Wide-field correlation spectroscopy of aqueous molecular solutions using CLiC imaging

John Ross

Department of Physics

McGill University, Montreal

August 15, 2018

A thesis presented to McGill University in partial fulfillment of the requirements of the degree of Masters of Science

ACKNOWLEDGMENTS

I would like to sincerely thank my supervisors Sabrina Leslie, Paul Wiseman, and Mark Sutton for providing me support and guidance throughout my Master's degree. This research would not have been possible without funding from NSERC and CREATE Training Program in Bionanomachines (CTPB), and I am truly thankful. I would like to thank my lab-mates with whom I have collaborated over the course of my degree, in particular Simon Sehayek, Zhiyue Zhang, Hugo Brandão, and Jason Leith. Hugo Brandão was especially helpful for code development with kICS. I would also like to thank Peter Grutter, Zach Abrams, Robin Beech, and Jen Weidner for their wisdom and understanding. Finally, I would like to thank my family, without whom none of this would be possible.

ABSTRACT

Many biochemical processes rely on weak and slow biomolecular interactions, such as interactions between protein and DNA molecules during DNA replication, repair, and transcription. Developing sensitive methodologies to characterize their energetics, kinetics, stoichiometries, and cooperativities in free solution is important for developing new biophysical insights. However, single-molecule experiments are typically limited to low concentrations and low signal to noise ratios when using standard confocal microscopy and total internal reflection fluorescence microscopy (TIRFM). This project introduces a new way of visualizing nanoscale interactions between biomolecules which combines Leslie's Convex Lens-induced Confinement (CLiC) imaging platform with Wiseman's wide-field k-space image correlation spectroscopy (kICS) techniques. CLiC confines a sample to a nanoscale wedge-shaped gap, meaning that the biomolecules are essentially restricted to two dimensions of movement. Due to the shallow sample depth, out-of-focus background fluorescence can be eliminated by a factor of 20 or more compared to TIRFM and confocal imaging. Similarly, per-molecule observation times of CLiC are increased by a factor of 10 000 compared to standard methods, meaning that very weak interactions and slow dynamics that were previously impossible to measure can now be analyzed. The kICS analysis also has inherent advantages, including correlation function insensitivity to quantum efficiency, optical system point spread function (PSF), and photophysical effects such as blinking. This project is primarily focused toward developing the CLiC kICS methodology using simple oligonucleotide systems. Analytical expressions are used to predict appropriate fitting ranges of the correlation function to extract characteristic rates. Furthermore, the single-dimension relationship between individual system parameters and correlation bias is calculated to assess appropriate imaging limits.

RÉSUMÉ

De nombreux processus biochimique, comme la réplication, réparation, et la transcription de l'ADN, sont dependents sur les interactions biomoléculaires faibles. C'est donc important de développer des méthodologies sensibles pour caractériser leur énergie, leur cinétique, leurs stoechiométries et leurs coopératives. Cependent, les expériences à seule-molécule sont généralement limitées aux faibles concentrations et de faibles rapports signal sur bruit en utilisant la microscopie confocale et la microscopie de fluorescence à réflexion interne (TIRFM). Ce projet introduit une nouvelle façon de visualiser les interactions nanométriques entre biomolécules qui combine la plateforme d'imagerie Confinement Convexe Lentille-induit (CLiC) de Leslie avec les techniques de spectroscopie de corrélation d'images en espace réciproque (kICS) de Wiseman. CLiC confine un échantillon à une hauteur à l'échelle nanométrique, réduisant ainsi les biomolécules à deux dimensions de mouvement. En raison de la faible profondeur de l'échantillon, la fluorescence de fond hors foyer peut être éliminée d'un facteur 20 ou plus par rapport à la TIRFM et à l'imagerie confocale. De même, les temps d'observation par molécule de CLiC sont multipliés par un facteur de 10 000 par rapport aux méthodes traditionelles, ce qui signifie que des interactions très faibles et les dynamiques lentes auparavant impossibles à mesurer peuvent maintenant être analysées. L'analyse kICS présente également des avantages inhérents, y compris l'insensibilité à l'efficacité quantique, la fonction d'étalement du point (PSF) et les effets photophysiques tels que le clignotement. Ce projet vise à établir la méthodologie CLiC-kICS en utilisant des systèmes oligonucléotidiques simples. Des expressions analytiques sont utilisées pour prédire des plages d'ajustement appropriées de la fonction de corrélation pour extraire les taux caractéristiques. En outre, la relation unidimensionnelle entre les paramètres individuels du système et le biais de corrélation est calculée pour évaluer les limites d'imagerie appropriées.

Contents

List of Figures

List of Acronyms

AFM - Atomic Force Microscopy

AWGN - Additive White Gaussian Noise

CLiC - Convex Lens-induced Confinement

DNA - Deoxyribonucleic Acid

EGF - Epidermal Growth Factor

EMCCD - Electron Multiplying Charge Coupled Device

FCS - Fluorescence Correlation Spectroscopy

FRET - Förster Resonant Energy Transfer

GFP - Green Fluorescent Protein

ICS - Image Correlation Spectroscopy

kICS - k-space Image Correlation Spectroscopy

LSFM - Light Sheet Fluorescence Microscopy

NA - Numerical Aperture

PSF - Point Spread Function

RICS - Raster Image Correlation Spectroscopy

SEM - Scanning Electron Microscopy

SPM - Scanning Probe Microscopy

STICS - Spatio-temporal Image Correlation Spectroscopy

STM - Scanning Tunneling Microscopy

TICS - Temporal Image Correlation Spectroscopy

TEM - Transmission Electron Microscopy

TIRFM - Total Internal Reflection Fluorescence Microscopy

Chapter 1

Chapter 1 - Motivation for Using kICS with CLiC

1.1 Introduction

Beyond the micrometre scale of analysis, studies on cellular systems have uncovered a vast network of molecular processes spanning orders of magnitude in both the spatial and temporal domains. These sub-cellular systems are responsible for regulating a wide range of mechanisms, including but not limited to: cell division, cell growth, apoptosis, DNA repair, replication, and recombination, and cell signaling events that are responsible for tissue repair and immunological responses [1]. These processes have been shown to be highly dependent on the physical and chemical properties of biomolecular agents; namely their mobility, interaction rates, cooperativity, and conformation changes.

The purpose of many nanoscale biochemical experiments that mimic physiological conditions are centered around quantifying these biomolecular properties. Although this regime is often difficult to study, recent advancements in microscopic technologies have given insight into biological systems at the molecular resolution level.

The focus of this chapter is to assess probing and analysis capabilities with the goal

of measuring physical and chemical properties of biomolecules. The following section provides a summary of non-optical and optical microscopy techniques in the context of cellular systems and single molecule experiments. This is followed by an overview of correlation-based analysis techniques that have been used to quantify dynamic systems.

1.2 Microscopy Techniques for Sub-Cellular Observation

The importance of developing accurate measurement techniques cannot be understated in the physical sciences. Every scientific advancement has originated from a new measurement or observation, which in turn validates or disproves existing hypotheses. In the field of cellular biology, sophisticated microscopy techniques have given insight into nanoscale processes which provides a more complete model of cellular functions. However, achieving these measurements is a challenging task for several reasons.

First and foremost, the small specimen size makes probing very difficult. To understand the relative scale of these systems, animal cells are 10-20 um in diameter, [1], while organelles are on the order of single microns [2], and proteins found in the human body are approximately 1.8 nm in diameter [3]. This means that all sub-cellular experiments must have micron scale resolution or finer, with single molecule experiments requiring a resolution in the nanoscale regime. For optical-based methods, this is a challenge due to the diffractions limit of light.

As light encounters a physical barrier, the interaction between the photon and the object produces a change in the light's phase, intensity, and direction. This phenomenon, known as diffraction, is responsible for airy disks seen around circular apertures, or the fringe pattern seen in optical slit experiments. In optical microscopy, the extent of the diffraction is based on the system's optical parameters such as the numerical aperture of the objective and the wavelength of light used. In the 19th Century, German physicist Ernst Abbe estimated the maximum spatial distance that can be resolved between two

objects:

$$d = \frac{\lambda}{2n \sin\theta} = \frac{\lambda}{2NA} \tag{1.1}$$

Where d is the spatial resolution limit, λ is the wavelength of light, n is the refractive index of the medium, θ is the angle of the spot convergence, and NA is the numerical aperture. This value, known as the Abbe diffraction limit [4], is the fundamental maximum spatial resolution limit of a system based on the wavelength nature of light. The reciprocal relationship between the source wavelength and the resolution limit will reappear multiple times in the following discussion for optical, X-ray, and electron microscopy techniques.

Another complication to imaging within this regime is that thermodynamic fluctuations and quantum effects must be considered more heavily. For example, the effect of vibrations that are caused by thermal expansion and contractions, nearby equipment, and the natural frequency of the buildings, can become a significant contribution to noise. This is especially true in raster scanning methods; scanning electron microscopes and transmission electron microscopes are particularly sensitive to horizontal vibrations, while scanning probe microscopy techniques are much more sensitive to vertical vibrations [5].

Another major barrier associated with probing cellular systems pertains to cell viability and the difficulty of probing without damaging the cell. For light-based methods, short wavelength UV light has been shown to damage adjacent thymine bases in DNA [6], and to be toxic for living cells [7] [8]. For non-optical methods, invasive probing techniques have been shown to cause irreparable cell damage if not carefully modulated [9] [10] [11]. Furthermore, there is a great amount of difficulty in performing experiments *in vivo*. In cell cultures, it is necessary to maintain consistent temperature and humidity conditions that support normal cellular behavior across the data acquisition time. As such, *in vitro* observation remains the only option for imaging times longer than a few hours [12].

Finally, assessing dynamic characteristics of sub-cellular systems presents a new set of challenges to overcome. Molecules in free solution are subject to Brownian motion, meaning that their position evolves over time according to their characteristic diffusion

rate. Since many optical and non-optical microscopy techniques cannot observe a wide spatial range, molecules may only be within an observation region for a brief period of time. Furthermore, single molecule localization experiments require an elevated level of precision to map the trajectory of a specific subset of molecules in a heterogeneous solution. These challenges will be explored further with respect to fluorescence microscopy in section 1.2.2.

Despite these obstacles, significant technological progress in detector efficiency, piezoelectric actuator speed and precision, probe design, and image processing techniques have significantly advanced microscopy capabilities. The current state of non-optical microscopy methods as well as some notable experimental results from these methods are discussed in the following section.

1.2.1 Non-Optical Microscopy

Biological microscopy can be divided into two sub-fields; optical and non-optical methods. While optical microscopy uses the interaction of visible light with the specimen for measurement, the domain of non-optical methods uses alternative matter interactions, such as the specimen's force deflection response or electron-based interactions. The advantage of non-optical methods is that the diffraction limit is either circumvented completely (in the case of scanning probe microscopy techniques) or is significantly reduced (in the case of X-ray and electron microscopy); thus, resulting in higher resolution capabilities. Although many types of microscopy techniques have been developed, few methods have been sufficiently developed to a degree of practicality. For instance, scanning tunnelling microscopes (STM) produce weak tunneling signals from poorly-conducting specimen such as DNA [13], and acoustic microscopes produce grainy scans that are often difficult to interpret [14]. As such, this section will be focusing primarily on more developed methods such as atomic force microscopy (AFM), electron microscopy variants, and X-ray probing.

Atomic Force Microscope

Atomic force microscopy is a type of scanning probe microscope (like STM) that uses a fine-tip probe to raster scan a specimen. The measured repulsive force generated between the probe and the sample can be used to characterize topography, defects, and surface roughness. In the context of biological samples, AFM scans have been used to observe size and conformation of molecules and aggregates [15]. While AFM experiments with biological specimens can cause structural damage, it also poses a lot of benefits over other microscopy methods. Besides the ability to surpass the diffraction limit and scan at nanometre resolution, AFM can operate in air and liquid, meaning that samples can be scanned in their native physiological buffer solutions. Furthermore, sample preparation is relatively simple, as it does not require any freezing, metallic coating or dyes [16]. As such, AFM is a powerful tool for gaining knowledge in sample structure and mechanical properties. Recent experiments in AFM have been able to observe protein unfolding [17], characterize cellular adhesion forces [18], and study drug interactions with supported lipid bi-layers [11].

Electron Microscopy and X-Ray Methods

Another variant of non-optical microscopy methods is the electron microscope, which uses signals from transmitted and scattered electrons from the specimen interaction to gain insight on cellular structures. Considering the wave-particle duality of electrons, resolution can be shown to be inversely proportional to the electron beam's acceleration voltage. As such, typical electron beam voltages produce signals that can achieve 1000-fold greater resolution than optical techniques [19]. However, a major disadvantage to this method is that electron microscopy requires a vacuum environment to reduce electron beam scattering. Moreover, biological samples generally require freezing; meaning that physiological conditions are impossible to reproduce with electron microscopy, and that dynamic experiments cannot be performed. Despite this, the strength in electron

microscopy lies in its ability to gain structural information on sub-cellular components. For example, scanning electron microscopy (SEM) has been used to visualize the DNA double helix in the chromatin of chicken [20]. Au-labeled epidermal growth factor (EGF) molecules bound to cellular receptors has been imaged in intact cells at the resolution of approximately 4 nm [21]; a feat that cannot be matched by other microscopy methods [22].

Closely related to this approach are X-ray based techniques that use X-ray interactions to generate a three-dimensional mapping. While high energy light such as X-rays increase the effective resolution of the sample due to the wavelength dependence of the diffraction limit, the incident light must remain in the "soft X-ray" bandwidth to avoid structural damage in the sample. X-ray crystallography has been used to study the structure DNA, as well the proteins myoglobin and hemoglobin at angstrom resolution [23]. Furthermore, small angle scattering (or solution X-ray scattering) has also been used to study protein folding [24].

Summary of Non-Optical Methods

The main advantage of the presented non-optical techniques over optical techniques is that resolution can be orders of magnitude greater in non-optical techniques. In AFM, the resolution is dependent on the mechanical properties of the probe, and nanometre to sub-nanometre resolutions have been demonstrated [25], while electron microscopy experiments are in the resolution range of $\geq$2 nm for biological specimen [26].

However, there are inherent disadvantages to these methods. As previously mentioned, biological samples are inherently fragile, and tend to be more likely to experience structural damage in non-optical techniques. Invasive AFM probes have been shown to cause deformations on the surface of cells (even in tapping mode), and the vacuum conditions and high energy electron beams in electron microscopy can be damaging to DNA samples. Moreover, the sample fixation process in electron microscopy has been shown to

alter the specimen microstructure [27]. Another disadvantage is that electron microscopy and AFM require fixed deposition of samples. Due to the nature of electron microscopy experiments, it is impossible to witness more than a snapshot of the specimen in time. With AFM, it is possible to generate a time series signal of the specimen, but it is necessary to exchange temporal resolution for spatial resolution. As a final note on this subject, high speed local raster scanning may be a promising solution to this challenge [28], but it currently cannot compare to the capabilities of wide-field imaging techniques.

1.2.2 Optical Microscopy

Another family of methods is the branch of optical microscopy which use the transmission, reflection, and emission of light in the visible spectrum to study specimens. Since their development, optical microscopes have played a vital role in biological research due to their ability to extract multidimensional structural and functional properties in a minimally invasive manner. Optical methods are diffraction limited; thereby putting the maximum attainable resolution at approximately half the probing wavelength, or 200 nm for blue light. Sophisticated equipment such as high numerical aperture objectives [29], light polarizing filters [30], and adaptive optical techniques [31] have helped reduce aberrations in optical systems, but capabilities generally fall short of the absolute diffraction limit. As an example, brightfield methods can only resolve cell outlines without additional phase-contrasting optics [32]. While there are many types of microscopy techniques that can be used for sub-cellular imaging, fluorescence-based microscopy techniques remain one of the most widely used due to its ability to provide high specificity and resolution in a minimally invasive manner [33] [34]. In order to understand the advantages that fluorescence microscopy present, it is necessary to investigate the broader topic of fluorescence and photophysics.

Fluorescence Overview and Applications in Single Molecule Experiments

Fluorescence is a quantum mechanical phenomenon whereby an electron in an excited singlet state undergoes non radiative vibrational relaxation on a ps time scale, after which it relaxes to the ground singlet state with emission of a photon of fluorescence. The opposite process, in which thermal or electromagnetic radiation promotes an electron to an excited state, is referred to as absorption. This absorption-emission cycle typically lasts on the order of nanoseconds and can occur repeatedly until the electron enters a permanent non-fluorescing state; in which case the molecule is said to be photobleached [35]. It should be noted that electrons generally experience thermal relaxation after entering a new state, causing the electron to enter a lower energy state. This energy bandgap difference produces a red-shift in the emission spectrum with respect to the absorption spectrum and is known as the Stokes shift [35].

These properties are exploited in the design of fluorescence microscopes. While imaging a fluorescent specimen, the molecule can be continuously promoted to an excited state using a light source that matches its absorption spectrum. The molecule then fluoresces, emitting a photon of lower energy that is collected through the same optical path way as the excitation profile. A series of filters can then be used to eliminate light from outside the emission bandwidth, and dichroic mirrors can be used to direct the signal to photon detectors [36]. In this way, it becomes possible to extract spatio-temporal information of the specimen and to infer concentrations and dynamic properties.

One of the more significant developments in fluorescence microscopy was the isolation of green fluorescent protein (GFP); a bioluminescent protein found in jellyfish with a naturally high quantum yield. The ability to fuse this GFP gene to genes for other non-fluorescent proteins made it possible to label and efficiently track cellular components in live cells over time [37]. Shortly after its isolation, variants of the GFP protein have been produced to fluoresce at an assortment of colours in the visible spectrum [38].

This ability to engineer naturally fluorescent species to suit experimental needs intro-

duced a level of precision and specificity that is unmatched in other microscopy methods [39]. Fluorescent probes could be designed to readily bond to conjugate sequences to their target molecules, allowing precision at the single molecule level. Prior to this development, traditional experiments on cellular systems were performed at an ensemble average level of the population; meaning that valuable information hidden within the bulk regime would be averaged out as noise. Single molecule studies have been able to give insight into processes such as protein folding, cell division mechanisms, signal transduction, and DNA replication and remodelling [39]. As a final note on single molecule methods, it is important to discuss some of the capabilities of Förster resonance energy transfer, or FRET. This mechanism occurs when energy is transfered non-radiatively from a donor fluorophore to an acceptor fluorophore. The resulting intensity signal can be used to accurately assess the distance between the molecules when it is within the 1-10 nm range [40]. As such, FRET is an invaluable tool for studying a range of molecular interactions, such as conformational changes in the Listeria monocytogenes and to monitor DNA synthesis in real time [41].

In reflecting over the challenges of biological microscopy from the beginning of this chapter, it is clear that fluorescence microscopy presents the most advantageous solution due to its non-invasive nature, its tagging specificity, and its capabilities for producing high resolution signals in dynamic systems. As such, fluorescence-based microscopy will be the sole focus for the remainder of this chapter and for this thesis.

Fluorescence Microscopy Techniques

One of the most significant challenges in fluorescence microscopy is its relatively poor signal to noise ratio. The intensity of the excitation beam is typically several hundred thousand to a million times brighter than the emitted fluorescence [42], and there are significant noise contributions from electronics in the detector and autofluorescence from the sample and substrate [43]. Furthermore, insufficient background suppression can

lead to out-of-focus fluorescence, thereby reducing the signal to noise ratio and causing premature photobleaching. This subsection describes various fluorescence techniques that use optical properties to selectively excite a sub-sample of fluorophores and increase the relative signal strength.

A widely used technique in sub-cellular analysis is confocal microscopy, which uses a small pinhole aperture in the confocal plane of the lens to limit out-of-focus fluorescence. By applying this pinhole filter, a very high background suppression is achieved since only light within the diffraction volume is permitted to pass through to the detector [44]. The size of the detection volume is dependent on the system's optical properties (i.e. pinhole size, numerical aperture of the objective) and beam parameters (i.e. wavelength, waist size), but it can be approximated as an ellipsoid with an approximate volume of one femtoliter [35]. Due to its high background suppression and small observation volume, confocal microscopy is used to generate high quality three dimensional mappings of biological specimens via raster scanning [45], and for assessing sample dynamics using intensity fluctuations within the observation volume [44].

While confocal systems can produce a composite image, it is insufficient for visualizing process pathways or localization-based experiments that require a wide-field view of the system. A commonly used wide-field method is total internal reflection fluorescence microscopy, or TIRFM. In this optical system, the excitation beam is oriented at an angle greater than the critical angle, resulting in an axially exponentially decaying evanescent wave near the surface of a cover-slip. As such, only fluorophores within a thin optical section of this surface will undergo fluorescence. TIRFM has been used to visualize receptor kinetics and study the roles of different proteins in exocytosis/endocytosis events in live cells [46].

Another example of a wide-field imaging platform is light sheet fluorescence microscopy (LSFM). In this method, a cylindrical lens is used to focus a thin sheet into the specimen, thereby selectively exciting a thin cross-section of the sample. While adoption

of this method is not currently as universal as confocal or TIRFM techniques, LSFM has been demonstrated to be an effective microscopic technique for in vivo samples. Two notable examples include whole-animal neural activity imaging [47], and for gaining insight into the development of embryos [48].

The fluorescence microscopy techniques presented in this section are very effective for their specific purpose; confocal microscopy can produce an excellent quality three-dimensional structural imaging; TIRFM is employed best with samples that are functionalized to the coverglass; and LSFM is effective for imaging samples in vivo. However, in the context of free solution dynamics, all these techniques are suboptimal, as molecules would tend to diffuse out of the illumination volume after a characteristic time. This in effect would influence the signal to noise ratio and would limit the concentration limit that would be practical for an experiment. With this in mind, it is necessary to introduce a relatively novel fluorescence imaging platform: convex lens induced confinement, or CLiC.

Convex Lens-Induced Confinement Microscopy

CLiC is a nano-fluidic fluorescence technique in which an aqueous sample is contained between two coverslips. A lens is used to apply pressure on the top coverslip, which variably deforms the height of the chamber (refer to Figure 1.1). In doing this, transferring molecules from a microfluidic loading environment to a nanofluidic environment becomes seamless. Additionally, the sample molecules are essentially confined to two spatial degrees of motion in the lateral directions, and thus remain within the focal plane for all imaging times. This aspect is what differentiates CLiC from other fluorescence techniques; not only is the sample physically confined to the focal plane, but noise due to out of focus fluorescence is eliminated. Furthermore, nano-lithographic techniques can be used to define the confinement space with greater precision. Some notable applications of CLiC include controlling polymer reaction rates [49], DNA conformation studies [50],

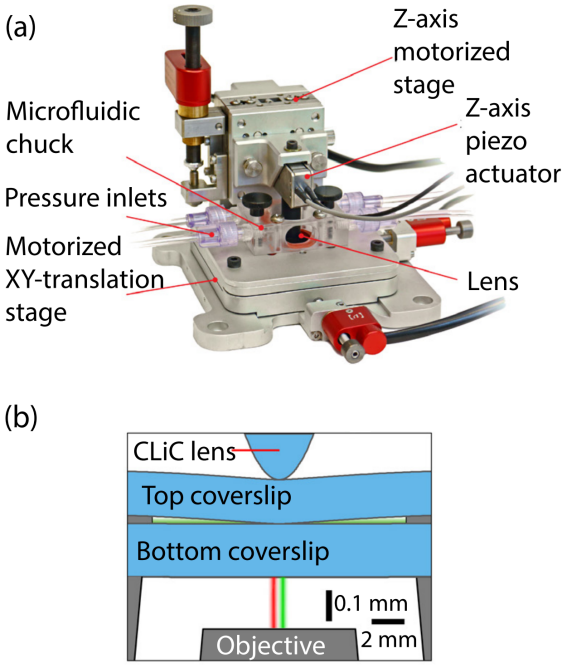

Figure 1.1: **a)** An overview of the CLiC instrumentation system. The vertical position of the CLiC lens is controlled by a piezoelectric actuator, and is used to apply pressure to the top coverslip. **b)** The size of the vertical confinement as well as the curvature of the coverslips is dependent on the force applied by the CLiC lens.

assessing macromolecule size [51], and visualizing topology-mediated interactions [52].

In the context of free solution analysis, CLiC is beneficial for several reasons. As a wide-field technique, CLiC enables a better statistical sampling of the dynamics, and captures information in the spatial domain. This provides more accurate results for studying diffusion, sample flow, and photophysics. Furthermore, the confinement of molecules to the focal plane guarantees three things: suppressed background noise, longer observation times, and higher concentration limits that resemble physiological conditions [51]. Free solution comparisons between confocal microscopy, TIRFM, and CLiC was performed by Leslie et al. [51] and confirms that CLiC can visualize samples with a higher concentration limit and can be observed for a longer time than the two alternative methods (refer to Figure 1.2).

Given the benefits of CLiC microscopy for free solution experiments, this platform

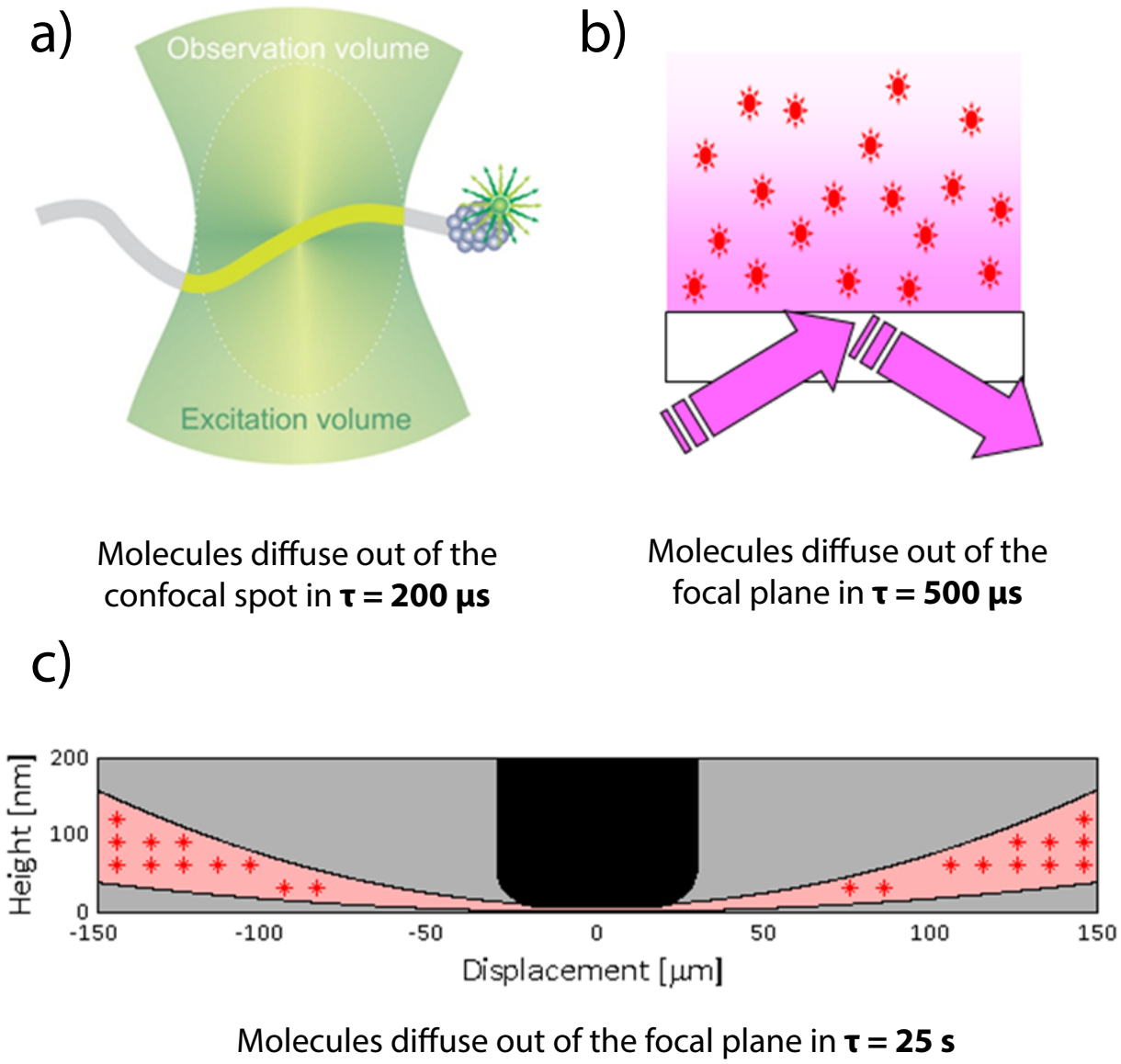

Figure 1.2: An overview of the imaging specifications of **a)** confocal microscopy; **b)** TIRF microscopy; and **c)** CLiC microscopy. Each microscopy method optimizes the signal-to-noise ratio in different ways, resulting in vastly different observation times for identical samples.

will be the sole technique considered for the remainder of this thesis. The first half of this review is now complete; given a suite of probing techniques, an appropriate method for examining dynamic sub-cellular systems has been found. It is now appropriate to investigate analysis techniques that suit this application. Given that fluorescence microscopy tends to produce relatively noisy data, correlation-based fluctuation techniques are widely employed to extract meaningful data. These techniques will be the sole focus for the remainder of this chapter.

1.3 Correlation-Based Analysis Techniques

Correlation functions can be considered a measure of self-similarity between two different signals (in the case of cross-correlation) or between a signal and itself (in the case of autocorrelation). A normalized intensity fluctuation autocorrelation function is defined as:

$$G(\tau) = \frac{\langle \delta i(t) \delta i(t+\tau) \rangle}{\langle i(t) \rangle^2} \tag{1.2}$$

Where the angular brackets $<>$ denote the average quantity, δ is the fluctuation from the mean, i is the signal intensity, τ is the lag time of light, t is the time, and G is the correlation function.

It should be noted that the correlation signal is based on the fluctuation of the intensity signal from the mean. This is purposely done, as additive white gaussian noise (AWGN) will not correlate with itself for non-zero lag times, while characteristic rates will still appear in the fluctuations. Furthermore, the averaging brackets demonstrate that the correlation function is solely a function of the lag time rather than the absolute time; a valid assumption for Weiner processes such as diffusion. Due to this averaging property, correlation functions are a useful tool for extracting characteristic rates from a signal while averaging out noise. As such, correlation functions are used in applications that produce noisy data, such as astronomy [53], medical physics [54], and geophysics [55].

In the early 1970's, Magde et al. developed fluorescence correlation spectroscopy (FCS), a technique that uses intensity fluctuations of mobile fluorescent molecules to calculate chemical reaction kinetics and the lateral diffusion rates [56] [57]. As fluorescence microscopy techniques became more sophisticated, FCS analysis has been adapted for different equipment such as confocal microscopy [58] and TIRFM [59]. Notable applications of FCS systems include studying conformational fluctuations of single nucleic acid molecules [60] and protein-ligand interactions in free solution [61].

While FCS applications typically deal with fluctuations in the temporal domain, the same principle can be applied to the spatial domain. Using independent spatial samples from a confocal scanning laser microscope, the correlated fluctuations in spatial intensities produce a characteristic particle density term and can be used to quantify aggregates. This technique, applied to an image rather than a time series, is aptly named image correlation spectroscopy (ICS) [62]. Expanding correlation analysis to the spatial domain has resulted in multiple variants of ICS. Raster image correlation spectroscopy (RICS) uses intensities gathered from a raster scan to gain insight on dynamics as a function of space and time (according to the scanning speed), thus achieving measurement of diffusion coefficients that is inaccessible from slower methods [63]. Temporal ICS (or TICS) calculates temporal correlation functions of spatially integrated fluctuations and is used to measure rotational and translational diffusion as well as number densities [64]. Spatio-temporal image correlation spectroscopy (STICS) analyzes lags across time and spatial dimensions, rendering it a very effective tool for measuring diffusion and velocity vectors.

The ICS variants presented thus far use a combination of spatial and temporal correlation analyses to extract system parameters. It is apparent that the insight gained from these methods is highly dependent on how the correlation is performed and should match the acquisition method of the experiment and physical parameters of the system. With this in mind, it is appropriate to introduce a final ICS variant: k-space ICS, or kICS. Rather than using spatial intensity fluctuations to infer physical rates, k-space ICS considers the

reciprocal space intensity of the image, which is the input for temporal correlation. This technique is useful for a several reasons. First of all, reciprocal space correlation functions can be analyzed to separate the dependence on the point spread function (PSF), the photophysical term, and the number of emitters. Secondly, the system of interest is a free solution experiment that is isotropic in nature; meaning that only the magnitude of the reciprocal space vector is important rather than the 'x' and 'y' components. Finally, the resulting analytical function for diffusion-based experiments can be analyzed to measure the characteristic diffusion rate. In this way, it is easier to understand the effect that various parameters have on the characteristic rate and could give insight into higher order corrections to the analytical expression. The mathematical justification for these claims is shown in the next chapter.

1.4 Contribution of Original Material

The methodology of using kICS correlation functions to analyze CLiC microscopy experiments is a very promising for analyzing biomolecules in free solution. However, given the novelty of both techniques, the methodology is still relatively underdeveloped. It is important to explore the subtle nuances of the system in order to identify sources of systematic bias in the measurement capabilities.

The original contributions to this thesis are presented in Chapter 3, and explore the single-dimensional relationship between various experimental parameters and the characteristic decay rate of the correlation function. By exploring how artifacts manifest in correlations, experimentalists can understand appropriate boundaries in parameter space and can apply higher order correction terms to account for this error. As such, this work is of value to the Leslie-lab and other groups who are adopting CLiC for free solution experiments.

The results in Chapter 3 are dependent on the theoretical framework that is laid out

in Chapter 2. Section 2.1.1 presents a methodology for accurately modeling free solution experiments. This is based on physical processes such as Brownian motion and imaging acquisition properties of the experimental instrumentation. The second section presents a theoretical framework for the system that is used as a basis for extracting the system's characteristic rate. This framework is used to guide and predict the most useful regions to interpret data and to identify sources of correlated biases in normalization.

The objective of this chapter was to understand the optimal way to measure biomolecular dynamics with sub-cellular resolution. As such, the chapter was broken into two distinct sections; one devoted to different microscopy platforms and one devoted to analysis techniques. The capabilities and limitations of modern-day biological microscopy are identified for optical and non-optical techniques in order to understand which method is most appropriate for the application. A relatively novel approach, CLiC, is also identified within this context, and is shown to be most appropriate for studying free-solution dynamics. The second section was devoted to the family of correlation-based analysis techniques as they have become the standard for dynamic system analysis. Variants of correlation techniques have been shown to be application specific, as characteristic rates manifest themselves differently based on how data is analyzed. For studying free diffusion with CLiC microscopy, k-space ICS was identified as the most optimal solution due to the isotropic nature of the system, the ease of extracting the population diffusion rate, and the independence of photophysical and transport terms.

Chapter 2

Guiding Simulations of Fluorescence Images Acquired with CLiC

2.1 Introduction

Simulations provide a relatively cheap, time efficient, and robust alternative to single molecule microscopy experiments. Unlike physical experiments, the simulated environment is very easily controlled and therefore provides a reliable method for capturing statistical representation of movies with modeled experimental parameters. Simulations are a useful tool in guiding experimental analysis, not only because the experimental parameter space can be easily explored, but also because non-idealities found in the experiment can be simulated and its corresponding effect on the analysis can be quantified. In order to gain confidence in the aforementioned analytical models, the theory was tested using computer modeling in MATLAB. The following sections provide an overview of the methods and results of the simulation investigations.

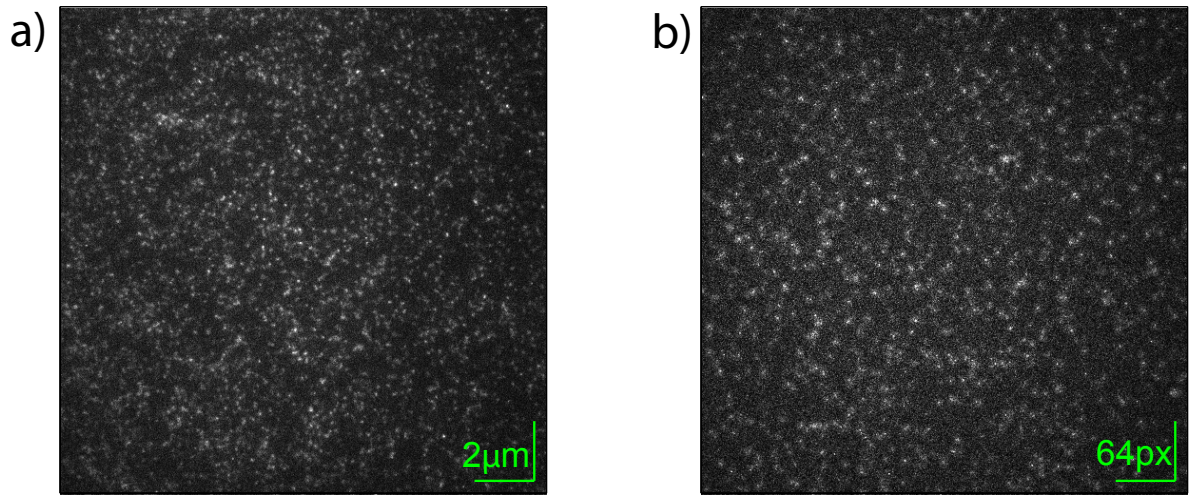

Figure 2.1: **This side by side comparison shows that the experimental environment can be accurately replicated in simulations. a)** A single frame snapshot of a solution of fluorescently tagged beads (size = 50 nm in diameter, excitation wavelength = 540 nm; Fluospheres from Life Technologies, lot number: 1241418). The beads were chosen to be large and bright to increase the signal-to-noise ratio. D = 9.6 μm^2/s, C = 8.2x10^{11}/mL. **b)** A simulated representation of a) with parameters chosen to match the experimental conditions. D = 2.7 pixel^2/frame, C = 0.015 pixel^{-2}.

2.1.1 Simulation Setup and Overview

While there are many benefits of performing a simulation analysis, the utility of the final result is dependent on how accurately the simulations replicate the physical processes. A sample side by side comparison of an experimental and simulated image is shown in Figure 2.1, while the distribution of intensities are shown in Figure 2.2.

In figure 2.1, both frames contain clusters of high intensity pixels showing emitter locations, and is contrasted against a low intensity but noisy background based on the time-integrated motion of the particles and by shot noise of the camera. In figure 2.2, the distribution of pixel intensities across these frames display very similar properties in terms of the first and second order moments of the distributions. Furthermore, fluores-

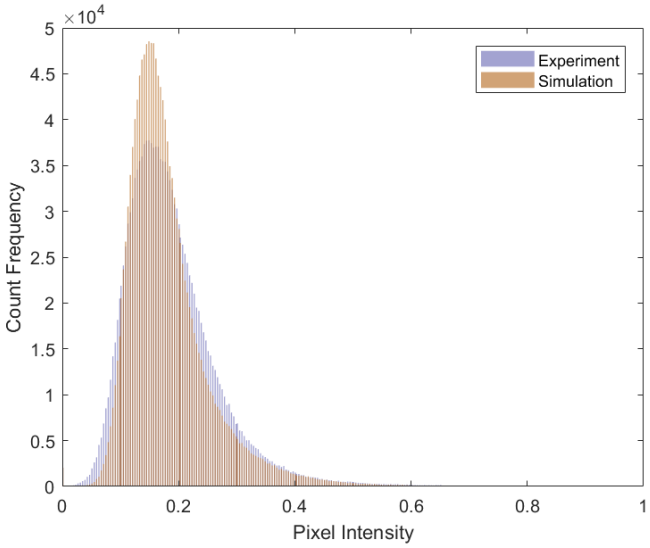

Figure 2.2: **A histogram showing the distribution of point intensities for the frames in Figure 2.1.** While the simulated output in this case has a slightly higher background noise contribution, it is evident that the experimental moments are captured in the simulation.

cent beads have been widely demonstrated to be accurately modeled by pure Brownian diffusion [65]. Given the similarities in the spatial domain, and the well documented accuracies in the temporal domain modeling, simulation results of this nature can be used as a guiding tool for experiments.

The computer modeling for this chapter is done entirely in MATLAB. Simulations for a specific combination of parameters are performed for multiple (approximately 5) replicates, although each individual movie is completely unique of any other replicate. As such, the statistical variability between simulation replicates is inherently dependent on MATLAB's uniform pseudo-random number generator. By averaging between replicate results, a good estimate of the measured diffusion coefficient and the corresponding error can be made.

In order to provide the most accurate results, a large list of system parameters were considered when designing the simulation structure. While changing each parameter would influence the resulting data in some way, only a few of these parameters (such

as concentration of emitters, acquisition time length, relative diffusion coefficient, and effective signal to noise) are discussed in the following sections.

Simulation of the Fluorescent Sample

For this system, a certain concentration of point emitters representing fluorescently tagged molecules undergo changes in space and time according to their physical attributes and acquisition parameters. Without considering the influence of the optical and electronic imaging systems, this system can be modeled as many point sources changing position and signal intensity over time as a result of Brownian motion, chemical interactions, and photophysical influence. Note that common fluorophores are generally on the order of nanometres in size; green fluorescent protein (GFP) is approximately 2.5 nm in diameter and 4 nm tall [38], while cyanine based dyes (composed of a polymethine bridge between two nitrogen atoms) are even smaller. Conversely, the spatial resolution of modern EM-CCD cameras with 100x objective is on the order of several hundred nanometres [66], meaning that the point source model is justified. In the case where multiple fluorophores are attached to a long polymer or to a large polystyrene bead, this simplification is no longer valid. In the former case, the polymer is modeled as many point sources with a displacement constraint that is dependent on the elasticity of the polymer. In the latter case, the bead is modeled as a very large point source occupying many pixels. In any and all cases, exact overlap of position is forbidden and is another displacement constraint given to the particles.

For a given set of experimental parameters, an empty two dimensional grid (corresponding to the size in pixels of the camera used for experimental data acquisition) is populated with particles. The initial positions of these particles are independent of any previous position and of each other, and therefore obey a uniform distribution with the previously defined position constraints. At this point, the position of each individual particle is dependent on its position and state in the previous frame, as well as the pre-

defined governing rates. The diffusive motion of a particle between adjacent frames is simulated as a a zero mean Gaussian random variable with spatial variance of twice the diffusion coefficient. The fluorophore state follows an exponential distribution based on the previous state of the fluorophore and the corresponding blinking and photobleaching rates (i.e. the Gillespie algorithm) [67]. The chemical reactions are similarly simulated by a Gillespie algorithm with the additional proximity condition.

Finally, boundary conditions must be considered. For the CLiC set-up, particles are free to diffuse out of the observation volume in the lateral directions, but not in the axial direction due to the flow cell confinement (refer to Figure 1.2). The confinement in the axial direction changes quadratically as a function of lateral position, much like the true geometry of the CLiC flow cell [51]. The lateral diffusion out of the observation volume can be properly modeled by purposely making the two dimensional grid larger than the spatial size (in pixels) that the camera can capture. This larger grid effectively serves as the entire flow cell, and from this the smaller observation grid can be selected out. The accuracy of this method is dependent on the size of this larger grid relative to the observation grid. While accuracy increases as the area of the effective flow cell grid increases, the computational efficiency decreases significantly since only a portion of the simulated data gets used. Other factors that influence the required size of the larger grid are the particle concentration and particle diffusion coefficients.

Simulation of the Imaging System

At this point in the simulation, the particles' trajectories have been calculated and are stored in a multidimensional array of data. In order to produce a single image, the position of every point source at that time frame is assigned to its pixel coordinate, and is then convolved with a Gaussian kernel. This simulates the point spread function (PSF) of the microscope, which is generally assumed to be a Gaussian profile. By the nature of MATLAB, this convolution is discrete, meaning that spatially binning the intensities into

pixels is inherent to this process. Since convolution is an linear operation, particles can be convolved one by one and be individually added to the image.

In order to accurately simulate the time integration property of camera data acquisition, each frame is calculated as the summation of many discrete 'sub-frames'. Each sub-frame can be considered a representation of the sample for a very small time range, and exhibiting a fraction of the overall brightness and particle diffusivity of one frame. In this way, the motion blurring exhibited at higher exposure times can be emulated. As with real movies, both the total brightness and the visible path of the fluorophore increase with exposure time.

It should be noted that time integration is a continuous property and is being modeled in simulations as a discrete process. For these simulations, each sub-frame corresponds to a time interval of 0.5 ms. This is considered a good approximation, as the upper limit of modern EMCCD camera's acquisition rate for via over-clocking and region cropping is approximately 1000 frames per second [66] for region sizes on the order of 128x128 pixels. Although higher rates up to 10 MHz can be achieved for smaller regions (on the order of 1000 pixels or a 32x32 grid), the limited spatial resolution corresponding to these rates render ICS an impractical tool, and are thus not considered in the scope of this thesis. As with real experiments, the fluorescent molecules still behave dynamically during the readout time when the camera does not capture any photons. This is modeled in simulations by calculating these new positions without recording the corresponding sub-frames as part of the frame summation. The length of this dead-time was evaluated at 0.367 ms per frame, and was calculated based on fitting the acquisition rates as a function of region size according to the published specifications.

The influence of EMCCD cameras in experiments both amplifies the intensity signal and introduces a source of noise to the resulting image. The noise contribution is well understood and is modeled in the simulations in three stages, as shown in 2.3. The stages are as follows:

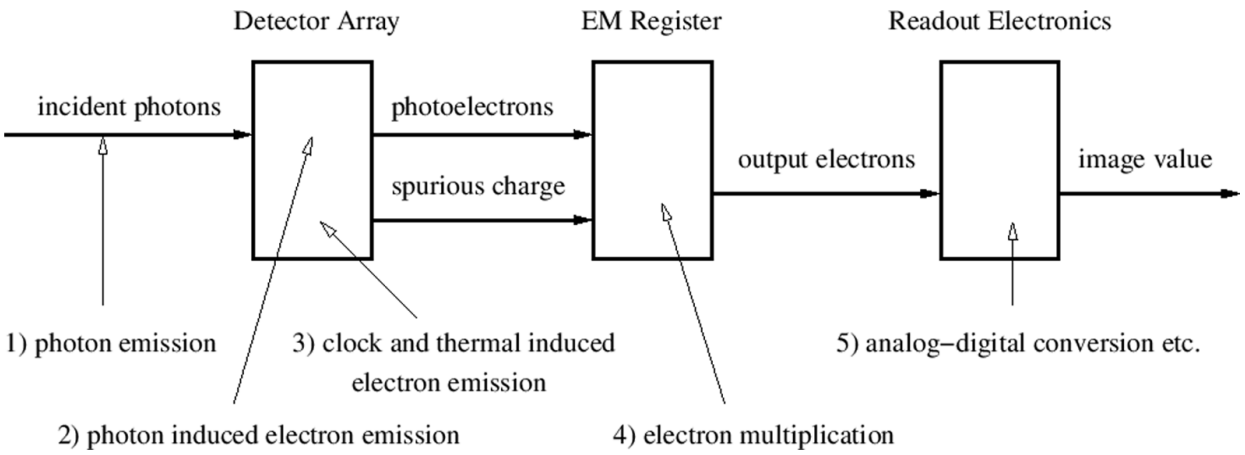

Figure 2.3: **A flow diagram demonstrating the various stages of image processing in an EMCCD camera, and the corresponding noise introduced at each stage. [68]**

- **The Detector:** photons from the fluorescent sample create a photocurrent in the register, which is dependent on the input photon intensity signal, quantum efficiencies, exposure time, and the baseline currents of the camera. The number of electrons detected at the register is governed by a Poisson distribution.

- **The Camera Register:** The output signal from the detector array (number of electrons) is amplified according to a Gamma distribution with scale parameter determined by the gain setting of the camera.

- **The Readout Analogue to Digital Conversion:** the error due to the read-out discretization of the signal can be modeled by a zero-mean additive Gaussian random variable.

This effectively introduces an intensity probability distribution to the output as a function of true position. In order to understand how this process influences the correlation analysis, it is first important to dissect the probability introduced at each stage.

The final stage creates a zero-mean additive Gaussian random variable to each pixel. Due to the linearity property of Fourier transforms and convolution operators, this term can be analyzed independently from the previous noise contributions. Since AWGN is

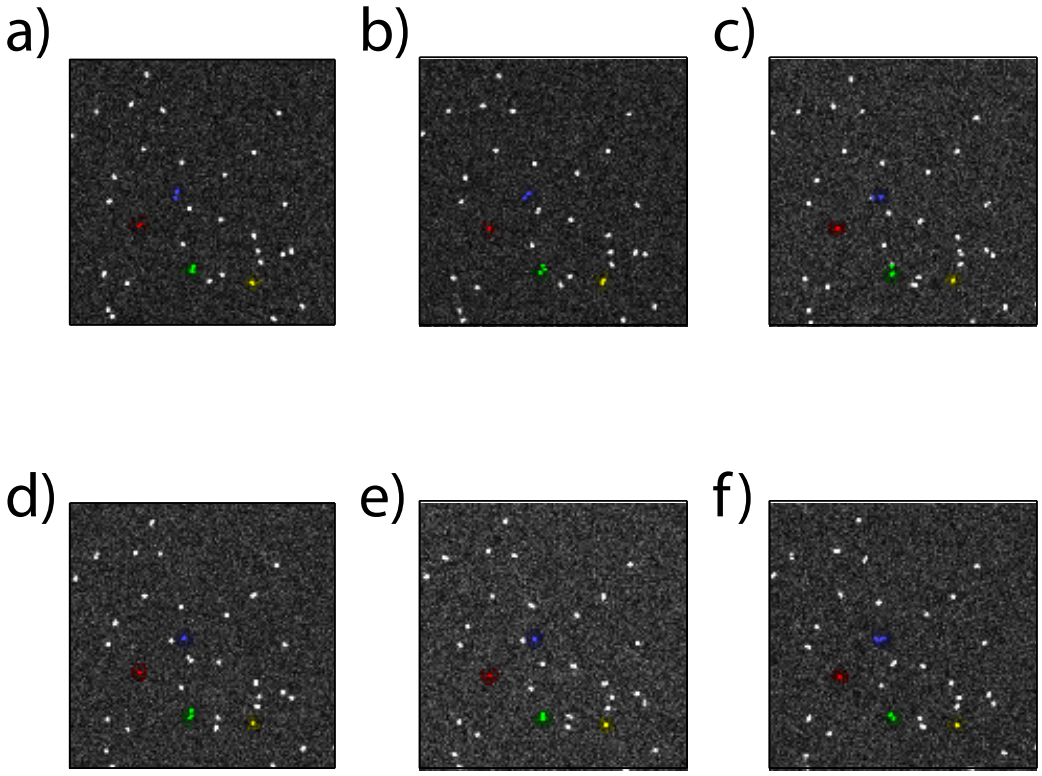

Figure 2.4: **This figure demonstrates the end result of a simulation or an experiment; fluorescently tagged particles evolving over time. It is from here that the kICS analysis begins.** The evolution of bead positions over time in a simulation is governed by the thermal fluctuations of the solution. Certain beads have been colour coded so that they can be tracked between time frames. Each image corresponds to a successive time frame, ranging from **a)** t = 0.02 s to **f)** t = 0.12 s. (D = 0.5 pixel^2/frame; t_{exp} = 0.02 s; T = 4096 frames; C = 1.95x10^{-3} pixel^{-2}; 1 pixel = 266.7 nm; 1 frame = 0.025 s).

by definition uncorrelated noise at non zero lags, this will have no influence on the final correlation expression. As such, only the first two stages need to be considered.

The first stage generates a random variable according to a Poisson distribution. Poisson distributions have the properties that the mean and variance are both equal to the rate parameter, which is the (input) photon count. The second stage turns the random variable from the first stage and creates a second random variable with a gamma distribution. The mean of this second random variable is the product of the camera gain and the value of the first stage random variable. As such, the mean value of the output is simply the input photon count multiplied by the camera gain. On average, there is no systematic bias due to the EMCCD camera.

However, the variance of this second random variable scales according to the gain parameter squared, and scales linearly with respect to photon count. As such, the influence of these stages to the total correlation function is more significant at high values of gain and exposure time. While this is still a contribution of error, there is no systematic effect to this contribution. Averaging over many simulations will result in a regression to the true value.

An example simulation of pure diffusion is shown in Figures 2.4 and 2.5. The trajectories of four particles can be mapped out over the time frames, thereby demonstrating the influence of the fluorescent sample simulation. Furthermore, the resulting intensities, spatial blurring, and background noise contributions demonstrate the influence of the imaging system simulation.

2.1.2 Correlation Analysis of the Simulations

For a given time series of images, the experimental parameters of interest are tied intrinsically to the movie, but may not be readily apparent. Correlation analysis allows these parameters to be extracted from noisy data.

In essence, correlations are a measure of similarity between signals, and are used both

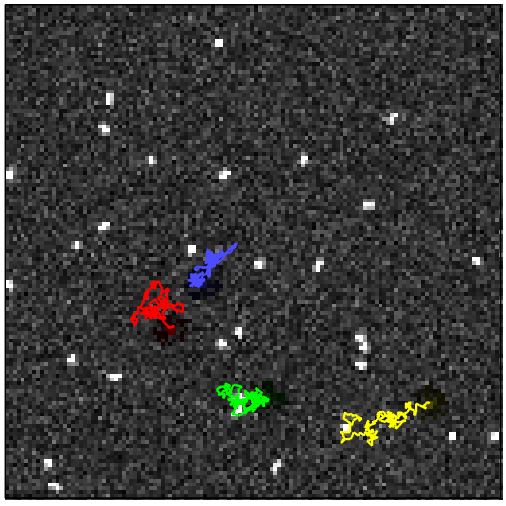

Figure 2.5: **The stochastic motion of the individual particles evolves over time according to the beads' diffusion coefficient.** The simulated Brownian movement of fluorescently tagged particles from the simulation in figure 2.4 is mapped out between frames, and over the course of 37 time steps (corresponding to the first 2000 ms of the movie). Part of the simulated movie can be seen at: <https://tinyurl.com/jtakkc9> (D = 0.5 pixel^2/frame; t_{exp} = 0.02 s; T = 4096 frames; C = 1.95x10^{-3} pixel^{-2}; 1 pixel = 266.7 nm; 1 frame = 0.025 s)

to analyze patterns at various length scales and for filtering noisy data. This type of signal processing technique is ideal for the experimental regime of fluorescence microscopy, as the signal to noise ratio is often poor, and the length scales (both in time and size) of the correlation curves correspond to different molecular dynamic and kinetic rates.

Before the mathematical model of this process is derived, some important physical considerations should be made:

- The size of a single fluorescent tag is approximately 3 nm in size. Given that the lower limit of the diffraction size in optical microscopy is roughly 200 nm, it is safe to model the biomolecules of interest as point sources. Thus, we can model a series of images containing 'N' particles as a summation of delta functions in both space and time.

- The optical point spread function of the microscope introduces a blur to the delta functions. Since spatial convolution is a linear operator, this can be expressed as a sum of 'N' independent convolutions.

- Although the experiment is analogue in nature, the acquisition of experimental data (and the simulated counterpart) is discrete in both time and space. Photon emissions are captured by an EMCCD detector array and are binned spatially, resulting in the appearance of discrete pixel intensities.

- Furthermore, the exposure time of the camera increases the signal to noise ratio; the rate of photon emission per tagged molecule is of the order of 100,000 per second, which is significantly more than the shot noise produced by the camera electronics.

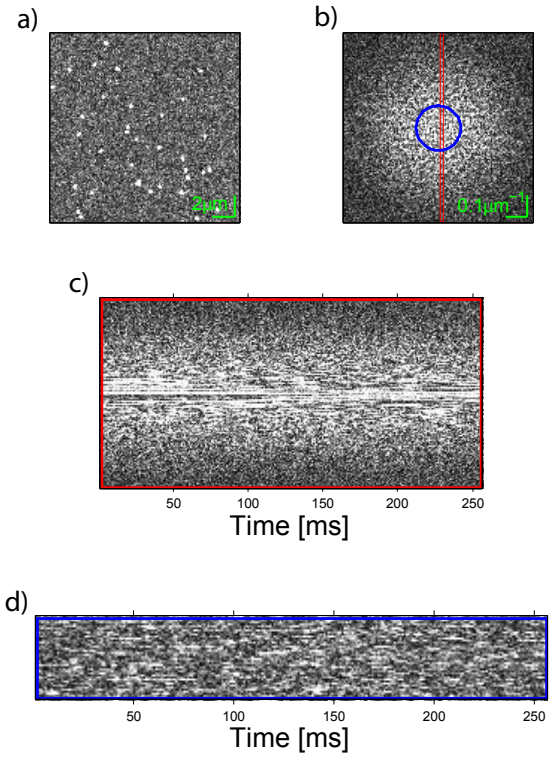

Figure 2.6: **A detailed breakdown of how an image gets transformed into k-space and some notable features of the k-space curves as a function of time. a)** A single 128x128 pixel image from the bead simulation movie shown in figure 2.4. The larger spots correspond to simulated particles, while the small regions of high intensity correspond to the noisy background of the image. **b)** The corresponding spatial Fourier Transform of the sample image in part a), demonstrating the magnitudes of the reciprocal spatial vectors. Note that the reciprocal spatial vectors decay exponentially from the centre of the image (i.e. the zero frequency component). The red column shows a cross section of the vector intensities for a given k_x value ($k_x = \frac{\pi}{64}$), while the blue ring corresponds to all reciprocal spatial vectors with length 0.1 μm^{-1} . **c)** A kymograph of the red column in part b) for the first 256 ms of the movie. Note that the centre portion (corresponding to the low spatial frequencies) have a higher magnitude and longer streaks of similar values. **d**) A kymograph of the blue ring from figure b) taken for the first 256 ms. Note that all values for the same radius have similar magnitude and streak patterns. The corresponding movie to this reciprocal space transform is available at: <https://tinyurl.com/zphn8p3>.

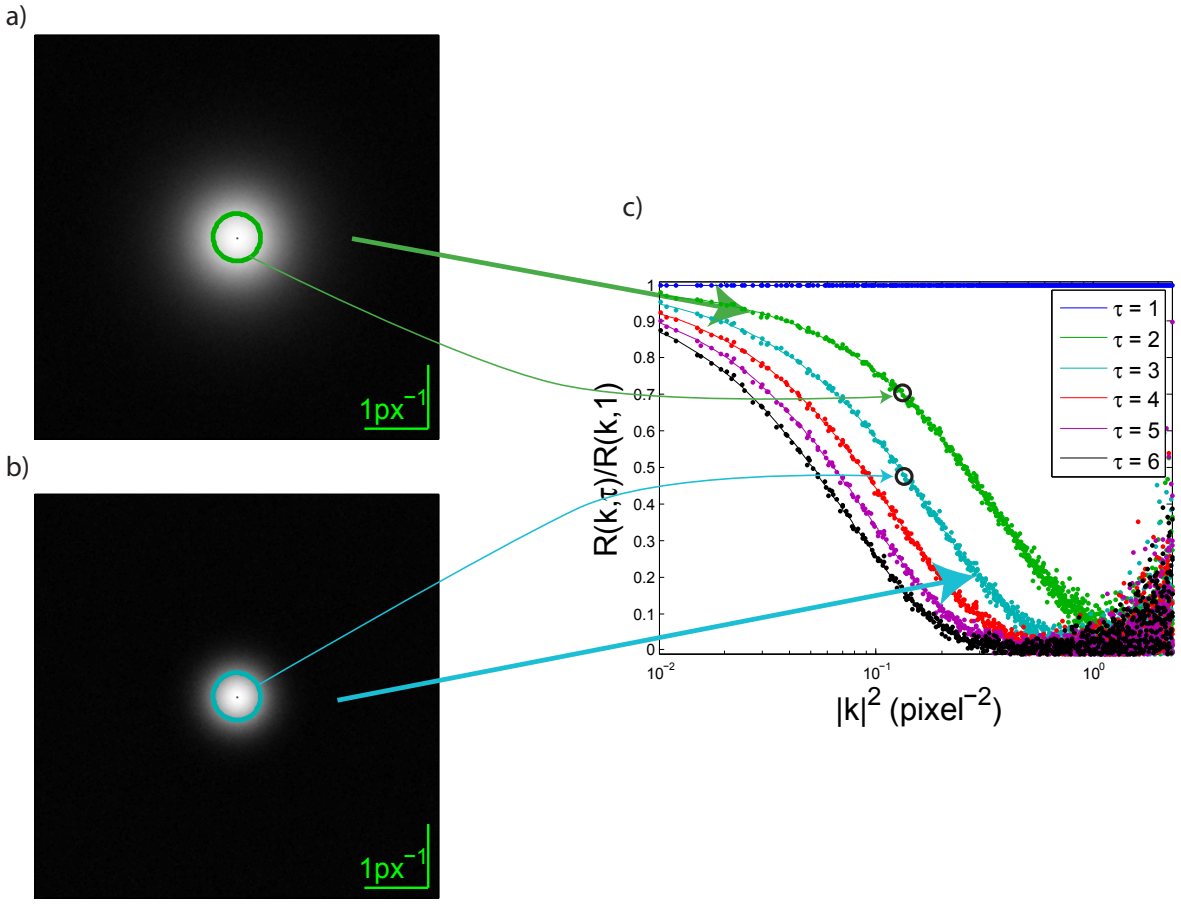

Figure 2.7: **This figure demonstrates how the kICS curves relate to the correlation slices of the movie from figure 2.4**. The different exponential curves in sub-figure c) correspond to individual lag decays of the correlation function (i.e. a single frame from the correlation movie). The same simulation from figure 2.4 is used. **a)**) The correlation slice for $\tau = 2$. **b)** The correlation slice for $\tau = 3$. **c)** The corresponding correlation plots shows how each lag value corresponds to each individual curve vs k^2. The correlation movie for this simulation can be found at: <https://tinyurl.com/jxk76pa>

2.2 Mathematical Derivation of the Fluorescence Correlation Analysis

Fluorescence correlation spectroscopy provides a useful tool to probe intermolecular processes. By studying the change in intensity as a function of space and time, it is possible to quantitatively extract information about the system's governing coefficients using correlation analysis. The original derivation for reciprocal correlation was done by Kolin et al. in 2006 [64]. For a series of fluorescent images, one can represent the intensity as a function of space and time according to the following equation:

$$i(\vec{r}, t) = I(\vec{r}) \circledast_r \rho(\vec{r}, t) \tag{2.1}$$

where $\vec{r}$ is the variable denoting spatial position, t is the variable corresponding to time, i is the frame intensity function, I is the point spread function of the imaging system, ρ is the effective particle density, and $\circledast_r$ is the convolution operator with respect to the variable $\vec{r}$. This expression shows that the overall intensity profile of an image is given by the spatial convolution of the point spread function of the microscope set-up and the particle density of sample. Intuitively, this make sense; the point spread function is the system's response of a perfect point source, and can be applied to the entire particle density since convolution is a linear operator.

As stated in the assumptions, the effective particle density can be modeled as a sum of individual delta particles. For a sample containing N particles, the corresponding expression is:

$$\rho(\vec{r}, t) = \sum_{n=1}^{N} q_n \theta_n(t) \delta_n(\vec{r} - \vec{r}_n(t)) \tag{2.2}$$

where q_n is the yield efficiency of the n^{th} fluorophore, δ is the Dirac delta-function which by definition corresponds to the particle density of a single particle, and θ is the photophysical term of the particle, defined as:

$$\theta_n(t) = \begin{cases} 1 & \text{particle } n \text{ is fluorescing at time } t \\ 0 & \text{otherwise} \end{cases}$$

It often becomes more convenient to analyze equation (2.1) by converting the expression to spatial frequency space, since the convolution operator becomes a multiplication. Let us define the variable $\vec{k}$ as the spatial frequency vector such that: $\widetilde{f}(\vec{k}) = \int e^{-i\vec{k}\cdot\vec{r(t)}} f(\vec{r}) d^3\vec{r}$. By applying this definition to equation (2.1), the intensity profile can be rewritten as:

$$\widetilde{i}(\vec{k}, t) = \widetilde{I}(\vec{k}) \sum_{n=1}^{N} q_n \theta_n(t) e^{-i\vec{k}\cdot\vec{r_n(t)}} \tag{2.3}$$

Note that both the yield efficiency and the photophysical term are identical in both the spatial domain and the inverse spatial domain, since both variables are independent of position. The resulting spatial frequency intensity is demonstrated in Figure 2.6-b.

It is useful at this point to define the reciprocal space time correlation function as follows:

$$R(\vec{k}, \tau) \equiv \langle \widetilde{i^*}(\vec{k}, t) \widetilde{i}(\vec{k}, t+\tau) \rangle_t \tag{2.4}$$

Where the angular brackets signify the statistical average or the time average of the intensity signals. It should be noted that since this process is time-shift invariant for equilibrium systems, only the lag variable τ is of importance. Since each time sample is equally weighted, equation (2.4) can be rewritten as:

$$R(\vec{k}, \tau) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \left(\widetilde{I^*}(\vec{k}, t) \sum_{n=1}^{N} q_n \theta_n(t) e^{i\vec{k}\cdot\vec{r_n}(t)} \right) \left(\widetilde{I}(\vec{k}, t+\tau) \sum_{m=1}^{N} q_m \theta_m(t+\tau) e^{-i\vec{k}\cdot\vec{r_m}(t+\tau)} \right)$$

where T corresponds to the total number of frames in the series. The discrete summation appears from the fact that time samples are collected by taking discrete snap shots of the sample in time. As such, t can be rewritten as a discrete index corresponding to the frame

number.

Assuming that the incident laser intensity does not change with respect to time, and that the quantum efficiency of all particles is equal, the expression can be rewritten as:

$$R(\vec{k},\tau) = \frac{q^2|\widetilde{I}(\vec{k})|^2}{T-\tau}\sum_{t=1}^{T-\tau}\sum_{n=1}^{N}\sum_{m=1}^{N}\theta_n(t)\theta_m(t+\tau)e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_m}(t+\tau)\right)}$$

The expression can be simplified if we consider the regime of very dilute concentrations. In this case, the contribution of unlike particles ($m \neq n$) to the sum is on average zero since they do not correlate. Note that for movie time lengths on the order of many diffusion times, it is possible for unlike particles to correlate. However, the contribution to the sum is insignificant in comparison to the contribution of like particles, because there is no characteristic time constant for which unlike particles will correlate at the same inverse spatial vector. For interacting species, this is not the case, and the expression must be reconsidered based on interaction rates between species.

The expression for the correlation can thus be simplified to a single sum over N particles:

$$R(\vec{k},\tau) = \frac{q^2|\widetilde{I}(\vec{k})|^2}{T-\tau}\sum_{t=1}^{T-\tau}\sum_{n=1}^{N}\theta_n(t)\theta_n(t+\tau)e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_n}(t+\tau)\right)}$$

For experiments of this nature, it is safe to assume that all N particles behave similarly in nature; all trajectories will evolve similarly over time, and all photophysical efficiency terms are similar between molecules. Much like Kolin et al. we will assume that the photophysical term is independent of the fluorophore's position, and thus can consider each correlation independently. Rewriting the terms:

$$\left\langle R(\vec{k},\tau)\right\rangle = \frac{Nq^2|\widetilde{I}(\vec{k})|^2}{T-\tau}\sum_{t=1}^{T-\tau}\left(e^{i\vec{k}\cdot\left(\vec{r}(t)-\vec{r}(t+\tau)\right)}\right)\left(\theta(t)\theta(t+\tau)\right) \tag{2.5}$$

The key to solving equation (2.5) lies in the sum over the time frames.

Starting with the exponential term:

$$\langle e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_n}(t+\tau)\right)}\rangle_t = \sum_{t=1}^{T-\tau} e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_n}(t+\tau)\right)} P(\vec{r_n}(t),\vec{r_n}(t+\tau))$$

where the function P denotes the probability of a particle n to occupy a specific position $\vec{r}$ at a specific time step t. Using conditional probability, the probability density function can be written as:

$$P\big(\vec{r_n}(t),\vec{r_n}(t+\tau)\big) = P\big(\vec{r_n}(t)\big)P\big(\vec{r_n}(t+\tau)|\vec{r_n}(t)\big)$$

The probability of a given particle to be at a certain position is equal for all positions. As such, the probability density function of a single particle is

$$P\big(\vec{r_n}(t)\big) = \frac{1}{V}$$

where V is the observation volume.

The second conditional probability term is the probability of the particle to be in a certain position after a certain amount of time τ passes. This probability can be expressed as

$$P(\vec{r_n}(t)|\vec{r_n}(t+\tau)) = \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} e^{\frac{-|\vec{r_n}(t+\tau)-\vec{r_n}(t)|^2}{4D\tau}} \tag{2.6}$$

which is exactly the solution to the diffusion equation:

$$\frac{\partial P(\vec{r_n}(t)|\vec{r_n}(t+\tau))}{\partial\tau} = D\nabla^2 P(\vec{r_n}(t)|\vec{r_n}(t+\tau))$$

The statistical average for the exponential term in equation (2.5) can thus be rewritten as:

$$\langle e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_n}(t+\tau)\right)}\rangle_t = \frac{1}{V(4\pi D\tau)^{\frac{3}{2}}} \int_V d^3\vec{r} \int_V d^3\vec{r'} e^{\frac{-|\vec{r'}(t+\tau)-\vec{r}(t)|^2}{4D\tau}} e^{i\vec{k}\cdot\left(\vec{r}(t)-\vec{r'}(t+\tau)\right)}$$

This expression can be solved analytically in the limit of $V \to \infty$. Using a change of

variables for the vector difference $\vec{r'}(t+\tau) - \vec{r}(t)$, the integral becomes a Fourier transform of a Gaussian function, and evaluates to:

$$\langle e^{i\vec{k}\cdot\left(\vec{r_n}(t)-\vec{r_n}(t+\tau)\right)}\rangle_t = e^{-k^2D\tau}$$

As expected, the statistical average is time-shift invariant and only depends on the lag time between time samplings.

Substituting this back into equation (2.5), the final form for the correlation curve for simple diffusion can be written as:

$$\langle R(\vec{k},\tau)\rangle = \frac{Nq^2|\widetilde{I}(\vec{k})|^2}{T-\tau}e^{-k^2D\tau}\sum_{t=1}^{T-\tau}\theta(t)\theta(t+\tau) \tag{2.7}$$

In this form, the correlation can easily be fitted for the parameters of interest. In the case of no blinking, $\theta_n(t) = 1$ for all t, meaning that the correlation function simplifies to:

$$\langle R(\vec{k},\tau)\rangle = Nq^2|\widetilde{I}(\vec{k})|^2e^{-k^2D\tau}$$

2.3 A Note On Normalization and Fitting

When fluorescence correlation spectroscopy was first introduced, a single photocurrent intensity profile was captured recording the presence of fluorescently labeled molecules in the observation volume. By autocorrelating this signal in time, chemical rates and diffusion coefficients could be extracted [57]. The improvement of wide-field imaging techniques such as TIRFM and LSM has allowed fluorescence to be tracked spatio-temporally, thereby giving a better understanding of the underlying physics in the system.

The major benefit in transforming the intensity profile to reciprocal space is in the decoupling of the photophysical term, the imaging parameters, and the system kinetic parameters. Taking the negative logarithm of both sides in equation (2.7), the resulting

plot of the correlation logarithm as a function of time lag and spatial frequency modulus square results in a linear plot with a slope equal to the diffusion coefficient for a single species and an intercept given by the imaging parameters (as shown in Figure 2.8).

A common strategy in analyzing the correlation function involves normalizing the data so that the resulting expression simplifies to a function of only chemical rates and diffusion coefficients. By dividing the general expression in equation (2.7) with the correlation function of a single time lag, only the kinetic parameters of the system survive, as shown below:

$$\frac{\langle R(\vec{k},\tau)\rangle}{\langle R(\vec{k}),1\rangle} = e^{-k^2 D(\tau-1)} \tag{2.8}$$

While in theory any characteristic time lag can be chosen, it is generally preferred to choose a small time lag to get a better statistical understanding of the correlations. Conversely, the zero time lag is not chosen because the variance of independent noise sources also contributes to this term. For this reason, a time lag of a single frame (which is generally on the order of 10 ms) is the optimal choice for this method of normalization.

The main fault with this method occurs when the signal is weak and dominated by noise. In this regime, the intensities do not readily correlate and are very noisy. As such, dividing all correlation curves by a characteristic time lag may produce a higher estimation of the noise floor, as shown in Figure 2.9 at the higher values of $|k|^2$. This can be limited to a certain extent by choosing a larger binning range for this regime, but it is generally best to ignore this regime for fitting purposes. Another issue with normalizing the correlation functions this way is that the function used for normalization greatly influences all correlation curves, and can therefore significantly bias the measurement of desired parameters. If the normalization curve deviates from theory significantly, this will propagate into the correlation curves of all time lags (refer to Figure 2.9). The likelihood of a significant deviation from theory decreases with better spatial and temporal

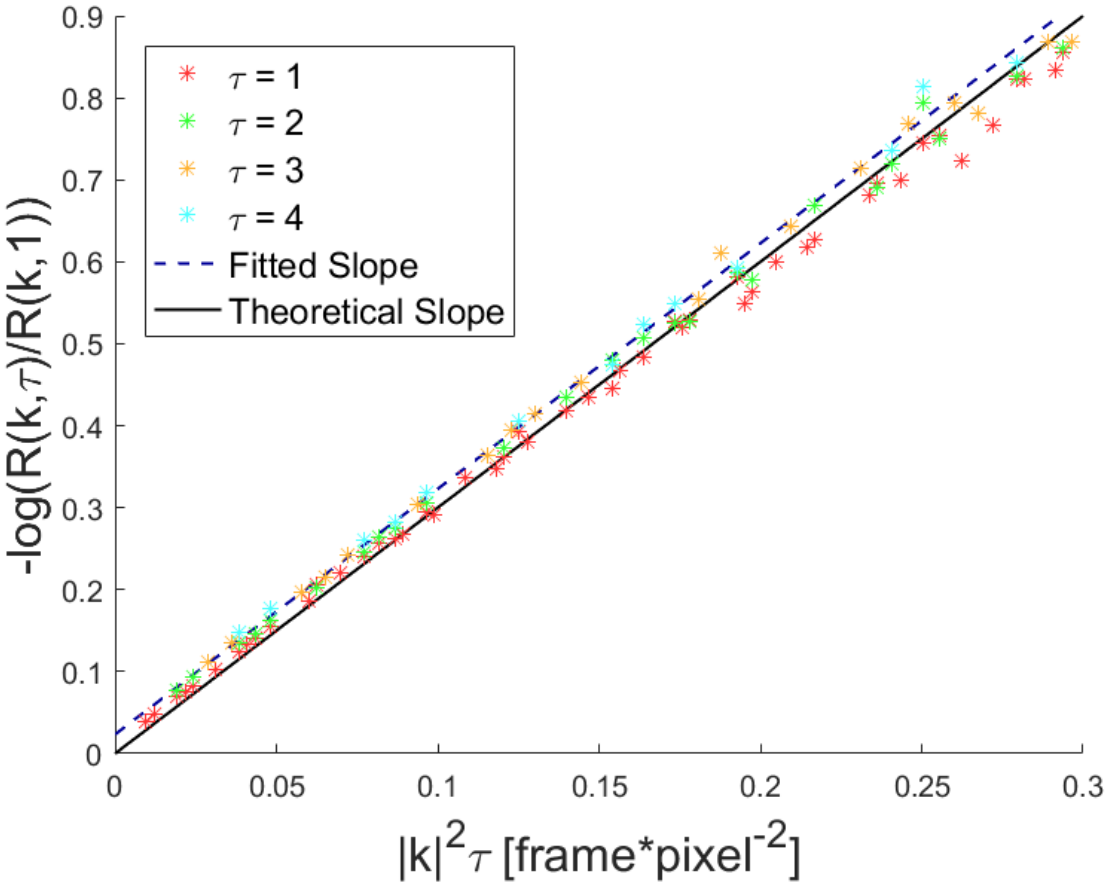

Figure 2.8: **A plot of the logarithmic normalized intensity correlation function $R(k,\tau)$, as a function of the product of τ and $|k|^2$.** In this case, the line of best fit would give a slope that is equal to the diffusion coefficient in inverse units to the x-axis. The range of the fit is chosen so that minimal noise at the higher correlation values due to large τ and $|k|^2$ values are chosen. In this figure, the diffusion coefficient was measured to be 2.9950 $pixel^2/frame$, which corresponds to an error of 1% when compared to the inputted value of 3 $pixel^2/frame$. The y-intercept offset corresponds to the dominance in noise and particle density terms at low values of $|k|^2\tau$.

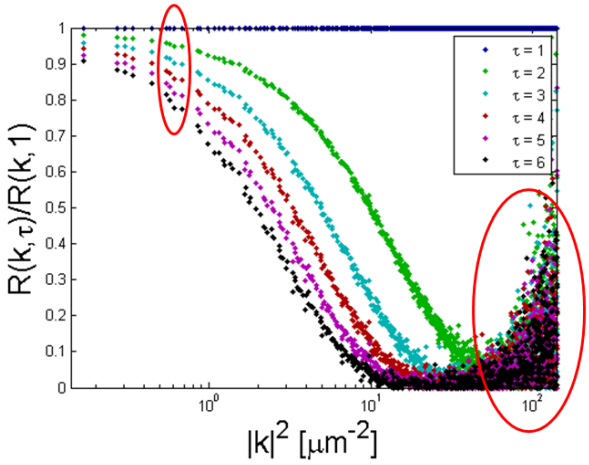

Figure 2.9: **Normalizing the correlation functions by a characteristic time lag $R(k,1)$ produces a notable bias in two ways.** Firstly, if a given point in the correlation function deviates significantly from the theoretical fit model, then this error will also appear in all correlation curves. For example, the red oval at the top-left portion of the figure shows similar behaviour in the correlation curves from $R(k,2)$ to $R(k,6)$, indicating that this behaviour likely originates from the normalization curve $R(k,1)$. The second manifestation is the over-emphasis of the noise-floor, as seen in the red oval at the bottom-right portion of the image. At high values of $|k|^2$, and τ, the signal intensities do not correlate, meaning that correlation curves at these values are random low values. As such, it is possible to get an over-estimation of the true noise-floor when dividing small random numbers with other small random numbers.

sampling, and will be investigated further in the next chapter.

Since this expression in equation (2.8) decays exponentially both with respect to the modulus squared of spatial frequency and time lag, it is equally valid to perform the normalization with respect to a characteristic spatial frequency. It should be noted, however, that for spatial resolution and acquisition rates corresponding to typical EMCCD cameras, the resolution in spatial frequency is significantly greater than the resolution in time. This can be shown in Figure 2.10 as the intensity with respect to time lag diminishes much more quickly than it does with respect to modulus squared of spatial frequency. Figure 2.10 also shows a good regime for fitting the experimental/simulated data. The lower intensity areas of the figure show the parameter regime for which the intensities do not readily correlate (the higher $|k|^2$ and τ values). As such, the fitting range is chosen to avoid this range, and typically corresponds to $|k|^2\tau < \alpha$ where 'α' is the number of characteristic decays, in units of $frames/pixel^2$. The value of this parameter is chosen to include the full decay of the correlation without including noisy data. Through trial and error, a value of $\alpha = 8$ was determined to be most optimal.

It should also be noted that the fitting region does not include very low values of time lag and reciprocal space vectors. For time lags, the correlation function corresponding to the zero time lag was not chosen to filter white noise. White noise does not correlate at any time lag except at $\tau = 0$, rendering this correlation function noisy. For reciprocal spatial vectors, very low values correspond to very high spatial separation. As such, the intensity correlations corresponding to these vectors contain information about the spatial distribution of emitters, and do not exhibit characteristic exponential decay that corresponds to the dynamic properties of the sample. This is the same reason that very large spatial vectors are typically not used in spatial correlation analysis.

The normalized expression of the correlation is truly spectacular; the function is independent of factors such as the optical point spread function, the incident laser intensity profile, and even the photophysical aspects of the fluorophores such as blinking and

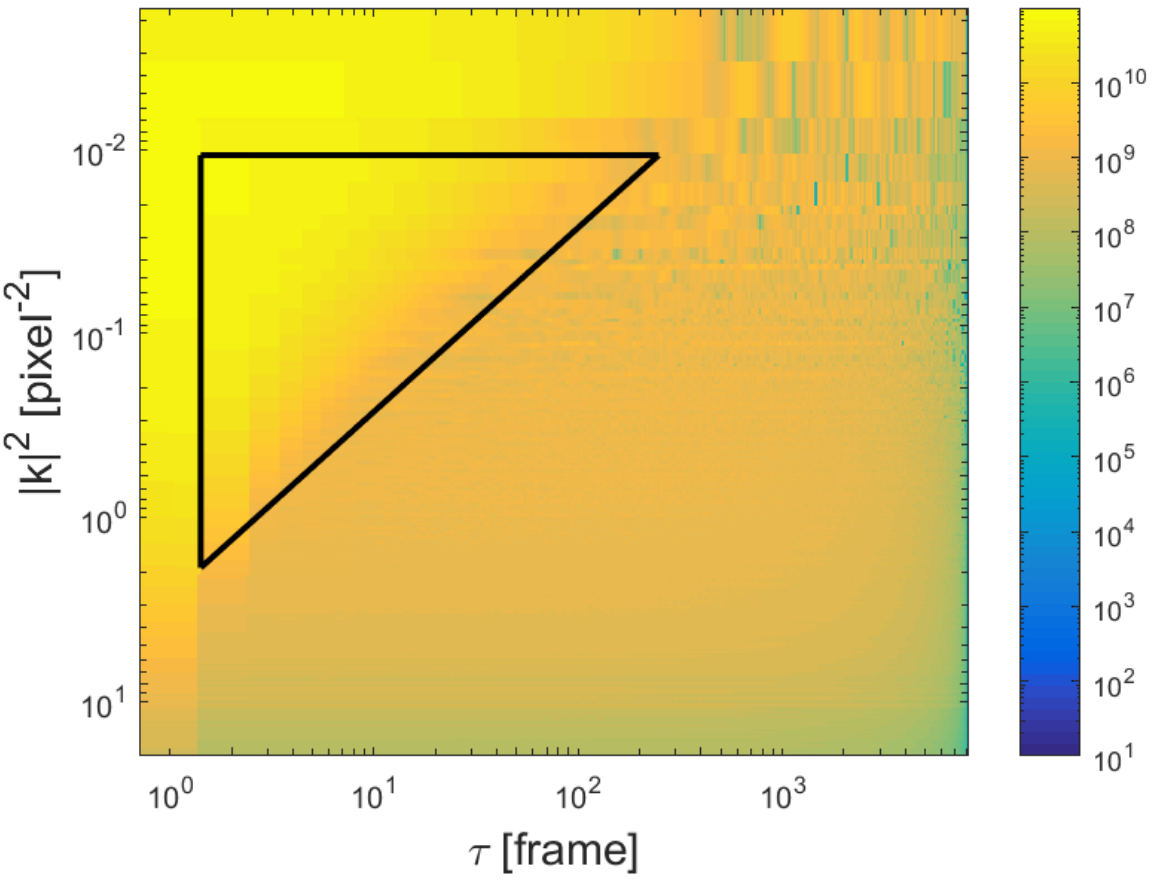

Figure 2.10: **A carpet plot of the logarithm of the unnormalized correlation as a function of temporal lag and inverse spatial vector square shows the correlation range with respect to both parameters.** As expected, the intensity decays exponentially in the temporal direction, while the inverse spatial vector dependence is also influenced on the optical parameters of the imaging system. In order to stay in the linear fitting regime and avoid the noisy portion of the correlation (beige regions), the fitting parameters were specified specifically to fit the higher magnitude regime (the bright yellow portion). Note that intensities corresponding to the $|k|^2 = 0$ vectors are not shown in order to provide contrast in the regions of interest. The black triangle corresponds to the fitting region used for a sample with D = 3 pixel^2/frame. All simulation parameters are identical to those in Figures 2.4 and 2.5.

photobleaching (as long as they are independent of space). However, it is important to understand how different methods of normalization introduce bias into parameter measurements. In this way, it is also valuable to analyze unnormalized correlation functions, as influences from non-idealities can be accurately quantified. The primary down-side of this method is that an accurate theoretical model is required to understand the properties of the correlation function, and is thus mostly useful for simulations rather than experiments. The next chapter uses unnormalized correlation functions for quantifying the bias in parameter space, with the theoretical model given by:

$$R(k,\tau) = \frac{Nq^2|\widetilde{I}(\vec{k})|^2}{T-\tau}e^{-k^2D\tau} + \beta \tag{2.9}$$

This result is based on the conclusions from the previous section, with the added term 'β' corresponding to the uncorrelated noise floor. This model readily agrees with unnormalized data, such as the carpet plot in Figure 2.10.

2.4 Concluding Remarks

This chapter outlined a simulation methodology that produces image series which closely resemble experiments of freely diffusing fluorescent molecules in solution. These simulated movies allow precise control over the system inputs and enable a thorough investigation of parameter space. A mathematical model based on reciprocal space correlations (kICS) was derived using appropriate assumptions about the freely diffusing system. This correlation technique was implemented on a simulation, and the diffusion coefficient was accurately calculated. Finally, important considerations on the noise, normalization, and fitting range of the correlation functions were made.

Chapter 3

Exploration of Parameter Space and Experimental Non-Idealities with Simulations

3.1 Introduction of Correlation Bias and Error

The previous chapter outlined a theory for measuring particle dynamics using reciprocal-space correlation functions, and provided a sample proof-of-concept simulation to regime for which an accurate result was measured. From here, it is important to consider the underlying assumptions made in the model, and evaluate the range in parameter space for which the model is useful. As such, the influence of each parameter change on the dynamics of correlation curves will be studied in terms of bias (i.e. systematic deviations from the true value) and in terms of error (the spread of the measurement distribution over many identical experiments or simulations). The accurate quantification of a parameter's influence in bias and error requires complete control over all other parameters; as such, simulations are used heavily in this chapter.

The theory of correlation analysis is based upon an idealistic model; the light intensity values are a perfect representation of the particle concentration, and can be perfectly measured for all time and space. This of course, is not the case; noise due to signal amplification, PSF blurring, pixel binning, photo-physical effects, spatial windowing and out of plane kinetics create non-idealities in the correlation curve. As such, the non-idealities are manifested as disturbances (or biases) in the curve, and influence the true nature of the measurement.

It should be noted that the bias calculated is done without any normalization, since this produces the best estimate of the pure correlation function and the errors/biases that influence them. The error bars produced in each parameter figure correspond to the standard deviation in the diffusion coefficient measurement between five simulations using identical imaging parameters. Unless otherwise specified, the conditions used in Figures 2.4 and 2.5 are used for all parameters in the bias figures.

A common theme in this parameter space exploration is the interdependence of different variables. Time can be expressed in terms of seconds, number of frames, or in correlation times (which is the amount of time required to diffuse across a focal spot). In this way, the temporal behaviour of samples is dependent on the length of the image series, the characteristic diffusion coefficient, and the exposure time. Similarly, the spatial variable is commonly expressed in terms of microns, pixels, or in size of the observation volume. As such, parameters such as the diffusion coefficient, size of the field of view, size of the Gaussian PSF, and particle density (when considering average distance between emitters) all influence correlation lengths. Finally, the overall signal intensity of the sample is determined by factors such as the particle density, yield efficiency of the fluorescent tags, camera gain, and exposure time. Because of this relative dependence it becomes easier to express quantities such as time and space in terms of relative parameters. This also makes the choice of relative units (pixels and frames rather than microns and seconds) a large benefit in terms of the versatility of analyzing real experiments (i.e.

two samples with different physical parameters may behave similarly in the context of the movie due to relative nature of parameters.).

Note that these simulations were done with the specific goal of replicating experiments. For a sample containing plasmids that diffuse at 10 $\mu m^2/s$ with 267 nm/pixel resolution at a frame rate of 47 frames per second would measure a diffusion coefficient of approximately 3 $pixel^2$/frame. While this may be an optimal point for considering the influence of individual parameters on the measured diffusion coefficient, it is not a perfect set-point for this analysis. As shown in the following sections, the pixel dwell time of a single point emitter is only $\frac{1}{6}$ of a frame, meaning that dynamics are under-sampled. While this bias is not significant (it is generally within 5% of the set diffusion coefficient and agrees within error of the set value), it is worth noting as this slight under-estimation appears consistently throughout the analysis (especially in Figure 3.3, where the systematic underestimation is most evident).

3.1.1 Bias due to the Number of Time Samples

As previously mentioned, two movies with different physical diffusers may be sampled differently in space and time such that the resulting image series are nearly identical. As such, a general relationship between bias and time can be determined in relative units, and the physical units can be applied to this relationship. To eliminate the dependence on particle speed, the units of time used in Figure 3.2 are expressed with respect to correlation time; i.e. the expected amount of time required for a particle to diffuse out of a pixel. From equation 2.6, it is clear that the expected distance traveled $\langle \Delta x \rangle$ for a particle in a given amount of time τ is

$$\langle \Delta x \rangle = \sqrt{2D\tau}$$

In the context of classic correlation spectroscopy, the correlation time was typically expressed with respect to the observation volume. For an excitation beam with e^{-2} radius

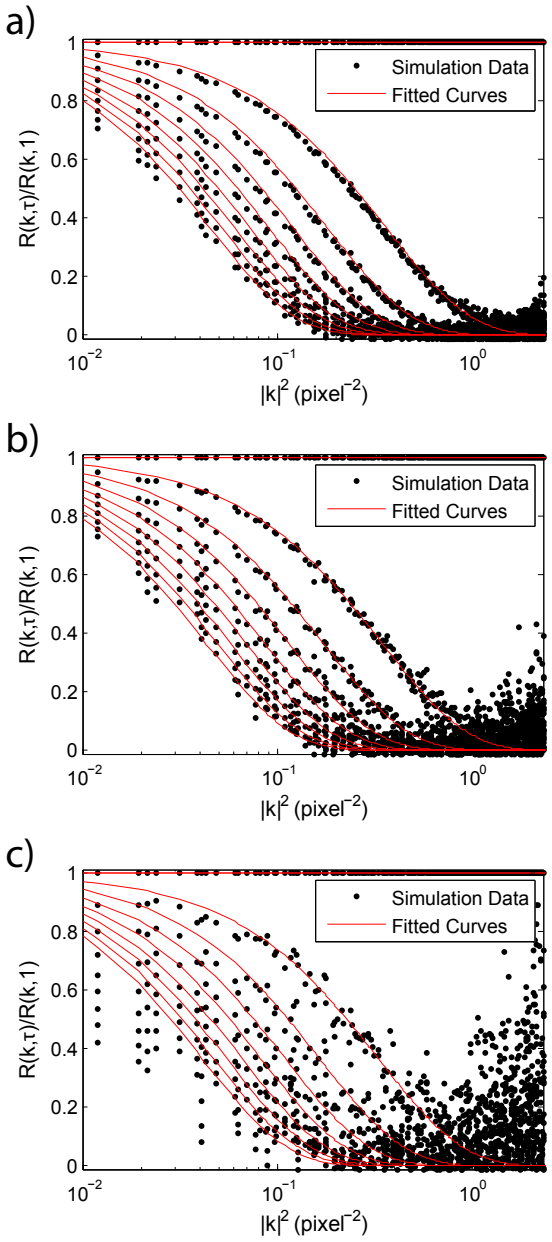

Figure 3.1: **As experimental conditions become more challenging, the correlation measurements deviate more from the theoretical curves leading to error and bias.** These curves (corresponding to τ = 1,2,3,..9) show the diffusion coefficient is measured for three different simulated movie lengths: **a)** 8192 frames; **b)** 1024 frames; **c)** 128 frames. All other simulated parameters are identical to the ones used in figure 2.4.

of ω, the characteristic time constant τ_c is given by

$$\tau_c = \frac{\omega^2}{4D}$$

However, in the context of wide-field microscopy such as CLiC, this quantity is not very helpful. Uniform illumination of the excitation beam is often preferred in order to eliminate spatial dependencies in the sample. In order to achieve this, the excitation beam often passes through beam expanders that render the radius ω very large. Because only a small subset of the beam area is captured by the camera, it renders the characteristic time constant a useless quantity. As such, a new unit of time τ_{pixel} which characterizes how long a fluorescent particle remains in a pixel area, becomes far more useful for analysis. Setting $\langle \Delta x \rangle$ to 1 pixel and solving for τ, a single pixel correlation time is given by

$$\tau_{pixel} = \frac{1}{2D}$$

Where D is in units of pixel^2/frame.

It is also important to note that since kICS is a correlation between reciprocal space vectors, the correlation curves provide insight into the change in spatial vectors between all emitters within a sample. This makes sense given the formulation given in equation 2.3, as the sum of complex exponentials will constructively and destructively interfere such that the resultant sum is a complex representation of particle separations. In this sense, it becomes easy to see that even if particles have very small pixel dwell times, their dynamics can be accurately captured by considering the spatial dimensions.

The correlation of the intensity profile for a given lag variable τ is given by equation 2.4. Since ergodicity is assumed, each correlation pair is weighted equally in the ensemble average, meaning that the equation can be rewritten as:

$$R(\vec{k}, \tau) = \frac{1}{T - \tau} \sum_{t=1}^{T-\tau} \widetilde{i}^*(\vec{k}, t) \widetilde{i}(\vec{k}, t + \tau)$$

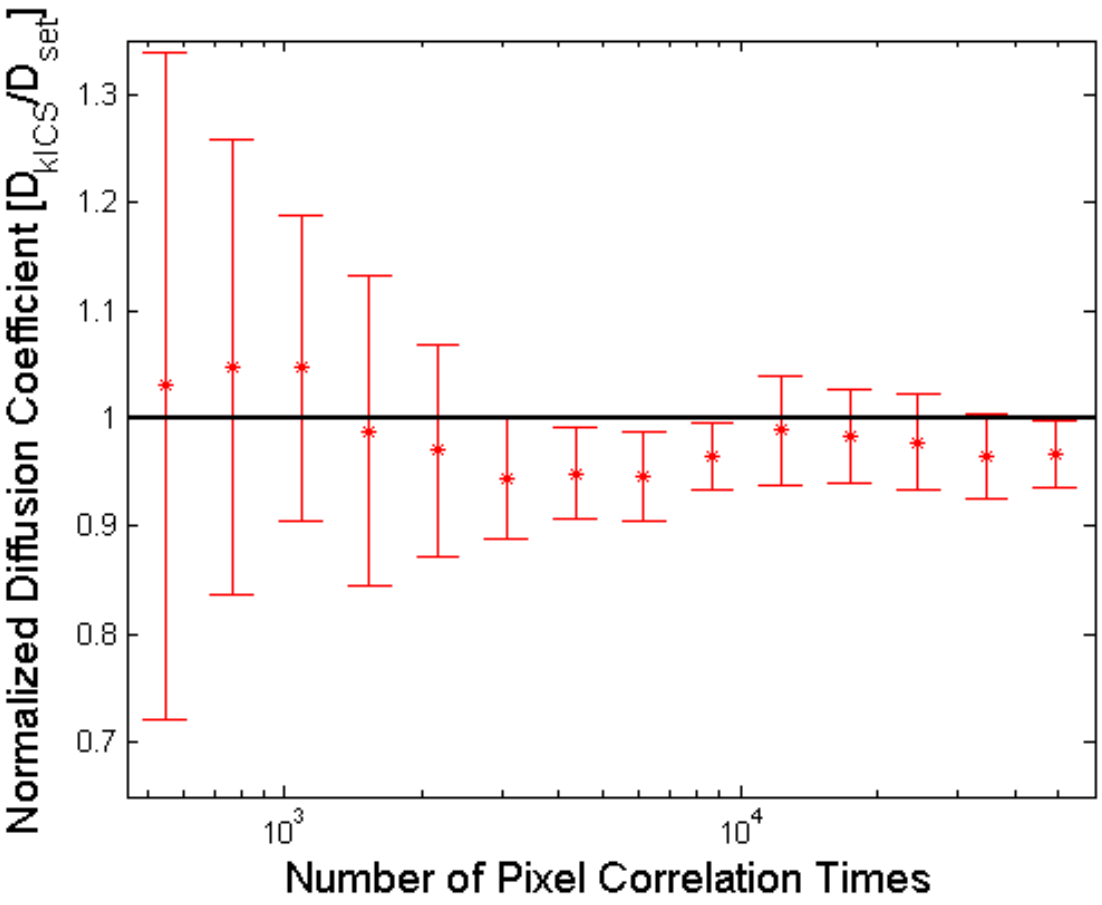

Figure 3.2: The bias of the measured kICS diffusion as a function of movie length (measured in multiples of pixel correlation times). **This figure shows that movie length does not create any bias in diffusion measurement, but the error of a single measurement increases with fewer time samples**. This result closely matches the result of Figure 3.1. (Averaged over n = 5 simulations; D = 0.5 pixel^2/frame; t_{exp} = 0.02 s; C = 1.95x10^{-3} pixel^{-2}; 1 pixel = 266.7 nm; 1 frame = 0.025 s; FOV = 256 pixels X 256 pixels)

In other words, each correlation pair calculated can be considered to have $\frac{1}{T-\tau}$ weighting. If there are many time samples in an image series, the relative fluctuation experienced by one correlation pair is very small. As the number of time samples decreases, the relative influence of each correlation pair increases, and the value of the parameters extracted from the correlation functions becomes significantly influenced by individual fluctuations in correlation pairs. This phenomenon can be seen in Figure 3.1; as the number of time samples decreases between sub-figure a) to sub-figure c), the correlation curves become much noisier. Similarly, this phenomenon also appears in Figure 3.2. As the relative influence correlation pair fluctuations increase, the measured diffusion coefficient is correspondingly influenced, and a greater range of diffusion coefficients are measured (leading to larger error bars). It should be noted that the size of the standard deviation as function of movie length exhibits behaviour that scales similarly to $\frac{1}{T-\tau}$. It is worth noting that although the error bars exhibit time-length dependence, the mean value measured does not exhibit any systematic bias. This is because the fluctuation in each intensity pair fluctuates about the true value; hence many measurements will result in a good estimation of the true diffusion coefficient.

One interesting artifact that is produced in Figure 3.2 is the semi-oscillating nature of the measured diffusion coefficient as a function of movie length. While Figures 3.3 through 3.5 required each point to be the result of analysis from distinct movies, the bias with respect to time samples could be performed by using the same 5 movies and selecting a different end frame for each point. In doing this, the measured normalized diffusion coefficient becomes in itself correlated across values in the x-axis. Simulating a set of 5 different movies for each point would have eliminated this artifact. While this oscillating property is not intrinsically useful, it gives insight on the convergence of the measurement to the correct value. Much like graphs in Figure 3.1, the fitting function to produce the diffusion coefficient will regress to the true value once enough time samples are used for correlation.

As such, the best way to interpret the results from Figure 3.2 is that a there is no systematic bias contributed directly from the number of time samples; each point should on average correspond to the true value (assuming there are no systematic biases from other sources).

3.1.2 Bias due to Various Diffusion Rates

In order to explain the results in 3.3, it is important to consider two different phenomena that bias extremely large diffusion coefficients and extremely small diffusion coefficients. First of all, the probability of correlating intensities for a given pixel decreases as the diffusion coefficient increases. This is a direct manifestation of equation 2.6. As such, a much wider range of emitter-to-emitter distances can be observed, which biases the correlation function towards slower characteristic decays. This bias becomes much more pronounced as the particle density increases, as the likelihood of measuring intensities between non-identical particles increases in this regime. This effect is seen on the right-hand portion of Figure 3.3, as the underestimation bias of the measured diffusion coefficient becomes more apparent for faster particles.

Secondly, it is important to consider correlation time; i.e. how long the diffusers stay in a given area. This quantity was previously defined both in terms of the excitation beam radius, and the length of a pixel. What is important to note is that in all cases, the correlation time scales inversely with the diffusion coefficient.

It is important to note that the measured diffusion coefficients in Figure 3.3 were determined from movies of identical length ($T = 4096$ frames). However in doing this, the number of correlation times is different between samples of different diffusion coefficients. Samples with a smaller diffusion coefficient contain few diffusion times in the movies than those with higher diffusion coefficients. According to Figure 3.2, this means that samples with a smaller diffusion coefficient should contain a larger error than those with a larger diffusion coefficient value. This completely agrees with the increased error

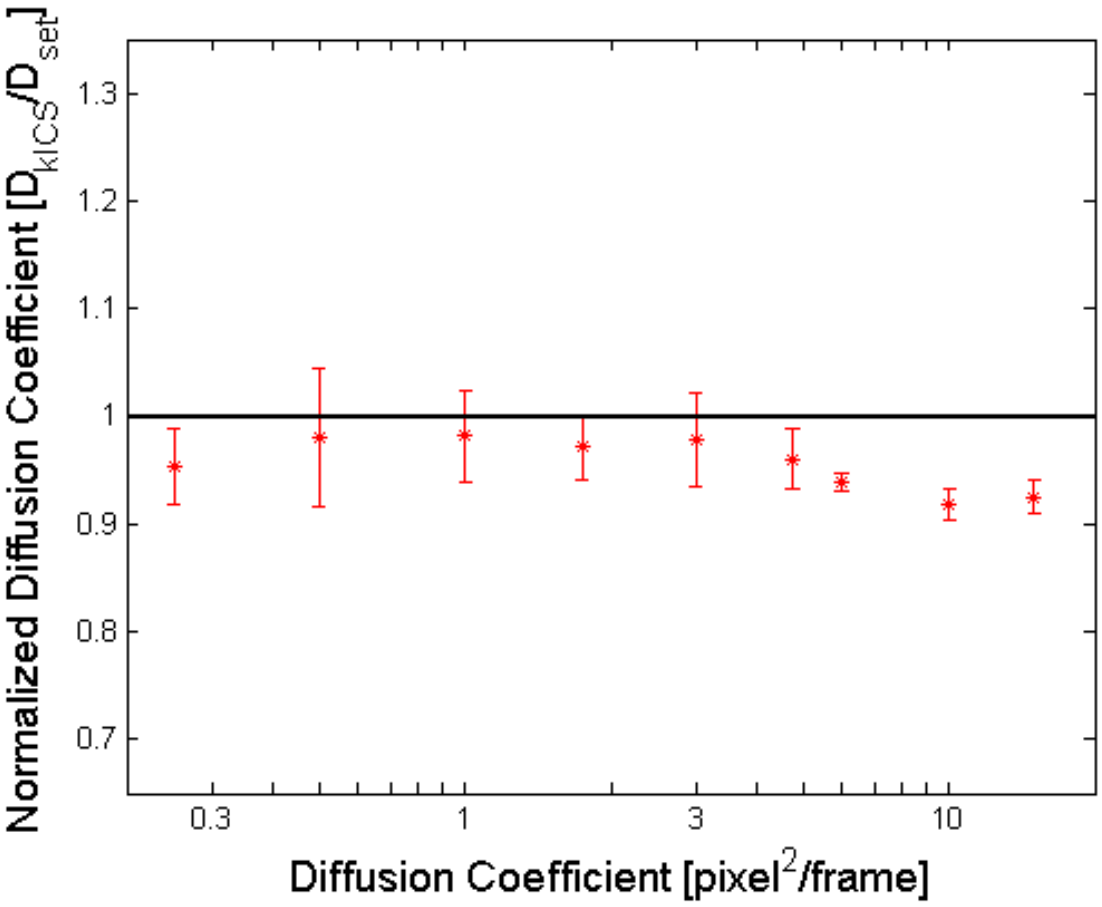

Figure 3.3: The bias of the measured kICS diffusion as a function of inputted diffusion coefficient. **This figure shows that it becomes more difficult to correlate intensities when the diffusion coefficient is very high**. As expected, there seems to be a range of values for which a reliable measurement can be made; a very small diffusion coefficient autocorrelates too readily while a very large diffusion coefficient increases the likelihood of cross-correlates. The ideal values lie between 0.5 pixel^2/frame and 5 pixel^2/frame. (Averaged over n = 5 simulations; t_{exp} = 0.02 s; T = 4096 frames; C = 1.95x10^{-3} pixel^{-2}; 1 pixel = 266.7 nm; 1 frame = 0.025 s; FOV = 256 pixels X 256 pixels)

in measured diffusion coefficient from simulations containing slower diffusers.

One final noteworthy point to be made is the influence of acquisition rate (related to the exposure time) compared to the influence of number of correlation times. In the temporal dimension, there are three times of interest: the exposure time of each frame, the correlation time of a given diffuser, and the total time samples of the image series. Since all relative units are used in the figure, the results can be interpreted in different ways depending on what is chosen as the absolute quantity. As such, rather than considering absolute acquisition rate and equal movie lengths between samples of different diffusers, Figure 3.3 can be considered a representation of samples with identical diffusers but with different acquisition rates and different movie lengths. In this case, the left hand portion of the x-axis can be considered the regime in which a high acquisition rate is used to take a short movie, while the right hand portion of the x-axis can be considered the regime of long movies with a low acquisition rate. In this sense, a logical conclusion is that having an acquisition rate that are within the same range of the dynamic rates of the sample. In the context of experiments, it is not uncommon to study oligonucleotides that have diffusion coefficients on the order of 50 $\mu m^2/s$, which may be very difficult to temporally sample without making significant sacrifices on the camera's field of view.

3.1.3 Bias due to Particle Density

While other spatial and temporal parameters directly contribute to the measurement of system dynamics, the particle density is independent of these rates. The final result calculated in equation 2.7 shows that the number of particles appears as a pre-factor to the exponentially decaying term. As such, the emitter density is essentially independent of parameter measurement if the correct normalization is made. By looking at the result in 3.4, it is clear that for a wide range of 3-500 particles per 128x128 pixel grid, the particle density does not influence the calculation of system dynamics. With this in mind, it is important to consider the extreme regimes: very small concentration, and very high con-

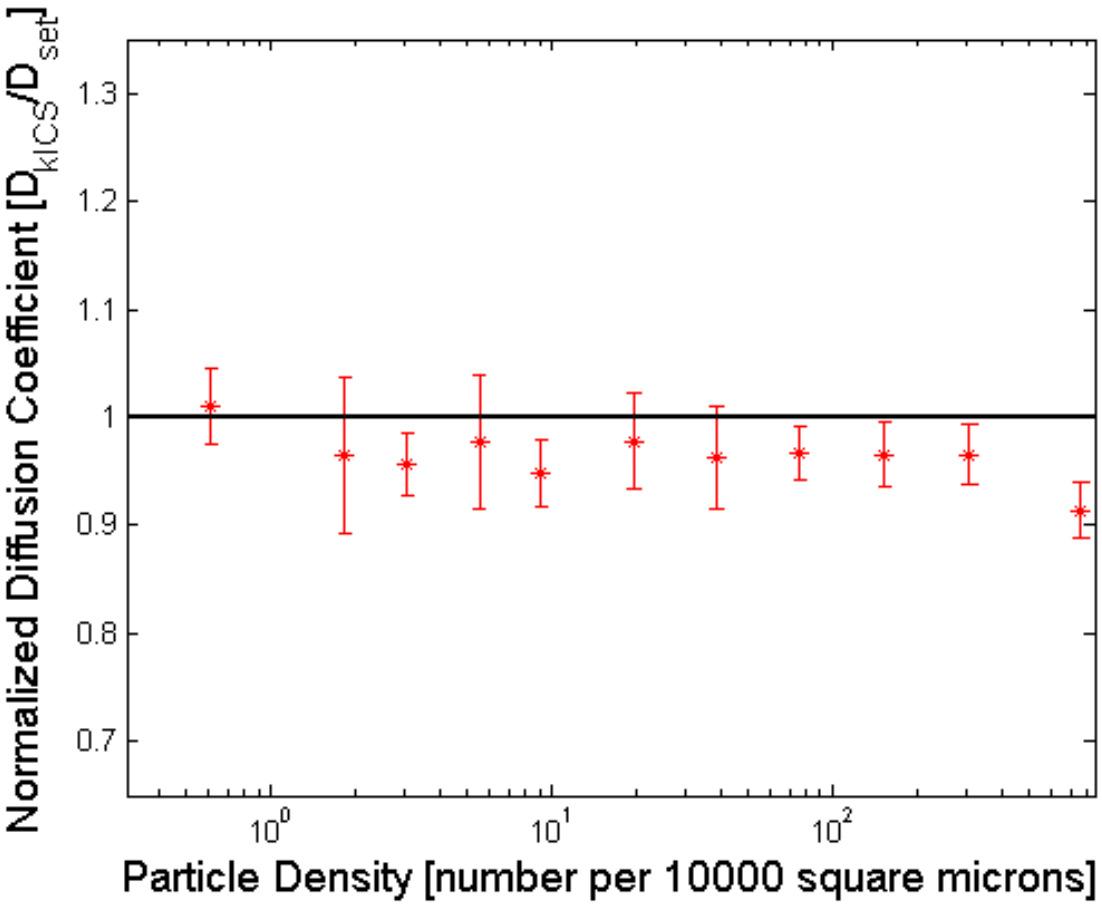

Figure 3.4: The bias of the kICS diffusion as a function of particle density. **The bias with respect to particle density is very small, as shown by the consistent estimation of the diffusion coefficient between 2×10^{-4} and 3×10^{-2} particles per pixel^2.** (Averaged over n = 5 simulations; D = 0.5 pixel^2/frame; t_{exp} = 0.02 s; T = 4096 frames; 1 pixel = 266.7 nm; 1 frame = 0.025 s; FOV = 256 pixels X 256 pixels)

centration. As derived in section 2.2, one of the main underlying assumptions in kICS is that unlike particles do not correlate. This stems from the fact that the probability density functions of different particles are independent, and thus the correlation pairs average to zero given a sufficiently large sample of data. At very large concentrations however, the change in pixel to pixel intensity becomes difficult to measure, resulting in relatively smaller fluctuations. This results in a lower signal to noise ratio, which would bias the measurement towards higher error. The bias in the right-most point in Figure 3.4 may be explained by considering that the average distance between emitters is very small. This means that the correlation corresponding to a high reciprocal spatial vector would be relatively large. This is effect, causes an underestimation of the dynamics when fitting the function to an exponential (since the correlation would not decay quickly at those intensities).

The other extreme regime is that the field of view only contains a few particles. While this maximizes the likelihood that unlike particles do not correlate, it creates an issue when coupled with the fact that the field of view is finite. As such, when there are no emitters in the field of view, the signal intensity essentially disappears. Correlating between successive frames of signal intensity to no intensity results in a large decay with respect to τ. As such, this manifests itself as a very steep correlation curve, and therefore, a high estimation of the system dynamics. Note that in order for this to happen, the signal intensity would essentially have to disappear. This is why this effect is only present with the left-most point in Figure 3.4, as that point corresponds to one particle in the 128x128 pixel field of view.

3.1.4 Bias due to size of the Field of View

kICS analysis is a correlation analysis tool using information from reciprocal space intensities. As such, a discrete Fourier transform is done in the field of view. A property of Fourier transforms in a limited data range is that the overall bandwidth remains un-

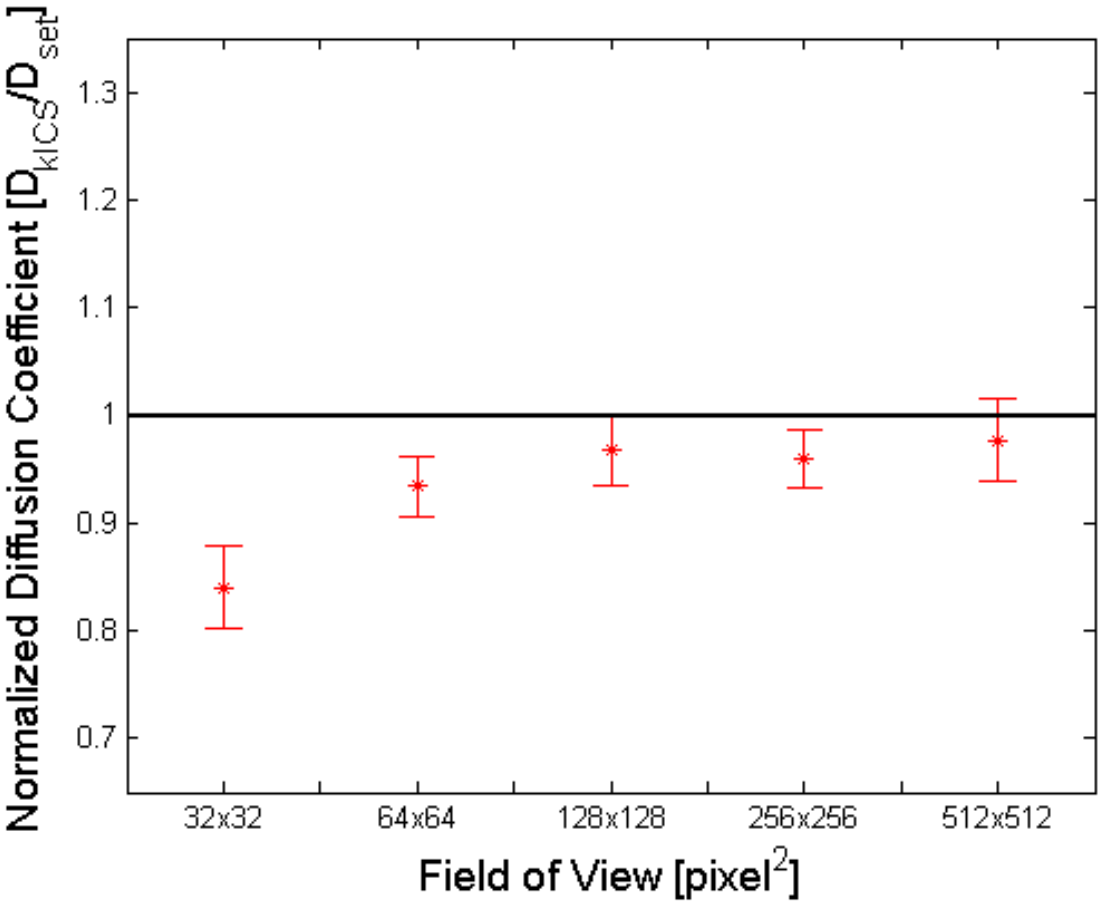

Figure 3.5: The bias of the kICS diffusion as a function of grid size. **The measurement of dynamics is most accurate when the size of the field of view is very large.** As shown in the figure, fields of view greater than 128x128 pixels is sufficient. (Averaged over n = 5 simulations; D = 0.5 pixel^2/frame; t_{exp} = 0.02 s; T = 4096 frames; C = 1.95x10^{-3} pixel^{-2}; 1 pixel = 266.7 nm; 1 frame = 0.025 s)

affected, but the resolution of the vectors in reciprocal space decreases significantly. As such, the overall behaviour of the correlation curve remains constant, although the number of points used for fitting decreases significantly.

In this sense, having an infinitely large field of view would be most optimal for analysis. As shown in Figure 3.5, a field of view greater than 128 $\times$ 128 pixels provides sufficient resolution in reciprocal space.

3.2 Concluding Remarks

Simulations provide a very useful tool for guiding experimental analysis, as it provides a reliable method of studying parametric relationships. This chapter explored correlation bias as a function of four independent parameters: number of time samples, diffusion rate, particle density, and domain size. The results of the analysis show appropriate regimes for experimentalists:

- Longer movies present more reliable imaging statistics and should be on the order of 1000 correlation times;
- Molecules with diffusion coefficient ranges between 0.5 pixel^2/frame and 5 pixel^2/frame (according to the experimental set-up) are appropriate for imaging;
- A wide-range of particle densities is appropriate for measurement 2×10^{-4} and 3×10^{-2} particles per $pixel^2$ and contributes minimally to correlation bias;
- A minimal field of view of 128x128 pixels is required for accurate reciprocal space correlation analysis;

The resulting biases calculated due to different parameters allow experimentalists to understand not only why certain biases occur, but also how to account for these biases, and which regimes in parameter space would be most optimal for a given experiment. Furthermore, these biases can be used to give insight into higher order correction terms to make the analysis more reliable.

Chapter 4

Conclusions

4.1 Summary of Contributions

The motivation of this thesis is to understand the optimal way to measure biomolecular dynamics with sub-cellular resolution. A review of the capabilities of modern-day microscopic methodologies indicated that CLiC fluorescence microscopy is most appropriate for analyzing free solution systems that mimic physiological concentrations. Furthermore, k-space Image Correlation Spectroscopy was identified as the most appropriate technical tool for analysis.

As the CLiC-kICS relationship is still being developed, a theoretical framework was established in Chapter 2. Section 2.1 presents a methodology for accurately modeling free solution experiments based physical processes such as Brownian motion and imaging acquisition properties of the experimental instrumentation. This was followed by a the derivation of an analytical correlation expression in section 2.2 and an investigation into appropriate normalization and fitting techniques in section 2.3.

Finally, Chapter 3 investigated explored correlation bias as a function of four independent parameters: number of time samples, diffusion rate, particle density, and domain size. All of these parameters can be controlled (within reason) by the experimental set-

up, thereby guiding experimentalists on the appropriate parameter boundaries to perform CLiC experiments with kICS analysis.

4.2 Future Work

While many relationships between the characteristic correlation rate and experimental parameters were explored, there are many extensions that can be made on this analysis. Many experimental parameters are not independent, meaning that bias relationships cannot be considered the product of independent parameter calculations. For example, the kICS derivation in section 2.2 operated under the assumption that a particle position in one frame will not correlate with a different particle's position in another frame. This condition is satisfied under a specific range of values for the emitter density, field of view, diffusion coefficient, and imaging rate. Changing one of these parameters also influences the acceptable range in the others. Investigating bias from a multi-parameter stand-point may provide more insight into system dynamics.

An additional avenue to explore is to replicate some of the biases with simple experiments. While there are many advantages to performing simulations, a major disadvantage is that there can be unknown or unaccounted effects in experiments. Comparing experimental and simulated results would help further refine the simulated model of the system and produce more accurate bias estimations.

As noted in Chapter 1, FRET can provide spatial resolution on the order of 1-10 nm based on the emission intensity of the acceptor probe. Due to CLiC's unrivaled noise-suppression abilities, the intensity signal can be interpreted very precisely to generate a positional tracking at the sub-pixel level. Moreover, cross-correlation techniques can provide insight into binding and conformation characteristics.

Finally, applying this type of analysis to other CLiC systems such as the one defined by Scott et al. [52] would prove to be very useful. The results of the multi-parametric analysis

would give insight into which biomolecules can be used for the microscopy system, and it would provide insight on how to correct first order biases.

Chapter 5

Appendix

5.1 Software Documentation

A general purpose software infrastructure for performing the analysis is described in Figure 5.1. The code was designed with maximum generality to support any specific purpose (for example, photophysical bias analysis, time integration effects, and chamber geometry). As jobs were submitted to a super-computing cluster, routine checkpoints were established to save data in case of memory limits or time-checkouts occurred. The code can be found at: <https://www.dropbox.com/sh/4xti2yj2h16pdft/AACPy0rPc4zylxkjLCwFOrQda?dl=0>.

5.2 Supporting Figures

In addition to Figure 2.10, the carpet plots in Figures 5.2 and 5.3 were used to motivate the fitting region.

The parameter space exploration reference in Chapter 4 was attempted, but gained insufficient to be practical. Gaining this methodology with the resolution acquired in Chapter 3 would prove to be very useful to future developments in the project.

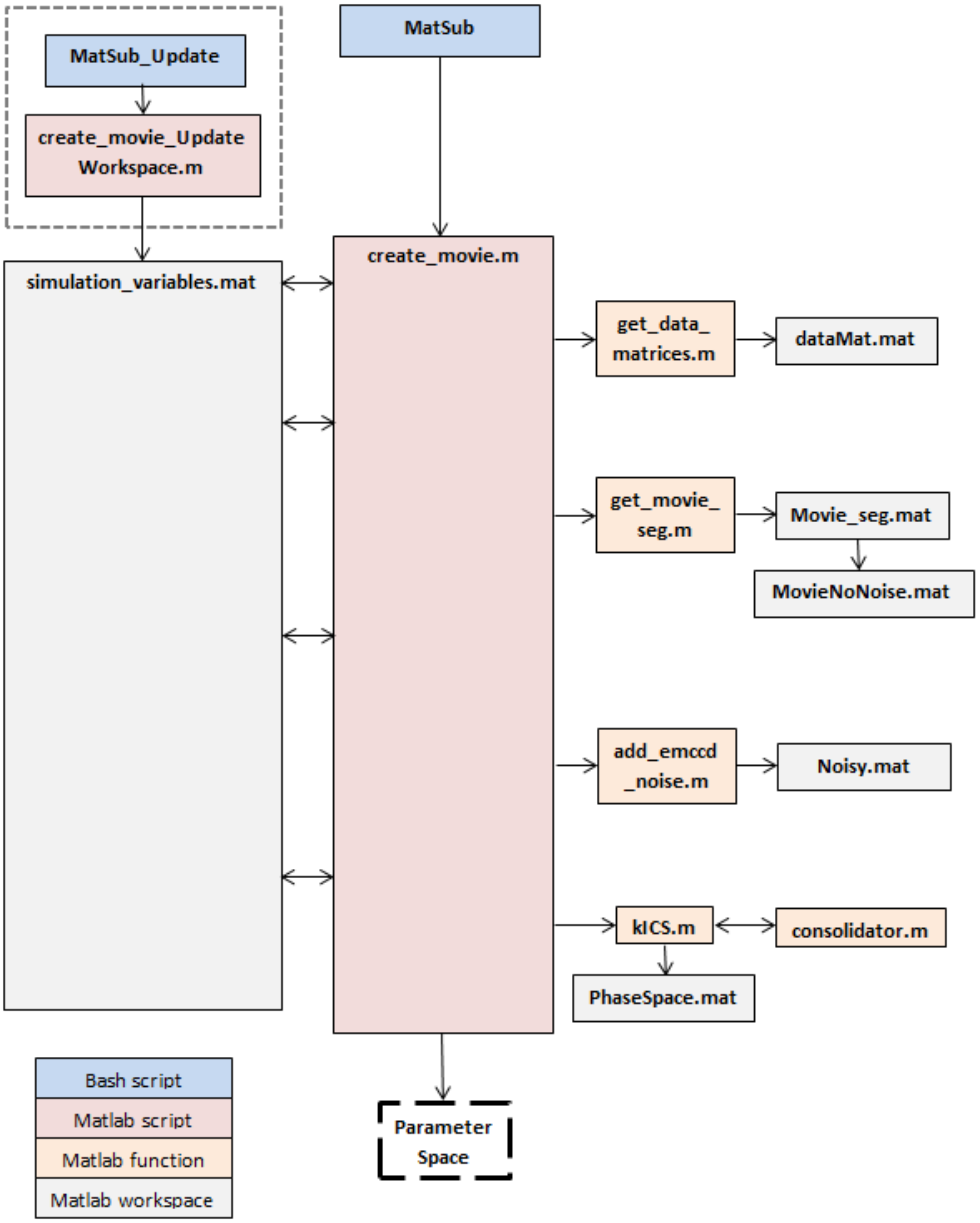

Figure 5.1: **A flow chart demonstrating how simulated movies are created.** Bash scripts are used to submit MATLAB jobs remotely. Sending batches of simulations to a super-computing cluster proved to be the most effective way to study bias.

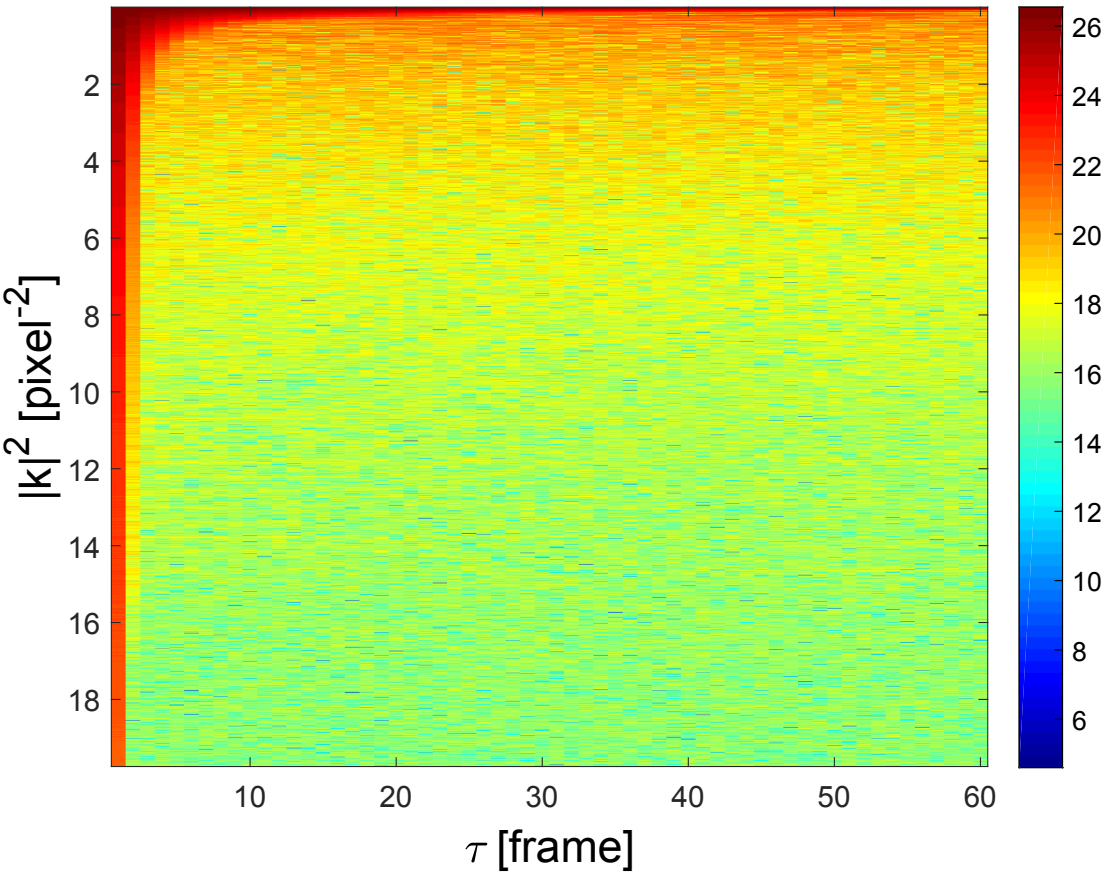

Figure 5.2: **The correlation function from Figure 2.10, plotted linearly against spatial vector and time lag. A smaller selection of τ is chosen to demonstrate the influence of the PSF in the $|k|^2$ dimension.**

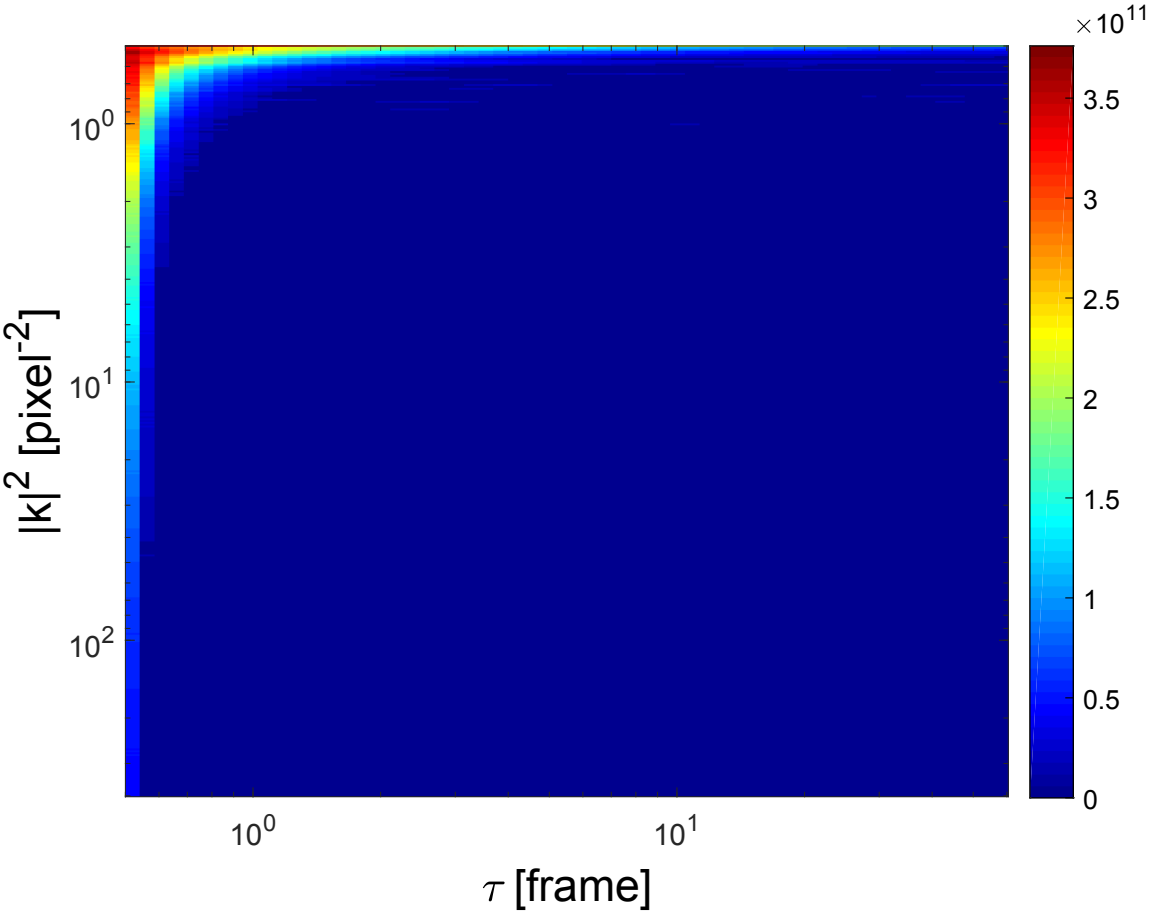

Figure 5.3: **The intensity contour of the fitting region used in the previous plots in Figures 3.2 through 3.4. Below a threshold intensity, the correlation function reaches the noise floor and can no longer be used for determining rate coefficients.** This threshold noise value is found by averaging the noisier regions of the correlation curve.

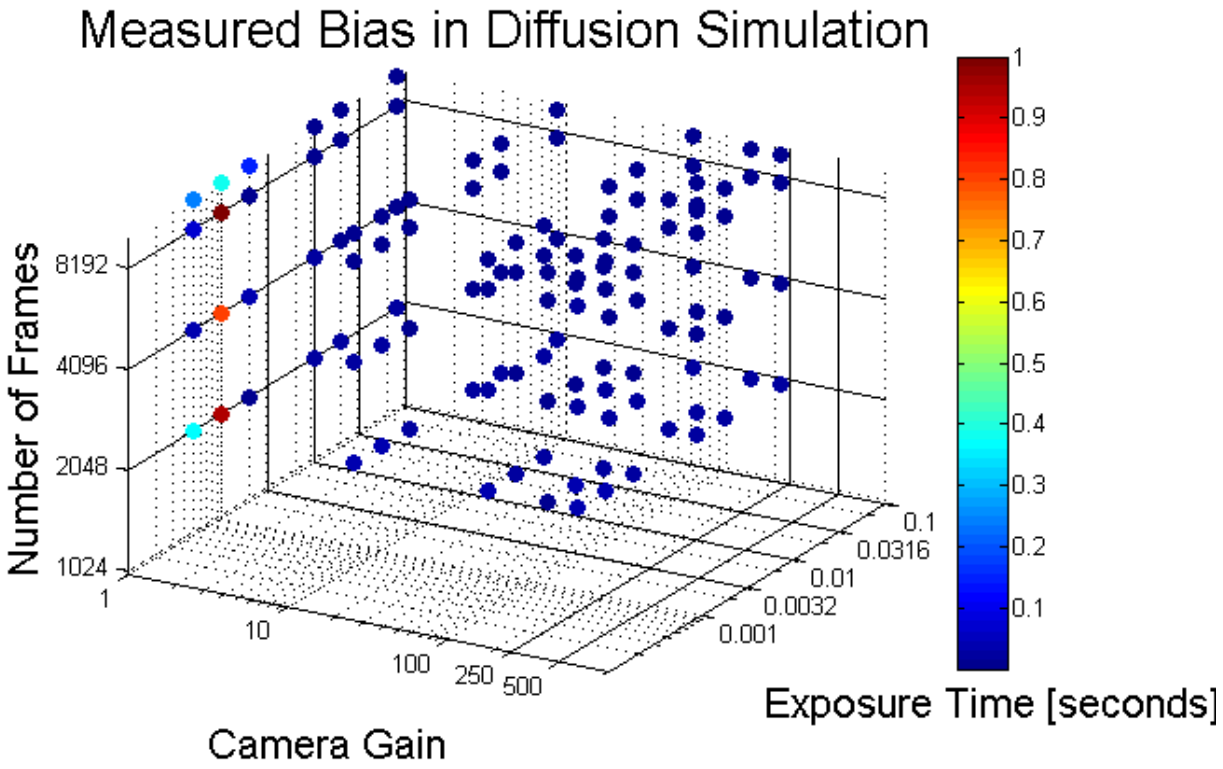

Figure 5.4: The deviation in the calculated kICS diffusion coefficient as a function of three parameters: gain, exposure time, and number of time samples. This parameter space graph shows that the measured coefficients readily agree with the inputted values, except at the more challenging conditions (low gain, low exposure time, and few time frames).

Bibliography

[1] Alberts B. *Molecular Biology of the Cell*. New York: Garland Science, https://www.ncbi.nlm.nih.gov/books/NBK21054/, 4 edition, 2002.

[2] Ron Milo. Useful fundamental numbers in molecular biology. bioNumbers.

[3] L. Brocchieri. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research*, 33(10):3390–3400, jun 2005.

[4] E. Abbe. Xv.—the relation of aperture and power in the microscope (continued)*. *Journal of the Royal Microscopical Society*, 3(6):790–812, 1883.

[5] Bert Voigtländer, Peter Coenen, Vasily Cherepanov, Peter Borgens, Thomas Duden, and F. Stefan Tautz. Low vibration laboratory with a single-stage vibration isolation for microscopy applications. *Review of Scientific Instruments*, 88(2):023703, feb 2017.

[6] Tobias Müller, Christian Schumann, and Annette Kraegeloh. STED microscopy and its applications: New insights into cellular processes on the nanoscale. *ChemPhysChem*, 13(8):1986–2000, feb 2012.

[7] Marian Morys and Daniel Berger. Accurate measurements of biologically effective ultraviolet radiation. In Knut H. Stamnes, editor, *Atmospheric Radiation*. SPIE, nov 1993.

[8] Malgorzata Rózanowska, John Jarvis-Evans, Witold Korytowski, Mike E. Boulton, Janice M. Burke, and Tadeusz Sarna. Blue light-induced reactivity of retinal age pigment. *Journal of Biological Chemistry*, 270(32):18825–18830, aug 1995.

[9] Robert Nowakowski, Paul Luckham, and Peter Winlove. Imaging erythrocytes under physiological conditions by atomic force microscopy. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1514(2):170 – 176, 2001.

[10] Kai-Chih Chang, Yu-Wei Chiang, Chin-Hao Yang, and Je-Wen Liou. Atomic force microscopy in biology and biomedicine. *Tzu Chi Medical Journal*, 24(4):162 – 169, 2012.

[11] Marie-Paule Mingeot-Leclercq, Magali Deleu, Robert Brasseur, and Yves F Dufrêne. Atomic force microscopy of supported lipid bilayers. *Nature Protocols*, 3(10):1654–1659, oct 2008.

[12] Stavroula Skylaki, Oliver Hilsenbeck, and Timm Schroeder. Challenges in long-term imaging and quantification of single-cell dynamics. *Nature Biotechnology*, 34(11):1137–1144, nov 2016.

[13] D. Rekesh, Y. Lyubchenko, L.S. Shlyakhtenko, and S.M. Lindsay. Scanning tunneling microscopy of mercapto-hexyl-oligonucleotides attached to gold. *Biophysical Journal*, 71(2):1079–1086, aug 1996.

[14] Jürgen Bereiter-Hahn. *Probing Biological Cells and Tissues with Acoustic Microscopy*, pages 79–115. Springer US, Boston, MA, 1995.

[15] Surena Vahabi, Bahareh Nazemi Salman, and Anahita Javanmard. Atomic force microscopy application in biological research: A review study. *Iran J Med Sci*, 38(2):76–83, Jun 2013. IJMS-38-076[PII].

[16] Kai-Chih Chang, Yu-Wei Chiang, Chin-Hao Yang, and Je-Wen Liou. Atomic force microscopy in biology and biomedicine. *Tzu Chi Medical Journal*, 24(4):162–169, dec 2012.

[17] William J. Greenleaf, Michael T. Woodside, and Steven M. Block. High-resolution, single-molecule measurements of biomolecular motion. *Annual Review of Biophysics and Biomolecular Structure*, 36(1):171–190, jun 2007.

[18] Nico Strohmeyer, Mitasha Bharadwaj, Mercedes Costell, Reinhard Fässler, and Daniel J. Müller. Fibronectin-bound alpha beta integrins sense load and signal to reinforce adhesion in less than a second. *Nature Materials*, 16(12):1262–1270, nov 2017.

[19] Ronald E. Gordon. Electron microscopy: A brief history and review of current clinical application. In *Histopathology*, pages 119–135. Springer New York, 2014.

[20] SEM images of DNA double helix and nucleosomes observed by ultrahigh-resolution scanning electron microscopy. *Journal of Electron Microscopy*, jun 1991.

[21] N. d. Jonge, D. B. Peckys, G. J. Kremers, and D. W. Piston. Electron microscopy of whole cells in liquid with nanometer resolution. *Proceedings of the National Academy of Sciences*, 106(7):2159–2164, jan 2009.

[22] Niels de Jonge and Frances M. Ross. Electron microscopy of specimens in liquid. *Nature Nanotechnology*, 6(11):695–704, oct 2011.

[23] Yigong Shi. A glimpse of structural biology through x-ray crystallography. *Cell*, 159(5):995–1014, nov 2014.

[24] Mikio Kataoka and Yuji Goto. X-ray solution scattering studies of protein folding. *Folding and Design*, 1(5):R107–R114, oct 1996.

[25] Sergio Santos, Victor Barcons, Hugo K. Christenson, Josep Font, and Neil H. Thomson. The intrinsic resolution limit in the atomic force microscope: Implications for heights of nano-scale features. *PLoS ONE*, 6(8):e23821, aug 2011.

[26] Griffiths Gareth. *Electron Microscopy in Cell Biology*. American Cancer Society, 2006.

[27] Rainer Heintzmann and Gabriella Ficz. Chapter 22 - breaking the resolution limit in light microscopy. In Greenfield Sluder and David E. Wolf, editors, *Digital Microscopy*, volume 114 of *Methods in Cell Biology*, pages 525 – 544. Academic Press, 2013.

[28] Peter I. Chang, Peng Huang, Jungyeoul Maeng, and Sean B. Andersson. Local raster scanning for high-speed imaging of biopolymers in atomic force microscopy. *Review of Scientific Instruments*, 82(6):063703, jun 2011.

[29] Robert R. Shannon and Brian J. Ford. Microscope - encyclopædia britannica, Apr 2007.

[30] R. Oldenbourg. Polarized light microscopy: Principles and practice. *Cold Spring Harbor Protocols*, 2013(11):pdb.top078600–pdb.top078600, nov 2013.

[31] Martin J Booth. Adaptive optical microscopy: the ongoing quest for a perfect image. *Light: Science & Applications*, 3(4):e165–e165, apr 2014.

[32] Kurt Thorn. A quick guide to light microscopy in cell biology. *Molecular Biology of the Cell*, 27(2):219–222, jan 2016.

[33] Christopher DuFort and Matthew Paszek. Nanoscale cellular imaging with scanning angle interference microscopy. In *Methods in Cell Biology*, pages 235–252. Elsevier, 2014.

[34] Kevin M. Dean, Philippe Roudot, Erik S. Welf, Theresa Pohlkamp, Gerard Garrelts, Joachim Herz, and Reto Fiolka. Imaging subcellular dynamics with fast and light-efficient volumetrically parallelized microscopy. *Optica*, 4(2):263–271, Feb 2017.

[35] Joseph R. Lakowicz, editor. *Introduction to Fluorescence*, pages 1–26. Springer US, Boston, MA, 2006.

[36] K. Bacia, E. Haustein, and P. Schwille. Fluorescence correlation spectroscopy: Principles and applications. *Cold Spring Harbor Protocols*, 2014(7):pdb.top081802–pdb.top081802, jul 2014.

[37] Helge Ewers. 4 - nano resolution optical imaging through localization microscopy. In Floris G. Wouterlood, editor, *Cellular Imaging Techniques for Neuroscience and Beyond*, pages 81 – 100. Academic Press, San Diego, 2012.

[38] S. James Remington. Green fluorescent protein: A perspective. *Protein Science*, 20(9):1509–1519, jul 2011.

[39] Sviatlana Shashkova and Mark C. Leake. Single-molecule fluorescence microscopy review: shedding new light on old problems. *Bioscience Reports*, 37(4):BSR20170031, jul 2017.

[40] Elizabeth A Jares-Erijman and Thomas M Jovin. FRET imaging. *Nature Biotechnology*, 21(11):1387–1395, nov 2003.

[41] Carel Fijen, Alejandro Montón Silva, Alejandro Hochkoeppler, and Johannes Hohlbein. A single-molecule FRET sensor for monitoring DNA synthesis in real time. *Physical Chemistry Chemical Physics*, 19(6):4222–4230, 2017.

[42] Monica Monici. Cell and tissue autofluorescence research and diagnostic application. 11:227–56, 02 2005.

[43] S. A. Haider, A. Cameron, P. Siva, D. Lui, M. J. Shafiee, A. Boroomand, N. Haider, and A. Wong. Fluorescence microscopy image noise reduction using a stochastically-connected random field model. *Scientific Reports*, 6(1), feb 2016.

[44] Elke Haustein and Petra Schwille. *Fluorescence Correlation Spectroscopy*, pages 637–675. Springer Netherlands, Dordrecht, 2008.

[45] James Jonkman and Claire M. Brown. Any way you slice it—a comparison of confocal microscopy techniques. *Journal of Biomolecular Techniques : JBT*, pages jbt.15–2602–003, jul 2015.

[46] Kenneth N. Fish. Total internal reflection fluorescence (tirf) microscopy. *Curr Protoc Cytom*, 0 12:Unit12.18–Unit12.18, Oct 2009. 19816922[pmid].

[47] Raghav K Chhetri, Fernando Amat, Yinan Wan, Burkhard Höckendorf, William C Lemon, and Philipp J Keller. Whole-animal functional and developmental imaging with isotropic spatial resolution. *Nature Methods*, 12(12):1171–1178, oct 2015.

[48] R. S. Udan, V. G. Piazza, C. w. Hsu, A.-K. Hadjantonakis, and M. E. Dickinson. Quantitative imaging of cell dynamics in mouse embryos using light-sheet microscopy. *Development*, 141(22):4406–4414, oct 2014.

[49] Gil Henkin, Daniel Berard, Francis Stabile, Marjan Shayegan, Jason S. Leith, and Sabrina R. Leslie. Manipulating and visualizing molecular interactions in customized nanoscale spaces. *Analytical Chemistry*, 88(22):11100–11107, oct 2016.

[50] Sabrina Leslie, Daniel Berard, Jason S. Leith, and Francois Michaud. Using tunable nanoscale confinement to image and manipulate DNA. In *Advanced Photonics*. OSA, 2014.

[51] Sabrina R. Leslie, Alexander P. Fields, and Adam E. Cohen. Convex lens-induced confinement for imaging single molecules. *Analytical Chemistry*, 82(14):6224–6229, 2010. PMID: 20557026.

[52] Shane Scott, Zhi Ming Xu, Fedor Kouzine, Daniel J Berard, Cynthia Shaheen, Barbara Gravel, Laura Saunders, Alexander Hofkirchner, Catherine Leroux, Jill Laurin,

David Levens, Craig J Benham, and Sabrina R Leslie. Visualizing structure-mediated interactions in supercoiled DNA molecules. *Nucleic Acids Research*, 46(9):4622–4631, apr 2018.

[53] Ramin Skibba, Ravi K. Sheth, Andrew J. Connolly, and Ryan Scranton. The luminosity-weighted or 'marked' correlation function. *Monthly Notices of the Royal Astronomical Society*, 369(1):68–76, may 2006.

[54] Jiang-Feng Chen, James A. Zagzebski, Fang Dong, and Ernest L. Madsen. Estimating the spatial autocorrelation function for ultrasound scatterers in isotropic media. *Medical Physics*, 25(5):648–655, may 1998.

[55] L. Klimeš. Correlation functions of random media. *Pure and Applied Geophysics*, 159(7-8):1811–1831, jul 2002.

[56] Douglas Magde, Elliot Elson, and W. W. Webb. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.*, 29:705–708, Sep 1972.

[57] Elliot L. Elson and Douglas Magde. Fluorescence correlation spectroscopy. i. conceptual basis and theory. *Biopolymers*, 13(1):1–27, 1974.

[58] R. Rigler, Ü. Mets, J. Widengren, and P. Kask. Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion. *European Biophysics Journal*, 22(3):169–175, Aug 1993.

[59] Kai Hassler, Marcel Leutenegger, Per Rigler, Ramachandra Rao, Rudolf Rigler, Michael Gösch, and Theo Lasser. Total internal reflection fluorescence correlation spectroscopy (tir-fcs) with low background and high count-rate per molecule. *Opt. Express*, 13(19):7415–7423, Sep 2005.

[60] Zeno Foldes-Pappn and Masataka Kinjo. *Fluorescence Correlation Spectroscopy in Nucleic Acid Analysis*, pages 25–64. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[61] Edmund Matayoshi and Kerry Swift. *Applications of FCS to Protein-Ligand Interactions: Comparison with Fluorescence Polarization*, pages 84–98. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[62] N.O. Petersen, P.L. Höddelius, P.W. Wiseman, O. Seger, and K.E. Magnusson. Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. *Biophysical Journal*, 65(3):1135–1146, sep 1993.

[63] Molly J. Rossow, Jennifer M. Sasaki, Michelle A. Digman, and Enrico Gratton. Raster image correlation spectroscopy in live cells. *Nature Protocols*, 5:1761 EP –, Oct 2010.

[64] David L. Kolin, David Ronis, and Paul W. Wiseman. k-space image correlation spectroscopy: A method for accurate transport measurements independent of fluorophore photophysics. *Biophysical Journal*, 91(8):3061–3075, oct 2006.

[65] Marco A. Catipovic, Paul M. Tyler, Josef G. Trapani, and Ashley R. Carter. Improving the quantification of brownian motion. *American Journal of Physics*, 81(7):485–491, jul 2013.

[66] Andor An Oxford Instruments Company. *iXon Ultra*, 2016.

[67] Radek Erban, Jonathan Chapman, and Philip Maini. A practical guide to stochastic simulations of reaction-diffusion processes, 2007.

[68] Michael Hirsch, Richard J. Wareham, Marisa L. Martin-Fernandez, Michael P. Hobson, and Daniel J. Rolfe. A stochastic model for electron multiplication charge-coupled devices – from theory to practice. *PLoS ONE*, 8(1):1–13, 01 2013.