410:259, 2001.

[54] M. W. Deem. Field-theoretic approximations for normal diffusion in random velocity fields. *Phys. Rev. E*, 51:4319, 1995.

[55] M. W. Deem and D. Chandler. Classical diffusion in strong random media. *J. Stat. Phys.*, 76:911, 1994.

[56] U. Deker. Perturbation theory for classical random processes with arbitrary perparation. *Phys. Rev. A*, 19:846, 1979.

[57] U. Deker and F. Haake. Fluctuation-dissipation theorems for classical processes. *Phys. Rev. A*, 11:2043, 1975.

[58] U. Deker and F. Haake. Mass and charge renormalization for classical processes with a fluctuation-dissipation theorem. *Phys. Rev. A*, 12:1629, 1975.

[59] A DeMasi, P.A. Ferrari, S. Goldstein, and W. D. Wick. An invariance principle for reversible markov processes: applications to random motions in random environments. *J. Stat. Phys.*, 55:787, 1989.

[60] L. A. Deschenes and D. A. Vanden Bout. Single-molecule studies of heterogeneous dynamics in polymer melts near the glass transition. *Science*, 292:255, 2001.

[61] R. M. Dickson, D. J. Norris, Y. L. Tzeng, and W. E. Moerner. Three-dimensional imaging of single molecules solvated in pores of poly(acrylamide) gels. *Science*, 274:966, 1996.

[62] M. Doi and S. F. Edwards. *Theory of Polymer Dynamics*. Cambridge University Press, Cambridge, 1986.

[63] B. Doliwa and A. Heuer. Cage effect. local anisotropies, and dynamic heterogeneities at the glass trasition: a computer study of hard spheres. *Phys. Rev. Lett.*, 88:4915, 1998.

[64] B. Doliwa and A. Heuer. Hopping in a supercooled lennard-jones liquid: Metabasins, waiting time distribution, and diffusion. *Phys. Rev. E*, 67:031506, 2003.

[65] C. Donati, S. C. Glotzer, and P. H. Poole. Growing spatial correlations of particle displacements in a simulated liquid on cooling toward the glass transition. *Phys. Rev. Lett.*, 82:627, 1999.

[66] C. Donati, S. C. Glotzer, P. H. Poole, W. Kob, and S. J. Plimpton. Spatial correlations of mobility and immobility in a glass-forming lennard-jones liquid. *Phys. Rev. E*, 60:3107, 1999.

[67] S. Doniach and P. Eastman. Protein dynamics simulations from nanoseconds to microseconds. *Curr. Opin. Struct. Biol.*, 9:157, 1999.

[68] P. Doruker, A. R. Atligan, and I. Bahar. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to alpha-amylase inhibitor. *Proteins: Struct. Func. and Gene.*, 40:512, 2000.

[69] P. S. Doyle, Benoit Ladoux, and J. L. Viovy. Dynamics of a tethered polymer in shear flow. *Phys. Rev. Lett.*, 84:4769, 2000.

[70] P. S. Doyle, E. S. G. Shaqfeh, and A. P. Gast. Dynamics simulation of freely draining flexible polymers in steady linear flows. *J. Fluid Mech.*, 334:251, 1997.

[71] P. S. Doyle, E. S. G. Shaqfeh, G. H. McKinley, and S. H. Spiegelberg. Relaxation of dilute polymer solutions following extensional flow. *J. Non-Newtonian Fluid Mech.*, 76:79, 1998.

[72] A. N. Drozdov and S. C. Tucker. Self-diffusion near the liquid-vapor critical point. *J. Chem. Phys*, 114:4912, 2001.

[73] I. T. Drummond, S. Duane, and R. R. Horgan. Scalar diffusion in simulated helical turbulence with molecular diffusivity. *J. Fluid. Mech.*, 128:75, 1984.

[74] M. D. Ediger. Spatially heterogeneous dynamics in supercooled liquids. *Annu. Rev. Phys. Chem.*, 51:99, 2000.

[75] L. Edman, U. Mets, and R. Rigler. Conformational transitions monitored for single molecules in solution. *Proc. Natl. Acad. Sci.*, 93:6710, 1996.

[76] C. Eggeling, J. R. Fries, L. Brand, R. Gunther, and C. A. M. Seidel. Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proc. Natl. Acad. Sci.*, 95:1556, 1998.

[77] E. L. Elson. Fluorescence correlation spectroscopy measures molecular transport in cells. *Traffic*, 2:789, 2001.

[78] E. L. Elson and H. Qian. Fluorescence correlation spectroscopy with high-order and dual-color correlation to probe nonequilibrium steady-states. *Proc. Natl. Acad. Sci. USA*, 101:2828, 2004.

[79] W. Feller. *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, 1970.

[80] R. P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769, 1949.

[81] M. Fixman and J. Kovac. Polymer conformational statistics .3. modified gaussian models of stiff chains. *J. Chem. Phys.*, 58:1564, 1973.

[82] O. Flomenbom, J. Klafter, and A. Szabo. What can one learn from two-state single-molecule trajectories? *Biophysical J.*, 88:3780, 2005.

[83] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254:1598, 1991.

[84] D. R. Fredkin and J. A. Rice. On aggregated markov processes. *J. Appl. Prob.*, 23:208, 1986.

[85] A. E. Garcia and G. Hummer. Conformational dynamics of cytochrome c: Correlation to hydrogen exchange. *Proteins: Stuct. Funct. and Gene.*, 36:175, 1999.

[86] P. H. Garthwaite. Assessment of prior distributions for regression-models - an experimental-study. *Commun. Stat.Simulat.*, 23:871, 1994.

[87] E. Geva and J. L. Skinner. Two state dynamics of single biomolecules in solution. *Chem. Phys. Lett.*, 288:225, 1998.

[88] J. D. Gezelter, E. Rabani, and B. J. Berne. Calculating the hopping rate for diffusion in molecular liquids: Cs2. *J. Chem. Phys*, 110:3444, 1999.

[89] S. C. Glotzer, V. N. Novikov, and T. B. Schroder. Time-dependent, four-point density correlation function description of dynamical heterogeneity and decoupling in supercooled liquids. *J. Chem. Phys.*, 112:509, 2000.

[90] C. Godreche and J. M. Luck. Anomalous self-diffusion in the ferromagnetic ising chain with kawasaki dynamics. *J. Stat. Phys.*, 104:489, 2001.

[91] G. Goodyear and S. C. Tucker. Glass-like behavior in supercritical fluids: the effect of critical slowing down on solute dynamics. *J. Chem. Phys.*, 111:9673, 1999.

[92] I. V. Gopich and A. Szabo. Statistical of transitions in single molecule kinetics. *J. Chem. Phys.*, 118:454, 2003.

[93] I. Goychuk and P. Hanggi. Ion channel gating: A first-passage time analysis of the kramers type. *Proc. Natl. Acad. Sci*, 99:3552, 2002.

[94] T. Ha, X. W. Zhuang, H. D. Kim J. W. Orr, J. R. Williamson, and S. Chu. Ligand-induced conformational changes observed in single rna molecules. *Proc. Natl. Acad. Sci.*, 96:9077, 1999.

[95] T. J. Ha, A. Y. Ting, J. Liang, W. B. Caldwell, A. A. Deniz, D. S. Chemla, P. G. Schultz, and S. Weiss. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc. Natl. Acad. Sci. USA*, 96:893, 1999.

[96] S. J. Hagen, J. Hofrichter, and W. A. Eaton. Geminate rebinding and conformational dynamics of myoglobin embedded in a glass at room temperature. *J. Phys. Chem.*, 100:12008, 1996.

[97] R. W. Hammond, X. L. Shi, and M. D. Morris. Dynamics of t2 dna during electrophoresis in entangles and ultradilute hydroxyethyl cellulose solutions. *J. Microcolumn Separations*, 44:713, 2000.

[98] K. G. Heinze, M. Jahnz, and P. Schwilley. Triple-color coincidence analysis: One step further in following higher order molecular complex formation. *Biophysical J.*, 86:506, 2–4.

[99] T. Higuchi. Frequency-domain characteristics of linear operator to decompose a time-series into the multi-components. *Ann. I Stat. Math.*, 43:469, 1991.

[100] G. Hinze and H. Sillescu. 2h nuclear magnetic resonance study of supercooled toluene: slow and fast processes above and below the glass transition. *J. Chem. Phys*, 104:314, 1996.

[101] M. E. A. Hodgson and P. J. Green. Bayesian choice among markov models of ion channels using markov chain monte carlo. *Proc. R. Soc. Lond. A*, 455:3425, 1999.

[102] J. S. Hur, E. S. G. Shaqfeh, and R. G. Larson. Brownian dynamics simulations of single dna molecules in shear flow. *J. Rheol.*, 44:713, 2000.

[103] M. Imielinski, C. Belta, A. Halasz, and H. Rubin. Investigating metabolite essentiality through genome-scale analysis of escherichia coli production capabilities. *Bioinformatics*, 21:2008, 2005.

[104] M. Ingelman, S. Ramaswamy, V. Niviére, M. Fontecave, and H. Eklund. Bayesian choice among markov models of ion channels using markov chain monte carlo. *Proc. R. Soc. Lond. A*, 455:3425, 1999.

[105] M. Ingelman, S. Ramaswamy, V. Niviére, M. Fontecave, and H. Eklund. Chrystal structure of nadph:flavin oxidoreductase from e. coli. *Biochemistry*, 28:7040, 1999.

[106] M. Ishikawa, J. Y. Ye, Y. Maruyama, and H. Nakatsuka. Triphenylmethane dyes revealing heterogeneity of their nanoenvironment: Femtosecond, picosecond, and single-molecule studies. *J. Phys. Chem. A*, 103:4319, 1999.

[107] Hans Karl Janssen. On a lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties. *Z. Physik B*, 23:377, 1976.

[108] R. M. Jendrejack, J. J de Pablo, and M. D. Graham. Stochastic simulations of dna in flow: Dynamics and the effects of hydrodynamic interactions. *J. Chem. Phys.*, 116:7752, 2002.

[109] R. V. Jensen. Functional integral approach to classical statistical dynamics. *J. Stat. Phys.*, 25:183, 1981.

[110] Y. Jia, A. Sytnik, L. Li, S. Vladimirov, B. S. Cooperman, and R. M. Hochstrasser. Nonexponential kinetics of a single trna(phe) molecule under physiological conditions. *Proc. Natl. Acad. Sci.*, 94:7932, 1997.

[111] J. Kanetakis, A. Tolle, and H. Sillescu. Rotational diffusion of colloid spheres in concentrated suspensions studied by deuteron nmr. *Phys. Rev. E*, 55:3006, 1997.

[112] M. Karplus. Aspects of protein reaction dynamics: Deviations from simple behavior. *J. Phys. Chem. B*, 104:11, 2000.

[113] A. Kasper, E. Bartsch, and H. Sillescu. Self-diffusion in concentrated colloid suspensions studied by digital video microscopy of core-shell tracer particles. *Langmuir*, 14:5004, 1998.

[114] H. Kim and K. J. Shin. Exact solution of the reversible diffusion-influences reaction for an isolated pair in three dimensions. *Phys. Rev. Lett.*, 82:1578, 1999.

[115] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes. Scaling concepts for the dynamics of viscous liquids near and ideal glassy state. *Phys. Rev. A*, 40:1045, 1989.

[116] J. Klafter and R. Silbey. Derivation of the continuous-time random-walk equation. *Phys. Rev. Lett.*, 44:55, 1980.

[117] G. R. Kneller. Quasielastic neutron scattering and relaxation processes in proteins: analytical and simulation-based models. *PHYSICAL CHEMISTRY CHEMICAL PHYSICS*, 7:2641, 2005.

[118] D. L Koch and E. S. G. Shaqfeh. Averaged-equation and diagrammatic approximations to the average concentration of a tracer dispersed by a gaussian random velocity field. *Phys. Fluids A*, 4:887, 1992.

[119] D. L. Kock and J. F. Brady. Anomalous diffusion due to long-range velocity fluctuations in the absence of a mean flow. *Phys. Fluids A*, 1:47, 1989.

[120] S. C. Kou and X. S. Xie. Generalized langevin equation with fractional gaussian noise: Subdiffusion within a single protein molecule. *Phys. Rev. Lett.*, 93:180603, 2004.

[121] S. C. Kou, X. S. Xie, and J. S. Lui. Bayesian analysis of the rna hairpin. *preprint*, 2003.

[122] R. H. Kraichnan. Diffusion by a random velocity field. *Theor. Phys. of Fluids*, 13:22, 1970.

[123] V. E. Kravtsov, I. V. Lerner, and V. I. Yudson. Random walks in media with constrained disorder. *J. Phys. A*, 18:L703, 1985.

[124] V. E. Kravtsov, I. V. Lerner, and V. I. Yudson. The einstein relation and exact gell-mann-low function for random walks in media with random drifts. *Phys. Lett.*, 119:203, 1986.

[125] J. Kuha. Aic and bic: Comparison of assumptions and performance. *Sociological Methods and Research*, 33:188, 2004.

[126] K. Kuwata, Y. O. Kamatari, K. Akasaka, and T. L. James. Slow conformational dynamics in the hamster prion protein. *Biochemistry*, 43:4439, 2004.

[127] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. Diffusion on a solid surface: Anomalous is normal. *Phys. Rev. Lett.*, 92:250601, 2004.

[128] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. From subdiffusion to superdiffusion of particles on solid surfaces. *Phys. Rev. E*, 70:051104, 2004.

[129] B. Ladoux and P. S. Doyle. Stretching tethered dna chains in shear flow. *Europhysics Letters*, 52:511, 2000.

[130] Y. P. Lai and K. Binder. Grafted polymer layers under shear: a monte carlo simulation. *J. Chem. Phys.*, 98:2366, 1993.

[131] R. G. Larson, T. T Perkins, D. E. Smith, and S. Chu. Hydrodynamics of a dna molecule in a flow field. *Phys. Rev. E*, 55:1794, 1997.

[132] A. I. Lee and J. P. Brody. Single-molecule enzymology of chymotrypsin using water-in-oil emulsion. *Biophysical J.*, 88:4303, 2005.

[133] O. Legeza and J. Solyom. Optimizing the density-matrix renormalization group method using quantum information entropy. *Proc. Natl. Acad. Sci.*, 94:7927, 1997.

[134] V. B. P. Leite, J. N. Onuchic, G. Stell, and J. Wang. Probing the kinetics of single molecule protein folding. *Biophysical Journal*, 87:3633, 2004.

[135] H. P. Lerch, R. Rigler, and A. S. Mikhailov. Functional conformational motions in the turnover cycle of cholesterol oxidase. *Proc. Natl. Acad. Sci. USA*, 102:10807, 2005.

[136] H. T. Li, X. J. Ren, L. M. Ying, S. Balasubramanian, and D Klenerman. Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy. *Proc. Natl. Acad. Sci. USA*, 101:14425, 2004.

[137] C. Z. Liu and I. Oppenheim. Enhanced diffusion upon approaching the kinetic glass transition. *Proc. Natl. Acad. Sci.*, 94:7927, 1997.

[138] H. P. Lu, L. Y. Xun, and X. S. Xie. Single molecule enzymatic dynamics. *Science*, 282:1877, 1998.

[139] V. Lubchenko and R. J. Silbey. Interrupted escape and the emergence of exponential relaxation. *J. Chem. Phys.*, 121:5958, 2004.

[140] J. B. Maddox and E. R. Bittner. Estimating bohm's quantum force using bayesian statistics. *J. Chem. Phys.*, 119:6465, 2003.

[141] A. H. Marcus, J. Schofield, and S. A. Rice. Experimental observations of non-gaussian behaviors and stringlike cooperative dynamics in concentrated quasi-two-dimensional colloidal liquids. *Phys. Rev. E*, 60:5725, 1999.

[142] J. F. Marko and E. D. Siggia. Stretching dna. *Macromolecules*, 28:8759, 1995.

[143] P. C. Martin, E. D. Siggia, and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev. A*, 8:423, 1973.

[144] J. A. McCammon. Protein dynamics. *Rep. Prog. Phys.*, 47:1, 1984.

[145] R. Metzler, E. Barkai, and J. Klafter. Anomalous diffusion and relaxation close to thermal equilibrium: A fractional fokker-planck equation approach. *Phys. Rev. Lett.*, 82:3563, 1999.

[146] R. Metzler, J. Klafter, and I. M. Sokolov. Anomalous transport in external fields: Continuous time random walks and fractional diffusion equations extended. *Phys. Rev. E*, 58:1621, 1998.

[147] L. Miao, H. Guo, and M. J. Zuckermann. Conformation of polymer brushes under shear: chain tilting and stretching. *Marcromolecules*, 29:2289, 1996.

[148] W. E. Moerner. A dozen years of single molecule spectroscopy in physics chemistry and biophysics. *J. Phys. Chem. B*, 106:910, 2002.

[149] W. E. Moerner and M. Orrit. Illuminating single molecules in condensed matter. *Science*, 283:1670, 1999.

[150] C. Monthus and J. P. Bouchaud. Models of traps and glass phenomenology. *J. Phys. A: Math. Gen.*, 29:3847, 1996.

[151] K. L. Montooth, J. H. Marden, and A. G. Clark. Mapping determinants of variation in energy metabolism, respiration and flight in drosophila. *Genetics*, 165:623, 2003.

[152] E. W. Montroll and G. H. Weiss. Random walks on lattices. *J. Math. Phys.*, 6:167, 1965.

[153] C. Mothus and J. P. Bouchaud. Models of traps and glass phenomenology. *J. Phys. A*, 29:3847, 1996.

[154] J. Muller and T. Hillen. Modulation equations and parabolic limits of reaction random-walk systems. *Math. Meth. in Appl. Sci.*, 21:1207, 1998.

[155] K. Murase, T. Fujiwara, Y. Umemura, K. Suzuki, R. Iino, H. Yamashita, M. Saito, Hideji Murakoshi, Ken Ritchie, and A. Kusumi. Ultrafine membrane compartments for molecular diffusion as revealed by single molecule techniques. *Biophys. J.*, 86:4075, 2004.

[156] R. F. Nalewajski. Entropic measures of bond multiplicity from the information theory. *J. Phys. Chem. A*, 104:11940, 2000.

[157] S. Nie and R. Zare. Optical detection of single molecules. *Ann. rev. Biophys. Biomol. Struct.*, 26:567, 1997.

[158] B. Okumus, T. J. Wilson, D. M. J. Lilley, and T. Ha. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophys. J.*, 87:2798, 2004.

[159] J. N. Onuchic, J. Wang, and P. Wolynes. Analyzing single molecule trajectories on complex energy landscapes using replica correlation functions. *Chem. Phys.*, 247:175, 1999.

[160] O. Parcollet and M. Barthelemy. Replica treatment of the effective elastic behavior of a composite. *Phys. Rev. B*, 53:3161, 1996.

[161] T. T. Perkins, S. R. Quake, D. E. Smith, and S. Chu. Relaxation of a single dna molecule observed by optical microscopy. *Science*, 258:1122, 1992.

[162] T. T. Perkins, D. E. Smith, and S. Chu. Single polymer dynamics in an elongational flow. *Science*, 276:2016, 1997.

[163] T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu. Stretching of a single tethered polymer in a uniform flow. *Science*, 268:83, 1995.

[164] V. Pham and M. W. Deem. Direct numerical observation of anomalous diffusion in random media. *J. Phys. A*, 31:7235, 1998.

[165] R. Phythian. Dispersion by random velocity-fields. *J. of Fluid Mech.*, 67:145, 1975.

[166] R. Phythian. The operator formalism of classical statistical dynamics. *J. Phys. A*, 8:1423, 1975.

[167] R. Phythian. Further application of the martin, siggia, rose formalism. *J. Phys. A*, 9:239, 1976.

[168] R. Phythian. The functional formalism of classical statistical dynamics. *J. Phys. A*, 10:777, 1977.

[169] R. Phythian. Effective long-time diffusivity for a passive scalar in a gaussian model fluid-flow. *J. Fluid Mech.*, 89:241, 1978.

[170] D. N. Politis. A simple information-theoretic proof of the maximum-entropy property of some gaussian random-fields. *IEEE Trans. Imag. Proc.*, 4:865, 1994.

[171] M. Qian and H. Qian. Pumped biochemical reactions, nonequilibrium circulation, and stochastic resonance. *Phys. Rev. Lett.*, 84:2271, 2000.

[172] F. Qin, A. Auerbach, and F. Sachs. Maximum likelihood estimation of aggregated markov processes. *Proc. R. Soc. Lond. B*, 264:375, 1997.

[173] S. R. Quake, H. Babcock, and S. Chu. The dynamics of partially extended single molecules of dna. *Nature*, 388:151, 1997.

[174] E. Rabani, J. D. Gezelter, and B. J. Berne. Calculating the hopping rate for self-diffusion on rough potential energy surfaces: cage correlations. *J. Chem. Phys*, 107:6967, 1997.

[175] S. A. Rice. Diffusion-limited reactions. In C. H. Bamford, C. F. H Tipper, and R. G. Compton, editors, *Comprehensive Chemical Kinetics*, volume 25. Elsevier, New York, 1985.

[176] K. Ritchie, X. Y. Shan, J. Kondo, K. Iwasawa, T. Fujiwara, and A. Kusumi. Detection of non-brownian diffusion in the cell membrane in single molecule tracking. *Biophys. J.*, 88:2266, 2005.

[177] C. P. Robert, T. Ryden, and D. M. Titterington. Bayesian inference in hidden markov models through the reversible jump markov chain monte carlo method. *J. Roy. Stat. Soc. B*, 62:57, 2000.

[178] P. H. Roberts. Analytical theory of turbulent diffusion. *J. Fluid. Mech.*, 11:257, 1961.

[179] A. E. Roitberg, R. B. Gerber, and M. A. Ratner. A vibrational eigenfunction of a protein: Anharmonic coupled-mode ground and fundamental excited states of bpti. *J. Phys. Chem. B*, 101:1700, 1997.

[180] S. De S, Y. Shapir, and E. H. Chimowitz. Scaling of self and fickian diffusion coefficients in the critical region. *Chem. Eng. Sci.*, 56:5003, 2001.

[181] H. Scher and E. W. Montroll. Anomalous transit-time dispersion in amorphous solids. *Phys. Rev. B*, 12:2455, 1975.

[182] S. Schnell and T. E. Turner. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog. Biophys. Mol. Bio.*, 85:235, 2004.

[183] A. Schob, F. Cichos, J. Schuster, and C. von Borczyskowski. Reorientation and translation of individual dye molecules in a polymer matrix. *Euro. Poly. J.*, 40:1019, 2004.

[184] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumke, and W. A. Eaton. Polyproline and the spectroscopic ruler revisited with single-molecule fluorescence. *Proc. Natl. Acad. Sci. USA*, 102:2754, 2005.

[185] K. Seki, M. Wojcik, and M. Tachiya. Recombination kinetics in subdiffusive media. *J. Chem. Phys.*, 119:7527, 2003.

[186] M. K. Sen and P. L. Stoffa. Rapid sampling of model space using genetic algorithms - examples from seismic wave-form inversion. *Geophys. J. Int.*, 108:281, 1992.

[187] M. Y. Shen and K. F. Freed. Long time dynamics of met-enkephalin: Tests of mode-coupling theory and implicit solvent models. *J. Chem. Phys.*, 118:5143, 2003.

[188] X. L. Shi, R. W. Hammond, and M. D. Morris. Dynamics of dna during pulsed-field electrophoresis in entangled and dilute polymer solutions. *Polymer*, 42:8483, 2001.

[189] J. E. Shore and R. Zwanzig. Dielectric-relaxation and dynamic susceptibility of a one-dimensional model for perpendicular-dipole polymers. *J. Chem. Phys.*, 63:5445, 1975.

[190] J. Skilling and R. K. Bryan. Maximum-entropy image-reconstruction - general algorithm. *Mon. Not. R. Astron. Soc.*, 211:111, 1984.

[191] D. E. Smith, H. P. Babcock, and S. Chu. Single polymer dynamics in steady shear flow. *Science*, 283:1724, 1999.

[192] D. E. Smith and S. Chu. Response of flexible polymers to a sudden elongational flow. *Science*, 281:1335, 1998.

[193] S. B. Smith, L. Finzi, and C. Bustamante. Direct mechanical measurements of the elasticity of single dna by using magnetic beads. *Science*, 258:1122, 1992.

[194] I. M. Sokolov, A. Blumen, and J. Klafter. Linear response in complex systems: Ctrw and the fractional fokker-planck equations. *Physica A*, 302:268, 2001.

[195] M. J. Solomon and Q. Lu. Rheology and dynamics of particles in viscoelastic media. *Current Opinion in Colloid & Interface Science*, 6:430, 2001.

[196] L. Song and K. L. Magleby. Testing for microscopic reversibility in the gating of maxi k+ channels using 2-dimensional dwell-time distributions. *Biophysical J.*, 67:91, 1994.

[197] B. J. STANLEY, K. TOPPER, and D. B. MARSHALL. Analysis of the heterogeneous rate of dissociation of cu(ii) from humic and fulvic-acids by statistical deconvolution. *Analytica Chimica Acta*, 287:25, 1994.

[198] P. J. Steinbach, K. Chu, H. Frauenfelder, J. B. Johnson, D. C. Lamb, G. U. Nienhaus, T. B. Sauke, and R. D. Young. Determination of rate distributions from kinetic experiments. *Biophysical J.*, 61:235, 1992.

[199] T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The elasticity of a single supercoiled dna molecule. *Science*, 271:1835, 1996.

[200] J. Sung, E. Barkai, R. Silbey, and S. Lee. Fractional dynamics approach to diffusion-assisted reactions in disordered media. *J. Chem. Phys.*, 116:2338, 2002.

[201] J. Sung and R. J. Silbey. Exact dynamics of a continuous time random walker in the presence of a boundary: Beyond the intuitive boundary condition approach. *Phys. Rev. Lett.*, 91:160601, 2003.

[202] S. Tan. Monte-carlo simulation of polymers in shear flow. *Polymer*, 40:695, 1999.

[203] G. Tarjus and D. Kielson. Breakdown of the stokes-einstein relation in super-cooled liquids. *J. Chem. Phys*, 103:3071, 1995.

[204] P. Thomen, U. Bockelmann, and F. Heslot. Rotational drag on dna: a single molecule experiment. *Phys. Rev. Lett.*, 88:8102, 2002.

[205] M. Tokuyama, Y. Enomoto, and I. Oppenheim. Slow dynamics of supercooled colloidal fluids: spatial heterogeneitites and nonequilibrium density fluctuations. *Physica A*, 270:380, 1999.

[206] A. Tolle and H. Sillescu. Rotational diffusion of colloid spheres in suspension investigated by deuteron nmr. *Langmuir*, 10:4420, 1994.

[207] S. C. Tucker. Solvent density inhomogeneities in supercritical fluids. *Chem. Rev.*, 99:391, 1999.

[208] R. S. Urdahl, K. D. Rector, D. J. Myers, P. H. Davis, and M. D. Fayer. Vibrational relaxation of a polyatomic solute in a polyatomic supercritical fluid near the critical point. *J. Chem. Phys*, 105:8973, 1996.

[209] R. van Zon and J. Schofield. Mode-coupling theory for multiple-point and multiple-time correlation functions. *Phys. Rev. E*, 65:011106, 2002.

[210] R. van Zon and J. Schofield. Multiple-point and multiple-time correlation functions in a hard-sphere fluid. *Phys. Rev. E*, 65:011107, 2002.

[211] L. Venkataramanan and F. J. Sigworth. Applying hidden markov models to the analysis of single ion channel activity. *Biophysical J.*, 82:1930, 2002.

[212] R. Verberk and M. Orrit. Photon statistics in the fluorescence of single molecules and nanocrystals: Correlation functions versus distributions of on- and off-times. *J. Chem. Phys.*, 119:2214, 2003.

[213] C. Verde, B. D. Howes, M.C. DeRosa, L. Raiola, G. Smulevich, R. Williams, B. Giardina, E. Parisi, and C. DiPrisco. Structure and function of the gondwanian hemoglobin of pseudaphritis urvillii, a primitive notothenioid fish of temperate latitudes. *Protein Science*, 13:2766, 2004.

[214] M. Wagner and J. Timmer. The effects of non-identifiability on testing for detailed balance in aggregated markov models for ion-channel gating. *Biophysical Journal*, 79:2918, 2000.

[215] J. Wang, , and P. Wolynes. Intermittency of single-molecule reaction dynamics in fluctuating environments. *Phys. Rev. Lett.*, 74:4317, 1995.

[216] J. Wang and P. Wolynes. Intermittency of single molecule reaction dynamics. *Phys. Rev. Lett.*, 74:4317, 1995.

[217] K. G. Wang. Long-range correlation effects, generalized brownian motion and anomalous diffusion. *J. Phys. A*, 27:3655, 1994.

[218] L. P. Watkins and H. Yang. Information bounds and optimal analysis of dynamic single molecule measurements. *Biophysical Journal*, 86:4015, 2004.

[219] Web of Science. based on search for articles, 2005.

[220] E. R. Weeks, J. C. Crocker, A. C. Levitt, A. Schofield, and D. A. Weitz. Three-dimensional direct imagining of stuctural relaxation near the colloidal glass transition. *Science*, 287:627, 2000.

[221] S. Weiss. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nature Struct. Biol.*, 7:724, 2000.

[222] J. B. Witkoskie and J. S. Cao. Scaling and universality of inherent structure simulations. *Phys. Rev. E*, 69:061108, 2004.

[223] J. B. Witkoskie and J. S. Cao. Single molecule kinetics. i. theoretical analysis of indicators. *J. Chem. Phys*, 121:6361, 2004.

[224] J. B. Witkoskie and J. S. Cao. Single molecule kinetics. ii. numerical bayesian approach. *J. Chem. Phys*, 121:6373, 2004.

[225] J. B. Witkoskie and J. S. Cao. Analysis of the entire sequence of a single photon counting experiment on a flavin protein. *in preparation*, 2005.

[226] J. B. Witkoskie and J. S. Cao. Classification of single molecules based on two dimensional data. *in preparation*, 2005.

[227] J. B. Witkoskie and J. S. Cao. Signatures of detailed balance violations in single molecule blinking sequences. *in preparation*, 2005.

[228] J. B. Witkoskie and J. S. Cao. Testing for renewal and detailed balance violations in single molecule blinking sequences. *in preparation*, 2005.

[229] J. B. Witkoskie and J. S. Cao. Two time interrupted diffusion propagator. *in preparation*, 2005.

[230] J. B. Witkoskie, Vas Lubchenko, , J. S. Cao, and R. J. Silbey. Coupling first order kinetics with fractional diffusion. *in preparation*, 2005.

[231] J. B. Witkoskie, J. Wu, and J. S. Cao. Basis set study of classical rotor lattice dynamics. *J. Chem. Phys.*, 120:5696, 2004.

[232] J. Wu and J. Cao. Gaussian factorization of hydrodynamic correlation functions and mode-coupling memory kernels. *Physical Review E*, 67:061116, 2003.

[233] X. Y. Xia and P. G. Wolynes. Microscopic theory of heterogeneity and nonexponential relaxations in supercooled liquids. *Phys. Rev. Lett.*, 86:5526, 2001.

[234] X. S. Xie. Single molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuations and enzymatic dynamics. *J. Chem. Phys.*, 117:11024, 2002.

[235] X. S. Xie and J. K. Trautman. Optical studies of single molecules at room temperature. *Ann. Rev. Phys. Chem.*, 49:441, 1998.

[236] Z. Xie, X. Fang, T. R. Sosnick, T. Pan, and N. F. Scherer. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *Biochemsitry*, 40:100, 2001.

[237] Z. Xie, N. Srividya, T. R. Sosnick, T. Pan, and N. F. Scherer. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. *P. Natl. Acad. Sci. USA*, 101:334, 2004.

[238] G. Xu, J. Ding, and Y. Yang. Monte carlo simulation of self-avoiding lattice chains subject to oscillatory shear flow in polymer solution. *Rheol. Acta*, 40:60, 2001.

[239] X. Xu and E. S. Yeung. Direct measurement of single-molecule diffusion and photodecomposition in free solution. *Science*, 275:1106, 1997.

[240] H. Yang, G. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie. Protein conformational dynamics probed by single-molecule electron transfer. *Science*, 302:262, 2003.

[241] H. Yang and X. S. Xie. Statistical approaches for probing single molecule dynamics. *Chem. Phys.*, 284:423, 2002.

[242] S. L. Yang and J. S. Cao. Two-event echos in single molecule kinetics: A signature of conformational fluctuations. *J. Phys. Chem. B*, 105:6536, 2001.

[243] S. L. Yang and J. S. Cao. Direct measurements of memory effects in single molecule kinetics. *J. Chem. Phys.*, 117:10996, 2002.

[244] J. Y. Ye, Y. Yamane, M. Yamauchi, H. Nakatsuka, and M. Ishikawa. Direct observation of the interaction of single fluorescent nucleotide analogue molecules with dna polymerase i. *Chem. Phys. Lett.*, 320:607, 2000.

[245] E. Zaccarelli, G. Foffi, F. Sciortino, P. Tartaglia, and K. A. Dawson. Gaussian density fluctuations and mode coupling theory for supercooled liqids. *Europhys. Lett.*, 55:157, 2001.

[246] X. W. Zhuang, L. E. Bartley, H. P. Babcock, R. Russell, T. J. Ha, D. Herschlag, and S. Chu. A single molecule study of rna catalysis and folding. *Science*, 288:2048, 2000.

[247] X. W. Zhuang, H. Kim, M. J. B. Pereira, H. P. Babcock, N. G. Walter, and S.Chu. Correlating structural dynamics and function of single ribozyme molecules. *Science*, 296:1473, 2002.

[248] B. H. Zimm. Extension in flow of a dna molecule tethered at one end. *Macromolecules*, 31:6089, 1998.

[249] R. Zwanzig. Diffusion in a dynamically disordered continuum. *Chem. Phys. Lett.*, 164:639, 1989.

488 Cambridge St. #3
Cambridge, MA 02141-1113
(617)876-9151
jbw@mit.edu

James B. Witkoskie

Education

2005   M.I.T.   Cambridge, MA

Ph. D. Physical Chemistry

Adviser: Jianshu Cao

Thesis: Modeling and Analysis of single molecule experiments

- Applied Bayesian, maximum entropy, Monte Carlo, and signal processing methods to determine validity of diffusion models for describing experiments on single molecule systems.
- Wrote Monte Carlo and molecular dynamics simulations to explore the physics of various systems.
- Analyzed the shortcomings of previous methods of analysis used in the single molecule field.
- Analyzed the physical repercussions of certain random walk and diffusion models, such as semi-Markov processes and passive diffusion in random field.

2000   University of Delaware   Newark, DE

B.S. Chemistry

Minor: Mathematics

DuPont Scholar: Full tuition/room/board/books

Degree with Distinction.

Honors Certificate.

Adviser: Douglas Doren

Thesis: Potential energy surface interpolation through ab initio calculations, neural networks, and Kalman filtering.

Developed a method of determining a potential energy surface for interaction of water with various solutes by sampling configurations from a molecular dynamics simulation with a psuedo-potential energy surface, determining the exact quantum energy of the configuration from the ab initio calculations, and updating the psuedo-potential energy surface through adaptive neural networks and signal processing techniques.

Experience

2001-2002   M.I.T.   Cambridge, MA

Teaching Assistant

Quantum Mechanics/Statistical Mechanics

- Held Recitations to review and reinforce fundamental concepts of chemical physics
- Wrote and graded homework and examines to determine student proficiency
- Helped to modify and update curriculum to address student needs

1998-2000   Univ. of Delaware   Newark, MA

Lab Technician

Analyzed chemical samples to identify components through mass spectroscopy.

1997   Merck and Co.   West Point, PA
Intern in Vaccine Engineering
Worked through Ross Resources; Plymouth Meeting, PA
- Process development
- Quality testing

Presentations

Sept. 2003   American Chemical Society   New York, NY
Applications of computer simulations and analytic methods to the study of
constrained dynamic systems

Nov. 2001   Hole Burning and Single Molecule   Taipei, Taiwan
Analysis of single molecule experiments.

May. 2000   Mid-Atlantic Conference   Bloomberg, PA
Determination of transition state geometry from potential energy surface
interpolation through neural networks and signal processing

Publications

Analysis of the entire photon emission sequence of a flavin protein. J.B. Witkoskie
and J.S. Cao. (in preparation) (2005)

Signatures of detailed balance violations in single molecule blinking sequences. J.B.
Witkoskie and J.S. Cao. (in preparation) (2005)

*Testing renewal and detailed balance violations in single molecule blinking
sequences.* J. B. Witkoskie and J.S. Cao. (in preparation) (2005)

Classification of single molecule sequences. J. B. Witkoskie and J.S. Cao. (in
preparation) (2005)

Aging in the interrupted fractional diffusion propagator. J.B. Witkoskie and J.S.
Cao. (in preparation) (2005)

Coupling fractional and first order kinetic processes to . V. Lubchenko, J.B.
Witkoskie, R.J. Silbey, and J.S. Cao. (in preparation) (2005)

Neural network models of potential energy surfaces: prototypical examples. J. B.
Witkoskie and D. J. Doren. Journal of Chemical Theory and Computation 1 XX
(2006)
Analysis of single molecule experiments II: Numerical Bayesian approach. J.B.
Witkoskie and J.S. Cao. Journal of Chemical Physics   121, 6373 (2004)

Analysis of single molecule experiments I: Analysis of indicators. J. B. Witkoskie
and J.S. Cao. Journal of Chemical Physics 121 6361 (2004)

Scaling and universality of inherent structure simulations. J.B. Witkoskie and J.S.
Cao. Physical Review E 69 061108 (2004)

Basis set study of classical rotor lattice dynamics. J.B. Witkoskie, J. Wu , and J.S.
Cao. Journal of Chemical Physics 120, 5695 (2004)

First-principle path integral study of DNA under hydrodynamic flows. S.L. Yang,
J.B. Witkoskie, and J.S. Cao. Chemical Physics Letters 377, 399 (2003)

Single-molecule dynamics of semi-flexible Gaussian chains. S.L. Yang, J.B.
Witkoskie, and J.S. Cao. Journal of Chemical Physics 117, 11010 (2002)

Brownian motion in dynamically disordered media. J. B. Witkoskie, S.L. Yang, J.S. Cao. Physical Review E 66 051111 (2002)