Analysis Methods for Single Molecule Fluorescence Spectroscopy

Kristofer Gryte

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy at the University of Oxford

St Cross College

University of Oxford

Trinity Term 2012

Disclaimer

I hereby declare that the work in this thesis is that of the candidate alone, except where indicated in the text, and as described below.

- Ludovic Le Reste, Johannes Hohlbein, and Ling Chin Hwang designed and built the microscope setup; see *Materials and Methods* section D.3.
- Seamus J. Holden wrote data processing software to extract fluorescence intensities from TIRF microscopy images (i.e., image registration, auto-detection, photometry); see *Materials and Methods* section D.5.
- Ling Chin Hwang and Nicole Robb labeled and purified the DNA used for the single-molecule studies of bacterial RNA polymerase (RNAP); see *Materials and Methods* section D.1.
- Thorben Cordes and Nicole Robb performed solution-based experiments of RNAP-promoter open complex and provided data for burst variance analysis (BVA) and for burst correlation analysis (BCA); see Chapters 4, 5, and 6.
- Diego Duchi performed surface-immobilized RNAP-promoter open complex experiments and provided data for mean variance analysis (MVA); see Chapter 7.

I state all contributions at each chapter's end. I, alone, accept responsibility for the material contained in this thesis. All errors are mine. I have benefited from the help, feedback, ideas, discussions, and questions of others; an individual is not an island. I could not have brought this thesis to fruition without the contributions of many. I therefore use the inclusive *we* to reflect that my achievement results from a collective effort.

Kristofer Gryte
October 2012

> *...we are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size.* – attributed to Bernard of Chartres

Analysis Methods for Single Molecule Fluorescence Spectroscopy

Kristofer Gryte, St Cross College
Biological Physics Research Group, Department of Physics

Thesis submitted for the degree of Doctor of Philosophy
University of Oxford
Trinity Term 2012

Abstract

This thesis describes signal analysis methods for single-molecule fluorescence data. The primary factor motivating method development is the need to distinguish single-molecule FRET fluctuations due to conformational dynamics from fluctuations due to distance-independent FRET changes.

Single-molecule Förster resonance energy transfer (FRET) promises a distinct advantage compared to alternative biochemical methods in its potential to relate biomolecular structure to function. Standard measurements assume that the mean transfer efficiency between two fluorescent probes, a donor and an acceptor, corresponds to the mean donor-acceptor distance, thus providing structural information. Accordingly, measurement analysis assumes that mean transfer efficiency fluctuations entail mean donor-acceptor distance fluctuations. Detecting such fluctuations is important in resolving molecular dynamics, as molecular function often necessitates structural changes. A problem arises, however, in that factors other than donor-acceptor distance changes may induce mean transfer efficiency fluctuations. We refer to these factors as distance-independent FRET changes.

We present analysis methods to detect distance-independent photophysical dynamics and to determine their correlation with distance-dependent FRET dynamics. First, we review a theory of photon statistics and show how we can use the theory to detect FRET fluctuations. Second, we extend the theory to alternating laser excitation (ALEx) measurements and demonstrate how fluorophore stoichiometry, a measure of fluorophore brightness, reports on distance-independent photophysical dynamics. Next, we provide a measure to determine the extent to which stoichiometry fluctuations account for FRET dynamics. Finally, we use a framework similar to the preceding along with recent advances in the theory of total internal reflection fluorescence (TIRF) microscopy FRET measurements to detect TIRF FRET fluctuations which occur on a timescale faster than the measurement temporal resolution. We validate our methods with simulations and demonstrate their utility in delineating RNA polymerase open complex conformational dynamics.

Acknowledgements

I am indebted to many for their support and assistance with this work.

I am grateful for financial support from the Clarendon Fund. Without this support, my study at the University of Oxford could not have happened. Thank you.

For believing in me and overseeing my growth and education, my teachers, professors, and mentors. A few stand above the rest: Gerard Protomastro, Brian Hopkins, Peter Cvek, David Surrey, Richard Thurston, William Evans, Regina Giuliani, Frances Raleigh, James Grant, Leonard Sciorra, Richard Loomis, Dewey Holten, Nancy McKinney, Karen Rowe, Brian Hart, and Coach Principi (P).

For his help, humour, and coding expertise, Gerhard Blab.

For his influence in helping me become a better writer, David Gryte.

For convincing me to stand steadfast and complete this work, Frank Krikhaar. You were right.

For friendship and helping me become who I am today, Derek Oldridge, Pinshane Huang, Aimel Rios, Nastassja Pugliese, Julian Summers, Stephen Layne, Max Thompson, Lee Neuharth, Bigyan Bista, Victoria Coleman, Melisa Garcia LLamas, Catalina Cardona, and Casey Dickemann.

For their help in all things science, their friendship, and the conversations, Alexandra Tomescu, Joseph Torella, Robert Crawford, and Johannes Hohlbein.

For his amazing work and dedication, Alistair Wardrope.

For the late nights, long discussions, and el pigskin, Stephan Uphoff and Ludovic Le Reste. For his friendship and constant support for all things RNA polymerase, Diego Duchi.

For his guidance and patience and for giving me the freedom to explore, Achillefs Kapanidis. Thank you for all you have done on my behalf.

For being who she is, Starr Leigh Gardner. If not in body, in spirit. A*.

For her support, advice, companionship, and love, Keren Simons. Once a Teefer, always a Teefer. Your icons will never be forgotten.

For his insight and editorial assistance and for making my mother laugh, Robert Gislason, a.k.a., Giz.

For her love and support, my mother, Barbara Gislason. You have always been there, and always will be. To you, I owe the most.

Kristofer Gryte
Oxford, United Kingdom
March 2013

Publications arising from DPhil

Crawford, R., Torella, J.P., Aigrain, L., Plochowietz, A., **Gryte, K.**, Uphoff, S., and A.N. Kapanidis (2013) "Long-lived intracellular single-molecule fluorescence using electroporated molecules", *Nat. Meth.*, (under review).

Robb, N.C., Cordes, T., Hwang, L.C., **Gryte, K.**, Duchi, D., Craggs, T.D., Santoso, Y., Weiss, S., Ebright, R.H., and A.N. Kapanidis (2013) "The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics: implications for transcription start-site selection", *J. Molec. Biol.*, 425(5):875-885.

Le Reste, L., Hohlbein, J., **Gryte, K.**, and A.N. Kapanidis (2012) "Characterization of dark quencher chromophores as nonfluorescent acceptors for single-molecule FRET", *Biophys. J.*, 11:2658-2668.

Uphoff, S., **Gryte, K.**, Evans, G., and A.N. Kapanidis (2011) "Improved temporal resolution and linked hidden Markov modeling for Switchable FRET", *ChemPhysChem*, 3:571-579.

Cordes, T., Santoso, Y., Tomescu, A.I., **Gryte, K.**, Hwang, L.C., Camará, B., Wigneshweraraj, S., and A.N. Kapanidis (2010) "Sensing DNA opening in transcription using quenchable FRET", *Biochem.*, 49:9171-9180.

Hohlbein, J., **Gryte, K.**, Heilemann, M., and A.N. Kapanidis (2010) "Surfing on a new wave of single-molecule fluorescence methods", *Phys. Biol.*, **7** 031001.

Conference Presentations

56th Annual Biophysical Society Meeting

San Diego, California, USA **February 25-29, 2012**

"Regime-changing hidden Markov modeling and statistical analysis for complex single-molecule time series": Poster presentation regarding new statistical methods for single-molecule data analysis.

55th Annual Biophysical Society Meeting

Baltimore, Maryland, USA **March 5-9, 2011**

"Real-time initial transcription by a multisubunit RNA polymerase": Poster presentation showing new quenchable FRET assay and RNAP RNA synthesis.

22nd Annual RNA Polymerase Workshop

University of Nottingham, Nottingham, UK **April 12-13, 2010**

"Real-time, single-molecule abortive transcription of RNAs": Presented research showing individual synthesis and release of abortive RNAs by RNAP.

21st Annual RNA Polymerase Workshop

University of Bristol, Bristol, UK **March 23-24, 2009**

"Real-time kinetics of initial transcription at near-base-pair resolution using single-molecule FRET": Presented research demonstrating kinetic behavior of RNAP RNA synthesis and release.

Contents

Chapter 1

Introduction

...and so we begin

Scientific breakthroughs are often accompanied by innovative techniques and instrumentation. Fluorescence[1] detection is one such class of techniques which has provided insight into fundamental physical and biological questions. New fluorescence probes, biological assays, and instruments have significantly advanced fluorescence detection. With the first observation of single fluorophores in aqueous solution [289], single-molecule fluorescence detection joined patch clamp techniques for single ion-channel recording [9], atomic force microscopy [64, 170, 236], tethered particle motion [19, 84, 212, 224, 280, 325, 325, 357], magnetic tweezers [170, 264], and optical tweezers [20, 304, 333] as a biophysical method capable of revealing biomolecular behaviors previously obscured by ensemble- and time-averaging inherent to conventional analysis. Single-molecule methods could study reaction pathways and asynchronous reactions in real time and could uncover short-lived intermediates. The additional capability to combine optical and mechanical measurements on single-molecules further proposed new avenues for scientific discovery [333].

Single-molecule fluorescence methods included detection of biomolecules both freely diffusing in solution and immobilized on surfaces. Some methods used a

[1] i.e., spontaneous photon emission by a molecule after electronic excitation. Electronic excitation occurs by photon absorption where the absorbed photon has an energy matching the energy gap between a ground and a higher ro-vibronic energy level.

point-detection methodology adapted from fluorescence correlation spectroscopy[2] [79, 121, 287]. A point-detection approach uses confocal optics and typically monitors only one molecule at a time but with possible picosecond temporal resolution. Other methods used wide-field microscopy, often involving evanescent electric fields [10], to excite immobilized molecules and to exploit simultaneous millisecond imaging of multiple individual molecules [367]. Initial point-detection and wide-field methods used single-laser excitation and collected fluorescence via a single emission detection channel. These methods could study singly-labeled molecules to report on fluorophore location, fluorescence lifetime, quantum yield, and diffusion coefficients [237, 339]. Higher complexity involved adding a second detection channel. Initial 'dual-channel' work performed polarization measurements to report on fluorophore rotational freedom and spectral shift (reviewed in [339]). Introduction of a second fluorophore enabled inter-fluorophore distance measurements within a single-molecule, either through high-resolution localization [110] or Förster resonance energy transfer (FRET)[3] [88–91]. The efficiency of FRET has a steep distance dependence, making FRET useful as a nanoscale ruler to study biomolecular structure and dynamics [284, 301].

The first reports of single-molecule detection were seminal proof-of-principle experiments. The measurements, however, were rather limited due to fluorophore photobleaching, complex photophysical behavior, an inability to work at fluorophore concentrations greater than 200 pM, complications involving multi-component biomolecules, and surface-immobilization challenges. Several recent solutions address these challenges, helping shift an emphasis on initial method development to an emphasis on answering biological questions. For example, protein encapsulation in lipid vesicles [24, 244, 265] allowed extended observation times without perturbation arising from direct surface immobilization. Alternating laser excitation (ALEx) methods permitted both FRET and stoichiometric observation and

[2] Fluorescence correlation spectroscopy (FCS) is a 'small-ensemble' method that reports on diffusion and reaction kinetics vis-à-vis temporal fluctuations in fluorescence intensity.

[3] FRET is a non-radiative energy transfer from one molecule, called a 'donor', to another molecule, called an 'acceptor'. Transfer occurs via a weak dipole-dipole coupling interaction.

thus facilitated experiments involving multi-component interaction [167,195,229]. Furthermore, our understanding of fluorophore photophysics, such as blinking, bleaching, and triplet states, has led to advances which improve our ability to influence fluorophore behavior [95,262,331]. The additional discovery of conditions that make fluorophores controllable photoswitches [12,13,124–126,296,329] set the stage for new super-resolution methods to investigate biomolecular structure and cellular organization. And lastly, methods developed using non-fluorescent FRET acceptors[4] now allow micromolar labeled analyte concentrations and thus enable the study of weakly interacting biomolecules [96,193,285,334].

New experimental methods, however, breed new data types, where we denote a new 'type' to reflect the constraints and prior probability distributions[5] determined by the experiment and our knowledge of the studied physical system. For example, we can describe photon statistics by a Poisson distribution, where photon counts belong to the set of natural numbers. From the detected photon counts, we typically calculate the FRET efficiency in order to reflect an energy transfer probability. As such, FRET efficiency distributions cannot physically assume values outside the interval from $[0, 1]$. Moreover, photon emission stochasticity introduces fundamental FRET efficiency resolution limits[6]. The combination of all constraints and priors makes each experimental data type unique and requires tailored methods for *data reduction*, i.e., the process of identifying and extracting meaningful information from data.[7]

The journey to scientific discovery does not, however, end with data reduction. Once meaningful information is at hand, determining what the data means and ascribing biological relevance introduces another layer of complexity. With regard to fluorescence detection measurements, FRET is one of several competing photophysical processes which can occur after electronic excitation of a donor fluorophore. Accordingly, complications arise when we attempt to unambiguously re-

[4] also known as 'dark quenchers'
[5] often simply called *priors*
[6] i.e., the ability to resolve two or more component distributions
[7] This is not to say that experimental data types do not share similarities, but to realize that we should not expect, or even desire, a one-size-fits-all approach to data analysis.

late a variable such as observed fluorescence to a specific biological function [272]. The story of nucleosome dynamics serves as one cautionary tale. Tomschik and co-workers reported a significant discovery when they used single-molecule FRET and found fast, long-range, reversible conformational fluctuations in nucleosomes [312]. Two years later, however, Koopmans and co-workers used single-molecule FRET and ALEx and found that distance-independent photophysics accounted for most of the observed dynamic fluctuations [183]. Tomschik and co-workers subsequently re-evaluated their work and found results consistent with Koopmans, largely nullifying their prior findings [311].

With the exponential growth of publications containing single-molecule fluorescence detection measurements [160, 225], researchers from the biological sciences will increasingly seek to answer questions using FRET and related methods. While a boon for the single-molecule field, the peril is that insufficient understanding of experimental complications will lead researchers to naive conclusions. As others have recognized [107], single-molecule data analysis lacks standardization, with each laboratory having its own data collection, storage, and analysis strategy. Furthermore, most works present new single-molecule analysis methods and make the implicit assumption that FRET fluctuations correspond to conformational dynamics [3, 25, 42, 102–105, 107, 161, 197, 204, 218, 338]; fewer works aim to identify distance-independent photophysics either experimentally [43, 44, 165] or statistically [21, 165, 310]. As the nucleosome dynamics example highlights, a need exists for experimental and statistical methods to delineate conformationally induced FRET fluctuations from fluctuations due to distance-independent photophysics. While alternating laser excitation [167, 195, 229], multiparameter fluorescence detection [44, 165], and TIRF-based spectra [43] methods provide experimental solutions, complementary and, perhaps more importantly, independent signal analysis methods remain to be developed. This thesis aims to provide such a contribution.

We note, however, that the personal motivation for this thesis would undoubt-

edly not have arisen had experimental work not been conducted on bacterial RNA polymerase in an attempt to better understand initial transcription. RNA polymerase (RNAP) is a highly dynamic and multifunctional multi-subunit enzyme, which acts in concert with associated transcription factors to orchestrate transcription. Transcription is the highly regulated first step in gene expression. RNAP transcribes DNA into mRNA which subsequently enters the pathway to protein synthesis. To initiate transcription, RNAP must first bind a transcription factor called a σ factor and then interact with double-stranded promoter DNA to form an initial transcription complex. The initial transcription complex must undergo several structural changes in order to become a mature transcription elongation complex and synthesize one strand of mRNA ready for ribosome translation. Our work focused on the RNAP open complex and attempted to provide structural and kinetic insight to RNAP dynamics. But as this thesis demonstrates, the quest for unambiguous conclusions remains allusive, with significant challenges confronted in distinguishing FRET fact from artifact.

This thesis is organized as follows:

Chapter 2 introduces photophysics and photochemistry and establishes the basis for all observed phenomena in this thesis.

Chapter 3 provides an overview of single-molecule signal analysis methods and helps motivate the present work.

Chapter 4 reviews one such signal analysis method, burst variance analysis (BVA), and demonstrates its use in investigating RNA polymerase open complex dynamics.

Chapter 5 extends BVA to fluorophore stoichiometry. Stoichiometry BVA identifies distance-independent photophysical dynamics and provides a measure for fluorophore fluorescence stability. After discussing the theory and validating the method through simulations, the chapter revisits the RNA polymerase open complex dynamics data of Chapter 4 and highlights the difficulty in distinguishing FRET fluctuations due to biomolecular conformational dynamics from distance-

independent photophysics.

Chapter 6 introduces a method termed burst correlation analysis (BCA) to correlate distance-independent photophysics with conformational dynamics FRET fluctuations. Simulations validate BCA, and, once again, RNA polymerase open complex dynamics data serves as the experimental test case. The chapter ends by arguing that the stoichiometric fluctuations in Chapter 5 can at least partially account for the FRET dynamics of Chapter 4.

Chapter 7 builds on recent work which derived the measurement uncertainty for single-molecule TIRF and applies a methodology similar to BVA to detect dynamics in single-molecule FRET time series. The method is termed mean-variance analysis (MVA), and simulations demonstrate the method's effectiveness in detecting FRET dynamics across different kinetic regimes. The chapter details MVA's application to surface-immobilized RNA polymerase open complex data and discusses inter-molecular heterogeneity within the context of the branched pathway hypothesis for transcription initiation.

Chapter 8 summarizes the present work and suggests directions for future development.

Chapter 2

Photophysics and Photochemistry

Abstract

Single-molecule fluorescence has emerged as a potentially powerful technique to probe biochemical and cellular processes. By monitoring fluorescence intensity fluctuations over time, we are able to deduce reaction pathways, link structure to function, and reveal the stochastic nature of molecular processes. Here, we introduce photophysics and photochemistry within the context of single-molecule spectroscopy. We begin with the theory of photon absorption and emission and then proceed to outline the various competing pathways to fluorescence. We discuss triplet states and intermolecular quenching in relation to fluorophore blinking, and we describe general photoswitching mechanisms. Finally, we present the theory of Förster resonance energy transfer (FRET) and motivate its application in studying single-molecule reaction dynamics. This chapter provides a solid foundation to understand and interpret single-molecule fluorescence data, and we expect the chapter will prove useful in guiding both experimental and analysis method design.

2.1 Introduction

Molecules which are capable of absorbing and emitting photons with wavelengths in the visible spectrum have become important elements in our quest to probe biochemical and cellular processes. Each application presents its own demands and challenges, and determining the appropriate molecular probes has become paramount. Fluorophores, i.e., chemical compounds that can absorb a photon of a specific wavelength and then re-emit a photon at a different but equally spe-

cific wavelength, each possess their own unique photophysical and photochemical characteristics. Some are bigger; others are smaller. Some are water soluble, and others not. Some are incredibly photostable in the presence of oxygen, while others destruct. Each fluorophore has its own mix of spectroscopic parameters, redox properties, and other qualities. While the combinations are endless, finding a fluorophore with the particular properties desired is challenging and demands significant characterization, optimization, and experimental controls.

In single-molecule spectroscopy, we are unable to optically resolve native biological molecules at the single-molecule level due to their small size. To permit fluorescence detection, we must first attach a fluorescent probe to the biomolecule. Ideally, probes must be capable of being linked to the biomolecule such that attachment introduces minimal perturbations and does not affect biological activity. For many applications, the probe should emit a high number of photons and be photostable (i.e., exhibit constant fluorescence intensity and be slow to photobleach) for the entire measurement duration. For other applications, such as single-molecule-based super-resolution imaging, variability in photon emission is desired, and, more specifically, a variability capable of being finely tuned [322, 323].

So what insight can a fluorophore attached to a biomolecule provide?

1. The location of a biomolecule. By monitoring the diffusional and directional movements of molecular motors, we can follow such biomolecules with nanometer precision using sophisticated single-molecule particle tracking algorithms. By introducing multiple fluorophore colors, we can colocalize and distinguish multiple biomolecules within the same complex and observe biomolecular interaction [133].

2. Information about the local environment. Changes in fluorescence lifetimes and intensities can report on local environment changes, such as protein proximity and quencher interaction [193].

3. Rotational dynamics. We can use polarized excitation and emission to monitor a biomolecule's orientational changes [269].

4. Conformational dynamics. Dipole-dipole interactions between adjacent fluorophores enable us to detect small distance changes, which can report on intramolecular structural changes and on intermolecular interactions between multiple molecules [50].

The growth of single-molecule fluorescence spectroscopy as an experimental technique underscores the need to understand the fundamental processes of fluorescence. Building a solid understanding will inform us as to the method's advantages and limitations and help guide our experimental design choices. Importantly, a principle aim in reviewing the theory of molecular photophysics is to help us develop an intuition as to the interpretation of single-molecule fluorescence data. As we will see in the chapters that follow, interpreting single-molecule fluorescence data is neither immediate nor straightforward.

2.2 Photon Absorption and Emission

We can describe a molecule's photophysical and photochemical state in terms of its *electronic state*. We can in turn characterize an electronic state by an electronic configuration, which we may further describe in terms of a molecule's *highest-energy occupied molecular orbital* (HOMO) and *lowest-energy unoccupied molecular orbital* (LUMO) and in terms of a characteristic spin configuration, either singlet (anti-parallel electron spins, $\uparrow\downarrow$) or triplet (parallel electron spins, $\uparrow\uparrow$) [319]. In studying fluorescence, we are not concerned with the absolute energy of a particular electronic state, but with the energy difference between molecular electronic states. We refer to this difference as the *energy gap*

$$\Delta E = |E_j - E_i| \quad \text{The energy gap between } E_j \text{ and } E_i \tag{2.1}$$

where E_j and E_i are the absolute energies for states j and i, respectively.

Now consider that a photon has energy

$$E_{photon} = h\nu \tag{2.2}$$

where h is Planck's constant (1.58×10^{-37} kcal·sec) and ν is the frequency of the photon (typically in units of $s^{-1} = Hz$). Molecular photon absorption transforms the photon energy, $h\nu$, into an electronic excitation energy such that

$$h\nu + |D\rangle \rightarrow |D^*\rangle$$

where, for a molecule D, $|D\rangle$ is the ground electronic state and $|D^*\rangle$ is an electronically excited state. The photon performs work by using its energy to either change electron orbital structure, vibrate nuclei, or precess molecular spins. With the absorbed photon energy, a molecule may make or break chemical bonds, or it may change a molecule's electronic configuration and hence the electronic distribution about the nuclei.

We refer to transitions between states as *photophysical processes*. To model these processes, we turn to quantum mechanics. According to the postulates of quantum mechanics [109, 219], we may compute the value of any observable property P_i of a state i, if the state's wave function Ψ_i and the mathematical operator $\hat{P}_i$ corresponding to the observable property are known.[1] For example, to measure

[1] The only observed measurement values possible must be eigenvalues that satisfy the eigenvalue equation

$$\hat{P}\Psi = P\Psi \tag{2.3}$$

Hence, for every measurable property P of a molecular system, a mathematical function $\hat{P}$ exists that operates on the eigenfunction Ψ to produce an eigenvalue P that corresponds to an experimental measurement of that system property [109, 219]. Equation 2.3 refers to a single measurement of a single molecule in an experiment. For a large number of experiments, we can calculate the property's average (or *expectation*) value by mathematical integration

$$\underbrace{\langle P \rangle}_{\text{Expectation Value}} = \int \Psi^* \hat{P} \Psi d\tau = \underbrace{\langle \Psi | \hat{P} | \Psi \rangle}_{\text{Matrix Element}} \tag{2.4}$$

where $\langle \Psi | \cdot | \Psi \rangle$ corresponds to the alternative notation to wave mechanics called *bra-ket notation*. In words, we extract information about a molecular system, i.e., the expectation value of an experimental measurement, by "acting" on a normalized wave function Ψ with an operator $\hat{P}$ and by computing the matrix element.

the energy we use the Hamiltonian operator $\hat{H}$ and its eigenvalue equation

$$\hat{H}\Psi = E\Psi$$

and solve for E. But suppose we are interested in the transition rates between an initial state Ψ_i and a second state Ψ_j. If we know the wave functions Ψ_i and Ψ_j, the postulates of quantum mechanics allow us to calculate the transition rate constant k, i.e., the transition probability per unit time, for $\Psi_i \rightarrow \Psi_j$ by computing the square of a matrix element that corresponds to the transition, if $\hat{P}_{i\rightarrow j}$, the operator for the interaction which triggers $\Psi_i \rightarrow \Psi_j$, is known.

$$k_{i\rightarrow j} \sim \langle \Psi_j | \hat{P}_{i\rightarrow j} | \Psi_i \rangle^2$$

where "$\sim$" indicates the omission of constants. Hence, for the transition $h\nu + |D\rangle \rightarrow |D^*\rangle$, $\hat{P}_{h\nu}$ is the operator which represents the interaction of D electrons with a photon, and thus,

$$k_{0\rightarrow 1} \sim \langle \Psi_1(D^*) | \hat{P}_{h\nu} | \Psi_0(D) \rangle^2$$

The interaction represented by $\hat{P}_{h\nu}$ causes wave function 'mixing' [319], where effective mixing of two waves only occurs upon *resonance* between $\Psi_0(D)$ and $\Psi_1(D^*)$. For example, if we consider $h\nu + |D\rangle \rightarrow |D^*\rangle$, we can imagine that the photon carries the energy and required interaction needed to achieve resonance and to cause the electronic transition from D to D^*. Expressed differently, the operator $\hat{P}_{h\nu}$ transforms the wave function $\Psi_0(D)$; if the transformed $\Psi_0(D)$ is sufficiently similar to $\Psi_1(D^*)$, a transition can occur.

Calculating molecular wave functions Ψ_n is not possible except for the simplest molecular systems. Instead, we rely on the *Born-Oppenheimer approximation*, which assumes that electron motion is much more rapid than nuclear vibration and that the distribution of low-mass, rapidly moving, negatively charged electrons instantaneously adjusts to the positive potential of relatively massive, slowly moving

nuclei.[2] As a consequence, an electronic transition between orbitals occurs while nuclei are effectively stationary. Using the Born-Oppenheimer approximation, we can approximate Ψ in terms of three separate components: an electronic wave function, a vibrational wave function, and a spin wave function.

$$\underbrace{\Psi}_{\text{"True" molecular wave function}} \approx \underbrace{\Psi_0 \cdot \chi \cdot \mathbf{S}}_{\text{electronic, vibrational, spin}} \tag{2.5}$$

The electronic wave function Ψ_0 is a zero-order approximation for electronic orbital motion and position about a stationary and positively charged nucleus. We can further approximate Ψ_0 by neglecting electron-electron interactions. If we calculate a solution to the wave function ϕ_i for a one-electron molecule, we can then approximate the molecular wave function as a product of multiple overlapping one-electron orbitals

$$\Psi_0 \approx \phi_1 \phi_2 \cdots \phi_m \; ; \quad i = 1, 2, \ldots, m \tag{2.6}$$

Along with these approximations, we can subsequently approximate the interaction represented by $\hat{P}_{h\nu}$ by only considering HOMO (HO) and LUMO (LU) molecular orbitals [332]

$$k_{0\to 1} \sim \left\langle (\phi_{HO}(D^*)\phi_{LU}(D^*)) \cdot \chi(D^*) \cdot \mathbf{S}(D^*) \middle| \hat{P}_{h\nu} \middle| (\phi_{HO}(D)\phi_{HO}(D)) \cdot \chi(D) \cdot \mathbf{S}(D) \right\rangle^2$$

The HOMO and LUMO molecular orbital approximation provides a sufficient starting point for most photochemical reactions [319].

Provided we have selected a good zero-order approximation Ψ_0, we can introduce minor refinements to our approximation. For instance, we may "mix" one-electron wave functions, allowing electron-electron interactions, such as electron-electron repulsion. If the mixing effect is small, i.e., we have a good Ψ_0, then we

[2] Two notes: (1) a proton's mass is $\sim 10^3$ times greater than electron's mass; and (2) an electron transition (or jump) is two orders of magnitude faster than nuclear vibrations: 10^{-15}-10^{-16} versus 10^{-13}-10^{-14}.

call such an effect a *perturbation*. The mathematical theory which uses such perturbations to better approximate the true system is called *perturbation theory*.

Importantly, weak perturbations[3], such as weak electronic, vibrational, or spin wave function "distortions", include electron interaction with electromagnetic fields, as well as HOMO-LUMO interaction between molecules, i.e., the interactions which lead to chemical reactions and energy transfer [319]. Hence, we can account for transitions of the type $h\nu + |D\rangle \rightarrow |D^*\rangle$ by weak perturbations, and we can calculate the transition rate from one energy eigenstate to accessible eigenstates which are capable of resonance through the perturbation $P'_{h\nu}$ using *Fermi's golden rule*

$$k_{0\rightarrow 1} = \frac{2\pi}{\hbar}\rho[\langle \Psi_1(D^*)|P'_{h\nu}|\Psi_0(D)\rangle]^2 \tag{2.7}$$

where ρ is the density of final states (number of states per unit energy) capable of effective mixing between $\Psi_0(D)$ and $\Psi_1(D^*)$ and $\hbar$ is the reduced Planck constant ($= h/2\pi$). In Equation 2.7, the perturbation relates to the transition dipole moment. As electrons interact with the electromagnetic field, they oscillate along the molecule's positive nuclear framework and distort the electron cloud. If the oscillation is large, i.e., the interaction generates a large transition dipole, the perturbation is strong, and the transition rate is fast.

For organic molecules, vibronic coupling, i.e., vibrational nuclear motion coupled to electron orbital motion, is the most important perturbation for mixing electronic wave functions [319]. Let us then denote the vibronic coupling operator $\hat{P}_{vib}$ to represent the perturbation which vibrationally mixes the electronic wave functions $\psi_0(D)$ and $\psi_1(D^*)$. When mixing sufficiently distorts $\psi_0(D)$ to look like $\psi_1(D^*)$, an electronic transition can occur. We can address the density of states ρ to which the molecule can transition by considering the overlap integral $\langle\cdot|\cdot\rangle$ of the vibrational wave functions $\chi_0(D)$ and $\chi_1(D^*)$, i.e., the extent of similarity in nuclear configuration and vibrational momentum at the instant of an electronic tran-

[3] A weak perturbation is a perturbation which does not significantly alter the energies associated with Ψ_0.

sition. The square of the vibrational overlap $\langle \chi_1 | \chi_0 \rangle^2$ is called the *Franck-Condon (FC) factor*.

$$k_{0\to 1} \sim \underbrace{\langle \psi_1(D^*) | \hat{P}'_{vib} | \psi_0(D) \rangle^2}_{\text{vibronic coupling}} \cdot \underbrace{\langle \chi_1(D^*) | \chi_0(D) \rangle^2}_{\text{FC factor}}$$

Accordingly, the *Franck-Condon principle* states that the most probable electronic state transitions occur when the initial vibrational state wave function χ_i most closely resembles the final vibrational state wave function χ_f. Figure 2.1 provides a visual representation of FC factors.[4]

The Born-Oppenheimer approximation indicates that the nuclear configuration in the excited state $|D^*\rangle$ is approximately the same as in the ground state $|D\rangle$ at the instant of electronic transition. This explains why electronic transitions are vertical arrows in Figure 2.1. Once an electron has transitioned, the nucleus experiences a new negative electronic force field which causes the nucleus to vibrate as it adopts a new nuclear geometry. The additional vibrational energy means zero-point transitions, i.e., direct transitions from the lowest ground vibronic state to the lowest vibrational level of an excited electronic state, do not typically occur. Instead, electronic transitions to excited states predominantly end at higher vibrational levels, and excited electronic state minima are located at increased internuclear separation [219].

Intramolecular *vibrational relaxation*, i.e., a non-radiative transition from a higher vibrational level to a lower vibrational level, occurs when the excess vibrational energy due to electronic transition transfers to nearby molecules. This form of

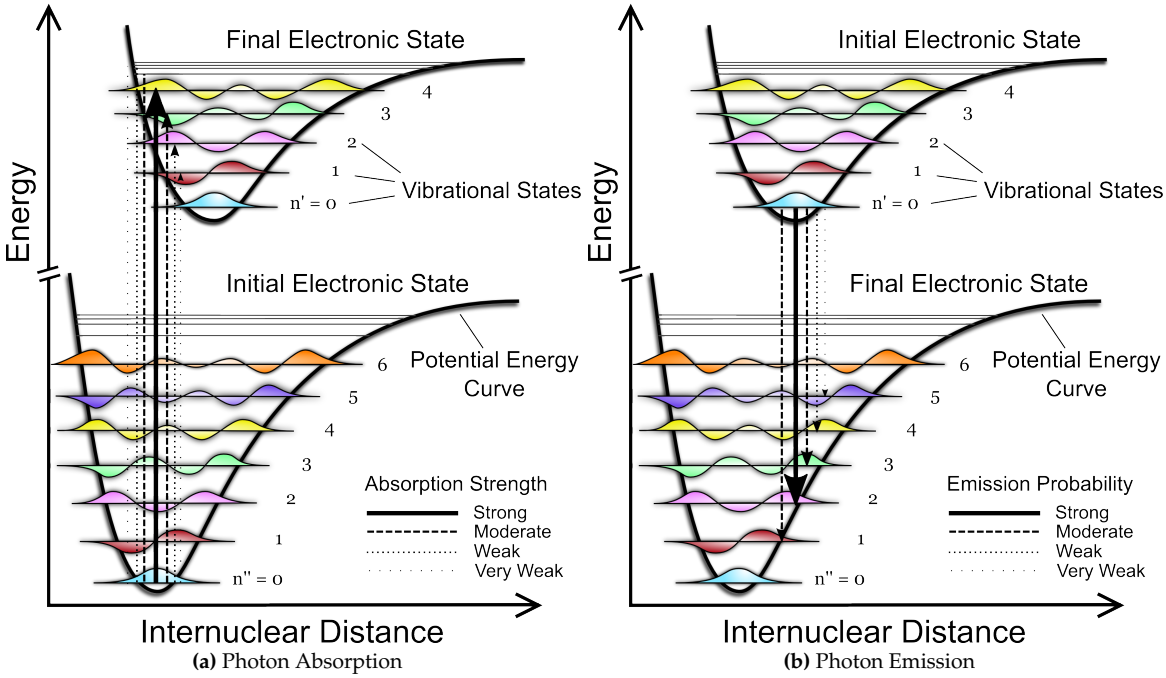
[4]Note that we only show transitions which begin from the lowest vibrational level of the ground electronic state. The energy gap between vibrational states is approximately 0.1eV. This energy is large compared to the thermal energy $k_B T \approx 0.025$ eV, where k_B is the Boltzmann constant and $T = 296$ K is room temperature in degrees Kelvin. We can describe the distribution of molecules over available vibrational levels by a *Boltzmann distribution* [278]

$$N_i \sim e^{-E_i/k_B T} \tag{2.8}$$

where N_i is the number of molecules with energy E_i. We can calculate the fraction of molecules in the ith energy level

$$f_i = \frac{N_i}{N} = \frac{e^{-E_i/k_B T}}{\sum_i e^{-E_i/k_B T}}$$

where $\sum_i N_i = N$. For $T = 296$ K, $f_i > 0.99$, meaning that nearly all molecules at room temperature are in the ground vibrational state.



(a) Photon Absorption **(b)** Photon Emission

Figure 2.1: Quantum Mechanical Representations of the Franck-Condon Principle for Photon Absorption and Photon Emission. (a) potential energy curves for initial and final electronic states. Transitions to excited electronic states predominantly begin from the lowest vibrational level. The probability of ending in a particular vibrational level of the final electronic state is proportional to the overlap integral between vibrational wave functions (shown in color). Here, we indicate the most probable transition with a thick black line. (b) similar to photon absorption in (a), photon emission also follows the Franck-Condon principle. Transitions to lower excited states predominantly begin from the lowest vibrational level of the excited electronic state. Excited electronic state internuclear separations frequently differ from the ground state due to nuclear adjustment to the excited state orbital configuration; the shift in potential energy minimum thus favors transitions to ground state vibrational levels greater than 0. The vertical lines represent $n'' \rightarrow n'$ and $n' \rightarrow n''$ vibronic transitions. We omit rotational states for clarity.

relaxation, on the order of 10^{-11}-10^{-13} seconds, is followed by a slower intermolecular vibrational relaxation, which occurs on the order of 10^{-10}-10^{-11} seconds. In solution, solvent molecules surround the ground state molecule to create a solvation shell in which dipole moment interaction induces an ordered distribution of solvent molecules. The transition dipole moment of the vibrationally excited molecule affects the surrounding solvent via van der Waals (London dispersive) forces, hydrogen bond rearrangement, polarization changes, and short-range repulsion between molecular electronic clouds; these effects cause the solvent molecules to re-orient (termed *solvent relaxation*) within the solvent-solute environment and allow for thermal equilibration.

Solvent acceptance of the excess vibrational energy allows the excited molecule to further relax to reach the lowest vibrational level of the excited electronic state. Accordingly, intra- and intermolecular relaxation combine to reduce the energy separation between the ground and excited electronic states. As solvent polarity increases, this produces a correspondingly larger reduction in excited state electronic energy, while a decrease produces the opposite effect. The same relation is true for the excited molecule: polar and charged excited molecules exert a greater effect on the solvent than non-polar molecules. Hence, a molecule's polarity determines the sensitivity of the excited state to solvent effects.

We can characterize a substance's propensity to undergo electronic transitions due to photon absorption based on two laws: Lambert's and Beer's Laws. *Lambert's Law* states that the proportion of photons absorbed by a medium is independent of the initial photon intensity I_0. *Beer's Law* states that the number of absorbed photons is proportional to the concentration of absorbing molecules in the photon path. Accordingly, we can combine both laws in a single equation called the *Beer-Lambert Law* [189]

$$\frac{I}{I_0} = 10^{-\varepsilon(\nu)Cl} \tag{2.9}$$

where I is the transmitted photon intensity, l is the optical path length (usually 1 cm), C is the sample concentration, and $\varepsilon(\nu)$ (with units $\text{cm}^2/\text{mol} = \text{cm}^{-1}/\text{M}$) is

the molar extinction coefficient as a function of photon energy ν. The coefficient ε is a fundamental molecular property and is the quantity which characterizes a substance's ability to absorb photons of energy ν.

Once a molecule transitions to an excited electronic state, vibrational relaxation is one of many competing pathways. One pathway, *photon emission*, is essentially the reverse of photon absorption: radiative emission (also called *fluorescence*) between an excited electronic state and a ground electronic state (see Figure 2.1(b) for a Franck-Condon principle interpretation). Once again, nuclei must adjust to the new (ground-state) electronic force field, causing excess vibrational energy. The excess energy dissipates, and the molecule vibrationally relaxes to the lowest energy level. Importantly, the energy gaps for the most likely electronic transitions from an excited state to a ground state are smaller than the energy gaps from a ground state to an excited state. We refer to the energy shift from higher energy photon absorption to lower energy photon emission as the *Stokes Shift*.

Figure 2.2 provides examples of the Stokes shift in fluorophore absorption and emission spectra. A *fluorophore*, i.e., a chemical compound that can emit a photon upon photon excitation, typically contains several combined aromatic groups or conjugated double bonds. For fluorophores Cy3B and ATTO647N, two fluorophores which are common in single-molecule fluorescence, conjugated double bonds create a de-localized π electron system. Conjugation reduces the energy gap [332] and permits absorption of photons with resonance energies (and hence, wavelengths) in the visible light spectrum; for chemical structures, see Figure 2.3.

One feature to note in absorption and emission spectra is that the spectra are smooth and continuous. Based on the resonance equation $\Delta E = h\nu$, where ΔE corresponds to the energy gap between electronic ground and excited states, we might expect discrete, or quantized, photon absorption energies matching the resonance condition. Such quantization would be manifest as sharp lines. We should, however, remember that electron motion is coupled to vibrational and rotational motion. The Franck-Condon diagram (Figure 2.1) shows electronic and

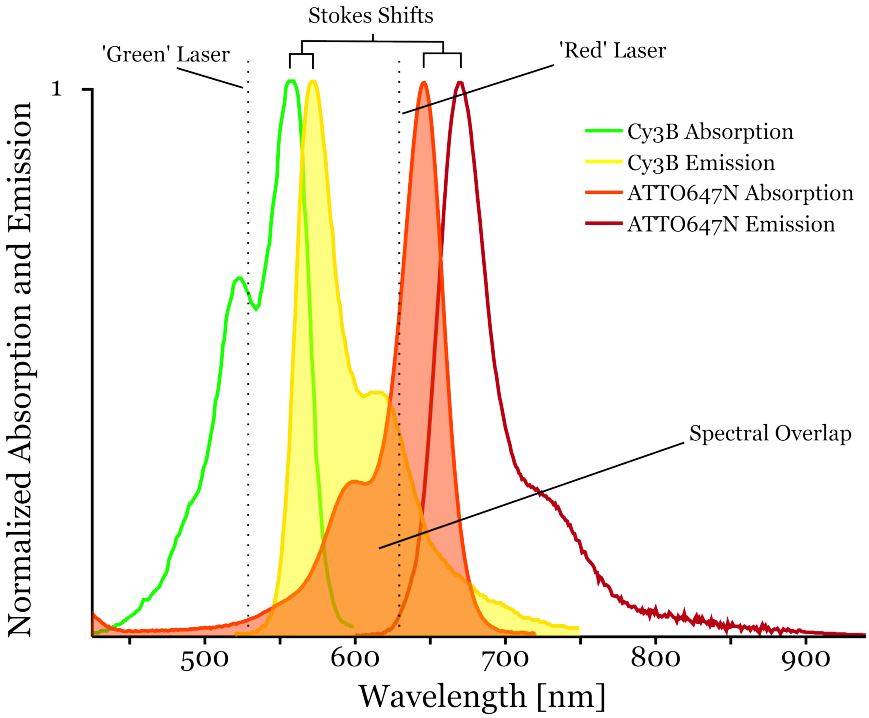
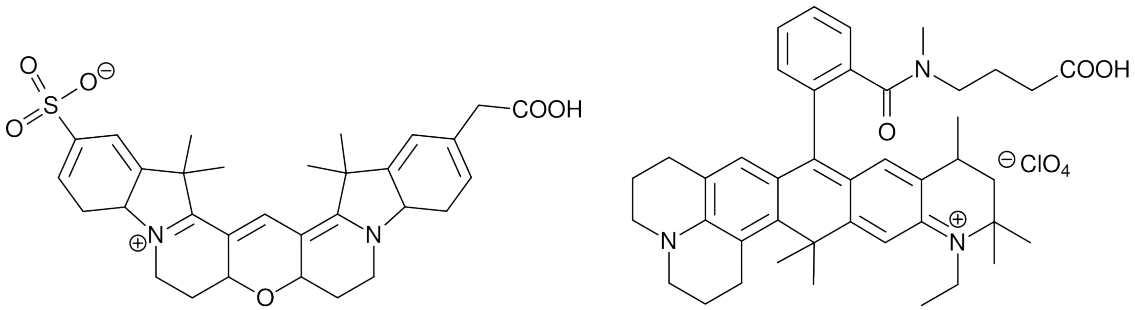


Figure 2.2: Absorption and Emission Spectra for Fluorophores Cy3B and ATTO647N. Relative to Cy3B, ATTO647N exhibits a greater Stokes shift. This may be due to ATTO647N's amphiphilicity, i.e., having a hydrophilic head and lipophilic tail, and its exerting London dispersive forces which induce greater reorganization of a more rigid solvation shell; see Figure 2.3 and main text. We highlight the significant spectral overlap between Cy3B emission and ATTO647N absorption spectra and further include laser excitation wavelengths ($\lambda_{green} = 532$nm and $\lambda_{red} = 638$nm) for reference.



(a) Cy3B **(b)** ATTO647N

Figure 2.3: Chemical Structures for Cy3B [49] and ATTO647N [75, 180].

vibrational states, but does not show that, for each vibrational state, many rotational, i.e., quantized angular momentum, states are possible.[5] The transition between $|D\rangle$ and $|D^*\rangle$ does not correspond to a single quantum energy, but instead, a range of energies are available due to an ensemble of possible nuclear configurations. Additionally, the solvation shell surrounding a molecule in solution may occupy one of many supramolecular configurations which are weakly coupled to the molecule. Each coupling will correspond to a slightly different energy gap for absorption or emission. The combination of ro-vibronic coupling and solvation effects means that the number of possible energy gaps trends toward infinity, resulting in a broadened, mostly featureless band with one, or possibly two, maxima.

2.3 Fluorophore Blinking

Photon emission is, however, not the only pathway which allows an excited molecule to return to its ground electronic state. Excited molecules may 'relax' by (i) emitting a photon, a process we refer to as *fluorescence*, (ii) participating in photochemical reactions to produce a reactive intermediate ($S_1 \rightarrow {}^1I$), (iii) undergoing a spin-allowed non-radiative transition termed *internal conversion* between states of the same spin ($S_1 \rightarrow S_0 + \text{heat}$), (iv) non-radiatively transferring energy to another molecule via *electronic exchange interactions*, and (v) undergoing a spin-forbidden non-radiative transition termed *intersystem crossing* between excited states of a different spin ($S_1 \rightarrow T_1 + \text{heat}$). Further pathways are possible once the excited molecule transitions to available triplet states T_i. The molecule may undergo intersystem crossing to the ground state ($T_1 \rightarrow S_0 + \text{heat}$), emit a photon in a process termed *phosphorescence*, or participate in a photochemical reaction to produce a reactive intermediate ($T_i \rightarrow {}^3I$). Each pathway has a characteristic rate constant

[5] Rotational quantum states have energies of approximately 0.001 eV. These energies are small compared to the thermal energy $k_BT \approx 0.025$ eV, where k_B is the Boltzmann constant and $T = 296$ K is room temperature in degrees Kelvin. Similar to vibrational states, we can describe rotational state occupancy by a Boltzmann distribution [219]. Consequently, a significant fraction of molecules will occupy rotational states higher than the lowest rotational state.

Process	Transition	Rate Constant	Rate [s^{-1}]
Singlet-Singlet Absorption	$S_0 \rightarrow S_i$	k_{exc}	10^{15}
Internal Conversion	$S_i \rightarrow S_1, T_i \rightarrow T_1$	k_{IC}	10^{14}-10^{10}
Vibrational Relaxation	$S_{1,n'=j} \rightarrow S_{1,n'=0}$	k_{vr}	10^{12}-10^{10}
Fluorescence	$S_1 \rightarrow S_0$	k_{fl}	10^{9}-10^{7}
Internal Conversion	$S_1 \rightarrow S_0$	k_{IC}	10^{7}-10^{6}
Quenching	$S_1 \rightarrow S_0$	k_Q	10^{7}-10^{5}
Intersystem Crossing	$S_1 \rightarrow T_1, S_i \rightarrow T_i$	k_{ST}	10^{10}-10^{5}
Triplet-Triplet Absorption	$T_1 \rightarrow T_i$	k_{exc}	10^{15}
Phosphorescence	$T_1 \rightarrow S_0$	k_{ph}	10^{3}-10^{-2}
Intersystem Crossing/Quenching	$T_1 \rightarrow S_0$	k_{TS}	10^{3}-10^{-2}

Table 2.1: Photophysical and Photochemical Processes and Their Associated Rates [189].

k_x, where the subscript x denotes the process. We summarize the main pathways using the *state energy diagram* in Figure 2.4 and provide their timescales in Table 2.1.

Before proceeding, we note a few characteristics of the state energy diagram (also called a *Jablonski diagram*). First, we omit electronic states higher than S_1 and T_1. Molecules excited to higher electronic states rapidly internally convert to S_1 or T_1 and further vibrationally relax to the lowest vibrational level. As a consequence, we only experimentally observe photon emission from either thermally equilibrated S_1 and T_1 states. This is known as *Kasha's rule*, which states that emitted photon energy is independent of excitation photon energy. Second, in contrast to the potential energy curves in Figure 2.1, state energy diagrams assume similar equilibrium geometries corresponding to the potential energy minima of $|D\rangle$ and $|D^*\rangle$. Formally, all species, i.e., energetic states, in the diagram are *isomers*: each has the same atomic composition and connectivity but each is chemically different from the other. For example, S_0 and S_1 are electronic isomers; whereas, S_1 and T_1 are spin-electronic isomers. The unique chemical nature of each isomer has important implications for reactivity and for an isomer's participation in pathways affecting fluorescence. We address these pathways next and emphasize their role in the phenomenon of *fluorophore blinking*, i.e., the temporary absence of fluorescence over a timescale greater than the excited state lifetime.

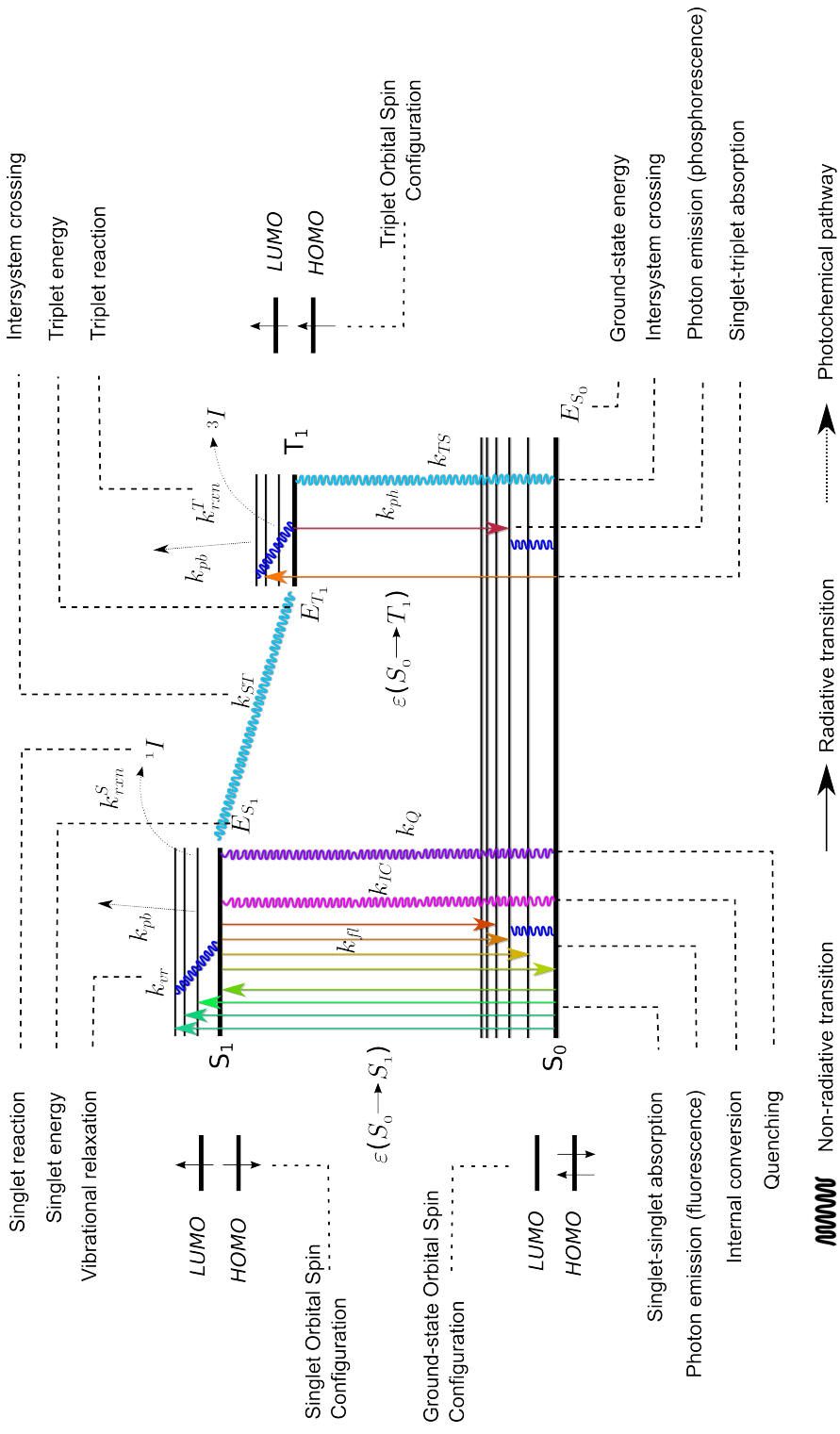


Figure 2.4: State Energy Diagram for Photophysical and Photochemical Processes.

2.3.1 Triplet States

Recall our previous discussion concerning vibronic coupling. We introduced an operator $\hat{P}'_{vib}$ to represent the perturbation which vibrationally mixed electronic wave functions ψ_i and ψ_j, The perturbation caused ψ_i to distort and appear like ψ_j, and we used Franck-Condon factors to determine the most probable vibronic transitions.

A *triplet* state refers to two interacting orbitally unpaired electrons with "parallel" spins. We can explain the scenario in which transitions involve a change in spin ($\mathbf{S_i} \neq \mathbf{S_j}$) similar to how we explained vibronic coupling. Here, the most important perturbation is *spin-orbit coupling*. By distorting the spin, i.e., twisting or flipping it, anti-parallel spins appear parallel, and vibronic coupling to the triplet state is thus "allowed". If we introduce the operator $\hat{P}'_{so}$ to represent the spin-orbit perturbation, we can modify the transition rate for vibronic coupling (see Frank Condon principle) such that

$$k_{S_1 \to T_1} \sim \underbrace{\langle \psi_{T_1} | \hat{P}'_{vib} | \psi_{S_1} \rangle^2}_{\text{vibronic coupling}} \cdot \underbrace{\langle \psi_{T_1} | \hat{P}'_{so} | \psi_{S_1} \rangle^2}_{\text{spin-orbit coupling}} \cdot \underbrace{\langle \chi_{T_1} | \chi_{S_1} \rangle^2}_{\text{FC factor}}$$

The additional spin distortion effectively introduces a transition prohibition, hence reducing the transition rate compared to spin allowed electronic transitions.

For atomic electron configuration, *Hund's rule* states that electrons will occupy different orbitals rather than pair in the same orbital when orbitals are of equal energy. We can similarly apply Hund's rule to molecular orbitals in which electrons first half-fill every orbital of equal energy before spin pairing. According to the *Pauli exclusion principle*, all electronic wave functions must be antisymmetric under pairwise exchange of identical electrons, which means that electrons having the same quantum numbers $n, l,$ and m_l must have opposite spins m_s, i.e., must be spin paired. Because electrons with parallel spins are forbidden to occupy the same space at the same time, these electrons must correlate their motions and positions in space. Paired electrons, however, can occupy the same region of

space, but they experience electron-electron repulsion forces due to their negative charge. As triplet state electrons minimize electron-electron repulsions, the triplet state has a reduced electronic energy. The higher energy requirement for singlet states means that, for molecules possessing two half-filled orbitals, one a HOMO and the other a LUMO, the triplet state is always of lower electronic energy than a corresponding singlet state, i.e., with the same $(\text{HOMO})^1(\text{LUMO})^1$ configuration.

Because spin-forbidden transition rates for organic fluorophores to and from the triplet state are typically three orders of magnitude smaller than fluorescence rates (see Table 2.1), intersystem crossing to a triplet state leads to a temporary absence in observed fluorescence.[6] External perturbations can enhance intersystem crossing rates. For example, molecular oxygen has two unpaired electrons [8] and predominantly resides in a energetically low-lying triplet state. The highly reactive triplet oxygen is an efficient quencher of excited singlet states, in which the mechanism may be one of several possibilities, including spin-orbit coupling, charge transfer, and energy transfer via electron exchange [29, 176, 249, 286]. Oxygen quenching efficiency depends on oxygen concentration, such that we may express the observed intersystem crossing rate as

$$k_{ST}^{obs} = k_{ST} + k_{ST}^{O_2}[O_2]$$

where $k_{ST}^{O_2}$ is a bimolecular rate constant for oxygen perturbation and $[O_2]$ is the oxygen concentration [189]. Typical values of $k_{ST}^{O_2}$ are 10^{10}-10^9 M^{-1} s^{-1} [298]. Given that the O_2 solubility in many organic solvents is 10^{-2} M [319], $k_{ST}^{O_2}[O_2] \approx 10^8$-$10^7$ s^{-1}, and hence, the effect of molecular oxygen on intersystem crossing becomes significant when $k_{ST} \leq 10^8$ s^{-1}. Interestingly, however, oxygen also effectively quenches the excited triplet state such that $k_{TS}^{O_2} > k_{ST}^{O_2}$. The net result is a decrease in the triplet population. Other intersystem crossing enhancers include molecules which, like molecular oxygen, have energetically low lying triplet states and often

[6] Reverse intersystem crossing ($T_1 \rightarrow S_1$) is possible and follows the Franck-Condon principle, but the transition can occur only from sparsely populated higher vibrational levels in T_1. Eggeling and co-workers found this transition to be very efficient for Cy5-labeled DNA [76].

contain heavy ions such as iodide [142].

For applications which require stable fluorescence and hence minimal blinking, reagents such as Trolox[7], β-mercaptoethanol, and β-mercaptoethylamine (MEA) are effective triplet state quenchers [51,262]. Other works have found that combining reducing and oxidizing reagents can also selectively quench triplet states [331]. The efficacy of these additives, however, is highly variable and depends on fluorophore structure and local environment. The search for fluorophores which are unlikely to transition to the triplet state, as well as reagents which quench triplet states is still ongoing.

2.3.2 Quenching

We have already discussed quenchers within the context of triplet states, but we can speak more generally about competing processes that induce non-radiative relaxation where the processes involved are intermolecular in nature. The unifying theme of intermolecular quenching processes is that they compete with fluorescence and can significantly reduce the photon emission rate. We can separate quenching into roughly three categories: collisional processes, static complex formation, and long-range energy transfer mechanisms.

Collisional (or dynamic) quenching occurs when a molecule collides with an excited state fluorophore and causes fluorophore relaxation to the ground electronic state. This type of quenching does not chemically alter either molecule. Molecular oxygen is one example of a collisional quencher, but other elements and compounds such as halogens, amines, and electron-deficient molecules can serve as quenchers, as well. Dynamic quenching mechanisms can include electron transfer, spin-orbit coupling, and intersystem crossing. For example, electron transfer can occur in the contact-mediated fluorophore-guanosine residue interaction which is possible when fluorophores with low reduction potentials, e.g., ATTO655, are attached to DNA and positioned near guanosine residues [69,70,113,127,188]. Tryp-

[7] 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid, a water soluble vitamin E analog

tophan quenches fluorophore fluorescence in a similar fashion when fluorophores are located within a protein environment [128].

Static quenching occurs when fluorophore-quencher complexes form which exhibit either limited or shifted absorption and emission [27]. Static quenching reduces the population of fluorescently active molecules and thus results in reduced observed fluorescence. Recent work implicated this type of quenching between fluorophores Cy3B and ATTO647N when both were closely separated on double-stranded DNA [50].

Long-range *energy transfer* mechanisms[8] can occur via a non-radiative dipole-dipole interaction between a fluorophore and another chromophore in close proximity. This type of interaction is strongly distance dependent and is most efficient for distances shorter than 10 nm. We will discuss long-range energy transfer at length below.

2.3.3 Photoswitching

While applications which require strong and stable fluorescence may seek to limit fluorophore blinking, recent work in *super-resolution microscopy* exploited the phenomenon to achieve optical resolution below the diffraction-limit [126, 271]. By controlling the factors which facilitate blinking, researchers turned fluorophores into optical switches [12, 13, 74, 124, 125, 156, 213, 324, 329]. As such, under a given set of experimental conditions, the fluorophores turn 'on' and 'off', i.e., entering a state from which we detect fluorescence and a different state from which we do not detect fluorescence, in a predictable manner, and we refer to this reversible modulation as *photoswitching*.

Photoswitching mechanisms can be either photochemical, e.g., reduction-oxidation reactions from excited electronic states [329], or light-induced transformations (also known as *photochromism*), e.g., cis-trans isomerizations [61]. The mechanisms drive

[8] We use the term 'long-range' to differentiate this mechanism from electron transfer which is also an energy transfer process, but occurs between molecules typically separated within 1 nm or less.

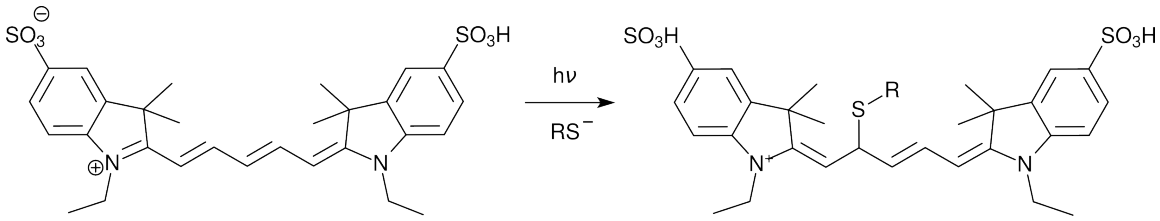


Figure 2.5: Schematic for Proposed Cy5 Photoswitching Mechanism [61]. Cy5 is irradiated with red laser light $h\nu$ and reacts with a thiol containing reducing agent RS^-. The introduction of a thiol disrupts the fully conjugated polymethine bridge connecting the two aromatic rings.

a molecule into a form which either has different resonance energies or differs in the number of photons it can absorb. While the molecule may still be fluorescent, the photophysical properties can change such that fluorescence is undetected. For instance, an absorption spectrum shift due to a loss in bond conjugation may prevent a molecule from absorbing at a particular excitation wavelength.

One example of photoswitching occurs in rhodamine compounds where a lactam ring opens and closes [179]. When the ring is closed, a fluorophore absorbs in the UV spectrum and does not fluoresce. Upon UV irradiation, the lactam ring can open. Ring opening allows the fluorophore to form a metastable state which absorbs in the visible spectrum. Visible-light irradiation or thermal relaxation allows the lactam ring to reform, and the cycle can then repeat. Another example is the cyclic photoswitching of carbocyanine fluorophore Cy5 [12,13,48,61,124,126,323]. Carbocyanine fluorophores undergo photo-induced cis-trans isomerization which leads to blinking on the microsecond timescale [142,143,145,159,331,340,344,345,356].[9] Red laser light, thiol-containing reducing agents, such as MEA, and the absence of oxygen cause Cy5 to enter an inactive state; see Figure 2.5. Green light irradiation and thiol removal recovers the active Cy5 state and Cy5 fluorescence. A third example is reversible switching of oxazines, e.g., ATTO655 [125,323,329,330]. Red laser excitation, a reducing agent, and molecular oxygen cause ATTO655 inactivation. Subsequent oxidation recovers the active form of ATTO655.

Most photoswitching mechanisms appear to be variations of a general reaction

[9] An exception is the fluorophore Cy3B, which differs from Cy3 by having a rigid backbone (via a tetrohydropyranyl) to prevent isomerization [49]. Cy3B is not known to exhibit light-dependent photoactivation.

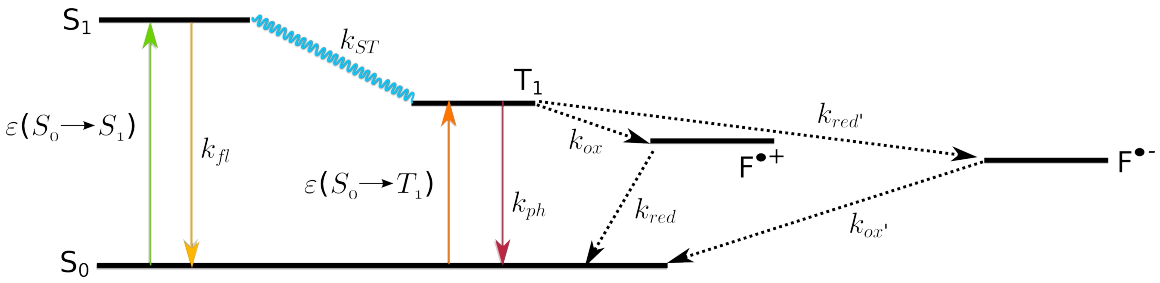


Figure 2.6: Extended State Energy Diagram for a General Photoswitching Mechanism. After intersystem crossing, oxidation/reduction to a radical ion state $F^{\bullet+}/F^{\bullet-}$ results in an absence of fluorescence. Subsequent reduction/oxidation recovers the ground electronic state.

sequence [113,125,202,296,329–331]; see Figure 2.6 for a state-energy diagram representation. The process involves reduction and oxidation reactions that reversibly switch the fluorophore between one of two states, commonly referred to as 'active' and 'inactive' states. For example, after intersystem crossing from an excited singlet state to an excited triplet state, a reducing agent such as ascorbic acid, Trolox, or MEA reduces the fluorophore to an inactive radical anion. An oxidizing agent such as methylviologen or molecular oxygen subsequently oxidizes the anion and recovers the singlet ground electronic state. Reducing and oxidizing agents do not generally react with the ground state, and thus photon absorption is necessary to drive the photoinduced electron transfer reactions. Accordingly, several factors control the active state lifetime: oxidizing agent concentration, reducing agent concentration, photon flux (i.e., laser excitation intensity), and solution pH (via hydrogen ion stabilization of the fluorophore's radical anion state [330]).

2.3.4 Photobleaching

Excited singlet and triplet states tend to be more reactive than the ground electronic state. First, electron cloud polarization is more likely for an excited molecule because the promoted electron is located farther away from the nucleus. A greater likelihood for polarization increases a molecule's reactivity. Second, absorbed resonance energy reduces the activation energy needed to drive photochemical reactions.

If we compare excited singlet and triplet reactivity, we find the triplet state is

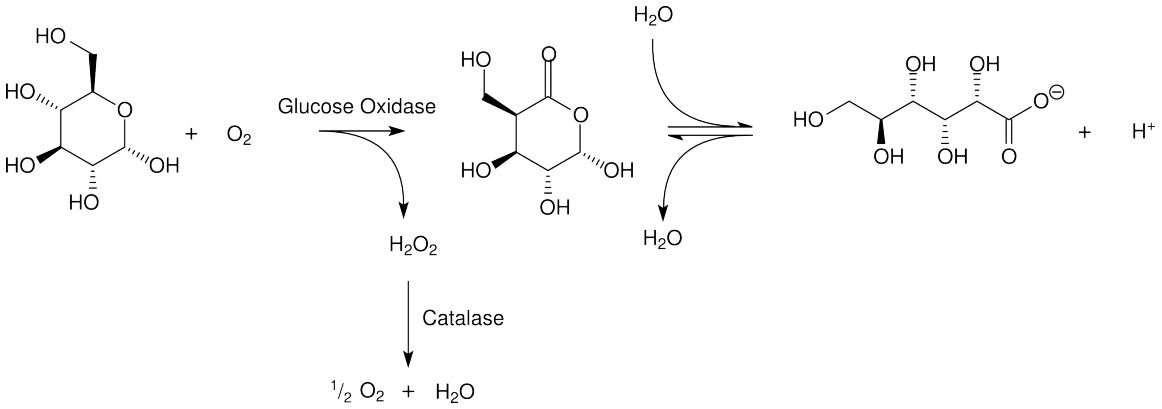


Figure 2.7: Schematic for Glucose Oxidase and Catalase Oxygen-scavenging System Chemical Reactions [2].

generally more reactive. Due to the slow rate of phosphorescence, internal conversion and other non-radiative pathways compete successfully with radiative relaxation of the triplet state. The probability of reactive intermediates is thus higher compared to the excited singlet state where rapid fluorescence depletion of the excited singlet state more effectively competes with reaction pathways producing such intermediates. In both cases, however, a fraction of the intermediates are involved in *photobleaching*, i.e., an irreversible reaction that dramatically changes a molecule's absorption and emission capabilities resulting in a permanent absence of fluorescence.

Recall that molecular oxygen played a role in intersystem crossing and triplet quenching. A by-product of molecular oxygen's role is a highly reactive higher energy singlet oxygen [146, 254, 263]. The close proximity of the by-product in relation to the ground state molecule means the singlet oxygen can easily diffuse and react with the molecule both in its ground state or, after photon absorption, its more reactive excited state within the excited state lifetime [189]. As part of the reaction, the molecule loses an electron due to oxidation and consequently loses its fluorescence capabilities.

Oxygen removal and alternative triplet quencher, e.g., Trolox, addition can improve fluorophore stability in solution. In single-molecule fluorescence, a common method to remove oxygen is an enzymatic oxygen scavenging system. For

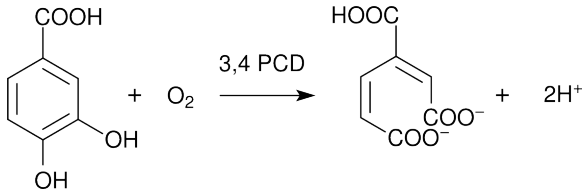


Figure 2.8: Schematic for protocatechuic acid (PCA)/protocatechuate-3,4-dioxygenase (PCD) Oxygen-scavenging System Chemical Reaction [2].

example, glucose oxidase and catalase (GOC) catalyze the reaction of oxygen and glucose to gluconic acid and water [2]; see Figure 2.7. Another system is a protocatechuic acid (PCA)/protocatechuate-3,4-dioxygenase (PCD) oxygen-scavenging system; see Figure 2.8. Note that both the GOC and PCA/PCD system produce hydrogen ions which will lower the solution pH [290]. A drop in pH can have a strong influence on fluorophore photophysics. A recent work provides an alternative enzymatic system promising pH stability, but its widespread applicability remains to be proven [305]. Microfluidics and nitrogen degassing provide alternatives to enzymatic oxygen scavenging systems [68], but their implementation is more involved and requires significant setup and optimization.

Excitation photon intensity further determines photobleaching probability [76]. As a molecule spends more time in a reactive excited state, the molecule has increased probability of entering a photobleaching pathway. Moreover, excitation wavelength and excitation scheme also have an effect on photobleaching rates. Both simultaneous laser excitation and rapid laser alternation increase the probability of higher excited states and photoionization [181].[10] While the exact mechanisms for photobleaching remain elusive, pathways often appear to involve several intermediates, and the triplet state seems to be the primary starting point for a majority of the proposed pathways [114]. Accordingly, an efficient triplet state quencher is vital for applications which require extended observation periods and stable fluorescence.

[10]Note that laser excitation coupled with energy transfer processes can also increase the probability of higher excited states. Under donor laser excitation, especially if in close proximity to a donor fluorophore, FRET acceptor molecules have a greater probability of higher excited states and, hence, have a decreased photobleaching lifetime $\tau_{ph} = 1/k_{ph}$.

2.4 Förster Resonance Energy Transfer

Let us first consider the following reaction sequence

$$|D^*\rangle \rightarrow |D\rangle + h\nu \qquad \text{(i)}$$

$$h\nu + |A\rangle \rightarrow |A^*\rangle \qquad \text{(ii)}$$

In this scenario, an excited molecule D^* serves as a photon source and emits a photon which is subsequently absorbed by another molecule A to yield A^*. The probability per unit time, i.e., the rate, of a radiative emission-absorption energy transfer mechanism will depend on several factors:

- $\Phi_{em}^{D^*}$, the *quantum yield* for photon emission from D^*

$$\Phi_{em}^{D^*} = \frac{\text{number of photons emitted}}{\text{total number of photons absorbed}} \tag{2.10}$$

- $\varepsilon_A(\nu)$, the molar extinction coefficient, which characterizes A's ability to absorb a photon of energy ν
- the overlap between the emission spectrum of D^* and the absorption spectrum of A

We may combine the latter two factors to define a *normalized spectral overlap integral*

$$J = \frac{\int_0^\infty f_D(\nu)\varepsilon_A(\nu)\, d\nu}{\int_0^\infty f_D(\nu)\, d\nu} \tag{2.11}$$

where $f_D(\nu)$ is the donor emission spectrum as a function of photon energy ν. For the case of poor overlap, $f_D(\nu)$ will close to zero when $\varepsilon_A(\nu)$ is large and vice versa, and hence $J \rightarrow 0$. For a molecule A which does not readily absorb photons, i.e., ε_A is small, J will be small no matter how well D^* emission and A absorption spectra overlap. Because the values for $\varepsilon_A(S_0 \rightarrow T_n)$, i.e., singlet-triplet absorption, are very small, only singlet-singlet absorption is likely to occur.

But now consider a non-radiative energy transfer reaction

$$|D^*\rangle + |A\rangle \rightarrow |D\rangle + |A^*\rangle \tag{2.12}$$

Instead of photon emission and absorption, the excited states of molecules D and A are somehow coupled. Although not explicitly mentioned or detailed, we have already encountered one such coupling in the form of electron exchange. Electron exchange involves molecular orbital overlap allowing for simultaneous hole-electron transfer [319]. This form of energy transfer is commonly known as *Dexter energy transfer* [67] and occurs when D-A separation is below 20 Å.

A different coupling mechanism not involving molecular orbital overlap or van der Waals contact involves a weak dipole-dipole interaction. This mechanism is known as *Förster resonance energy transfer* (FRET) [47, 88, 89, 91] and, relative to Dexter electron transfer, is the predominant energy transfer mechanism when D-A separation is between 20 Å and 1 μm.[11] FRET is the non-radiative energy transfer of an electronic excitation from one molecule termed a *donor* to another molecule termed an *acceptor*. We can refer to FRET as a *photosensitization reaction*, as the acceptor molecule becomes 'sensitized', not by a photon absorbed while in its original electronic state, but by a photon absorbed by the donor molecule. By definition, a chemical compound, capable of photon emission upon receiving energy from another molecule excited previously in the chemical reaction, is a *sensitized* molecule. Accordingly, we term the donor molecule the *sensitizer*.

Recall our previous discussion (section 2.2) which concerned the interaction between a molecule D and a photon $h\nu$. As the photon approaches D, D's electrons interact with the electromagnetic field and begin to oscillate along the molecule's positive nuclear framework, thus distorting the electron cloud. The strength of this perturbation dictates the magnitude of the transition dipole and the likelihood of electronic excitation.

[11] Beyond 1 μm, the weak dipole-dipole coupling of FRET is negligible, and energy transfer is radiative, i.e., photon emitted and photon absorbed, rather than virtual photon mediated; see [4] for a unifying principle which accounts for these energy transfer mechanisms.

Upon electronic excitation, we may classically model an excited electron as undergoing periodic harmonic oscillations along the molecular framework with a natural frequency ν_0. These oscillations will create an oscillating electric dipole and thus an oscillating electric field in the space around D^*. We can then envision the oscillating electric field having a similar effect to that of a photon's electric field. If the frequency ν_0 matches the energy gap $\Delta E(S_0^A \rightarrow S_1^A)$ for electronic excitation, then ν_0 satisfies the Bohr frequency (i.e., resonance) condition. Moreover, if the intermolecular distance between D^* and A is sufficiently small such that D^* can interact with A's electrons and the perturbations induced by D^* are sufficiently strong, then energy transfer is possible between D^* and A. Several factors therefore determine the efficiency of non-radiative resonance energy transfer:

- ν_0
- R_{DA}, the distance between D^* and A
- μ_D, the strength of the oscillating electric dipole
- θ, the relative orientation of the transition dipole moments $\boldsymbol{\mu}_{D^*}$ and $\boldsymbol{\mu}_{A^*}$
- $\varepsilon_A(\nu_0)$, which, in the classical theory, refers to how readily A's electrons will begin to oscillate

Interestingly, A is unable to distinguish between photon absorption and electron oscillation of another molecule as the cause of its own oscillations. The literature sometimes refers to the perturbation as that of a "virtual photon" [4, 281, 319].

According to the postulates of quantum mechanics, the interaction strength is directly proportional to the matrix element which corresponds to the interaction (see Equation 2.4). The rate of energy transfer k_{ET} is proportional to the square of the strength of the interaction. As both electron exchange and dipole-dipole

interactions are weak interactions, we can use perturbation theory to express k_{ET}

$$k_{ET} \sim \alpha \underbrace{\left\langle \Psi_0(D)\Psi_1(A^*)|\hat{H}'_{ex}|\Psi_1(D^*)\Psi_0(A) \right\rangle^2}_{\text{electron exchange}} + \beta \underbrace{\left\langle \Psi_0(D)\Psi_1(A^*)|\hat{H}'_{dd}|\Psi_1(D^*)\Psi_0(A) \right\rangle^2}_{\text{dipole-dipole interaction}} \tag{2.13}$$

where $\hat{H}'_{ex}$ is the perturbation operator for electron exchange, $\hat{H}'_{dd}$ is the perturbation operator for dipole-dipole energy transfer, and α and β are the relative contributions of the two terms ($\alpha + \beta = 1$). $\hat{H}'_{ex}$ follows an exponential decay[12] $\exp(-R_{DA})$; whereas, $\hat{H}'_{dd}$ represents an electrostatic (Coulombic) interaction potential which we express using a multipole expansion [240]

$$V_{D^*A} \sim \frac{q_{D^*}q_{A^*}}{R_{DA}} + \frac{q_{D^*}\langle \boldsymbol{\mu}_{A^*}, \mathbf{R}_{DA}\rangle}{R_{DA}^3} - \frac{q_{A^*}\langle \boldsymbol{\mu}_{D^*}, \mathbf{R}_{DA}\rangle}{R_{DA}^3} + \frac{R_{DA}^2\langle \boldsymbol{\mu}_{D^*}, \boldsymbol{\mu}_{A^*}\rangle - 3\langle \boldsymbol{\mu}_{D^*}, \mathbf{R}_{DA}\rangle\langle \boldsymbol{\mu}_{A^*}, \mathbf{R}_{DA}\rangle}{R_{DA}^5} + \cdots \tag{2.14}$$

where $\langle \mathbf{x}, \mathbf{y}\rangle$ is the inner product, q_x is the molecular charge, $\boldsymbol{\mu}_x$ is the transition dipole moment, $\mathbf{R}_{xy}$ is the vector beginning at x and ending at y, and $R_{xy} = \|\mathbf{R}_{xy}\|$ is the separation vector norm. We can include additional terms such as quadrupole-charge, quadrupole-dipole, and quadrupole-quadrupole interactions, but these terms are of much shorter range. For neutral molecules, the charge-charge and charge-dipole interactions (first, second, and third terms) are zero. We can express the inner product in terms of its geometric interpretation to obtain

$$\begin{aligned} V_{D^*A} &\sim \frac{R_{DA}^2\|\boldsymbol{\mu}_{D^*}\|\|\boldsymbol{\mu}_{A^*}\|\cos\theta_{DA} - 3R_{DA}^2\|\boldsymbol{\mu}_{D^*}\|\|\boldsymbol{\mu}_{A^*}\|\cos\theta_D\cos\theta_A}{R_{DA}^5} \\ &\sim \frac{\|\boldsymbol{\mu}_{D^*}\|\|\boldsymbol{\mu}_{A^*}\|}{R_{DA}^3}(\cos\theta_{DA} - 3\cos\theta_D\cos\theta_A) \\ &\sim \frac{\kappa\,\|\boldsymbol{\mu}_{D^*}\|\|\boldsymbol{\mu}_{A^*}\|}{R_{DA}^3} \end{aligned} \tag{2.15}$$

[12] electronic wave functions decay exponentially as a function of distance

where θ_{xy} is the angle between the two transition dipole moment vectors, θ_x is the angle between the transition dipole moment vector for x and the intermolecular separation vector $\mathbf{R}_{xy}$, and κ is the *orientation factor*. Provided both the donor and acceptor exhibit high rotational freedom, orientational averaging on the measurement timescale gives $\kappa^2 = 2/3$ [54, 90], with κ^2 having limits of 0 and 4. In the case of perpendicular dipoles, $\kappa^2 = 0$, and, for co-linear dipoles, $\kappa^2 = 4$.

The interaction potential strength thus decays as R_{DA}^3, and hence, as the energy transfer rate is proportional to the square of the interaction strength,

$$k_{ET} \propto \frac{1}{R_{DA}^6}$$

In the regime where electron exchange is negligible compared to energy transfer ($> 20\mathring{A}$), we can use Fermi's golden rule (Equation 2.7) to account for the weak electronic interaction involved in dipole-dipole coupling

$$\begin{aligned} k_{ET} &= \frac{2\pi}{\hbar}\left\langle \Psi_0(D)\Psi_1(A^*)|\hat{P}'_{D^*\to A^*}|\Psi_1(D^*)\Psi_0(A)\right\rangle^2 \rho \\ &\sim \frac{2\pi}{\hbar}\left\langle \Psi_0(D)\Psi_1(A^*)|\hat{P}'_{D^*\to A^*}|\Psi_1(D^*)\Psi_0(A)\right\rangle^2 \cdot \langle\chi_0(D)|\chi_1(D^*)\rangle^2\langle\chi_1(A^*)|\chi_0(A)\rangle^2 \end{aligned} \tag{2.16}$$

where ρ is the density of states capable of effective mixing between the four electronic states and which are also coupled by the dipole-dipole interaction; $\hbar = h/2\pi$ is the reduced Planck constant; $\hat{P}'_{D^*\to A^*}$ is the perturbation operator which mixes the wave functions $\Psi_1(D^*)$ and $\Psi_0(A)$ of the initial system state with the wave functions $\Psi_0(D)$ and $\Psi_1(A^*)$ of the final system state; $\langle\chi_i(x)|\chi_j(y)\rangle$ is the overlap integral of the respective vibrational wave functions for x and y. The square of the overlap integral is the Franck-Condon factor, which indicates the most probable electronic state transitions.

One of Förster's main insights was to relate the theoretical quantities involved in non-radiative electronic energy transfer, e.g., transition dipole moments $\boldsymbol{\mu}_x$, to

experimental quantities, e.g., molar extinction coefficients ε_x [88,92].[13] We provide his main result here:

$$k_{ET} = \underbrace{\left(\frac{9000 \ln 10}{128\pi^5 N_A}\right)}_{\text{constants}} \cdot \underbrace{\left(\eta^4\right)}_{\substack{\text{solvent} \\ \text{effects}}} \cdot \underbrace{\left(\frac{\Phi_D}{\tau_D}\right)}_{\substack{\text{donor} \\ \text{emission} \\ \text{properties}}} \cdot \underbrace{\left(\frac{\kappa^2}{R_{DA}^6} \frac{\int_0^\infty \frac{f_D(\bar{\nu})\varepsilon_A(\bar{\nu})}{\bar{\nu}^4} d\bar{\nu}}{\int_0^\infty f_D(\bar{\nu}) d\bar{\nu}}\right)}_{\substack{\text{dipole-dipole interaction} \\ \text{and} \\ \text{spectral overlap}}} \tag{2.17}$$

where η is the refractive index of the solvent, N_A is Avogadro's number, τ_D and Φ_D are the inverse excited state deactivation rate and the fluorescence quantum yield, respectively, for the donor molecule in the *absence* of the acceptor, $f_D(\nu)$ is the donor emission spectrum, $\varepsilon_A(\nu)$ is the molar extinction coefficient (units: M^{-1}cm^3), and $\bar{\nu} = 1/\lambda$ is the wavenumber (units: cm^{-1}). Let us define R_0 to be the intermolecular separation distance at which $k_{ET} = 1/\tau_D$.

$$R_0 = \frac{9000 \ln 10 \, \kappa^2 \Phi_D}{128\pi^5 N_A \eta^4} \underbrace{\left(\frac{\int_0^\infty \frac{f_D(\bar{\nu})\varepsilon_A(\bar{\nu})}{\bar{\nu}^4} d\bar{\nu}}{\int_0^\infty f_D(\bar{\nu}) d\bar{\nu}}\right)}_{J(\bar{\nu})} \tag{2.18}$$

where $J(\bar{\nu})$ is known as the *spectral density integral*. We may thus express Equation 2.17 more compactly

$$k_{ET} = \frac{1}{\tau_D}\left(\frac{R_0}{R_{DA}}\right)^6 \tag{2.19}$$

We refer to R_0 as the characteristic *Förster distance* for a donor-acceptor FRET pair, and we typically consider the quantity constant.[14]

A donor molecule's excited state lifetime in the absence of an acceptor depends

[13] Note: Förster's original publications describing FRET included a typographic error in Equation 2.17 which was not corrected until 1965.

[14] R_0 can fluctuate based on four quantities: η, Φ_D, κ^2, and $\varepsilon_A(\bar{\nu})$. Protein induced fluorescence enhancement (PIFE) is one example of local viscosity changes affecting fluorophore fluorescence [85,147,207,234,293]. Φ_D and $\varepsilon_A(\bar{\nu})$ may both fluctuate due to spectral shift [43,44,113,165]. Temporary fluorophore interaction with nearby nucleic/amino acids or lipid membrane intercalation may result in restricted rotational freedom and cause κ^2 fluctuations on the measurement timescale [38,69,149,165,187,235]. Typically, only when absolute distance measurements are sought is κ^2 an important parameter [323]; otherwise, one is primarily interested in relative changes in the transfer rate. In which case, we need only assume that κ^2 is constant. For most cases, $\kappa^2 = 2/3$ remains a reasonable approximation for single-molecule studies in which the measurement timescale is greater than the timescale for dipole randomization; see [54,71,352] for a discussion on the importance of κ, particularly in structural biology.

on several competing processes, each having a characteristic rate constant; see Figure 2.4. The deactivation rate, τ_D^{-1}, is equal to the sum of all such rate constants

$$\tau_D^{-1} = k_{IC} + k_{ST} + k_Q + k_{rxn} + k_{fl} \tag{2.20}$$

where k_{IC} is the rate of internal conversion, i.e., ro-vibrational relaxation, k_{ST} is the rate of intersystem crossing, k_Q is the rate of static and collisional quenching processes, k_{rxn} is the rate for alternative reaction pathways, and k_{fl} is the rate of photon emission. In the presence of an acceptor, dipole-dipole interaction provides an additional deactivation pathway which we previously characterized by the rate constant k_{ET}. We thus modify Equation 2.20

$$\tau_{DA}^{-1} = k_{IC} + k_{ST} + k_Q + k_{rxn} + k_{fl} + k_{ET} \tag{2.21}$$

We may define the energy transfer efficiency, E, between a donor-acceptor pair as the ratio of k_{ET} to all deactivating processes

$$E = \frac{k_{ET}}{k_{IC} + k_{ST} + k_Q + k_{rxn} + k_{fl} + k_{ET}} \tag{2.22}$$

Substituting Equations 2.19 and 2.20 into Equation 2.22, we obtain

$$\begin{aligned} E &= \frac{1}{1 + \left(\frac{R_{DA}}{R_0}\right)^6} \\ &= \frac{R_0^6}{R_0^6 + R_{DA}^6} \end{aligned} \tag{2.23}$$

The energy transfer efficiency scales with the inverse sixth power of the donor-acceptor separation, R_{DA}. FRET predominantly occurs at the nanometer scale. Most organic fluorophore donor-acceptor FRET pairs have Förster distances on the order of 50 Å [97]). Given that the typical diameter of biological macromolecules is between 10 Å and 1000 Å, [15] the application of Equation 2.23 to donor-acceptor

[15]The average protein diameter is 50 Å. Core bacterial RNA polymerase, a multifunctional

FRET pairs attached to biomolecules can provide a highly sensitive molecular length probe [295, 300, 301].

If we assume that the donor excited state lifetime, donor emission spectrum, acceptor absorption spectrum, and transition dipole moment strength are not perturbed due to donor-acceptor interaction [281], we can manipulate Equation 2.23 to obtain several variations in which E may be calculated experimentally [47].

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \tag{2.24a}$$

$$E = 1 - \frac{\Phi_{DA}}{\Phi_D} \tag{2.24b}$$

$$E = 1 - \frac{I_{DA}}{I_D} \tag{2.24c}$$

where $\{\tau_{DA}, \Phi_{DA}, I_{DA}\}$ and $\{\tau_D, \Phi_D, I_D\}$ are the excited state lifetimes, fluorescence quantum yields, and fluorescence intensities of the donor in the presence and absence of the acceptor, respectively.

Detected intensities are ultimately a convolution of the emitted photon flux with fluorescence quantum yields Φ (which rarely equal unity) and experimental detection efficiencies ξ

$$\xi = \frac{\text{number of photons detected}}{\text{number of photons emitted}}$$

Accordingly, $I_x^{detected} = I_x^{true}\Phi_x\xi_x$, where x corresponds to the photon emission detection channel. With this in mind,

$$\begin{aligned} E &= \frac{I_A^{true}}{I_{DA}^{true} + I_A^{true}} \\ &= \frac{I_A^{det}/\Phi_A\xi_A}{\frac{I_{DA}^{det}}{\Phi_{DA}\xi_{DA}} + \frac{I_A^{det}}{\Phi_A\xi_A}} \\ &= \frac{I_A^{det}}{\frac{\Phi_A\xi_A}{\Phi_{DA}\xi_{DA}} I_{DA}^{det} + I_A^{det}} \end{aligned} \tag{2.25}$$

enzyme responsible for RNA transcription, is 150 Å long and 110 Å wide [33].

Let $\gamma = \Phi_A \xi_A / \Phi_{DA} \xi_{DA}$, such that

$$E = \frac{I_A^{det}}{\gamma I_{DA}^{det} + I_A^{det}} \tag{2.26}$$

where I_A^{det} is the acceptor fluorescence upon electronic excitation due to FRET and γ is a correction factor accounting for the different detection efficiencies and fluorescence quantum yields of the donor-acceptor pair. This expression of E in terms of acceptor and donor fluorescence intensities is commonly referred to as a *ratiometric* approach [53, 63].

If we assume γ is equal to unity (for discussion, see [284] and [38, 216]), we can further simplify the calculated efficiency; we refer to this form of E as the 'apparent', 'raw', or 'uncorrected' FRET.

$$E_{raw} = \frac{I_A}{I_{DA} + I_A} \tag{2.27}$$

Equation 2.27 exhibits the expected behavior: i) as R_{DA} decreases, k_{ET} will increase and result in increased acceptor fluorescence I_A and decreased donor fluorescence I_{DA} in the presence of the acceptor: $E_{raw} \to 1$. ii) As R_{DA} increases, k_{ET} will decrease and result in a reduced acceptor fluorescence intensity and an increased donor fluorescence intensity: $E_{raw} \to 0$.

FRET promises significant advantages in the investigation of individual biological macromolecules. First, Equation 2.27 is a ratiometric measure which is able to calculate internal molecular distance within the molecular reference frame as opposed to the laboratory reference frame. This consequently avoids complications due experimental noise and instrumental drift [270]. Second, for $R_{DA} \in [0.5R_0, 2R_0]$, Equation 2.23 scales approximately linearly with changes in the donor-acceptor separation R_{DA} relative to the Förster distance R_0. Linearity allows a straightforward measure to detect distance fluctuations. Third, because both ensemble and single-molecule FRET techniques do not require the comparatively harsh experimental conditions of crystallographic and cryo-electron methods, we

can monitor biological molecules under conditions similar to the molecules' native environment in real-time, and we can even possibly extend *in vitro* assays to *in vivo* conditions. Lastly, we can investigate individual molecules over the entire duration of a reaction process. Such extended observation times allow the recovery of kinetic pathways, intermediates, static and dynamic heterogeneity, conformational fluctuations, and structure.

2.5 Summary and Outlook

This chapter provided an introduction to photophysics and photochemistry. We learned how quantum mechanics helps us build an intuition about why some processes are kinetically favored over others. We further discussed how environmental factors, such as molecular oxygen, pH, solvent polarity, and oxidizing and reducing agents, can affect photophysical kinetics. Many applications seek to maximize photon emission and photostability by minimizing the effect of processes which compete with fluorescence. We learned how work in molecular photoswitching has revealed general mechanisms which permit us to tune the behavior of fluorescent molecules.

We ended our introduction with a review of Förster resonance energy transfer (FRET) and highlighted the phenomenon's suitability as a molecular ruler for biomolecular organization and structural determination. Notably, single-molecule FRET provides an opportunity to monitor individual biomolecular conformational dynamics in real-time and with minimal perturbation. The competing photophysical and photochemical processes which we described in this chapter present, however, significant challenges and limit the potential of single-molecule FRET. Both experimental and statistical methods are needed to harness the technique's sensitivity and provide unambiguous determination of system dynamics. We outline such methods in the chapters which follow.

Contributions

All writing and figures are my own.

Chapter 3

Single-Molecule Signal Analysis

Abstract

Single-molecule fluorescence methods can monitor molecular behavior in real-time, at high spatial resolution, and with great sensitivity. Time-dependent fluorescence intensity fluctuations can reveal both kinetic and structural information about a molecular process. A method's practical value, however, depends on the ability to extract information content contained in experimental data. Here, we review signal analysis methods within the context of single-molecule spectroscopy. We first describe in detail single-molecule fluorescence methods to understand the data types obtained. Next, we discuss how we can extract information from equilibrium distributions and generate models which account for the data. We then address time series analysis. Two problems are common in single-molecule analysis: i) identifying time series correlations, and ii) modeling the statistical properties of a generative process. Accordingly, we focus on correlation methods and Markov models which provide solutions to these problems. We note each method's strengths and limitations. Finally, we follow with a brief discussion about software implementation and discuss its increasingly important role in facilitating single-molecule data analysis. This chapter helps to motivate the analysis methods introduced in this thesis, and we expect the chapter will prove useful in informing experimental design to ensure reliable and robust information content extraction.

3.1 Introduction

Before we can understand the information content of single-molecule fluorescence experiments, we must first understand the experimental methods we use to obtain single-molecule fluorescence data. We typically classify experimental meth-

ods as either solution- or surface-based experiments. In solution experiments, a fluorescent analyte freely diffuses and is able to explore an entire sample volume; whereas, in surface experiments, biomolecules are surface-immobilized and thus limited to tethered motion.

We will first introduce the basic principles of experimental methods relevant to single-molecule fluorescence. We highlight what parameters we can extract, and we describe the characteristics of the data obtained. We then discuss analysis methods appropriate to solution- and surface-based measurements. We follow with a brief discussion about software implementation and state a case for a common analysis platform. Finally, we conclude by noting the strengths and limitations of current approaches and by motivating the analysis methods which are the subject of this thesis.

3.2 Experimental Methods

3.2.1 Solution-Based

Solution-based experiments allow us to monitor biomolecular interactions between diffusing molecules. We typically perform these measurements on a fluorescence confocal microscope [252], in which a Gaussian laser beam is focused approximately 1 micron above an objective-glass interface. Diffraction limits the focal spot size [122]. We characterize the focal spot based on the minimum laser cross section radius called the *beam waist* w_0 and the distance over which the laser cross-sectional area doubles called the *Rayleigh range*

$$z_R = \frac{\pi w_0^2}{\lambda} \tag{3.1}$$

where λ is the laser wavelength. We refer to a microscope's optical response to a point source at the center of the focus as the *point spread function* (PSF). Our observation of a PSF is typically a 3-dimensional diffraction pattern brought into focus by an optical lens. Accordingly, the observed image is a convolution of the

true object and the PSF.

To improve signal-to-noise, we use a pinhole to reject out-of-focus light, and consequently, we superimpose a detection volume on the focal volume. We can approximate the observation volume, i.e., the overlap between the sample, excitation, and detection volumes, as a function of position vector $\mathbf{r}$ by a *prolate ellipsoid*, which is a 3-dimensional Gaussian with different lateral and axial dimensions.

$$W(\mathbf{r}) = I_0(\mathbf{0})\exp\left[-\frac{2(x^2+y^2)}{w_r^2} - \frac{2z^2}{w_z^2}\right] \tag{3.2}$$

where w_r and w_z are the radial and axial distance, respectively, from the PSF center to where the intensity $I_0(0)$ has decreased by $1/e^2$. The confocal observation volume is approximately 1 fL (10^{-15}L).

If we assume freely diffusing, non-interacting fluorescent analytes, which do not undergo photophysical effects, we can determine the measured fluorescence intensity as a function of time [152].

$$F(t) = \xi\sigma_{abs}\Phi_{fl}\int W(\mathbf{r})C(\mathbf{r},t)\,d\mathbf{r} \tag{3.3}$$

where ξ is the overall system detection efficiency, σ_{abs} is the analyte absorption cross section at the laser excitation wavelength, Φ_{fl} is the fluorescence quantum yield, and $C(\mathbf{r},t)$ is the analyte concentration at position $\mathbf{r}$ and time t. The product

$$B = \xi\sigma_{abs}\Phi_{fl}W(\mathbf{0})$$

gives the molecular brightness at the PSF center.

As fluorescent analytes diffuse into and out of the observation volume, we observe a stochastic series of temporary fluorescence *bursts*, in which detected photon counts exceed a background photon count rate; see Figure 3.1. Background photon counts arise from elastic, i.e., Rayleigh, and inelastic, i.e., Raman, scattering, fluorescent impurities, and detector and other electronic dark counts. Despite

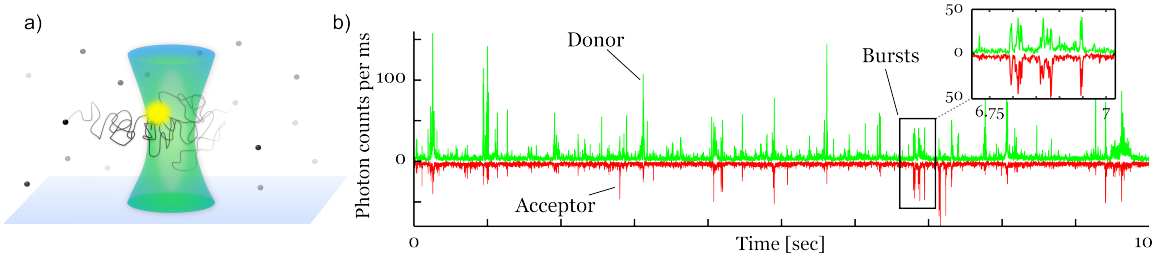


Figure 3.1: Confocal Microscopy of Diffusing Molecules in Solution. a) A focused laser beam excites a fluorescent analyte, represented by the yellow sphere, as it diffuses through a small excitation (green cone)/detection volume ($<$ 1fL). b) We can subsequently bin collected photons into 1 ms bins. Here, we show a doubly labeled analyte where the fluorescent labels are a donor-acceptor FRET pair. The donor photon count is green and acceptor count is red. Bursts (see inset) are sequences of photon counts which exceed the background photon count rate. Rayleigh and Raman scattering, fluorescence impurities, and detector and electronic noise contribute to the observed background. Sample: RNAP-promoter open complex, RP_o .

the method's simplicity, determining the information content is not straightforward and requires sophisticated analyses. Bursts often consist of < 100 photons, and thus, shot noise dominates photon count fluctuations. *Shot noise* is a statistical fluctuation which arises because of randomness in photon emission, as well as the quantum mechanical nature of photon-detector interaction. An incident photon will only probabilistically result in an output count. Hence, photon detection is a stochastic, i.e., random, process. Additionally, an analyte may take any path through the observation volume, which has a spatially dependent excitation intensity and spatially dependent collection efficiency. The result is a burst sequence, where bursts vary in terms of width and photon count rate.

From the theory of diffusion controlled reactions, we can estimate the frequency $f_{burst} = \tau_{burst}^{-1}$ that a diffusing analyte encounters the observation volume [78]

$$\tau_{burst} \sim \mathcal{O}\left(\frac{1}{4\pi w_r C D_{\text{coeff}}}\right) \tag{3.4}$$

where $\mathcal{O}$ indicates 'on the order of', D_{coeff} is the analyte diffusion coefficient, w_r is the PSF radius, and C is the analyte sample concentration. For sufficiently low, i.e., single-molecule, concentrations (on the order of pM), bursts will be well separated. We can similarly estimate the burst duration $\tau_{duration}$

$$\tau_{duration} \sim \mathcal{O}\left(\frac{w_r^2}{3D_{\text{coeff}}}\right) \tag{3.5}$$

The average protein diameter is on the order of 50-100 Å. If we estimate $w_r = 0.5$ μm, then $\tau_{duration} \approx$ 1-2 ms.[1] For an exact burst size distribution description, one must account for triplet kinetics, fluorescence and detection saturation, multi-molecule events, and photobleaching [94].

For a labeled analyte with multiple spectrally distinct fluorescent probes, we spectrally separate photon emission based on photon wavelength and define photon emission streams for each emission detection channel. Spectral separation is not, however, perfect, with photons from one probe 'leaking' into the detection channel of another probe. Moreover, fluorescent probe absorption spectra can overlap and laser excitation may excite multiple probes, whose absorption efficiency would depend on their respective molar extinction coefficient $\varepsilon_x(\nu)$. Hence, detected photon emission streams are complex convolutions which involve multiple factors. For example, for a doubly-labeled analyte where the fluorescent probes are a donor-acceptor FRET pair, we define

$$F_{D_{exc}}^{D_{em}} = \sigma_{D_{exc}}^{D} \Phi_D \xi_{D_{em}}^{D} (1 - E) \int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\, dt + b_D \tag{3.7}$$

$$F_{D_{exc}}^{A_{em}} = \Bigg(\underbrace{\sigma_{D_{exc}}^{D} \Phi_D \xi_{A_{em}}^{D} (1 - E)}_{\text{leakage}} + \underbrace{\sigma_{D_{exc}}^{A} \Phi_A \xi_{A_{em}}^{A}}_{\text{direct excitation}} + \underbrace{\sigma_{D_{exc}}^{D} \Phi_A \xi_{A_{em}}^{A} E}_{\text{FRET}} \Bigg) \int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\, dt + b_A \tag{3.8}$$

where D stands for 'donor' and A for 'acceptor'; $I_{D_{exc}}$ is the D-excitation laser intensity as a function of the analyte path $\vec{r}(t)$ through the confocal volume; $\Delta t = t_j - t_i$ is the total time the analyte spent in the confocal volume; $\sigma_{D_{exc}}^{D}$ and $\sigma_{D_{exc}}^{A}$ are the absorption cross sections of D upon D-excitation and of A upon D-excitation, respectively; Φ_D and Φ_A are fluorescence quantum yields of D and A, respectively;

[1] For reference, RNA polymerase, a multi-subunit enzyme involved in RNA transcription, is 150 Å long and 110 Å wide [33]. The *Stokes-Einstein relation* allows us to calculate the diffusion coefficient for a spherical particle

$$D_{\text{coeff}} = \frac{k_B T}{6\pi \upsilon r} \tag{3.6}$$

where $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, T is the absolute temperature, υ is the viscosity, and r is the particle radius. For $T = 293$ K, $\upsilon_{H_2O} = 1 \times 10^{-3}$ Pa·s, and $r \approx 60$ Å, $D_{\text{coeff}} \approx 3.5 \times 10^{-11}$ m/s. Using $w_r = 0.5$ μm, then $\tau_{duration} \approx 2$ ms.

$\xi^D_{D_{em}}$, $\xi^D_{A_{em}}$, and $\xi^A_{A_{em}}$ are detection efficiencies of D-emission in the D-detection channel, of D-emission in the A-detection channel, and of A-emission in the A-detection channel, respectively; b_D and b_A are the background counts during the interval Δt in the D- and A-detection channels, respectively; and E is the energy transfer efficiency, i.e., FRET.

For basic solution-based confocal measurements, we record detected photons according to photon arrival time and detection channel, which results in an $N \times 2$ matrix, where N is the total detected photon number. More sophisticated experimental techniques such as time-correlated single photon counting (TCSPC) [22] and separating photons based on polarization [257,341] introduce additional information. Photon arrival and burst analysis methods attempt to uncover temporal correlations and average burst properties.

3.2.2 Surface-Based

We typically perform single-molecule surface-based measurements in one of two realizations: confocal scanning microscopy and total internal reflection fluorescence (TIRF) microscopy. Both realizations have advantages and limitations, and we discuss each in turn.

3.2.2.1 Confocal Scanning

Confocal scanning microscopy differs only slightly from solution-based confocal microscopy. The difference is two-fold. First, because we surface-immobilize biomolecules, we position the laser focus closer to the immobilization surface. Second, instead of positioning the focus once and collecting population statistics, we must actively scan across a surface for immobilized biomolecules.

Confocal scanning has the advantage that we can collect individual photons for an individual biomolecule over an extended time. In solution-based measurements, diffusion limits observation times to milliseconds. With confocal scanning and recent advances yielding photostable organic fluorophores, we can observe a

single biomolecule for tens of seconds. Hence, confocal scanning extends our ability to probe reaction timescales, without a loss in temporal resolution. Moreover, confocal scanning does not, as in TIRF, involve complex image analysis or introduce additional camera noise. We can thus adapt, in many cases, analysis methods we develop for solution-based confocal measurements to confocal scanning photon arrival sequences.

The method, however, has two significant limitations. First, we can only observe one molecule at a time.[2] This means we cannot simultaneously observe interactions for multiple immobilized biomolecules. As a corollary, the method suffers from low-throughput. To obtain sufficient statistics, one must sequentially record individual photon sequences, which, for photobleaching lifetimes approaching minutes, means extended total measurement times. Second, the promise of confocal scanning lies in its ability to capture fast reaction kinetics, but to achieve an adequate signal-to-noise ratio to detect fast (or small) fluctuations, confocal scanning requires high laser excitation intensities. As mentioned discussed previously (see Section 2.3.4), however, higher laser excitation intensities may affect fluorophore photostability and introduce photophysical complications [76].[3]

3.2.2.2 Total Internal Reflection Fluorescence

Whereas confocal microscopy is a point-detection method, total internal reflection fluorescence (TIRF) microscopy is a widefield method, in which we image an extended surface area on the order $10 - 100$ μm^2 on a sensitive camera, such as an electron multiplying charged coupled device (EMCCD). When a light beam propagating through a incident medium with index of refraction η_i (e.g., glass

[2]Note that holographic techniques can overcome this limitation [206], but most laboratories currently lack the ability to realize such implementations.

[3]Additionally, compared to TIRF, confocal scanning suffers from increased solution background, thus mitigating one of its advantages (i.e., no camera noise). TIRF illuminates a smaller sample volume ($\lll$ 1fL), and thus excites fewer fluorescence impurities and generates less Rayleigh and Raman scattering. Lower background offsets the additional camera noise. In our experience, we have found that confocal scanning does not achieve a better signal-to-noise ratio compared to TIRF at millisecond temporal resolution and, inferred based on detected photon count rate, similar laser excitation intensities.

$\eta_i = 1.51$), encounters a planar interface with another medium with index of refraction η_t (e.g., water $\eta_t = 1.33$), the beam will undergo total internal reflection for incidence angles (measured relative to the normal) greater than a critical angle θ_c [10,226]

$$\theta_c = \arcsin\left(\frac{\eta_t}{\eta_i}\right) \tag{3.9}$$

When $\eta_t/\eta_i \geq 1$, total internal reflection does not occur. For an incidence angle $\theta < \theta_c$, Snell's law determines the refraction angle at which light propagates through the transmitting medium. For $\theta > \theta_c$, all light reflects back into the incident medium, but, nevertheless, some incident energy in the form of an electric field transmits through the interface and propagates parallel to the interface plane [122]. The transmitted field is called an *evanescent field* (or wave). The evanescent field intensity I is proportional to the square of the electric field amplitude and exponentially decays according to the following relation

$$I(z) = I(0)e^{-z/d} \tag{3.10}$$

where $I(0)$ is the intensity at the interface and d is the characteristic decay distance

$$d = \frac{\lambda_0}{2\pi\sqrt{\eta_i^2 \sin^2\theta - \eta_t^2}} \tag{3.11}$$

λ_0 is the incident light wavelength in vacuum. In the limit that $\theta \to \theta_c$, $d \to \infty$, but, in practice, $\theta > \theta_c$ such that $d < \lambda_0$ (usually on the order of 100 nm). We illustrate the concept of TIRF microscopy in Figure 3.2.

The optical depth for TIRF is relatively shallow compared to confocal microscopy, differing by an order of magnitude. Importantly, a shallow optical depth reduces background and thus allows analytes to be resolved at lower laser excitation intensities. Moreover, in confocal microscopy, fluorescent analyte concentrations must be less than 100 pM to ensure that only one analyte traverses the confocal volume at a time. A shorter optical depth translates to a reduced illumination volume and thus allows higher fluorescent analyte concentrations (up to 100 nM). The ability

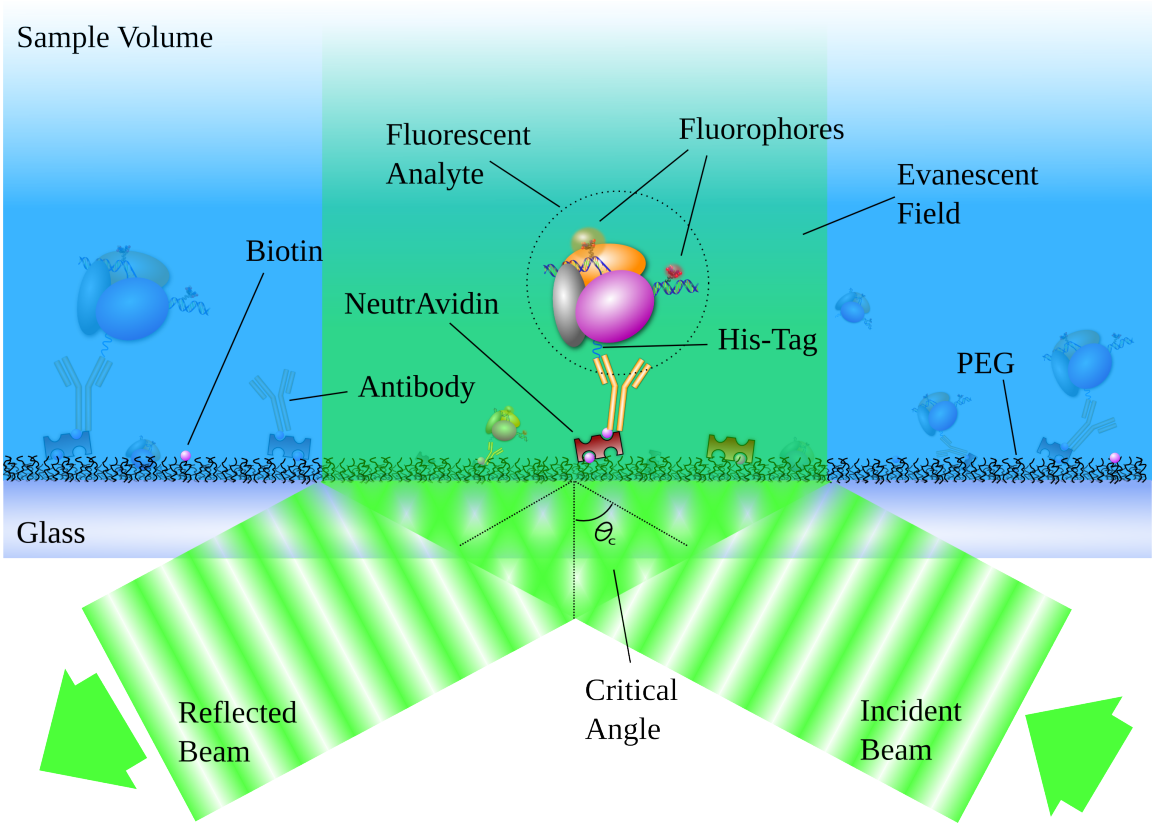


Figure 3.2: Total Internal Reflection Fluorescence (TIRF) Microscopy Schematic. A light beam totally reflects when two conditions are satisfied: i) the transmitting medium, here the sample volume, is of lower refractive index than that of the incident medium, and ii) the light beam is incident at an angle greater than a certain critical angle. While the beam is reflected back into the incident medium, an electric field propagates into the transmitting medium in a direction perpendicular to the medium interface. TIRF microscopy exploits this electric field to excite fluorophores located typically within ∼100 nm of the interface. The shallow optical depth minimizes sample background contributions. The evanescent field illuminates multiple fluorescent analytes, while maintaining spatial distinction, and thus, TIRF microscopy permits highly parallel imaging. Note: elements are not drawn to scale.

to measure at higher fluorescent analyte concentrations is beneficial for substrate binding studies, in which we observe freely diffusing fluorescent analytes bind an immobilized non-fluorescent substrate.[4] The other advantage of TIRF microscopy is its relatively high throughput compared to confocal scanning microscopy. We are able to simultaneously observe multiple individual reaction trajectories (on the order of 100 molecules).

TIRF microscopy is not, however, without limitations. Two limitations feature most prominently. First, TIRF microscopy introduces additional camera noise and requires complex image analysis [135]. We describe in detail the origins of such noise and how we recover the underlying fluorescent signal in Chapter 7. Second, camera electronics limit the measurement temporal resolution. Currently, TIRF microscopy is unable to access kinetics regimes which occur on timescales faster than the single millisecond timescale, and, at the highest temporal resolution, one must usually sacrifice imaging area size and thus reduce throughput.

While we can temporally bin confocal scanning photon arrival data at discrete sampling intervals, we are not required to do so. Because we have access to the raw photon arrival data, we can arrange and sort the data as we desire, and thus, we have great flexibility in choosing analysis methods. We are free to use either *event-driven* (time inhomogeneous) or uniform sampling (time homogeneous) time series approaches. In TIRF microscopy, we do not have access to photon arrival information. Instead, image-extracted single-molecule TIRF intensity data represents a discrete-interval time series, $X = \{x_0, x_1, \ldots, x_{N-1}\}$, where N is the total number of sampling intervals.[5] This limitation thus places constraints on the types of analysis we can employ and translates to reduced information content.

[4] Zero mode waveguides further extend this advantage of TIRF, allowing even higher diffusing fluorescent analyte concentrations [77, 184, 199, 366].

[5] The sampling interval f_s^{-1} is the inverse of the sampling frequency (or camera frame rate) f_s.

3.2.3 Excitation Schemes

We have already observed that we can increase information content by introducing multiple spectrally distinct fluorescent probes and multiple fluorescence emission detection channels (Equations 3.7 and 3.8). Access to multiple photon streams allows us to examine the relationships between those streams, such as the extent to which variation in one stream co-varies with the variation of a different stream. For the case of spectral separation, we created multiple variables in a time-independent manner.

We might ask if we can create variables which are time-dependent, possibly by temporally interleaving them together. Laser excitation schemes provide such a possibility and introduce additional information content variation. We now discuss two such schemes, single-laser excitation and alternating-laser excitation, within the context of single-molecule FRET.

3.2.3.1 Single-Laser Excitation

Initial single-molecule FRET experiments used single-laser excitation and a single donor-acceptor FRET pair [111]. While several early seminal works demonstrated the ability of single-pair FRET to study protein, nucleic acid, and protein-nucleic acid complex dynamics [112, 283, 368], these works could only identify relative distance fluctuations and associated kinetics. Single-pair FRET is not a general framework for quantitative structural analysis. Additional information is needed to measure accurate FRET efficiencies and to detect experimental artifacts which arise from fluorophore photochemistry and non-FRET photophysics.

Moreover, single-pair FRET is not a general framework for quantitative analysis of molecular interactions. Consider the following reaction

$$\mathrm{E^D} + \mathrm{S^A} \underset{k_{10}}{\overset{k_{01}}{\rightleftharpoons}} \mathrm{E^D S^A}$$

where E^D is a donor-labeled enzyme and S^A is an acceptor-labeled substrate.

Single-pair FRET cannot accurately quantify the reaction species. The first problem is that single-pair FRET only produces an appreciable FRET signal when the inter-fluorophore separation is sufficiently close ($R_{DA} < 1.5R_0$). For $R_{DA} > 1.5R_0$, single-pair FRET cannot distinguish between E^D and $E^D S^A$ species. The next problem is that inactive acceptor fluorophores will not participate in FRET; their inactivity will result in an apparent increase in free E^D species. Furthermore, single-pair FRET cannot typically detect (or quantify) S^A species at donor-laser excitation wavelengths. Finally, single-pair FRET cannot reliably identify stoichiometries other than 1:1 (e.g., $E^D[S^A]_{n>1}$).

If we combine these limitations with sub-stoichiometric fluorophore labeling (i.e., labeling ratios other than 1:1), non-FRET photophysics, photochemistry, and aggregation phenomena, we realize that we need additional means to increase information content. Finding such means is particularly important if we want to relate fluorescence intensity fluctuations to biological structure and function. Nevertheless, in non-FRET contexts, such as protein-induced fluorescence enhancement (PIFE) [85, 147, 207, 234, 293], a single-laser single-fluorophore system may suffice to extract the information needed.

3.2.3.2 Alternating-Laser Excitation

To address the challenges of single-laser single-pair FRET, Kapanidis and co-workers introduced alternating-laser excitation (ALEx) spectroscopy [168] (reviewed in [274]). The method's basic premise is as follows: interleave short laser excitation pulses and directly excite each spectrally distinct fluorophore which has its own dedicated emission channel. For example, for a single donor-acceptor FRET pair, rapidly alternate between donor-excitation and acceptor-excitation lasers. Due to time-dependent segmentation, we can recover four different photon emission

streams:

$$F_{D_{exc}}^{D_{em}} = \sigma_{D_{exc}}^{D}\Phi_D\xi_{D_{em}}^{D}(1-E)\int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\,dt + b_{D_{exc}}^{D_{em}} \tag{3.12}$$

$$F_{D_{exc}}^{A_{em}} = \left(\sigma_{D_{exc}}^{D}\Phi_D\xi_{A_{em}}^{D}(1-E) + \sigma_{D_{exc}}^{A}\Phi_A\xi_{A_{em}}^{A} + \sigma_{D_{exc}}^{D}\Phi_A\xi_{A_{em}}^{A}E\right)\int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\,dt + b_{D_{exc}}^{A_{em}} \tag{3.13}$$

$$F_{A_{exc}}^{D_{em}} \approx 0 + b_{A_{exc}}^{D_{em}} \tag{3.14}$$

$$F_{A_{exc}}^{A_{em}} = \sigma_{A_{exc}}^{A}\Phi_A\xi_{A_{em}}^{A}\int_{t_i}^{t_j} I_{A_{exc}}(\vec{r}(t))\,dt + b_{A_{exc}}^{A_{em}} \tag{3.15}$$

where $\sigma_{A_{exc}}^{A}$ is the absorption cross section of A upon A-excitation; $I_{A_{exc}}$ is the A-excitation laser intensity as a function of the analyte path $\vec{r}(t)$ through the confocal volume; and b_X^Y is the background count during the interval $\Delta t = t_j - t_i$ in the Y-detection channel upon X-excitation. Typically, FRET pairs are matched with laser excitation wavelengths such that a FRET donor is not appreciably excited by direct acceptor-laser excitation; hence, background counts will be the primary contributor to the $F_{A_{exc}}^{D_{em}}$ photon stream.

Examination of the four ALEx photon streams reveals the following relations:

$$\begin{aligned}\text{leakage} &= \sigma_{D_{exc}}^{D}\Phi_D\xi_{A_{em}}^{D}(1-E)\int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\,dt \\ &= F_{D_{exc}}^{D_{em}}\frac{\xi_{A_{em}}^{D}}{\xi_{D_{em}}^{D}}\end{aligned} \tag{3.16}$$

$$\begin{aligned}\text{direct excitation} &= \sigma_{D_{exc}}^{A}\Phi_A\xi_{A_{em}}^{A}\int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\,dt \\ &= F_{A_{exc}}^{A_{em}}\frac{\sigma_{D_{exc}}^{A}\int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\,dt}{\sigma_{A_{exc}}^{A}\int_{t_i}^{t_j} I_{A_{exc}}(\vec{r}(t))\,dt}\end{aligned} \tag{3.17}$$

We can readily determine the leakage using a background-corrected ratio $F_{D_{exc}}^{A_{em}}/F_{D_{exc}}^{D_{em}}$ for a donor-only species and the direct excitation using a background-corrected ratio $F_{D_{exc}}^{A_{em}}/F_{A_{exc}}^{A_{em}}$ for an acceptor-only species.

We calculate a raw FRET efficiency E as discussed previously (Equation 2.27)

$$E_{raw} = \frac{F_{D_{exc}}^{A_{em}}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}} \tag{3.18}$$

and we now introduce a new parameter called the *stochiometry*. We define stoichiometry S as the ratio of donor-laser excitation fluorescence to the total detected fluorescence, such that

$$S_{raw} = \frac{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}} + F_{A_{exc}}^{D_{em}} + F_{A_{exc}}^{A_{em}}} \tag{3.19}$$

A few points to consider: i) both E and S are ratiometric random variables which exist on the interval between [0,1]. ii) E is independent of excitation intensity; S is dependent on excitation intensities. The ratio of donor- and acceptor-laser excitation intensities affects the relative molecular brightness. For example, if we increase A-excitation relative to D-excitation, $F_{A_{exc}}^{A_{em}}$ will increase, and $S \to 0$. Excitation intensity dependence also means that S is affected by confocal excitation volume mismatch. iii) Both E and S are sensitive to confocal emission volume mismatch. iv) E is dependent on inter-fluorophore separation R_{DA}; S is distance-independent[6].

S subsequently allows us to identify different stoichiometric ratios and species brightnesses, i.e., number and variation of fluorescent components. S sensitivity to fluorophore brightness follows from its sensitivity to species brightness, which is important for identifying non-FRET photophysics. Changes in fluorescence quantum yield due to absorption/emission shifts, quenching, or restricted rotational freedom will therefore be manifest in changes in S. S thus provides a means to exclude non-FRET photophysics as the origin of E fluctuations. We can visualize resolved subpopulations by plotting the transfer efficiency versus stoichiometry in a two dimensional histogram; see Figure 3.3.

[6]Not strictly true. In its current form, Equation 3.19 is sensitive to changes in the sum $F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}$. We can write a corrected (or accurate) form similar to E, which includes the detection correction factor γ [195]. Nevertheless, even for $\gamma \neq 1$, the distance dependence is weak.

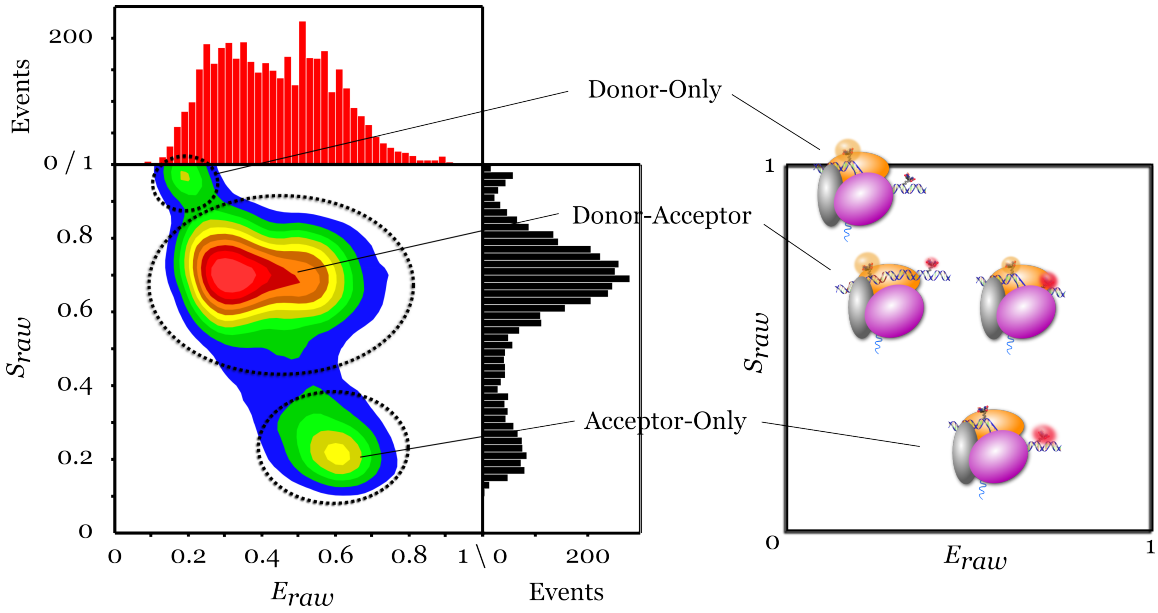


Figure 3.3: Alternating-Laser Excitation (ALEx) Spectroscopy and Virtual Molecule Sorting. By creating a two-dimensional histogram from the ratiometric random variables E, the energy transfer efficiency, and S, the stoichiometry, we can resolve multiple sub-populations buried in the respective marginal distributions (E is in red; S is in black). We can thus filter donor-only and acceptor-only subpopulations to isolate the donor-acceptor species. Experimental sample: RNAP-promoter DNA open complex RP_o . All-photon burst search [239] with parameters: $T = 500\ \mu$s$, M = 7$ neighbor photons$, L = 10$ photons. Additional per burst filtering: $F^{A_{em}}_{A_{exc}} > 10$ photons and $F^{D_{em}}_{D_{exc}} + F^{A_{em}}_{D_{exc}} > 25$ photons. The spreads connecting the donor-only and acceptor-only subpopulations with the donor-acceptor subpopulation indicates photoblinking and/or photobleaching during analyte transit through the confocal volume. Identifying such non-FRET photophysics within bursts is the focus of later chapters.

In its initial incarnation, laser alternation occurred on the microsecond timescale [168, 195]. Subsequent extensions to the nanosecond regime (nanosecond ALEx [192] and pulsed interleaved excitation (PIE) [229]) allow the laser excitation source for each photon emission to be known. Time-correlated single photon counting (TCSPC) combines with PIE to provide access to fluorescence lifetimes. Further extensions which involve three or more spectrally distinct fluorophores and multiple FRET pairs greatly extend our ability to maximize information content [194, 361]. Finally, multiparameter fluorescence detection (MFD) methods can determine up to 14 independent parameters from a single fluorophore [257, 341].

Unfortunately, however, many developed methodologies which are capable of identifying and resolving photophysical processes remain out-of-reach, both experimentally (due to costs and setup) and analytically. Determining 14 parameters is not particularly helpful if we do not have analytical methods which relate each parameter to one another. With this in mind, we now address the problem of actually extracting meaningful information from single-molecule fluorescence data.

3.3 Analysis Methods

A *time series* is an ordered measurement sequence, such that data points belong to successive time instants. Both photon arrival sequences and discrete-interval TIRF intensity data qualify as time series. TIRF intensity data is *time-homogeneous*, which means that each temporally ordered data point is equidistant in time from its two neighboring data points. The majority of time series analysis methods address time-homogeneous data sets; e.g., Markov models, auto- and cross-correlation, and many change-point detection algorithms. Photon arrival sequences are *time-inhomogeneous*; each photon arrival time, i.e., data point, is stochastic, where the waiting times separating successive data points obey some form of exponential distribution, which, itself, may time vary. Fewer time series analysis methods

specifically address time-inhomogeneous, or event-driven, time series.[7]

Here, we review analysis methods for single-molecule fluorescence time series data reduction, i.e., the process of identifying and extracting meaningful information from single-molecule data. We first review how we identify fluorescence bursts in photon arrival data. Next, we discuss how we can extract information from equilibrium distributions and generate models which account for the data. We then explicitly address time series analysis. We focus on two method types: correlation methods and Markov models. Correlation methods quantify time series self- and cross-similarity and describe relations between random variables. Markov models characterize a time series' statistical properties and enable us to forecast a system's future behavior. We conclude by offering an assessment of current methods and suggest future directions.

3.3.1 Burst Analysis

A fluorescent analyte diffusing into and out of a confocal observation volume generates a stochastic series of fluorescent bursts, in which the detected photon count rate exceeds a background level; see Figure 3.1. Recall that despite the data's apparent simplicity, the data's stochastic nature requires sophisticated analysis. Excitation intensity and detection efficiency are both spatially dependent. This dependence results in a burst size distribution where both burst length and photon count varies. Moreover, photon shot noise typically contributes significantly to data variation because of low photon numbers.

In order to analyze the burst size distribution and probe the properties of diffusing molecules, we must first identify bursts. We address burst detection next.

[7]Often, when we want to apply time-homogeneous time series analysis methods, we transform an event-driven time series into a time-homogeneous time series, e.g., by counting the number of data points which fall into equally spaced bins. This transformation, however, reduces information content–primarily the waiting times between data points which themselves may be important information.

3.3.1.1 Burst Detection

The most straightforward burst detection methods use photon thresholds [73, 82, 239]. First, define a temporal window T; e.g., $T = 500$ μs. Next, beginning from the measurement onset t_0, incrementally slide T in time, where the increment is equal to the detector resolution; e.g., APD dead time. Record the number of photons in each window. If a window contains more photons than a specified threshold M, mark the first photon as the beginning of a burst. Continue to slide T until a window fails to exceed threshold M, and mark the previous window's last photon as the burst end. Compute the burst total photon number N_i. We may optionally apply another threshold L. If the photon number for burst i exceeds L, retain the burst; otherwise, discard. Continue to slide the window and identify bursts until the measurement end t_f.

Many laboratories, including our own, use this threshold approach. We can apply the method to one or multiple channels, either singly or jointly. The particular application depends on the experimental sample. Threshold burst search algorithms are sensitive to fluorophore brightness, analyte diffusion coefficients, and confocal optics. For applications which require accurate burst detection, an individual should ensure that threshold parameters are appropriate by using experimental controls and/or performing simulations. For applications in which we can afford to be more liberal in our burst selection criteria, we can specify minimal threshold parameters, e.g., $M = 7, L = 10$, effectively detecting anything which vaguely resembles a fluorescence burst, and then further filter and sort bursts after visualizing burst data, e.g., using ALEx E-S histograms.

Accurate burst detection can be essential in several applications, such as biosensing [208] and determining *in vivo* copy numbers [313]. Several works have developed statistical methodologies to identify bursts which are more robust than threshold algorithms. Grange and co-workers developed a parameter-free method, which relies on iterative estimation of the background distribution [106]. Recent work by Barber and co-workers introduced a Bayesian method capable of detect-

ing bursts with signal amplitudes only 10% above background [11]. Other methods exploit fluorescence lifetime information to distinguish background photons from fluorescent analytes [81, 94, 256].

We should note, however, that inter-burst intervals are not random. In solution, a focused laser beam introduces a biasing potential and creates a weak optically induced trap [246]. The effect is only significant for electronically excited molecules. Consequently, a focused laser biases diffusing fluorescent analytes to both remain in the observation volume and to re-enter the volume after exiting.

Recurrence analysis of single particles (RASP) exploits the biasing effect to extend the measurement timescale we can observe a single diffusing analyte [132]. We expect fluorescence bursts from different and non-interacting analytes to be uncorrelated; however, we might surmise that we would observe correlation if bursts originate from the same molecule. RASP computes the probability that successive bursts belong to the same analyte in a manner similar to a correlation function and then groups those bursts based on their fluorescence properties, e.g., E values. Hoffman and co-workers demonstrated that, by systematically extending a recurrence interval about a burst to include bursts from different analytes, they could extend the measurement timescale for a single analyte to ~ 100 ms and access interconversion kinetics beyond the diffusion time.

3.3.2 Distribution Analysis

One method to extract information from single-molecule fluorescence data is to fit functions to parameter distributions. For example, to determine the mean energy transfer efficiency and, hence, obtain relative distance information, a common practice among researchers is to fit Gaussian functions to E histograms. While fitting histograms is quick and straightforward, such a method often lacks statistical rigor, especially if we consider that the fitting function choice, histogram bin width, and number of components are often arbitrary. Moreover, while we can obtain additional fit parameters beside the mean transfer efficiency, such as E dis-

tribution variance and other moments, the fit does not inform us as to the physical significance of these parameters.

A problem faced within the single-molecule field is resolving sample heterogeneity, which is frequently related to sample variance. *Static heterogeneity* refers to inter-molecular variation among sample molecules, and *dynamic heterogeneity* refers to intra-molecular variation of an individual molecule's statistical qualities over time. For example, a sample which contains multiple DNA standards, each internally labeled with the same donor-acceptor FRET pair separated at different inter-fluorophore distances, would exhibit a broad E distribution. Such distribution broadening would be due to static heterogeneity. In contrast, consider a single species which dynamically fluctuates on the measurement timescale between two distinct conformations. If we were to label this species with a single donor-acceptor FRET pair such that different E values reported on the different conformations, the E distribution would also broaden, but, here, the broadening would be due to dynamic heterogeneity.

One question to ask is: how we would distinguish between static and dynamic heterogeneity based on parameter distributions alone? A parameter's marginal distribution removes temporal information, and thus, we cannot discern *a priori* whether distribution broadening is static and/or dynamic in origin. Several works have attempted to address this problem and have provided detailed and statistically rigorous analysis methods to extract information from parameter distributions and to determine the origins of heterogeneity.

3.3.2.1 Photon Counting Histograms

Photon counting histogram (PCH) analysis analyzes the distribution of fluorescence intensity during a specified sampling interval. Several works introduce the theory and provide rigorous mathematical derivation [34, 35, 230], including accounting for non-ideal detection and extending to multiple detection channels [36, 130, 131]. A slightly different method termed fluorescence intensity distribu-

tion analysis (FIDA) seeks to achieve the same aim as PCH, but differs in its mathematical methodology [171, 172, 220]. We briefly review the main results here.

PCH uses two parameters to characterize fluorescence fluctuation data: $\langle N \rangle$, the average number of molecules present in the confocal volume, and B, the molecular brightness. We describe the photon detection shot noise by a Poisson distribution, which gives the probability of detecting k photon counts

$$p(k;\mu) = \frac{\mu^k e^{-\mu}}{k!} = \frac{(\zeta_I I)^k e^{-\zeta I}}{k!} \tag{3.20}$$

where μ is the mean count rate, ζ_I is a detection efficiency proportionality constant, and I is the fluorescence intensity. Equation 3.20 holds for a perfectly steady light source. We must modify this equation to account for the spatial dependence of I within the confocal volume [35]

$$p(k) = \int_0^\infty \frac{(\zeta_I I)^k e^{-\zeta I}}{k!} p(I)\, dI \tag{3.21}$$

where $p(I)$ is the intensity probability distribution. Recall that the mean of a Poisson distribution is equal to its variance: $\langle k \rangle = \langle k^2 \rangle - \langle k \rangle^2$. The basic idea for PCH is then as follows: an observed variance greater than $\langle k \rangle$ must be due to additional fluctuations accounted for by $p(I)$. One source of such fluctuations is analyte diffusion within the confocal volume.

We may address analyte diffusion by either theoretically [35] or empirically [172] modeling the confocal volume. By measuring the PSF and the brightness, i.e., the mean photon count per sampling interval per molecule, where we position the molecule at the PSF center, we can relate $I = B \cdot \mathrm{PSF}(\mathbf{r})$, where the PSF is normalized to unity at its center. An analyte located at each position within the confocal volume will contribute a Poisson number of counts based on that position. If we assume that each position is equally likely, i.e., $p(\mathbf{r}) = 1/V$, where V is

the confocal volume, then we may express Equation 3.21 for a single analyte as

$$p(k;V,B) = \frac{1}{V}\int_V \frac{(B\cdot \mathrm{PSF}(\mathbf{r}))^k e^{-B\cdot\mathrm{PSF}(\mathbf{r})}}{k!}\,d\mathbf{r} \tag{3.22}$$

To implement PCH, for each fluorescence burst, we bin photon counts into 10-40 μs sampling intervals [35, 172, 253] and then fit the photon count distribution using Equation 3.22. Deviations from the Poisson fit are indicative of sample heterogeneity [230], which may be either of static or dynamic nature. We can resolve the number of species by fitting a convolution of single analyte distributions, each with a different brightness parameter B_i.

3.3.2.2 Shot Noise Comparison

While PCH models shot noise to predict the photon count distribution for an emission detection channel, other works have sought a shot noise prediction for the energy transfer efficiency. Dahan and co-workers [53, 63] showed that a Beta distribution could approximate an E distribution. In which case,

$$\sigma(E) = \sqrt{\frac{\langle E\rangle(1-\langle E\rangle)}{\langle N\rangle + 1}} \tag{3.23}$$

where $\langle N\rangle = \langle F^{D_{em}}_{D_{exc}}\rangle + \langle F^{A_{em}}_{D_{exc}}\rangle$ is the mean photon count during a specified time interval. By comparing an observed E variance to that predicted by theory, Dahan and co-workers could infer the presence of sample heterogeneity.

Gopich and Szabo subsequently provided a rigorous theory of photon statistics in single-molecule FRET [101]. They established an upper bound for E distribution variance

$$\sigma(E;\langle E\rangle, N > N_T) < \sqrt{\frac{\langle E\rangle(1-\langle E\rangle)}{N_T}} \tag{3.24}$$

where N_T is a minimum photon threshold for a specified time interval. They subsequently extended their theory to include conformational dynamics [103,105] and fluorescence lifetime information [102].

Nir and co-workers recognized that, if one knows the number of photons in a burst, a binomial distribution with success probability $p = E$ describes the expected number of acceptor-emitted photons [239]. Hence, we can use error propagation to obtain the observed E^* distribution standard deviation

$$\sigma(E^*; E, N) = \sqrt{\frac{E(1-E)}{N}} \tag{3.25}$$

where E is the true probability. If we know the experimental burst size distribution, which we readily obtain using burst search algorithms, we can generate a predicted shot noise distribution using Monte Carlo techniques. For each burst i, draw a random variable X_i from a binomial distribution with parameters E and N_i, where N_i is the number of $F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}$ photons. Compute the simulated FRET efficiency $\epsilon = X_i/N_i$, and add to a histogram. Upon comparing the shot noise prediction with the observed distribution, we attribute any dispersion beyond the shot noise limit to heterogeneity.

Probability distribution analysis (PDA) takes shot noise prediction a step further [5, 163, 164, 277]. In this approach, we generate E distributions according to shot noise prediction and a kinetic model and then compare the predicted E distribution to the observed data. By performing a global kinetic parameter search and minimizing the sum of squared error between the data and the PDA prediction, we can find kinetic models which agree with the observed data [277].

While potentially useful, PDA cannot comment on whether a static or dynamic prediction agrees better with experimental data. According to PDA, as long as both a multiple static species simulation and a dynamic species simulation predict an E distribution equally well, the predictions are equally valid. Consequently, Torella and co-workers developed burst variance analysis (BVA) to address exactly this issue. BVA uses the insight of Nir and co-workers to detect, for each burst, E fluctuation variance beyond the shot noise prediction. In combination with PDA, BVA can distinguish whether a static or dynamic prediction better accounts for distribution dispersion. In subsequent chapters, we will discuss in detail both

shot noise prediction and BVA.

3.3.3 Time Series Analysis

3.3.3.1 Time Homogeneous

Fluorescence Correlation Spectroscopy Fluorescence correlation spectroscopy (FCS) is an experimental technique to extract information from time-dependent fluorescence intensity fluctuations [79, 152, 189]. As a fluorescent analyte diffuses through a confocal observation volume, the analyte's fluorescence intensity may fluctuate due to a variety of factors, such as translational diffusion [209,266], ligand-macromolecule binding [353], rotational diffusion [173], conformational dynamics [72, 316], intersystem crossing [342, 343], and excited-state reactions [190, 345]. The goal of FCS is to detect the existence of temporal correlations within photon arrival data. Different processes will generate correlations on different timescales. We may describe each timescale using analytical expressions. When used to fit correlation data, these expressions allow us to extract parameters which provide meaningful information about an underlying process.

The basis for FCS is the normalized temporal correlation function, $R(\tau)$

$$R(\tau) = \frac{\left\langle (X(t) - \langle X \rangle)(Y(t+\tau) - \langle Y \rangle) \right\rangle}{\langle X \rangle \langle Y \rangle} \tag{3.26}$$

where $X = \{x_0, x_1, \ldots, x_{T-1}\}$, $Y = \{y_0, y_1, \ldots, y_{T-1}\}$, T is the total number of data points, and $\langle \cdot \rangle$ denotes the average value. We refer to τ as the *lag*. Commonly, we express Equation 3.26 in terms of fluctuations $\Delta X(t) = X(t) - \langle X \rangle$. $R(\tau)$ then becomes

$$R(\tau) = \frac{\langle \Delta X(t) \Delta Y(t+\tau) \rangle}{\langle X \rangle \langle Y \rangle} \tag{3.27}$$

A correlation function is a similarity measure between two time series as a function of time. When $Y = X$, we refer to Equation 3.26 as the *temporal autocorrelation function* (ACF), which we normalize by the time series average squared. In the case

of autocorrelation, $R(\tau)$ measures self-similarity. For example, consider an analyte diffusing through a confocal observation volume. In the limit of an infinitely small displacement during the interval $\Delta t = t_j - t_i$, i.e., the analyte is effectively stationary during the time interval, we expect to measure similar photon count rates at t_j and t_i. As $\Delta t \to \infty$, the analyte will diffuse away from its position at t_i. Eventually, at some time t_j, we should no longer expect similar photon count rates at t_i and t_j, e.g., if the analyte diffused out of the observation volume. Accordingly, we expect Equation 3.27 to have a maximum value for $\tau = 0$. By computing the fluctuation similarity as a function of time, we can determine how long correlations persist. In our example, correlations would inform us about diffusion coefficients.

If we assume a 3D Gaussian observation volume (see Equation 3.2), we can analytically solve the fluctuation component of Equation 3.27 [7]

$$R_{\text{diffusion}}(\tau) = \frac{\delta}{CV}\left(\frac{1}{1+\frac{\tau}{\tau_{\text{diffusion}}}}\right)\left(\frac{1}{1+(\frac{w_r}{w_z})^2\frac{\tau}{\tau_{\text{diffusion}}}}\right)^{1/2} \tag{3.28}$$

where $\tau_{\text{diffusion}}$ is the analyte's mean diffusion time through the observation volume, C is the analyte sample concentration, V is the observation volume size, w_r and w_z are the PSF radial and axial distance, respectively, and δ is a correction factor which depends on observation volume geometry; for a 3D Gaussian, $\delta = 1/\sqrt[3]{2}$. By inspection, we see that

$$R_{\text{diffusion}}(0) \propto \frac{1}{CV} \tag{3.29}$$

where $\langle N \rangle = CV$ is the average number of analytes in the observation volume at any time.

While we can model the observation volume V theoretically, we typically use FCS to determine the volume's dimensions using a control sample with a well-characterized diffusion coefficient D_{coeff}. We fit Equation 3.28 for parameters $\tau_{\text{diffusion}}$ and $a = w_r/w_z$. From the mean diffusion time $\tau_{\text{diffusion}} = w_0^2/4D_{\text{coeff}}$, we obtain the radial distance w_0 [189]. The axial distance follows from $z_0 = aw_0$, and, if we assume a 3D Gaussian observation volume, we calculate the volume size as

$V = \pi^{3/2} w_0^2 z_0$. We subsequently perform FCS measurements for our experimental sample and fit Equation 3.28, leaving only $\tau_{\text{diffusion}}$ unconstrained.

For other processes which generate fluorescence intensity fluctuations, we need to derive, similar to that of diffusion (Equation 3.28), appropriate models which describe the processes and yield analytical solutions. One such process is triplet-state photophysics. Widengren and co-workers derived a triplet state component for the ACF, which assumed that the fluorescence rate is much faster than the rate of intersystem crossing and triplet relaxation [344].

$$R_{\text{triplet}}(\tau) = 1 + \frac{\langle f_{\text{triplet}} \rangle}{1 - \langle f_{\text{triplet}} \rangle} e^{-\tau/\tau_{\text{triplet}}} \tag{3.30}$$

where $\langle f_{\text{triplet}} \rangle$ is the equilibrium fraction of molecules in the triplet state and τ_{triplet} is the singlet-triplet relaxation lifetime. If we assume that the diffusion time is well-separated from the triplet timescale, the overall correlation is proportional to the product of the respective components

$$R(\tau) = R_{\text{diffusion}}(\tau) R_{\text{triplet}}(\tau) \tag{3.31}$$

Several works have further developed FCS capabilities to address multiple channel cross-correlation [76, 175, 261, 287, 316], in vivo measurement [178], and total internal reflection fluorescence [**?**, 198]; see [97, 121, 189] for further advances. Nevertheless, FCS has limitations. First, FCS is sensitive to optical artifacts which distort observation volume shape and generate artificial fluctuations [129]. Second, FCS requires process timescales to be well-separated. For example, process timescales which overlap with diffusion timescale fluctuations can be difficult to model and identify. Third, FCS is termed a small-ensemble method, in which analyte concentrations are typically > 1 nM. Nanomolar concentrations are too high for standard single-molecule confocal detection and therefore limit our ability to identify sample heterogeneity and isolate molecular subpopulations in complex samples [168, 221].

Although FCS is a versatile experimental and analysis technique, the limitations of FCS help motivate the single-molecule signal analysis methods we develop in later chapters to quantify diffusion timescale dynamics.

Correlation We, in the single-molecule field, frequently cite donor-acceptor anti-correlated fluorescence intensity changes as evidence for FRET-based donor-acceptor interaction and against non-FRET photophysics as the cause of E fluctuations. Few works, however, provide statistical support for their assertion. Most researchers manually search for time series correlations and, depending on the application and additional criteria, retain anti-correlated donor-acceptor time series for further analysis. Such an approach is not ideal.

One method to determine correlation between two random variables X and Y is the Pearson product-moment linear correlation coefficient

$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \sigma_Y} = \frac{\mathrm{E}[(X-\mu_x)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{3.32}$$

where σ_X and σ_Y are the respective standard deviations, and μ_X and μ_Y are the respective mean values. Equation 3.32 is a normalized measure of X-Y covariance. We can modify Equation 3.32 to calculate a sample correlation coefficient r between two time series $X = \{x_0, x_1, \ldots, x_{N-1}\}$ and $Y = \{y_0, y_1, \ldots, y_{N-1}\}$ of finite length N

$$r_{XY} = \frac{\sum_{i=0}^{N-1}(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum_{i=0}^{N-1}(X_i-\bar{X})^2}\sqrt{\sum_{i=0}^{N-1}(Y_i-\bar{Y})^2}} \tag{3.33}$$

where $\bar{X}$ and $\bar{Y}$ are the respective sample means. Equation 3.33 becomes ill-defined, however, for measurements which involve noise and where the corresponding signal-to-noise ratio is small. Consider the situation in which X and Y are $N = 100$ data points in length and uncorrelated except for five data points where X and Y are negatively correlated. If the negative correlations arise from fluctuations which only weakly deviate from the mean, then these negative correlations will have little impact on r_{XY}. In this scenario, the correlations are buried

in the noise, and, if negative correlations are what we hope to identify, r_{XY} would likely fail to provide confirmatory evidence.

The work of Wang and Lu [205, 335] attempted to address this shortcoming and proposed a method called 2-dimensional regional correlation analysis. The basic premise is to compute the cross-correlation between time series segments segmented by lag τ.

$$R_{XY}(\tau, t_i : t_j) = \sum_{t_i}^{t_j} X(t)Y(t+\tau) \tag{3.34}$$

where t_i is the index defining the start of a segment and t_j is the index defining the end of a segment. This calculation is repeated for multiple values of τ to generate a 2-dimensional correlation amplitude histogram. We provide the generated output in Figure 3.4.

We offer a few comments on the method. First, for uncorrelated time series, the method identifies significant correlations. False positives arise due to the fact that the method computes correlation values between short segments, and shorter segments translate to reduced statistics and a greater susceptibility to random correlations. Second, if we consider perfectly anti-correlated time series (see Figure 3.4(b)), the method fails to detect negative correlations. False negatives arise due to the method's bias toward detecting significant deviations from mean values. For small and/or transient fluctuations, local 3-pixel averaging obscures correlations. In other tests, we found the method did identify negative correlations when the fluctuations were large and the time series exhibited clear state behavior (data not shown). These tests, however, beg the question as to the utility of the method, as other methods are arguably more appropriate for time series which show clear states. Finally, the method does not inform us as to correlation significance, but only provides amplitudes. If we simply want an automatic way to identify and isolate time series with significant correlations, the method, in its current form, is unable to provide such a framework.

Hanson and Yang [117, 118] sought to develop a framework which could pro-

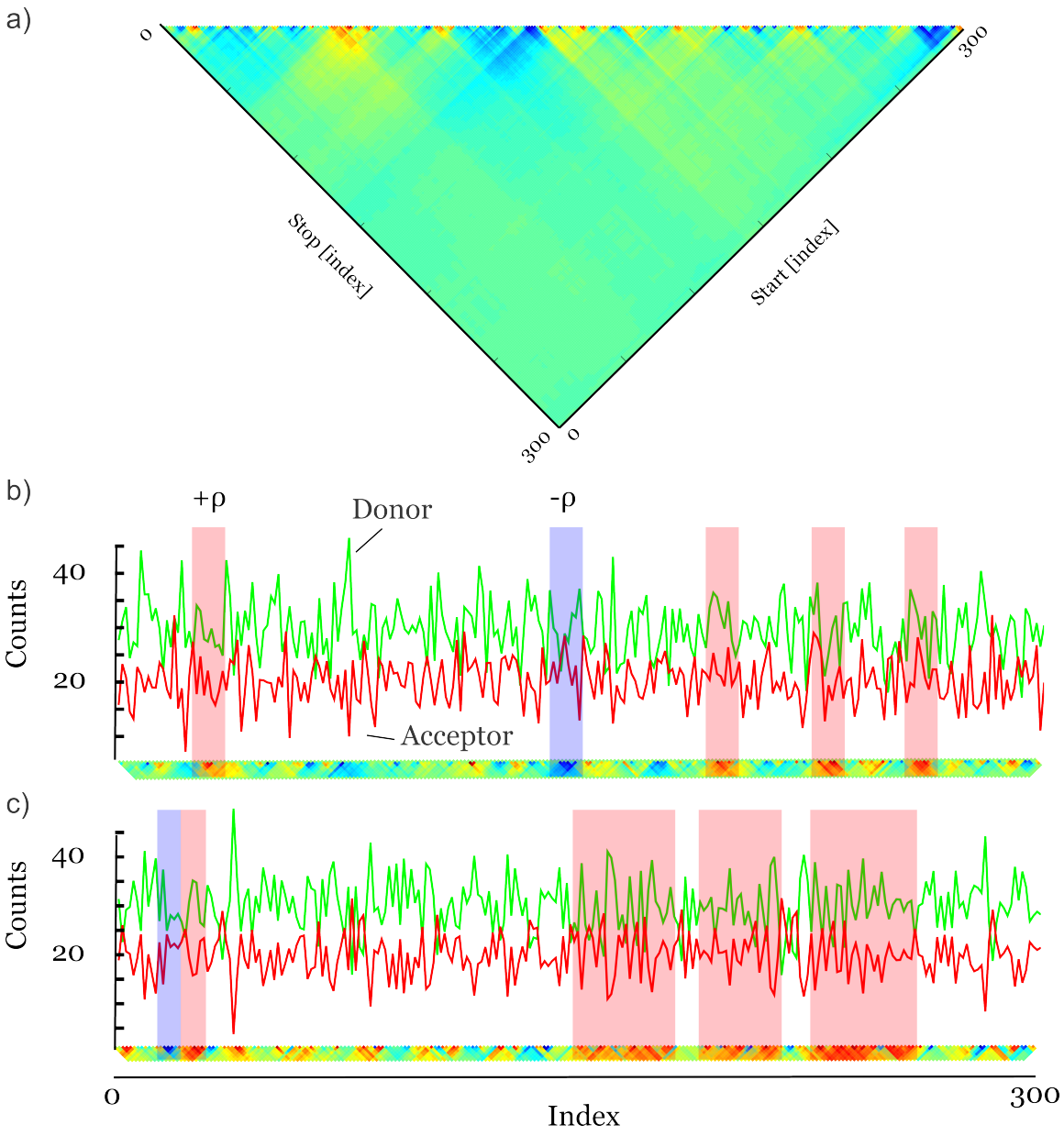


Figure 3.4: 2D Regional Correlation Analysis. We simulated correlated Poisson time series and applied 2-dimensional regional correlation analysis, which calculates a modified cross-correlation function to identify local correlations [335]. a) 2-dimensional correlation amplitude histogram. Cross-correlation between time series segments separated by lag τ (= Stop_{index} - Start_{index}). Cross-correlation measures can suggest both local and lagged correlation. Green indicates no correlation; blue indicates negative correlation; red indicates positive correlation. b) uncorrelated time series, $\rho = 0$. The truncated heat map located below the time series indicates correlations within time series segments which are separated by small τ. Full histogram calculation is omitted as such calculation is not needed for identifying local cross-correlation. For uncorrelated time series, the method frequently detects positive correlations. c) anti-correlated time series, $\rho = -1$. For perfectly anti-correlated time series, the method does not detect anti-correlations, but primarily detects positive correlations. In general, the method is not well-suited to detect transient correlations buried in noise.

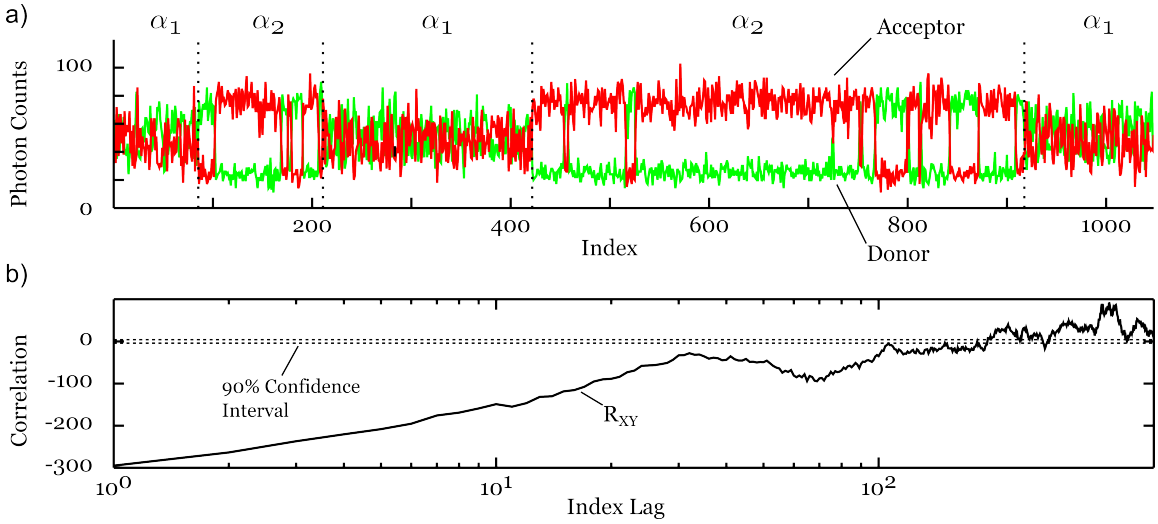


Figure 3.5: Statistical Test for Time Series Cross-Correlation. We simulated a Markov-switching process and performed a statistical test for time series cross-correlation [117, 118]. a) simulated Markov-switching process. Simulated donor photon counts are in green and acceptor counts are in red. b) statistical test for cross-correlation. The cross-correlation function R_{XY} (shown as a black line) exceeds the expected 90% confidence intervals (shown as dotted lines), and thus, the correlation is significant.

vide a statistical test for correlations. Building on the work of Watkins and co-workers [336], their insight was to determine the uncertainties in cross-correlation which arise due to finite-length time series. By determining the uncertainties, they generated confidence bounds against which they could hypothesis test for significant correlations. We demonstrate the results of their approach in Figure 3.5.

We note three qualities of the time series in Figure 3.5(a). i) They show clear anti-correlated behavior. ii) The fluctuation magnitude is large and signal-to-noise ratio large. iii) The time series demonstrate multiple distinct interconversion timescales (in the example, a rapid interconversion regime, α_1, and another slower interconversion regime, α_2). We found these three aspects were required in order for the statistical test to consistently confirm significant correlation. We performed simulations similar to those used to test 2D regional correlation analysis (Figure 3.4), and the statistical test failed to reliably identify correlation. Thus, while providing a statistical test which can identify correlated time series in an automatic fashion, improvements to the method are needed to identify correlations at the measurement temporal resolution.

Finally, we mention the role of correlation functions in single-molecule enzy-

mology. Intra-molecular dynamic heterogeneity arises when an enzyme changes conformation and the conformational change results in altered kinetics. For example, consider the case of an immobilized enzyme binding to diffusing substrates. Suppose the enzyme fluctuates between two conformations which affect its binding (k_{+1}) and unbinding (k_{-1}) rates. Interconversion can occur from either the bound or unbound forms. We denote the rates for enzyme conformation interconversion by γ.

$$\begin{array}{ccc} \mathrm{E}^{\alpha} + \mathrm{S} & \underset{k^{\alpha}_{-1}}{\overset{k^{\alpha}_{+1}}{\rightleftharpoons}} & \mathrm{E}^{\alpha}\mathrm{S} \\ \gamma_{+1} \downarrow \uparrow \gamma_{-1} & & \gamma_{+2} \downarrow \uparrow \gamma_{-2} \\ \mathrm{E}^{\beta} + \mathrm{S} & \underset{k^{\beta}_{-1}}{\overset{k^{\beta}_{+1}}{\rightleftharpoons}} & \mathrm{E}^{\beta}\mathrm{S} \end{array}$$

where $k^{\alpha}_{+1} \neq k^{\beta}_{+1}$ and $k^{\alpha}_{-1} \neq k^{\beta}_{-1}$. A donor-acceptor FRET pair which monitors the enzyme will yield an E time series whose statistical properties change over time (similar to the time series shown in Figure 3.5(a)). The phenomenon which describes enzymatic rate fluctuations over time is called *dynamic disorder*. One way to evaluate dynamic disorder is to calculate the dwell time autocorrelation function [346]

$$R(i) = \frac{\sum_i \Big(\tau_{\text{dwell}}(i) - \langle \tau_{\text{dwell}} \rangle \Big) \Big(\tau_{\text{dwell}}(i+j) - \langle \tau_{\text{dwell}} \rangle \Big)}{\sum_i (\tau_{\text{dwell}}(i) - \langle \tau_{\text{dwell}} \rangle)^2} \tag{3.35}$$

where τ_{dwell} is the dwell time, i is the dwell index, and j is the dwell index lag. In the absence of dynamic disorder, $R(0) = 1$ and $R(i) = 0$ for $i > 0$. In the presence of dynamic disorder, $R(i)$ shows a decay. The initial ($i = 0$) amplitude reflects the variance in rates among the different conformations. The decay time yields the interconversion timescale. Repeated interconversion on the measurement timescale is manifest as a recurring change in a time series' statistical properties over time; this recurrence is called a *memory effect*.

The ability to detect and investigate memory effects remains elusive. In order to accurately calculate an autocorrelation function for a single time series,

many enzymatic turnovers are needed, thus posing stringent measurement requirements. Such requirements exponentially scale with each additional conformational isomer and additional reaction step. Currently, organic fluorophore photobleaching lifetimes are too short to adequately investigate enzymatic dynamic disorder, but continuous improvements in fluorophore photostability, as well as new experimental methods [41], suggest such investigations may be possible in the near future.

Markov Models We have thus far considered analysis methods which generally exploit a time series' known specific properties, e.g., Poisson noise, to estimate time series parameters, e.g., photon count rate. These methods, however, cannot accurately predict a time series' future behavior, as the methods do not learn anything about the generative process. In order to generate predictive models, we must turn to another class of methods which seeks to understand properties of the underlying system generating the time series. Such methods generally fall under the heading of *machine learning*, which is a rapidly growing field at the forefront of computational analysis.

Here, we focus on a subset of machine learning algorithms which we use to learn *Markov models*. Several works have introduced Markov models for single-molecule fluorescence time series analysis [23, 161, 218, 359, 360, 362]. The strength of Markov models lies in the ability to impart our prior knowledge about the system to help guide, structure, and improve analysis performance. Before we apply Markov models to single-molecule data, we first introduce the concept of a Markov model and the machine learning methods we use to learn a time series' statistical properties.

A *model* for a stochastic process, $X = \{X_0, X_1, \ldots, X_{T-1}\}$, is a set of rules characterizing the joint probability distribution between all random variables derived from that process. We define a *stochastic process* as a finite or infinite sequence $X = \{X_t\}_{t\geq 0} = \{X_0, X_1, \ldots\}$ of random variables taking values in a discrete state-space $\mathcal{S}$. The elements of $\mathcal{S}$ are called *states*, $\{S_i\}_{0\leq i\leq N-1}$. X is a Markov chain if

the conditional properties between the outcomes at different t satisfy the Markov property, which we now explain. Consider a time t and the event $X_t = q_t$

$$\{X_0 = q_0, \ldots, X_{t-1} = q_{t-1}, X_t = q_t\}$$

for some sequence of states $q_0, q_1, \ldots, q_t$, where $q_t = S_i$. This sequence is a record of the process history up to and including time t. The conditional probability

$$P(X_{t+1} = q_{t+1}|\underbrace{X_t = q_t, X_{t-1} = q_{t-1}, \ldots, X_0 = q_0}_{\text{entire history}})$$

represents the event probability one time step into the future beyond time t conditioned on the entire past of the process up to t. The conditional probability

$$P(X_{t+1} = q_{t+1}|X_t = q_t)$$

represents the event probability of a future event given only the present. The *Markov property* is satisfied when

$$P(X_{t+1} = q_{t+1}|X_t = q_t, X_{t-1} = q_{t-1}, \ldots, X_0 = q_0) = P(X_{t+1} = q_{t+1}|X_t = q_t) \quad (3.36)$$

for every sequence of elements of $\mathcal{S}$ $(q_0, \ldots, q_t, q_{t+1})$ and for every $t \geq 0$. A random variable sequence with the Markov property is called a *Markov chain*. If X is a Markov chain, and i and j are indexes for states in $\{\mathcal{S}\}$, the conditional probability

$$a_{ij}(t) = P(X_{t+1} = S_j|X_t = S_i) \quad (3.37)$$

is called the *transition probability* from i to j at time t. If the transition probabilities do not depend on time, then $a_{ij}(t) = a_{ij},\ i, j \in \mathcal{S}$, and we say the Markov chain is *time-homogeneous*.

$$a_{ij}(t) = a_{ij}, \qquad \forall\, t \quad (3.38)$$

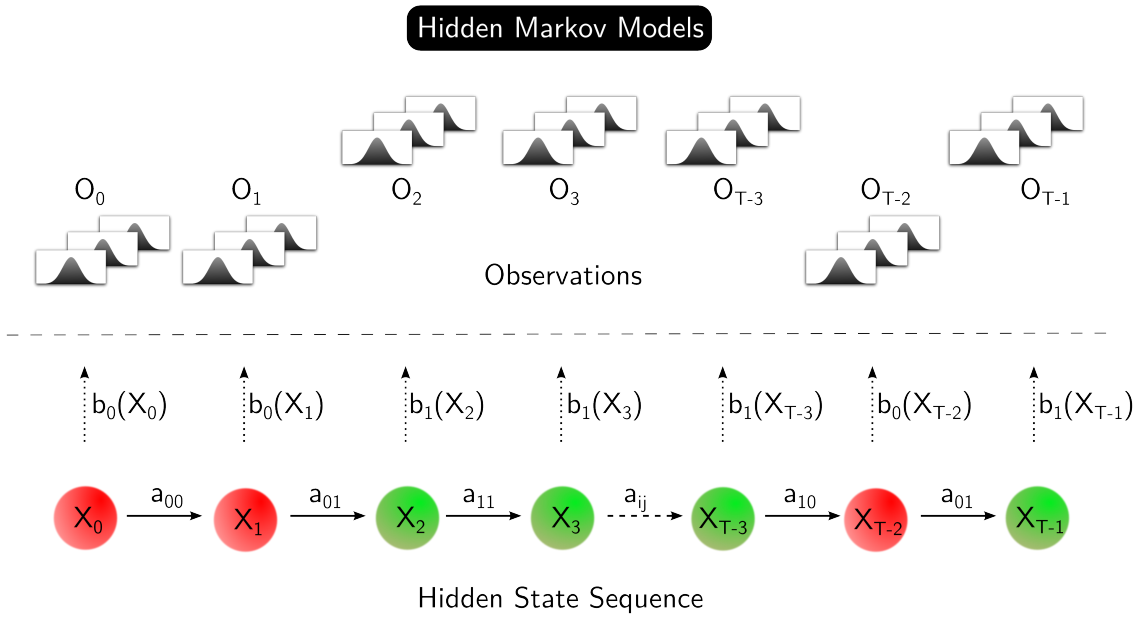


Figure 3.6: Hidden Markov Models (HMMs). A Markov chain consisting of two states (red and green) undergoes state changes according to the transition probabilities $\{a_{ij}\}$. The state sequence $\{X_t\}$ is hidden from the observer (indicated by the dashed line), and the observation sequence is a probabilistic function of the state sequence with state emission probabilities $\{b_j(O_t)\}$. Here, the observation output $\{O_t\}$ is multivariate, with densities given by Gaussian distributions.

The Markov property states that, once an event X_t at the present time is known, all past sequence events have no relevance to conditional probabilities of future events. Or stated more poetically: *the future is independent of the past given the present*.

A *hidden Markov model* (HMM) extends the Markov model to map an observation sequence $\mathcal{O} = \{O_t\}_{t\geq 0}$ to a hidden (or *latent*) state sequence $\mathcal{Q} = \{q_t\}_{t\geq 0}$; see Figure 3.6. The HMM models the observation sequence as a probabilistic function of an underlying Markov chain X. In its most basic form, an HMM assumes each observation O_t is independent of prior observations and solely depends on the hidden state $X_t = q_t$ at time t, $p(O_t|q_t = S_i)$.

We parameterize an HMM as follows [260]

- **initial probability distribution**: $\boldsymbol{\pi} = [\pi_i]$

$$\pi_i = P(X_0 = S_i), \qquad 0 \leq i \leq N-1 \tag{3.39}$$

$\boldsymbol{\pi}$ is a probability vector, and thus satisfies the conditions $0 \leq \pi_i \leq 1 \; \forall \; i$ and $\sum_i \pi_i = 1$. π_i is the probability that the Markov chain starts in state S_i.

- **transition probability matrix**: $\mathbf{A} = [a_{ij}]$, where $\sum_i a_{ij} = 1$, i.e., the matrix is row stochastic.

- **observation probability distribution**: $\mathbf{B} \sim \mathfrak{N}(\mathcal{O}, \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm})$ where $0 \leq j \leq N-1$, $\mathfrak{N}$ is a log-concave symmetric probability density function, $\mathcal{O}$ is the observation sequence to be modeled, and $\boldsymbol{\mu}_{jm}$ and $\boldsymbol{\Sigma_{jm}}$ are the mean vector and covariance matrix, respectively, for the mth mixture component in state j.

Given a model parameterized by $\lambda = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{B})$, we calculate the probability of $\mathcal{O}$ by computing the sum over all possible hidden state sequence probabilities

$$\mathcal{L}(\mathcal{O}|\lambda) = \sum_{\mathcal{Q}} \pi_{q_0} a_{q_0 q_1} b_{q_1}(O_1) a_{q_1 q_2} b_{q_2}(O_2) \cdots a_{q_{T-2} q_{T-1}} b_{q_{T-1}}(O_{T-1}) \tag{3.40}$$

We traditionally decompose learning HMMs into three problems [260]

- Efficient computation of $\mathcal{L}(\mathcal{O}|\lambda)$.

- Selecting an 'optimal' state sequence $\mathcal{Q}$ which explains the observation sequence $\mathcal{O}$.

- Determining model parameters λ which maximize $\mathcal{L}(\mathcal{O}|\lambda)$.

We efficiently solve the first problem using the *forward-backward algorithm* [15]. We solve the second problem via the *Viterbi algorithm* [87, 326]. The Viterbi algorithm determines the state sequence $\mathcal{Q}$ which maximizes $p(\mathcal{Q}|\mathcal{O}, \lambda)$. We compute a maximum likelihood estimate for the model parameters λ using the *Baum Welch algorithm* [14, 16, 17]. The Baum-Welch algorithm belongs to the class of *expectation-maximization algorithms* [62, 255]. Figure 3.7 shows HMM analysis results for simulation data and demonstrates that an HMM can correctly infer a hidden state trajectory.

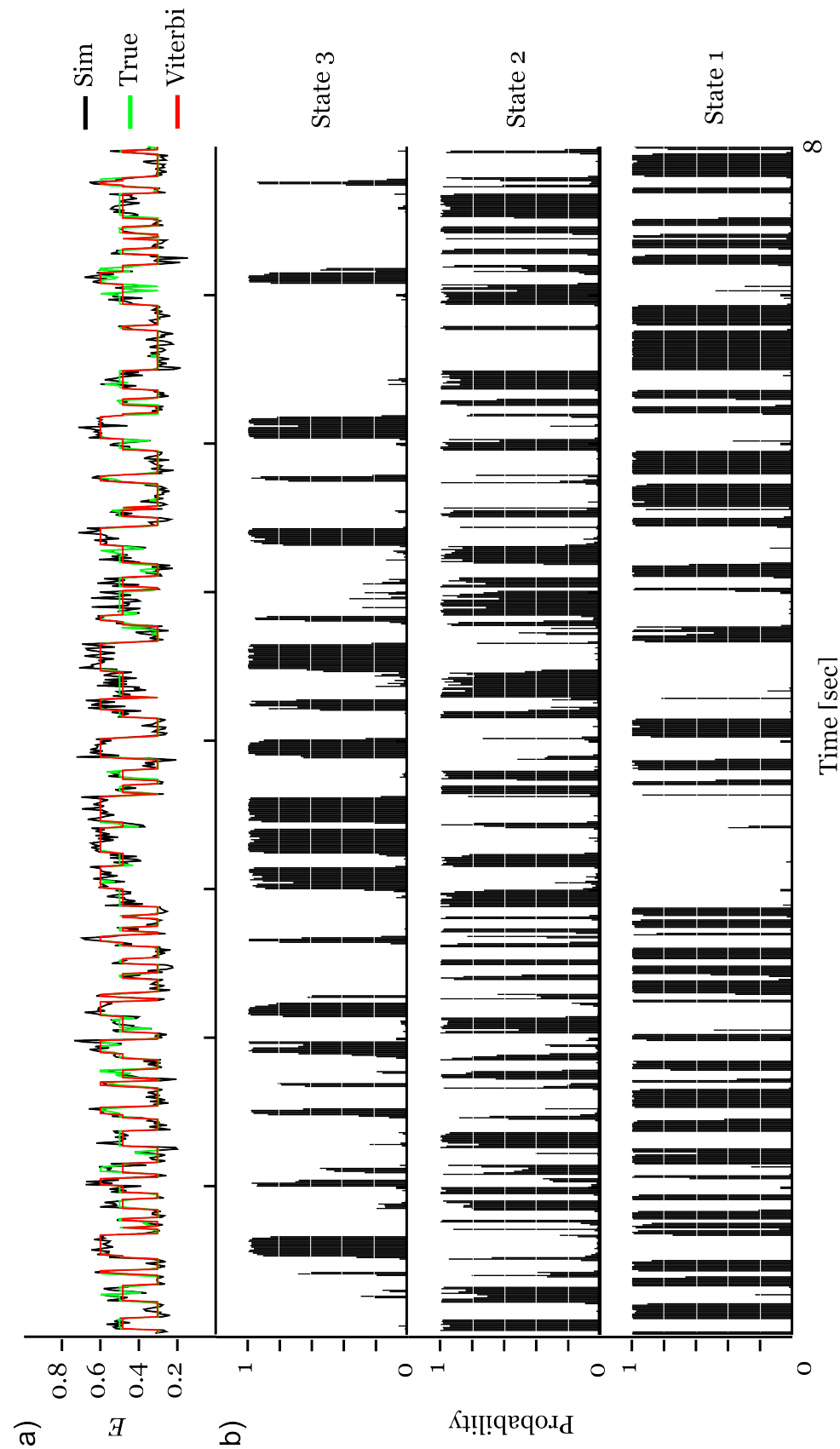


Figure 3.7: Hidden Markov Model (HMM) Analysis. We simulated a three-state Markov process using kinetic Monte Carlo methods: $\{(E, \sigma(E)) : (0.3, 0.04), (0.5, 0.06), (0.6, 0.05)\}$, $k_{01} = 25\ \text{s}^{-1}, k_{10} = 33\ \text{s}^{-1}, k_{12} = 33\ \text{s}^{-1}, k_{20} = 25\ \text{s}^{-1}$, $T = 2000$ data points, and 10 ms temporal resolution. a) HMM analysis. Simulated trajectory is green; Gaussian noise corrupted time series is black; inferred Viterbi (global decoding) path is red. The HMM is able to identify most transitions correctly. Missed transitions are transient fluctuations. For kinetics near the measurement temporal resolution, HMM performance degrades significantly. b) Local decoding. The forward-backward algorithm scores transitions. An HMM is confident when a state probability is close to unity for a data index. For fast transitions, we see that probabilities are more evenly distributed between multiple states, reflecting uncertainty in state assignment.

An additional problem is model selection, i.e., determining the correct number of states. A standard means to determine model complexity is to use an information criterion [197, 218]. Multiple information criteria exist, but our tests indicate that the *Bayesian information criteria* (BIC) generally fares best

$$\text{BIC} = -2 \ln \mathcal{L}(\mathcal{O}|\lambda) + k \ln T \tag{3.41}$$

where k is the number of free parameters and T is the data length. The second parameter acts as a penalty term for increased complexity. As more parameters k are introduced, the likelihood that the model λ generated the observations $\mathcal{O}$ must be sufficient to offset the increased penalty incurred. Accordingly, the best model is that model which minimizes the BIC. An alternative method for model selection is to use the recently proposed *variational Bayes expectation maximization algorithm* [18] which maximizes the model evidence and obviates the need for an information criterion. The best model is that model which maximizes the evidence [25, 242]. More recent works have developed HMMs for multivariate inference [204, 322] and model comparison [107].

While powerful, HMM learning has disadvantages. i) HMMs are most effective when state lifetimes are 10x greater than the measurement sampling interval. For fast kinetics near the measurement temporal resolution, HMM performance degrades significantly, as we will show in a subsequent chapter. ii) States must be clearly resolved. Basic HMMs are unable to distinguish degenerate states [174]. iii) HMM learning is a global method and not amenable to real-time analysis. iv) HMM computation scales quadratically with the number of states. For complex systems, HMM analysis quickly becomes computationally expensive. v) HMM performance is best when learning is seeded with appropriate prior probability distributions (*priors*). Poor initial parameters can lead an HMM to become trapped in a local, rather than global, maximum. For situations in which little is known about the data's statistical properties before analysis, other methods should be performed to obtain parameters to inform model priors. vi) HMM learning is most

effective when states are populated at least 10 times during a time series. For slow kinetics, fluorophore photobleaching is limiting. vii) Simultaneous HMM analysis across multiple time series assumes kinetic homogeneity. For a kinetically heterogeneous population, a global HMM will learn the mean population properties. The inferred parameters may lead researchers to draw incorrect conclusions regarding system kinetics [107].

As with all analysis methods, we should exercise care in how we apply HMM learning. We strongly make three recommendations. First, run experimentally accurate simulations to understand the limitations and requirements of HMMs. Second, perform complementary analysis methods, such as change-point detection and correlation analysis. Third, perform experimental controls. This is particularly important to ensure that distance-independent photophysics do not contribute to FRET fluctuations. While an HMM may infer a state change, an HMM is silent as to the transition's physical meaning [21]. Inference difficulties of this kind are discussed at length in subsequent chapters.

3.4 Software

Time series analysis has a rich history in statistics, signal processing, econometrics, and mathematical finance. This history has benefited single-molecule researchers, who have adopted analysis methods developed elsewhere to analyze single-molecule data, such as clustering techniques [321, 338], maximum likelihood methods [102, 104, 116, 137, 250, 325, 348], Bayesian inference [11, 26, 185, 187, 217, 308, 336, 337], Markov models [3, 19, 45, 46, 153, 201, 222, 223, 231, 258, 259, 292, 299, 306], wavelet techniques [307, 308, 351], image analysis [66, 227, 309], and stellar photometry [135]. While these methods have advanced the field, new experimental techniques and data types require that we continually develop new analysis methodologies to manage, process, analyze, and visualize the increasingly complex data sets. Importantly, we must continue to refine existing, as well as develop new, analysis methods which address the unique characteristics and complica-

tions of single-molecule fluorescence data, such as fluorophore photophysics and photochemistry, sources of experimental noise, and reaction trajectory inference.

Single-molecule approaches provide unprecedented access to biological systems; however, their utility is limited if we lack methods and platforms to extract meaningful information. A significant challenge going forward is developing a standard framework across which single-molecule time series are visualized and analyzed. Within the field of single-molecule fluorescence, researchers employ a patchwork of data analysis methods, and too few works publicly release software implementing the methods. This patchwork varies from researcher-to-researcher, preventing reproducibility and comparison.

While single-molecule tailored software packages do exist [25,107,135,174,218, 223, 282, 307], they fail to provide comprehensive solutions, each performing an often singular task without interfacing with other packages or platforms. Additionally, existing software fails to be intuitive, especially for the non-technical user, and often requires users to be comfortable operating at the command-line–something which cannot be expected as single-molecule techniques expand to fields less computational. Third, most software is neither interoperable or open-source. Proprietary software platforms such as LabVIEW and MATLAB® are closed development systems and restrict open collaboration and engagement with the much larger computing and development community. Finally, the vast majority of single-molecule researchers lack sufficient training in computing, design, and software development to produce quality software solutions. These obstacles subsequently prevent proposed solutions from being scalable.

The lack of satisfactory solutions for single-molecule time series analysis led us to develop extensive in-house software. Previous work in our laboratory applied stellar photometry techniques for single-molecule identification, detection channel association, and processing [135]. This work solved the problem of data extraction, but we further needed visualization and analysis methods to extract information content and infer underlying reaction trajectories. I designed and wrote

Seneca, a single-molecule fluorescence data analysis package, to address our needs. *Seneca* implements many of the techniques discussed above, includes simulation facilities, enables extensive visualization, and provides a user interface to access the methods.

Future work will involve design and development of a web interface, which will enable analysis to occur in a web browser and be device agnostic. We aim to leverage the burgeoning open-source development community and facilitate broader engagement within and beyond the single-molecule field. While not discussed in this thesis, our work will form the basis of a future publication.

3.5 Summary and Outlook

The information buried in single-molecule time series has immense potential to reveal fundamental mechanistic insights into biological behavior, both structurally and kinetically. In this chapter, we reviewed experimental and analysis techniques to extract information content. Each method presented its advantages and disadvantages, particularly in terms of data requirements necessary for information extraction.

For example, hidden Markov methods require i) ergodicity such that the underlying kinetic topology is manifest as a connected graph without missing edges, ii) multiple state transitions allowing for repeated single state observation, and iii) prolonged state occupancy. These constraints impose significant experimental burdens in terms of total observation times, system kinetics, and emitted fluorescence intensity, and further place an upper bound on accessible kinetic regimes. With regard to HMMs, current fluorophores do not emit sufficient photon numbers to access kinetic timescales beyond 1 transition every 10 milliseconds.

Watkins and Yang used information theory to demonstrate that the best temporal resolution one can hope to achieve in order to resolve single-molecule dynamics with a 10% relative measurement error is 1ms [337]. If we use the Rayleigh criterion as a proxy for our ability to distinguish conformations having different

inter-fluorophore separations, this means that, for 1 millisecond kinetics, conformational states must differ by at least 2nm (i.e., $\Delta E \approx 30\%$). In their work on protein folding, Chung and co-workers went beyond the mentioned limits to infer a kinetic upper bound of 200μs by exploiting unique features of their experimental design and thus extended measurement resolution by an order of magnitude [43]. As a general rule, however, sub-millisecond kinetics cannot be directly resolved.

As a consequence, the capacity for confocal scanning microscopes to record photon arrival times does not currently provide a distinct advantage over TIRF microscopy in its ability to resolve conformational dynamics at kinetic timescales below 1ms. For both experimental designs, inferring reaction trajectories through time series idealization techniques requires millisecond kinetics. Nevertheless, as methods such as PDA and E-BVA show, with only 5 photons, we can still infer conformational dynamics which occur at timescales below the millisecond regime. Such methods reside at the limit of our ability to extract information content and to probe single-molecule structural and kinetic behavior.

While several works have provided analysis methods capable of data reduction and parameter extraction, each method has its strengths and limitations, and further analysis improvements and developments are necessary. In particular, methods are needed which better address single-molecule data complexities and more fully account for physical constraints and processes. Complex time series arise due to multiple factors. Multiple interconverting structural and chemical states and intermolecular interactions generate time series fluctuations. Competing pathways to fluorescence, such as intersystem crossing, quenching, and photobleaching, lead to abrupt changes in a time series' statistical properties. Discrete data sampling, either experimentally (TIRF) or through binning, leads to complicating quantization effects. In Förster resonance energy transfer (FRET) measurements, distinguishing true distance-related FRET fluctuations from non-distance-related photophysics presents significant challenges.

Experimental design must minimize complications in order to detect molec-

ular processes and to reliably extract information content. For example, single-molecule fluorescence experiments should exploit the information contained in crystal structures to optimize fluorophore placement. New technologies now permit chemical- and site-selective biomolecule-labeling using unnatural amino acids [31, 32, 39, 279]. These methods inform experimental design such that fluorescent probes i) minimally perturb biomolecular activity, ii) have unrestricted rotational freedom, and iii) in FRET measurements, are optimally separated to maximize distance-related FRET fluctuations.

Nevertheless, even if we minimize complications and ensure maximum signal-to-noise, we still need analysis methods for data reduction. As experimental data becomes more complex, richer, and more varied, we need analysis methods which are better informed about the underlying processes which generate observed data. This is particularly true in order to deconvolve fluorescence intensity fluctuations which report on biological activity from photophysical- and photochemical-related fluctuations that obscure the relevant signal. To this end, we present in the chapters which follow analysis methods that detect and identify time series fluctuations and their photophysical origins.

Contributions

I performed all simulations and evaluated method performance. All writing and figures are my own.

Chapter 4

Burst Variance Analysis (Efficiency)

Abstract

Burst variance analysis (BVA) is a method that detects FRET dynamics by comparing the time-dependent FRET standard deviation of individual molecules diffusing through a confocal volume to the standard deviation expected from shot noise alone. Recent work [314] used both simulations and experiments to demonstrate that BVA can distinguish between static heterogeneity, i.e., multiple shot-noise-limited species, and dynamic heterogeneity, i.e., individual species fluctuating between two or more distinct FRET states. Here, we use BVA to analyze RNA polymerase (RNAP) open complex FRET distributions and to detect open complex conformational dynamics. Our results confirm FRET dynamics, but we report a significant limitation of BVA in its ability to distinguish between FRET fluctuations due to conformational dynamics and FRET fluctuations due to distance-independent photophysics, such as those attributable to fluorophore blinking. We discuss possible sources of FRET distribution broadening beyond the shot noise limit and state the need for both experimental and statistical measures capable of distinguishing FRET fact from artifact.

4.1 Introduction

Burst variance analysis (BVA) attempts to detect dynamic fluctuations in Förster Resonance Energy Transfer (FRET) as individual molecules emitting donor and acceptor fluorescence diffuse through a confocal volume. Hereafter, we refer to this method as E-BVA to distinguish the method from additional BVA applications introduced in this thesis. Motivations for E-BVA reside in the attempt to distinguish between static and dynamic heterogeneity and to explain distribution broadening

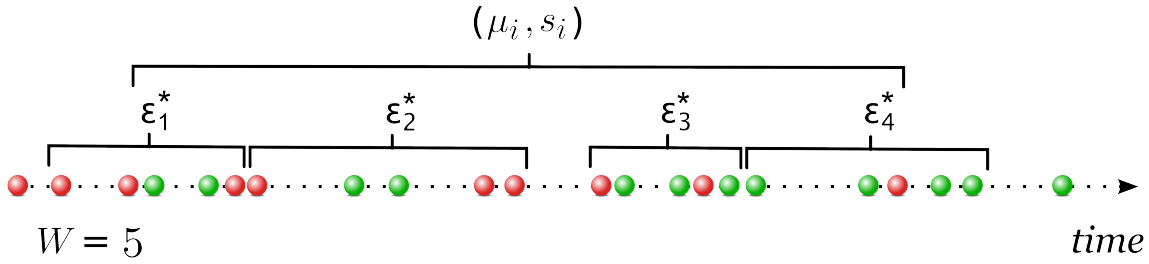


Figure 4.1: FRET Burst Variance Analysis (E-BVA) schematic. FRET sample mean, μ_i, and standard deviation, s_i, calculation. We take M samples of n consecutive photons (in the diagram, $M = 4, n = 5$) from each burst. We then calculate FRET, ε_j^*, for each n-photon window, W, and μ_i and s_i are the sample mean and standard deviation, respectively, of ε_j^* over all M. Hence, for each burst, i, we have a single mean and standard deviation pair (μ_i, s_i). Green sphere: donor photon; red sphere: acceptor photon.

in confocal histograms. Existing methods proved unable to adequately address the static-dynamic distinction, necessitating a burst-by-burst analysis [314].

As detailed by Torella and co-workers [314], E-BVA employs a statistical criterion to compare observed experimental standard deviations to simulated standard deviations generated according to theoretical prediction. Observed systems exceeding the strict confidence bounds are considered dynamic, where dynamics are presumed to arise due to state interconversion for a sample species. We note that E-BVA is not conclusive, however, in ascertaining whether observed dynamics arise due to conformational fluctuations alone.

Depending on experimental conditions and setup, fluorescent dye quenching, dye isomerization, or optical alignment introduce artifacts which can obfuscate the presence of underlying conformational dynamics. This stated, what E-BVA can validate is distribution broadening due to species dynamics, whatever those dynamics might be, as opposed to static heterogeneity, where sample species are in accordance with theoretical prediction. Hence, E-BVA is most effective as one of a suite of tools and experimental controls designed to determine system heterogeneity and effectively eliminate artifacts.

4.2 Algorithm

E-BVA implementation is relatively straightforward (see Figure 4.1):

1. For each burst, bin photons originating during donor excitation by arrival time into non-overlapping bins where the bin size, i.e., number of photons, is defined by window W.[1]

2. Calculate the FRET, ε^*_{ij}, for each bin j in burst i.

$$\varepsilon^*_{ij} = \frac{n_a}{n_d + n_a} \tag{4.1}$$

   $n_x =$ number of photons; a: acceptor; d: donor

3. Calculate the sample mean, μ_i, and standard deviation[2], s_i, of ε^* across all bins, M_i.

$$\mu_i = \frac{1}{M_i} \sum_{j=0}^{M_i-1} \varepsilon^*_{ij} \tag{4.2}$$

$$s_i = \sqrt{\frac{1}{M_i - 1} \sum_{j=0}^{M_i-1} (\varepsilon^*_{ij} - \mu_i)^2} \tag{4.3}$$

4. For all bursts with similar mean FRET $\{\mu(E^*) : a \leq \mu(E^*) < b\}$, calculate the partition (cluster) standard deviation, $\sigma(E^*)$, across all bins for all considered bursts.

$$\mu_k(E^*) = \frac{1}{M_k} \sum_{\substack{i \text{ where} \\ a \leq E < b}} \sum_{j=0}^{M_i-1} \varepsilon^*_{ij} \tag{4.4}$$

$$\sigma_k(E^*) = \sqrt{\frac{1}{M_k - 1} \sum_{\substack{i \text{ where} \\ a \leq E < b}} \sum_{j=0}^{M_i-1} (\varepsilon^*_{ij} - \mu_k(E^*))^2} \tag{4.5}$$

[1] E-BVA is suitable for either continuous wave or ALEx measurements. Note, however, for E-BVA, we disregard ALEx acceptor excitation photons, concatenate all donor excitation photons, and bin neighbors regardless of arrival time. This means artificial change-points may occur (e.g., during one donor excitation, $\varepsilon \sim 0.8$ and, during the next donor excitation, $\varepsilon \sim 0.2$). While omitting arrivals times is an issue if we want to determine an underlying state transition sequence, here, we only want to determine if a molecule is dynamic. E-BVA is not designed to provide exact timescales, but determines only whether dynamics are present over the timescale of interest (the time a molecule spends diffusing through the confocal volume).

[2] Note: Torella and co-workers [314] propose using a biased estimate for the sample standard deviation (although they do not justify its use). Here, I use the appropriate unbiased estimate throughout. This means that my sample variance calculations will be higher than as computed by their algorithm, but my conclusions should remain the same, however, as I use unbiased estimators for both the observed sample variance and the expected sampling simulation.

where $M_k = \sum M_i$ is the total number of windows for bursts satisfying the conditions of the kth partition.

5. Use Monte Carlo techniques[3] to simulate the expected sampling distribution.

$$\mu_k(E) = \frac{1}{M_k} \sum_{\substack{i \text{ where} \\ a \leq E < b}} \sum_{j=0}^{M_i - 1} \frac{F_A^{ij}}{n} \tag{4.6}$$

$$\sigma_k(E) = \sqrt{\frac{1}{M_k - 1} \sum_{\substack{i \text{ where} \\ a \leq E < b}} \sum_{j=0}^{M_i - 1} \left(\frac{F_A^{ij}}{n} - \mu_k(E) \right)^2} \tag{4.7}$$

where F_A^{ij} is a random variable drawn from a binomial distribution with n trials (i.e., the number of acceptor photons in a window of n photons; $n = W$) and success probability E; define the resulting distribution as $P_{MC}(\sigma(E))$.

6. Use a Bonferroni correction for multiple hypothesis testing [1] and compare the upper tail confidence interval s_E^{CI} to the observed $\sigma(E^*)$.

$$\left(1 - \frac{\alpha}{K}\right) = \int_0^{s_E^{CI}} P_{MC}(\sigma(E))\, d\sigma \tag{4.8}$$

where K is the number of partitions (clusters) and α is the confidence level ($\alpha = 0.001$).

7. Determine if the observed standard deviation exceeds the strict confidence bounds for static behavior and is thus indicative of dynamics.

4.3 Biological System

Transcription is the process of synthesizing RNA from a DNA template. The process is the first step in gene expression and a highly regulated process. In *Escherichia coli* and other bacteria, transcription is initiated after RNA polymerase

[3] Monte Carlo methods are a class of computational algorithms which rely on repeated random sampling, e.g., repeatedly drawing samples from a proposed population distribution.

(RNAP) binds to a DNA promoter sequence, a process facilitated by transcription factors known as sigma (σ) factors. For bacterial promoters controlled by the main sigma factor, σ^{70} , the RNAP-σ^{70} complex (referred to as the *holoenzyme*) binds to -10 and -35 promoter elements [120], melts ~14 base-pairs (bp) in the DNA surrounding the main transcription start site, and forms the catalytically active RNAP-DNA open complex (RP_o). See Figure 4.2.

RNAP can initiate transcription from multiple positions nearby to the main transcription start site, both in prokaryotes and eukaryotes [30, 158, 203]. In bacteria, RNAP predominantly initiates at purines located 4-to-12 bp downstream of the -10 element[4]. Start site preferences seem suitable targets for gene regulation, as 5′ transcript end variation apparently affects both secondary structure formation inhibiting translational initiation [58, 59] and transcript stability [80]. Moreover, start site variation may affect the extent of abortive initiation or transcriptional slippage, both of which could influence promoter initiation frequency [30, 166, 169, 264, 347].

Early observations of *lac* operon mRNA 5′ end heterogeneity led to the hypothesis that RNAP initiates from multiple start sites due to transient DNA template strand binding [30], where RP_o single-stranded DNA (ssDNA) flexibility and dynamics permit transcription start site sampling by the polymerase active site. Subsequent structural work on RNAP-σ^{70} and RP_o [232] did not reveal any obvious RNAP conformational changes that could account for start site sampling, consistent with the ssDNA hypothesis. Additionally, proposed DNA scrunching mechanisms suggest ssDNA flexibility associated with possible looped DNA structures within the RP_o transcription bubble [169, 264].

Although proposed thirty years ago, the ssDNA dynamics hypothesis has no direct evidence confirming its validity. Obtaining direct evidence is challenging for several reasons: (1) experimental difficulty in observing subtle ssDNA conformational changes within the RP_o transcription complex, (2) uncertainty in the timescale of conformational changes, and (3) RP_o complex preparation hetero-

[4]The -10 element extends from -7 to -12 bp upstream of the +1 start site.

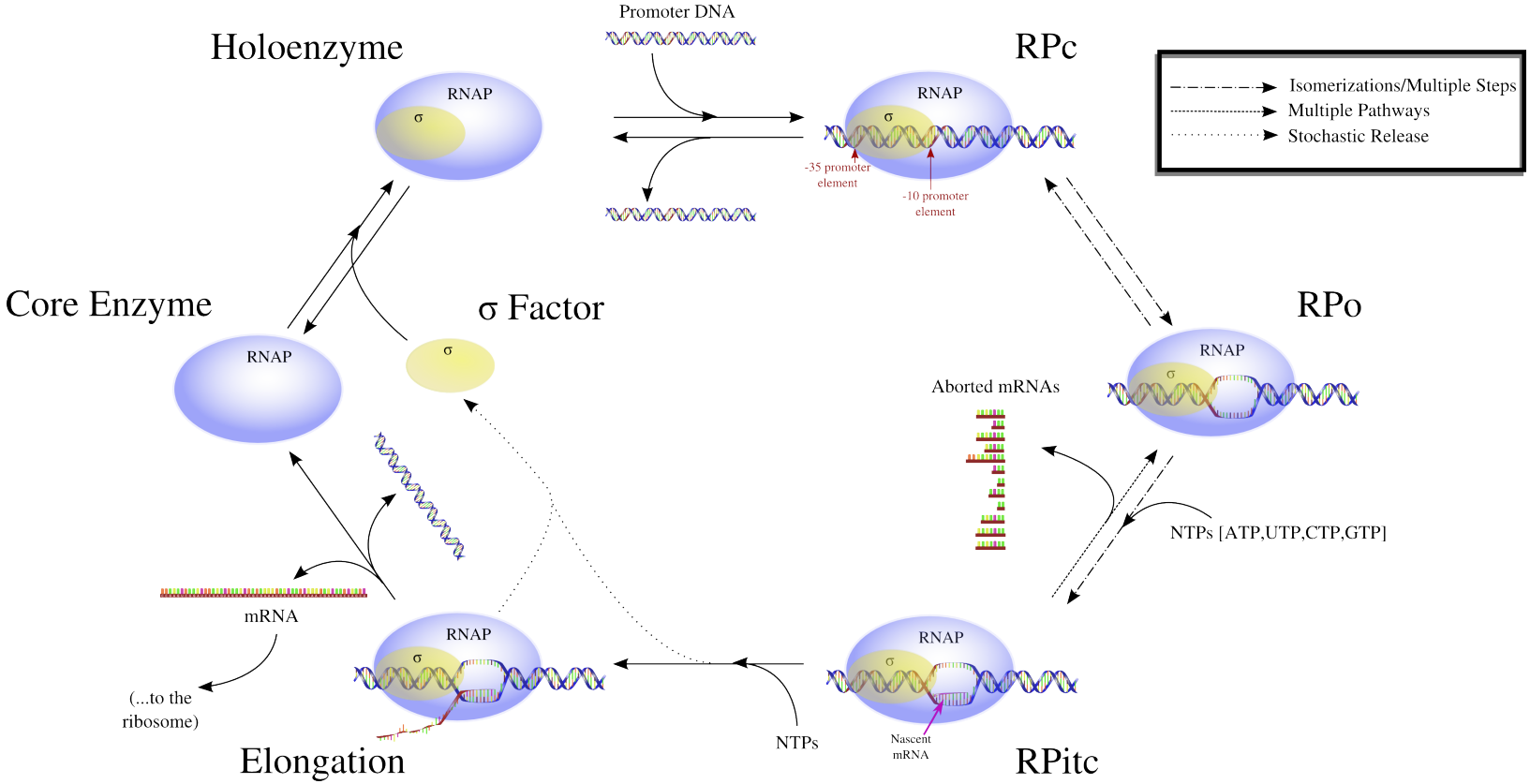


Figure 4.2: Bacterial Transcription Cycle. i) RNA polymerase (RNAP) binds a transcription factor termed a σ factor to form holoenzyme. Holoenzyme subsequently binds to promoter DNA, forming a closed DNA complex, RP_c. ii) Structural isomerizations lead to promoter DNA unwinding from -12 to +2 to create a 14 basepair transcription bubble. The resulting RNAP-promoter open complex is termed RP_o. iii) As the holoenzyme remains fixed on the promoter, the complex proceeds to produce aborted RNA transcripts which are typically less than 9 nucleotides. This reiterative process is termed abortive initiation, and the initial transcribing complex is termed RP_{itc}. iv) After synthesizing a sufficiently long transcript, the σ factor breaks contact with the promoter, and RNAP enters elongation. The complex is now termed an elongation complex, RD_e, and proceeds to transcribe the gene by translocating along the DNA template. At some point during RNAP release and escape, the σ factor stochastically dissociates and is not needed during elongation. v) At a gene's end, RNAP encounters a termination signal, causing RNAP to dissociate and thereby release the DNA template and RNA transcript. RNAP is then free to bind another σ factor and re-enter the transcription pathway. Key: RNAP is blue; σ factor is yellow; -35 and -10 promoter elements are red on the promoter DNA sequence; the RNA transcript is red; arrows indicate structural changes and reaction pathways.

geneity. Single-molecule Förster resonance energy transfer (smFRET) methods, such as confocal microscopy of diffusing molecules and total internal reflection fluorescence (TIRF) microscopy of surface-immobilized molecules, promise to address these challenges by characterizing nanometer-scale dynamics at the millisecond timescale [275,277,314].

Previous studies used fluorophore FRET pairs to monitor transcription initiation conformational changes [169]. We adopt a similar methodology by placing fluorophore pairs at different regions within promoter DNA in hope of detecting transcription bubble formation and DNA movements associated with start site sampling. With one fluorophore upstream of the main transcription start site and another downstream, transcription bubble expansion and compaction as RP_o samples nearby upstream and downstream start sites should, in principle, result in detectable FRET changes, provided the changes are large enough[5]. And if sampling occurs at the millisecond timescale, E-BVA provides a means to account for distribution broadening beyond the shot noise limit, either due to multiple static species[6] or dynamic species[7].

To probe RP_o dynamics, we designed three double-stranded DNA (dsDNA) promoter sequences, differing only in the fluorophore labeling positions; see Figure 4.3.[8] For the promoter, we used a *lac*CONS+2 promoter, a derivative of the *E. coli lac*UV5 promoter, with the +2 corresponding to two differences: a single bp substitution in the -35 element and a single bp deletion in the spacer region located between the -35 and -10 elements.[9] We designed two constructs, (-15Cy3B, +15ATTO647N)*lac*CONS+2 and (-5Cy3B, -3ATTO674N)*lac*CONS+2 , to directly sense RP_o dynamics [50,169], and a third construct, (-15Cy3B, -25ATTO647N)*lac*CONS+2

[5]Note that, in the linear FRET regime, i.e., $E \in [0.2, 0.8]$, a $\Delta E \approx 0.05$ corresponds to a $\Delta R \approx 3\mathring{A}$.

[6]e.g., RP_o does not actually sample, but rather, different RP_o have preferred start sites and the start site distribution previously observed is just the distribution of the static species

[7]i.e., RP_o actively samples start sites, inter-converting between multiple conformations

[8]For sequences and labeling schemes, see Appendix D.1.

[9]These two differences create consensus sequences [228].

, to serve as a static control [169].[10,11] We hereafter refer to these constructs as (-15,+15), (-5,-3), and (-15,-25).

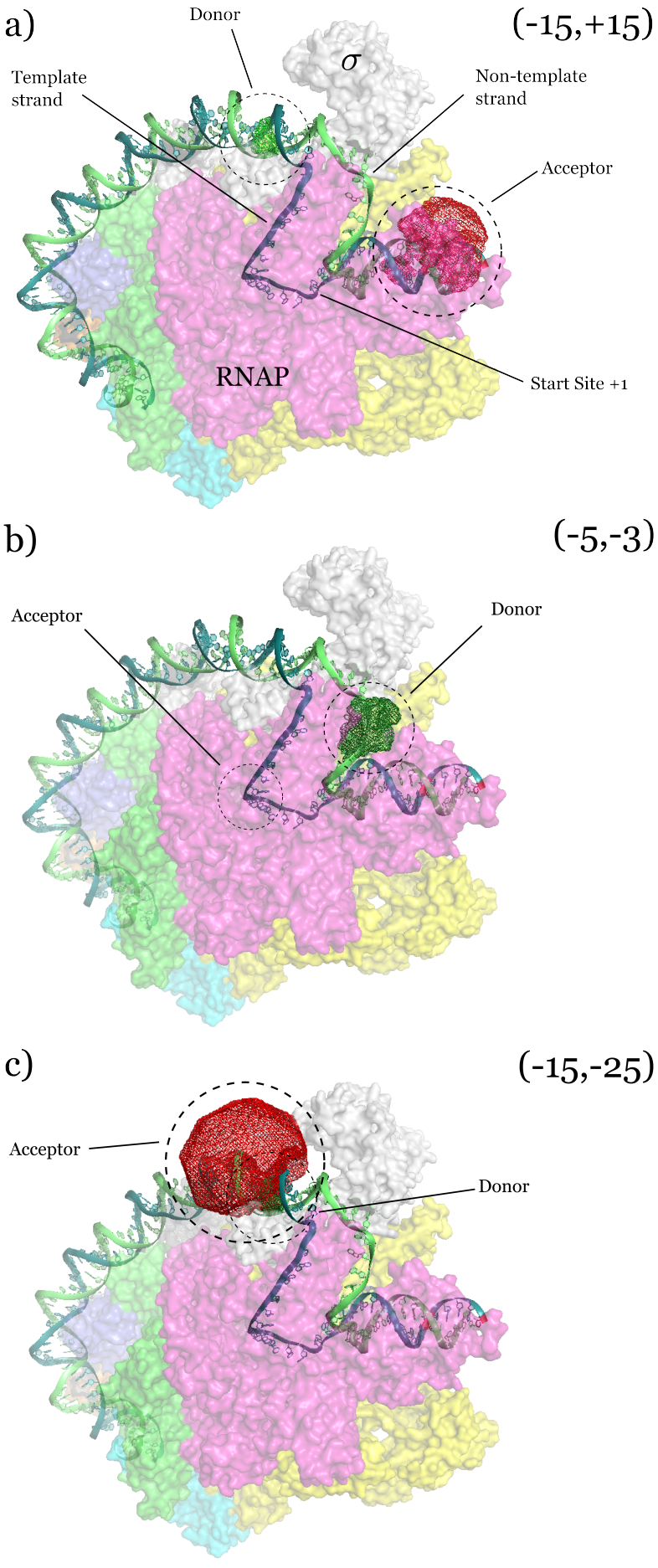
We note several considerations with regard to our constructs. Ideally, a fluorophore labeling position should ensure minimal fluorophore-protein interaction and provide a stable local environment. This precaution aims to promote resonance energy transfer at the expense of other photophysical and photochemical pathways. When such precautions are not possible due to distance constraints or protein functionality, we should be cautious in our interpretation of FRET observations and control for distance-independent photophysics which may obscure a conformational dynamics FRET signal.

According to accessible volume (AV) modeling, each construct has one or more fluorophores which contact the polymerase. Hence, although our labeling posi-

[10]Key: ($\pm x$ D, $\pm y$ A) are the labeling positions for the FRET pair. x and y are the positions relative to the main transcription start site +1, with negative positions upstream and positive positions downstream. D and A correspond to donor and acceptor fluorophore, respectively.

[11]Dynamics are not known or suggested to occur within the promoter spacer region, i.e., the region between -10 and -35 promoter elements. Therefore, FRET pairs placed within this region are not expected to exhibit dynamic fluctuations and should provide a static control.

Figure 4.3 ***(following page)*****:** Monitoring RNAP-Promoter Open Complex Dynamics. Fluorophore labeling schemes to monitor RNAP-promoter open complex (RP_o) dynamics, where the green mesh is the donor (Cy3B) accessible volume and the red mesh is the acceptor (ATTO647N) accessible volume. A larger accessible volume (AV) suggests greater rotational freedom and, possibly, a decreased likelihood for fluorophore-protein interaction. With regard to the latter, we note that ATTO647N is amphiphilic. a) (-15,+15) labeling scheme. ATTO647N has considerable rotational freedom compared to Cy3B, which appears oriented toward the protein. Restricted Cy3B rotational freedom should not necessarily impact the FRET efficiency via κ^2, but this does not preclude contact-induced fluorescence quenching. While ATTO647N appears mobile, ATTO647N contacts the β-β' subunits as the downstream dsDNA is pulled toward the active site. Such interaction may also promote photophysical and photochemical pathways other than resonance energy transfer. b) (-5,-3) labeling scheme. Cy3B has greater rotational freedom than in (a), but fluorophore-protein interaction still appears likely. ATTO647N is buried within the protein, and, according to AV modeling which uses a static structure, ATTO647N is completely restricted (hence, the absence of an acceptor mesh). Despite the high probability for fluorophore-protein interaction, gel assays did not suggest that enzyme reactivity is significantly perturbed [50]. c) (-15,-25) labeling scheme. The AV for Cy3B is the same as in (a). ATTO647N is oriented away from the enzyme and has a large AV. We note that fluorophore linker lengths are approximately 2nm, and, for this dsDNA construct, Cy3B and ATTO647N are approximately 3.5nm apart. Thus, we cannot exclude the possibility for fluorophore-fluorophore interaction and contact-induced fluorescence quenching. Subunits: β is magenta; β' is yellow; αI is green; αII is cyan; α C-terminal domains dark purple and orange; ω not shown; σ^{70} is grey. Promoter DNA: non-template strand is light green; template strand is dark green and, within the transcription bubble, purple. Structure based on [232] and [364]. DNA structure and model based on [144]. Accessible volume calculation [**?**] and molecular modeling by Diego Duchi.


a)
(-15,+15)
Donor
σ
Template strand
Non-template strand
Acceptor
RNAP
Start Site +1
b)
(-5,-3)
Acceptor
Donor
c)
(-15,-25)
Acceptor
Donor

tions may provide access to FRET signals corresponding to conformational dynamics or the lack thereof, we must treat E-BVAresults with suspicion and use appropriate biological controls and statistical analysis methods to account for possible FRET dynamics artifacts.

4.4 Experimental Results

Single-molecule FRET spectroscopy with alternating-laser excitation (ALEx) [168] of diffusing dsDNA and RP_o molecules resulted in FRET histograms and E-BVA plots for molecules having a 1:1 labeling stoichiometry (one donor and one acceptor); see Appendix D.3 and D.4 and Figure 4.4. For each dsDNA construct, we performed measurements in the absence and presence of RNAP holoenzyme.

Beginning with (-15,+15) , in the absence of holoenzyme, we expect the DNA to behave as a rigid rod, resulting in a unimodal distribution close to the shot noise limit. Upon holoenzyme addition, RP_o formation results in downstream dsDNA positioned closer to upstream dsDNA . As the acceptor fluorophore is closer to the donor fluorophore, we expect to observe increased FRET. (-15,+15) should consequently report on the movement of downstream dsDNA toward the transcription start site. Figure 4.4(a) shows this to be the case, as indicated by the bimodal distribution, with the higher E population assigned to RP_o .

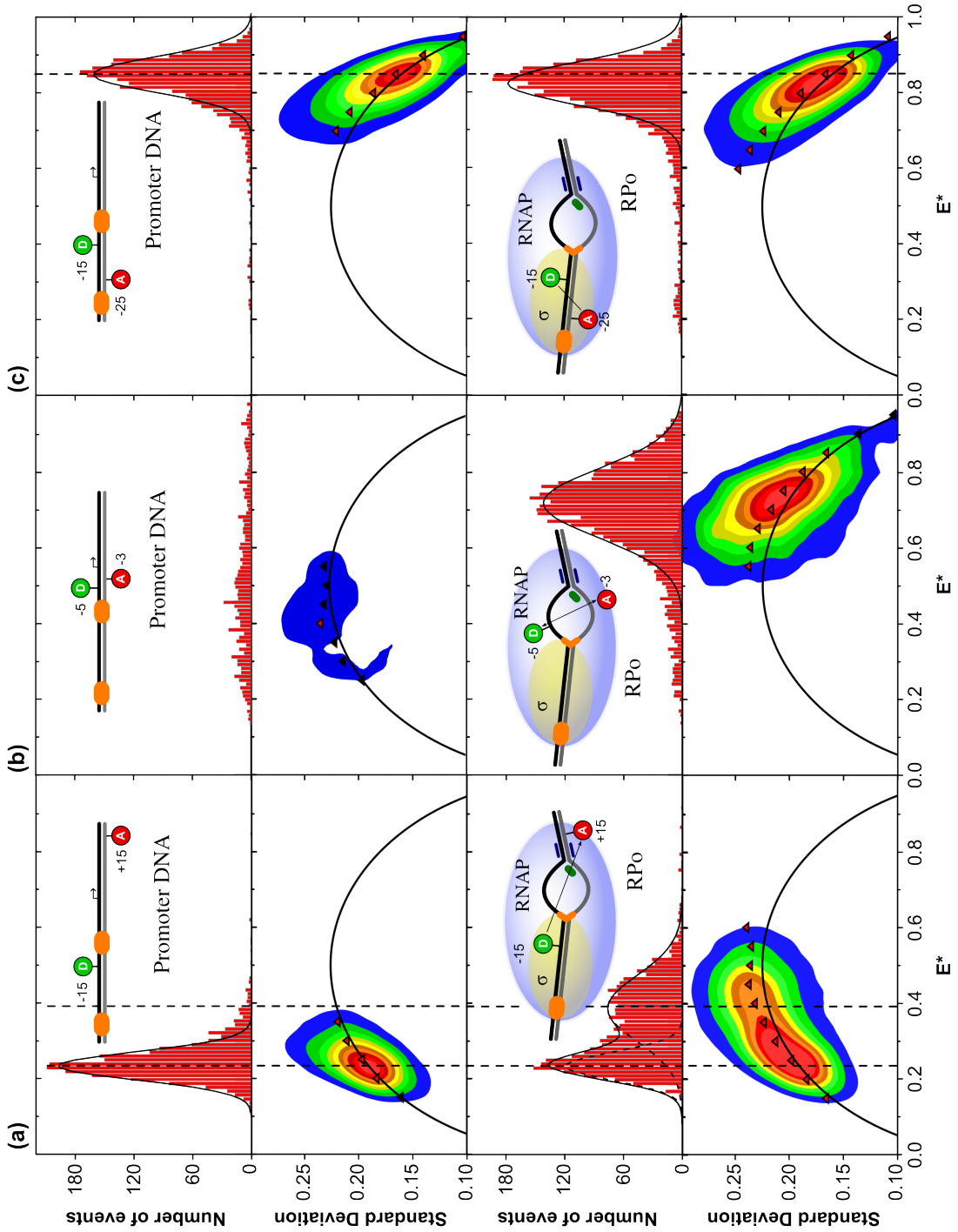
Next, for (-5,-3) , in the absence of holoenzyme, close fluorophore proximity results in contact-mediated fluorescence quenching [50, 113] and random E fluctuations without a defined distribution. Holoenzyme addition relieves quenching, as the fluorophores are pulled apart during transcription bubble formation. Hence, (-5,-3) should report on transcription bubble expansion and compaction. Figure 4.4(b) shows RP_o having a high E value centered around $E \sim 0.75$.

Previously, Kapanidis and co-workers [169] used a construct similar to (-15,-25) and concluded that the promoter spacer region did not undergo substantial conformational changes upon transcription initiation. We presumed this construct to be a suitable static control in the presence of holoenzyme and transcription ini-

tiation. The absence of dynamics would also support RP_o -mediated downstream DNA movement accounting for any possible RP_o dynamics. Figure 4.4(c) confirms that fluorophore separation in (-15,-25) does not substantially change upon RP_o formation, with the distribution maintaining the same relative shape, mean, and standard deviation.

We performed E-BVA for each measurement as specified by Torella et al [314] and outlined above, with $W = 5$ and $\alpha = 0.001$. Ideally, static controls, such as dsDNA and (-15,-25) RP_o , would not exhibit dynamics in E-BVA. But as may be observed in Figure 4.4, every sample but the sparsely populated (-5,-3) dsDNA construct show 'dynamics', as indicated by the red triangles representing the cluster E standard deviation. While some may argue that the *dynamic score*, i.e., the distance between the cluster standard deviation value and a strict upper confidence bound, is greater for RP_o samples, the case for E-BVA supporting RP_o dynamics is weak and inconclusive. Hypothesis testing for significance in the difference between dynamic scores for control and RP_o samples failed to reject the null hypothesis that the dynamic scores are effectively the same (data not shown). E-BVA

Figure 4.4 *(following page)*: Burst Variance Analysis (E-BVA) of RP_o Dynamics. We performed E-BVA on promoter dsDNA constructs that differed only in the fluorophore labeling positions. (a) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -15 and acceptor at +15, relative to main transcription start site. The upper-left histogram shows the DNA alone and corresponding E-BVA plot below. Upon addition of RNAP-σ^{70} (holoenzyme), the distribution is bimodal, where the lower E distribution is assigned to either unbound dsDNA or the RNA polymerase (RNAP) closed complex (RP_c) and the higher E distribution to RP_o . The bottom plot shows E-BVA for the entire distribution. (b) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -5 and acceptor at -3. Due to dynamic quenching [50], we see few dsDNA events. Upon addition of holoenzyme, quenching is removed due to RP_o formation, resulting in a high E population. E-BVA plots for both dsDNA and RP_o are below their corresponding histograms. (c) *lac*CONS+2 dsDNA labeled with a donor fluorophore at -15 and acceptor at -25. We expect the region between the fluorophores to remain rigid during RP_o formation [169] and thus do not expect a significant difference in histograms. E-BVA plots are below their corresponding histograms. Key: orange regions on the promoter DNA are the -10 and -35 promoter elements to which σ^{70} tightly binds; the arrow indicates the main transcription start site and transcription direction; donor fluorophore is green; acceptor fluorophore is red; the blue oval represents RNAP; the yellow oval represents σ^{70} ; the green marker inside the transcription bubble marks the main transcription start site; the blue lines downstream of the transcription bubble are the downstream dsDNA contacts for RNAP; histogram fits (black lines) are Gaussian fits; the black line in the E-BVA plots is the shot noise prediction ($W = 5$); black triangles indicate burst E fluctuations in accordance with a static prediction; red triangles indicate E fluctuations exceeding the static prediction and are evidence of dynamic interconversion. Except for the (-5,-3) dsDNA construct, *both* promoter DNA control experiments *and* RP_o show dynamics.


(a)
(b)
(c)
Promoter DNA
RNAP
RPo
-15
-25
-5
-3
+15
σ
Number of events
Standard Deviation
E*
0.0
0.2
0.4
0.6
0.8
1.0
0
60
120
180
0.10
0.15
0.20
0.25

for additional control samples also demonstrate dynamics (data not shown).

One possible explanation for dynamics is distance-independent photophysics. Achieving the shot noise limit has proven difficult to achieve [5, 38, 165, 239, 283]. Recent work suggests significant variability in fluorophore fluorescence due to dye stacking [148, 149], DNA intercalation [165], guanine induced quenching [70, 315], *cis-trans* isomerization [124, 143, 159, 345], oxygen presence [2, 369], availability of reduction-oxidation agents [331], photochemistry [245], and protein induced fluorescence enhancement [136, 207, 293]. Fluorophores, including ATTO647N, are known to undergo significant spectral shifts and to be prone to dsDNA sticking [43, 44, 69, 113, 149, 165, 187, 235, 331]. We repeated dsDNA control experiments using Trolox for triplet state quenching [51, 262], and a glucose, glucose oxidase, and catalase [119, 290] system for oxygen removal. These samples also showed dynamics. And as evident in Figure 4.4(c), holoenzyme binding to promoter DNA apparently increases the dynamic score for the (-15,-25) static control.

Taken by themselves, these results fail to conclusively support active sampling by holoenzyme in RP_o and, instead, suggest the possibility that distance-independent photophysics may be obscuring the true underlying signal. Quite possibly, ATTO647N state behavior is exacerbated by the presence of RNAP (e.g., transient fluorophore-protein interaction), giving rise to an increased dynamic score, irrespective of actual RP_o dynamics. Differentiating between distance-independent photophysical artifacts and FRET changes reporting on conformational changes remains a significant challenge in single-molecule fluorescence, and the single-molecule field has yet to establish a robust statistical framework for addressing single-molecule complexities.

4.5 Summary

E-BVA can be an effective method to detect dynamics in single-molecule data [239, 314]. Used in conjunction with other analytical methods, such as probability distribution analysis (PDA) [5, 277, 310], fluorescence lifetime analysis [102],

recurrence analysis [132], photon counting histograms [35, 172], multiparameter fluorescence detection [73, 164, 291, 341], correlation analysis [6], maximum likelihood [104], and hidden Markov and semi-Markov models [23, 185, 359, 360], E-BVA provides a statistical method to hypothesis test for dynamics. E-BVA helps to eliminate those models, such as multiple static species, which yield similar distributions to that observed but which are incompatible with fluctuations identified within individual bursts. We do note, however, that 'dynamics' is a broad term, including not only conformational fluctuations, but also photophysical effects resulting from quantum yield changes, dye sticking/stacking (restricted orientation), and other events which obfuscate the underlying biological behavior. Hence, from E-BVA alone, one cannot attribute fluctuations solely to conformational dynamics. Appropriate experimental controls and optimization are needed to mitigate the effects of artifacts and conclusively explain distribution broadening. As such, E-BVA builds upon previous work investigating shot-noise-limited distributions and provides a quick test for static versus dynamic heterogeneity as the origin of excess distribution width.

Contributions

Torella and co-workers [314] first proposed E-BVA. Yusdi Santoso, Joseph Torella, and Johannes Hohlbein provided initial E-BVA software implementations. I, alone, wrote the publicly available E-BVA software. Thorben Cordes and Nicole Robb conducted all RP_0 start-site selection measurements, and, on these measurements, I performed E-BVA. All writing and figures are my own, except where indicated.

Chapter 5

Burst Variance Analysis (Stoichiometry)

Abstract

Distinguishing distance-independent photophysics, such as fluorophore blinking, from conformational dynamics FRET fluctuations remains a significant challenge in single-molecule fluorescence. Accounting for distance-independent photophysics is critical if we want to draw conclusions concerning the biological significance of FRET dynamics. Several recent works have shown distance-independent photophysics, such as quantum yield variation [44, 165, 216], spectral shift [43, 44], and restricted fluorophore orientation [38], can lead to increased FRET fluctuation dynamics and distribution broadening beyond the shot noise limit. We use alternating laser excitation (ALEx) spectroscopy to calculate fluorophore stoichiometry. The fluorophore stoichiometry is a ratiometric variable which contains information about a fluorophore's photophysical state and allows us monitor the state's evolution over time. We derive an expression for stoichiometry similar to the analytical expression for FRET shot noise and use simulations to validate its use in detecting distance-independent photophysics. The expression provides a simple theory which predicts the expected stoichiometry distribution width defined by photon statistics. We additionally provide a statistical framework analogous to burst variance analysis (BVA) to hypothesis test for dynamic heterogeneity, i.e., individual species fluctuating between two or more distinct stoichiometry states, against the null hypothesis of static heterogeneity, i.e., multiple shot-noise-limited species. We then apply our statistical method to analyze RNA polymerase (RNAP) open complex stoichiometry distributions. Our results confirm distance-independent photophysical fluctuations, but their causal relation to FRET dynamics remains speculative.

5.1 Motivation

We previously concluded that E-BVA is unable to attribute fluctuations exclusively to conformational dynamics. FRET-induced acceptor fluorophore blinking may cause anti-correlated donor intensity changes, thus creating illusory FRET dynamics. Similarly, donor fluorophore blinking will result in FRET being a ratio of background levels, possibly resulting in apparent FRET fluctuations. The additional information from ALEx leads us to ask whether we can develop a method capable of identifying distance-independent photophysical fluctuations.

For E-BVA, ALEx is not necessary, as we only use the photon stream originating from donor excitation (D_{ex}) and ignore the acceptor excitation photon stream (A_{ex}). As discussed previously (see Section 3.2.3.2 and references [167, 195]), by using both D_{ex} and A_{ex}, we can calculate the fluorescence intensity stoichiometry (S), which gives us a handle on non-conformational dynamic fluctuations. For instance, assuming a constant D_{ex} ($\gamma \approx 1$), S fluctuations may be due to either the donor or acceptor fluorophore entering a non-emitting fluorescent state. Accordingly, if we can develop a method similar to E-BVA for stoichiometry, we will have a method to detect possible non-conformational FRET fluctuations.

5.2 Theory

During donor excitation, the probability that a detected photon was emitted by an acceptor fluorophore is given by the FRET efficiency, E, where we assume the absence of direct acceptor excitation, leakage, and background and have 100% detection efficiency in both channels. For a simple two-color experiment, a detected photon is either one of two colors: 'green' or 'red'. Let green and red be defined by the binary values 0 and 1, respectively. And let $p = E$ be the probability of a red photon, given that we have detected a photon. For a discrete random variable

X with support $k = \{0, 1\}$ (equivalent to photon colors), we have

$$P(X = 1) = 1 - P(X = 0) = 1 - q = p$$

The probability that X is exactly equal to either 0 or 1 is then given by the probability mass function (pmf)

$$f_X(k; p) = \begin{cases} p & \text{if } k = 1 \\ 1 - p & \text{if } k = 0 \end{cases}$$

which may also be expressed as

$$f_X(k; p) = p^k \, (1 - p)^{1-k} \text{ for } k \, \epsilon \, \{0, 1\} \tag{5.1}$$

An experiment whose event is a random variable with two possible outcomes is termed a *Bernoulli trial*. For a fixed number N independent Bernoulli trials $X_0, \ldots, X_{n-1}$, each with probability of success p, the number of successes, m, has a binomial distribution $Binom(N, p)$

$$P(m) = \binom{N}{m} p^m \, (1 - p)^{N-m} \tag{5.2}$$

where

$$\binom{N}{m} = \frac{N!}{m!(N - m)!}$$

is the binomial coefficient. The binomial coefficient accounts for the number of different ways of distributing m successes in a sequence of N trials. A binomial distribution has mean and variance

$$\text{E}[X] = Np \tag{5.3}$$

$$\text{Var}[X] = Np(1 - p) \tag{5.4}$$

Hence, if we want to know the probability that we observe m acceptor photons within N total photons, we know that the probability will follow a binomial distribution with standard deviation

$$\sigma(F_A) = \sqrt{\text{Var}[F_A]} = \sqrt{NE(1-E)} \tag{5.5}$$

where F_A represents a random value corresponding to detected acceptor fluorescence.

But what about the observed FRET distribution, E^*? Defining a function $g(X)$

$$g(F_A) = E^* = F_A/N \approx g(\text{E}[F_A]) + (F_A - \text{E}[F_A])g'(\text{E}[F_A])$$

where we approximate E^* with the first two terms of a Taylor series expansion, we use the *delta method* [241] to calculate the expected standard deviation of E^* as

$$\begin{aligned}
\text{Var}[E^*] &\approx \text{Var}[F_A - \text{E}[F_A]] \cdot \left(\frac{\partial g(F_A)}{\partial F_A}\bigg|_{F_A=\text{E}[F_A]} \right)^2 \\
&\approx \sigma(F_A)^2 \cdot \left(\frac{\partial g(F_A)}{\partial F_A}\bigg|_{F_A=\text{E}[F_A]} \right)^2 \\
&\approx \frac{\sigma(F_A)^2}{N^2}
\end{aligned} \tag{5.6}$$

hence,

$$\sigma(E^*) \approx \frac{\sigma(F_A)}{N} = \sqrt{\frac{E(1-E)}{N}} \tag{5.7}$$

where we recall that E is the true probability. In practice, E is not known, but we can estimate E with μ_{E^*} by fitting observed E^* values to a Gaussian distribution and using the fit mean [53].

Thus far we have only discussed a single binomial sampling, but what about the case where we have multiple samplings, each with a different number of independent trials? For multiple samplings, we define a random variable Y as the sum of binomial random variables $X_0 + \ldots + X_{\Omega-1}$, each with the same sample probability p. Y is then also a binomial random variable.

$$Y \sim Binom\left(\sum\{n_i\}, p\right)$$

(That this is the case follows intuitively by considering a story proof involving two experiments: one involving 10 coin toss trials followed by another having 5 trials. The expected mean would be $\mathrm{E}[X] = \sum_{i=1}^{2} n_i p = 15 \cdot 0.5 = 7.5; (p = 0.5)$.)

Suppose, then, that we have a mixture of different sample sizes, $\{n_i\}$. Let Ω be the number of samples and ξ be the number of occurrences of sample size N within our set $\{n_i\}$. The probability that we will observe m successes over all samples is given by

$$P(m) = \sum_{N=\min(\{n_i\})}^{\max(\{n_i\})} \left(\frac{\xi}{\Omega}\right) \frac{N!}{m!(N-m)!} p^m \, (1-p)^{N-m} \tag{5.8}$$

where $0 \leq m \leq N$ and $P_N(m) = 0$ for $N < m$. If we once again let $F_A = m$ and now let $\{n_i\}$ be the experimental distribution of burst sizes from confocal solution data, Equation 5.8 instructs us as to how to build a *shot noise* distribution, $f_{E^*_{SN}}(F_A)$, which accounts for the statistical fluctuations in E^* due to small sample sizes.

$$f_{E^*}(F_A, s_i) = \frac{s_i!}{F_A!(s_i - F_A)!} \frac{(E^*)^{F_A}(1-E^*)^{s_i - F_A}}{S} \tag{5.9}$$

where s_i is the burst size for burst i and S is the total number of bursts. For each burst size s_i, we randomly generate a binomial variable F_A with parameters $p = E^*$ and $n_i = s_i$ and calculate $f_{E^*}(F_A, s_i)$. By building a probability histogram and multiplying the histogram bin height by the total burst number S, we generate the expected shot-noise-limited distribution.[1] For the special case where all $s_i = N$, a constant, Equation 5.9 reduces to a single binomial distribution with variance σ^2_{SN}.

$\sigma(E^*) = \sigma_{SN}$ represents the expected distribution width for a single static species with FRET probability E. Observed distribution widths that exceed the shot noise limit may arise due to a number of factors: multiple static species, excitation/emission volume mismatch, restricted fluorophore orientation, distance-

[1] In practice, one uses an oversampling factor k to ensure rare burst sizes are sufficiently sampled, and then one divides the final histogram by k to recover the expected distribution [239].

independent photophysics, and conformational dynamics corresponding to FRET reported distance changes. E-BVA consequently exploits Equation 5.7 for a fixed sample size N to exclude multiple static species as the basis of distribution broadening beyond the shot noise limit.

While acceptor photon counts under donor excitation follow a binomial distribution given a known number of total photons, which makes FRET a binomial process, we may wonder whether stoichiometry is also a binomial process. Recall, the stoichiometry equation (see Section 3.2.3.2, Equation 3.19) is

$$S = \frac{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}} + F_{A_{exc}}^{D_{em}} + F_{A_{exc}}^{A_{em}}}$$

where, in the absence of background, the photon counts for each burst $F_{X_{exc}}^{Y_{em}}$ associated with excitation X_{exc} and emission Y_{em} channels are as follows:

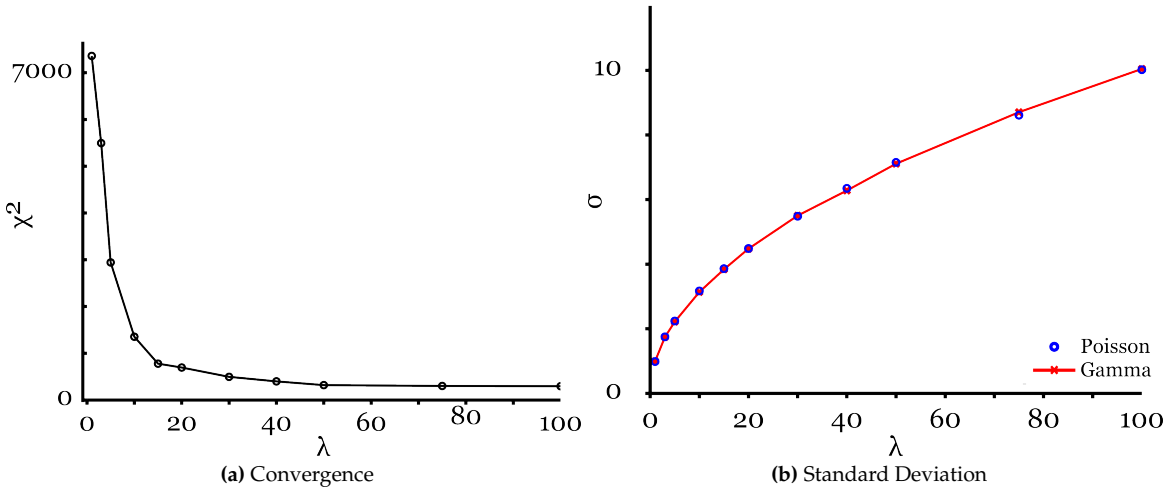
$$F_{D_{exc}}^{D_{em}} = \sigma_{D_{exc}}^{D} \Phi_D \xi_{D_{em}}^{D} (1 - E) \int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\, dt$$

$$F_{D_{exc}}^{A_{em}} = \left(\underbrace{\sigma_{D_{exc}}^{D} \Phi_D \xi_{A_{em}}^{D} (1 - E)}_{\text{leakage}} + \underbrace{\sigma_{D_{exc}}^{A} \Phi_A \xi_{A_{em}}^{A}}_{\text{direct excitation}} + \underbrace{\sigma_{D_{exc}}^{D} \Phi_A \xi_{A_{em}}^{A} E}_{\text{FRET}} \right) \int_{t_i}^{t_j} I_{D_{exc}}(\vec{r}(t))\, dt$$

$$F_{A_{exc}}^{D_{em}} \approx 0$$

$$F_{A_{exc}}^{A_{em}} = \sigma_{A_{exc}}^{A} \Phi_A \xi_{A_{em}}^{A} \int_{t_i}^{t_j} I_{A_{exc}}(\vec{r}(t))\, dt$$

where D stands for 'donor' and A for 'acceptor'; $I_{D_{exc}}$ and $I_{A_{exc}}$ are the D- and A-excitation laser intensities, respectively, as a function of the particle path $\vec{r}(t)$ through the confocal volume; $\Delta t = t_j - t_i$ is the total time we observe a fluorescent molecule in the confocal volume, i.e., the burst duration; $\sigma_{D_{exc}}^{D}$, $\sigma_{D_{exc}}^{A}$, and $\sigma_{A_{exc}}^{A}$ are the absorption cross sections of D upon D-excitation, of A upon D-excitation, and of A upon A-excitation, respectively; Φ_D and Φ_A are fluorescence quantum yields of D and A, respectively; $\xi_{D_{em}}^{D}$, $\xi_{A_{em}}^{D}$, and $\xi_{A_{em}}^{A}$ are detection efficiencies of D-emission in the D-detection channel, of D-emission in the A-detection channel, and of A-emission in the A-detection channel, respectively; and E is FRET efficiency.



(a) Convergence **(b)** Standard Deviation

Figure 5.1: Poisson and Gamma Distribution Convergence. We simulated Poisson ($\lambda = \{1, 3, 5, 10, 15, 20, 30, 40, 50, 75, 100\}$) and Gamma ($\alpha = \lambda, \beta = 1$) distributions. a) χ^2 statistic. The χ^2 statistic suggests rapid convergence by $\lambda = 10$, with remaining disparity attributable to fluctuation differences in the distribution tails. Number of simulations: $N = 10,000$. b) standard deviation comparison. The standard deviation σ is the same for both Poisson and Gamma distributions, following the same $\sqrt{\lambda}$ dependence. The blue circles represent the Poisson distribution simulation, and the red line with red x's the Gamma distribution simulation. Number of simulations: $N = 10,000$.

Previously, Dahan and co-workers [53] modeled FRET efficiency as a random variable $E = F_{DA}/(F_{DD} + F_{DA})$ given by a Beta distribution. In their work, they assumed F_{DD} and F_{DA} to be two *independent* Poisson variables with rates $\lambda_{DD}, \lambda_{DA} \geq 5$.

$$E \sim \frac{Poiss(\lambda_{DA})}{Poiss(\lambda_{DD}) + Poiss(\lambda_{DA})}$$

In the limit that $\lambda \to \infty$, a Poisson distribution approaches a normal distribution with $\mu = \lambda$ and $\sigma^2 = \lambda$. And similarly, a Gamma distribution, with parameters α and β, in the limit that $\alpha \to \infty$, approaches a normal distribution with $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$. For $\beta = 1$ and $\alpha = \lambda$, we may therefore infer that Poisson and Gamma distributions should also approach each other (see Figures 5.1a, 5.1b, and 5.2).

Dahan and co-workers approximated F_{DD} and F_{DA} along these lines for sufficient photon numbers (≥ 5) and described F_{DD} and F_{DA} by Gamma distributions ($\alpha = \lambda \geq 5, \beta = 1$):

$$P(X = k; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} k^{\alpha-1} e^{-\beta k} = \frac{k^{\lambda-1} e^{-k}}{\Gamma(\lambda)} \tag{5.10}$$

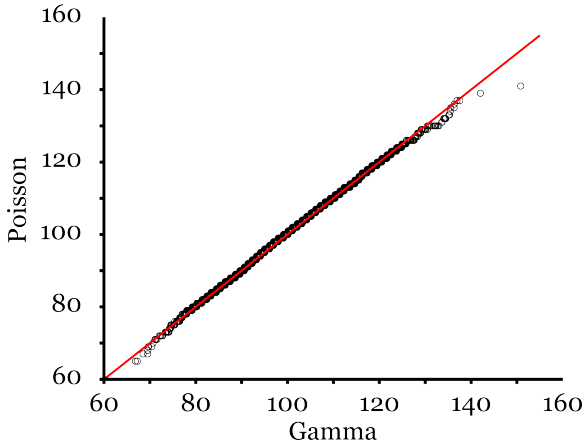


Figure 5.2: Poisson and Gamma Distribution QQ-plot. Same simulation method as in Figure 5.1a, but only showing one Poisson ($\lambda = 100$) and Gamma ($\alpha = \lambda, \beta = 1$) distribution. The black circles are the sorted pairs (x_i, y_i), where x_i is a random value drawn from a Gamma distribution and y_i a random value from a Poisson distribution. The red line is $x = y$, and indicates distribution equivalence. The main discrepancy between the Poisson and Gamma distributions lies in the tails of the distributions. The Poisson distribution is slightly fatter toward lower values, and the Gamma toward higher values. Number of simulations: $N = 1\text{E}6$.

where k is the number of donor (acceptor) photons and

$$\Gamma(\lambda) = \int_0^\infty k^{\lambda-1} e^{-k} dk$$

E is then given by

$$E \sim \frac{\Gamma(\alpha_2, \beta)}{\Gamma(\alpha_1, \beta) + \Gamma(\alpha_2, \beta)}$$

with $\alpha_1 = \lambda_{DD}$ and $\alpha_2 = \lambda_{DA}$ for the time interval of interest. For two Gamma distributions with equal scale parameters, β, the ratio distribution $X/(X+Y)$ is given by a Beta distribution [294]

$$P(\varepsilon; \alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \varepsilon^{\alpha_1 - 1}(1-\varepsilon)^{\alpha_2 - 1} \tag{5.11}$$

where

$$\mathrm{B}(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} = \int_0^1 \varepsilon^{\alpha_1 - 1}(1-\varepsilon)^{\alpha_2 - 1} d\varepsilon$$

is the Beta function, which serves as a normalization constant. The first moment

of the Beta distribution is given by

$$\langle E \rangle = \frac{\lambda_{DA}}{\lambda_{DD} + \lambda_{DA}} \tag{5.12}$$

with the square root of its variance given by

$$\sigma(E) = \sqrt{\frac{\lambda_{DD}\lambda_{DA}}{(\lambda_{DD} + \lambda_{DA})^2(\lambda_{DD} + \lambda_{DA} + 1)}} \tag{5.13}$$

In practice, we may not know the population rates λ_{DD} and λ_{DA}; however, given a sample of m measured values k_i, we can estimate the population value of the Poisson parameter λ. The maximum likelihood estimate (MLE) is

$$\hat{\lambda}_{MLE} = \frac{1}{m}\sum_{i}^{m} k_i \tag{5.14}$$

which is the sample mean. If we then let $\hat{\lambda}_{DD} = \langle F_{DD} \rangle$ and $\hat{\lambda}_{DA} = \langle F_{DA} \rangle$, with $\langle \cdot \rangle$ representing the average, our estimate for the first moment becomes

$$\begin{aligned} \langle E \rangle &= \frac{\hat{\lambda}_{DA}}{\hat{\lambda}_{DD} + \hat{\lambda}_{DA}} \\ &= \frac{\langle F_{DA} \rangle}{\langle F_{DD} \rangle + \langle F_{DA} \rangle} \end{aligned} \tag{5.15}$$

and similarly

$$\langle (1 - E) \rangle = 1 - \langle E \rangle = \frac{\langle F_{DD} \rangle}{\langle F_{DD} \rangle + \langle F_{DA} \rangle} \tag{5.16}$$

Using Equations 5.15 and 5.16, and defining $N_E = \langle F_{DD} \rangle + \langle F_{DA} \rangle$, we find that

$$\sigma(E) = \sqrt{\frac{\langle E \rangle (1 - \langle E \rangle)}{N_E + 1}} \tag{5.17}$$

which has a similar form to that of Equation 5.7.

Although similar in appearance, Equations 5.7 and 5.17 represent two different experiments. In Equation 5.7, we know the number of trials (photons) and we ask how many 'successes' we can expect given the number of trials and the probability

of success. What we lose in Equation 5.7 is the temporal information, notably the waiting times between photon arrival. On the other hand, in Equation 5.17, we are agnostic on the number of photons, but do know something about the rates, and we ask how many photons we can expect within a time interval given the rates. In fact, in Equation 5.17, N_E is the average number of total photons in the donor and acceptor channel, which is dependent on the rates and is not an arbitrary positive integer. So when Nir and co-workers [239] claimed that Equation 5.17 is wrong, they were mistaken: Dahan and co-workers answered a different question. Nevertheless, Equation 5.7 does have two distinct advantages: 1) it provides an *analytical* solution, and 2) E-BVA implementation is straightforward.

Although Equation 5.17 is not exact for E, its logic may prove useful in describing stoichiometry. To simplify notation, let the set $\{F_{X_{exc}}^{Y_{em}}\}$ be represented by the random variables $X, Y, W,$ and Z, respectively. In which case,

$$S = \frac{X + Y}{X + Y + W + Z}$$

The ratio S is the result of two consecutive independent experiments: we first draw random values, X and Y, during donor excitation; we then draw W and Z during acceptor excitation. Let D be a random variable equaling the sum of X and Y and, similarly, $A = W + Z$.

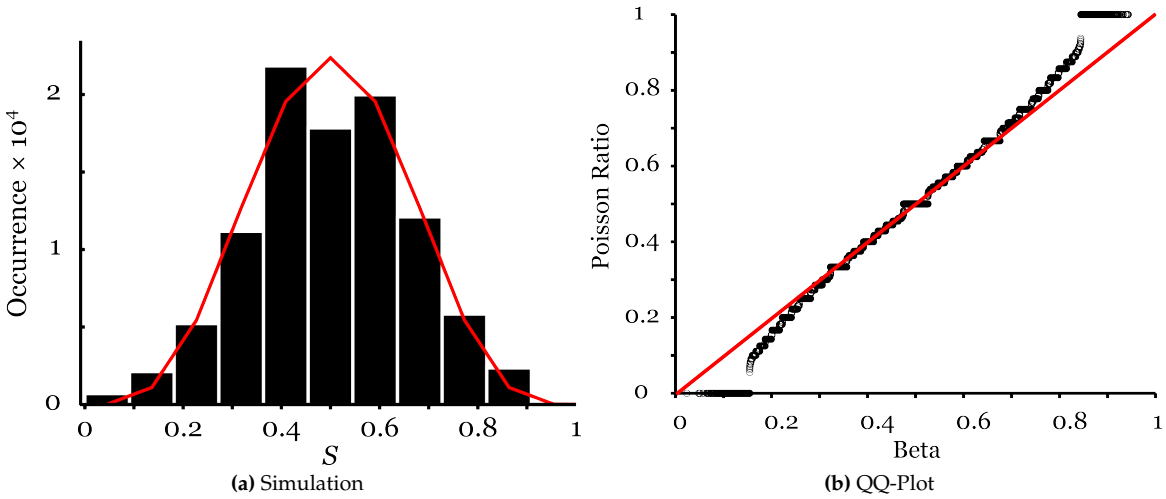
$$S = \frac{D}{D + A}$$

We immediately notice the similarity between S and E. Let $D \sim Poiss(\lambda_D)$ and $A \sim Poiss(\lambda_A)$ [101]. Provided that D and A are *independent* Poisson random variables (proof provided in the Appendix B.1.1), we can then describe S as a ratio distribution given by

$$S \sim \frac{Poiss(\lambda_D)}{Poiss(\lambda_D) + Poiss(\lambda_A)} \tag{5.18}$$

where we follow the logic used previously in approximating E to describe S.

$$S \sim \beta(\lambda_D, \lambda_A) \tag{5.19}$$



(a) Simulation **(b)** QQ-Plot

Figure 5.3: Poisson Ratio and Beta Distribution Comparison. We simulated S both as a Poisson ratio distribution using Equation 5.18 and as a Beta distribution. Simulation parameters: $\lambda_D = 5$; $\lambda_A = 5$; total simulations = 10,000. As evident in (b), the Poisson ratio distribution is more heavily populated in the distribution tails. As $\lambda \to \infty$, this discrepancy becomes negligible, but for low rates, we should expect that Equation 5.21 will underestimate the distribution spread. The red line in (a) is the simulated Beta distribution. The representation in (b) is the same as Figure 5.2, but with x_i being a random value drawn from a Beta distribution and with y_i drawn from a Poisson ratio.

and

$$\langle S \rangle = \frac{\langle D \rangle}{\langle D \rangle + \langle A \rangle} \tag{5.20}$$

$$\sigma(S) = \sqrt{\frac{\langle S \rangle (1 - \langle S \rangle)}{N_S + 1}} \tag{5.21}$$

where $N_S = \langle D \rangle + \langle A \rangle$ is the total number of photons over both donor and acceptor excitation.

To confirm $\sigma(S)$ as a suitable approximation, we simulated both a ratio distribution of Poisson random variables and a Beta distribution. We used Monte Carlo techniques to build distributions (10,000 random values from each distribution), with $N_S = 10$, $\lambda_D = 5$, and $\lambda_A = 5$. Figure 5.3 shows the results of the simulation.

As shown in the QQ-plot (Figure 5.3b), the Poisson ratio distribution is more heavily populated in the distribution tails. This is evident in the Beta distribution overlay atop the Poisson distribution histogram (Figure 5.3a), where more events occur toward extreme values. We can thus conclude that, for low rates, Equation 5.21 will underestimate σ_S. This conclusion is confirmed in Figure 5.4,

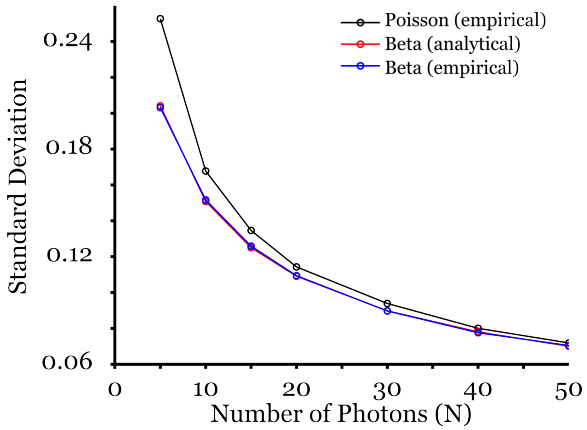


Figure 5.4: Poisson Ratio and Beta Distribution Comparison. We simulated S both as a Poisson ratio distribution using Equation 5.18 and as a Beta distribution. Simulation parameters: $\lambda_D = \lambda_A = \{2.5, 5, 7.5, 10, 15, 20, 25, 37.5, 50, 75, 100, 175, 250, 500\}$; total simulations = 10,000. For low photon numbers, the Beta distribution underestimates σ_S, but rapidly converges to the Poisson ratio distribution for $\lambda_D + \lambda_A = N_S > 50$. The black dashed line plots the calculated sample standard deviation for the Poisson ratio distribution simulation; the blue line the calculated sample standard deviation for the Beta distribution simulation; and the red line is the analytical solution according to Equation 5.21. Note: we only show $N_S \leq 50$, as beyond this photon number the difference in standard deviations is negligible.

where we simulated across a range of rates (λ_D, λ_A varied between $[2.5, 500]$). By $\lambda_D = 25$ ($N_S = 50$), the difference between the sample standard deviation for the Poisson ratio and Equation 5.21 is negligible.

But now consider the scenario in which, within a temporal period T, we find $D + A = N$ photons. What is the distribution of D given this information? Let X and Y be independent Poisson random variables with rates λ_X and λ_Y, respectively. Then $X + Y = N$ is Poisson with rate $\lambda = \lambda_X + \lambda_Y$. The question then becomes: what is the probability that $X = k$, given N? We recall the definition of conditional probability

$$P(X = k|X + Y = N) = \frac{P(X = k, X + Y = N)}{P(X + Y = N)}$$

for $0 \leq k \leq N$. Because X and Y are independent, we can factor the joint probability

$$\begin{aligned} &= \frac{P(X=k)P(Y=N-k)}{P(X+Y=N)} \\ &= \frac{p(k;\lambda_X)p(N-k;\lambda_Y)}{p(N;\lambda_X+\lambda_Y)} \end{aligned}$$

where $p(n;\lambda)$ is given by a Poisson distribution parameterized by rate λ

$$p(n;\lambda) = e^{-\lambda}\frac{\lambda^n}{n!}$$

Hence,

$$\begin{aligned} P(X=k|X+Y=N) &= \frac{e^{-\lambda_X}\frac{\lambda_X^k}{k!} \cdot e^{-\lambda_Y}\frac{\lambda_Y^{N-k}}{(N-k)!}}{e^{-(\lambda_X+\lambda_Y)}\frac{(\lambda_X+\lambda_Y)^N}{N!}} \\ &= \frac{e^{-(\lambda_X+\lambda_Y)}}{e^{-(\lambda_X+\lambda_Y)}} \cdot \frac{N!}{k!N-k!} \cdot \frac{\lambda_X^k\lambda_Y^{N-k}}{(\lambda_X+\lambda_Y)^N} \\ &= \binom{N}{k}\frac{\lambda_X^k}{(\lambda_X+\lambda_Y)^k} \cdot \frac{\lambda_Y^{N-k}}{(\lambda_X+\lambda_Y)^{N-k}} \\ &= \binom{N}{k}\left(\frac{\lambda_X}{\lambda_X+\lambda_Y}\right)^k\left(\frac{\lambda_Y}{\lambda_X+\lambda_Y}\right)^{N-k} \end{aligned}$$

Let $p = \lambda_X/(\lambda_X+\lambda_Y)$, and hence, $1-p = \lambda_Y/(\lambda_X+\lambda_Y)$. In which case,

$$P(X=k|X+Y=N) = \binom{N}{k}p^k(1-p)^{N-k}$$

which means that, conditional on N, X is a binomial random variable, with conditional expectation Np and variance $Np(1-p)$.

This is a rather remarkable result. Even though D and A may be separated in time, e.g., by half an alternation cycle in ALEx, D is a binomial random variable with probability S, so long as we know the total number of photons N for a given time interval. Hence, for a fixed N,

$$\sigma(D) = \sqrt{\text{Var}[D]} = \sqrt{NS(1-S)} \tag{5.22}$$

and, similar to Equation 5.7,

$$\sigma(S^*) \approx \frac{\sigma(D)}{N} = \sqrt{\frac{S(1-S)}{N}} \tag{5.23}$$

where S is the true probability.

If we consider an ALEx experiment, two successive measurements separated in time, the thought that we can 'choose' D photons seems counter-intuitive. After all, for $D > 0$, we know that the first photon n_0 *has to be* a D photon; we cannot 'choose' anything because only one permutation is possible: D photons followed by A photons. But if we consider the scenario in which we take two Poisson processes, D and A, both starting from $t = 0$, and merge them together to create a single 'mixed' process beginning from $t = 0$, we suddenly can 'choose' whether a photon belongs to D or A. This is essentially the reverse of splitting a Poisson process, as discussed in Appendix B.1.1. Thus, in knowing N, rather than asking how many D and A photons we expect in a time interval T, we can ask, out of N total photons, how many photons are D photons and how many A photons, without considering arrival order.

5.3 Algorithm

E-BVA ignores photon arrival times within a burst, except in their temporal ordering.[2] All D_{ex} photons from each donor-acceptor excitation alternation period, T_A, are concatenated and placed into five-photon bins, whereupon ε^* is then calculated. For S-BVA, we first address the scenario in which we consider photon arrival times.

If we simply used a window W and binned photons from D_{ex} and A_{ex}, without considering T_A, we cannot guarantee that the photons within W are equally sampled from both D_{ex} and A_{ex}; see Figure 5.5. For example, suppose D_{ex} has 7 photons and A_{ex} 5 photons. If $W = 10$, then S is biased toward 1. Similarly,

[2]The actual time that a photon arrived is not as important as that one photon arrived before another.

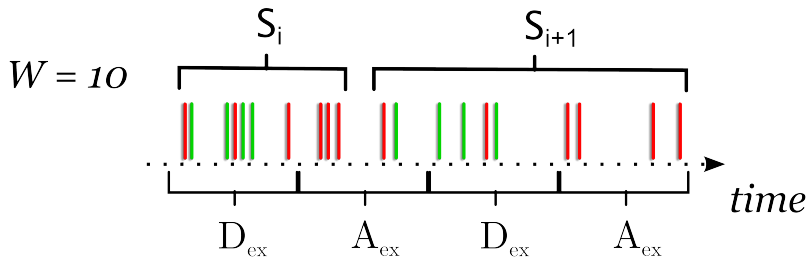


Figure 5.5: Problem with Applying Burst Variance Analysis Algorithm for FRET Efficiency E to Stoichiometry S. If we do not consider photon arrival time relative to laser excitation, we cannot guarantee equal photon sampling from both donor (D_{ex}) and acceptor (A_{ex}) excitation. Here, the first alternation period defined by $D_{ex} + A_{ex}$ contains 7 D_{ex} photons, indicated by the green lines, and 5 A_{ex} photons, indicated by the red lines, yet the sliding window ($W = 10$ photons) contains 7 D_{ex} photons and only 3 A_{ex} photons. The calculated S_i is biased toward 1. Shifting W to the next 10 photons includes the remaining 2 A_{ex} photons and the next 8 photons equally sampled from D_{ex} and A_{ex} . The calculated S_{i+1} is biased toward 0. If we calculated a sample standard deviation based on S_i and S_{i+1}, we would infer greater S fluctuations than the data actually supports.

suppose we were to use the previous example and move W to include the next 10 photons, where the following D_{ex} has 4 photons and the following A_{ex} has 4 photons. In this case, the number of photons from D_{ex} is 4 and A_{ex} is 6 (2 from previous + 4 from next), and S is biased toward 0. Using both of these windows to calculate the sample standard deviation s_S, we would infer greater S fluctuations than the data actually supports.

Hence, a simple photon window W does not work for S-BVA, and we need to use a temporal window to ensure equal temporal contribution from both D_{ex} and A_{ex}. While we could define the set of $\{W_j\}$ to be the sequence of alternation periods T_A within a burst (donor excitation + acceptor excitation; $n =$ number of donor-acceptor alternation cycles; $\sum W_j = nT_A <$ burst length), this would not guarantee a constant photon number, thus preventing straightforward shot noise prediction. And while we could only accept those T_A with exactly N photons, this would limit the amount of data available for us to use because many bursts only last upward of 20-30 alternation periods (2-3 ms).

To overcome these problems, we redefine W as a temporal window which is an integer multiple of the alternation period cT_A, where $c \ll n$. Using an integer multiple thus ensures equal (in time) contributions from D_{ex} and A_{ex}. Starting

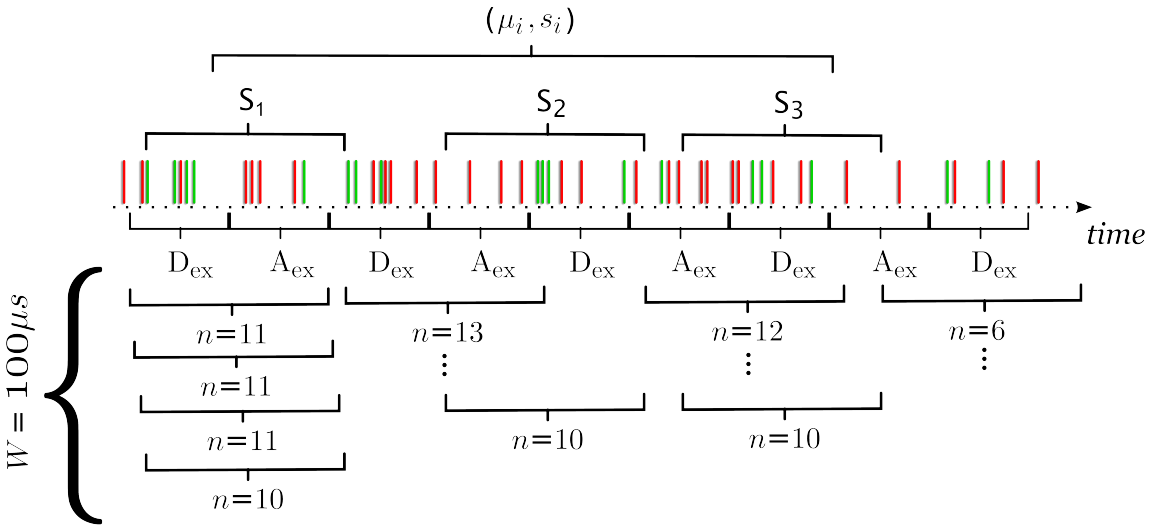


Figure 5.6: Stoichiometry Burst Variance Analysis (S-BVA) schematic. Stoichiometry sample mean, μ_i, and standard deviation, s_i, calculation. We take M samples of n consecutive photons (in the diagram, $M = 3, n = 10$) found within a sliding temporal window W from each burst. We then calculate Stoichiometry, S_j, for each non-overlapping n-photon window. μ_i and s_i are the sample mean and standard deviation, respectively, of S_j over all M. Hence, for each burst, i, we have a single mean and standard deviation pair (μ_i, s_i). Green line: donor photon; red line: acceptor photon; D_{ex} : donor excitation; A_{ex} : acceptor excitation; $D_{ex} + A_{ex}$ defines one alternation period, T_A $(= 50\mu s + 50\mu s)$. By restricting $W = cT_A$, where c is an integer greater than 0, we ensure equal (in time) photon contribution from D_{ex} and A_{ex} .

from the first photon's arrival time within a burst, we shift W in 100ns increments.[3] If W contains N photons, we calculate the stoichiometry S and move W to a time immediately after the photon arrival time of the last photon in W. We repeat our search for a window containing N photons until we reach the end of the current burst (see Figure 5.6). Similarly, as for E-BVA, this algorithm is then repeated for each burst.

Using a time window approach as defined above ensures a constant N, while maximizing the number of windows containing N photons within a burst. The choice of W and N depends on burst photon statistics. For high photon counts per T_A, c should be closer to unity, while low photon counts per T_A necessitate larger values for c. For example, if the mean photon counts per T_A is 10 and $N = 10$, then c should equal 1. For $c > 1$, few windows will have exactly 10 photons. If the mean photon count is 3, then c should equal 3. By adjusting parameters W and N,

[3]The increment choice is three orders of magnitude shorter than T_A for our experimental setup. This parameter may be adjusted if a more finely grained photon search is desired, but at the cost of increased computation time.

one can ensure that one obtains the maximum statistics per burst.

Using the standard E-BVA algorithm, we cannot ensure that W samples photons consistently across the same timescale. For example, suppose we perform a Seidel burst search [73, 239] on $A_{ex}A_{em}$, thus detecting all acceptor-labeled diffusing molecules in which the acceptor is active. If the donor fluorophore enters a dark state during a burst, some alternation periods will have few $D_{ex}D_{em}$ and $D_{ex}A_{em}$ photons. Suppose further that we then perform E-BVA with $W = 5$. In the first alternation period, we have 7 D_{ex} photons detected from a molecule in a particular conformational state within the first 20μs, in which case 5 of these photons will belong to a window, with 2 left over. During the second alternation period, the donor enters a dark state, and we detect zero D_{ex} photons. During the third alternation period, the donor is once again active, and we detect 3 photons in the final 10μs. Whereas the first window sampled 5 photons within 20μs, the second window samples 5 photons within 220μs, if we account for laser alternation. In general, the E-BVA algorithm is more likely to detect dynamics, as windows are more likely to contain transitions. Conversely, a fixed temporal window is less likely to contain transitions for inverse rate constants greater than the temporal sampling.

Alternatively, if we are comfortable sampling across varying timescales, we can use the insight set forth at the end of the previous section to construct a merged Poisson process, allowing a S-BVA algorithm nearly identical to the standard E-BVA algorithm. We proceed as follows:

1. Isolate D_{ex} and A_{ex} excitation photon streams.

2. Condense each photon arrival stream such that $t = 0$ marks the start of the first respective excitation pulse and such that each subsequent excitation pulse is shifted to begin at the end of the previous pulse.

3. Merge the two condensed photon arrival streams and sort the merged stream by condensed arrival time.

4. Run a non-overlapping window W (e.g., $W = 10$ photons) along the sorted merged stream, calculating the stoichiometry S^*_{ij}

$$S^*_{ij} = \frac{n_{DD} + n_{DA}}{n_{DD} + n_{DA} + n_{AA} + n_{AD}} \tag{5.24}$$

where $n_x =$ number of photons; DD: donor excitation donor emission; DA: donor excitation acceptor emission; AA: acceptor excitation acceptor emission; AD: acceptor excitation acceptor emission.

5. For each burst i, calculate the stoichiometry mean and sample standard deviation (μ_i, s_i) across all windows

$$\mu_i = \frac{1}{M_i} \sum_{j=0}^{M_i - 1} S^*_{ij} \tag{5.25}$$

$$s_i = \sqrt{\frac{1}{M_i - 1} \sum_{j=0}^{M_i - 1} (S^*_{ij} - \mu_i)^2} \tag{5.26}$$

where M_i is the number of windows W for burst i.

6. For all bursts with similar mean stoichiometry $\{\mu(S^*) : a \leq \mu(S^*) < b\}$, calculate the partition (cluster) standard deviation, $\sigma(S)$, across all windows for all considered bursts

$$\mu_k(S^*) = \frac{1}{M_k} \sum_{\substack{i \text{ where} \\ a \leq S^* < b}} \sum_{j=0}^{M_i - 1} S^*_{ij} \tag{5.27}$$

$$\sigma_k(S^*) = \sqrt{\frac{1}{M_k - 1} \sum_{\substack{i \text{ where} \\ a \leq S^* < b}} \sum_{j=0}^{M_i - 1} (S^*_{ij} - \mu_k(S^*))^2} \tag{5.28}$$

where $M_k = \sum M_i$ is the total number of windows belonging to bursts satisfying the conditions of the kth partition.

7. Use Monte Carlo techniques to simulate the expected sampling distribution

$$\mu_k(S) = \frac{1}{M_k} \sum_{\substack{i \text{ where} \\ a \leq S < b}} \sum_{j=0}^{M_i - 1} \frac{(F_{DD} + F_{DA})_{ij}}{n} \tag{5.29}$$

$$\sigma_k(S) = \sqrt{\frac{1}{M_k - 1} \sum_{\substack{i \text{ where} \\ a \leq S < b}} \sum_{j=0}^{M_i - 1} \left(\frac{(F_{DD} + F_{DA})_{ij}}{n} - \mu_k(S) \right)^2} \tag{5.30}$$

where $(F_{DD} + F_{DA})_{ij}$ is a random variable drawn from a binomial distribution with n trials, i.e., the number of D_{ex} photons in a window of n photons ($n = W$), and success probability S. Define the resulting distribution as $P_{MC}(\sigma(S))$.

8. Use a Bonferroni correction for multiple hypothesis testing [1] and compare the upper tail confidence interval s_S^{CI} to the observed $\sigma(S)$

$$\left(1 - \frac{\alpha}{K}\right) = \int_0^{s_S^{CI}} P_{MC}(\sigma(S))\, d\sigma \tag{5.31}$$

where K is the number of partitions (clusters) and α is the confidence level ($\alpha = 0.001$).

9. Determine if the observed standard deviation exceeds the strict confidence bounds for static behavior and is thus indicative of dynamics.

Figure 5.7 provides a visual reference for expected S-BVA results.

5.4 Simulation Results

To validate S-BVA , we simulated diffusion-based confocal measurements; however, rather than simulate *in silico*, we chose a photon recoloring method, similar to Gopich and Szabo's proposed method for generating model histograms [104]. Accurately reproducing experimental conditions *in silico*, including molecule translation, dye photophysics, photon detection, and optics, is challenging, particularly

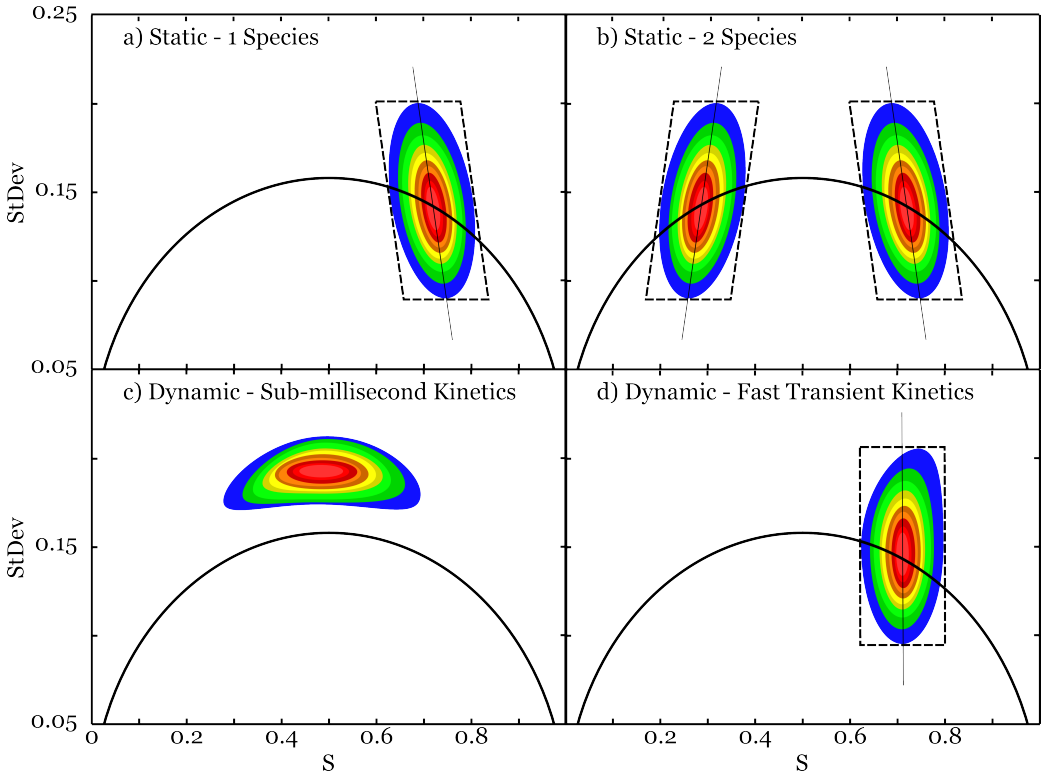


Figure 5.7: Stoichiometry Burst Variance Analysis (S-BVA): a Visual Guide. a) single static species. The default behavior for a single static species is an elliptical distribution angled toward $S = 0.5$ and centered on the shot noise prediction (solid black arc). b) multiple static species. Species which do not inter-convert on the diffusion timescale appear as individual static species, centered on the shot noise prediction. c) dynamic species. Species which inter-convert on or near the diffusion timescale result in a distribution centered above the shot noise prediction. The distribution center along the S-axis is dependent on the relative occupancy of the inter-converting statics and thus a weighted average of the individual species S values. d) fast kinetics involving a transient state. The slope and shape of the distribution can suggest transient state dynamics. For example, sub-millisecond acceptor blinking corresponds to a transient state at $S = 1$. For $S > 0.5$, transient occupancy of this state results in a vertically oriented distribution with a shoulder toward higher S values. Distribution features such as changes in slope and shape do not provide rigorous confirmation of fast dynamics, but do suggest their presence and motivate experiments which can either access the kinetic timescales or remove their presence.

when many parameters may be unknown. The advantage of recoloring experimental photon distributions and arrival times is that we need only concern ourselves with photon colors, independent of other experimental processes.

A photon sequence emitted by a diffusing molecule through a laser focal volume is characterized by arrival times and photon colors (DD, DA, AD, AA). For fixed molecules, the waiting times between successive photon arrivals are assumed to be independent via the memoryless property of exponential waiting times (see Appendix A.1.1) and to obey Poisson statistics. The Poisson distribution is completely characterized by photon count rates $\lambda_{DD},\ \lambda_{DA},\ \lambda_{AD}$, and λ_{AA}, where λ_{XY}

is the number of photons n_{XY} detected under X excitation in the Y channel per unit time. For diffusing molecules, however, photon arrival times are correlated based on a molecule's location within the focal volume. The waiting time distribution for diffusing molecules, regardless of photon color, is thus determined by a location-dependent total photon count rate and does not obey Poisson statistics. Assuming the interphoton times do not depend on conformation (i.e., $\gamma = 1$), arrival times provide information about diffusion through the focal volume and not about conformational dynamics.

But now consider photon colors. The ratio of λ_{DD} and λ_{DA} count rates is independent of photon arrival time, but dependent on molecule conformational state, made manifest by FRET efficiency E. Similarly, the ratio of $(\lambda_{DD} + \lambda_{DA})$ to the total photon count rate is also independent of photon arrival time, but dependent on relative fluorophore emission fluorescence under alternating laser excitation, which is made apparent by stoichiometry S. As long as E and S do not depend on molecule location, photon colors depend solely on conformation and relative brightness. Hence, changes in photon color patterns are location independent and must relate to changes in conformation and/or relative brightness.

Our concern is not with accurately modeling interphoton arrival times, but with simulating the pattern of photon colors. The latter is computationally more tractable and straightforward. We proceed as follows:

1. Obtain an experimental distribution of photon arrival times.

2. Perform a burst search [73, 239] to identify molecule transits through a confocal volume.

3. Obtain the burst size distribution and calculate background photon count rates $\lambda_{B_{XY}}$.

4. For each burst i, perform two continuous-time Markov chain simulations (one for conformational dynamics and one for distance-independent photo-

physical dynamics)

$$\boldsymbol{\pi}_E = [p_{E_0}, p_{E_1}, \ldots, p_{E_{N-1}}] \qquad \boldsymbol{\pi}_S = [p_{S_0}, p_{S_1}, \ldots, p_{S_{M-1}}] \tag{5.32}$$

$$\mathbf{Q}_E = \begin{bmatrix} -\sum_j k_{0j} & k_{01} & \cdots & k_{0j} & \cdots & k_{0(N-1)} \\ k_{10} & -\sum_j k_{1j} & \cdots & k_{1j} & \cdots & k_{1(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{i0} & k_{i1} & \cdots & -\sum_j k_{ij} & \cdots & k_{i(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{(N-1)0} & k_{(N-1)1} & \cdots & k_{(N-1)j} & \cdots & -\sum_j k_{(N-1)j} \end{bmatrix} \tag{5.33}$$

$$\mathbf{Q}_S = \begin{bmatrix} -\sum_j k_{0j} & k_{01} & \cdots & k_{0j} & \cdots & k_{0(M-1)} \\ k_{10} & -\sum_j k_{1j} & \cdots & k_{1j} & \cdots & k_{1(M-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{i0} & k_{i1} & \cdots & -\sum_j k_{ij} & \cdots & k_{i(M-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{(M-1)0} & k_{(M-1)1} & \cdots & k_{(M-1)j} & \cdots & -\sum_j k_{(M-1)j} \end{bmatrix} \tag{5.34}$$

where $\boldsymbol{\pi}$ is an initial probability matrix specifying the probability of starting in state i with $\sum_n p_{E_n} = \sum_m p_{S_m} = 1$ (e.g., the equilibrium probability distribution); N is the total number of conformational states; M is the total number of distance-independent photophysical states; $\mathbf{Q}$ is a rate matrix with rates k_{ij} and diagonal equal to the negative sum of transition rates from a particular state i to all states j. Retain the state number sequence (0,1,0,2,1,...) and absolute transition times. Assume photons are emitted from a defined conformational and distance-independent photophysical state (i.e., confor-

mational and distance-independent photophysical transitions are instantaneous), and label each photon by states n and m.

5. The actual number of emitted photons is unknown, being convolved with background photons. For each burst, draw a random number B from a Poisson distribution with rate parameter $\lambda_B = \sum \lambda_{B_{XY}}$. Determine the probability of a background photon p_B.

6. For each photon t, draw a random number u from a uniform distribution on the interval $[0, 1]$. If $u < p_B$, the photon is background, and colored according to the relative probabilities based on rates $\lambda_{B_{XY}}$. Proceed to next photon $t+1$. Otherwise, the photon is an emitted photon.

7. For each emitted photon, draw a random number u from a uniform distribution on the interval $[0, 1]$. If $u \geq S(t)$, where $S(t)$ is the stoichiometry probability determined by the current distance-independent photophysical state, color the photon 'black' and proceed to the next photon $t + 1$. Otherwise, the photon is colored 'green'.

8. For each 'green' photon, draw a random number u from a uniform distribution on the interval $[0, 1]$. If $u < E(t)$, where $E(t)$ is the FRET efficiency probability determined by the current conformational (FRET) state, color the photon 'red'. Proceed to the next photon $t + 1$.

The above algorithm assumes that coloring photons green and red are independent processes. Fluorophore blinking violates that independence assumption. For example, acceptor fluorophore blinking corresponds to an anti-correlated change as $E \to 0$ and $S \to 1$. Donor fluorophore blinking means E is undefined and $S \to 0$. To incorporate E-S dependence, we can define a separate state vector $\mathbf{v}_m = [\mu_0, \mu_1, \ldots, \mu_{N-1}]$ for each state m in the paired process. For example, suppose $\{S : 0.7, 1.0\}$, where $S_1 = 1.0$ is an acceptor dark state. We define a matrix $\mathbf{V}$

containing the corresponding E states

$$\mathbf{V} = [\mathbf{v}_0 \;\; \mathbf{v}_1] = \begin{bmatrix} 0.3 & 0.0 \\ 0.5 & 0.0 \end{bmatrix}$$

where we define each $\mathbf{v}_m$ as a column vector. Hence, for an acceptor dark state, regardless of the conformation, the FRET is 0. To incorporate E-S dependence in our simulation, we modify Step 8 such that E is both a function of the current conformational state and of the photophysical state $E(t, m)$. Similarly, we can modify the recoloring method to address distance-independent photophysical dependence on conformation, as, e.g., for $\gamma \neq 1$, making S a function of conformational and photophysical states $S(t, n)$.

The one obvious disadvantage of the recoloring simulation is our inability to model diffusion; thus, we are unable to accurately manipulate the burst size distribution and interphoton arrival times. If we wanted to generate longer photon sequences, we would need experimental samples with smaller diffusion coefficients. Obtaining such samples is obviously harder than simply changing simulation parameters. What we lose in absolute control, however, we make up for in simulation speed; we can generate 30 minute diffusion simulations under experimental conditions in under 2 minutes, as opposed to 3 hours.

For the recoloring simulations which follow, we used the same diffusion-based confocal measurements investigating RP_o dynamics as those presented earlier for E-BVA (see Figure 4.4). Three 10 minute recordings for (-5,-3) dsDNA-RP_o dynamics on provided the experimental burst size distribution and burst photon arrival times. We identified 3×10^4 bursts using a Seidel burst search [73, 239] in $A_{ex}A_{em}$ where we define a temporal window $T = 500\mu$s in which we search for $M = 7$ photons marking the start (or end) of a burst and set a minimum total photon threshold of $L = 10$ photons. We then filtered for bursts having a minimum of 30 $A_{ex}A_{em}$ photons, removed acceptor-only species by setting a minimum stoichiometry $S = 0.5$ (main population centered at $S \approx 0.7$), and isolated the main E

population by setting a minimum $E = 0.55$. After filtering, we retained 2778 total bursts.

We calculate background rates for each detection channel under donor and acceptor laser excitation by determining the number of photon counts per unit time for all photons not belonging to a burst. For the RP_o dynamics experimental sample, $\lambda_{DD} \approx \lambda_{DA} = 10$kHz, $\lambda_{AD} = 500$Hz, and $\lambda_{AA} = 5$kHz.

Using the burst size distribution for the 2778 bursts and the calculated background rates, we define global Markov chain parameters $\boldsymbol{\pi}$ and $\mathbf{A}$ and perform a recoloring simulation for each burst. We then perform S-BVA on each recoloring, using a 10 photon window, $W = 10$.

5.4.1 S-BVA Distinguishes Static and Dynamic Heterogeneity

To test whether S-BVA can distinguish between static and dynamic heterogeneity, we first tested whether S-BVA accurately predicts a single static species ($S = 0.7$) as shot-noise-limited. Indeed, as Figure 5.8(a) shows, the distribution is centered about the shot noise prediction (Equation 5.23), and no dynamics are detected (black triangles).

We next simulated two equimolar static species $\{S : 0.5, 0.7\}$. S-BVA correctly identifies both species as static, as all $\sigma_k(S)$ fall below the predicted 99.9% upper confidence bound (Figure 5.8(b)).

We then simulated a single dynamic species inter-converting between the same two states used in the static case before $\{S : 0.5, 0.7\}$ with rates $k_{01} = k_{10} = 1000\text{s}^{-1}$. The S-BVA distribution shows a clear shift toward higher standard deviation (red triangles), indicative of S fluctuation dynamics. We note that the molecule diffusion time is ≈ 1ms, on the same timescale as S fluctuations. Accordingly, due to the stochastic nature of the fluctuations, some molecules will not fluctuate while diffusing through the confocal volume. 'Non-fluctuating' molecules appear as static species occupying one of the S states. Other molecules will fluctuate multiple times. Such fluctuations lead to an intermediate S value and large standard de-

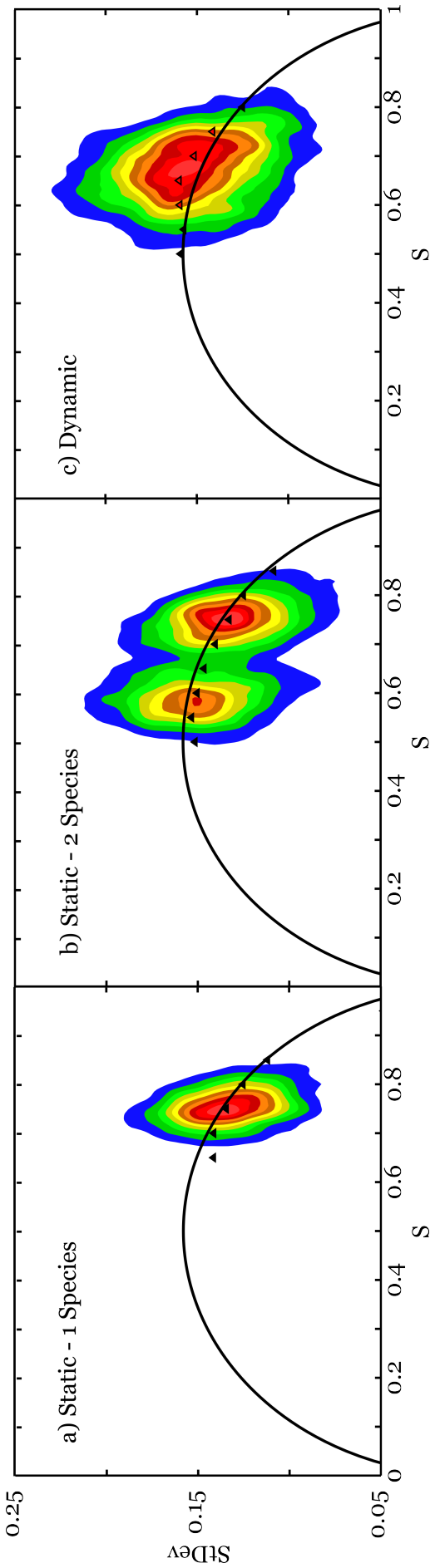


Figure 5.8: Stoichiometry Burst Variance Analysis (S-BVA) can discriminate between static and dynamic heterogeneity. Photon recoloring simulation of 2778 bursts and an S-BVA window $W = 10$. (a) single static species with $S = 0.7$. S-BVA accurately identifies a shot-noise-limited species. (b) two static species with $\{S : 0.5, 0.7\}$. S-BVA correctly identifies both species as shot-noise-limited. (c) single inter-converting species with $\{S : 0.5, 0.7\}$ and $k_{01} = k_{10} = 1000\text{s}^{-1}$. S-BVA detects millisecond dynamics, with fluctuations exceeding the shot noise prediction according to Equation 5.23; black triangles are the partition standard deviations falling below the upper confidence bound ($\alpha = 0.001$) for bursts binned along the x-axis (20 bins on the interval [-0.025, 1.025], with bin width $\Delta S = 0.05$); red triangles are the partition standard deviations exceeding the upper confidence bound and indicate dynamics.

viation. The presence of both types of bursts explains the intermediate S values which exceed the shot noise upper confidence bound, while S values close to the actual state values are in accordance with static prediction.

In conclusion, S-BVA distinguishes between static and dynamic heterogeneity.

5.4.2 S-BVA Detects Timescale-Dependent Dynamics

Similar to E-BVA [314], we should expect S-BVA to be sensitive to dynamic timescales. Kinetics much faster than the millisecond timescale will result in a single shot-noise-limited species with an observed value which is an average of the inter-converting states. Kinetics much slower will result in apparently static species.

To confirm S-BVA fluctuation detection dependence on kinetic timescale, we simulated the same two-state Markov chain $\{S : 0.5, 0.7\}$ with differing kinetics. We simulated 'slow' (compared to the diffusion timescale) kinetics $k_{01} = k_{10} = 1\text{s}^{-1}$, millisecond kinetics $k_{01} = k_{10} = 1000\text{s}^{-1}$, and 'fast' kinetics $k_{01} = k_{10} = 2.5 \times 10^4\text{s}^{-1}$. Figures 5.9(a-c) show the results.

As predicted, slowly inter-converting states appear as static species, while rapidly inter-converting species appear as a single species centered about the mean value of the two states. The fast kinetics are barely resolved, thus setting an upper bound to S-BVA's ability to detect fast kinetic systems. S-BVA clearly detects millisecond dynamics (Figure 5.9(b)). As in Torella et al [314], we expect S-BVA to be sensitive to the diffusion timescale. Slower diffusion times mean longer bursts and increased chances of a burst containing one or more transitions, while faster diffusion times are biased toward snapshots of inter-converting states. As diffusion times are dependent on focal volume dimensions, S-BVA will be sensitive to the experimental optics.

We simulated two additional scenarios: in the first, one state is sparsely populated (one state shows fast kinetics); in the second, one state is long-lived (slow kinetics). Figures 5.9(d-e) show the results.

While neither distribution is shifted above the shot noise upper confidence

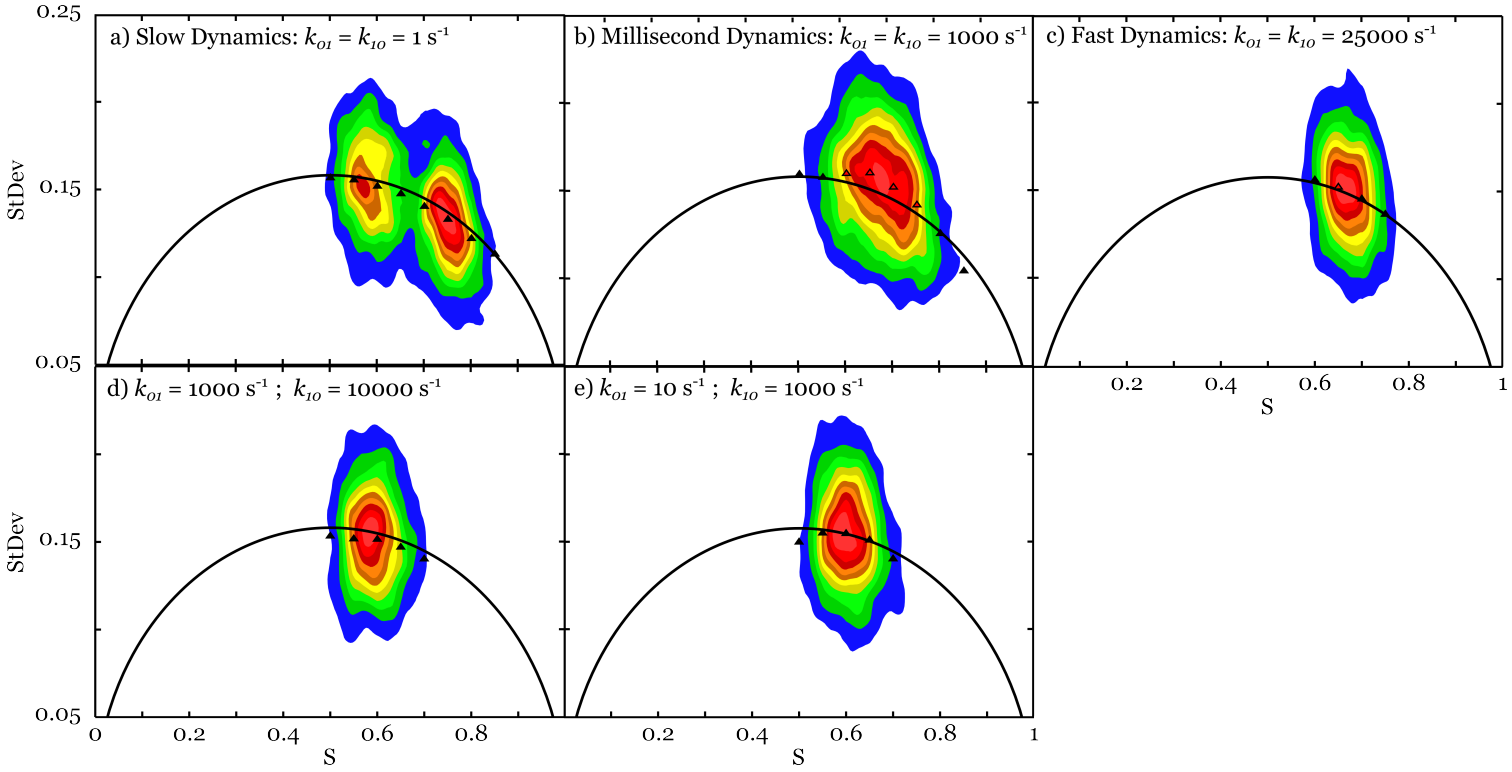


Figure 5.9: Stoichiometry Burst Variance Analysis (S-BVA) is sensitive to millisecond dynamics. Photon recoloring simulation of 2778 bursts, $\{S : 0.5, 0.7\}$, and an S-BVA window $W = 10$. (a) 'slow' dynamics on the observation timescale with $k_{01} = k_{10} = 1\text{s}^{-1}$. Slowly inter-converting states appear as static species in S-BVA, as few bursts contain state transitions. (b) dynamics on the observation timescale with $k_{01} = k_{10} = 1000\text{s}^{-1}$. With an average of 1 state transition per ms, S-BVA has sufficient statistics to identify inter-converting states. (c) 'fast' dynamics on the observation timescale with $k_{01} = k_{10} = 25000\text{s}^{-1}$. Rapidly inter-converting states (>25 transitions per ms) means each W will likely contain at least 1 transition, resulting in an average state value. Here, we are at the kinetic resolution limit for $\Delta S = 0.2$, as S-BVA shows one partition exceeding the upper confidence bound, indicating dynamics. (d) short-lived states with $k_{01} = 1000\text{s}^{-1}$ and $k_{10} = 10000\text{s}^{-1}$. S-BVA has difficulty identifying short lived states for closely spaced states. While no partitions exceed the upper confidence bound, dynamics are still evident in the distribution shape, with asymmetry up-and-to-the-right, suggesting an arc to a higher S state. (e) long-lived states with $k_{01} = 10\text{s}^{-1}$ and $k_{10} = 1000\text{s}^{-1}$. Long-lived states reduce the number of observed transitions, with the majority of bursts in accordance with the shot noise prediction. As in (d), however, distribution asymmetry suggests an arc to a higher S state. The black curve is the shot noise prediction according to Equation 5.23; black triangles are the partition standard deviations falling below the upper confidence bound for bursts binned along the x-axis (20 bins on the interval [-0.025, 1.025], with bin width $\Delta S = 0.05$); red triangles are the partition standard deviations exceeding the upper confidence bound and indicate dynamics.

bound (only black triangles), we can still qualitatively infer dynamics based on the distribution shape. S-BVA distributions as in Figure 5.8(a) are typically angled toward $S = 0.5$.In the sparsely populated scenarios, the S-BVA distributions become more vertical and show asymmetries, such as the slight arcs up-and-to-the-right. The presence of such arcs suggests transient excursions to short-lived states.

We infer from the simulation results that S-BVA is most sensitive to kinetics on the millisecond timescale. Interconversion rates much slower than the diffusion time result in apparently static species. Interconversion rates much faster than the diffusion time result in single, apparently static, species. In conclusion, S-BVA detects timescale-dependent fluctuation dynamics.

5.4.3 Effect of Background on S-BVA

We next sought to determine the effect of background on S-BVA results. We simulated three scenarios: i) background-free, ii) experimental background, with background rates equal to the rates calculated from the RP_o dynamics sample, and iii) high background, with background rates double that of (ii). Figure 5.10 shows the results.

Background has two effects. First, background shifts distribution peaks. This shift is predictable, however. When donor excitation background counts exceed acceptor excitation background counts, $S_{bkgd} < 0.5$. Hence, a convolution of background counts with $S_{true} > 0.5$ will result in the observed S peaks shifted toward lower S values. The magnitude of the shift is dependent on i) the separation of S_{bkgd} and S_{true} and ii) the background rates. Second, background reduces the observed dynamics, which is indicated by a smaller distance between partition (red triangles) exceeding the upper confidence bound and the shot noise prediction.

Accordingly, we conclude that minimizing background is important for S-BVA to detect fluctuation dynamics, especially for closely spaced states.

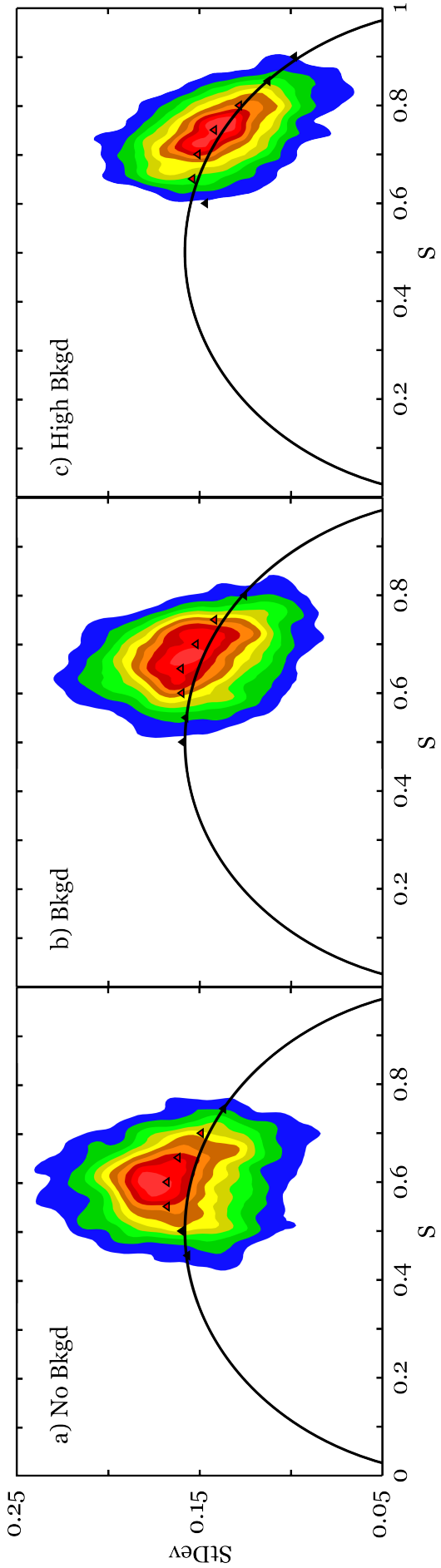


Figure 5.10: Effect of Background on Stoichiometry Burst Variance Analysis (S-BVA). Photon recoloring simulation of 2778 bursts, $\{S : 0.5, 0.7\}$, $k_{01} = k_{10} = 1000\text{s}^{-1}$, and an S-BVA window $W = 10$. Background has two effects: (i) shifts the distribution in S and (ii) reduces the observed dynamics (as indicated by the reduced distance between the partitions (red triangles) exceeding the shot noise prediction and the shot noise prediction (black curve)). Hence, minimizing background is important in enabling S-BVA to identify dynamic fluctuations, especially for closely spaced states. (a) background-free state inter-conversion. (b) experimental (normal) background: $\lambda_{DD} \approx \lambda_{DA} = 10\text{kHz}$, $\lambda_{AD} = 500\text{Hz}$, and $\lambda_{AA} = 5\text{kHz}$. (c) high background: $\lambda_{DD} \approx \lambda_{DA} = 20\text{kHz}$, $\lambda_{AD} = 1\text{kHz}$, and $\lambda_{AA} = 10\text{kHz}$.

5.4.4 S-BVA Detects Non-FRET Photophysics

While E-BVA is designed to detect conformational dynamics, S-BVA is intended to detect distance-independent photophysical changes. Detecting distance-independent photophysical changes is important because we can thereby ensure that observed E fluctuations are due to conformational dynamics alone and not fluorophore blinking (i.e., temporary inactivation, a.k.a. a 'dark', or non-photon emitting, state) or fluorescence quenching. To demonstrate the ability of S-BVA to detect distance-independent photophysical fluctuations, we simulated donor and acceptor blinking.

Fluorophore blinking is often characterized by short excursions to a non-emitting state, commonly the triplet state. Triplet state lifetimes range from 2-3μs (Cy3B) to longer than 1ms (ATTO647N), depending on buffer conditions and triplet quencher availability [51]. The observed fluorophore occupancy of a non-emitting state varies considerably. Some fluorophores enter this state repeatedly; some only occasionally. We sought to reproduce the varied behavior in Figure 5.11.

We first simulated longer-lived but relatively infrequent donor blinking with $k_{blink} = 200\text{s}^{-1}$, $k_{recover} = 10000\text{s}^{-1}$, and $S_{blink} = 0$; see Figure 5.11a. While the weight of the population is located near the shot noise prediction, we find a high arc of partition values toward lower S values. The fluctuation magnitude, which, in the figure, is the distance between red triangles and black curve, is due to the large separation between S states. Even though S_{blink} is sparsely populated, S-BVA still detects donor blinking.

We increased the frequency of donor blinking $k_{blink} = 10000\text{s}^{-1}$ and recovery $k_{recover} = 1 \times 10^5\text{s}^{-1}$; see Figure 5.11b. The high arc is no longer present because the inactive donor state is only transiently populated. The increased blinking frequency and large state separation, however, has shifted the entire S-BVA distribution toward high standard deviation. This indicates distance-independent photophysical dynamics. Here, however, we would not be able to infer the exact cause of the dynamics–we could attribute observed dynamics to either rapid donor blink-

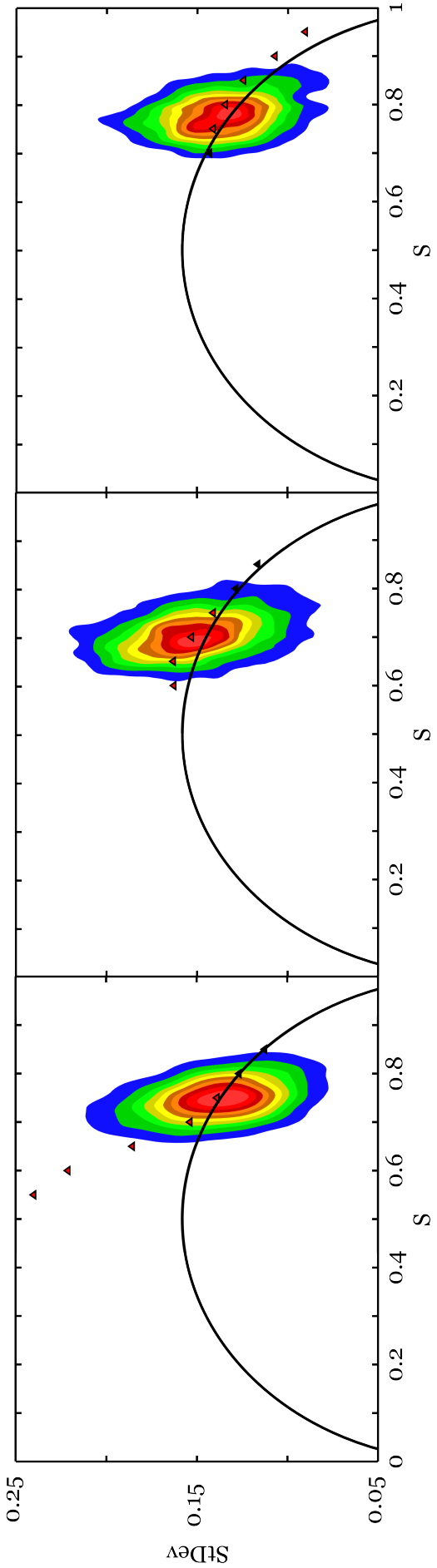


Figure 5.11: Stoichiometry Burst Variance Analysis (S-BVA) Detects Non-FRET Photophysics. Photon recoloring simulation of 2778 bursts, $S = 0.7$, and an S-BVA window $W = 10$. (a) donor blinking with $k_{blink} = 200\text{s}^{-1}$, $k_{recover} = 10000\text{s}^{-1}$. The large separation between S states, $S_{blink} = 0$ and $S = 0.7$, enables S-BVA to identify relatively infrequent and short-lived donor blinking. The high arc (red triangles) indicates infrequent but rapid inter-conversion to a low S state. (b) donor blinking with $k_{blink} = 1 \times 10^4\text{s}^{-1}$, $k_{recover} = 1 \times 10^5\text{s}^{-1}$. Frequent short-lived donor blinking shifts the distribution weight above the shot noise prediction and toward lower S values. (c) acceptor blinking with $k_{blink} = 1000\text{s}^{-1}$ and $k_{recover} = 1 \times 10^4\text{s}^{-1}$. Millisecond acceptor blinking ($S_{blink} = 1$) shifts the distribution weight above the shot noise prediction and toward higher S values. S-BVA can thus detect distance-independent photophysical dynamics, providing a diagnostic tool for fluorophore fluorescence stability.

ing or another low S state with different kinetics.

We simulated acceptor blinking on the diffusion timescale with $k_{blink} = 1000\text{s}^{-1}$, $k_{recover} = 10000\text{s}^{-1}$, and $S_{blink} = 1.0$; see Figure 5.11c. Transient acceptor blinking shifts the S-BVA distribution toward higher S values, as well as toward higher standard deviation. The trail marked on the figure by the red triangles toward high S values suggests acceptor blinking, as we sought to demonstrate.

In conclusion, S-BVA can detect distance-independent photophysical fluctuation dynamics on a variety of timescales. Calibrating relative brightness such that S is positioned at 0.5 facilitates easy identification of either donor or acceptor blinking and increases the timescales over which we can detect these sources of S dynamics.

5.4.5 S-BVA Dynamic Resolution

We have already observed that fast kinetics and background noise limit our ability to resolve closely spaced states using S-BVA (Figures 5.9 and 5.10). We subsequently asked what is the resolution limit under ideal circumstances? Figure 5.8 demonstrates S-BVA can resolve states separated by $\Delta S = 0.20$, even in the presence of background. We consequently simulated two-state inter-conversion with millisecond dynamics, $k_{01} = k_{10} = 1000\text{s}^{-1}$, and reduced the state separation in 0.05 increments until S-BVA failed to show dynamics.

In the absence of background, S-BVA clearly suggests dynamics for $\{S : 0.55, 0.70\}$ and $\Delta S = 0.15$. See Figure 5.12(a). If we decrease the separation to $\Delta S = 0.10$ with states $\{S : 0.60, 0.70\}$, S-BVA still resolves dynamics (Figure 5.12(b)) but barely and not consistently. Repeated simulation runs find dynamics only part of the time (data not shown). The distribution shape, however, remained at least qualitatively suggestive of dynamics. If we decrease the state separation to $\Delta S = 0.05$ with states $\{S : 0.65, 0.70\}$, the result is that S-BVA fails to detect dynamics, even in terms of distribution shape (Figure 5.12(c)).

If we include background with experimental background rates as calculated

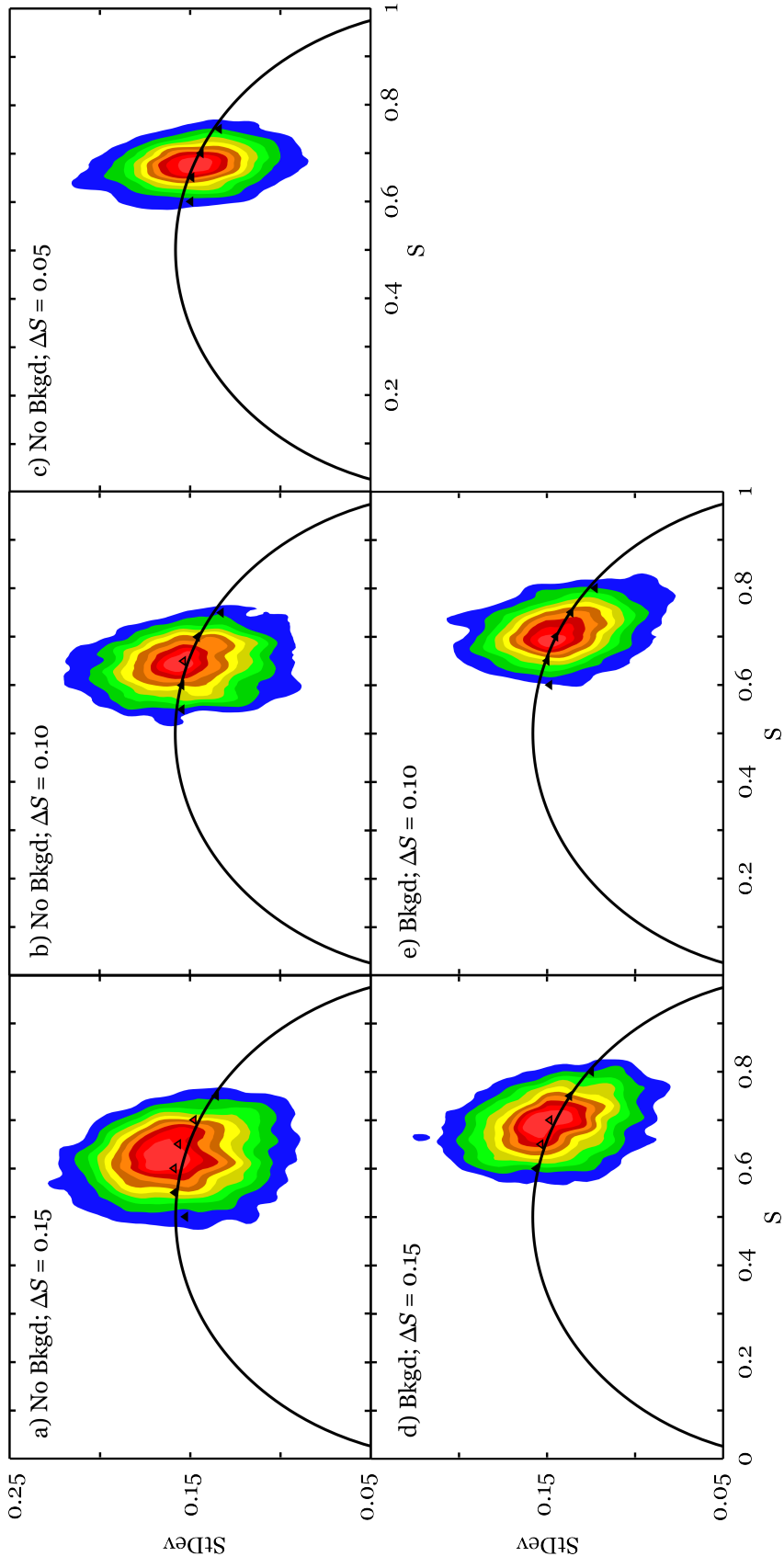


Figure 5.12: Stoichiometry Burst Variance Analysis (S-BVA) Resolves Well-Separated States. Photon recoloring simulation of 2778 bursts, $k_{01} = k_{10} = 1000\text{s}^{-1}$, and an S-BVA window $W = 10$. (a) background-free state inter-conversion with $\{S : 0.55, 0.70\}$. For large fluctuation amplitudes and low background, S-BVA readily identifies dynamics. (b) background-free state inter-conversion with $\{S : 0.60, 0.70\}$. $\Delta S = 0.10$ is barely resolved as dynamic, setting a limit to the ability of S-BVA to identify state inter-conversion. (c) background-free state inter-conversion with $\{S : 0.65, 0.70\}$. S-BVA identifies closely separated states as a single shot-noise-limited state, centered about the average state value. (d) state inter-conversion as in (a) and experimental background. As in Figure 5.10, background shifts the distribution weight toward higher S values (for the case where green excitation background is higher than red excitation background) and reduces the dynamic score, i.e., distance between partition standard deviations (red triangles) and shot noise prediction upper confidence bound. (e) state inter-conversion as in (b) and experimental background. While barely resolved in (b), a fluctuation amplitude $\Delta S = 0.10$ is not resolved as dynamic in the presence of background. Hence, minimizing background is key in resolving closely spaced S states.

from the RP_o dynamics experimental sample, a state separation of $\Delta S = 0.15$ shows an S-BVA distribution situated close to the shot noise prediction, but the sample is still detected as dynamic. Reducing the separation to $\Delta S = 0.10$, however, results in an S-BVA distribution that resides along the shot noise prediction. Consequently, we would conclude that the sample is static. While the distribution shape is varied, the inference that we observe dynamics is more tenuous than the background-free case.

In conclusion, S-BVA requires a state separation $\Delta S \geq 0.10$, especially in the presence of background, in order to resolve fluctuation dynamics. By slowing the diffusion time, thereby increasing photon statistics, and eliminating background, the resolution limit can be overcome, but resolving small fluctuation amplitudes ($\Delta S \leq 0.05$) is nearly impossible with S-BVA . We cannot even qualitatively infer dynamics based on distribution shape.[4] The same conclusion holds for E-BVA .

5.5 Experimental Results

We previously performed E-BVA in an attempt to probe RP_o dynamics. We designed three double-stranded DNA (dsDNA) promoter sequences (see Appendix D.1 and Figure D.1), differing only in the fluorophore labeling positions. We designed two constructs, (-15,+15) and (-5,-3) to directly sense RP_o dynamics [50,169], and a third construct, (-15,-25) , to serve as a static control [169].[5]

We found, however, that E-BVA was unable to conclusively support active sampling by RNAP in the promoter open complex. Instead, we observed 'dynamics' even in our control experiments, suggesting that sources other than RP_o sampling may contribute to fluctuations beyond the shot noise prediction. One possible source is distance-independent photophysics, such as donor or acceptor

[4]Note: this means, even for $\gamma \ll 1$ or $\gamma \gg 1$, we should not expect to detect stoichiometry fluctuations due to conformational/FRET changes, i.e., due to γ. For $\gamma = 0.5$, $F_{DD} + F_{DA} = 10$, $F_{AA} = 5$, and $\Delta E = 0.33$, $\Delta S \approx 0.056$, which is below S-BVA resolution. Hence, any fluctuations detected by S-BVA are almost certainly due to relatively large scale distance-independent photophysical dynamics.

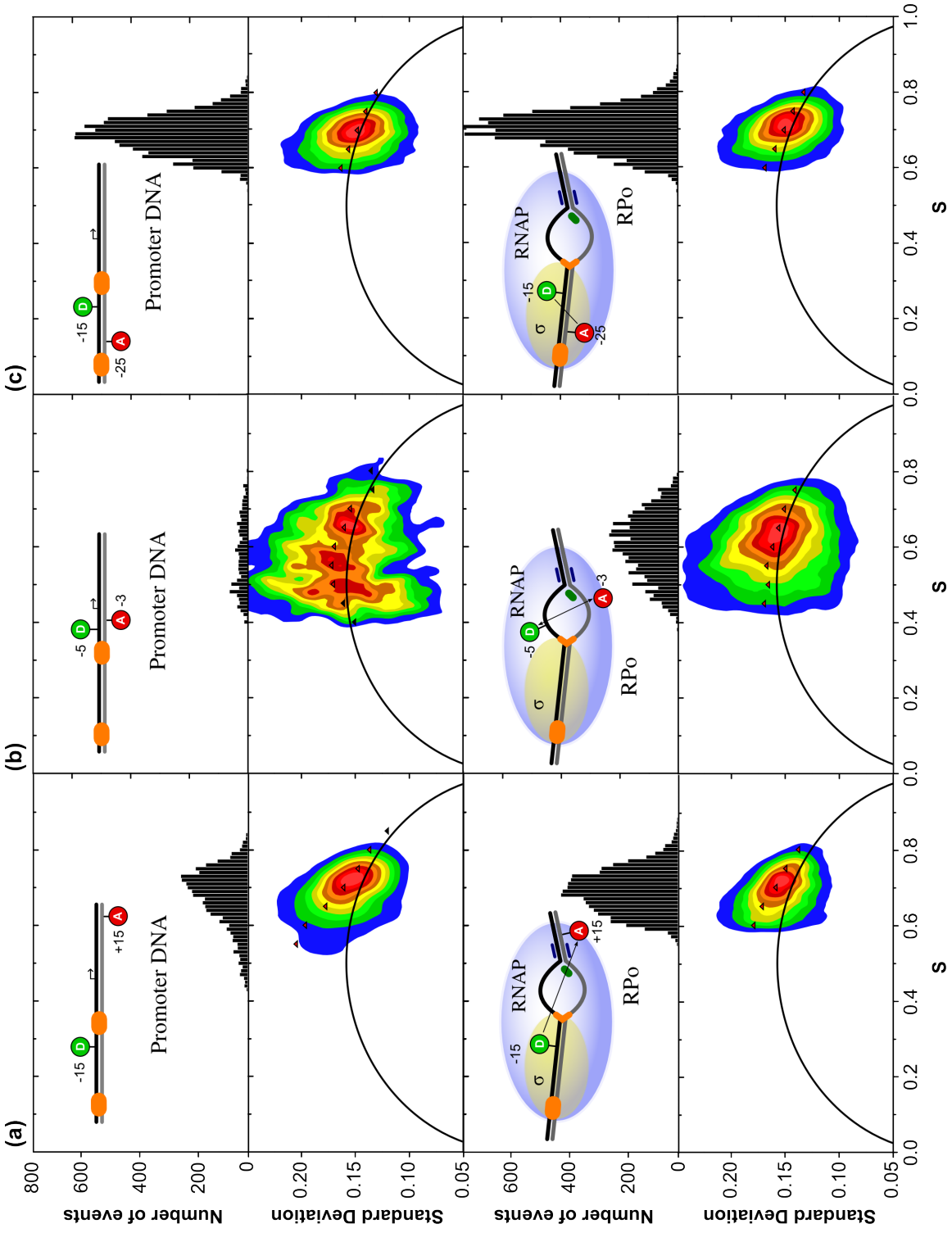
[5]For construct naming convention and further details, see Section 4.4.

blinking, with acceptor blinking being particularly problematic as apparent E fluctuations would be due to distance-independent photophysics and not exclusively inter-fluorophore distance changes.

To provide evidence for the presence of millisecond S dynamics, we performed S-BVA on the same experimental data sets as in Figure 4.4 with $W = 10$ and $\alpha = 0.001$. As a reminder, for each dsDNA construct, we performed measurements in both the absence and presence of RNAP holoenzyme (RNAP-σ^{70}). Upon holoenzyme addition, RP_o formation occurs, resulting in dsDNA melting, transcription bubble formation, and downstream dsDNA being pulled closer to the transcription start site.

Beginning with (-15,+15) (Figure 5.13(a)), in the case of dsDNA only, we expect stable donor and acceptor micro-environments and shot-noise-limited S fluctuations. The dsDNA histogram and S-BVA suggest otherwise, with the histogram showing a broad distribution and S-BVA detecting dynamics exceeding the shot noise upper confidence bound, as indicated by the red triangles. The same is true

Figure 5.13 ***(following page)*****:** Stoichiometry Burst Variance Analysis (S-BVA) of RP_o Dynamics. We performed S-BVA on the same promoter DNA constructs previously analyzed with E-BVA (Figure 4.4), where the constructs differ only in the fluorophore labeling positions. (a) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -15 and acceptor at +15, relative to main transcription start site. The upper-left histogram shows the DNA alone and corresponding S-BVA plot below. Upon addition of RNAP-σ^{70} (holoenzyme), the distribution weight remains similar to the dsDNA alone, with only one distinct major population. S-BVA suggests, however, distance-independent photophysical fluctuations due to donor blinking. Donor blinking kinetics change slightly (possibly to faster timescales) in the presence of holoenzyme. (b) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -5 and acceptor at -3. Due to dynamic quenching [50], we see few dsDNA events. Upon addition of holoenzyme, quenching is removed due to RP_o formation, resulting in a high S population. S-BVA plots for both dsDNA and RP_o are below their corresponding histograms. (c) *lac*CONS+2 dsDNA labeled with a donor fluorophore at -15 and acceptor at -25. S-BVA plots are below their corresponding histograms. The distribution weight remains the same, with increased S fluctuations in the presence of holoenzyme. Key: orange regions on the promoter DNA are the -10 and -35 promoter elements to which σ^{70} tightly binds; the arrow indicates the main transcription start site and transcription direction; donor fluorophore is green; acceptor fluorophore is red; the blue oval represents RNAP; the yellow oval represents σ^{70} ; the green marker inside the transcription bubble marks the main transcription start site; the blue lines downstream of the transcription bubble represent the downstream dsDNA contacts for RNAP; the black line in the S-BVA plots is the shot noise prediction ($W = 10$); black triangles indicate burst S fluctuations in accordance with a static prediction; red triangles indicate S fluctuations exceeding the static prediction and are evidence of dynamic inter-conversion. All promoter DNA control experiments *and* RP_o show distance-independent photophysical dynamics, primarily donor blinking; however, acceptor blinking appears present as well, for all samples, except for the (-5,-3) dsDNA construct, with dynamics extending from the main population toward a higher S state.


(a)
(b)
(c)
Promoter DNA
RNAP
σ
RPo
-15
-25
-5
-3
+15
D
A
Number of events
Standard Deviation
s
0
200
400
600
800
0.05
0.10
0.15
0.20
0.0
0.2
0.4
0.6
0.8
1.0

upon holoenzyme addition and RP_o formation. S-BVA shows both a high arc toward low S and a tail toward high S. We previously saw the high arc toward low S for the case of donor blinking (Figure 5.11(a-b)). For low background counts, E-BVA will primarily sample photons only when the donor is active. Thus, while donor blinking is not desired, E-BVA dynamics may still be attributed to conformational dynamics.

The tail toward high S is more problematic, however, particularly for RP_o where dynamics toward high S easily exceed the shot noise prediction. If acceptor blinking occurs on the millisecond timescale during donor excitation, E-BVA will detect dynamics. While we cannot rule out conformational dynamics, S-BVA suggests distance-independent photophysical fluctuations may contribute to fluctuations in E.

S-BVA for the (-5,-3) construct also detects S fluctuations exceeding the shot noise prediction (Figure 5.13(b)). For dsDNA alone, we expect S fluctuations due to dynamic quenching [50]. Possibly interesting, S-BVA seems to detect three populations. Whether these populations are significant or merely aberrations is unknown. Upon RP_o formation, a broad S distribution results. Similar to the (-15,+15) promoter construct, S-BVA detects fluctuations exceeding shot noise prediction toward both higher and lower S values. The detected dynamics, however, appear closer to prediction than for (-15,+15). This difference could be due to a difference in kinetic timescales (either more or less frequent blinking). While we expect the donor and acceptor fluorophore micro-environments for (-15,+15) to be more stable, as the fluorophores are positioned outside the protein [169], the micro-environments for the (-5,-3) construct are unknown and likely to significantly vary due to their being within the transcription bubble, possibly facilitating protein-fluorophore interaction. Other data sets hint at two closely spaced S states comprising the rather large distribution center (data not shown).

The third promoter dsDNA construct (-15,-25) also shows S dynamics for both dsDNA alone and RP_o (Figure 5.13(c)). While the distribution centers remain the

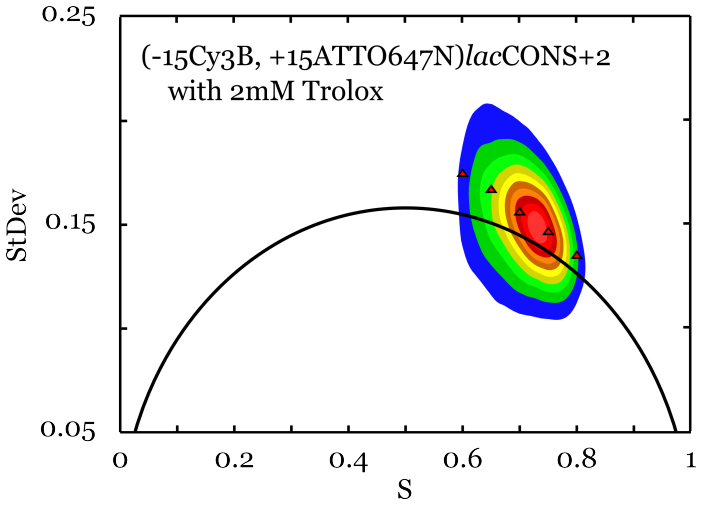


Figure 5.14: Stoichiometry Burst Variance Analysis (S-BVA) and the Effect of Trolox. We examined the effect of 2mM Trolox in minimizing distance-independent photophysical dynamics for (-15,+15) dsDNA. In comparison with Figure 5.13(a), Trolox does not appear to have any effect, with the distribution center unchanged and dynamics (red triangles) extending toward both lower and higher S values. That Trolox, a triplet quencher [51], does not reduce dynamics suggests that intersystem crossing to the triplet state is not the primary source of fluctuations.

same, S dynamics toward both higher and lower S values upon RP_o formation further exceed the shot noise prediction than for dsDNA alone. Based on crystal structures [169], we expect fluorophores positioned at -15 and -25 to reside outside the transcription bubble, but these labeling positions do not preclude protein interaction.

To determine if S dynamics could be mitigated using triplet quenchers, we performed measurements with 2mM Trolox, an antibleaching, antiblinking agent which quenches triplet states via electron transfer and recovers via a complementary redox reaction [51, 262]. Figure 5.14 shows that Trolox failed to have an effect for (-15,+15) , and other constructs showed similar results (data not shown). That Trolox did not decrease dynamics suggests Cy3B and ATTO647N intersystem crossing to the triplet state is not the primary cause of S fluctuations. Other possible causes may be quenching by nearby guanines [70, 315] (see Figure D.1), presence of oxygen (although oxygen, a triplet quencher, is more relevant for photobleaching) [2,365,369], or ATTO647N DNA interaction or spectral shift [113,165, 331].

We further explored the possibility of removing fluorophore blinking, if present,

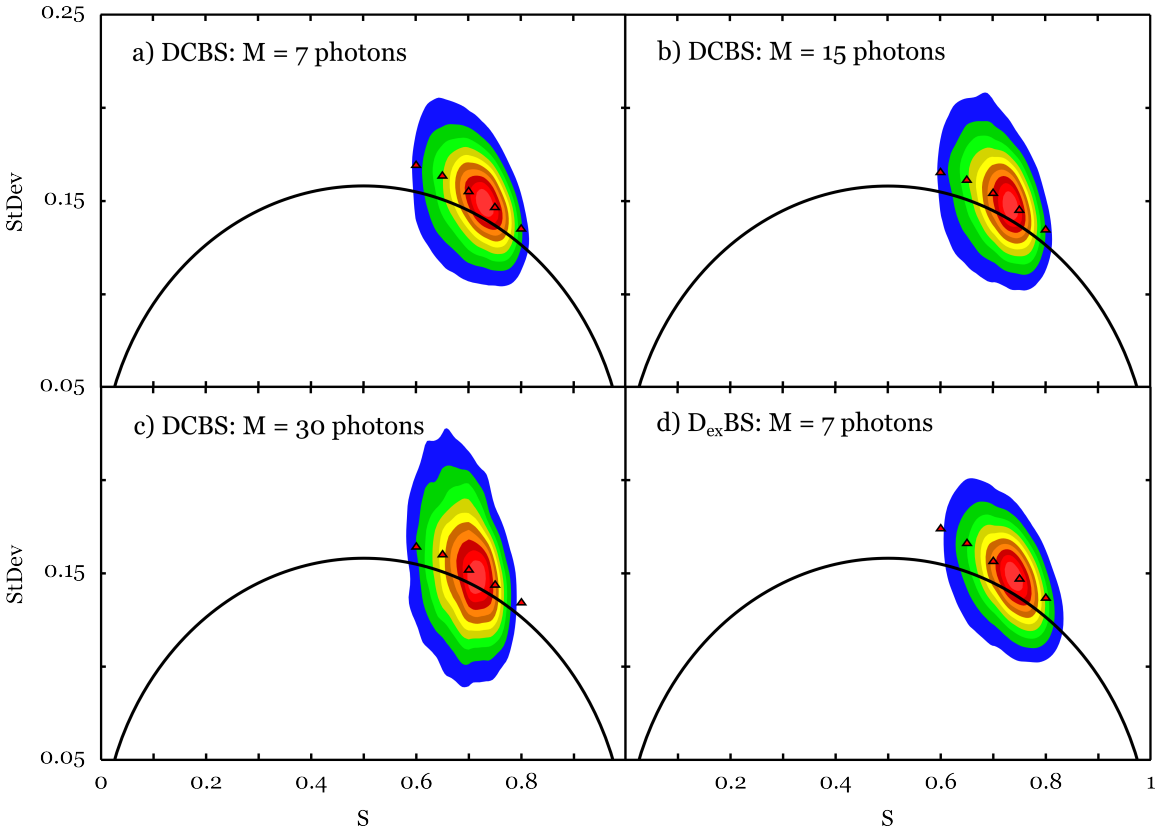


Figure 5.15: Stoichiometry Burst Variance Analysis (S-BVA) and the Effect of Burst Search Method. We investigated how the burst search method [239] affected S-BVA results. Analyzed data is the same as Figure 5.14. (a) Dual-channel burst search (DCBS) with $M = 7$ photons. (b) DCBS with $M = 15$ photons. (c) DCBS with $M = 30$ photons. (d) Donor excitation channel burst search (D_{ex}BS) with $M = 7$ photons. Figures (a-d) indicate that the choice of burst search method does not significantly affect S-BVA results for RP_o dynamics. Interestingly, requiring more neighboring photons, i.e., increasing M, does not change the extent of dynamic fluctuations. These results suggest, if blinking is present, rapid fluorophore recovery from a non-emitting state.

by applying different burst search methods [239]. Analysis has thus far used an acceptor-excitation acceptor-emission burst search ($A_{ex}A_{em}$BS), which finds bursts based solely on the $A_{ex}A_{em}$ photon emission stream. Alternatively, ALEx permits other burst search methods. For example, a donor-excitation burst search (D_{ex}BS) finds bursts based on the D_{ex} photon emission stream. A dual-channel burst search (DCBS) performs a separate burst search on $D_{ex}D_{em}$ and $D_{ex}A_{em}$ photon emission streams and only retains bursts found in both $D_{ex}D_{em}$ and $D_{ex}A_{em}$ channels. Each burst search method is biased to include particular types of bursts. $A_{ex}A_{em}$BS ensures bursts where molecules have an acceptor; however, $A_{ex}A_{em}$BS is biased to include bursts with donor, but not acceptor, blinking. D_{ex}BS ensures bursts where molecules have a donor; however, D_{ex}BS is biased to include bursts with acceptor,

but not donor, blinking. DCBS ensures bursts where molecules have both a donor and an acceptor and excludes bursts with either donor or acceptor blinking.

Recall that a photon sequence is a burst if, within a time window of $T = 500\mu s$, we find M photons. Accordingly, if we decrease T, we increase our temporal resolution, i.e., we can detect shorter bursts, but at the risk of possibly splitting bursts due to molecules diffusing around the confocal volume. If we increase M, we bias burst selection toward molecules diffusing through the confocal volume center, i.e., where the laser excitation intensity is greatest, and exclude molecules diffusing toward the confocal volume periphery. But also, by either decreasing T or increasing M, we require increased fluorophore brightness and select against fluorophore blinking.

Figure 5.15 explores the extent to which the burst search method affects S-BVA. We explored both alternative methods, D_{ex}BS and DCBS, and different burst search parameters. For all tested methods and parameters, S-BVA results remained consistent and did not significantly differ from $A_{ex}A_{em}$BS (for comparison to $A_{ex}A_{em}$BS , see Figure 5.14). Interestingly, increasing M did not significantly reduce the extent of dynamic fluctuations in S. We would expect that we would select against bursts exhibiting sufficiently slow fluorophore blinking. As dynamics persist, we surmise that blinking must entail rapid recovery from a non-emitting state.

In conclusion, all samples show both E and S fluctuations that exceed the shot noise prediction. As $\gamma \neq 1$ is unlikely to produce detectable S fluctuations for a dynamic E, S fluctuations are likely due to distance-independent photophysics. While S-BVA cannot explicitly address whether E-BVA fluctuations are due to distance-independent photophysics (e.g., blinking), we should nevertheless doubt that E-BVA dynamics are genuine. We saw for both E-BVA and S-BVA that the presence of RNAP seemed to induce greater fluctuations; this suggests fluorophore-protein interaction. Additional experimental design is thus needed to stabilize fluorophore micro-environments while still ensuring sufficient sensi-

tivity to conformational changes of interest.

5.6 Summary

We have shown that S-BVA can be an effective method for detecting millisecond distance-independent photophysical dynamics in single-molecule data. Used in conjunction with fluorescence lifetime analysis [165], ALEx-2CDE [310], and other methods [42, 44], S-BVA provides a statistical method to differentiate static and dynamic distance-independent photophysical heterogeneity. Distinguishing distance-independent photophysics from FRET fluctuations due to conformational dynamics remains a significant challenge in single-molecule fluorescence, and further experimental and statistical methods are needed to remove unwanted photophysics and isolate species of interest.

We note in passing possible additional applications for S-BVA, beyond detecting unwanted photophysical fluctuations. Recent work exploits protein induced fluorescent enhancement (PIFE) to detect protein binding and dynamics [85, 147, 207, 234, 293]. By using PIFE, PET, or any other assay relying on fluorescence enhancement or quenching, in combination with a spectrally distinct non-FRET fluorophore pair, S-BVA can hypothesis test for conformational dynamics, similar to E-BVA . While fluorophore blinking would remain an issue, S-BVA would still distinguish between multiple static species and dynamically inter-converting species.

S-BVA builds upon previous work investigating shot-noise-limited distributions and provides a straightforward test for static versus dynamic stoichiometry heterogeneity as the origin of excess distribution width. We expect broad application of S-BVA in detecting non-FRET fluorescence dynamics.

Contributions

Joseph Torella first proposed Equation 5.23 in his work on stoichiometry shot noise (unpublished). The theory and implementation of S-BVA is my own. I performed all simulations validating S-BVA . Thorben Cordes and Nicole Robb conducted all RP_0 start-site selection measurements, and, on these measurements, I performed S-BVA. All writing and figures are my own.

Chapter 6

Burst Correlation Analysis

Abstract

Single-molecule Förster resonance energy transfer (FRET) proposes a distinct advantage in its promise to relate molecular structure to function. The method, however, is not without significant obstacles, particularly in delineating conformational dynamics and distance-independent photophysics. While multiparameter fluorescence detection (MFD) methods [44,165] allow for identifying the presence of distance-independent photophysical dynamics via changes in fluorescence lifetimes, the experimental systems place significant demands in terms of outlay and setup. Here, we exploit alternating laser excitation (ALEx) spectroscopy to propose a simple statistical method called burst correlation analysis (BCA) that measures the extent to which distance-independent photophysics affect FRET dynamics. By correlating FRET and fluorophore stoichiometry for individual molecules over time, where both FRET and stoichiometry are ratiometric variables which are independent of a molecule's 3D position within a confocal volume, we can determine whether distance-independent photophysics can account for FRET fluctuations. We use simulations to validate BCA and describe its utility in probing fluorophore fluorescence stability. We then use BCA to analyze ALEx data of RNA polymerase (RNAP) open complex. Our results confirm that distance-independent photophysics can at least partially account for apparent RNAP open complex FRET dynamics. We are not, however, able to determine the exact source of photophysical correlations.

6.1 Introduction

Burst variance analysis (BVA) attempts to detect dynamic fluctuations as individual molecules emitting donor and acceptor fluorescence diffuse through a confocal volume. Thus far, BVA has only considered one random variable in isolation, i.e.,

E-BVA or S-BVA , and has not addressed correlated dynamics between multiple variables. E-BVA is limited to fluctuations in E, and S-BVA is limited to S. Although S-BVA dynamics are likely to physically correlate with E-BVA dynamics, S fluctuations do not entail E fluctuations, or vice versa (see Appendix C.1). For a physical example in which E and S are uncorrelated, consider the case in which an acceptor fluorophore may exhibit state-dependent fluctuations under acceptor excitation, while maintaining a constant fluorescence intensity under donor excitation and FRET.[1] $A_{ex}A_{em}$ states would affect stoichiometry, but not efficiency, and E and S fluctuations would be uncorrelated. Simply because S-BVA detects S dynamics, we cannot infer distance-independent photophysics as the cause of E dynamics detected by E-BVA .

We thus sought to develop a statistical method for detecting correlated E-S dynamics, within the same general BVA framework. Such a method would possibly permit us to isolate bursts free of S dynamics. We could then more confidently assert that E dynamics are due to conformational dynamics alone. We propose burst correlation analysis (BCA) as a method capable of detecting E-S correlations, and discuss its merits below.

6.2 Theory

Whereas variance measures the dispersion of a random variable X, covariance measures the extent to which two random variables X and Y change together. The definition of covariance between two random variables X and Y is

$$\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] \tag{6.1}$$

where $\mathrm{E}[X]$ is the expected value of X [191]. We can simplify Equation 6.1 using

[1] We observe such laser-excitation-dependent state behavior frequently under total internal reflection fluorescence (TIRF) microscopy, especially for ATTO647N.

the linearity property of expectations such that

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] \tag{6.2}$$

An important property of covariance is that, if X and Y are independent, their covariance is zero. This follows immediately from the definition of independence

$$\mathrm{E}[XY] = \mathrm{E}[X]\,\mathrm{E}[Y] \qquad iff\ X \text{ and } Y \text{ are independent}$$

If the values of X demonstrate similar behavior to those of Y, both rising and falling together, we expect positive covariance; we correspondingly expect negative covariance when their behaviors diverge, i.e., when one rises, the other falls. While the sign of the covariance expresses the tendency of two random variables to show similar or divergent behavior, the interpretation of the magnitude is not straightforward. The normalized version of the covariance, the *correlation coefficient*, provides a more intuitive measure concerning linear relation strength.

Correlation methods are frequently used in time series analysis to evaluate relationships between two or more variables [76, 100, 117, 118, 198, 248, 318, 351]. The Pearson product-moment correlation coefficient, ρ, is one such measure of the linear dependence between two random variables X and Y. Here, we define ρ as the covariance of X and Y divided by the product of their standard deviations:

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{6.3}$$

where μ_X and μ_Y are the population means. We can determine Pearson's correlation coefficient for a sample using estimates for the covariance and individual variances , thereby obtaining

$$r_{X,Y} = \frac{\sum_{j=0}^{N-1}(X_j - \bar{X})(Y_j - \bar{Y})}{\sqrt{\sum_{j=0}^{N-1}(X_j - \bar{X})^2}\sqrt{\sum_{j=0}^{N-1}(Y_j - \bar{Y})^2}} \tag{6.4}$$

where $\bar{X}$ and $\bar{Y}$ are the sample means and N the sample size. We can further

express Equation 6.4 in terms of normalized values. We transform the random variables X and Y to have zero mean and unit variance

$$r_{Z_X,Z_Y} = \frac{1}{N-1}\sum_{j=0}^{N-1}\left(\frac{X_j - \bar{X}}{s_X}\right)\left(\frac{Y_j - \bar{Y}}{s_Y}\right) \tag{6.5}$$

where s_X and s_Y are the sample standard deviations ($s_X = \sigma_X/\sqrt{N}$). Equation 6.5 allows us to represent the distance between a value X_j and the sample mean in units of standard deviation.

In single-molecule FRET measurements, we typically concern ourselves with identifying negatively correlated (or *anti-correlated*; $r \to -1$) intensity changes in donor and acceptor fluorescence. Such changes suggest inter-fluorophore distance changes associated with, e.g., conformational dynamics. For random variables E and S (FRET efficiency and stoichiometry, respectively), however, we seek the *absence* of correlated changes because S is a means to monitor fluorophore distance-independent photophysics and E-S correlations suggest distance-independent photophysical fluctuations.

When considering E-BVA and S-BVA , we want to determine the extent to which fluctuations in E covary with fluctuations in S. If E fluctuations are do not covary with S, we can more confidently attribute fluctuations to inter-fluorophore distance changes, commonly associated with conformational dynamics. On the other hand, if E fluctuations do covary with S, we should be wary of associating E fluctuations with biological activity.[2]

To apply E-S correlation analysis within the BVA framework, our first intuition would be to calculate Pearson's correlation coefficient r_i for each burst i. r_i would then provide a measure for the strength of E-S correlations over the entire burst. But what about the case in which an (E, S) state is only visited transiently once or, for a long burst, a few times? We know from Appendix C.1 that E and S are independent random variables for constant Poisson rate parameters λ_{DD}, λ_{DA}, and

[2] An exception would be the scenario in which we expect E-S correlations; e.g., a FRET measurement convolved with protein-induced fluorescence enhancement (PIFE).

λ_{AA}. Accordingly, for a majority of a burst duration, we expect $\mathrm{Cov}(E, S) \approx 0$, and if the transient correlations are sufficiently short or between closely spaced states, then any correlations may be effectively reduced by the much larger uncorrelated region.

Furthermore, especially for short sample E-S time series, E-S independence does not guarantee that the sample will not show correlations. Only for sufficiently large N will the covariance, and therefore r, converge to zero. For small N, E-S will often show $r \pm 1$. For example, if a molecule undergoes anti-correlated fluctuations in E and S, but, by chance, the remaining (E_{ij}, S_{ij}) pairs are strongly positively correlated, then r_i may very well be close to zero or even positive, depending on the number and magnitude of anti-correlated transitions. In this case, we may wrongly conclude that burst i shows uncorrelated, or even positively correlated, E-S behavior.

An alternative method would be to compute a 'scaled' correlation coefficient, where scale refers to a restricted sampling range [238]. As with BVA, we restrict the sampling range using a non-overlapping sliding window C over the (E_{ij}, S_{ij}) pairs, and compute

$$r_{im}(C) = \frac{\sum_{k=j}^{j+C-1}(E_{ik} - \bar{E}_m)(S_{ik} - \bar{S}_m)}{\sqrt{\sum_{k=j}^{j+C-1}(E_{ik} - \bar{E}_m)^2}\sqrt{\sum_{k=j}^{j+C-1}(S_{ik} - \bar{S}_m)^2}} \tag{6.6}$$

where $\bar{E}_m$ and $\bar{S}_m$ are the window sample means for window m. We define the number of non-overlapping windows M by

$$M = \mathrm{floor}\left(\frac{N}{C}\right)$$

We can then compute the average scaled correlation $\bar{r}_i$ at scale C

$$\bar{r}_i(C) = \frac{1}{M}\sum_{m=0}^{M-1} r_{im}(C) \tag{6.7}$$

For a segmented time series where the mean and variance remain constant across

all segments, $\bar{r}_i(C)$ is identical to the global burst correlation coefficient r_i [238].

For each $r_{im}(C)$, we can test the null hypothesis that the true correlation coefficient ρ is zero using re-sampling techniques. One such technique is a *permutation test*, which proceeds as follows:

1. From the original (E_{ij}, S_{ij}) paired data, draw a permutation ξ on the set $\{j : 0, \ldots, N-1\}$, from a population with equal probabilities for all $N!$ permutations.[3]

2. Calculate Pearson's correlation coefficient $r_\xi(C)$ for the randomized data $(E_{ij'}, S_{ij})$.

3. Repeat Steps 1 and 2 for $\Xi \ggg 1$.

4. Calculate the *p-value*[4] for the permutation test. The *p-value* is equivalent to the proportion of $r_\xi(C)$ greater than $r_{im}(C)$, where 'greater' means either in absolute magnitude (two-sided test) or in signed value (one-sided test). For example, if only five $r_\xi(C)$ are greater than $r_{im}(C)$ out of $\Xi = 1000$ trials, then $p = 0.005$; such a value is highly significant at the 99% confidence level ($\alpha = 0.01$).

We note that an equally valid approach for hypothesis testing and constructing confidence intervals uses *bootstrapping*, which is re-sampling with replacement.

While we can use re-sampling methods to also determine the p-value for $\bar{r}_i(C)$, for large Ξ, many bursts, and large M, re-sampling methods quickly become computationally expensive. The *fixed effects* method, however, allows us to determine the significance of average correlations based on the standard error SE [83, 115, 123]. Here, we assume $r_{im}(C)$ is normally distributed and estimate the SE[5] of

[3] This is equivalent to drawing each member j randomly, without replacement, from the set $\{0, \ldots, N-1\}$.

[4] i.e., the probability of obtaining a value at a given confidence level by chance alone

[5] Note that the concept of standard error (SE) differs from that of the standard deviation (SD). The standard error of the mean determines how well the sample mean estimates the population mean; whereas the standard deviation of the mean is the degree to which we expect individual measurements within the sample to differ from the sample mean.

$$SE = \frac{s}{\sqrt{N}}$$

$\bar{r}_i(C)$ from the number of windows M for burst i and window size C

$$\text{SE}(\bar{r}_i(C)) = \sqrt{\frac{1}{M(C-3)}} \tag{6.8}$$

Using the SE, we compute a normalized z-score

$$z_{\bar{r}_c} = \frac{\bar{r}_c}{\text{SE}(\bar{r}_c)} \tag{6.9}$$

and obtain a p-value by evaluating at $-|z_{\bar{r}_c}|$, where $|\cdot|$ is the absolute value, the cumulative distribution of a normal distribution with 0 mean and unit variance.

By calculating scaled correlation coefficients, we obtain information about both fast correlations and per burst correlation. Such information facilitates E-S covariance identification and lays the groundwork for isolating those bursts with independent E and S.

6.3 Simulation Results

To test burst correlation analysis (BCA), we performed recoloring simulations, as done previously for S-BVA (see Section 5.4) using the same data set under the same data reduction parameters.

6.3.1 Static S Species

To validate BCA, we performed initial simulations on static species. For infinitely long time series, we expect E and S to be uncorrelated $\bar{r}_c = 0$; however, this need not be true for random samples drawn from the global population. As $C \to 0$,

$$SD = \frac{\sigma}{\sqrt{N}}$$

where s is the sample standard deviation, i.e., a sample-based estimate of the population standard deviation σ, and N is the sample size. For example, consider the set of observations $\{X_0, X_1, \ldots, X_C\}$. Suppose, due to chance, the observations are highly dispersed, meaning a large sample standard deviation s. For comparatively small σ relative to s, SE would be large compared to SD. Accordingly, we would conclude that the sample mean is likely a poor estimator of the population mean and further surmise that the measured correlation is unreliable, thereby reducing its significance.

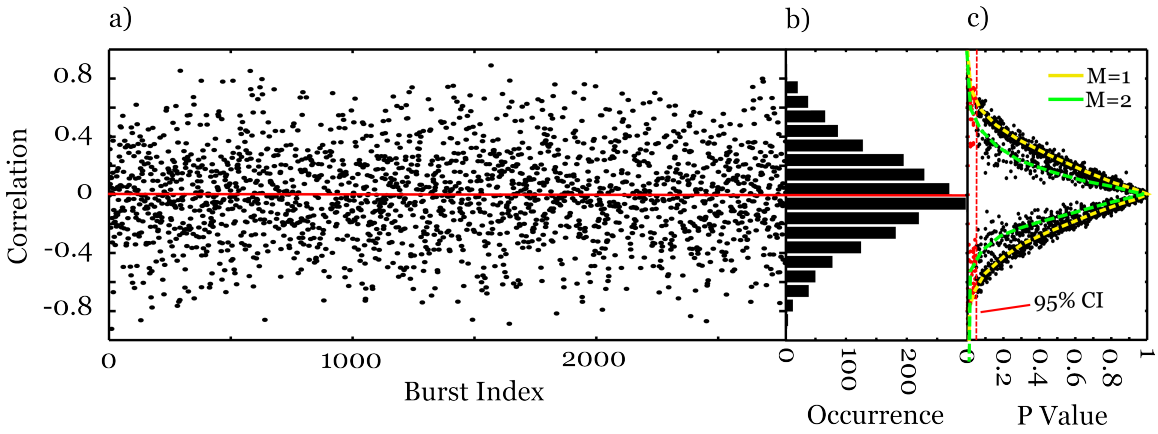


Figure 6.1: Burst Correlation Analysis (BCA) of a Static Species. Photon recoloring simulation of a single static species with $E = 0.5$, $S = 0.7$, a BCA window $C = 10$, and 2778 bursts. (a) per burst correlation. By chance alone, bursts will exhibit both high and low E-S correlation; however, the distribution will be approximately normally distributed with 0 mean (on average, E and S are uncorrelated). Red line marks $r_c = 0$. (b) correlation histogram for the correlations in (a). For static species, we expect E and S to be uncorrelated, but sampled correlation values will have a standard error of 0.378 ($= \sqrt{1/M(C-3)}$; $\langle M \rangle = 1$). c) correlation versus significance (p-values). p-values follow the shape of the correlation histogram, with more extreme correlations having increased significance. We notice that several bursts have $p \ll 0.05$ (shown in red), indicating significant correlations. This result follows from the definition of a p-value; for ~ 3000 bursts, we expect at least 150 bursts to exhibit significant correlations at the 95% confidence level by chance alone. We obtained p-values using permutation tests for $N = 120$. Longer bursts (≥ 200 photons; $M = 2$) will have greater statistical support and a decreased likelihood to exhibit extreme correlation values (Law of Large Numbers). These bursts form the inner p-value distribution (in green). For shorter bursts (< 200 photons; $M = 1$), we expect a greater spread in correlation. These bursts form the outer p-value distribution (in yellow).

we expect r_c values to fluctuate by chance alone about a global mean. Figure 6.1 confirms that even uncorrelated species will exhibit significant correlations.

We used a permutation re-sampling test to determine the significance of the average correlations $\bar{r}_c$. We expect that, for a 95% confidence interval, 150 out of 3000 bursts will show significant correlation. Should we use BCA as a filtering method, we should bear in mind the rate of false positive identification.

We next tested BCA for a shot-noise-limited $S = 0.7$ and a dynamic E. We simulated two inter-converting conformational states with FRET values $\{E : 0.4, 0.6\}$ and rates $k_{01} = k_{10} = 1000\text{s}^{-1}$. Similar to the case of a single shot-noise-limited E, we do not expect E-S correlations. For a constant total intensity under donor excitation ($F^{D_{em}}_{D_{exc}} + F^{A_{em}}_{D_{exc}} =$ constant), S will fluctuate independently of E. Figure 6.2 confirms that E and S are uncorrelated for a dynamic E, as the correlation distribution has 0 mean.

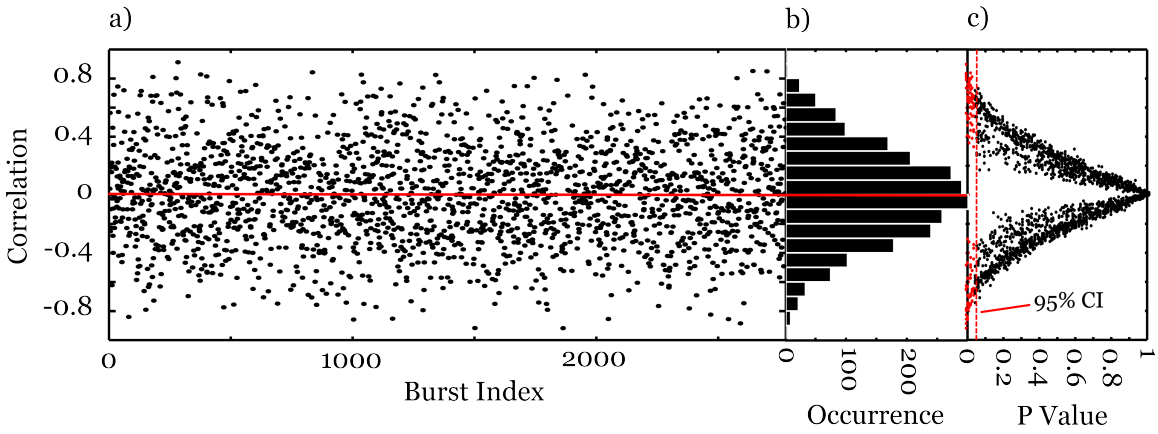


Figure 6.2: Burst Correlation Analysis (BCA) of a Conformationally Dynamic Species. Photon recoloring simulation of a conformationally dynamic species with $\{E : 0.4, 0.6\}$, $k_{01} = k_{10} = 1000\text{s}^{-1}$, $S = 0.7$, a BCA window $C = 10$, and 2778 bursts. (a-b) per burst correlation and distribution. E-S correlations appear similar to the static case, as we expect for uncorrelated E and S. c) correlation versus significance (p-values). The number of significant correlations does not change for a dynamic E. We obtained p-values using permutation tests for $N = 120$.

6.3.2 The Effect of γ

While molecules with $\gamma = 1$ and dynamic E lack E-S correlations, we should not expect E and S to be uncorrelated for $\gamma \neq 1$. $\gamma < 1$ effectively discards acceptor photons, such that $F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}$ decreases with increasing E. The decrease in D_{ex} intensity results in $S \to 0$. The opposite is true for $\gamma > 1$. Therefore, E and S will show anti-correlated changes upon E dynamics. This is confirmed in Figure 6.3, in which more bursts show negative average correlations for $\gamma = 0.5$.

Interestingly, even though S-BVA fails to detect S dynamics for $\gamma \neq 1$, BCA is able to detect subtle concomitant changes in E and S. While the shift in r_c values is small, the shift is detectable if we hold everything else constant.

6.3.3 Dynamic S: Non-FRET Photophysics

The goal of BCA is to determine if E-S fluctuations are correlated. Correlated E-S fluctuations would argue against E fluctuations exclusively attributable to inter-fluorophore distance changes (and associated conformational dynamics). We thus sought to examine BCA results for S dynamics and distance-independent photophysics. We first investigated the effect of donor blinking, performing a recoloring

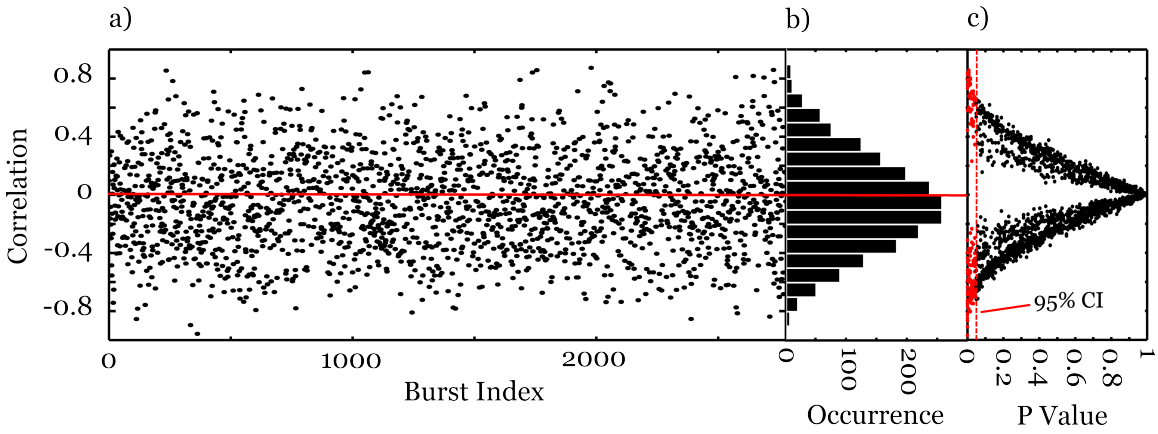


Figure 6.3: Burst Correlation Analysis (BCA) of a Conformationally Dynamic Species and $\gamma \neq 1$. Photon recoloring simulation of a conformationally dynamic species with $\{E : 0.4, 0.6\}$, $k_{01} = k_{10} = 1000\text{s}^{-1}$, $S = 0.7$, $\gamma = 0.5$, a BCA window $C = 10$, and 2778 bursts. (a-b) per burst correlation and distribution. E-S correlations shift toward a negative average correlation $\bar{r}_c$. The shift reflects $\gamma < 1$ effectively discarding acceptor photons, resulting in a variable total $F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}$ intensity, such that $S_{E=0.4} > S_{E=0.6}$. c) correlation versus significance (p-values). The number of significant correlations increases for negative r_c and decreases for positive r_c. We obtained p-values using permutation tests for $N = 120$.

simulation with blinking rate $k_{blink} = 500\text{s}^{-1}$ and recovery rate $k_{recover} = 25000\text{s}^{-1}$.

As evident in Figure 6.4, we should not necessarily expect E-S correlation for donor blinking as $\Delta E = |E - E_{bkgd}| \approx 0$. Once the donor blinks, background photon counts B_{DD} and B_{DA} determine E, and accordingly, E will not correlate with S. For sufficiently large $\Delta E = |E - E_{bkgd}| > 0$, we may expect to see a slight correlation distribution shift, whose sign is determined by the relation of E to E_{bkgd}, and possibly be similar in magnitude to the effect of $\gamma \neq 1$. For low background counts, we can modify the BCA algorithm to detect when E is undefined $(0/0)$, and then correlate $S \rightarrow 1$ with donor blinking. This modification is straightforward, and we will be investigate its use in further work; however, the ability to link donor blinking to E-BVA dynamics remains tenuous because E fluctuations due to donor blinking are likely to be limited for the same ΔE issue mentioned above.

In contrast to donor blinking, we do expect acceptor blinking to result in E-S correlations when the acceptor blinking lifetime is on the order of or greater than the $D_{ex}|A_{ex}$ alternation period; e.g., $\tau_{blink} = 1\mu$s is unlikely to result in significant correlation. In this scenario, $E \rightarrow 0$ and $S \rightarrow 1$, resulting in strong E-S anti-

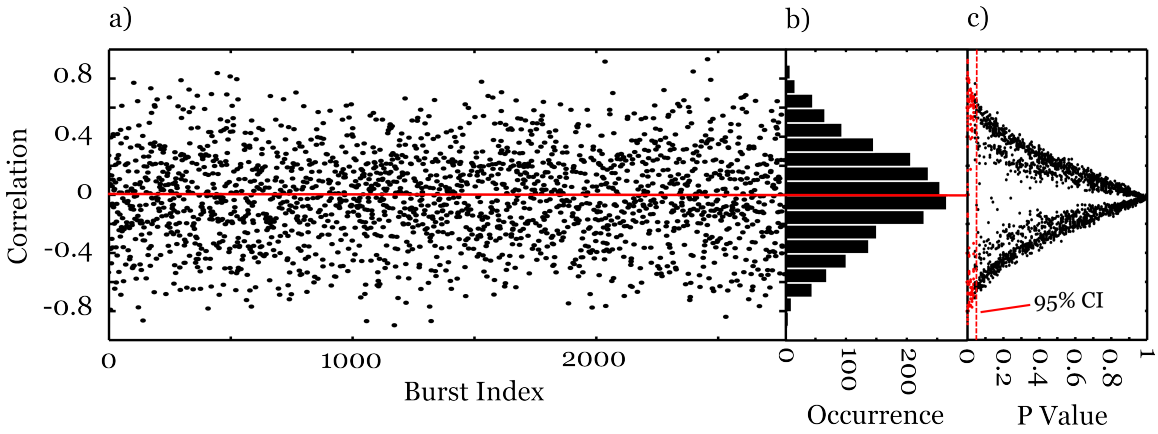


Figure 6.4: Burst Correlation Analysis (BCA) and the Effect of Donor Blinking. Photon recoloring simulation of a species exhibiting donor blinking, $E = 0.6$, $\{S : 0.7, 1.0\}$, $k_{blink} = 500\text{s}^{-1}$, $k_{recover} = 25000\text{s}^{-1}$, a BCA window $C = 10$, and 2778 bursts. (a-b) per burst correlation and distribution. E-S correlations appear the same as for the static case. For background counts such that $E_{bkgd} \approx 0.5$, we are not likely to detect E fluctuations for small ΔE because of how we calculate the correlation using a sliding photon window. That the correlation does not change, however, is advantageous from the standpoint that we can differentiate between donor and acceptor blinking. c) correlation versus significance (p-values). The number of significant correlations does not change for the case of donor blinking. We obtained p-values using permutation tests for $N = 120$.

correlation. We simulated millisecond acceptor blinking, with $k_{blink} = k_{recovery} = 1000\text{s}^{-1}$.

Figure 6.5(a) confirms E-S anti-correlation, with the distribution shifted toward negative correlation. While we expect on average one blinking event per ms, and hence per burst, we should not expect every correlation to be negative. Recall our earlier discussion and Figure 6.1 in which uncorrelated molecules will show strong positive and negative correlations by chance alone. Now consider the scenario in which a molecule randomly samples strongly positively correlated E-S values, but also contains at least one strongly negatively correlated event. The result will be a reduced correlation value, thus affecting the correlation strength for the entire burst. Accordingly, not only will BCA identify false-positives, i.e., uncorrelated E and S randomly showing strong correlation, but also will include false-negatives, i.e., correlated E and S which randomly sample values correlated with opposite sign. The black arrows in Figure 6.5(a) highlight this occurrence, in which r_c values move along the significance curve.

We also simulated the scenario in which acceptor recovery occurs on a faster timescale, $k_{recovery} = 10000\text{s}^{-1}$. The more transient excursion to the blinking state

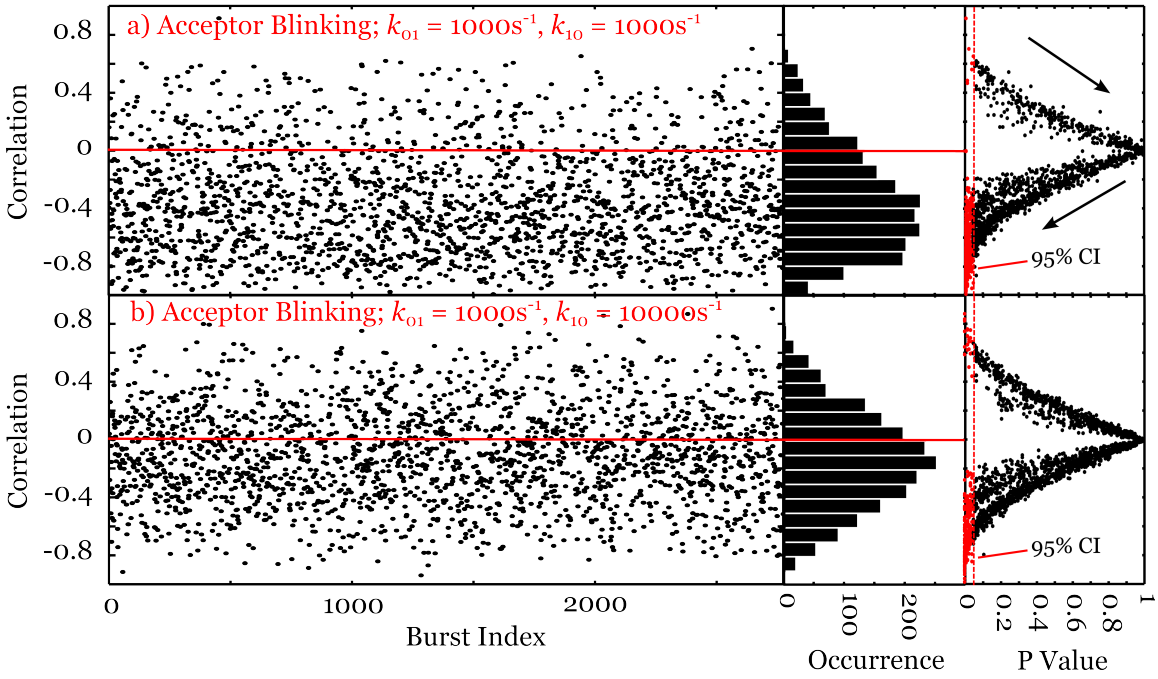


Figure 6.5: Burst Correlation Analysis (BCA) and the Effect of Acceptor Blinking. Photon recoloring simulation of a species exhibiting acceptor blinking, $E = 0.6$, $\{S : 0.0, 0.7\}$, a BCA window $C = 10$, and 2778 bursts. a) per burst correlation, distribution, and significance for millisecond acceptor blinking: $k_{01} = k_{blink} = 1000\text{s}^{-1}$, $k_{10} = k_{recover} = 1000\text{s}^{-1}$. E-S correlations due to millisecond acceptor blinking are highly negatively correlated. Negative window correlations can pass undetected, however, when bursts randomly exhibit high positive correlation. The result for an individual burst is an average $\bar{r}_c \to 0$. We indicate the shift in $\bar{r}_c$ along the significance curve with black arrows. b) per burst correlation, distribution, and significance for millisecond acceptor blinking: $k_{01} = k_{blink} = 1000\text{s}^{-1}$, $k_{10} = k_{recover} = 10000\text{s}^{-1}$. E-S correlations due to millisecond acceptor blinking and fast recovery are negatively correlated, but the correlation distribution shift is not as drastic. Nevertheless, we can detect E dynamics due to transient acceptor blinking.

means the molecule spends a majority of time in a bright state. This has the effect of weighting correlations toward an uncorrelated value, drowning out infrequent rapid anti-correlations. Nevertheless, the correlation distribution does shift. In this case, we can conclude E dynamics are at least partly attributable to acceptor blinking.

6.3.4 Photophysics within a Heterogeneous Population

The simulations have thus far assumed sample homogeneity, i.e., all diffusing molecules exhibit the same behavior. Experimental samples, however, are rarely homogeneous, with some molecules undergoing conformational fluctuations and others not, and some fluorophores exhibiting photophysical characteristics dis-

tinct from others within the population. That samples are inherently heterogeneous is a significant limitation of ensemble methods, which can only report on average population behavior. For biological systems where heterogeneity is fundamentally important [41,140,186,302,303,327], the primary molecules of interest, possibly even the ones driving biological activity, are the molecules exhibiting deviant behavior. An ensemble average obscures single-molecule renegades[6]. We thus sought to investigate how photophysical sub-populations affect BCA.

We simulated two sub-populations, one static S species and the other species exhibiting acceptor blinking with $k_{blink} = 1000\text{s}^{-1}$ and $k_{recover} = 10000\text{s}^{-1}$, under different mixing proportions. Intuitively, we can predict the outcome. For a smaller acceptor blinking sub-population, we expect the uncorrelated behavior of the major population to dominate and pull the distribution toward $\langle \bar{r}_c \rangle = 0$. As the acceptor blinking sub-population grows, more molecules will show E-S anti-correlations and pull the distribution toward $\langle \bar{r}_c \rangle = -1$. Figure 6.6 confirms our intuition.

We note, however, that we can still identify the acceptor blinking population for the 20% mixture in a way additional to the slight shift toward negative correlation. Acceptor blinking molecules will sample negative correlations more frequently than non-blinking molecules, and importantly, acceptor blinking molecules will more frequently sample the more extreme negative correlations. The result will be a change in distribution skewness, as evident in the distribution asymmetry and the longer tail toward negative correlations. Consequently, we could introduce additional hypothesis tests for distribution normality and skewness to determine the existence of photophysical sub-populations. Of course, such an analysis would become more complicated for a sample containing many sub-populations, or as is more likely, a continuum of heterogeneous behavior.

[6] I am allowed literary bravado.

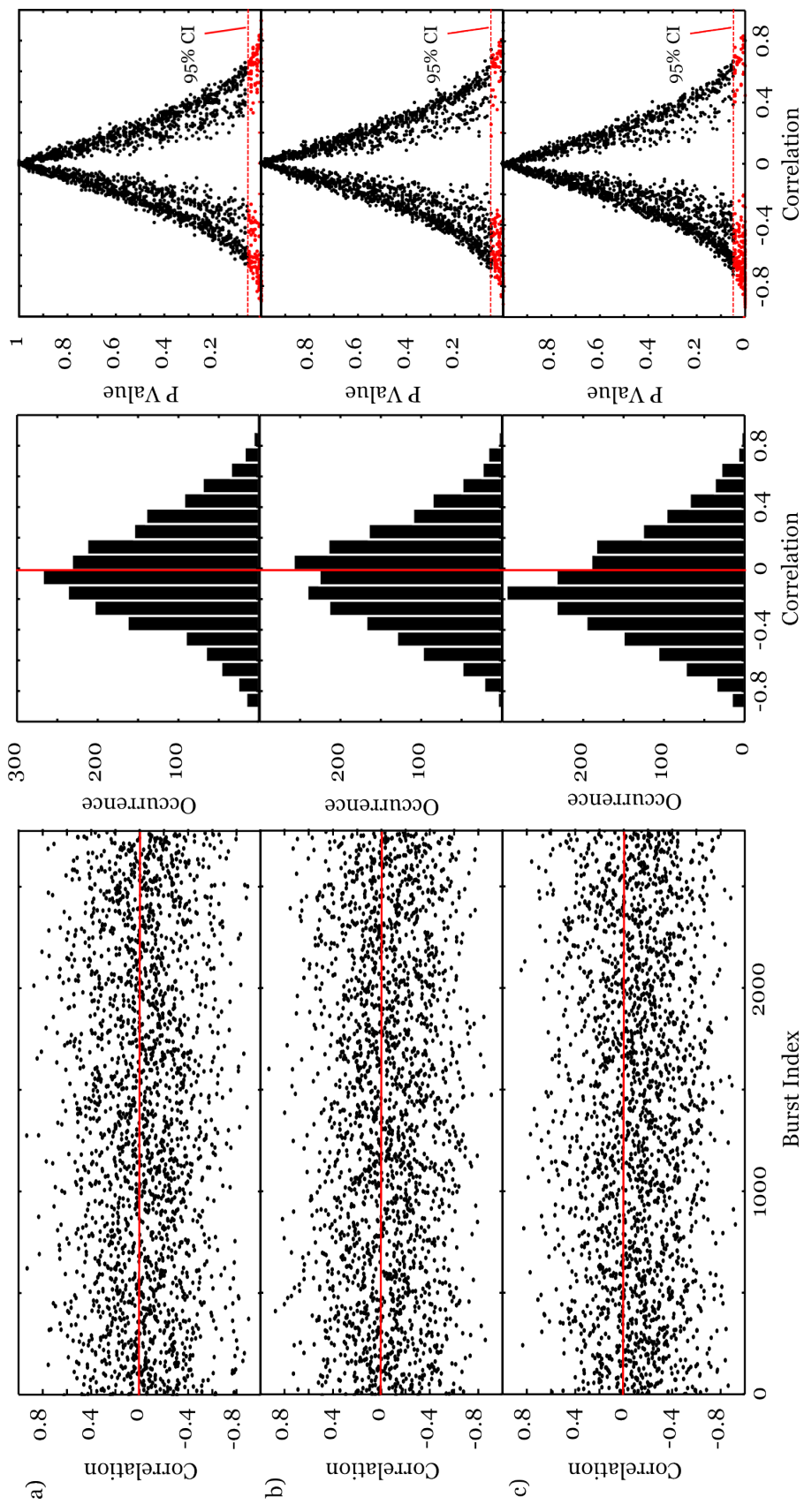


Figure 6.6: Burst Correlation Analysis (BCA) and the Effect of Acceptor Blinking within Sub-populations. We analyzed BCA sensitivity to only a fraction of molecules exhibiting S dynamics, using a photon recoloring simulation of two species: one static with $S = 0.7$ and $E = 0.6$, and the other exhibiting millisecond acceptor blinking, $E = 0.6$, $\{S : 0.0, 0.7\}$, and $k_{blink} = 1000\text{s}^{-1}$, $k_{recover} = 1000\text{s}^{-1}$. a) per burst correlation, distribution, and significance for 20% of molecules exhibiting acceptor blinking. b) 50% of molecules exhibiting acceptor blinking. c) 80% of molecules exhibiting acceptor blinking. As expected, as the fraction of molecules exhibiting acceptor blinking increases, the individual correlations shift toward negative values, thus shifting the overall distribution center. But even for small fractions, millisecond blinking is evident in the distribution skewness, with the distribution in (a) showing a longer tail toward more significant negative correlations. $C = 10$ and 2778 bursts.

6.4 Experimental Results

Having validated burst correlation analysis (BCA) as a means to measure E-S correlations, we sought to apply the method to investigate RP_o dynamics and build upon previous E-BVA and S-BVA studies. Recall that both E-BVA and S-BVA detected dynamics, but we could not determine whether E and S co-varied or determine whether distance-independent photophysics could account for E dynamics. Figure 6.7 shows BCA results.

Beginning with (-15,+15) in Figure 6.7(a), we notice significant deviations from normality, with heavily populated distribution tails. While the distribution has mean $\mu \approx 0$, the extracted standard deviation ($\sigma = 0.1994$) from the fit does not adequately capture the distribution dispersion. For this construct, acceptor blinking can possibly account for the significant negative correlations, where $S_{blink} \approx 1$ and $E_{blink} \approx 0.16$ (cross-talk). Additionally, the mean FRET efficiency for the ds-DNA subpopulation is $E \approx 0.2$. Hence, given relatively high background rates, with $E_{bkgd} \approx 0.5$, we expect E-S anti-correlations for donor blinking molecules within this subpopulation. The significant positive correlations are more difficult to explain. One possibility is a shift in the ATTO647N absorption spectrum, such that $F_{D_{exc}}^{A_{em}}$ increases and $F_{A_{exc}}^{A_{em}}$ decreases [43, 44, 113].

An absorption spectrum shift is not altogether implausible if we once again consider the E-BVA results (see Figure 4.4(a)). Surface-immobilized RP_o measurements reveal dynamic interconversion between a closed complex RP_c with E corresponding to that of dsDNA alone and an open complex RP_o with $E \approx 0.45$ (see Chapter 7). While only a fraction ($< 50\%$) of the immobilized molecules show RP_c -RP_o dynamics, ΔE is sufficiently large for E-BVA to detect dynamics even for relatively few inter-converting molecules. Such inter-conversion likely explains the detected dynamics exceeding the shot noise prediction in the E region between dsDNA/RP_c and RP_o . If we consider that RP_c to RP_o conversion requires conformational rearrangement, ATTO647N-RP_o interaction is not altogether unlikely, possibly leading to a transient change in photophysical properties. The dynamic

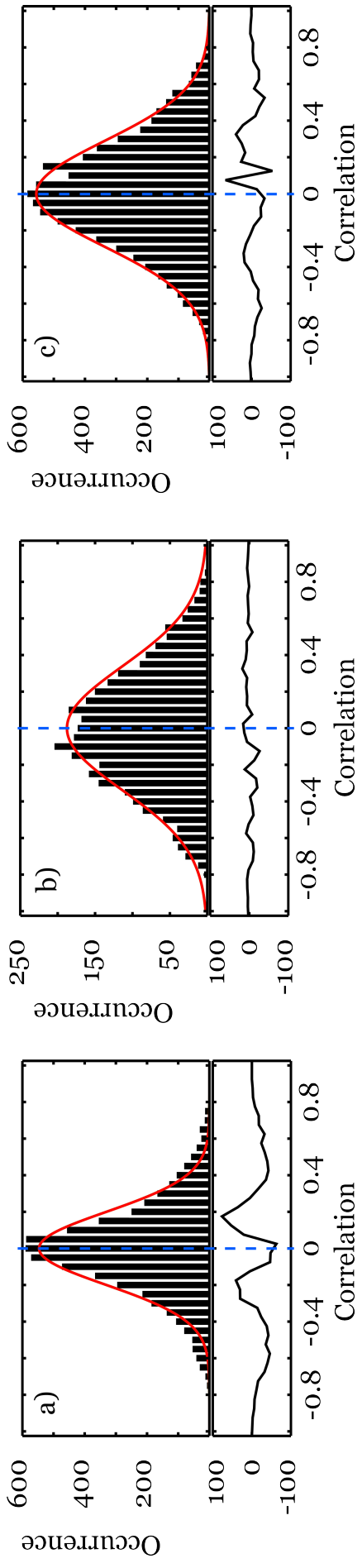


Figure 6.7: Burst Correlation Analysis of RP_o Dynamics. We performed BCA on RP_o complexes which formed on the same promoter DNA constructs we previously analyzed with E-BVA and S-BVA (Figures 4.4 and 5.13), where the constructs differ only in the fluorophore labeling positions. Here, we show BCA results for RP_o samples. For each correlation histogram, we fit a single Gaussian function to extract the fit mean μ and standard deviation σ for comparison with the standard error SE prediction. (a) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -15 and acceptor at +15, relative to main transcription start site. The distribution mean is centered about 0, with $\sigma = 0.1994$. The standard error prediction is 0.1719 for $\langle M \rangle = 4.8324$ The fit residuals show significant deviations in the distribution tails, suggesting E fluctuations are at least partly accounted for by distance-independent photophysics. (b) *lac*CONS+2 dsDNA labeled with a donor fluorophore at position -5 and acceptor at -3. The distribution mean is 0, with $\sigma = 0.3378$. The standard error prediction is 0.3147 for $\langle M \rangle = 1.4426$. The correlation distribution is slightly skewed toward negative correlations, suggesting acceptor blinking could at least partially account for E fluctuations. (c) *lac*CONS+2 dsDNA labeled with a donor fluorophore at -15 and acceptor at -25. The distribution mean is centered about 0, with $\sigma = 0.2889$. The standard error prediction is 0.2677 for $\langle M \rangle = 1.9940$. The correlation distribution is slightly skewed toward negative correlations, supporting acceptor blinking, as found in S-BVA . Red lines are single Gaussian fits.

interconversion for immobilized molecules occurs on the millisecond-to-second timescale and matches the observed timescale in BVA and BCA.

The (-5,-3) promoter construct also shows deviant behavior from normality, although the effect is not as pronounced. The correlation histogram (Figure 6.7(b)) is slightly skewed toward negative correlations and indicates fluorophore blinking. Given that the fluorophores are located within the transcription bubble, we might expect increased distance-independent photophysics for this promoter dsDNA; however, we do not observe drastic deviations from uncorrelated E-S . In fact, initial tests using BCA as a filter to remove bursts with significant correlation demonstrated S-BVA detected dynamics could be reduced, or even eliminated, while still retaining fluctuations beyond the static E prediction (data not shown). BCA filtering, however, was not consistent across samples for reasons outlined earlier with regard to significant correlations arising from limited random sampling.

We also analyzed the (-15,-25) promoter construct (see Figure 6.7(c)). Similar to (-5Cy3B, -3ATTO674N), we observe a slight skewness toward negative correlations. Recall that S-BVA showed increased S dynamics for RP_o compared to dsDNA alone, suggesting that fluorophore-protein interaction may be affecting fluorescence stability. The (-15Cy3B, +15ATTO647N) and (-15Cy3B, -25ATTO647N) promoter construct BCA results support this hypothesis; however, more advanced experimental methods are needed to determine the underlying cause.

Based on the BCA results for RP_o dynamics, all three dsDNA promoter constructs show E-S correlations. These results suggest that we be wary in attributing inter-fluorophore distance changes as the source of E dynamics. We cannot, however, determine the exact source of E-S correlations. We are limited to a global perspective of individual molecule behavior. Analysis methods capable of de-convolving distance-independent photophysical fluctuations from legitimate E dynamics are still under active development.

6.5 Summary

Determining sources of heterogeneity remains a prominent goal within single molecule fluorescence data analysis. The ability to isolate species of interest while having legitimate physical or statistical justification to exclude other species is a significant challenge, and several published works have tried to provide solutions. The methods, however, fail in one or more respects, often due to limited applicability or difficulty in interpreting results. While many published works have provided detailed mathematical treatments [100–105, 185, 354, 355], few have provided robust frameworks which are easy to implement and intuitive to understand.

One of the strengths of BCA is its extension of BVA, in which results are, while more qualitative than quantitative, straightforward and built on statistical concepts with which many in the field are familiar. Nevertheless, BCA has several shortcomings. Notably, in its current form, BCA makes both type I and type II errors, thus creating significant difficulty in isolating uncorrelated and correlated behavior. One possible method to overcome BCA's shortcomings is *independent component analysis* (ICA), which allows us to remove components shared by E and S and to retain independent behavior. For example, suppose E fluctuations are a mixture of both inter-fluorophore distance changes and distance-independent photophysics. By removing the photophysical component of E via S, we can isolate the inter-fluorophore distance changes component. Initial efforts are underway to test whether we can use ICA as an effective filter, and results are thus far promising.

Contributions

The theory and implementation of BCA for confocal solution data is my own. I performed all simulations. Thorben Cordes and Nicole Robb conducted all RP_0 start-site selection measurements, and, on these measurements, I performed BCA. All writing and figures are my own.

Chapter 7

Mean-Variance Analysis

Abstract

Single-molecule Förster resonance energy transfer (FRET) in conjunction with total-internal reflection fluorescence (TIRF) microscopy allows simultaneous observation of multiple biomolecules with high temporal and spatial resolution. While single-molecule FRET TIRF affords extensive observation periods amenable to time series analysis, FRET histogram analysis remains prominent in determining the presence of FRET sub-populations. Recent work derived an expression for the expected shot-noise-limited distribution for single-molecule FRET TIRF [135] and thus provided a metric for distribution broadening. Here, we combine mean-variance analysis (MVA) [251] with a statistical approach used previously for diffusing molecules [314] to detect fluctuations within FRET time series. We use simulations to validate our method and show that MVA can detect FRET dynamics faster than the measurement timescale. MVA thus overcomes a barrier to probing faster kinetic regimes not directly accessible to conventional time series analysis methods. We further demonstrate how MVA provides an easy-to-interpret method for measuring inter-molecular heterogeneity, especially for short time series. We use MVA to analyze RNA polymerase (RNAP) open complex FRET TIRF distributions. Our results find a broad spectrum of RNAP behavior and support the presence of a stable intermediate on the pathway between RNAP closed and open complex. We discuss the intermediate's possible relevance within the context of transcription regulation and suggest additional avenues for testing an initial transcription branched pathway hypothesis.

7.1 Introduction

Many single-molecule fluorescence signal analysis techniques originated in previous work in ion channels. Burst variance analysis is no exception, falling under the general heading of *mean-variance analysis* (MVA). The work of Patlak [251] demonstrated MVA's potential to extract kinetics of complex single ion channel data. As shown above, MVA is performed by calculating mean and variance values within windows consisting of N consecutive data samples. Repeated window movement and calculation produces mean-variance pairs which can be assembled into two-dimensional histograms. Low variance regions are assigned to defined state values, while high variance regions represent dynamic interconversion between states.

Chapters 4-6 have thus far offered a cautionary tale. E-BVA presented a method to detect FRET fluctuations for diffusing molecules, but the method could not unequivocally assert whether these fluctuations resulted from changes in inter-fluorophore separation or distance-independent photophysics, as S-BVA and BCA analyses demonstrated. Hence, analyses employing E-BVAwithout regard for possible distance-independent photophysics could reach incorrect conclusions and ascribe biological significance to fluctuations where, in fact, such fluctuations are artifacts.

As Chapter 6 demonstrated, solution-based photon arrival data of molecules diffusing on the millisecond timescale presents difficulties for delineating distance-dependent and distance-independent FRET fluctuations. Critically, burst analysis suffers from insufficient statistics. Bursts frequently contain less than 200 photons, which, for a sliding 10-photon window, corresponds to at most 20 (E, S) data pairs. 20 data points are too few for most model-based methods, e.g., hidden Markov models (HMMs), which might allow for hypothesis testing to identify behavior deviant from a *prior* model. Further, small photon windows translate to large E/S variance. Hence, actually visualizing burst E-S time series with an aim to take advantage of human pattern recognition does not provide any benefit, as

legitimate fluctuations are easily buried in noise.

One possible approach to overcome these difficulties is to extend the diffusion timescale. Santoso and co-workers combined gel-electrophoresis with alternating-laser excitation (ALEx) to probe dynamically active biomolecules [276]. As the authors showed, in-gel ALEx can yield bursts which last longer than 10 seconds and contain over 1000 photons. Additionally, Zarrabi and co-workers used vesicle encapsulation to slow diffusion times and applied HMMs to burst time series to resolve dynamic fluctuations [360]. Such methods, however, are cumbersome (in-gel ALEx) or involved (vesicle encapsulation). Ideally, we would use other experimental methods to probe biomolecules over extended timescales with minimal perturbation.

Total-internal reflection fluorescence (TIRF) microscopy provides one such method. Importantly, TIRF microscopy allows for comparatively high photon numbers ($>$ 200 per (E, S) pair and extended observation ($>$ 20 seconds). As a consequence, fluctuations exceeding the predicted shot-noise are easily detected by the human eye. Molecules exhibiting clear S dynamics can thus be readily removed from analysis. Consequently, we can isolate molecules whose E fluctuations show no obvious relation to distance-independent photophysics. Such isolation obviates the imperative for S-BVA and BCA TIRF time series analyses, which were needed previously as visual filtering of diffusion-based FRET data was not possible.[1]

With these thoughts in mind, we might wonder if an analysis akin to E-BVA is applicable to the TIRF-FRET context. Single-molecule literature abounds with time series analyses of data containing clear step-like transitions. Arguably, signal analysis methods such as HMMs [26, 218] and change-point detection approaches [338] appear more appropriate. We note, however, that even for systems where the majority of molecules show clear state transitions, we sample only a

[1] We do note that we can still apply S-BVA and BCA to TIRF time series, but these statistical methods are not as critical when visual filtering methods provide sufficient controls. Although not discussed here, we do recommend S-BVA for TIRF time series when performing MVA to ensure that S fluctuations are in accordance with the shot-noise prediction and are not contributing to E variance. This is particularly important when trying to access dynamic timescales faster than the measurement temporal resolution.

subset of the molecular transition rates [107]. Transition rates faster or slower than the measurement timescale are not accessible. Hence, we may ask: are molecules which appear 'static' truly static or actually dynamic, rapidly transitioning between two or more states and exhibiting a single state value which is the weighted average of the transitioning states? MVA can help answer this question.

BVA's novel addition to MVA was inclusion of a theoretical model for identifying dynamics, vis-á-vis the shot noise prediction. Recent work by Holden and co-workers [134, 135, 321] investigated the theoretical limits of TIRF-based FRET and determined the expected uncertainty in TIRF-FRET measurements. Accordingly, we sought to extend E-BVA to TIRF and build on the work of Patlak [251] to identify dynamics in immobilized single-molecule time series.

7.2 Theory

We can approximate fluorescence photon counts for surface-immobilized fluorescently-labeled molecules by a Poisson process [101]. For doubly-labeled species, having a FRET pair consisting of a donor and acceptor fluorophore, the Poisson process is split into two sub-processes by a parameter process with probability p, equal to the ratio of mean acceptor to donor plus acceptor photon counts (see Appendix B.1.1). We assign p to the FRET efficiency.

We collect fluorescence emission photons from a surface-immobilized biomolecule on an electron multiplying CCD (EMCCD) camera. For FRET measurements, we spectrally separate photons and focus them onto one of two channels, the 'donor' channel or 'acceptor' channel. The photons collected on a single pixel within the integration time for a single image frame generate electrons amplified in the camera electron multiplying gain register, reducing the device read noise [139]. An analogue-to-digital converter converts the amplified electron count to digital units (DU), which form the final fluorescence emission image. For an individual chan-

nel, we relate DUs and photon count by

$$N_{ij}^{DU} = \frac{G}{S} N_{ij} \tag{7.1}$$

where N_{ij}^{DU} is the number of DU for pixel ij, N_{ij} the photon number for pixel ij, G the electron multiplying (EM) gain (electrons per photon, a stochastic variable), and S is the CCD sensitivity (electrons per DU). The CCD sensitivity is available from the camera manufacturer, and G is adjustable, being scaled by the operator, and is typically 100-1000 [267]; however, due to EM gain register aging, G frequently differs from the specified value. For photon-counting applications, the ratio G/S should be measured directly using standard methods [320]. Knowing G/S, we can obtain N_{ij}.

For completeness, we will now recap the work of Holden and co-workers [134, 135, 321], which establishes the equivalent of the solution-based 'static limit' for surface-immobilized biomolecules. We can predict the photon-counting error for a surface-immobilized molecule located at (x_0, y_0), with total expected count N_0 and independent Gaussian distributed errors, using the least-squares fitting criterion that the sum of squared errors, χ^2, is minimized:

$$\chi^2(x, y, N) = \sum_{ij} \frac{\Big(n_{ij} - N_{ij}(x, y, N)\Big)^2}{s_{ij}^2} \tag{7.2}$$

where n_{ij} is the observed photon count, $N_{ij}(\cdot)$ the expected photon count from a model point-spread function (PSF) located at (x, y) with total photon count N, and s_{ij}^2 is the observed photon count sample variance (uncertainty). For square pixels of width a, one can numerically evaluate the expected photon-counting error as

$$\langle \sigma^2(N) \rangle_{x_0, y_0} = a^{-2} \int_0^a \int_0^a \left[\sum_{i,j=-\infty}^{i,j=\infty} \left(\frac{1}{\sigma_{i,j}} \frac{\partial N_{ij}}{\partial N} \right)^2 \right]^{-1} dx_0 \, dy_0 \tag{7.3}$$

where we can use the observed variance s_{ij}^2 as an estimator for the expected vari-

ance σ_{ij}^2. While exact, Equation 7.3 is a bit unwieldy. If we neglect pixellation effects (i.e., the uncertainty arising from photon arrival location within a finite-sized pixel) and instead modify the result of Thompson and co-workers [309] (Equation 19) for the expected variance of the measured photon count on a single pixel

$$\sigma^2(N) = N + \frac{4\pi s^2}{a^2} b^2 \tag{7.4}$$

to include electron multiplication (EM) effects, we have

$$\sigma^2(N) = f_G^2 N + \frac{4\pi s^2}{a^2} b^2 \tag{7.5}$$

where b is the background count (photons per pixel), including measurement background and on-camera noise such as read and dark noise, and f_G is the excess noise factor arising from EM.[2] Using error propagation [191], where E equals[3]

$$E = \frac{F_{D_{exc}}^{A_{em}}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}}}$$

the apparent FRET uncertainty is

$$\begin{aligned} \mathrm{Var}(E) &\approx \left[\frac{\partial E}{\partial F_{D_{exc}}^{D_{em}}}\bigg|_{\langle E \rangle} \right]^2 \mathrm{Var}(F_{D_{exc}}^{D_{em}}) + \left[\frac{\partial E}{\partial F_{D_{exc}}^{A_{em}}}\bigg|_{\langle E \rangle} \right]^2 \mathrm{Var}(F_{D_{exc}}^{A_{em}}) \\ &\approx \frac{1}{(F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}})^2} \left(\langle E \rangle^2 \,\mathrm{Var}(F_{D_{exc}}^{D_{em}}) + (1 - \langle E \rangle)^2 \,\mathrm{Var}(F_{D_{exc}}^{A_{em}}) \right) \end{aligned} \tag{7.6}$$

where $\langle E \rangle$ is the FRET value. We can obtain a full prediction $\sigma(E)$ by numerically integrating Equation 7.3 for each channel and substituting in Equation 7.6. But if we are content neglecting pixellation and interpolating between high background and high shot noise limiting cases as in Equation 7.5 [134,309], we arrive at a more

[2] b is empirically modeled. Rigorous assessment would require de-convolving individual noise sources and determining what is amplified, with appropriate modeling. Here, we take a phenomenological approach based on the observed background, assigning this value units of photons per pixel.

[3] Note: the fluorescence emission variables are similar to those of solution-based confocal measurements, except the intensity is not a function of a particle's path through a confocal volume, but a function of the evanescent electric field and camera integration time.

convenient approximation[4] using $\sigma^2(F_{D_{exc}}^{A_{em}})$ and $\sigma^2(F_{D_{exc}}^{D_{em}})$ approximations

$$\mathrm{Var}(E) \approx \frac{f_G^2\langle E\rangle(1-\langle E\rangle)}{F_{D_{exc}}^{D_{em}}+F_{D_{exc}}^{A_{em}}} + \frac{4\pi}{a^2(F_{D_{exc}}^{D_{em}}+F_{D_{exc}}^{A_{em}})^4}\left((F_{D_{exc}}^{A_{em}})^2 s_D^2 b_D^2 + (F_{D_{exc}}^{D_{em}})^2 s_A^2 b_A^2\right) \tag{7.7}$$

where s_D and s_A are the PSF widths (in nanometers) in the donor and acceptor channels, respectively; b_D and b_A are the background noise standard deviations (photons per pixel) for each channel; a is the pixel size (in nanometers); f_G is the excess noise factor introduced by electron multiplication ($f_G \approx \sqrt{2}$) [267]. For our experimental setup, $s_D = 132$nm, $s_A = 150$nm, $a = 94$nm, $b_D = b_A = 2.9$ photons per pixel, and measured G/S ratio of ~ 6.

Note that Equation 7.7 is an approximation. Numerically integrating Equation 7.3, substituting into Equation 7.6, and comparing to Equation 7.7 reveals that Equation 7.7 fails to fully account for observed fluctuations, underestimating the standard deviation by $\sim 6\%$ for moderate photon counts ($N = 390$ photons) [134, 135]. The effect is, however, consistent and reproducible, scaling linearly with photon counts and converging to 0 as $N \to \infty$. Hence, in the context of MVA, as long as we remember this offset, we can determine whether observations exceed the static limit.

Depending on the PSF profile-fitting technique, additional uncertainty may arise affecting the ability of Equation 7.7 to account for observed dispersion. Different techniques, such as aperture photometry [138, 150], ordinary least-squares (OLS) minimization [65, 297], weighted least-squares (WLS) minimization [309], and maximum likelihood methods [227], introduce different additional excess uncertainty. Holden and co-workers previously implemented OLS for image extraction and data reduction and found OLS introduced a constant excess uncertainty of $\sim 30\%$ beyond Equation 7.7 [135, 227, 309]. As OLS remains the implemented

[4]The following equation differs from Equation 6 in [134, 135], where $F_{D_{exc}}^{A_{em}}$ and $F_{D_{exc}}^{D_{em}}$ are switched. My equation is correct, as the reader may confirm. I have communicated this error, amongst others (Equations 4, 5, 7, S7, S15, and S22 in [135]), to Holden and co-workers (Stephan Uphoff, personal communication, 2012).

method in our laboratory, we calculate the apparent uncertainty as

$$\sigma_{app}(E) = 1.30\sqrt{\mathrm{Var}(E)} \tag{7.8}$$

Hence, within the context of MVA, in order for experimental time series fluctuations to be considered 'dynamic', they would need to exceed the predicted limit of Equation 7.8.[5]

By propagating errors, we can derive an expression similar to Equation 7.7 for the stoichiometry S. Let $N = F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}} + F_{A_{exc}}^{A_{em}}$.

$$\mathrm{Var}(S) \approx \left[\left.\frac{\partial S}{\partial F_{D_{exc}}^{D_{em}}}\right|_{\langle S\rangle}\right]^2 \mathrm{Var}(F_{D_{exc}}^{D_{em}}) + \left[\left.\frac{\partial S}{\partial F_{D_{exc}}^{A_{em}}}\right|_{\langle S\rangle}\right]^2 \mathrm{Var}(F_{D_{exc}}^{A_{em}}) + \left[\left.\frac{\partial S}{\partial F_{A_{exc}}^{A_{em}}}\right|_{\langle S\rangle}\right]^2 \mathrm{Var}(F_{A_{exc}}^{A_{em}}) \tag{7.9}$$

where we defined S in Equation 5.2.

$$\mathrm{Var}(S) \approx \frac{1}{N^2}\left((1-\langle S\rangle)^2\,\mathrm{Var}(F_{D_{exc}}^{D_{em}}) + (1-\langle S\rangle)^2\,\mathrm{Var}(F_{D_{exc}}^{A_{em}}) + \langle S\rangle^2\,\mathrm{Var}(F_{A_{exc}}^{A_{em}})\right) \tag{7.10}$$

Substituting Equation 7.5 gives the approximate stoichiometry measurement uncertainty

$$\mathrm{Var}(S) \approx \frac{f_G^2\langle S\rangle(1-\langle S\rangle)}{N} + \frac{4\pi}{a^2N^4}\left((F_{A_{exc}}^{A_{em}})^2(s_D^2b_D^2 + s_A^2b_A^2) + (F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{A_{em}})^2 s_A^2 b_A^2\right) \tag{7.11}$$

Now having expressions for measurement uncertainty, we can now turn to MVA.

MVA is similar to BVA, as outlined above, except we now follow Patlak [251] in using overlapping sliding windows.[6] The mean-variance calculations are straightforward. Let $X = \{X_0, X_1, \ldots, X_{T-1}\}$ be a time series defined by the sequence of observations X_t. For each X_t, we calculate the mean μ_t and sample variance s_t^2

[5] For a different fitting technique, Equation 7.8 would need to be modified accordingly. Recent work [227] suggests that, for maximum likelihood profile fitting, $\sigma_{app}(E) = 1.0\sqrt{\mathrm{Var}(E)}$, meaning the fitting technique does not introduce any excess heterogeneity.

[6] We will address later in this chapter why overlapping windows are advantageous for single-molecule time series.

within a window including that observation and the following $(W-1)$ observations, where W is the observation number defining the sliding window width. The mean and sample variance are defined as

$$\mu_t = \frac{1}{W} \sum_{i=t}^{t+W-1} x_i \tag{7.12}$$

$$s_t^2 = \frac{1}{W-1} \sum_{i=t}^{t+W-1} (x_i - \mu_t)^2 \tag{7.13}$$

Repeated computation produces $(T-W+1)$ mean-variance pairs (μ_t, s_t^2), where T is the number of observations. Figure 7.1 provides a schematic illustrating MVA for an ideal time series.

To obtain a sense of timescales for dynamics, we construct MVA histograms for multiple window widths, e.g., $W = \{3, 5, 10, 15, 20, 30, 50, 100\}$. To minimize the number of arithmetic operations needed to analyze data segments, we use a simple trick by maintaining a rolling array. For each W, build a vector of W observations. Compute the sum and the square of the sum for all window observations, and place the paired values in a separate two-column array $\mathbf{m}_t = (\mu_t, s_t^2)$. Slide the window one data point. Add the new observation and its square to the previous values and subtract the observation immediately prior to the current window and its square, placing the new result on the next row of the two-column array $\mathbf{m}_{t+1} = (\mu_{t+1}, s_{t+1}^2)$.

In contrast to Figure 7.1, data in the wild is never so ideal. Observations will vary about their true mean, and alas, the variance will never be zero. Furthermore, as observations within a window W constitute a small sample of the parent population, these observations will be dispersed about the population mean, even when a window is centered over a set of observations that shows relative stability. Assuming Gaussian noise, the distribution of mean estimates μ_t will also be Gaussian, with variance

$$\sigma_\mu^2 = \frac{\sigma_X^2}{W} \tag{7.14}$$

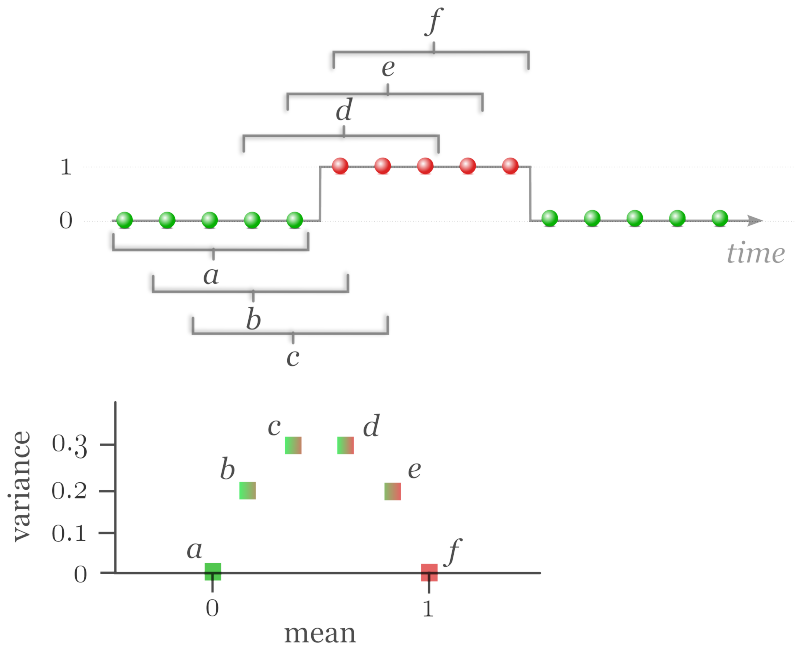


Figure 7.1: Mean-Variance Analysis (MVA) schematic. For an ideal time series (no background, instantaneous transitions) consisting of two states, represented here by green (state 0) and red (state 1) circles, the variance is maximum when the windows (labeled brackets) are centered about a transition point. Hence, windows c and d show the highest variance; whereas, for windows centered over data of constant value (windows a and f), the variance is zero and the mean value is the true state mean. The letters in the lower plot illustrating the mean-variance relation correspond to the letters labeling the windows in the time series above.

where σ_X^2 is the population variance. The sample variance estimates, s_t^2, will have a χ^2 distribution with $\nu = W - 1$ degrees of freedom [211], such that

$$\frac{s_t^2}{\sigma_X^2} = \frac{\chi_\nu^2}{\nu} \tag{7.15}$$

In a two-dimensional mean-variance histogram, low-variance regions will correspond to observations that characterize individual system states. By fitting these regions while keeping Equations 7.14 and 7.15 in mind, we can determine the population mean and variance for those states. Additionally, the region volume will correspond to the total number of observations during which the entire window is completely centered about that region, and hence, the volume will be a function of (a) window width, (b) the frequency at which a state is populated, and (c) state dwell time distributions.

For Markovian kinetics, in which dwell times obey exponential distributions,

Patlak [251] provides a rubric for extracting rate constants and fitting kinetic models. We will not concern ourselves here with pursuing this approach, but mention in passing this extension for possible further development.[7]

7.3 Simulation Results

We performed simulations to test and validate MVA for surface immobilized single-molecule time series. In contrast to single-molecule confocal spectroscopy measurements where the output data from avalanche photo-diodes (APDs) directly corresponds, more or less, to detected photons, single-molecule TIRF data is the result of a multi-step process. Detected photons are first convolved with camera noise (read and dark noise) and background noise, then amplified, and finally converted to digital units. In order to first understand how raw emitted photon counts are 'corrupted', we simulated each step of the camera data acquisition process.

7.3.1 Raw Photon Counts

We first modeled a physical system as a single-state ($E = 0.3$) Markov chain with Poisson emissions using kinetic Monte Carlo (KMC) techniques and realistic photon counts; $N = 250$ photons per camera integration period (Figure 7.2(a)). We then simulated background noise in accordance with Equation 7.4 (term 2) by adding Poisson noise approximated by a Gaussian distribution $\mathcal{N}(\mu = \lambda, \sigma^2 = \lambda)$, with $s_A = 150$nm, $s_D = 132$nm, and $a = 94$nm. We used the experimentally determined value $b = 2.9$ photons per pixel to account for scattering, incomplete filtering, thermal noise, and clock-induced charge (Figure 7.2(b)). While background increases channel signal variance, σ_E^2 decreases due to increased photon statistics, in accordance with Equation 5.7. We note, however, that background contribution is proportional to PSF area which is proportional to wavelength, and, as acceptor emission is of longer wavelength, background is relatively greater in the acceptor channel. This artificially shifts $E \approx 0.4$.

[7] Note that a similar extension could be made for BVA.

7.3.2 Electron Multiplication

We next explicitly simulate electron multiplication for a gain register with $n = 512$ elements and total effective gain[8] $U = 6$ (Figure 7.2(c)). The probability for electron multiplication in each stage is

$$p_e = U^{\frac{1}{n}} - 1 \tag{7.16}$$

Whether an electron is multiplied, i.e., doubled[9], is determined by drawing a uniform random number $r \in [0, 1]$ for each element in the register. If $r < p_e$, then an electron is added. The output signal has a U-fold higher value, featuring multiplicative noise which increases the signal variance by a factor of 2 ($= f_G^2$) [267].

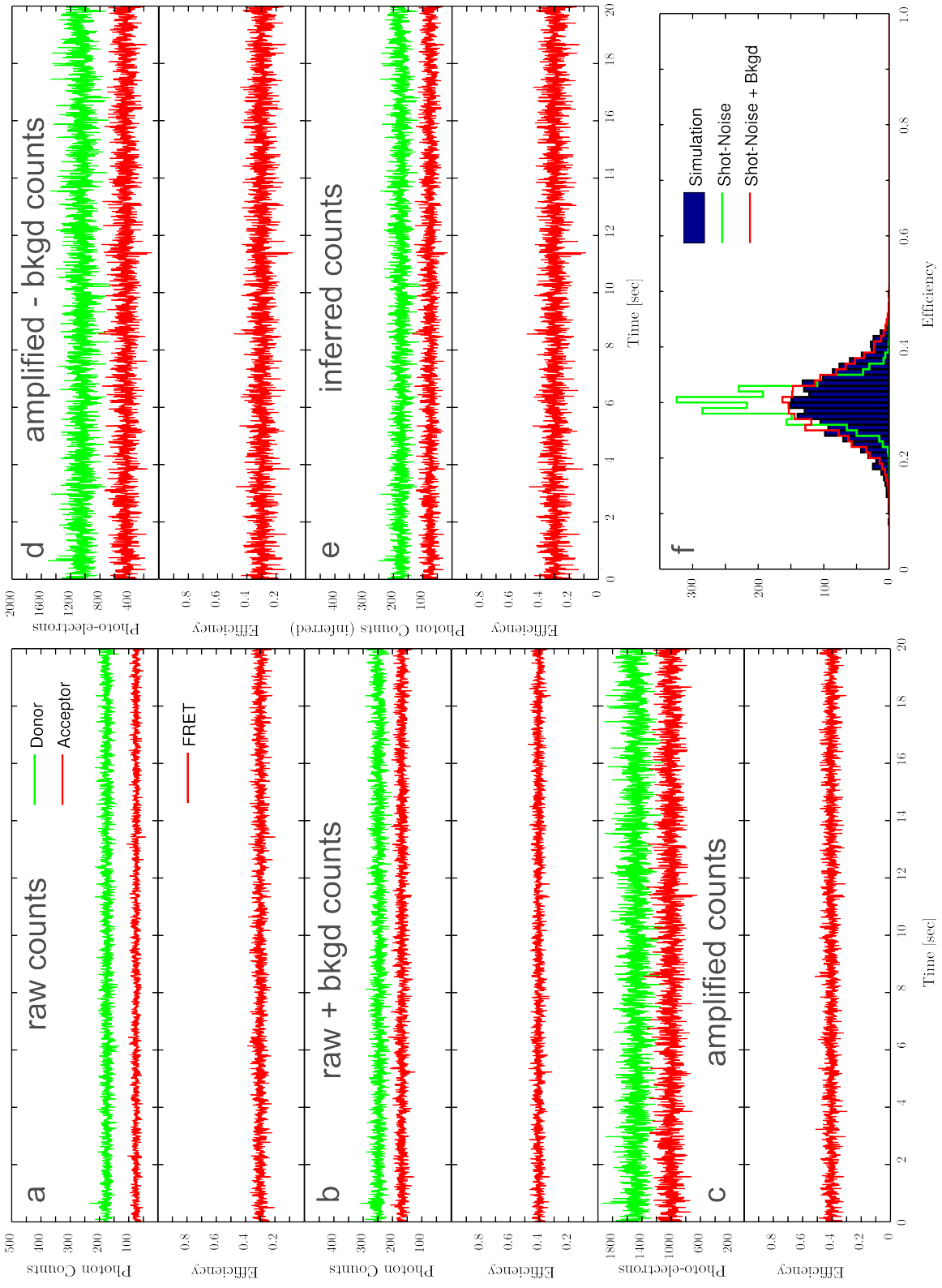
7.3.3 Signal Recovery

Standard fitting techniques utilize background subtraction to recover the 'true' underlying signal. Once again, we use the second term in Equation 7.4, but this

[8] Measured by Stephan Uphoff, 2012.

[9] We assume that an electron is not multiplied more than once within a gain register element.

Figure 7.2 ***(following page)*****:** Shot Noise in Total Internal Reflection Fluorescence (TIRF) microscopy. An electron multiplying CCD (EMCCD) collects fluorescence emission photons from surface-immobilized biomolecules. Collected photons are convoluted with camera noise. Electron multiplication of detected photons (raw emission + background) results in an amplified signal, reducing camera read noise. An analogue-to-digital converter converts photo-electrons to digital units, the final format given to the end user. (a) raw photon counts. Poisson statistics for both donor (green) and acceptor (red) result in raw shot noise, with variance in either channel equal to the mean channel signal. (b) addition of camera noise. Poisson noise increases channel variance, but decreases FRET variance (due to increased photon statistics). Note that the background contribution is proportional to the point spread function (PSF) area; in which case, as PSF width is proportional to wavelength ($\lambda/2$), the acceptor channel is affected more than the donor channel. This results in FRET being artificially increased. (c) amplified counts (photo-electrons). Photons incident on the camera undergo electron multiplication (EM), dependent on gain factor G. EM introduces an additional noise factor ($f_G \approx \sqrt{2}$) in both channels. (d) background subtraction from amplified signal. To minimize the effect of background and in an attempt to recover the 'true' signal, we subtract the mean background level from the amplified counts. The decrease in counts results in increased FRET variance, but background subtraction recovers the true FRET value. (e) By determining the effective camera gain (Equation 7.1), we can recover photon counts. Note that channel noise remains greater than that shown in (b), with FRET variance still significantly greater than in (a), despite similar photon statistics. (f) shot noise histogram. The simulation (blue) is significantly broader than the predicted shot noise distribution in the absence of background (green). Inclusion of background generates a shot noise distribution (red) which nearly accounts for the simulation. The additional $\sim 10\%$ difference ($\sigma_{obs} = 0.0563$ vs $\sigma_{SN} = 0.0509$) is due to Equation 7.7 being an approximation, interpolating between high shot noise and high background limits and failing to fully account for the region between. The observed difference is in general agreement with [134, 135]; here, the simulated photon count is $N = 250$ with $E = 0.3$.


a
raw counts
Donor
Acceptor
FRET
b
raw + bkgd counts
c
amplified counts
d
amplified - bkgd counts
e
inferred counts
f
Simulation
Shot-Noise
Shot-Noise + Bkgd
Photon Counts
Photo-electrons
Photon Counts (inferred)
Efficiency
Time [sec]
Efficiency

time we multiply by U. Background subtraction recovers the original E value of 0.3, but increases the FRET variance due to decreased photo-electron count (Figure 7.2d).

To recover photon counts, we divide the background-subtracted, amplified counts by U (Figure 7.2(e)). Channel noise is significantly greater than in Figure 7.2b because of the excess noise introduced by multiplication, while FRET variance is significantly greater than in Figure 7.2(a) despite having similar photon statistics.

7.3.4 Shot Noise Prediction

To compare theory and simulation, we generated the predicted shot noise distribution using a procedure adapted from Nir and co-workers [239]. For the case of no background, for each image frame t

1. Set N to the photon number of t, and draw a binomially distributed random number A with parameters N and E.

2. Compute A/N and store in a vector $\mathbf{v}$.

3. Repeat Steps 1 and 2 m times, where m is an oversampling factor accounting for insufficient sampling of photon numbers in the distribution tails.

4. Histogram the vector $\mathbf{v}$ and divide the counts by m to obtain the predicted shot noise distribution.

Figure 7.2(f) shows the background-free shot noise prediction in green. The predicted distribution fails to account for the observed simulated distribution width. To include background and multiplicative noise, we modify the above algorithm as follows:

1. For each image frame t, generate Poisson distributed noise for both channels (B_D, B_A) according to the second term in Equation 7.5, where we approximate the noise as Gaussian $B \sim \mathcal{N}(\mu = \lambda, \sigma^2 = \lambda)$, with $\lambda = \frac{4\pi s^2}{a^2} b^2$.

2. Set $N = N_0 - B_D - B_A$, where N_0 is the total photon count for frame t, and draw a binomially distributed random number A with parameters N and E.[10]

3. Compute A/N and store in vector $\mathbf{v}$.

4. Repeat Steps 1-3 m times, where m is an oversampling factor accounting for insufficient sampling of photon numbers.

5. Histogram the vector $\mathbf{v}$ and divide the counts by m to obtain the predicted shot noise distribution.

Figure 7.2(f) also shows the shot noise prediction including background in red. The predicted distribution nearly accounts for the observed simulated distribution width; $\sigma_{obs} = 0.0563$ versus $\sigma_{SN} = 0.0509$. The approximately 10% difference is due to Equation 7.7 being an approximation, which interpolates between high shot noise and high background limits and neglects pixellation effects. The observed difference is consistent for all values of E and in general agreement with [134, 135]; here, the simulated photon count is $N = 250$ with $E = 0.3$. Consequently, by adding a 10% correction offset to Equation 7.7, any distribution broadening beyond the (offset) predicted limit is indicative of additional heterogeneity that is not attributable to photon statistics or data acquisition alone.

7.3.5 Validating the Predicted Shot Noise Limit

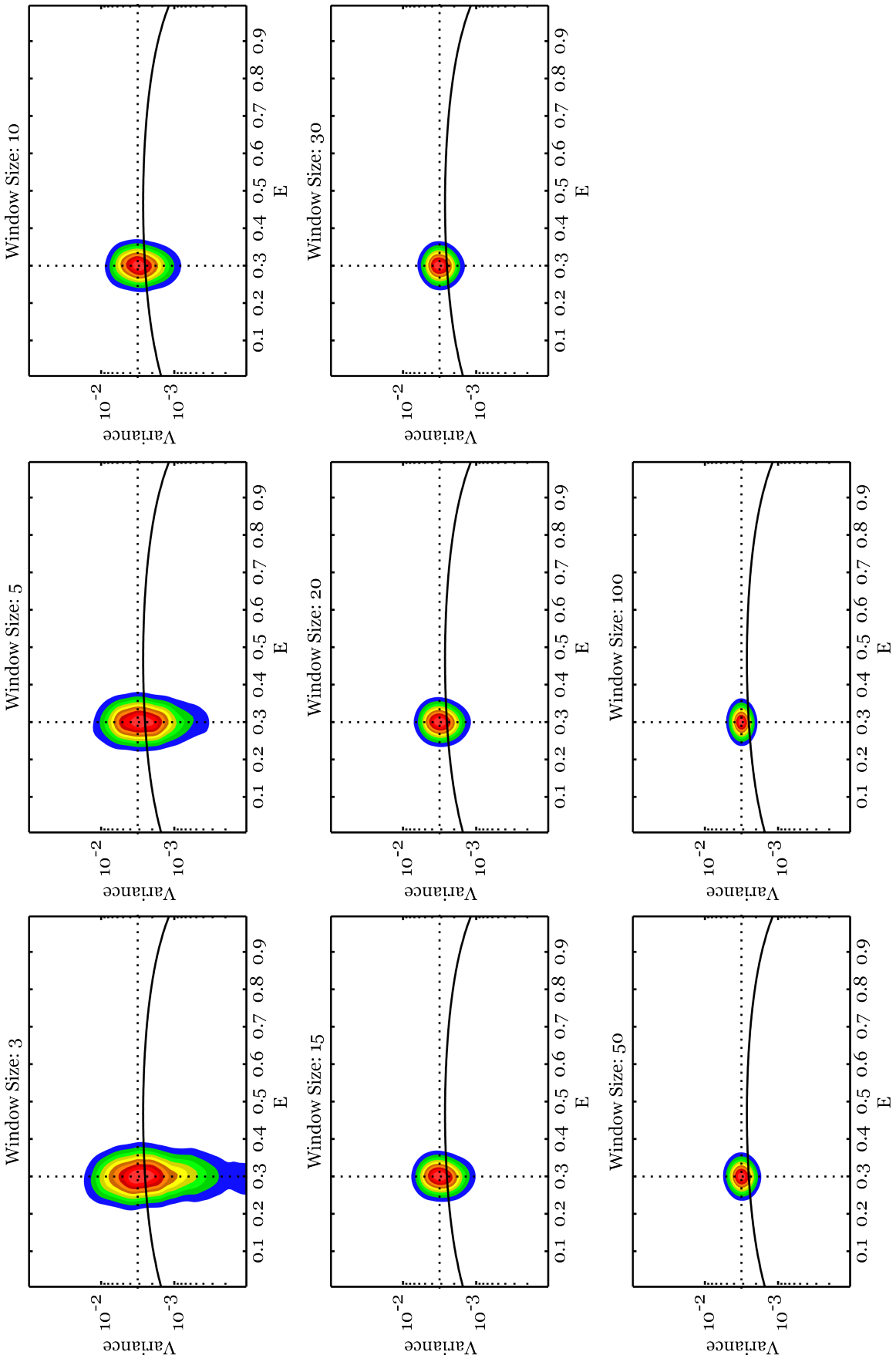
Having determined the shot noise limit, we sought to validate Equation 7.7 for MVA. We simulated a single-state Markov chain of length $T = 2000$ with Poisson emissions, mean photon count $N = 250$, and FRET efficiency $E = 0.3$. We performed MVA for several sliding window sizes $\{W : 3, 5, 10, 15, 20, 30, 50, 100\}$. For each observation x_t, we computed the sample mean μ_t and variance s_t^2 within a window W, before sliding W to the next data point x_{t+1}. We then bin $T - W + 1$ mean-variance pairs (μ_t, s_t^2), building a histogram where contour maps indicate

[10] By reducing the photon count, we increase σ_E.

relative density (Figure 7.3). By comparing the density to the shot noise prediction, which, in the figure, is a solid black line, we can distinguish between *static* (non-interconverting species) and *dynamic* (interconverting) *heterogeneity*.

We note a few key MVA features. First, smaller windows W demonstrate greater spread in variance because small windows are poor estimators of population variance. As W increases, sample variance estimates become more reliable population variance estimators. Second, for static species, the density is symmetric about the mean along the x-axis, with density width not changing for increasing W. And third, static species are located above the predicted shot noise curve. The offset is approximately 10%, as observed in Figure 7.2. Hence, by shifting the predicted curve by a 10% correction factor, we validate Equation 7.7 as a reliable predictor for static species in MVA. The dotted lines indicate the population mean and variance, and their intersection should coincide with the density center for static species. For multiple static species that have different mean state values within the same sample, each species will be scattered along the predicted curve in accordance with the static limit. For dynamic inter-converting species, we expect the variance to be greater than the shot noise prediction, with the density center located above the black curve.

Figure 7.3 *(following page)*: Mean-Variance Analysis: Simulation (Control). To demonstrate the concept of mean-variance analysis (MVA), we simulated a time series of length $T = 2000$, with FRET efficiency $E = 0.3$ and mean photon count $N = 250$. We generate mean-variance histograms for several sliding window sizes $\{W : 3, 5, 10, 15, 20, 30, 50, 100\}$. For each observation, x_t, we compute the sample mean μ_t and variance s_t^2 within a window including the observation x_t and the next $W - 1$ observations, thus generating $T - W + 1$ mean-variance pairs (μ_t, s_t^2). We then plot each pair in the histogram, using contour maps to indicate the relative density. The solid black line represents the predicted shot noise curve, while dotted black lines correspond to the population mean and variance. If the dotted lines cross in the distribution center, the distribution does not have bias, indicating the absence of dynamics. As may be observed, the distribution center is located above the shot noise prediction, but the discrepancy is attributed to Equation 7.7 being an approximation. By taking into account the $\sim 10\%$ offset to account for pixellation and interpolation effects, we find Equation 7.7 a suitable approximation for single-molecule TIRF shot noise.


Window Size: 3
Window Size: 5
Window Size: 10
Window Size: 15
Window Size: 20
Window Size: 30
Window Size: 50
Window Size: 100
Variance
E
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
10^-2
10^-3

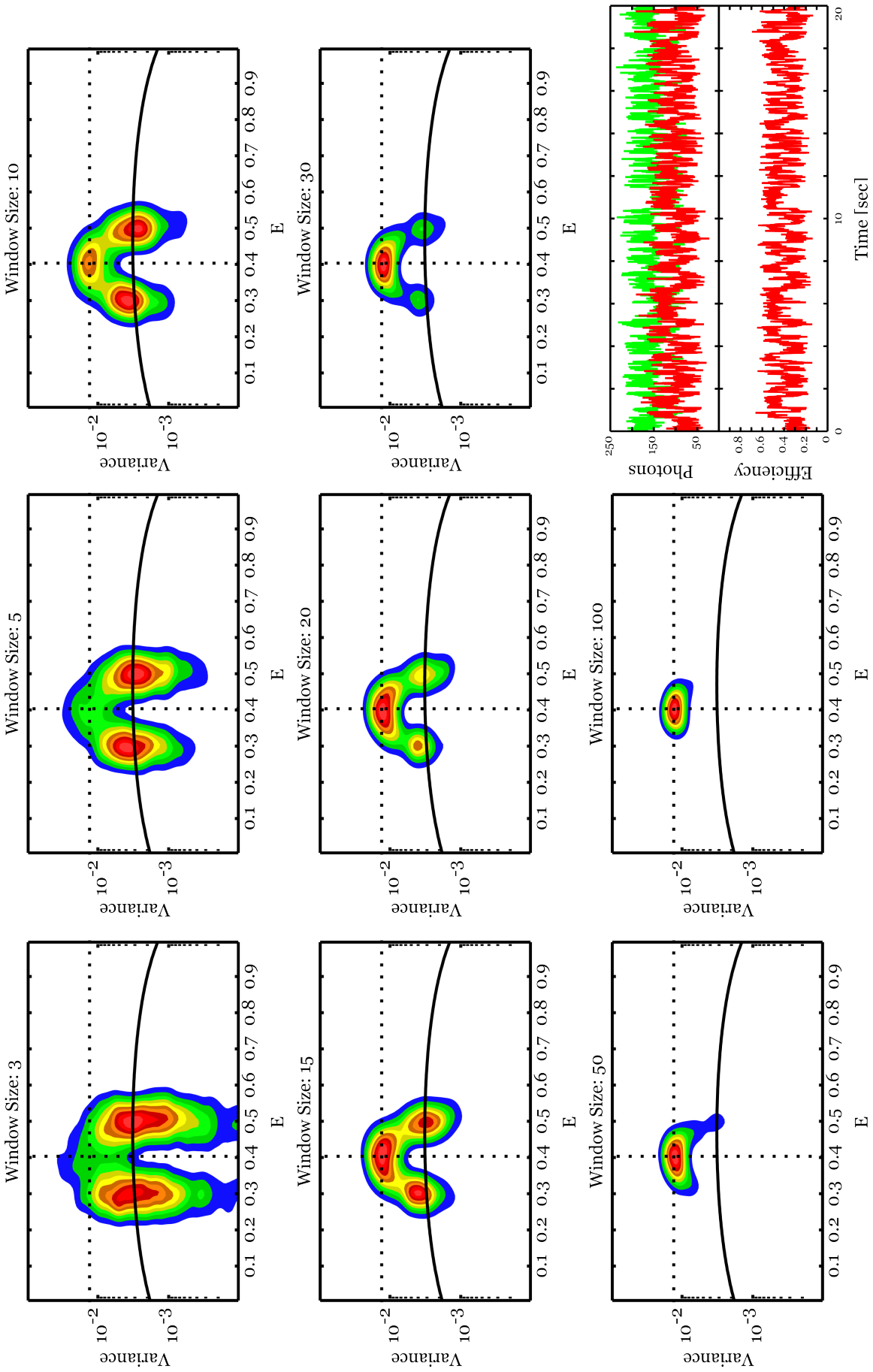
7.3.6 MVA Detects Clear Dynamics

Next, we simulated a two-state Markov chain of length $T = 2000$ with Poisson emissions, mean photon counts $N = 250$, and FRET efficiencies $\{E : 0.3, 0.5\}$. To provide a reference for detecting dynamics using MVA, we required that the system have clear dynamics, where transitions occurred every 20 data points on average for a total of about 100 transitions (see Figure 7.4). The histograms are clearly distinguished from those in Figure 7.3, with a high variance band connecting the two inter-converting states, the band attributable to windows including time series change-points.[11] Accordingly, the variance is maximal when high and low FRET states contribute equally. As W increases, windows are more likely to contain transitions, thus having higher variance. Increasing W has two further consequences: (i) as W increases, low variance populations are depopulated, and (ii) distribution centers converge to a single population located at the weighted average of the individual population means (e.g., here, $(0.5)(0.3) + (0.5)(0.5) = 0.4$). By observing how the MVA histogram changes as W increases, one obtains an approximate timescale for dynamics.

While Figure 7.4 is useful for obtaining a baseline for dynamics, other techniques, such as change-point analysis [338, 348] and hidden Markov modeling (HMM) [26, 218], are perhaps better suited for time series showing such clear dynamics. Where other techniques fail, however, is in detecting fast dynamics

[11] All MVA histograms including shot noise predictions from here forward include the 10% offset. Notice that the low-variance regions are centered about the shot noise prediction, suggesting the appropriateness of the offset factor.

Figure 7.4 ***(following page)*****:** Mean-Variance Analysis: Simulation (Dynamics). To demonstrate the concept of mean-variance analysis (MVA) for detecting dynamics, we simulated a Poisson Markov chain of length $T = 2000$, with state means $\{E : 0.3, 0.5\}$, mean dwell times $\{\tau : 20, 20\}$ (data points), and average photon count $N = 250$. The spread connecting the two main populations is indicative of dynamics. For mean dwell times greater than the window size ($\tau \gg W$), low-variance regions in the mean-variance histogram are heavily populated by component species. As W increases, windows are more likely to contain a transition, and the histogram becomes more populated in high-variance regions connecting the species. For $W \gg \tau$, most windows will contain at least one transition and thus have higher than predicted variance. The bottom right plot shows the simulated donor (green) and acceptor (acceptor) photon emissions, as well as the calculated FRET efficiency. Histogram: solid black lines correspond to the offset shot noise prediction (shifted up by 10%; see main text); dotted black lines correspond to global mean and variance.


Window Size: 10
Window Size: 30
Window Size: 5
Window Size: 20
Window Size: 100
Window Size: 3
Window Size: 15
Window Size: 50
Variance
E
Photons
Efficiency
Time [sec]

at, or just below, the observation timescale. While rapidly transitioning systems will not provide easily identifiable change-points, instead showing a continuum of intermediate values, the distribution width should increase, especially for well-separated states.

7.3.7 MVA Detects Fast Dynamics

To test the ability of MVA to detect fast dynamics between closely spaced states, we simulated a two-state Markov chain of length $T = 2000$ with Poisson emissions, $N = 250$, and mean FRET states $\{E : 0.3, 0.4\}$. The states inter-converted on average twice for each data point. Figure 7.5 shows a simulated time series, where the green line is the underlying 'true' state trajectory. The black line shows that even powerful techniques such as HMMs are unable to accurately analyze the simulated time series. We performed maximum likelihood HMM for 1 to 3 states and used the Bayesian information criterion (BIC) to determine the best model. BIC found the best data fit to be a single state HMM.

We generated the predicted shot noise distribution, which failed to fully account for the simulated distribution width (Figure 7.5(b)). We then performed MVA. Even after including the 10% offset, the population density remains significantly higher than the predicted shot noise curve across all window sizes (see Figure 7.5(c); showing $\{W : 3, 20\}$). This indicates the presence of dynamic interconversion. Note that a single distribution center is located about $E \approx 0.35$, informing us that the inter-conversion rate is faster than once every 3 frames. Additionally, the distribution is symmetric (see $W = 20$), which indicates equal state occupancy. We might predict an asymmetry would indicate disproportionate time spent in one or the other of the two states.

7.3.8 MVA Detects Transient Dynamics

Figure 7.5 assumed equal rate constants. We relaxed this assumption, hoping to test the sensitivity of MVA in detecting transient dynamics. We simulated a Pois-

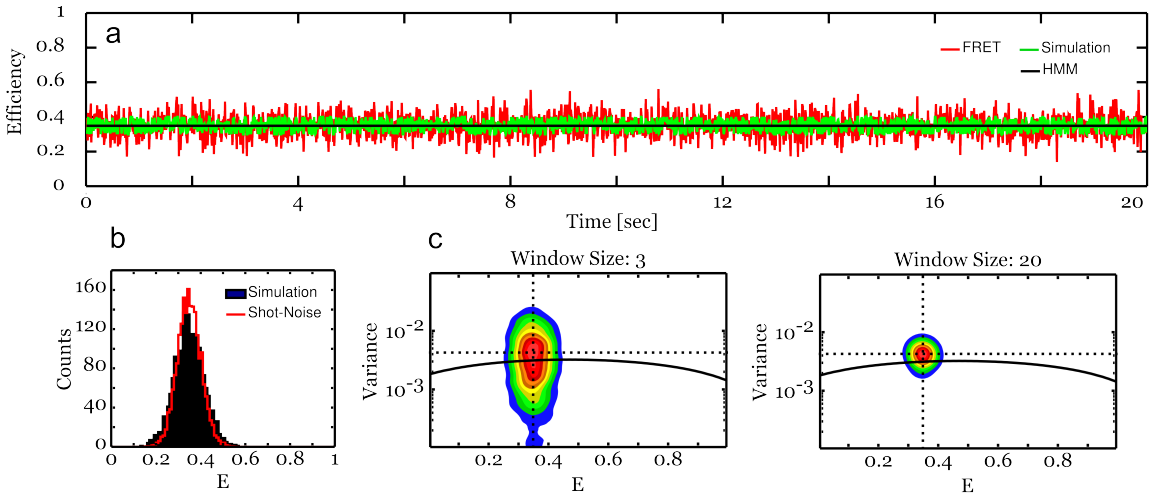


Figure 7.5: Mean-Variance Analysis: Simulation (Dynamics). To test the ability of mean-variance analysis (MVA) to detect fast dynamics between closely spaced states, we simulated a Poisson Markov chain of length $T = 2000$, with state means $\{E : 0.3, 0.4\}$, mean dwells times $\{\tau : 0.5, 0.5\}$ (data points), and average photon count $N = 250$. (a) simulated time series. The underlying 'true' trajectory (green) rapidly inter-converts between states, often multiple times within the camera integration period resulting in intermediate FRET values. Hidden Markov modeling (HMM) failed to detect two states (BIC criterion used for model selection: 1 to 3 states). (b) simulation histogram. The predicted shot noise distribution fails to account for the simulated distribution due to the simulation's fatter tails. (c) mean-variance histograms for $\{W : 3, 20\}$. Even after offsetting the shot noise prediction by 10%, the simulation distribution remains significantly higher than predicted across window sizes, indicating the presence of dynamic interconversion. (Compare to Figure 7.6.)

son Markov chain of length $T = 2000$, having state means $\{E : 0.3, 0.5\}$, $N = 250$, and two distinct kinetic states: one transitioning once every 100 data points and the other transitioning within one data point. The green line shows the simulated 'true' trajectory, and the black line the HMM fit. Once again, the HMM finds only a single state. The predicted shot noise distribution (see Figure 7.6(b)) fails to account for the simulated distribution, which is skewed toward higher E values.

In contrast to Figure 7.5, the main density is centered about the shot noise prediction. This result makes sense because the system spends most of its time in a shot-noise-limited state. We note, however, that the density is asymmetric, indicated by the pink arrow (see Figure 7.6(c)); the density shoulder moves up-and-to-the-right. We previously saw a similar, but more dramatic, shoulder in Figure 7.4 for clear dynamics. Here, because the high FRET state is only sparsely populated, we only see the transition, but not the occupancy. Depending on the dynamic timescale, the appearance and disappearance of a shoulder across W gives a clue as to the frequency at which such excursions occur. While not quantitative, MVA

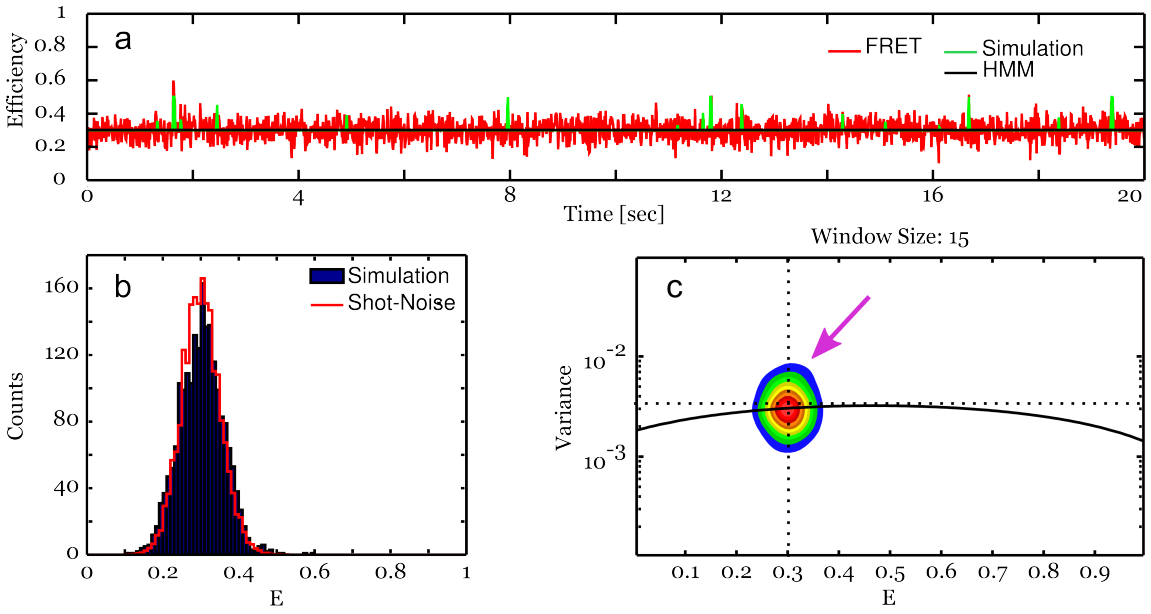


Figure 7.6: Mean-Variance Analysis: Simulation (Dynamics). To test the sensitivity of mean-variance analysis (MVA) for dynamics, we simulated a Poisson Markov chain of length $T = 2000$, with state means $\{E : 0.3, 0.5\}$, mean dwell times $\{\tau : 100, 1\}$ (data points), and average photon count $N = 250$. (a) simulated time series. The underlying 'true' trajectory (green) only sparsely populates the short-lived high FRET state. Hidden Markov modeling (HMM) failed to detect the second state (BIC criterion used for model selection: 1 to 3 states). (b) simulation histogram. The predicted shot noise distribution closely matches the simulated distribution except for the slight skewness toward higher FRET values. (c) mean-variance histogram. In the absence of dynamics (Figure 7.3), for $W > 5$, an observed distribution is symmetric about its major and minor axes; however, in the presence of dynamics, asymmetries arise. The pink arrow indicates such an asymmetry, which is reproducible for multiple simulations having the same parameters. The appearance and disappearance of the asymmetry across window sizes tells something about the timescale of dynamics.

can at least provide a qualitative hint of transient state dynamics and motivate higher temporal resolution experiments.

7.3.9 Multi-state Systems

Next, we investigated MVA performance for multi-state systems. We simulated 3-state Poisson Markov chains of length $T = 2000$, with different kinetic topologies and average photon count $N = 250$. Initially, we considered a model where two of the three states rapidly inter-convert.

$$\mathrm{A} \underset{k_{10}}{\overset{k_{01}}{\rightleftharpoons}} \mathrm{B} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \mathrm{C}$$

States A and B are closely separated states, $E = 0.3$ and $E = 0.36$, respectively, and

$k_{01} = k_{10} = 500\text{s}^{-1}$. State A cannot directly reach state C ($E = 0.5$), but must first pass through B, and the inter-conversion between B and C is more than two orders of magnitude slower than from A to B, $k_{12} = 5\text{s}^{-1};\ k_{21} = 10\text{s}^{-1}$. Investigating kinetic schemes with vastly different rates is difficult, as one typically optimizes temporal resolution for rates residing within an order of magnitude of the resolution. For example, if inter-conversions occur on average once every 50ms, then 5ms is a reasonable temporal resolution. But for the case with multiple and vastly different kinetic regimes, any single temporal resolution is sub-optimal.

With MVA, however, we can optimize for one kinetic regime to achieve better signal-to-noise under reduced laser powers by integrating over a longer data acquisition period, and yet, we can still infer the existence of faster regimes. Figure 7.7(a) provides an example. Rapid inter-conversion results in a low-FRET region located above the shot noise prediction, indicating dynamics, while the slower excursions from B to C result in the blue spread which connects the FRET states. We observe that an HMM, which, in the figure, is the black line, is unable to resolve the third state. The HMM fits a low-FRET state which is the average of A and B. While a fit of the dwell time histogram in the HMM's 'lower'-FRET state may show deviation from mono-exponential kinetics, indicating dynamic disorder, MVA provides additional visual confirmation of multiple states.

But now consider the model

$$\text{A} \underset{k_{20}}{\overset{k_{02}}{\rightleftharpoons}} \text{C} \underset{k_{12}}{\overset{k_{21}}{\rightleftharpoons}} \text{B}$$

where states $A, B,$ and C have the same mean values as above, but where the topology is such that the high FRET state is the 'intermediate' and the lower FRET states have similar kinetics. Such a topology may model a process which fluctuates between one of two or more reactive states. The MVA histograms in Figure 7.7(b) do not support an inference of inter-conversion between A and B, as the low-FRET region is centered on the shot noise prediction. The histograms provide additional support that the low-FRET dynamics in Figure 7.7(a) are 'real'. Once again, the

HMM fails to resolve the third state. The HMM fits a low-FRET state as the average of A and B. Inferring dynamic disorder for the HMM's low-FRET state may prove difficult, and MVA does not provide clear cut evidence for multiple low-FRET states; however, the low-FRET region may possibly exhibit a greater than expected dispersion in E for the low E low-variance density.

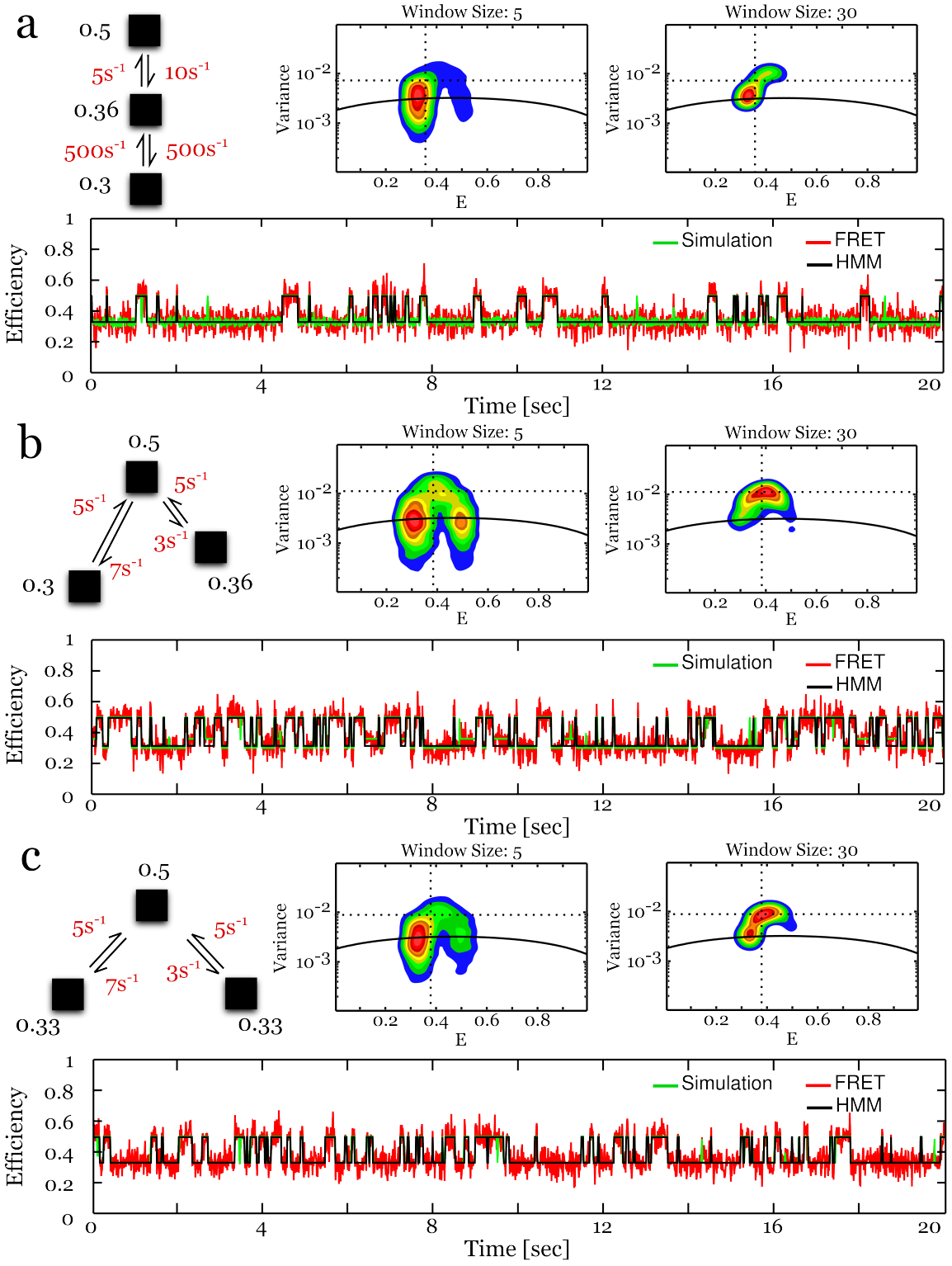
We can at least visually confirm that the dispersion is greater than expected by simulating a 3-state Poisson Markov chain with degenerate states, $E_A = E_B = 0.33$ and the same kinetic scheme as in Figure 7.7(b). We notice that the low-FRET low-variance region in Figure 7.7(c) exhibits reduced dispersion in E along the shot noise prediction, and thus differs from the distribution in Figure 7.7(b).[12]

7.3.10 MVA Detects Population Dynamics

Thus far, our discussion has focused on individual molecule behavior, but can MVA detect dynamics at the population level? Population behavior in experimental samples is rarely homogeneous. Observation times in smFRET measure-

[12] I would not bet my life inferring kinetic topologies from MVA histograms such as those in Figure 7.7 (a) and (b), but we can at least get a crude sense of what different kinetic schemes may look like, allowing us to visually eliminate certain topologies.

Figure 7.7 ***(following page)*****:** Mean-Variance Analysis: Simulation (3 States). To test MVA performance for a multiple state system, we simulated Poisson Markov chains of length $T = 2000$, with multiple state means, varying kinetics, and average photon count $N = 250$. (a) simulated time series with states $\{E : 0.3, 0.36, 0.5\}$ and kinetics as shown. We simulated the lower E states to rapidly convert faster than the experimental time resolution (10ms), with slower excursions between the higher two E states. The MVA histograms indicate variance beyond the shot noise prediction and suggest inter-state conversion within the low E low-variance region. The hidden Markov model (HMM) fit is unable to resolve 3 states. The HMM fits the mean of the two lower E states, while MVA detects the presence of multiple lower E states. (b) simulated time series with states $\{E : 0.3, 0.36, 0.5\}$ and kinetics as shown. To demonstrate that the low-FRET dynamics in (a) are real, we simulated a similar system, but where the two lower E states do not inter-convert. The MVA histograms show that the lower E low-variance region is shot-noise-limited, which indicates the absence of inter-conversion. The HMM fit is unable to resolve 3 states; MVA does not resolve 3 states either, except for a possible greater than expected dispersion in E for the low E low-variance density. (c) simulated time series with states $\{E : 0.33, 0.5\}$ and kinetics as shown. To distinguish the 3 state model in (b) from a 3 state model with one degenerate state, we simulated both lower E states as the average of the two low E states in (a) and (b). As in (b), the lower E states do not inter-convert. The low-variance region is, in fact, different from (b), with mean value dispersion in (a) reduced compared to (b). HMM correctly identifies two states; neither (straightforward) HMM or MVA distinguish degenerate states. Key: kinetic rates in red, states are black squares with mean state values in black, underlying 'true' state trajectory in green, HMM fit in black, calculated FRET in red.


a
0.5
5s-1
10s-1
0.36
500s-1
500s-1
0.3
Window Size: 5
Window Size: 30
Variance
E
Efficiency
Simulation
FRET
HMM
Time [sec]
b
0.5
5s-1
5s-1
3s-1
7s-1
0.3
0.36
c
0.5
5s-1
5s-1
3s-1
7s-1
0.33
0.33

ments are highly constrained due to fluorophore photobleaching lifetimes being on the order of only 30-60 seconds for low laser powers. Due to factors such as uneven TIRF illumination intensity and inter-molecular heterogeneity, individual state baselines are often highly variable. Inter-molecular heterogeneity also extends to kinetics, where some molecules exhibit kinetics that range over several orders of magnitude. Hence, detecting dynamics at the population level is not straightforward, as the population will be a convolution of multiple, often vastly different, behaviors, thus masking any one particular behavior.

Nevertheless, applying MVA to populations may be useful to at least suggest that some molecules exhibit dynamics, which would allow, for example, optimization of experimental temporal resolution for dynamics detection. We thus simulated 100 2-state Poisson Markov chains of variable length T, varying state means, varying kinetics, and low average photon count $N = 250$. Letting T vary simulates photobleaching, which we modeled as a single geometric decay.[13] We define the probability of 'success' p (where success is photobleaching) as the inverse expected number of data points. Hence, for $\langle T \rangle = 2000$, $p = 0.0005$. To generate a random geometric variable X, we draw a random number r from a uniform distribution $\mathrm{U}(0, 1)$ and calculate

$$X = \text{int}\left(\frac{\ln r}{\ln(1-p)}\right) + 1 \tag{7.17}$$

where the `int` function converts the quantity in braces to an integer. If we let E vary for both states, this simulates baseline shift and inter-molecular heterogeneity with regard to conformational dynamics. E had mean state values $\{\langle E \rangle : 0.3, 0.4\}$. The state means varied with Gaussian errors and standard deviation $\sigma_1 = \sigma_2 = 0.05$, which is greater than found by Holden and co-workers [135] for DNA standards ($\sigma \sim 0.03$), but in accordance with both my own and others' experimental protein dynamics work[14]. To generate state means, we defined a random variable μ according to

$$\mu_i = \langle E \rangle_i + (\sigma_i \cdot r) \tag{7.18}$$

[13] The geometric distribution is the discrete analogue of the exponential distribution.
[14] Geraint Evans, personal communication, 2012.

where i is the state index and r is a random number drawn from a standard normal distribution $r \sim \mathcal{N}(0, 1)$.

Letting the rates k_{ij} vary simulates the inter-molecular kinetic heterogeneity observed between experimental time series. We can approximate the logarithm of a kinetic rate distribution by a normal distribution [86, 107, 218], such that $\ln k \sim \ln \mathcal{N}(\mu, \sigma^2)$. A log-normal random variable X is given by

$$X = e^{\mu + \sigma Z} \tag{7.19}$$

where Z is a standard normal random variable with zero mean and unit variance. Hence, we generate a kinetic rate random variable by drawing a random number Z from a standard normal distribution and calculating

$$k_{ij} = \exp[\ln \langle k_{ij} \rangle + \sigma^2 \cdot Z] \; ; \qquad i \neq j \tag{7.20}$$

which reduces to $k_{ij} = \langle k_{ij} \rangle$ when the rates do not vary. Note that σ^2 roughly controls the orders of magnitude over which k_{ij} vary. We consequently simulated kinetics on the order of the simulated temporal resolution (10ms) $\langle k_{01} \rangle = \langle k_{10} \rangle = 50\text{s}^{-1}$ and defined $\sigma^2 = 1$. Rates varied between 1 and 500s^{-1}.

Figure 7.8 shows the simulation parameters and results. While the average mean values of the states differed by $\Delta E = 0.1$, which is equivalent to 3σ separation, baseline variation leads to a unimodal distribution with a peak centered about the average of the two simulated states. The data histogram does not suggest multiple states within the population; the MVA histogram, however, indicates otherwise. The distribution dispersion and asymmetry suggests the presence of state inter-conversion for time series within the population. Hence, for complex heterogeneous populations, MVA provides visual affirmation of dynamics.

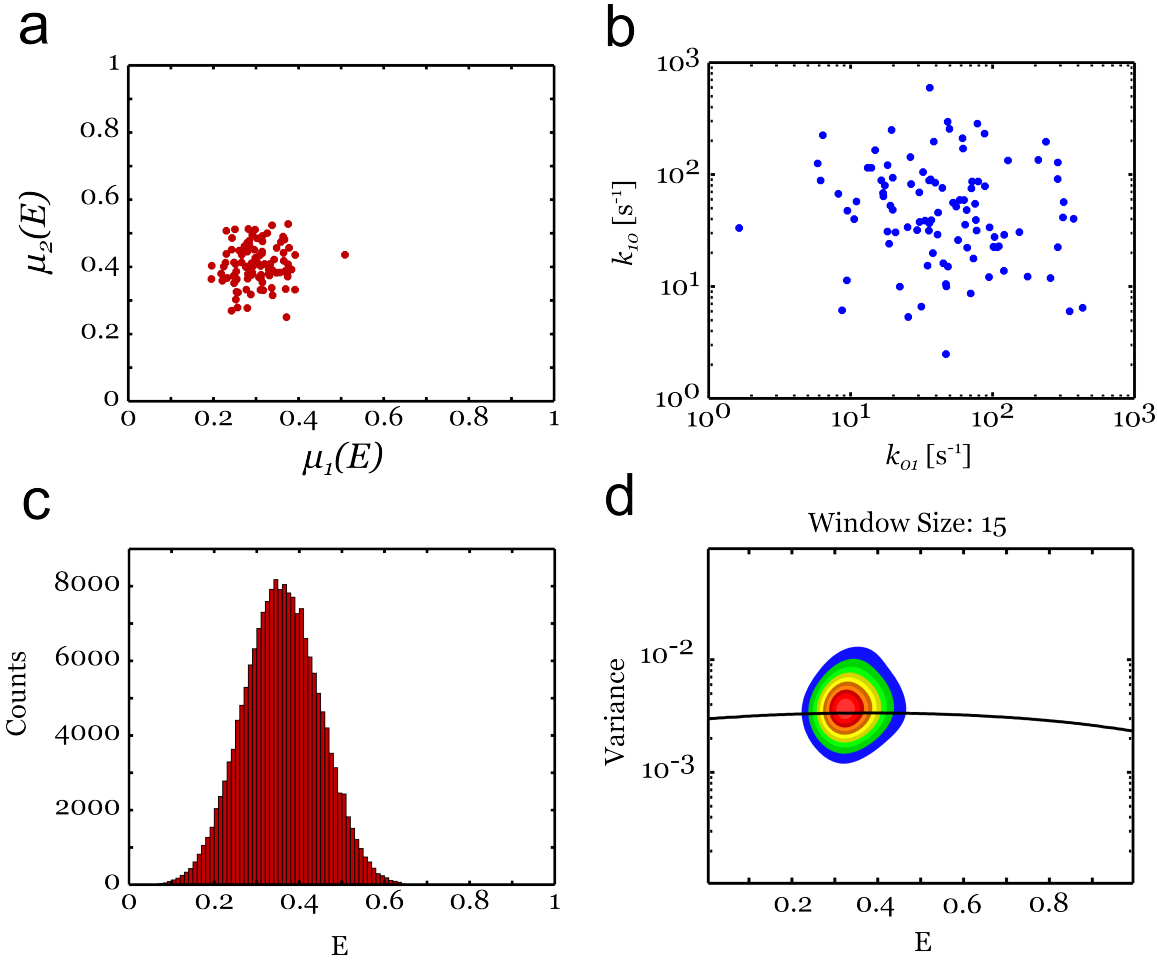


Figure 7.8: Mean-Variance Analysis: Simulation (Multiple Time Series). To test MVA performance for multiple time series where samples exhibit inter-molecular heterogeneity, we simulated 100 2-state Poisson Markov chains of variable length T, varying state means $\{\langle E \rangle : 0.3, 0.4\}$, varying kinetics $\{\langle k_{01} \rangle = \langle k_{10} \rangle = 50\text{s}^{-1}\}$ and average photon count $N = 250$. (a) to account for experimental baseline variation, we simulated time series with varying mean state values, where we allowed the values to independently vary with standard deviation $\{\sigma_1 = \sigma_2 = 0.05\}$. (b) to account for experimental samples having a distribution of kinetic behavior, both kinetic rates in each time series varied according to a log-normal distribution $\ln \mathcal{N}(\mu = \ln \langle k \rangle, \sigma^2 = 1)$, resulting in rates varying between 1 and 500 per second. (c) E distribution for all time series. The histogram appears unimodal with a peak centered about the average of the two simulated states and does not suggest multiple states. (d) MVA histogram for all time series. While the population density is close to the shot noise prediction, the density asymmetry and dispersion toward higher E suggests at least two inter-converting states. Hence, for complex heterogeneous populations, MVA provides visual affirmation of dynamics.

7.3.11 Global MVA of Limited Data Sets

We performed one additional test for MVA. Global analysis techniques, such as HMM, often require long time series training sets to ensure sufficient sampling of states. While simultaneous HMM analysis of many short training sets is possible with HMM, such analysis assumes the time series exhibit homogeneous kinetic behavior, i.e., do not have time-varying transition probabilities; see Chapter 3. We know inter-molecular kinetic homogeneity is rarely the norm. What we frequently desire is simply a macroscopic overview of many molecules.

To this end, we simulated 100 short 2-state Poisson Markov chains of length $T = 100$, with varying state means $\{\langle E \rangle : 0.3, 0.4\}$, varying kinetics $\{\langle k_{01} \rangle = 50\text{s}^{-1};\ \langle k_{10} \rangle = 10\text{s}^{-1}\}$, and high average photon count $N = 1000$. Mean state values E_i and kinetic rates k_{ij} varied as above. By increasing the photon count, we increase our photon stream signal-to-noise and FRET resolution, allowing us to better resolve E states despite having limited data. Figure 7.9 shows the results of the simulation.

Based on system kinetics and expected variation, we expect three types of behavior: (i) time series which do not transition within T, due to relatively slow kinetics; (ii) time series which transition rapidly, resulting in one or more transitions per data point and a continuum of intermediate states; and (iii) time series with kinetics 'optimal' for the temporal resolution and with clear state transitions. A mixture of all three behaviors comprises the MVA histograms in Figure 7.9. Most time series are shot-noise-limited, with dispersion along the shot noise curve. An example is the time series at the bottom right. A second population appears most clearly in the MVA histogram for $W = 50$, exhibiting high variance. Time series from this population show rapid fluctuations that cannot be accurately resolved, but which cause significant dispersion beyond the shot noise prediction. An example is the middle time series in the right column. A third population belongs to time series with clear step-like transitions. These time series show the characteristic arc toward higher variance, indicating windows containing a single transition.

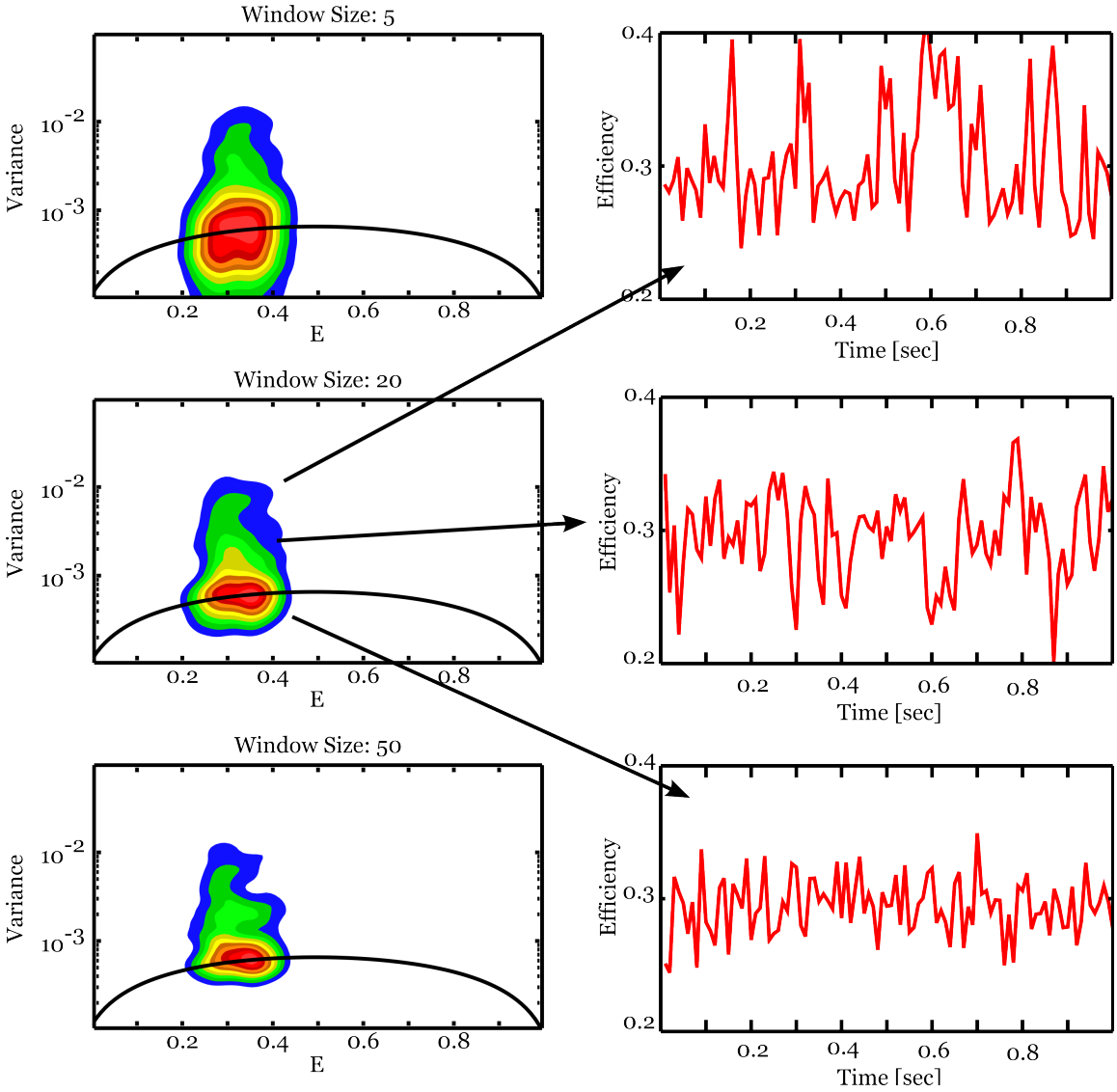


Figure 7.9: Mean-Variance Analysis: Simulation (Short Time Series). To test MVA performance for multiple short time series where samples exhibit inter-molecular heterogeneity, we simulated 100 2-state Poisson Markov chains of length $T = 100$, varying state means $\{\langle E \rangle : 0.3, 0.4\}$, varying kinetics $\{\langle k_{01} \rangle = 50\text{s}^{-1}; \langle k_{10} \rangle = 10\text{s}^{-1}\}$ and average photon count $N = 1000$. The left column shows MVA histograms for $\{W : 5, 20, 50\}$, and the right column shows three simulated time series illustrating the behavior of similar time series comprising the corresponding population (arrows). We find three distinct populations in the histograms: (i) shot-noise-limited species, with dispersion along the shot noise curve accounting for both $\langle E \rangle = 0.3$ and $\langle E \rangle = 0.4$ species not fluctuating within T; an example is the time series at the bottom right. (ii) rapidly inter-converting species, with high variance. Due to fast kinetics (faster than experimental temporal resolution), directly identifying transitions is difficult, but we indirectly infer their presence due to increased dispersion. An example is the middle time series. (iii) species inter-converting on time scales slower than experimental temporal resolution. These species show clear state transitions and occupancies. Accordingly, not only do the species have high variance, we expect an arc in the histogram connecting the species, as in Figure 7.4. An example is the time series at the top right.

An example is the time series at the top right.

While simulated time series allow confident assignment of populations and time series, MVA histograms for experimental data can be 'reverse engineered' in the sense that we can identify density regions of interest and then investigate the individual molecule behavior. Hence, MVA histograms provide an intuitive and simple visualization tool for data reduction and sorting.

7.4 Experimental Results

Having tested and validated MVA using simulations, we sought to apply MVA to experimental data. We subsequently revisit RNAP, which we previously investigated using BVA (see above, Section 4.4). Transcription initiation is a multi-step process [28]. After forming the RNAP-σ^{70} holoenzyme, binding promoter DNA, and unwinding $\sim$ 14 bp of DNA surrounding the transcription start site, RNAP begins RNA synthesis as an RNAP-promoter initial transcribing complex (RP_{itc}) before escaping the promoter and entering productive RNA synthesis as an RNAP-DNA elongation complex (RD_e). RP_{itc} does not immediately transition to RD_e , however, but instead reiteratively synthesizes and aborts short nascent RNAs ($\lessapprox$ 9-11 nucleotides (nt)) while remaining firmly bound to the -35 and -10 promoter elements. This reiterative synthesis is termed *abortive initiation*.

Previous work using confocal microscopy and magnetic tweezers elucidated the mechanism of abortive initiation [169, 264]. Margeat and co-workers subsequently immobilized heparin resistant RNAP-DNA complexes on passivated surfaces and attempted to monitor initial transcribing complexes and their kinetic behavior using single-molecule TIRF microscopy and ALEX spectroscopy [214]. While qualitatively informative, Margeat and co-workers lacked the temporal resolution to provide a quantitative assessment of initial transcribing complexes. Using brighter fluorophores and an experimental setup with higher temporal resolution, we sought to better understand RP_{itc} kinetics.

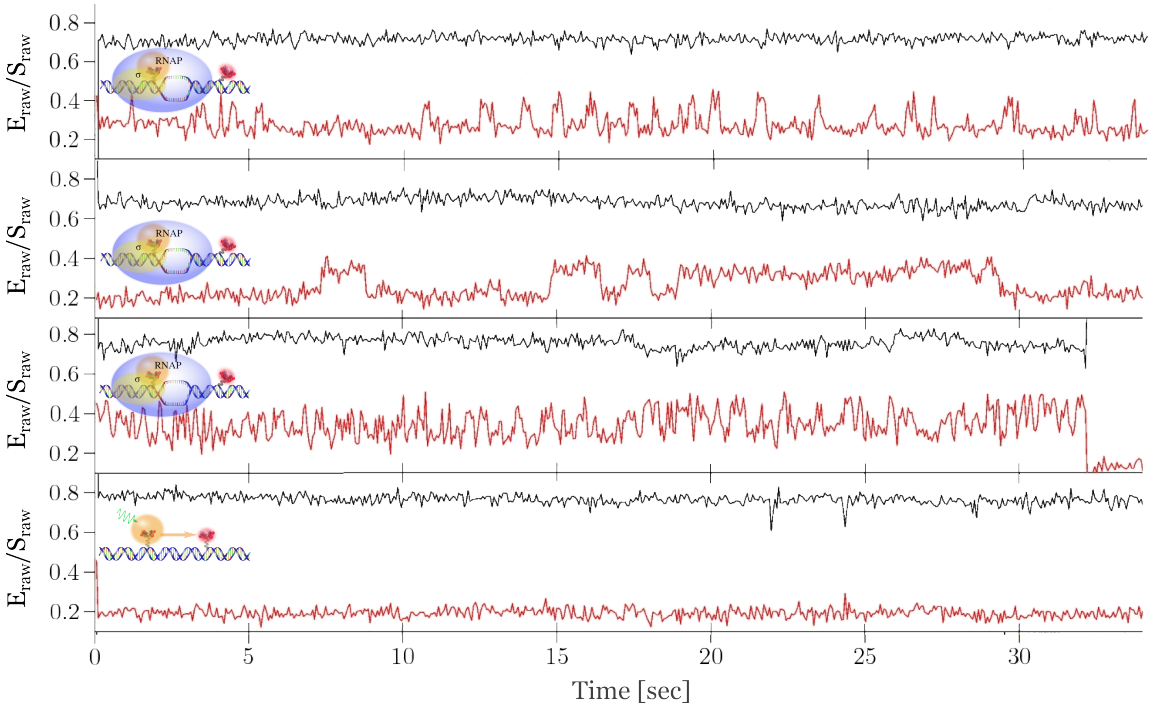


Figure 7.10: Surface-immobilized Initial Transcription Complexes. We immobilized RNAP initial transcription complexes on a passivated surface. The bottom plot shows a DNA only control, representing a lack of fluctuations in FRET efficiency E (red) in the absence of RNAP. The top three plots show time series in the presence of RNAP. We note that the stoichiometry time series (black) does not show significant fluctuations beyond shot noise, suggesting that the fluctuations observed in E should not be attributed to distance-independent photophysics. We notice three types of kinetic behavior at 60ms temporal resolution: (i) intermediate state inter-conversion rates (top); (ii) slow inter-conversion rates (second from top); (iii) fast inter-conversion rates (third from top).

7.4.1 Immobilized dsDNA

We surface-immobilized RNAP-promoter complexes using the (-15,+15) dsDNA construct (see Figure D.1). Initial immobilization strategies used biotinylated ds-DNA, to which RNAP was bound. Figure 7.10 shows acquired time series.

In the absence of RNAP holoenzyme, we do not observe E fluctuations (bottom of Figure 7.10), but, once bound to RNAP, we observe a range of kinetic behavior. At 60ms temporal resolution, some immobilized molecules slowly inter-convert between E states at $\sim$ 0.25 and $\sim$ 0.4, while others show fluctuations at the experimental temporal resolution. We initially performed these measurements in the presence of nucleotides, but subsequent measurements without nucleotides demonstrated similar behavior, suggesting that complexes are not engaged in abortive initiation. Instead, complexes appear to inter-convert between

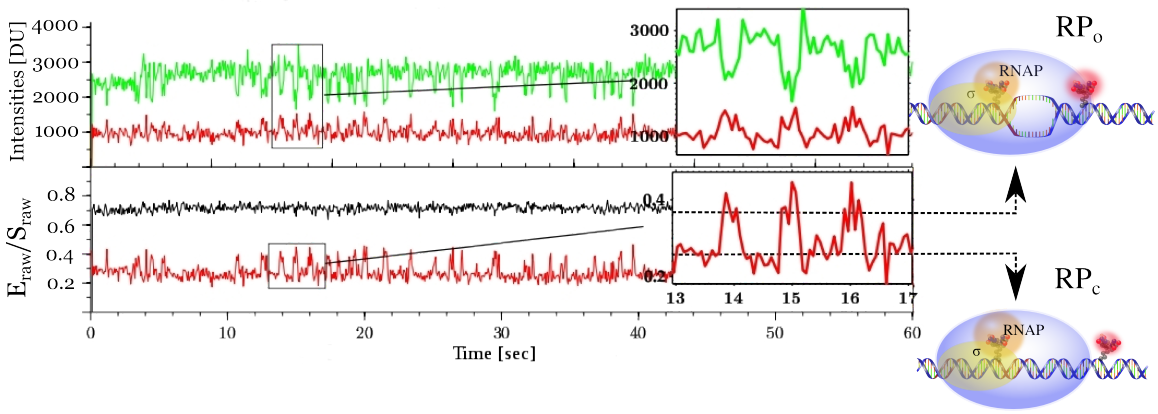


Figure 7.11: Surface-immobilized Initial Transcription Complex. Further to Figure 7.10, we provide raw intensity time series for a complex undergoing dynamic fluctuations. Donor and acceptor time series show clear anti-correlated fluctuations, suggesting that donor blinking is not the cause of E fluctuations. The E states appear centered about $\{\langle E \rangle : 0.25, 0.4\}$. We attribute these states to RP_c and RP_o , respectively. Note: raw intensities are in digital units [DU]. Key: $D_{ex}D_{em}$ is green; $D_{ex}A_{em}$ is red (top plot); S is black; E is red (bottom plot). Temporal resolution: 60ms.

a 'closed' complex, in which dsDNA is roughly linear (compare the E value in dsDNA only control in Figure 7.10), and a state in which downstream DNA (and consequently the acceptor) is pulled closer to upstream DNA. A plausible explanation is transcription bubble opening and closing, whereby RNAP samples closed RP_c and open RP_o states.

In addition to the dsDNA control, Figure 7.11 argues against the conclusion that distance-independent photophysics are the cause of fluctuations. We observe clear anti-correlated fluctuations between donor and acceptor time series, and we further observe nearly constant stoichiometry (in the figure, S is in black), suggesting that distance-independent acceptor photophysics are not the cause of E fluctuations.

7.4.2 Immobilized RNAP Holoenzyme

While the fluctuations appear real, we struggled with reproducibility. The yields for surface-immobilized complexes remained low, despite significant optimization attempts and complex formation strategies, e.g., *in situ* and using different dsDNA constructs. Subsequent work in our laboratory has pursued two alternative approaches for surface-immobilization: vesicle encapsulation [24,151,204,243,

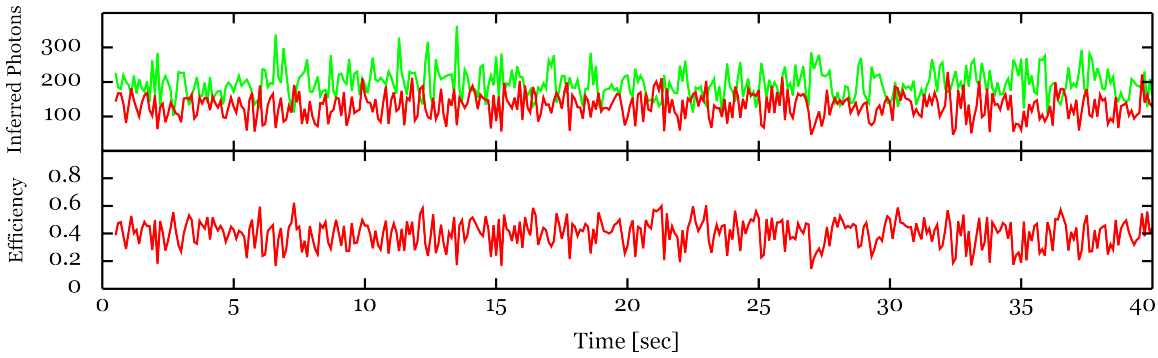


Figure 7.12: Surface-immobilized RP_o . RP_o complexes immobilized using a his-tag antibody surface attachment show similar behavior to the immobilized dsDNA strategy in Figure 7.13. Kinetics appear faster due to decreased temporal resolution, but the complexes exhibit similar state behavior ($\{\langle E \rangle : 0.2, 0.45\}$) as previously. Temporal resolution: 100ms.

244,270] and his-tag antibody surface attachment [154,155,247]. These approaches have met with much greater success.

Accordingly, we focus on the most recent work, which used hig-tag antibody surface attachment. With his-tag antibody attachment, the enzyme, rather than the dsDNA, is surface attached, meaning unbound dsDNA is only immobilized through non-specific absorption. This technique has the added benefit that dsDNA cannot be confused with RP_c . As seen in Figure 7.10, dsDNA and RP_c are sufficiently close in E to be considered the same given experimental uncertainty, e.g., due to variations in baseline. With careful pegylation, we minimize non-specific absorption, and thus we can be confident that observed immobilized biomolecules are RNAP-DNA complexes. We can then attribute any E fluctuations not attributable to distance-independent photophysics to enzyme conformational dynamics.

Figure 7.12 provides a sample time series for his-tag immobilized complexes taken at 100ms temporal resolution[15]. The time series shows similar state interconversion behavior as before, with faster apparent kinetics as expected with decreased temporal resolution. Out of a sample population of 105 complexes, 56 were analyzable, i.e., free of distance-independent photophysics and focal drift, at

[15]Note that the 100ms temporal resolution his-tag experiments differ from the 60ms dsDNA immobilized experiments for reasons of convenience and as both sets of experiments were preliminary. As we were able to resolve dynamics at 100ms, we used this resolution for subsequent experiments throughout the rest of this chapter.

FRET Behavior	Number
$\langle E \rangle < 0.3$	23
$0.3 \leq \langle E \rangle < 0.4$	7
$0.4 \leq \langle E \rangle < 0.5$	10
$0.5 \leq \langle E \rangle$	2
Dynamic ($\Delta E \geq 0.2$)	14

Table 7.1: RNAP Initial Transcription Dynamics. We categorized molecules based on their state and kinetic behavior. We placed molecules showing stable E on the measurement timescale (~ 30 sec) in groups based on their mean FRET. Any molecules showing persistent, i.e., non-blinking fluctuations beyond $\Delta E \geq 0.2$, we categorized 'dynamic'.

least 100 data points, and having a single photobleaching step, and 14 of those 56 showed fluctuations similar to Figure 7.12, with varying kinetics as in Figure 7.10.

Previously for the dsDNA immobilization strategy, dsDNA and RP_c complexes were too closely spaced to permit virtual molecule sorting, i.e., to distinguish dsDNA from a closed RNAP complex, and population statistics were too low to allow saying anything meaningful about individual molecule behavior. Using his-tag antibody attachment, however, we overcame both barriers and could identify behaviors consistent among molecules and within E subpopulations. In addition to the 14 identified dynamic complexes, we provide a summary in Table 7.1.

We then asked if we could say anything about the 'dynamics' of the 42 molecules which did not show obvious fluctuations between well-separated E states. Two possibilities may explain the heterogeneity: different kinetics, either faster or slower, or different E state values. For example, the 7 molecules with $0.3 \leq \langle E \rangle < 0.4$ could be rapidly inter-converting between the 0.20 and 0.45 states in Figure 7.12 at rates faster than the experimental temporal resolution. Such rapid inter-conversion would result in an average (intermediate) E value. Alternatively, these molecules could be a stable species with a distinct characteristic E state value differing from the lower and higher E states found in fluctuating molecules. We turned to MVA to provide insight into sorted species and their behavior.

We first generated a shot noise prediction for the entire sample population under the assumption that each molecule is static. Figure 7.13 shows that the shot noise prediction (in red) fails to account for increased dispersion between the main

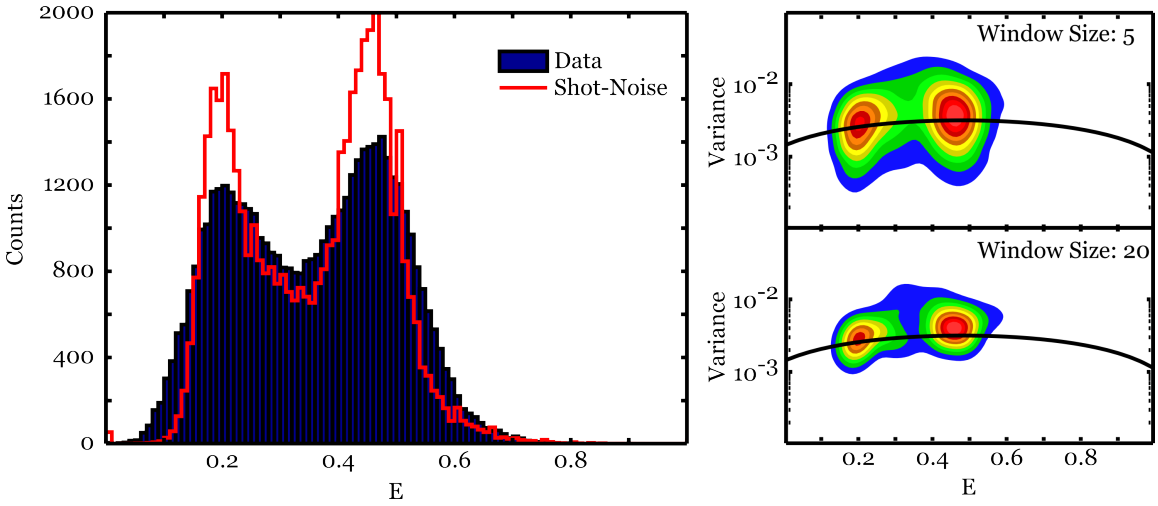
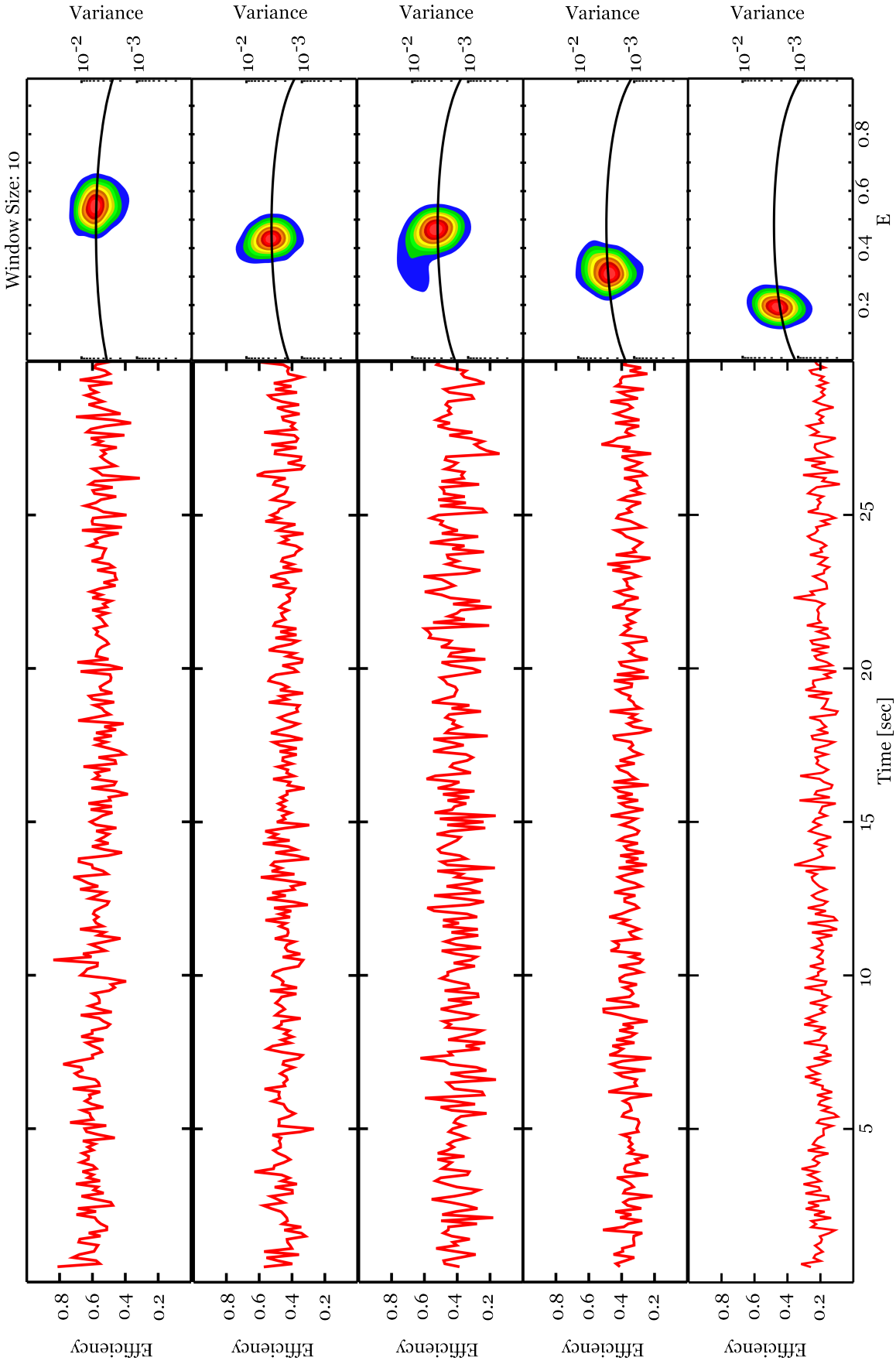


Figure 7.13: Mean-Variance Analysis of RP_o : All Molecules. The shot noise prediction (left; red) for the case where each time series within the sample population is shot-noise-limited fails to account for increased dispersion between the main peaks centered around $E = 0.2$ and $E = 0.45$, as well as in the data distribution tails. MVA histograms show the two main populations evident in the 1D histogram, but also asymmetries connecting the two distributions, suggesting dynamic inter-conversion, e.g., as in Figure 7.12.

distribution peaks centered around $E = 0.2$ and $E = 0.45$, as well in the distribution tails. Accordingly, we performed MVA for the entire sample population. We observe that the two main populations are close to the shot noise prediction; however, the significant spread connecting the populations and the density asymmetries suggest dynamic inter-conversion. Interestingly, for $W = 20$, we notice that the lower density appears to connect with a more sparsely populated density centered at $E = 0.3$. We can also surmise that the high variance region in blue located at $E = 0.35$ belongs to molecules showing clearing transitions and mean state lifetimes $\tau \approx W/2$, as in Figure 7.12.

We separately analyzed each group of molecules listed in Table 7.1. Figure 7.14 shows the MVA results. Dynamic molecules show high-variance, with kinetics favoring the higher E state (middle row). The spread toward high-variance

Figure 7.14 *(following page)*: Mean-Variance Analysis of RP_o : Sorted Molecules. We performed MVA for each group of molecules listed in Table 7.1. The left column provides representative E time series for molecules belonging to each group, while the right column shows the MVA histogram for each group. Dynamic molecules show high-variance as expected (middle row), with kinetics favoring the higher E state. Higher E populations show asymmetries toward lower E populations and vice versa, suggesting inter-conversion between the two. Interestingly, the intermediate E population ($\langle E \rangle \approx 0.35$) shows asymmetries toward both lower and higher E values, while its main density resides along the shot noise prediction suggesting that the population is a legitimate E state and not simply the result of high-low E averaging.


Variance
Variance
Variance
Variance
Variance
Window Size: 10
E
Efficiency
Efficiency
Efficiency
Efficiency
Efficiency
Time [sec]

explains the spread in Figure 7.13 for all molecules, as we predicted. Based on the inter-converting E state values, this group of molecules appears to be actively sampling both RP_c and RP_o . High E molecules (top row) appear to be shot-noise-limited, with some dispersion along the shot noise curve, possibly due to the slight baseline drift, as in the example time series. These molecules appear to remain primarily as RP_o . The mid-E ($\langle E \rangle \approx 0.45$) and low-$E$ ($\langle E \rangle \approx 0.2$) populations show the beginnings of high arcs, suggestive of fast dynamics not captured at the experimental temporal resolution. See, for example, between 10-15 sec in the second row, evident from anti-correlation in donor/acceptor intensity time series (data not shown). The low-E molecules we attribute to molecules primarily as RP_c . Perhaps more interestingly, the intermediate-E molecules ($\langle E \rangle \approx 0.35$) show asymmetries toward both lower and higher E values while also having the main density centered about the shot noise prediction. This finding supports a hypothesis that the intermediate-E state is a legitimate state and not simply the average E value for molecules rapidly inter-converting between high and low E states at timescales faster than the experimental temporal resolution.

7.5 Discussion

The observations of a stable intermediate complex between RP_c and RP_o along with dynamic interconversion between RP_c and RP_o are the first single-molecule observations of active sampling and of a long-lived intermediate for bacterial RNAP holoenzyme. Chen and co-workers [33] previously performed Brownian dynamics simulations of the $RP_c \rightarrow RP_o$ transition in *Thermus aquaticus* (*Taq*) RNAP holoenzyme, a closely related homolog of *E. coli* RNAP [232, 233, 363]. As RNAP sequentially melts double-stranded promoter DNA [37,60], the non-template strand makes stable interactions with RNAP, and RNAP pulls downstream base-pairs into itself, a process termed *scrunching* [169, 264]. After promoter DNA melting and bubble formation, DNA bending and positioning inside the main channel relieves strain. Template strand entry into the active site allows the RNAP-promoter

complex to form a catalytically-competent transcription complex.

The simulations of Chen and co-workers revealed two types of trajectories. In the comparatively fast trajectory, local promoter DNA melting began from the -10 element and then sequentially unwound until +2, which is the base immediately downstream of the transcription start site +1. In the slow trajectory, stochastic melting downstream of the -10 element occurred first which prevented sequential base-pair melting. Slow trajectories subsequently formed relatively long-lived metastable intermediate states which lasted for tens of microseconds. Only after resetting to the initial RP_c state and then sequentially melting base-pairs from the -10 element could the RNAP-promoter complex form RP_o.

Previous experimental work has found intermediates for phage (T7), eukaryotic RNA polymerase II, and mitochondrial RNA polymerases [177, 293, 317, 350]. Moreover, bacterial RNAP initial transcription complexes involving other transcription factors (σ^{54}) form stable closed complex intermediates [32, 93, 98]. Gries and co-workers used fast permanganate footprinting of bacterial RNAP at a phage promoter (λP_R) to uncover intermediates during transcription bubble opening and stabilization [108]; however, their work could only infer intermediate existence without direct observation. While several additional works have provided experimental evidence for intermediates involving σ^{70} [268, 273, 288], the intermediates were unstable and thus short-lived. To our knowledge, no work has found long-lived, i.e., on the order of tens of seconds, heparin-resistant bacterial RNAP-promoter closed and intermediate complexes.

What might be the origin of these stable intermediates? One possibility is that the intermediates are experimental artifacts. We use short dsDNA promoter constructs[16], which provide limited upstream RNAP-dsDNA interaction. Several works have demonstrated that upstream dsDNA contacts are needed for RNAP to form stable downstream contacts, possibly due to allosteric and/or network effects [57, 273]. Davis and co-workers found that RP_o formation at λP_R is greatly

[16](-15,+15) is 64 base-pairs, which is less than half the DNA persistence length of ∼150 base-pairs [56].

accelerated by DNA upstream of -47. They further found that an intermediate formed at an upstream-truncated λP_R promoter exhibits different downstream contacts [57]. Saecker and co-workers hypothesized that upstream DNA contacts beyond -47 may remove an obstacle in the main channel cleft and contribute to the difference in downstream dsDNA cleft insertion [273]. In the absence of upstream elements, early complex stability is dependent on DNA wrapping concomitant with downstream contact formation and DNA entry into the main channel [268]. Rogozina and co-workers observed a branched pathway in which an off-pathway intermediate formed that did not efficiently isomerize to establish the downstream contacts important for RP_o complex stability [32]. They proposed that a different DNA orientation within the jaws obstructed downstream contact accessibility [268]. The slow rate of back-isomerization led to off-pathway complex accumulation. One hypothesis, then, is that the off-pathway intermediate may account for the unproductive initiation complexes referred to as moribund [303]. Taken together, these studies suggest that our promoter DNA constructs may facilitate off-pathway intermediates due to insufficient upstream contacts.

We note, however, that surface perturbations are not likely to explain the presence of intermediates. While distance-independent photophysics complicates E-BVA analysis of RP_o diffusing in solution, a spread connects a low-E and mid-E population for the (-15,+15) promoter construct; see Figure 4.4(a). This spread is consistent with the dynamic interconversion observed in surface measurements. In addition, we have encapsulated RP_o complexes inside lipid vesicles and observed RP_c -RP_o interconversion with kinetics similar to that of his-tag antibody attachment (data not shown). Hence, we conclude that the experimental method is unlikely to be the source of intermediates.

While our work supports the hypothesis of multiple intermediates [108] and supports our hypothesis that RNAP can actively sample multiple states before, and possibly during, abortive initiation, we are currently reproducing our results using longer promoter DNA constructs. Nevertheless, if we suppose that the sta-

ble intermediate is legitimate, its presence motivates several further experiments to test its importance within transcription initiation.

What might be the physiological relevance of active sampling and stable intermediates along the pathway between promoter binding and RP_o ? Several works have proposed a branched pathway for ternary initial transcription complexes RP_{itc} – one branch leading to a moribund complex making only abortive products and another branch leading to a complex primarily making full length transcripts [186,302,303,328]. For the promoters previously investigated, the overall abortive-to-productive ratio is large, meaning that most synthesized RNA are aborted products, and promoter-sequence dependent [328]. Abortive transcripts are made in vivo [99], but their cellular function as small RNAs remains a matter of open debate [196,215].

The downstream mobile domains in the β' subunit, including the clamp and jaw, can form tight downstream dsDNA contacts [182], helping to stabilize RP_o. Hsu and co-workers observed that downstream DNA from +1 to +20 plays an important role in determining initiation-to-elongation transition efficiency [141]. Tight contacts would mean an increased energetic barrier for the complex to overcome in order to escape into elongation. Therefore, increased complex stability via downstream contacts could favor abortive initiation.

Interestingly, the clamp/jaw contacts downstream DNA only *after* DNA opening and unwinding is complete [32]. Complexes favoring conformational states closer to RP_c , and consequently weaker downstream contacts, would have lower energetic barriers to elongation and productive synthesis. An intriguing possibility then is that lower FRET intermediates and complexes actively sampling both low and high FRET states may be more capable of productive synthesis, whereas stable high FRET species are only capable of abortive initiation. Nucleotide availability and regulatory ligands should further affect the relative proportion of molecules in a branched pathway, suggesting additional avenues for transcription regulation.

A next step for our work would then be to observe how the relative proportion of molecules belonging to low, intermediate, and higher FRET states changes upon the addition of limiting nucleotides. By preventing complexes from entering elongation, we bias molecules into abortive initiation even if not initially moribund. If we compare the distribution of complexes when providing limiting nucleotides to the relative proportion when providing all nucleotides, we will have direct evidence for the branched pathway hypothesis and the importance of partially closed intermediates in productive initiation.

7.6 Summary

We demonstrated that MVA presents a qualitative summary of single-molecule time series data through which signal-to-noise ratio, individual state variance, baseline drift, and state number can be accessed. By comparing observed time series dispersion to a shot noise prediction, MVA provides an easy-to-interpret method to identify dynamics and to delineate deviation from theory, even when dynamics occur on timescales faster than experimental temporal resolution. While several researchers have investigated the shot noise limits of solution-based confocal measurements, few works have presented analysis methods addressing the shot noise limit in surface measurements.

While limited as a change-point detection and time series segmentation method, MVA does allow for dynamics identification in real-time. Rather than having to perform a global fit, as in, for example, hidden Markov methods, MVA is performed locally. By comparing MVA distributions to the shot noise prediction, we can tell, even for a single state system, something about distance-independent photophysics and system stability. Further, determining raw data quality used for analysis is particularly difficult *ex post facto* because literature presenting time series analyses supply only extracted dwell time histograms and a handful of "representative" raw traces. MVA is a means for summarizing time series data and demonstrating data suitability for change-point detection analysis. Also, because

MVA is essentially model free, excluding the shot noise prediction, MVA provides a fast and convenient first-pass method. One rarely performs more complex state detection model-based methods without first having a sufficient handle on data quality and characteristics for informing model *priors*. MVA provides one such means for data-reduction and identifying those molecules appropriate for later analysis involving model-based methods.

Contributions

Holden and co-workers [134,135,321] derived the uncertainty for surface-immobilized biomolecule TIRF FRET measurements and performed simulations validating its prediction. I implemented MVA for single-molecule FRET time series and performed all simulations. I conducted initial experiments of surface-immobilized initial transcription complexes. Diego Duchi continued the experiments, adopting different immobilization strategies (vesicle encapsulation and his-tag antibody attachment) to optimize experimental conditions. Analyzed data after Figure 7.11 is Diego's. I performed all MVA analysis. All writing and figures are my own.

Chapter 8

Conclusion

Single-molecule fluorescence techniques combine high spatial and temporal resolution to probe biochemical and cellular processes. These techniques allow us to study reaction pathways and asynchronous reactions in real time and uncover short-lived intermediates. By monitoring fluorescence intensity fluctuations over time, we can deduce structural mechanisms and reaction kinetics, and, in so doing, we can link structure to function. While ensemble methods allow us to investigate equilibrium distributions, single-molecule methods reveal the stochastic nature of molecular processes. Stochasticity is inherent to single-molecules and has important implications for cell-to-cell variability, where random fluctuations in gene expression provide the flexibility needed for cells to adapt to changing environments and to respond to sudden stresses [40, 55, 162, 210].

Single-molecule fluorescence data is not, however, without complications. Many photophysical and photochemical processes compete with fluorescence. This competition poses a significant challenge to our ability to unambiguously relate a variable, such as observed fluorescence, to a specific biological function. Careful experimental design and the use of techniques such as alternating laser excitation (ALEx) or multiparameter fluorescence detection (MFD) can help overcome this challenge, but a need for signal analysis methods which account and control for competing processes remains paramount.

In this thesis, we presented methods to detect non-distance-related fluores-

cence intensity fluctuations. First, we developed the theory of fluorophore stoichiometry S, a ratiometric variable which contains information about a fluorophore's photophysical state and allows us to monitor the state's evolution over time (Chapter 5). We used stoichiometry to extend burst variance analysis (BVA) to account for dynamics in the energy transfer efficiency E which did not directly depend on inter-fluorophore separation dynamics. Next, we proposed burst correlation analysis (BCA) to measure the extent to which S fluctuations could account for E dynamics (Chapter 6). We concluded that S fluctuations could at least partially account for apparent RNA polymerase (RNAP) promoter open complex E dynamics. Our work highlighted the importance of statistical methods capable of distinguishing FRET fact from artifact.

We further applied the theoretical predictions for total internal reflection fluorescence (TIRF) microscopy measurements of FRET and developed mean-variance analysis (MVA) to detect unresolved dynamics in surface-immobilized biomolecule time series [135] (Chapter 7). We demonstrated that we could identify FRET dynamics at and just below the measurement timescale. We used MVA to analyze RNAP promoter open complex dynamics. Our results found heterogeneous behavior and suggested a stable intermediate on the pathway between RNAP promoter closed and open complex. The existence of such an intermediate may have important implications in transcription regulation, and we proposed experiments to test an initial transcription branched pathway hypothesis [303].

In addition, we wrote and released publicly available software to facilitate single-molecule data analysis. We expect that such contributions will be of immense benefit to the single-molecule field and will encourage others to make their own contributions toward developing a common analysis framework for single-molecule time series.

8.1 Future Work

Our work has several opportunities for extension and improvement. S is more versatile than simply an internal control for E dynamics. Arguably, S is a more versatile molecular ruler than E. Many processes other than FRET (typically $20-80$ Å) are involved in fluorescence quenching, such photo-induced electron transfer ($<$ 20Å) [27, 70, 353] and nanometal surface energy transfer (20-200 Å) [157, 358]. Moreover, for molecular assemblies, FRET-based inter-fluorophore separation constraints are obviated. We can use two-color coincidence and S-BVA to study dynamic assembly of multiple fluorescent components, where fluorescent probes are beyond the Förster radius. Accordingly, we can use S-BVA to analyze any process which involves dynamic fluorophore quenching [200, 207]. Work is already underway to use S-BVA to study the dynamic interaction of DNA polymerase I (Pol I) with double-flap overhang substrates. These substrates are important in lagging strand DNA synthesis, as well as DNA repair pathways including base excision repair and nucleotide excision repair [349]. The combination of FRET and protein induced fluorescent enhancement (PIFE) [147] will enable us to determine substrate binding kinetics and to link enzyme-substrate organization to reaction chemistry.

We are further working to improve BCA using independent component analysis (ICA). Our goal is to use ICA to extract E fluctuations which are independent of S fluctuations. This improvement will enable us to unambiguously detect inter-fluorophore separation dynamics.

We have also extended the BVA statistical framework to develop a change-point detection algorithm. The binomial filter provides a straightforward method to detect change-points in solution-based single-molecule time series.

We previously developed hidden Markov model algorithms which use E-S information to delineate fluorophore stoichiometry dynamics from conformational dynamics in Holliday junctions [322]. We have subsequently improved the method and developed *regime-changing hidden Markov models* to model S-reported intramolec-

ular heterogeneity. While correlation methods can detect the presence of memory effects (Section 3.3.3.1), they cannot determine their location within a time series and thus are limited in their ability to quantify interconversion kinetics. Regime-changing HMM overcomes this limitation and provides a method to monitor ligand mediated dynamic disorder and detect both kinetic and conformational heterogeneity. This work will form the basis of a future publication.

Lastly, while we have developed several algorithms for single-molecule data analysis, the implementing software is not unified. We are currently compiling all our analysis algorithms into a single comprehensive software analysis package. Single-molecule researchers will soon be able to access our work through their web browser for fast, efficient, and robust data visualization and analysis.

Contributions

I developed the algorithm for linked hidden Markov models (HMMs) and performed all simulations. Stephan Uphoff supplied Holliday junction data for linked HMM. Timothy Craggs is performing Pol I measurements, and I am analyzing Pol I data using S-BVA. Alistair Wardrope and I jointly developed regime-changing HMM. I developed the binomial filter. I wrote all algorithms and software used in this thesis. All writing is my own.

Appendix A

Distributions

A.1 Geometric Distribution

Consider a series of independent trials, each having one of two outcomes: success or failure, i.e., a series of Bernoulli trials. Let $p = P(\text{trial ends in success})$. Define a random variable X as the trial at which the *first* success occurs. The probability mass function (PMF) is given by

$$\begin{aligned} p_X(k) &= P(X = k) \\ &= P(\text{first success occurs on the } k\text{th trial}) \\ &= P(\text{first } k-1 \text{ trials end in failure and } k\text{th trial ends in success}) \\ &= \underbrace{P(\text{first } k-1 \text{ trials end in failure}) \cdot P(k\text{th trial ends in success})}_{\text{independence}} \\ &= (1-p)^{k-1}p, \qquad k = 1, 2, \ldots \end{aligned} \tag{A.1}$$

We call the probability model in Equation A.1 a *geometric distribution* with parameter p. Regardless of its association with independent trials, the function $p_X(k) = (1-p)^{k-1}p,\ k = 1, 2, \ldots$ qualifies as a discrete PDF because (1) $p_X(k) \geq 0\ \forall\ k$ and (2) $\sum_{\forall\ k} p_X(k) = 1$:

$$\begin{aligned}\sum_{k=1}^{\infty} p_X(k) &= \sum_{k=1}^{\infty}(1-p)^{k-1}p \\ &= p\sum_{j=0}^{\infty}(1-p)^{j} \\ &= p\left(\frac{1}{1-(1-p)}\right) \\ &= 1\end{aligned}$$

The distribution's mean and variance are given by

$$E[X] = \frac{1}{p} \tag{A.2}$$

$$\text{Var}[X] = \frac{1-p}{p^2} \tag{A.3}$$

A.1.1 Memoryless Property

Statement: The probability of needing an additional $X = k$ trials (steps) to obtain the first success (transition out of state S_i) is unaffected by however many trials $n-1$ have already been observed.

$$P(X = n-1+k | X > n-1) = P(X = k) \tag{A.4}$$

Proof: From the definition of conditional probability,

$$\begin{aligned}P(X = n-1+k | X > n-1) &= \frac{P((X = n-1+k) \cap (X > n-1))}{P(X > n-1)} \\ &= \frac{P(X = n-1+k)}{1 - P(X \leq n-1)}\end{aligned}$$

where

$$P(X \leq n-1) = \sum_{k=1}^{n-1}(1-p)^{k-1}p = p\sum_{j=0}^{n-2}(1-p)^{j}$$

The sum is a partial sum for a geometric random variable and is equal to

$$p\sum_{j=0}^{n-2}(1-p)^{j} = p\frac{-[(1-p)^{n-1}-1]}{1-p-1} = 1-(1-p)^{n-1}$$

Hence, if we substitute for the numerator and denominator in our conditional probability

$$\begin{aligned} P(X = n-1+k|X > n-1) &= \frac{p(1-p)^{n-1+k-1}}{1-(1-(1-p)^{n-1})} \\ &= p(1-p)^{k-1} \end{aligned}$$

which equals $P(X = k)$. Q.E.D.

This is a rather counter-intuitive result. Given a geometric distribution with parameter p, if we consider the trial history with respect to the current instance, we can describe the probability by the same distribution: a sequence of trials geometrically distributed with parameter p until the first success. And further, the remaining number of trials until the first success has the same geometric distribution irrespective of the number of trials which have already occurred. Hence, the stochastic characterization of the trial process is the same in both cases. Accordingly, no matter when we begin to observe a process, the number of trials needed until a success is given by the same distribution, which is equivalent to starting anew each time we enter a queue.

Appendix B

Stochastic Processes

B.1 Poisson Process

B.1.1 Splitting a Poisson Process

Merging two independent Poisson processes with rates λ_1 and λ_2 is known to yield a Poisson process with rate $\lambda = \lambda_1 + \lambda_2$. But does the converse also hold true? Namely, can we split a Poisson process into two (or more) processes which are also *independent* Poisson processes? Intuitively, we would imagine the answer is yes. But if so, what is the rate of each process?

Let the set of random variables $\{X_i\}$ (inter-arrival times) form a Poisson process with rate λ, and let this process be the parent process we would like to 'split' into two separate processes. To make our case more concrete, let each event be a photon arrival, and let p be the probability of splitting off an event (photon) and placing it in sub-process 1 and $(1-p)$ the probability of placing an event (photon) in sub-process 2. Note that 'splitting' is itself a random process, with probability p.[1]

Let the set of random variables $\{Y_i^1\}$ form a stochastic process of independently, identically exponentially distributed random variables with rate λ_1. Hence,

[1] If splitting was not random, but instead involved dependent selection (e.g., even event numbers in sub-process 1 and odd event numbers in sub-process 2), splitting would be conditional. Merging two dependent Poisson processes does not necessarily result in a Poisson process, and neither is non-random splitting guaranteed to yield Poisson sub-processes.

Y_i^1 are the inter-arrival times for a Poisson process, and let this be sub-process 1. In which case, let N be a random variable equal to the number of events until an event is split into sub-process 2. A moment's reflection will reveal that N is a geometrically distributed random variable with

$$P(N = k) = p^{k-1}(1 - p), \qquad \text{for } k = 1, 2, \ldots$$

where k is the number of events until an event is placed in sub-process 2.[2]

Let $\{Y_j^2\}$ be the inter-arrival times of sub-process 2. Each Y_j^2 is a sum of X_i, the number of which is geometrically distributed. By construction, Y_j^2 are identically distributed and independent (vis-á-vis Y_i^1). But are the Y_j^2 also exponentially distributed?

Let $Z = Y_j^2 = \sum_{i=1}^{N} X_i$. So the probability that $Z \leq z$ is given by

$$P(Z \leq z) = \sum_{k=1}^{\infty} P(Z \leq z | N = k) P(N = k)$$

Substitute for Z

$$= \sum_{k=1}^{\infty} P\left(\sum_{i=1}^{N} X_i \leq z | N = k\right) P(N = k)$$

Use the condition that $N = k$, modifying the indexes of our sum...

$$= \sum_{k=1}^{\infty} P\left(\sum_{i=1}^{k} X_i \leq z\right) P(N = k)$$

[2]Each event has probability p of being placed in sub-process 1. The probability that $k-1$ events are placed in sub-process 1 is p^{k-1}. The probability that event number k is placed in sub-process 2 is $(1 - p)$. Hence, the total probability is $p^{k-1}(1 - p)$.

Substitute for $P(N = k)$ with the geometric distribution provided earlier

$$= \sum_{k=1}^{\infty} P\left(\sum_{i=1}^{k} X_i \leq z\right) p^{k-1}(1-p)$$
$$= (1-p) \sum_{k=1}^{\infty} P\left(\sum_{i=1}^{k} X_i \leq z\right) p^{k-1}$$

Now we use the fact that the sum of k exponentially distributed random variables, each with parameter λ is Gamma (Erlang) distributed with shape parameter k, scale (inverse rate) parameter θ, and PDF

$$f_{\Gamma(k,\lambda)}(z) = \frac{1}{\theta^k (k-1)!} z^{k-1} e^{-k/\theta}$$

in which case, setting $\theta = 1/\lambda$

$$P(Z \leq z) = (1-p) \sum_{k=1}^{\infty} \left(\int_{t=0}^{z} f_{\Gamma(k,\theta)}(z)\, dt \right) p^{k-1}$$

Substitute and rearrange

$$= \frac{(1-p)}{\theta} \int_{t=0}^{z} \left(\sum_{k=1}^{\infty} \frac{p^{k-1}}{\theta^{k-1}(k-1)!} t^{k-1} e^{-t/\theta} \right) dt$$

Shift the index for our sum

$$= \frac{(1-p)}{\theta} \int_{t=0}^{z} \left(\sum_{k=0}^{\infty} \frac{p^{k}}{\theta^{k}(k)!} t^{k} e^{-t/\theta} \right) dt$$

Consolidate terms

$$= \frac{(1-p)}{\theta} \int_{t=0}^{z} e^{-t/\theta} \left(\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{pt}{\theta} \right)^{k} \right) dt$$

The sum is a Maclaurin series converging to e^x, with $x = pt/\theta$

$$= \frac{(1-p)}{\theta} \int_{t=0}^{z} e^{-t/\theta} e^{pt/\theta} \, dt$$

which, after rearranging terms, equals

$$= \int_{t=0}^{z} \frac{(1-p)}{\theta} \, e^{-t(1-p)/\theta} \, dt$$

But this is just the integral for an exponentially distributed random variable with rate $\lambda^\star = (1-p)/\theta = \lambda(1-p)$. Hence, Y_j^2 are exponentially distributed and form the inter-arrival times of a Poisson process with rate $\lambda_2 = \lambda(1-p)$. Accordingly, sub-process 1 has rate $\lambda_1 = \lambda p$, with $\lambda = \lambda_1 + \lambda_2$. Q.E.D.

This result proves convenient in the case of $D_{ex}D_{em}$ and $D_{ex}A_{em}$ photon streams. If we imagine beginning with an unlabeled photon stream D_{ex} and then, dependent on probability p, assigning each photon to a separate sub-stream, we replicate the FRET process, where $p = E$. In this thought experiment, $D_{ex}D_{em}$ and $D_{ex}A_{em}$ are simply our labeling (or coloring) of photons which land in either sub-stream. Accordingly, not only are $D_{ex}D_{em}$ and $D_{ex}A_{em}$ independent Poisson processes, but then so is their sum $D_{ex} = D_{ex}D_{em} + D_{ex}A_{em}$. See Figure B.1 for a schematic.[3]

A similar result follows for $A_{ex}D_{em}$ and $A_{ex}A_{em}$ and their sum $A_{ex} = A_{ex}D_{em} + A_{ex}A_{em}$, where $p = dir$ (the direct excitation of the donor via acceptor excitation $A_{ex}D_{em}$; in practice, $dir \approx 0$).

[3]Note that, while we have chosen a Bernoulli process, independent of $N(t)$, as our means to split $N(t)$, we can further generalize splitting using a multinomial process.

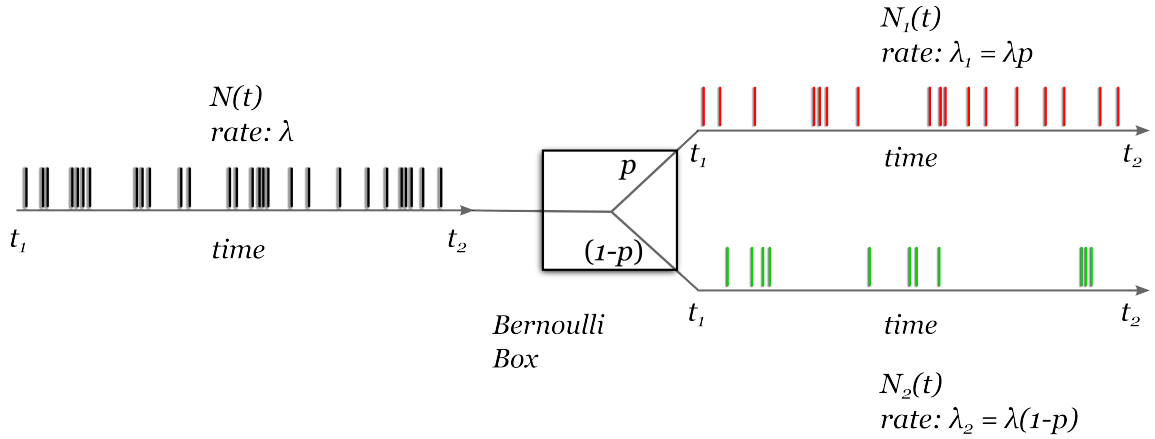


Figure B.1: Splitting a Poisson Process Schematic. Given a Poisson process $N(t)$ with rate λ, we can split the process into two (or more) *independent* sub-processes $N_1(t)$ and $N_2(t)$, with rates λ_1 and λ_2, respectively. How we split the processes, however, must be random. In this example, we run the parent process $N(t)$ through a Bernoulli box, in which photons are randomly selected with probability p to join sub-process $N_1(t)$. The remaining photons are placed in sub-process $N_2(t)$ (equivalently, we could reverse the order, selecting photons to join sub-process $N_2(t)$ with probability $(1-p)$, and the general result would be same). Those photons selected for $N_1(t)$ are colored red, indicating acceptor photons $D_{ex}A_{em}$; for $N_2(t)$, photons are colored green, indicating donor photons $D_{ex}D_{em}$. In the case of FRET, $p = E$.

Appendix C

Burst Variance Analysis

C.1 E-S Covariance

We define the FRET efficiency E and stoichiometry S as

$$E = \frac{F_{D_{exc}}^{Fret}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{Fret}}; \qquad S = \frac{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{Fret}}{F_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{Fret} + F_{A_{exc}}^{D_{em}} + F_{A_{exc}}^{A_{em}}}$$

To simplify notation, let us define $F_{D_{exc}}^{D_{em}}$, $F_{D_{exc}}^{Fret}$, and $F_{A_{exc}}^{A_{em}}$ by the independent Poisson random variables X, Y, and Z, respectively, with rates λ_X, λ_Y, and λ_Z, and omit $F_{A_{exc}}^{D_{em}} \approx 0$. And further, let $g(X,Y) = E$ and $h(X,Y,Z) = S$, where $X+Y+Z > 0$ and $g(X,Y) = 0$ if $Z = N = X+Y+Z$.

Suppose we know the value of random variable $N = X+Y+Z$. Our goal is to find the conditional variance of functions of random variables $g(X,Y)$ and $h(X,Y,Z)$ given N. We begin by defining the conditional covariance:

$$\begin{aligned}\mathrm{Cov}(g(X,Y), h(X,Y,Z)|N) = \mathrm{E}\Big[g(X,Y)h(X,Y,Z)\Big|N\Big] \\ - \mathrm{E}\Big[g(X,Y)\Big|N\Big]\ \mathrm{E}\Big[h(X,Y,Z)\Big|N\Big] \quad \text{(C.1)}\end{aligned}$$

First recall we previously proved a convenient property of Poisson random variables (see Section 5.2). Namely that, given two independent Poisson random vari-

ables X and Y, the distribution of Y conditioned on the sum of X and Y is binomial with parameters $n = X + Y$ and $p = \lambda_Y/(\lambda_X + \lambda_Y)$.

Now, consider $\mathrm{E}[X|N]$ ...

$$\begin{aligned}\mathrm{E}[X|X+Y+Z] &= \mathrm{E}[X|X+W] \\ &= \frac{\lambda_X}{\lambda_X+\lambda_W}(X+W) \\ &= \frac{\lambda_X}{\lambda_X+\lambda_Y+\lambda_Z}(X+Y+Z) \\ &= \frac{\lambda_X}{\lambda_X+\lambda_Y+\lambda_Z}N\end{aligned}$$

where we substituted having recognized that the sum of two independent Poisson random variables $Y + Z$ is also a Poisson random variable, say, W, whose rate is the sum $\lambda_Y + \lambda_Z$.

Now, consider the joint expectation

$$\mathrm{E}\left[g(X,Y)h(X,Y,Z)\middle|N\right] = \mathrm{E}\left[\frac{X}{X+Y+Z}\mathbf{1}_{X+Y+Z\neq 0}\middle|X+Y+Z\right]$$

We can use the property that $\mathrm{E}[aX] = a\,\mathrm{E}[X]$, and hence,

$$\mathrm{E}\left[g(X,Y)h(X,Y,Z)\middle|N\right] = \frac{\lambda_X}{\lambda_X+\lambda_Y+\lambda_Z}\mathbf{1}_{X+Y+Z\neq 0}$$

Now, consider $\mathrm{E}[X+Y|N]$ ...

$$\begin{aligned}E[X+Y|X+Y+Z] &= \mathrm{E}[W|W+Z] \\ &= \frac{\lambda_W}{\lambda_W+\lambda_Z}(W+Z) \\ &= \frac{\lambda_X+\lambda_Y}{\lambda_X+\lambda_Y+\lambda_Z}N\end{aligned}$$

where we made use of another W substitution. Similar to above,

$$\begin{aligned}\mathrm{E}\left[h(X,Y,Z)\middle|N\right] &= \mathrm{E}\left[\frac{X+Y}{X+Y+Z}\mathbf{1}_{X+Y+Z\neq 0}\middle|X+Y+Z\right] \\ &= \frac{\lambda_X+\lambda_Y}{\lambda_X+\lambda_Y+\lambda_Z}\mathbf{1}_{X+Y+Z\neq 0}\end{aligned}$$

Now, one more term remains $\mathrm{E}\left[g(X,Y)\middle|N\right]$. We can use the *Law of Iterated Expectations*

$$\mathrm{E}[X|N] = \mathrm{E}\left[\mathrm{E}[X|M]\middle|N\right]$$

where the value of N is determined by M. In our case, N is determined by Z and $X+Y$, both existing on the same probability space. So

$$\mathrm{E}\left[g(X,Y)\middle|N\right] = \mathrm{E}\left[\mathrm{E}[g(X,Y)|Z,X+Y]\middle|N\right]$$

But $g(X,Y)$ is independent of Z. In which case, focusing on the inner expectation,

$$\begin{aligned}\mathrm{E}\left[g(X,Y)\middle|Z,X+Y\right] &= \mathrm{E}\left[g(X,Y)\middle|X+Y\right] \\ &= \mathrm{E}\left[\frac{X}{X+Y}\mathbf{1}_{X+Y\neq 0}\middle|X+Y\right]\end{aligned}$$

where we have already seen this above.

$$\mathrm{E}\left[\frac{X}{X+Y}\mathbf{1}_{X+Y\neq 0}\middle|X+Y\right] = \frac{\lambda_X}{\lambda_X+\lambda_Y}\mathbf{1}_{X+Y\neq 0}$$

If we insert this result into the outer expectation, we have

$$\mathrm{E}\left[g(X,Y)\middle|N\right] = \mathrm{E}\left[\frac{\lambda_X}{\lambda_X+\lambda_Y}\mathbf{1}_{X+Y\neq 0}\middle|N\right] = \frac{\lambda_X}{\lambda_X+\lambda_Y}\mathbf{1}_{X+Y\neq 0}$$

where we used the property that the expectation of a constant is equal to the constant ($\mathrm{E}[b] = b$).

Putting the three terms together, we arrive at our solution for the covariance:

$$\begin{aligned}\mathrm{Cov}(g(X,Y),h(X,Y,Z)|N) &= \frac{\lambda_X}{\lambda_X+\lambda_Y+\lambda_Z} - \left(\frac{\lambda_X}{\lambda_X+\lambda_Y}\cdot\frac{\lambda_X+\lambda_Y}{\lambda_X+\lambda_Y+\lambda_Z}\right) \\ &= 0 \end{aligned} \tag{C.2}$$

So while the functions $g(X,Y)$ and $h(X,Y,Z)$, i.e., E and S, appear clearly dependent, they are not expected to co-vary[1]. Q.E.D.

This result may seem counter-intuitive, as we can write $E = Y/(X+Y) = Y/NS$. While one can manipulate the equations for stoichiometry (see Equation 5.2) to show that E and S are, in fact, physically *independent*, understanding why the covariance is zero when expressing the equations in terms of random variables remains elusive. But consider the following.

Fix $N \geq 1$ and consider N uncolored photons. Paint each uncolored photon green with probability

$$p_G = \frac{\lambda_X+\lambda_Y}{\lambda_X+\lambda_Y+\lambda_Z}$$

Now repaint color each green photon red with probability

$$p_R = \frac{\lambda_Y}{\lambda_X+\lambda_Y}$$

A moment's thought and we realize that the mean proportion of red photons amongst green photons is independent of the overall proportion of green photons. If we apply this example for N random with Poisson distribution $Poiss(\lambda_N)$, with $\lambda_N = \lambda_X+\lambda_Y+\lambda_Z$, then the number of red, green, and uncolored photons are independent and Poisson with parameters λ_X, λ_Y, and λ_Z, respectively.

Note that the same result holds for measured (uncorrected) E:

[1] If X and Y are independent, then the covariance of X and Y is 0, as independence implies $\mathrm{E}[XY] = \mathrm{E}[X]\cdot\mathrm{E}[Y]$. The converse, as we have just seen, is *not* generally true.

$$E_{raw} = \frac{F_{D_{exc}}^{Fret} + dF_{A_{exc}}^{A_{em}} + lF_{D_{exc}}^{D_{em}}}{\gamma F_{D_{exc}}^{D_{em}} + \underbrace{lF_{D_{exc}}^{D_{em}}}_{\text{leakage}} + F_{D_{exc}}^{Fret} + \underbrace{dF_{A_{exc}}^{A_{em}}}_{\text{direct excitation}}} \tag{C.3}$$

and measured (uncorrected) S:

$$S_{raw} = \frac{F_{D_{exc}}^{Fret} + dF_{A_{exc}}^{A_{em}} + lF_{D_{exc}}^{D_{em}} + \gamma F_{D_{exc}}^{D_{em}}}{\gamma F_{D_{exc}}^{D_{em}} + lF_{D_{exc}}^{D_{em}} + F_{D_{exc}}^{Fret} + dF_{A_{exc}}^{A_{em}} + F_{A_{exc}}^{A_{em}}} \tag{C.4}$$

where $F_{D_{exc}}^{A_{em}}$ is a convolution of FRET photons ($F_{D_{exc}}^{Fret}$), donor photons leaking into the acceptor emission channel ($lF_{D_{exc}}^{D_{em}}$), and direct excitation by the laser exciting the donor ($dF_{A_{exc}}^{A_{em}}$), and γ is the detection-correction factor, accounting for differences in fluorescence quantum yield and emission detection efficiency. If we assume correction factors $d, l,$ andγ are constants, then we may express the uncorrected quantities in terms of random variables as follows:

$$g(X,Y) = \frac{qX + rY}{sX + rY}\,; \qquad h(X,Y,Z) = \frac{sX + rY}{sX + rY + Z} \tag{C.5}$$

with constants $q = l$, $r = a + d$ (a being the fraction of total detected acceptor photons minus leakage from FRET), and $s = l + \gamma$. If we then know $N = sX + rY + Z$, one can proceed as above to find

$$\begin{aligned} \mathrm{Cov}(g(X,Y), h(X,Y,Z)|N) &= \frac{q\lambda_X + r\lambda_Y}{s\lambda_X + r\lambda_Y + \lambda_Z} - \left(\frac{q\lambda_X + r\lambda_Y}{s\lambda_X + r\lambda_Y} \cdot \frac{s\lambda_X + r\lambda_Y}{s\lambda_X + r\lambda_Y + \lambda_Z}\right) \\ &= 0 \end{aligned} \tag{C.6}$$

which holds for a known N.

Appendix D

Materials and Methods

D.1 DNA and Reagents

Amino-modified oligonucleotides (IBA, Germany) were internally labeled, as previously described [50], with fluorophores Cy3B (Invitrogen, USA) and ATTO647N (ATTOTEC, Germany) and purified using gel electrophoresis. Single-stranded DNA (ssDNA) were annealed in hybridization buffer (50mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM NaCl). See Figure D.1 for DNA sequences and labeling schemes.

D.2 RNAP Open Complex Formation

RNAP open complexes (RP_o) were formed according to published procedures [50,167,169,214]. 10 nM dsDNA was mixed with 50 nM *E. coli* RNAP holoenzyme (Epicentre, USA) in KG7 buffer (40 mM HEPES-NaOH pH 7, 100 mM potassium glutamate, 10 mM $MgCl_2$, 1 mM DTT, 100 μg/mL BSA, 5% glycerol, 1 mM mercaptoethylamine) for a total volume of 20 μL. The mixture was incubated for 15 minutes at 37°C. After incubation, 1mg/mL heparin sepharose (GE Healthcare) was added to disrupt non-specific RNAP-DNA complexes and to remove free RNAP. After incubating for 30 seconds at 37°C, samples were centrifuged, and 13μL of supernatant was transferred to a pre-warmed tube.

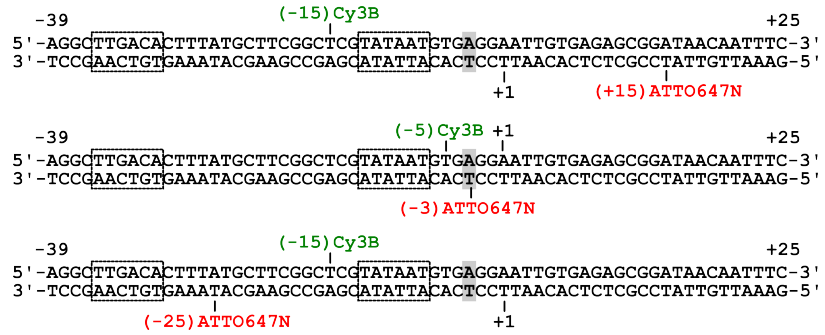


Figure D.1: DNA Sequences and Labeling Schemes. The donor fluorophore Cy3B is green, and the acceptor fluorophore ATTO647N is red. Base pair substitutions different from the lacCONS+2 sequence are grey. The boxes indicate the -35 and -10 consensus sequence promoter elements.

D.3 Single-molecule Fluorescence Spectroscopy

D.3.1 Confocal Microscopy

A custom-built confocal microscope was used for single-molecule FRET experiments as described previously [276,277]. The setup was modified to allow alternating-laser excitation (ALEx) of donor and acceptor fluorophores [167, 168]. The fiber-coupled outputs of a 532 nm green (Samba, Cobolt, Sweden) and 638 nm red (Cube Coherent, USA) laser were alternated with a 10 kHz modulation frequency. Both beams were spatially filtered and coupled into an IX71 inverted confocal microscope (Olympus, Germany). Average excitation intensities were 250 μW at 532 nm and 60 μW at 638 nm. Fluorescence was collected through the same objective as excitation. Fluorescence emission was separated from excitation light by a dichroic mirror (630 DRLP, Chroma). The separated emission was then focused onto a 200 μm pinhole and subsequently spectrally split with two distinct filters (green: 585DF70; red: 650LP) onto two SPCM-AQR-14 avalanche photodiodes (APDs; PerkinElmer, UK). Photon arrival times were recorded by a PCI-6602 board (National Instruments). Custom-made LabVIEW software was used to register and evaluate the detected signal. For all experiments, the sample temperature was 37$\pm$1°C via a custom-made heated objective collar connected to a heating bath.

D.3.2 Total Internal Reflection Fluorescence Microscopy

Imaging was performed on a custom-built total internal reflection fluorescence (TIRF) microscope using ALEX [167,168], between a directly modulated diode 635 nm (red) laser (Cube Coherent, USA) and a continuous-wave solid state 532 nm (green) laser (Samba, Cobolt, Sweden) modulated using an acousto-optical modulator (AA Optics). The alternation period was 5 ms in all experiments used for analysis, unless otherwise indicated. The green and red laser beams were coupled into the same single-mode optical fiber. At the fiber output, the laser beams were collimated, directed into an IX71 inverted microscope (Olympus, Germany) and focused (100× oil-immersion objective, numerical aperture (NA) 1.4, Olympus, Germany) onto the sample under an angle allowing for total internal reflection. The fluorescence emission was collected through the same objective, separated from excitation light by a dichroic mirror (545 nm/650 nm, Semrock) and additional filters (545 nm LP, Chroma; and 633/25 nm notch filter, Semrock). To allow dual-color detection, the emission light was split into red and green channels (630 nm DRLP; Omega) and projected onto two separate regions of an electron-multiplying charge-coupled device (EMCCD) camera (iXon, Andor). Laser powers used were 1.5 mW (532 nm) and 0.5 mW (635 nm) for Cy3B and ATTO647N (intensities measured before the beams entered the TIRF objective). Surfaces were passivated with a biotin and biotin-PEG mixture, incubated with NeutrAvidin, and then incubated with antibody (Qiagen), as specified previously [155]. Heparin resistant RP_o complexes were then deposited.

D.4 Confocal Data Analysis

Fluorescence photons were assigned to donor (D_{ex}) or acceptor (A_{ex}) excitation based on photon arrival time and to donor (D_{em}) or acceptor (A_{em}) emission based on the detection channel [276,277]. Bursts were identified via a Seidel burst search [73]. Bursts were characterized by three parameters ($M,\ T,$ and L). A pho-

ton sequence is considered a burst if a total of L photons, having M neighbor photons, arrive at the detector within a time interval of T seconds. Unless otherwise indicated, the burst search was performed on the $A_{ex}A_{em}$ photon stream with parameters $M = 7,\ T = 500\mu\text{s},$ and $L = 10$. Subsequent filtering after the burst search selected only bright donor-acceptor molecules ($A_{ex}A_{em} > 30$ photons).

Two characteristic ratios, fluorophore stoichiometry S and apparent FRET efficiency E^*, were calculated for each burst which met the applied thresholds. S is the ratio of D_{ex} fluorescence to total D_{ex} and A_{ex} fluorescence. S describes the ratio of donor-to-acceptor fluorophores and the relative fluorophore brightness. Apparent, i.e., uncorrected, E is commonly employed to measure fluorophore proximity. A 1-dimensional E^* distribution for donor-acceptor species was obtained after applying a threshold on S (typically $S > 0.45$) which removed acceptor-only species. These E^* distributions could be fit with a Gaussian function to determine the distribution mean and standard deviation.

D.5 Tirf Data Analysis

To extract single-molecule intensity time series, TIRF movies were analyzed with custom written software in MATLAB® (Mathworks) by applying methods adapted from astronomy [135, 150]; the software is publicly available. The red and green emission channels were aligned based on calibration movies of immobilized fluorescence beads. The emission of single immobilized labeled DNA molecules in the red and green channel was automatically detected and linked using a peak-finding algorithm [52]. Peaks that did not meet ellipticity or nearest-neighbor criteria were excluded from further analysis ($\varepsilon < 0.25, NN > 8$ pixels), thereby minimizing artifacts resulting from molecules with overlapping point spread functions. The emission intensity of each molecule was estimated by fitting a two-dimensional Gaussian to its point spread function. This yields $F_{D_{exc}}^{D_{em}}$ (donor emission upon donor excitation), $F_{D_{exc}}^{A_{em}}$ (acceptor emission upon donor excitation) and $F_{A_{exc}}^{A_{em}}$ (acceptor emission upon acceptor excitation). These intensities were divided by an

EMCCD effective gain factor of 6 to obtain photon counts.

Bibliography

[1] H. Abdi. The Bonferonni and Šidák Corrections for Multiple Comparisons. *Cognition*, 1:1–9, 2007.

[2] C. E. Aitken, R. A. Marshall, and J. D. Puglisi. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical Journal*, 94(5):1826–35, 2008.

[3] M. Andrec, R. M. Levy, and D. S. Talaga. Direct Determination of Kinetic Rates from Single-Molecule Photon Arrival Trajectories Using Hidden Markov Models. *The Journal of Physical Chemistry A*, 107(38):7454–7464, 2003.

[4] D. L. Andrews and D. S. Bradshaw. Virtual photons, dipole fields and energy transfer: a quantum electrodynamical approach. *European Journal of Physics*, 25(6):845–858, 2004.

[5] M. Antonik, S. Felekyan, A. Gaiduk, and C. A. M. Seidel. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *The Journal of Physical Chemistry B*, 110(13):6970–6978, 2006.

[6] S. R. Aragón and R. Pecora. Fluorescence correlation spectroscopy and Brownian rotational diffusion. *Biopolymers*, 14(1):119–137, 1975.

[7] S. R. Aragon. Fluorescence correlation spectroscopy as a probe of molecular dynamics. *The Journal of Chemical Physics*, 64(4):1791, Feb. 1976.

[8] P. Atkins, T. Overton, J. Rourke, M. Weller, and F. Armstrong. *Inorganic Chemistry*. Oxford University Press, Oxford, fourth edition, 2006.

[9] A. Auerbach and F. Sachs. Patch clamp studies of single ionic channels. *Annual review of biophysics and bioengineering*, 13:269–302, Jan. 1984.

[10] D. Axelrod. Total internal reflection fluorescence microscopy in cell biology. *Traffic Copenhagen Denmark*, 2(11):764–774, 2001.

[11] P. R. Barber, S. M. Ameer-Beg, S. Pathmananthan, M. Rowley, and A. C. C. Coolen. A Bayesian method for single molecule, fluorescence burst analysis. *Biomedical optics express*, 1(4):1148–1158, Jan. 2010.

[12] M. Bates, T. R. Blosser, and X. Zhuang. Short-range spectroscopic ruler based on a single-molecule optical switch. *Physical Review Letters*, 94(10):108101, 2005.

[13] M. Bates, B. Huang, G. T. Dempsey, and X. Zhuang. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, 317(5845):1749–1753, 2007.

[14] L. E. Baum. An inequality and an associated maximization technique in statistical estimation of probabilistic functions of a Markov process. *Inequalities*, 3(1):1–8, 1972.

[15] L. E. Baum and J. A. Eagon. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73(3):360–363, 1967.

[16] L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics*, 37(6):1554–1563, 1966.

[17] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1):164–171, 1970.

[18] M. J. Beal. *Variational algorithms for approximate Bayesian inference*. PhD thesis, University College London, 2003.

[19] J. F. Beausang and P. C. Nelson. Diffusive hidden Markov model characterization of DNA looping dynamics in tethered particle experiments. *Physical biology*, 4(3):205–19, Sept. 2007.

[20] R. M. Berry. Absence of a barrier to backwards rotation of the bacterial flagellar motor demonstrated with optical tweezers. *Proceedings of the National Academy of Sciences*, 94(26):14433–14437, Dec. 1997.

[21] M. Blanco and N. G. Walter. *Analysis of Complex Single-Molecule FRET Time Trajectories*, volume 472 of *Methods in Enzymology*. Elsevier Inc., 2010.

[22] M. Böhmer, M. Wahl, H.-J. Rahn, R. Erdmann, and J. Enderlein. Time-resolved fluorescence correlation spectroscopy. *Chemical Physics Letters*, 353(5-6):439–445, Feb. 2002.

[23] M. Börsch and J. Wrachtrup. Improving FRET-based monitoring of single chemomechanical rotary motors at work. *Chemphyschem A European Journal Of Chemical Physics And Physical Chemistry*, 12(3):542–553, 2011.

[24] E. Boukobza, A. Sonnenfeld, and G. Haran. Immobilization in Surface-Tethered Lipid Vesicles as a New Tool for Single Biomolecule Spectroscopy. *The Journal of Physical Chemistry B*, 105(48):12165–12170, 2001.

[25] J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, Jr., and C. H. Wiggins. Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *Biophysical Journal*, 97(12):3196–3205, 2009.

[26] J. E. Bronson, J. M. Hofman, J. Fei, R. L. Gonzalez, and C. H. Wiggins. Graphical models for inferring single molecule dynamics. *BMC Bioinformatics*, 11(Suppl 8):S2, 2010.

[27] R. Brune, S. Doose, and M. Sauer. Analyzing the influence of contact-induced quenching processes on Förster resonance energy transfer. In *Proceedings of SPIE*, volume 6633, pages 66331M–66331M–11. SPIE, June 2007.

[28] H. Buc and T. Strick. *RNA Polymerases as Molecular Motors*. Royal Society of Chemistry, 2009.

[29] S. Camyshan, N. Gritsan, V. Korolev, and N. Bazhin. Quenching of the luminescence of organic compounds by oxygen in glassy matrices. *Chemical Physics*, 142(1):59–68, Mar. 1990.

[30] A. J. Carpousis, J. E. Stefano, and J. D. Gralla. 5′ nucleotide heterogeneity and altered initiation of transcription at mutant lac promoters. *Journal of Molecular Biology*, 157(4):619–633, 1982.

[31] A. Chakraborty, D. Wang, Y. W. Ebright, and R. H. Ebright. Azide-specific labeling of biomolecules by Staudinger-Bertozzi ligation phosphine derivatives of fluorescent probes suitable for single-molecule fluorescence spectroscopy. *Methods in enzymology*, 472(null):19–30, Jan. 2010.

[32] A. Chakraborty, D. Wang, Y. W. Ebright, Y. Korlann, E. Kortkhonjia, T. Kim, S. Chowdhury, S. Wigneshweraraj, H. Irschik, R. Jansen, B. T. Nixon, J. Knight, S. Weiss, and R. H. Ebright. Opening and Closing of the Bacterial RNA Polymerase Clamp. *Science*, 337(6094):591–595, Aug. 2012.

[33] J. Chen, S. A. Darst, and D. Thirumalai. Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28):12523–8, July 2010.

[34] Y. Chen, J. D. Müller, K. M. Berland, and E. Gratton. Fluorescence fluctuation spectroscopy. *Methods (San Diego, Calif.)*, 19(2):234–52, Oct. 1999.

[35] Y. Chen, J. D. Müller, P. T. So, and E. Gratton. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophysical Journal*, 77(1):553–567, 1999.

[36] Y. Chen, M. Tekmen, L. Hillesheim, J. Skinner, B. Wu, and J. D. Müller. Dual-color photon-counting histogram. *Biophysical journal*, 88(3):2177–92, Mar. 2005.

[37] Y. F. Chen and J. D. Helmann. DNA-melting at the Bacillus subtilis flagellin promoter nucleates near -10 and expands unidirectionally. *Journal of molecular biology*, 267(1):47–59, Mar. 1997.

[38] D. I. Cherny, I. C. Eperon, and C. R. Bagshaw. Probing complexes with single fluorophores: factors contributing to dispersion of FRET in DNA/RNA duplexes. *European biophysics journal EBJ*, 38(4):395–405, 2009.

[39] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, and P. G. Schultz. Addition of p -Azido- l -phenylalanine to the Genetic Code of Escherichia c oli. *Journal of the American Chemical Society*, 124(31):9026–9027, Aug. 2002.

[40] P. J. Choi, L. Cai, K. Frieda, and X. S. Xie. A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science (New York, N.Y.)*, 322(5900):442–6, Oct. 2008.

[41] Y. Choi, I. S. Moody, P. C. Sims, S. R. Hunt, B. L. Corso, I. Perez, G. A. Weiss, and P. G. Collins. Single-molecule lysozyme dynamics monitored by an electronic circuit. *Science (New York, N.Y.)*, 335(6066):319–24, Jan. 2012.

[42] H. S. Chung, I. V. Gopich, K. McHale, T. Cellmer, J. M. Louis, and W. A. Eaton. Extracting rate coefficients from single-molecule photon trajectories and FRET efficiency histograms for a fast-folding protein. *The journal of physical chemistry A*, 115(16):3642–3656, 2011.

[43] H. S. Chung, J. M. Louis, and W. A. Eaton. Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories. *Proceedings of the National Academy of Sciences of the United States of America*, 106(29):11837–11844, 2009.

[44] H. S. Chung, J. M. Louis, and W. A. Eaton. Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments. *Biophysical Journal*, 98(4):696–706, 2010.

[45] S. H. Chung, V. Krishnamurthy, and J. B. Moore. Adaptive processing techniques based on hidden Markov models for characterizing very small channel currents buried in noise and deterministic interferences. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, 334(1271):357–384, 1991.

[46] S. H. Chung, J. B. Moore, L. G. Xia, L. S. Premkumar, and P. W. Gage. Characterization of single channel currents using digital signal processing techniques based on Hidden Markov Models. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, 329(1254):265–285, 1990.

[47] R. M. Clegg. Fluorescence resonance energy transfer and nucleic acids. *Methods in Enzymology*, 211(1971):353–388, 1992.

[48] N. R. Conley, J. S. Biteen, and W. E. Moerner. Cy3-Cy5 covalent heterodimers for single-molecule photoswitching. *The journal of physical chemistry. B*, 112(38):11878–80, Sept. 2008.

[49] M. Cooper, A. Ebner, M. Briggs, M. Burrows, N. Gardner, R. Richardson, and R. West. Cy3B: improving the performance of cyanine dyes. *Journal of fluorescence*, 14(2):145–50, Mar. 2004.

[50] T. Cordes, Y. Santoso, A. I. Tomescu, K. Gryte, L. C. Hwang, B. Camará, S. Wigneshweraraj, and A. N. Kapanidis. Sensing DNA opening in transcription using quenchable Forster resonance energy transfer. *Biochemistry*, 49(43):9171–9180, 2010.

[51] T. Cordes, J. Vogelsang, and P. Tinnefeld. On the mechanism of Trolox as antiblinking and antibleaching reagent. *Journal of the American Chemical Society*, 131(14):5018–5019, 2009.

[52] J. Crocker. Methods of Digital Video Microscopy for Colloidal Studies. *Journal of Colloid and Interface Science*, 179(1):298–310, Apr. 1996.

[53] M. Dahan, A. A. Deniz, T. Ha, D. S. Chemla, P. G. Schultz, and S. Weiss. Ratiometric measurement and identification of single diffusing molecules. *Chemical Physics*, 247(1):85–106, 1999.

[54] R. E. Dale, J. Eisinger, and W. E. Blumberg. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophysical Journal*, 26(2):161–193, 1979.

[55] R. P. das Neves, N. S. Jones, L. Andreu, R. Gupta, T. Enver, and F. J. Iborra. Connecting variability in global transcription rate to mitochondrial variability. *PLoS biology*, 8(12):e1000560, Jan. 2010.

[56] M. Daune. *Molecular Biophysics: Structures in Motion*. OUP Oxford, 1999.

[57] C. A. Davis, C. A. Bingman, R. Landick, M. T. Record, and R. M. Saecker. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by Escherichia coli RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 104(19):7833–8, May 2007.

[58] M. H. De Smit and J. Van Duin. Control of prokaryotic translational initiation by mRNA secondary structure. *Progress in nucleic acid research and molecular biology*, 38:1–35, 1990.

[59] M. H. De Smit and J. Van Duin. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 87(19):7668–7672, 1990.

[60] P. L. DeHaseth, M. L. Zupancic, and M. T. Record. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *Journal of bacteriology*, 180(12):3019–25, June 1998.

[61] G. T. Dempsey, M. Bates, W. E. Kowtoniuk, D. R. Liu, R. Y. Tsien, and X. Zhuang. Photoswitching Mechanism of Cyanine Dyes. *Journal of the American Chemical Society*, 131(51):18192–18193, 2009.

[62] A. Dempster, N. Laird, and D. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1 – 38, 1977.

[63] A. A. Deniz, T. A. Laurence, M. Dahan, D. S. Chemla, P. G. Schultz, and S. Weiss. Ratiometric single-molecule studies of freely diffusing biomolecules. *Annual Review of Physical Chemistry*, 52(January):233–253, 2001.

[64] A. A. Deniz, S. Mukhopadhyay, and E. A. Lemke. Single-molecule biophysics: at the interface of biology, physics and chemistry. *Journal of the Royal Society, Interface / the Royal Society*, 5(18):15–45, Jan. 2008.

[65] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, volume 16 of *Classics in Applied Mathematics*. Prentice-Hall, 1983.

[66] M. C. DeSantis, S. H. DeCenzo, J.-L. Li, and Y. M. Wang. Precision analysis for standard deviation measurements of immobile single fluorescent molecule images. *Optics express*, 18(7):6563–76, Mar. 2010.

[67] D. L. Dexter. A Theory of Sensitized Luminescence in Solids. *The Journal of Chemical Physics*, 21(5):836, 1953.

[68] N. Di Fiori and A. Meller. Automated system for single molecule fluorescence measurements of surface-immobilized biomolecules. *Journal of visualized experiments : JoVE*, (33), Jan. 2009.

[69] N. Di Fiori and A. Meller. The Effect of dye-dye interactions on the spatial resolution of single-molecule FRET measurements in nucleic acids. *Biophysical journal*, 98(10):2265–72, May 2010.

[70] S. Doose, H. Neuweiler, and M. Sauer. Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules. *Chemphyschem A European Journal Of Chemical Physics And Physical Chemistry*, 10(9-10):1389–1398, 2009.

[71] C. G. Dos Remedios and P. D. Moens. Fluorescence resonance energy transfer spectroscopy is a reliable "ruler" for measuring structural changes in proteins. Dispelling the problem of the unknown orientation factor. *Journal of Structural Biology*, 115(2):175–185, 1995.

[72] L. Edman and R. Rigler. Memory landscapes of single-enzyme molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 97(15):8266–71, July 2000.

[73] C. Eggeling, S. Berger, L. Brand, J. R. Fries, J. Schaffer, A. Volkmer, and C. A. Seidel. Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *Journal of biotechnology*, 86(3):163–180, 2001.

[74] C. Eggeling, M. Hilbert, H. Bock, C. Ringemann, M. Hofmann, A. C. Stiel, M. Andresen, S. Jakobs, A. Egner, A. Schönle, and S. W. Hell. Reversible photoswitching enables single-molecule fluorescence fluctuation spectroscopy at high molecular concentration. *Microscopy Research and Technique*, 70(12):1003–1009, 2007.

[75] C. Eggeling, C. Ringemann, R. Medda, G. Schwarzmann, K. Sandhoff, S. Polyakova, V. N. Belov, B. Hein, C. von Middendorff, A. Schönle, and S. W. Hell. Direct observation of the nanoscale dynamics of membrane lipids in a living cell. *Nature*, 457(7233):1159–62, Feb. 2009.

[76] C. Eggeling, J. Widengren, L. Brand, J. Schaffer, S. Felekyan, and C. A. M. Seidel. Analysis of photobleaching in single-molecule multicolor excitation and Förster resonance energy transfer measurements. *The journal of physical chemistry A*, 110(9):2979–95, 2006.

[77] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. Dewinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, and S. Turner. Real-time DNA sequencing from single polymerase molecules. *Science (New York, N.Y.)*, 323(5910):133–8, Jan. 2009.

[78] M. Eigen. Sorting Single Molecules: Application to Diagnostics and Evolutionary Biotechnology. *Proceedings of the National Academy of Sciences*, 91(13):5740–5747, June 1994.

[79] E. L. Elson and D. Magde. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[80] S. A. Emory, P. Bouvet, and J. G. Belasco. A 5'-terminal stem-loop structure can stabilize mRNA in Escherichia coli. *Genes & Development*, 6(1):135–148, 1992.

[81] J. Enderlein, P. M. Goodwin, A. Van Orden, W. Patrick Ambrose, R. Erdmann, and R. A. Keller. A maximum likelihood estimator to distinguish single molecules by their fluorescence decays. *Chemical Physics Letters*, 270(5-6):464–470, May 1997.

[82] J. Enderlein, D. L. Robbins, W. P. Ambrose, and R. A. Keller. Molecular Shot Noise, Burst Size Distribution, and Single-Molecule Detection in Fluid Flow: Effects of Multiple Occupancy. *The Journal of Physical Chemistry A*, 102(30):6089–6094, July 1998.

[83] A. P. Field. Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological methods*, 6(2):161–180, 2001.

[84] L. Finzi and D. D. Dunlap. Single-molecule Approaches to Probe the Structure, Kinetics, and Thermodynamics of Nucleoprotein Complexes That Regulate Transcription. *The Journal of Biological Chemistry*, 285(25):18973–18978, 2010.

[85] C. J. Fischer, N. K. Maluf, and T. M. Lohman. Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA. *Journal of molecular biology*, 344(5):1287–309, 2004.

[86] D. L. Floyd, S. C. Harrison, and A. M. Van Oijen. Analysis of Kinetic Intermediates in Single-Particle Dwell-Time Distributions. *Biophysical Journal*, 99(2):360–366, 2010.

[87] G. Forney. The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278, 1973.

[88] T. Förster. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik*, 437(1-2):55–75, 1948.

[89] T. Förster. Experimentelle und theoretische Untersuchung des zwischenmolekularen Uebergangs von Elektronenanregungsenergie. *Zeitschrift für Naturforschung*, 4(5):321–327, 1949.

[90] T. Forster. 10th Spiers Memorial Lecture. Transfer mechanisms of electronic excitation. *Discussions Of The Faraday Society*, 27(10):7, 1959.

[91] T. Förster. Transfer Mechanisms of Electronic Excitation Energy. *Radiation Research Supplement*, 2(10):326, 1960.

[92] T. Forster. Delocalized excitation and excitation transfer. In O. Sunanoglu, editor, *Modern Quantum Chemistry*, pages 93–137. Academic Press, New York, 1965.

[93] L. J. Friedman and J. Gelles. Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell*, 148(4):679–89, 2012.

[94] J. R. Fries, L. Brand, C. Eggeling, M. Köllner, and C. A. M. Seidel. Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy. *The Journal of Physical Chemistry A*, 102(33):6601–6613, Aug. 1998.

[95] T. Funatsu, Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, 374(6522):555–559, 1995.

[96] A. Gaiduk, M. Yorulmaz, P. V. Ruijgrok, and M. Orrit. Room-temperature detection of a single molecule's absorption by photothermal contrast. *Science (New York, N.Y.)*, 330(6002):353–6, Oct. 2010.

[97] C. Gell, D. Brockwell, and A. Smith. *Handbook of Single Molecule Fluorescence Spectroscopy*. OUP Oxford, 2006.

[98] T. Ghosh, D. Bose, and X. Zhang. Mechanisms for activating bacterial RNA polymerase. *FEMS microbiology reviews*, 34(5):611–27, Sept. 2010.

[99] S. R. Goldman, R. H. Ebright, and B. E. Nickels. Direct detection of abortive RNA transcripts in vivo. *Science*, 324(5929):927–928, 2009.

[100] I. Gopich and A. Szabo. Fluorophore-quencher distance correlation functions from single-molecule photon arrival trajectories. *The journal of physical chemistry. B*, 109(14):6845–8, Apr. 2005.

[101] I. Gopich and A. Szabo. Theory of photon statistics in single-molecule Förster resonance energy transfer. *The Journal of chemical physics*, 122(1):14707, 2005.

[102] I. V. Gopich. Likelihood functions for the analysis of single-molecule binned photon sequences. *Chemical Physics*, 396:53–60, Mar. 2012.

[103] I. V. Gopich and A. Szabo. Single-molecule FRET with diffusion and conformational dynamics. *The Journal of Physical Chemistry B*, 111(44):12925–12932, 2007.

[104] I. V. Gopich and A. Szabo. Decoding the pattern of photon colors in single-molecule FRET. *The Journal of Physical Chemistry B*, 113(31):10965–10973, 2009.

[105] I. V. Gopich and A. Szabo. FRET efficiency distributions of multistate single molecules. *The journal of physical chemistry B*, 114(46):15221–15226, 2010.

[106] W. Grange, P. Haas, A. Wild, M. A. Lieb, M. Calame, M. Hegner, and B. Hecht. Detection of transient events in the presence of background noise. *The journal of physical chemistry. B*, 112(23):7140–4, June 2008.

[107] M. Greenfeld, D. S. Pavlichin, H. Mabuchi, and D. Herschlag. Single Molecule Analysis Research Tool (SMART): An Integrated Approach for Analyzing Single Molecule Data. *PLoS ONE*, 7(2):e30024, 2012.

[108] T. J. Gries, W. S. Kontur, M. W. Capp, R. M. Saecker, and M. T. Record. One-step DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. *Proceedings of the National Academy of Sciences of the United States of America*, 107(23):10418–10423, 2010.

[109] D. J. Griffiths. *Introduction to Quantum Mechanics (2nd Edition)*. Benjamin Cummings, 2004.

[110] T. Ha, T. Enderle, D. S. Chemla, and S. Weiss. Dual-molecule spectroscopy: molecular rulers for the study of biological macromolecules, 1996.

[111] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proceedings of the National Academy of Sciences of the United States of America*, 93(13):6264–8, 1996.

[112] T. Ha, I. Rasnik, W. Cheng, H. P. Babcock, G. H. Gauss, T. M. Lohman, and S. Chu. Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature*, 419(6907):638–41, Oct. 2002.

[113] T. Ha and P. Tinnefeld. Photophysics of Fluorescent Probes for Single-Molecule Biophysics and Super-Resolution Imaging. *Annual Review of Physical Chemistry*, (January):1–23, 2012.

[114] T. Ha and J. Xu. Photodestruction intermediates probed by an adjacent reporter molecule. *Physical review letters*, 90(22):223002, June 2003.

[115] A. R. Hafdahl and M. A. Williams. Meta-analysis of correlations revisited: attempted replication and extension of Field's (2001) simulation studies. *Psychological methods*, 14(1):24–42, 2009.

[116] M. Hajdziona and A. Molski. Maximum likelihood-based analysis of single-molecule photon arrival trajectories. *The Journal of chemical physics*, 134(5):054112, 2011.

[117] J. A. Hanson and H. Yang. A general statistical test for correlations in a finite-length time series. *The Journal of chemical physics*, 128(21):214101, 2008.

[118] J. A. Hanson and H. Yang. Quantitative evaluation of cross correlation between two finite-length time series with applications to single-molecule FRET. *The journal of physical chemistry B*, 112(44):13962–13970, 2008.

[119] Y. Harada, K. Sakurada, T. Aoki, D. D. Thomas, and T. Yanagida. Mechanochemical coupling in actomyosin energy transduction studied by in vitro movement assay. *Journal of Molecular Biology*, 216(1):49–68, 1990.

[120] S. P. Haugen, W. Ross, and R. L. Gourse. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature reviewsMicrobiology*, 6(June):507–519, 2008.

[121] E. Haustein and P. Schwille. Fluorescence correlation spectroscopy: novel variations of an established technique. *Annual Review of Biophysics and Biomolecular Structure*, 36(1):151–169, 2007.

[122] E. Hecht. *Optics (4th Edition)*. Addison-Wesley, 2001.

[123] L. V. Hedges and I. Olkin. *Statistical methods for meta-analysis*, volume 20. Academic Press, 1985.

[124] M. Heilemann, E. Margeat, R. Kasper, M. Sauer, and P. Tinnefeld. Carbocyanine dyes as efficient reversible single-molecule optical switch. *Journal of the American Chemical Society*, 127(11):3801–3806, 2005.

[125] M. Heilemann, S. Van De Linde, A. Mukherjee, and M. Sauer. Super-resolution imaging with small organic fluorophores. *Angewandte Chemie International Edition*, 48(37):6903–6908, 2009.

[126] M. Heilemann, S. van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie (International ed. in English)*, 47(33):6172–6, Jan. 2008.

[127] T. Heinlein, J.-P. Knemeyer, O. Piestert, and M. Sauer. Photoinduced Electron Transfer between Fluorescent Dyes and Guanosine Residues in DNA-Hairpins. *The Journal of Physical Chemistry B*, 107(31):7957–7964, Aug. 2003.

[128] S. Henkenjohann and M. Sauer. Photoinduced electron transfer probes for the observation of enzyme activities. In *Proceedings of SPIE*, volume 7185, pages 718503–718503–10. SPIE, Feb. 2009.

[129] S. T. Hess and W. W. Webb. Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophysical journal*, 83(4):2300–17, Oct. 2002.

[130] L. N. Hillesheim and J. D. Müller. The photon counting histogram in fluorescence fluctuation spectroscopy with non-ideal photodetectors. *Biophysical journal*, 85(3):1948–58, Sept. 2003.

[131] L. N. Hillesheim and J. D. Müller. The dual-color photon counting histogram with non-ideal photodetectors. *Biophysical journal*, 89(5):3491–507, Nov. 2005.

[132] A. Hoffmann, D. Nettels, J. Clark, A. Borgia, S. E. Radford, J. Clarke, and B. Schuler. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Physical Chemistry Chemical Physics*, 13(5):1857–1871, 2011.

[133] J. Hohlbein, K. Gryte, M. Heilemann, and A. N. Kapanidis. Surfing on a new wave of single-molecule fluorescence methods. *Physical Biology*, 7(3):031001, 2010.

[134] S. J. Holden. *Improved methods for sub-diffraction-limit single-molecule fluorescence measurements*. Phd thesis, University of Oxford, 2011.

[135] S. J. Holden, S. Uphoff, J. Hohlbein, D. Yadin, L. Le Reste, O. J. Britton, and A. N. Kapanidis. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophysical Journal*, 99(9):3102–11, 2010.

[136] M. Honda, J. Park, R. A. Pugh, T. Ha, and M. Spies. Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. *Molecular Cell*, 35(5):694–703, 2009.

[137] R. Horn and K. Lange. Estimating kinetic constants from single channel data. *Biophysical Journal*, 43(2):207–223, 1983.

[138] S. B. Howell. Two-dimensional aperture photometry - Signal-to-noise ratio of point-source observations and optimal data-extraction techniques. *Publications of the Astronomical Society of the Pacific*, 101:616–622, 1989.

[139] S. B. Howell. Handbook of CCD Astronomy. *Library*, page 222, 2006.

[140] L. M. Hsu. Monitoring abortive initiation. *Methods San Diego Calif*, 47(1):25–36, 2009.

[141] L. M. Hsu, N. V. Vo, C. M. Kane, and M. J. Chamberlin. In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry*, 42(13):3777–3786, 2003.

[142] Z. Huang, D. Ji, S. Wang, A. Xia, F. Koberling, M. Patting, and R. Erdmann. Spectral identification of specific photophysics of Cy5 by means of ensemble and single molecule measurements. *The Journal of Physical Chemistry A*, 110(1):45–50, 2006.

[143] Z. Huang, D. Ji, A. Xia, F. Koberling, M. Patting, and R. Erdmann. Direct observation of delayed fluorescence from a remarkable back-isomerization in Cy5. *Journal of the American Chemical Society*, 127(22):8064–8066, 2005.

[144] B. P. Hudson, J. Quispe, S. Lara-González, Y. Kim, H. M. Berman, E. Arnold, R. H. Ebright, and C. L. Lawson. Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proceedings of the National Academy of Sciences of the United States of America*, 106(47):19830–5, Nov. 2009.

[145] P. A. Hunt and M. A. Robb. Systematic control of photochemistry: the dynamics of photoisomerization of a model cyanine dye. *Journal of the American Chemical Society*, 127(15):5720–6, Apr. 2005.

[146] C. G. Hubner, A. Renn, I. Renge, and U. P. Wild. Direct observation of the triplet lifetime quenching of single dye molecules by molecular oxygen. *The Journal of Chemical Physics*, 115(21):9619, Dec. 2001.

[147] H. Hwang, H. Kim, and S. Myong. Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, 108(18):7414–8, May 2011.

[148] A. Iqbal, S. Arslan, B. Okumus, T. J. Wilson, G. Giraud, D. G. Norman, T. Ha, and D. M. J. Lilley. Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32):11176–81, 2008.

[149] A. Iqbal, L. Wang, K. C. Thompson, D. M. J. Lilley, and D. G. Norman. The structure of cyanine 5 terminally attached to double-stranded DNA: implications for FRET studies. *Biochemistry*, 47(30):7857–7862, 2008.

[150] M. Irwin. Detectors and Data Analysis Techniques for Wide Field Optical Imaging. *Instrumentation for large telescopes VII Canary*, 1997.

[151] Y. Ishitsuka, B. Okumus, S. Arslan, K. H. Chen, and T. Ha. Temperature-independent porous nanocontainers for single-molecule fluorescence studies. *Analytical chemistry*, 82(23):9694–9701, 2010.

[152] S. Ivanchenko and D. C. Lamb. Fluorescence Correlation Spectroscopy: Principles and Developments. In J. Brnjas-Kraljević and G. Pifat-Mrzljak, editors, *Supramolecular Structure and Function 10*, chapter 10, pages 1–30. Springer Netherlands, Dordrecht, 2011.

[153] M. Jäger, A. Kiel, D.-P. Herten, and F. A. Hamprecht. Analysis of single-molecule fluorescence spectroscopic data with a Markov-modulated Poisson process. *Chemphyschem a European journal of chemical physics and physical chemistry*, 10(14):2486–2495, 2009.

[154] A. Jain, R. Liu, B. Ramani, E. Arauz, Y. Ishitsuka, K. Ragunathan, J. Park, J. Chen, Y. K. Xiang, and T. Ha. Probing cellular protein complexes using single-molecule pull-down. *Nature*, 473(7348):484–488, 2011.

[155] A. Jain, R. Liu, Y. K. Xiang, and T. Ha. Single-molecule pull-down for studying protein interactions. *Nature protocols*, 7(3):445–52, 2012.

[156] E. A. Jares-Erijman and T. M. Jovin. FRET imaging. *Nature biotechnology*, 21(11):1387–95, Nov. 2003.

[157] T. L. Jennings, J. C. Schlatterer, M. P. Singh, N. L. Greenbaum, and G. F. Strouse. NSET molecular beacon analysis of hammerhead RNA substrate binding and catalysis. *Nano letters*, 6(7):1318–24, July 2006.

[158] W. Jeong and C. Kang. Start site selection at lacUV5 promoter affected by the sequence context around the initiation sites. *Nucleic Acids Research*, 22(22):4667–4672, 1994.

[159] K. Jia, Y. Wan, A. Xia, S. Li, F. Gong, and G. Yang. Characterization of photoinduced isomerization and intersystem crossing of the cyanine dye Cy3. *The Journal of Physical Chemistry A*, 111(9):1593–1597, 2007.

[160] C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha. Advances in single-molecule fluorescence methods for molecular biology. *Annual Review of Biochemistry*, 77(1):51–76, 2008.

[161] S. Jung and R. M. Dickson. Hidden markov analysis of short single molecule intensity trajectories. *The Journal of Physical Chemistry B*, 113(42):13886–13890, 2009.

[162] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature reviews. Genetics*, 6(6):451–64, June 2005.

[163] S. Kalinin, S. Felekyan, M. Antonik, and C. A. M. Seidel. Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer. *The journal of physical chemistry. B*, 111(34):10253–62, Aug. 2007.

[164] S. Kalinin, S. Felekyan, A. Valeri, and C. A. M. Seidel. Characterizing multiple molecular States in single-molecule multiparameter fluorescence detection by probability distribution analysis. *The journal of physical chemistry B*, 112(28):8361–8374, 2008.

[165] S. Kalinin, E. Sisamakis, S. W. Magennis, S. Felekyan, and C. A. M. Seidel. On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. *The Journal of Physical Chemistry B*, 114(18):6197–6206, 2010.

[166] W. Kammerer, U. Deuschle, R. Gentz, and H. Bujard. Functional dissection of Escherichia coli promoters: information in the transcribed region is involved in late steps of the overall process. *the The European Molecular Biology Organization Journal*, 5(11):2995–3000, 1986.

[167] A. N. Kapanidis, T. A. Laurence, N. K. Lee, E. Margeat, X. Kong, and S. Weiss. Alternating-laser excitation of single molecules. *Accounts of Chemical Research*, 38(7):523–533, 2005.

[168] A. N. Kapanidis, N. K. Lee, T. A. Laurence, S. Doose, E. Margeat, and S. Weiss. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24):8936–8941, 2004.

[169] A. N. Kapanidis, E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss, and R. H. Ebright. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*, 314(5802):1144, 2006.

[170] A. N. Kapanidis and T. Strick. Biology, one molecule at a time. *Trends in biochemical sciences*, 34(5):234–43, May 2009.

[171] P. Kask, K. Palo, N. Fay, L. Brand, U. Mets, D. Ullmann, J. Jungmann, J. Pschorr, and K. Gall. Two-dimensional fluorescence intensity distribution analysis: theory and applications. *Biophysical journal*, 78(4):1703–13, Apr. 2000.

[172] P. Kask, K. Palo, D. Ullmann, and K. Gall. Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. *Proceedings of the National Academy of Sciences of the United States of America*, 96(24):13756–13761, 1999.

[173] P. Kask, P. Piksarv, M. Pooga, U. Mets, and E. Lippmaa. Separation of the rotational contribution in fluorescence correlation experiments. *Biophysical journal*, 55(2):213–20, Feb. 1989.

[174] D. Kelly, M. Dillingham, A. Hudson, and K. Wiesner. A new method for inferring hidden markov models from noisy time sequences. *PloS one*, 7(1):e29703, Jan. 2012.

[175] U. Kettling, A. Koltermann, P. Schwille, and M. Eigen. Real-time enzyme kinetics monitored by dual-color fluorescence cross-correlation spectroscopy. *Proceedings of the National Academy of Sciences*, 95(4):1416–1420, Feb. 1998.

[176] K. Kikuchi, C. Sato, M. Watabe, H. Ikeda, Y. Takahashi, and T. Miyashi. New aspects of fluorescence quenching by molecular oxygen. *Journal of the American Chemical Society*, 115(12):5180–5184, June 1993.

[177] H. Kim, G.-Q. Tang, S. S. Patel, and T. Ha. Opening-closing dynamics of the mitochondrial transcription pre-initiation complex. *Nucleic acids research*, 40(1):1–11, 2011.

[178] S. A. Kim, K. G. Heinze, and P. Schwille. Fluorescence correlation spectroscopy in living cells. *Nature methods*, 4(11):963–73, Nov. 2007.

[179] K.-H. Knauer and R. Gleiter. Photochromism of Rhodamine Derivatives. *Angewandte Chemie International Edition in English*, 16(2):113–113, Feb. 1977.

[180] K. Kolmakov, V. N. Belov, C. A. Wurm, B. Harke, M. Leutenegger, C. Eggeling, and S. W. Hell. A Versatile Route to Red-Emitting Carbopyronine Dyes for Optical Microscopy and Nanoscopy. *European Journal of Organic Chemistry*, 2010(19):3593–3610, July 2010.

[181] X. Kong, E. Nir, K. Hamadani, and S. Weiss. Photobleaching pathway in single-molecule FRET. *Journal of the American Chemical Society*, 129(15):4643–4654, 2007.

[182] W. S. Kontur, M. W. Capp, T. J. Gries, R. M. Saecker, and M. T. Record. Probing DNA binding, DNA opening, and assembly of a downstream clamp/jaw in Escherichia coli RNA polymerase-lambdaP(R) promoter complexes using salt and the physiological anion glutamate. *Biochemistry*, 49(20):4361–4373, 2010.

[183] W. J. A. Koopmans, A. Brehm, C. Logie, T. Schmidt, and J. Van Noort. Single-Pair FRET Microscopy Reveals Mononucleosome Dynamics. *Journal of Fluorescence*, 17(6):785–795, 2007.

[184] J. Korlach, P. J. Marks, R. L. Cicero, J. J. Gray, D. L. Murphy, D. B. Roitman, T. T. Pham, G. A. Otto, M. Foquet, and S. W. Turner. Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proceedings of the National Academy of Sciences of the United States of America*, 105(4):1176–81, Jan. 2008.

[185] S. C. Kou, X. S. Xie, and J. S. Liu. Bayesian analysis of single-molecule experimental data. *Journal of the Royal Statistical Society Series C Applied Statistics*, 54(3):469–506, 2005.

[186] T. Kubori and N. Shimamoto. A branched pathway in the early stage of transcription by Escherichia coli RNA polymerase. *Journal of Molecular Biology*, 256(3):449–457, 1996.

[187] W. Kügel, A. Muschielok, and J. Michaelis. Bayesian-Inference-Based Fluorescence Correlation Spectroscopy and Single-Molecule Burst Analysis Reveal the Influence of Dye Selection on DNA Hairpin Dynamics. *Chemphyschem a European journal of chemical physics and physical chemistry*, 13(4):1–11, 2012.

[188] M. Kumbhakar, A. Kiel, H. Pal, and D.-P. Herten. Single-molecule fluorescence studies reveal long-range electron-transfer dynamics through double-stranded DNA. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 10(4):629–33, Mar. 2009.

[189] J. R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer, 2010.

[190] D. C. Lamb, A. Schenk, C. Rcker, and G. U. Nienhaus. Determining chemical rate coefficients using time-gated fluorescence correlation spectroscopy. *Journal of Physical Organic Chemistry*, 13(10):654–658, Oct. 2000.

[191] R. J. Larsen and M. L. Marx. *An Introduction to Mathematical Statistics and Its Applications*. Prentice Hall, 2006.

[192] T. A. Laurence, X. Kong, M. Jäger, and S. Weiss. Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48):17348–53, Nov. 2005.

[193] L. Le Reste, J. Hohlbein, K. Gryte, and A. Kapanidis. Characterization of Dark Quencher Chromophores as Nonfluorescent Acceptors for Single-Molecule FRET. *Biophysical Journal*, 102(11):2658–2668, 2012.

[194] N. K. Lee, A. N. Kapanidis, H. R. Koh, Y. Korlann, S. O. Ho, Y. Kim, N. Gassman, S. K. Kim, and S. Weiss. Three-color alternating-laser excitation of single molecules: monitoring multiple interactions and distances. *Biophysical Journal*, 92(1):303–12, 2007.

[195] N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, and S. Weiss. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophysical Journal*, 88(4):2939–2953, 2005.

[196] N. K. Lee, H. R. Koh, K. Y. Han, J. Lee, and S. K. Kim. Single-molecule, real-time measurement of enzyme kinetics by alternating-laser excitation fluorescence resonance energy transfer. *Chemical Communications*, 46(26):4683–4685, 2010.

[197] T.-H. Lee. Extracting kinetics information from single-molecule fluorescence resonance energy transfer data using hidden markov models. *The Journal of Physical Chemistry B*, 113(33):11535–11542, 2009.

[198] M. Leutenegger, H. Blom, J. Widengren, C. Eggeling, M. Gösch, R. A. Leitgeb, and T. Lasser. Dual-color total internal reflection fluorescence cross-correlation spectroscopy., 2006.

[199] M. J. Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, and W. W. Webb. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science (New York, N.Y.)*, 299(5607):682–6, Jan. 2003.

[200] H. Li, X. Ren, L. Ying, S. Balasubramanian, and D. Klenerman. Measuring single-molecule nucleic acid dynamics in solution by two-color filtered ratiometric fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 101(40):14425–30, Oct. 2004.

[201] Y. Li, G. J. Augustine, and K. Weninger. Kinetics of complexin binding to the SNARE complex: correcting single molecule FRET measurements for hidden events. *Biophysical Journal*, 93(6):2178–2187, 2007.

[202] S. Linde, R. Kasper, M. Heilemann, and M. Sauer. Photoswitching microscopy with standard fluorophores. *Applied Physics B*, 93(4):725–731, Oct. 2008.

[203] J. Liu and C. L. Turnbough. Effects of transcriptional start site sequence and position on nucleotide-sensitive selection of alternative start sites at the pyrC promoter in Escherichia coli. *Journal Of Bacteriology*, 176(10):2938–2945, 1994.

[204] Y. Liu, J. Park, K. A. Dahmen, Y. R. Chemla, and T. Ha. A comparative study of multivariate and univariate hidden Markov modelings in time-binned single-molecule FRET data analysis. *The Journal of Physical Chemistry B*, 114(16):5386–5403, 2010.

[205] H. P. Lu. Revealing time bunching effect in single-molecule enzyme conformational dynamics. *Physical chemistry chemical physics : PCCP*, 13(15):6734–49, Apr. 2011.

[206] P. M. Lundquist, C. F. Zhong, P. Zhao, A. B. Tomaney, P. S. Peluso, J. Dixon, B. Bettman, Y. Lacroix, D. P. Kwo, E. McCullough, M. Maxham, K. Hester, P. McNitt, D. M. Grey, C. Henriquez, M. Foquet, S. W. Turner, and D. Zaccarin. Parallel confocal detection of single molecules in real time. *Optics Letters*, 33(9):1026, Apr. 2008.

[207] G. Luo, M. Wang, W. H. Konigsberg, and X. S. Xie. Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31):12610–12615, 2007.

[208] K. Lymperopoulos, R. Crawford, J. P. Torella, M. Heilemann, L. C. Hwang, S. J. Holden, and A. N. Kapanidis. Single-molecule DNA biosensors for protein and ligand detection. *Angewandte Chemie (International ed. in English)*, 49(7):1316–20, Feb. 2010.

[209] D. Magde, E. L. Elson, and W. W. Webb. Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers*, 13(1):29–61, Jan. 1974.

[210] N. Maheshri and E. K. O'Shea. Living with noisy genes: how cells function reliably with inherent variability in gene expression. *Annual review of biophysics and biomolecular structure*, 36:413–34, Jan. 2007.

[211] J. Mandel. *The Statistical Analysis of Experimental Data*, volume 61. Dover Publications, 1984.

[212] C. Manzo and L. Finzi. Quantitative analysis of DNA-looping kinetics from tethered particle motion experiments. *Methods in enzymology*, 475:199–220, Jan. 2010.

[213] S. Mao, R. K. P. Benninger, Y. Yan, C. Petchprayoon, D. Jackson, C. J. Easley, D. W. Piston, and G. Marriott. Optical lock-in detection of FRET using synthetic and genetically encoded optical switches. *Biophysical journal*, 94(11):4515–24, June 2008.

[214] E. Margeat, A. N. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, and S. Weiss. Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes. *Biophysical Journal*, 90(4):1419–1431, 2006.

[215] J. Matsumoto. Evolutionary role of abortive transcript as a primer for DNA replication. *Journal of Molecular Evolution*, 39(6):620–624, 1994.

[216] J. J. McCann, U. B. Choi, L. Zheng, K. Weninger, and M. E. Bowen. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophysical journal*, 99(3):961–970, 2010.

[217] K. McHale, A. J. Berglund, and H. Mabuchi. Bayesian estimation for species identification in single-molecule fluorescence microscopy. *Biophysical journal*, 86(6):3409–22, June 2004.

[218] S. A. McKinney, C. Joo, and T. Ha. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical Journal*, 91(5):1941–51, 2006.

[219] D. A. McQuarrie. *Quantum Chemistry*. University Science Books,U.S., 2007.

[220] F. Meng and H. Ma. Extended photon counting histogram and fluorescence intensity distribution analysis approaches for optically biased photon counting statistics. *The Journal of Physical Chemistry B*, 110(19):9667–9673, 2006.

[221] U. Meseth, T. Wohland, R. Rigler, and H. Vogel. Resolution of fluorescence correlation measurements. *Biophysical journal*, 76(3):1619–31, Mar. 1999.

[222] T. C. Messina, H. Kim, J. T. Giurleo, and D. S. Talaga. Hidden Markov model analysis of multichromophore photobleaching. *The Journal of Physical Chemistry B*, 110(33):16366–16376, 2006.

[223] L. S. Milescu, C. Nicolai, A. Yildiz, P. R. Selvin, and F. Sachs. Hidden Markov model applications in QuB: analysis of nanometer steps in single molecule fluorescence data and ensemble ion channel kinetics. *Biophysical Journal*, 84(2):124A–124A, 2003.

[224] J. N. Milstein, Y. F. Chen, and J.-C. Meiners. Bead size effects on protein-mediated DNA looping in tethered-particle motion experiments. *Biopolymers*, 95(2):144–50, Feb. 2011.

[225] W. E. Moerner. New directions in single-molecule imaging and analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31):12596–602, July 2007.

[226] W. E. Moerner and D. P. Fromm. Methods of single-molecule fluorescence spectroscopy and microscopy. *Review of Scientific Instruments*, 74(8):3597, Aug. 2003.

[227] K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods*, 7(5):377–381, 2010.

[228] J. Mukhopadhyay, A. N. Kapanidis, V. Mekler, E. Kortkhonjia, Y. W. Ebright, and R. H. Ebright. Translocation of sigma(70) with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell*, 106(4):453–463, 2001.

[229] B. K. Müller, E. Zaychikov, C. Bräuchle, and D. C. Lamb. Pulsed interleaved excitation. *Biophysical Journal*, 89(5):3508–3522, 2005.

[230] J. D. Müller, Y. Chen, and E. Gratton. Resolving heterogeneity on the single molecular level with the photon-counting histogram. *Biophysical journal*, 78(1):474–86, Jan. 2000.

[231] F. E. Müllner, S. Syed, P. R. Selvin, and F. J. Sigworth. Improved hidden Markov models for molecular motors, part 1: basic theory. *Biophysical Journal*, 99(11):3684–3695, 2010.

[232] K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, and S. A. Darst. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, 296(5571):1285–1290, 2002.

[233] K. S. Murakami, S. Masuda, and S. A. Darst. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science*, 296(5571):1280–1284, 2002.

[234] S. Myong, S. Cui, P. V. Cornish, A. Kirchhofer, M. U. Gack, J. U. Jung, K.-P. Hopfner, and T. Ha. Cytosolic viral sensor RIG-I is a 5'-triphosphate-dependent translocase on double-stranded RNA. *Science New York NY*, 323(5917):1070–1074, 2009.

[235] H. Neubauer, N. Gaiko, S. Berger, J. Schaffer, C. Eggeling, J. Tuma, L. Verdier, C. A. M. Seidel, C. Griesinger, and A. Volkmer. Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy. *Journal of the American Chemical Society*, 129(42):12746–55, Oct. 2007.

[236] K. C. Neuman and A. Nagy. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods*, 5(6):491–505, June 2008.

[237] S. Nie and R. N. Zare. Optical detection of single molecules. *Annual review of biophysics and biomolecular structure*, 26:567–96, Jan. 1997.

[238] D. Nikolić, R. C. Murean, W. Feng, and W. Singer. Scaled correlation analysis: a better way to compute a cross-correlogram. *The European journal of neuroscience*, 35(October 2008):742–762, 2012.

[239] E. Nir, X. Michalet, K. M. Hamadani, T. A. Laurence, D. Neuhauser, Y. Kovchegov, and S. Weiss. Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *The Journal of Physical Chemistry B*, 110(44):22103–24, 2006.

[240] L. Novotny and B. Hecht. *Principles of Nano-Optics*. Cambridge University Press, 2006.

[241] G. W. Oehlert. A Note on the Delta Method. *American Statistician*, 46(1):27–29, 1992.

[242] K. Okamoto and Y. Sako. Variational Bayes Analysis of a Photon-Based Hidden Markov Model for Single-Molecule FRET Trajectories. *Biophysical journal*, 103(6):1315–24, Sept. 2012.

[243] B. Okumus, S. Arslan, S. M. Fengler, S. Myong, and T. Ha. Single Molecule Nanocontainers Made Porous Using a Bacterial Toxin. *Journal of the American Chemical Society*, 131(41):14844–14849, 2009.

[244] B. Okumus, T. J. Wilson, D. M. J. Lilley, and T. Ha. Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. *Biophysical journal*, 87(4):2798–2806, 2004.

[245] M. Orrit. Chemical and physical aspects of charge transfer in the fluorescence intermittency of single molecules and quantum dots. *Photochemical photobiological sciences Official journal of the European Photochemistry Association and the European Society for Photobiology*, 9(5):637–642, 2010.

[246] M. A. Osborne, S. Balasubramanian, W. S. Furey, and D. Klenerman. Optically Biased Diffusion of Single Molecules Studied by Confocal Fluorescence Microscopy. *The Journal of Physical Chemistry B*, 102(17):3160–3167, Apr. 1998.

[247] P. Pal, J. F. Lesoine, M. A. Lieb, L. Novotny, and P. A. Knauf. A Novel Immobilization Method for Single Protein spFRET Studies. *Biophysical journal*, 89(2):L11–L13, 2005.

[248] S. Papadimitriou, J. Sun, and P. Yu. Local Correlation Tracking in Time Series. *Sixth International Conference on Data Mining ICDM06*, pages 456–465, 2006.

[249] C. S. Parmenter. Fluorescence Quenching in Aromatic Hydrocarbons by Oxygen. *The Journal of Chemical Physics*, 51(5):2242, Sept. 1969.

[250] V. P. Pastushenko and H. Schindler. Level detection in ion channel records via idealization by statistical filtering and likelihood optimization. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, 352(1349):39–51, 1997.

[251] J. B. Patlak. Measuring kinetics of complex single ion channel data using mean-variance histograms. *Biophysical Journal*, 65(1):29–42, 1993.

[252] J. Pawley. *Handbook of Biological Confocal Microscopy*. Springer, 2006.

[253] T. D. Perroud, B. Huang, and R. N. Zare. Effect of bin time on the photon counting histogram for one-photon excitation. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 6(5):905–12, May 2005.

[254] H. Piwoski, R. Koos, A. Meixner, and J. Sepio. Optimal oxygen concentration for the detection of single indocarbocyanine molecules in a polymeric matrix. *Chemical Physics Letters*, 405(4-6):352–356, Apr. 2005.

[255] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2007.

[256] M. Prummer, C. G. Hübner, B. Sick, B. Hecht, A. Renn, and U. P. Wild. Single-Molecule Identification by Spectrally and Time-Resolved Fluorescence Detection. *Analytical Chemistry*, 72(3):443–447, Feb. 2000.

[257] M. Prummer, B. Sick, A. Renn, and U. P. Wild. Multiparameter microscopy and spectroscopy for single-molecule analytics. *Analytical chemistry*, 76(6):1633–40, Mar. 2004.

[258] F. Qin, A. Auerbach, and F. Sachs. A direct optimization approach to hidden Markov modeling for single channel kinetics. *Biophysical Journal*, 79(4):1915–1927, 2000.

[259] F. Qin, A. Auerbach, and F. Sachs. Hidden Markov modeling for single channel kinetics with filtering and correlated noise. *Biophysical Journal*, 79(4):1928–1944, 2000.

[260] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[261] M. Rarbach, U. Kettling, A. Koltermann, and M. Eigen. Dual-color fluorescence cross-correlation spectroscopy for monitoring the kinetics of enzyme-catalyzed reactions. *Methods (San Diego, Calif.)*, 24(2):104–16, June 2001.

[262] I. Rasnik, S. A. Mckinney, and T. Ha. Nonblinking and long- lasting single-molecule fluorescence imaging. *Nature Methods*, 3(11):891–893, 2006.

[263] A. Renn, J. Seelig, and V. Sandoghdar. Oxygen-dependent photochemistry of fluorescent dyes studied at the single molecule level. *Molecular Physics*, 104(3):409–414, Feb. 2006.

[264] A. Revyakin, C. Liu, R. H. Ebright, and T. R. Strick. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*, 314(5802):1139–1143, 2006.

[265] E. Rhoades, E. Gussakovsky, and G. Haran. Watching proteins fold one molecule at a time. *Proceedings of the National Academy of Sciences of the United States of America*, 100(6):3197–3202, 2003.

[266] R. Rigler, U. Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *European Biophysics Journal*, 22(3), Aug. 1993.

[267] M. S. Robbins and B. J. Hadwen. The noise performance of electron multiplying charge-coupled devices. *IEEE Transactions on Electron Devices*, 50(5):1227–1232, 2003.

[268] A. Rogozina, E. Zaychikov, M. Buckle, H. Heumann, and B. Sclavi. DNA melting by RNA polymerase at the T7A1 promoter precedes the rate-limiting step at 37 degrees C and results in the accumulation of an off-pathway intermediate. *Nucleic acids research*, 37(16):5390–404, Sept. 2009.

[269] S. A. Rosenberg, M. E. Quinlan, J. N. Forkey, and Y. E. Goldman. Rotational motions of macro-molecules by single-molecule fluorescence microscopy. *Accounts of chemical research*, 38(7):583–93, July 2005.

[270] R. Roy, S. Hohng, and T. Ha. A practical guide to single-molecule FRET. *Nature Methods*, 5(6):507–516, 2008.

[271] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, 3(10):793–5, Oct. 2006.

[272] C. R. Sabanayagam, J. S. Eid, and A. Meller. Long time scale blinking kinetics of cyanine fluorophores conjugated to DNA and its effect on Förster resonance energy transfer. *The Journal of chemical physics*, 123(22):224708, Dec. 2005.

[273] R. M. Saecker, M. T. Record, and P. L. Dehaseth. Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of molecular biology*, 412(5):754–71, Oct. 2011.

[274] Y. Santoso, L. C. Hwang, L. Le Reste, and A. N. Kapanidis. Red light, green light: probing single molecules using alternating-laser excitation. *Biochemical Society Transactions*, 36(Pt 4):738–744, 2008.

[275] Y. Santoso, C. M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. P. Torella, N. D. F. Grindley, and A. N. Kapanidis. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2):715–720, 2010.

[276] Y. Santoso and A. N. Kapanidis. Probing biomolecular structures and dynamics of single molecules using in-gel alternating-laser excitation. *Analytical Chemistry*, 81(23):9561–9570, 2009.

[277] Y. Santoso, J. P. Torella, and A. N. Kapanidis. Characterizing single-molecule FRET dynamics with probability distribution analysis. *Chemphyschem A European Journal Of Chemical Physics And Physical Chemistry*, 11(10):2209–2219, 2010.

[278] M. Sauer, J. Hofkens, and J. Enderlein. *Handbook of Fluorescence Spectroscopy and Imaging*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, Jan. 2011.

[279] E. Saxon. Cell Surface Engineering by a Modified Staudinger Reaction. *Science*, 287(5460):2007–2010, Mar. 2000.

[280] D. A. Schafer, J. Gelles, M. P. Sheetz, and R. Landick. Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature*, 352(6334):444–8, 1991.

[281] G. D. Scholes. Long-range resonance energy transfer in molecular systems. *Annual review of physical chemistry*, 54:57–87, Jan. 2003.

[282] G. F. Schroder and H. Grubmuller. Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. *The Journal of Chemical Physics*, 119(18):9920, Nov. 2003.

[283] B. Schuler, E. A. Lipman, and W. A. Eaton. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*, 419(6908):743–747, 2002.

[284] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumke, and W. A. Eaton. Polyproline and the spectroscopic ruler revisited with single-molecule fluorescence. *PNAS*, 102(8):9754–9759, 2005.

[285] J. J. Schwartz and S. R. Quake. Single molecule measurement of the "speed limit" of DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 106(48):20294–9, Dec. 2009.

[286] C. Schweitzer and R. Schmidt. Physical mechanisms of generation and deactivation of singlet oxygen. *Chemical reviews*, 103(5):1685–757, May 2003.

[287] P. Schwille, F. J. Meyer-Almes, and R. Rigler. Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophysical Journal*, 72(4):1878–1886, 1997.

[288] B. Sclavi, E. Zaychikov, A. Rogozina, F. Walther, M. Buckle, and H. Heumann. Real-time characterization of intermediates in the pathway to open complex formation by Escherichia coli RNA polymerase at the T7A1 promoter. *Proceedings of the National Academy of Sciences of the United States of America*, 102(13):4706–11, Mar. 2005.

[289] E. B. Shera, N. K. Seitzinger, L. M. Davis, R. A. Keller, and S. A. Soper. Detection of single fluorescent molecules. *Chemical Physics Letters*, 174(6):553–557, 1990.

[290] X. Shi, J. Lim, and T. Ha. Acidification of the Oxygen Scavenging System in Single-Molecule Fluorescence Studies: In Situ Sensing with a Ratiometric Dual-Emission Probe. *Analytical Chemistry*, 82(14):6132–6138, 2010.

[291] E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. M. Seidel. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods in enzymology*, 475(10):455–514, 2010.

[292] D. A. Smith, W. Steffen, R. M. Simmons, and J. Sleep. Hidden-Markov methods for the analysis of single-molecule actomyosin displacement data: the variance-Hidden-Markov method. *Biophysical Journal*, 81(5):2795–2816, 2001.

[293] M. Sorokina, H.-R. Koh, S. S. Patel, and T. Ha. Fluorescent lifetime trajectories of a single fluorophore reveal reaction intermediates during transcription initiation. *Journal of the American Chemical Society*, 131(28):9630–1, 2009.

[294] M. Springer. *The algebra of random variables*. Wiley, New York, 1979.

[295] I. Z. Steinberg. Long-range nonradiative transfer of electronic excitation energy in proteins and polypeptides. *Annual Review of Biochemistry*, 40(1):83–114, 1971.

[296] C. Steinhauer, C. Forthmann, J. Vogelsang, and P. Tinnefeld. Superresolution microscopy on the basis of engineered dark states. *Journal of the American Chemical Society*, 130(50):16840–16841, 2008.

[297] P. B. Stetson. DAOPHOT - A computer program for crowded-field stellar photometry. *Publications of the Astronomical Society of the Pacific*, 99:191, 1987.

[298] B. Stevens and B. E. Algar. Photoperoxidation of unsaturated organic molecules. I. Relaxation and oxygen-quenching parameters of the sensitizer singlet state. *The Journal of Physical Chemistry*, 72(7):2582–2587, July 1968.

[299] J. Stigler and M. Rief. Hidden markov analysis of trajectories in single-molecule experiments and the effects of missed events. *Chemphyschem A European Journal Of Chemical Physics And Physical Chemistry*, 13(4):1079–86, 2012.

[300] L. Stryer. Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemistry*, 47(5):819–46, 1978.

[301] L. Stryer and R. P. Haugland. Energy transfer: a spectroscopic ruler. *Proceedings of the National Academy of Sciences*, 58(2):719–726, 1967.

[302] M. Susa, T. Kubori, and N. Shimamoto. A pathway branching in transcription initiation in Escherichia coli. *Molecular Microbiology*, 59(6):1807–1817, 2006.

[303] M. Susa, R. Sen, and N. Shimamoto. Generality of the branched pathway in transcription initiation by Escherichia coli RNA polymerase. *The Journal of Biological Chemistry*, 277(18):15407–15412, 2002.

[304] K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, 365(6448):721–7, Oct. 1993.

[305] M. Swoboda, J. Henig, H.-M. Cheng, D. Brugger, D. Haltrich, N. Plumeré, and M. Schlierf. Enzymatic Oxygen Scavenging for Photostability without pH Drop in Single-Molecule Experiments. *ACS nano*, 6(7):6364–9, June 2012.

[306] S. Syed, F. E. Müllner, P. R. Selvin, and F. J. Sigworth. Improved hidden Markov models for molecular motors, part 2: extensions and application to experimental data. *Biophysical Journal*, 99(11):3696–3703, 2010.

[307] J. N. Taylor and C. F. Landes. Improved resolution of complex single-molecule FRET systems via wavelet shrinkage. *The journal of physical chemistry. B*, 115(5):1105–14, Feb. 2011.

[308] J. N. Taylor, D. E. Makarov, and C. F. Landes. Denoising single-molecule FRET trajectories with wavelets and Bayesian inference. *Biophysical Journal*, 98(1):164–173, 2010.

[309] R. E. Thompson, D. R. Larson, and W. W. Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5):2775–2783, 2002.

[310] T. E. Tomov, R. Tsukanov, R. Masoud, M. Liber, N. Plavner, and E. Nir. Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophysical journal*, 102(5):1163–73, Mar. 2012.

[311] M. Tomschik, K. Van Holde, and J. Zlatanova. Nucleosome dynamics as studied by single-pair fluorescence resonance energy transfer: a reevaluation. *Journal of Fluorescence*, 19(1):53–62, 2009.

[312] M. Tomschik, H. Zheng, K. Van Holde, J. Zlatanova, and S. H. Leuba. Fast, long-range, reversible conformational fluctuations in nucleosomes revealed by single-pair fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9):3278–3283, 2005.

[313] J. P. Torella. *Confocal single-molecule fluorescence as a tool for investigating biomolecular dynamics in vitro and in vivo*. Dphil, University of Oxford, Oxford, 2010.

[314] J. P. Torella, S. J. Holden, Y. Santoso, J. Hohlbein, and A. N. Kapanidis. Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophysical Journal*, 100(6):1568–1577, 2011.

[315] M. Torimura, S. Kurata, K. Yamada, T. Yokomaku, Y. Kamagata, T. Kanagawa, and R. Kurane. Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Analytical sciences the international journal of the Japan Society for Analytical Chemistry*, 17(1):155–160, 2001.

[316] T. Torres and M. Levitus. Measuring conformational dynamics: a new FCS-FRET approach. *The journal of physical chemistry. B*, 111(25):7392–400, June 2007.

[317] B. Treutlein, A. Muschielok, J. Andrecka, A. Jawhari, C. Buchen, D. Kostrewa, F. Hög, P. Cramer, and J. Michaelis. Dynamic Architecture of a Minimal RNA Polymerase II Open Promoter Complex. *Molecular Cell*, 46(2):136–146, 2012.

[318] R. S. Tsay. *Analysis of Financial Time Series Second Edition*, volume 543 of *Wiley Series in Probability and Statistics*. Wiley-Interscience, 2005.

[319] N. J. Turro, J. C. Scaiano, and V. Ramamurthy. *Principles of Molecular Photochemistry: An Introduction*. University Science Books, 2008.

[320] M. H. Ulbrich and E. Y. Isacoff. Subunit counting in membrane-bound proteins. *Nature Methods*, 4(4):2006–2008, 2007.

[321] S. Uphoff. *Probing Molecular and Cellular Structure using Single Molecule Photoswitching*. Master's thesis, University of Oxford, 2009.

[322] S. Uphoff, K. Gryte, G. Evans, and A. N. Kapanidis. Improved temporal resolution and linked hidden Markov modeling for switchable single-molecule FRET. *Chemphyschem A European Journal Of Chemical Physics And Physical Chemistry*, 12(3):571–579, 2011.

[323] S. Uphoff, S. J. Holden, L. Le Reste, J. Periz, S. Van De Linde, M. Heilemann, and A. N. Kapanidis. Monitoring multiple distances within a single molecule using switchable FRET. *Nature Methods*, 7(10):831–836, 2010.

[324] S. van de Linde, A. Löschberger, T. Klein, M. Heidbreder, S. Wolter, M. Heilemann, and M. Sauer. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature protocols*, 6(7):991–1009, July 2011.

[325] F. Vanzi, L. Sacconi, and F. S. Pavone. Analysis of kinetics in noisy systems: application to single molecule tethered particle motion. *Biophysical journal*, 93(1):21–36, July 2007.

[326] A. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm: Information Theory, IEEE Transactions on. *IEEE Transactions on Information Theory*, 13(2):260–269, 1967.

[327] N. V. Vo, L. M. Hsu, C. M. Kane, and M. J. Chamberlin. In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 2. Formation and characterization of two distinct classes of initial transcribing complexes. *Biochemistry*, 42(13):3787–3797, 2003.

[328] N. V. Vo, L. M. Hsu, C. M. Kane, and M. J. Chamberlin. In vitro studies of transcript initiation by Escherichia coli RNA polymerase. 3. Influences of individual DNA elements within the promoter recognition region on abortive initiation and promoter escape. *Biochemistry*, 42(13):3798–3811, 2003.

[329] J. Vogelsang, T. Cordes, C. Forthmann, C. Steinhauer, and P. Tinnefeld. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 106(20):8107–8112, 2009.

[330] J. Vogelsang, T. Cordes, and P. Tinnefeld. Single-molecule photophysics of oxazines on DNA and its application in a FRET switch. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology*, 8(4):486–96, Apr. 2009.

[331] J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, and P. Tinnefeld. A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angewandte Chemie International Edition*, 47(29):5465–5469, 2008.

[332] L. S. G. Wade. *Organic Chemistry (6th Edition)*. Prentice Hall, 2005.

[333] N. G. Walter, C.-Y. Huang, A. J. Manzo, and M. A. Sobhy. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature methods*, 5(6):475–89, June 2008.

[334] T.-H. Wang, Y. Peng, C. Zhang, P. K. Wong, and C.-M. Ho. Single-molecule tracing on a fluidic microchip for quantitative detection of low-abundance nucleic acids. *Journal of the American Chemical Society*, 127(15):5354–9, Apr. 2005.

[335] X. Wang and H. P. Lu. 2D regional correlation analysis of single-molecule time trajectories. *The journal of physical chemistry. B*, 112(47):14920–6, Nov. 2008.

[336] L. P. Watkins, H. Chang, and H. Yang. Quantitative single-molecule conformational distributions: a case study with poly-(L-proline). *The journal of physical chemistry. A*, 110(15):5191–203, Apr. 2006.

[337] L. P. Watkins and H. Yang. Information bounds and optimal analysis of dynamic single molecule measurements. *Biophysical journal*, 86(6):4015–29, June 2004.

[338] L. P. Watkins and H. Yang. Detection of intensity change points in time-resolved single-molecule measurements. *The journal of physical chemistry B*, 109(1):617–628, 2005.

[339] S. Weiss. Fluorescence spectroscopy of single biomolecules. *Science*, 283(5408):1676–1683, 1999.

[340] S. S. White, H. Li, R. J. Marsh, J. D. Piper, N. D. Leonczek, N. Nicolaou, A. J. Bain, L. Ying, and D. Klenerman. Characterization of a single molecule DNA switch in free solution. *Journal of the American Chemical Society*, 128(35):11423–32, Sept. 2006.

[341] J. Widengren, V. Kudryavtsev, M. Antonik, S. Berger, M. Gerken, and C. A. M. Seidel. Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Analytical chemistry*, 78(6):2039–2050, 2006.

[342] J. Widengren, U. Mets, and R. Rigler. Fluorescence correlation spectroscopy of triplet states in solution: a theoretical and experimental study. *The Journal of Physical Chemistry*, 99(36):13368–13379, Sept. 1995.

[343] J. Widengren, R. Rigler, and U. Mets. Triplet-state monitoring by fluorescence correlation spectroscopy. *Journal of Fluorescence*, 4(3):255–258, Sept. 1994.

[344] J. Widengren, E. Schweinberger, S. Berger, and C. A. M. Seidel. Two New Concepts to Measure Fluorescence Resonance Energy Transfer via Fluorescence Correlation Spectroscopy: Theory and Experimental Realizations. *The Journal of Physical Chemistry A*, 105(28):6851–6866, July 2001.

[345] J. Widengren and P. Schwille. Characterization of Photoinduced Isomerization and Back-Isomerization of the Cyanine Dye Cy5 by Fluorescence Correlation Spectroscopy. *The Journal of Physical Chemistry A*, 104(27):6416–6428, 2000.

[346] S. Xie. Single-Molecule Approach to Enzymology. *Single Molecules*, 2(4):229–236, Dec. 2001.

[347] X. F. Xiong and W. S. Reznikoff. Transcriptional slippage during the transcription initiation process at a mutant lac promoter in vivo. *Journal of Molecular Biology*, 231(3):569–580, 1993.

[348] C. S. Xu, H. Kim, C. C. Hayden, and H. Yang. Joint statistical analysis of multichannel time series from single quantum dot-(Cy5)n constructs. *The journal of physical chemistry B*, 112(19):5917–5923, 2008.

[349] Y. Xu, N. D. Grindley, and C. M. Joyce. Coordination between the polymerase and 5'-nuclease components of DNA polymerase I of Escherichia coli. *The Journal of biological chemistry*, 275(27):20949–55, July 2000.

[350] M. Yan and J. D. Gralla. Multiple ATP-dependent steps in RNA polymerase II promoter melting and initiation. *The EMBO journal*, 16(24):7457–7467, 1997.

[351] H. Yang. Detection and characterization of dynamical heterogeneity in an event series using wavelet correlation. *The Journal of chemical physics*, 129(7):074701, 2008.

[352] H. Yang. The Orientation Factor in Single-Molecule Forster-Type Resonance Energy Transfer, with Examples for Conformational Transitions in Proteins. *Israel Journal of Chemistry*, 49(3-4):313–321, June 2009.

[353] 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 (New York, N.Y.)*, 302(5643):262–6, Oct. 2003.

[354] H. Yang and X. Xie. Statistical approaches for probing single-molecule dynamics photon-by-photon. *Chemical Physics*, 284(1-2):423–437, Nov. 2002.

[355] H. Yang and X. S. Xie. Probing single-molecule dynamics photon by photon. *The Journal of Chemical Physics*, 117(24):10965, Dec. 2002.

[356] H.-C. Yeh, C. M. Puleo, Y.-P. Ho, V. J. Bailey, T. C. Lim, K. Liu, and T.-H. Wang. Tunable blinking kinetics of Cy5 for precise DNA quantification and single-nucleotide difference detection. *Biophysical journal*, 95(2):729–37, July 2008.

[357] H. Yin, R. Landick, and J. Gelles. Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophysical journal*, 67(6):2468–78, Dec. 1994.

[358] C. S. Yun, A. Javier, T. Jennings, M. Fisher, S. Hira, S. Peterson, B. Hopkins, N. O. Reich, and G. F. Strouse. Nanometal surface energy transfer in optical rulers, breaking the FRET barrier. *Journal of the American Chemical Society*, 127(9):3115–9, Mar. 2005.

[359] N. Zarrabi. *Hidden Markov Modelle für Einzelmoleküldaten*. Phd thesis, Universität Stuttgart, Holzgartenstr. 16, 70174 Stuttgart, 2010.

[360] N. Zarrabi, M. G. Düser, R. Reuter, S. D. Dunn, J. Wrachtrup, and M. Börsch. Detecting substeps in the rotary motors of FoF1-ATP synthase by Hidden Markov models. *Proceedings of SPIE*, (49):64440E–64440E–12, 2007.

[361] N. Zarrabi, S. Ernst, M. G. Dueser, A. Golovina-Leiker, W. Becker, R. Erdmann, S. D. Dunn, and M. Borsch. Simultaneous monitoring of the two coupled motors of a single FoF1-ATP synthase by three-color FRET using duty cycle-optimized triple-ALEX. *Proceedings of the SPIE*, 7185(49):14, 2009.

[362] N. Zarrabi, T. Heitkamp, J.-C. Greie, and M. Börsch. Monitoring the conformational dynamics of a single potassium transporter by ALEX-FRET. In *Proceedings of SPIE*, volume 6862, pages 68620M–68620M–11. SPIE, Jan. 2008.

[363] G. Zhang, E. A. Campbell, L. Minakhin, C. Richter, K. Severinov, and S. A. Darst. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. *Cell*, 98(6):811–24, Sept. 1999.

[364] Y. Zhang, Y. Feng, S. Chatterjee, S. Tuske, M. X. Ho, E. Arnold, and R. H. Ebright. Structural basis of transcription initiation. *Science (New York, N.Y.)*, 338(6110):1076–80, Nov. 2012.

[365] Q. Zheng, S. Jockusch, Z. Zhou, R. B. Altman, J. D. Warren, N. J. Turro, and S. C. Blanchard. On the Mechanisms of Cyanine Fluorophore Photostabilization. *The Journal of Physical Chemistry Letters*, 3(16):2200–2203, Aug. 2012.

[366] P. Zhu and H. G. Craighead. Zero-mode waveguides for single-molecule analysis. *Annual review of biophysics*, 41:269–93, Jan. 2012.

[367] X. Zhuang, L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu. A Single-Molecule Study of RNA Catalysis and Folding. *Science*, 288(5473):2048–2051, 2000.

[368] X. Zhuang, H. Kim, M. J. B. Pereira, H. P. Babcock, N. G. Walter, and S. Chu. Correlating structural dynamics and function in single ribozyme molecules. *Science (New York, N.Y.)*, 296(5572):1473–6, May 2002.

[369] R. Zondervan, F. Kulzer, S. B. Orlinskii, and M. Orrit. Photoblinking of Rhodamine 6G in Poly(vinyl alcohol): Radical Dark State Formed through the Triplet. *The Journal of Physical Chemistry A*, 107(35):6770–6776, 2003.