ENABLING MULTIVARIATE INVESTIGATION OF
SINGLE-MOLECULE DYNAMICS IN SOLUTION BY
COUNTERACTING BROWNIAN MOTION

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL
ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Quan Wang
November 2014
© 2014 by Quan Wang. All Rights Reserved.
Re-distributed by Stanford University under license with the author.

This work is licensed under a Creative Commons Attribution-
Noncommercial 3.0 United States License.
http://creativecommons.org/licenses/by-nc/3.0/us/

This dissertation is online at: http://purl.stanford.edu/wr470yr0497
I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

William Moerner, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Shanhui Fan

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Olav Solgaard

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.
Abstract

Although single-molecule fluorescence spectroscopy has developed into a mature toolkit that allows one to follow the dynamical behaviors of individual biomolecules in exquisite detail, its application in the solution phase is often limited by Brownian diffusion of the nanometer-sized biomolecules to only a 1-millisecond observation window. To remove this limitation while avoiding potential perturbation via immobilization, a device known as the Anti-Brownian ELectrokinetic (ABEL) trap has been developed to counteract Brownian motion that combines position-sensitive fluorescence detection and closed-loop feedback based on electrokinetic forces in a microfluidic geometry. This dissertation describes my work on recent methodology developments and new applications of the ABEL trap technology.

First, we describe an optimal redesign of the feedback control strategy that operates close to the ultimate physical limits imposed by diffraction, shot-noise, photobleaching and information bandwidth provided by fluorescence emission. This advanced system greatly improves the capability to hold individual proteins and oligonucleotides in buffer solution and enables multivariate interrogation of single-molecule dynamics for as long as 30 seconds in a non-perturbative aqueous environment.

Next, we combine multimodal fluorescence detection with the ABEL trap and illustrate its power using single organic fluorophores and a photosynthetic protein as examples. Direct observation of the synchronized dynamics of different fluorescence parameters reveals the internal states of a single nano-emitter and the interconversions between these states on a 1-second timescale in solution.

Last, we show that by applying machine-learning techniques to the ABEL trap’s photon data stream, a single molecule’s diffusive and electric-field-induced motion parameters
can be extracted in real-time during its residence in the trap. We demonstrate how this new measurement paradigm enables molecule-by-molecule sensing of size and charge heterogeneity and their fluctuations at the nanoscale. Two biophysical applications are illustrated: resolving the heterogeneous molecular mixture along a multimeric protein dissociation pathway and real-time visualization of the binding/unbinding interactions between single DNA strands.
Acknowledgement

Looking back, the past seven years that I spent at Stanford can be best described as fun. During these years, I’m grateful for many wonderful people I had a chance to interact with along the way.

W.E. Moerner has been an amazing adviser in many ways. He is the perfect role model for a young experimental scientist. When things don’t work as expected, he advises me with sound intuition, out-of-the-box thinking and lists of potential resources. When things do work, he reminds me to put things in perspective and acknowledge the limitation of the measurements before challenging me to “go for another 10× improvement” and “explore a new regime”. He is always enthusiastic about new ideas and supports them with big-picture advice and encouragement. He never hesitates to explore ways to make funding work or ask for new collaborators for the problems I wanted to work on. Moreover, he puts a big emphasis on developing a self-sustaining infrastructure to foster the development of writing and communication skills in scientific research, which I have benefited immensely from.

I feel extremely fortunate to have overlapped with a number of past and present Moerner lab members. Alexandre Fuerstenberg taught me a lot about basic wet lab practices and helped me get settled in the group. Yan Jiang and Randall Goldsmith were two extremely skillful and talented ABEL trappers who introduced me to the wet-lab wizardry of surface chemistry and microfluidic bonding. Jianwei Liu taught me every detailed steps of protein labeling and purification. Gabriela Schlau-Cohen introduced me to the fascinating research world of photosynthesis. Sam Lord, Whitney Duim, Matthew Lew, Steven Lee have all helped me, in one way or the other, reside on the winning side in the everlasting battle with Murphy’s laws. I enjoyed interactions with Lana Lau, Anika Kinkhabwala, Steffen Sahl,
Yoav Schetman in the underground office of Rm 15. I thank all Moerner group members, in addition to the aforementioned, Alex Diezmann, Mikael Backlund, Luc Weiss, Colin Comerci for the insightful feedback and suggestions on my work and presentations. Finally, I am privileged to mentor Hsiang-Yu Yang, a smart and knowledgeable graduate student who will receive the baton and continue to explore exciting new measurements with the ABEL trap.

I want to acknowledge Profs. Shanhui Fan, Olav Solgaard and Hideo Mabuchi for serving on my dissertation committee and providing thought-provoking feedback on my work. I thank Prof. Robert Byer for agreeing to serve as the chair of my thesis committee. I am thankful to the administrative staff in both Chemistry and EE, namely, Kathi Robbins, Ann Olive and Amy Duncan.

I am also indebted to a number of people in the University of New Mexico where I started my first two years of graduate education. Prof. Wolfgang Rudolf has been a phenomenal instructor that taught me all I know about laser and ultrafast physics and fostered my interest in experimental optical physics. Prof. James Thomas, having been my research mentor for a year, introduced me to the realm of biophysics and supported my own research ideas as well as my decision to apply to Stanford. I thank Mr. Boye “Mickey” Odom for giving me a chance to teach the undergrad labs. I would also like to thank Prof. Keith Lidke and his group for taking me as a visiting student in from Oct. 2013 till Jan. 2014.

I thank friends from both New Mexico and the bay area for support: Xiang He, Chengao Wang, Chang-Yi Lin, Chia-Ye Li, Chengyong Feng, Xudong Sun, Xing Chao and Qing Shao. Special thanks go to Claire Noonan and Peter Landsberger for renting me a lovely studio in Los Altos throughout my graduate school.

Last, I dedicate this thesis to my family, whose love and support I will eternally cherish. My parents have always taught me to be an independent thinker and a diligent worker, and supported every one of my decisions after high school. The work presented here would not have been possible without their financial support for my first two quarters at Stanford. I was lucky to have met my wife Xiaozhen Xu during my second year in New Mexico while becoming immediately California-bound to attend Stanford. Little did I know what a long distance relationship means at the time, only to find out that it will not be until 6 years and 238 days later that we finally reunite. I cannot thank her enough for her unfaltering
patience, love and support, incredible wisdom of life, as well as the sacrifice she has to make during all these years. I’m also grateful to my daughter Nuo-Jia “Nova” Wang (1 year old when this dissertation is submitted), who has added more joy and meaning to my life.
## Contents

Abstract ........................................ iv  

Acknowledgement ................................. vi  

1 Introduction ................................. 1  

1.1 Single-molecule fluorescence spectroscopy ........................ 2  

1.1.1 Basic principles ........................................ 2  

1.1.2 Single-molecule label as a probe for biomolecules .............. 5  

1.1.3 Methods for measuring single freely-diffusing molecules ....... 8  

1.1.4 Methods for measuring single immobilized molecules ......... 11  

1.2 The Anti-Brownian Electrokinetic trap ............................ 12  

1.2.1 Basic principle .......................................... 13  

1.2.2 Electrokinetics .......................................... 16  

1.3 Feedback tracking of a single object ............................... 21  

1.4 Scope and organization of this dissertation ......................... 21  

1.5 Publications from graduate work .................................. 24  

2 Theory: optimal tracking and parameter learning .................. 26  

2.1 Background: tracking single-molecule motion with high spatio-temporal resolution ........................................ 26  

2.2 Optimal strategy to track Brownian motion with photon-by-photon time resolution ........................................ 28  

2.2.1 Beam scanning on a square lattice and photon-by-photon mapping ........................................ 30  

2.2.2 Quantifying the information provided by one photon ........... 31
2.2.3 A graphical representation of the tracking problem . . . . . . . . . . 33
2.3 The optimal (minimum variance) online estimator . . . . . . . . . . . . . 35
2.4 The optimal offline estimator . . . . . . . . . . . . . . . . . . . . . . . . 39
2.4.1 The case of Gaussian measurement noise: Kalman filter/smooother . . 41
2.4.2 The case of non-Gaussian measurement noise: expectation-
propagation based message passing . . . . . . . . . . . . . . . . 45
2.5 Learning the diffusion coefficient and electrokinetic mobility . . . . . . . 49
2.5.1 Adaptive trapping by innovation whitening . . . . . . . . . . . . . . . 49
2.5.2 The expectation-maximization based maximum-likelihood approach 56

3 Making them all work: ABEL trap instrumentation

3.1 System overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 69
3.2 Excitation path . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 70
3.2.1 Acousto-optic deflectors . . . . . . . . . . . . . . . . . . . . . . . . . 70
3.2.2 Relay optics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74
3.3 Detection optics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 76
3.4 Electronics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
3.4.1 Connection diagram . . . . . . . . . . . . . . . . . . . . . . . . . . . 81
3.4.2 High-voltage amplifier . . . . . . . . . . . . . . . . . . . . . . . . . . 81
3.5 Control algorithm and FPGA programming . . . . . . . . . . . . . . . . . 83
3.6 Trap calibration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 86
3.6.1 The quality of the fast scanning pattern . . . . . . . . . . . . . . . . . 86
3.6.2 Creating the scan pattern: position calibration and brightness ho-
mogenization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87
3.6.3 Static trap response . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
3.7 The microfluidic chip . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
3.7.1 Quartz cell fabrication, bonding and cleaning . . . . . . . . . . . . . 91
3.7.2 Surface passivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
3.8 Other measurement modalities with the ABEL trap . . . . . . . . . . . . . 95
3.8.1 Fluorescence lifetime . . . . . . . . . . . . . . . . . . . . . . . . . . . 96
3.8.2 Fluorescence emission spectrum . . . . . . . . . . . . . . . . . . . . . 100
5.5 Methods ......................................................... 157
5.5.1 FRET model of the monomer ......................... 157

6 Beyond fluorescence parameters: single-molecule transport properties 161
6.1 Real-time estimation of single-molecule diffusion coefficient ($D$) and electrokinetic mobility ($\mu$) ......................... 162
6.2 Validation on different single molecules .................. 164
6.3 Resolving a three-component mixture molecule-by-molecule ........................................ 168
6.4 Application: following the dissociation pathway of Allophycocyanin in solution ........................................ 170
6.5 Application: binding kinetics of short DNA strands........ 172
6.6 Summary ....................................................... 178
6.7 Materials and methods ........................................ 178

A A list of ABEL trap electronics 182

B Protocols ......................................................... 184
B.1 Quartz cell cleaning procedure ............................ 184
B.2 Coating the interior of a quartz cell with polyelectrolyte multilayers (PEMs) 186
B.3 Preparing the protocatechuic acid (PCA)/protocatechuate-3,4-dioxygenase (PCD) oxygen-scavenging system .................. 188
B.4 Labeling amino-modified oligonucleotides with NHS dyes ...................... 189

C Joule heating in the ABEL trap .............................. 192

D Example Matlab code for processing the ABEL trap data 194

Bibliography ...................................................... 196
# List of Tables

1.1 Diffusion coefficients of example biomolecules .................................. 7  
2.1 Success rate of convergence, with a 12,000 photon processing window ... 63  
3.1 Acousto-optic properties of TeO$_2$ ................................................. 70  
3.2 AOD system summary of Trap A. .................................................. 72  
3.3 AOD system summary of Trap A. .................................................. 72  
3.4 Managing spherical aberration in the detection path ......................... 78  
3.5 Calibration of $g$ factor for single-molecule anisotropy measurements. .... 112  
6.1 $D - \mu$ measurements of different single molecule samples. ............ 167  
A.1 Electronics components .......................................................... 183
## List of Figures

1.1 Jablonski diagram of the energy levels and processes involved in fluorescence emission. ............................................................... 4
1.2 Comparing free diffusion, surface immobilization and ABEL trap based single molecule measurements. .......................................... 6
1.3 Basic principles of the Anti-Brownian ELectrokinetic (ABEL) trap ........ 14
1.4 The electric double-layer ........................................................................ 18
1.5 The electrokinetic response is a combined effect of electrophoresis (EP) and electroosmosis (EO). ..................................................... 19
1.6 Structure and organization of this dissertation. ....................................... 22
2.1 Position sensing by a revolving laser excitation beam. .......................... 28
2.2 From scanning on a circle to scanning on a 2D grid. ............................... 30
2.3 Example photon-by-photon position mapping. ........................................ 31
2.4 Design of the excitation beam scanning pattern. ..................................... 32
2.5 Graphical model representation of photon-by-photon single-molecule tracking. .............................................................................. 34
2.6 The Kalman filter scheme. ...................................................................... 35
2.7 Simulated trapping trace with beam scanning and Kalman filter. ........... 38
2.8 Comparison of the knight’s tour scanning, Kalman filter scheme and the revolving beam, lock-in detection scheme by simulation. ............... 40
2.9 Message passing pathways (arrows) for Kalman filter/smooth and EP-MP. 42
2.10 Moment matching examples. See text for details. .................................... 46
2.11 Simulation of adaptive trapping by innovation whitening. ........................ 51
2.12 Adaptive trapping of fluorescent polystyrene beads of different sizes. .... 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.13</td>
<td>Convergence of the EM algorithm.</td>
<td>62</td>
</tr>
<tr>
<td>2.14</td>
<td>Accuracy of the EM algorithm</td>
<td>64</td>
</tr>
<tr>
<td>2.15</td>
<td>Precision of the EM algorithm.</td>
<td>65</td>
</tr>
<tr>
<td>2.16</td>
<td>Benchmark of the EM algorithm</td>
<td>66</td>
</tr>
<tr>
<td>3.1</td>
<td>Overview of the ABEL trap system.</td>
<td>69</td>
</tr>
<tr>
<td>3.2</td>
<td>Acousto-optic interaction in the Bragg regime.</td>
<td>71</td>
</tr>
<tr>
<td>3.3</td>
<td>Relay optics in the ABEL trap excitation path.</td>
<td>75</td>
</tr>
<tr>
<td>3.4</td>
<td>Detection optics of Trap B.</td>
<td>77</td>
</tr>
<tr>
<td>3.5</td>
<td>Wiring diagram of Trap A.</td>
<td>80</td>
</tr>
<tr>
<td>3.6</td>
<td>Wiring diagram of Trap B.</td>
<td>82</td>
</tr>
<tr>
<td>3.7</td>
<td>Circuit diagram of the home-built high-voltage amplifier.</td>
<td>84</td>
</tr>
<tr>
<td>3.8</td>
<td>Hardware architecture of NI FPGA boards.</td>
<td>85</td>
</tr>
<tr>
<td>3.9</td>
<td>Block diagram of ABEL trap implementation on the FPGA.</td>
<td>87</td>
</tr>
<tr>
<td>3.10</td>
<td>Direct imaging of the scanning beam at different speeds.</td>
<td>88</td>
</tr>
<tr>
<td>3.11</td>
<td>Calibration of the position sensitivity of the ABEL trap.</td>
<td>90</td>
</tr>
<tr>
<td>3.12</td>
<td>ABEL trap microfluidic chip.</td>
<td>92</td>
</tr>
<tr>
<td>3.13</td>
<td>ABEL trap setup with multi-parameter fluorescence detection.</td>
<td>96</td>
</tr>
<tr>
<td>3.14</td>
<td>Representative spectra of the supercontinuum light under different pumping conditions.</td>
<td>97</td>
</tr>
<tr>
<td>3.15</td>
<td>Example measurements of the APD instrument response function.</td>
<td>98</td>
</tr>
<tr>
<td>3.16</td>
<td>Spectral calibration of the Amici prism-based spectrometer.</td>
<td>102</td>
</tr>
<tr>
<td>3.17</td>
<td>Typical background spectrum with 594 nm excitation.</td>
<td>103</td>
</tr>
<tr>
<td>3.18</td>
<td>Comparing center-of-mass analysis and fitting for spectral quantification using simulated spectra.</td>
<td>105</td>
</tr>
<tr>
<td>3.19</td>
<td>Modeling fluorescence anisotropy with a high NA objective.</td>
<td>109</td>
</tr>
<tr>
<td>3.20</td>
<td>Example anti-bunching histogram from background.</td>
<td>114</td>
</tr>
<tr>
<td>3.21</td>
<td>Example anti-bunching data in the ABEL trap.</td>
<td>116</td>
</tr>
<tr>
<td>3.22</td>
<td>Two-color trapping with one pair of acousto-optic deflectors.</td>
<td>118</td>
</tr>
<tr>
<td>3.23</td>
<td>Architecture of ABEL trap data processing system, from acquisition to visualization.</td>
<td>120</td>
</tr>
</tbody>
</table>
4.1 Experimental setup for the single-dye experiments. ......................... 122
4.2 Example traces of trapped single fluorophores. ......................... 125
4.3 Representative traces of trapped single fluorophores with simultaneously recorded excited-state lifetime and emission spectrum. ............... 127
4.4 Brightness, lifetime and emission spectrum analysis of a single Atto 647N molecule with one unique emissive state. ......................... 129
4.5 Brightness, lifetime and emission spectrum analysis of a single Atto 647N molecule that switches between multiple emissive states. ............... 130
4.6 Clustering analysis of Atto 647N emission states. ....................... 132
4.7 Transitions in single Atto 647N molecules. ....................... 134
4.8 Rates of spectral fluctuation in Atto 647N. ....................... 135
4.9 Example photodynamics of Atto 633 at the single-molecule level. ........ 137
4.10 Clustering analysis of the emissive states of Atto 633. .................. 139
4.11 Alexa 647 lacks significant photodynamics. ..................... 141

5.1 Experimental design to probe the individual pigment sites on allophyco- cyanin in solution. .................................................. 146
5.2 Example single-molecule traces of APC monomers and trimers. .......... 150
5.3 Multi-parameter mapping of the different emissive states along the photo- bleaching pathway of APC monomers. .................................. 153
5.4 Photophysical model of the APC monomer. .......................... 156

6.1 Real-time estimation of single-molecule diffusion coefficient and electroki- netic mobility. ................................................... 163
6.2 $D - \mu$ mapping of different single molecules. ....................... 165
6.3 Resolving a three-component mixture by single-molecule transport coeffi- cients. ...................................................... 168
6.4 Resolving monomers and trimers of APC in solution along the dissociation pathway. ........................................... 171
6.5 Single-molecule classification of the trimer-monomer mixture of APC. .. 173
6.6 Real-time sensing of DNA binding/unbinding using fluctuating motion pa- rameters. ...................................................... 174
6.7 DNA hybridization kinetics visualized by dynamic changes in transport coefficients in solution. ............................................ 176
6.8 Extracting binding kinetics from dwell-time analysis and correlation functions of the fluctuating transport coefficients. ........................ 177
6.9 Accessible parameter space of binding kinetics. ................................................. 179
B.1 Verifying the PCD activity. ................................................................. 189
C.1 Joule heating in the ABEL trap. ........................................................... 192
D.1 Return data structure from the FPGA and example Matlab code to read the data files. ................................................................. 195
Chapter 1

Introduction

At the molecular level, the processes of biology are carried out by biomolecules in an environment filled with water. These molecules, mostly proteins, RNA and DNA, are highly dynamic entities that change shapes, move from place to place and communicate with one another as they fulfill their biological roles. A deep understanding of the physical and chemical principles behind these processes can provide keys to understand biology. In the early 1990s, biophysicists invented tools to study biology at the single-molecule level. Such a breakthrough made it possible to directly watch the dynamic evolution of individual biomolecules when they are at work (Noji et al., 1997; Yildiz et al., 2003). The trick is simply to make the object fluorescent so that emission from the object can be detected above any potential interfering backgrounds (Betzig et al., 1993). Then one can study a natively fluorescent biomolecule like a green fluorescent protein (Dickson et al., 1997). Or one can attach a tiny fluorescent molecule to the biomolecule of interest: the probe acts as a light source that allows visualization and sensing of the minuscule amount of motion and disturbance at the molecular scale.

Although single-molecule methods have enjoyed a near-exponential growth in the past two decades (Moerner, 2007), one important technical aspect calls for improvement for wider biological applications: Most of the existing techniques rely on attaching the molecule of interest to a surface or confining it somehow in a vesicle or other tiny volume because otherwise, nanometer-sized molecules undergo rapid Brownian motion in aqueous solution that precludes extended investigation. However, the act of immobilization not
only puts the molecule into an unnatural environment but can also perturb its biological function.

This dissertation focuses on developing methods to probe the dynamical behaviors of single molecules directly in aqueous solution using a feedback-based molecular trap to suppress Brownian motion (Cohen et al., 2005a; Cohen et al., 2006). The methodology developed here not only allows fluorescence spectroscopy to be performed on single molecules without surface perturbation or diffusion-time limitations, but also enables new variables to be extracted that probes interactions between individual molecules, an aspect which is difficult to sense by other single-molecule means. Each method is demonstrated by applications to biological systems.

This chapter aims to provide the general background and motivation of my graduate work.

1.1 Single-molecule fluorescence spectroscopy

Since the initial breakthrough by Moerner et al. (1989) and Orrit et al. (1990), single-molecule optical spectroscopy has quickly evolved from a low-temperature physical tool that unveils new dynamical processes (e.g. spectral diffusion (Ambrose et al., 1991), antibunching (Basché et al., 1992), magnetic resonance of a single spin (Wrachtrup et al., 1993; Kohler et al., 1993)) on individual molecules, to a set of easy-to-use, ubiquitous methods (Walter et al., 2008) that probe nano-scale physics (Moerner et al., 1999), chemistry (Cordes et al., 2013) and biology (Weiss, 2000) at the fundamental limit. Recently, single-molecule imaging has enabled a new wave of technological advances in imaging beyond the diffraction limit (Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006).

1.1.1 Basic principles

Among the vastly diverse single-molecule methods, optical detection of fluorescence provides a simple (Moerner et al., 2003) yet immensely powerful route to observe individual molecules. Fluorescence emission is a three-step process (Figure 1.1): first, the molecule absorbs a photon from the excitation source and thereby is promoted from the electronic
singlet ground state ($S_0$) to the first singlet excited state manifold. Second, fast internal relaxation brings the molecule to the lowest level of the electronic excited state ($S_1$). Third, spontaneous emission returns the molecule to the ground-state manifold and a fluorescent photon that is red-shifted from the excitation laser is generated. Note that this radiative decay process competes kinetically with a number of pathways, including inter-system crossing (ISC) to a triplet state and non-radiative decay to the ground state, so to have a strongly fluorescent molecule, these competing pathways should not be dominant. Once in the ground state, the molecule is ready to be excited again and repeat the whole process. On the other hand, a single molecule will not be able to repeat the excitation-emission cycle forever and inevitably loses the ability to emit light permanently. This process is known as photo-bleaching and can be caused by irreversible alteration of the chemical structure (for example, via intersystem crossing to the long-lived $T_1$ followed by further excitation to a higher, reactive triplet state $T_n$). A fluorescent molecule can be parameterized by the absorption cross-section ($\sigma_{01}$), excited-state lifetime ($\tau_{10}$, typically $\sim 1$ ns), the fluorescence quantum yield ($\Phi$, i.e. the branching ratio of the radiative decay rate, $\Phi = k_{10}^R / (k_{10}^R + k_{10}^{NR})$, typically $0.1 - 1$), the inter-system crossing rate ($k_{ISC}$), the triplet lifetime ($\tau_T$, typically $\sim \mu$s-ms) and the photo-bleaching rate ($k_{bleach}$). A “good” fluorophore for single-molecule studies generally has (at room temperature) a large absorption cross-section ($\sigma_{01} \sim 10^{-16}$ cm$^2$ or equivalently, a molar absorptivity $\varepsilon_{01} \sim 10^6$ M$^{-1}$ cm$^{-1}$), a high quantum yield ($\Phi > 0.5$) and a low photobleaching rate ($k_{bleach} \ll 1 / \tau_{10}$).

To detect a single molecule by its emitted fluorescence, two conditions must be met (Moerner et al., 1999). First, the concentration of the sample needs to be low enough so that the average number of fluorescent molecules in the probing volume is on the order of one. Second, the background signal, which arises from a) residual light from the excitation laser and b) millions or billions of surrounding host molecules, needs to be kept sufficiently low. The key enabling factor leading to single-molecule sensitivity for fluorescence detection is the red shift in detection wavelength compared to the excitation wavelength, so that long-wavelength-pass filters can be used to sufficiently suppress any scattered excitation photons. The required suppression ratio of the pumping laser can be estimated as the follows. Given a laser beam with a spot size of area $S$ illuminating a single molecule, the fraction of light absorbed is given by $\sigma_{01}/S$. With a typical absorption cross-section of
Figure 1.1: Jablonski diagram of the energy levels and processes involved in fluorescence emission (Lakowicz, 2006).
$\sigma_{01} \sim 10^{-16} \text{cm}^2$ and a diffraction-limited spot size ($S \sim 0.1 \mu\text{m}^2$), we have $\sigma_{01}/S \sim 10^{-7}$. In other words, even if every absorption event produces a fluorescent photon (i.e. $\Phi = 1$), the signal only contributes to about ten-millionth of the optical field at the sample plane. This means that a filter must suppress the laser field to $10^{-7}$ and pass the fluorescence emission from the molecule with near-unity transmission. Today’s interference filters adequately fulfill this requirement. On the other hand, it is much more difficult to suppress the background component that has a emission spectrum similar to that of the molecule, for example, autofluorescence and/or Raman scattering of the mounting medium. We discuss this issue in Section 1.1.3.

1.1.2 Single-molecule label as a probe for biomolecules.

One of the early observations of single-molecule optical spectroscopy was the ultrasensitivity to nano-environments. For example, interactions with the host molecules can lead to discrete jumps of the zero phonon line absorption (Ambrose et al., 1991; Reilly et al., 1993) at low temperatures. Recently, it was reported that local mechanical stress induces shifts in the emission spectrum of a single molecule at room temperature (Stöttinger et al., 2014). A single molecule can thus be utilized as a nanoscale reporter in diverse applications from the physics of phase transitions (Zondervan et al., 2007) to the conformational fluctuations of a biomolecule (Weiss, 2000).

As a classical example of using a single-molecule label to probe bio-molecules, Yang et al. (2003) used the time-dependent excited-state lifetime of a fluorescent cofactor to probe distance fluctuations within the protein flavin reductase. In their work, the excited-state lifetime of the fluorescent cofactor (flavin) directly senses the distance to a nearby tryptophan residue as a result of electron transfer-induced fluorescence quenching. From the distribution of measured lifetimes, the authors directly estimates the energy landscape along the designed probing degree of freedom. Further, by analyzing the timescale of the dynamic lifetime fluctuations, the authors examined the mode of conformational fluctuations and concluded that the data supports the physical picture of a rugged energy landscape.
Figure 1.2: Measurement scheme (left) and example data (right) of (a) free diffusion, (b) surface immobilization (side view) and (c) ABEL trap-based single-molecule experiments. Red arrows indicate photobleaching events. In the ABEL trap case, the blue arrow indicates diffusion into the trapping region.
### Table 1.1: Diffusion coefficients of example biomolecules

<table>
<thead>
<tr>
<th>Reference object</th>
<th>Molecular weight (kDa)</th>
<th>D (µm$^2$/s$^{-1}$)</th>
<th>Diffusion rms distance (1 ms window)$^\dagger$</th>
<th>Diffusion time$^\ddagger$ (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 nm sphere</td>
<td></td>
<td>50$^a$</td>
<td>0.45</td>
<td>1.3</td>
</tr>
<tr>
<td>10 nm sphere</td>
<td></td>
<td>25$^a$</td>
<td>0.32</td>
<td>2.5</td>
</tr>
<tr>
<td>100 nm sphere</td>
<td></td>
<td>2.5$^a$</td>
<td>0.10</td>
<td>25.2</td>
</tr>
<tr>
<td><strong>Proteins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lysozyme</td>
<td>14</td>
<td>112$^b$</td>
<td>0.67</td>
<td>0.56</td>
</tr>
<tr>
<td>Allophycocyanin</td>
<td>104</td>
<td>49$^c$</td>
<td>0.44</td>
<td>1.3</td>
</tr>
<tr>
<td>IgG</td>
<td>160</td>
<td>40$^b$</td>
<td>0.40</td>
<td>1.6</td>
</tr>
<tr>
<td>Ferritin</td>
<td>450</td>
<td>36$^b$</td>
<td>0.38</td>
<td>1.7</td>
</tr>
<tr>
<td>Ribosomes 30S</td>
<td>900</td>
<td>30$^b$</td>
<td>0.35</td>
<td>2.1</td>
</tr>
<tr>
<td><strong>Oligonucleotides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λ-phage DNA</td>
<td>48502 bp</td>
<td>0.8$^d$</td>
<td>0.06</td>
<td>78</td>
</tr>
<tr>
<td>10nt-dsDNA</td>
<td>~6</td>
<td>98$^e$</td>
<td>0.63</td>
<td>0.64</td>
</tr>
<tr>
<td>10nt-ssDNA</td>
<td>~3</td>
<td>205$^f$</td>
<td>0.91</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Fluorescent labels</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantum Dot</td>
<td></td>
<td>20$^g$</td>
<td>0.28</td>
<td>3.1</td>
</tr>
<tr>
<td>eGFP</td>
<td>33</td>
<td>95$^h$</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td>Alexa 647</td>
<td>1.3</td>
<td>330$^i$</td>
<td>1.1</td>
<td>0.19</td>
</tr>
<tr>
<td>Atto655</td>
<td>0.63</td>
<td>426$^i$</td>
<td>1.3</td>
<td>0.15</td>
</tr>
</tbody>
</table>

$^a$ Equation 1.1 $^b$ (Tyn et al., 1990) $^c$ (MacColl et al., 2003) $^d$ (McHale et al., 2009) $^e$ (Bonifacio et al., 1997) $^f$ (Stellwagen et al., 2003) $^g$ (McHale et al., 2007) $^h$ (Petrášek et al., 2008) $^i$ (Müller et al., 2008) $^j$ (Dertinger et al., 2007) $^\dagger$ calculated using $\Delta x = \sqrt{4D\Delta t}$, for $\Delta t = 1$ ms, two-dimensional diffusion assumed $^\ddagger$ calculated using $\Delta t = d^2/4D$, for $d = 0.5 \mu$m, the width of a diffraction limited laser spot
1.1.3 Methods for measuring single freely-diffusing molecules

Most biological processes take place in an aqueous-like environment so naturally, methods to detect single molecules as they freely diffuse in solution have been developed (Shera et al., 1990). In its most common realization (Figure 1.2a), a laser beam is tightly focused into a solution sample and a sensitive detector monitors signal from the focal region (Shera et al., 1990; Nie et al., 1994; Goodwin et al., 1996). The concentration of the fluorescent molecules is kept extremely low (~10 pM) so that most of the time, there is no molecule in the excitation region. However, occasionally diffusion or Brownian motion randomly drives a single molecule into the excitation spot and fluorescence photons are emitted as a result of the molecule-laser interactions. If the required signal-to-background conditions are met by the detection system, those single molecule events can be observed. However, also due to Brownian motion, the molecule will soon escape the excitation region so the signal from a single molecule in this mode of measurement is a short spike in brightness (Figure 1.2a example trace). The typical duration of these spikes can be estimated by the object’s diffusion coefficient and the spot size. A biomolecule’s diffusion coefficient (see Table 1.1 for typical values) is related to size by the Stokes-Einstein relation,

\[ D = \frac{k_B T}{6\pi \eta r} \]  

where \( r \) is the hydrodynamics radius, \( k_B \) is the Boltzmann constant, \( \eta \) is the solvent viscosity and \( T \) is the Kelvin temperature. Thus, for \( D = 50 \mu \text{m}^2 \text{s}^{-1} \) and a diffraction-limited spot size of 0.5 \( \mu \text{m} \) (~1 \( \mu \text{m}^3 \) excitation volume), we get a typical (2D) transit time of \( \Delta t = d^2/4D = 1 \text{ms} \) (Table 1.1).

One millisecond is a very short time. As far as biology is concerned, many processes happen on the seconds timescale so that a 1 ms observation window cannot follow the dynamics of those slow processes but only yields static snapshots. On the other hand, from a technical point of view, the amount of photon information from a single 1 ms burst is extremely limited (typically < 100 photons). Consequently, even for fast processes (e.g. ns to \( \mu \text{s} \)) that a 1 ms window can in principle observe, dynamic information is often buried in a high level of shot noise.
CHAPTER 1. INTRODUCTION

Correlation-based analysis  A solution to the signal-to-noise problem is to use correlation analysis to extract information averaged over many “statistically similar” molecules. Such measurements can be viewed as the single-molecule limit of Fluorescence Correlation Spectroscopy (FCS), a well-established technique (Magde et al., 1972) in the 1970s.

Over the years, a number of technical improvements in FCS in the single-molecule context have made it an easy-to-use and powerful technique (Haustein et al., 2004). For example, by analyzing the averaged transit time of successive single molecules passing through the laser focus, the diffusion coefficient can be extracted (Dertinger et al., 2007; Petrášek et al., 2008). In a similar manner, flow velocities can be characterized (Magde et al., 1978). Brightness information can be obtained by a photon counting histogram approach (Chen et al., 1999; Huang et al., 2004) and spectral information can also be learned (Previte et al., 2008; Bestvater et al., 2010). FCS is a suitable technique to study the fast (μs) fluorescence dynamics that might originate from photophysics (Schwille et al., 2000), base-pairing (Wallace et al., 2000) or protein conformation changes (Nettels et al., 2007). Nevertheless, it is important to keep in mind that all these examples only extract ensemble-averaged properties.

Burst-by-burst analysis  To avoid averaging over many molecules and really study dynamics of single copies in solution, we need to extract information from each diffusing burst. To increase the amount of information in individual bursts, one strategy is to increase the spot size of the laser so that a single molecule can spend more time in the excitation volume. However, as discussed previously, doing so incurs a penalty on the signal-to-background ratio. To estimate the maximum allowable excitation volume, consider Raman scattering of the water molecules as the dominant source of background. (In practice, Stokes Raman photons are difficult to filter out spectrally, since these are also red-shifted, like the fluorescence photons.) For single-molecule sensitivity, we want the collective Raman signal from all the water molecules in the excitation volume to be smaller than the fluorescence signal from a single molecule, i.e.

\[
\frac{V_f \cdot \rho_w \cdot N_A \cdot \sigma_w^R}{M_w} \leq \sigma_{01}^f \cdot \Phi
\]  

(1.2)
where $\sigma^R_w \approx 10^{-28} \text{cm}^2$ is the cross-section of the 3400 cm$^{-1}$ Raman process of water (Faris et al., 1997), $N_A$ is Avogadro’s number, $V_f$ is the excitation volume, $\rho_w$ and $M_w$ are the density and molar mass of water. If we assume a single molecule with an absorption cross-section $\sigma_{01}^f \approx 10^{-16} \text{cm}^2$ and a quantum yield $\Phi = 1$, we have $V_f \leq 30 \mu \text{m}^3$. In reality, there are other sources of fluorescent background which further lowers this upper bound. Consequently, there is not much room for increasing the excitation volume in the confocal geometry before losing single-molecule sensitivity. An example of this approach is provided by Diez et al. (2004), which used a 7 $\mu \text{m}^3$ excitation volume to study the millisecond rotation dynamics of the FoF1 ATP synthase embedded in ~120 nm lipid vesicles. Although the enlarged laser spot allowed the authors to look at a few events ~100 ms in duration, these events were extremely rare.

Given the limitations of diffusion-based single-molecule detection to study slow (>10 ms) dynamics, most studies focus on getting static snapshots of individual bursts. In these measurements, the excitation power is generally set high in order to extract the maximum amount of photons from each transient. Coupled with other spectroscopic readouts, this mode of measurement provides a powerful tool to separate (statically) heterogeneous populations. For example, Eggeling et al. (1998) was the first to extract the excited-state lifetime of each burst and used it to characterize a three component mixture. Later, the same group pioneered the use of fluorescence anisotropy from individual bursts to identify molecules of different rotational mobilities (Schaffer et al., 1999). Forster energy transfer (FRET) between an donor dye and an acceptor dye on the same molecule provides a distance sensitive readout for molecule-by-molecule classification in solution (Deniz et al., 1999) and is particularly useful in studying protein folding (Schuler et al., 2002). The capabilities of burst-integrated fluorescence methods have been reviewed by (Kühnemuth et al., 2001). Nevertheless, it is important to emphasize that these measurements can only resolve static (not dynamic) heterogeneity of single molecules.

In summary, free-diffusion based single-molecule detection is limited to either extracting ensemble-averaged, fast dynamics or characterizing the static heterogeneity of a sample.
1.1.4 Methods for measuring single immobilized molecules

The most common approach to remove the ~ms time limitation imposed by diffusion is to immobilize single molecules of interests to a surface. In this way, one can continuously monitor the behavior of individual molecules for a long time until photobleaching (Figure 1.2b), either by confocal imaging of one spot or by wide-field fluorescence imaging. A variety of immobilization methods exists in the literature, including nonspecific electrostatic interactions (Bopp et al., 1999), polymer-assisted adsorption (Krüger et al., 2010), porous gels (Dickson et al., 1997), polymer matrices (Garcia-Parajo et al., 2000), biotin-streptavidin linkers (Rasnik et al., 2005) and surface attachment of oligonucleotides.

Immobilization methods face the question of whether they perturb the functions of the restrained biomolecules. Although in many investigations, the effect of the surface was concluded to be minimum (Rasnik et al., 2005; Lamichhane et al., 2010), immobilization can still be deleterious in many cases. In the work by Friedel et al. (2006) using molecular dynamics simulations, it was shown that the folding dynamics of a model protein were significantly altered upon tethering and that the effect depends on the location of the tether. Experimentally, loss of enzyme activities was frequently found in surface-based immunoassays (Butler, 2000) and recently, in a systematic study of β-Galactosidase using different immobilization strategies (Liu et al., 2013b). These finding are not really surprising given that proteins are delicate molecular machines whose function critically depend on a correct three-dimensional structure, proper solvation and, in the case of an enzyme, an unobstructed active site. The act of immobilization could affect function in a variety of ways such as reducing (or eliminating) rotational mobility, altering the hydrogen bonding network, partial denaturation and introducing steric hindrance. Another effect relevant to single-molecule studies is that immobilization could induce heterogeneity into the sample. For instance, one could imagine that molecules on the surface are embedded in slightly different nano-environments, with different orientations or secondary contacts to the surface. These effects would artificially broaden the measured single-molecule distributions. In the Moerner group, we have found that the behavior of many photosynthetic antenna proteins are significantly more homogeneous in solution compared to previous measurements on surfaces (for example, see Goldsmith et al. (2010)).
Nano-container encapsulation  Immobilization of a single molecule can also be achieved by trapping it in nano-sized containers. In this way, the molecule’s translational and rotational degrees of freedom are minimally affected and the physical constraint ensures long-time probing of single-molecule behavior. Different types of encapsulation methods have been developed, including lipid vesicles (Boukobza et al., 2001; Cisse et al., 2007), nanofabricated nano-containers (Rondelez et al., 2005; Rissin et al., 2006; Shon et al., 2012) or optically trapped water droplets (Reiner et al., 2006). Another advantage of nano-containers is the capability of studying biomolecular interactions at the single-molecule level (Keller et al., 2012; Shon et al., 2012; Cisse et al., 2012). On the other hand, nano-entrapment has a number of disadvantages, including low yield of capturing only one molecule, difficulty of providing buffer exchanges and nano-confinement effects (Rubinovich et al., 2013) which depend on the degree of confinement and shifts single-molecule kinetics away from bulk values.

1.2 The Anti-Brownian Electrokinetic trap

An effective strategy to avoid immobilization of single molecules is to “trap” them in aqueous solution. Indeed, numerous methods have been developed to meet this challenge. Perhaps the most well known example is the optical tweezers approach (Ashkin et al., 1986; Moffitt et al., 2008), which uses the optical gradient forces from the intense electromagnetic radiation of a focused laser beam to hold a single object in place. However, the restoring forces in an optical trap are proportional to the polarizability of the trapped object, which scales as $\propto d^3$ (with $d$ the diameter of the object). This steep scaling law makes trapping sub-50 nm objects (e.g. biomolecules) extremely challenging. It was not until recently that several groups successfully trapped relatively large biomolecules using enhanced local fields from slot waveguides (Yang et al., 2009; Chen et al., 2012) and plasmonic structures (Pang et al., 2012; Berthelot et al., 2014). Another limitation of applying optical tweezers on single biomolecules is the potential heating produced by the intense trapping laser (Peterman et al., 2003).

Having realized that the polarizability-induced optical forces are too weak for manipulating nanometer-sized objects, other forces have been exploited for trapping. For example,
direct electrostatic interactions can be used to levitate charged nanoscale objects in nano-
channels (Carlson et al., 2010; Krishnan et al., 2010). Trapping of ~100 nm objects has also
been demonstrated using thermophoretic (Braun et al., 2013) and dielectrophoresis (Kuzyk
et al., 2008) forces in solution. Finally, electrokinetic forces, which result from the direct
action of an external electric field on charged molecules (including both the target and its
surrounding solvent molecules), can be used to manipulate nano-objects in solution, and
this forms the basis of the ABEL (anti-Brownian Electrokinetic) trap technology, the focus
of this dissertation.

1.2.1 Basic principle

The basic idea of the ABEL trap is remarkably simple: we continuously track the spatial
position of the molecule, calculate its Brownian displacement from a predefined trap center,
and manipulate the molecule back to the center by electrokinetic forces. In other words, if
thermal motion drives the target molecule to the right, we apply an electric field in solution
to move the molecule back to the left, and vice versa. Rapid updating of the feedback cycle
(1,000-50,000 times per second) keeps the target molecule near the center of the probing
region until the target ceases to emit. The closed-loop feedback mechanism effectively
creates a trapping (quadratic) potential (Cohen, 2005) similar in shape but can be much
stronger in magnitude compared to optical tweezers forces for nanoscale objects.

A simulated trap behavior is illustrated in Figure 1.2c. In conventional solution-phase
single-molecule detection methods (plotted in Figure 1.2a for comparison), the ephemeral
interactions between the diffusing target molecule and a fixed excitation laser spot result
in transient intensity spikes (~1 ms in duration) on the detector, as previously described.
Note that the height of those intensity spikes is not a good measure of object brightness
and neither is the number of those spikes a good measure to count molecules (because the
same molecule can re-enter the laser spot and create additional bursts). Due to the inho-
mogeneous Gaussian profile of a focused laser beam (green spot) and the randomness of
the trajectory (red), even a single molecule with a constant brightness can produce multiple
spikes of vastly different amplitudes, as shown in the simulated trace of Figure 1.2a. In the
Figure 1.3: Basic principles of the Anti-Brownian ELectrokinetic (ABEL) trap. The ABEL trap works as a closed-loop feedback system consisting of three major components: a microfluidic chamber (3D cartoon in blue, 2D projection of the center region on left), a position sensitive detector and a feedback control module. The system continuously tracks the position of a single target molecule (represented here by a piece of double-stranded DNA) via fluorescence emission (red cloud) and calculates a pair of voltages that, when applied to the microfluidic chamber along both Cartesian coordinates, produce electrokinetic forces (black arrow, with the individual components as shaded arrows) that push the molecule towards the center (black “+”). This feedback action is repeated (four example time slices are shown inside the feedback control electronics box) at high speed to maintain trapping of a single target molecule.
ABEL trap, once a single molecule diffuses into the trapping region, the feedback mechanism keeps it under the interrogation of the excitation laser until complete photobleaching, allowing the molecule to be probed for ~1 s (1,000 times longer than the case in Figure 1.2a, limited by photophysics). Careful design of the excitation optics creates a time-averaged flat-top excitation profile so that brightness of the object can be reliably inferred from the intensity traces (black).

Three key elements enable the ABEL trap technology: the microfluidic chamber, the position estimator and electrokinetics. Below I give a non-technical overview the microfluidic chamber and position estimator. Details will be discussed in Chapter 3. Electrokinetic forces is discussed in Section 1.2.2.

**Microfluidic chamber**  The ABEL trap experiments are carried out in a microfluidic chip fabricated from PDMS or quartz and sealed with a coverslip (Figure 1.3). The center region of the cell holds the liquid sample in a thin sheet with dimensions of 20µm × 20µm (in x-y) × 0.5µm (in z). This means that in the current implementation, confinement in the z-direction is provided by the upper and lower walls of the chip rather than by feedback forces, although trapping in all three dimensions has recently been reported by King et al. (2013). Four electrodes are inserted into the deep (~15 µm) reservoirs millimeters away from the “trapping region”. It has been shown that this arrangement of electrodes creates uniform x-y electric fields across the central trapping region (Cohen et al., 2005b).

**Position estimation**  The initial incarnation of the trap used a high sensitivity camera to localize the molecule by centroid fitting of its fluorescent image (Cohen et al., 2005a; Cohen et al., 2006). Although straightforward to implement, the “software-based” trap was limited by the delay time (~4 ms, image acquisition and processing) to only trap relatively large objects (>20 nm in diameter). Later, an improved sensing scheme using a revolving excitation laser was implemented to have a delay as short as 25 µs (Cohen et al., 2008). The position estimation method was further revamped as part of the work described in this thesis (Chapter 2). I developed a real-time tracking and feedback strategy that has minimum delay (<1 µs) and achieves information-limited estimation accuracy.
1.2.2 Electrokinetics

When a voltage is applied to a microfluidic channel, a corresponding flow will develop. This phenomena, known as electrokinetics (Hiemenz et al., 1997), provides the feedback kicks in the ABEL trap system. Electrokinetics is a combined effect of electrophoresis and electroosmosis, which are now described in more detail.

Electrophoresis  Electrophoresis describes the movement of a charged particle relative to its surrounding solvent under the influence of an externally applied electric field in solution. Initially, the particle feels the electrostatic force and begins to accelerate. Meanwhile, the viscous drag, which is proportional to particle velocity, starts to increase and eventually balances the Coulomb force, i.e.

\[ qE = 6\pi \eta rv_e \]  \hspace{1cm} (1.3)

where the right-hand-side of Equation 1.3 is the expression for the Stokes drag (the same force that defines diffusion in Equation 1.1). The steady-state velocity of the particle is

\[ v_e = \left( \frac{q}{6\pi \eta r} \right) E \]  \hspace{1cm} (1.4)

where the proportionality factor between velocity and electric field is the electrophoretic mobility \( \mu_e \). If we recall from fundamental physics that the electric potential at the surface of a charged particle is \( \zeta = q/(4\pi \varepsilon r) \), we can express the electrophoretic mobility as

\[ \mu_e = \frac{2\varepsilon \zeta}{3\eta} \]  \hspace{1cm} (1.5)

The above derivation ignores the behavior of the mobile ions in the buffer solution. In reality, a charged particle would slightly distort its nearby ion distribution by attracting ions of the opposite charge and repelling ions of the same charge. This local gradient of ion imbalance, called the “electrical double layer”, decays exponentially at increasing distances away from the particle (Figure 1.4a). The characteristic length scale of the double-layer is called the “Debye length” \( \kappa^{-1} \), which is an important parameter in shaping the electrophoretic response. In aqueous solution, the Debye length is controlled by the ionic
strength

\[ \kappa^{-1} (\text{nm}) = \frac{0.307}{\sqrt{I(M)}} \]  

(1.6)

and in typical physiological conditions (~100 mM, is about ~1 nm (Figure 1.4b). Basically, the counterions within the Debye length create a “retardation force” by moving in the opposite direction of the particle and they shield the electric potential of the particle. By taking into all these effects, the electrophoretic mobility can be modeled as (Henry, 1931)

\[ \mu_e = \frac{2 \epsilon \zeta}{3\eta} f_H(r\kappa) \]  

(1.7)

where \( f_H(r\kappa) \) is a correction factor that depends on the particle size and the Debye length and \( \zeta \) is the electric potential at the boundary of the double layer (the “zeta” potential), which itself is a complex function of the Debye length and the amount of charges that the particle carries (Doane et al., 2012). In practice, it is extremely challenging to predict electrophoretic mobility from first principles, but it is generally accepted to calculate the zeta potential from measured mobility. Also, it is difficult to calculate the amount of surface charges given a calculated zeta potential. For more details, see the textbook by Hiemenz et al. (1997).

**Electroosmosis** Electroosmosis describes the movement of the bulk liquid relative to charged surfaces under the influence of an external electric field. Here, the immobile surface charges in a microfluidic geometry create a gradient of ion imbalance near the interface and when an electric field is applied, the excess amount of counterions within the Debye layer feel a net Coulomb force. As the counterions begin to move, they first drag the liquid within the double layer and eventually the whole bulk liquid with them. The result is a flow profile shown in Figure 1.5: the flow velocity is zero at the solid-liquid interface due to the no-slip boundary condition, gradually increases within the double layer and is a constant in the bulk liquid. Note that this profile is very different from a pressure-driven flow, which has a parabolic shape.
In bulk liquid (away from the double layer), the electroosmosis velocity is homogenous and proportional to the electric field. The mobility in this case is given by

\[ \mu_{eo} = \frac{\varepsilon \zeta}{\eta} \]  

(1.8)

where \( \zeta \) is now the electric potential at one Debye-length away from the interface. Equation 1.8 is simply the special case of Equation 1.7 with \( r \to \infty \) \( (f(r\kappa) \approx 3/2) \). The flow velocity within the double layer, on the other hand, is highly inhomogeneous. In ABEL trap applications, it is desirable to keep the double layer thickness much smaller than the channel depth (~600 nm), so that the flow velocity is maximally independent of the particle’s axial position. From Figure 1.4b, this requires keeping the ionic strength higher than about 1 mM, which corresponds to \( \kappa^{-1} < 10 \text{nm} \). (On the other hand, Joule heating limits the maximum ionic strength in the buffer, which will be discussed in Appendix C.) An interesting case arises when the ionic strength is extremely low (e.g. in deionized water,
CHAPTER 1. INTRODUCTION

Figure 1.5: The electrokinetic response is a combined effect of electrophoresis (EP) and electroosmosis (EO).

\[ I \approx 10^{-7} \text{M}, \] in which the two double layers from the upper and lower surfaces overlap (\( \kappa^{-1} \approx 0.9 \mu\text{m} \)).

**Transient response of electrokinetics** I now briefly discuss the response time of electrokinetic flow: when we apply an electric field to push the molecule towards the target position, what is the time delay before the object acquires the steady-state velocity?

In electroosmosis flow, the transient response consists of the following steps: first, an electric double layer forms near the interface. This step has a characteristic time given by the diffusion of ions over the length of the double layer,

\[ t_{dl} \sim \frac{2D_{\text{ion}}}{(\kappa^{-1})^2} \quad (1.9) \]

which is as fast as 100 ns for a 10 nm double layer and \( D_{\text{ion}} = 500 \mu\text{m}^2\text{s}^{-1} \). Next we consider the time to transfer momentum from the double layer to the bulk fluid. This step can be described by the start up of Couette flow (Kirby, 2010; Minor et al., 1997) and has a characteristic time of

\[ t_{EO} \sim \frac{\rho d^2}{\eta} \quad (1.10) \]

where \( \rho \) and \( \eta \) are the density and viscosity of the liquid phase and \( d \) is the size of the channel. With \( d \approx 0.5 \mu\text{m} \), we have \( t_{EO} \approx 2 \mu\text{s} \). Although this is still a very fast response time,
note that this is only true for a sub-micron channel size. As the channel depth increases, this hydrodynamic relaxation time increases quadratically! Finally, the momentum needs to be transferred to the particle, this takes

\[ t_p \sim \frac{\rho r^2}{\eta} \]  

which is extremely fast (~10 ns for a \( r = 100 \text{nm} \) particle).

Now we turn to electrophoretic response. Here, the same double-layer formation time (Equation 1.9, ~100 ns) applies as well. A slower process that has to do with balancing the concentration polarization (Minor et al., 1997) subsequently follows with a characteristic time

\[ t_{cp} \sim \frac{r^2}{D_{ion}} \]  

which is about 100 ns for a 10 nm objects. Thus, combining all those effects, the electrokinetic response in the ABEL trap takes place within about a \( \mu \text{s} \) and can be treated as instantaneous compared to the time between measurements (20~200 \( \mu \text{s} \)).

**Practical considerations** The ABEL trap can utilize either electrophoretic or electroosmotic motion to drive particles. In many cases, however, we have the situation where the two types of flows can counteract each other. Figure 1.5 illustrate one such example. If the two flows are of comparable strengths, the effective mobility of the object is near zero, which makes the object difficult to trap. It is sometimes desirable to eliminate electroosmosis for trapping charged objects. On the other hand, the electroosmotic flow is of particular importance to the ABEL trap technology as it enables trapping of neutral objects. It is also possible to engineer both the polarity and magnitude of the electroosmotic flow in the microfluidic chip by coating the surfaces with charged polymers. Tuning of electroosmotic flow will be discussed in Chapter 3.

Another practical consideration is resistive heating of the buffer solution, which has been found to be significant in electrokinetically driven microfluidic devices (Erickson et al., 2003), especially when the buffer contains high concentrations of salt. The amount of temperature rise in the ABEL trap was assessed by experiments described in Appendix C.
CHAPTER 1. INTRODUCTION

1.3 Feedback tracking of a single object

Feedback tracking is another class of methods designed to investigate single diffusing molecules in solution without immobilization. Here the basic idea is similar to the ABEL trap: a position sensing apparatus monitors the motion of a target object in real time and instead of pushing the object back, the feedback mechanism translates the detection field-of-view along with the target. This approach can be traced back to Howard Berg’s apparatus (Berg, 1971) to track individual swimming bacteria. Parallel to the development of the ABEL trap, efforts by many groups (Berglund et al., 2005; Cang et al., 2006; Lessard et al., 2007; Juette et al., 2010) pushed the feedback tracking sensitivity to the sub-100 nm regime using either fluorescence emission or scattering.

The most charming feature of feedback tracking is the straightforward extension to live cells, as beautifully demonstrated recently (Welsher et al., 2014; Keller et al., 2014). On the other hand, the technology is still limited to relatively large objects (>20 nm), because it is difficult to move a massive piezoelectric stage fast enough to follow nanoscale objects.

1.4 Scope and organization of this dissertation

My contribution to the ABEL trap technology started with designing an optimal feedback strategy to enable trapping of fast and dim objects (Chapter 2, Publication 1). This work was inspired by the theoretical treatment of the feedback tracking apparatus in the Mabuchi lab (Berglund et al., 2004; Berglund et al., 2006). I showed by simulation analysis that the new strategy would represent a major boost in the trap’s ability to handle sub-10 nm objects.

I then built two ABEL traps with an improved optical design and with the optimal feedback algorithm implemented on an Field-Programmable-Gate-Array (FPGA) platform, and then demonstrated that the new algorithm is indeed significantly better than previous versions (Chapter 3, Publication 2). I also developed a simple method to probe transport properties of individual trapped objects. Around the same time, the Cohen lab published their own ABEL trap system with a feedback algorithm similar to my initial proposal and beautifully demonstrated trapping of single fluorophores and transport analysis of individual trapped objects (Fields et al., 2011).
Figure 1.6: Structure and organization of this dissertation with chapter number in parenthesis.
CHAPTER 1. INTRODUCTION

After three years of instrumentation work, I began to think more about the unique scientific problems that can be addressed by the apparatus. I started working on the photodynamics of the photosynthetic antenna protein Allophycocyanin (APC, a project initiated by Dr. Randall Goldsmith). In order to fully elucidate the molecular mechanism behind the switching behavior of APC, I wanted to extract more information from multiple fluorescence parameters. With this goal in mind, I enriched the ABEL trap setup with multi-dimensional fluorescence recording capability and essentially to test the apparatus, demonstrated synchronized dynamical measurement of multiple fluorescence parameters on single trapped fluorophores (Chapter 4, Publication 4). After developing an appreciation for the power of multiparameter measurements, I returned to measure the behavior of individual APC molecules. Finally, synthesizing information from five different fluorescence parameters, I constructed a model of the organization of the individual pigments in the complex that satisfactorily describes the data, which will be the subject of a future publication (Chapter 5, Publication 9).

Interactions between biomolecules at the single-molecule level is something I always wanted to study when I started in the Moerner Lab. I began a side project during my fourth year trying to visualize binding/unbinding events in the trap using FRET with a dark quencher. Although I had some initial success on the measurements, a flaw in the choice of the quencher quickly abolished the project, because the quencher had its own dynamics, i.e., its quenching capability fluctuated with time (Holzmeister et al., 2014b)! Nevertheless, this pilot experiment strengthened my belief that single-molecule interaction events can be observed and can become a unique application of the ABEL trap. At the same time, I had some success in developing an accurate and robust statistical algorithm for extracting transport properties from trapped molecules (Chapter 2) and I started to focus on sensing interactions by monitoring time-dependent changes in transport properties. Eventually, I succeeded in using this new method to resolve protein dissociation distributions and binding/unbinding of short oligonucleotides (Chapter 6, Publication 6). The relations between the different chapters are summarized in Figure 1.6.

I believe that my contribution of the ABEL trap technology has made it a more robust and versatile tool in biophysical research with many exciting future applications down the road.
1.5 Publications from graduate work

Here is a list of publications resulting from my graduate work, two of which (3 and 7 below) are reviews of the work in the Moerner lab. Those marked with a (*) are primary topics of this dissertation.


Chapter 2

Theory: optimal tracking and parameter learning

In this chapter I discuss the theoretical framework of the advanced ABEL trap which I have developed to address the following two challenges. First, I explore the optimal strategy to track the position of a fast-diffusing single molecule in aqueous solution using fluorescence emission. Second, I discuss methods to infer the transport properties of single molecules within the framework of the optimal tracking and trapping schemes. For a description of the theoretical underpinning of previous incarnations of the trap, please see Adam Cohen’s thesis (Cohen, 2006).

2.1 Background: tracking single-molecule motion with high spatio-temporal resolution

Tracking the motion of single molecules has long been a powerful technique to study the properties of individual molecules and their nano-environment (Saxton et al., 1997). Recently single-particle tracking (SPT) has enjoyed a booming development in both experimental techniques (Cang et al., 2008; Dupont et al., 2011) and theoretical analysis (Michalet et al., 2012; Türkcan et al., 2012; Calderon et al., 2013). However, the majority of the existing SPT studies focus on slowly diffusing single molecules (i.e.
$D < 1 \mu m^2 s^{-1}$), which can be tracked using frame-by-frame, camera-based single-molecule localization. The ABEL trap, on the other hand, requires tracking the position of a single fast diffusing target (i.e. $D \sim 100 \mu m^2 s^{-1}$) using fluorescence emission, a task that is usually difficult to achieve by conventional CCD imaging. This is mainly due to the excess amount of motion blur (Deschout et al., 2012; Michalet, 2010) that makes localization difficult and inaccurate. To see this, consider a typical protein molecule with a diffusion coefficient of $D = 100 \mu m^2 s^{-1}$. During a camera frame of 10 ms (a typical value used because of signal-to-background considerations), it would have been blurred by diffusion with a RMS distance of $\sqrt{4D\Delta t} = 2 \mu m$, which is almost an order of magnitude larger than the standard deviation of the (unblurred) point spread function (~200 nm). In other words, to keep the amount of motion blur to be on the same order of the localization precision requires an exposure time of 100 µs (10,000 fps), which is challenging to achieve with today’s EMCCD or sCMOS technology. Before introducing our tracking scheme, I first review some other work in the literature treating SPT with high temporal resolution.

Juette et al. (2010) used feedback tracking and a de-scanning technique to keep the region of interest small (5 × 5 pixels), permitting extremely fast frame rates (~3,000 fps) as allowed by state-of-the-art EMCCD technology. Yasuda et al. (2001) took advantage of the strong scattering signal of a gold nanoparticle to capture the sub-millisecond rotational motion of a molecular motor (~8,000 fps). Ritter et al. (2010) recorded the motion of fast diffusing ($D \sim 100 \mu m^2 s^{-1}$) single molecules by running the camera at ~2 ms frame rate in a light sheet microscope.

Having realized that the slow frame rate of the EMCCD is generally the bottleneck, many camera-less tracking schemes have also emerged. Lessard et al. (2007) used a four-fiber bundle arranged in a tetrahedron geometry for sensing the 3D motion of a single quantum dot in solution. Similar in spirit, Sahl et al. (2010) used a set of three offset detectors in a confocal microscope to track dye-labeled lipid molecules on a membrane in 2D with sub-millisecond time resolution. In these two schemes, position information is encoded in the relative signal strengths of the detectors. Enderlein (2000) proposed an alternative scheme to infer position information by a point detector. In his scheme, a laser beam scans in a circular motion at high speeds (~kHz) around the target molecule of interest. It can be shown that radial distance information is encoded on the amplitude of the
fluorescence signal at the modulation frequency, while the azimuth angle of displacement can be retrieved by analyzing the phase of the modulated fluorescence with respect to the revolving beam (Figure 2.1). Thus, by demodulation (e.g. lock-in) of the detected fluorescence signal, the coordinates of the molecule can be estimated with high bandwidth without the use of a camera. This “revolving beam” scheme is an example of using the temporal waveform of the fluorescence signal to encode position information, and the scheme has been successfully adopted in the tracking apparatus developed in the labs of Gratton (Levi et al., 2005), Mabuchi (Berglund et al., 2004; Berglund et al., 2005; McHale et al., 2007), Koehler (Ernst et al., 2012) and Lamb (Katayama et al., 2009) as well as in the previous generation of “all-hardware” ABEL trap (Cohen et al., 2008).

2.2 Optimal strategy to track Brownian motion with photon-by-photon time resolution

By using simulations to systematically examine the parameters of the revolving beam scheme (Jiang et al., 2008), I realized that this approach has a couple of disadvantages
when used to track fast and dim objects. This section describes my work on the development of a much improved target tracking and feedback strategy for the ABEL trap. The majority of this section is based on Wang et al. (2010).

First, we define “fast and dim” objects as those that emit fluorescence at a rate lower than the required bandwidth for trapping. Mathematically, this condition is expressed as $\bar{\Gamma} \leq 4D/(a_0/3)^2$, where $\bar{\Gamma}$ is the average emission rate and $a_0$ is the radius of the trapping region. (The factor of three in the denominator expresses the empirically observation that trap loss is likely to occur when $3\sigma$ of the position fluctuation matches the size of the trapping region.) In other words, this condition means that between photon detection events the target molecule has undergone a significant amount of motion. For $D = 100\mu m^2 s^{-1}$ and $a_0 = 0.6\mu m$, we have $\bar{\Gamma}_{\text{max}} \approx 10\text{kHz}$. On the other hand, to avoid premature photobleaching, it is advantageous to keep the average photon detection rate to be <10 kHz–20 kHz. This means that most of the molecules in the ABEL trap are “fast and dim” (or at least close to this regime), in which case it is disadvantageous to average over many measurements but to compute and apply feedback on every detected photon as soon as possible. Thus, the up-to-25 µs delay present in the initial hardware trap design (Cohen et al., 2008) limited trapping performance.

It is also illuminating to see how much position information a single photon carries in the revolving beam scheme. For a single photon detected at time $t_0$, the in-phase and in-quadrature outputs are (Berglund et al., 2004)

$$\hat{x}(\rho, \theta) = \int \delta(t - t_0) \cos(\omega_{\text{rot}}t) \, dt = \cos(\omega_{\text{rot}}t_0)$$
$$\hat{y}(\rho, \theta) = \int \delta(t - t_0) \sin(\omega_{\text{rot}}t) \, dt = \sin(\omega_{\text{rot}}t_0)$$ (2.1)

where $\rho$ and $\theta$ are the radial distance and azimuthal angle with respect to the origin and $\omega_{\text{rot}}$ is the revolving frequency of the laser beam. Equation 2.1 clearly shows that with a single photon, the measured “position” collapses to a point on a unit circle. In other words, a single detected photon in this scheme contains only information about the azimuthal angle (Jiang et al., 2008) but no information about the target molecule’s distance from the center. The implication of this observation is the following: every single photon generates a feedback kick with a constant amplitude but a different direction. Oddly, in the revolving
beam scheme, when a molecule is already near the origin, any detected photon would push it outwards!

### 2.2.1 Beam scanning on a square lattice and photon-by-photon mapping

Having realized that every detected photon contains little position information at the single photon level in the revolving beam scheme, I proposed an alternative beam scanning trajectory (Figure 2.2): instead of traveling in a circular motion, the beam scans on a square lattice. Since the effect of a single photon is to collapse the measurement to one of the positions on the laser scanning trajectory, the resulting single-photon information directly reports on target position in a Cartesian coordinates. A hypothetical example is illustrated in Figure 2.3. In other words, the beam location on the grid at the photon detection event serves as a good estimate of the target’s position. This idea can also be viewed as the single photon limit of a confocal laser scanning microscope, where the image is constructed by mapping the integrated signal on a point detector to a laser scanning position.

The next question is, how to traverse a 2D grid like the one shown on the right of Figure 2.2? Here, we want the beam to survey a large portion of the trapping region in a short amount of time while still covering every point on the grid for fine position resolution. Consequently, a raster scanning scheme, where the beam sweeps from left to right on each
vertical line, should be avoided. I use a “knight’s tour” (as on a chessboard) pattern to cover the 32 reference positions on the grid. Basically, the beam moves two grid spacings in one direction and one grid spacing in the orthogonal direction. I came up with this scheme because it is well known that the knight on a chessboard can complete a closed tour (visiting each position once while starting and ending at the same place) on a $6 \times 6$ grid. The one particular pattern I ended up using is sketched in Figure 2.4. It can be decomposed into four symmetric sets of sub-trajectories (Figure 2.4b), each of which covers $\sim 75\%$ of the total trapping region. With this arrangement, the time interval between consecutive scans over the same region in space is about one fourth of the complete frame period. Importantly, the time-averaged excitation profile of this design pattern is flat within a square-like region around the center (Figure 2.4a), which is critical for accurately measuring target brightness.

### 2.2.2 Quantifying the information provided by one photon

Now that we’ve established a beam scanning scheme which encodes much more position information on each detected photon, we will derive a real-time controller that optimally processes the measurements (Section 2.3). Doing so requires a probability model of the single photon measurement process, which we now derive.
When the object’s position is at $x$ and the scanning beam docks at position $y_k$, the probability of emitting $n$ photons during an observation interval $\delta t$, predominantly arising from the Gaussian excitation spot, is given by Poisson distribution,

$$p_n(x, y_k) = \frac{(\lambda(x, y_k))^n \exp(-\lambda(x, y_k))}{n!}$$

(2.2)

$\lambda$ is the expected number of detected photons during the interval $\delta t$ and can be expressed as

$$\lambda(x, y_k) = \left( G_0 \exp \left( - (x - y_k)^T R^{-1} (x - y_k) \right) + G_b \right) \delta t$$

(2.3)

where $G_0$ is the peak photon detection rate when the object is at the center of the laser beam, $G_b$ is the detection rate from background and $R$ is a $2 \times 2$ matrix describing the (2D) spatial profile of the laser beam. Direct computation of Equation 2.2 by plugging in Equation 2.3 is rather involved. Nevertheless, we note that the probability of zero photon detection is

$$p_0(x, y_k) = \exp(-\lambda(x, y_k))$$

(2.4)
so the probability of at least one photon detection is

\[ p_{n>0}(x,y_k) = 1 - \exp(-\lambda(x,y_k)) \] (2.5)

Since \( \lambda(x,y_k) \ll 1 \) (\( \delta t \) is on the order of hundreds of nanoseconds), we keep up to the second term of the Taylor expansion

\[ p_{n>0}(x,y_k) \approx 1 - (1 - \lambda(x,y_k)/2) = \lambda(x,y_k)/2 \] (2.6)

The conditional probability \( p(x|y_k) \) is related to Equation 2.6 by

\[ p(y_k|x) = p(x,y_k)/p(x) \] (2.7)

The denominator in Equation 2.7 is a constant because the beam uniformly samples the trapping region. We can ignore all proportionality constants in Equation 2.7 to arrive at the measurement likelihood function, since it does not need to be normalized.

\[ l(y_k|x) = G_0 \exp\left(-\left(x - y_k\right)^T R^{-1} \left(x - y_k\right)\right) + G_b \] (2.8)

and \( R \) can be explicitly written as

\[ R = \begin{pmatrix} \left(w_x/2\right)^2 & w_{xy} \\ w_{xy} & \left(w_y/2\right)^2 \end{pmatrix} \] (2.9)

where \( w_x \) and \( w_y \) are the beam waist radii along the two Cartesian coordinates (1/e² intensity), \( w_{xy} \) characterizes the tilt on the x-y plane. Note that Equation 2.8 is not Gaussian due to the background term, yet the expression matches our intuitive expectation that the shape of the laser spot at position \( k \) primarily controls the probability of pumping the molecule.

### 2.2.3 A graphical representation of the tracking problem

The photon-by-photon tracking problem can be intuitively represented by a graphical model as shown in Figure 2.5. Here the \( x \) nodes represent the real position of the target.
Figure 2.5: Graphical model representation of photon-by-photon single-molecule tracking. Shaded circles: positions of the target object at instances of measurements. Open circles: raw measurements (scanning beam positions at photon detection events). Solid arrows: Gaussian transition densities. Dashed arrows: non-Gaussian measurement likelihoods.

Object and the y nodes represents photon-stamped measurements of the laser spot position when photons are detected. The solid arrows between the real positions illustrate the evolution driven by diffusion and electrokinetic transport, which can be described by the following discrete-time Langevin equation

$$x_{k+1} = x_k + \mu U_k \Delta \tau_k^{fb} + \sqrt{2\Delta t_k D} b$$  \hspace{1cm} (2.10)

where the subscript $k$ denotes the $k$-th observation, $U_k$ is the applied voltage since the previous observation with duration $\Delta \tau_k^{fb}$ and $\Delta t_k$ is the time between observation $k-1$ and $k$. ($\Delta \tau_k^{fb} \leq \Delta t_k$, because the feedback voltage (hence $fb$ superscript) is turned off if no photon has arrived for a long time.) Here, position $x_k$ is a $2 \times 1$ vector representing the 2D Cartesian position of the object. The mobility $\mu$ should be thought of as a $2 \times 2$ tensor while the diffusion coefficient $D$ is considered to be isotropic along the two Cartesian directions (We define a matrix form of the diffusion coefficient as $D \equiv DI$). $b$ is a Gaussian distributed, random $2 \times 1$ vector that describes the stochasticity of Brownian motion ($b \sim N(0, I)$, $I$ is the $2 \times 2$ identity matrix). Note that real positions are not known directly but are “hidden
variables” in the problem (as represented by shaded circles, instead). All we measure in
the experiment are a series of photon-stamped beam positions that sample the object’s real
coordinates with high levels of measurement noise. Each measurement is a probabilistic
realization of the measurement model derived in Section 2.2.2. Once a $y_k$ is observed, it
provides information to infer $x_k$. The varying lengths between the position nodes capture
the uneven time intervals between detected photons (measurements).

Given the graphical model of Figure 2.5), we define three problems. First, we want to
estimate $x_k$ in real-time from a series of measurements up to time point $k$, given a model
of transport coefficients ($D$ and $\mu$). This is needed for generating accurate feedback forces
and will be discussed in Section 2.3. Second, we want to reconstruct the full trajectory of
particle trajectory $\{x_1, x_2, ..., x_k, ..., x_N\}$ during post processing from all measurements (Sec-
tion 2.4). Third, we want to estimate the parameters of the model, namely the diffusion
coefficient ($D$) and electrokinetic mobility ($\mu$). This problem will be solved in Section 2.5.

2.3 The optimal (minimum variance) online estimator

For real-time feedback applications, the goal is to estimate object position at the $k$-th time
point given all past and present measurements ($y_{1:k}$, meaning from time point 1 to $k$). First,
we look at the case where background photons are insignificant ($G_b \approx 0$ in Equation 2.8).
We thus have a Gaussian measurement likelihood function. If we assume that the excitation beam is symmetric ($w_x = w_y = w$ and $w_{xy} = w_{yx} = 0$), $R$ becomes diagonal and the estimation problem along the two Cartesian coordinates can be separated.

Along each dimension, the Kalman filter provides the minimum variance estimation of object position. To see this, we first outline the well-known recursive update rules of the Kalman filter (for a pedagogical discussion, see Welch et al. (1995)). In the first step, a prediction (or prior estimate, superscript -) is made according to

$$
\hat{x}_k^- = \hat{x}_{k-1}^+ + u_{k-1} \mu \Delta t_{k-1}
$$

(2.11)

where $\hat{x}_{k-1}^+$ is the posterior estimate from the previous time step (to be defined later). The prior represents our “best guess” of the object’s position immediately before a new measurement is received and Equation 2.11 says that the prior at time $k$ is related to the posterior at time $k-1$ by the electrokinetic translation in between, as intuitively expected.

In the second step, the prior estimate is refined by incorporating the newest measurement and the result is our “best guess” of the object’s position immediately after a new measurement is received, or the posterior estimate (superscript +)

$$
\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - \hat{x}_k^-)
$$

(2.12)

Equation 2.12 expresses the posterior as a weighted average between the prior and the measurement with a weighting factor $K_k$ (also called the Kalman gain, takes values between 0 and 1). The value of $K_k$ determines our confidence level of the (noisy) measurement $y_k$. When $K_k = 1$, $\hat{x}_k^+ = y_k$ and we trust only the measurement. On the other hand, when $K_k = 0$, $\hat{x}_k^+ = \hat{x}_k^-$ and the measurement is discarded completely. In general, neither of those two extreme cases are desirable and an ideal value of $K_k$ exists that optimally balances past and present information.

What value of the Kalman gain is optimal? To find out, we consider the estimation error variances associated with the prior and posterior estimates (defined in the 1D case as $P_k^\pm = E\left[\left(x_k - \hat{x}_k^\pm\right)^2\right]$). It is straightforward to derive the propagation rules of the variances. For the prior,

$$
P_k^- = P_{k-1}^+ + 2D(t_k - t_{k-1})
$$

(2.13)
which intuitively expresses the broadening of the estimation uncertainty due to diffusion between measurements. For the posterior, applying the standard error propagation rules to Equation 2.12 results in

$$P_k^+ = (1 - K_k)^2 \cdot P_k^- + K_k^2 \cdot \left( \frac{w}{2} \right)^2$$  \hspace{1cm} (2.14)

Now, if we define “optimal” as the posterior estimator that achieves minimum error variance, that is

$$K_{opt}^k = \arg\min_{K_k} P_k^+$$  \hspace{1cm} (2.15)

the optimal Kalman gain can be found by taking the derivative with respect to $K_k$ in Equation 2.14 and setting it to zero. We find

$$K_{opt}^k = \frac{P_k^-}{P_k^- + \left( \frac{w}{2} \right)^2}$$  \hspace{1cm} (2.16)

It is important to realize that since the time delay between measurements is a random variable in our case (Figure 2.5), $K_{opt}^k$ is different for each measurement instant. The Kalman filter thus dynamically optimizes the weighting factor of the new measurement in real time for every detected photon. For example, when the photon waiting time is large, $2D(t_k - t_{k-1}) \gg \left( \frac{w}{2} \right)^2$, $K_{opt}^k \rightarrow 1$, which means the previous estimates is considered to be obsolete and the filter becomes “aggressive” and uses the newest measurement directly. On the other hand, when $P_k^- \ll \left( \frac{w}{2} \right)^2$, $K_{opt}^k \rightarrow 0$, which means that the filter becomes “conservative” and trusts mostly the prior. With the optimal Kalman gain, the posterior error variance is

$$P_k^+ = (1 - K_{opt}^k) \cdot P_k^- = \left( \frac{w}{2} \right)^2 \frac{P_k^-}{P_k^- + \left( \frac{w}{2} \right)^2}$$  \hspace{1cm} (2.17)

So the posterior estimation error variance is always reduced compared to that of the prior. Essentially, the Kalman filter can be regarded as a weighted moving averager that uses a prior knowledge about the target’s transport properties to select the optimal weighting factors of past measurements.

Since the posterior given by Equation 2.12 (with $K_{opt}^k$) represents the optimal estimation of position at time point $k$, it is used to calculate the feedback voltage to drive the object.
Figure 2.7: Simulated trapping trace with beam scanning and Kalman filter.

toward the center of the trap

\[ u_k = -g\hat{x}_k^+ \]  

(2.18)

where \( g \) is the user supplied feedback gain. In practice, it is convenient to set the duration of the feedback kick to be

\[ \Delta t_k = \min\left( t_{k+1} - t_k, \Delta t_{\text{max}} \right) \]  

(2.19)

where \( \Delta t_{\text{max}} \) is a user defined upper bound.

Figure 2.7 shows a simulated trace of the Kalman filter output, together with the real positions and the photon-stamped beam positions (raw measurements). Clearly, the raw measurements are very noisy compared to the Kalman filtered outputs, which closely tracks the real position of the target. I now compare the trapping performance using the Kalman filter based grid-scanning scheme with the revolving beam scheme with lock-in detection using synthetic data. In the simulation, the beam scanning pattern is updated with a dwell time of 1.05 \( \mu \)s per reference position. The \( 1/e^2 \) beam waist radius is set to be \( w = 0.4 \mu \)m and the spacing of the 2D grid is 0.3 \( \mu \)m. Photon detection is simulated as a Poisson process with a time-dependent rate determined by the distance between the excitation beam and the object plus a constant background (Equation 2.3). Typical values for single-molecule
fluorescence detection are used (14 kHz signal and a 2 kHz background). Each detected photon triggers the Kalman filter update equations outlined previously. In the rare cases when more than one photon is detected within one dwell time of the scan beam, only the first photon is retained. Parameters used to simulate the revolving beam scheme can be found in the figure caption. As can be seen from Figure 2.8a, the Kalman filter scheme is capable of trapping the objects much more tightly, with up to 30% suppression of position fluctuation in the trap. It is also illuminating to examine the outputs of the two different control schemes. In the revolving beam case, most outputs localize on a circle at this count rate (Figure 2.8b and c), which indicates phase-only information (thus confirming Equation 2.1). In sharp contrast, the outputs of the Kalman filter (Figure 2.8d and e) are Gaussian distributed. This observation suggests that the Kalman filter scheme encodes more detailed position information per each detected photon and the posterior estimates accurately sample the true position trajectory of the object.

The Kalman filter is highly computationally efficient and is suitable to be implemented on customizable hardware that can operates in real time with minimum calculation delay. The implementation of the trapping scheme described here on a Field-Programmable-Gate-Array (FPGA) will be described in Chapter 3.

2.4 The optimal offline estimator

In many cases, it is desirable to reconstruct the object’s trajectory after an experiment. Mathematically, we seek after the probability distribution of position at each time point, given a complete series of measurements (i.e. \( p(x_k|y_{1:N}) \)). This task is different from (but related to) the online estimator discussed in Section 2.3, which can only access measurements up to time \( k \) (i.e. \( p(x_k|y_{1:k}) \)). In the following, I first focus on the scenario where the background term \( (G_b) \) in Equation 2.8 can be omitted and we end up with a Gaussian measurement likelihood. In this case, the object’s full trajectory can be efficiently reconstructed by the Kalman filter (forward) and smoother (backward) algorithm. I will re-introduce the Kalman filter using the formalism of message-passing on a graphical model, which serves as the theoretical foundation to discuss the more general case in which the background term \( (G_b) \) can not be dismissed in Equation 2.8. I then focus on an approximation scheme to
Figure 2.8: Comparison of trapping performance between the revolving beam, lock-in detection scheme (black) and the knight’s tour scanning, Kalman filter scheme (red). The setup and time-averaged intensity of the revolving beam scheme are depicted in the inset of (a). The revolving beam is simulated at 26 kHz. Feedback voltages are generated after each rotation cycle and last throughout the subsequent cycle. For both algorithms, 300 ms of trapping data was generated for characterizing the position fluctuations (σ) of the object in the trap and a total of 20 trapping traces were simulated for a specific diffusion coefficient. For each algorithm and each $D$ value, the feedback gain ($g$) was fine tuned to minimize the position variance. (a) Standard deviation of (1-D) position fluctuation as a function of diffusion coefficient. Outliers at high $D$ arose from trapping runs during which the object escaped, and were not included in calculation of the averaged standard deviations. The solid lines are linear fits to the data. (b) Distribution of the lock-in outputs in the revolving beam scheme during a trapping event of a $D = 100 \mu m^2 s^{-1}$ object. Most outputs are located on the circle with a radius of 0.85 µm. (c) Radial distance histogram of the data in (b). (d) Distribution of the posterior estimates in the Kalman filter scheme during a trapping event of a $D = 100 \mu m^2 s^{-1}$ object. (e) One dimensional histogram of the data in (d). Reprinted with permission from Wang et al. (2010) Copyright Springer-Verlag.
treat the non-Gaussian measurement likelihood in the reconstruction problem. The material in this section is modified from the supplementary information of Wang et al. (2014b).

### 2.4.1 The case of Gaussian measurement noise: Kalman filter/smoother

Here, I formulate the problem using a probabilistic graphical model, which, in the case of the ABEL trap estimation and inference problem (Figure 2.5), has the same topology and mathematical structure as the widely-used hidden Markov model (Rabiner, 1989). To construct a posterior probability density for the object’s position at measurement $k$, it is helpful to examine the dependency between $x_k$ and other neighboring variables. From the graphical representation in Figure 2.9, we can divide the graph into “past”, “present”, and “future” segments by using Bayes’ rule (Minka, 1999)

$$
p(x_k|y_{1:N}) \propto p(B^T_k, x_k) l(y_k|x_k) p(F^T_k|x_k) \tag{2.20}
$$

where $B^T_k \equiv \{y_1, y_2, \ldots, y_{k-1}\}$ represents all observations before time $k$ (past) and $F^T_k \equiv \{y_{k+1}, y_{k+2}, \ldots, y_N\}$ represents all observations that are received after time $k$ (future). Intuitively, Equation 2.20 means the fact that all past (1st term), present (2nd term), future (last term) information help constrain the posterior density at time $k$. These terms are also called “forward”, “measurement” and “backward” messages. Computation of $p(x_k|y_{1:N})$ is equivalent to computing all three messages.

Given Gaussian transition densities ($p(x_k|x_{k-1})$) and Gaussian measurement likelihood ($l(y_k|x_k)$), the Kalman forward/backward recursion is an efficient recursive algorithm to compute these messages. When such assumptions are met, all probability density functions along the graph are Gaussian so that only the first two moments need to be updated, giving rise to the algorithm’s highly praised efficiency.

In the forward pass (Kalman filter, Figure 2.9a red arrows, Algorithm 2.1 and 2.2), we start with the first node in the graph and go node by node forward in time, incorporating both forward and measurement messages. In the backward pass (Kalman smoother, Figure 2.9b red arrows, Algorithm 2.3), we start with the last node in the graph and go node by node backward in time, incorporating the backward messages. After one completion of
Figure 2.9: Message passing pathways (arrows) for Kalman filter/smoother and EP-MP. B: past, F: future. (a) Forward pass. (b) Backward pass. In EP-MP (blue arrows), previous (Gaussian) likelihood is deleted from the posterior as the first step. Then, the posterior is re-calculated using the true likelihood and projected to a Gaussian. Finally, the Gaussian-approximated likelihood is obtained.

Algorithm 2.1 Kalman prediction (forward message incorporation)

**Goal:** compute the mean and covariance matrix of the forward prior distribution at node $k$, by receiving a forward message from node $k - 1$

\[
\begin{align*}
\mu_k &= \mu_{k-1} + \mu u_{k-1} \\
\Sigma_k &= \Sigma_{k-1} + 2 \Delta t_{k-1} D
\end{align*}
\]
Algorithm 2.2 Kalman filtering (measurement message incorporation)

**Goal:** compute the mean and covariance matrix of the forward posterior at node \( k \), by incorporation of a Gaussian measurement message with mean \( y_k \) and covariance matrix \( V_k \).

\[
V_k^\circ = \left( (V_k^\circ)^{-1} + R^{-1} \right)^{-1} \\
m_k^\circ = V_k^\circ \left( (V_k^\circ)^{-1} m_k^\circ + R^{-1} y_k \right)
\]

the forward and backward passes, we acquire the posterior densities of position at every time point. To facilitate further discussion, we adopt the notation used in Qi et al. (2007): superscript \( \triangleright, \circ \) and \( < \) represent the “forward”, “measurement” and “backward” messages, respectively. For example, \( m_k^{\triangleright, \circ} \) represents the mean of \( p(x_k|y_{1:N}) \), the posterior density of position node \( k \) after incorporating all three messages; \( V_{k-1}^{\triangleright} \) denotes the covariance matrix of \( p(x_{k-1}|y_{1:k-1}) \), the density of position node \( k - 1 \), when only the “forward” and “measurement” messages are incorporated.

It is straightforward to see that the minimum variance derivation of the Kalman filter in Section 2.3 is equivalent to the combination of the “forward” and “measurement” message passing step using Bayes inference on the graphical model (Equation 2.20). First, we start with the posterior density at time \( k - 1 \), \( p(x_{k-1}|y_{1:k-1}) \) which is a Gaussian with mean \( m_{k-1}^{\triangleright} \) and covariance matrix \( V_{k-1}^{\triangleright} \). To compute the prior at time \( k \), we integrate over all possible intermediate position \( x_{k-1} \)

\[
p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1}) p(x_{k-1}|y_{1:k-1}) dx_{k-1}
\]  

(2.21)

From Equation 2.10, the transition probability density \( p(x_k|x_{k-1}) \) is a Gaussian with mean \( x_{k-1} + \mu u_{k-1} \) and covariance matrix \( 2\Delta t_k^{-1} D \). The integration of Equation 2.21 is essentially a convolution of two Gaussian densities and the result is a Gaussian with mean \( m_{k-1}^{\triangleright} + \mu u_{k-1} \) and covariance \( V_{k-1}^{\triangleright} + 2\Delta t_k^{-1} D \), which is summarized in Algorithm 2.1. To compute the posterior density, we use Baye’s rule and Markov properties on the graph.
Algorithm 2.3 Kalman Smoothing (backward message incorporation)

Goal: compute the mean, covariance and lag one covariance of the posterior at node k, by receiving a backward message from node \( k + 1 \)

\[
\begin{align*}
J_k &= V_k^{\varrho \omega} (V_{k+1}^{\varrho \omega})^{-1} \\
m_{k^{\varrho \omega}} &= m_k^{\varrho \omega} + J_k (m_{k+1}^{\varrho \omega} - m_k^{\varrho \omega}) \\
V_k^{\varrho \omega} &= V_k^{\varrho \omega} + J_k (V_{k+1}^{\varrho \omega} J_k^T - V_k^{\varrho \omega}) \\
V_{k,k-1}^{\varrho \omega} &= J_{k-1} V_k^{\varrho \omega}
\end{align*}
\]

Algorithm 2.4 Multiplication/Division of two Gaussian densities

Given two multivariate Gaussian distributions \( p_1(x) = N(m_1, \Sigma_1) \) and \( p_2(x) = N(m_2, \Sigma_2) \), the product density remains Gaussian \( p_1(x) p_2(x) = N(m_3, \Sigma_3) \), where \( m_3 = \Sigma_3 (\Sigma_1^{-1} m_1 + \Sigma_2^{-1} m_2) \)

\[
\begin{align*}
\Sigma_3 &= (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} \\
m_3 &= \Sigma_3 (\Sigma_1^{-1} m_1 + \Sigma_2^{-1} m_2)
\end{align*}
\]

Similarly, division of two Gaussian densities remains Gaussian \( p_1(x) / p_2(x) = N(m_4, \Sigma_4) \), where

\[
\begin{align*}
\Sigma_4 &= (\Sigma_1^{-1} - \Sigma_2^{-1})^{-1} \\
m_4 &= \Sigma_4 (\Sigma_1^{-1} m_1 - \Sigma_2^{-1} m_2)
\end{align*}
\]

The denominator, which is the probability of the observation conditioned on all previous observations, is a normalization factor for the posterior density function. The mean and covariance matrix of the posterior can be calculated using the multiplication rule of two Gaussian densities (Algorithm 2.4). The result is summarized in Algorithm 2.2 and the 1D special case is exactly the online Kalman filter derived in Equations 2.12 and 2.17. Similarly, the backward pass can be derived using conditional probabilities, see, for example Minka (1999).
A useful byproduct of the derivation is the denominator in Equation 2.22, which can be explicitly calculated by

\[ p(y_k | y_{1:k-1}) = \int p(y_k | x_k) p(x_k | y_{1:k-1}) dx_k \]  
\[ (2.23) \]

We can use Equation 2.23 to compute the likelihood function of all observations

\[ l(y_1, y_2, ..., y_N) = \prod_{k=1}^{N} p(y_k | y_{k-1}) \]
\[ = \prod_{k=1}^{N} p(y_k | y_{1:k-1}) \]  
\[ (2.24) \]

I gave an example of using this formalism to estimate the diffusion coefficient of a trapped object in Wang et al. (2010).

2.4.2 The case of non-Gaussian measurement noise: expectation-propagation based message passing

When the measurement likelihood is not Gaussian, probability density functions lose their Gaussian shapes as well. Consequently, the Kalman filter/ smoother algorithm is not strictly applicable. In principle one needs to calculate all moments of the density function in Equation 2.22 and propagate them throughout the graph, a task that is generally computationally infeasible. Many approximation methods designed to treat non-Gaussian measurement likelihood (and/or nonlinear dynamics) have been developed, for example, the particle filter (Arulampalam et al., 2002), the variational Bayes methods (Beal, 2003) and the method used here: expectation propagation based message passing (EP-MP, Qi, 2004; Minka, 2001).

The EP-MP scheme, originally developed by Minka (2001) and later applied to the graphical models by Qi (2004) incorporates two basic ideas. First, it recognizes the fact that even in the presence of non-Gaussian likelihood, posterior densities are generally unimodal and can be well approximated by Gaussian functions (Minka, 2001). Second, it realizes that if we approximate each measurement likelihood as a Gaussian, the Kalman
forward-backward procedure can still be used to form the posterior densities. The important question is: what is the best way to approximate those non-Gaussian likelihoods while utilizing all the information they contain?

The key insight by Tom Minka in his PhD thesis is that the best way to approximate the likelihood depends on the function that it is about to be multiplied with, or “context”. As a one dimensional example, consider the the three density functions in Figure 2.10a. Here $g(x)$ is a Gaussian prior, $l(x)$ is a non-Gaussian likelihood (a mixture of two Gaussian densities in this example) and $h(x)$ is the resulting (non-Gaussian) posterior density ($h(x) = g(x)l(x)/\int g(x)l(x)dx$). We want to find the best Gaussian approximation ($\tilde{h}(x)$) to $h(x)$ and the best Gaussian approximation ($\tilde{l}(x)$) to represent $l(x)$. In doing so, we match the moments (mean and variance) of $\tilde{h}(x)$ and $h(x)$,

$$
\begin{align*}
  m_{\tilde{h}} &\leftarrow m_h = \frac{\int x \cdot g(x) \cdot l(x)dx}{\int g(x)l(x)dx} \\
  V_{\tilde{h}} &\leftarrow V_h = \frac{\int (x - m_h)^2 \cdot g(x) \cdot l(x)dx}{\int g(x)l(x)dx}
\end{align*}
$$

(2.25)
The exact and Gaussian-approximated posterior are plotted in the middle panel in Figure 2.10a. Evidently, the simple \( \tilde{h}(x) \) does a reasonable job of representing the true posterior. To see what Gaussian likelihood is needed to achieve such an approximation, we divide the Gaussian-approximated posterior \( \tilde{h}(x) \) by the prior \( g(x) \),

\[
\tilde{I}(x) = \frac{\tilde{h}(x)}{g(x)}
\]  

(2.26)

The result is a Gaussian whose moments can be obtained using Algorithm 2.4. The right panel in Figure 2.10a plots \( \tilde{I}(x) \) together with the true likelihood. We notice that if we just replace the true likelihood with the much simpler Gaussian \( \tilde{I}(x) \), we arrive at the approximated posterior that is pretty accurate. This is an intuitive result since the prior overlaps largely with only the positive component of the likelihood. The “context” (here the prior) determines how the likelihood should be approximated.

If now, for some reason, we decide that our prior is wrong and instead should be represented by the red curve in Figure 2.10b (left panel), we need to re-approximate the likelihood to reflect such a change in context. Not surprisingly, when we use the negative Gaussian component of the likelihood (Figure 2.10b, right panel), we achieve an approximated posterior that is almost indistinguishable from the true posterior (Figure 2.10b, middle panel).

The above example illustrates precisely the two critical steps behind EP-MP. First, we use Gaussian density to approximate the likelihood function depending on its context. Second, we constantly re-approximate the likelihood when an updated context is available. Figure 2.9 shows the message passing pathways of EP-MP. The forward pass is the same as the Kalman filter. During the backward pass, in addition to propagating information from end to start, we re-approximate the likelihood using an updated context (Figure 2.9b blue arrows). In contrast to the Gaussian likelihood case (Section 2.4.1), which needs only one forward and one backward iteration to obtain the posterior, EP-MP requires multiple forward/backward passes to extract the most information from the measurement likelihood functions and the accuracy of the approximation increases with each additional iteration until convergence (Minka, 2001).
Algorithm 2.5 EP-MP algorithm

**Goal:** Using the parameter set $\Theta = \{D, \mu\}$, and measured $y_{1:N}$, $u_{1:N}$, compute $m^\bowtie_k$, $V^\bowtie_k$, $V^\bowtie k,k-1$, which are the mean, variance and lag-one cross covariance of the posterior distribution.

- Loop $k = 2 : N$
  1. Compute the forward message $q^\bowtie (x_k)$, using the Kalman filter forward equation (Algorithm 2.1)
  2. Incorporate the observation likelihood $p(y_k | x_k)$, by moment matching (Algorithm 2.6)
      
      $$q^{\bowtie \circ} (x_k) = \text{proj} (q^\bowtie (x_k) l (y_k | x_k))$$
  3. Compute the observation message (Algorithm 2.4)
      $$q^\circ (x_k) = q^{\bowtie \circ} (x_k) / q^\bowtie (x_k)$$

- Loop until convergence or NumPass iterations
  1. (Skip in first iteration) Loop $k = 2 : N$
     a. Compute the forward message $q^\bowtie (x_k)$ according to the Kalman forward equation (Algorithm 2.1).
     b. Incorporate the observation message using Algorithm 2.2.
        $$q^{\bowtie \circ} (x_k) = q^\bowtie (x_k) q^\circ (x_k)$$
  2. Loop $k = N - 1 : 1$
     a. Compute the posterior $q^{bow \bowtie} (x_k)$ by incorporating the backward message via the Kalman smoothing equations (Algorithm 2.3).
     b. Re-approximate the measurement likelihood
        i. Message deletion: remove previous observation message message to get the partial belief (Algorithm 2.4) $q^{bow \bowtie} (x_k) = q^{bow \bowtie} (x_k) / q^\circ (x_k)$
        ii. Re-approximate the likelihood function and project the posterior to a Gaussian (Algorithm 2.6)
        $$q^{bow \bowtie} (x_k) = \text{proj} (q^{bow \bowtie} (x_k) l (y_k | x_k))$$
     iii. Update the observation message (Algorithm 2.4)
        $$q^\circ (x_k) = q^{bow \bowtie} (x_k) / q^{bow \bowtie} (x_k)$$
I give pseudo-code of the EP-MP procedures in Algorithm 2.5. For clarity, we define the projection operator whose function is to complete the moment matching steps in Equation 2.25.

\[ \tilde{h}(x) = \text{proj} (g(x) \cdot l(x)) \]  

(2.27)

Steps to calculate the projected Gaussian posterior given a Gaussian prior and measurement likelihood (Equation 2.8) are derived in Algorithm 2.6. For a detailed derivation of the algorithm, see the original work by Qi (2004).

2.5 Learning the diffusion coefficient and electrokinetic mobility

So far, we have been focusing on the problem of estimating the target’s position. We now turn to estimating the parameters of the dynamic system. In the ABEL trap, we are particularly interested in the target’s diffusion coefficients \( D \) and electrokinetic mobility \( \mu \) since they directly report on the object’s hydrodynamic radius and the amount of charge. I describe two approaches, one based on adaptive Kalman filtering and the other based on maximum likelihood estimation via the EM algorithm.

2.5.1 Adaptive trapping by innovation whitening

The basic idea behind adaptive Kalman filtering is the following: during the trapping of a single nano-object, we continuously tune the \( D/\mu \) parameters in real-time and monitor a response function (to be defined below) that is sensitive to modeling error. I use a combination of simulation and semi-analytical modeling to guide the tuning process so that the model eventually matches the physical reality. The work presented in this section was published in Wang et al. (2011).

I formulate the adaptive trapping algorithm based on the innovation approach. For clarity, we restrict the following discussion to one Cartesian component on the 2D plane. The “innovation” (or measurement residue) sequence (defined as the difference between raw
**Algorithm 2.6** Moment matching

**Goal:** compute the Gaussian projection of the posterior formed by a Gaussian prior and a likelihood function given in Equation 2.8. i.e.

\[ N(m^+, \Sigma^+) = \text{proj}(N(m, \Sigma) | l(y_k|x_k)) \]

- Calculate the normalization factor (partition function)

\[
Z = \int l(y_k|x_k) N(m, \Sigma) dx = A + G_b
\]

where

\[
A = G_0 \frac{|R|^{1/2}}{|\Sigma + R|^{1/2}} \exp \left( -\frac{1}{2} (y_k - m)^T (\Sigma + R)^{-1} (y_k - m) \right)
\]

- Calculate the probability that such a measurement is not from background

\[ r = A/Z \]

- Calculate partial derivatives of the partition function

\[
g = \nabla_m \ln Z = r (\Sigma + R)^{-1} (y_k - m)
\]

\[
G = \nabla_\Sigma \ln Z = -\frac{1}{2} r (\Sigma + R)^{-1} + \frac{1}{2} r \left( (\Sigma + R)^{-1} (y - m) (y - m)^T (\Sigma + R)^{-1} \right)
\]

- The mean and covariance of the projected Gaussian are

\[
m^+ = m + \Sigma g
\]

\[
\Sigma^+ = \Sigma - \Sigma \left( gg^T - 2G \right) \Sigma
\]

These results agrees with (Herbrich, 2005)
Figure 2.11: Adaptive trapping by whitening the autocorrelation of the innovation sequence. (a) Long time correlation of the innovation sequence as a function of model deviation showing the primary dependence on mobility error. (b) Short time correlation of the innovation sequence as a function of model deviation showing the primary dependence on the diffusion error. (c) Example autocorrelations of the innovation sequences. Grey circle: accurate model. Orange circle: $D$ overestimated by 45%, $\mu$ overestimated by 25%. (d) Simulated adaptation trajectory on the parameter space by correlation whitening using parameters similar to the experiment in Figure 2.12, where the grey and red dots correspond to the two cases in c. Reprinted with permission from Wang et al. (2011) Copyright American Chemical Society.
measurement and prior estimate, \( r_k = y_k - \dot{x}_k^- \) represents the new information available in measurement \( y_k \), i.e., the new information arising from photon detection at time \( k \) when the beam is at position \( y \). The optimality of the Kalman filter is reflected in the fact that if the model is accurate, the innovation at time \( k \) is statistically independent of all the past or future innovations, i.e. \( \{r_1, r_2, \ldots, r_N\} \) is an uncorrelated white noise sequence. Mathematically, we have, with \( E \) the expectation,

\[
E(r_k \cdot r_{k+j}) = 0, \text{ for } j > 0 \tag{2.28}
\]

Any error in modeling would result in a correlated (non-white) innovation sequences. It can be shown that in the case of the ABEL trap, an inaccurately modeled \( D \) (parameterized by the fractional error \( \alpha = \Delta D / D \)) would cause correlations in the innovation sequence approximated by,

\[
E(r_k \cdot r_{k+j}) \approx -\frac{\alpha \cdot (2D\Delta t)}{\overline{K} \cdot (2 - \overline{K})} \cdot (1 - \overline{K})^j \tag{2.29}
\]

where \( \Delta t \) is the average duration of the feedback kicks, and \( \overline{K} \) is the averaged Kalman gain during trapping. Similarly, a modeling error in the mobility \( \mu \) (parameterized by the fractional error \( \beta = \Delta \mu / \mu \)) would produce a correlated innovation sequence approximated by

\[
E(r_k \cdot r_{k+j}) \approx \beta \cdot (g\mu \Delta t) \cdot (2D\Delta t) \cdot (1 - \mu g \Delta t)^{j-1} \tag{2.30}
\]

where \( g \) is the feedback gain. It can be seen from Equations 2.29 and 2.30 that nonzero autocorrelation in the innovation sequence contains information about the underlying true model parameters. Moreover, due to the different time dependence of displacements due to electrokinetic flows (\( \propto \tau \), the time between photon detection events) and due to Brownian motion (\( \propto \sqrt{\tau} \)), the effect of modeling error in \( D \) and \( \mu \) can be separated, when the autocorrelation is examined at different time lags (see below). Specifically, short time correlations are mainly caused by errors in the modeled diffusion coefficient, while long time correlations are mostly related to errors in modeling the mobility. Derivations of Equations 2.29 and 2.30 are presented in the SI of Wang et al. (2011).

Figure 2.11 shows Monte-Carlo simulation results of the short lag time (defined in the simulation as \( \tau < 100\mu s \), termed \( stCorr \), Figure 2.11a) and the long lag time (300 < \( \tau < \)
2000µs, termed \(ltCorr\), Figure 2.11b) correlations of the innovation sequence as a function of fractional modeling error. It can be seen that the contours are almost orthogonal near the origin in the parameter space, reflecting good separation of the effect caused by the two parameters. Moreover, those contours only depend weakly on the specific definitions of \(stCorr\) and \(ltCorr\). The adaptive algorithm works by real-time tuning of \(D\) and \(\mu\) in the direction that respectively minimizes \(stCorr\) and \(ltCorr\), as detailed in the Algorithm 2.7, a process which we term “correlation whitening”. Figure 2.11c shows the comparison between the autocorrelation of the innovation sequence for the correct model (grey upper) and for a situation where \(D\) is underestimated and \(\mu\) is overestimated (red lower). Figure 2.11d illustrates the adaptation process on the parameter space from an actual experimental implementation described below in Figure 2.12a.

To demonstrate the ability of the adaptive trap to provide real-time measurement of diffusion coefficient and electrokinetic mobility, we used fluorescently doped polystyrene beads (Invitrogen) as a proof-of-principle, because of their known sizes and superior signal to background. Figure 2.12a shows data from one 100 nm diameter bead in the trap. In this experiment, the Kalman filter initially operates with user supplied model parameters \((D = 11.2 \mu m^2 s^{-1}, \mu_x = \mu_y = 500 \mu m s^{-1} V^{-1})\), until at 4.7 s, the adaptation algorithm is enabled. As can be seen from the corresponding plot of real time diffusion coefficient and the two mobility components (red and black), all three adapted model parameters responded, in this case, to a lower value \((D\sim 3.8 \mu m^2 s^{-1}, \mu_x \sim 350 \mu m s^{-1} V^{-1}, \mu_y \sim 250 \mu m s^{-1} V^{-1})\). The autocorrelation of the innovation sequence, calculated from before (red region) and after (blue region) the adaptation, are plotted in Figure 2.12b. A much “whitened” autocorrelation is observed as the result of the adaptive tuning. Good qualitative agreement between experiment and simulation (Figure 2.11c, generated from parameters used in Figure 2.12a) can be seen. It is interesting to note that in the case of this particular bead, \(\mu_x\) and \(\mu_y\) adapt to different values, possibly reflecting the flow inhomogeneity along the two dimensions in the microfluidic environment. A distribution of measured diffusion coefficients from this 100 nm bead sample, collected from 30 beads, is plotted in Figure 2.12d (blue curve), with a mean value of 3.8 \(\mu m^2 s^{-1}\), in reasonably good agreement with a value calculated from the Stokes-Einstein equation, which is 4.4 \(\mu m^2 s^{-1}\).
Figure 2.12: Adaptive trapping of fluorescent polystyrene beads of different sizes. (a) Trapping signals from one 100 nm bead. The simultaneously recorded intensity trace (upper) and calculated diffusion coefficient (middle) and electrokinetic mobility (lower) are shown. At around 4.7 s, adaptation is enabled, before which user supplied initial parameters were used (see text for details). (b) Autocorrelation of measured innovation sequences, from corresponding regions shown in a. (c) Fluorescence emission from several trapped beads with nominal size of 26 nm, with real-time extracted diffusion coefficient and electrokinetic mobility. Each bead is actively released to allow another bead to occupy the trap (except at around 37 s, where a brighter bead diffused in and replaced the previous occupant). Update rates in $D$ and $\mu$ correlate with intensity since adaptive tuning is performed for every fixed number of detected photons. (d) Histograms of the measured diffusion coefficients for the 100 nm (blue) and 26 nm beads (orange). Reprinted with permission from Wang et al. (2011) Copyright American Chemical Society.
Algorithm 2.7 Adaptive Kalman filtering

1. Initialize the model with $\tilde{D} = D_0$, $\tilde{\mu} = \mu_0$

2. Initialize long time and short time correlations to zero, $stCorr = 0$, $ltCorr = 0$

3. While the $k$-th photon is detected and an object trapped,
   
   (a) Run Kalman filter (Algorithms 2.1 and 2.2) with the model $\{ \tilde{D}, \tilde{\mu} \}$
   
   (b) Use innovation $r_k$ to update $stCorr$, $ltCorr$
   
   (c) If $numStCorrPairs$ is reached, tune $\tilde{D} \leftarrow \tilde{D} \cdot (1 - g_{st} \cdot stCorr)$
   
   (d) If $numLtCorrPairs$ is reached, tune $\tilde{\mu} \leftarrow \tilde{\mu} \cdot (1 + g_{lt} \cdot ltCorr)$

4. Reset $\tilde{D} = D_0$, $\tilde{\mu} = \mu_0$, for next object in the trap, repeat 2-4.

Similar measurements were performed on 26 nm diameter beads in solution, and Figure 2.12c shows a snapshot, where several beads with different diffusion coefficients and brightness are captured by the trap. We found that bright objects generally correlate with lower diffusion coefficients, presumably due to the fact that larger beads contain more dye molecules. To rule out the possibility that the measured diffusion coefficient is affected by the object’s brightness, we trapped a single bead and modulated its brightness by changing the excitation power, and minimal influence of brightness on the measured diffusion coefficient is seen. In contrast to the 100 nm bead case, the 26 nm beads appear to be less homogeneous in size, indicated by the broader distribution of measured diffusion coefficients (Figure 2.12d, orange curve). This inhomogeneity agrees qualitatively with the manufacturer’s specification (20% size polydispersity for the 26 nm beads versus 5% for 100 nm beads). The measured electrokinetic mobility of the beads varies slightly between trapping cells, but the average value of $250 \mu m s^{-1} V^{-1}$, acquired from the adaptive trapping of the 26 nm beads, agrees favorably with a bulk measurement by FCS, under identical surface preparation conditions.

The precision of the diffusion coefficient/mobility values obtained by our adaptive Kalman filtering method depends on several factors: first of all, the precision of measuring the scan grid spacing ($a$) and the laser beamwaist ($w$). The grid spacing defines the length scale so that if it is not characterized accurately, systematic errors in estimating both
$D$ and $\mu$ will occur. The beamwaist defines the error variance of individual measurements and is critical in the calculation of the Kalman gain ($K_k$). An inaccurate beamwaist would thus affect the estimation precision of the diffusion coefficient (Equation 2.29), but not the mobility. Secondly, and perhaps more importantly, the presence of background photons would greatly affect the estimation of the diffusion coefficient. Intuitively, the fact that every detected photon has a certain probability to originate from background would increase the measurement noise accordingly (the measurement error variance would be greater than $(w/2)^2$). This effect is qualitatively similar to the case of an underestimated beamwaist, as discussed above. A significant amount of background photons would also introduce non-Gaussian statistics to the problem, in which case the Kalman filter ceases to be a truly optimal solution and complete treatment is generally beyond efficient implementation. To retain the favorable computational complexity of the Kalman filter while taking background photons into account, we used a built-in “innovation filter” to reject measurements with large residues (over a user defined threshold), since these are most likely to be generated by background photons (Grewal et al., 2008). While the extracted $D$ and $\mu$ from this approach can be particularly helpful in sensing changes in these parameters in real time, the absolute accuracy of both depend on the effects mentioned above, and measurements on a calibration object may be necessary for precise values.

### 2.5.2 The expectation-maximization based maximum-likelihood approach

In practice, the adaptive Kalman filtering approach works well in situations when the background is insignificant compared to the signal, but significantly underestimates the diffusion coefficient when the signal-to-background ratio decreases. This is not surprising since the Kalman filter assumes no background in the measurement likelihood (Section 2.3). This also means that, to accurately estimate $D$ and $\mu$, we need to explicitly model background and use the full measurement likelihood function in Equation 2.8. To address this, I subsequently developed a maximum-likelihood estimator based on the expectation-maximization (EM) formalism. The method presented here was published in the supplementary material of Wang et al. (2014b).
Our goal is to compute the maximum likelihood estimate (MLE) of the parameters \( \Theta = \{D, \mu\} \), given a series of \( N \) observations \( y_{1:N} \equiv \{y_1, y_2, ..., y_N\} \) and feedback inputs \( u_{1:N} \equiv \{u_k \equiv U_k \Delta \tau_{fb}^k\} \).

\[
\Theta = \arg\max_{\Theta} p(y_{1:N} | u_{1:N}, \Theta)
\]

(2.31)

In doing so, we use the Expectation-Maximization (EM) algorithm (Bilmes, 1997; Do et al., 2008; Moon, 1996; Ghahramani et al., 1996; Shumway et al., 1982), which is an intuitive and robust alternative to gradient-based likelihood optimization (Fields et al., 2012).

In the following, we first derive the measurement likelihood function in the ABEL trap problem and then derive the E-step and M-step of the EM algorithm. A similar formalism was recently developed to infer a single-molecule energy landscape from FRET measurements (Haas et al., 2013).

**Intuitive description of the EM algorithm** Although mathematically daunting at first sight, the intuition behind the EM algorithm is quite simple. Our goal is to estimate the transport coefficients of the trapped object (\( D \) and \( \mu \)). If, hypothetically, real positions are known at all measurement instances, we can estimate the mobility (\( \mu \)) by a linear regression between the displacements (\( x_k - x_{k-1} \)) and their corresponding voltages since the Brownian motion contributions to the displacements are zero mean Gaussians. Once we’ve inferred \( \mu \), we can then subtract the voltage-induced translation from the displacements to get the Brownian motion component, which can be subsequently used, by the mean-square-displacement (MSD) analysis for example, to estimate the diffusion coefficient (\( D \)).

In reality, object positions are not known precisely but are measured with high noise levels. We can still describe positions with probability density functions (pdf) and integrate over the position variable when estimating \( D \) and \( \mu \). However, in order to precisely compute the density functions at all measurement instances, we need to know \( D \) and \( \mu \) themselves, since they are parameters of the dynamic model (Equation 2.10). The EM algorithm solves this “chicken-and-egg” problem by a divide-and-conquer strategy. We start with some guesses of the parameters (\( D_0 \) and \( \mu_0 \)), and compute the pdf of positions (E-step). We subsequently
estimate the parameters \((D_1 \text{ and } \mu_1)\) using these position densities (M-step). Then, we iterate and use the newly estimated parameters \((D_1 \text{ and } \mu_1)\) to compute the pdf and run the M-step to update the parameters. This iteration procedure is repeated until convergence. It has been shown that EM converges to the maximum likelihood estimates of the parameters (Bilmes, 1997).

The complete-data likelihood function To derive the EM steps, we start with the complete-data likelihood (pretending that we know the positions \(x_k\)). This is easily formulated by utilizing conditional dependencies on the graphical model (Figure 2.5),

\[
L(\Theta|y_1:N, x_1:N) = p(x_1:N, y_1:N|\Theta) = p(x_1) \prod_{k=1}^{N} p(x_{k+1}|x_k) \prod_{k=1}^{N} l(y_k|x_k) \tag{2.32}
\]

where \(\Theta = \{D, \mu\}\) is a short hand notation for the parameters. The first factor on the second line is the probability distribution of the initial position, the second factor represents the state (position) transition probabilities and the third factor represents the measurement likelihood (Equation 2.8). The state transition probabilities, arising from drift and diffusion, are strictly Gaussian as shown in Equation 2.10

\[
p(x_{k+1}|x_k) = \frac{1}{2\pi (2\Delta t_k)^{1/2}|D|^{1/2}} \exp \left[ -\frac{1}{2(2\Delta t_k)} (x_{k+1} - x_k - \mu u_k)^T D^{-1} (x_{k+1} - x_k - \mu u_k) \right] \tag{2.33}
\]
CHAPTER 2. THEORY: OPTIMAL TRACKING AND PARAMETER LEARNING

Algorithm 2.8 Maximum Likelihood estimation of \( \{D, \mu\} \) by EM

**Goal:** given a series of raw measurements \( y_{1:N} \) and feedback voltages \( u_{1:N} \), compute the maximum likelihood estimate of \( \{D, \mu\} \) (and the posterior distribution of positions)

- Start with initial guesses of the parameters \( \{D^0, \mu^0\} \)
- Loop \( i = 1: \text{NumEMpass} \) iterations or stop when the change in the parameters is smaller than a given precision
  1. Compute sufficient statistics (Equations 2.41-2.46) by conducting the EP-MP smoothing (Algorithm 2.5), using the parameters \( \{D^i, \mu^i\} \)
  2. Update parameters using Equations 2.48 and 2.49. (M-step)

where \( D = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \) is the covariance matrix of the multivariate Gaussian distribution.

The complete-data likelihood function can be written explicitly as

\[
L(\Theta|y_{1:N}, x_{1:N}) = p(x_1) \\
\prod_{k=1}^{N} \frac{1}{2\pi (2\Delta t_k)^{1/2}|D|^{1/2}} \cdot \exp \left[ -\frac{1}{2(2\Delta t_k)} (x_{k+1} - x_k - \mu u_k)^T D^{-1} (x_{k+1} - x_k - \mu u_k) \right] \\
\prod_{k=1}^{N} l(y_k|x_k) 
\]

We prefer to deal with the log-likelihood \( \mathcal{L} \),

\[
\mathcal{L}'(\Theta|y_{1:N}, x_{1:N}) \equiv -2\ln L = \sum_{k=1}^{N} \ln (2\Delta t_k|D|) \\
+ \sum_{k=1}^{N} \frac{1}{2\Delta t_k} (x_{k+1} - x_k - \mu u_k)^T D^{-1} (x_{k+1} - x_k - \mu u_k) \\
+ \text{TDNDP} \tag{2.35}
\]

(TDNDP = terms that do not depend on the parameters).
The Expectation (E) step In the E-step of the EM algorithm, the expectation of the complete data log-likelihood is computed,

\[ Q \left( \Theta | \Theta^{(j-1)} \right) = \mathbb{E} \left\{ \mathcal{L} \left( \Theta | \Theta^{(j-1)}, y_{1:N} \right) \right\} \]  

(2.36)

where \( \Theta^{(j-1)} = \{D^{(j-1)}, \mu^{(j-1)}\} \) summarizes the parameters from the previous iteration, \( \mathbb{E} \) is the expectation operator. In order to calculate the expectation values, we need to calculate the following quantities, using the parameter set \( \Theta^{(j-1)} \)

\[ x_N^k \equiv \mathbb{E} \{ x_k | y_{1:N} \} \]  

(2.37)

\[ P_{N,k} \equiv \mathbb{E} \left\{ (x_k - x_N^k) (x_k - x_N^k)^T | y_{1:N} \right\} \]  

(2.38)

\[ P_{N,k,k+1} \equiv \mathbb{E} \left\{ (x_{k+1} - x_N^{k+1}) (x_k - x_N^k)^T | y_{1:N} \right\} \]  

(2.39)

Equation 2.37 is the posterior mean of the positions given all measurements \( y_{1:N} = \{y_1, y_2, ..., y_N\} \) and . Equations 2.38 and 2.39 are the posterior variance and lag-one cross-variances, respectively. These are nothing but the outputs of the EP-MP algorithm (Algorithm 2.5) in Section 2.4.2.

Once we have collected the above statistics, the expectation of the complete data likelihood has the following form, this is done by expanding the quadratic form in Equation 2.35

\[ Q \left( \Theta | \Theta^{(j-1)} \right) = \sum_{k=1}^{N} \ln \left( 2 \Delta t_k | D_k | \right) \]  

\[ + \quad Tr \left\{ D^{-1} \left( S_{11} + S_{00} - S_{10} - S_{10}^T - \mu (R_{10} - R_{00}) - (R_{10} - R_{00})^T \mu^T + \mu U_{00} \mu^T \right) \right\} \]  

(2.40)
where

\begin{align*}
S_{11} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} \left( P_{k,k+1}^N + x_{k+1}^N (x_{k+1}^N)^T \right) \\
S_{00} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} \left( P_{k,k}^N + x_k^N (x_k^N)^T \right) \\
S_{10} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} \left( P_{k,k+1}^N + x_{k+1}^N (x_k^N)^T \right) \\
R_{10} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} u_k (x_{k+1}^N)^T \\
R_{00} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} u_k (x_k^N)^T \\
U_{00} &= \sum_{k=1}^{N} \frac{1}{2\Delta t_k} u_k u_k^T
\end{align*}

for derivation, please see Chapter 6 of Shumway et al. (2010).

**The Maximization (M) step** In the M-step, the expectation of the log-likelihood is maximized with respect to the parameters (which is equivalent to minimizing Equation 2.40), i.e.,

\[
\Theta^{(j)} = \arg \min Q \left( \Theta | \Theta^{(j-1)} \right)
\]  

(2.47)

Given the analytical expression in Equation 2.40, the M-step is equivalent to taking the partial derivative of Equation 2.40 with respect to \( \mathbf{D} \) and \( \mu \) and setting them to zero. For \( \mu \),

\[
\frac{\partial Q \left( \Theta | \Theta^{(j-1)} \right)}{\partial \mu} = 0
\]

\[
\mu^{(j)} = (R_{10} - R_{00}) U_{00}^{-1}
\]

(2.48)
which is basically a linear regression given applied voltages and posterior estimates of
translations, as intuitively expected. For $D$,

$$\frac{\partial Q(\Theta | \Theta^{(j-1)})}{\partial D} = 0$$

$$D^{(j)} = \frac{1}{N-1} \left( S_{11} - S_{10} - \mu^{(j)} R_{10} \right)$$  (2.49)

The off-diagonal components of $D$ arise from noise and are generally small compared to
the diagonal elements. We simply set the off-diagonal elements to zero before using $D$
for next iteration. The physical interpretation of Equation 2.49 is also straightforward: we
subtract off the deterministic component of the MSD in order to calculate the diffusion
coefficient. The EM algorithm alternates between the E-step and the M-step to iteratively
estimate $\{D, \mu\}$. The complete EM algorithm is summarized in Algorithm 2.8.

**Validation of the EM algorithm by simulation**

I validated Algorithm 2.8 using simulated ABEL trap data. In particular, I focused on the converging properties, the estimation
accuracy and precision and the execution speed of the algorithm.
The EM algorithm almost always converges, except under conditions of marginal trapping (high diffusion coefficient, low signal-to-background ratio), as shown in Table 2.1. In Figure 2.13a, I show the output of the algorithm after each iteration using experimental data of a single trapped Atto 647N molecule (Chapter 4). Convergence (small white “X”) is achieved in fewer than 20 iterations and is independent of initial guesses of the parameters (different colors).

In Figure 2.13b, I illustrate the effectiveness of the EP-MP algorithm (Section 2.4.2) in extracting non-biased information from non-Gaussian likelihood functions. Here, the final result of the EM algorithm ($D$ only) is plotted against the number of EP-MP passes using simulated data (black symbols). The dotted red line indicates the true $D$ value in simulation. Apparently, a single forward-backward pass (i.e. “the Assumed Density Filter” as in Fields et al. (2012)) produced a large bias in the end result. This is because the information in a series of non-Gaussian measurement likelihood functions has not been fully incorporated to reconstruct the trajectory in the E-step. A second pass through the data greatly improved the result and with more than three passes, the “true” diffusion coefficient is recovered. It is worth noting that the time complexity (Figure 2.16) scales linearly with the number of EP-MP passes so there is a trade-off between accuracy and execution time.

I tested the accuracy of the EM algorithm over a large parameter space ($D$ spanning over one order of magnitude and $\mu$ varying by a factor of four) and under different signal-to-background ratio ($SBR$) conditions (Figure 2.14). Here, each black dot represents the ($D, \mu$) estimated from one simulated molecule with 12,000 photons (measurements). Red squares indicate parameter values used in the simulations (“ground truth”). As shown, the algorithm recovers the accurate parameter values in all cases. In the worst case, the diffusion coefficient is slightly underestimated (~3%) for $D = 320 \mu m^2 s^{-1}$ and a $SBR$ of 3. Evident from the scatter plot is the observation that the spread in estimated $D$ varies
for different parameter settings: for the same $D$, the spread gets smaller with increased $SBR$, as intuitively expected. More interestingly, for the same $SBR$, the uncertainty in estimated $D$ decreases with increasing values of $D$. This is probably due to the fact that the overall movement of the object becomes increasingly dominated by diffusive motion at higher $D$. There is also a slight correlation between the estimated diffusion coefficient and mobility. We suspect this is due to the algorithm’s limited capability to completely separate out the deterministic voltage-induced transport from the stochastic Brownian motion in the presence of high measurement noise. Interestingly, the positive correlation decreases in cases where the diffusion coefficient is high. This could also be explained by the dominance of Brownian motion in these cases on the photon detection time scale.

It is of practical importance to see how the relative error in $D$ scales with the number of signal photons ($N_{\text{sig}} = N_{\text{tot}} \times SBR/(1+SBR)$, as an attempt to normalize different $SBR$). This information can help determine, for example, the resolution of the method to differentiate a mixture. I define relative error as the standard deviation over the mean. In Figure 2.15a, I show simulation results for different diffusion coefficients with varying $SBR$ on a log-log scales. The standard deviation of estimated $D$ was estimated by running the simulation with the same parameter setting 500 times. In all cases, the relative uncertainty

![Figure 2.14: Accuracy of the EM algorithm. Black dots: estimated transport coefficients from simulated data sets. Red squares: ground truth.](image)
Figure 2.15: Precision of the EM algorithm. (a) Simulation results for $D = 20\mu m^2 s^{-1}$ (black), $D = 100\mu m^2 s^{-1}$ (red) and $D = 320\mu m^2 s^{-1}$ (blue) under different signal-to-background ratios (circular: $SBR = 9$, square: $SBR = 3$). Lines are power law fits to the data. (b) Comparison of experimental data to simulation. Black squares: experimental data on allophycocyanin (see Chapter 6 for experimental details), Black line: simulation with $D = 50\mu m^2 s^{-1}$. Red circle: experimental data on 10 nucleotide single-stranded DNA with a single Alexa 647 label. Red line: simulation with $D = 150\mu m^2 s^{-1}$.
in $D$ follows the $N^{(0.48\pm0.02)}$ scaling, as expected from shot-noise-limited measurements. In Figure 2.15b, we plot experimentally determined relative errors (symbols) together with lower bounds determined by simulation (solid lines). In the case of 10nt-ssDNA, the experimental data track the simulation very well. In the case of APC, I think the deviation comes from intrinsic sample heterogeneity (90% XL and 10% NXL APC). Also worth noting is the observation that high accuracy can be achieved by our method with relatively few numbers of photons. For example, consider the task of resolving protein monomers and dimers in solution. The relative diffusion coefficients of the two species can be estimated by $D_{\text{dimer}}/D_{\text{mono}} = (1/2)^{1/3} \approx 0.8$. Such a difference is too small to be detected by fluorescence correlation spectroscopy (Meseth et al., 1999; Guo et al., 2012). Using the ABEL trap, as long as we can harvest $\sim 4000$ photons from each single molecule (a modest amount for a good single-molecule dye or fluorescent protein), we can probe $D$ with an error of $\sim 10\%$ (for $D = 50\mu m^2 s^{-1}$, Figure 2.15b), thus being capable of resolving monomers from dimers. In principle, an arbitrary precision in the transport coefficient can be obtained given enough number of detected photons (e.g. multiple labeling). This could potentially lead to detection of a protein conformation change or the generation of a single charge.
In many cases it is desirable to measure the transport coefficient \((D, \mu)\) of single molecules in real time. For example, so far as trapping performance is concerned, accurate knowledge of both the diffusion coefficient and mobility increases the accuracy of the online Kalman filter in generating optimal feedback voltages (Section 2.5.1). One can also imagine applications such as single-molecule filtering and sorting if \((D, \mu)\) can be estimated in real time. Moreover, such a capability directly visualizes dynamic changes in transport coefficients, as will be shown in the DNA hybridization example in Chapter 6.

Although Algorithm 2.8 is derived as a batch algorithm, we can transform it to an online algorithm by chopping the incoming data stream into small, non-overlapping “windows” and running the batch EM algorithm for each data window. In this way, whether the algorithm retains real-time efficiency depends on the throughput of the calculation compared to the rate of photon collection in the experiment. The time complexity of Algorithm 2.8 scales linearly with the number of EM iterations, the number of EP-MP passes and the size of the processing window. The computation times (tested on Intel i7-2670QM processor with 8G RAM) with two different configurations are plotted here (blue symbol for four EP-MP passes and red symbols for two EP-MP passes) as a function of online processing window sizes. Also plotted are lines of constant photon detection rates (dashed black lines). Shaded areas indicate count rate regimes of a certain configuration that achieves real-time efficiency. As illustrated by Figure 2.16, the configuration with two (2) EP-MP passes (reasonably accurate as shown in Figure 2.13b) can handle photon counting rates as high as 40 kHz while still maintaining real-time efficiency.
Chapter 3

Making them all work: ABEL trap instrumentation

Over the years I’ve designed, constructed and maintained two ABEL trap setups. The 9a front (Trap A) setup started as an FPGA version of the revolving beam trap (Cohen et al., 2008) in late 2008 and was used to measure the conformational dynamics of single GPCR receptors in solution by Sam Bockenhauer (Bockenhauer et al., 2011). I started to implement a new feedback scheme using a knight’s tour scanning beam and the Kalman filter in early 2010. Later that year, I added two-color capability to the setup, which allows simultaneous trapping (at 660 nm) with a trapping dye and probing (488 nm) of a FRET pair (Section 3.8.5). This modification also added flexibility to the setup, as new excitation pump wavelengths can be quickly configured using an existing path as the guide. Other people have also contributed to the optical configuration of Trap A. Sam Bockenhauer aligned the 515 nm pulse laser and implemented lifetime measurements. Dr. Gabriela Schlau-Cohen added the capability to use light from the Ti:Saph laser and the supercontinuum as the excitation source and added a grating spectrometer to the setup.

I began to assemble the 8b front (Trap B) setup in mid-2010 with the goal of establishing a new platform for more advanced technical development, as Trap A became increasingly involved with scientific experiments. Trap B also featured much faster beam scanning
deflectors and a more advanced FPGA board. Features such as polarization-resolved measurements (Section 3.8.3) and EM-based estimation of transport coefficients (Section 2.5.2 and Chapter 6) were subsequently developed on Trap B.

In this chapter, I describe how to implement an advanced ABEL trap, including optics, electronics, feedback programming, microfluidic handling and data post-processing. Rather than giving step-by-step, easy-to-follow protocols on how to put things together, I will focus on the physical principles and logic behind the implementation.

### 3.1 System overview

An ABEL trap system can be divided into seven functional components. These are depicted in Figure 3.1 and will be discussed in detail.
<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refractive index $n$</td>
<td>2.26</td>
</tr>
<tr>
<td>Density $\rho$ (10$^3$ kg m$^{-3}$)</td>
<td>6.0</td>
</tr>
<tr>
<td>Transverse wave velocity $v_a$ (km s$^{-1}$)</td>
<td>0.6</td>
</tr>
<tr>
<td>Longitudinal wave velocity $v_a$ (km s$^{-1}$)</td>
<td>4.2</td>
</tr>
<tr>
<td>Aperture time $t_a$ (L-wave, ns mm$^{-1}$)</td>
<td>150</td>
</tr>
</tbody>
</table>

Table 3.1: Acousto-optic properties of TeO$_2$

3.2 Excitation path

3.2.1 Acousto-optic deflectors

The acousto-optic deflector (AOD) is the most critical element for generating a programmable beam scanning pattern. I choose the AO technology because it enables extremely fast (> MHz) angular modulation, which is orders of magnitude faster than resonant galvo mirrors, yet is simpler and cheaper than an electro-optic deflector implementation. In an AOD system, an ultrasonic transducer produces an traveling acoustic wave in a transparent material with a frequency $f$. The acoustic wave induces index modulation in the crystal which scatters the laser beam. When the acoustic frequency is sufficiently high ($\geq$ 100 MHz) and the interaction length is sufficiently long (> 10 acoustic wavelengths), the interaction is in the “Bragg regime” and distinct deflected beams appear on both sides of the zeroth order beam (Liu, 2005). When the Bragg condition is satisfied, the +1 (or the −1) order is enhanced due to constructive interference. This can be best understood by considering the following physical picture (Figure 3.2), the optical path length difference between successive rays are given by $2\Lambda \sin \alpha$ ($\Lambda$ is the wavelength of the acoustic wave) and constructive interference happens when this difference equals integer multiples of the optical wavelength $\lambda$. The first order Bragg angle is then given by

$$\sin \alpha_B = \frac{\lambda}{2\Lambda} = \frac{\lambda f}{2v_a}$$

(3.1)

where the acoustic wavelength is expressed as the sound velocity $v_a$ divided by frequency $f$. The total angular displacement of the beam is twice the Bragg angle (Figure 3.2). Because the Bragg angle is typically small for AODs, the deflection angle of the first order beam
Figure 3.2: Acousto-optic interaction in the Bragg regime. See text for definition of symbols.

can be well approximated by

\[ \Delta \theta = 2\alpha_B \approx \frac{\lambda f}{v_a} \]  
(3.2)

The basic idea of the AOD naturally follows: when the acoustic frequency \( f \) is modulated, the beam wiggles around the first order beam and the amount of wiggle is given by

\[ \Delta (\Delta \theta) = \frac{\lambda \Delta f}{v_a} \]  
(3.3)

We subsequently denote \( \Delta (\Delta \theta) \) as \( \delta \theta \) for clarity.

A number of design considerations are important for building the optimal beam scanning system for the ABEL trap. First, we want the scan speed to be faster than the characteristic diffusive motion of the molecule, so that the position of the molecule can be sufficiently sampled before the molecule moves appreciably. The fundamental limit of the scanning speed is set by the crystal aperture time. Simply put, when a new RF frequency is applied to the crystal, the sound wave needs a finite amount of time (\( \approx 2w_a/v_a \), \( w_a \) is the laser beam radius) to propagate across the laser beam. For TeO\(_2\), a popular AOD material in the visible (Table 3.1), the aperture time is about 150 ns mm\(^{-1}\). Intuitively, the smaller the...
### Table 3.2: AOD system of Trap A. Crystal model: NEOS 46080-3-LTD. Controller model: NEOS 21060-100-2ASVCO-2

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aperture size $a$ (mm)</td>
<td>3 × 4</td>
</tr>
<tr>
<td>Center frequency $f_0$ (MHz)</td>
<td>80</td>
</tr>
<tr>
<td>Modulation bandwidth $\Delta f_{\text{max}}$ (MHz)</td>
<td>40</td>
</tr>
<tr>
<td>Number of resolvable spots $N_{sp}$</td>
<td>28</td>
</tr>
<tr>
<td>RF generation</td>
<td>VCO</td>
</tr>
<tr>
<td>RF electronics switching time (µs)</td>
<td>$&gt; 1.2$</td>
</tr>
</tbody>
</table>

### Table 3.3: AOD system of Trap B. Crystal model: AA MT110-B54A1.5-VIS. Controller model: AA DDSPA2X-D8b15b-34

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aperture size $a$ (mm)</td>
<td>1.5 × 1.5</td>
</tr>
<tr>
<td>Center frequency $f_0$ (MHz)</td>
<td>110</td>
</tr>
<tr>
<td>Modulation bandwidth $\Delta f_{\text{max}}$ (MHz)</td>
<td>54</td>
</tr>
<tr>
<td>Number of resolvable spots $N_{sp}$</td>
<td>10</td>
</tr>
<tr>
<td>RF generation</td>
<td>DDS</td>
</tr>
<tr>
<td>RF electronics switching time (µs)</td>
<td>$&lt; 0.04$</td>
</tr>
</tbody>
</table>

Beam diameter at the crystal, the faster the scanning speed. However, the beam size cannot be decreased indefinitely because doing so reduces the number of resolvable spots at the sample plane. This is because as the beam shrinks in size, its intrinsic angular spread along the propagation direction (due to diffraction) increases and could eventually “swamp” the small angular displacement produced by the AOD. Mathematically, we can calculate the number of resolvable spots by

$$N_{sp} = \frac{\delta \theta}{2 \Delta \Theta} = \frac{\lambda \Delta f / v_a}{2 \lambda / (\pi w_a)} = \frac{\pi}{4} \frac{\Delta f}{v_a} \frac{2 w_a}{v_a}$$  \hspace{1cm} (3.4)$$

where $2 \Delta \Theta$ is the full divergence angle of a Gaussian beam of waist radius $w_a$. From Equation 3.4, $N_{sp}$ is proportional to the beam size on the AO crystal given a fixed modulation bandwidth $\Delta f$. More importantly, since $N_{sp}$ is directly proportional to the Helmhotz invariant $\delta \theta \times w_a$, it is preserved by the downstream relay optics. The scan pattern (Section 2.2) based on a $6 \times 6$ grid translates to at least 6 resolvable spots along one dimension,
CHAPTER 3. MAKING THEM ALL WORK: ABEL TRAP INSTRUMENTATION 73

so we want to control the beam size on the AOD to be the smallest value that guarantees 6 spots. Using Equation 3.4, we have

\[ N_{sp} = \frac{\pi}{4} \Delta f \frac{2}{v_a} w_a \geq 6 \] (3.5)

Given a \( \Delta f \) of about 30 MHz, slightly smaller than the maximum value specified by the manufacturer (Tables 3.2 and 3.3) and the acoustic velocity of TeO\(_2\) (Table 3.1), we calculate the minimum beam waist radius to be \(-0.5\) mm. This beam size corresponds to an aperture time of \(\sim 150\) ns, which can be regarded as the lower bound of the transition time between scan points.

In practice, whether this physical limit can be reached also depends on the RF switching speed. The commercially available AOD controllers generally come in two varieties: the analog, voltage-controlled-oscillator VCO-based RF synthesizer and the digital, direct-digital-synthesis DDS-based RF synthesizer. In a VCO, a certain RF frequency is generated given an analog voltage and in a DDS, the RF waveform is synthesized given a digital code. After being generated, the RF is amplified to about 1 W for driving the transducer on the AO crystal. We have a VCO controller for Trap A and a DDS controller for Trap B. The DDS technology is supposed to offer much improved frequency switching speed and waveform fidelity compared to its analogue counterpart. In our case the DDS driver is preferred mainly due to interfacing issues with the FPGA: The digital-to-analog converters on National Instruments R series FPGA require about 1 \(\mu\)s update time and this feature limits the frequency update time to \(>2\) \(\mu\)s in practice when using an analog-based VCO controller. On the the hand, the DDS-type driver programming information is communicated via the digital lines on the FPGA, where the update time is only limited by the on-board clock (~25 ns).

Another practical issue with AOD is reduced diffraction efficiency near the limits of the modulation bandwidth \((f_0 \pm \Delta f_{\text{max}}/2)\). Experimentally, I find that the beam at the four vertices of the squared scan area retains about 50% of the intensity of the beam near the center. This is consistent with the specifications from the manufacturer \((-4\) dB for two AODs). The inhomogeneity of illumination intensity makes a single molecule more detectable near
the center of the trapping region. (In other words, the object’s radial displacement is always underestimated.) By controlling the RF power at each scan position, the pattern can be homogenized at the expense of reduced diffraction efficiency near the center. I describe a detailed procedure for this in Section 3.6.2.

### 3.2.2 Relay optics

I now discuss the relay optics (Figure 3.3), which transforms the angular modulation created by the AODs to a grid scanning pattern in the sample plane of the microscope. First, L1 and L2 form a beam expansion/reduction system that controls the size of the laser beam ($w_a$) that interacts with AODx. As discussed in Section 3.2.1, $w_a$ is chosen to be about 0.5 mm for the optimal balance between speed and number of resolvable spots. Starting from the first AOD (x), the goal of the relay system is twofold. First, it adjusts the size of the beam before the objective in order to achieve a certain spot size at the scanning plane. The second and less obvious purpose is to map the pivot plane of AODx to the pivot plane of AODy and eventually, also to the back focal plane of the objective lens. In this way, wiggling of the laser beam by the AODs will translate to only angular shifts (without lateral translation) on the back focal plane. I achieve the relay with three pairs of Keplerian telescope systems (Figure 3.3 inset). The beauty of a Keplerian system is that it allows independent control of beam size (without changing beam divergence) and conjugate plane mapping. The beam expansion ratio is controlled by the transverse magnification $f_b/f_a$. Meanwhile, the Keplerian system also images the focal plane of the first lens to the focal plane of the second lens (this configuration is sometimes referred to as a “telecentric” or “afocal” imaging system). However, FPa and FPb are not the only conjugate planes in the system, other pairs of conjugate planes along the optical axis can be found by the following equation (Klein et al., 1986)

$$
(z_b - f_b) = - \left( \frac{f_b}{f_a} \right)^2 (z_a - f_a)
$$

(3.6)

where $z_a$ and $z_b$ are the distances between the lenses and their respective conjugate planes (see definition in Figure 3.3 inset). Equation 3.6 reveals the fact that this second pair of
CHAPTER 3. MAKING THEM ALL WORK: ABEL TRAP INSTRUMENTATION

Figure 3.3: Relay optics in the ABEL trap excitation path. Top: schematics. The dashed lines indicate planes in space that are optically conjugated to one another. Inset: beam transformation by a Keplerian telescope system. Bottom: photo of excitation optics.
conjugate planes (dashed blue, Figure 3.3 inset) is related to the conjugate pair of focal planes (dashed black, Figure 3.3 inset) via the longitudinal magnification \((f_b/f_a)^2\).

The telescope made up of L3 and L4 maps the pivot point of AODx, which controls beam scanning along the horizontal direction, to the pivot point of AODy, which controls scanning along the orthogonal direction. Here, the same spot size on AODx needs to be preserved on AODy because we want the number of resolvable spots to be the same along the two axes. We thus choose \(f_3 = f_4 = 50\) mm and place the two AODs at the pair of conjugated focal planes. This configuration is preferred over stacking the two crystals back-to-back, because it completely decouples the degrees of freedom of the two deflectors and allows beam modulation along each dimension to be optimized independently. The next two telescope systems (L5 and L6, L7 and L8) then expand the beam and map the beam scanning pivot plane to the back focal plane of the objective lens. The desired expansion ratio can be calculated from the desired spot size at the sample plane. For a reasonable beam waist radius of \(w_s = 0.5\) µm at the focal plane, the input beam needs to be (Verdeyen, 1995)

\[
w_b = \frac{\lambda f_{obj}}{\pi w_s}
\]

where \(w_b\) is the beam radius entering the objective and \(f_{obj}\) is the focal length of the objective. Given \(\lambda = 600\) nm and \(f_{obj} = 3\) mm (for a 60× Olympus oil-immersion objective with an \(f = 180\) mm tube lens), we have \(w_b = 1.1\) mm, which corresponds to expanding the beam by ~2.2 fold by lenses L5-8. Finally, the AODy-to-L5 and L6-to-L7 distances need to be carefully chosen using Equation 3.6 in order to map the pivot plane to the back focal plane of the objective. Although in principle one telescope would suffice, using two telescopes provides the added flexibility to work around some physical constraints (e.g. a lens cannot be easily placed inside the microscope, etc).

### 3.3 Detection optics

Compared to the excitation path, the detection side is relatively simple. Fluorescence from the trapped single molecule is collected by the same high-NA objective lens that focuses the excitation light. A tube lens then focuses the light and forms an intermediate image.
In Trap A, the tube lens equipped in the microscope base (Olympus IX71, \( f = 180 \text{mm} \)) was used without modification. In Trap B (Figure 3.4), I built a customized path to bypass any optics in the microscope base (Nikon Diaphot 200), because the internal optics only allows 80% of the signal at the side port. I selected a cemented doublet with \( f = 75 \text{mm} \) from Melles Griot as the tube lens. A pinhole is placed at the intermediate image plane to restrict the detection volume. In our case, the size of the pinhole must be large enough to encompass the full scan region in the sample plane. This is why the sample design for ABEL trap studies must reduce fluorescent background more than is required in a conventional confocal microscope. In Trap A, when a 60× objective is used, a \( 3 \times 3 \mu\text{m}^2 \) scanning region requires a pinhole with a diameter of \( 3 \times 60 = 180 \mu\text{m} \). In Trap B, given a tube lens of \( f = 75 \text{mm} \), a pinhole with a diameter of 150 \( \mu\text{m} \) corresponds to about a 6 \( \mu\text{m} \) region in the sample plane. After the pinhole, the fluorescence light is re-collimated, spectrally filtered as required for the fluorophore under study, and softly focused down to the surface of a photon-counting avalanche diode using a 10× objective (or an aspheric lens for better transmission in the NIR). In cases of multidimensional fluorescence detection (Section 3.8), other optical elements can be inserted to sort the fluorescence photons based on their physical properties (see Figure 3.13).

**Correcting quartz-induced spherical aberration** Although a quartz coverslip can significantly reduce the autofluorescence background, it introduces severe spherical aberration (SA) to the detection optics when an oil-immersion objective lens is used. This is because...
### Table 3.4: Managing spherical aberration in the detection path.

<table>
<thead>
<tr>
<th>Objective lens</th>
<th>Coverslip</th>
<th>Optimum pinhole position</th>
<th>Optimal transmission</th>
<th>Correction collar</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>60×Oil NA = 1.4</td>
<td>glass #1 (~150 µm)</td>
<td>1</td>
<td>99.5%</td>
<td></td>
<td>perfectly matched</td>
</tr>
<tr>
<td>60×Oil NA = 1.4</td>
<td>quartz (~160 µm)</td>
<td>4</td>
<td>75%</td>
<td></td>
<td>severe S.A.</td>
</tr>
<tr>
<td>60×water NA = 1.3</td>
<td>quartz (~160 µm)</td>
<td>2</td>
<td>94%</td>
<td>140 µm</td>
<td>compensated S.A.</td>
</tr>
<tr>
<td>60×water NA = 1.3</td>
<td>glass #1 (~150 µm)</td>
<td></td>
<td>99%</td>
<td>150 µm</td>
<td>close to matched</td>
</tr>
<tr>
<td>60×water NA = 1.3</td>
<td>glass #1 (~150 µm)</td>
<td>3</td>
<td>87%</td>
<td>190 µm</td>
<td>induced S.A.</td>
</tr>
</tbody>
</table>

Inset figure: ray tracing diagram of a quartz-mounted single emitter (with an oil-immersion objective) that is consistent with all experimental observations.
the oil-immersion objective is only aberration-corrected with a glass coverslip with a refractive index of $n_G = 1.518$, and quartz has an index that is much lower ($n_Q = 1.458$). This SA problem is most relevant in a phase-sensitive imaging scheme. For example it causes a near-complete distortion of the double-helix point-spread-function (Matthew Lew, private communication). In a confocal detection scheme, such as the one used in the ABEL trap, the SA causes an apparent focus shift at the pinhole plane that reduces photon throughput. Here I discuss this issue in some detail and provide a simple remedy.

To access the severity of the SA, I measured the transmission efficiency of the pinhole with different mounting configurations (Table 3.4). In the inset figure, the numbered vertical lines indicate four possible pinhole positions. Plane 1 is the optimal position when the sample is mounted on a glass coverslip (99.8% transmission, row 1) and can be regarded as the paraxial focal plane of the tube lens. With quartz, the pinhole (equivalent to a circular region of ~4 µm diameter in the sample plane) at plane 1 cuts off about 70% of the fluorescence signal. Moving the pinhole closer to the tube lens restored some of the signal, with a maximum of only 75% at plane 4 (row 2). This observation is consistent with positive SA caused by the index mismatch and plane 4 corresponds to the circle of least confusion. One can compensate the SA in a system by introducing SA of the opposite sign. I realized that since the correction collar in a water immersion (WI) objective tunes system SA to accommodate a range of coverglass thickness, it can also be exploited to correct for mismatch-induced SA. Indeed, experimentally I found that setting the collar to “underestimate” the coverslip thickness achieves this goal. Nevertheless, the specific WI objective (Plan Apo 60×/1.20) has a limited correction range and the best result (94%, row 3) was achieved at the end of travel (140 µm). I later found out that similar tricks have been used in other single-molecule experiments for correction of SA (Bacia et al., 2007; Gambin et al., 2011; Reihani et al., 2011).
Figure 3.5: Wiring diagram of Trap A. Each circle denotes a BNC connector. See text for details.
3.4 Electronics

3.4.1 Connection diagram

The wiring of the system is pretty straightforward. The FPGA board is hosted in a Windows computer via a PCI (Trap A) or PCIe (Trap B) socket. To allow easy access to the I/O interface of the FPGA, I used a customizable breakout box (CA1000) from National Instruments. Only a subset of all the available I/Os are needed for the ABEL trap and these are wired to BNC panelettes for interfacing with other components. In Trap A (Figure 3.5), two analog outputs (AO1 and AO2) are connected to the frequency control ports of the AOD driver. The voltage-to-frequency map can be found in the datasheet provided by the manufacturer. Two additional analog outputs (AO3 and AO4) are connected to a bi-polar high-voltage amplifier (Section 3.4.2), which drives the electrokinetic flows in the microfluidic cell. On the digital side corresponding to inputs to the FPGA, the TTL outputs of the APD(s) are connected to the digital I/O ports (DIO1 and DIO2). Each TTL pulse signals a detected photon and drives the feedback action in the trap.

In Trap B (Figure 3.6), the wiring is slightly more complicated because of the digital AOD driver. Our DDS driver has a 15-bit RF frequency control and a 8-bit RF amplitude control (which regulates the efficiency of the AO interactions as described above), which, in addition to the enable bit per channel, consumes a total of 24 digital lines per channel. I use a 50-wire ribbon cable (Digikey MC50R-5-ND) to organize these connections. To minimize crosstalk between the lines at MHz frequencies, the wires that carry signal are interleaved with wires that are pulled to ground. One end of the ribbon cable is soldered into a DB44 connector that plugs into the controller while the other end directly connects to the DIOs through the CA1000 breakout box. I use a strain relief connector panelette (NI 184721-01) to secure the ribbon cable on the front panel of the breakout box.

3.4.2 High-voltage amplifier

The high-voltage amplifier increases the capability of the output voltages of the DACs on the FPGA to drive the electrokinetic response of low-mobility species. It also acts as a
Figure 3.6: Wiring diagram of Trap B. Each circle denotes a BNC connector. See text for details.
buffer between the precious FPGA board and the microfluidic load. Although many commercial products (e.g. Kron-Hite 7602, Trek 2100 HF) can be used, they have many advanced features (e.g. high driving current, capability to handle capacitive load, etc) which seem to be an overkill for the ABEL trap (let alone being expensive). Here, all we need is a simple amplifier with a bipolar output (to avoid quadrupole electric fields (Cohen et al., 2005b)). I built an amplifier circuit around the OPA453\(^1\), which is a high-voltage op-amp chip from Texas Instruments, to provide an external gain of \(16 \times\). OPA453 has a number of desirable features. First, it provides up to 80 V of bipolar voltage and 50 mA driving current, which is sufficient for most trapping applications. Second, with a gain-bandwidth-product of 7.5 MHz and a slew rate of \(\sim 25 \text{ V} \mu \text{s}^{-1}\), it is a very fast op-amp. Third, it is cheap (~$3) and widely available. The circuit shown in Figure 3.7 is made up of a non-inverting amplifier (gain = 8) stage and an unity-gain inverting amplifier (gain = \(-1\)) stage (Mancini, 2003). In the actual circuit, R2 and R5 are potentiometers to allow fine tuning of the gain. Also, C3 is a critical capacitor which stabilizes the feedback loop of the unity-gain inverting stage and its value of 1 nF was found empirically.

### 3.5 Control algorithm and FPGA programming

The photon-by-photon position mapping and the Kalman filter scheme (Section 2.3) are implemented on a field-programmable-gate-array (FPGA). The FPGA is a piece of reconfigurable digital signal processing electronics that offers high-speed (~10 MHz) computation of sophisticated logic where the wiring and control timing sequence are defined by software.

We used FPGA boards from National Instruments (NI) as the “brain” of the ABEL trap. NI’s FPGA board is branded under the “R-series multifunction reconfigurable IO” devices. These boards have a number of analog and digital inputs/outputs integrated around the FPGA chip (Figure 3.8) and easily plug into a PCI/PCIe slot on a PC. Unlike traditional DAQ devices, which are designed to carry out certain data acquisition tasks (e.g. record an analog waveform or count TTL pulses), the function of a R-series board is completely customizable via programming the FPGA chip. Due to its enormous on-the-fly computing

\(^1\)http://www.ti.com/lit/ds/symlink/opa453.pdf
power, configuration flexibility and affordability, the FPGA has found applications in many corners of experimental physics (Stockton et al., 2002; King et al., 2009; Kubanek et al., 2009; Comstock et al., 2011; Mocsár et al., 2012).

The most important reason for us to choose the FPGA boards from National Instruments is the programming interface. Traditionally, FPGAs are programmed by the Hardware Description Language (HDL) such as VHDL and Verilog. These methods, although necessary to unleash the full power and flexibility offered by the FPGA chip, require detailed knowledge of the underlying hardware. On the other hand, NI has created an FPGA add-in module to its LabVIEW platform that allows users to complete the design using the high-level, graphical programming language (LabVIEW). After the design is complete, the LabVIEW FPGA module then translates the design to HDL and uploads the resulting low-level implementation to the FPGA chip. This feature by NI greatly reduces the learning curve and speeds up the prototyping cycle.

Figure 3.9 illustrates the block diagram of the ABEL trap implementation. Here, each square block is a loop structure operating at high speed (either at fixed rates indicated by
Figure 3.8: Hardware architecture of NI PCI-7833R/PCIe-7842R FPGA devices. From “NI R Series Intelligent DAQ User Manual”.
the number at the top-left corner or varying rates depending on the task to process). Communications between these asynchronous loops are carried out by FIFO queues (i.e. First In, First Out, dashed arrows in Figure 3.9). The Kalman filter calculation is broken up into multiple sub-steps to maximize the degree of parallelism on the FPGA (termed “pipelining”, see NI’s white paper “Optimizing your LabVIEW FPGA VIs: Parallel Execution and Pipelining”\(^2\) for more details). Because the FPGA can only handle integer math, all variables are scaled and rounded to integers of appropriate numbers of bits (e.g. 1 µm is represented by \(2^{12} = 4096\), see Appendix D) and division operations are implemented as look-up tables.

It is important to realize that all the position estimation and feedback generation steps are performed by the FPGA hardware alone, with no involvement of the PC. The computer is needed for updating the parameters of the trap (e.g. feedback gain, data acquisition time, etc...), recording data from the FPGA and also providing configuration and data visualization interfaces for the user. To stream data from the FPGA, data are bit-packaged into three 64-bit words (Appendix D) and transferred to computer memory via the three DMA (direct memory access) channels on the FPGA. Details of the trapping code will be released in the future.

### 3.6 Trap calibration

This section describes experimental procedures that are needed before successful trapping experiments can be performed. The overall goal of these calibration steps is two-fold. First, we make sure that the actual laser scanning pattern indeed follows the design. Second, we measure the actual “phase” of the scanning pattern in order to map each detected photon to a reference position.

#### 3.6.1 The quality of the fast scanning pattern

The optical design predicts that it takes ~150 ns to form a stable sound wave across the laser beam as it interacts with one AO deflector. To see how fast the scanning speed can be

\(^2\)http://www.ni.com/white-paper/3749/en/
in reality, I directly imaged the reflection of a circular scanning pattern (discretized to 16 points) at different point dwell times. Figure 3.10 displays the scan pattern at 1 µs per point (left) and 500 ns per point (right). In both cases, individual dwell positions were clearly resolved, although the contrast of the faster scan is slightly lower. No qualitative changes in the reflective image were observed when the dwell time is longer than 1 µs and smearing between the points started to appear for dwell times shorter than 500 ns. I thus selected the shortest beam dwell time to be 600 ns.

3.6.2 Creating the scan pattern: position calibration and brightness homogenization

To accurately program the scanning beam to hit each desired spot position, I designed two calibration strategies. The first method uses a CCD camera mounted on the microscope (conjugated to the eyepiece) to image the reflection of the laser beam. The position of the beam is extracted from fitting the reflected image and then compared to the desired target position on the grid. The errors are then used to iteratively steer the beam towards the
target. This procedure is repeated for each grid position. The result is two $6 \times 6$ matrices of the RF frequencies (along X and Y) required to position the beam on the grid. This method is easy to implement and generally works well in practice.

However, there are two practical issues that this first method ignores in the calibration process. First, for acousto-optic deflectors, the efficiency of the deflected beam decreases as the acoustic frequency deviates from the center frequency. This creates a “vignetting” effect on the square-shaped scanning pattern (e.g. the beam at the four corners of the grid is weaker). To correct for this effect, we need to homogenize the scan pattern. Second, for unbiased position tracking of a single molecule, we need to ensure that the detection probability is the same within the square-shaped ROI in the confocal detection scheme outlined in Section 3.3.

To properly account for these practical issues, I designed a second calibration scheme. First, a single fluorescent nanosphere (~100 nm in diameter) is placed on a piezoelectric scanning stage and moved to a target position. Subsequently, a feedback loop is activated to steer the excitation beam (by tuning the RF frequencies) on top of the bead, so that the fluorescent intensity recorded by the APD in the confocal path is maximized. While the beam stays locked to the bead, the RF power on the AO crystal is adjusted to tune the intensity of the beam. The process is repeated for each grid position, where the scan stage is used to move the bead to each new desired position at the same time that the beam is moved.
to that position. The result is three $6 \times 6$ matrices: the first two contain the RF frequencies (along X and Y) required to position the beam on the grid and the third contains the RF power values required to even out the illumination intensity within the ROI. Note that the possible “vignetting” effect created by the confocal pinhole is also properly accounted for by this calibration method.

### 3.6.3 Static trap response

The final step in the calibration process is to run the trapping algorithm on a static fluorescent object, scan the object around the trapping region and record the feedback response. Figure 3.11 illustrates an example. Here, the time averaged intensity (panel a) matches the design (Figure 2.4) quite well and the estimated XY positions (panel b) are orthogonal and agree with the real positions of the bead (controlled by the piezo stage) near the center of the trapping region. During this step, it is critical to adjust the “phase” of the look-up table for accurate position mapping. This is because there is a fixed amount of time delay in the system between beam scanning and photon detection and the “correct phase” can be found by the observation of a non-tilting or rectilinear position contour (like the one in Figure 3.11b). The trap is thus ready to be used for capturing single fluorescent molecules in solution.

### 3.7 The microfluidic chip

The microfluidic chip is the most critical component of all ABEL trap experiments. Not only does it hold the miniscule amount of buffer solution that provides a non-perturbative environment for single-molecule investigation, its unique geometry also facilitates the generation of electrokinetic forces at the micro-meter scale. Two different kinds of chips are being used in the Moerner lab: One variety is made up of a patterned PDMS (polydimethylsiloxane) block and a glass coverslip and the two pieces are irreversibly bonded by a short plasma etching treatment (a detailed protocol is summarized in Bockenhauer (2013)). The other incarnation of the chip is made entirely of quartz, in which a patterned piece is bonded with a quartz coverslip using sodium silicate chemistry. The quartz cell has a fluorescence
Figure 3.11: Calibration of the position sensitivity of the ABEL trap. (a) Experimental time-averaged intensity map, acquired by scanning a 100 nm fluorescent bead over the trapping region. (b) Position sensitivity of the apparatus, shown as contour plots of measured position (interval: 0.25 µm) of the fluorescent bead in b. Position response is linear and orthogonal in an area of 2 µm × 2 µm near the center. Reprinted with permission from Wang et al. (2011) Copyright American Chemical Society.
CHAPTER 3. MAKING THEM ALL WORK: ABEL TRAP INSTRUMENTATION

background that is at least 10× lower than that of the PDMS cell but is much more expensive to fabricate and more challenging to work with. The design of the microfluidic chip was originally developed by Adam Cohen (Cohen et al., 2005b; Cohen et al., 2008) in the early years of the ABEL trap technology and has been used without further modification. On the other hand, numerous technical refinements have been made by fellow ABEL trap users, mostly regarding the bonding, cleaning and surface passivation of the quartz cells. Here I discuss some of these technical issues.

3.7.1 Quartz cell fabrication, bonding and cleaning

The patterned quartz pieces (Figure 3.12b) are fabricated on a 4-inch wafer using a two-step etching process in the Stanford nano-fabrication facility by Eric Perozziello. In the first step, the deep channels are formed by a 49% HF etch for 9.5 min. Using a stylus profiler from the Quake lab, I measured the depth of the deep channels to be ~12 µm. This is consistent with the established etching rate of 1.2 µm min\(^{-1}\) on the material. In the second lithography step, the shallow trapping region (light yellow in Figure 3.12) is formed by the slower and more controllable plasma etching process (6 sccm O\(_2\) and 85 sccm CHF\(_3\), 40 mTorr for 21 min). I measured the actual depth to be ~480 nm by the profiler.

After the second lithography step, the wafer is first post-processed to remove the photoresist and the polysilicon stencil, diced and drilled before chip assembly. To bond the fabricated piece to a quartz coverslip, we use sodium silicate as an intermediate layer. The protocol, detailed in Jiang (2012), is first adapted from Wang et al. (1997) by Adam Cohen and later refined by Yan Jiang and Randall Goldsmith. My experiences with this procedure have been very positive. The bonding is reproducible and shows high resistance to harsh cleaning procedures such as Piranha and KOH sonication. In principle, direct bonding (Jacobson et al., 1995), which involves annealing the two pieces placed in contact at 1100° could also be used for an even higher bonding strength (Wang et al., 1997), but we have not yet tried this.

Unlike the PDMS cell, which is disposed after each experiment, the quartz cell is used repeatedly until the bonding comes apart. This requires thorough cleaning of the interior
Figure 3.12: ABEL trap microfluidic chip. (a) Idealized 3D model (b) Optical micrograph of the patterned quartz wafer after the second lithography step.

of the chip before each experiment, which turns out to be a significant challenge. Generally speaking, the goal of the cleaning step is to restore the interior surface to the uncontaminated and hydrophilic state. The current cleaning procedure, developed by Randall Goldsmith and Yan Jiang, uses UV-ozone clean followed by vigorous sonication in 1 M KOH (see Appendix A.2 of Jiang (2012)). Although this procedure generally works well, I have noticed a number of drawbacks. First, sonication in 1 M KOH weakens the bonding and randomly generates pellet-like debris in the chip that is difficult to remove. I suspect that the debris is released from partially damaged bonding areas. Second, I have found that UV-ozone cleaning of freshly bonded cells generates a fluorescence background that is readily detectable with 594 nm excitation. This fluorescence background can be gradually eliminated with KOH sonication and is not present in cells that has underwent many KOH sonication cycles. I suspect that the origin of the background is from UV-ozone activated fluorescent impurities in the sodium silicate bonding layer. To avoid potential problems from UV-ozone radiation and sonication, I started to use Piranha clean (3 : 1 mixture of H₂SO₄ and H₂O₂) followed by 1 M KOH rinse (no sonication). This alternative cleaning procedure works well and seems to have extended the longevity of a bonded chip. I include a step-by-step protocol in Appendix B.
3.7.2 Surface passivation

Nonspecific adsorption of biomolecules to the quartz/glass/PDMS surfaces is a critical problem when applying the ABEL trap for biophysical studies. In fact, most proteins stick to an untreated or KOH-activated glass/quartz surfaces in aqueous buffer, presumably due to a combination of electrostatic and hydrophobic interactions. Due to the enormous surface-to-volume ratio of the microfluidic geometry, adsorption is the primary cause for protein loss in the trap. To eliminate nonspecific binding of biomolecules, the interior of the cell surface needs to be “passivated”. Despite the huge amount of literature on “surface passivation of microfluidic devices” (Horvath et al., 2001; Zhou et al., 2010), very few methods actually work well in the ABEL trap scenario due to the more stringent requirements of the passivation method. First, to be compatible with single-molecule detection, the reagents used in the passivation procedure needs to be free of fluorescent impurities. Second, due to the difficulty of performing complex chemical reactions inside the microfluidic chip, in situ or solution-based methods are desirable. Third, the passivation needs to be extremely effective in the sense that it completely eliminates a biomolecules’s affinity to surfaces. This is because when a molecule is trapped in the shallow channel, its chances of interacting with the (upper and lower) surfaces are increased due to geometry restrictions (500 nm) along the axial dimension, as compared to experiments done in less confined geometries. Over the years, the ABEL trap team has identified and perfected a number of effective strategies for surface passivation and these are now discussed.

Dynamic coating  One method in which we have found a lot of success is adding a small amount of coating agent to the final sample mixture before injection into the trap. The idea is that the coating agent, in great excess compared to the molecules of interest, would cover the surface and eliminate adsorption sites. Three different coating agents have been proven to be effective. The first is a polymer solution from Applied Biosystems (now part of Life Technologies) called POP6 (performance optimized polymer 6, without denaturant). It comes as a viscous colorless solution and adding the POP6 stock solution to ~2-10% of the final volume effectively eliminates sticking. Although in our hands this works for many proteins and lipid vesicles, we really don’t know what the actual composition of this proprietary mixture is (other than it is a linear polyacrylamide). Moreover, I found that the
recent commercial batches of POP6 has an unacceptable amount of fluorescent impurities with 488 nm excitation.

The second coating agent is the protein casein (EMD Millipore 70955-3), which is widely used as a superior blocking agent in western blot. Inspired by Whitney Duim’s positive experience of using casein to block the sticky Huntington proteins (Duim, 2012), I tried casein as a dynamic coating agent in the ABEL trap and it worked quite well for allophycocyanin and, as reported by other ABEL trap users, the peridinin-chlorophyll-protein (Bockenhauer et al., 2013) and LHCII (Hsiang-yu Yang, personal communication). Problems associated with casein are a) unacceptable amount of fluorescent impurities with 488 nm excitation and b) incompatibility with the cleaning procedure of the quartz cell, as 1 M KOH makes the protein form large aggregates that tend to block the trapping region.

The third coating agent is the polymer polyvinylpyrrolidone (PVP, Sigma-Aldrich PVP40), which is widely used in the electrophoresis community (Horvath et al., 2001; Milanova et al., 2012). I use the 40 kDa PVP at a final concentration of 0.1% (by weight) and it is generally very effective in eliminating sticking for proteins and proteoliposomes. At the same time, it effectively eliminates electro-osmotic flow, which is desirable for trapping highly charged analytes, such as DNA and proteoliposomes. On the other hand, the loss of electro-osmosis mobility makes trapping near-neutral proteins difficult. I found PVP to contain very few fluorescent impurities at either 488 nm or 594 nm excitation.

**Polyelectrolyte multilayer (PEM)** One drawback of dynamic coating in general is the need to mix in high concentrations of the coating agent that might interfere with the single molecules of interest. To remedy this, Randall Goldsmith successfully adapted the polyelectrolyte multilayer (PEM) procedure to the ABEL trap chip and I have subsequently optimized the protocol. PEM eliminates non-specific interactions by the electrostatic assembly of multiple charged polymer layers on top of a negatively charged Si/SiO$_2$ surface (Decher, 1997). Although many different polymer combinations can work, we use polyethylenimine (PEI) as the positive polymer and polyacrylic acid (PAA) as the negative polymer (Kartalov et al., 2003). To deposit PEM onto the interior surface, the microfluidic cell is incubated first with a positively charged polymer (PEI) followed by a negatively charged polymer (PAA) and this procedure is repeated to enhance the effectiveness of the
coating. I find that two pairs of coatings (four layers in total) are generally sufficient to prevent sticking but three pairs are sometimes needed in the case of an aging quartz chip. I summarize a detailed protocol of the PEM procedures in Appendix B. Due to the high surface charge density of the terminating layer, PEM brings strong electroosmotic flow, which is particular advantageous for trapping biomolecules that are near neutral in buffer.

**Future approaches** Surface modification with polyethylene glycol (PEG) is another popular method in the single-molecule field to prevent non-specific adsorption (Roy et al., 2008). Although PEG has not been extensively tried in the ABEL trap geometry, I want to mention two other approaches that seem promising. One approach is based on a polylysine (PLL) grafted PEG polymer (Kenausis et al., 2000; Lee et al., 2005), which utilizes the electrostatic interaction between the positively charged lysine side chains and the negatively charged surface to achieve surface PEGylation. This approach is extremely simple to use and has been successfully used in a number of single-molecule experiments (Valentine et al., 2006; Dunn et al., 2007; Mathiasen et al., 2014). The other approach is based on a one-step silanization reaction with 2-[methoxy(polyethylenxy)propyl]-trimethoxysilane (Sui et al., 2006) and has been applied in the recent single-molecule work by Tyagi et al. (2014), whose microfluidic chip shares some similarities to the ABEL trap.

### 3.8 Other measurement modalities with the ABEL trap

Simultaneous measurements of multiple fluorescence parameters can be extremely useful in single-molecule spectroscopy to separate sub-populations and gain physical insights into the different fluorescent states. The ABEL trap is an ideal platform to conduct multi-dimensional spectroscopy of single molecules in aqueous solution. Here I describe how to implement the different measurement modalities with the ABEL trap and other technical considerations.
3.8.1 Fluorescence lifetime

**Apparatus** Time correlated single photon counting (TCSPC) has been a widely used technique in measuring the excited-state lifetime of a fluorescent sample and is straightforward to apply in single-molecule experiments. To perform TCSPC, a pulsed source (typically ps or fs at 10’s of MHz repetition rate) and a fast detector capable of timing each photon with <100 ps time resolution are needed. Since the new ABEL trap design uses beam modulation and a point detector to sense single-molecule position, it is directly compatible with TCSPC if a pulsed laser is used. In this mode, the timing information of each fluorescent photon is analyzed in two different timescales: the sub-nanosecond delay between the signal photon and the laser pulse is recorded to extract excited-state lifetime, while the microsecond arrival time is used to map the position of the scanning laser beam for position estimation.

It is not always straightforward to find a pulsed source in the visible spectral window, where most of the “single-molecule suitable” dyes are in resonance. Within the scope of
this thesis, I have used three different laser sources for pulsed excitation: a frequency-doubled fiber laser emitting at 515 nm, a Ti:Sapph-pumped supercontinuum-generating fiber and an optical parametric oscillator (OPO). a) Mode-locked fiber lasers have evolved to become a cheap alternative to Ti:Sapph technology. Our model (Mercury 2000-200-MOD, Polaronyx, San Jose) outputs ~100 mW of 1030 nm light and is frequency doubled using a PPLN crystal to 515 nm (about 5 mW). Details of using this laser is documented in the PhD thesis of Bockenhauer (2013). b) Supercontinuum generation is a relatively economical solution for producing pulses with tunable wavelength. I use a nonlinear photonic crystal fiber by Newport (SCG-800) which is optimized for pumping with a Ti:Sapph oscillator. Using the Mira, I can routinely generate ~90 mW of continuum light with ~500 mW pumping at around 780 nm. The output spectrum depends strongly on input power and wavelength and representative spectra are shown in Figure 3.14. Given a power density of ~0.5 mWnm$^{-1}$, the supercontinuum is best suited for experiments that does not demand a lot of power (i.e. <1 mW) at a single wavelength. Selection of the exact wavelength required is performed by a narrow band excitation filter. c) Finally, the OPO provides both...
Figure 3.15: Example measurements of the APD instrument response function. A tight focus on the detector is needed to achieve the optimal timing resolution.

power and tunability but is an expensive and complex solution. Our system (Coherent Mira-OPO) is capable of generating ~ 100 mW light with a tuning range from 530 nm to 620 nm.

On the detection side, the critical components are the avalanche photon-detector (Si-APD) and a high-resolution timing module. For the detector, I use the Perkin-Elmer SPCM series. These typically have a dark count rate of ~300 cnts/s and a timing resolution of ~300 ps and for a long time, were the only choice for single-experiments. I have noticed that some of these units started to show increased dark count levels (to about 2000 cnts/s) and increased after-pulsing probabilities after about 10 years of use. Recently, other suppliers (e.g. PicoQuant, MPD) have entered the market. I have tested a pair of \(\tau\)-SPAD (Kell et al., 2011) detectors from PicoQuant and found them to be superior in performance compared to the Perkin-Elmer ones. However, the long term stability of these new detectors has yet to be tested.

For quantitative analysis of experimental TCSPC data, the instrument response function (IRF) of the APD needs to be characterized carefully. This is commonly done by recording a decay histogram using highly attenuated emission from the pulsed laser, either via scattering or back reflection from a glass coverslip. However, care must be taken when carrying out this characterization step, due to some peculiar features of the APD. First, the temporal position of the IRF depends on wavelength ("color shift"), so that the IRF should be measured with a source that is spectrally close to the fluorescence signal of interest.
(preferably with the fluorescence from a dye that has a lifetime that is much shorter than the expected IRF). This is sometimes difficult so a color shift term is explicitly included in the fitting function (see below). Second, on some units/models, the IRF depends on the count rate (Kell et al., 2011). Consequently, the IRF needs to be measured at a count rate similar to the real experiment. Third, the width of the IRF critically depends on the size of the focusing spot, the bigger the spot, the broader the response. Although the size of the active area is generally specified to be about 150 µm, I found empirically that a sub-300 ps IRF can only be achieved under tight focusing conditions (estimated spot size <50 µm). Examples of IRF, measured with loose and tight focusing, are shown in Figure 3.15. Of course, one should be careful to not focus so tightly as to burn the APD surface. This is unlikely with single-molecule emission, but if other more intense beam are under study, one should be careful.

I use the PicoHarp 300 (PicoQuant) timing module to register (time-tag) each detected photon to picosecond time resolution. For recording lifetime data, the PicoHarp is set to operate in “pt3” mode, with channel 0 connected to the sync pulses of the laser (usually from an internal fast diode) and channel 1 connected to the APD. In “pt3” mode, the relative delay time (~ns) between the two channels is recorded. Those delays correspond the time lag between absorption and fluorescence photon emission and can be histogrammed during post-processing to extract the excited-state lifetime. The PicoHarp 300 has two input channels and recording multi-channel lifetime decays (using the same laser sync) can be achieved using a router (PHR 800).

Analysis To extract excited-state lifetime, the first step is to construct a photon-pulse delay time histogram. I generally use a bin time of 80 ps, which is a good tradeoff between shot-noise and time resolution. The histogram is fitted with the following functional form:

\[
g(t; \tau, c) = (1 - \gamma) \left[ \text{IRF} \left( t - c \right) \otimes \exp \left( -\frac{t}{\tau} \right) \right] + \gamma g_{BG}(t) \tag{3.8}
\]

where \( \gamma \) is the background fraction, \( g_{BG}(t) \) is the experimentally measured decay histogram from background photons, \( \text{IRF} \) is the instrument response function, \( c \) is the temporal shift of the IRF and \( \tau \) is the excited state lifetime. Fitting is performed by a maximum likelihood approach which correctly models Poissonian statistics in photon counting (Zander et al.,
The likelihood of observing a particular decay histogram is the product of the likelihood of individual bins $l_k$

$$L(\{n_1, n_2, \ldots, n_K\}; \tau, c) = \prod_{k=1}^{K} l_k$$

(3.9)

which, given a Poisson noise model, can be written as

$$l_k(n_k; \tau, c) = \frac{[N \tilde{g}(t_k; \tau, c)]^{n_k} e^{-N\tilde{g}(t_k; \tau, c)}}{n_k!}$$

(3.10)

where $\tilde{g}$ is the discretized version of Equation 3.8, $k$ is the bin index, $n$ is the number of photon in a particular bin and $N$ is the total number of photons in the decay. By taking the logarithm of Equation 3.8 and dropping constants, we can greatly reduce the computational burden during the maximum likelihood search. Fitting errors in $\tau$ (68% confidence interval) can be determined by inversion of the observed Fisher information matrix (Pawitan, 2001),

$$\Delta \tau = \sqrt{(J^{-1})_{11}}$$

(3.11)

where $J$ is the observed Fisher information matrix (the negative of the Hessian matrix during likelihood maximization).

### 3.8.2 Fluorescence emission spectrum

**Apparatus** The fluorescence emission spectrum was one of the first physical properties measured for single molecules on a surface at room temperature (Trautman et al., 1994; Macklin et al., 1996; Lu et al., 1997). It contains rich information about an individual molecule’s excited-state energy landscape and the molecule’s immediate nanoenvironment. Generally, the emission spectrum is generated by dispersing the fluorescence photons using a prism or a grating. Comparing the two, a prism generally has a lower insertion loss while gratings are designed to have more dispersive power. At room temperature, the spectral features of a single molecule are broadened by phonons and as a result, a spectrometer with a ~1 nm resolution would generally be sufficient. On the other
hand, with standard EMCCD technology, photons captured by an array detector would lose their nanosecond timing resolution. This means that, using the current ABEL trap feedback strategy, the photons used to determine spectral information can not be used for feedback actuation. Consequently, high transmission efficiency is strongly desirable in the spectrometer for the ABEL trap. A number of possible solutions exist in the literature (e.g. holographic grating (Mirzov et al., 2006), spherical prism (Cutler et al., 2013)) and we chose the Amici prism, which has about 90% transmittance across the visible spectrum, as the dispersion element. The Amici prism has the added benefit of being easy to configure in a direct-view geometry (Hagen et al., 2011), which means that it can be conveniently inserted into an imaging optical path in situ when the spectral information is needed. The Amici prism is widely used in hand-held spectrometers and its use in single-molecule spectroscopy first appeared in a study of the peridinin–chlorophyll–protein complex (Wörmke et al., 2007).

In our setup, the fluorescence emission spectra of trapped single molecules are recorded by diverting 30%-50% of the collected signal photons to the Amici prism (Edmund Optics, NT42-586) before projecting onto an EMCCD camera (Andor DU860E). One minor problem associated with direct insertion of the Amici prism into the imaging path is the astigmatism that degrades spectral resolution. I address this problem by placing a lens (f = 250 mm) before the prism and tuning the lens-prism distance so that the sagittal focal plane is imaged on the EMCCD. In other words, the function of the lens is to make sure that the line-shaped single-wavelength-response is aberrated along the direction perpendicular to the dispersion axis (as opposed to the direction along the dispersion axis), so that spectral resolution does not deteriorate because of the astigmatism. Both the prism and the lens are mounted on magnetic bases and are inserted when needed. Without them, the camera is conjugated to the sample plane for direct imaging of single molecules and as a diagnostic tool. Two additional measures are taken to optimize the signal-to-noise ratio of the single-molecule spectrum. First, the spectra are recorded in “multi-track” configuration using the frame transfer mode and only lines with > 5% of the peak intensity are vertically binned. Second, the EMCCD is set with a “vertical shift speed” of 0.1 µs. I noticed that setting the number higher than this induces a significant amount of spurious charges that deteriorate
Figure 3.16: Spectral calibration of the Amici prism-based spectrometer. (a) A spectrum of the HeNe discharge taken by the prism spectrometer. (b) A spectrum of the HeNe discharge taken by the Ocean Optics spectrometer (red: raw data, black: after convolution with a Gaussian kernel with a standard deviation of 1.6 nm). (c) Result of pixel-to-wavelength mapping from matching the peak positions in (a) and (b) (black squares: peak positions, red line: fit by a third-order polynomial). Inset: raw spectrum of the 594 nm laser line and a Gaussian fit to extract the standard deviation.

Calibration Before any analysis, the raw spectral data needs to be calibrated. I use a discharge from a HeNe tube as the reference light source. The tube is taken from a cheap 632 nm HeNe laser with the cavity mirrors removed. To perform pixel-to-wavelength calibration, I first measure the spectrum of the HeNe tube discharge with the Amici-prism on the microscope (Figure 3.16a) and then compare it to the spectrum taken with a commercial spectrometer (Ocean Optics USB4000, Figure 3.16b). Due to the much higher spectral resolution of the commercial spectrometer, its recorded spectrum is convolved with a Gaussian function ($\sigma = 1.8$ nm, Figure 3.16b) before a detailed comparison. Figure 3.16c plots the pixel-to-wavelength mapping by matching the peak positions between the two spectra. To map an arbitrary pixel to wavelength, the data is fitted with a third order polynomial

---

with the following form

\[ \lambda (p_x) = a (p_x - p_{594})^3 + b (p_x - p_{594})^2 + c (p_x - p_{594}) + d \]  

(3.12)

where \(a, b, c, d\) are fitting parameters, \(p_x\) is the camera pixel to be mapped and \(p_{594}\) is the pixel position of the 594 nm laser line. The 594-line is re-measured every time before each spectral recording to compensate for drift of the spectrometer optics. The width (1 s.d.) of the measured 594.1 nm laser line (Figure 3.16c inset) can be regarded as the resolution of the spectrometer (Sparrow criterion), which is about 1 nm.

**Analysis** The first step in analyzing the spectral data is background subtraction. In the experiments using 594 nm laser excitation and an all-quartz microfluidic cell, the background spectrum is dominated by two water Raman peaks around 640 nm and 740 nm (Figure 3.17). The background spectrum is calculated by averaging the spectral frames during which is trap is not occupied by a single molecule.

After background subtraction, the spectrum can be fitted by a model function. I tried both Gaussians and skewed Gaussians (Krüger et al., 2010) and found that they both gave satisfactory results. I eventually choose to fit the spectra with a sum of two Gaussians for the single dye data (Chapter 3) and the APC data (Chapter 4). On the other hand, because the signal-to-noise ratio of our spectra are generally low, I also use center-of-mass (CoM) as a simple yet robust measure of the spectral position. The CoM of a spectrum is defined

![Figure 3.17: Typical background spectrum with 594 nm excitation. The two peaks correspond well to the 1600 cm\(^{-1}\) and 3380 cm\(^{-1}\) water Raman peaks when excited at 594 nm.](image)
as the following
\[
\bar{\lambda}_m = \frac{\sum I_k \lambda_k}{\sum I_k}
\]  
(3.13)

where \( I_k \) is the signal of the wavelength bin centered at \( \lambda_k \). To see that the CoM measure is more robust at low signal conditions, I compare it with 2-Gaussian fitting on simulated spectra at different total number of photons (Figure 3.18). Note that the two estimators do not converge to the same value and the CoM estimator has a much smaller uncertainty given the same number of photons. The disadvantage of the CoM measure is that all shape information is lost by condensing the full spectrum into one number.

I now derive an analytical form of the measurement uncertainty associated with the CoM estimator given Poissonian photon counting noise. Applying the standard error propagation rule to Equation 3.13, we get

\[
\text{var} (\bar{\lambda}_m) = \sum_k \left( \frac{\lambda_k \sum I_k - \sum I_k \bar{\lambda}_k}{(\sum I_k)^2} \right)^2 \text{var} (I_k)
\]  
(3.14)

and the variance of each spectral bin on the EMCCD can be calculated by converting the acquired signal from ADU to photons and accounting for the factor of two excess noise (i.e., \( \text{var} (I_k) \approx 2(I_k + b_k) \) and \( b_k \) is the background contribution).

### 3.8.3 Fluorescence anisotropy (polarization)

Fluorescence anisotropy (polarization) is a classic biophysical technique based on the simple physical picture that the process of fluorescence emission involves two steps: an excitation dipole absorbing a photon and an emission dipole radiating a photon. An angular offset between the two dipoles causes depolarization, which could be the result of a number of physical processes. For example, depolarization could be caused by internal relaxation (e.g. exciting at \( S_0 \rightarrow S_2 \) and emitting from \( S_1 \rightarrow S_0 \)), rotational diffusion during the time delay between the absorption and emission event (fluorescence lifetime) or resonant energy transfer (i.e. the excitation and emission dipoles are on different molecules). Due to
its direct sensitivity to these processes, fluorescence anisotropy has been a powerful tool in a wide range of disciplines (Akimoto et al., 2007; LiCata et al., 2008; Hiller et al., 2003).

Measurement of fluorescence anisotropy (polarization) generally involves exciting the sample with linearly polarized light and detecting alone two orthogonal polarization directions: one parallel ($I_\parallel$) and the other perpendicular ($I_\perp$) to the excitation polarization. The (steady-state) anisotropy ($r$) and polarization ($pol$) are defined as

$$ r = \frac{I_\parallel - I_\perp}{I_\parallel + 2I_\perp} $$  \hspace{1cm} (3.15)

$$ pol = \frac{I_\parallel - I_\perp}{I_\parallel + I_\perp} $$  \hspace{1cm} (3.16)
The two definitions differ in the normalization factor. Anisotropy uses the total intensity (Lakowicz, 2006) as the denominator and is the more “rigorous” definition. Mathematically, polarization (Equation 3.16) is related to anisotropy by

\[ \text{pol} = \frac{3r}{r + 2} \]
\[ r = \frac{2 \cdot \text{pol}}{3 - \text{pol}} \]

In an ensemble measurement, the linearly polarized excitation light interacts with many molecules and a distribution of excitation dipoles are “preselected” with a probability that scales with \( \cos^2 \theta \) (where \( \theta \) is the angle between the polarization of the source and the molecular excitation dipole). This photo-selection process is the reason that the maximum anisotropy obtainable in an ensemble measurement is \( r_{\text{max}} = 0.4 \) instead of 1. Additional depolarization processes further decreases the anisotropy value from \( r_{\text{max}} \). We discuss two common cases here. First, if there is fast (compared to fluorescent lifetime) energy migration from dipole \( A \) to dipole \( B \), the resulting anisotropy becomes (Lakowicz, 2006)

\[ r(A \rightarrow B) = 0.4 \left( \frac{3 \cos^2 \Theta_{AB} - 1}{2} \right) \]

where \( \Theta_{AB} \) is the angular displacement between the excitation dipole of \( A \) and the emission dipole of \( B \). In the second case, consider tumbling motion in solution with a rotational correlation time \( t_{\text{rot}} \), the resulting steady-state anisotropy becomes (Lakowicz, 2006)

\[ r(t_{\text{rot}}) = 0.4 \left( \frac{1}{1 + \tau/t_{\text{rot}}} \right) \]

where \( \tau \) is the fluorescence lifetime.

The single-molecule analog of the fluorescence anisotropy measurement is straightforward to conceive. However, photo-selection entails a different interpretation: instead of creating a biased molecular population, the selection acts upon a single molecule’s own rotational ensemble. Consequently, if measurement is performed over a time interval much longer than the rotational correlation time (so that the rotational degree of freedom is fully
sampled), the same theory should apply. Indeed, such measurements have been successfully performed (Schaffer et al., 1999; Ha et al., 1999) and applied to identify molecules of different sizes in solution (Schaffer et al., 1999). However, in a free diffusion experiment, the short (~1 ms) observation window and the paucity of detected photons thwarted further development of single-molecule anisotropy measurements.

The ABEL trap allows freely-rotating single molecules to be observed for seconds and thus provides an ideal platform to measure fluorescence anisotropy on individual molecules. Moreover, it enables dynamics of single-molecule anisotropy to be measured for the first time. I demonstrate the power of this technique in Chapter 5.

It is also helpful to compare solution-based single-molecule anisotropy measurements with surface-based, polarization-resolved measurements. In the latter, single molecules are immobilized and the fluorescence is split into two orthogonal polarization channels. On surface, since the dipole orientation of a single molecule is fixed, there is no photo-selection and the maximum polarization contrast is 1. Moreover, the excitation polarization can be modulated to gain more information (Ha et al., 1996; Ying et al., 1998; Peterman et al., 2001; Rosenberg et al., 2005). On the other hand, the measured polarization contrast depends strongly on the orientation of the molecule. The immobilization-induced heterogeneity of dipole orientation makes it difficult to compare results between molecules for statistical analysis. In an solution-based experiment, a molecule’s orientation information is rapidly averaged out due to rotational diffusion so that molecules are experimentally equivalent and can be directly compared. The resulting polarization contrast only depends on the relative orientations between the absorption and emission dipoles.

**Apparatus** To perform fluorescence polarization measurements on single trapped molecules, the excitation laser is converted to linear polarization before entering the microscope. I measured an extinction ratio of ~1000:1 before the objective lens. A high NA lens is known to distort the polarization at its focus (Novotny et al., 2001). However, in our case this distortion is mitigated by the fact that the scanning excitation beam does not fill the whole back aperture of the objective lens (effective focusing NA ~ 0.8). Polarization sensitive detection is achieved by inserting a polarizing beam splitter in the collection path and counting photons along the two orthogonal polarization directions. I measured the
extinction ratio of the beam splitter to be ~100:1. Since fluorescence is detected using the full NA of the objective lens, the depolarization effect in the detection path needs to be accounted for (see below).

**Analysis** Single-molecule fluorescence polarization is calculated using the following equation

\[
pol = \frac{(S_{||} - b_{||}) - (S_{\perp} - b_{\perp})}{(S_{||} - b_{||}) + (S_{\perp} - b_{\perp})} / g \tag{3.20}
\]

Here, \( S \) and \( b \) are the number of signal and background photons, respectively. \( || \) and \( \perp \) denote the parallel and perpendicular channels and \( g \) is a parameter that corrects for the uneven detection efficiency of the two channels. The theory of steady-state single-molecule fluorescence polarization is similar to its bulk counterpart, namely, the polarization contrast is a result of photo-selection and angular displacement between the excitation and emission dipoles during the spontaneous emission process. One bit of complication arises from the high NA collection optics used in single-molecule schemes: the maximum anisotropy of 0.4 assumes paraxial optics and will be lower in a high-NA system. This effect has been well known (Axelrod, 1979; Koshioka et al., 1995) and corrected for (Koshioka et al., 1995; Schaffer et al., 1999) in previous measurements. Here I develop a simple quantitative model to account for the depolarization effect.

**Depolarization by a high-NA objective** To start, we consider the polarization state of a fixed dipole on the sample plane (Figure 3.19a). We assume that the orientation of a single dipole emitter is described by \((\Theta, \Phi)\), where \( \Theta \) is the polar angle with respect to the optical axis (z), and \( \Phi \) is the angle with respect to the polarization axis (in this case x). If this dipole emitter is being imaged by a high-NA objective, the emission intensities after the objective lens, decomposed into parallel and perpendicular components \((I_{||} \text{ and } I_{\perp})\) have been derived by Fourkas (2001) to be,

\[
I_{\|}^{\Theta,\Phi} = I_{tot} (A + B \sin^2 \Theta + C \sin^2 \Theta \cos 2\Phi) \tag{3.21}
\]
\[
I_{\perp}^{\Theta,\Phi} = I_{tot} (A + B \sin^2 \Theta - C \sin^2 \Theta \cos 2\Phi) \tag{3.22}
\]
where the factors $A$, $B$ and $C$ are functions of the objective NA (parameterized by the collection half angle $\alpha$).

\[
A = \frac{1}{6} - \frac{1}{4} \cos \alpha + \frac{1}{12} \cos^3 \alpha \\
B = \frac{1}{8} \cos \alpha - \frac{1}{8} \cos^3 \alpha \\
C = \frac{7}{48} - \frac{1}{16} \cos \alpha - \frac{1}{16} \cos^2 \alpha - \frac{1}{48} \cos^3 \alpha
\] (3.23) (3.24) (3.25)

This result is similar to the pioneering work of Axelrod and has also been used to analyze the polarization state of a single molecule immobilized on a surface (Börner et al., 2012).

Next, we perform averaging by photo-selection. Assuming that the excitation polarization is along the $x$ axis, a weighting factor that is proportional to $|\hat{\mu} \cdot \vec{E}|^2 \propto \cos^2 \Phi \sin^2 \Theta$ needs to be applied before integration over the full solid angle.

\[
I_\parallel = \int_0^{2\pi} \int_0^\pi I_{\parallel}^{\Theta, \Phi} \cos^2 \Phi \sin^2 \Theta \sin \Theta d\Theta d\Phi \\
I_\perp = \int_0^{2\pi} \int_0^\pi I_{\perp}^{\Theta, \Phi} \cos^2 \Phi \sin^2 \Theta \sin \Theta d\Theta d\Phi
\] (3.26) (3.27)
From Equations 3.26 and 3.27 we can calculate the maximum anisotropy in the presence of depolarization. As shown in Figure 3.19b, the value of 0.4 is valid only at small NAs (NA < 0.2). For a water immersion objective with a NA of 1.2 (α ≈ 63), $r_{\text{max}}$ is reduced to about 0.33.

We now consider the case where there is an angular displacement between the excitation and emission dipoles. Solving this problem requires the use of Euler angles (Axelrod, 1979; Amerongen et al., 2000). First, we define a molecular reference frame $(a, b, c)$. In this frame, the absorption dipole is expressed as $\hat{\mu} = c$ and the emission dipole is $\hat{\nu} = a \hat{a} + b \hat{b} + c \hat{c}$ (where $a, b, c$ are the direction cosines in the molecular frame). To get the coordinates in the lab frame, we left multiply the following transformation matrix by the molecular frame coordinates,

$$M = \begin{bmatrix}
\cos \psi \cos \phi \cos \theta - \sin \psi \sin \phi & -\sin \psi \cos \phi \cos \theta - \cos \psi \sin \phi & \cos \phi \sin \theta \\
\cos \psi \sin \phi \cos \theta - \sin \psi \cos \phi & -\sin \psi \sin \phi \cos \theta - \cos \psi \cos \phi & \sin \phi \sin \theta \\
-\cos \psi \sin \theta & \sin \psi \sin \theta & \cos \theta
\end{bmatrix}$$

(3.28)

where the $\theta, \phi, \psi$ are the Euler angles. Consequently, the emission dipole in the lab frame can be expressed as

$$\begin{bmatrix}
\sin \theta' \cos \phi' \\
\sin \theta' \sin \phi' \\
\cos \theta'
\end{bmatrix} = M \begin{bmatrix}
a \\
b \\
c
\end{bmatrix}$$

(3.29)

Specifically

$$\theta' = \arccos \left( -a \cos \psi \sin \theta + b \sin \psi \sin \theta + c \cos \theta \right)$$

(3.30)

$$\phi' = \arctan \left( \frac{a \left( \cos \psi \sin \phi \cos \theta - \sin \psi \cos \phi \right) - b \left( \sin \psi \sin \phi \cos \theta + \cos \psi \cos \phi \right) + c \sin \phi \sin \theta}{a \left( \cos \psi \cos \phi \cos \theta - \sin \psi \sin \phi \right) - b \left( \sin \psi \cos \phi \cos \theta + \cos \psi \sin \phi \right) + c \cos \phi \sin \theta} \right)$$

We then perform orientational averaging on the excitation dipole

$$I_\parallel = \int_0^{2\pi} d\psi \int_0^{2\pi} d\phi \int_0^\pi I_{\parallel, \phi'} \cos^2 \phi \sin^2 \theta \sin \theta d\theta$$

(3.31)

$$I_\perp = \int_0^{2\pi} d\psi \int_0^{2\pi} d\phi \int_0^\pi I_{\perp, \phi'} \cos^2 \phi \sin^2 \theta \sin \theta d\theta$$

(3.32)
Equations 3.31 and 3.32 have to be evaluated numerically. Figure 3.19c plots the polarization contrast with a high-NA objective and compares it to Equation 3.18 (converted to polarization using Equation 3.17).

**Calibration**  In Equation 3.20, the $g$ factor, which is the ratio of photon counting efficiencies between the two polarization channels, needs to be calibrated experimentally. I calibrate it by measuring the apparent polarization contrast between the two channels at two orthogonal excitation polarizations with a bulk fluorescent sample. The idea and the procedure are now presented. For clarity, I use $p$ and $s$ to denote the polarization axes defined by the polarizing beam splitter ($p =$ parallel to optical table) and $\|$ and $\bot$ to specify the signal components (after the objective lens) with respect to the excitation polarization. In the first step, I set the excitation polarization to be along $x$ (Figure 3.13). In this case, $\|$ = $p$ and $\bot$ = $s$. The signals on the $p$ and $s$ detectors are $L_p = T_pI_\|$ and $L_s = T_pI_\bot$. Here, $T$ denotes the transmission efficiency of a detection channels. The apparent polarization contrast is

$$a = \frac{L_p - L_s}{L_p + L_s} = \frac{I_\| - gI_\bot}{I_\| + gI_\bot}$$ (3.33)

where $g = T_s/T_p$ is the definition of the correction factor. Similarly, when the excitation polarization is rotated 90° to be along $y$ (Figure 3.13), the apparent polarization contrast for the same sample is (now $\|$ = $s$ and $\bot$ = $p$)

$$b = \frac{L_s - L_p}{L_s + L_p} = \frac{T_sI_\| - T_pI_\bot}{T_sI_\| - T_pI_\bot} = \frac{I_\| - gI_\bot}{I_\| + gI_\bot}$$ (3.34)

If we assume that the ratio between $I_\|$ and $I_\bot$ is the same in the two measurements, we have

$$\frac{I_\|}{I_\bot} = \frac{g(1 + a)}{1 - a} = \frac{1 + b}{g(1 - b)}$$ (3.35)

so the $g$ factor can be obtained using

$$g = \sqrt{\frac{(1 - a)(1 + b)}{(1 + a)(1 - b)}}$$ (3.36)
Table 3.5: Calibration of $g$ factor for single-molecule anisotropy measurements. See text for the definition of $a$ and $b$. The rows in the table refer to the samples in Chapter 5.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$a$</th>
<th>$b$</th>
<th>$g$</th>
<th>corrected pol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atto 647N</td>
<td>0.096</td>
<td>-0.02</td>
<td>0.89</td>
<td>0.027</td>
</tr>
<tr>
<td>XL-APC</td>
<td>0.12</td>
<td>0.0034</td>
<td>0.89</td>
<td>0.055</td>
</tr>
<tr>
<td>NXL-APC</td>
<td>0.37</td>
<td>0.23</td>
<td>0.86</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Table 3.5 summarizes the three samples I used to characterize the $g$ factor of the setup. An average value of $g = 0.87$ is used in the analysis. I found that the 50:50 beam splitter (Figure 3.13) used to conduct simultaneous spectral measurements accounted for most of the dichroism in the system.

The uncertainty in the single-molecule polarization measurements can be estimated as follows. Applying the standard error propagation rules to Equation 3.20, we get

$$
\text{var}(p) = \left( \frac{\partial p}{\partial S_{\parallel}} \right)^2 \text{var}(S_{\parallel}) + \left( \frac{\partial p}{\partial S_{\perp}} \right)^2 \text{var}(S_{\perp}) + \left( \frac{\partial p}{\partial b_{\parallel}} \right)^2 \text{var}(b_{\parallel}) + \left( \frac{\partial p}{\partial b_{\perp}} \right)^2 \text{var}(b_{\perp})
$$

(3.37)

where we’ve assumed that the $g$ factor is known without uncertainty. If we define $I_{\text{tot}} \equiv (S_{\parallel} - b_{\parallel}) + (S_{\perp} - b_{\perp}) / g$ and assume that the photon counting uncertainty follows Poisson statistics (i.e. $\text{var}(S) = S$), we have

$$
\text{var}(p) = \frac{4(S_{\perp} - b_{\perp})^2}{I_{\text{tot}}^4 g^2} (S_{\parallel} + vb_{\parallel}) + \frac{4(S_{\parallel} - b_{\parallel})^2}{I_{\text{tot}}^4 g^2} (S_{\perp} + vb_{\perp})
$$

(3.38)

where $\text{var}(b) = vb \approx 0$, because background levels can be determined accurately by averaging over many traces. A similar equation has been derived previously by Lidke et al. (2005).

### 3.8.4 Fluorescence anti-bunching

Anti-bunching refers to a unique feature of a single-molecule emitter: the probability of emitting two photons within a time interval shorter than the fluorescence lifetime is zero. This simply represents the fact that after emitting a photon, a single molecule needs to
undergo two sequential Poisson processes before another emission event is possible: re-absorbing a photon and spontaneous (radiative) decay back to the ground state. Consequently, the second emission event can not be infinitely close to the first one but is spaced out by at least the fluorescence lifetime. Fluorescence anti-bunching is regarded as the ultimate proof that a single emitter is being observed (Basché et al., 1992; Lounis et al., 2000a). Because of the anti-bunching behavior, the emission from a single molecule is highly non-classical and has been exploited as an “on-demand” photon source (Lounis et al., 2000b) for quantum optics research.

Photon anti-bunching is generally visualized as a “dip” in the inter-photon delay histogram near zero delay (or equivalently, from the intensity correlation function near nanosecond lag time). Recent developments in single-molecule spectroscopy has extended the application of anti-bunching measurements to probe a number of interesting molecular properties. One type of analysis focused on the degree of the anti-bunching dip to extract the number of independent emitters in a molecular complex (Weston et al., 2002; Hollars et al., 2003). Interestingly, even systems that contain multiple emitters sometimes show a near-complete dip. This phenomenon can be reconciled by the presence of efficient singlet-singlet annihilation pathways, which has been found in a number of synthetic multi-chromophoric systems (Hofkens et al., 2003) and multimeric fluorescent proteins (Sánchez-Mosteiro et al., 2004). Another application of anti-bunching type measurement is studying nanosecond dynamics. For example, conformational dynamics can be extracted by analyzing the nanosecond photon correlation of a FRET system (Berglund et al., 2002; Nettels et al., 2007; Hillger et al., 2008). Spectral dynamics (Sallen et al., 2010) can be studied in a similar fashion by the anti-bunching signal from two color channels.

It is important to realize that the fluorescence anti-bunching is easier to observe with a near-saturation pumping rate. This is because only absorption-emission cycles that are closely spaced in time (~ns) contribute to the anti-bunching contrast and these events are rare under non-saturating conditions. As a result, anti-bunching experiments using freely-diffusing molecules needed ~hours of integration time (e.g. 50 h in Nettels et al. (2007)) and the final anti-bunching curve was averaged over a large number of molecules (> 10,000). The advantage of performing anti-bunching type measurements in the ABEL trap is the greatly reduced data collection time (~ minutes). It is also possible to measure
anti-bunching signal on singles in the trap, without ensemble averaging, although this re-
quires a stable emitter that can withstand the high (near saturation) pumping intensities. 
Another subtle advantage of performing anti-bunching measurements in the ABEL trap, 
when the signal from a single molecule is not enough for anti-bunching analysis, is the 
ability to conduct sub-population averaging. Here, sub-populations may be identified by 
other fluorescence variables (e.g. lifetime, spectrum, brightness, etc). We use this type of 
analysis in Chapter 5 where we look at the photophysical properties of Allophycocyanin, 
an essential antenna protein.

**Apparatus**  For anti-bunching experiments, I adopt the classic Hanbury Brown and Twiss 
configuration. The detection path consists of a 50:50 non-polarizing beam splitter and a pair 
of photon counters (Fore et al., 2007). A second detector is needed in this case because 
current single photon counters have a ~50 ns dead time. It turns out that the state-of-the-
art single photon counters have another peculiar feature: they emit light after detecting a 
photon (Dravins et al., 2000; Kurtsiefer et al., 2001). Referred to in the community as “after 
glow”, this phenomenon is thought to originate from the avalanche process. This spurious 
emission from the detector induces “false” coincidences which are much stronger than the 
anti-bunching signal from single molecules. I found that by placing bandpass filters right 
before each detector, the crosstalk can be completely eliminated. However, this method 
might not work if the spectrum of interest overlaps with the spectrum of the after glow 
(which was determined to be in the near infrared (Kurtsiefer et al., 2001), but possibly
device dependent). Other methods can be used to solve the crosstalk problem (Becker, 2005).

I use the PicoHarp 300 for the anti-bunching measurements. Both the “pt3” and “pt2” modes can be used in this case. In “pt3” mode, the delay times between the two detectors are recorded and can be directly histogrammed to yield the anti-bunching curve. In “pt2” mode, the arrival time of each detected photon in both detectors are recorded and the anti-bunching curve can be constructed by cross-correlating the two photon arrival records at ~ ns lags. During the recording, a cable delay of ~20 ns is introduced in one of the detectors to visualize the negative delays. I use the algorithm developed by Laurence et al. (2006) to calculate the cross-correlations.

**Analysis**  The anti-bunching curve is fitted using the following model function

\[
g(t,n,c) = \left[ 1 - \frac{1}{n} \exp\left(-\frac{|t-c|}{\tau}\right) \right] \otimes IRF
\]

where \( n \) is the number of independent emitters, \( \tau \) is the rise time constant, \( c \) is the overall shift of the whole curve and \( IRF \) is the experimentally measured instrument response function. Here, the response function is taken from a control experiment using a 515 nm femtosecond laser as the start-stop signal (Figure 3.21 inset). Note that Equation 3.39 does not contain a background term. This is because that under the ABEL trap measurement conditions, background photons contribute minimally to the acquired anti-bunching histogram. A typical coincidence histogram of the background is shown in Figure 3.20.

Two example anti-bunching measurements are shown in Figure 3.21. The sample used 10-nucleotide single-stranded DNA molecules labeled with one fluorescent probe. Two different labels, Cy5 (Figure 3.21a and b) and Atto 647N (Figure 3.21c and d) were tried and compared. As illustrated in Figure 3.21a, molecules in solution were caught one-by-one by the ABEL trap and their emission signal was analyzed to produce the anti-bunching histogram (intensity spikes indicate transient co-occupancy of the trap by two molecules). In Figure 3.21, the histograms were generated using the emitted photons from
Figure 3.21: Example anti-bunching data in the ABEL trap. (a) Example intensity traces of trapped ssDNA-Cy5 molecules. Inset: instrument response function acquired from recording a start-stop histogram of femtosecond laser pulses. (b) The anti-bunching histogram of ssDNA-Cy5, averaged over ~100 molecules and a fit to Equation 3.39. (c) The anti-bunching histogram of ssDNA-Atto647N molecules, showing a delay range of −20 to 200 nanoseconds. (d) A fit to the anti-bunching dip for the ssDNA-Atto647N sample.

~100 molecules. Nevertheless, given the extreme stability of some red dyes under optimized buffer conditions (Vogelsang et al., 2008), it should be possible to obtain the anti-bunching signal from singles. From quantitative analysis, the differences in fluorescence lifetime between Cy5 and Atto 647N was clearly resolved and the number of independent emitters was extracted to be close to one, as expected. Note that the number of coincidence photons at zero delay is not truly zero in these measurements and this is due to IRF broadening and the large time bins (~128 ps) used in the calculation.
3.8.5 Two-color setup and FRET

Here I discuss the implementation of a two-color ABEL trap. This setup is equipped with two excitation lasers of different wavelengths and both excitation lasers are simultaneously modulated by the same pair of acousto-optic deflectors. Three measurement modes are available for this setup: trapping with the emitted light produced with either color, trapping with one color while probing with the other (with uniform excitation intensity) and trapping on the emitted photons produced by both lasers. The setup first originated when Dr. Jianwei Liu and I decided to measure, using FRET, the conformational dynamics of the protein Rhodanase as it interacts with a chaperonin protein. Later, it found extensive use in the FoF1 ATP synthase experiments carried out by Sam Bockenhauer and Prof. Michael Börsch (see Bockenhauer (2013)).

The key challenge here is that the acousto-optic deflector is a color-dependent dispersive element so that different colors will not travel colinearly after interaction with the very first deflector. I found that using the afocal relay system between the two deflectors (Section 3.2.2), it is possible to obtain reasonably good 2D scan quality for two distinct colors. The trick is to first align one of the laser lines (red in Figure 3.22), optimize the two AO deflectors, and then pre-compensate the second laser line (blue in Figure 3.22) in the vertical tilt direction (using DM1) so that the Bragg condition on the Y (AOMy) deflector can be fulfilled for the second laser. With this technique, I could simultaneously deflect ~80% of the first laser and ~50% of the second laser to the (1, 1) order and obtain good beam quality. However, after the vertical AO deflector, the two laser lines still have to be spectrally separated (using DM2), individually collimated and recombined before the microscope in order to generate overlapping foci at the sample plane. To make sure that the two beams scan in phase, it is critical to ensure that the number of reflective surfaces on both paths has the same parity.

3.9 Data processing software

During the development of the hardware, I’ve also written a number of routines to process the trapping data acquired from the FPGA. Most of these functions are written in C and
Figure 3.22: Two-color trapping with one pair of acousto-optic deflectors. The two laser lines (red and blue) are combined by a dichroic mirror (DM1) and coupled to the same pair of acousto-optic deflectors (AOMx and AOMy) for phase-locked modulation. After AOMy, the two lines are separated using another dichroic mirror (DM2), individually collimated and expanded (not shown), and recombined before the microscope using a third dichroic mirror (not shown).
compiled into Matlab using the MEX interface. On top of those core routines I then wrote Matlab-based GUIs to visualize and analyze the data. The organization chart of the software is illustrated in Figure 3.23. Sample codes are given in Appendix D. The full set of data processing routines will be made available in the future.
Figure 3.23: Architecture of ABEL trap data processing system, from acquisition to visualization. This diagram corresponds to version TrapIIb.
Chapter 4

Photodynamics of single fluorophores in solution

In this chapter, I use multi-parameter fluorescence detection in the ABEL trap to study the photodynamics of single fluorophores in solution. To this end, I first demonstrate successful trapping of individual fluorophores for ~1 s in aqueous solution. This represents the ultimate limit for the ABEL trap technology and also pave the way for studying photodynamics that happen at ~10 ms to 1000 ms timescale in aqueous buffer. Three different red dyes that are commercially available, Atto 647N, Atto 633 and Alexa 647 were chosen as the target. Interestingly, I observed brightness dynamics in Atto 647N and Atto 633 but not in Alexa 647. To better characterize these dynamics, I simultaneously measured brightness, lifetime and emission spectrum of individual trapped molecules. I discovered that dynamics in the different parameters were correlated and the photodynamics can be well-described by transitions between a small number of “emissive states”. In the case of Atto 647N, three states were observed. These results represent a significant advance in the ability to identify and characterize different dynamical states of single molecules in aqueous solution with high precision and millisecond time resolution. The work presented in this chapter was published in Wang et al. (2013) and Wang et al. (2014a).
Figure 4.1: Experimental setup illustrating microfluidic cell, optical excitation and single-photon detection as well as spectral recording, and FPGA-based electrokinetic feedback. A 32-point “knight’s tour” scanning pattern (upper right) was used to quickly determine the position of diffusing objects in real-time. Both beam size ($1/e^2$ waist radius) and the spacing of the grid were about 0.5 µm. See Chapter 3 for a detailed description of the instrumentation. SM/PM = single mode/polarization maintaining. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.

4.1 Multi-parameter fluorescence detection in solution using the ABEL trap

It was realized early on in the single-molecule field that different physical properties of the emitted fluorescence photons can be measured to extract different information about a single molecule or its local environment. For example, the polarization state of the fluorescence photons can be analyzed to report on the orientation of the molecular dipole (Rosenberg et al., 2005; Sosa et al., 2001). By measuring the ns-delayed emission times of the fluorescence photons with respect to a short excitation pulse, the excited state lifetime can be obtained which reports on the character of de-excitation pathways (Yang et al., 2003).
CHAPTER 4. PHOTODYNAMICS OF SINGLE FLUOROPHORES IN SOLUTION

Emission spectra, obtained by binning each photon by its frequency, can be used to infer aspects of the excited state energy landscape of the system (Oijen et al., 1999; Blum et al., 2001).

Simultaneous measurements of multiple parameters maximize the information obtained from each single molecule (Macklin et al., 1996; Vanden Bout et al., 1997). Additionally, analyzing the correlated changes between different variables in a time-dependent fashion on a molecule-by-molecule basis provides insights into dynamic processes that are difficult to extract from ensemble-averaged measurements. For example, the photophysics of fluorescent protein DsRed has been elucidated by simultaneous intensity, lifetime and emission spectrum measurements (Cotlet et al., 2001).

Most previous experiments using multi-parameter fluorescence detection were performed on surfaces (Macklin et al., 1996; Vanden Bout et al., 1997; Tinnefeld et al., 2001; Prummer et al., 2004; Xu et al., 2007; Börner et al., 2012), where additional heterogeneity could be introduced by the act of immobilization (Friedel et al., 2006). In many cases, solution phase experiments are preferred. However, in free solution, multi-parameter measurements of single molecules, although useful, have information limits due to the paucity of detected photons per molecule (Widengren et al., 2006). Analysis usually involves statistical averaging between many molecular events (Chen et al., 1999; Widengren et al., 2006). Fluorescence lifetimes from freely diffusing molecules during the brief ms-long transit through a tight laser focus have been determined (Edman et al., 1996; Eggeling et al., 1998), but usually with large uncertainties. Full emission spectra of single diffusing molecules have not been determined in solution. Also, in diffusion-based measurement schemes, slow dynamics of individuals (>10 ms) are inaccessible due to the ~ms observation window. Here, I demonstrate and apply multi-parameter spectroscopy in solution using the ABEL trap. This can be achieved by incorporating multi-modal detection optics into the ABEL trap setup. Details of the configuration to measure lifetime and emission spectrum are discussed in Chapter 3 and the setup used for the current study is sketched in Figure 4.1. I then use simultaneously obtained information about every single molecule’s brightness, life time and emission spectrum to characterize the photodynamics of trapped single fluorophores.
CHAPTER 4. PHOTODYNAMICS OF SINGLE FLUOROPHORES IN SOLUTION

4.2 Experimental Details

4.2.1 Sample preparation

All single dye trapping experiments were performed in 1X PBS (pH 7) buffer containing a triplet quencher (1 mM Trolox) (Rasnik et al., 2006), a protocatechuate oxygen scavenger system (50 nM protocatechuate-3,4-dioxygenase and 1 mM protocatechuic acid) (Aitken et al., 2008), and 10% glycerol (to avoid rapid sample evaporation during experiments). Final concentration of the dyes ranged between 1 pM to 5 pM. Dyes (Atto 647N and Atto 633 (Atto-Tech GMBH), Alexa 647 (Invitrogen)) with a N-hydroxysuccinimide functional group were used in this study without further purification. To avoid labeling the surface lysines of protocatechuate 3,4-dioxygenase, dye solutions were hydrolyzed by dilution into 1X PBS buffer at ~5 µM concentration and stored at 4 °C overnight before use.

4.2.2 ABEL trap algorithm and microfluidic environment

The single dye trapping experiments were conducted using the “knight’s tour” scanning, Kalman filter-based trap (version TrapIIa, detailed in Chapter 3) in an all-quartz microfluidic cell with a channel depth of ~600 nm in the axial z-direction (Cohen et al., 2008). The cell cleaning procedure in this case started with a 30 min UV ozone clean, followed by 30 min sonication in 1 M KOH and nanopure rinse. No surface passivation method was needed in this study. Electro-osmosis was determined to be the dominant force in feedback actuation.

4.2.3 Data analysis

All data analysis was performed with home-written software in Matlab. Before single-molecule identification and extraction steps, the intensity (brightness) versus time data was preprocessed to determine the background level. First, the photon-by-photon arrival time data was converted to a ‘mcs’ trace (counts/bin time) with a 5 ms bin window. A change point (CP) finding algorithm (Watkins et al., 2005) with a large critical value was then performed on the 50 ms binned mcs trace as a smoothing step. A k-means algorithm with
Figure 4.2: Single fluorophore trapping. (a) Example fluorescence intensity traces of Atto 647N molecules before and after feedback is switched on, with the feedback voltage records (summed every 10 ms, bottom strip, red: x, blue: y. (b) Example traces of Atto 633, infrequent, transient dips were likely due to blinking. (c) Voltage autocorrelation function from the shaded region in b and fit to an exponential decay with a time constant of 120 µs. (d) Trapping duration histogram of Atto 647N and fit to an exponential decay. Average trapping time from this data set is 1.3 s. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.
CHAPTER 4. PHOTODYNAMICS OF SINGLE FLUOROPHORES IN SOLUTION

N clusters (N > 3) subsequently classified every CP bin to a certain cluster. Background is considered to be represented by the cluster with the lowest averaged intensity. With every intensity bin now classified as either background or not, single-molecule events were identified as starting from the first non-background bin and ending at the last bin before a background bin. Only non-sticking molecules that last more than 200 ms were selected for further analysis. Sticking events (to quartz surfaces) were easily identified by the presence of non-symmetric feedback voltages and were rare (<3% for Atto 647N and Atto 633, <1% for Alexa 647) under the buffer conditions used (~100 mM Na+ salt (Daniels et al., 2010)). With selected single molecules, the intensity CP algorithm was performed once more with a small critical value to capture all possible intensity dynamics. Lifetime analysis and emission spectrum analysis were performed on each CP intensity level, defined as data between two adjacent change points.

4.2.4 Simultaneous fluorescence lifetime and emission spectrum measurements

The experimental details are described in Chapter 3.

4.3 Results

4.3.1 Trapping single fluorophores in solution

Before feedback was switched on, short intensity spikes were observed due to diffusing molecules undergoing transient interactions with the confocal scanning pattern (Figure 4.2a, left). After feedback was enabled, single dye molecules, upon random diffusion into the beam scanning region, became localized near the center of the trap. Due to the time-averaged homogeneous excitation profile provided by the beam scan pattern, near constant emission intensities were observed, exemplified by typical traces shown in Figure 4.2a. Several experimental observations provided confirmation that the objects trapped were indeed single dye molecules. First, when the feedback is off, the interaction time with the
Figure 4.3: Representative traces of trapped single fluorophores with simultaneously recorded excited-state lifetime and emission spectrum. The same excitation power and spot size was used in all three cases so the differences in measured intensities reflect the relative brightness of the single emitter with 594 nm excitation. Reprinted with permission from Wang et al. (2014a) Copyright (2014) Society of Photo-Optical Instrumentation Engineers.
4 µm scanning pattern, as estimated from the height of the fluorescence spikes, is only several milliseconds. This short dwell time is consistent with an object with a large diffusion coefficient diffusing through the trapping region ($\sqrt{4D\Delta t} = 2.5\,\mu m$ for $D \approx 300\mu m^2 s^{-1}$, a typical single dye value, and $\Delta t = 5\,ms$). Second, stable trapping of single fluorophores could only be achieved when the detected photon count rate was higher than around 30 kHz, while stable trapping of single-dye labeled proteins (~100 kDa, roughly eight times slower than a single dye) tolerated a much lower count rate of 3 kHz to 5 kHz. This is consistent with the design principle of the ABEL trap that an increase of diffusion coefficient by $k$ requires an increase in feedback bandwidth by the same factor in order to maintain comparable trapping strength (Moerner, 2007). Third, the measured fluorescence lifetime and emission spectrum in the trap (see below) were in excellent agreement with the manufacturer’s specifications and literature values for the three different dyes (Atto 647N, Atto 633 and Alexa 647) in aqueous solution.

Zero mean, rapidly fluctuating trapping voltage records (Figure 4.2a and b bottom) suggest that the long dwell times in the trap were not due to surface sticking. Sub-millisecond autocorrelation of the applied voltages showed exponentially decaying behavior (with a decay time of ~100 µs, Figure 4.2c), further confirming the lack of short-time surface interactions. Since applied voltages are proportional to position estimates, the decay constant of voltage correlation can be used to estimate the trapping strength (Fields et al., 2011). Under our experimental conditions (594 nm pulsed excitation, oxygen removal and Trolox as the triplet quencher), we achieved the best trapping with Atto 647N. In the best experimental run (131 molecules), the average trapping duration was 1.3 seconds (Figure 4.2d). We harvested an average of 51700 photons per single molecule (max: 247000 median: 32800). An example trace showing trapping of Atto 633 (cw excitation) is shown in Figure 4.2b.

For the dye Atto 647N, some molecules showed one unique intensity level during their residence time, while others switched between different intensity states (Figure 4.2a). We next investigated these different intensity state behaviors in more detail with simultaneous intensity/lifetime/emission spectrum measurements in the ABEL trap.
CHAPTER 4. PHOTODYNAMICS OF SINGLE FLUOROPHORES IN SOLUTION

Figure 4.4: Simultaneous intensity, lifetime and emission spectrum measurements of a single Atto 647N molecule in solution. This molecule contains only one unique intensity level. (a) Histogram of lifetimes (blue trajectory in b, 250 photon binned) with a Gaussian fit (solid black, standard deviation $\sigma = 0.39 \text{ ns}$) (b) Intensity (red), lifetime (blue) and spectral (purple) trajectories of the trapped Atto 647N molecule. Intensity is displayed in 5 ms bins, lifetime is constructed by successive fitting of every 250 photons. Spectrum is represented by the center of mass of each 50 ms spectral frame. (c) Auto-correlation of lifetime, calculated with the 250 photon binned data in a. (d) Lifetime determined from all photons from the intensity level. (e) Intensity auto-correlation; fluctuation at short time scale is due to beam scanning motion. (f) Emission spectrum determined from all photons from the intensity level. Dashed purple line indicates the peak position. Position of the black cross indicates the center of mass of the spectrum, which is the quantity plotted in a.

Figure 4.5: Example of an Atto 647N molecule showing dynamics. (a) Intensity, lifetime and spectral dynamics. Intensity (red) is displayed in 5 ms bins. Both lifetime (blue) and spectrum (center of mass, purple) are determined from every 50 ms spectral frame. Intensity levels identified by the change-point finding algorithm are drawn in black. (b-d) Fluorescence lifetime determined for each color coded region in a. (e-g) Emission peaks determined for each color coded region in a. The different colors correspond to the clusters in Figure 4.6. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.
4.3.2 Simultaneous measurement of brightness/lifetime/emission spectrum in the ABEL trap

For every molecule held in the ABEL trap, fluorescence lifetime and emission spectrum were measured simultaneously along with the brightness (fluorescence intensity). Representative traces of trapped single fluorophores are shown in Figure 4.3. Figure 4.4b shows an example of a multi-parameter measurement of a trapped Atto 647N molecule. In this particular molecule, only one unique intensity level is present. Intensity is quantified by the number of photon counts per 5 ms and is plotted in red. Intensity autocorrelation (Figure 4.4e) showed no blinking at short timescales (down to 100 μs). From the time-tagged photons, a fluorescence lifetime trajectory (blue, left axis) was constructed by successively fitting every 250 photons. The distribution of lifetimes in this molecule was well-fitted by a Gaussian function (Figure 4.4a). The width of the lifetime distribution (s.d. 0.39 ns) was found to be comparable to the average of the fitting uncertainty (1 s.d., 0.38 ns), suggesting no lifetime dynamics. With a photon counting rate of 52 kHz in this case, lifetime dynamics can be probed with 4.7 ms time resolution with every 250 photons. Autocorrelation analysis of the lifetime trajectory (Figure 4.4c) was performed and further confirmed the absence of significant dynamics. An emission spectrum trajectory, represented by the center of mass of each spectral frame (50 ms time resolution), is plotted in purple (Figure 4.4b). No apparent spectral dynamics during the residence time of the molecule was observable. Due to the general lack of dynamics of all three parameters within a single intensity level, we pooled all the photons during that level and determined its time-average intensity (black line in Figure 4.4b), lifetime (Figure 4.4d and blue dotted line in Figure 4.4b) and spectrum (Figure 4.4f and purple dotted line in Figure 4.4b), where the expected presence of a vibrational sideband was evident.

4.3.3 Atto 647N switches between different intensity levels with distinct spectra and lifetimes

Surprisingly, a number of Atto 647N molecules were observed to switch between distinct intensity levels during their residence in the trap. Figure 4.5a provides an example of a
Figure 4.6: Clustering analysis of Atto 647N emission states. Distribution of all the intensity levels, projected onto (a) intensity-spectrum (b) lifetime-spectrum (c) intensity-lifetime spaces. Different states (clusters) are color coded for clarity. Marginal densities along each variable are shown on the sides. Each curve of the marginal density has an integrated area that is proportional to the total number of levels in that cluster. (d) Average spectrum from each cluster, normalized by the peak. The sharp drop near 745 nm is due to filter cutoff. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.
molecule that visited multiple intensity states. Lifetime (blue) and spectral (purple) dynamics are displayed above the intensity trajectory and are particularly interesting. Here, lifetime is determined for every spectral frame (50 ms) with error bars determined by the Fisher information analysis (Chapter 3). Within one intensity level, changes in both lifetime and emission spectrum are generally small (except the region following a blinking event at 208 s, where lifetime lingers sometimes below and sometimes above the level average while no significant changes were present in intensity/spectrum), consistent with the molecule shown in Figure 4.2. However, abrupt changes in both lifetime and spectrum happen along with an intensity switch. Changes between lifetime and intensity are correlated (decrease in intensity accompanies decrease in lifetime and vice versa) while lifetime and spectrum changes are anti-correlated (red shift accompanies reduction in lifetime and vice versa). A total of three intensity states are present in this particular molecule (exemplified by the color coded regions). Figure 4.5b-g show lifetime (blue) and emission spectrum fits of the color coded regions. The emission spectra from the three levels differ in shape as well.

In order to determine whether this switching behavior is common, we conducted measurements on a total of 683 Atto 647N molecules. For every molecule, we identify its intensity levels by the change-point finding algorithm (1153 levels in total). We then extract the fluorescence lifetime and emission peak wavelength of each intensity level and plot it as a point in the intensity-lifetime-emission peak 3D parameter space. Figure 4.6a-c show 2D distributions of all the levels as different projections of the 3D parameter space. Three distinct clusters can be identified in intensity/spectrum and intensity/lifetime spaces. Representative examples of these states are shown in Figure 4.5 and Figure 4.2. As an attempt to quantitatively describe these clusters, we used the k-means algorithm in 3D parameter space to identify and classify all levels into 3 populations. As a result, the levels are color coded by the cluster to which they belong. Marginal densities of every cluster along each parameter dimension are plotted on the sides of the distribution. Specifically, we have a low-intensity (cluster averaged mean: 165 cnts/5ms), low-lifetime (3.20 ns) and red-shifted (672 nm) state (red, 19.3% of total levels); a high-intensity (233 cnts/5ms), high-lifetime (3.60 ns) and blue-shifted (658 nm) state (blue, 15.5%) and finally a state (green, 65.2%) with intermediate intensity (202 cnts/5ms), lifetime (3.46 ns) and emission peak (667 nm).
Figure 4.7: Transitions in single Atto 647N molecules. (a) Changes in intensity versus changes in emission peak, positive values indicate red shifts. (b) Changes in lifetime versus changes in emission peak. (c) Changes in intensity versus changes in lifetime. (d) Transition probabilities between states. The relative positions of the three states correspond qualitatively to Figure 4.6a. Black arrows indicate transitions between states. Unfilled arrows indicate trap loss/photobleaching. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.
CHAPTER 4. PHOTODYNAMICS OF SINGLE FLUOROPHORES IN SOLUTION

135

Figure 4.8: Rates of spectral fluctuation in Atto 647N. (a) Dwell time histogram before a blue shift is fitted to an exponential decay. (b) Dwell time histogram before a red shift is fitted to an exponential decay. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.

The most abundant “green” state has properties that are similar to bulk (lifetime 3.5 ns, emission peak 669 nm) (Kolmakov et al., 2010). Solution phase identification of the “red” and “blue” states is only possible by our new technique. The spread in lifetime and emission peak positions within a cluster are larger than the fitting errors, suggesting additional inhomogeneity within the cluster. Figure 4.6d shows the cluster averaged spectra of the three states. The “green” and “blue” states are similar in shape despite the ~10 nm shift in peak positions while the “red” state is broader and has an altered vibronic structure.

4.3.4 State transitions in Atto 647N

Our cluster analysis has unveiled three emission states in Atto 647N in solution. From the example in Figure 4.5 and the ability to follow the same molecule for a long time, we also know that transitions happen between states. To better visualize the state transitions, we plot the correlated parameter changes which occur with each identified change point (Figure 4.7a-c). Every transition is color coded by the ending state. As shown in these plots, the types of transitions seen in Figure 4.5 are prevalent among Atto 647N molecules: upon shifting to the “blue” state, the molecule undergoes increase in intensity, blue shift in emission peak, and an increase in lifetime; upon shifting to the “red” state, the molecule undergoes decrease in intensity, red shift in emission peak and a decrease in lifetime. Using
all the transitions observed, Figure 4.7d illustrates the transition probabilities of state transitions \( P(y|x) \), defined as the probability of going into state \( y \) when it is at state \( x \). Unfilled arrows indicate photo-bleaching or trap loss, which is the dominant pathway for all states. Interestingly, the “red” and “blue” states do not interconvert and both tend to transition back to the “green” state. From the “green” state, the molecule can go to either “red” or “blue” states, although going to the high-intensity, blue shifted state is more likely. Self-transition (indicated by an arrow pointing back to the starting state) of the “green” state is higher than the other two, suggestive of possible substructure within that state.

### 4.3.5 Rates of spectral fluctuation in solution

It is of practical interest to characterize the timescale of spectral fluctuations of Atto 647N in solution. Before we proceed, however, it is worth noting that in our case, trap loss/photo-bleaching happened at a rate around \( 1 \text{ s}^{-1} \), so slow dynamics were not sufficiently sampled. Our result should thus be regarded as an upper bound of the rates. In Figure 4.8, dwell times before a red/blue spectral shift are histogrammed and fit to exponential decays to extract the rates. A spectral shift is defined as when the difference in emission peaks is larger than 1.5 times the maximum fitting error. A slightly higher rate for red-shifting was obtained.

### 4.3.6 Photodynamics of Atto 633

We observed similar photodynamics in Atto 633. A representative Atto 633 molecule that switched emissive states during its residence in the trap is shown in Figure 4.9a. This molecule entered the trap at a lower intensity with slightly shortened excited-state lifetime and red-shifted emission spectrum compared to bulk values. At \( \sim 79.7 \text{ s} \), its brightness increased (\( \sim 20\% \)) with a simultaneous increase in lifetime (\( \sim 14\% \)) and a blue shift (\( \sim 18 \text{ nm} \)) in emission spectrum. There was also a noticeable change in the shape of the emission spectrum. To better characterize the different emissive states of Atto 633, we conducted measurements on 638 molecules. A total of 1182 intensity levels were identified using the change-point finding algorithm (Figure 4.9a solid black line). Similar to the case with Atto 647N, we extracted the excited-state lifetime and the center of mass of the emission spectrum.
Figure 4.9: Example photodynamics of Atto 633. (a) Intensity trace. (b, d) Lifetime and spectrum for the red time interval; (c, e) for the blue time interval. Reprinted with permission from Wang et al. (2014a) Copyright (2014) Society of Photo-Optical Instrumentation Engineers.
spectrum for each intensity level, and plotted it on a three-dimensional (intensity-lifetime-spectral position) parameter space as a scatter plot. 2D projections of this data set are presented in Figure 4.10a-c. Apparently, a significant amount of heterogeneity exists in the emissive properties of single Atto 633 molecules in solution as well. The majority of the levels can be described by three clusters: a red (~675 nm, spectral center of mass), dim (~120 cnts/5ms) state with low (~2.9 ns) lifetimes, a bulk-like state with intermediate brightness (~170 cnts/5ms), lifetime (~3.2 ns) and spectrum (~669 nm) and a slightly blue-shifted (~666 nm) state with increased brightness (~200 cnts/5ms) and lifetime (~3.6 ns). The three states of Atto 633 highly parallel those observed for Atto 647N, suggesting that the two dyes might share similarities of conformational configurations. On the other hand, in contrast to Atto 647N, Atto 633 has a number of additional rarely populated but distinctive emissive sub-states (e.g. Figure 4.10 red circle). Population-averaged spectra are shown for the “blue” state and the most red-shifted sub-state in Figure 4.10d.

4.3.7 Alexa 647 is more homogeneous

Similar measurements and analysis were performed on Alexa 647 (220 molecules, Figure 4.11a), a dye structurally similar to the widely studied fluorophore, Cy5. Unlike the two Atto dyes, only one population with a relatively narrow emission peak distribution is observed. There is some asymmetry in the lifetime and intensity distributions, but no obvious sub-populations can be identified. The average fluorescence lifetime is determined to be 1.13 ns, consistent with measurement done at the ensemble level. The 10% glycerol present in our buffer slightly increases the lifetime of Alexa 647 from 1.0 ns (Buschmann et al., 2003). An example intensity trace, together with its spectral and lifetime is shown in Figure 4.11b-d.

4.4 Discussion

The ability to discover and characterize molecule-to-molecule heterogeneity has been one of the most celebrated feats of single-molecule spectroscopy (Moerner, 2010). However, in many cases, heterogeneity may be induced by sample preparation. For instance, when
Figure 4.10: Emissive states of Atto 633. (a) intensity-spectrum, (b) lifetime-spectrum and (c) intensity-lifetime mapping of all intensity levels extracted from 638 trapped Atto 633 molecules (insets: density plots). Dashed circles indicate certain (sub)populations in all three 2D projections of the parameter space. (d) Averaged emission spectrum of the (sub)populations indicated by dashed circles. Reprinted with permission from Wang et al. (2014a) Copyright (2014) Society of Photo-Optical Instrumentation Engineers.
molecules are immobilized on a surface, factors such as differences in dielectric environment, hydrogen bonding network, electrostatic interactions with the surface, molecular orientation, denaturation and restricted degrees of rotational freedom might all affect the emission properties of the molecule. One of the great advantages of conducting solution phase measurements with the ABEL trap is the ability to suppress surface-induced artifactual heterogeneity, while harvesting enough photons for precise measurement of intrinsic properties of single molecules. An alternative strategy, encapsulation in a vesicle (Benitez2008NL; Cisse et al., 2007), can provide similar advantages, but the number of molecules which can be studied is larger with the ABEL trap. In our experiments, every molecule experiences the same aqueous environment, orientation and polarization effect are rapidly averaged out by rotational diffusion. Indeed, our measured lifetime/emission peak distributions of Alexa 647 are much narrower than previous reported surface immobilized experiments on Cy5 (Tinnefeld et al., 2001). Similar effects of reduced heterogeneity in solution have been observed in the antenna protein allophycocyanin (Goldsmith et al., 2010).

The distinct emission states of Atto 647N could result from different conformations of the molecule. Previous experiments on another rhodamine, sulforhodamine 101, on surfaces (Lu et al., 1997; Yip et al., 1998) reported correlated dynamics of intensity and emission peak. In their case, higher intensity corresponds to a blue shift (~200 cm\(^{-1}\)) in emission spectrum. We have observed a similar trend and shifting magnitude for Atto 647N in solution. In the case of sulforhodamine 101, Lu et al. (1997) found two timescales for spectral fluctuation, a relatively fast decay component of 1.9 s\(^{-1}\), which was independent of excitation power and a slow component of 0.022 s\(^{-1}\), which had a quasi-linear dependence on excitation power (Lu et al., 1997). The fast component was assigned to spontaneous (thermally driven) fluctuations while the slow component was likely light-induced. We have not done intensity-dependent characterization of the spectral fluctuations of Atto 647N, although we made an observation which indicates that transition to the “blue” state is likely light-driven while transition to the “red” state is less so. When molecules first enter the trap (initial level), relative populations were 20.7% (red), 74.2% (green) and 5.1% (blue). However, the relative populations shifted to 19.8% (red), 64.1% (green) and 16.1% (blue) right before molecules left the trap. Many molecules were already in the “red” state
Figure 4.11: Alexa 647 lacks significant photodynamics. (a) Mapping of each intensity level on lifetime-spectrum and intensity-spectrum spaces. (b) Example intensity trajectory of a trapped Alexa 647 molecule, its concomitant lifetime and spectrum. Reprinted with permission from Wang et al. (2013) Copyright (2013) American Chemical Society.
while very few were in the “blue” state before they entered the trap, suggestive of a “red” and “green” only equilibrium before laser illumination. Apparently, interaction with the laser deposited more population to the “blue” state, shifting the equilibrium. It is also possible that since some molecules are lost from the trap before photobleaching, some of the molecules which first enter the trap had been previously irradiated.

Subtle differences in the shapes of emission spectra are observed for the three states (Figure 4.5 and Figure 4.6). Several vibronic sidebands are visible, and their relative contributions compared to the main emission band change for the various states. Although it is not possible in this initial study to fully determine the precise mechanism, these spectral changes might be due to enhanced coupling to certain sets of vibronic modes after slight rearrangement of intra-molecular nuclear coordinates. Of course, whatever coordinate changes occur, the overall highly emissive $\pi$ electron core structure of the molecule is not altered dramatically, so one can only surmise that the changes could be due to conformational changes of more distant functional groups. Recently, Lee et al. (2013) found that a number of red-emitting dyes become brighter in heavy water D$_2$O (or other solvents less capable of hydrogen-bond formation compared to water). In their study, increases in brightness are accompanied by increases in excited-state lifetime and changes in emission spectrum (~10 nm). Although this effect is most prominent for oxazine dyes and Cy7 (Klehs et al., 2014), nearly all dyes they tested show measureable brightness enhancement in heavy-water (Atto 647N: 30%). They suggested a model in which hydrogen-bonding facilitates non-radiative pathways that quenches fluorescence emission. Given the similarities between the emissive characteristics of Atto 647N in heavy water and the “blue” state found here, light-driven transitions into different hydrogen-bonding configurations offer one possible explanation of the observed behavior here. We do not think that the switching behavior was caused by the feedback electric fields, because the rapid varying applied fields seem incompatible with the slow (~100 ms) switching observed here.

As noted in the early days of single-molecule spectroscopy, intensity fluctuations associated with spectral changes can be explained by shifts in absorption profile (Trautman et al., 1994). We have also observed lifetime changes with spectral/intensity changes. In
particular, lower intensity, redder states tend to have lower lifetimes, indicating that additional quenching pathways activated after a conformational switch, might also play a role in reducing the observed brightness.

Recently, Le Reste et al. (2012) noticed that Atto 647N, when covalently attached to ssDNA, interconverts slowly between two or more intensity states. The predominant dimmer state is 20% lower in intensity. This finding corresponds well to the “red” state (18.5% dimmer than the “green” state) discovered in our case. Vogelsang et al. (2008) also mentioned two emission states that are “spectrally distinct but otherwise photophysically similar” when labeled on oligonucleotides. The conversion rates are similar in these studies (~10 s) despite order of magnitude differences in excitation power (28 W cm$^{-2}$ by Le Reste et al. and 2000 W cm$^{-2}$ by Vogelsang et al., both with 638 nm excitation), which also indicates that the transition to the dimmer state is not light driven. Our excitation rate (4000 W cm$^{-2}$ at 594 nm) is similar to that used by Vogelsang et al., but we have observed a much faster conversion rate (~0.5 s). It might be possible that the negatively charged backbone of the oligonucleotides has a stabilizing effect on the conformational switching dynamics of the fluorophore. It is also possible that the rate of dynamics depends on the buffer conditions, especially the oxygen scavenging agent and triplet quencher.

Both Atto 647N and Atto 633 have become increasingly popular in single-molecule studies due to their excellent brightness/photostability. Observation of intrinsic brightness fluctuations of these dyes suggests that care must be exercised when using their intensities as reporters in single-molecule assays (Di Fiori et al., 2010; Ha et al., 2012; Chung et al., 2010).

4.5 Summary

In this chapter, we have demonstrated simultaneous intensity, lifetime, and spectral measurements of single fluorophores in the ABEL trap. In aqueous solution, such measurements have not been possible previously. Our apparatus is capable of capturing objects as small as a single fluorophore when the photon detection rate exceeds about 300000 s$^{-1}$, needed for fast sampling of position dynamics before diffusive escape. The high photon emission rates also allow us to measure fluorescence lifetime and emission spectra with a
high time resolution (4.7 ms for lifetime and 50 ms for spectrum in our case). I used the apparatus to study the intrinsic photodynamics of Atto 647N, Atto 633 and Alexa 647. Three distinct emission states with slightly different intensities, lifetimes and emission peaks were observed in the dye Atto 647N and Atto 633. Transitions were observed to happen between a “bulk-like” state, either to a red shifted, low intensity and low lifetime state or to a blue shifted, high intensity and high lifetime states. Both forward and backward transitions were observed. We characterized these states in the intensity-lifetime-spectrum parameter space and quantified the transition probabilities between the states. We speculate that these states correspond to slight rearrangements of the fluorophore conformation, possibly in the hydrogen-bonding network. Unlike Atto 647N and Atto 633, no apparent dynamics were observed in the cyanine dye Alexa 647 on timescale of 1 ms to 1000 ms. Our technique allows an unprecedented amount of information to be obtained for single molecules in solution and will be applicable to a variety of systems undergoing biophysical investigations.
Chapter 5

Dissecting pigment architecture of individual allophycocyanin antenna complexes in solution

In this chapter I describe the use of multi-parameter fluorescence detection with the ABEL trap to watch the photodynamics of allophycocyanin (APC) in solution. APC is a photosynthetic pigment-protein complex in cyanobacteria and red algae with fascinating light-harvesting properties. In particular, APC’s well-tuned absorption behavior is a result of proper arrangement of six pigment molecules within a triangular-shaped trimer structure. The optical properties of the individual pigment sites, as well as the organization principles between these sites are not well understood in either the monomeric or the trimeric form of the protein. Utilizing stochastic photo-bleaching of the individual emitters and multi-parameter fluorescence detection (Chapters 3 and 4) in the ABEL trap, the single-molecule measurements presented here provided extensive photophysical characterizations of the individual pigment sites in the protein and revealed previously unknown organization principles of this pigment-protein-complex.
CHAPTER 5. PIGMENT ARCHITECTURE OF ALLOPHYCOCYANIN COMPLEXES

5.1 Introduction

The first step in photosynthesis concerns capturing solar radiation by light harvesting antennae. These are pigment-protein complexes evolutionarily tuned to convert electromagnetic energy to molecular excited state energy and subsequently to transfer the energy to the reaction center to extract chemical energy. Nature uses a number of strategies to effectively cover the broad solar spectrum (Croce et al., 2014). At the molecular level, chemically distinct pigments are bio-synthesized to provide spectral variety. At the antenna level, a pigment’s spectral properties can be tuned by exploiting interactions between a pigment and the amino acid side chains in its nearby protein nano-environment. For example, a major pigment in cyanobacteria, phycocyanobilin (PCB) shifts its absorption spectrum from ~600 nm in the β155 site of C-phycocyanin (CPC) to ~620 nm in the β84 site of the same protein (Debreczeny et al., 1995) and further to ~650 nm when bound to the photosensory protein phytochrome (Hughes et al., 1997).
Protein aggregation can be exploited as another effective mechanism for shaping the optical properties of the antenna complexes. The example we choose to study here is allophycocyanin (APC) (MacColl, 2004), whose optical properties and \textit{in vivo} functions critically depend on the protein’s \textit{trimeric} aggregation state. APC is a key peripheral antenna protein in the phycobilisome (MacColl, 1998), which is the primary light-harvesting apparatus of cyanobacteria and red-algae. APC serves as the terminal energy funnel component between the phycobilisome and the photosystem (Liu et al., 2013a) during productive photosynthesis and, as recently discovered, the primary site for photoprotective quenching in cyanobacteria (Tian et al., 2011). The self-assembly hierarchy of APC starts with two structurally homologous polypeptides (the $\alpha$ and $\beta$ subunits in the monomer). One phycoerythrobilin (PCB) pigment is covalently attached to each subunit via post-translational modification (Scheer et al., 2008). The ligated $\alpha$ and $\beta$ subunits then form a heterodimer (Figure 5.1a), which is the monomeric form of the protein (“monomer”). Once formed, the monomer is very stable (Toole et al., 1998) and further assembles into a trimer, which contains six pigments. The crystal structure of the trimer (Figure 5.1a) shows an equilateral triangular arrangement of the three monomers with the $\alpha$ pigment brought into close proximity (~2 nm) to the $\beta$ pigment of a neighboring monomer (Brejc et al., 1995).

The fascinating feature about APC is that trimer formation induces a large red shift in the optical properties (~30 nm in linear absorption and ~20 nm in fluorescence emission). This spectral shift is critical in defining APC’s biological role in harvesting sunlight. Not only does the shift make APC trimer the only absorber in the 620-650 nm spectral window inside the phycobilisome, helping to fill up the “green gap”, it also optimizes the spectral overlap between the antenna complex and chlorophyll a pigments in the photosystem, facilitating energy migration from the antenna to the reaction center. Intriguingly, in a related biliprotein C-phycocyanin, which binds chemically identical pigments in a near-identical 3D arrangement, trimerization does not produce a large spectral shift. The origin of this bathochromic shift remains elusive. Specifically, it is not known what roles the $\alpha$ and $\beta$ pigments play in generating the spectral shift during the process of trimer assembly.

To elucidate the organization principles of APC and more specifically, the roles of the individual $\alpha$ and $\beta$ pigments before and after self-assembly, we developed a bottom-up, single-molecule approach. First, we study the emissive properties of APC monomers at the
single-molecule level. Individual \( \alpha \) and \( \beta \) pigments are resolved by stochastic photobleaching from an intact monomer and investigated by multi-parameter fluorescence mapping (Figure 5.1b). We then probe the sequential photodegradation of individual APC trimers to reveal the emission characteristics of the pigments in the trimer configuration. In contrast to immobilizing molecules on a surface, as is commonly done in single-molecule studies, we employ an aqueous solution-based device known as the Anti-Brownian ELectrokinetic (ABEL) trap, in which a single molecule can be continuously monitored in solution without immobilization-induced perturbations (Wang et al., 2012; Schlau-Cohen et al., 2014). In this Chapter we focus on the behavior of the monomer. Details regarding the trimer data and modeling will appear in a future publication.

5.2 Results

**Polarization-resolved photodynamics of single light-harvesting complexes in solution using an anti-Brownian trap** The ABEL trap provided a non-perturbative solution environment for probing individual pigment-protein complexes, one molecule at a time (Goldsmith et al., 2010; Wang et al., 2012; Schlau-Cohen et al., 2014). Here, single molecules of APC, when kept near the center region of a microfluidic chamber (Figure 5.1b) by an electrokinetically-actuated feedback mechanism, were investigated in aqueous solution for durations limited by complete photobleaching of the whole complex (~1 second). During this time, each underwent a series of photoinduced dynamical changes, which, due to the lack of ensemble averaging, contains a record of the life-history of the molecule. To obtain detailed information regarding pigment photophysics and organizations within the protein complex, we measured five independent fluorescence parameters from trapped single molecules (Figure 5.1b). Of these, fluorescence brightness probes the combined rates of absorption and radiative decay; emission spectrum reports on electronic and vibronic energy structure of the pigments which could be modulated by the surrounding protein nano-environments; and excited-state lifetime is sensitive to quenching pathways. In this study, we extended the instrument to also measure fluorescence polarization, which is the single-molecule analog of fluorescence anisotropy measurements (Lakowicz, 2006) to probe energy transfer pathways (Chapter 3.8), and photon anti-bunching, which has been employed
to study excited-state processes such as singlet-singlet annihilation (SSA) (Hofkens et al., 2000; Hübner et al., 2003). In particular, brightness, polarization and spectrum were recorded simultaneously, as correlated dynamics among single-molecule variables offer additional insights (Prummer et al., 2004; Börner et al., 2012; Wang et al., 2013; Tinnefeld et al., 2005).

**Monomer photodynamics at the single-pigment resolution** In phosphate buffer with anti-Brownian trapping, monomers were typically observed for $0.1 - 5$ seconds with clear step-wise dynamics in brightness (Figure 5.2). This was in sharp contrast to the previously reported behavior of surface-attached APC monomers, which underwent fast photobleaching that limited single-molecule investigations (Ying et al., 1998). It is possible that the solvent-exposed bilin pigments in the monomer are susceptible to perturbations introduced by the immobilization procedure. Our solution-based measurement scheme avoided those perturbations and directly enabled single-molecule studies of the monomer state of the APC protein.

Monomers were observed to show digital transitions between distinct brightness levels with small changes in concomitant polarization and emission spectrum. Figure 5.2 shows two typical monomers in the trap. For each brightness level, the average values of the polarization and spectrum can be computed for the same time interval. To better characterize the different intermediates along the photodegradation pathway of the monomer, we mapped each observed level to a point in a multi-dimensional parameter space using its measured brightness, polarization and emission spectrum (Figure 5.3a). Densities in the parameter space was computed using the 2D kernal density estimator developed by Botev et al. (2010). Further, correlated brightness and excited-state lifetime dynamics of the monomers were measured and aligned along the brightness axis to provide additional insight for state identification.

We observed four states or clusters in the multi-dimensional parameter space. These states separate mainly by brightness and lifetime, and show small but distinctive differences in spectra and polarizations. Given that a monomer protein covalently binds one $\alpha$ pigment and one $\beta$ pigment, we naturally expect three observable states as a result of sequential photo-degradation: monomer with intact $\alpha+\beta$ pigments, monomer with only the $\alpha$ pigment
Figure 5.2: Example single-molecule traces of APC monomers (left) and trimers (right). For each trapped molecule, time-dependent changes in brightness, emission polarization and spectrum were recorded simultaneously. Grey lines indicate parameter values averaged over the change-point levels (Watkins et al., 2005) identified on the intensity trace. The symbols in the polarization and spectrum panels show the corresponding parameters calculated from every 250 photons.
(from bleaching of the $\beta$ pigment) and monomer with only the $\beta$ pigment (from bleaching of the $\alpha$ pigment). Given that the majority (~60%) of the single molecules entered the trap with an emission brightness around 28 cnts/5ms, a spectral center-of-mass around 647.6 nm and a lifetime of ~1.28 ns, we assigned this most luminous population to be the $\alpha+\beta$ state. Interestingly, this state has a slightly lower polarization (~0.36) compared to the other three states, a clear signature of non-parallel energy transfer between the $\alpha$ and $\beta$ pigments (to be analyzed below). The three low intensity states all have polarization values (~0.40) close to the theoretical maximum of a single dipole emitter in solution (~0.40, considering depolarization effects by the high NA optics and rotational diffusion, Figure 3.19), which suggests that each represents a state with only one pigment.

We next assigned the second brightest cluster (~22 cnts/5ms) to be the $\alpha$ state and the dimmest cluster (~12 cnts/5ms) to be the $\beta$ state. This assignment is based on comparison with previous ensemble spectroscopic studies using biochemically purified $\alpha$ and $\beta$ subunits (Cohen-Bazire et al., 1977; Füglistaller et al., 1987), where it was established that the $\alpha$ subunit is about twice as bright as the $\beta$ subunit for APCs from a variety of strains. Note that based on this assignment, the $\beta$ pigment has a slightly different (red-shifted) emission spectrum (647.8 nm) compared to the $\alpha$ pigment (646.2 nm), which is reasonable given the differences in amino acid composition of the two pigment binding pockets and the sensitivity of PCB’s optical properties to its immediate environment. Comparable dissimilarities between the $\alpha$ and $\beta$ emission spectra were also found in re-natured subunits from both APC (Füglistaller et al., 1987) and a structurally similar protein C-phycocyanin (Debreczeny et al., 1993), albeit in non-physiological buffer conditions. Also based on this assignment, the $\alpha$ pigment has a much longer excited-state lifetime (~1.7 ns) compared to the $\beta$ pigment (~1.1 ns), which is another manifestation of the differences in their local environments.

Interestingly, between the $\alpha$ and $\beta$ populations there exists another cluster. This fourth population, compared with the $\alpha$ pigment, has nearly identical spectral position and high polarization (indicative of a single pigment) but reduced excited-state lifetime. Based on those observations, we identified this state to be the $\alpha$ pigment in a quenched form ($\alpha^Q$). Moreover, this $\alpha^Q$ state is a single-step photoproduct of the intact monomer, as 72% of this state results from direct transitions originated from the $\alpha+\beta$ state. As a result, we attributed
the source of quenching to be the photoproduct(s) of the $\beta$ pigment acting as a non-emissive FRET acceptor.

5.3 Discussion

Pigment-specific spectroscopic properties of the monomer  Using photobleaching to spectroscopically “knock out” the individual pigments on the protein allows us to examine the emissive properties of the individual $\alpha$ and $\beta$ pigment sites. Specifically, by performing analysis on the individual $\alpha$ and $\beta$ clusters defined by the multiparameter fluorescence mapping, we extract the relative brightness, excited-state lifetime, emission polarization, full emission spectrum (Figure 5.3d) and photobleaching rate (at intermediate to high excitation powers, Figure 5.3e) of the two pigments. Our data reveal that in APC monomers, the brighter $\alpha$ pigment has a longer excited-state lifetime, a slightly blue-shifted emission spectrum and enhanced photostability compared to the $\beta$ pigment. We attribute these differences to the nuances in the pigment-protein interactions in the PCB binding pockets. It is well known that the rigidification of the PCB provided by the protein cavity greatly reduces the number of non-radiative pathways in the excited state (Göller et al., 2005), thus extending the excited-state lifetime to ~ns (from <40 ps of free PCB in solution). We thus hypothesize that the photophysical differences revealed here suggest that the relative degree of (remaining) conformational flexibility of the $\beta$ pigment is higher compared to that of the $\alpha$ pigment.

Interestingly, the power dependence of the photobleaching rate for both pigments shows a significant nonlinear component (Figure 5.3e) and this nonlinear dependence is more pronounced for the $\beta$ pigment. This observation suggests that photobleaching of the individual pigments can proceed either through the first electronic excited state ($S_1$) or from a higher excited state ($S_n$) via absorption from $S_1$. Indeed, previous femtosecond pump-probe measurements performed on the $\alpha$ subunit of C-phycocyanin (Riter et al., 1996) and APC monomer (Shiu et al., 2002) both revealed significant components of $S_1$ absorption with favorable spectral overlap to the $S_1$ emission.
Figure 5.3: Multi-parameter mapping of the different emissive states along the photobleaching pathway of APC monomers. (a) State distribution in the brightness-polarization, brightness-spectrum and brightness-lifetime parameter spaces. Identifications of those states are illustrated at the bottom, see text for details. (b) Distribution of the measured emission polarization for the α+β state and the β only state with Gaussian fits. (c) Calculated fluorescence polarization of the α+β state (normalized by the single pigment state) as a function of the angular separation between the α and β pigment dipoles on the monomer. Experimentally measured polarization value is shown as a horizontal gray line, with the 95% confidence intervals indicated as dashed gray lines. The projection of the intercept onto the angular axis gives the estimated angular separation supported by our measurements. (d) Emission spectra of the individual α and β pigment sites. (e) Photobleaching rates of the individual α and β pigment sites. (f) Black: fluorescence anti-bunching signal of the intact monomer (the α+β state) with a fit to Equation 3.39. Gray: fluorescence anti-bunching signal from a dsDNA molecule labeled with two Cy5 dyes.
Coupling between the two pigments on the monomer  We now examine the coupling between the $\alpha$ and $\beta$ pigments on the monomer. Here, given the ~5 nm separation between the two pigments, Forster resonance energy transfer (FRET) is expected to be the dominant coupling mechanism. As a first step, it is helpful to compare the brightness of the intact monomer ($\alpha+\beta$ state) with the sum of the site brightness ($\alpha$ state + $\beta$ state). As revealed by our multi-parameter fluorescence mapping (Figure 5.3a), the monomer is only 83% as bright as the sum of its individual constituent pigments. This observation can be understood by asymmetric energy transfer between the two pigments that results in a shift of the excited-state probability distribution towards the dimmer $\beta$ site. To further quantify the degree of asymmetry in the bidirectional energy flow of the monomer, we constructed a quantitative FRET model (see the method section below) using parameters acquired through our site-specific multi-parameter fluorescence mapping. If we assume equal absorptivity of the two pigments, we calculated the $\alpha$-to-$\beta$ transfer rate ($k_{\alpha\rightarrow\beta}$) to be 1.8 times faster than the $\beta$-to-$\alpha$ transfer rate ($k_{\beta\rightarrow\alpha}$). In other words, the lower energy pigment $\beta$ acts as an energy sink that traps about 65% of the excitations. If we assume that the two pigments are geometrically positioned as revealed by the monomer subunit in the trimer crystal structure (i.e. a distance of 4.96 nm and a $\kappa^2$ value of 2.7), we get $k_{\alpha\rightarrow\beta} = 5.4 \text{ ns}^{-1}$ and $k_{\beta\rightarrow\alpha} = 3.0 \text{ ns}^{-1}$ (Figure 5.4). Given these parameters, the emissive properties (brightness, lifetime and emission spectrum) of the intact monomer can be derived from the individual $\alpha$ and $\beta$ pigments. Those calculated properties of the $\alpha+\beta$ state all compare favorably with experiments. Additionally, our quantitative model directly predicts an equilibration time of ~120 ps between the two pigments, which is similar to the experimentally measured value of ~140 ps (Beck et al., 1992).

Using this quantitative FRET model, we can further estimate the angular separation between the $\alpha$ and $\beta$ pigments on the monomer from the measured fluorescence polarization. For the $\alpha+\beta$ state, the degree of emission polarization is lower (Figure 5.3b) than that of the single-pigment states, presumably due to bi-directional energy transfer between two non-parallel pigment dipoles. The precise value of the polarization depends on a) the angular separation between the two pigments and b) the percentage of excitation-emission events that start and end on different sites. In Figure 5.3b, we plot the expected degree of emission polarization of the monomer (normalized to that of a single pigment), predicted by the
FRET model, as a function of the angular separation between the \( \alpha \) and \( \beta \) pigments. Using this approach, the measured emission polarization corresponds to \( 29.5 \pm 1.5^\circ \) between the two pigment dipoles. Although this value is very different from an early estimation based on the crystal structure of the trimer (Brejc et al. (1995), \( \sim 0^\circ \)), we believe that two factors might explain the discrepancy. First, the previous estimation of the pigment dipole orientation based on fitting a line to the conjugated portion of the pigment is of limited accuracy. Second, the arrangement of the two pigments in the monomer (not yet crystallized) is likely different from that in the trimer.

To further investigate if the excited states of the two pigments are coupled via FRET, we conducted photon-antibunching measurements on APC monomers. Photon-antibunching measures the coincidence rate of detecting two photons at the same time and is particular sensitive to singlet-singlet-annihilation (SSA) processes (FRET between two excited states followed by rapid internal relaxation, \( S_{1}^{a} + S_{1}^{b} \rightarrow S_{0}^{a} + S_{n}^{b} \rightarrow S_{0}^{a} + S_{1}^{b} \)). This is because SSA effectively lowers the coincidence rate at zero delay when the system contains two emitters (Tinnefeld et al., 2002; Hofkens et al., 2003; Hübner et al., 2003). Here, we constructed the antibunching curve for the intact monomer (\( \alpha + \beta \) state, Figure 5.3f), as defined on the multi-parameter fluorescence map (Figure 5.3a, by correlating the emission times of all photon pairs for this state). Intriguingly, we found a near-zero coincident probability near zero delay (when properly accounting for the bandwidth of the detector, Section 3.8.4). This observation suggests efficient SSA between the \( \alpha \) and \( \beta \) pigments, when both are in the excited state. Note that the presence of efficient SSA on the monomer is directly supported by the strong red-shifted ESA signal in previous pump-probe measurements (Riter et al., 1996; Shiu et al., 2002) (the spectral overlap between \( S_{1} \) absorption and \( S_{1} \) emission is almost perfect as revealed by Riter et al. (1996)) and is also consistent with the non-linear bleaching behavior discussed previously. Similarly, efficient SSA processes were also found in the tetrameric fluorescent protein dsRed using photon anti-bunching measurements (Sánchez-Mosteiro et al., 2004).

The \( \alpha \) pigment is actively quenched in the monomer Because the \( \alpha \)-to-\( \beta \) energy transfer pathway is more efficient (Figure 5.4), the \( \alpha \) pigment in the intact monomer is effectively quenched. In other words, the probability of the \( \alpha \) pigment being in the excited state
is reduced due to (more) rapid energy transfer to the β pigment. We calculated that the emission rate of the α site on the monomer is reduced by ~43% in the presence of the β pigment. Moreover, our measurements revealed a previously unknown state (α₉) of the APC monomer, in which only the quenched emission of the α pigment is observed. We hypothesize that this quenched α state is produced by the β pigment entering into a non-emissive but still absorptive state, likely a photoproduct of the PCB degradation. We noticed that this α₉ state is fairly common, as quenching happens in ~55% of the cases where β emission terminates. In this quenched state, the brightness of the α pigment is reduced by ~25%.

By comparing the measured lifetime between the quenched and unquenched forms of the α pigment, we extracted a quenching rate of 0.2 ns⁻¹. Note that this number is much smaller than \( k_{α→β} = 5.4 \text{ ns}^{-1} \) in an intact monomer, which could be due to reduced oscillator strength and/or shifted absorption spectrum of the β photoproduct. The finding that the α pigment is actively quenched in an intact monomer, even after partial photodegradation of the β pigment is interesting and could have functional implications. On the other hand, our data did not show a quenched β state, suggesting that the photoproduct of the α pigment does not act as a quencher.
5.4 Conclusions

We have presented a detailed spectroscopic study of the APC monomer, an important photosynthetic antenna protein, at the single-molecule level. Here, the ABEL trap offers the unique capability to measure individual copies of these intrinsically fluorescent pigment-protein complexes in a biologically relevant solution environment. This proved to be critical in the case of APC monomers, as previous single-molecule investigations on these proteins on a surface achieved limited success (Ying et al., 1998).

For the very first time, the combination of single-pigment photobleaching and multi-parameter fluorescence detection allowed us to take a bottom-up approach in understanding the excited-state properties of the pigment-binding protein. Specifically, the methodology presented here enabled a detailed characterization of the individual pigment sites on the protein and then used these as the basis to unravel the coupling between the pigments.

We found that the two PCB pigments of APC monomer have distinct emission properties and photostabilities, likely due to differences in conformation flexibility in the binding pockets. The emission of the full monomer is a result of asymmetric, bi-directional energy transfer between the two pigments. Most interestingly, we found that the $\alpha$ pigment is quenched in the protein by the $\beta$ pigment acting as an emissive trap and continues to be quenched by the $\beta$ pigment’s photoproducts.

5.5 Methods

5.5.1 FRET model of the monomer

Energy transfer rates were calculated using the standard FRET equation (Parson, 2007)

$$k_{D \rightarrow A} = \frac{\Phi_D k_D^{10} \kappa^2}{R_{DA}^6} \left( \frac{9000 \ln 10}{128 \pi^2 N_A n^4} \right) \int \frac{F_D(\lambda) e_A(\lambda) \lambda^2 d\lambda}{\int F_D(\lambda) \lambda^{-2} d\lambda}$$

(5.1)

where $\Phi_D$ and $k_D^{10}$ are the quantum yield and the excited state decay rate of the donor, $\kappa^2$ is the orientation factor, $R_{DA}$ is the distance between the donor and acceptor dipoles.
CHAPTER 5. PIGMENT ARCHITECTURE OF ALLOPHYCOCYANIN COMPLEXES

and the third factor calculates the spectral overlap given the spectral shapes of the donor fluorescence ($F_D(\lambda)$) and acceptor absorption ($\varepsilon_A(\lambda)$).

To calculate the FRET rates on an APC monomer, the excited state decay rates ($k_{10}^D$), and the donor fluorescence spectral shapes ($F_D(\lambda)$) of each pigment type were directly measured using “state purification” on the multi-dimensional parameter space (i.e. averaging the parameter of interest over all the data points clustered as a particular population in Figure 5.3a). The absorption profiles of the individual pigments were assumed to be identical and derived from the bulk absorption measurement. The quantum yields ($\Phi_D$) were measured by comparing the brightness of the respective pigment with that of a free dye with a known quantum yield (Atto 647N, 0.65).

The spectroscopic properties of the monomer can be calculated from the individual sites and their FRET coupling as the following. First, we calculate the probability ($p(t)$) of finding the quanta of excitation energy in each of the system’s $N$ sites ($N = 2$ in this case) by solving the master equation,

$$\frac{dp(t)}{dt} = Mp(t) \quad (5.2)$$

where $p(t)$ is a $2 \times 1$ vector that specifies the probability of finding the excitation quanta in each site at time $t$, $M$ is the transition matrix with off-diagonal elements denoting FRET

$$M_{ij} = k_{j \rightarrow i} \quad (i \neq j) \quad (5.3)$$

and diagonal elements

$$M_{ii} = -\left(k_{10}^i + \sum_{j,i \neq j} k_{i \rightarrow j}\right) \quad (5.4)$$

Equation 5.2 was solved by consecutively initializing the system at the individual sites (i.e. first using $p^1(0) = [1,0]^T$ and then using $p^2(0) = [0,1]^T$). Each solution set (e.g. $p^1(t)$) represents the probability of finding the excitation between the two sites on the monomer as a function of time after the excitation event. Note that this treatment assumes that at any given time, only one quanta of excitation can be created in the system, which is reasonable considering an estimated double excitation probability of only ~1% under the single-molecule experimental conditions.
CHAPTER 5. PIGMENT ARCHITECTURE OF ALLOPHYCOCYANIN COMPLEXES

Having solved for \( \mathbf{p}^s(t) \), the brightness of the monomer can be calculated by

\[
B = \sum_{i=1}^{2} B_i \tag{5.5}
\]

where \( B_i \) is the site brightness calculated by

\[
B_i = \sum_{j=1}^{2} A_j P^j_i \tag{5.6}
\]

where \( A_j \) is the absorption probability of site \( j \) and \( P^j_i \) is the probability of emitting from site \( i \) when the initial excitation happened at site \( j \), give by

\[
P^j_i = \int_{0}^{+\infty} p^j_i(t) k^{10}_{i} \Phi_i dt \tag{5.7}
\]

where \( p^j_i(t) \) is the \( i \)-th component of the vector \( \mathbf{p}^j(t) \), \( k^{10}_{i} \) and \( \Phi_i \) are the excited state decay rate and quantum yield of site \( i \).

Similarly, fluorescence anisotropy of the monomer can be calculated by considering all possible site combinations that a quanta of excitation energy enters and leaves (via fluorescence emission) the system (Lounis et al., 2001)

\[
r = \sum_{j} A_j \left( \frac{\sum r_{ji} P^j_i}{\sum P^j_i} \right) / \sum A_j \tag{5.8}
\]

where \( r_{ji} \) is the fundamental anisotropy that links site \( i \) and \( j \). For \( i = j \), \( r_{ii} = 0.4 \), for \( i \neq j \), \( r_{ij} \) is determined by the angle between the two pigment dipoles and is varied to match experimental observation (Figure 5.3c). Fluorescence polarization is subsequently calculated by using Equation 3.17.

Finally, the monomer emission spectrum is simply the average of the individual sites, weighted by their relative brightness (Equation 5.6). Time-resolved fluorescence decay can be constructed by averaging the excited-state probability decay of the two emissive sites in
the system, weighted by their relative brightness

\[ g(t) = \sum_i \Phi_i k_i^{10} \left( \frac{\sum_j A_j p_j^i(t)}{\sum_j A_j} \right) \]  

(5.9)
Chapter 6

Beyond fluorescence parameters: single-molecule transport properties

In the previous two chapters, I have demonstrated how the ABEL trap enables synchronous dynamics of multiple fluorescence parameters to be recorded on single molecules in solution. These measurements reveal the internal states of a nano-emitter and the interconversion between these states on a 100 ms timescale and longer. However, up to this point all the information is indirectly extracted from fluorescence variables and is limited to reporting properties of the probe. To sense dynamical changes of a biomolecule, the probe often needs to be specially engineered (Oijen, 2011; Ulbrich et al., 2007) so that its emission properties are modulated by the biomolecule (e.g. FRET requires two probes to be placed at specific locations on a biomolecule so that distance changes are reflected in the relative brightness of the donor and acceptor). On the other hand, to broaden the utility of single-molecule measurements, it is much more desirable to directly sense the intrinsic physical properties (e.g. size, shape, charge, etc) of the single biomolecules themselves, but this has been an extremely challenging experimental task for biophysical research.

In this chapter, I demonstrate how the ABEL trap can estimate parameters related to size and charge of single biomolecules in solution. Size and charge are fundamental physical properties of nano-objects that are orthogonal to fluorescence emission. I will show that they can be used as an intrinsic contrast mechanism to identify molecules. In addition,
interactions between biomolecules are generally accompanied by size and charge modifications at the nanometer scale. The ability to sense dynamics in these physical parameters thus provides a powerful means to study biomolecular interactions at the single-molecule level without the need to use engineered fluorescence contrast. The work presented in this chapter was published in Wang et al. (2014b).

To sense size and charge, we can simply analyze how a molecule moves in the ABEL trap. This relies on the fact that the (residual) motion of a trapped molecule is governed by a combination of diffusion, which depends on the size, and electrokinetic drift, which is sensitive to charge. From fundamental physics (Equation 1.1 and Equation 1.7), alteration in a molecule’s size and charge (e.g. from a change in association state) can be detected from modulations in transport properties (i.e. an increase in size slows down diffusion; an increase in charge enhances mobility). We aim to precisely determine the (time-dependent) single-molecule diffusion coefficient \( D \) and electrokinetic mobility \( \mu \) in solution\(^1\).

### 6.1 Real-time estimation of single-molecule diffusion coefficient \( D \) and electrokinetic mobility \( \mu \)

The theoretical framework and a computational procedure for estimating the transport coefficients \( D \) and \( \mu \) are discussed in Chapter 2. Here, I briefly summarize the problem and its algorithmic solution.

To extract transport coefficients of single trapped molecules, we utilized the fact that a “trapped” molecule still exhibits residual motion within the ~4\( \mu \)m \( \times \) 4\( \mu \)m region near trap center. This motion reflected the limited bandwidth of the feedback system and can be tracked to infer \( D \) and \( \mu \). To track the fast motion of a nanometer-sized biomolecule \( (D \sim 100\mu m^2 s^{-1}) \) in aqueous solution, we utilized beam scanning (Section 2.2) to circumvent the speed limitations of imaging-based tracking schemes. Briefly, an excitation beam rapidly (600 ns per point) surveyed the trapping region via a deterministic pattern on a 32-point grid; each detected photon was mapped to a corresponding beam position based on

\(^1\)For interpretation of mobility changes as charge changes, the trap is operated with electro-osmosis suppressed.
Figure 6.1: Real-time estimation of single-molecule diffusion coefficient and electrokinetic mobility. The ABEL trap achieved ~16 seconds observation of a single 10nt-ssDNA in solution by photon-by-photon electrokinetic feedback that approximately cancels Brownian motion. To estimate the diffusion coefficient ($D$) and electrokinetic mobility ($\mu$) of the trapped molecule, a series of photon-stamped position estimates (shaded circles: beam scan positions, black dots: real positions of molecule, orange dots: photon-stamped position estimates) and corresponding feedback voltages were subjected to a statistical learning procedure (EM). Estimation of $D$ and $\mu$ was conducted in real-time on windows (green slice) containing ~5,000 photons each. Reprinted with permission from Wang et al. (2014b) Copyright Nature Publishing Group.
its arrival time (Figure 6.1). These photon-stamped beam positions were recorded as estimates of molecule positions. In this way, we followed the motion of a single biomolecule in solution with photon-by-photon time resolution (~20 µs for a 50 kHz count rate).

\( D \) and \( \mu \) were estimated by fitting the photon-by-photon tracking trajectories to a simple Langevin-type model (Section 2.2.3, Figure 6.1). To do so, we developed an estimator based on Bayesian machine learning. Specifically, a recursive, distributed approximation scheme was employed to treat the non-Gaussian measurement noise (arising from background photons) with unbiased statistics (Section 2.4), and the maximum-likelihood estimates of \( D \) and \( \mu \) were obtained by an Expectation-Maximization (EM) search (Section 2.5.2). We verified by simulation that the algorithm properly converges (Figure 2.13) and achieves unbiased (Figure 2.14) and precise (Figure 2.15) estimation of transport coefficients over a large parameter space and different signal-to-background conditions. Moreover, the algorithm runs in real-time (Figure 2.16) under typical count rate conditions for single-molecule detection.

Real-time calculation of transport parameters was achieved by chopping the incoming data stream into 5,000-photon windows and running the C-coded EM algorithm with two EP-MP passes (Figure 2.13b) on each window during data acquisition. For a photon detection rate of 25 kHz (typical for single-molecule detection), real-time calculation of the motion parameters achieved a time resolution of ~200 ms. Alternatively, \( D \) and \( \mu \) can be (re-)estimated with higher accuracy (i.e. six EP-MP passes, Figure 2.13b) and time resolution (usually fixed 50 ms time windows) during post-processing of the data set.

### 6.2 Validation on different single molecules

We first validated our method experimentally by extracting transport coefficients of single fluorophores (Atto 647N, <1 nm), 10 nucleotide (10nt) and 18nt single-stranded DNA (ssDNA, ~2 nm), natively fluorescent 104 kDa photosynthetic antenna proteins (Allophyco-cyanin, the subject of Chapter 5, ~10 nm disk) and fluorescent polystyrene beads (~26 nm). Particularly revealing for molecular identification is a scatter-plot display of the \( D - \mu \) values for each measured single molecule (Figure 6.2, Table 6.1). For the various samples studied, the \( D - \mu \) estimates for each molecule form clusters characteristic of each
Figure 6.2: $D - \mu$ mapping of different single molecules. Each symbol represents parameters extracted from one molecule, averaged over its entire residence in the trap. The slight correlation between $D$ and $\mu$ is characteristic of the learning algorithm (Figure 2.14). Shaded region: 10nt-ssDNA with glycerol concentrations of 0, 10, 15, 20, 30 percent (direction of arrow) in water. Relative viscosities of the water-glycerol mixtures were extracted that agree with literature. Direct proportionality (dashed line: fit of center-of-mass values of clusters) between $D$ and $\mu$ confirmed the Einstein–Smoluchowski relation. Reprinted with permission from Wang et al. (2014b) Copyright Nature Publishing Group.
molecule type and the sample-averaged diffusion coefficients compare favorably with literature (Table 6.1).
<table>
<thead>
<tr>
<th>Sample</th>
<th>MW (Da)</th>
<th>Buffer</th>
<th>Electro-kinetic mechanism</th>
<th># of molecules measured</th>
<th>Average # of photons per molecule</th>
<th>Average SBR</th>
<th>Average $D$ ($\mu m^2 s^{-1}$)</th>
<th>Literature $D$ ($\mu m^2 s^{-1}$) and method</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atto647N-NHS</td>
<td>843</td>
<td>$1 \times$ PBS + 10% glycerol</td>
<td>EO + EP</td>
<td>49</td>
<td>58400</td>
<td>5.0</td>
<td>281±1</td>
<td>285 (2fFCS&lt;sup&gt;a·†&lt;/sup&gt;)</td>
<td>-1.5%</td>
</tr>
<tr>
<td>10nt-Alexa647</td>
<td>4047</td>
<td>20 mM HEPES</td>
<td>EP</td>
<td>390</td>
<td>52900</td>
<td>3.6</td>
<td>162.0±0.2</td>
<td>205 (CE&lt;sup&gt;b&lt;/sup&gt;) 173 (FCS&lt;sup&gt;c&lt;/sup&gt;)</td>
<td>-20%</td>
</tr>
<tr>
<td>18nt-Alexa647</td>
<td>6628</td>
<td>20 mM HEPES + 5% glycerol</td>
<td>EP</td>
<td>169</td>
<td>46300</td>
<td>3.6</td>
<td>113.6±0.5</td>
<td>119 (CE&lt;sup&gt;d·‡&lt;/sup&gt;) 129 (CE&lt;sup&gt;e·◦&lt;/sup&gt;)</td>
<td>-4%</td>
</tr>
<tr>
<td>APC trimers</td>
<td>104k</td>
<td>$1 \times$ PBS</td>
<td>EO</td>
<td>398</td>
<td>13900</td>
<td>13±6 (s.d.)</td>
<td>46.3±0.2</td>
<td>49 (HYDROPRO&lt;sup&gt;f&lt;/sup&gt;) 41 (FCS&lt;sup&gt;g&lt;/sup&gt;) 48.6 (DLS&lt;sup&gt;h&lt;/sup&gt;)</td>
<td>-5%</td>
</tr>
<tr>
<td>26 nm polystyrene beads</td>
<td>&gt;1M</td>
<td>Nanopure water</td>
<td>EO + EP</td>
<td>37</td>
<td>59100</td>
<td>27±17 (s.d.)</td>
<td>13.4±0.6</td>
<td>18 (Stokes-Einstein)</td>
<td>-25%</td>
</tr>
</tbody>
</table>

Errors indicate s.e.m. unless otherwise specified


<sup>†</sup>: Atto6655-maleimide, corrected for 10% glycerol used in this study

<sup>‡</sup>: Corrected for 5% glycerol used in this study, no fluorescent label in original study

<sup>◦</sup>: 20 base pair ssDNA, corrected for 5% glycerol used in this study, no fluorescent label in original study

<sup>a</sup> (Müller et al., 2008)<sup>b</sup> (Stellwagen et al., 2003)<sup>c</sup> (Doose et al., 2007)<sup>d</sup> (Nkodo et al., 2001)<sup>e</sup> (Stellwagen et al., 2002)<sup>f</sup> (Ortega et al., 2011)<sup>g</sup> (Fatin-Rouge et al., 2003)<sup>h</sup> (MacColl et al., 2003)

Table 6.1: Measurement statistics of Figure 6.2.
As a separate verification of the method, we conducted a series of glycerol titration experiments on 10nt ssDNA (Figure 6.2 shaded box). The acquired diffusion coefficient and mobility at different solvent viscosities can be well-fitted by a line of direct proportionality. This result confirms the Einstein–Smoluchowski relation, which reveals that the viscous drag that slows down Brownian motion also impedes directed (drift) transport (the same $\eta$ term appears in both Equation 1.1 and Equation 1.7). Relative viscosities extracted from these measurements (data not shown) also agree with literature (Segur et al., 1951).

### 6.3 Resolving a three-component mixture molecule-by-molecule

We next demonstrated identification of different molecular populations in solution using measured single-molecule transport properties. A mixture of 10nt-ssDNA and 18nt-ssDNA was loaded into the trap and molecule-by-molecule analysis of transport coefficients was performed (in a deoxygenated buffer containing 10 mM HEPES, 3 mM Trolox and 10%...
glycerol). A representative data trace is shown in Figure 6.3a. Importantly, the molecules are indistinguishable using the intensity data (upper panel) but can be well-resolved by their diffusion coefficients (lower panel). This example clearly illustrates the additional information revealed by transport parameters. To be sure there is no confusion, no molecule was in the trap during the ~6-8 s interval, and one molecule exchanged for another at ~13 s.

When all the single molecules measured in the mixture were plotted in the $D - \mu$ parameter space, three distinct populations can be resolved (Figure 6.3b). Now, to quantify the exact composition of the mixture, we applied a Gaussian-mixture-model (GMM) classifier (Murphy, 2012) to the problem, described as follows. First we fitted the scatter-plot data to a three-component GMM:

$$ p(x_i|\theta) = \sum_{k=1}^{3} \pi_k N(x_i|m_k, \Sigma_k) $$  \hspace{1cm} (6.1)

where $x_i$ is a single-molecule data point in $D - \mu$ space, $\theta$ represents all model parameters, and $\pi_k$, $m_k$ and $\Sigma_k$ are the mixing weight, mean and covariance matrix of the $k$-th Gaussian component, with $N$ denoting a Normal distribution. After the model is constructed, classification of each data point can be performed by first evaluating the probability of $x_i$ generated by mixture $k$ by Bayes’ rule,

$$ p(z_i = k|x_i, \theta) = \frac{p(z_i = k|\theta) p(x_i|z_i = k, \theta)}{\sum_{k'}^{3} p(z_i = k'|\theta) p(x_i|z_i = k', \theta)} $$  \hspace{1cm} (6.2)

and selecting the cluster that gives the highest probability. The colors in Figure 6.3b illustrate the final classification result for this data set.

The species with only a few molecules (magenta) is identified to be the free dye (Alexa 647) left from purification, because of the high diffusion coefficient associated with this population. Among the two densely packed clusters, the species with a lower diffusion coefficient and higher mobility (green) is identified to be the 18nt-ssDNA, since the longer strand is expected to have a larger hydrodynamic radius and a higher magnitude of the zeta-potential (more negatively charged, electroosmosis is suppressed in this experiment using the PVP polymer). The cluster colored in red is identified to be the 10nt-ssDNA.
Quantitatively, the mixture contains 51% 18nt-ssDNA, 45% 10nt-ssDNA and 4% free dye, as resolved by the GMM classifier.

It is worth noting that once the mixture model is constructed, we can perform online identification in a subsequent experiment by evaluating Equation 6.2 in real-time. This capability raises the prospect of conducting real-time single-molecule filtering: one can selectively trap and probe certain molecular species while quickly releasing unwanted molecules.

6.4 Application: following the dissociation pathway of Allophycocyanin in solution

Having demonstrated the capability to resolve a multi-component molecular mixture using single-molecule transport coefficients, we applied the method to follow the dissociation process of trimeric metastable protein allophycocyanin (APC). APC dissociation is a critical process in understanding the stability of the photosynthetic apparatus of cyanobacteria but so far could only be studied in non-physiological conditions (with destabilization agents) (Huang et al., 1987). Here, to visualize dissociation of wild-type (wt) APC in an aqueous environment (PBS buffer), we measured transport coefficients protein-by-protein (Figure 6.4b and c) and plotted distributions of \( D - \mu \) obtained at different time points after dilution to single-molecule concentration (Figure 6.4a). Chemically cross-linked trimers (XL), incapable of dissociation, were used as a control. In the case of XL-APC, we observed one population with \( \langle D \rangle = 46.3 \pm 0.2 \mu m^2 s^{-1} \) (mean ± s.e.m) and \( \langle \mu \rangle = 263 \pm 1 \mu m s^{-1} V^{-1} \). In sharp contrast, wt-APC, measured ~1 hour after dilution, partitioned into two groups. One group exhibited transport coefficients similar to the XL sample (\( \langle D \rangle = 50 \pm 0.5 \mu m^2 s^{-1} \) and \( \langle \mu \rangle = 269 \pm 1 \mu m s^{-1} V^{-1} \)), which we identified as APC trimers not yet dissociated. The other group had a higher diffusion coefficient (\( \langle D \rangle = 68.5 \pm 0.5 \mu m^2 s^{-1} \)) and a slightly lower mobility (\( \langle \mu \rangle = 269 \pm 1 \mu m s^{-1} V^{-1} \)), assigned to be dissociated monomers. Such an assignment is mainly based on the expected decrease in hydrodynamic radius and an altered electric zeta potential from newly solvent-exposed residues on the protein surface following the trimer-to-monomer transformation.
Figure 6.4: Resolving monomers and trimers of APC in solution along the dissociation pathway. (a) Distribution of monomers and trimers at different time points after dilution, resolved by molecule-by-molecule $D - \mu$ mapping. Black dots show transport coefficients extracted from each trapped protein (number of molecules measured: N = 499, 500, 500, 384 from left to right). Continuous densities are visualized by kernel density estimation. White schematics show species assignments. (b) Four example single molecules from the cross-linked (XL) APC experiment with fluorescence intensity and measured transport coefficients (~500 photons per estimation, offline). (c) Four example single molecules from the wide type APC experiment with fluorescence intensity and measured transport coefficients (~500 photons per estimation, offline). Red dotted lines: single-molecule averages of the transport coefficients. Reprinted with permission from Wang et al. (2014b) Copyright Nature Publishing Group.
Confirmation of this identification comes from the observation that the molecules in this second group collectively showed three intensity levels (Figure 6.5c), which is expected from monomeric APC as it carries two chromophores with unequal brightness (Cohen-Bazire et al., 1977). Repeating the experiment at ~3 hours after dilution revealed the same two populations although the monomer population grew in relative abundance. At ~15 hours, dissociation into monomers was complete. Further, we used the GMM classifier to quantify the relative monomer composition of the mixture at different time points (e.g. Figure 6.5a). At 1, 3 and 15 hours after dilution, the percentages of monomer proteins are 58%, 73% and 100%, respectively. We thus can roughly estimate a dissociation rate of $1 \times 10^{-4} \text{s}^{-1}$. Not surprisingly, this number is about 2,000 times slower than stop-flow measurements using destabilizing agents (Huang et al., 1987). Interestingly, the XL sample contained 9% monomers, suggesting limited efficiency of the crosslinking reaction. Finally, although it would be interesting to observe dimers, we did not find such a population, which means the stability of a dimer if present is shorter than our observation time (~50 ms).

6.5 Application: binding kinetics of short DNA strands

Finally, we used transport coefficients to sense hybridization and melting dynamics of single DNA molecules. The experiment involved trapping a single fluorescently-labeled ssDNA (10nt, 0.75 pM concentration) molecule in the presence of a high concentration (~µM) of its complementary strand (10-base complementary region with a 14-base overhang, ‘24nt-10comp’). Note that the complementary strand is not labeled and thus cannot be detected in a fluorescence microscope under normal conditions.

When a single ssDNA strand is trapped in the presence of its complementary strand, we monitored its diffusion coefficient and mobility in real-time (Figure 6.6b-e). Remarkably, while the intensity channel displayed shot-noise-limited stable emission, we observed digital switching in both measured diffusion coefficient and mobility. More interestingly, changes in $D$ and $\mu$ happened simultaneously between two levels in an anti-correlated fashion (Figure 6.7c and Figure 6.8f). In sharp contrast, 10nt-ssDNA alone showed only one state with high $D$ and low $\mu$ values (Figure 6.6a). We thus identified the new state (low
Trimers show a complicated intensity structure, due to the presence of a large number of interaction pathways between the six chromophores (Chapter 5). Monomers, on the other hand, show a much simpler intensity structure that can be well described by a sum of three Gaussian distributions (solid line). This observation validates the transport-based classification result.
Figure 6.6: Real-time sensing of DNA binding/unbinding using fluctuating motion parameters.
CHAPTER 6. SINGLE-MOLECULE TRANSPORT PROPERTIES

$D$, high $\mu$) that appeared with complementary strand present to be the transiently formed double-stranded DNA (dsDNA). This assignment is consistent with the expanded hydrodynamic radius and increased negative charge density following duplex formation.

As shown in Figure 6.6a-e, when the concentration of the complementary strand went up, the single DNA molecule spent more time in the duplex state and the kinetics accelerated. To analyze the data quantitatively, we used the 50 ms binned $D/\mu$ estimates acquired during post-processing. First, a binding curve can be constructed by calculating the percentage time in dsDNA as a function of complementary strand concentration (Figure 6.7b). Fitting the binding curve yielded a dissociation constant of $K_d = 2.5 \pm 0.2 \mu\text{M}$ (0.95 C.I.) that agrees well with free energy calculations (Markham et al., 2005) for this particular pair ($2.0 \mu\text{M}$).

Such a single-molecule perspective on the hybridization reaction yields unique kinetic information. We used hidden Markov modeling (HMM) to uncover the most likely state evolution (Figure 6.7a) and the melting and hybridization rates (Figure 6.7d-e). We estimated a hybridization rate constant of $k_{on} = (2.99 \pm 0.12) \times 10^5 \text{M}^{-1} \text{s}^{-1}$ (0.95 C.I.) and a melting rate of $k_{off} = (0.70 \pm 0.05) \text{s}^{-1}$ from the HMM analysis. Alternatively, similar rates can be extracted by analyzing the dwell times or the correlation functions of the fluctuating transport coefficients (Figure 6.8). These values are consistent with previous bulk and single-molecule measurements (Henry et al., 1999; Jungmann et al., 2010).

We subsequently studied the destabilizing effect of a single mismatch at the 3’ end of the ssDNA (‘24nt-9comp’, Figure 6.7a and Figure 6.6f-g). We observed similar anti-correlated switching of $D$ and $\mu$ between the same two states (compared with the fully-matched partner strand). The dynamics, on the other hand, were much faster (Figure 6.7c). This observation indicates that the single terminal mismatch did not appreciably affect the hydrodynamic radius or the charge density of the formed duplex, but substantially altered the energetics of the interaction. Quantitatively, from a detailed analysis of the kinetic parameters, we concluded that the terminal mismatch reduced the binding affinity by a factor of 5.1, which could be mostly accounted for by a corresponding 4.7-fold increase in melting rate (Figure 6.7d-e). These findings are consistent with single-molecule binding assays that used engineered fluorescence perturbation (Cisse et al., 2012; Jungmann et al., 2010).
Figure 6.7: DNA hybridization kinetics visualized by dynamic changes in transport coefficients in solution. (a) Top: reaction scheme. Example traces showing intensity (counts per millisecond) and estimated transport coefficients (post-processed every 50 ms) of single trapped 10nt-ssDNA molecules in the presence of its binding partner. Red: 24nt-10comp, 2 µM. Blue: 24nt-9comp, 6.6 µM. Black lines: states identified by HMM. (b) Equilibrium analysis of binding. For 24nt-10comp, \( \langle N \rangle = 72 \) (min: 36, max: 110) molecules were measured per concentration; for 24nt-9comp, \( \langle N \rangle = 55 \) (min: 37, max: 125) molecules were measured per concentration. Average measurement time per molecule was \( \sim 10 \) s. Color maps (top and bottom): density distributions (50 ms binned transport coefficients) in a \( D - \mu \) parameter space, at noted concentrations of the complementary strand. Middle: extracted duplex fractions and fit to a bimolecular binding model. (c) Cross-correlation of \( D \) and \( \mu \), calculated from the cases in a, and fitted (dashed lines) with a single exponential rise. (d) Melting and (e) hybridization rates extracted by HMM analysis. Reprinted with permission from Wang et al. (2014b) Copyright Nature Publishing Group.
CHAPTER 6. SINGLE-MOLECULE TRANSPORT PROPERTIES

Figure 6.8: Extracting binding kinetics from dwell-time analysis and correlation functions of the fluctuating transport coefficients. (a-e) Dwell-time analysis, (f-g) correlation analysis. (a-c) Representative on and off time histograms for different binding conditions, using the states identified by the HMM analysis. Here “on time” is defined as the time in the high $D$, low $\mu$ state (ssDNA) before transition to the low $D$, high $\mu$ state (dsDNA), “off time” is defined as the time in the low $D$, high $\mu$ state (dsDNA) before transition to the high $D$, low $\mu$ state (ssDNA). Both the on and off times can be well-fitted by single exponential decays, suggesting that both binding and unbinding are single-step processes within our experimental time resolution. (d-e) Rates extracted at different complementary strand concentrations. For 24nt-10comp (red), extracted parameters are $k_{on} = (3.34 \pm 0.25) \times 10^5$ M$^{-1}$ s$^{-1}$ and $k_{off} = (0.70 \pm 0.06)$ s$^{-1}$. For 24-9comp (blue), extracted parameters are $k_{on} = (2.55 \pm 0.28) \times 10^5$ M$^{-1}$ s$^{-1}$ and $k_{off} = (3.47 \pm 0.33)$ s$^{-1}$. (f) Cross-correlation between $D$ and $\mu$, defined as $\text{corr}^{D,\mu}(\tau) = \langle D(t) - \langle D \rangle \rangle \langle \mu(t+\tau) - \langle \mu \rangle \rangle / \langle D \rangle \langle \mu \rangle$, calculated at noted concentrations of 24nt-10comp. (g) Extracting kinetic parameters from the correlation analysis. The total rate is the sum of a forward (hybridization) rate, which scales linearly with concentration, and a backward (melting) rate ($k_{all} = c \times k_{on} + k_{off}$). The total rate ($k_{all}$) can be extracted from the cross-correlation by fitting it with a single rising exponential. $k_{all}$ at different concentrations are fitted to a linear function. The slope of the line gives the hybridization rate constant ($k_{on}$) and the y-intercept gives the melting rate constant ($k_{off}$). For 24nt-10comp (red), extracted parameters are $k_{on} = (3.06 \pm 0.75) \times 10^5$ M$^{-1}$ s$^{-1}$ and $k_{off} = (0.60 \pm 0.24)$ s$^{-1}$. For 24-9comp (blue), extracted parameters are $k_{on} = (3.0 \pm 2.2) \times 10^5$ M$^{-1}$ s$^{-1}$ and $k_{off} = (3.5 \pm 1.2)$ s$^{-1}$. 
6.6 Summary

In summary, transport coefficients provide a new contrast mechanism to identify single molecules in solution. Fluorescence emission is only needed to detect and track single molecules; the measured quantities are independent of the probe but are directly linked to the physical properties of the biomolecules. For example, APC displayed substantial intensity fluctuations due to sequential photobleaching of its pigments and complicated photophysics (Figure 6.4b-c), but the measured transport coefficients remained constant for every trapped molecule and were independent of intensity. On the other hand, in the case of DNA hybridization, the probe intensity remained constant (Figure 6.6 and 6.7a) while it was the fluctuating transport coefficients that reported on binding/unbinding dynamics. Our method provides a general access to single-protein oligomerization states with a photon-limited resolving power (Figure 2.15). It also enables visualization of bi-molecular interactions having physiologically-relevant affinities ($K_d = 1\text{ nM} - 100\mu\text{M}$, Figure 6.9) (Holzmeister et al., 2014a) with possible extensions to multiple interactions or to buffers containing crowding agents or cell lysate. Moreover, all measurements are conducted in solution without surface attachment, which helps preserve a biomolecule’s conformational, rotational and binding degrees of freedom (Tyagi et al., 2014). It is intriguing that so much can be learned simply by precisely measuring how a single molecule moves in solution in a feedback-based trap.

6.7 Materials and methods

Sample preparation  Before single-molecule trapping experiments, both cross-linked *(Spirulina sp., Prozyme)* and wild-type *(Spirulina sp., Invitrogen)* allophycocyanin were buffer exchanged into $1\times$ PBS with a P6 spin column (Bio-rad) and subsequently diluted to ~6 pM. All ssDNA samples were purchased from Integrated DNA Technologies with the following sequences: 5’-ATC ATA CTA A -3’ (10nt), 5’-TTA GTA TGA TCC AAA AAA AAA AAA-3’ (24nt-10comp), 5’- TTA GTA TGA CCC AAA AAA AAA AAA -3’ (24nt-9comp, mismatch indicated by bold font). Amino modified (3’) 10nt-ssDNA was labeled with Alexa647 succinimidyl ester (Invitrogen) using a 5:1 dye-to-oligo ratio and purified by ethanol precipitation followed by reverse phase HPLC (C18 column, 10%-70%
Figure 6.9: Accessible parameter space of binding kinetics. The fundamental limit of our technique to study bi-molecular interaction kinetics in equilibrium can be estimated from the following constraints. Let $T_{\text{trap}}$ be the typical trapping time of a single molecule, $T_{\text{min}}$ be the time resolution of the transport coefficient measurement and $c$ be the concentration of the binding partner, we must have $T_{\text{min}} < k_{\text{off}}^{-1} < T_{\text{trap}}$ and $T_{\text{min}} < (k_{\text{on}} \cdot c)^{-1} < T_{\text{trap}}$, in order to resolve the kinetics. In the ABEL trap, we can achieve $T_{\text{trap}} \approx 10\,s$, $T_{\text{min}} \approx 50\,\text{ms}$ and $c < 50\mu\text{M}$ (the concentration of the unlabeled binding partner is limited by the amount of residual fluorescent impurities). The corresponding region of parameter space where the kinetics of the interaction can be studied by our technique is unshaded. The two DNA hybridization examples in Figure 6.7 are denoted with colored circles.
acetonitrile gradient over 40 min on a GE Akta Purifier 100/10). Hybridization experiments were conducted in a buffer containing 25 mM HEPES, 100 mM NaCl, 3 mM Trolox (Rasnik et al., 2006) (Sigma-Aldrich) and 10% glycerol. The concentration of the labeled strand was kept at ~0.75 pm to avoid another molecule entering the trap during a trapping event. Oxygen was removed by a combination of a protocatechuate system (Aitken et al., 2008) (50 nM of protocatechuate 3,4-dioxygenase and 2 mM of 3,4-dihydroxybenzoic acid, both from Sigma-Aldrich) and an argon blanket.

**ABEL trap instrumentation** Optical excitation was provided by a 594 nm HeNe laser. Rapid beam scanning was achieved by an orthogonal pair of acousto-optic deflectors (Trap B in Chapter 3). A Kalman filter, implemented on dedicated FPGA hardware, provided refined online estimation of the object’s position. Feedback voltages (<5 V in this study) were applied to the microfluidic cell via four platinum electrodes. Joule heating of buffer solution was measured to be insignificant (<1 °C, Appendix C) for the experimental conditions used in this study. Average excitation intensity was 300 W cm\(^{-2}\) in the APC dissociation experiments. At this excitation level, the trapping time of APC was limited to ~1 s by photobleaching or blinking (with a dark period >0.01 s) because increasing the excitation intensity led to shortened trapping events. An excitation intensity of ~2.3 kW cm\(^{-2}\) was used in DNA hybridization experiments. In this case, the limiting factor for trap loss was determined to be photobleaching as well, since no notable changes in average trapping duration were observed when different amounts of glycerol were added to slow down diffusion.

Trapping experiments were performed in all-quartz microfluidic cells fabricated with a channel depth of ~600 nm (Section 3.7). Prior to sample injection, the microfluidic device was cleaned using piranha solution (3:1 mixture of sulfuric acid and hydrogen peroxide) and subsequently incubated in 1 M potassium hydroxide for 20 minutes. To eliminate non-specific adsorption, which would otherwise influence the motional dynamics of trapped molecules, the interior of the microfluidic cell was rigorously passivated. For APC experiments, this was accomplished by pre-coating the cell with two pairs of polyelectrolyte multilayers (Section 3.7.2). For experiments involving DNA, a final concentration of 0.1% (weight) polyvinylpyrrolidone (PVP, 40 kDa, Sigma-Aldrich, see Section 3.7.2) was added.
to the final solution. PVP also efficiently eliminated electroosmotic flow produced by surface charges so that measured mobility directly reflected electrophoresis.

**Kernel density estimation**  Continuous density in the $D - \mu$ parameter space was estimated by substituting each data point on a scatter plot with a 2D Gaussian kernel (kernel density estimator). The covariance matrix of the kernel was chosen by an optimal criterion (Botev et al., 2010). Kernel estimated density helps to visualize the underlying distribution which can be otherwise overshadowed by the densely-packed symbols in a scatter plot.

**Hidden Markov model**  A hidden Markov model (HMM) was used to analyze the joint $D - \mu$ time trajectories in the DNA binding/unbinding experiment. The underlying transitions were assumed to happen between a bound and an unbound state with a $2 \times 2$ transition matrix. Each observation (50 ms binned $D$ and $\mu$) was modeled as being drawn from a two-dimensional Gaussian distribution in $D - \mu$ space. The transition matrix was estimated using the Baum-Welch algorithm and the most probable state trajectory was found by the Viterbi algorithm. We used the HMM module in the Probabilistic Modeling Toolkit (pmtk3, Murphy (2012)) for these calculations. Before the HMM optimization procedure, the observation densities were first estimated by a Gaussian-mixture-model and used as initial conditions.
Appendix A

A list of ABEL trap electronics
## APPENDIX A. A LIST OF ABEL TRAP ELECTRONICS

<table>
<thead>
<tr>
<th>Part</th>
<th>Manufacturer</th>
<th>Model number</th>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPGA</td>
<td>National Instruments</td>
<td>PCIe-7842R</td>
<td>1</td>
<td>R Series Multifunction RIO With Virtex-5 LX50 FPGA</td>
</tr>
<tr>
<td>Shielded cable that interfaces with the FPGA</td>
<td>National Instruments</td>
<td>SHC68-68-RMIO</td>
<td>1</td>
<td>Shielded cable designed specifically for NI R Series multifunction devices for MIO (mixed) connection</td>
</tr>
<tr>
<td>Shielded cable that interfaces with the FPGA</td>
<td>National Instruments</td>
<td>SHC68-68-RDIO</td>
<td>1</td>
<td>Shielded cable designed specifically for NI R Series multifunction devices for DIO (digital) connection</td>
</tr>
<tr>
<td>Breakout box</td>
<td>National Instruments</td>
<td>CA-1000</td>
<td>1</td>
<td>Customizable breakout box</td>
</tr>
<tr>
<td>Connector blocks inside the breakout box</td>
<td>National Instruments</td>
<td>CB-68LPR</td>
<td>2</td>
<td>Unshielded 68-pin I/O Connector Block</td>
</tr>
</tbody>
</table>

### Beam scanning control

<table>
<thead>
<tr>
<th>Part</th>
<th>Manufacturer</th>
<th>Model number</th>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acousto-optics deflector</td>
<td>AA Opto-electronic</td>
<td>MT110-B54A1.5-VIS</td>
<td>2</td>
<td>TeO$_2$ longitudinal mode AO crystal with 10 resolvable spots</td>
</tr>
<tr>
<td>AO controller/driver</td>
<td>AA Opto-electronic</td>
<td>DDSPA2X-D8b15b-34</td>
<td>1</td>
<td>Dual-channel digital RF synthesizer and amplifier</td>
</tr>
<tr>
<td>Male DB44 connector</td>
<td>Digi-key</td>
<td>180-M4401MN-ND</td>
<td>2</td>
<td>Interface on the AO controller</td>
</tr>
<tr>
<td>DB44 connector backshell</td>
<td>Digi-key</td>
<td>970-25NE-ND</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Ribbon cable</td>
<td>Digi-key</td>
<td>MC50R-5-ND</td>
<td>2</td>
<td>50-wire 5-ft shielded between AO controller and FPGA</td>
</tr>
</tbody>
</table>

### Electrokinetics generation on the microfluidic chip

<table>
<thead>
<tr>
<th>Part</th>
<th>Manufacturer</th>
<th>Model number</th>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HV op-amp</td>
<td>Digi-key</td>
<td>OPA453TA-1-ND</td>
<td>4</td>
<td>80 V, 50 mA op-amp with slew rate 38 V $\mu$s$^{-1}$</td>
</tr>
<tr>
<td>HV amp power supply</td>
<td>Instek</td>
<td>GPC-6030D</td>
<td>1</td>
<td>0-60V/0-3A two channel DC power supply</td>
</tr>
<tr>
<td>Platinum electrodes</td>
<td>Alfa Aesar</td>
<td>43288</td>
<td>10 cm</td>
<td>0.5 mm diameter, annealed, 99.95% (metals basis)</td>
</tr>
</tbody>
</table>

Table A.1: Electronics components
Appendix B

Protocols

B.1 Quartz cell cleaning procedure

Reagents

• Hydrogen peroxide

• Sulfuric acid

• 1 M potassium hydroxide solution

• Nanopure water

Equipment

• Fume hood

• Stirring hot plate

• Desiccated nitrogen stream

• Microscope
Procedures

1. Fish cell out of storage in nanopure water, dry with a desiccated stream of nitrogen.

2. Place cell under the microscope and examine the degree of contamination in the trapping region.

3. Clean the cell using the Piranha bath protocol (CAUTION: read the Piranha Solution SOP before proceeding, the following is NOT a substitute for the SOP)

   (a) Clean a glass beaker (~20 mL) with nanopure water, put in a stirring bar and have dried cell(s) around in the hood

   (b) Place the cleaned, dried beaker in clamp inside a glass secondary container, on a stirring hot plate in the fume hood.

   (c) With gloves and proper eye protection gear, pour in ~15 mL of sulfuric acid, turn stirring on (120 rpm). Put sulfuric acid back to cabinet.

   (d) Have a glass pipette ready, get hydrogen peroxide into the hood and pour in ~5 mL (so that $\text{H}_2\text{SO}_4 : \text{H}_2\text{O}_2 = 3 : 1$).

   (e) Within one minute, quickly withdraw some of the highly reactive mixture using the glass pipette and put down a drop on one of the four ports in the microfluidic cell.

   (f) Wait until the Piranha solution fills more than 80% of the cell interior and transfer the cell into the beaker.

   (g) Let the cell sit in the beaker for 20 min to 1 h, depending on the degree of debris contamination.

4. Fish the cell out and put it in a beaker full of nanopure water.

5. Clear out the Piranha solution that remains inside of the cell, this can be done by letting the cell sit in water overnight, or repeatedly flushing the cell with nanopure water. Dry the cell with the nitrogen stream. Examine under a microscope, the trapping region should be clear.
6. Filter ~15 mL 1 M KOH into another glass beaker. Fill the cell with KOH using a 20 µL pipette and immerse the cell in the KOH solution for ~10 min. Avoid sonication.

7. Rinse the cell with nanopure water. This is done by repeatedly (sequentially going through the four ports) drying the cell with nitrogen stream and injecting a drop of nanopure water. At this point, the interior of the cell should be very hydrophilic, if not, repeat the Piranha procedure.

8. After the last drying step, the cell is ready for sample injection or surface modification. This procedure prepares the interior of the cell to be negatively charged (deprotonated silanol groups).

### B.2 Coating the interior of a quartz cell with polyelectrolyte multilayers (PEMs)

**Reagents**

- Poly(acrylic acid, sodium salt) solution (MW~15,000 35% wt in H₂O, Sigma-Aldrich #416037, PAA)
- Poly(ethyleneimine) solution (MW~750,000 by LS, 50 wt. % in H₂O, Sigma-Aldrich #181978, PEI)

**Equipment**

- Fume hood
- pH meter
- Stirring hot plate
Procedures

1. Making stock PEM solutions

(a) PEI (+)
   i. Weigh ~8 g PEI solution and add nanopure water to make 500 mL.
   ii. Using HCl/NaOH, adjust the pH to 8.0 (starting pH ~ 10.5).
   iii. Filter with a 0.2 µm membrane filter.

(b) PAA (-)
   i. Weigh ~10 g PAA solution and add nanopure water to make 500 mL.
   ii. Using HCl/NaOH, adjust the pH to 8.0 (starting pH ~9.5).
   iii. Filter with a 0.2 µm membrane filter.

2. Layer-by-layer coating

(a) Make sure that the microfluidic cell is thoroughly cleaned and hydrophilic (See Appendix B.1).

(b) Using a 20 µL pipette, draw ~5 µL stock PEI solution and place a droplet on one of the four ports on the microfluidic chip. Allow the capillary forces to fill the interior of the cell. After all the channels are filled, place droplets on the remaining three ports to prevent rapid evaporation.

(c) Cover the cell with a petri dish on top and incubate for 10 min.

(d) Use a desiccated nitrogen stream to dry the cell and then rinse the channels with nanopure water two times.

(e) Deposit the PAA layer on top of the PEI layer following steps (b)-(d).

(f) Repeat steps (b)-(e) until the desired number of layers is reached.

3. The PEM layers can be removed by either Piranha bath or UV ozone clean.
B.3 Preparing the protocatechuic acid (PCA)/protocatechuate-3,4-dioxygenase (PCD) oxygen-scavenging system

This protocol is adapted from Aitken et al. (2008).

Reagents

- Protocatechuate-3,4-dioxygenase from *Pseudomonas sp.* (Sigma P8279, PCD)
- 3,4-Dihydroxybenzoic acid (Aldrich 37580, PCA)

Equipments

- Fume hood
- pH meter
- Stirring hot plate

Procedures

1. Make storage buffer: 10 mM Tris-HCl + 1 mM EDTA + 50% glycerol

2. Prepare the PCD stock solution
   (a) Dissolve the lyophilized protein powder in the buffer prepared in Step 1.
   (b) (Optional) determine the concentration of PCD by UV/VIS. The extinction coefficient of PCD is estimated to be $7.4 \times 10^5 \text{M}^{-1} \text{cm}^{-1}$ (Patil et al., 2000).
   (c) Make aliquots with a concentration of ~2 µM (40×), store at −20 °C.

3. Prepare the PCA stock solution
   (a) Dissolve in nanopure water to make ~20 mL, 100 mM solution (15.4 mg/mL).
   (b) Adjust the pH to 8 using HCl/KOH (starting pH ~ 2.6).
   (c) Filter with a 0.2 µm membrane filter.
Figure B.1: The enzymatic activity of PCD can be verified by monitoring the loss of the A290 peak of the substrate (PCA).

(d) Make aliquots (40×) and store at −20°C.

4. (Optional) verify the activity of the enzyme by monitoring the disappearance of the A290 peak of PCA in the presence of PCD and ambient oxygen. See Aitken et al. (2008) for details and an example measurement is shown in Figure B.1

B.4 Labeling amino-modified oligonucleotides with NHS dyes

Reagents

- Amino-modified oligonucleotides (Integrated DNA Technologies)
- Reactive dyes with the NHS functional group (stored in anhydrous DMSO to prevent hydrolysis)
- Ethanol (“200 proof” grade >99.5%)
• Reaction buffer (freshly made 0.1 M sodium carbonate buffer, pH 9)
• Storage buffer (1× TE: 10 mM Tris, 1 mM EDTA, pH 8)
• 3 M NaCl

Equipment

• −80 °C freezer
• Centrifuge
• HPLC with a C18 reverse phase column

Procedures

1. Dissolve the oligonucleotide in 1× TE buffer with a typical concentration of ~100 µM.

2. Store the high concentration sample at −20 °C, prepare 50 µL, ~5 µM for the labeling reaction.

3. Purification before labeling: ethanol precipitation

   (a) Starting with the 50 µL sample, add 5 µL 3 M NaCl, mix well, and add 150 µL 100% cold ethanol (−20 °C), mix well.

   (b) Store the Eppendorf tube in the −80 °C freezer for 1 h.

   (c) Centrifuge at 0 °C, 12,000 g for 30 min (be sure to pre-cool the centrifuge and carefully weigh the counter balance), remember the orientation of the tube for easy identification of the pellet in the next step.

   (d) Carefully remove the supernatant with a pipette tip. Take care not to disturb the precipitated pellet of the oligonucleotide (which may be difficult to see). Use the pipette tip to remove any drops of fluid that adhere to the walls of the tube. Save and label the supernatant.
(e) Add 200 µL 70% ethanol (−20 °C) and recentrifuge at 12,000 g for 5 min at 4 °C.

(f) Repeat Step (d)

(g) Let the remaining solvent evaporate in a clean hood.

4. Labeling with a reactive dye (e.g. NHS-Alexa 647)

   (a) Dissolve the pellet in the reaction buffer (freshly made 0.1 M sodium carbonate buffer, pH 9), add the reactive dye (in DMSO), so that the dye to oligo ratio is about 5 : 1. Add reaction buffer to a volume of 50 µL.

   (b) Incubate at room temperature, on a rotator overnight.

5. Purify away the reactive dyes

   (a) Ethanol precipitation of the labeled oligonucleotides, see Step 3.

   (b) On an HPLC, run a 5-90% acetonitrile gradient on a C18 column over 40 minutes, monitor A260 and A650 (dye absorption peak), collect the fraction with the highest A260/A650 ratio. Alternatively, ethanol precipitation 3 times also seemed to work very well, although it is more time consuming and lossy.
Appendix C

Joule heating in the ABEL trap

Resistive heating of the buffer solution can be significant in electrokinetically driven microfluidic devices (Erickson et al., 2003), especially when the buffer contains high concentrations of salt. To assess the severity of Joule heating in the ABEL trap, I conducted excited-state lifetime measurements on a temperature-sensitive dye Rhodamine B (RhodB), as a function of applied voltages. The fluorescence lifetime of RhodB is known to decrease with increasing temperature at an approximate rate of $-0.03 \text{ ns K}^{-1}$ in the range of 20-40 °C.

In conducting these measurements, ~100 nM of RhodB was dissolved in a buffer containing 30 mM HEPES, 100 mM NaCl and 30 mM MgCl$_2$ (similar ionic strength as the hybridization buffer in Chapter 6 and the 1× PBS used in Chapters 4 and 5) and injected into the microfluidic trapping device. Time-correlated single-photon counting was performed with 515 nm pulsed excitation to measure fluorescence lifetime. Different magnitudes of applied voltages, in the form of 100 Hz square waves, were applied to the device.

![Figure C.1: Probing voltage induced heating in the ABEL trap using the temperature sensitive dye Rhodamine B. See text for details.](image-url)
As shown by Figure C.1, clear drops in measured lifetime were observed when voltages exceeding 20 V were applied to the microfluidic channels. Higher voltages gave rise to larger decreases in lifetime. Moreover, the measured lifetime immediately restored to the initial value after applied voltages were switched off. Such observed behavior indicates voltage-induced resistive heating. Fortunately, operation of the ABEL trap only requires very small voltages (e.g. the standard deviation of applied feedback voltages is about ~4 V in the case of electrophoretic trapping of ssDNA and ~2 V in the case of PEM-assisted electro-osmotic trapping of proteins). We estimated a temperature rise of < 0.5 °C for APC experiments (Chapter 5) and < 1 °C for the binding experiments (Chapter 6).
Appendix D

Example Matlab code for processing the ABEL trap data
APPENDIX D. EXAMPLE CODE

DMA 0

<table>
<thead>
<tr>
<th>photon time tag</th>
<th>ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>31</td>
<td>47</td>
</tr>
<tr>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

DMA 1

<table>
<thead>
<tr>
<th>rX</th>
<th>kgainX</th>
<th>mX</th>
<th>rY</th>
<th>kgainY</th>
<th>mY</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>15</td>
<td>31</td>
<td>47</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

DMA 2

<table>
<thead>
<tr>
<th>rawX</th>
<th>rawY</th>
<th>fbttLag</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>47</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

% length scaling, digital equivalence of 1 um
posScale = 4096;

% time resolution of the photon time tags, in milliseconds
ttTmRes = 12.5e-6;

% time resolution of the feedback kicks
fbTmRes = 96*ttTmRes;

% max number of photons to read
maxPhoton = fileSize(apdChPath);

% read the integer time tag of each photon
apdtt_raw = fread(apdttPt,maxPhoton,'uint32');

% read channel of each photon
apdCh = fread(apdChPt,maxPhoton,'uint8');

% fix wrap
apdtt_ms = unWrapTTmx(apdtt_raw)*ttTmRes;

% read the fbttlag (time since previous feedback)
fbttlag_ms = fread(fbttlagPt,maxPhoton,'uint16')*fbTmRes;

% read the posterior estimates
mX = fread(mXPt,maxPhoton,'int16')/posScale;

mY = fread(mYPt,maxPhoton,'int16')/posScale;

% read the raw measurements
rawX = fread(rawXPt,maxPhoton,'int16')/posScale;

rawY = fread(rawYPt,maxPhoton,'int16')/posScale;

Figure D.1: Return data structure from the FPGA and example Matlab code to read the data files. Data is transferred from the FPGA to the host computer’s memory via DMA (direct memory access), unpackaged and written to binary files in LabVIEW.
Bibliography


BIBLIOGRAPHY


imaging of transient binding on DNA origami”. *Nano Lett.* **10** (11), 4756–4761. DOI: 10.1021/n1103427w (see p. 175).


measurement using photon correlation”. Nature Photon. 4 (10), 696–699. DOI: 10.1038/nphoton.2010.174 (see p. 113).


Trautman, J. K., Macklin, J. J., Brus, L. E., and Betzig, E. (1994). “Near-field spectroscopy of single molecules at room-temperature”. *Nature* **369** (6475), 40–42. DOI: [10.1038/369040a0](http://dx.doi.org/10.1038/369040a0) (see pp. 100, 142).


without immobilization”. *Nat. Methods* **11** (3), 297–300. DOI: [10.1038/nmeth.2809](https://doi.org/10.1038/nmeth.2809) (see pp. 95, 178).


