RESOLUTION AND LOCALIZATION IN SINGLE MOLECULE MICROSCOPY

by

SRIPAD RAM

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2007

To my family.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervising professors Drs. Raimund J. Ober and E. Sally Ward for their invaluable advice, constant support and for their faith in me even when things (or I) seemed hopeless during my doctoral studies. I have learnt a lot in the last six years and I thank them for giving me the opportunity to work in their lab. I would like to thank my committee members Prof. Peter Antich, Dr. Ege Kavalali, Dr. Wen-hong Li and Dr. George Alexandrakis for taking the time to serve on my committee.

I would like to extend my appreciation to all current and former members in the laboratory of Sally Ward and Raimund Ober who have created a very productive, enjoyable work environment. In particular, I would like to thank the lab manager Carlos Vaccaro for excellent logistical support, Cruz Martinez for helping me get started with fluorescence microscopy experiments and Sylvia Wanjie for helping me with mammalian transfection protocols and for her (mostly cynical) passing comments, which I always enjoyed. Special thanks are due to all my student colleagues Prashant, Silvia, Jerry, Zhuo, Anish and Keerthi, who have been a great support during my PhD.

I would like to thank my department chairman Professor Khosrow Behbehani for helping me with the STEM doctoral fellowship and for providing me travel support to attend conferences.

Finally I would like to express my deep gratitude to my family, especially my second brother Srinath, his wife Sowmya and their kids Naren and Shreya. Besides free food, wine and board that I received for a significant portion of my PhD, they looked out for me, which allowed me to completely focus on my thesis. Thank you folks.

July 5, 2007

ABSTRACT

RESOLUTION AND LOCALIZATION IN SINGLE MOLECULE MICROSCOPY

Sripad Ram, Ph.D.

The University of Texas at Arlington, 2007

Supervising Professors: Raimund J. Ober & E. Sally Ward

Rayleigh's criterion is extensively used in optical microscopy to determine the resolution of microscopes. Despite its widespread use, it is well known that this criterion is based on heuristic notions and can be surpassed in a regular optical microscope. The inadequacy of Rayleigh's criterion has necessitated a reassessment of the resolution limits of optical microscopes. The thesis proposes a new resolution criterion that overcomes the limitations of Rayleigh's criterion. The new result predicts that there is no resolution limit, but that the resolvability depends on the number of detected photons. Analytical tools are introduced to estimate the distance from microscopy images. By imaging fluorescently labeled DNA nano-rulers, it is shown that distances as small as 12 nm can be measured from experimental data with an accuracy as predicted by the new resolution criterion. The new result is derived by adopting a stochastic framework and using the theory concerning the Fisher information matrix. This approach is generalized to a wide variety of estimation problems in optical microscopy by deriving expressions for the limits to the accuracy of the parameter estimates. As an application, the thesis addresses the location estimation problem. Analytical formulae are derived that provide a limit to the accuracy with which the location of a microscopic object can be determined. These results are illustrated by considering specific image profiles that describe the image of a single molecule.

Another contribution of this thesis is the development of a new microscopy technique called multifocal plane microscopy for tracking single molecules/particles in 3D. An important property of this technique is its improved depth discrimination capability, which in turn enables accurate determination of the axial location of the particle especially when it is close to the plane of focus.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT v

LIST OF FIGURES x

LIST OF TABLES xii

1. INTRODUCTION 1

1.1 Introduction 1

1.2 Single molecule microscopy 2

1.3 Techniques for imaging single molecules 3

1.4 Challenges in single molecule imaging 5

1.4.1 Single molecule localization accuracy 6

1.4.2 3D single molecule tracking 7

1.4.3 Rayleigh's criterion and single molecule imaging 8

1.5 Overview of the thesis 9

2. GENERAL RESULTS 14

2.1 Introduction 14

2.2 General stochastic framework 14

2.3 Effects of reduced detector size 20

2.4 Effects of pixelation 22

2.4.1 Loss of information due to pixelation 25

2.4.2 Effect of image addition 27

2.5 Appendix to Chapter 2 29

2.5.1 Sum of a discrete and a continous random variable 37

3. LOCALIZATION ACCURACY PROBLEM 41

3.1 Introduction 41

3.2 Performance limits and Fisher information matrix 42
3.3 Image function 49
3.3.1 Examples 53
3.3.2 Upper and lower bounds to the performance limits 58
3.3.3 Examples 63
3.4 Effects of pixelation and noise 68
3.4.1 Examples and applications 72
3.5 Appendix to Chapter 3 82
3.5.1 Simulations and parameter values 82
3.5.2 Maximum likelihood estimation 82
3.5.3 Random number generation 83
4. MULTIFOCAL PLANE MICROSCOPY & 3D PARTICLE TRACKING 85
4.1 Introduction 85
4.2 Principle of MUM 87
4.2.1 Construction of MUM 89
4.2.2 Demonstration of simultaneous multifocal plane imaging 90
4.2.3 Determination of focal plane separation 90
4.3 Depth discrimination and 3D localization measure 92
4.4 Verification of improved depth discrimination capability 94
4.5 Appendix to Chapter 4 96
4.5.1 Determination of appropriate plane spacing 96
4.5.2 Imaging experiments 98
4.5.3 Cells, transfection and reagents 98
4.5.4 Data analysis 98
4.5.5 Fisher information matrix for a MUM setup 100
5. REDEFINING THE RESOLUTION LIMITS IN OPTICAL MICROSCOPY 105

5.1 Introduction . . . 105
5.2 Fundamental resolution measure (FREM) . . . 106
5.3 Practical resolution measure (PREM) . . . 109
5.4 Improving the resolution measure . . . 111
5.5 Experimental verification . . . 111
5.6 Generalization of the FREM . . . 113
5.7 Discussion . . . 114
5.8 Appendix to Chapter 5 . . . 119
5.8.1 Single molecule microscopy . . . 119
5.8.2 Derivation of the new resolution measure . . . 119
5.8.3 Effects of pixelation and noise . . . 122
5.8.4 Additional spatial information . . . 127
5.8.5 Maximum-likelihood estimation . . . 128
REFERENCES . . . 134
BIOGRAPHICAL STATEMENT . . . 144

LIST OF FIGURES

Figure Page

1.1 Image of single DNA molecules 4

2.1 Schematic of an optical microscope based imaging setup 15

3.1 Fundamental limit to the localization accuracy of x_0 for a GFP single molecule 57

3.2 Effect of finite detector size on the limit of the 2D localization accuracy 67

3.3 Effect of magnification on the limit of the 2D localization accuracy 73

3.4 Effects of pixel array size and extraneous noise sources on the limit of the 2D localization accuracy 74

3.5 Effect of pixel size on the limit of the 2D localization accuracy 79

3.6 Effect of single molecule loaction on the limit of the 2D localization accuracy 80

4.1 Principle of MUM 87

4.2 Schematic of a MUM setup that can image upto four distinct planes 88

4.3 Calibration of MUM 91

4.4 Improved depth discrimination capability of MUM 92

4.5 Results of defocus level estimation for simulated and experimental data 95

5.1 Behavior of the FREM as a function of distance and photon count 108

5.2 Behavior of the PREM as a function of distance and photon count 131

5.3 Results of distance estimation for closely spaced Cy5 molecules 132

5.4 Results of distance estimation for fluorescently

labeled DNA nanorulers 133

5.5 Coordinate system used for the derivation of the generalized PREM 133

LIST OF TABLES

Table Page

3.1 Verification of the limit of the 2D localization accuracy - practical imaging conditions 77

4.1 Results of global estimation from MUM data 97

5.1 Experimental verification of the new resolution measure 113

CHAPTER 1

INTRODUCTION

1.1 Introduction

The imaging of cellular events with optical microscopes has been an indispensable tool in biological research. Ever since the invention of the first compound microscope over three centuries ago, optical microscopy has undergone several revolutionary changes. In the more recent past, a variety of technological developments have led to a significant expansion of the capabilities of optical microscopes. On the one hand, the advent of highly sensitive photon counting imaging detectors, high speed computer control of imaging experiments and automated image processing techniques have led to major improvements in microscopy hardware. On the other hand, the introduction of genetic labeling approaches, in particular the use of green fluorescent protein (GFP) fusion constructs, have allowed the specific labeling of proteins for imaging their behavior in a live-cell environment. These advances have enabled several groups to image biomolecular interactions at the single molecule level even within a cellular environment (see [1, 2, 3, 4, 5, 6, 7]).

The study of cellular processes at the single molecule level is motivated by several considerations. Traditionally, cellular imaging studies using fluorescently labeled proteins were bulk studies, since an accumulation of proteins was required to yield a detectable fluorescence signal. However, the information from such studies is averaged over ensembles of molecules (i.e., hundreds and thousands of molecules) thereby masking subtle individual variations. These variations become particularly important when the system under study is heterogeneous as in the case of live cells. For example, heterogeneity arises due to the dynamic nature of protein-protein interactions wherein a cell membrane protein could exist in multimeric states (i.e., monomer, dimer, trimer etc). Single molecule experiments remove the ensemble averaging that is characteristic of bulk measurements and provide information on the behavior of individual molecules.

Single molecule studies also remove the need for synchronization of many copies of individual biomolecules that are involved in a time dependent process. For instance, in the study of protein conformation within a cellular environment the individual copies of a protein could exist in different confirmations due to being in different folded states. Thus it is believed that imaging biomolecular interactions at the single molecule level holds the promise that significant new insights can be gained.

1.2 Single molecule microscopy

One of the initial studies on single molecule detection was carried out in the late 1980s by Moerner and colleagues who observed single molecules in a solid material ([8]). In the last two decades there has been tremendous progress in this field leading to the development of several imaging approaches to enable single molecule detection in biological samples at room temperature (see, for example, [9, 10] and references therein). Common to all these techniques is that the biomolecule of interest has to be fluorescently labeled in order to be imaged. The fluorescent dye is excited with a light source at a specific wavelength range. As a result of the excitation, the dye emits photons at a higher wavelength, which is then detected as the signal of interest. Classically, biomolecules were chemically coupled to fluorescent dye molecules. In a cellular context, this approach, however, lacked specificity and moreover subjected the cells to harsh conditions. Hence they were not suitable for live cell imaging studies. On the other hand, the labeling technology based on GFP provides a highly specific approach to label the protein of interest. GFP is a naturally occurring fluorescent protein which was first isolated from the jelly fish *Aequorea Victoria* ([11]). By using DNA cloning techniques, GFP can be expressed in cells as a fusion protein that is tagged to the protein of interest. This obviates the need for chemical coupling of fluorescent dyes and paves the way for live cell imaging. Presently several mutants of GFP exist which fluoresce at spectrally distinct wavelengths (for example, Red Fluorescent Protein (RFP), Cyan Fluorescent Protein (CYP), Yellow Fluorescent Protein (YFP) etc.).

Single molecule experiments place several requirements on the properties of fluorophores ([9, 12]). They must 1) be bright (i.e., have high extinction coefficient and high quantum yield) so that their signal is above background, 2) absorb and emit light in the visible to near-infrared wavelengths, 3) not perturb or affect the functionality of the biomolecule to which they are tagged, 4) be available in a form suitable for covalent conjugation to the molecule of interest, 5) show little fluctuations in the emitted light intensity during the course of experiment, and 6) be photostable, i.e., emit photons upon excitation for long periods of time. Most fluorescent dyes are susceptible to photobleaching, which is a phenomenon in which subsequent to several rounds ($10^5 - 10^9$) of excitation and emission, the fluorophores lose their ability to emit photons when excited. This means that after a certain amount of time, the fluorescent dye and hence the labeled biomolecule can no longer be imaged. As a result of this, proteins can only be tracked for relatively short periods of time, which are typically in the order of a few seconds in a live cell environment, when the sample is continuously illuminated. Recently, a new class of fluorescent labels, known as quantum dots, has been developed to overcome the limitations of conventional fluorescent dyes ([13, 14]). Quantum dots are extremely photostable under continuous illumination (they emit photons for hours) and are also significantly brighter when compared to conventional fluorescent dyes ([14]). However, unlike GFP, quantum dots have to be tagged to the protein of interest through chemical coupling or through avidin-biotin labeling. Moreover, many problems need to be overcome for quantum dots to be usable in a routine fashion, such as the multivalency of the current streptavidin coating/labeling approach, non-specific binding, blinking of the quantum dots etc.

1.3 Techniques for imaging single molecules

The different techniques used for imaging single molecules can be classified as confocal and widefield. In confocal methods, the light source is focused to a diffraction limited spot illuminating a very small volume ($10^{-10} - 10^{-12}$ cm^3) within the sample, from which the fluorescence signal is collected. For example, fluorescence correlation spectroscopy (FCS) uses this

method to estimate diffusion coefficients of the labeled molecules ([15]). Confocal techniques typically use detectors that have very high temporal resolution (10^{-9} s). For imaging applications, the focused spot is sequentially scanned at each point on the sample. Hence when the sample is being scanned at a particular location, important events can be missed in the other parts of the sample. Thus, confocal imaging techniques do not provide spatial information.

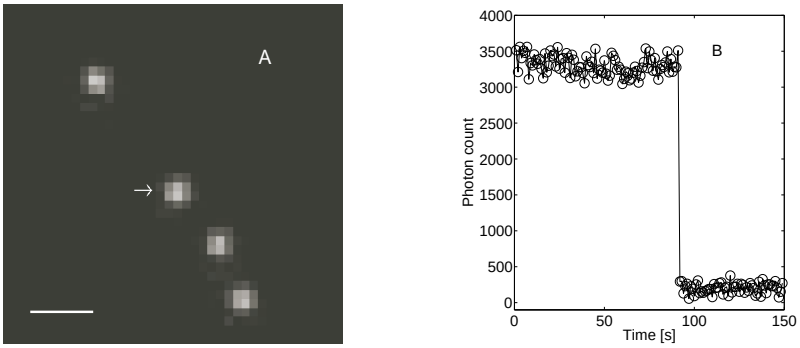

Figure 1.1. Image of single DNA molecules.
Panel A shows the image of four DNA molecules labeled with Cy5 dye. A cooled CCD camera (ORCA-ER, Hamamatsu, Bridgewater, NJ) was used to acquire the image with an exposure time of 1 s. Scale bar equals 1 μm. Panel B shows the single step photobleaching behavior of the single molecule that is indicated by an arrow in Panel A. The photon count is calculated by summing the pixel values of a 5 $\times$ 5 array that is centered at the brightest pixel in the image of the single molecule.

Widefield imaging methods use high resolution, high sensitivity imaging detectors such as cooled CCD cameras, which are capable of imaging multiple single molecules simultaneously that are in the field of view (see Fig. 1.1A). In cellular imaging applications, due to the low signal level, CCD cameras are often used in conjunction with image intensifiers ([16]). More recently, CCDs based on electron multiplication technology have shown significant promise ([17, 18]). In a live cell imaging application, widefield imaging techniques acquire the image of the entire cell thereby providing important spatial information. Although the temporal resolution of CCD based imaging detectors is limited ($10^{-3} - 10^{-1}$ s), it is usually high enough to record typical cellular events and to track single molecules in live cells.

Total internal reflection fluorescence microscopy (TIRFM) ([19]) is a widefield imaging technique that is used to image cellular events close to the plasma membrane. In this technique, the illumination intensity decreases exponentially with the distance from the cover glass that supports the cell. This has significant advantages in the study of membrane events. Often intracellular components are very strongly labeled. With a standard illumination approach these intracellular components produce strong signals that can completely overwhelm events at the plasma membrane. With TIRFM, only the membrane adjacent to the cover glass is illuminated. Most fluorophores in the intracellular compartments are not excited and can therefore not contribute to background signal at the plasma membrane. This is often essential to be able to detect single molecule events at the plasma membrane proximal to the coverslip.

One of the important aspects of single molecule imaging is the verification that the detected signal does indeed arise from a single molecule rather than from the accumulation of a large number of molecules. There are a number of ways of carrying out this verification (see [20]). The most common approach is to use the single step photobleaching property of single molecules. For a single molecule a constant signal level for the emitted fluorescence is expected that drops in one (quantum) step to the background level at photobleaching (see Fig. 1.1B). This analysis can be complicated to some extent due to the *blinking* behavior that some fluorophores exhibit, such as GFP ([21]). In this behavior the fluorophore switches back and forth between emitting photons and the dark state.

1.4 Challenges in single molecule imaging

Single molecule experiments place extraordinary demands on the capabilities of current microscope setups and data analysis tools. There are severe technical problems that limit the study of single molecules in particular in a live cell situation. The main reason for this limitation is the high level of background noise that obscures the signal from the single molecule. The background noise can be attributed to a variety of sources such as autofluorescence of the cell, buffer etc. In addition, scattered photons that randomly reflect off optical components in the

light path also contribute to the background signal. Aside from this, the dark current and the readout noise that originates from the detector also contribute to the noise in the acquired data (see [22]). Even without any noise sources the analysis of single molecule images would not be devoid of problems, since the acquired data is stochastic due to photon emission being a random process. Thus this means that detecting single molecules requires a very careful experimental setup. However, even under the best circumstances the signal to noise ratio is very unfavorable ([16]). Therefore there is a considerable need for advanced signal and image processing methods that allow the available information to be extracted from the experimental data as accurately as possible. This is particularly so if quantitative results are to be obtained. In the following sections some of the key problems encountered in single molecule imaging are discussed.

1.4.1 Single molecule localization accuracy

One of the fundamental problems in single molecule imaging concerns the accuracy with which the location of a single molecule can be determined. Specifying the accuracy with which the location of a single molecule can be established is not only of importance to be able to characterize the level of accuracy that is achievable in single-molecule microscopy. The accuracy with which a single molecule can be localized has significant influence on the type of studies that can be carried out using single molecule microscopy. It is also of significance in the analysis of single-molecule data. For instance, it has been shown in [23] that the accuracy of the location estimates has to be taken into account when analyzing the diffusion behavior of single molecules. Otherwise noisy measurements of the single molecule locations could lead to the erroneous interpretation that subdiffusional behavior is present even though this is not the case. Earlier approaches to the characterization of the localization accuracy mainly relied on an approach by Bobroff ([24]) in which the localization accuracy problem was examined using the least-squares criterion (see, for example, [25, 26, 27, 28, 29]). This criterion is ideally suited to estimate parameters from data that has a Gaussian probability distribution. It does, however, appear problematic to assume that single molecule data is in fact Gaussian distributed, since

the photon detection process is typically modeled as a shot-noise (Poisson) process ([30, 31]). Aside from the reliance on the least-squares algorithm other approximations are made in [24] in the derivation of the result that are often difficult to verify. Moreover, in the application of those results to single molecule microscopy the image profile of a single molecule predicted by standard diffraction theory is often replaced by a Gaussian profile.

1.4.2 3D single molecule tracking

The study of bio-molecular interactions in cells is of fundamental importance to understanding cellular processes. The advent of single molecule/particle imaging technology has generated significant interest in studying these interactions at the individual biomolecule level ([9]). Presently, widefield microscopy techniques are extensively used for imaging single molecules within a cellular environment ([32, 33]). These techniques are well suited to study fast moving single molecules in two dimensions. In the imaging modality that is typically used at present only one focal plane is imaged at any given point in time. Single particle tracking in a 3D environment such as a cell is therefore problematic since the particle can easily move out of the focal plane that is currently being imaged. This is of particular relevance for the tracking of fast intracellular dynamics that are not confined to one plane. Another shortcoming of widefield microscopes is their poor depth discrimination capability ([34, 35]). For instance, when imaging cellular specimens in a widefield microscope the acquired image contains photons from the region that is in focus with respect to the objective lens as well as from regions that are out of focus with respect to the objective lens. The in focus region appears sharp and clear while the out of focus regions appear hazy, which deteriorates the image quality and obscures the in focus component. Due to this, images acquired in a widefield microscope provide inadequate depth information ([35]). In the case of single particle tracking, this inadequacy becomes pronounced. It has been shown that in a widefield microscope, there exists significant uncertainty in determining the axial location of the single particle, especially when it is close to the plane of focus ([36, 37]).

Several techniques have been introduced in the past to address these and related problems in the context of single particle tracking in three dimensions. One type of approach is based on the scanning of the sample in 3D ([38, 39, 40, 41]). Successful implementations have been reported, but they are typically limited to the tracking of one or very few particles. In addition the approach has limitations in speed and the implementation is reliant on sophisticated online processing schemes. Another type of approach is based on changing the point spread function of the acquisition system and thereby either producing a much larger depth of focus or by encoding the 3D position of the particle ([42, 43]). Increasing the depth of focus essentially makes it even more difficult to obtain information on the axial position of the particle. The approach based on encoding the 3D position requires repetitive defocusing and refocusing of the sample, which compromises the temporal resolution and could also lead to alignment problems.

1.4.3 Rayleigh's criterion and single molecule imaging

Rayleigh's criterion is extensively used in optical microscopy for determining the 2D resolution of microscopes. According to this criterion, the minimum resolvable distance between two identical point sources that lie in the plane of focus is $0.61\lambda/n_a$, where λ denotes the wavelength of the point source and n_a denotes the numerical aperture of the objective lens. The advent of single molecule microscopy has generated significant interest in studying nanoscale biomolecular interactions. Classically, fluorescence resonance energy transfer-based methods have been used to probe interactions in the distance range of 1 to 10 nm ([44]). It is widely believed that Rayleigh's criterion precludes the resolution of two single molecules at distances of 10-200 nm, which leaves a gap in the distance range of 10 to 200 nm that is vital for the study of many biological processes with an optical microscope. However, it is well known that Rayleigh's criterion is based on intuitive notions ([16]). For example, Rayleigh's criterion is formulated in a deterministic setting that neglects the photon statistics of the acquired data. Hence it does not take into account the number of detected photons, which, in turn, raises concern over the use of Rayleigh's criterion in photon-counting techniques such as single-molecule microscopy. It

has been suggested that Rayleigh's resolution limit can be superseded when objects are imaged with modern detectors and when apriori information in conjunction with parameter estimation approaches are used to analyze the acquired data ([45, 46, 47]). Not surprisingly, recent single molecule experiments have shown that distances well below Rayleigh's limit can be measured in an optical microscope setup ([48, 49, 50]). Thus Rayleigh's resolution limit is inadequate for modern imaging techniques. This inadequacy has necessitated a reassessment of the resolution limits of optical microscopes. In particular, for proper planning of an experiment, it is important to have a methodology available to be able to assess with what accuracy the distance between two point sources (single molecules) can be determined.

1.5 Overview of the thesis

In single molecule microscopy, parameter estimation approaches play an important role in data analysis. Examples relate to determining the location of single molecules ([24, 26, 28, 51]), determining the distance between two or more single molecules ([48, 49, 50, 52]), determining the photon detection rate, estimating the defocus level of single molecules ([36, 37, 43, 53]) etc. Common to all these problems is that it is helpful for the experimenter to have an analytical method to assess with what accuracy the various parameters can be estimated. This in turn has generated significant interest in quantifying the capabilities of optical microscopes. The process of image acquisition in microscopy is inherently a stochastic one. For example, upon excitation a fluorophore emits photons in a stochastic manner. Therefore the data acquired in a fluorescence microscopy experiment is stochastic in nature. As a result questions concerning the capabilities of microscopes are best addressed using a stochastic approach.

Here, by adopting a rigorous stochastic framework, analytical expressions are obtained to calculate performance limits that quantify the capabilities of optical microscopes. The underlying approach to calculating the performance limits is based on the theory concerning the Fisher information matrix ([54, 55]), which plays a central role in the analysis of estimation algorithms.

The organization of the thesis is as follows. Chapter 2 contains the detailed mathematical derivation of results for a general parameter estimation problem encountered in optical microscopy. The scope of these results are broad and are applicable to a wide variety of imaging conditions such as TIRF illumination, polarized excitation/emission etc., and can be extended to several microscopy modalities such as confocal/multi-photon microscopy, brightfield microscopy, phase-contrast/DIC microscopy etc. Analytical expressions for the Fisher information matrix are derived for image models that allow both stationary and moving objects. The effects of deteriorating experimental factors such as pixelation of the detector and additive noise sources ([30, 31]) are also investigated and general expressions for the Fisher information matrix are derived that take into account these factors. The results derived in Chapter 2 can be used to obtain limits to the accuracy for a large class of estimation problems. These results are illustrated in the subsequent chapters by considering specific problems in single molecule microscopy and data analysis.

Chapter 3 deals with the localization accuracy problem, which is concerned with how accurately the location of a microscopic object (for example, single molecules) can be determined when imaged through a fluorescence (optical) microscope. Here, specific emphasis is placed on the derivation of conditions that guarantee block diagonal or diagonal Fisher information matrices, as this has several implications in parameter estimation (see Section 3.2 for details). Throughout the chapter, examples are provided to illustrate the results by considering specific image profiles that describe the image of an in focus single molecule. In particular, two image profiles are considered, namely, the Airy profile and the 2D Gaussian profile. According to optical diffraction theory, the image of an in focus point source that emits incoherent, unpolarized light is given by the Airy profile ([56]). On the other hand, in single molecule applications the 2D Gaussian profile is widely used to approximate the Airy profile ([24, 28, 26]). An unexpectedly simple formula is derived that provides a fundamental limit to the 2D localization accuracy of a single molecule (see Corollary 3.3.2 for derivation). A detailed analysis is presented on the effects of experimental parameters such as pixel size, magnification, detector size

and noise sources, on the limit of the 2D localization accuracy. The implications of these results in single molecule imaging are discussed and guidelines for experimental design are provided. A maximum likelihood estimation algorithm is presented for determining the 2D location of single molecules and is tested on simulated data. It is shown that the standard deviation of the maximum likelihood estimates of the location comes consistently close to the predicted limit of the localization accuracy for a wide variety of imaging conditions.

Chapter 4 discusses the development of a new microscope called multifocal plane microscope (MUM), and the application of this microscope for 3D single particle tracking. The design and construction of MUM is described. One of the important properties of MUM is its improved depth discrimination capability. As a result, the axial location of the single particle can be determined with relatively high accuracy. An estimation technique is introduced for determining the 3D particle locations from MUM images, and is tested on simulated and experimental data. Using the results derived in Chapter 2, analytical expressions are obtained for the limits to the accuracy with which the 3D particle locations can be estimated from MUM data. These analytical expressions validate the improved depth discrimination capability of the MUM and provide guidelines for optimizing the MUM setup for 3D single particle tracking.

Chapter 5 deals with the resolution problem. Here, a new 2D resolution measure is derived which overcomes the limitation of classical Rayleigh's criterion. The new result is referred to as the fundamental resolution measure (FREM). Unlike Rayleigh's criterion, the FREM predicts that there is no resolution limit, but that the resolvability depends on the number of photons collected from the point sources. Analytical expressions for the new resolution measure are obtained that take into account the effect of deteriorating experimental factors such as pixelation of the detector and extraneous noise sources. Experimental verification of the resolution measure is carried out by imaging closely spaced single molecule pairs and measuring distances that are above and well below Rayleigh's limit. In particular by imaging fluorescently labeled DNA nanorulers, distances as small as 12 nm are measured with an accuracy as predicted by the new resolution measure. The implications of the new resolution measure for single

molecule imaging are discussed. Numerous examples are provided to illustrate how the new result can be used to design and optimize single molecule experimental setups for distance measurement.

The following is a list of papers published during the course of the PhD work.

Journal Publications

J1. R. J. Ober, **S. Ram** and E. S. Ward, Localization accuracy in single molecule microscopy, *Biophysical Journal*, vol. 85, pp. 1185-1200, 2004.

J2. P. Prabhat, **S. Ram**, E. S. Ward and R. J. Ober, Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions, *IEEE Transactions on Nanobioscience*, vol. 3, pp. 237-242, 2004.

J3. **S. Ram**, E. S. Ward and R. J. Ober, A stochastic analysis of the performance limits of optical microscopes, *Multidimensional Systems and Signal Processing*, vol. 17, pp. 27-57, 2006.

J4. **S. Ram**, E. S. Ward and R. J. Ober, Resolution beyond Rayleigh's criterion: a new resolution measure with applications to single molecule microscopy, *Proceedings of the National Academy of Sciences*, vol. 103, pp. 4457-4462, 2006.

J5. P. Prabhat, Z. Gan, J. Chao, **S. Ram**, C. Vaccaro, S. Gibbons, R. J. Ober and E. S. Ward, Elucidation of intracellular recycling pathways leading to exocytosis of the Fc receptor, FcRn, by using multifocal plane microscopy, *Proceedings of the National Academy of Sciences*, vol. 104, pp. 5889-5894, 2007.

Conference Publications

C1. **S. Ram**, E. S Ward and R. J. Ober, How accurately can a single molecule be localized when imaged through an optical microscope?, *Proceedings of the IEEE International Symposium on Biomedical Imaging*, vol. 2, pp. 1087-1090, 2004.

C2. **S. Ram**, E. S. Ward and R. J. Ober, How accurately can a single molecule be localized in three dimensions using a fluorescence microscope?, *Proceedings of the SPIE*, vol. 5699, 426-435, 2005.

C3. P. Prabhat, **S. Ram**, E. S. Ward and R. J. Ober, Simultaneous imaging of several focal planes in fluorescence microscopy for the study of cellular dynamics in 3D, *Proceedings of the SPIE*, vol. 6090, pp. L1-L7, 2006.

C4. **S. Ram**, E. S. Ward and R. J. Ober, A novel stochastic resolution criterion for fluorescence microscopes, *Proceedings of the SPIE*, vol. 6090, pp. J1-J6, 2006.

C5. **S. Ram**, E. S. Ward and R. J. Ober, A novel resolution measure for optical microscopes: stochastic analysis of the performance limits, *Proceedings of the IEEE International Symposium on Biomedical Imaging*, vol. 3, pp. 770-773, 2006.

C6. **S. Ram**, J. Chao, P. Prabhat, E. S. Ward and R. J. Ober, A novel approach to determining the three-dimensional location of microscopic objects with applications to 3D particle tracking, *Proceedings of the SPIE*, vol. 6443, pp. D1-D7, 2007.

C7. **S. Ram**, A. V. Abraham. E. S. Ward and R. J. Ober, A novel 3D resolution measure with applications to single molecule imaging, *Proceedings of the SPIE*, vol. 6444, pp. D1-D9, 2007.

CHAPTER 2

GENERAL RESULTS

2.1 Introduction

This chapters deals with the detailed derivation of results for a general parameter estimation problem in optical microscopy. The results given in this chapter are broad in scope and provide a framework to calculate the performance limit for a wide variety of estimation problems. The applications of these results are given in the subsequent chapters. The organization of this chapter is as follows. Section 2.2 describes the rigorous stochastic framework that is used to model the acquired data and contains the main result, i.e., the derivation of the Fisher information matrix for a general parameter estimation problem in optical microscopy. Section 2.3 deals with the effect of detector size and investigates how this affects the performance limits. Section 2.4 considers the effect of deteriorating experimental factors such as pixelation and noise sources. Analytical expressions for the Fisher information matrix are obtained that take into account these factors. In Section 2.4.1, a closed form expression is obtained that quantifies the loss of information due to pixelation and noise sources. In many practical cases, the acquired images are summed to improve the signal content. Section 2.4.2, investigates the effect of image addition and analytical expressions of the Fisher information matrix are derived for a general parameter estimation problem when using summed data. The results presented in this chapter have been published in [51, 57].

2.2 General stochastic framework

A basic optical microscope setup consists of an object located in the object space, a lens system and a detector in the image space that captures the image of the object (see Fig. 2.1). For example, the object could be an individual point source (for example, a fluorescent single molecule or a fluorescent nano-particle), a collection of two or more point sources, or

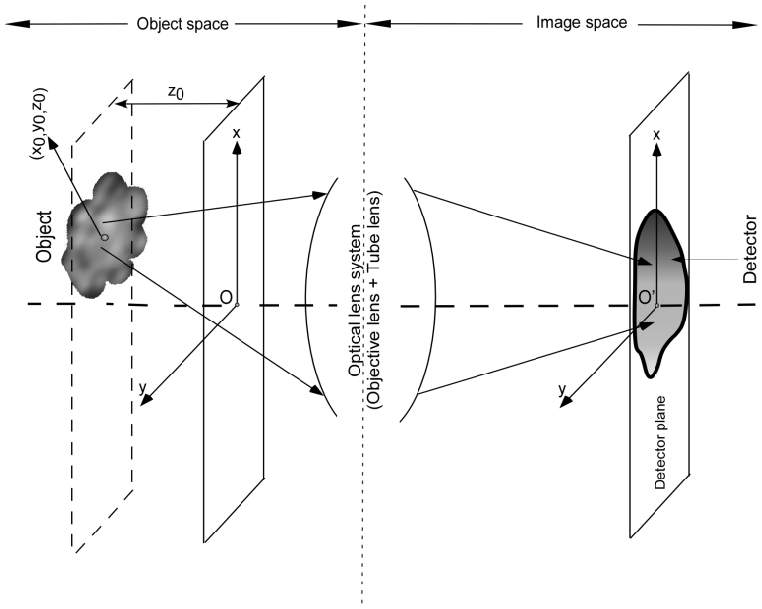

Figure 2.1. Schematic of an optical microscope based imaging setup.
Here, an object located in the object space is imaged by an optical lens system and the image of the object is captured by the detector that is located in the image space.

a fluorescently labelled cellular organelle. Here, we are primarily interested in experiments in which the detector detects photons from the object of interest for a fixed acquisition time. Since the photon detection process is inherently a random phenomenon (see, for example, [58]), the recorded image of the object is stochastic in nature.

We assume that the acquired data consists of the spatial coordinates of the arrival location of the detected photons on the detector and the time points at which the photons are detected. In a typical quantitative experiment, some attributes of the object such as the location, distance of separation from other objects, orientation, size etc., are determined from the acquired data by using a specific estimation procedure. The accuracy of the estimates can be determined by calculating the standard deviation of the estimates of this attribute assuming repeated experiments.

In any estimation problem, it is important to know whether the specific estimation technique used to estimate the desired attribute indeed comes close to the best possible accuracy. This can be determined by calculating the Fisher information matrix ([55, 59]) for the un-

derlying random process that characterizes the acquired data. According to the Cramer-Rao inequality ([55, 59]), the (co)variance (matrix) of any unbiased estimator $\hat{\theta}$ of an unknown vector parameter θ is bounded from below by the inverse of the Fisher information matrix $\mathbf{I}(\theta)$, i.e.,

$$\mathrm{Cov}(\hat{\theta}) \geq \mathbf{I}^{-1}(\theta).$$

Since we have defined the accuracy of an estimator in terms of its standard deviation, the square root of the inverse Fisher information matrix provides a lower bound to the best possible accuracy. It is important to note that the Fisher information matrix is independent of the estimation procedure used to estimate the parameter θ and only depends on the statistical nature of the acquired data. For instance, if the desired attribute is the location of an object, then the above equation implies that for any (asymptotically) unbiased estimator of the location, the accuracy of its location estimates can never be smaller than the square root of the inverse Fisher information matrix. Therefore, the square root of the inverse Fisher information matrix provides a *limit* to the accuracy with which the location of the object can be determined. Generalizing this, in an optical microscope the performance limit in determining a specific attribute of an object is defined as the square root of the inverse Fisher information matrix calculated for that attribute.

Due to its stochastic nature, the acquired data is modeled as a space-time random process (see, for example, [60]) which we refer to as the image detection process $\mathcal{G}$. The temporal part of $\mathcal{G}$ describes the time points of the detected photons and is modeled as a temporal Poisson process with intensity function Λ_θ. The spatial part of $\mathcal{G}$ describes the spatial coordinates of the arrival location of the detected photons and is modeled as a family of mutually independent random variables $\{U_\tau\}_{\tau \geq t_0}$ with probability densities $\{f_{\theta,\tau}\}_{\tau \geq t_0}$ defined on the detector $\mathcal{C}$, where τ denotes the time point of a detected photon. The time dependence of the random variables $\{U_\tau\}_{\tau \geq t_0}$ denotes the fact that the spatial distribution of the detected photons can change with time. For example, this is the case when photons from a moving object are detected. In some applications the spatial part of $\mathcal{G}$ is independent of τ and in that case the random

variables are independent and identically distributed. In all cases, we assume that the spatial and temporal parts of $\mathcal{G}$ are mutually independent of each other. We note that the probability density $f_{\theta,\tau}$ satisfies certain regularity conditions that are necessary for the calculation of the Fisher information matrix (see [55]).

Definition 2.2.1 *Let $\mathcal{C}$ denote a* **detector**, *i.e., an open subset of $\mathbb{R}^2$ with non-zero Lebesgue measure. Let Θ denote the* **parameter space** *that is an open subset of $\mathbb{R}^n$ and let $t_0 \in \mathbb{R}$. For $\theta \in \Theta$, an* **image detection process** *$\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ is defined as a spatio-temporal process whose temporal part describes the time points of the photons detected on the detector $\mathcal{C}$ and the spatial part describes the spatial coordinates of the arrival location of the photons detected on the detector $\mathcal{C}$.*

The temporal part is modeled as a Poisson process $\{Z(\tau); \tau \geq t_0\}$ with intensity Λ_θ, called the **photon detection rate**, *such that*

C1. *$\Lambda_\theta(\tau)$ is piecewise continuously differentiable with respect to θ for each $\tau \geq t_0$.*

C2. *$\Lambda_\theta(\tau)$ is piecewise continuous with respect to τ for each $\theta \in \Theta$.*

Let $\mathcal{F}_\Theta$ be the set of probability densities f_θ on $\mathcal{C}$ parameterized by θ that satisfy the following regularity conditions

C3. *$\partial f_\theta(r)/\partial\theta_i$ exists for $r \in \mathcal{C}$, $i = 1, \ldots, n$ and $\theta \in \Theta$.*

C4. *$\int_{\mathcal{C}} |\partial f_\theta(r)/\partial\theta_i| dr < \infty$ for $i = 1, \ldots, n$ and $\theta \in \Theta$.*

C5. *The integral $\int_{\mathcal{C}} \frac{1}{f_\theta(r)} \frac{\partial f_\theta(r)}{\partial\theta_i} \frac{\partial f_\theta(r)}{\partial\theta_j} dr$ exists and is finite for $i, j = 1, \ldots, n$ and $\theta \in \Theta$.*

The spatial part of the image detection process is modeled as a family of mutually independent random variables $\{U_\tau\}_{\tau \geq t_0}$ that is assumed to be independent of $\{Z(\tau); \tau \geq t_0\}$. The corresponding family of probability densities $\{f_{\theta,\tau}\}_{\tau \geq t_0} \subseteq \mathcal{F}_\Theta$ is called the **photon distribution profile**, *if U_τ with probability density $f_{\theta,\tau}$ describes the spatial distribution of the location of the point of detection of a photon on the detector $\mathcal{C}$ that is detected at time τ, $\tau \geq t_0$.*

We next derive an expression for the Fisher information matrix of the image detection process $\mathcal{G}$.

Theorem 2.2.1 *Let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ be an image detection process. Then for $\theta \in \Theta$ the Fisher information matrix $\mathbf{I}(\theta)$ of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by*

$$\mathbf{I}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right)^T \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right) d\tau + \int_{t_0}^{t}\int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right)^T \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right) dr d\tau$$

$$= \int_{\mathcal{C}}\int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta}\right)^T \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta}\right) dr d\tau.$$

Proof: This result is a generalization of the Fisher information matrix for a spatio-temporal random process whose temporal component is a Poisson process and the spatial component is independent of the time points (see [60, pg 213]). The proof of this theorem can be found in the appendix of Chapter 2 (see Section 2.5). •

In deriving the above result we made no specific assumptions about the geometry of the imaging setup or the analytical expression for the photon distribution profile f_θ. Hence the above theorem provides a general result to calculate the Fisher information matrix for a wide range of situations. Note that the two-term expression of $\mathbf{I}(\theta)$ shows explicitly the dependence of $\mathbf{I}(\theta)$ on the temporal and spatial components of the image detection process.

We next consider the superposition of two image detection processes. In many concrete situations the detected photons originate from different sources. For example, the detected photons can result from a background component in addition to those detected from the object of interest. In an incoherent imaging setup, such as in fluorescence microscopy, the photon detection process that describes the collection of all the detected photons is then the superposition of the object and the background image detection process.

Theorem 2.2.2 *Let $\mathcal{G}^1(\Lambda^1_\theta, \{f^1_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ and $\mathcal{G}^2(\Lambda^2_\theta, \{f^2_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ be two independent image detection processes. Then the superposition process is an image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ whose photon detection rate Λ_θ is given by*

$$\Lambda_\theta(\tau) = \Lambda^1_\theta(\tau) + \Lambda^2_\theta(\tau), \quad \tau \geq t_0, \quad \theta \in \Theta, \tag{2.1}$$

and the photon distribution profile $\{f_{\theta,\tau}\}_{\tau\geq t_0}$ is given by

$$f_{\theta,\tau}(r) = \epsilon^1_\theta(\tau) f^1_{\theta,\tau}(r) + \epsilon^2_\theta(\tau) f^2_{\theta,\tau}(r), \quad r \in \mathcal{C}, \quad \theta \in \Theta, \quad \tau \geq t_0, \tag{2.2}$$

where $\epsilon_\theta^1(\tau) = \Lambda_\theta^1(\tau)/\Lambda_\theta(\tau)$, $\epsilon_\theta^2(\tau) = \Lambda_\theta^2(\tau)/\Lambda_\theta(\tau)$, $\theta \in \Theta$, $\tau \geq t_0$.

Proof: This result is analogous to the result of the superposition of Poisson processes (see, for example, [60]). The expression for the photon detection rate Λ_θ (eq. 2.1) immediately follows from the fact that the time points of all the detected photons are described by a Poisson process which is a superposition of the Poisson processes of $\mathcal{G}^1$ and $\mathcal{G}^2$. Let us suppose that the image detection process $\mathcal{G}^1$ ($\mathcal{G}^2$) describes the detected photons from source 1 (source 2). For $\tau_0 \geq t_0$ and $\delta > 0$, let I_{τ_0} denote the time interval $[\tau_0 - \delta, \tau_0 + \delta]$ and let $\mathcal{A} \subseteq \mathcal{C}$. Let D_{τ_0} denote the event that a photon is detected by the detector during the time interval I_{τ_0} and $D_{\tau_0}^1$ ($D_{\tau_0}^2$) denote the event that a photon from source 1 (source 2) is detected by the detector during the interval I_{τ_0}. By definition $D_{\tau_0} = D_{\tau_0}^1 \cup D_{\tau_0}^2$ and $P(D_n(\tau)) = P(D_{1,n}(\tau)) + P(D_{2,n}(\tau))$, since the events $D_{\tau_0}^1$ and $D_{\tau_0}^2$ are mutually exclusive. U_τ (U_τ^1, U_τ^2) be the random variable that describes the arrival location of a detected photon (from source 1, source 2) detected at time τ on the detector $\mathcal{C}$. Given that a photon is detected at time τ_0, the probability that the arrival location of this detected photon is in the set $\mathcal{A}$ is given by

$$
\begin{aligned}
P[\mathcal{U}_{\tau_0} \in \mathcal{A}] &= \lim_{\delta \to 0} P[(U_\tau \in \mathcal{A}, \tau \in I_{\tau_0}) \mid D_{\tau_0}] = \lim_{\delta \to 0} \frac{P[(U_\tau \in \mathcal{A}, \tau \in I_{\tau_0}) \cap D_{\tau_0}]}{P[D_{\tau_0}]} \\
&= \lim_{\delta \to 0} \frac{P[(U_\tau \in \mathcal{A}, \tau \in I_{\tau_0}) \cap \left(D_{\tau_0}^1 \cup D_{\tau_0}^2\right)]}{P[D_{\tau_0}]} \\
&= \lim_{\delta \to 0} \frac{P[(U_\tau \in \mathcal{A}, \tau \in I_{\tau_0}) \cap D_{\tau_0}^1]}{P[D_{\tau_0}]} + \lim_{\delta \to 0} \frac{P[(U_\tau \in \mathcal{A}, \tau \in I_{\tau_0}) \cap D_{\tau_0}^2]}{P[D_{\tau_0}]} \\
&= \lim_{\delta \to 0} P[(U_\tau^1 \in \mathcal{A}, \tau \in I_{\tau_0})] \frac{P[D_{\tau_0}^1]/\delta}{P[D_{\tau_0}]/\delta} + \lim_{\delta \to 0} P[(U_\tau^2 \in \mathcal{A}, \tau \in I_{\tau_0})] \frac{P[D_{\tau_0}^2]/\delta}{P[D_{\tau_0}]/\delta} \\
&= P[U_{\tau_0}^1 \in \mathcal{A}] \frac{\Lambda_\theta^1(\tau_0)}{\Lambda_\theta(\tau_0)} + P[U_{\tau_0}^2 \in \mathcal{A}] \frac{\Lambda_\theta^2(\tau_0)}{\Lambda_\theta(\tau_0)} \\
&= \left(\int_{\mathcal{A}} f_{\theta,\tau_0}^1(r) m(dr) \right) \epsilon_\theta^1(\tau_0) + \left(\int_{\mathcal{A}} f_{\theta,\tau_0}^2(r) m(dr) \right) \epsilon_\theta^2(\tau_0), \qquad \tau_0 \geq t_0, \qquad (2.3)
\end{aligned}
$$

where m denotes the Lebesgue measure in $\mathbb{R}^2$ and by definition of the Poisson process, for $\theta \in \Theta$, $\tau_0 \geq t_0$,

$$
\lim_{\delta \to 0} P[D_{\tau_0}^i]/\delta = \lim_{\delta \to 0} \Big(\int_{\tau_0 - \delta}^{\tau_0 + \delta} \Lambda_\theta^i(\sigma) d\sigma \Big) \exp\Big(- \int_{\tau_0 - \delta}^{\tau_0 + \delta} \Lambda_\theta^i(\sigma) d\sigma \Big)/\delta = \Lambda_\theta^i(\tau_0), \quad i = 1, 2
$$

and similarly $\lim_{\delta\to 0} P[D_{\tau_0}]/\delta = \Lambda_\theta(\tau_0)$. Note that eq. 2.3 holds for every $\mathcal{A} \subseteq \mathcal{C}$ and $\tau_0 \geq t_0$. Hence the probability measure $P[U_{\tau_0} \in \mathcal{A}]$ is absolutely continuous with respect to the Lebesgue measure m and there exists a probability density function $f_{\theta,\tau}$ such that

$$f_{\theta,\tau}(r) = \epsilon^1_\theta(\tau) f^1_{\theta,\tau}(r) + \epsilon^2_\theta(\tau) f^2_{\theta,\tau}(r), \quad r \in \mathcal{C}.$$

Since, by definition $f^1_{\theta,\tau}$ and $f^2_{\theta,\tau}$ satisfy conditions **C3 - C5** of Definition 2.2.1, it can be verified that $f_{\theta,\tau}$ also satisfies these conditions. •

2.3 Effects of reduced detector size

The results derived in the previous section assume that the photon distribution profile $\{f_{\theta,\tau}\}_{\tau\geq t_0}$ is defined on the detector $\mathcal{C}$ which is used to acquire the data. In many cases, the photon distribution profile is defined on the two dimensional Euclidean space ($\mathbb{R}^2$). However, in practice, microscopy images are acquired with finite sized detectors. Moreover, when analyzing microscopy images typically only small regions of interest are used. This raises the question of how the detector size or the region of interest influences the performance limits.

In the following proposition we show how an image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ has to be adjusted when instead of the detector $\mathcal{C}$, the photons are detected on a reduced part $\mathcal{C}^{rd}$ of $\mathcal{C}$, i.e., on an open subset $\mathcal{C}^{rd}$ of $\mathcal{C}$.

Proposition 2.3.1 *Let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ be an image detection process and let $\mathcal{C}^{rd} \subseteq \mathcal{C}$ be open. For $\theta \in \Theta$ and $\tau \geq t_0$, let $\alpha_{\theta,\tau} := \int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r)dr$. The time points and the spatial coordinates of the arrival location of the photons detected on the reduced detector $\mathcal{C}^{rd}$ are described by an image detection process $\mathcal{G}^{rd}$ whose photon detection rate Λ^{rd}_θ and photon distribution profile $f^{rd}_{\theta,\tau}$ are given by*

$$\Lambda^{rd}_\theta(\tau) = \alpha_{\theta,\tau}\Lambda_\theta(\tau), \quad \tau \geq t_0, \quad \theta \in \Theta, \qquad f^{rd}_{\theta,\tau}(r) = \frac{1}{\alpha_{\theta,\tau}} f_{\theta,\tau}(r), \quad r \in \mathcal{C}^{rd}, \quad \theta \in \Theta, \quad \tau \geq t_0.$$

Proof: By definition of the image detection process $\mathcal{G}$, the time points of the detected photons on the detector $\mathcal{C}$ are modeled as a Poisson process with intensity function Λ_θ. It then follows

that the time points of the detected photons on the detector $\mathcal{C}^{rd}$ form a Poisson process with intensity function $\alpha_{\theta,\tau}\Lambda_\theta$, $\tau \geq t_0$, $\theta \in \Theta$ ([61, pg 381]).

Let $\mathcal{A} \subseteq \mathcal{C}^{rd}$. Let U_τ denote the random variable that describes the arrival location of a photon that is detected on the detector $\mathcal{C}^{rd}$ at time τ, $\tau \geq t_0$. Then the probability that the arrival location of the detected photon is in the set $\mathcal{A}$ given that the arrival location is in the detector $\mathcal{C}^{rd}$ is given by

$$P[U_\tau \in \mathcal{A} \mid U_\tau \in \mathcal{C}^{rd}] = \frac{P[(U_\tau \in \mathcal{A}) \bigcap (U_\tau \in \mathcal{C}^{rd})]}{P[U_\tau \in \mathcal{C}^{rd}]} = \frac{P[U_\tau \in \mathcal{A}]}{P[U_\tau \in \mathcal{C}^{rd}]}$$
$$= \frac{\int_{\mathcal{A}} f_{\theta,\tau}(r)m(dr)}{\int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r)m(dr)} = \frac{\int_{\mathcal{A}} f_{\theta,\tau}(r)m(dr)}{\alpha_{\theta,\tau}},$$

where m denotes the Lebesgue measure in $\mathbb{R}^2$. Since the above equation holds for every $\mathcal{A} \subseteq \mathcal{C}^{rd}$ and $\tau \geq t_0$, the term $P[U_\tau \in \mathcal{A} \mid U_\tau \in \mathcal{C}^{rd}]$ is absolutely continuous with respect to m. Hence there exists a probability density function $f^{rd}_{\theta,\tau}$ such that

$$f^{rd}_{\theta,\tau}(r) = \frac{1}{\alpha_{\theta,\tau}} f_{\theta,\tau}(r), \quad r \in \mathcal{C}^{rd}, \quad \theta \in \Theta, \quad \tau \geq t_0.$$

Since, by definition $f_{\theta,\tau}$ satisfies conditions **C3 - C5** of Definition 2.2.1, it can be verified that $f^{rd}_{\theta,\tau}$ also satisfies these conditions. •

We refer to the image detection process $\mathcal{G}^{rd}$ as the **reduced version** of $\mathcal{G}$ corresponding to the detector $\mathcal{C}^{rd}$. We next derive a general expression for the Fisher information matrix of $\mathcal{G}^{rd}$. We also derive a formula to calculate the loss of information when a detector of reduced size is used.

Theorem 2.3.1 *Let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ be an image detection process and $\mathcal{G}^{rd}$ be the reduced version of $\mathcal{G}$ corresponding to the detector $\mathcal{C}^{rd}$, where $\mathcal{C}^{rd} \subseteq \mathcal{C}$. For $\theta \in \Theta$, let $\mathbf{I}(\theta)$ denote the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$. Then for $\theta \in \Theta$,*

1. the Fisher information matrix of $\mathcal{G}^{rd}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}_{rd}(\theta) = \int_{t_0}^{t} \int_{\mathcal{C}^{rd}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right) dr d\tau,$$

2. $$\Delta\mathbf{I}(\theta) := \mathbf{I}(\theta) - \mathbf{I}_{rd}(\theta) = \int_{t_0}^{t} \int_{\mathcal{C} \setminus \mathcal{C}^{rd}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right) dr d\tau,$$

3. $\mathbf{I}(\theta) \geq \mathbf{I}_{rd}(\theta)$.

Proof: 1. For $\theta \in \Theta$ and $\tau \geq t_0$, let $\alpha_{\theta,\tau} = \int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r)dr$. For the image detection process $\mathcal{G}^{rd}$, by Proposition 2.3.1 the photon detection rate $\Lambda_\theta^{rd}(\tau) = \alpha_{\theta,\tau}\Lambda_\theta(\tau)$ and the photon distribution profile $f_{\theta,\tau}^{rd}(r) = (1/\alpha_{\theta,\tau})f_{\theta,\tau}(r)$ for $r \in \mathcal{C}^{rd}$, $\theta \in \Theta$, $\tau \geq t_0$. Substituting for Λ_θ^{rd} and $f_{\theta,\tau}^{rd}$ in Theorem 2.2.1 the result immediately follows.

2. The result immediately follows by using the expressions for the Fisher information matrix of $\mathcal{G}$ and $\mathcal{G}^{rd}$ that are given in Theorem 2.2.1 and in part 1 of this Theorem, respectively.

3. The integrand in the integral expression of $\Delta\mathbf{I}(\theta)$ given in result 2 of this theorem is non-negative. This implies that $\Delta\mathbf{I}(\theta)$ is positive semidefinite for $\theta \in \Theta$ and from this the result follows. •

From result 1 of the above theorem we see that the expression for the Fisher information matrix of $\mathcal{G}^{rd}$ is analogous to that of $\mathcal{G}$ (see Theorem 2.2.1) with the only difference being that the region of integration of the spatial integral is now the reduced detector $\mathcal{C}^{rd}$.

2.4 Effects of pixelation

In all our results so far the detector $\mathcal{C}$ is such that the acquired data consist of the time points and the spatial coordinates of the detected photons. However, in the presence of pixelation the acquired data consist of the number of detected photons at each pixel. We next show how this data can be described in terms of the photon distribution profile and the photon detection rate of an image detection process. Let $\mathcal{G}^1(\Lambda_\theta^1, \{f_{\theta,\tau}^1\}_{\tau \geq t_0}, \mathcal{C})$ denote an image detection process that models the detected photons from the object of interest. The pixelated version of the detector $\mathcal{C}$ is defined as a collection $\{C_1, \ldots, C_{N_p}\}$ of open, disjoint subsets of $\mathbb{R}^2$ such that $\cup_{k=1}^{N_p} C_k = \mathcal{C}$, where N_p denotes the total number of pixels. For $k = 1, \ldots, N_p$ and $t \geq t_0$, assume that n_k photons are detected in the pixel C_k during the time interval $[t_0, t]$. Let K denote the total number of detected photons from the object of interest, i.e., $\sum_{k=1}^{N_p} n_k = K$. Then it can be shown that for $k = 1, \ldots, N_p$, n_k is independently Poisson distributed with

mean $\mu_\theta(k,t) = \int_{t_0}^{t}\int_{C_k}\Lambda_\theta^1(\tau)f_{\theta,\tau}^1(r)drd\tau$, $\theta\in\Theta$ (see Lemma 2.5.3 in Appendix of Chapter 2). Similarly, the number of detected photons at the k^{th} pixel during the time interval $[t_0,t]$ from a background component $\mathcal{G}^2(\Lambda^2,\{f_\tau^2\}_{\tau\geq t_0},\mathcal{C})$ is independently Poisson distributed with mean $\beta(k,t) = \int_{t_0}^{t}\int_{C_k}\Lambda^2(\tau)f_\tau^2(r)drd\tau$, $\theta\in\Theta$. Hence the acquired data in the time interval $[t_0,t]$ from a pixelated detector can be described by a collection $\{\mathcal{I}_{\theta,1},\ldots,\mathcal{I}_{\theta,N_p}\}$ of random variables given by

$$\mathcal{I}_{\theta,k} = S_{\theta,k} + B_k, \quad \theta\in\Theta, \quad k=1,\ldots,N_p.$$

Using the standard expression for the Fisher information matrix of a Poisson distribution ([60]), the Fisher information matrix for $\{\mathcal{I}_{\theta,1},\ldots,\mathcal{I}_{\theta,N_p}\}$ corresponding to the time interval $[t_0,t]$ is given by

$$\mathbf{I}(\theta) = \sum_{k=1}^{N_p}\frac{1}{\mu_\theta(k,t)+\beta(k,t)}\left(\frac{\partial\mu_\theta(k,t)}{\partial\theta}\right)^T\frac{\partial\mu_\theta(k,t)}{\partial\theta}. \tag{2.4}$$

In a pixelated detector the acquired image contains measurement noise, which, for example, arises due to the readout process ([31]). At each pixel this can be modeled as a Gaussian random variable W_k with mean η_k and variance $\sigma_{w,k}^2$, $k=1,\ldots,N_p$. The acquired image is then given by $\mathcal{I}_{\theta,k} = S_{\theta,k}+B_k+W_k$, $\theta\in\Theta$, $k=1,\ldots,N_p$. To derive the Fisher information matrix for the present case, we first note that $\mathcal{I}_{\theta,k}$ is a sum of a Poisson and independent Gaussian random variable, and its probability density function is given by (see [31], also see Lemma 2.5.6 in Appendix of Chapter 2)

$$p_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\sigma_{w,k}}\sum_{l=0}^{\infty}\frac{[\nu_\theta(k,t)]^l e^{-\nu_\theta(k,t)}}{l!}e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad z\in\mathbb{R}, \quad k=1,\ldots,N_p, \tag{2.5}$$

where $\nu_\theta(k,t) := \mu_\theta(k,t)+\beta(k,t)$, $k=1,\ldots,N_p$. If $\{n_1,\ldots,n_{N_p}\}$ denotes the acquired data, then the log likelihood function is given by $\mathcal{L}(\theta\mid n_1,\ldots,n_{N_p}) := \sum_{k=1}^{N_p}\ln[p_{\theta,k}(n_k)]$, $\theta\in\Theta$ and the partial derivative of the log-likelihood function with respect to θ is given by

$$\frac{\partial\mathcal{L}(\theta\mid n_1,\ldots,n_{N_p})}{\partial\theta} = \sum_{k=1}^{N_p}\left[\frac{\partial\mu_\theta(k,t)}{\partial\theta}(\zeta_{\theta,k}(n_k)-1)\right], \quad \theta\in\Theta, \tag{2.6}$$

where

$$\zeta_{\theta,k}(z) := \frac{\sum_{l=1}^{\infty}\frac{l[\nu_\theta(k,t)]^{l-1}e^{-\nu_\theta(k,t)}}{(l-1)!}\frac{1}{\sqrt{2\pi}\sigma_{w,k}}e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{k,w}}\right)}}{p_{\theta,k}(z)}, \quad \theta\in\Theta, \quad k=1,\ldots,N_p, \quad z\in\mathbb{R}.$$

It can be shown that $E[\zeta_{\theta,k}(n_k)] = 1$ for $\theta \in \Theta$ and $k = 1, \ldots, N_p$. Further, it can be verified that the random variables $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ are mutually independent. Using these results and eq. 2.6, the Fisher information matrix for $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = E\left[\left(\frac{\partial \mathcal{L}(\theta \mid n_1, \ldots, n_{N_p})}{\partial \theta}\right)^T \frac{\partial \mathcal{L}(\theta \mid n_1, \ldots, n_{N_p})}{\partial \theta}\right]$$

$$= E\left[\sum_{k=1}^{N_p}\sum_{m=1}^{N_p}\left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(m,t)}{\partial \theta}\left(\zeta_{\theta,k}(n_k)\zeta_{\theta,m}(n_m) - \zeta_{\theta,k}(n_k) - \zeta_{\theta,m}(n_m) + 1\right)\right]$$

$$= \sum_{k=1}^{N_p}\left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta}\left(E[\zeta_{\theta,k}^2(n_k)] - 1\right)$$

$$+ \sum_{k\neq m,k,m=1}^{N_p}\left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(m,t)}{\partial \theta}\left(E[\zeta_{\theta,k}(n_k)]E[\zeta_{\theta,m}(n_m)] - 1\right)$$

$$= \sum_{k=1}^{N_p}\left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \times$$

$$\left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}\right)^2}{p_{\theta,k}(z)} dz - 1\right), \ \theta \in \Theta. \tag{2.7}$$

The above expressions are valid for general intensities $\mu_\theta(k,t)$ ($\nu_\theta(k,t)$), $k = 1, \ldots, N_p$, $t > t_0$. These intensities depend on the photon distribution profiles $f_{\theta,\tau}$, $\tau \geq t_0$, through the above identities. The approach is therefore generally applicable to a large class of photon distributions profiles/image functions. Note that in the above expressions no assumptions are made about the size or shape of the pixels. In [62] an expression was derived that is essentially a special case of eq. 2.4, i.e. for a two-dimensional estimation problem of the location parameters for a Gaussian photon distribution profile in a stationary imaging scenario for square pixels in the absence of Gaussian noise.

2.4.1 Loss of information due to pixelation

In the description of the acquired data given in Section 2.2, we considered an ideal detector that provides the time points and the spatial coordinates of the detected photons. However, in a pixelated detector the acquired data consists of the number of detected photons at each pixel. That is, the image acquired by a pixelated detector is a discretized version of the actual image. We next show that the Fisher information matrix for a pixelated detector is smaller than the Fisher information matrix for a non-pixelated detector. An implication of this result is that the Cramer-Rao lower bound for a pixelated detector is bigger than that for a non-pixelated (ideal) detector.

Lemma 2.4.1 *Let $\Theta \subseteq \mathbb{R}^n$ be the parameter space, $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ denote an image detection process and $\{C_1, \ldots, C_{N_p}\}$ be a pixelated detector where for $j \neq k$, $j, k = 1, \ldots, N_p$, $C_j \bigcap C_k = \emptyset$ and $\bigcup_{k=1}^{N_p} C_k = \mathcal{C}$.*

1. For $\theta \in \Theta$ and $k = 1, \ldots, N_p$, let $\kappa_\theta(k,t) := \frac{1}{m(C_k)} \int_{t_0}^{t} \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau$, where m denotes the Lebesgue measure. Then for $\theta \in \Theta$,

$$\int_{t_0}^{t} \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\kappa_\theta(k,t)]}{\partial \theta} \right)^T \times$$

$$\left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\kappa_\theta(k,t)]}{\partial \theta} \right) \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \geq 0. \tag{2.8}$$

2. Let $\mathbf{I}(\theta)$ be the Fisher information matrix given by Theorem 2.2.1 corresponding to the detector $\mathcal{C}$ and $\mathbf{I}_p(\theta)$ be the Fisher information matrix given by eq. 2.4 for the pixelated detector $\{C_1, \ldots, C_{N_p}\}$ in the absence of additive Poisson noise i.e., $\beta(k,t) = 0$, $k = 1, \ldots, N_p$, $t \geq t_0$. Then for $\theta \in \Theta$, $\mathbf{I}(\theta) \geq \mathbf{I}_p(\theta)$.

Proof: 1. For $\theta \in \Theta$, $\tau \geq t_0$ and $r \in \mathbb{R}^2$, define

$$\Psi_\theta(r,\tau) := \frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\kappa_\theta(k,t)]}{\partial \theta}.$$

Without loss of generality let θ denote an n-dimensional vector parameter and let $v \in \mathbb{R}^n$ denote a column vector. By definition $\Lambda_\theta(\tau) f_{\theta,\tau}(r) \geq 0$, $r \in \mathbb{R}^2$, $\tau \geq t_0$ and $\theta \in \Theta$ and $\Psi_\theta^T(r,\tau) \Psi_\theta(r,\tau) \Lambda_\theta(\tau) f_{\theta,\tau}(r)$ is positive semidefinite, since for every $v \in \mathbb{R}^n$, $r \in \mathbb{R}^2$, $\tau \geq t_0$ and

$\theta \in \Theta$, $\Psi_\theta(r,\tau)v$ is a scalar and $(v\Psi_\theta(r,\tau))^T(v\Psi_\theta(r,\tau)) \geq$. From this the result immediately follows.

2. For $\theta \in \Theta$ and $k = 1, \ldots, N_p$, define $\mu_\theta(k,t) := \int_{t_0}^t \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau$ and let

$$\mathbf{I}_k(\theta) := \int_{t_0}^t \int_{C_k} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} dr d\tau,$$

$$\mathbf{I}_{p,k}(\theta) := \frac{1}{\mu_\theta(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta}.$$

From the above equations, it follows that $\mathbf{I}(\theta) = \mathbf{I}_1(\theta) + \cdots + \mathbf{I}_{N_p}(\theta)$, $\theta \in \Theta$ and $\mathbf{I}_p(\theta) = \mathbf{I}_{p,1}(\theta) + \cdots + \mathbf{I}_{p,N_p}(\theta)$, $\theta \in \Theta$. From result 1 of this Lemma, we have

$$\begin{aligned}
0 &\leq \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\kappa_\theta(k,t)]}{\partial \theta} \right)^T \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\kappa_\theta(k,t)]}{\partial \theta} \right) \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&= \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\frac{\mu_\theta(k,t)}{m(C_k)}]}{\partial \theta} \right)^T \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\frac{\mu_\theta(k,t)}{m(C_k)}]}{\partial \theta} \right) \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&= \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \right)^T \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} - \frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \right) \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&= \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&\quad - \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&\quad - \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \right)^T \frac{\partial \ln[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&\quad + \int_{t_0}^t \int_{C_k} \left(\frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \right)^T \frac{\partial \ln[\mu_\theta(k,t)]}{\partial \theta} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau \\
&= \int_{t_0}^t \int_{C_k} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} dr d\tau - \frac{1}{\mu_\theta(k,t)} \int_{t_0}^t \int_{C_k} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T dr d\tau \frac{\partial \mu_\theta(k,t)}{\partial \theta} \\
&\quad - \frac{1}{\mu_\theta(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \int_{t_0}^t \int_{C_k} \frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} dr d\tau + \frac{\int_{t_0}^t \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau}{\mu_\theta^2(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta},
\end{aligned}$$

where $\theta \in \Theta$. Dividing the above expression by $\mu_\theta(k,t)$ $\left(= \int_{t_0}^t \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau\right)$ we get

$$\begin{aligned}
0 &\leq \frac{1}{\mu_\theta(k,t)} \int_{t_0}^t \int_{C_k} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} dr d\tau - \frac{1}{\mu_\theta^2(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \\
&\quad - \frac{1}{\mu_\theta^2(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} + \frac{1}{\mu_\theta^2(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \\
&= \mathbf{I}_k(\theta) - \mathbf{I}_{p,k}(\theta), \quad \theta \in \Theta. \qquad (2.9)
\end{aligned}$$

Hence $\mathbf{I}_k(\theta) \geq \mathbf{I}_{p,k}(\theta)$, $\theta \in \Theta$. Since this is true for every $k = 1, \ldots, N_p$, it follows that $\mathbf{I}_1(\theta) + \mathbf{I}_2(\theta) + \cdots + \mathbf{I}_{N_p}(\theta) \geq \mathbf{I}_{p,1}(\theta) + \mathbf{I}_{p,2}(\theta) + \cdots + \mathbf{I}_{p,N_p}(\theta)$, $\theta \in \Theta$ and from this the result follows. •

The above result considers the scenario when the data acquired by a pixelated detector is not corrupted by additive noise sources. However, if these noise sources are taken into account, then a further deterioration would be expected in the limit to the accuracy of the unknown parameter θ. The verification of this will be carried out in the next chapter through a concrete example.

2.4.2 Effect of image addition

The Fisher information matrix derived in eqs. 2.4 and 2.7 are obtained for the problem of estimating θ from a single image acquired by a pixelated detector. However, in many applications, the analysis is carried out on a composite image, which is a sum of N distinct images. In the following theorem, an expression is derived for the probability density function of the detected photon count at each pixel of the composite image.

Theorem 2.4.1 *For $\theta \in \Theta$ and $k = 1, \ldots, N$, let $\{\mathcal{I}^k_{\theta,1}, \ldots, \mathcal{I}^k_{\theta,N_p}\}$ denote a sequence of independent random variables that model the k^{th} image acquired by a pixelated finite detector, where N denotes the total number of acquired images. Let $\mathcal{I}^i_{\theta,k}$ be given by*

$$\mathcal{I}^i_{\theta,k} := S^i_{\theta,k} + B^i_k + W^i_k, \quad k = 1, \ldots, N_p, \quad \theta \in \Theta, \quad i = 1, \ldots, N,$$

where $S^i_{\theta,k}$ (B^i_k) is a Poisson random variable with mean $\mu^i_\theta(k)$ ($\beta^i(k)$) and W^i_k is a Gaussian random variable with mean η^i_k and variance $(\sigma^i_{w,k})^2$, for $k = 1, \ldots, N_p$ and $i = 1, \ldots, N$. Assume that for $\theta \in \Theta$ and $i = 1, \ldots, N$, $\{S^i_{\theta,1}, \ldots, S^i_{\theta,N_p}\}$, $\{B^i_1, \ldots, B^i_{N_p}\}$ and $\{W^i_1, \ldots, W^i_{N_p}\}$ are mutually independent and independent of each other. For $k = 1, \ldots, N_p$ and $\theta \in \Theta$, define $Y_{\theta,k} := \sum_{i=1}^N \mathcal{I}^i_{\theta,k}$.

1. For $k = 1, \ldots, N_p$ and $\theta \in \Theta$, the probability density function of $Y_{\theta,k}$ is given by

$$\tilde{p}_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\sigma_{w,k}} \sum_{l=0}^{\infty} \frac{[\mu_\theta(k) + \beta(k)]^l e^{-[\mu_\theta(k)+\beta(k)]}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad z \in \mathbb{R}^2,$$

where

$$\mu_\theta(k) := \sum_{i=1}^N \mu^i_\theta(k), \qquad \beta(k) := \sum_{i=1}^N \beta^i(k), \qquad \theta \in \Theta, \quad k = 1, \ldots, N_p, \tag{2.10}$$

$$\eta_k := \sum_{i=1}^{N} \eta_k^i, \qquad \sigma_{w,k}^2 := \sum_{i=1}^{N} [\sigma_{w,k}^i]^2, k = 1, \ldots, N_p, \quad \theta \in \Theta. \tag{2.11}$$

2. For $\theta \in \Theta$ and $k = 1, \ldots, N_p$, assume that $\mu_\theta^1(k) = \cdots = \mu_\theta^N(k)$, $\beta^1(k) = \cdots = \beta^N(k)$, $\eta_k^1 = \cdots = \eta_k^N$, and $\sigma_{w,k}^1 = \cdots = \sigma_{w,k}^N$. Then for $k = 1, \ldots, N_p$ and $\theta \in \Theta$, the probability density function of $\mathcal{I}_{\theta,k}$ is given by

$$\tilde{p}_{\theta,k}(z) := \frac{1}{\sqrt{2\pi N}\sigma_{w,k}^1} \sum_{l=0}^{\infty} \frac{[N(\mu_\theta^1(k) + \beta^1(k))]^l e^{-N[\mu_\theta^1(k)+\beta^1(k)]}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-N\eta_k^1}{N\sigma_{w,k}^1}\right)^2}, \quad z \in \mathbb{R}^2.$$

Proof: For $\theta \in \Theta$, $k = 1, \ldots, N_p$ and $i = 1, \ldots, N$ define $X_{\theta,k}^i := S_{\theta,k}^i + B_k^i$. Then $\mathcal{I}_{\theta,k}^i = X_{\theta,k}^i + W_k^i$, where $X_{\theta,k}^i$ is a Poisson random variable with mean $\mu_\theta^i(k) + \beta^i(k)$, $i = 1, \ldots, N$, $k = 1, \ldots, N_p$ and $\theta \in \Theta$. Consider the term

$$Y_{\theta,k} = \sum_{i=1}^{N} \mathcal{I}_{\theta,k} = (X_{\theta,k}^1 + W_k^1) + (X_{\theta,k}^2 + W_k^2) + \cdots + (X_{\theta,k}^N + W_k^N)$$

$$= (X_{\theta,k}^1 + X_{\theta,k}^2 + \cdots + X_{\theta,k}^N) + (W_k^1 + W_k^2 + \cdots + W_k^N) = \underbrace{\sum_{i=1}^{N} X_{\theta,k}^i}_{T_1} + \underbrace{\sum_{i=1}^{N} W_k^i}_{T_2}.$$

The sum term T_1 is a sum of N independent Poisson random variables, which is a Poisson random variable with mean $\mu_\theta(k) + \beta(k)$, where $\mu_\theta(k) = \mu_\theta^1(k) + ... + \mu_\theta^N(k)$ and $\beta(k) = \beta^1(k) + ... + \beta^N(k)$. Similarly, the sum term T_2 is a Gaussian random variable with mean $\eta_k = \eta_k^1 + ...\eta_k^N$ and variance $[\sigma_{w,k}]^2 = [\sigma_{w,k}^1]^2 + ... + [\sigma_{w,k}^N]^2$. Further, the sum terms T_1 and T_2 are statistically independent. Hence $\mathcal{Y}_{\theta,k}$ is the sum of a Poisson and independent Gaussian random variable. Since by definition $\mathcal{I}_{\theta,k}^i$ is also a sum of a Poisson and an independent Gaussian random variable whose probability density function is given by eq. 2.5, substituting $\mu_\theta(k)$, $\beta(k)$, η_k and $[\sigma_{w,k}]^2$ in eq. 2.5, the result immediately follows.

2. The result immediately follows from result 2. •

From the above theorem we see that for a composite image the expression for the probability density function of the photon count in the k^{th} pixel is analogous to that of a single image, which is given in eq. 5.9. From this it can be deduced that the expression of Fisher

information matrix for the problem of estimating θ from a composite image is also analogous to that for a single image (eqs. 2.4 and 2.7).

2.5 Appendix to Chapter 2

Lemma 2.5.1 *Proof of Theorem 2.2.1: the expression for the Fisher information matrix of the image detection process $\mathcal{G}$ corresponding to a time interval $[t_0, t]$.*

Proof: By definition of the temporal process $\{Z(\tau), \tau \geq t_0\}$ of $\mathcal{G}$, the random variable $Z(t)$ is Poisson distributed with mean $\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau$ and denotes the total number of detected photons in the time interval $[t_0, t]$. Let $\{w_1, \ldots, w_K\}$ denote a realization of the image detection process in the time interval $[t_0, t]$, where $w_k := (r_k, \tau_k)$, $r_k = (x_k, y_k) \in \mathcal{C}$ denotes the spatial coordinates of the arrival location of the k^{th} detected photon, $k = 1, \ldots, K$, $t_0 \leq \tau_1 \leq \cdots \leq \tau_K \leq t$ denotes the time points at which the photons are detected and K denotes the total number of detected photons that is a realization of the random variable $Z(t)$.

Let $T_k : \tilde{\Omega} \to [t_0, \infty)$, $k = 1, 2, \ldots$, denote a random variable that describes the time point of the k^{th} detected photon, where $\tilde{\Omega}$ denotes the sample space and let $\mathcal{T}_K$ denote the event $\{T_1 = \tau_1, \ldots, T_K = \tau_K\}$. Using the fact that the spatial and temporal components of the image detection process are independent, that the spatial coordinates $\{r_1, \ldots, r_K\}$ are mutually independent and that the probability density function of the spatial coordinates of the k^{th} detected photon only depends on the time point τ_k, $k = 1, \ldots, K$, the sample function density is given by

$$
\begin{aligned}
& p[U_1 = r_1, \ldots, U_K = r_K, T_1 = \tau_1, \ldots, T_K = \tau_K, Z(t) = K] \\
&= p[U_1 = r_1, \ldots, U_K = r_K \mid \mathcal{T}_K, Z(t) = K] p[\mathcal{T}_K \mid Z(t) = K] p[Z(t) = K] \\
&= p[U_1 = r_1 \mid \mathcal{T}_K] p[U_2 = r_2 \mid \mathcal{T}_K] \ldots p[U_K = r_K \mid \mathcal{T}_K] p[\mathcal{T}_K \mid Z(t) = K] p[Z(t) = K] \\
&= p[U_1 = r_1 \mid T_1 = \tau_1] p[U_2 = r_2 \mid T_2 = \tau_2] \ldots p[U_K = r_K \mid T_K = \tau_K] \times \\
&\qquad p[\mathcal{T}_K \mid Z(t) = K] p[Z(t) = K] \\
&= \left(\prod_{k=1}^{K} f_{\theta, \tau_k}(r_k) \right) p[\mathcal{T}_K \mid Z(t) = K] p[Z(t) = K]
\end{aligned}
$$

$$= \left(\prod_{k=1}^{K} f_{\theta,\tau_k}(r_k)\right) \frac{K! \prod_{k=1}^{K} \Lambda_\theta(\tau_k)}{\left(\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau\right)^K} \left(\frac{1}{K!} e^{-\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau} \left(\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau\right)^K\right)$$

$$= \left(\prod_{k=1}^{K} f_{\theta,\tau_k}(r_k)\right) \left(\prod_{k=1}^{K} \Lambda_\theta(\tau_k)\right) \exp\left(-\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau\right),$$

where $f_{\theta,\tau_k}(r_k)$ denotes the density function of the spatial coordinates r_k of the k^{th} detected photon, $p[\mathcal{T}_K \mid Z(t) = K] = p[T_1 = \tau_1, \ldots, T_K = \tau_K \mid Z(t) = K]$ denotes the conditional probability density function of the time points of the detected photons and $p[Z(t) = K]$ denotes the probability of detecting K photons in the time interval $[t_0, t]$. In the above equation the expression for $p[\mathcal{T}_K \mid Z(t) = K]$ pertains to the ordered sequence $t_0 \leq \tau_1 \leq \cdots \leq \tau_K \leq t$ of the occurrence times (see e.g., [60]).

The log likelihood function is given by

$$\mathcal{L}(\theta | w_1, \ldots, w_K) := \ln(p[U_1 = r_1, \ldots, U_K = r_K, T_1 = \tau_1, \ldots, T_K = \tau_K, Z(t) = K])$$

$$= \sum_{k=1}^{K} \ln[f_{\theta,\tau_k}(r_k)] + \sum_{k=1}^{K} \ln[\Lambda_\theta(\tau_k)] - \int_{t_0}^{t} \Lambda_\theta(\tau) d\tau, \quad \theta \in \Theta.$$

By definition of Λ_θ, the integral $\int_{t_0}^{t} \Lambda_\theta(\tau) d\tau$ exists and is finite for $\theta \in \Theta$, $t \geq t_0$ and $\partial \Lambda_\theta(\tau)/\partial\theta$ is piecewise continuous with respect to θ for every $\theta \in \Theta$ and $\tau \geq t_0$. Hence $(\partial/\partial\theta) \int_{t_0}^{t} \Lambda_\theta(\tau) d\tau = \int_{t_0}^{t} (\partial \Lambda_\theta(\tau)/\partial\theta) d\tau$, $\theta \in \Theta$. For $\theta \in \Theta$ and $\tau \geq t_0$, define

$$\mathcal{J}_{\theta,\tau}(r) := \frac{1}{f_{\theta,\tau}(r)} \frac{\partial f_{\theta,\tau}(r)}{\partial\theta}, \quad r \in \mathcal{C}, \qquad \mathcal{M}_\theta(\sigma) := \frac{1}{\Lambda_\theta(\sigma)} \frac{\partial \Lambda_\theta(\sigma)}{\partial\theta}, \quad \sigma \geq t_0.$$

Then the derivative of the log likelihood function with respect to $\theta \in \Theta$ is given by

$$\frac{\partial \mathcal{L}(\theta | w_1, \ldots, w_K)}{\partial\theta} = \sum_{k=1}^{K} \frac{1}{f_{\theta,\tau_k}(r_k)} \frac{\partial f_{\theta,\tau_k}(r_k)}{\partial\theta} + \sum_{k=1}^{K} \frac{1}{\Lambda_\theta(\tau_k)} \frac{\partial \Lambda_\theta(\tau_k)}{\partial\theta} - \frac{\partial}{\partial\theta} \int_{t_0}^{t} \Lambda_\theta(\tau) d\tau$$

$$= \sum_{k=1}^{K} \mathcal{J}_{\theta,\tau_k}(r_k) + \sum_{k=1}^{K} \mathcal{M}_\theta(\tau_k) - \int_{t_0}^{t} \frac{\partial \Lambda_\theta(\tau)}{\partial\theta} d\tau.$$

The Fisher information matrix is given by ([55])

$$\mathbf{I}(\theta) = E\left[\left(\frac{\partial \mathcal{L}(\theta | w_1, \ldots, w_K)}{\partial\theta}\right)^T \left(\frac{\partial \mathcal{L}(\theta | w_1, \ldots, w_K)}{\partial\theta}\right)\right]$$

$$
\begin{aligned}
&= E\left[\left(\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}^{T}(r_k)+\sum_{k=1}^{K}\mathcal{M}_{\theta}^{T}(\tau_k)-\int_{t_0}^{t}\left(\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}\right)^{T}d\tau\right)\times\right.\\
&\qquad\left.\left(\sum_{l=1}^{K}\mathcal{J}_{\theta,\tau_l}(r_l)+\sum_{l=1}^{K}\mathcal{M}_{\theta}(\tau_l)-\int_{t_0}^{t}\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}d\tau\right)\right]\\
&=\underbrace{E\left[\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}^{T}(r_k)\sum_{l=1}^{K}\mathcal{J}_{\theta,\tau_l}(r_l)\right]}_{T_1}+\underbrace{E\left[\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}^{T}(r_k)\sum_{l=1}^{K}\mathcal{M}_{\theta}(\tau_l)\right]}_{T_2^T}\\
&\quad-\underbrace{E\left[\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}^{T}(r_k)\right]}_{T_3^T}\int_{t_0}^{t}\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}d\tau+\underbrace{E\left[\sum_{k=1}^{K}\mathcal{M}_{\theta}^{T}(\tau_k)\sum_{l=1}^{K}\mathcal{J}_{\theta,\tau_l}(r_l)\right]}_{T_2}\\
&\quad+\underbrace{E\left[\sum_{k=1}^{K}\mathcal{M}_{\theta}^{T}(\tau_k)\sum_{l=1}^{K}\mathcal{M}_{\theta}(\tau_l)\right]}_{T_4}-\underbrace{E\left[\sum_{k=1}^{K}\mathcal{M}_{\theta}^{T}(\tau_k)\right]}_{T_5^T}\int_{t_0}^{t}\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}d\tau\\
&\quad-\int_{t_0}^{t}\left(\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}\right)^{T}d\tau\left(\underbrace{E\left[\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}(r_k)\right]}_{T_3}\right.\\
&\quad\left.+\underbrace{E\left[\sum_{k=1}^{K}\mathcal{M}_{\theta}(\tau_k)\right]}_{T_5}-\int_{t_0}^{t}\frac{\partial\Lambda_\theta(\tau)}{\partial\theta}d\tau\right),\quad\theta\in\Theta. \qquad (2.12)
\end{aligned}
$$

Let $\mathcal{N}$ denote any function that depends on $\{z_1,\ldots,z_K\}$. Then using the fact that the spatial and temporal processes of $\mathcal{G}$ are independent, and that the spatial coordinates of the detected photons are mutually independent, we have

$$
\begin{aligned}
E[\mathcal{N}(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K)] &= \sum_{K=1}^{\infty}\int_{t_0}^{t}\cdots\int_{t_0}^{t}\int_{\mathcal{C}}\cdots\int_{\mathcal{C}}\mathcal{N}(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K)\times\\
&\qquad p(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K,Z(t)=K)dr_1\ldots dr_K d\tau_1\ldots d\tau_K\\
&= \sum_{K=1}^{\infty}\left(\int_{t_0}^{t}\cdots\int_{t_0}^{t}\left[\int_{\mathcal{C}}\cdots\int_{\mathcal{C}}\mathcal{N}(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K)\times\right.\right.\\
&\qquad\left.\left.p(r_1,\ldots,r_K\mid\tau_1,\ldots,\tau_K,Z(t)=K)dr_1\ldots dr_K\right]\times\right.
\end{aligned}
$$

$$p(\tau_1,\ldots,\tau_K \mid Z(t)=K)d\tau_1\ldots d\tau_K\Bigg)P(Z(t)=K)$$

$$=\sum_{K=1}^{\infty}\Bigg(\int_{t_0}^{t}\ldots\int_{t_0}^{t}\Bigg[\int_{\mathcal{C}}\ldots\int_{\mathcal{C}}\mathcal{N}(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K)f_{\theta,\tau 1}(r_1)\ldots f_{\theta,\tau_K}(r_K)dr_1\ldots dr_K\Bigg]$$

$$\times p(\tau_1,\ldots,\tau_K \mid Z(t)=K)d\tau_1\ldots d\tau_K\Bigg)P(Z(t)=K). \tag{2.13}$$

Now, consider a special case where $\mathcal{N}$ is given by $\mathcal{N}(r_1,\ldots,r_K,\tau_1,\ldots,\tau_K)=\sum_{k=1}^{K}\mathcal{U}(r_k,\tau_k)$ and $\mathcal{U}$ is a real-valued vector function defined on $\mathcal{C}\times[t_0,\infty)$. Note that for fixed K, a random reordering of the acquired data $(r_1,\tau_1),\ldots,(r_K,\tau_K)$ does not change the value of the summation $\sum_{k=1}^{K}\mathcal{U}(r_k,\tau_k)$ and hence the value of $\mathcal{N}$ would remain the same for the given data. However, with this reordering the occurrence times can treated as independent and identically distributed random variables with probability density p_θ that is given by (see [60, pg 65], [63, pg 272])

$$p_\theta(\tau)=\frac{\Lambda_\theta(\tau)}{\int_{t_0}^{t}\Lambda_\theta(\tau)d\tau},\quad \tau\in[t_0,t],\quad \theta\in\Theta.$$

This implies that in evaluating the expectation value of the above special case of $\mathcal{N}$, the probability density $p(\tau_1,\ldots,\tau_K \mid Z(t)=K)$ can be replaced by $p_\theta(\tau_1)\ldots p_\theta(\tau_K)$ in eq. 2.13 and we have

$$E\Bigg[\sum_{k=1}^{K}\mathcal{U}(r_k,\tau_k)\Bigg]=\sum_{K=1}^{\infty}\Bigg(\int_{t_0}^{t}\ldots\int_{t_0}^{t}\Bigg[\int_{\mathcal{C}}\ldots\int_{\mathcal{C}}\Bigg(\mathcal{U}(r_1,\tau_1)+\cdots+\mathcal{U}(r_K,\tau_K)\Bigg)\times$$

$$f_{\theta,\tau_1}(r_1)\ldots f_{\theta,\tau_K}(r_K)dr_1\ldots dr_K\Bigg]p_\theta(\tau_1)\ldots p_\theta(\tau_K)d\tau_1\ldots d\tau_K\Bigg)P(Z(t)=K)$$

$$=\sum_{K=1}^{\infty}\Bigg(\int_{t_0}^{t}\ldots\int_{t_0}^{t}\Bigg[\int_{\mathcal{C}}\mathcal{U}(r,\tau_1)f_{\theta,\tau_1}(r)dr\int_{\mathcal{C}}f_{\theta,\tau_2}(r)dr\ldots\int_{\mathcal{C}}f_{\theta,\tau_K}(r)dr\quad+\quad\ldots$$

$$+\int_{\mathcal{C}}f_{\theta,\tau_1}(r)dr\int_{\mathcal{C}}f_{\theta,\tau_2}(r)dr\ldots\int_{\mathcal{C}}\mathcal{U}(r,\tau_K)f_{\theta,\tau_K}(r)dr\Bigg]\times$$

$$p_\theta(\tau_1)\ldots p_\theta(\tau_K)d\tau_1\ldots d\tau_K\Bigg)P(Z(t)=K)$$

$$=\sum_{K=1}^{\infty}\Bigg(\int_{t_0}^{t}\ldots\int_{t_0}^{t}\Bigg[\sum_{k=1}^{K}\int_{\mathcal{C}}\mathcal{U}(r,\tau_k)f_{\theta,\tau_k}(r)dr\Bigg]p_\theta(\tau_1)\ldots p_\theta(\tau_K)d\tau_1\ldots d\tau_K\Bigg)P(Z(t)=K)$$

$$= \sum_{K=1}^{\infty} \left(\int_{t_0}^{t} \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau_1) f_{\theta,\tau_1}(r) dr \right] p_\theta(\tau_1) d\tau_1 \int_{t_0}^{t} p_\theta(\tau_2) d\tau_2 \ldots \int_{t_0}^{t} p_\theta(\tau_K) d\tau_K \quad + \quad \ldots \right.$$

$$\left. + \quad \int_{t_0}^{t} p_\theta(\tau_1) d\tau_1 \int_{t_0}^{t} p_\theta(\tau_2) d\tau_2 \ldots \int_{t_0}^{t} \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau_K) f_{\theta,\tau_K}(r) dr \right] p_\theta(\tau_K) d\tau_1 \right) P(Z(t) = K)$$

$$= \sum_{K=1}^{\infty} \left(K \int_{t_0}^{t} \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau) f_{\theta,\tau}(r) dr \right] p_\theta(\tau) d\tau \right) P(Z(t) = K)$$

$$= \sum_{K=1}^{\infty} KP(Z(t) = K) \int_{t_0}^{t} \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau) f_{\theta,\tau}(r) dr \right] p_\theta(\tau) d\tau$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) d\tau \int_{t_0}^{t} \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau) f_{\theta,\tau}(r) dr \right] p_\theta(\tau) d\tau$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau) f_{\theta,\tau}(r) dr \right] d\tau, \tag{2.14}$$

where we have substituted for p_θ in the last step and the term $\sum_{K=1}^{\infty} KP(Z(t) = K) = \sum_{K=0}^{\infty} KP(Z(t) = K)$ is the mean of the Poisson random variable $Z(t)$. Note that for $K = 0$, the term $KP(Z(t) = K)$ is zero and hence does not change the infinite sum. If $\mathcal{U}$ is only a function of τ, then eq. 2.14 becomes $\int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{U}(\tau) f_{\theta,\tau}(r) dr \right] d\tau = \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{U}(\tau) \left[\int_{\mathcal{C}} f_{\theta,\tau}(r) dr \right] d\tau = \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{U}(\tau) d\tau$.

Consider another special case of $\mathcal{N}$ given by

$$\mathcal{N}(r_1, \ldots, r_K, \tau_1, \ldots, \tau_K) = \sum_{k=1}^{K} \mathcal{U}(r_k, \tau_k) \sum_{l=1}^{K} \mathcal{V}^T(r_l, \tau_l),$$

where $\mathcal{V}$ is also a real-valued vector function defined on $\mathcal{C} \times [t_0, \infty)$. Here again, a random reordering of the acquired data for a fixed value of K leaves the value of $\mathcal{N}$ unchanged. Hence replacing the term $p(\tau_1, \ldots, \tau_K \mid Z(t) = K)$ by $p_\theta(\tau_1) \ldots p_\theta(\tau_K)$ in eq. 2.13 and using eq. 2.14 we have

$$E\left[\sum_{k=1}^{K} \mathcal{U}(r_k, \tau_k) \sum_{l=1}^{K} \mathcal{V}^T(r_l, \tau_l) \right]$$

$$= E\left[\sum_{k=1}^{K} \mathcal{U}(r_k, \tau_k) \mathcal{V}^T(r_k, \tau_k) \right] + E\left[\sum_{k \neq l, k, l=1}^{K} \mathcal{U}(r_k, \tau_k) \mathcal{V}^T(r_l, \tau_l) \right]$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{U}(r,\tau) \mathcal{V}^T(r,\tau) f_{\theta,\tau}(r) dr \right] p_\theta(\tau) d\tau$$

$$+\sum_{K=1}^{\infty}\left(\int_{t_0}^{t}\dots\int_{t_0}^{t}\left[\int_{\mathcal{C}}\dots\int_{\mathcal{C}}\left(\sum_{k\neq l,k,l=1}^{K}\mathcal{U}(r_k,\tau_k)\mathcal{V}^T(r_l,\tau_l)\right)f_{\theta,\tau_1}(r_1)\dots f_{\theta,\tau_K}(r_K)dr_1\dots dr_K\right]\times\right.$$

$$\left.p_\theta(\tau_1)\dots p_\theta(\tau_K)d\tau_1\dots d\tau_K\right)P(Z(t)=K)$$

$$=\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau$$

$$+\sum_{K=1}^{\infty}\left(\int_{t_0}^{t}\dots\int_{t_0}^{t}\left[\sum_{k\neq l,k,l=1}^{K}\int_{\mathcal{C}}\mathcal{U}(r,\tau_k)f_{\theta,\tau_k}(r)dr\int_{\mathcal{C}}\mathcal{V}^T(r,\tau_l)f_{\theta,\tau_l}(r)dr\right]\times\right.$$

$$\left.p_\theta(\tau_1)\dots p_\theta(\tau_K)d\tau_1\dots d\tau_K\right)P(Z(t)=K)$$

$$=\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau+\sum_{K=1}^{\infty}(K^2-K)P(Z(t)=K)$$

$$\times\left(\int_{t_0}^{t}\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau\int_{t_0}^{t}\left[\int_{\mathcal{C}}\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau\right)$$

$$=\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau$$

$$+\left(\int_{t_0}^{t}\Lambda_\theta(\tau)d\tau\right)^2\int_{t_0}^{t}\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau\int_{t_0}^{t}\left[\int_{\mathcal{C}}\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]p_\theta(\tau)d\tau$$

$$=\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]d\tau$$

$$+\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{U}(r,\tau)f_{\theta,\tau}(r)dr\right]d\tau\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{V}^T(r,\tau)f_{\theta,\tau}(r)dr\right]d\tau,\quad\theta\in\Theta,\tag{2.15}$$

where we have substituted for p_θ in the last step and the term $\sum_{K=1}^{\infty}(K^2-K)P(Z(t)=K)=E[K^2]-E[K]=\mathrm{Var}(K)+(E[K])^2-E[K]=\left(\int_{t_0}^{t}\Lambda_\theta(\tau)d\tau\right)^2$ with $E[\cdot]$ denoting the expectation operator.

Using Lemma 2.5.2, we $\int_{\mathcal{C}}\mathcal{J}_{\theta,\tau}(r)f_{\theta,\tau}(r)dr=\int_{\mathcal{C}}(\partial f_{\theta,\tau}(r)/\partial\theta)dr=0$ for $\theta\in\Theta$ and $\tau\geq t_0$. Using this and eq. 2.14 the term T1 is given by

$$E\left[\sum_{k=1}^{K}\mathcal{J}_{\theta,\tau_k}^T(r_k)\sum_{l=1}^{K}\mathcal{J}_{\theta,\tau_l}(r_l)\right]=\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{J}_{\theta,\tau}^T(r)\mathcal{J}_{\theta,\tau}(r)f_{\theta,\tau}(r)dr\right]d\tau$$

$$+\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{J}_{\theta,\tau}^T(r)f_{\theta,\tau}(r)dr\right]d\tau\int_{t_0}^{t}\Lambda_\theta(\tau)\left[\int_{\mathcal{C}}\mathcal{J}_{\theta,\tau}(r)f_{\theta,\tau}(r)dr\right]d\tau$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \frac{1}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta} \right)^T \frac{1}{f_{\theta,\tau}(r)} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta} f_{\theta,\tau}(r) dr \right] d\tau$$

$$= \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta} \right)^T \frac{\partial f_{\theta,\tau}(r)}{\partial \theta} d\tau dr, \quad \theta \in \Theta. \tag{2.16}$$

Similarly, the term T_2 is given by

$$E\left[\sum_{k=1}^{K} \mathcal{M}_\theta^T(\tau_k) \sum_{l=1}^{K} \mathcal{J}_{\theta,\tau_l}(r_l) \right] = \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{M}_\theta^T(\tau) \mathcal{J}_{\theta,\tau}(r) f_{\theta,\tau}(r) dr \right] d\tau$$

$$+ \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{M}_\theta^T(\tau) f_{\theta,\tau}(r) dr \right] d\tau \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{J}_{\theta,\tau}(r) f_{\theta,\tau}(r) dr \right] d\tau$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta^T(\tau) \left[\int_{\mathcal{C}} \mathcal{J}_{\theta,\tau}(r) f_{\theta,\tau}(r) dr \right] d\tau$$

$$+ \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta^T(\tau) \left[\int_{\mathcal{C}} f_{\theta,\tau}(r) dr \right] d\tau \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{J}_{\theta,\tau}(r) f_{\theta,\tau}(r) dr \right] d\tau$$

$$= 0, \quad \theta \in \Theta. \tag{2.17}$$

By definition of $\mathcal{M}_\theta$, $\int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta(\tau) d\tau = \int_{t_0}^{t} (\partial \Lambda_\theta(\tau)/\partial \theta) d\tau$, $\theta \in \Theta$. Using this, the term T4 is given by

$$E\left[\sum_{k=1}^{K} \mathcal{M}_\theta^T(\tau_k) \sum_{l=1}^{K} \mathcal{M}_\theta(\tau_l) \right]$$

$$= \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta^T(\tau) \mathcal{M}_\theta(\tau) d\tau + \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta^T(\tau) d\tau \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta(\tau) d\tau$$

$$= \int_{t_0}^{t} \frac{\Lambda_\theta(\tau)}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta} \right)^T \frac{1}{\Lambda_\theta(\tau)} \frac{\partial \Lambda_\theta(\tau)}{\partial \theta} d\tau + \int_{t_0}^{t} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta} \right)^T d\tau \int_{t_0}^{t} \frac{\partial \Lambda_\theta(\tau)}{\partial \theta} d\tau$$

$$= \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta} \right)^T \frac{\partial \Lambda_\theta(\tau)}{\partial \theta} d\tau + \int_{t_0}^{t} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta} \right)^T d\tau \int_{t_0}^{t} \frac{\partial \Lambda_\theta(\tau)}{\partial \theta} d\tau, \quad \theta \in \Theta. \tag{2.18}$$

Using eq. 2.14 the term T_3 is given by

$$E\left[\sum_{k=1}^{K} \mathcal{J}_{\theta,\tau_k}(r_k) \right] = \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{J}_{\theta,\tau}(r) f_{\theta,\tau}(r) dr \right] d\tau = 0, \quad \theta \in \Theta. \tag{2.19}$$

Finally, using eq. 2.14 the term T_5 is given by

$$E\left[\sum_{k=1}^{K} \mathcal{M}_\theta(\tau_k) \right] = \int_{t_0}^{t} \Lambda_\theta(\tau) \left[\int_{\mathcal{C}} \mathcal{M}_\theta(\tau) f_{\theta,\tau}(r) dr \right] d\tau = \int_{t_0}^{t} \Lambda_\theta(\tau) \mathcal{M}_\theta(\tau) d\tau$$

$$= \int_{t_0}^{t} \frac{\partial \Lambda_\theta(\tau)}{\partial \theta} d\tau, \ \theta \in \Theta. \tag{2.20}$$

Substituting eqs. 2.16 - 2.20 in eq. 2.12 and using the fact that $\int_{\mathcal{C}} f_{\theta,\tau}(r)dr = 1$, $\theta \in \Theta$, $\tau \geq t_0$, we get

$$\mathbf{I}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right)^T \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right) d\tau + \int_{\mathcal{C}} \int_{t_0}^{t} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right)^T \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right) dr d\tau$$

$$= \int_{t_0}^{t} \int_{\mathcal{C}} \frac{f_{\theta,\tau}(r)}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right)^T \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right) dr d\tau + \int_{\mathcal{C}} \int_{t_0}^{t} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right)^T \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right) dr d\tau$$

$$+ \underbrace{\int_{t_0}^{t} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right)^T d\sigma \int_{\mathcal{C}} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right) dr}_{S_1} + \underbrace{\int_{\mathcal{C}} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial \theta}\right)^T dr \int_{t_0}^{t} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta}\right) d\tau}_{S_2}$$

$$= \int_{t_0}^{t} \int_{\mathcal{C}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta}\right)^T \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta}\right) dr d\tau, \quad \theta \in \Theta.$$

Since by Lemma 2.5.2 $\int_{\mathcal{C}} (\partial f_{\theta,\tau}(r)/\partial \theta) dr = 0$, $\theta \in \Theta$, $\tau \geq t_0$, adding the terms S_1 and S_2 does not change the expression for the Fisher information matrix and aids in writing it in a compact form. •

Lemma 2.5.2 *Let $\mathcal{C}$ be a detector and Θ denote a parameter set that is an open subset of $\mathbb{R}^n$. Let $\mathcal{F}_\Theta$ be the set of probability densities f_θ defined on $\mathcal{C}$ that satisfy conditions* **C3-C5** *of Definition 2.2.1. Then*

$$\int_{\mathcal{C}} \frac{\partial f_\theta(r)}{\partial \theta_i} dr = 0, \quad i = 1, \ldots, n, \quad \theta \in \Theta.$$

Proof: See [55, pg 182-183] for proof. •

Lemma 2.5.3 *Let $\Theta \subseteq \mathbb{R}^n$ be open. For $\theta \in \Theta$, let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C}\}$ be an image detection process and $\{C_1, \ldots, C_{N_p}\}$ denote a pixelated detector, where N_p denotes the total number of pixels. Then the number of photons detected at the k^{th} pixel, $k = 1, \ldots, N_p$ during the time interval $[t_0, t]$ is independently Poisson distributed with $\mu_\theta(k,t)$, which is given by*

$$\mu_\theta(k,t) := \int_{t_0}^{t} \int_{C_k} \Lambda_\theta(\tau) f_{\theta,\tau}(r) dr d\tau, \quad k = 1, \ldots, N_p, \quad \theta \in \Theta, \quad t \geq t_0.$$

Proof: Using Proposition 2.3.1, it can be shown that the time points and the spatial coordinates of the photons detected at the k^{th} pixel C_k, $k = 1, \ldots, N_p$, can be described by an image detection process $\mathcal{G}^k(\Lambda_\theta^k, \{f_{\theta,\tau}^k\}_{\tau \geq t_0}, C_k)$, which is the reduced version of $\mathcal{G}$ corresponding to the detector C_k, where

$$\Lambda_\theta^k(\tau) := \Lambda_\theta(\tau) \int_{C_k} f_{\theta,\tau}(r)dr, \quad \theta \in \Theta, \quad \tau \geq t_0, \quad k = 1, \ldots, N_p, \tag{2.21}$$

$$f_{\theta,\tau}^k(r) = \begin{cases} \frac{1}{\int_{C_k} f_{\theta,\tau}(r)} f_{\theta,\tau}(r), & r \in C_k, \quad \theta \in \Theta, \quad \tau \geq t_0, \quad k = 1, \ldots, N_p, \\ 0 & \text{otherwise.} \end{cases}$$

This implies that the time points of the detected photons at the kth pixel C_k is described by a Poisson process with intensity function Λ_θ^k. Moreover, by definition of the pixelated detector the pixels are disjoint and hence the image detection processes $\mathcal{G}^1$, $\mathcal{G}^2$,..., $\mathcal{G}^{N_p}$ are mutually independent. Finally by using the well known result that for a Poisson process with intensity Λ, the number of detected photons detected over the time interval $[t_1, t_2]$ is Poisson distributed with mean $\int_{t_1}^{t_2} \Lambda(\tau)d\tau$, the result immediately follows. •

2.5.1 Sum of a discrete and a continous random variable

In this section we derive an analytical expression for the probability density function of a random variable Z which is the sum of a discrete and a continuous random variable. First, we state without proof two well known results that will be used in the derivation of the density function.

Lemma 2.5.4 *Let $S \subseteq \mathbb{R}$ and $\{f_n\}$ be a sequence of functions defined on S. If $\{M_n\}$ is a sequence of nonnegative numbers such that*

$$0 \leq |f_n(x)| \leq M_n, \quad n = 0, 1, \ldots, \quad x \in S,$$

then $\sum_{n=0}^{\infty} f_n(x)$ converges uniformly on S if $\sum_{n=1}^{\infty} M_n$ converges.

Proof: See [64, pg 223] for proof. •

Lemma 2.5.5 *Let $S \subset \mathbb{R}$ be an open interval and let $\{f_n\}$ be a real valued function defined on S such that the derivative $f_n^{'}(x)$ exists for every $x \in S$. Assume that for at least one*

point x_0 in S, $\sum_{n=1}^{\infty} f_n(x)$ converges. Assume further that there exists a function g such that $\sum_{n=1}^{\infty} f_n^{'}(x) = g(x)$ uniformly on S. Then

1. For every $x \in S$ there exists a function $f(x)$ such that $\sum_{n=1}^{\infty} f_n(x) = f(x)$ uniformly on S.

2. For every $x \in S$ the derivative $f^{'}(x)$ exists and equals $g(x)$.

Proof: See [64, pg 230, theorem 9.14] for proof. •

Lemma 2.5.6 *Let $\Theta \subseteq \mathbb{R}^p$ be open and let $(\Omega, \mathcal{F}, P)$ denote a probability space. Define $X : \Omega \to Z^+$ to be a discrete random variable with finite first moment where Z^+ denotes the set of non negative integers (including zero) and define $Y : \Omega \to \mathbb{R}$ to be a continuous random variable. Let p_Y denote the probability density function of Y such that $0 \leq p_Y \leq V_0 < \infty$ except on a Lebesgue null set and let $p_{X,\theta}$ denote the probability mass function of X that depends on a vector parameter $\theta \in \Theta$. Then*

1. $$\frac{\partial}{\partial z}\left[\sum_{n=0}^{\infty} p_{X,\theta}(n) \int_{-\infty}^{z-n} p_Y(y)dy\right] = \sum_{n=0}^{\infty} p_{X,\theta}(n) p_Y(z-n), \quad z \in \mathbb{R}, \quad \theta \in \Theta,$$

except on a Lebesgue null set.

2. If $Z = X + Y$ and X and Y are independent, then Z is a real valued random variable with density function $p_{Z,\theta}$ given by

$$p_{Z,\theta}(z) = \sum_{n=0}^{\infty} p_{X,\theta}(n) p_Y(z-n), \quad z \in \mathbb{R}, \quad \theta \in \Theta, \text{ except on a Lebesgue null set.}$$

3. For $\theta \in \Theta$,

$$\int_{-\infty}^{\infty} \left(\sum_{n=0}^{\infty} n p_{X,\theta}(n) p_Y(z-n)\right) dz = \sum_{n=0}^{\infty} n p_{X,\theta}(n).$$

4. For $\theta \in \Theta$,

$$\int_{-\infty}^{\infty} \left(\sum_{n=0}^{\infty} p_{X,\theta}(n) p_Y(z-n)\right) dz = \sum_{n=0}^{\infty} p_{X,\theta}(n).$$

5. Suppose X is a Poisson random variable with parameter μ_θ that depends on a vector parameter $\theta \in \Theta$ such that $\partial\mu_\theta/\partial\theta_i$ exist and is finite for every $\theta \in \Theta$, $i = 1, \ldots, p$, then

$$\frac{\partial}{\partial\theta_i}\left[\sum_{n=0}^{\infty} p_{X,\theta}(n) p_Y(z-n)\right] = \sum_{n=0}^{\infty} \frac{\partial p_{X,\theta}(n)}{\partial\theta_i} p_Y(z-n), \quad i = 1, \ldots, p \quad z \in \mathbb{R}, \quad \theta \in \Theta,$$

except on a Lebesgue null set.

Proof: 1. Let $F_{n,\theta}$ be given by

$$F_{n,\theta}(z) := p_{X,\theta}(n) \int_{-\infty}^{z-n} p_Y(y)dy = p_{X,\theta}(n)P_Y(z-n), \quad z \in \mathbb{R},$$

where $n = 0, 1, \ldots$, $\theta \in \Theta$ and P_Y denotes the distribution function of Y. Consider the term

$$\sum_{n=0}^{\infty} F_{n,\theta}(z) = \sum_{n=0}^{\infty} p_{X,\theta}(n)P_Y(z-n) \leq \sum_{n=0}^{\infty} p_{X,\theta}(n) = 1,$$

since $0 \leq P_Y \leq 1$. Thus by Lemma 2.5.4 the series $F_{n,\theta}(z)$ converges uniformly for $z \in \mathbb{R}$. Let $f_{n,\theta}$ be given by

$$f_{n,\theta}(z) := \frac{\partial F_{n,\theta}(z)}{\partial z} = p_{X,\theta}(n)\frac{\partial}{\partial z}\int_{-\infty}^{z-n} p_Y dy = p_{X,\theta}p_Y(z-n), \quad n = 0, 1, \ldots, \quad z \in \mathbb{R}, \quad \theta \in \Theta.$$

Consider the term

$$\sum_{n=0}^{\infty} f_{n,\theta}(z) = \sum_{n=0}^{\infty} p_{X,\theta}(n)p_Y(z-n) \leq V_0 \sum_{n=0}^{\infty} p_{X,\theta}(n) \leq V_0, \tag{2.22}$$

since $0 \leq p_Y \leq V_0$ by definition. Using Lemma 2.5.4 $\sum_{n=0}^{\infty} f_{n,\theta}(z)$ converges uniformly for $z \in \mathbb{R}$. Hence using Lemma 2.5.5 we have

$$\frac{\partial}{\partial z}\left[\sum_{n=0}^{\infty} p_{X,\theta}(n) \int_{-\infty}^{z-n} p_Y(y)dy\right] = \sum_{n=0}^{\infty} p_{X,\theta}(n)p_Y(z-n), \quad z \in \mathbb{R}, \quad \theta \in \Theta.$$

2. Let n be a non negative integer, $y, z \in \mathbb{R}$. Consider the term $P[Z \leq z] = P[X + Y \leq z]$. In terms of conditional probability $P[X + Y \leq Z]$ is given by

$$P[X+Y \leq z] = \sum_{n=0}^{\infty} P[X+Y \leq z|X=n]P[X=n] = \sum_{n=0}^{\infty} P[Y \leq z-n|n]P[X=n]$$

$$= \sum_{n=0}^{\infty} P[Y \leq z-n]P[X=n] = \sum_{n=0}^{\infty} p_{X,\theta}(n) \int_{-\infty}^{z-n} p_Y(y)dy,$$

where we have the fact that X and Y are independent in the third step. Differentiating the above equation with respect to z and using result 1 we get the desired result.

3. Let $\gamma = \{0, 1\}$ and let $f_{n,\theta}$, f_θ be given by

$$f_{n,\theta}(z) := \sum_{k=0}^{n} k^\gamma p_{X,\theta}(k)p_Y(z-k), \quad n = 0, 1, \ldots, \quad z \in \mathbb{R}, \quad \theta \in \Theta,$$

$$f_\theta(z) := \sum_{k=0}^{\infty} k^\gamma p_{X,\theta}(k) p_Y(z-k), \quad n = 0, 1, \ldots, \quad z \in \mathbb{R}, \quad \theta \in \Theta.$$

Then $f_{n,\theta}(z) \to f_\theta(z)$ for all $z \in \mathbb{R}$, $\theta \in \Theta$ except possibly on a Lebesgue null set. Also $f_{n,\theta}(z) > f_{m,\theta}(z)$ when $n > m$ since by definition $p_{X,\theta}$ and p_Y are non negative. Hence by Monotone Convergence Theorem we have

$$\int_{\mathbb{R}} f_\theta(z)dz = \int_{\mathbb{R}} \lim_{n\to\infty} f_{n,\theta}(z)dz = \lim_{n\to\infty} \int_{\mathbb{R}} f_{n,\theta}(z)dz = \lim_{n\to\infty} \int_{\mathbb{R}} \sum_{k=0}^{n} k^\gamma p_{X,\theta}(k) p_Y(z-k)dz$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} k^\gamma p_{X,\theta}(k) \int_{\mathbb{R}} p_Y(z-k)dz = \lim_{n\to\infty} \sum_{k=0}^{n} k^\gamma p_{X,\theta}(k) = \sum_{k=0}^{\infty} k^\gamma p_{X,\theta}(k), \quad \theta \in \Theta.$$

Setting $\gamma = 1$ we get result 3 and settting $\gamma = 0$ we get result 4.

4. See above for proof.

5. We prove this result similar to result 1. Let $f_{n,\theta}$ be given by

$$f_{n,\theta}(z) = p_{X,\theta}(n) p_Y(z-n), \quad z \in \mathbb{R}, \quad \theta \in \Theta, \quad n = 0, 1, \ldots$$

From eq. 2.22 we know $\sum_{n=0}^{\infty} f_{n,\theta}(z)$ converges uniformly for every $z \in \mathbb{R}$. Let $g_{n,\theta}$ be given by

$$g_{n,\theta}(z) = \frac{\partial f_{n,\theta}(z)}{\partial \theta_i} = \frac{\partial \mu_\theta}{\partial \theta_i} \left[\frac{n}{\mu_\theta} - 1 \right] p_{X,\theta}(n) p_Y(z-n), \quad z \in \mathbb{R}, \quad \theta \in \Theta, \quad n = 0, 1, \ldots, \quad (2.23)$$

where $\theta_i \in \theta$ and $i = 1, \ldots, p$. Consider the term

$$\sum_{n=0}^{\infty} g_{n,\theta}(z) = \sum_{n=0}^{\infty} \frac{\partial p_{X,\theta}}{\partial \theta_i} p_Y(z-n) = \frac{\partial \mu_\theta}{\partial \theta_i} \sum_{n=0}^{\infty} \left[\frac{n}{\mu_\theta} - 1 \right] p_{X,\theta}(n) p_Y(z-n)$$

$$\leq \frac{\partial \mu_\theta}{\partial \theta_i} \sum_{n=0}^{\infty} \frac{n}{\mu_\theta} p_{X,\theta}(n) p_Y(z-n) \leq \frac{V_0}{\mu_\theta} \frac{\partial \mu_\theta}{\partial \theta_i} \sum_{n=0}^{\infty} n p_{X,\theta}(n) = V_0 \frac{\partial \mu_\theta}{\partial \theta_i}, \quad z \in \mathbb{R}, \quad \theta \in \Theta,$$

since $0 \leq p_Y \leq V_0$ and $i = 1, \ldots, p$. Hence using Lemma 2.5.4 the series $\sum_{n=1}^{\infty} g_{n,\theta}(z)$ converges uniformly for $z \in \mathbb{R}$ and $\theta \in \Theta$ and using Lemma 2.5.5 we get the desired result. •

CHAPTER 3

LOCALIZATION ACCURACY PROBLEM

3.1 Introduction

The localization accuracy problem is concerned with how accurately the location of a microscopic object (for example, single molecule) can be estimated from data acquired in an optical microscope. Here, the results derived in Chapter 2 are used to obtain analytical expressions of the Fisher information matrix for the location estimation problem. The organization of this chapter is as follows. In Section 3.2 a general expression for the Fisher information matrix is derived corresponding to the location estimation problem. Conditions that guarantee block diagonality or diagonality of the Fisher information matrix are also derived.

In Section 3.3 the *image function* is introduced. Previously, the derivation of the Fisher information matrix made no specific assumptions about the functional form of the image profile of the object. In many practical cases, the image profile can be assumed to be shift-invariant. In such cases, the image of the object can be described as a scaled and shifted version of the image function. In this section, analytical expressions of the Fisher information matrix are obtained in terms of the image function. These results are illustrated by considering specific image profiles that describe the image of a fluorescent point source (for example, single molecule). In particular by using the Airy profile, which, according to optical diffraction theory, describes the image of an infocus point source, a simple formulae is derived that provides a *fundamental* limit to the 2D localization accuracy of a single molecule.

Section 3.4 considers the effects of pixelation and noise sources. Here analytical expressions for the limit of the 2D localization accuracy are derived for an Airy profile. The results are extensively illustrated by investigating the effect of various parameters such as magnification, pixel size, pixel array size, photon detection rate etc., on the limit of the 2D localization accuracy. Concrete examples are given to show how these results can be used to design and op-

timize single molecule imaging setups from the point of view of achieving maximum accuracy in location estimation. A maximum likelihood estimation algorithm is introduced for determining the 2D location of a single molecule from optical microscopy data. Throughout the chapter the algorithm is tested for different imaging configurations. It is shown that the maximum likelihood estimator can correctly determine 2D location of single molecules with an accuracy that is consistently close to the theoretical result given by the limit of the 2D localization accuracy.

3.2 Performance limits and Fisher information matrix

For the location estimation problem the parameter vector $\theta := (\theta_l, \theta_a, \theta_\Lambda)$ is decomposed into three components, where θ_l denotes the location component, θ_a denotes an auxiliary component and θ_Λ denotes the rate component. The location component θ_l typically consists of the x and y coordinates of the object location. The auxiliary component θ_a, if present, may consist of other relevant parameters such as the z coordinate of the object location. In single molecule microscopy it is often reasonable to assume that the photon detection rate is a constant, i.e., $\Lambda_\theta(\tau) = \Lambda_0$, $\tau \geq t_0$, and that the intensity level Λ_0 needs to be estimated. In this case θ_Λ would consist of the parameter Λ_0. For example, in tracking problems for vesicles, due to photobleaching effects the photon detection rate can often be modeled as $\Lambda_\theta(\tau) = \Lambda_0 \exp(-(\tau - t_0)k_b)$, $\tau \geq t_0$. In this case θ_Λ would consist of the parameters Λ_0 and k_b.

In the following theorem, we consider two independent image detection processes $\mathcal{G}^1$ and $\mathcal{G}^2$. The image detection process $\mathcal{G}^1$ is such that its photon detection rate Λ^1_θ only depends on the rate component θ_Λ, and its photon distribution profile $f^1_{\theta,\tau}$ only depends on the location component θ_l and the auxiliary component θ_a. The image detection process $\mathcal{G}^2$ is such that its photon detection rate Λ^2 and photon distribution profile f^2_τ are independent of θ. For example, $\mathcal{G}^1$ can model the detected photons from the object of interest, whereas $\mathcal{G}^2$ might model a background component.

Theorem 3.2.1 *Let $\mathcal{G}^1(\Lambda^1_\theta, \{f^1_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$ and $\mathcal{G}^2(\Lambda^2, \{f^2_\tau\}_{\tau\geq t_0}, \mathcal{C})$ be two independent image detection processes such that $\mathcal{G}^2$ is independent of θ, $\theta \in \Theta$. Let $\mathcal{G}$ be the superposition of $\mathcal{G}^1$*

and $\mathcal{G}^2$. Assume that

A1. *for $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$ and $\tau \geq t_0$, $\partial f^1_{\theta,\tau}(r)/\partial\theta_\Lambda = 0$, $r \in \mathcal{C}$, $\partial\Lambda^1_\theta(\tau)/\partial\theta_l = 0$, $\partial\Lambda^1_\theta(\tau)/\partial\theta_a = 0$.*

Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & \mathbf{I}_{l,a}(\theta) & \mathbf{I}_{l,\Lambda}(\theta) \\ \mathbf{I}^T_{l,a}(\theta) & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ \mathbf{I}^T_{l,\Lambda}(\theta) & \mathbf{I}^T_{a,\Lambda}(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix}, \tag{3.1}$$

where, for $\theta \in \Theta$,

$$\mathbf{I}_{\alpha,\beta}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_\alpha}\right)^T \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_\beta}\right) dr d\tau, \quad \alpha, \beta \in \{l, a\}, \tag{3.2}$$

$$\mathbf{I}_{\alpha,\Lambda}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r)}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_\alpha}\right)^T \left(\frac{\partial\Lambda^1_\theta(\tau)}{\partial\theta_\Lambda}\right) dr d\tau, \quad \alpha \in \{l, a\}, \tag{3.3}$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{(f^1_{\theta,\tau}(r))^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \left(\frac{\partial\Lambda_\theta(\tau)}{\partial\theta_\Lambda}\right)^T \left(\frac{\partial\Lambda^1_\theta(\tau)}{\partial\theta_\Lambda}\right) dr d\tau. \tag{3.4}$$

Proof: By Theorem 2.2.2, for the image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$, $\Lambda_\theta(\tau) := \Lambda^1_\theta(\tau) + \Lambda^2(\tau)$ and $f_{\theta,\tau}(r) := \epsilon^1_\theta(\tau) f^1_{\theta,\tau}(r) + \epsilon^2_\theta(\tau) f^2_\tau(r)$ for $r \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, where $\epsilon^1_\theta(\tau) := \Lambda^1_\theta(\tau)/(\Lambda^1_\theta(\tau) + \Lambda^2(\tau))$ and $\epsilon^2_\theta(\tau) := \Lambda^2(\tau)/(\Lambda^1_\theta(\tau) + \Lambda^2(\tau))$, $\tau \geq t_0$, $\theta \in \Theta$. Thus we have $\Lambda_\theta(\tau) f_{\theta,\tau}(r) = (\Lambda^1_\theta(\tau) + \Lambda^2(\tau))\ (\epsilon^1_\theta(\tau) f^1_{\theta,\tau}(r) + \epsilon^2_\theta(\tau) f^2_\tau(r)) = \Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)$, $r \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$. Substituting this in Theorem 2.2.1, using condition **A1** and the fact that Λ^2 and f^2_τ are independent of θ, we get

$$\begin{aligned} \mathbf{I}(\theta) &= \int_{\mathcal{C}} \int_{t_0}^t \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial\theta}\right)^T \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial\theta}\right) dr d\tau \\ &= \int_{t_0}^t \int_{\mathcal{C}} \frac{1}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \begin{bmatrix} \Lambda^1_\theta(\tau) \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_l}\right)^T \\ \Lambda^1_\theta(\tau) \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_a}\right)^T \\ f^1_{\theta,\tau}(r) \left(\frac{\partial\Lambda^1_\theta(\tau)}{\partial\theta_\Lambda}\right)^T \end{bmatrix} \times \\ &\qquad \begin{bmatrix} \Lambda^1_\theta(\tau) \frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_l} & \Lambda^1_\theta(\tau) \frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_a} & f^1_{\theta,\tau}(r) \frac{\partial\Lambda^1_\theta(\tau)}{\partial\theta_\Lambda} \end{bmatrix} dr d\tau, \quad \theta \in \Theta. \end{aligned}$$

From the above equation the result immediately follows. •

In many practical situations it is important to know whether the Fisher information matrix $\mathbf{I}(\theta)$ is diagonal, as the diagonality of $\mathbf{I}(\theta)$ has several implications. For example, it is well known that under certain conditions the maximum likelihood estimator of a vector parameter θ is asymptotically Gaussian distributed with asymptotic mean θ and covariance $\mathbf{I}^{-1}(\theta)$ (see e.g., [55]). Here, if $\mathbf{I}(\theta)$ is diagonal, this implies that the components of the maximum likelihood estimate of θ are asymptotically independent. Note that if an efficient estimator of θ exists (i.e., an unbiased estimator whose covariance matrix is equal to $\mathbf{I}^{-1}(\theta)$, $\theta \in \Theta$), then a diagonal $\mathbf{I}(\theta)$ ensures that the estimates of θ are uncorrelated.

In general, for $\theta = (\theta_1, \ldots, \theta_n) \in \Theta$, if $\mathbf{I}(\theta)$ is diagonal, then this implies that the *limit* of the accuracy of the unbiased estimates of θ_i, $i = 1, \ldots, n$, does not depend on the other unknown parameters in θ. For an unbiased estimator of the object location, this means that the limit of the localization accuracy of the x coordinate of the location is the same whether or not the y and z coordinates of the location are known.

We next investigate the conditions under which the Fisher information matrix given in Theorem 3.2.1 is block diagonal. As will be shown, for the parameter vector $\theta = (\theta_l, \theta_a, \theta_\Lambda)$, it turns out that $\mathbf{I}(\theta)$ is block diagonal when the detector $\mathcal{C}$ and the photon distribution profiles $f^1_{\theta,\tau}$ and f^2_τ satisfy certain symmetry conditions. Furthermore, for some special cases of θ, $\mathbf{I}(\theta)$ becomes fully diagonal (see Corollary 3.3.1). We next define a symmetric detector and a symmetric function.

Definition 3.2.1 *1. A detector $\mathcal{C}$ is said to be* **symmetric** *if there exists a point $(c_x, c_y) \in \mathbb{R}^2$, known as the* **center** *of $\mathcal{C}$, such that for every $(x, y) \in \mathcal{C}$, $(2c_x - x, y) \in \mathcal{C}$, $(x, 2c_y - y) \in \mathcal{C}$ and $(2c_x - x, 2c_y - y) \in \mathcal{C}$.*

2. Let $\mathcal{C}$ be a symmetric detector with center (c_x, c_y). A function $f : \mathcal{C} \to \mathbb{R}$ is said to be **symmetric (antisymmetric) along the x axis with respect to c_x** *if for every $(x, y) \in \mathcal{C}$, $f(x, y) = f(2c_x - x, y)$ $(f(x, y) = -f(2c_x - x, y))$. If f is symmetric along both the x and y*

axes with respect to c_x and c_y respectively, then f is said to be **symmetric with respect to the center of** $\mathcal{C}$.

In the following theorem we assume the location component θ_l to be $\theta_l = (\theta_1, \theta_2)$. Further, we assume that $(\partial f^1_{\theta,\tau}(r)/\partial\theta_1) = -\vartheta_x(\partial f^1_{\theta,\tau}(r)/\partial x)$ and $(\partial f^1_{\theta,\tau}(r)/\partial\theta_2) = -\vartheta_y\ (\partial f^1_{\theta,\tau}(r)/\partial y)$, $r = (x, y) \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, where ϑ_x and ϑ_y are constants that are independent of $(x, y) \in \mathcal{C}$. This assumption is satisfied if the photon distribution profile $f^1_{\theta,\tau}$ is a function of $(x/\vartheta_x - \theta_1, y/\vartheta_y - \theta_2)$, where $(x, y) \in \mathcal{C}$ (see Section 3.3).

Theorem 3.2.2 *Let $\mathcal{G}^1(\Lambda^1_\theta, \{f^1_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$, $\mathcal{G}^2(\Lambda^2, \{f^2_\tau\}_{\tau\geq t_0}, \mathcal{C})$ and $\mathcal{G}$ be image detection processes as given in Theorem 3.2.1. Assume that*

A1. *for $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, $\theta_l = (\theta_1, \theta_2)$, $\theta_a = (\theta_3, \ldots, \theta_k)$ and $\theta_\Lambda = (\theta_{k+1}, \ldots, \theta_n)$, $\partial f^1_{\theta,\tau}(r)/\partial\theta_\Lambda = 0$, $r \in \mathcal{C}$, $\tau \geq t_0$, $\partial\Lambda^1_\theta(\tau)/\partial\theta_l = 0$, $\tau \geq t_0$ and $\partial\Lambda^1_\theta(\tau)/\partial\theta_a = 0$, $\tau \geq t_0$,*

A2. *$\partial f^1_{\theta,\tau}(r)/\partial\theta_1 = -\vartheta_x(\partial f^1_{\theta,\tau}(r)/\partial x)$, $\partial f^1_{\theta,\tau}(r)/\partial\theta_2 = -\vartheta_y(\partial f^1_{\theta,\tau}(r)/\partial y)$, $r = (x, y) \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, where ϑ_x and ϑ_y are constants that are independent of x and y,*

A3. *the detector $\mathcal{C}$ is symmetric, and*

A4. *$f^1_{\theta,\tau}$ and f^2_τ are symmetric with respect to the center of the detector $\mathcal{C}$ for $\theta \in \Theta$ and $\tau \geq t_0$.*

1. Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & \mathbf{I}_{l,a}(\theta) & 0 \\ \mathbf{I}^T_{l,a}(\theta) & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ 0 & \mathbf{I}^T_{a,\Lambda}(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where for $\theta \in \Theta$, $\mathbf{I}_{l,a}(\theta)$ and $\mathbf{I}_{a,a}(\theta)$ are given by eq. 3.2, $\mathbf{I}_{a,\Lambda}(\theta)$ is given by eq. 3.3, $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ is given by eq. 3.4 and

$$\mathbf{I}_{l,l}(\theta) := \begin{bmatrix} \vartheta_x^2 \int_{t_0}^t \int_{\mathcal{C}} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial x}\right)^2 dr d\tau & 0 \\ 0 & \vartheta_y^2 \int_{t_0}^t \int_{\mathcal{C}} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial y}\right)^2 dr d\tau \end{bmatrix}. \quad (3.5)$$

2. In addition to conditions **A1 - A4**, *assume that*

A5. *all the elements of the vector $\partial f^1_{\theta,\tau}(r)/\partial\theta_a$, $r \in \mathcal{C}$, are symmetric with respect to the center of the detector $\mathcal{C}$ for $\theta \in \Theta$ and $\tau \geq t_0$.*

Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ 0 & \mathbf{I}_{a,\Lambda}^T(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where all the non-zero entries of $\mathbf{I}(\theta)$ are given in result 1.

3. In addition to conditions **A1 - A5**, *assume that*

A6. *the photon detection rate of $\mathcal{G}^2$ is zero, i.e., $\Lambda^2(\tau) = 0$, $\tau \geq t_0$.*

Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & 0 \\ 0 & 0 & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix}, \tag{3.6}$$

where for $\theta \in \Theta$,

$$\mathbf{I}_{l,l}(\theta) := \begin{bmatrix} \vartheta_x^2 \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial x} \right)^2 dr d\tau & 0 \\ 0 & \vartheta_y^2 \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial y} \right)^2 dr d\tau \end{bmatrix}, \tag{3.7}$$

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} \right)^T \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} dr d\tau, \tag{3.8}$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^t \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right)^T \frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} d\tau. \tag{3.9}$$

Proof: 1. Using condition **A1** and the fact that Λ^2 and f_τ^2 are independent of θ, we can show that the general expression for the Fisher information matrix is given by Theorem 3.2.1 (see eq. 3.1). Consider the matrix $\mathbf{I}_{l,l}(\theta)$ that is given by eq. 3.2. Using condition **A2** we immediately obtain the integral expressions of $[\mathbf{I}_{l,l}(\theta)]_{11}$ and $[\mathbf{I}_{l,l}(\theta)]_{22}$ that are given in eq. 3.5. To obtain the desired result, we need to show that the off-diagonal terms of $\mathbf{I}_{l,l}(\theta)$ (i.e., $[\mathbf{I}_{l,l}(\theta)]_{12}$ and $[\mathbf{I}_{l,l}(\theta)]_{21}$) and all the terms of $\mathbf{I}_{l,\Lambda}(\theta)$ (i.e., $[\mathbf{I}_{l,\Lambda}(\theta)]_{ij}$, $i = 1, 2$, $j = k+1, \ldots, n$) are zero.

Let (c_x, c_y) denote the center of the detector $\mathcal{C}$ and define $T_X : \mathcal{C} \to \mathbb{R}^2$, $(x, y) \mapsto (2c_x - x, y)$. By condition **A4**, $f^1_{\theta,\tau}(x, y) = (f^1_{\theta,\tau} \circ T_X)(x, y)$, $(x, y) \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$. Using the chain rule of differentiation we get

$$\frac{\partial f^1_{\theta,\tau}(x,y)}{\partial x} = \frac{\partial[(f^1_{\theta,\tau} \circ T_X)(x,y)]}{\partial x} = \frac{\partial f^1_{\theta,\tau}(2c_x - x, y)}{\partial x}\frac{\partial(2c_x - x)}{\partial x} = -\left(\frac{\partial f^1_{\theta,\tau}}{\partial x} \circ T_X\right)(x,y), \tag{3.10}$$

where $(x, y) \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$. Similarly, by using condition **A4** we can show that for $\theta \in \Theta$ and $\tau \geq t_0$,

$$\frac{\partial f^1_{\theta,\tau}(x,y)}{\partial y} = \frac{\partial[(f^1_{\theta,\tau} \circ T_X)(x,y)]}{\partial y} = \left(\frac{\partial f^1_{\theta,\tau}}{\partial y} \circ T_X\right)(x,y), \quad (x,y) \in \mathcal{C}, \tag{3.11}$$

$$\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(x,y) + \Lambda^2(\tau) f^2_\tau(x,y) = \left((\Lambda^1_\theta(\tau) f^1_{\theta,\tau} + \Lambda^2(\tau) f^2_\tau) \circ T_X\right)(x,y), \quad (x,y) \in \mathcal{C}. \tag{3.12}$$

Consider the term $[\mathbf{I}(\theta)]_{12}$ that is given by eq. 3.2. Hence using this result, condition **A2** and eqs. 3.10 - 3.12, we get for $\theta \in \Theta$,

$$\begin{aligned}
[\mathbf{I}_{l,l}(\theta)]_{12} &= \int_{t_0}^{t}\int_{\mathcal{C}} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(r) + \Lambda^2(\tau) f^2_\tau(r)} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_1}\frac{\partial f_{\theta,\tau}(r)}{\partial \theta_2} dr d\tau \\
&= \vartheta_x \vartheta_y \int_{t_0}^{t}\int_{\mathcal{C}} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(x,y) + \Lambda^2(\tau) f^2_\tau(x,y)} \frac{\partial f^1_{\theta,\tau}(x,y)}{\partial x}\frac{\partial f^1_{\theta,\tau}(x,y)}{\partial y} dx dy d\tau \\
&= -\vartheta_x \vartheta_y \int_{t_0}^{t}\int_{\mathcal{C}} \left(\left(\frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau} + \Lambda^2(\tau) f^2_\tau} \frac{\partial f^1_{\theta,\tau}}{\partial x}\frac{\partial f^1_{\theta,\tau}}{\partial y}\right) \circ T_X\right)(x,y) dx dy d\tau \\
= -\vartheta_x \vartheta_y \int_{t_0}^{t}\int_{\mathcal{C}} &\frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(x,y) + \Lambda^2(\tau) f^2_\tau(x,y)} \frac{\partial f^1_{\theta,\tau}(x,y)}{\partial x}\frac{\partial f^1_{\theta,\tau}(x,y)}{\partial y} dx dy d\tau = -[\mathbf{I}_{l,l}(\theta)]_{12},
\end{aligned} \tag{3.13}$$

where we have used the Theorem on change of variables (see [65, pg 153-155]) in the final step. Similarly we can show that $[\mathbf{I}_{l,l}(\theta)]_{21} = -[\mathbf{I}_{l,l}(\theta)]_{21}$, $\theta \in \Theta$. Thus we have $[\mathbf{I}_{l,l}(\theta)]_{12} = [\mathbf{I}_{l,l}(\theta)]_{21} = 0$, $\theta \in \Theta$.

If $T_Y : \mathcal{C} \to \mathbb{R}^2$, $(x, y) \mapsto (x, 2c_y - y)$, then similar to eqs. 3.10 and 3.12, we can show that for $\theta \in \Theta$ and $\tau \geq t_0$,

$$\frac{\partial f^1_{\theta,\tau}(x,y)}{\partial y} = \frac{\partial[(f^1_{\theta,\tau} \circ T_Y)(x,y)]}{\partial y} = -\left(\frac{\partial f^1_{\theta,\tau}}{\partial y} \circ T_Y\right)(x,y), \quad (x,y) \in \mathcal{C},\ \tau \geq t_0,\ \theta \in \Theta, \tag{3.14}$$

$$\Lambda^1_\theta(\tau) f^1_{\theta,\tau}(x,y) + \Lambda^2(\tau) f^2_\tau(x,y) = \left((\Lambda^1_\theta(\tau) f^1_{\theta,\tau} + \Lambda^2(\tau) f^2_\tau) \circ T_Y\right)(x,y), \quad (x,y) \in \mathcal{C}. \tag{3.15}$$

Hence by using eqs. 3.10, 3.12, 3.14 and 3.15 and condition **A2**, we can show that $[\mathbf{I}_{l,\Lambda}(\theta)]_{ij} = -[\mathbf{I}_{l,\Lambda}(\theta)]_{ij}$, for $i = 1, 2$, $j = k+1, \ldots, n$ and $\theta \in \Theta$. Hence $\mathbf{I}_{l,\Lambda}(\theta) = 0$, $\theta \in \Theta$. From this the result follows.

2. By condition **A5**, for $\zeta \in \{X, Y\}$, $\partial f^1_{\theta,\tau}(x,y)/\partial\theta_a = \partial[(f^1_{\theta,\tau} \circ T_\zeta)(x,y)]/\partial\theta_a = \Big((\partial f^1_{\theta,\tau}/\partial\theta_a) \circ T_\zeta\Big)(x,y)$, $(x,y) \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$, where T_X and T_Y are defined above. Hence using this result, eqs. 3.10 - 3.12, eqs. 3.14 - 3.15, condition **A2** and by taking an approach similar to that of eq. 3.13, we can show that $[\mathbf{I}_{l,a}(\theta)]_{ij} = -[\mathbf{I}_{l,a}(\theta)]_{ij}$, $i = 1, 2$, $j = 3, \ldots, k$, $\theta \in \Theta$. From this it follows that $\mathbf{I}_{l,a}(\theta) = 0$, $\theta \in \Theta$. Substituting this in result 1 of this Theorem the result follows immediately.

3. Substituting for Λ^2 in eq. 3.5 we immediately obtain the expression for $\mathbf{I}_{l,l}(\theta)$ that is given by eq. 3.7. Consider the term $\mathbf{I}_{a,\Lambda}(\theta)$ that is given by eq. 3.3 (see Theorem 3.2.1). Since $f^1_{\theta,\tau}$ is a density function that satisfies conditions **C3 - C5** of Definition 2.2.1, $\int_{\mathcal{C}} (\partial f^1_{\theta,\tau}(r)/\partial\theta) = 0$, $\theta \in \Theta$ and $\tau \geq t_0$ ([55, pg 182-183]). Using this result and substituting for Λ^2, we get

$$\mathbf{I}_{a,\Lambda}(\theta) = \int_{t_0}^{t} \left(\int_{\mathcal{C}} \frac{\partial f^1_{\theta,\tau}(r)}{\partial\theta_a} \right)^T dr \frac{\partial \Lambda_\theta(\tau)}{\partial\theta_\Lambda} d\tau = 0, \quad \theta \in \Theta.$$

Finally, consider the term $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ that is given by eq. 3.4. Substituting for Λ^2 and using the fact that $\int_{\mathcal{C}} f^1_{\theta,\tau}(r) dr = 1$, $\tau \geq t_0$, $\theta \in \Theta$, we obtain the desired expression that is given in eq. 3.9. •

Remark 3.2.1 *Note that when condition* **A6** *is satisfied, the photon detection rate and the photon distribution profile of* $\mathcal{G}$ *are equal to that of* $\mathcal{G}^1$*. In this case, the condition that (see condition* **A4***)* f^2_τ *is symmetric with respect to the center of the detector is no longer required.*

From the above theorem we see that the number of symmetry conditions imposed on the partial derivatives of $f^1_{\theta,\tau}$ (i.e., $\partial f^1_{\theta,\tau}(r)/\partial\theta_l$ and $\partial f^1_{\theta,\tau}(r)/\partial\theta_a$) determines the number of off-diagonal terms that are zero in the Fisher information matrix $\mathbf{I}(\theta)$. Note that the photon detection rate Λ^2 also plays a crucial role in making $\mathbf{I}(\theta)$ block diagonal. Consider a special case, where the parameter vector $\theta = (\theta_l, \theta_\Lambda)$ only consists of the location θ_l and the rate θ_Λ

components. In this case the Fisher information matrix given in result 1 of the above theorem is block diagonal, even when the photon detection rate $\Lambda^2(\tau) \neq 0$, $\tau \geq t_0$. Finally, we note that in result 3 of the above theorem if the auxiliary component θ_a and the rate component θ_Λ are scalars, then $\mathbf{I}(\theta)$ is diagonal.

3.3 Image function

In the previous sections we made no assumptions about the specific functional form of the photon distribution profile. In an optical microscope, the image of an object can often be considered to be invariant with respect to shifts in the object location ([66]). Hence the photon distribution profile $f^1_{\theta,\tau}$ can be expressed as a scaled and shifted version of the image of the object. For example, in the case of a moving object, $f^1_{\theta,\tau}$ can be written as $f^1_{\theta,\tau}(x,y) = \frac{1}{M^2} q_{\theta_e}(\frac{x}{M} - x_{0,\tau}, \frac{y}{M} - y_{0,\tau})$, $(x,y) \in \mathbb{R}^2$, $\theta \in \Theta$, $\tau \geq t_0$, where q_{θ_e} denotes an image function, $M > 0$ denotes the lateral magnification and $(x_{0,\tau}, y_{0,\tau})$ denotes the time dependent $x-y$ location of the object. An image function q_{θ_e} describes the image of a fixed object on the detector plane at unit lateral magnification when the object is located along the z axis in the object space. Here, θ_e is a vector that parameterizes the image function. For example, θ_e could be the z position of the object and/or the angles that specify the 3D orientation of the object. In some applications the θ_e parameterization is not required and in such cases the image function is denoted as q. Since $f^1_{\theta,\tau}$ is a probability density function that satisfies conditions **C3 - C5** of Definition 2.2.1, to express $f^1_{\theta,\tau}$ in terms of q_{θ_e} we impose appropriate conditions on the image function that are given below.

Definition 3.3.1 *Let $\Theta_e \subseteq \mathbb{R}^m$ be a parameter space. For $\theta_e = (\theta_{e,1}, \ldots, \theta_{e,m}) \in \Theta_e$, we define $q_{\theta_e} : \mathbb{R}^2 \to [0,\infty)$ to be an* **image function** *if the following properties are satisfied.*

1. $\int_{\mathbb{R}^2} q_{\theta_e}(x,y)dxdy = 1$,
2. $\frac{\partial q_{\theta_e}(x,y)}{\partial x}$, $\frac{\partial q_{\theta_e}(x,y)}{\partial y}$ *and* $\frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}}$ *exist for every* $(x,y) \in \mathbb{R}^2$,
3. $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial x}\right| dxdy < \infty$, $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial y}\right| dxdy < \infty$ *and* $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}}\right| dxdy < \infty$, *and*
4. $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)} \frac{\partial q_{\theta_e}(x,y)}{\partial \zeta_k} \frac{\partial q_{\theta_e}(x,y)}{\partial \zeta_l} dxdy$, $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)} \frac{\partial q_{\theta_e}(x,y)}{\partial \zeta_k} \frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}} dxdy$ *and* $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)} \frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}} \frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,j}}$

dxdy exist and are finite, where $\zeta_1 = x$, $\zeta_2 = y$, $k, l = 1, 2$ *and* $i, j = 1, \ldots, m$.

The image function q_{θ_e} *and its derivative* $(\partial q_{\theta_e}/\partial \theta_{e,i})$ *are said to be symmetric if, for* $\theta_e \in \Theta_e$, $q_{\theta_e}(x, y) = q_{\theta_e}(-x, y) = q_{\theta_e}(x, -y)$ *and* $\frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}} = \frac{\partial q_{\theta_e}(-x,y)}{\partial \theta_{e,i}} = \frac{\partial q_{\theta_e}(x,-y)}{\partial \theta_{e,i}}$, *respectively, for* $(x, y) \in \mathbb{R}^2$.

In several applications the image of the object can be considered to be invariant with respect to time, for example, when the object is not moving during the acquisition of its image. In such cases, the expression for the photon distribution profile will be independent of time. In the following Corollary we derive a general expression for the Fisher information matrix for such applications. Here, the parameter vector is set to be $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, where $\theta_l = (x_0, y_0)$ denotes the $x-y$ location of the object and $\theta_\Lambda = \Lambda_0$ is a scalar parameter that characterizes the photon detection rate Λ^1_θ. We assume that the photon distribution profile $f^1_{\theta,\tau}$ is given in terms of a symmetric image function. We also assume the detector to be infinite, i.e., $\mathcal{C} = \mathbb{R}^2$. An infinite detector provides the best case scenario, where all the photons that reach the detector plane are detected by the detector. Hence the square root of the inverse Fisher information matrix for an infinite detector provides the *fundamental limit* to the accuracy with which the components of the parameter vector θ can be determined. We then consider a special case, where θ_l and θ_Λ are as given above and the auxiliary component $\theta_a = \epsilon_0$ is a scalar. For this special case we show that the Fisher information matrix is diagonal. Finally, we assume the photon distribution profile $f^1_{\theta,\tau}$ to be independent of θ_a and the parameter vector to be $\theta = (\theta_l, \theta_\Lambda)$ with θ_l and θ_Λ as given above.

Corollary 3.3.1 *Let* $\Theta \subseteq \mathbb{R}^n$ *be a parameter space. Let* $\mathcal{G}^1(\Lambda^1_\theta, \{f^1_{\theta,\tau}\}_{\tau \geq t_0}, \mathbb{R}^2)$ *and* $\mathcal{G}^2(\Lambda^2, \{f^2_\tau\}_{\tau \geq t_0}, \mathbb{R}^2)$ *be two independent image detection processes such that* $\mathcal{G}^2$ *is independent of* θ. *Let* $\mathcal{G}$ *be the superposition of* $\mathcal{G}^1$ *and* $\mathcal{G}^2$. *For* $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, *let* $\theta_l = (x_0, y_0)$ *and* $\theta_\Lambda = \Lambda_0$, *where* x_0, y_0 *and* Λ_0 *are scalar parameters. Assume that for* $\theta \in \Theta$,

A1. $\partial \Lambda^1_\theta(\tau)/\partial \theta_l = 0$, $\partial \Lambda^1_\theta(\tau)/\partial \theta_a = 0$ *and* $\Lambda^2(\tau) = 0$ *for* $\tau \geq t_0$,

A2. *there exists a symmetric image function q_{θ_a} such that for $M > 0$, the photon distribution profile $f^1_{\theta,\tau}$ is given by*

$$f^1_{\theta,\tau}(x,y) = \frac{1}{M^2} q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right), \quad (x,y) \in \mathbb{R}^2, \quad \tau \geq t_0,$$

A3. *all the elements of the vector $(\partial q_{\theta_a}(x,y)/\partial \theta_a)$ are symmetric, $(x,y) \in \mathbb{R}^2$.*

1. Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & 0 \\ 0 & 0 & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where for $\theta \in \Theta$,

$$\mathbf{I}_{l,l}(\theta) = \begin{bmatrix} \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial x}\right)^2 dxdy & 0 \\ 0 & \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial y}\right)^2 dxdy \end{bmatrix},$$

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a}\right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dxdy, \quad \theta \in \Theta,$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^t \frac{1}{\Lambda^1_\theta(\tau)} \left(\frac{\partial \Lambda^1_\theta(\tau)}{\partial \Lambda_0}\right)^2 d\tau, \quad \theta \in \Theta.$$

2. If $\theta_a = \epsilon_0$ is scalar, then for $\theta = (x_0, y_0, \epsilon_0, \Lambda_0) \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial x}\right)^2 dxdy & 0 & 0 & 0 \\ 0 & \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial y}\right)^2 dxdy & 0 & 0 \\ 0 & 0 & \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial \epsilon_0}\right)^2 dxdy & 0 \\ 0 & 0 & 0 & \int_{t_0}^t \frac{1}{\Lambda^1_\theta(\tau)} \left(\frac{\partial \Lambda^1_\theta(\tau)}{\partial \Lambda_0}\right)^2 d\tau \end{bmatrix}. \tag{3.16}$$

3. If $\theta = (x_0, y_0, \Lambda_0) \in \Theta$, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial x}\right)^2 dxdy & 0 & 0 \\ 0 & \int_{t_0}^t \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial y}\right)^2 dxdy & 0 \\ 0 & 0 & \int_{t_0}^t \frac{1}{\Lambda^1_\theta(\tau)} \left(\frac{\partial \Lambda^1_\theta(\tau)}{\partial \Lambda_0}\right)^2 d\tau \end{bmatrix}. \tag{3.17}$$

Proof: 1. Let $M > 0$. By conditions **A1** and **A2**, $\partial f^1_{\theta,\tau}(r)/\partial\theta_\Lambda = 0$, $\partial\Lambda^1_\theta(\tau)/\partial\theta_l = 0$ and $\partial\Lambda^1_\theta(\tau)/\partial\theta_a = 0$ for $r \in \mathbb{R}^2$, $\theta \in \Theta$ and $\tau \geq t_0$. Further, from condition **A2** we can verify that $\partial f^1_{\theta,\tau}(r)/\partial x_0 = -M(\partial f^1_{\theta,\tau}(r)/\partial x)$ and $\partial f^1_{\theta,\tau}(r)/\partial y_0 = -M(\partial f^1_{\theta,\tau}(r)/\partial y)$ for $r = (x, y) \in \mathbb{R}^2$, $\theta \in \Theta$ and $\tau \geq t_0$. Since $(Mx_0, My_0) \in \mathbb{R}^2$, for every $(x, y) \in \mathbb{R}^2$, $(2Mx_0 - x, y) \in \mathbb{R}^2$, $(x, 2My_0 - y) \in \mathbb{R}^2$ and $(2Mx_0 - x, 2My_0 - y) \in \mathbb{R}^2$. Hence $\mathbb{R}^2$ is symmetric with respect to the point (Mx_0, My_0) (see Definition 3.2.1). From conditions **A2 - A3** we can easily verify that $f^1_{\theta,\tau}(x, y)$ and all the elements of the vector $\partial f^1_{\theta,\tau}(x, y)/\partial\theta_a = (1/M^2)(\partial q_{\theta_a}\ (x/M - x_0, y/M - y_0)/\partial\theta_a)$ are symmetric with respect to the point (Mx_0, My_0), where $(x, y) \in \mathbb{R}^2$, $\theta \in \Theta$ and $\tau \geq t_0$. Finally we note that by condition **A1**, $\Lambda^2(\tau) = 0$, $\tau \geq t_0$.

Thus from the above and from Remark 3.2.1, we see that the photon distribution profile and the photon detection rate of $\mathcal{G}^1$ and $\mathcal{G}^2$ satisfy all of the conditions of result 3 of Theorem 3.2.2. Hence for the present case, the Fisher information matrix $\mathbf{I}(\theta)$ is block diagonal (see eq. 3.6).

From eq. 3.7 we see that $\mathbf{I}_{l,l}(\theta)$ is diagonal and we evaluate its diagonal terms $[\mathbf{I}_{l,l}(\theta)]_{11}$ and $[\mathbf{I}_{l,l}(\theta)]_{22}$. Substituting for Λ^1_θ and $f^1_{\theta,\tau}$ in the integral expression of $[\mathbf{I}_{l,l}(\theta)]_{11}$, we get

$$[\mathbf{I}_{l,l}(\theta)]_{11} = M^2 \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda^1_\theta(\tau)}{f^1_{\theta,\tau}(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial x} \right)^2 dr d\tau$$

$$= M^2 \int_{t_0}^{t} \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{\frac{1}{M^2} q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)} \left(\frac{1}{M^2} \frac{\partial q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x} \right)^2 dx dy$$

$$= \int_{t_0}^{t} \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{\Lambda^1_\theta(\tau)}{q_{\theta_a}(u, v)} \left(\frac{\partial q_{\theta_a}(u, v)}{\partial u} \frac{1}{M} \right)^2 (M du)(M dv)$$

$$= \int_{t_0}^{t} \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x, y)} \left(\frac{\partial q_{\theta_a}(x, y)}{\partial x} \right)^2 dx dy, \quad \theta \in \Theta,$$

where in the pen-ultimate step $u := \frac{x}{M} - x_0$ and $v := \frac{y}{M} - y_0$. Similarly, we can show that

$$[\mathbf{I}_{l,l}(\theta)]_{22} = \int_{t_0}^{t} \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x, y)} \left(\frac{\partial q_{\theta_a}(x, y)}{\partial y} \right)^2 dx dy,$$

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \Lambda^1_\theta(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{f^1_{\theta,\tau}(r)} \left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial \theta_a} \right)^T \frac{\partial f^1_{\theta,\tau}(r)}{\partial \theta_a} dr$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} \right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dxdy,$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right)^T \frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} d\tau = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \Lambda_0} \right)^2 d\tau,$$

where $\theta \in \Theta$, and $\mathbf{I}_{a,a}(\theta)$ and $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ are given by eqs. 3.8 and 3.9 respectively. From this the result follows.

2. If $\theta_a = \epsilon_0$ is scalar, then for $\theta \in \Theta$,

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} \right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dxdy$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial \epsilon_0} \right)^2 dxdy.$$

Substituting this in result 1 of this Corollary we obtain the desired result.

3. The result immediately follows from result 2 of this Corollary. •

From the above Corollary we see that the Fisher information matrix $\mathbf{I}(\theta)$ is independent of (x_0, y_0) and only depends on the image function and its partial derivatives. Moreover, $\mathbf{I}(\theta)$ is diagonal when $\theta = (x_0, y_0, \epsilon_0, \Lambda_0)$ and $\theta = (x_0, y_0, \Lambda_0)$. Note that if $\epsilon_0 = z_0$ denotes the z coordinate of the object location, then $\mathbf{I}(\theta)$ that is given in result 2 of the above Corollary can be used to calculate the three dimensional fundamental limit of the localization accuracy of the object. In [51], we recently reported integral expressions for $\mathbf{I}(\theta)$ that are analogous to eqs. 3.16 and 3.17, where the parameter vector was set to be $\theta = (x_0, y_0)$, Λ_0 was assumed to be known and $\mathbf{I}(\theta)$ was a 2×2 diagonal matrix. We note that eqs. 3.16 and 3.17 are generalizations of our earlier result and show that the diagonality of $\mathbf{I}(\theta)$ is preserved even when additional parameters such as ϵ_0 and Λ_0 are assumed to be unknown.

3.3.1 Examples

We now illustrate the results derived in this section by considering specific image functions that describe the image of a fixed point source. According to optical diffraction theory, when a point source is in focus with respect to the detector, the intensity distribution of the image of

the point source is described by the Airy profile ([56, pg 440]). The 2D Gaussian profile, on the other hand, has been widely used to approximate the Airy profile, for example, in the analysis of data from single molecule fluorescence experiments ([28, 26, 67]). In the following Corollary, the parameter vector is set to be $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and the photon distribution profile $f^1_{\theta,\tau}$ is assumed to be given in terms of an image function q. The photon detection rate is assumed to be a constant, i.e., $\Lambda^1_\theta(\tau) = \Lambda_0$, $\tau \geq t_0$. For each image function, we derive a simple formula for the fundamental limit of the localization accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{11}}$ ($\sqrt{[\mathbf{I}^{-1}(\theta)]_{22}}$) of x_0 (y_0) and for the fundamental limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$ of Λ_0. We note that the following results are extensions of the results reported in [51].

Corollary 3.3.2 *Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space. Let $\mathcal{G}^1$, $\mathcal{G}^2$ and $\mathcal{G}$ be image detection processes that are given in Corollary 3.3.1. For $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$, let $\Lambda^2(\tau) = 0$, $\Lambda^1_\theta(\tau) = \Lambda_0$ and for $M > 0$, assume that there exist a symmetric image function q such that $f^1_{\theta,\tau}(x, y) = (1/M^2)q(x/M - x_0, y/M - y_0)$, $(x, y) \in \mathbb{R}^2$.*

1. **Airy profile:** *If, for $n_a, \lambda > 0$, q is given by*

$$q(x,y) = \frac{J_1^2(\frac{2\pi n_a}{\lambda}\sqrt{x^2+y^2})}{\pi(x^2+y^2)}, \quad (x,y) \in \mathbb{R}^2, \tag{3.18}$$

then for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \frac{(2\pi n_a)^2 \Lambda_0 (t-t_0)}{\lambda^2} & 0 & 0 \\ 0 & \frac{(2\pi n_a)^2 \Lambda_0 (t-t_0)}{\lambda^2} & 0 \\ 0 & 0 & \frac{t-t_0}{\Lambda_0} \end{bmatrix}.$$

Further, the fundamental limit of the localization accuracy $\delta^{2d}_{x_0}$ ($\delta^{2d}_{y_0}$) of x_0 (y_0) and the fundamental limit of the accuracy $\delta^{2d}_{\Lambda_0}$ of Λ_0 are given by

$$\delta^{2d}_{x_0} = \delta^{2d}_{y_0} = \frac{\lambda}{2\pi n_a \sqrt{\Lambda_0 (t-t_0)}}, \quad \delta^{2d}_{\Lambda_0} = \sqrt{\frac{\Lambda_0}{(t-t_0)}}. \tag{3.19}$$

2. **2D Gaussian profile:** *If, for $\sigma > 0$, q is given by*

$$q(x,y) := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (x,y) \in \mathbb{R}^2, \tag{3.20}$$

then for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ *the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$ *is given by*

$$\mathbf{I}(\theta) = \begin{bmatrix} \frac{\Lambda_0(t-t_0)}{\sigma^2} & 0 & 0 \\ 0 & \frac{\Lambda_0(t-t_0)}{\sigma^2} & 0 \\ 0 & 0 & \frac{t-t_0}{\Lambda_0} \end{bmatrix}.$$

Further, the fundamental limit of the localization accuracy $\delta_{x_0}^{gau}$ ($\delta_{y_0}^{gau}$) *of* x_0 (y_0) *and the fundamental limit of the accuracy* $\delta_{\Lambda_0}^{gau}$ *of* Λ_0 *are given by*

$$\delta_{x_0}^{gau} = \delta_{y_0}^{gau} = \frac{\sigma}{\sqrt{\Lambda_0(t-t_0)}}, \quad \delta_{\Lambda_0}^{gau} = \sqrt{\frac{\Lambda_0}{(t-t_0)}}. \tag{3.21}$$

Proof: 1. It can be verified that the Airy profile is a symmetric image function. By definition, for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$, $\Lambda^2(\tau) = 0$, $\partial\Lambda_\theta^1(\tau)/\partial x_0 = \partial\Lambda_\theta^1(\tau)/\partial y_0 = 0$, $\partial f_{\theta,\tau}^1(r)/\partial\Lambda_0 = 0$, $r \in \mathbb{R}^2$, and $f_{\theta,\tau}^1$ is expressed as a shifted and scaled version of q. Hence for the present case result 3 of Corollary 3.3.1 holds and the Fisher information matrix $\mathbf{I}(\theta)$ is diagonal (see eq. 3.17). Using eq. 3.17 we can easily show that $[\mathbf{I}(\theta)]_{33} = (t-t_0)/\Lambda_0$. Let $\alpha := 2\pi n_a/\lambda$. Using the identity $(\partial/\partial x)\left[x^{-n}J_n(x)\right] = -x^{-n}J_{n+1}(x)$, $x \in \mathbb{R}$ (see [68, pg 18]) with $n = 1$, we can show that

$$\frac{\partial}{\partial\zeta}\frac{J_1^2(\alpha\sqrt{x^2+y^2})}{x^2+y^2} = -2\zeta\alpha\frac{J_1(\alpha\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}}\frac{J_2(\alpha\sqrt{x^2+y^2})}{x^2+y^2}, \quad (x,y) \in \mathbb{R}^2, \quad \zeta \in \{x,y\}. \tag{3.22}$$

Hence by using this result and the integral identity $\int_0^\infty (J_n^2(t)/t)dt = 1/(2n)$ ([68, pg 405]) with $n = 2$, we can show that $[\mathbf{I}(\theta)]_{11} = [\mathbf{I}(\theta)]_{22} = (2\pi n_a)^2\Lambda_0(t-t_0)/\lambda^2$ (see also [51]).

2. We can easily verify that the 2D Gaussian profile is a symmetric image function. Further, we can show that for the present situation the Fisher information matrix is diagonal and is given by eq. 3.17 (see proof of result 1). Substituting for Λ_θ^1 and q in eq. 3.17 we get $[\mathbf{I}(\theta)]_{33}$ $= (t-t_0)/\Lambda_0$, and it can be shown that $[\mathbf{I}(\theta)]_{11} = [\mathbf{I}(\theta)]_{22} = \Lambda_0(t-t_0)/\sigma^2$ (see [51]).

In both cases, the fundamental limit of the localization accuracy of x_0 (y_0) and the fundamental limit of the accuracy of Λ_0 are obtained by inverting the Fisher information matrix and taking the square root of the corresponding leading diagonal elements. •

The Airy profile depends on the term α that is given by $\alpha = 2\pi n_a/\lambda$, where n_a denotes the numerical aperture of the objective lens and λ denotes the wavelength of the detected photons. For a given experimental configuration, the numerical values of n_a and λ are known and hence α is known. On the other hand, the 2D Gaussian profile depends on the term σ that needs to be empirically determined from calibration experiments (see [28, 26]).

We use the term *fundamental* here to describe the fact that the model which underlies this expression does not take into account any deteriorating effects in the acquisition system such as pixelation of the detector and the various noise sources that typically occur in experimental settings. This expression only takes into consideration the basic optical and stochastic phenomena that are inherent in any single molecule experiment.

The result given in eq. 3.18 provides a simple analytical expression that quantitatively exhibits the dependence of the limit of the localization accuracy on the optical properties of the microscope and the photophysical properties of the single molecule. The fundamental limit exhibits an inverse square root dependence on the expected number of detected photons, which is in agreement with previously published results, see e.g. [27, 28]. The result, in particular, implies that to improve the limit of the localization accuracy by a factor of two (i.e halve the value of $\delta_{x_0}^{2d}$), we either need to double the numerical aperture of the objective lens, or increase the photon detection rate by a factor of four, or halve the emission wavelength of the single molecule. This means that the location of a single molecule emitting blue light can be more accurately determined than one that is emitting red light, provided all other factors remain the same. It should be noted that the fundamental limit is independent of the magnification M of the optical system.

Eq. 3.18 provides a limit for the smallest possible value of the standard deviation of a reasonable estimator of the location of a single molecule. It is therefore important to know whether an estimator exists whose standard deviation comes close to this limit. It is well known from large sample statistics that the variance of a large class of estimators asymptotically approaches the inverse of the Fisher information matrix ([54, 55]). We therefore consider one

such estimator, namely the maximum likelihood estimator (see Section 3.5 for details). Fig. 3.1 shows the standard deviations of the maximum likelihood estimates of the single molecule location for two different acquisition methods, one when the acquisition time is fixed and the other when the total number of detected photons is fixed. In both cases the standard deviation of the maximum likelihood estimates approaches the fundamental limit as the expected (total) number of detected photons increases.

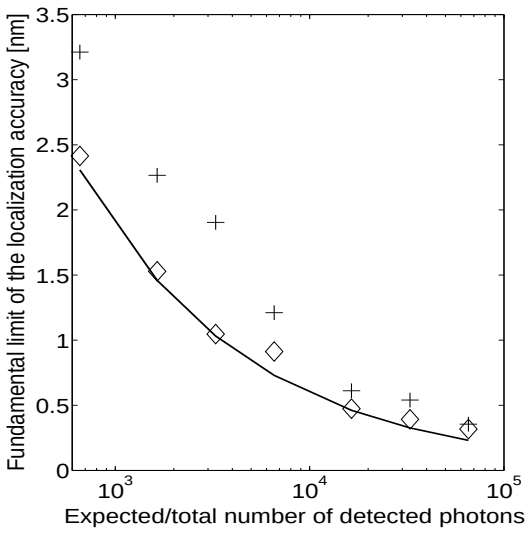

Figure 3.1. Fundamental limit to the localization accuracy of x_0 for a GFP single molecule. The figure shows the behavior of the fundamental limit to the localization accuracy of the x_0 coordinate of a single molecule with experimental parameters similar to those for a GFP molecule as a function of the expected number of detected photons $\Lambda_0(t-t_0)$. The x axis range corresponds to an acquisition time range of $t = 0.01$ s to $t = 1$ s (with $t_0 = 0$) when the photon detection rate is $\Lambda_0 = 66000$ photons/s. The figure also shows the standard deviation of the maximum likelihood estimates of the single molecule position as a function of the expected number $\Lambda_0(t-t_0)$ of detected photons (+) and as a function of the total number of detected photons ($\diamond$). The standard deviations of both the estimates approach the fundamental limit as the expected (total) number of detected photons increases. Note that the standard deviation for the latter case is uniformly closer to the fundamental limit than for the former case.

It can be shown that the maximum likelihood estimator of the photon detection rate Λ_0 is given by $\hat{\Lambda}_0 := N_{tot}/(t-t_0)$, where N_{tot} denotes the total number of detected photons and $t-t_0$ denotes the acquisition time (e.g., see [60, pg 74-75]), and the standard deviation of $\hat{\Lambda}_0$

is given by $\sqrt{\Lambda_0/(t-t_0)}$. From the above results, we see that for both the Airy profile and the 2D Gaussian profile, the performance limit to determining the parameter Λ_0 is given by $\sqrt{\Lambda_0/(t-t_0)}$, which is equal to the standard deviation of $\hat{\Lambda}_0$. Thus for the above scenario, the maximum likelihood estimator of the photon detection rate is an efficient estimator.

3.3.2 Upper and lower bounds to the performance limits

In Section 3.3.1 the integral expressions of the Fisher information matrix $\mathbf{I}(\theta)$ for an infinite detector reduced to simple formulae. However, in a practical situation the calculation of $\mathbf{I}(\theta)$ can become cumbersome, for example, due to the shape of the finite sized detector $\mathcal{C}$. Hence determining the limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$ for the components of θ can become difficult. We next address this concern by deriving integral expressions for matrices $\mathbf{I}_u(\theta)$ and $\mathbf{I}_l(\theta)$ that provide an upper and lower bound to the Fisher information matrix $\mathbf{I}(\theta)$, respectively, i.e., $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta)$. Note that if $\mathbf{I}_l(\theta)$ is invertible, then it can be shown that $\sqrt{[\mathbf{I}_l^{-1}(\theta)]_{ii}}$ and $\sqrt{[\mathbf{I}_u^{-1}(\theta)]_{ii}}$ provide an upper and lower bound to $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$, respectively, i.e., $\sqrt{[\mathbf{I}_u^{-1}(\theta)]_{ii}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}} \leq \sqrt{[\mathbf{I}_l^{-1}(\theta)]_{ii}}$, $i = 1, \ldots, n$. This is of particular relevance since in a number of situations the upper and lower bounds are diagonal matrices whose diagonal entries can be analytically calculated. We will show that this is the case for the 2D Gaussian profile and the Airy profile if the 'bounding detectors' are circular with center at the center of the image profile. The integral expression for $\mathbf{I}_u(\theta)$ ($\mathbf{I}_l(\theta)$) is derived in such a way that its integrand is identical to that of $\mathbf{I}(\theta)$ and its spatial integral is evaluated over a circular region $\mathcal{B}_{r_c}(u)$ ($\mathcal{B}_{r_c}(l)$) known as the upper (lower) circular bounding detector that is centered at a point $r_c \in \mathcal{C}$ with radius u (l). The circular bounding detectors are defined below.

Definition 3.3.2 *Let $\mathcal{C}$ be a detector and $r_c = (r_{c,x}, r_{c,y}) \in \mathcal{C}$. For $\rho > 0$, let $\mathcal{B}_{r_c}(\rho) := \{(x,y) \mid (x-r_{c,x})^2 + (y-r_{c,y})^2 < \rho^2, (x,y) \in \mathbb{R}^2\}$ denote a circular region centered at r_c with radius ρ and let $\mathcal{B}_{r_c}(\infty) := \mathbb{R}^2$. We define $\mathcal{B}_{r_c}(u)$ and $\mathcal{B}_{r_c}(l)$ to be the* **upper and lower circular bounding detectors of** $\mathcal{C}$*, respectively, if $l = sup\{\rho \mid \mathcal{B}_{r_c}(\rho) \subseteq \mathcal{C}\}$ and $u = inf\{\rho \mid \mathcal{C} \subseteq \mathcal{B}_{r_c}(\rho)\}$.*

In Section 2.3, we discussed the relationship between an image detection process described on a large, possibly infinite, detector and a more realistic smaller detector (see Proposition 2.3.1). In the derivation of the upper and lower bounds for the Fisher information matrix it will be useful to have the notion of an 'extended' version of an image detection process. We refer to any image detection process $\mathcal{G}^e$ as an extended version of $\mathcal{G}$ if $\mathcal{G}$ is the reduced version of $\mathcal{G}^e$.

Theorem 3.3.1 *Let $\Theta \subseteq \mathbb{R}^n$ be a parameter space and let $\mathcal{G}$ be an image detection process that is defined over the detector $\mathcal{C}$. Assume that $\mathcal{G}^e(\Lambda^e_\theta, \{f^e_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C}^e)$ is an extended version of $\mathcal{G}$. For $r_c \in \mathcal{C}$, let $\mathcal{B}_{r_c}(l)$ and $\mathcal{B}_{r_c}(u)$ denote the circular bounding detectors of $\mathcal{C}$. Let $\mathbf{I}(\theta)$ be the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$. If $\mathcal{C} \subseteq \mathcal{B}_{r_c}(u) \subseteq \mathcal{C}^e$, then*

$$1. \qquad \mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta), \quad \theta \in \Theta, \tag{3.23}$$

where

$$\mathbf{I}_\beta(\theta) := \int_{t_0}^{t} \int_{\mathcal{B}_{r_c}(\beta)} \frac{1}{\Lambda^e_\theta(\tau) f^e_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda^e_\theta(\tau) f^e_{\theta,\tau}(r)]}{\partial \theta} \right)^T \left(\frac{\partial [\Lambda^e_\theta(\tau) f^e_{\theta,\tau}(r)]}{\partial \theta} \right) dr d\tau, \quad \beta \in \{u, l\}. \tag{3.24}$$

2. Further, if $\mathbf{I}_l(\theta)$ is invertible, then

$$[\mathbf{I}_u^{-1}(\theta)]_{ii} \leq [\mathbf{I}^{-1}(\theta)]_{ii} \leq [\mathbf{I}_l^{-1}(\theta)]_{ii}, \quad i = 1, \ldots, n, \quad \theta \in \Theta. \tag{3.25}$$

Proof: 1. Since $\mathcal{G}^e$ is an extended extension of $\mathcal{G}$, it follows that $\mathcal{G}$ is the reduced version of $\mathcal{G}^e$ and the expression for the Fisher information matrix $\mathbf{I}(\theta)$ of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by result 1 of Theorem 2.3.1. Note that $\mathcal{B}_{r_c}(\beta)$ is open and $\mathcal{B}_{r_c}(\beta) \subseteq \mathcal{C}^e$, $\beta \in \{u, l\}$. Hence from Proposition 2.3.1 it follows that the detected photons on the circular bounding detector $\mathcal{B}_{r_c}(u)$ ($\mathcal{B}_{r_c}(l)$) can be modeled as an image detection process $\mathcal{G}^{rd}_u$ ($\mathcal{G}^{rd}_l$), which is the reduced version of $\mathcal{G}^e$ corresponding to $\mathcal{B}_{r_c}(u)$ ($\mathcal{B}_{r_c}(l)$). If the Fisher information matrix of $\mathcal{G}^{rd}_u$ ($\mathcal{G}^{rd}_l$) corresponding to the time interval $[t_0, t]$ is denoted as $\mathbf{I}_u(\theta)$ ($\mathbf{I}_l(\theta)$), then from result 1 of Theorem 2.3.1 we obtain the desired integral expression that is given by eq.

3.24. Further, since $\mathcal{B}_{r_c}(l) \subseteq \mathcal{C} \subseteq \mathcal{B}_{r_c}(u)$, from result 3 of Theorem 2.3.1 it can be deduced that $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta)$, $\theta \in \Theta$.

2. If $\mathbf{I}_l(\theta)$ is invertible for $\theta \in \Theta$, then by result 1 of this Theorem $\mathbf{I}(\theta)$ and $\mathbf{I}_u(\theta)$ are also invertible. It then follows that $\mathbf{I}_u^{-1}(\theta) \leq \mathbf{I}^{-1}(\theta) \leq \mathbf{I}_l^{-1}(\theta)$ ([69, pg 169]) and from this the result follows. •

The above result provides a general formula to calculate the upper and lower bound to the (inverse) Fisher information matrix $\mathbf{I}(\theta)$. We next consider a special case in which the extended version $\mathcal{G}^e$ is defined over $\mathbb{R}^2$. Here, the parameter vector is given by $\theta = (x_0, y_0, \Lambda_0) \in \Theta$, the photon distribution profile $f^e_{\theta,\tau}$ of $\mathcal{G}^e$ is assumed to be given in terms of a symmetric image function q and the photon detection rate Λ^e_θ of $\mathcal{G}^e$ is independent of x_0 and y_0 for $\theta \in \Theta$. We assume that the circular bounding detectors of $\mathcal{C}$ are centered at the point $r_c := (Mx_0, My_0) \in \mathcal{C}$, where $M > 0$ denotes the lateral magnification. For this case, we show that the matrices $\mathbf{I}_l(\theta)$ and $\mathbf{I}_u(\theta)$ are diagonal.

Corollary 3.3.3 *Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space and let $\mathcal{G}$ be an image detection process that is defined over the detector $\mathcal{C}$. Let $M > 0$. Assume that $\mathcal{G}^e(\Lambda^e_\theta, \{f^e_{\theta,\tau}\}_{\tau \geq t_0}, \mathbb{R}^2)$ is an extended version of $\mathcal{G}$ and that there exist a symmetric image function q such that $f^e_{\theta,\tau}(r) = \frac{1}{M^2} q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)$ for $r = (x, y) \in \mathbb{R}^2$, $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$. Assume that $\partial\Lambda^e_\theta(\tau)/\partial x_0 = \partial\Lambda^e_\theta(\tau)/\partial y_0 = 0$ for $\tau \geq t_0$ and $\theta \in \Theta$. Let $\theta \in \Theta$, assume that $r_c := (Mx_0, My_0) \in \mathcal{C}$, and let $\mathcal{B}_{r_c}(l)$ and $\mathcal{B}_{r_c}(u)$ denote the circular bounding detectors of $\mathcal{C}$. Then*

$$1. \qquad \mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta). \tag{3.26}$$

where

$$\mathbf{I}(\theta) = \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda^e_\theta(\tau)}{f^e_{\theta,\tau}(r)} \begin{bmatrix} M\frac{\partial f^e_{\theta,\tau}(r)}{\partial x} \\ M\frac{\partial f^e_{\theta,\tau}(r)}{\partial y} \\ -\frac{f^e_{\theta,\tau}(r)}{\Lambda_\theta(\tau)}\frac{\partial\Lambda^e_\theta(\tau)}{\partial\Lambda_0} \end{bmatrix} \begin{bmatrix} M\frac{\partial f^e_{\theta,\tau}(r)}{\partial x} & M\frac{\partial f^e_{\theta,\tau}(r)}{\partial y} & -\frac{f^e_{\theta,\tau}(r)}{\Lambda_\theta(\tau)}\frac{\partial\Lambda^e_\theta(\tau)}{\partial\Lambda_0} \end{bmatrix} dr d\tau, \tag{3.27}$$

$$\mathbf{I}_\beta(\theta) = Diag\left(\int_{t_0}^{t} \Lambda^e_\theta(\tau) d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial x}\right)^2 dxdy,\right.$$

$$\int_{t_0}^{t} \Lambda_\theta^e(\tau) d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x,y)} \left(\frac{\partial q(x,y)}{\partial y}\right)^2 dxdy,$$
$$\int_{t_0}^{t} \frac{1}{\Lambda_\theta^e(\tau)} \left(\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}\right)^2 d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} q(x,y) dxdy \Bigg), \tag{3.28}$$

with $\mathcal{B}_0(\beta/M) = \{(x,y) \mid \sqrt{x^2+y^2} < \frac{\beta}{M}\}$ *and* $\beta \in \{u,l\}$.

2. Further, if $\mathbf{I}_l(\theta)$ *is invertible, then*

$$[\mathbf{I}_u^{-1}(\theta)]_{ii} \leq [\mathbf{I}^{-1}(\theta)]_{ii} \leq [\mathbf{I}_l^{-1}(\theta)]_{ii}, \quad i = 1,2,3. \tag{3.29}$$

Proof: Let $\theta \in \Theta$. Eqs. 3.26 and 3.29 immediately follow by noting that $\mathcal{B}_{r_c}(l) \subseteq \mathcal{C} \subseteq \mathcal{B}_{r_c}(u) \subseteq \mathbb{R}^2$ and that the results of Theorem 3.3.1 hold for the present case. In rest of this proof we derive the integral expressions for $\mathbf{I}(\theta)$ and $\mathbf{I}_\beta(\theta)$ that are given in eqs. 3.27 and 3.28, respectively.

By assumption, $\partial\Lambda_\theta^e(\tau)/\partial x_0 = \partial\Lambda_\theta^e(\tau)/\partial y_0 = \partial f_{\theta,\tau}^e(r)/\partial\Lambda_0 = 0$ and it can be shown that $\partial f_{\theta,\tau}^e(r)/\partial x_0 = -M(\partial f_{\theta,\tau}^e(r)/\partial x)$ and $\partial f_{\theta,\tau}^e(r)/\partial y_0 = -M(\partial f_{\theta,\tau}^e(r)/\partial y)$, for $r = (x,y) \in \mathbb{R}^2$ and $\tau \geq t_0$. Using these results and substituting for Λ_θ^e and $f_{\theta,\tau}^e$ in result 1 of Theorem 2.3.1, we obtain the expression for $\mathbf{I}(\theta)$ that is given in eq. 3.27. By definition, for $\beta \in \{u,l\}$, $\mathcal{B}_{r_c}(\beta) = \{(x,y) \mid \sqrt{(x/M - x_0)^2 + (y/M - y_0)^2} < \beta/M\}, (x,y) \in \mathbb{R}^2\}$. Substituting for $f_{\theta,\tau}^e$ and Λ_θ^e in eq. 3.24, we have for $\beta \in \{u,l\}$,

$$\mathbf{I}_\beta(\theta) = \int_{t_0}^{t} \int_{\mathcal{B}_{r_c}(\beta)} \frac{1}{\Lambda_\theta^e(\tau) f_{\theta,\tau}^e(r)} \begin{bmatrix} \Lambda_\theta^e(\tau) \frac{\partial f_{\theta,\tau}^e(r)}{\partial x_0} \\ \Lambda_\theta^e(\tau) \frac{\partial f_{\theta,\tau}^e(r)}{\partial y_0} \\ f_{\theta,\tau}^e(r) \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} \times$$
$$\begin{bmatrix} \Lambda_\theta^e(\tau) \frac{\partial f_{\theta,\tau}^e(r)}{\partial x_0} & \Lambda_\theta^e(\tau) \frac{\partial f_{\theta,\tau}^e(r)}{\partial y_0} & f_{\theta,\tau}^e(r) \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} dr d\tau$$
$$= \frac{1}{M^2} \int_{t_0}^{t} \Lambda_\theta^e(\tau) \int_{\mathcal{B}_{r_c}(\beta)} \frac{1}{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)} \begin{bmatrix} -M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x} \\ -M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial y} \\ \frac{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} \times$$
$$\begin{bmatrix} -M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x} \\ -M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial y} \\ \frac{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix}^T dxdyd\tau$$

$$= \frac{1}{M^2}\int_{t_0}^{t}\Lambda_\theta^e(\tau)\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(u,v)}\begin{bmatrix} M\frac{\partial q(u,v)}{\partial u}\frac{\partial u}{\partial x} \\ M\frac{\partial q(u,v)}{\partial v}\frac{\partial v}{\partial y} \\ -\frac{q(u,v)}{\Lambda_\theta^e(\tau)}\frac{\partial\Lambda_\theta^e(\tau)}{\partial\Lambda_0}\end{bmatrix}\times$$

$$\begin{bmatrix} M\frac{\partial q(u,v)}{\partial u}\frac{\partial u}{\partial x} \\ M\frac{\partial q(u,v)}{\partial v}\frac{\partial v}{\partial y} \\ -\frac{q(u,v)}{\Lambda_\theta^e(\tau)}\frac{\partial\Lambda_\theta^e(\tau)}{\partial\Lambda_0}\end{bmatrix}^T (Mdu)(Mdv)d\tau$$

$$= \int_{t_0}^{t}\Lambda_\theta^e(\tau)\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(x,y)}\begin{bmatrix} \frac{\partial q(x,y)}{\partial x} \\ \frac{\partial q(x,y)}{\partial y} \\ -\frac{q(x,y)}{\Lambda_\theta^e(\tau)}\frac{\partial\Lambda_\theta^e(\tau)}{\partial\Lambda_0}\end{bmatrix}\begin{bmatrix} \frac{\partial q(x,y)}{\partial x} & \frac{\partial q(x,y)}{\partial y} & -\frac{q(x,y)}{\Lambda_\theta^e(\tau)}\frac{\partial\Lambda_\theta^e(\tau)}{\partial\Lambda_0}\end{bmatrix} dxdyd\tau, \tag{3.30}$$

where in the third step $u := \frac{x}{M} - x_0$ and $v := \frac{y}{M} - y_0$, for $(x,y) \in \mathbb{R}^2$, $\theta \in \Theta$. Since the image function q is symmetric, it can be shown that $\partial q(x,y)/\partial x = -\partial q(-x,y)/\partial x$ and $\partial q(x,y)/\partial y = \partial q(-x,y)/\partial y$ for $(x,y) \in \mathbb{R}^2$ (see eqs. 3.10 - 3.11). Thus we have

$$[\mathbf{I}_\beta(\theta)]_{12} = [\mathbf{I}_\beta(\theta)]_{21} = \int_{t_0}^{t}\Lambda_\theta^e(\tau)d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(x,y)}\frac{\partial q(x,y)}{\partial x}\frac{\partial q(x,y)}{\partial y}dxdy$$

$$= -\int_{t_0}^{t}\Lambda_\theta^e(\tau)d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(-x,y)}\frac{\partial q(-x,y)}{\partial x}\frac{\partial q(-x,y)}{\partial y}dxdy$$

$$= -\int_{t_0}^{t}\Lambda_\theta^e(\tau)d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(u,y)}\frac{\partial q(u,y)}{\partial u}\frac{\partial u}{\partial x}\frac{\partial q(u,y)}{\partial y}(-du)dy \qquad (u := -x)$$

$$= -\int_{t_0}^{t}\Lambda_\theta^e(\tau)d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{q(u,y)}\frac{\partial q(u,y)}{\partial u}\frac{\partial q(u,y)}{\partial y}dudy = -[\mathbf{I}_l(\theta)]_{12} = -[\mathbf{I}_l(\theta)]_{21}, \quad \beta \in \{u,l\}.$$

Hence $[\mathbf{I}_\beta(\theta)]_{12} = [\mathbf{I}_\beta(\theta)]_{21} = 0$, $\beta \in \{u,l\}$. Similarly, we can show that $[\mathbf{I}_\beta(\theta)]_{13} = [\mathbf{I}_\beta(\theta)]_{31} = 0$, $\beta \in \{u,l\}$. Further, by using the symmetry property of q we can also show that $\partial q(x,y)/\partial y = -\partial q(-x,y)/\partial y$ for $(x,y) \in \mathbb{R}^2$ (see eq. 3.14). From this it follows that $[\mathbf{I}_\beta(\theta)]_{23} = [\mathbf{I}_\beta(\theta)]_{32} = 0$, $\beta \in \{u,l\}$. Substituting these in eq. 3.30 the result follows. •

For the localization accuracy problem the upper and lower bounds of the limit of the localization accuracy of (x_0, y_0) are referred to as the **localization accuracy bounds**.

3.3.3 Examples

We now illustrate the results derived in this section by considering specific image functions that describe the image of a point source. Here, the parameter vector is set to be $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and the photon distribution profile $f^e_{\theta,\tau}$ is assumed to be given in terms of a symmetric image function q. Further, the photon detection rate is assumed to be a constant, i.e., $\Lambda^e_\theta(\tau) = \Lambda_0$, $\tau \geq t_0$. For each image function, we derive the expression for the Fisher information matrix $\mathbf{I}(\theta)$ corresponding to the detector $\mathcal{C}$ and also derive the upper and lower bound for $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$, $i = 1, 2, 3$, which denotes the limit of the accuracy of the components of θ.

Corollary 3.3.4 *Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space and let $\mathcal{G}$ be an image detection process that is defined over the detector $\mathcal{C}$. Let $M > 0$. Assume that $\mathcal{G}^e(\Lambda^e_\theta, \{f^e_{\theta,\tau}\}_{\tau \geq t_0}, \mathbb{R}^2)$ is an extended version of $\mathcal{G}$ and that there exists a symmetric image function q such that $f^e_{\theta,\tau}(r) = \frac{1}{M^2} q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)$ for $r = (x, y) \in \mathbb{R}^2$, $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$. Assume that $\Lambda^e_\theta(\tau) = \Lambda_0$ for $\tau \geq t_0$ and $\theta \in \Theta$. Let $\theta \in \Theta$, assume that $r_c := (Mx_0, My_0) \in \mathcal{C}$, and let $\mathcal{B}_{r_c}(l)$ and $\mathcal{B}_{r_c}(u)$ denote the circular bounding detectors of $\mathcal{C}$.*

1. **Airy profile:** *If q is given by eq. 3.18, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by*

$$\mathbf{I}(\theta) = \Lambda_0(t - t_0) \int_{\mathcal{C}} \frac{1}{\frac{J_1^2(a||r-r_c||)}{\pi||r-r_c||^2}} \mathcal{Q}^T_\theta(r) \mathcal{Q}_\theta(r) dr,$$

where $a = 2\pi n_a/(\lambda M)$, $||r - r_c|| := \sqrt{(x - Mx_0)^2 + (y - My_0)^2}$ and

$$\mathcal{Q}^T_\theta(r) := \frac{J_1(a||r-r_c||)}{\pi||r-r_c||} \begin{bmatrix} \frac{2aM(x-Mx_0)}{||r-r_c||^2} J_2(a||r-r_c||) \\ \frac{2aM(y-My_0)}{||r-r_c||^2} J_2(a||r-r_c||) \\ -\frac{1}{\Lambda_0} \frac{J_1(a||r-r_c||)}{||r-r_c||} \end{bmatrix}.$$

Moreover, if u and l are as defined above, then for $i = 1, 2$,

$$\frac{\lambda/\left(2\pi n_a \sqrt{\Lambda_0(t-t_0)}\right)}{\sqrt{1 - \left(J_0^2(au) + 2J_1^2(au) + J_2^2(au)\right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}} \leq \frac{\lambda/\left(2\pi n_a \sqrt{\Lambda_0(t-t_0)}\right)}{\sqrt{1 - \left(J_0^2(al) + 2J_1^2(al) + J_2^2(al)\right)}}, \tag{3.31}$$

$$\frac{\sqrt{\Lambda_0/(t-t_0)}}{\sqrt{1-\left(J_0^2(au)+J_1^2(au)\right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{33}} \leq \frac{\sqrt{\Lambda_0/(t-t_0)}}{\sqrt{1-\left(J_0^2(al)+J_1^2(al)\right)}}, \tag{3.32}$$

where J_n denotes the n^{th} order Bessel function of the first kind, $n = 0, 1, 2$.

2. **2D Gaussian profile:** *If q is given by eq. 3.20, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by*

$$\mathbf{I}(\theta) = \Lambda_0(t-t_0)\int_{\mathcal{C}} \frac{2\pi(M\sigma)^2}{e^{-\frac{||r-r_c||^2}{2(M\sigma)^2}}} \mathcal{Q}_\theta^T(r)\mathcal{Q}_\theta(r)dr,$$

where $\sigma > 0$, $||r - r_c|| := \sqrt{(x - Mx_0)^2 + (y - My_0)^2}$ and

$$\mathcal{Q}_\theta(r) := \frac{1}{2\pi(M\sigma)^2} e^{-\frac{||r-r_c||^2}{2(M\sigma)^2}} \begin{bmatrix} \frac{(x-Mx_0)}{M\sigma^2} & \frac{(y-My_0)}{M\sigma^2} & -\frac{1}{\Lambda_0} \end{bmatrix}.$$

If u and l are as defined above, then for $i = 1, 2$,

$$\frac{\sigma/\sqrt{\Lambda_0(t-t_0)}}{\sqrt{1-e^{-\frac{1}{2}\left(\frac{u}{M\sigma}\right)^2}\left(1+\frac{1}{2}\left(\frac{u}{M\sigma}\right)^2\right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}} \leq \frac{\sigma/\sqrt{\Lambda_0(t-t_0)}}{\sqrt{1-e^{-\frac{1}{2}\left(\frac{l}{M\sigma}\right)^2}\left(1+\frac{1}{2}\left(\frac{l}{M\sigma}\right)^2\right)}}, \tag{3.33}$$

$$\frac{\sqrt{\Lambda_0/(t-t_0)}}{\sqrt{1-e^{-\frac{1}{2}\left(\frac{u}{M\sigma}\right)^2}}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{33}} \leq \frac{\sqrt{\Lambda_0/(t-t_0)}}{\sqrt{1-e^{-\frac{1}{2}\left(\frac{l}{M\sigma}\right)^2}}}. \tag{3.34}$$

Proof: We can show that the Airy profile and the 2D Gaussian profile satisfy the properties of a symmetric image function. Let $\theta \in \Theta$. It can be verified that for the present case the results of Corollary 3.3.3 hold. Then the expressions for the Fisher information matrix $\mathbf{I}(\theta)$ immediately follow by substituting the corresponding image function in eq. 3.27. Further, we can also verify that $\mathbf{I}^{-1}(\theta)$ exists for each image function. In rest of this proof we derive the expressions for the upper and lower bounds of $\mathbf{I}(\theta)$.

1. Let $\alpha := 2\pi n_a/\lambda$. Substituting for Λ_θ^e and q in the integral expression of $[\mathbf{I}_\beta(\theta)]_{11}$ ($[\mathbf{I}_\beta(\theta)]_{22}$) that is given in eq. 3.28, we have for $\beta \in \{u, l\}$,

$$[\mathbf{I}_\beta(\theta)]_{11} = [\mathbf{I}_\beta(\theta)]_{22} = \int_{t_0}^{t} \Lambda_0 d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{\frac{J_1^2(\alpha\sqrt{x^2+y^2})}{\pi(x^2+y^2)}} \left[\frac{\partial}{\partial x}\left(\frac{J_1^2(\alpha\sqrt{x^2+y^2})}{\pi(x^2+y^2)}\right)\right]^2 dxdy$$

$$= \frac{4\alpha^2}{\pi}\Lambda_0(t-t_0)\int_{\{(x,y)|\sqrt{x^2+y^2}<\beta/M\}} \frac{J_2^2(\alpha\sqrt{x^2+y^2})}{(x^2+y^2)^2} dxdy$$

$$= \frac{4\alpha^2}{\pi}\Lambda_0(t-t_0)\int_0^{2\pi}\cos^2\phi d\phi\int_0^{\frac{\beta}{M}}\frac{J_2^2(\alpha\rho)}{\rho}d\rho$$

$$= 4\alpha^2\Lambda_0(t-t_0)\int_0^{a\beta}\frac{J_2^2(w)}{w}dw$$

$$= 4\alpha^2\Lambda_0(t-t_0)\left(\int_0^{\infty}\frac{J_2^2(w)}{w}dw - \int_{a\beta}^{\infty}\frac{J_2^2(w)}{w}dw\right)$$

$$= \frac{1-\left(J_0^2(a\beta)+2J_1^2(a\beta)+J_2^2(a\beta)\right)}{\lambda^2/((2\pi n_a)^2\Lambda_0(t-t_0))},$$

where $x=\rho\cos\phi$, $y=\rho\sin\phi$, $a=\alpha/M$, the partial derivative of q with respect to x is given in eq. 3.22, and the integral expressions in the final step are evaluated by using the integral identities $\int_0^{\infty}(J_n^2(t)/t)dt = (1/2n)$ ([68, pg 405]) and $n\int_x^{\infty}(J_n^2(t)/t)dt = (1/2)J_0^2(x)+J_1^2(x)+\cdots+J_{n-1}^2(x)+(1/2)J_n^2(x)$ ([70, pg 95]) with $n=2$ and $x\in\mathbb{R}$. Similarly, for $\beta\in\{u,l\}$, we have

$$[\mathbf{I}_\beta(\theta)]_{33} = \frac{1}{\pi}\int_{t_0}^{t}\frac{1}{\Lambda_0}d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{J_1^2(\alpha\sqrt{x^2+y^2})}{x^2+y^2}dxdy = \frac{(t-t_0)}{\pi\Lambda_0}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M}}\frac{J_1^2(\alpha\rho)}{\rho}d\rho$$

$$= \frac{2(t-t_0)}{\Lambda_0}\int_0^{a\beta}\frac{J_1^2(w)}{w}dw = \frac{1-(J_0^2(a\beta)+J_1^2(a\beta))}{\Lambda_0/(t-t_0)}.$$

Using result 2 of Corollary 3.3.3 the result follows.

2. Substituting for Λ_θ^e and q in the integral expression of $[\mathbf{I}_\beta(\theta)]_{11}$ ($[\mathbf{I}_\beta(\theta)]_{22}$) that is given in eq. 3.28, we have for $\beta\in\{u,l\}$,

$$[\mathbf{I}_\beta(\theta)]_{11} = [\mathbf{I}_\beta(\theta)]_{22} = \int_{t_0}^{t}\Lambda_0 d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{1}{\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}}\left(\frac{\partial}{\partial x}\left(\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\right)^2 dxdy$$

$$= \frac{\Lambda_0(t-t_0)}{2\pi\sigma^2}\int_{\mathcal{B}_0(\frac{\beta}{M})}\frac{x^2}{\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\frac{dxdy}{\sigma^2} = \frac{\Lambda_0(t-t_0)}{2\pi\sigma^2}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M\sigma}}\rho^3 e^{-\frac{\rho^2}{2}}d\rho$$

$$= \frac{\Lambda_0(t-t_0)}{4\sigma^2}\int_0^{\frac{\beta}{M\sigma}}\rho^2 e^{-\frac{\rho^2}{2}}2\rho d\rho = \frac{\Lambda_0(t-t_0)}{4\sigma^2}\int_0^{\left(\frac{\beta}{M\sigma}\right)^2}we^{-\frac{w}{2}}dw$$

$$= \frac{1-e^{-\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2}\left(1+\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2\right)}{\sigma^2/(\Lambda_0(t-t_0))},$$

$$[\mathbf{I}_\beta(\theta)]_{33} = \frac{1}{2\pi\sigma^2}\int_{t_0}^{t}\frac{1}{\Lambda_0}d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}e^{-\frac{x^2+y^2}{2\sigma^2}}dxdy = \frac{(t-t_0)}{\Lambda_0 2\pi}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M\sigma}}\rho e^{-\frac{\rho^2}{2}}d\rho$$

$$= \frac{(t-t_0)}{2\Lambda_0}\int_0^{\frac{\beta}{M\sigma}} 2\rho e^{-\frac{\rho^2}{2}}\,d\rho = \frac{(t-t_0)}{2\Lambda_0}\int_0^{\left(\frac{\beta}{M\sigma}\right)^2} e^{-\frac{w}{2}}\,dw = \frac{1-e^{-\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2}}{\Lambda_0/(t-t_0)}.$$

Using result 2 of Corollary 3.3.3 we obtain the desired result. •

From the above Corollary we see that for both image functions, the localization accuracy bounds for x_0 (y_0) and the bounds for the limit of the accuracy of Λ_0 reduce to simple formulae. Note that the above results for the upper and lower bounds hold only if the point (Mx_0, My_0) is located on the detector $\mathcal{C}^{rd}$. In most experimental situations this condition is satisfied.

We now discuss the results derived in Corollary 3.3.4 by considering a finite-sized square detector. Fig. 3.2A (Fig. 3.2B) shows the behavior of the limit of the localization accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{11}}$ for x_0 for a square detector corresponding to the Airy profile (2D Gaussian profile) as a function of detector size. A reduced detector detects relatively lower number of photons than an infinite detector. Hence the limit of the localization accuracy for the reduced detector is greater (worse) than the fundamental limit of the localization accuracy that is calculated for the infinite detector (see Corollary 3.3.2). As the detector size increases, more photons are detected by the reduced detector and the limit of the localization accuracy approaches the fundamental limit. Fig. 3.2A (Fig. 3.2B) also shows the behavior of the localization accuracy bounds given by eq. 3.31 (eq. 3.33) for a square detector corresponding to the Airy profile. Here, we see that the localization accuracy bounds provide a tight bound, as they are consistently close to the limit of the localization accuracy for the square detector.

Note that the behavior of the limit of the localization accuracy for a square detector also depends on the functional form of the image function. In the case of the 2D Gaussian profile (see Fig. 3.2B), the limit of the localization accuracy for a square detector with side length 80 μm is close to the fundamental limit of the localization accuracy. However, this is not the case for the Airy profile (see Fig. 3.2A), where even for a square detector with side length 140 μm, the limit of the localization accuracy does not come close to the fundamental limit. In a practical application such as single molecule data analysis, the above observation provides guidelines for choosing the optimal size of the region of interest on the acquired image.

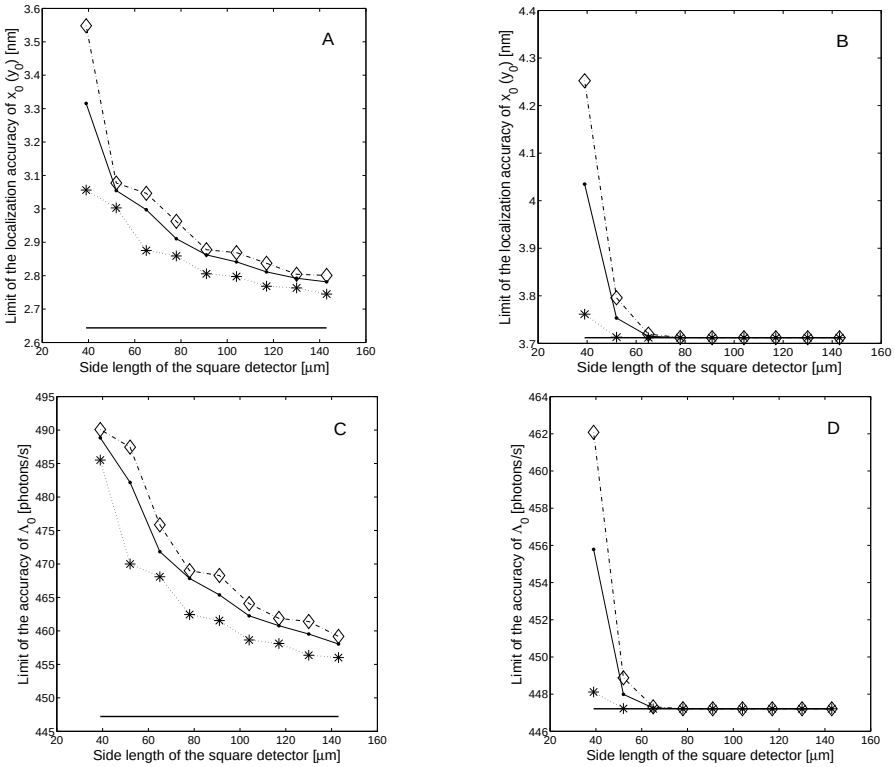

Figure 3.2. Effect of finite detector size on the limit of the 2D localization accuracy. Panel A shows the behavior of the limit of the localization accuracy of x_0 (y_0) for a square detector corresponding to the Airy profile ($\bullet$) as a function of detector size and Panel B shows the same for the 2D Gaussian profile ($\bullet$). Panel C shows the behavior of the limit of the accuracy of Λ_0 for a square detector corresponding to the Airy profile ($\bullet$) as a function of detector size and Panel D shows the same for the 2D Gaussian profile ($\bullet$). In all the panels the corresponding fundamental limit of the accuracy (—) and the upper ($\diamond$) and lower ($*$) bound to the limit of the accuracy for the square detector are shown. For a square detector with side length s, the position of the point source on the detector is set to be $r_c = (s/2, s/2)$, the radius of the lower circular bounding detector $\mathcal{B}_{r_c}(l)$ is $l = s/2$ and the radius of the upper circular bounding detector $\mathcal{B}_{r_c}(u)$ is $u = s/\sqrt{2}$. For the Airy profile, the numerical aperture is set to be $n_a = 1.4$ and the wavelength of the detected photons is set to be $\lambda = 0.52$ μm. The parameter σ corresponding to the Gaussian profile is set to be $\sigma = 0.083$ μm and is determined by fitting a 2D Gaussian profile to a Airy profile ($n_a = 1.4$, $\lambda = 0.52$ μm) through the least squares criterion. For all the plots, $\Lambda_0 = 10^4$ photons/s, the acquisition time is set to be $t = 50$ ms (with $t_0 = 0$) and the magnification is set to be $M = 100$.

Moreover, it also shows the importance of using the correct image function, as this influences the behavior of the limit of the localization accuracy.

Fig. 3.2C (Fig. 3.2D) shows the variation of the limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$ of Λ_0 for a square detector corresponding to the Airy profile (2D Gaussian profile) as a function of detector size. Analogous to the behavior of the limit of the localization accuracy, the limit of the accuracy of Λ_0 approaches the fundamental limit of the accuracy of Λ_0 ($= \sqrt{\Lambda_0/(t-t_0)}$, see eqs. 3.19 and 3.21) as the detector size increases. Moreover, the behavior of the limit of the accuracy of Λ_0 also depends on the functional form of the image function. Fig. 3.2C (Fig. 3.2D) also shows the behavior of the upper and lower bounds to the limit of the accuracy of Λ_0 that given by eq. 3.32 (eq. 3.34) for a square detector corresponding to the Airy profile (2D Gaussian profile). Similar to the localization accuracy bounds (Figs. 3.2A and 3.2B), we see that the upper and lower bounds for the limit of the accuracy of Λ_0 provide a tight bound.

3.4 Effects of pixelation and noise

In this section we derive expressions for the limit of the 2D localization accuracy when the data is acquired by a pixelated detector and is corrupted by extraneous noise sources. Here, we consider an Airy profile and assume the unknown parameter to be $\theta = (x_0, y_0)$.

Theorem 3.4.1 *Let $\{C_1, \ldots, C_{N_p}\}$ denote a pixelated detector and $\Theta \subseteq \mathbb{R}^2$ be open. Let q be given by*

$$q(x,y) := \frac{J_1^2(\alpha\sqrt{x^2+y^2})}{\pi(x^2+y^2)}, \quad (x,y) \in \mathbb{R}^2,$$

where $\alpha := 2\pi n_a/\lambda$. For $\theta \in \Theta$ and $\tau \geq t_0$, define

$$h_\theta(k) := \frac{1}{M^2}\int_{C_k} q\Big(\frac{x}{M} - x_0, \frac{y}{M} - y_0\Big)dxdy, \quad k = 1, \ldots, N_p.$$

For $\theta \in \Theta$, $t \geq t_0$ and $k = 1, \ldots, N_p$, let $\mu_\theta(k,t) := \Lambda_0(t-t_0)h_\theta(k)$ and $\beta(k,t) = b_k$, where $\Lambda_0 \geq 0$ ($b_k \geq 0$) denotes the photon detection rate of the single molecule (background component).

1. The limit of the 2D localization accuracy for a pixelated detector in the absence of extraneous noise sources is given by

$$\delta_{x_0}^{pix} = \frac{\lambda}{4\pi n_a \sqrt{\Lambda_0(t-t_0)}} \left[\sum_{k=1}^{N_p} \frac{\mathcal{J}_x^2(k)}{h_\theta(k)} - \frac{\left(\sum_{k=1}^{K} \frac{\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_\theta(k)} \right)^2}{\sum_{k=1}^{N_p} \frac{\mathcal{J}_y^2(k)}{h_\theta(k)}} \right]^{-\frac{1}{2}},$$

$$\delta_{y_0}^{pix} = \frac{\lambda}{4\pi n_a \sqrt{\Lambda_0(t-t_0)}} \left[\sum_{k=1}^{N_p} \frac{\mathcal{J}_y^2(k)}{h_\theta(k)} - \frac{\left(\sum_{k=1}^{K} \frac{\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_\theta(k)} \right)^2}{\sum_{k=1}^{N_p} \frac{\mathcal{J}_x^2(k)}{h_\theta(k)}} \right]^{-\frac{1}{2}},$$

where for $k = 1, \ldots, N_p$, $\mathcal{J}_x(k)$ and $\mathcal{J}_y(k)$ are given by

$$\mathcal{J}_x(k) := \int_{C_k} (x - Mx_0) \cdot \frac{J_1(a||r-r_0||)}{\pi||r-r_0||} \cdot \frac{J_2(a||r-r_0||)}{||r-r_0||^2} dr, \tag{3.35}$$

$$\mathcal{J}_y(k) := \int_{C_k} (y - My_0) \cdot \frac{J_1(a||r-r_0||)}{\pi||r-r_0||} \cdot \frac{J_2(a||r-r_0||)}{||r-r_0||^2} dr, \tag{3.36}$$

with $r_0 = M(x_0, y_0)$ and $a := \alpha/M$.

2. The limit of the 2D localization accuracy for a pixelated detector in the presence of additive Poisson noise is given by

$$\delta_{x_0,p}^{pix} = \frac{\lambda}{4\pi n_a \sqrt{\Lambda_0(t-t_0)}} \left[\sum_{k=1}^{N_p} \frac{\mathcal{J}_x^2(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}} - \frac{\left(\sum_{k=1}^{N_p} \frac{\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}} \right)^2}{\sum_{k=1}^{N_p} \frac{\mathcal{J}_y^2(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}}} \right]^{-\frac{1}{2}},$$

$$\delta_{y_0,p}^{pix} = \frac{\lambda}{4\pi n_a \sqrt{\Lambda_0(t-t_0)}} \left[\sum_{k=1}^{N_p} \frac{\mathcal{J}_y^2(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}} - \frac{\left(\sum_{k=1}^{N_p} \frac{\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}} \right)^2}{\sum_{k=1}^{N_p} \frac{\mathcal{J}_x^2(k)}{h_\theta(k) + \frac{b_k}{\Lambda_0}}} \right]^{-\frac{1}{2}},$$

where $\mathcal{J}_x$ and $\mathcal{J}_y$ are given in eqs. 3.35- 3.36.

3. The limit of the 2D localization accuracy for a pixelated detector in the presence of additive Poisson and Gaussian noise is given by

$$\delta_{x_0,p,g}^{pix} = \frac{\lambda}{4\pi n_a \Lambda_0(t-t_0)} \left[\sum_{k=1}^{N_p} \mathcal{J}_x^2(k)\Psi(k) - \frac{\left(\sum_{k=1}^{K} \mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k) \right)^2}{\sum_{k=1}^{N_p} \mathcal{J}_y^2(k)\Psi(k)} \right]^{-\frac{1}{2}},$$

$$\delta^{pix}_{y_0,p,g} = \frac{\lambda}{4\pi n_a \Lambda_0(t-t_0)} \left[\sum_{k=1}^{N_p} \mathcal{J}_y^2(k)\Psi(k) - \frac{\left(\sum_{k=1}^{N_p} \mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k)\right)^2}{\sum_{k=1}^{K} \mathcal{J}_x^2(k)\Psi(k)} \right]^{-\frac{1}{2}},$$

where $\mathcal{J}_x(k)$ and $\mathcal{J}_y(k)$ are given in eqs. 3.35- 3.36, $\Psi(k)$ is given by

$$\Psi(k) := \left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2} \right)^2}{p_{\theta,k}(z)} dz - 1 \right), \quad k = 1, \ldots, N_p, \tag{3.37}$$

with $\nu_\theta(k) := \Lambda_0(t-t_0)\mu_\theta(k) + \beta(k)$, $k = 1, \ldots, N_p$ and $p_{\theta,k}$ is given by eq. 2.5.

Proof: Note that result 1 is a special case of result 2 with $b_k = 0$ for $k = 1, \ldots, N_p$. So we first derive result 2. For $k = 1, \ldots, N_p$, define

$$\mathcal{J}_x(k) := \int_{C_k} (x - Mx_0) \cdot \frac{J_1(a||r-r_0||)}{\pi||r-r_0||} \cdot \frac{J_2(a||r-r_0||)}{||r-r_0||^2} dr,$$

$$\mathcal{J}_y(k) := \int_{C_k} (y - My_0) \cdot \frac{J_1(a||r-r_0||)}{\pi||r-r_0||} \cdot \frac{J_2(a||r-r_0||)}{||r-r_0||^2} dr,$$

where $r_0 := M(x_0, y_0)$. Consider the term

$$\begin{aligned}
\frac{\partial \mu_\theta(k)}{\partial x_0} &= \Lambda_0(t-t_0) \int_{C_k} \frac{\partial f_{\theta,\tau}(r)}{\partial x_0} dr = -\Lambda_0(t-t_0)M \int_{C_k} \frac{\partial f_{\theta,\tau}(r)}{\partial x} dr \\
&= -\frac{\Lambda_0 t}{M} \int_{C_k} \frac{\partial q(\frac{x}{M} - x_0, \frac{y}{M} - y_0)}{\partial x} dxdy \\
&= -\frac{\Lambda_0(t-t_0)}{M} \int_{C_k} \frac{-2a(x - Mx_0)}{\pi} \frac{J_1(a||r-r_0||)}{||r-r_0||} \frac{J_2(a||r-r_0||)}{||r-Mr_0||^2} dxdy \\
&= 2\Lambda_0(t-t_0)\alpha \mathcal{J}_x(k), \quad k = 1, \ldots, N_p.
\end{aligned}$$

Similarly we have

$$\frac{\partial \mu_\theta(k)}{\partial y_0} = 2\Lambda_0(t-t_0)\alpha \mathcal{J}_y(k), \quad k = 1 \ldots, N_p.$$

Substituting the above partial derivatives in the Fisher information matrix given by eq. 2.4 we get

$$\mathbf{I}(\theta) = \sum_{k=1}^{K} \frac{1}{\Lambda_0(t-t_0)(h_{\theta,\tau}(k) + \frac{b_k}{\Lambda_0})} \times$$

$$\begin{bmatrix} (2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x^2(k) & (2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k) \\ \\ (2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k) & (2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_y^2(k) \end{bmatrix}$$

and

$$\mathbf{I}^{-1}(\theta) = \frac{1}{\Delta(\theta)} \begin{bmatrix} \sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_y^2(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} & -\sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} \\ \\ -\sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} & \sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_x^2(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} \end{bmatrix},$$

where $\Delta(\theta)$ denotes the determinant of $\mathbf{I}(\theta)$ and is given by

$$\Delta(\theta) = \sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_x^2(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} \sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_y^2(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}} - \left[\sum_{k=1}^{K} \frac{4\Lambda_0(t-t_0)\alpha^2\mathcal{J}_x(k)\mathcal{J}_y(k)}{h_{\theta,\tau}(k)+\frac{b_k}{\Lambda_0}}\right]^2.$$

Substituting for $\alpha = 2\pi n_a/\lambda$ in $\mathbf{I}^{-1}(\theta)$ we get result 2. Setting $b_k = 0$, $k = 1, \ldots, N_p$ we get result 1.

3. Substituting for $\frac{\partial \mu_\theta(k,t)}{\partial x_0}$ and $\frac{\partial \mu_\theta(k,t)}{\partial y_0}$ in eq. 2.7 and making use of the terms $\mathcal{J}_x(k)$ and $\mathcal{J}_y(k)$ that are defined above, we have

$$\mathbf{I}(\theta) = \begin{bmatrix} \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x^2(k)\Psi(k) & \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k) \\ \\ \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k) & \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_y^2(k)\Psi(k) \end{bmatrix},$$

where $\Psi(k)$ is given by eq. 3.37. Inverting the above matrix we get

$$\mathbf{I}^{-1}(\theta) = \frac{1}{\Delta(\theta)} \begin{bmatrix} \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_y^2(k)\Psi(k) & -\sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k) \\ \\ -\sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k) & \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x^2(k)\Psi(k) \end{bmatrix},$$

where $\Delta(\theta)$ denotes the determinant of $\mathbf{I}(\theta)$ and is given by

$$\Delta(\theta) = \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x^2(k)\Psi(k) \sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_y^2(k)\Psi(k)$$

$$- \left[\sum_{k=1}^{K}(2\Lambda_0(t-t_0)\alpha)^2\mathcal{J}_x(k)\mathcal{J}_y(k)\Psi(k)\right]^2$$

Substituting for α in the above equation, we get the desired result. •

From the above theorem, we see that the expression for the limit of the localization accuracy for a pixelated detector is a modification of the fundamental limit δ_u given in eq. 3.18. In fact, the expression involves the fundamental limit $\delta_{x_0}^{2d}$ and a correction term (given in parentheses) that expresses the deterioration of the limit due to pixelation and noise.

3.4.1 Examples and applications

We next illustrate the results derived in Theorem 3.4.1 by showing how various experimental aspects such as magnification, pixel array size, pixel dimensions and noise levels influence the accuracy with which the location of a single molecule can be determined. The fundamental limit (eq. 3.18) serves as an important reference point to establish how closely the specific experimental implementation approaches the best possible localization accuracy.

For the numerical illustrations we chose parameters that are typical values for single molecule experiments with GFP molecules, see also [26] where similar values were used. In all figures, unless otherwise specified, the photon detection rate of the single molecule is assumed to be $\Lambda_0 = 66000$ photons/s. The emission wavelength is set to be $\lambda = 520$ nm corresponding to a GFP molecule, the numerical aperture is set to be $n_a = 1.4$ and the magnification is set to be $M = 100$. The single molecule is assumed to be located at the center of the pixel array. We also assume that the detector consists of square pixels with no dead space region between adjacent pixels.

We first consider the effect of pixelation on the localization accuracy in the absence of noise. Figs. 3.3a and 3.3c illustrate the limit of the localization accuracy for a 11×11 pixel array as a function of different magnification values for $t = 0.01$ s and $t = 1$ s respectively. For very low magnification values the image of the single molecule is to a large extent concentrated on one pixel. Therefore there is little information in the data about the location of the single molecule on the pixel. By increasing the magnification, the image of the single molecule spreads out over the pixel array and the localization accuracy improves. However, due to the finite size

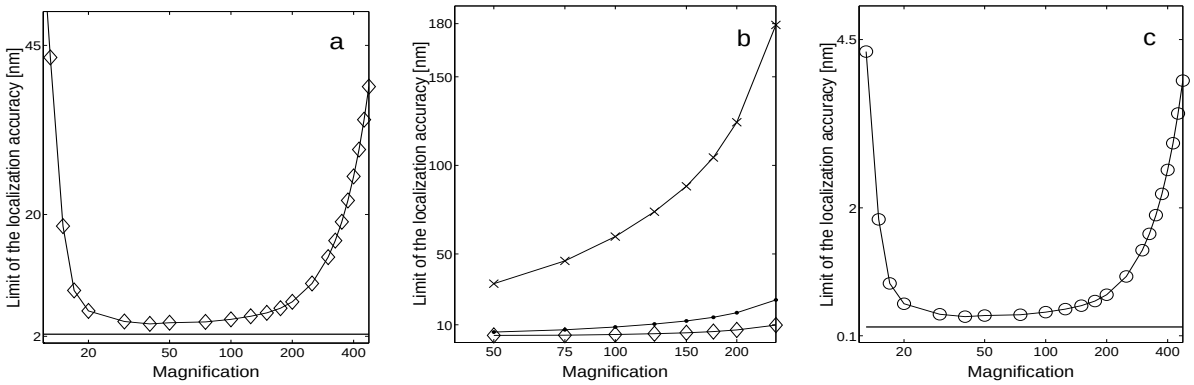

Figure 3.3. Effect of magnification on the limit of the 2D localization accuracy. The figure shows the limit of the localization accuracy for the x_0 coordinate of a single molecule with experimental parameters similar to those for a GFP molecule for a 11×11 pixel array ($5\ \mu m \times 5\ \mu m$ pixel size) as a function of magnification for different acquisition times and noise levels. Panels A and C show results in the noise free case for $t = 0.01$ s ($\diamond$) and $t = 1$ s ($\circ$), respectively. In both Panels, the fundamental limit (—) is also shown for reference. Panel B shows results for two different sets of noise parameter values. Here, ($\times$) corresponds to a scattering rate (b_k) of 6600 photons/pixel/s and a readout noise (σ_k) of 57 e^-/pixel (rms), ($\bullet$) corresponds to a scattering rate of 660 photon/pixel/s and a readout noise of 7 e^-/pixel (rms). In both cases the acquisition time is 10 ms.

of the pixel array, for larger magnification values only a small fraction of the image of the single molecule is detected by the pixel array which results in a deterioration of the localization accuracy. This shows that if data is only acquired or analyzed for a fixed pixel array, care has to be taken to match the pixel array and magnification. With an appropriate choice of magnification it is, however, possible to come close to the fundamental limit.

We next consider the effect of pixel array size on the limit of the localization accuracy. Fig. 3.4a shows the effect of the number of pixels on the limit of the localization accuracy in the noise free case for a 5×5 array and for a 21×21 array. In both cases the pixel size is fixed to $6.8\ \mu m \times 6.8\ \mu m$. By increasing the pixel array size from a 5×5 array to a 21×21 array the limit of the localization accuracy comes closer to the fundamental limit. As is to be expected, increasing the size of the pixel array improves the localization accuracy by increasing the amount of data that is available for analysis. However, in practical situations it is not always

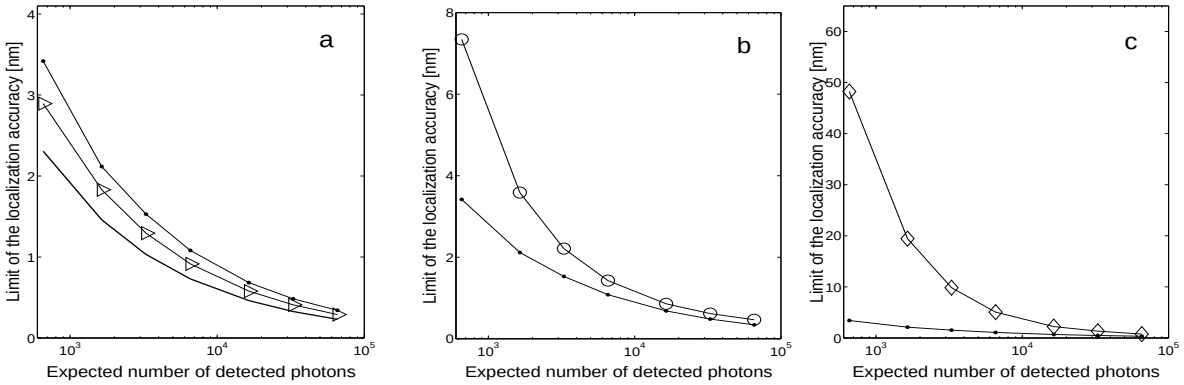

Figure 3.4. Effects of pixel array size and extraneous noise sources on the limit of the 2D localization accuracy.
The figure shows the limit of the localization accuracy of the x_0 coordinate of a single molecule with experimental parameters similar to those for a GFP molecule as a function of the expected number of detected photons for a pixelated detector in the presence of different noise levels. Panel A shows the results in the noise free case for a 5×5 pixel array ($\bullet$) and for a 21×21 pixel array ($\triangleright$). The fundamental limit (—) is also shown for reference. Panel B shows the limit of the localization accuracy ($\circ$) in the presence of noise with a scattering rate (b_k) of 660 photons/pixel/s and a readout noise (σ_k) of 7 e^-/pixel (rms) for a 5×5 pixel array. Similarly, Panel C shows the limit of the localization accuracy ($\diamond$) with a scattering rate (b_k) of 6600 photons/pixel/s and a readout noise (σ_k) of 57 e^-/pixel (rms). For all the plots the pixel size is fixed to 6.8 $\mu m \times 6.8$ μm and the x axis range corresponds to an acquisition time range of $t = 0.01$ s to $t = 1$ s. In Panels A and C, the limit of the localization accuracy in the noise free case ($\bullet$) for a 5×5 pixel array is also shown for reference.

possible to arbitrarily increase the size of the pixel array as often other elements are present in the image. This limits the number of pixels that can be used to determine the location of the single molecule unless a significant effort is made to model these other elements in the image.

The expression for the limit of the localization accuracy in the presence of noise sources allows for two noise sources, Gaussian noise as it arises, for example, in the readout process of the CCD camera and Poisson noise which can be used, for example, to model dark current in the CCD chip, scattered photons and autofluorescence. In order to study the effect of noise sources on the limit of the localization accuracy, we consider two sets of noise parameter values. In one, we set the standard deviation (σ_k) of the Gaussian noise, e.g. the readout noise, to be 7 e^-/pixel (rms). In our simulations we assume that the mean of the Gaussian noise is zero. This value for the readout noise is on the lower level of reported noise levels for current

CCD cameras. For the low noise simulations we assume that the Poisson noise has a rate b_k of $0.01\Lambda_0 = 660$ photons/pixel/s. This means, for example, that we assume that in each pixel scattered photons are collected at a rate that is 1 percent of the rate at which the photons emitted by the single molecule arrive in the detector plane. In the second set of noise parameters we consider parameters that correspond to high noise levels, in particular for the readout noise. In this case we set the standard deviation of the readout noise to be 57 e^-/pixel (rms) and the scattering rate to be $0.1\Lambda_0 = 6600$ photons/pixel/s. This level of readout noise is towards the high end on the scale of readout noise levels for current CCD cameras. The smaller value for the scattering rate is typically observed when imaging single molecules in solution ([71]), whereas the larger value is observed when imaging single molecules in a cellular environment ([26]). In all figures we assume that the noise statistics are the same for all pixels.

The dramatic effect that noise can have on the limit of the localization accuracy is shown in Fig. 3.4b, where the limit of the localization accuracy is plotted as a function of the expected number of detected photons for low scattering and measurement noise levels for a 5×5 pixel array with 6.8 $\mu m \times 6.8$ μm pixel size. Fig. 3.4c shows the same for high scattering and measurement noise levels. The effect is especially pronounced for low photon count numbers where the limit of the localization accuracy can be an order of magnitude larger than in the noise free case (see Fig. 3.4c). However, by increasing the total number of detected photons it is possible to come close to the fundamental limit even at high noise levels.

A similar effect of the noise sources on the limit of the localization accuracy is shown in Fig. 3.3b, where the limit of the localization accuracy is plotted as a function of magnification for different measurement and scattering noise levels and for the noise free case for a 11×11 pixel array with a pixel size of 5 $\mu m \times 5$ μm. For a given magnification value, the presence of high noise levels can deteriorate the best possible localization accuracy by an order of magnitude when compared to the noise free case. Also, at high noise levels the limit of the localization accuracy deteriorates by a factor of four when the magnification varies from $50\times$ to $200\times$. For example, consider the image of a GFP single molecule centered on a 11×11 pixel array with a pixel

size of 5 $\mu m \times 5$ μm. At 50× magnification ($n_a = 1.4$), the image of the GFP single molecule in the pixel array will contain 93% of the expected number of detected photons. At 100× magnification ($n_a = 1.4$), the image of the GFP single molecule in the pixel array will contain 87% of the expected number of detected photons. Although by increasing the magnification from 50× to 100× we only lose about 6% of the total number of detected photons, the limit of the localization accuracy significantly deteriorates from 38 nm at 50× magnification to 60 nm at 100× magnification at high noise levels.

As mentioned earlier, it is important to determine whether an estimation algorithm can attain the limit of the localization accuracy. In Table 1 we list the standard deviations of the maximum likelihood estimates of the single molecule location for different experimental conditions typically reported in the single molecule microscopy literature. The table also lists the limits of the localization accuracy that is calculated using Theorem 3.4.1. From the table we see that the standard deviations of the maximum likelihood estimates come reasonably close to the limit of the localization accuracy under the various experimental conditions. However, there are differences and in some cases the standard deviation of the estimates is even lower than the limit of the localization accuracy. This points to an important aspect of the theory that underlies the approach presented here. Whereas the theoretical derivations are based on considerations of the standard deviations of random variables, the simulations in Table 1 report estimates of those standard deviations, which can differ from the actual values.

Table 3.1. Verification of the limit of the 2D localization accuracy - practical imaging conditions

Parameter set#	Pixel dimensions $\mu m \times \mu m$	Readout noise e^-/pixel (rms)	Scattering noise photons/pixel/s	Pixel array size	Acquisition time ms	Fundamental limit nm	Limit of the localization accuracy in the pixelated case in the absence of noise sources nm	Limit of the localization accuracy in the pixelated case with readout and scattering noise nm	Standard deviation of the maximum likelihood estimator nm
1	6.45×6.45	6	660	13×13	10	2.3010	2.9152	6.6469	6.6903
2	6.8×6.8	18	660	12×12	10	2.3010	2.9377	14.8733	14.8175
3	13×13	4	660	7×7	10	2.3010	3.4338	4.4387	4.6287
4	15×15	32	660	6×6	10	2.3010	3.9379	13.5599	13.0138
5	6.45×6.45	6	6600	13×13	100	0.7277	0.9219	2.5138	2.6835
6	6.45×6.45	6	660	13×13	100	0.7277	0.9219	1.4241	1.4059
7	6.45×6.45	6	0	13×13	100	0.7277	0.9219	1.1951	1.4174
8	6.8×6.8	18	660	12×12	100	0.7277	0.9290	1.9855	1.8245
9	13×13	4	660	7×7	100	0.7277	1.0859	1.2319	1.3837
10	15×15	32	660	6×6	100	0.7277	1.2453	1.9365	1.7012

The table shows the limit of the localization accuracy for the x_0 coordinate of a single molecule with parameters similar to those of a GFP molecule under typical experimental conditions. The pixel dimensions and readout noise correspond to CCD cameras often used to image single molecules. For all calculations we set the photon detection rate to be $\Lambda_0 = 66000$ photons/s, the magnification to be $M = 100$, the numerical aperture to be $n_a = 1.4$, the emission wavelength to be $\lambda = 520$ nm and the mean of the measurement noise to be zero. For a given pixel dimension, the pixel array is chosen such that 92% of the photons that reach the image plane are collected. For all calculations and simulations we assume that all pixels have the same noise statistics and the single molecule is positioned at the center of the pixel array. The standard deviation of the maximum likelihood estimator for each parameter set was calculated based on maximum likelihood estimates of the single molecule location in 300 simulated images.

In Table 1, for a given pixel size, the pixel array was chosen such that approximately 92% of the photons that reach the detect or plane are collected by the pixel array. Despite this, we see that the limit of the localization accuracy varies widely for different experimental conditions, especially for short acquisition times ($t = 10$ ms, parameter sets 1-4). If the acquisition time is increased ($t = 100$ ms, parameter sets 6, 8-10), we see that the variation of the limits of the localization accuracy diminishes for the different experimental conditions. In addition, the limits of the localization accuracy also come close to the fundamental limit. So far we have shown that noise sources can significantly deteriorate the limit of the localization accuracy. It is instructive to investigate the contribution of the different noise types to the deterioration of the localization accuracy. In parameter sets 5-7 of Table 1, we show that when the scattering noise parameter is decreased from 6600 photons/pixel/s to 0 photons/pixel/s with measurement noise fixed to 6 e^-/pixel (rms), the limit of the localization accuracy decreases from 2.5138 nm to 1.1951 nm. However, in parameter set 7 we see that if the measurement noise is also set to 0, then the limit of the localization accuracy in the noise free case reduces to 0.9219 nm, which is significantly closer to the fundamental limit of 0.7277 nm.

Fig. 3.5 shows the effect of pixel size on the limit of the localization accuracy for a 1000 $\mu m \times 1000$ μm pixel array at different measurement noise levels with the scattering noise parameters set to zero, i.e. $b_k = 0$. The figure also shows the results in the noise free case and the fundamental limit. We consider a 1000 $\mu m \times 1000$ μm pixel array to ensure that a sufficient number of photons are detected in the case of large pixels. We note that the measurement noise is independent of the number of detected photons and only depends on the readout process of the CCD camera. Hence it is kept fixed when the pixel size is varied. However, since the number of pixels decreases as the pixel size increases, less noise is added to the total accumulated data. In the noise free case the limit of the localization accuracy decreases with decreasing pixel size, since with reduced pixel sizes the effect of pixelation diminishes and the limit of the localization accuracy approaches the fundamental limit. However, in the presence of measurement noise the limit of the localization accuracy first decreases but then increases with decreasing pixel

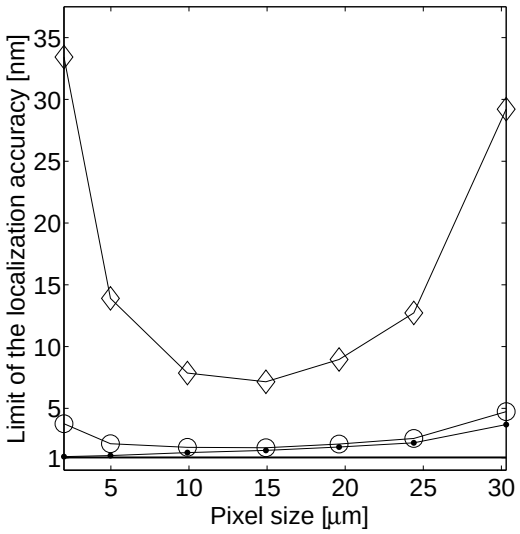

Figure 3.5. Effect of pixel size on the limit of the 2D localization accuracy.
The figure shows the limit of the localization accuracy of the x_0 coordinate of a single molecule with experimental parameters similar to those for a GFP molecule as a function of pixel size for a pixelated detector in the presence of measurement noise. ($\diamond$) corresponds to a readout noise (σ_k) of 57 e^-/pixel (rms), ($\circ$) corresponds to a readout noise (σ_k) of 7 e^-/pixel (rms) and the scattering rate (b_k) is set to 0 in both the cases. The limit of the localization accuracy in the noise free case ($\bullet$) and the fundamental limit (—) are also shown for reference. For all the plots the acquisition time is set to be $t = 0.05$ s and the pixel array size is 1000 $\mu m \times 1000$ μm. The pixel sizes were chosen such that the pixel array consists of an odd number of rows and columns.

size, since by decreasing the pixel size the number of detected photons in each pixel decreases, while the measurement noise remains the same.

We recall that a similar behavior was observed in Fig. 3.3 due to the variation in magnification. However, the present effect is different from that shown in Fig. 3.3, since by varying the magnification the size of the single molecule image was varied and this in turn affected the number of photons captured by the pixel array. In the present case, the pixel array is fixed to 1000 $\mu m \times 1000$ μm ensuring that the same number of photons are captured by the pixel array for all pixel sizes. From this we deduce that for a given experimental setup, the limit of the localization accuracy depends not only on the total number of detected photons but also on the number of photons captured in each pixel in the pixel array. We note that an analogous

behavior was reported in [28], where the effect of pixel size on the localization accuracy was discussed for a specific estimation procedure.

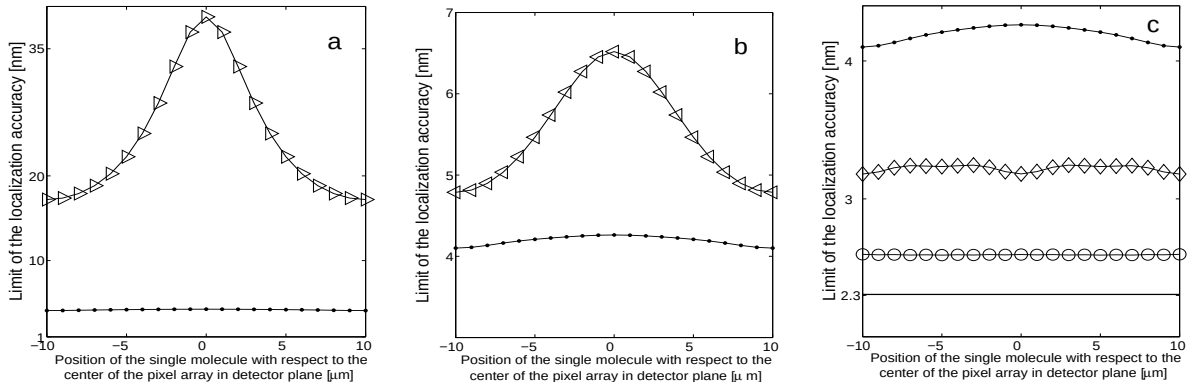

Figure 3.6. Effect of single molecule loaction on the limit of the 2D localization accuracy. The figure shows the limit of the localization accuracy of the x_0 coordinate of a single molecule with experimental parameters similar to those for a GFP molecule for a pixelated detector as a function of the single molecule position for different noise levels and pixel sizes. Panel A ($\triangleright$) shows results in the presence of noise with a scattering rate (b_k) of 6600 photons/pixel/s and a readout noise (σ_k) of 57 e^-/pixel (rms) for a 5×5 pixel array with a pixel size of 20 $\mu m \times 20$ μm. Panel B shows the same for a scattering rate (b_k) of 660 photons/pixel/sec and a readout noise (σ_k) of 7 e^-/pixel (rms) ($\triangleleft$). Panel C shows results in the noise free case for a 10×10 pixel array with 10 $\mu m \times 10$ μm pixels ($\diamond$) and for a 50×50 pixel array with 2 $\mu m \times 2$ μm pixels ($\circ$). The fundamental limit (—) is also shown for reference. In all three plots the acquisition time is $t = 0.01$ s, the pixel array size is 100 $\mu m \times 100$ μm and ($\bullet$) shows the limit of the localization accuracy in the noise free case for a 5×5 pixel array with 20 $\mu m \times 20$ μm pixels. The x axis of the plots denotes the position of the single molecule with respect to the center of the pixel array (in the detector plane). The single molecule is moved in steps of 10 nm in the specimen plane which corresponds to 1 μm steps in the detector plane. All movements are parallel to the pixel edges. For a 20 $\mu m \times 20$ μm pixel this corresponds to moving the single molecule from one edge of the central pixel to the opposite edge of the pixel, while for a 10 $\mu m \times 10$ μm pixel this corresponds to moving the single molecule over a pair of pixels that are centrally located on the detector and for a 5 $\mu m \times 5$ μm pixel this corresponds to moving the single molecule over a set of 4 pixels centrally located on the detector.

We next consider the effect of the location of the single molecule with respect to the pixel array on the limit of the localization accuracy. Fig. 3.6a shows the variation of the limit of the localization accuracy as a function of the single molecule position for a 5×5 pixel array with a pixel size of 20 $\mu m \times 20$ μm for high measurement and scattering noise levels. Fig. 3.6b

shows the same for low measurement and scattering noise levels. In both figures, the result for the noise free case is also shown for reference. From Figs. 3.6a and 3.6b, we observe that the localization accuracy varies periodically as the single molecule is moved along the x direction. The best, i.e. the smallest, values for the limit of the localization accuracy are achieved when the image of the single molecule is centered on the edge of a pixel. This is due to the fact that small changes in the location of the single molecule at the edge of a pixel lead to significant changes in the collected data. The worst, i.e. largest, values for the limit of the localization accuracy are achieved when the image of the single molecule is located at the center of a pixel. From Figs. 3.6a and 3.6b, we see that the worst case value can be anywhere between 10% - 80% higher than the best case value depending on the noise level. Note that the variation of the limit of the localization accuracy is particularly pronounced for large pixel sizes. This periodicity for a pixelated detector is in contrast to the fundamental limit (eq. 3.18) which is independent of the location of the single molecule. As shown in Fig. 3.6c, by reducing the pixel size the effect of pixelation is diminished and hence the periodic variation in the localization accuracy also decreases. An immediate implication of the above result is that moderate variations in the single molecule position within a pixel can lead to substantially different localization accuracy values in the presence of high noise levels. This analysis also provides an explanation of the phenomenon that was reported in [29], where it was observed using numerical investigations that the localization accuracy depends on the location of the single molecule with respect to the pixels. It should be noted that other published expressions for the localization accuracy ([26, 28]) do not show a dependence on the magnification and on the location of the single molecule. This is due to approximations that were used in the model for the acquired signal in the pixels.

The results presented in this section give an indication as to the type of phenomena that can be investigated with our approach. Further applications are easily conceivable such as the evaluation of the effect of pixel shape or the presence of dead space regions on the detector (e.g. due to the presence of anti blooming gates) on the limit of the localization accuracy.

While the expressions for the limit of the localization accuracy in the pixelated case are not as straightforward to analyze as the expression for the fundamental limit they can be numerically evaluated in a relatively straightforward way.

3.5 Appendix to Chapter 3

3.5.1 Simulations and parameter values

All simulations and calculations were carried out in the MATLAB programming environment ([72]). We assume the fluorescent single molecule to have an emission wavelength of 520 nm. For all calculations, unless explicitly stated, the numerical aperture is set to be $n_a = 1.4$, the magnification is set to be $M = 100$ and the acquisition time is in the range of $t = 0.01$ s to $t = 1$ s (with $t_0 = 0$). We set the photon detection rate to be $\Lambda_0 = 66000$ photons/s, which is in the range of values typically observed in single molecule experiments ([71, 25, 26, 73]). In the figures given in Section 3.4, we assume square pixels with no dead space between adjacent pixels and unless otherwise stated, the single molecule is positioned at the center of the pixel array.

3.5.2 Maximum likelihood estimation

In the non-pixelated case, maximum likelihood estimation was carried out for two different acquisition methods, one when the acquisition time was fixed and the other when the total number of detected photons was fixed. In the former case, due to the stochastic nature of photon emission, the total number of detected photons varied for every image, while in the latter case, the number of detected photons remained the same.

For the first acquisition method, a Poisson random number N_1 with mean $\Lambda_0(t - t_0)$ (denoting the expected number of detected photons) was generated and N_1 random vectors were generated (see Random number generation) that describe the spatial coordinates of the detected photons. The maximum likelihood estimation was carried out using a gradient based search algorithm (Optimization Toolbox of Matlab [74]). For every value of $\Lambda_0(t - t_0)$, 300

estimates of position were computed from which the standard deviation was calculated. For the second acquisition method the same procedure was followed except that no Poisson random number was generated since the number of detected photons was fixed.

In the pixelated case, the maximum likelihood estimation was performed for the fixed acquisition time method. For a given pixel array size, pixel dimensions, single molecule location and $\Lambda_0(t - t_0)$, 300 images were simulated by first generating a noise free pixelated image and then adding Poisson and Gaussian noise to the pixel values. Using the simulated data, maximum likelihood estimation was carried out using an algorithm analogous to the one that was mentioned above. The standard deviation of the estimates of the single molecule location was then calculated.

3.5.3 Random number generation

The simulation of the two-dimensional distribution corresponding to the point spread function can be carried out by reducing the simulation to that of two one-dimensional distributions. Let Φ denote a uniform random variable with density function $f_\Phi(\phi) = 1/2\pi$, $0 \leq \phi \leq 2\pi$ and let R denote a one-dimensional continuous random variable with density function $f_{R,a}(r) = \frac{2J_1^2(ar)}{r}$, $r \geq 0$, where $a = (2\pi n_a)/(\lambda_{em} M)$. Let R and Φ be independent of each other. Define $X := R\cos\Phi + Mu$ and $Y := R\sin\Phi + Mv$, where u, v denote the coordinates of location of the single molecule and M denotes the magnification of the objective lens. Then the joint density function of X and Y is given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi r} f_{R,a}(\sqrt{(x-Mu)^2 + (y-Mv)^2}) = \frac{J_1^2(a\sqrt{(x-Mu)^2 + (y-Mv)^2})}{\pi((x-Mu)^2 + (y-Mv)^2)},$$

where $-\infty < x, y < \infty$. To generate a random vector (x, y) that describes the spatial coordinates of the detected photons on the detector, we first generate a uniform random number ϕ between 0 and 2π, then generate a random number r with density function $f_{R,a}$ and set $x := r\cos\phi + Mu$, $y := r\sin\phi + Mv$. The uniform random number ϕ is generated using a standard random number generator ([72]). The random number r is generated by the trans-

formation method ([63]) in conjunction with a numerical inversion of the distribution function corresponding to $f_{R,a}$ using a look-up table.

CHAPTER 4

MULTIFOCAL PLANE MICROSCOPY & 3D SINGLE PARTICLE TRACKING

4.1 Introduction

Fluorescence microscopy is an important tool to study cellular trafficking pathways in live cells. Significant progress has been made in the understanding of cellular events by conducting new types of experiments with fluorescent proteins. This has been possible due to advances in fluorescent labeling techniques ([75]) along with the use of highly sensitive detectors ([16]). The standard microscope design is very well suited to image cellular events in one focal plane. For example, the advent of total internal reflection fluorescence (TIRF) microscopy ([19]) has enabled the study of trafficking dynamics on the cellular membrane with high sensitivity and in great detail. In TIRF imaging, the illumination intensity decreases exponentially with the distance from the cover glass. Hence the out of focus fluorescence typically does not interfere with the emitted signal. This has allowed events at the plasma membrane to be imaged with unprecedented sensitivity. However, for example in the study of exocytic events, the advantage of TIRF microscopy also has a negative aspect. Due to the shallow illumination layer, exocytic events cannot be related to intracellular events. On the other hand, in epifluorescence microscopy, the cell is uniformly illuminated, which enables the imaging of events that occur in the cell interior. However, in epifluorescence imaging the out-of-focus haze on the plasma membrane of the intracellular fluorescence typically overwhelms the signals from the plasma membrane and thereby makes high sensitivity imaging of events at the plasma membrane impossible.

The above limitations of current imaging techniques make it difficult to image relatively rapid intracellular trafficking processes that are not confined to one focal plane. Such trafficking events are of central interest to cell biologists. For example, the imaging of rapidly moving tubules or vesicles on the recycling pathway leading from sorting endosomes to exocytosis on the plasma membrane has not been possible with the existing microscopy techniques. A central

problem with the current microscope design is that only one focal plane can be imaged at a time. A classical solution to overcome this problem is to sequentially scan the sample at different focal planes by using a z-focusing device (for example, piezo nano-positioner). However, such focusing devices are relatively slow, typically taking tens of milliseconds to move the focal plane even by a fairly short distance. As a result, when the cell-sample is imaged at one plane important events occurring in the other planes can be missed.

This chapter discusses the design and implementation of a new microscope that overcomes the above problems by simultaneously imaging two or more distinct focal planes within the specimen. This is achieved by a modification of the emission pathway of a standard fluorescence microscope, where multiple detectors are placed at specific, calibrated distances from the tube lens. The new microscope is referred to as **mu**ltifocal plane **m**icroscope or MUM for short. The construction of MUM is simple and can be done with readily available off-the-shelf components. As an application, the chapter discusses the use of MUM for tracking single molecules/particles in 3D within a cellular environment. Analytical tools are introduced to estimate the 3D particle locations and to characterize the accuracy with which the 3D locations can be determined. By using the results derived in Chapter 2, it is shown that the MUM has improved depth discrimination capability when compared to a regular optical microscope. As a result, the axial location of the single particle can be determined with relatively high accuracy even when the particle is close to the plane of focus. This capability is verified through simulations and from experimental data.

The organization of this chapter is as follows. Section 4.2 deals with the design, construction and calibration of the MUM. Section 4.3 discusses the improved depth discrimination capability of the MUM and the implications for 3D single particle tracking. In Section 4.4, a global estimation algorithm is introduced for determining the 3D location of single particles from MUM data. The design, implementation and calibration of the MUM have been published in [76, 77, 78]. The analysis pertaining to the improved depth discrimination capability has been published in [36, 53].

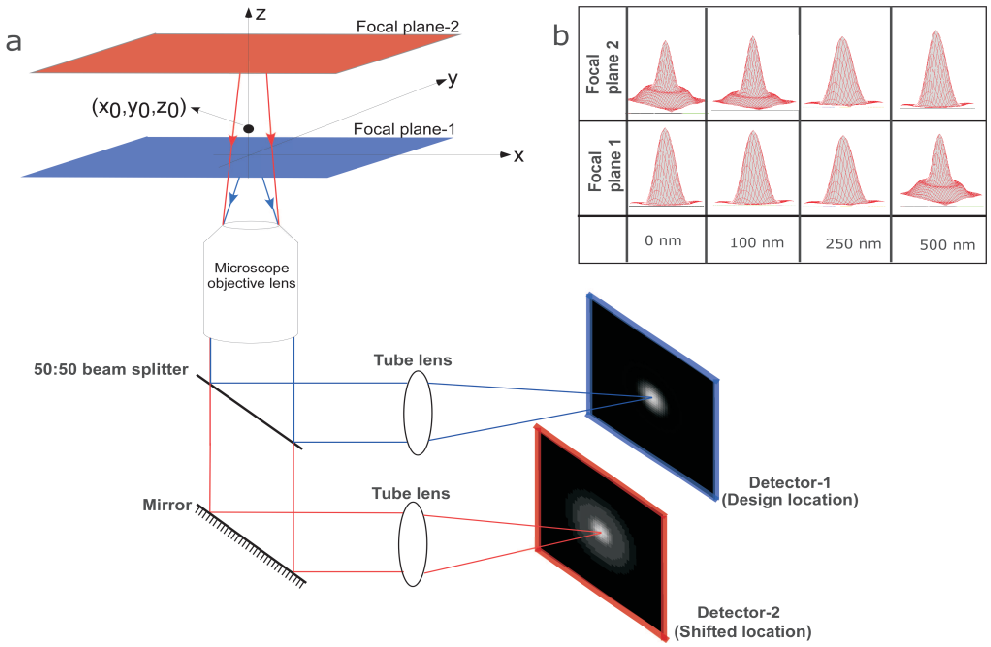

Figure 4.1. Principle of MUM.
Panel A shows the principle and schematic of a multifocal focal plane microscope. The figure illustrates the effect of changing the position of the detector relative to the tube lens, which results in imaging a plane that is distinct from the plane that is imaged by the detector positioned at the design location. Panel B shows simulated 3D images of single molecules acquired at different defocus levels when imaged through a multifocal plane microscope. Here, the defocus levels are specified with respect to focal plane 1 and the distance between the two focal planes in the object space is assumed to be 500 nm.

4.2 Principle of MUM

Figure 4.1A shows a schematic illustrating the principle behind MUM. Here, the sample is illuminated in widefield mode and the light collected from the sample is split into two paths by a beam splitter. In each path the split light is focused onto a detector by a tube lens. The detector is located at a specific calibrated distance from the tube lens, and this distance is different in every light path. As a result of this, the detectors capture images of distinct planes within the specimen. For example, if one of the detectors is located at the design position (i.e., focal plane of tube lens) and the other detector is shifted towards (away from) the tube lens, then the detector at the shifted location will image a plane in the specimen that is above (below) the plane imaged by the detector at the design position. In the schematic shown in Figure 4.1A,

two distinct planes can be simultaneously imaged. By introducing additional beam splitters in each light path, the design can be modified to image more than two planes. Presently, a multifocal plane microscope has been constructed that can image up to four distinct planes (see Fig. 4.2).

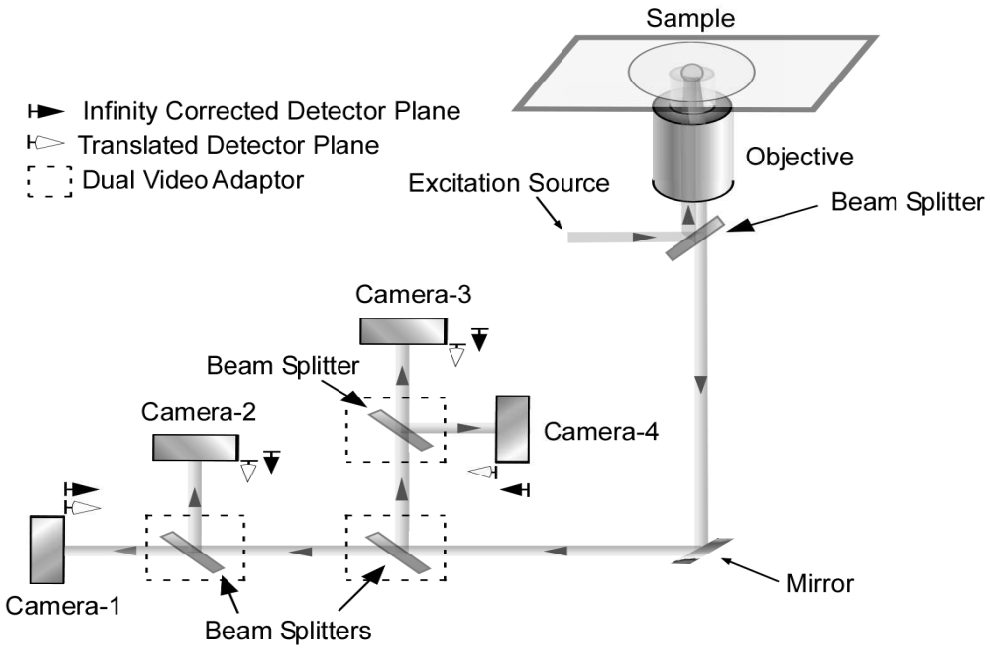

Figure 4.2. Schematic of a MUM setup that can image upto four distinct planes.

Now, consider a point source which lies between the two focal planes that are simultaneously imaged. Because the point source is out of focus with respect to both focal planes, each of the detectors will record a defocused image. Figure 4.1B shows the intensity profile of simulated images at various defocus levels for a two plane imaging setup. Here, the defocus level is specified with respect to focal plane 1. If we only have images acquired from focal plane 1, which pertains to the imaging scenario in a regular widefield microscope, then we see that for a range of defocus levels (0 - 250 nm), the images show negligible change in their shape thereby providing very little information about the depth of the point source. On the other hand, if we also consider images acquired from focal plane 2 then for the same range of defocus, the images

acquired at focal plane 2 show significant change in the shape thereby providing relatively more information about the depth of the point source when compared to images acquired at focal plane 1.

4.2.1 Construction of MUM

The implementation of MUM is straightforward and can be done in any widefield microscope. Here, the details of the implementation that was carried out with Zeiss microscopes are described. Two different multifocal plane imaging configurations were implemented. The first configuration supported simultaneous imaging of two distinct planes within the specimen. A Zeiss dual video adaptor (Part # 1058640000) was attached to the bottom port of a Zeiss Axiovert 100 microscope and two EM CCD cameras (iXon DV887, Andor Technologies, South Windsor, CT) were used. Here, one of the cameras was attached to the video adaptor through a standard Zeiss camera-coupling adaptor (Part# 4561059901). The other camera was attached to the video adaptor by using C-mount/spacer rings (Edmund Industrial Optics, Barrington, NJ) and a custom machined camera-coupling adaptor (which is identical to but shorter than a standard Zeiss camera-coupling adaptor).

The second configuration supported simultaneous imaging of up to four distinct planes within the specimen. Here, a Zeiss video adaptor was first attached to the side port of a Zeiss Axiovert 200 microscope. Then two Zeiss video adaptors were concatenated by attaching each of them to the output ports of the first Zeiss video adaptor. Four high resolution CCD cameras (2 ORCA-ERs, 2 C8484-05, Hamamatsu, Bridgewater, NJ) were attached to the output ports of the concatenated video adaptors by using C-mount/spacer rings and custom machined camera coupling adaptors.

To image more than four planes, the procedure described above can be repeated by concatenating additional video adaptors. It should be pointed out that the design of the input and output ports of the Zeiss video adaptor allows them to be concatenated. If dual

video adaptors from other vendors are used, then concatenation of video adaptors may not be possible, thereby limiting the implementation to image only two distinct planes.

4.2.2 Demonstration of simultaneous multifocal plane imaging

To verify if the above proposed design can image distinct planes, an imaging experiment was carried out. Images of 100 nm beads (Polysciences Inc. Warrington, PA) were acquired in a multifocal plane microscope that was configured to image two distinct planes within the specimen. Here, one of the cameras (camera-1) was positioned at the focal plane of the tube lens, i.e., the detector position in a regular microscope setup, and the other camera (camera-2) was positioned 8 mm towards the tube lens from the parfocal position. During the acquisition, the objective lens was moved in 25 nm increments by a piezo focusing device (Polytec PI, Auburn, MA). Fig. 4.3A shows a montage of bead images acquired with the MUM. The bead is clearly seen to be in focus at different positions of the objective z-focus. This shows that the imaging system indeed produces images at different focal planes.

4.2.3 Determination of focal plane separation

An important aspect of MUM is the calibration of the distances between the different focal planes that are being simultaneously imaged. To do this, a sample containing 1μm beads (Polysciences Inc, Warrington, PA) in water was imaged in a MUM setup. Here a two plane imaging setup was used where one of the cameras (camera-1) was positioned at the focal plane of the tube lens and the other camera (camera-2) was moved towards the tube lens in steps of 1 mm. At each position of camera-2, a 3D image stack of an isolated bead was simultaneously acquired in both cameras while the z-focus of the objective lens was changed in 25 nm steps by a piezo focusing device (Polytec PI, Auburn, MA). For each camera, the fluorescence intensity of the bead was plotted as a function of the piezo z-position and a low order polynomial was fit around the maxima of the intensity plot to find the peak (see Fig. 4.3B). Then the difference between the z-positions of the peaks for each camera was calculated and this was plotted as a function of camera-2 position. The resulting plot provides a calibration graph which relates

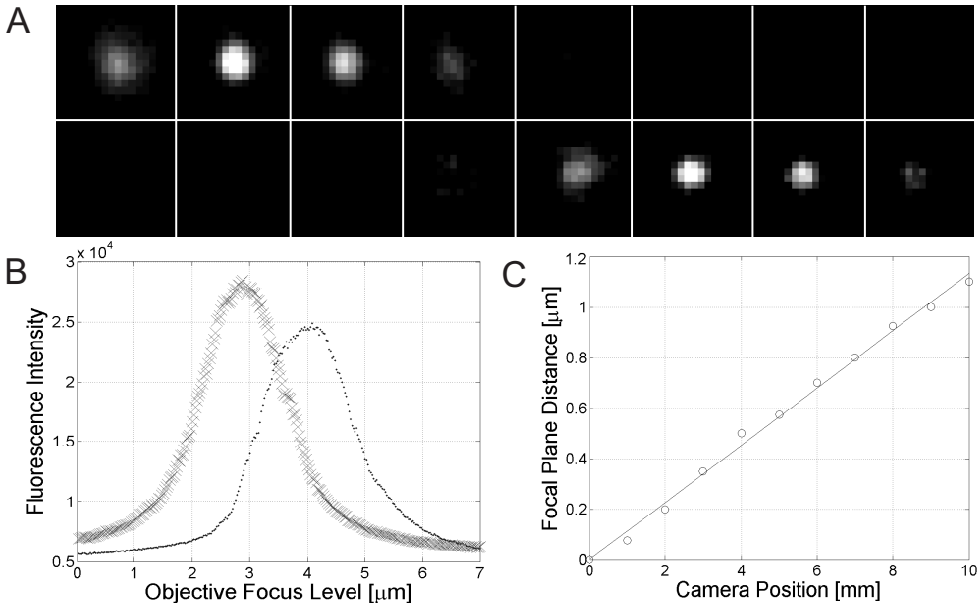

Figure 4.3. Calibration of MUM.
Panel A shows a montage of images taken of a 100-nm-diameter fluorescent bead with a MUM setup capable of imaging two distinct planes. One of the cameras (Camera 1) was positioned at the focal plane of the tube lens, whereas the other camera (camera 2) was positioned 8.8 mm away from the focal position toward the tube lens corresponding to a 1 μm distance between the two focal planes (based on calibration data obtained from Panel C). The top row shows images acquired by camera 2 whereas the bottom row shows images acquired by camera 1. The images were taken at different positions of the z-focus of the objective. It is clearly seen that the bead is in focus at different planes for the two cameras, thereby confirming that the MUM setup allows the imaging of different focal planes. Panel B shows the plot of the fluorescent intensities of the images acquired as in Panel A and plotted against objective z-focus levels. Here images were acquired at 25-nm increments. The fluorescence intensity plots have different peaks indicating that the focal planes for the two cameras are different. Panel C shows the calibration graph obtained for the MUM setup. For a given position of camera-1, the measurement of the distance between the peaks of the plots shown in Panel B gives a measurement of the difference in the focal planes between camera 1 and camera 2. For different positions of camera 2, analyses of the images were carried out as in Panel B that result in estimates of the difference in focal planes between the two cameras. The current plot reveals a linear relationship between the position of camera 2 (X mm) and the difference in focal planes between the two cameras (Y μm), given by Y = 0.113X.

the distance between the two focal planes in the object space to the location of camera-2 in the image space (see Fig.4.3C).

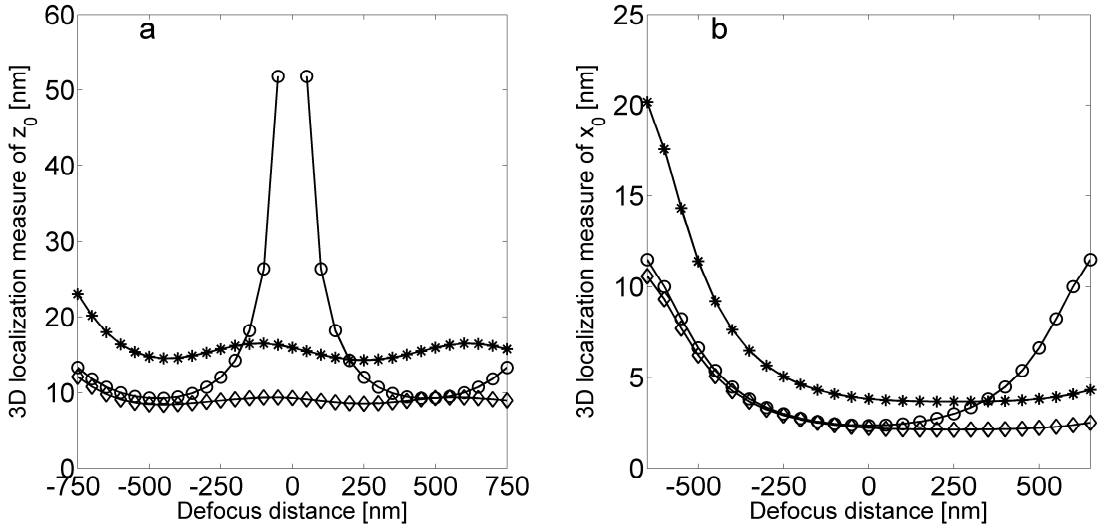

Figure 4.4. Improved depth discrimination capability of MUM.
Panel A (B) shows the variation of the 3D localization measure of z_0 (x_0/y_0) as a function of the defocus level for a conventional widefield microscope ($\circ$) and for a multifocal microscope that can image two distinct planes ($*$,$\diamond$). In all the plots the numerical aperture of the objective lens is set to $n_a = 1.45$, the wavelength is set to $\lambda = 0.665$ μm, the pixel array size is set to 11×11, the pixel size is set to 16 $\mu m \times 16$ μm, the X-Y location coordinates of the particle are assumed to coincide with the center of the pixel array, the background level is set to 300 photons/pixel/s, the exposure time is set to 1 s, the mean and standard deviation of the readout noise is set to 0 e^- and 8 e^- rms, respectively. For the conventional widefield microscope, the photon detection rate is set to 5000 photons/s. For the multifocal plane microscope, the photon detection rate is either set to 2500 photons/s ($*$) or set to 5000 photons/s ($\diamond$) and the distance between the two planes in the object space was set to be 500 nm.

4.3 Depth discrimination and 3D localization measure

In Section 4.2, qualitative explanation was provided for the improved depth discrimination capability of MUM (4.1B). Here, we provide a quantitative evaluation of the improvements from the point of view of 3D tracking. One of the fundamental questions in 3D particle tracking experiments concerns the accuracy with which the location of a particle/object of interest can

be determined. By making use of the results derived in Chapter 2, analytical expressions are derived to calculate the 3D localization measure, i.e., the limit to the 3D localization accuracy, of the single particle that is imaged through a MUM (see Section 4.5.5 for details).

Figure 4.4A shows the 3D localization measure for the z_0 coordinate as a function of the defocus level for a regular microscope setup. Here, we assume imaging conditions typical of single-molecule experiments and consider a point object emitting 665 nm light that is imaged by a 1.45 NA objective lens. From the figure we see that as the point object comes into focus, the 3D localization measure deteriorates and becomes infinitely large thereby implying that there exists significant uncertainty in determining the z_0 (axial) coordinate of a point object when it is close to the plane of focus ([36, 37]). For instance, our result predicts that defocus levels of 200 nm, 100 nm, 50 nm and 10 nm can be determined with an accuracy no better than 14.2 nm, 26.4 nm, 51.8 nm and 357.2 nm, respectively, when 5000 photons are collected from the object. This implies that defocus levels of 100 nm and 200 nm can be determined with moderate to high accuracy. In contrast, for defocus levels of 10 nm and 50 nm, the predicted accuracy is either equal to or much greater than the corresponding defocus level itself, thereby implying poor accuracy in determining the defocus level. As a result of this, it is problematic to carry out 3D single-particle tracking in a regular widefield microscope especially when the single particle is near the plane of focus. Figure 4.4A also shows the 3D localization measure of z_0 for the MUM that can simultaneously image two distinct planes. In stark contrast to a regular microscope, the 3D localization measure of z_0 for the MUM does not deteriorate as the point object comes into focus. In fact, the 3D localization measure of z_0 remains relatively constant for a range of defocus levels. For example, our result predicts that defocus levels of 200 nm, 100 nm, 50 nm and 10 nm, can be determined with an accuracy no better than 14.4 nm, 15 nm, 15.5 nm and 15.8 nm, respectively, with the MUM when 2500 photons are collected in each plane. This implies that the axial location of the point object can be determined with relatively the same level of accuracy for a range of defocus levels using the MUM. It should be pointed out that the 3D localization measure depends on the photon count, as a result of which the predicted

accuracy for the MUM can be further improved by collecting more photons (see Figure 4.4A). We have also investigated the 3D localization measure of the x_0/y_0 coordinates and we observe a similar behavior to that of the z_0 coordinate (see Fig. 4.4B). The investigation of the 3D localization measure presented here, can be extended for multifocal plane configurations that can simultaneously image more than two planes (see Chapter 4 Appendix for details).

4.4 Verification of improved depth discrimination capability

The results given in the previous section are predictions based on the theory concerning the Fisher information matrix. To verify if these predictions can be attained in practice with the MUM we carried out test experiments on simulated data for a GFP and a QD label that is imaged in a two plane imaging setup. Here, images of the GFP/QD label were simulated for different defocus levels. A global estimation algorithm was implemented to extract the defocus level from simulated data. The details of the estimation algorithm are given in the Chapter Appendix (see *Data Analysis* Section). The results of the analysis are shown in Fig. 4.5 and table 4.1 lists the mean and standard deviation of the defocus level estimates along with predicted 3D localization measure (see Chapter Appendix for details). From the table we see that the standard deviation of the defocus-level estimates comes consistently close to the predicted 3D localization measure thereby validating the improvement in depth discrimination. For example, for the GFP label, defocus levels of 10 nm, 20 nm, 50 nm, 100 nm and 200 nm, are estimated with an accuracy of 15.37 nm, 15.17 nm, 14.09 nm, 15.17 nm and 13.91 nm, respectively. These values are in reasonable agreement to the predicted 3D localization measure, which, for the above defocus levels are, 14.21 nm, 14.25 nm, 14.34 nm, 14.43 nm and 14.51 nm, respectively. Note that the defocus level estimates for the GFP label show greater variability than the QD label (See Fig. 4.5). That is the standard deviation of the defocus level estimates for the GFP label is greater than that for the QD label. This behavior is due to the fact that the total number of detected photons for the GFP images wase assumed to be 1500

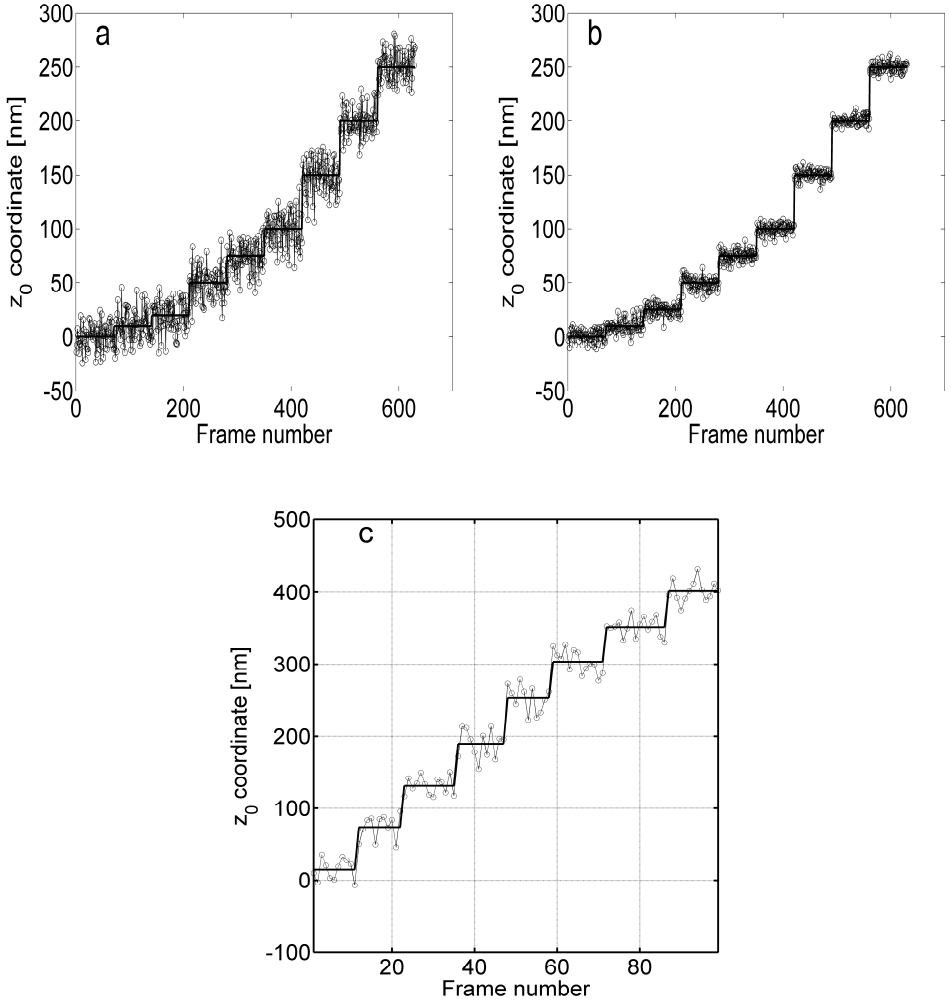

Figure 4.5. Results of defocus level estimation for simulated and experimental data. The figure shows the results of the global estimation of the defocus levels from multifocal plane data. Panels A and B show results from simulated data for GFP and QD label, respectively, and Panel C shows results from experimental data. In all panels, the imaging data is acquired from a two plane imaging setup. The distance between the two focal planes in the object space is set to 350 nm in panel A, 500 nm in panel B and 300 nm in panel C. For data simulation, the pixel size is set to $16\mu m \times \mu m$, the pixel array size is set to 11×11, the numerical aperture is set to $n_a = 1.45$, the wavelength is set to $\lambda = 520$ nm for GFP label and $\lambda = 665$ nm for the QD label, the background photon count is set to 150 photons/pixel/s for the GFP label and 300 photons/pixel/s for the QD label, the mean and standard deviation of the readout noise is set to 0 e^-/pixel and 8 e^- /pixel, respectively, the photon detection rate is set to 1500 photon/s for the GFP label and 10000 photons/s for the QD label, the expsoure time is set to 1 s, the X-Y coordinates is assumed to concide with the center of the pixel array and the noise statistics is assumed to be the same for all pixels.

per focal plane image, whereas for the QD label it was assumed to be 10,000 per focal plane image.

We next conducted test experiments on experimental data. Here, quantum dot labeled IgG (QD-IgG) molecules were pulsed in human endothelial cells that were transiently transfected with the IgG receptor, FcRn [79]. The presence of FcRn resulted in a three dimensional distribution of QD-IgG molecules within the cell. The cell sample was fixed, mounted on a microscope slide and imaged in a two plane imaging setup. To obtain images of the QD label at different defocus levels, the objective lens was moved with a piezo-nanopositioner in 50 nm steps and at each piezo position several images of the two planes were simultaneously captured. The images of the QD label acquired at the two focal planes were then analyzed with the maximum global estimation algorithm that was used above. Fig. 4.5C shows the results of the analysis for one of the QD labels. From the figure we see that the estimation algorithm is able to follow the stepwise movement of the piezo nanopositioner. For instance, defocus levels of 15 nm, 73 nm, 130 nm, and 194 nm, are estimated with an accuracy of 14.84 nm, 17.6 nm, 13.6 nm and 13.9 nm, respectively. For the above estimated defocus levels, the predicted 3D localization measure values are 8.8 nm, 8.7 nm, 8.65 nm and 7.98 nm, respectively. Unlike simulated data, the accuracy of the defocus level estimates for experimental data is not very close to the predicted 3D localization measure. This behavior can be attributed to the fact that in the experimental data the standard deviation is calculated from a relatively small number of defocus level estimates (N = 12), whereas for the simulated data the standard deviation is calculated from a large number of defocus level estimates (N = 70).

4.5 Appendix to Chapter 4

4.5.1 Determination of appropriate plane spacing

For tracking experiments, the distance between the membrane plane and the top plane was determined in the following manner. First the 3D localization measure for different defocus levels was calculated at various focal plane spacings. Then the appropriate plane spacing was

Table 4.1. Results of global estimation from MUM data

Defocus level z_0 [nm]	Mean of z_0 estimates [nm]	Std-dev of z_0 estimates [nm]	3D localization measure of z_0 [nm]	Label
0	0.63	12.01	14.16	GFP
10	9.38	15.37	14.21	GFP
20	18.99	15.37	14.25	GFP
50	49.28	14.09	14.34	GFP
75	72.85	14.77	14.40	GFP
100	98.36	15.17	14.43	GFP
150	148.85	15.11	14.46	GFP
200	199.53	13.91	14.51	GFP
250	251.57	13.67	14.45	GFP
0	0.77	5.50	5.66	QD
10	10.37	5.97	5.65	QD
25	24.90	5.25	5.64	QD
50	49.86	5.95	5.60	QD
75	74.83	5.32	5.54	QD
100	98.83	4.98	5.48	QD
150	149.95	4.65	5.35	QD
200	200.75	4.49	5.25	QD
250	250.05	5.01	5.21	QD

The table lists the results of the global estimation for simulated multifocal plane data for a GFP and a QD label. The data is simulated for a two plane imaging setup, where the focal plane spacing is 350 nm for the GFP label and 500 nm for the QD label. For each defocus level, the standard deviation is calculated from 70 estimates of z_0, which are shown in fig. 4.5. For the GFP (QD) label, the photon detection rate is set to 1500 (10000) photons/s, the wavelength is set to 520 nm (665 nm), the background is set to 150 (300) photons/pixel/s, the numerical aperture is set to 1.45, the exposure time is set to 1 s, the pixel size is set to 16 $\mu m \times 16$ μm, the ROI size is set to 11×11, the mean and standard deviation of the readout noise is set to 0 $e-$ and 8 e^-/pixel, respectively, the magnification is set to 100, the X-Y coordinates is set to be equal to the center of the ROI for all the focal plane images and the noise statistics is set to be the same for all pixels. The global estimation is carried out by estimating the 3D location coordinates (x_0, y_0, z_0) from the simulated MUM data.

set to be that value at which the 3D localization measure was constant for a range of defocus levels.

4.5.2 Imaging experiments

Images were acquired on Axiovert microscopy (Zeiss, Thornwood, NY) imaging workstations that were modified to simultaneously image different focal planes within the specimen. with a 100x, 1.45 NA α-plan Fluar Zeiss objective lens. For cellular imaging experiments the sample was illuminated in epifluorescence with a 488 nm laser line (Reliant 150M, Laser Physics, Salt Lake City, UT).

4.5.3 Cells, transfection and reagents

Human microvasculature endothelial cell line HMEC1.CDC ([80]), generously provided by F. Candal at the Centers for Disease Control (Atlanta, GA), was used for all experiments. Plasmids to express fusion protein constructs of wild type human FcRn (pHluorin-hFcRn) and mutant human FcRn (GFP-mut_hFcRn and mut_hFcRn-mRFP), and the construct of human beta2 microglobulin (hb2m) have been described previously ([81, 78]). HMEC cells were transiently transfected with protein expression plasmids using Nucleofector technology (Amaxa Systems, Cologne, Germany). Quantum dot 655 coated with streptavidin (QD) and alexa555 labeled transferrin were purchased from Invitrogen Corporation (Carlsbad, CA). Mutant human IgG1 (MST-HN) was expressed and purified as described previously ([82]). QD-IgG complexes were prepared by mixing QD and site-specifically biotinylated IgG at a molar ratio of 1:0.5 of QD to IgG.

4.5.4 Data analysis

All data processing was carried out in MATLAB and viewed using the Microscopy Image Analysis Tool (MIATool) software package[1]. To determine the defocus level of the quantum dots, a small region of interest containing the QD image was selected from each focal plane

[1] http://www4.utsouthwestern.edu/wardlab/miatool

image. Prior to curve fitting, the pixel values in the selected ROIs were converted to photon counts by subtracting the constant offset from each pixel value and then multiplying it by the conversion factor. The constant offset and the conversion factor are specific to a given camera and are typically reported by the camera manufacturer. The defocus level of the QD was then determined by fitting image profiles based on 3D point spread function models to the selected ROIs through a global estimation algorithm (`lsqnonlin`) implemented using the optimization toolbox of Matlab.

For a two plane imaging setup, the 3D PSF based image profiles ν_θ^1 and ν_θ^2 are given by

$$\nu_\theta^1(k,t) = \frac{\alpha^2 \Lambda_0 t}{\pi M_1^2} \iint\limits_{C_k} \left| \int_0^1 J_0\left(\frac{\alpha}{M_1}\sqrt{(x-M_1x_{01})^2+(y-M_1y_{01})^2}\rho\right) \exp(jW_{z_{01}}(\rho))\rho d\rho \right|^2 dxdy + B_1, \tag{4.1}$$

$$\nu_\theta^2(l,t) = \frac{\alpha^2 c \Lambda_0 t}{\pi M_2^2} \iint\limits_{C_l} \left| \int_0^1 J_0\left(\frac{\alpha}{M_2}\sqrt{(x-M_2x_{02})^2+(y-M_2y_{02})^2}\rho\right) \exp(jW_{z_0-\delta z_f}(\rho))\rho d\rho \right|^2 dxdy + B_2, \tag{4.2}$$

where C_k (C_l) denotes the region on the detector plane occupied by the k^{th} (l^{th}) pixel, $k = 1,\ldots,N_1$, $l = 1,\ldots,N_2$, and N_1 and N_2 denote the total number of pixels in the ROIs selected from plane 1 and plane 2, respectively.

In the above expressions z_0 denotes the axial location of the point source, (x_{01}, y_{01}) and (x_{02}, y_{02}) denote the lateral (X-Y) location of the point source corresponding to focal plane 1 and focal plane 2, respectively, Λ_0 denotes the photon detection rate, t denotes the exposure time, c is a constant, δz_f denotes the distance between the two focal planes in the object space, $\alpha := 2\pi n_a/\lambda$, M_1 and M_2 (B_1 and B_2) denote the lateral magnification (background component) corresponding focal plane 1 and focal plane 2, respectively, and $\theta = (x_{01}, y_{01}, x_{02}, y_{02}, z_0, \alpha, A)$ denotes the parameters that are estimated. The magnification M_1 is set to the magnification of the objective lens, and the magnification M_2 is determined as described in Section 4.5.4.1 (see below). The above expressions of μ_θ^1 and μ_θ^2 make use of the Born and Wolf model of the 3D point spread function ([83]) for which the phase aberration term W_{z_0} is given by $W_{z_0}(\rho) := \frac{\pi n_a^2 z_0}{\lambda n_{oil}}\rho^2$ for $\rho \in [0,1]$. The constant c specifies the fraction of photons detected at focal plane 2 relative to focal plane 1. Here the value of c is set to during curve fitting.

4.5.4.1 Magnification correction

In a multifocal plane microscope the lateral magnifications of the focal planes are different ([76]). To determine the magnification for each plane, an experiment was carried out where z-stack images of 100 nm tetraspeck fluorescent beads (Invitrogen, Carlsbad, CA) were acquired on the mutlifocal plane microscope setup. One of the focal planes was arbitrarily chosen as reference and for each focal plane the frame that contains the in-focus image of the beads was selected. Then for a given image, the X-Y location of arbitrarily selected beads was determined by independently fitting a 2D Airy profile to their image and the distance between the two beads was then determined. The ratio of the distance between the two beads in a given focal plane to the distance between the same two beads in the reference plane was calculated. This was repeated for several bead pairs and the average of all the ratios provided the (de)magnification factor for that focal plane.

4.5.5 Fisher information matrix for a MUM setup

In a multifocal plane microscope, N distinct planes within the specimen are imaged and the acquired data consists of N images. The N images can be assumed to be statistically independent of each other. Therefore, the analytical expression of the Fisher information matrix corresponding to a general parameter estimation problem in a multifocal plane microscope is given by

$$\mathbf{I}_{tot}(\theta) := \mathbf{I}_{plane_1}(\theta) + \cdots + \mathbf{I}_{plane_N}(\theta), \quad \theta \in \Theta, \tag{4.3}$$

where $\mathbf{I}_{plane_k}(\theta)$, $k = 1, \ldots, N$, denotes the Fisher information matrix pertaining to the data acquired from the k^{th} plane and the expression for $\mathbf{I}_{plane_k}(\theta)$ is analogous to that given for a conventional microscope. In the present work, the 3D location estimation for quantum dots is carried out by simultaneously imaging two distinct planes within the specimen. For this

configuration, eq. 4.3 becomes $\mathbf{I}_{tot}(\theta) := \mathbf{I}_{plane_1}(\theta) + \mathbf{I}_{plane_2}(\theta),\ \theta \in \Theta$, where the general expression for $\mathbf{I}_{plane_1}(\theta)$ and $\mathbf{I}_{plane_2}(\theta)$ is given by

$$\mathbf{I}_{plane_j}(\theta) = \sum_{k=1}^{N_p} \left(\frac{\partial \mu_\theta^j(k,t)}{\partial \theta} \right)^T \left(\frac{\partial \mu_\theta^j(k,t)}{\partial \theta} \right) \times$$

$$\left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta^j(k,t)]^{l-1} e^{-\nu_\theta^j(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}^j} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k^j}{\sigma_{w,k}^j}\right)^2} \right)^2}{\frac{1}{\sqrt{2\pi}\sigma_{w,k}^j} \sum_{l=0}^{\infty} \frac{[\nu_\theta^j(k,t)]^{l} e^{-\nu_\theta^j(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k^j}{\sigma_{w,k}^j}\right)^2}} dz - 1 \right), \tag{4.4}$$

where $\nu_\theta^j(k,t) := \mu_\theta^j(k,t) + \beta^j(k,t),\ k = 1,\ldots,N_p,\ \theta \in \Theta,\ t \geq t_0$ and $j = 1,2$. Here, $[t_0, t]$ denotes the exposure time interval, $\mu_\theta^1(k,t)$ and $\mu_\theta^2(k,t)$ ($\beta^1(k,t)$ and $\beta^2(k,t)$) denote the mean detected photon count of the object of interest (background component) at the k^{th} pixel in the image of focal plane 1 and focal plane 2, respectively, $k = 1,\ldots,N_p,\ t \geq t_0$, and η_k^1 and η_k^2 ($\sigma_{w,k}^1$ and $\sigma_{w,k}^2$) denote the mean (standard deviation) of the readout noise at the k^{th} pixel in the image of focal plane 1 and focal plane 2, respectively.

4.5.5.1 Calculation of the Fisher information matrix

In this section we present results for the calculation of the Fisher information matrix for a multifocal plane microscope that can image up to two distinct planes (N = 2). The analytical expressions for μ_θ^1 and μ_θ^2 are given by

$$\mu_\theta^1(k,t) = \frac{\Lambda_0 t}{M_1^2} \int_{C_k} q_{z_0}\left(\frac{x}{M_1} - x_{01}, \frac{y}{M_1} - y_{01} \right) dxdy, \quad k = 1,\ldots,N_p, \quad t \geq t_0, \quad \theta \in \Theta,$$

$$\mu_\theta^2(k,t) = \frac{c\Lambda_0 t}{M_2^2} \int_{C_k} q_{z_0 - \delta z_f}\left(\frac{x}{M_2} - x_{02}, \frac{y}{M_2} - y_{02} \right) dxdy, \quad k = 1,\ldots,N_p, \quad t \geq t_0, \quad \theta \in \Theta,$$

where q_{z_0} denotes the image function of the object, c is a constant, Λ_0 denotes the photon detection rate of the object t denotes the exposure time, δz_f denotes the plane spacing between the two focal planes in the object space, and M_1 and M_2 denote the lateral magnification in the two focal planes. The image of a self-luminous point source (e.g., fluorescent single particle)

that is located at $(0, 0, z_0)$ in the object space and imaged by a fluorescence microscope is given by ([84])

$$I_{z_0}(x, y) = \left| C \int_0^1 J_0 \left(\frac{2\pi n_a}{\lambda} (\sqrt{x^2 + y^2}) \rho \right) \exp(j W_{z_0}(\rho)) \rho d\rho \right|^2, \tag{4.5}$$

where $(x, y) \in \mathbb{R}^2$ denotes an arbitrary point on the detector plane, C is a constant with complex amplitude, λ denotes the wavelength of the detected photons, n_a denotes the numerical aperture of the objective lens, J_0 denotes the zeroth order Bessel function of the first kind and $W_{z_0}(\rho)$, $\rho \in [0, 1]$, denotes the phase aberration term. We note that eq. 4.5 provides a general expression for several 3D point spread function models ([84]) which describe the image of a point-source/single-molecule and are based on scalar diffraction theory. Rewriting eq. 4.5 in terms of an image function, we have

$$q_{z_0}(x, y) = \frac{1}{C_{z_0}} \left(U_{z_0}^2(x, y) + V_{z_0}^2(x, y) \right), \quad (x, y) \in \mathbb{R}^2, \quad z_0 \in \mathbb{R}, \tag{4.6}$$

where

$$\begin{aligned} U_{z_0}(x, y) &:= \int_0^1 J_0 \left(\frac{2\pi n_a}{\lambda} (\sqrt{x^2 + y^2}) \rho \right) \cos(W_{z_0}(\rho)) \rho d\rho, \quad (x, y) \in \mathbb{R}^2, \quad z_0 \in \mathbb{R}, \\ V_{z_0}(x, y) &:= \int_0^1 J_0 \left(\frac{2\pi n_a}{\lambda} (\sqrt{x^2 + y^2}) \rho \right) \sin(W_{z_0}(\rho)) \rho d\rho, \quad (x, y) \in \mathbb{R}^2, \quad z_0 \in \mathbb{R}, \\ C_{z_0} &= \int_{\mathbb{R}^2} (U_{z_0}^2(x, y) + V_{z_0}^2(x, y)) dx dy, \quad z_0 \in \mathbb{R}. \end{aligned} \tag{4.7}$$

In the above equation U_{z_0} (V_{z_0}) denotes the real (imaginary) part of I_{z_0} given in eq. 4.5. The term C_{z_0} is the normalization constant, and the $1/C_{z_0}$ scaling in eq. 4.6 ensures that

$$\frac{1}{M^2} \int_{\mathbb{R}^2} q_{z_0} \left(\frac{x}{M} - x_0, \frac{y}{M} - y_0 \right) dx dy = 1, \quad (x_0, y_0, z_0) \in \Theta,$$

where M denotes the lateral magnification of the objective lens. Through numerical calculations, it was found that the normalization constant is approximately equal to the ratio π/α^2, where $\alpha = 2\pi n_a/\lambda$ for z_0 values in the range of $[0, 2\mu m]$.

Although, not shown explicitly, it can be verified that q_{z_0} and the partial derivative of q_{z_0} with respect to z_0 are laterally symmetric along the x and y axes with respect to $(0, 0)$, for $z_0 \in \mathbb{R}$.

To calculate the 3D point spread function, we require an explicit analytical expression for the phase aberration term W_{z_0} and here, we set W_{z_0} to be

$$W_{z_0}(\rho) := \frac{\pi(n_a)^2 z_0}{n_{oil}\lambda}\rho^2, \quad \rho \in [0,1], \quad z_0 \in \mathbb{R}, \tag{4.8}$$

where n_a denotes the numerical aperture of the objective lens, n_{oil} denotes the refractive index of the immersion oil and z_0 denotes the axial coordinate of the single molecule in the object space. The above expression for W_{z_0} corresponds to the classical 'Born and Wolf' 3D point spread function model ([56]).

To calculate the Fisher information matrix, we also require the partial derivatives of μ_θ^j, $j = 1, 2$, with respect to the components of θ and and these are given below.

$$\frac{\partial \mu_\theta^1(k,t)}{\partial x_{01}} = \frac{2\alpha^3 At}{\pi M_1}\times$$

$$\left[\iint\limits_{C_k} \frac{x - M_1 x_{01}}{r_1}\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_1(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\rho^2 d\rho\right)dxdy\right.$$

$$\left. + \iint\limits_{C_k} \frac{x - M_1 x_{01}}{r_1}\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)sin(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_1(\frac{\alpha}{M_1}r_1\rho)\sin(W_{z_0}(\rho))\rho^2 d\rho\right)dxdy\right],$$

$$\frac{\partial \mu_\theta^1(k,t)}{\partial y_{01}} = \frac{2\alpha^3 At}{\pi M_1}\times$$

$$\left[\iint\limits_{C_k} \frac{y - M_1 y_{01}}{r_1}\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_1(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\rho^2 d\rho\right)dxdy\right.$$

$$\left. + \iint\limits_{C_k} \frac{y - M_1 y_{01}}{r_1}\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)sin(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_1(\frac{\alpha}{M_1}r_1\rho)\sin(W_{z_0}(\rho))\rho^2 d\rho\right)dxdy\right],$$

$$\frac{\partial \mu_\theta^1(k,t)}{\partial z_0} = -\frac{2\alpha^3 At}{\pi M_1}\times$$

$$\left[\iint\limits_{C_k}\left(\int_0^1 J_0(\frac{\alpha}{M}r_1\rho)\cos(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)\sin(W_{z_0}(\rho)\frac{\partial W_{z_0}(\rho)}{\partial z_0}\rho d\rho\right)dxdy\right.$$

$$\left. - \iint\limits_{C_k}\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)sin(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\frac{\partial W_{z_0}(\rho)}{\partial z_0}\rho d\rho\right)dxdy\right],$$

$$\frac{\partial \mu_\theta^1(k,t)}{\partial \alpha} = \frac{2}{\alpha}(\mu_\theta^1(k,t) - B_1) - \frac{2\alpha^2 At}{\pi M_1^3}\times$$

$$\left[\iint\limits_{C_k} r_1\left(\int_0^1 J_0(\frac{\alpha}{M_1}r_1\rho)\cos(W_{z_0}(\rho))\rho d\rho\right)\left(\int_0^1 J_1(\frac{\alpha}{M}r_1\rho)\cos(W_{z_0}(\rho))\rho^2 d\rho\right)dxdy\right.$$

$$+\iint_{C_k} r_1 \left(\int_0^1 J_0(\frac{\alpha}{M_1} r_1 \rho) sin(W_{z_0}(\rho)) \rho d\rho\right) \left(\int_0^1 J_1(\frac{\alpha}{M_1} r_1 \rho) \sin(W_{z_0}(\rho)) \rho^2 d\rho\right) dxdy\Bigg],$$

$$\frac{\partial \mu_\theta^1(k,t)}{\partial A} = \frac{1}{A} \mu_\theta^1(k,t),$$

where $\theta = (x_0, y_0, z_0) \in \Theta$, $k = 1, \ldots, N_p$, $\alpha := 2\pi n_a/\lambda$, $r_1 := \sqrt{(x - M_1 x_{01})^2 + (y - M_1 y_{01})^2}$, $W_{z_0}(\rho) := \frac{\pi n_a^2 \rho^2 z_0}{\lambda n_{oil}}$, $\rho \in [0,1]$, and $\frac{\partial W_{z_0}(\rho)}{\partial z_0} := \frac{\pi n_a^2 \rho^2}{\lambda n_{oil}}$, $\rho \in [0,1]$. The expression for the partial derivative of μ_θ^2 is analogous to that of μ_θ^1 expect that in the above equations M_1 is replaced by M_2, x_{01} and y_{01} are replaced by x_{02} and y_{02}, respectively, r_1 is replaced by r_2, where $r_2 := \sqrt{(x - M_2 x_{02})^2 + (y - M_2 y_{02})^2}$, z_0 is replaced by $z_0 - \delta z_f$, and A is replaced by cA.

CHAPTER 5

REDEFINING THE RESOLUTION LIMITS IN OPTICAL MICROSCOPY

5.1 Introduction

Rayleigh's resolution criterion, although extensively used, is well known to be based on heuristic notions and is inadequate for current microscopy techniques. This inadequacy has necessitated a reassessment of the resolution limits of optical microscopes. By using the general results given in Chapter 2, a new resolution measure is obtained that overcomes the shortcomings of Rayleigh's criterion and provides a quantitative measure of a microscope's ability to determine the distance between two point sources. The new result is referred to as the fundamental resolution measure (FREM). Unlike Rayleigh's criterion, the FREM predicts that the resolution of a microscope can be improved by increasing the number of photons collected from the point sources. The FREM is given in terms of quantities such as the expected number of detected photons, the numerical aperture of the objective lens and the wavelength of the detected photons. The effect of various experimental factors on the FREM is also investigated. The new resolution measure is experimentally verified by measuring distances of closely spaced single molecules. These results show that distances well below Rayleigh's resolution limit can be determined with an accuracy as specified by the new resolution measure.

The organization of this chapter is as follows. Section 5.2 discusses the FREM, which provides the best case scenario for resolving two infocus point sources that emit incoherent, unpolarized light. Section 5.3 discusses how deteriorating experimental factors such as pixelation of the detector and noise sources affect the FREM. The derivation of the FREM assumes that the acquired data contains photons detected from both single molecules. In many situations, the single molecule pair exhibits double-step photobleaching behavior. In such cases, the distance of separation between the two single molecules can be estimated with relatively high accuracy when using the data collected after the first photobleaching step. In Section 5.4,

a detailed analysis is presented, which shows how the use of additional information (i.e., data collected after the first photobleaching step) improves the resolution measure. Section 5.5 deals with the experimental verification of the new resolution measure. The derivation of the FREM in Section 5.2 considers a specific imaging condition, i.e., two infocus point sources that emit unpolarized incoherent light. Section 5.6 deals with the generalization of the FREM (g-FREM) and an analytical expression is given which is applicable to a wide variety of imaging scenarios. Section 5.7 discusses the implications of the new resolution measure for single molecule experiments. Specific examples are provided to illustrate how the new result can be used to evaluate the feasibility of carrying out single molecule imaging studies that involve distance determination between two single molecules.

The detailed derivation of the results given in Sections 5.2 - 5.6 can be found in the Appendix at the end of this chapter. All the results presented in this chapter have been published in [52].

5.2 Fundamental resolution measure (FREM)

Our approach to the derivation of the resolution measure is to obtain a bound/limit to the accuracy with which the distance between two point sources can be estimated based on the acquired data. Analogous to Rayleigh's criterion, we consider an optical microscope setup that images two identical, self-luminous, in-focus point sources emitting unpolarized, incoherent light. The analytical expression of the fundamental resolution measure (FREM) for this imaging condition is given by (see Chapter Appendix for derivation)

$$\delta_d := \frac{1}{\sqrt{4\pi \cdot \Lambda_0 \cdot (t - t_0) \cdot \Gamma_0(d)}} \cdot \frac{\lambda}{n_a}, \qquad \textbf{(FREM)} \qquad (5.1)$$

where λ denotes the emission wavelength of the detected photons, n_a denotes the numerical aperture of the objective lens, Λ_0 denotes the photon detection rate (intensity) per point source,

$[t_0, t]$ denotes the acquisition time interval, and $\Gamma_0(d)$ is given by (see Appendix at the end of this chapter for derivation)

$$\Gamma_0(d) := \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha r_{01})}{r_{01}^2} + \frac{J_1^2(\alpha r_{02})}{r_{02}^2}} \times$$

$$\left((x + \frac{d}{2}) \frac{J_1(\alpha r_{01}) J_2(\alpha r_{01})}{r_{01}^3} - (x - \frac{d}{2}) \frac{J_1(\alpha r_{02}) J_2(\alpha r_{02})}{r_{02}^3} \right)^2 dxdy, \tag{5.2}$$

with $r_{01} := \sqrt{(x + d/2)^2 + y^2}$, $r_{02} := \sqrt{(x - d/2)^2 + y^2}$, J_n denoting the n^{th} order Bessel function of the first kind and $\alpha := 2\pi n_a/\lambda$. According to Rayleigh's criterion, the minimum resolvable distance between two point sources is given by $0.61\lambda/n_a$. The FREM, on the other hand, provides a more complex expression, which, in addition to the dependence on the ratio λ/n_a, exhibits an inverse square root dependence on other factors, i.e., the expected number of detected photons $(\Lambda_0 \cdot (t - t_0))$ and the term $\Gamma_0(d)$ given by eq. 5.2. Note that the FREM depends on the distance of separation d through the term $\Gamma_0(d)$. Moreover, the presence of the ratio λ/n_a in $\Gamma_0(d)$ through the term α $(= 2\pi n_a/\lambda)$ shows that the FREM exhibits a non-linear dependence on λ/n_a.

The stochastic framework used to obtain the FREM models the photon emission (detection) process as a random process (shot noise process). The spatial locations at which the photons hit the detector are assumed to be randomly distributed according to the image profiles of the point-sources/single-molecules. This framework considers an optical microscope setup in which the detector provides the time points and the spatial coordinates of every detected photon without adding any extraneous noise. For any imaging condition, this can be thought of as an idealization of current imaging detectors in which the presence of finite-sized pixels and measurement noise deteriorates the acquired data. Thus the resolution measure derived within this framework provides a result that is *fundamental* for the given imaging condition. In the present context, the FREM is obtained for imaging conditions analogous to those of Rayleigh's criterion. Hence the spatial distribution of the detected photons from each point source is described by the Airy profile ([56]).

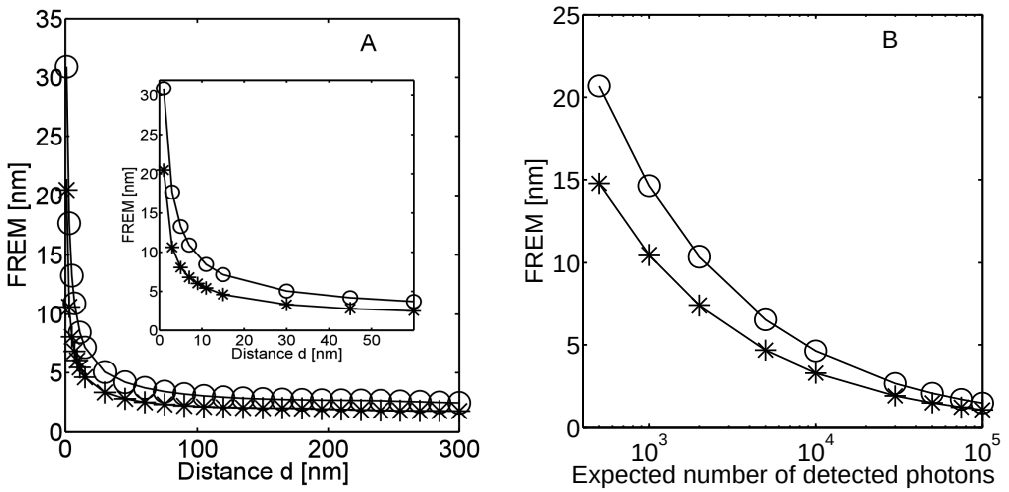

Figure 5.1. Behavior of the FREM as a function of distance and photon count. Panel A shows the FREM as a function of the distance of separation between two point-sources/single-molecules. The inset shows the same for a distance range of $1 - 50$ nm. Panel B shows the PREM as a function of the expected number of detected photons per molecule for a distance of separation of 10 nm. In both panels, the PREM is calculated for a pair of GFP molecules ($*$) and for a pair of Cy5 molecules ($\circ$). For all the plots, the numerical aperture is set to be $n_a = 1.45$ and the wavelength of the detected photons from the GFP (Cy5) molecule is set to be $\lambda = 520$ nm ($\lambda = 690$ nm). In Panel A, the photon detection rate Λ_0 of each GFP/Cy5 molecule is set to be $\Lambda_0 = 3000$ photons/s and the acquisition time is set to be 1 s.

The new resolution measure FREM predicts how accurately the distance d between two point sources can be resolved. A small numerical value for the FREM predicts a high accuracy in determining d, while a large numerical value of the FREM predicts a low accuracy in determining d. Fig. 5.1A shows the behavior of the FREM as a function of the distance of separation between a pair of GFP molecules ($\lambda = 520$ nm) and for a pair of Cy5 molecules ($\lambda = 690$ nm) that are imaged with an objective lens of numerical aperture 1.45. In the figure, it is assumed that the expected photon count is the same for both the fluorophores. For the GFP molecules, Rayleigh's criterion predicts the smallest resolvable distance to be about 220 nm ($\approx 0.61\lambda/n_a$). In contrast, Fig. 5.1A shows that the FREM has a small numerical value for distances in the range of 50 - 220 nm, which are well below Rayleigh's criterion. For distances less than 50 nm, however, the FREM deteriorates (i.e. increases) significantly with decreasing distance of separation (see Fig. 5.1A inset). In particular, as the distance of separation decreases to zero, the FREM

becomes infinitely large, since the term $\Gamma_0(d)$ (eq. 5.2), which appears in the denominator of the FREM, tends to zero. An analogous behavior of the FREM is also seen for the Cy5 molecules. Note that the numerical value of the FREM for the Cy5 molecules is consistently larger than that of the GFP molecules for the same expected photon count per fluorophore. For example, the FREM predicts that a distance of 10 nm between two GFP molecules can be determined with an accuracy not better than ±6 nm when the expected photon count per GFP molecule is 3000. On the other hand, for the same distance of separation and photon count per molecule, the FREM predicts an accuracy not better than ±9 nm for the Cy5 molecules. In the case of the Cy5 molecules, however, the numerical value of the FREM is comparable to the distance of separation itself. Since the FREM exhibits an inverse square root dependence on the expected number of detected photons, this deterioration can be compensated for by increasing the expected number of detected photons, as shown in Fig. 5.1B. Thus in the above example, if we increase the expected photon count per Cy5 dye molecule to 10^4, then the FREM predicts that a distance of 10 nm can be determined with an accuracy not better than ±5 nm.

5.3 Practical resolution measure (PREM)

The FREM provides the best case scenario for a microscope setup, where experimental factors that potentially deteriorate the acquired data were not taken into account. We next investigate how the resolution measure is affected by such experimental factors. Here we obtain an analytical expression for the resolution measure that takes into account these experimental factors. We refer to this result as the practical resolution measure (PREM). The PREM can be thought of as an extension to the FREM. For instance, the PREM takes into account the presence of additive noise sources, namely Poisson and Gaussian noise. Poisson noise is used to model the spurious photons in the acquired image which, for example, arises due to autofluorescence of the sample and dark current of the detector ([31]). Gaussian noise is used to model the measurement noise in the acquired data which, for example, arises during the readout process in the detector ([31]). The additive Poisson noise considered here is distinct

from the shot noise, which describes the statistics of the photon detection process from the single molecules and this is already accounted for by the FREM. Aside from these extraneous noise sources, the PREM also takes into account the effect of pixelation of the detector (see Chapter Appendix for the details).

Fig. 5.2A shows the behavior of the PREM as a function of the distance between two Cy5 molecules in the presence and absence of noise sources for a pixelated detector. The figure also shows the FREM for reference. Note that even in the absence of extraneous noise sources the numerical value of the PREM is consistently greater than that of the FREM due to the pixelation of the detector. Moreover, in the presence of noise sources this behavior of the PREM becomes more pronounced. In particular, for very small distances ($\leq$ 50 nm), the numerical value of the PREM is at least 3-5 times greater than that of the FREM (see Fig. 5.2A inset). Analogous to Fig. 5.1B, the deterioration of the PREM at very small distances can be compensated for by collecting more photons from the point sources (Fig. 5.2B). In contrast, for distances in the range of 100 - 250 nm, which are below Rayleigh's criterion ($\approx 0.61\lambda/n_a \approx 290$ nm), the numerical value of the PREM approaches that of the FREM even in the presence of noise, as shown in Fig. 5.2A. As an application of these results, consider a practical scenario in which we require distances in the range of 50 nm - 200 nm to be resolved between two Cy5 molecules with an accuracy of at least 5 nm. From Fig. 5.2B, we know that to estimate a distance of 50 nm with an accuracy not better than 5 nm, the PREM predicts the expected number of detected photons per single molecule to be at least 15000. On the other hand, from Fig. 5.2A we see that to estimate a distance of 200 nm with similar accuracy, the PREM predicts the expected number of detected photons per single molecule to be at least 2500. Hence to resolve distances in the range of $50 - 200$ nm between two Cy5 molecules with an accuracy not better than 5 nm, on average at least 15000 photons must be collected per single molecule.

5.4 Improving the resolution measure by using additional spatial information

It was shown in Figs. 5.1B and 5.2B that the resolution measure can be improved by increasing the number of detected photons from each of the point sources. In single molecule experiments this is not always possible, since the fluorophores may photobleach. However, for a single-molecule pair that exhibits a double-step photobleaching behavior, additional information can be obtained from the photons collected from the fluorophore that remains after the first photobleaching event ([48]). Fig. 3.2C shows the behavior of the resolution measure for a pair of Cy5 molecules spaced 10 nm apart by taking into account the number of photons collected before and after the first photobleaching event. Here the resolution measure is determined for an imaging condition with numerical values analogous to those used in Fig. 3.2A. From the figure we see that the resolution measure predicts an accuracy not smaller than ± 6 nm to determine a distance of 10 nm, when on average 5000 photons are collected from each fluorophore before and after the first photobleaching event. This is in contrast to the case when the additional information obtained from after the first photobleaching event is not used. In this case, the resolution measure predicts an accuracy not smaller than ± 40 nm to resolve a distance of 10 nm for the same photon count per fluorophore.

5.5 Experimental verification

The resolution measure provides a bound/limit to the smallest possible standard deviation of any unbiased estimator of the distance between the point sources. To verify if this can be attained in experiments, images of closely spaced Cy5 molecules were collected and their distances of separation were estimated by using the maximum likelihood estimator. According to Rayleigh's criterion, the minimum resolvable distance is given by $0.61\lambda/n_a$, which, in the present case is about 290 nm. Table 5.1 lists the results of distance estimation along with the predicted resolution measure for two pairs of single molecules. One of the single molecule pairs has a mean distance of separation of 293 nm (data analysis 1), which is close to Rayleigh's criterion, and the other single molecule pair has a mean distance of separation of 207 nm (data

analysis 2), which is below Rayleigh's criterion (see Fig. 5.3). From the Table we see that for each data analysis the standard deviation of the maximum likelihood estimates of the distance comes close to the resolution measure. Note that the numerical values of the standard deviations are themselves only estimates based on the acquired data. With larger data sets, the agreement with the resolution measure is expected to increase further.

The above data sets were also analyzed by estimating the distances of separation through the global maximum likelihood estimator, which used the additional spatial information available in the acquired data, i.e. the images collected after the first photobleaching event. Table 5.1 lists the results of the distance estimates (data analyses 3 and 4) for the Cy5 single-molecule pairs analyzed above. From the table we see that the accuracy of the distance estimates obtained by using the additional spatial information is consistently smaller than the accuracy obtained when the additional information is not used (data analyses 1 and 2). This is also in agreement with the resolution measure for the data sets. For example, in the case when the additional spatial information is not used, the standard deviation of the distance estimates for the Cy5 single-molecule pair with a mean distance of separation of 207 nm is equal to ± 10.1 nm (data analysis 2). On the other hand, for the same single-molecule pair, when additional spatial information is used the standard deviation of the distance estimates is equal to ± 3.8 nm (data analysis 4). Table 5.1 also lists the mean end-to-end distance estimates of the DNA molecular ruler, which are determined by using the global estimation approach (also see Fig. 5.4). From the table it can be seen that for each DNA data set the standard deviation of the global maximum likelihood estimator comes close to the resolution measure. For example, in data analysis 6 the standard deviation of the distance estimates is equal to ± 8.7 nm and the resolution measure predicts an accuracy not smaller than ± 7.39 nm to resolve the distance of 12 nm. We note that in the same data set, if the additional spatial information is not used, then the resolution measure predicts an accuracy not smaller than ± 52 nm to resolve a distance of 12 nm.

Table 5.1. Experimental verification of the new resolution measure

Data analysis	Data set no.	Exposure time (s)	Estimation method	N_1	N_2	Mean value of d (nm)	Std-dev of d (nm)	Resolution measure for a pixelated detector (nm)
1	1	3	direct	1	0	293	3.60	2.81
2	2	1	direct	1	0	207	10.10	7.01
3	1	3	global	3	3	293	1.87	1.36
4	2	1	global	2	2	211	3.80	4.41
5	3	1	global	6	6	12.5	5.20	6.85
6	4	1	global	4	4	14.8	8.77	7.39

Data sets 1 and 2 correspond to two closely spaced Cy5 molecules, and data sets 3 and 4 correspond to the DNA molecular ruler. The experimental data used for estimating the distances consists of time-lapse images of single-molecule pairs that exhibit a double-step photobleaching behavior. In the direct estimation method each distance estimate is obtained from an image that is acquired before the first photobleaching event. In the global estimation method, each distance estimate is obtained from two summed images. One of the summed images is obtained by adding N_1 frames that are acquired before the first photobleaching event, and the other summed image is obtained by adding N_2 frames that are acquired after the first photobleaching event. The data sets used in analyses 1 and 3 (2 and 4) are the same. For each data analysis, the resolution measure is calculated for a pixelated detector in the presence of noise sources.

5.6 Generalization of the FREM

The FREM given in eq. 5.1 was derived for imaging conditions analogous to those of Rayleigh's criterion, which considered two equal intensity, in-focus point sources that emit unpolarized, incoherent light. However, in several applications these conditions are not met, for example, when using polarized illumination and detection ([9]). We now consider a situation where the point sources can potentially have unequal intensities that vary as a function of time, and where the image profiles of the point sources can be distinct. The expression for the 'generalized' fundamental resolution measure (**g-FREM**) is given by (see Appendix at the end of this chapter for derivation)

$$\left[\frac{1}{4}\int_{t_0}^{t}\int_{\mathbb{R}^2}\frac{1}{\Lambda_1(\tau)q_1(x+\frac{d}{2},y)+\Lambda_2(\tau)q_2(x-\frac{d}{2},y)}\times\right.$$

$$\left(\Lambda_1(\tau) \frac{\partial q_1(x+\frac{d}{2},y)}{\partial x} - \Lambda_2(\tau) \frac{\partial q_2(x-\frac{d}{2},y)}{\partial x} \right)^2 dxdyd\tau \Bigg]^{-\frac{1}{2}}, \tag{5.3}$$

where q_1 and q_2 denote the image functions of the point sources and Λ_1 and Λ_2 denote the intensities of the point sources. In many situations the image of the point source significantly differs from the Airy profile, for example, due to the defocus in the objective lens ([85]), or due to the different orientations of the point-source emission dipole ([9, 86]) or due to the aberrations present in the imaging setup ([87]). Moreover, depending upon the nature of illumination, the intensity of the point sources can be unequal when their emission dipole orientations are different ([9, 86]). We note that eq. 5.3 provides a general expression for the FREM that is applicable to a wide variety of imaging conditions including the above mentioned scenarios. If in the above equation we set the intensities to be constant and identical, i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$, and assume the image functions to be given by the Airy profile, then we immediately obtain the expression for the FREM given in eq. 5.1. Analogous to the g-FREM, an expression has also been derived for the 'generalized' practical resolution measure (see Chapter Appendix).

The stochastic framework used to derive the g-FREM models the photon detection process for each point source as a Poisson process (shot noise process). Recently, there have been reports of the generation of non-classical states of light from a fluorescent light source in which the photon statistics deviate from the classical shot noise process ([88, 89, 90]). In order to take into account this deviation, a further generalization of the FREM can be obtained by modeling the photon detection process as a general counting process ([60], see Chapter Appendix).

5.7 Discussion

The advent of single molecule microscopy has generated significant interest in studying nanoscale biomolecular interactions. Classically, fluorescence resonance energy transfer based methods have been used to probe interactions in the distance range of $1 - 10$ nm ([44]). It is widely believed that Rayleigh's criterion precludes the resolution of two single molecules at distances of less than 200 nm. This leaves a gap in the distance range of $10 - 200$ nm that

is vital for the study of many biological processes with an optical microscope. It has been suggested that Rayleigh's resolution limit can be superseded if the distance between two point sources is determined by curve fitting the image with the sum of two point-source image profiles ([46]). In fact, by adopting this approach several groups have shown that Rayleigh's limit can be surpassed in experiments ([48, 49, 50]).

The FREM (eq. 5.1) is a resolution measure that overcomes several deficiencies of Rayleigh's criterion. It gives a bound for the accuracy with which the distance between two point sources can be estimated when the acquired data is not affected by deteriorating experimental factors. An important property of the FREM is that it provides a quantitative assessment of how the optical characteristics of the experimental setup and the photon budget influence the resolution performance in determining a particular distance of separation. Fig. 5.1A shows that the numerical value of the FREM for a pair of GFP molecules is consistently smaller than that for a pair of Cy5 molecules when the expected photon count per fluorophore is 3000 in both cases. For example, to resolve distances of 8 nm, 50 nm and 200 nm between a pair of GFP molecules, the FREM predicts an accuracy not smaller than ± 6.5 nm, ± 2.7 nm and ± 1.9 nm, respectively. On the other hand, for a pair of Cy5 molecules with the same expected photon count per fluorophore and distances, the FREM predicts an accuracy not smaller than ± 10 nm, ± 4 nm and ± 2.67 nm, respectively. For the 50 nm and 200 nm distances, the numerical value of the FREM for the GFP and the Cy5 molecules are significantly smaller than the corresponding actual distances. This implies that the FREM predicts a relatively high accuracy in resolving distances in the range of $50 - 200$ nm between single molecules. For the 8 nm distance, however, the numerical value of the FREM for the GFP and the Cy5 molecules are either comparable to or greater than the actual distance. This suggests that even in the best case scenario, i.e. in the absence of deteriorating experimental factors, distances of less than 8 nm are difficult to resolve between the GFP/Cy5 molecules, unless a higher than average number of photons are detected. In single molecule experiments typically 3000 photons can be collected before a GFP molecule irreversibly photobleaches (see e.g., [26, 91]), while for

Cy5 molecules typically greater than 10^4 photons can be collected (see [92]). Thus, to resolve distances of 8 nm, 50 nm and 200 nm with an expected photon count of 10^4 per Cy5 molecule, the FREM predicts an accuracy not smaller than ± 5.5 nm, ± 2.2 nm and ± 1.5 nm, respectively, i.e. about a two-fold improvement from the case when on average 3000 photons are collected per Cy5 molecule. Note that the localization accuracy of a single molecule, i.e. the accuracy with which the position of a single molecule can be determined also depends on the number of collected photons ([28, 51]).

The practical resolution measure (PREM) derived here extends the results of the FREM by illustrating how the resolution measure is deteriorated by experimental factors such as pixelation of the detector and extraneous noise sources. The PREM, i.e. the bound on the accuracy with which the distance can be estimated, for typical imaging conditions (see Fig 2A) is given by ± 31.6 nm, ± 5.3 nm and ± 2.2 nm for the case when the Cy5 single molecules are 8 nm, 50 nm and 200 nm apart, respectively, and the expected photon count per fluorophore is 10^4. For the same distances, if the expected photon count is 3000 per Cy5 molecule, the PREM is significantly higher at ± 76.5 nm, ± 12.5 nm and ± 4.5 nm, respectively. Similarly, for a pair of GFP molecules with an expected photon count of 3000 per molecule, the PREM predicts an accuracy not smaller than ± 42 nm, ± 7.4 nm and ± 3 nm to resolve distances of 8 nm, 50 nm and 200 nm, respectively. This shows that especially for small distances the predicted resolution measure is probably not acceptable for many applications. This deterioration in the limit of the accuracy with which the distance can be measured is due to the fact that the data acquired by a pixelated detector is a discretized version of the actual image, and the presence of extraneous noise sources corrupts the acquired data (e.g. scattered photons, noise in the acquisition electronics). Moreover, a comparison with the FREM illustrates that control of the noise sources is also of great importance to improve the accuracy of the estimated distance parameter in a practical scenario (Fig. 5.2A).

The above results suggest that the distance between two single molecules can be estimated with a reasonable level of accuracy, depending on the photon count, certainly for distances

above 50 nm, but possibly also for smaller distances. For distances around 10 nm, however, the predicted resolution measures are typically worse. To be able to resolve such distances the number of detected photons would have to be increased substantially, which is typically not possible due to photobleaching. For instance, if a PREM of 7 nm is to be achieved for a Cy5 single molecule pair spaced 10 nm apart, then under the noisy imaging conditions of Fig. 5.2A at least 1.5×10^5 photons need to be detected per Cy5 molecule. This underscores the importance of the development of brighter and more photostable fluorescent markers to carry out such studies ([9, 93]).

In [94] it was shown that GFP single molecule pairs typically photobleach together. However, in the case where two step photobleaching occurs, additional information can be used by imaging the remaining single molecule ([48]). The development of the FREM and the PREM discussed so far was based on the case when both point sources/single molecules do not bleach during the acquisition of the image. The approach can, however, also be applied to investigate the accuracy that can be achieved when additional information is taken into account from the remaining single molecule that did not bleach in the first photobleaching step. For example, consider a pair of GFP molecules that are 8 nm, 15 nm, 25 nm or 50 nm apart. Assume that 2000 photons are collected from the single molecules before and after the first photobleaching event. If the photons collected before and after the first photobleaching event are taken into account, and assuming the experimental conditions for pixelation and noise sources of Fig. 2A, then an accuracy of no better than ± 8.1 nm, ± 7.9 nm, ± 7.6 nm, and ± 6.7 nm, respectively, can be expected for these distances. Under the present assumptions on the expected number of detected photons, this shows that for distances up to 15 nm probably unreasonably large errors would be incurred in the estimation. However, for distances above 25 nm an error level of less than 30% could be expected.

The resolution measure provides a bound to the accuracy/standard-deviation with which the distance between two point sources can be estimated. This was experimentally verified by imaging closely spaced Cy5 molecules and estimating their distance of separation from the

acquired data. Here, the maximum likelihood estimator was used, since it possesses favorable properties for estimating parameters ([55]). In general, other estimation algorithms can also be used for determining the distance of separation between two point sources. However, a question that arises is which of the different estimation algorithms is the most suitable for estimating the distance of separation. In such a scenario, the knowledge of the resolution measure becomes crucial, since it can be used as a standard to compare the performance of the different estimation algorithms. The experimental results presented in Table 5.1 show that the standard deviations of the maximum likelihood distance estimates come close to the bound predicted by the resolution measure, thereby validating the choice of this estimator.

The FREM was derived for imaging conditions that were analogous to those assumed in Rayleigh's criterion. In some single molecule experiments, however, the conditions are different to those assumed in the derivation of Rayleigh's criterion and which formed the basis for the derivation of the FREM. Whereas the FREM assumes that the image of a point source is given by an Airy profile, the generalized FREM (g-FREM) was derived so that more complex image profiles can be analyzed. Such profiles could arise, for example, due to out of focus conditions ([85]), the presence of aberrations ([87]), or the use of polarized illumination ([9]). We note that the g-FREM can also be used to calculate the resolution measure for determining the distance of separation between any two (distinct) objects such as cellular organelles, provided the intensities and the image functions of the objects are known. The resolution limit of Rayleigh's criterion is thought to arise due to the finite width of the central peak of the point-source image, i.e. the point spread function ([16]). This led to the development of microscopy techniques such as 4Pi confocal microscopy ([95]), stimulated emission depletion (STED) microscopy ([96]) and image interference microscopy (or I^5M) ([97]) in which the width of the central peak of the point spread function is smaller than that of the conventional optical microscope. These techniques have reported improvement in resolving features that are typically unresolvable in conventional optical microscopes. However, for some of the techniques it was reported that this was achieved at a severe cost to the signal (i.e., number of collected photons) ([98]). This

illustrates the importance of considering the photon/light budget when discussing resolution performance, especially in fluorescence imaging applications that typically use photobleachable fluorophores. In the present context, the expression for the FREM/g-FREM explicitly shows the tradeoff between the intensities of the point sources, which determine the photon budget, and the image functions of the point sources, which determine the point spread function shape.

5.8 Appendix to Chapter 5

5.8.1 Single molecule microscopy

A molecular ruler consisting of a 30 base pair DNA duplex was used. The oligonucleotide $5'$-ATC TCG GTG CGT AAT ACT CAC GGG CAG GAC-$3'$ ([99] and personal communication, Dr. D. Holowka) and its complementary sequence (both labeled with Cy5 at the $5'$ end) were purchased from Synthegen (Houston, TX). The oligonucleotides were annealed in 200 mM Tris-HCl, 10 mM $MgCl_2$, pH 8.0 and stored at 4°C. Molecular modeling by the vendor of the DNA duplex labeled with Cy5 dye at both ends predicts the distance between the two Cy5 dyes to be 12 nm. Cy5 dye purchased from GE Healthcare (Piscataway, NJ) was used for calibration purposes. To image single DNA molecules, a cleaned glass bottomed dish (MatTek Corp, MA) was coated with a layer of Poly-Lysine (0.01% solution) followed by the fluorescent sample at a concentration of 1 pM. Imaging experiments were carried out in a custom setup that was built on a Zeiss Axiovert S100 fluorescence microscope. The setup consisted of a 643 nm laser (Research Electro-Optics, Boulder, CO), a cooled CCD Camera (ORCA-ER, Hamamatsu, Bridgewater, NJ), and an αPlan-FLUAR (NA 1.45, 100x) Zeiss objective lens. The sample was illuminated with circularly polarized light in wide-field mode, the camera was operated in 2×2 binning mode, and the exposure time was either 1 s or 3 s.

5.8.2 Derivation of the new resolution measure

By making use of the image detection process described in Section 2.2, we derive analytical expression for the new resolution measure. The image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$

describes the data acquired by the detector. For the resolution problem, the object of interest is a pair of point sources, Λ_θ and $f_{\theta,\tau}$ can be written as

$$\Lambda_\theta(\tau) := \Lambda_1(\tau) + \Lambda_2(\tau), \tag{5.4}$$

$$f_{\theta,\tau}(r) := \frac{\epsilon^1_\theta(\tau)}{M^2} q_1\left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M}\right) + \frac{\epsilon^2_\theta(\tau)}{M^2} q_2\left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M}\right), \tag{5.5}$$

where $r := (x, y) \in \mathbb{R}^2$, $\tau \geq t_0$, $\theta \in \Theta$, Θ denotes the parameter space and $\epsilon^i_\theta(\tau) := \Lambda_i(\tau)/\Lambda_\theta(\tau)$, $\tau \geq t_0$, $\theta \in \Theta$, $i = 1, 2$. In Eqs. 5.4 and 5.5, Λ_1 and Λ_2 denote the photon detection rate of the two point sources, M denotes the total lateral magnification of the microscope setup, d denotes the distance of separation between the point sources, and q_1 and q_2 denote the image functions of the two point sources (see Section 3.3 for definition of image function). In Eq. 5.5 we consider an arrangement (potentially after a suitable translation of the coordinate axes) in which the point sources lie along the x axis in the specimen plane and are equidistant from the origin of the coordinate axes. Substituting for $f_{\theta,\tau}$ and Λ_θ in Theorem 2.2.1 the Fisher information matrix is given by

$$\begin{aligned}
I(\theta) &= \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\frac{\Lambda_1(\tau)}{M^2} q_1\left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M}\right) + \frac{\Lambda_2(\tau)}{M^2} q_2\left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M}\right)} \times \\
&\qquad \left(\frac{\Lambda_1(\tau)}{M^2} \frac{\partial q_1\left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M}\right)}{\partial d} + \frac{\Lambda_2(\tau)}{M^2} \frac{\partial q_2\left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M}\right)}{\partial d} \right)^2 dx dy d\tau \\
&= \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau) q_1\left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M}\right) + \Lambda_2(\tau) q_2\left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M}\right)} \times \\
&\qquad \left(\frac{\Lambda_1(\tau)}{2} \frac{\partial q_1\left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M}\right)}{\partial x} - \frac{\Lambda_2(\tau)}{2} \frac{\partial q_2\left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M}\right)}{\partial x} \right)^2 dx dy d\tau \\
&= \frac{1}{4} \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau) q_1(x + \frac{d}{2}, y) + \Lambda_2(\tau) q_2(x - \frac{d}{2}, y)} \times \\
&\qquad \left(\Lambda_1(\tau) \frac{\partial q_1(x + \frac{d}{2}, y)}{\partial x} - \Lambda_2(\tau) \frac{\partial q_2(x - \frac{d}{2}, y)}{\partial x} \right)^2 dx dy d\tau.
\end{aligned} \tag{5.6}$$

Inverting Eq. 5.6 and taking the square root, we obtain the expression for the g-FREM.

5.8.2.1 Derivation of the FREM

Rayleigh's resolution criterion considers two identical, self-luminous, in-focus point sources that are imaged with a conventional wide-field optical microscope. Here we derive the expression of the fundamental resolution measure with similar assumptions. We assume the two point sources to have equal, constant intensities i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$, and identical image functions i.e., $q_1 = q_2$. According to optical diffraction theory ([56]), the image of an in-focus point source is described by the Airy profile, which is given by

$$q_i(x,y) := \frac{J_1^2(\alpha\sqrt{x^2+y^2})}{\pi(x^2+y^2)}, \quad (x,y) \in \mathbb{R}^2, \quad i = 1,2, \tag{5.7}$$

where J_1 denotes the first order Bessel function of the first kind, $\alpha := 2\pi n_a/\lambda$, n_a denotes the numerical aperture of the objective lens, and λ denotes the wavelength of the detected photons. Using the well known recurrence relations for Bessel functions (see e.g., ref. 5, pp. 17 and 18), the partial derivative of q_i with respect to x is given by $\partial q_i(x,y)/\partial x = -2\alpha x J_1(\alpha\sqrt{x^2+y^2})J_2(\alpha\sqrt{x^2+y^2})/(\pi(x^2+y^2)^{\frac{3}{2}})$, $(x,y) \in \mathbb{R}^2$, $i = 1,2$, where $\alpha = 2\pi n_a/\lambda$, and J_2 denotes the second-order Bessel function of the first kind. Substituting for $q_i(x,y)$ and $\partial q_i(x,y)/\partial x$ in Eq. 5.6 and setting $\Lambda_i(\tau) = \Lambda_0$, for $i = 1,2$, we get

$$\mathbf{I}(d) = \frac{\Lambda_0 \cdot (t-t_0)}{4} \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha\sqrt{(x+\frac{d}{2})^2+y^2})}{\pi((x+\frac{d}{2})^2+y^2)} + \frac{J_1^2(\alpha\sqrt{(x-\frac{d}{2})^2+y^2})}{\pi((x-\frac{d}{2})^2+y^2)}} \times$$

$$\left[-2\alpha(x+\frac{d}{2}) \frac{J_1(\alpha\sqrt{(x+\frac{d}{2})^2+y^2})J_2(\alpha\sqrt{(x+\frac{d}{2})^2+y^2})}{\pi\left((x+\frac{d}{2})^2+y^2\right)^{\frac{3}{2}}} \right.$$

$$\left. -(-2\alpha)(x-\frac{d}{2}) \frac{J_1(\alpha\sqrt{(x-\frac{d}{2})^2+y^2})J_2(\alpha\sqrt{(x-\frac{d}{2})^2+y^2})}{\pi\left((x-\frac{d}{2})^2+y^2\right)^{\frac{3}{2}}} \right]^2 dxdy$$

$$= \frac{\alpha^2 \Lambda_0 (t-t_0)}{\pi} \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha r_{01})}{r_{01}^2} + \frac{J_1^2(\alpha r_{02})}{r_{02}^2}} \times$$

$$\left[(x+\frac{d}{2}) \frac{J_1(\alpha r_{01})J_2(\alpha r_{01})}{r_{01}^3} - (x-\frac{d}{2}) \frac{J_1(\alpha r_{02})J_2(\alpha r_{02})}{r_{02}^3} \right]^2 dxdy$$

$$= \frac{4n_a^2}{\lambda^2}\pi \cdot \Lambda_0 \cdot (t - t_0) \cdot \Gamma_0(d),$$

where $r_{01} := \sqrt{(x + d/2)^2 + y^2}$, $(x, y) \in \mathbb{R}^2$, $r_{02} := \sqrt{(x - d/2)^2 + y^2}$, $(x, y) \in \mathbb{R}^2$ and Γ_0 be given by

$$\Gamma_0(d) = \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha r_{01})}{r_{01}^2} + \frac{J_1^2(\alpha r_{02})}{r_{02}^2}} \times$$

$$\left((x + \frac{d}{2})\frac{J_1(\alpha r_{01})J_2(\alpha r_{01})}{r_{01}^3} - (x - \frac{d}{2})\frac{J_1(\alpha r_{02})J_2(\alpha r_{02})}{r_{02}^3}\right)^2 dxdy.$$

Note that the Fisher information matrix $\mathbf{I}(d)$ is a scalar quantity. The FREM is obtained by taking the square root of the inverse Fisher information matrix (i.e, $1/\mathbf{I}(d)$), and is given by

$$\delta_d := \frac{1}{\sqrt{\mathbf{I}(d)}} = \frac{1}{\sqrt{4\pi \cdot \Lambda_0 \cdot (t - t_0) \cdot \Gamma_0(d)}} \frac{\lambda}{n_a}.$$

5.8.2.2 Extension to Non-Poissonian Statistics

The derivation of the Fisher information matrix given in Eq. 5.6 assumes the time points of the detected photons in the acquired data to be Poisson distributed. We next consider the scenario in which the times points of the detected photons are described by a general counting process $\{N(\tau), \tau \geq t_0\}$ that has finite first and second moment, i.e., $0 \leq E[N(\tau)], E[N^2(\tau)] < \infty$. Analogous to eq. 5.6, the spatial and temporal components of the acquired data are assumed to be independent of each other. The general expression of the FREM for the case of non-Poissonian photon statistics is then given by

$$\left[\frac{E[N(t)]}{4}\int_{\mathbb{R}^2} \frac{1}{q_1(x + \frac{d}{2}, y) + q_2(x - \frac{d}{2}, y)} \left(\frac{\partial q_1(x + \frac{d}{2}, y)}{\partial x} - \frac{\partial q_2(x - \frac{d}{2}, y)}{\partial x}\right)^2 dxdy\right]^{-\frac{1}{2}}, \quad (5.8)$$

where q_1 and q_2 denote the image functions of the two point sources and d denotes the distance of separation.

5.8.3 Effects of pixelation and noise

In the derivation of the FREM/g-FREM, it was assumed that the detector records the time points and the spatial coordinates of the detected photons, which was described by an

image detection process. However, current imaging detectors have pixels, and the acquired data only consists of the number of detected photons at each pixel. For a pixelated detector $\{C_1, \ldots, C_{N_p}\}$ with N_p pixels, the photon count at the kth pixel is independently Poisson distributed. We consider two types of additive noise sources, namely Poisson and Gaussian noise source. Poisson noise is used to model the effect of spurious light sources such as autofluorescence, and Gaussian noise is used to model measurement noise such as readout noise in the detector.

Hence the data acquired by a pixelated detector during the time interval $[t_0, t]$ is described by a sequence of independent random variables $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$, where $\mathcal{I}_{\theta,k} := S_{\theta,k} + B_k + W_k$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, and $S_{\theta,k}$, B_k and W_k are random variables such that $\{S_{\theta,1}, \ldots, S_{\theta,N_p}\}$, $\{B_1, \ldots, B_{N_p}\}$ and $\{W_1, \ldots, W_{N_p}\}$ are mutually independent and independent of each other. The random variable $S_{\theta,k}$ is Poisson distributed with mean $\mu_\theta(k,t)$ and describes the total number of detected photons at the kth pixel from the two point sources. The random variable B_k is Poisson distributed with mean $\beta(k,t)$ and describes the total number of detected photons at the kth pixel from spurious sources. The random variable W_k is Gaussian distributed with mean η_k and standard deviation $\sigma_{w,k}$ and describes the measurement noise at the kth pixel. We assume that $\beta(k,t)$, η_k and $\sigma_{w,k}$ are independent of θ, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$.

5.8.3.1 Fisher Information Matrix for a Pixelated Detector

In the absence of Gaussian noise (i.e., $W_k = 0$, $k = 1, \ldots, N_p$), the Fisher information matrix for a pixelated detector corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) := \sum_{k=1}^{N_p} \frac{1}{\mu_\theta(k,t) + \beta(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta}, \quad \theta \in \Theta,$$

where $\mu_\theta(k,t)$ ($\beta(k,t)$) denotes mean number of detected photons at the kth pixel from the two point (spurious) sources. Setting $\beta(k,t) = 0$ in the above equation, we obtain an expression for the Fisher information matrix of a pixelated detector in the absence of additive noise sources.

In the presence of Gaussian noise, the Fisher information matrix is given by (see ref. 3 for details)

$$\mathbf{I}(\theta) := \sum_{k=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \times$$

$$\left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2} \right)^2}{p_{\theta,k}(z)} dz - 1 \right),$$

where $\theta \in \Theta$, $\nu_\theta(k,t) := \mu_\theta(k,t) + \beta(k,t)$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, μ_θ and β are as given above, and

$$p_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\sigma_{w,k}} \sum_{l=0}^{\infty} \frac{[\nu_\theta(k,t)]^l e^{-\nu_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \tag{5.9}$$

Analogous to the result given in Theorem 2.2.1, the above expressions of the Fisher information matrix for a pixelated detector are applicable to a wide variety of imaging conditions. To calculate the Fisher information matrix in the present context, we require the analytical expression for $\mu_\theta(k,t)$ (and $\partial\mu_\theta(k,t)/\partial\theta$), which is given in Eq. 5.10. In addition, the numerical values of the noise parameters $\beta(k,t)$, η_k and $\sigma_{w,k}$ need to be known, which depend on the experimental setup.

5.8.3.2 The Generalized PREM

For the derivation of the g-PREM, we consider a geometry shown in Fig. 5.5, where the two point sources are located at arbitrary locations P_1 and P_2 on the specimen plane. Here (x_0, y_0) denotes the coordinates of the point P_1, d denotes the distance of separation between the point sources and ϕ denotes the angle of inclination of the line segment joining the two point sources with respect to the x axis. The coordinates of the point P_2 are given by $(x_0 + d\cos\phi, y_0 + d\sin\phi)$. In a practical situation, in addition to d, the other parameters, namely x_0, y_0 and ϕ are also unknown and therefore must be estimated along with d. Hence

the unknown parameter vector is given by $\theta := (x_0, y_0, d, \phi)$. The general expression for μ_θ is given by

$$\mu_\theta(k,t) := \mu_\theta^1(k,t) + \mu_\theta^2(k,t), \quad \theta \in \Theta, \quad k = 1, \ldots, N_p, \tag{5.10}$$

where, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$,

$$\mu_\theta^1(k,t) := \frac{1}{M^2} \int_{t_0}^{t} \Lambda_1(\tau) d\tau \int_{C_k} q_1 \left(\frac{x}{M} - x_0, \frac{y}{M} - y_0 \right) dxdy, \tag{5.11}$$

$$\mu_\theta^2(k,t) := \frac{1}{M^2} \int_{t_0}^{t} \Lambda_2(\tau) d\tau \int_{C_k} q_2 \left(\frac{x}{M} - x_0 - d\cos\phi, \frac{y}{M} - y_0 - d\sin\phi \right) dxdy. \tag{5.12}$$

In eqs. 5.10 - 5.12, C_k denotes the kth pixel, $[t_0, t]$ denotes the acquisition time interval, M denotes the magnification of the microscope setup, $\Lambda_1(\tau)$ and $\Lambda_2(\tau)$, $\tau \geq t_0$, denote the photon detection rates of the point sources, and q_1 and q_2 denote the image functions of the point sources.

Because the parameter θ is a 1×4 vector, by definition the Fisher information matrix $\mathbf{I}(\theta)$ is a 4×4 matrix. The g-PREM is given by the square root of the leading diagonal entry in $\mathbf{I}^{-1}(\theta)$ that corresponds to the distance parameter d. In the present case, this is the third leading diagonal entry, i.e., $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$, as d is the third component of θ.

5.8.3.3 Derivation of the PREM

The PREM is a special case of the g-PREM in which the photon detection rates of the point sources are assumed to be a constant, i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$, and image functions of the point sources are assumed to be the Airy profile. In the Section *Analytical expression of* μ_θ (*Analytical expression of* $\partial\mu_\theta/\partial\theta$), we give the analytical expression for $\mu_\theta(k,t)$ ($\partial\mu_\theta(k,t)/\partial\theta$) in terms of the Airy profile, which is required for the calculation of the PREM.

5.8.3.4 Analytical expression of μ_θ

For the Airy profile, we have

$$\mu_\theta(k,t) := \mu_\theta^1(k,t) + \mu_\theta^2(k,t), \quad k = 1, \ldots, N_p, \quad \theta \in \Theta, \tag{5.13}$$

where, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$,

$$\mu_\theta^1(k,t) := \Lambda_0(t-t_0) \int_{C_k} \frac{J_1^2(a\sqrt{(x-Mx_0)^2+(y-My_0)^2})}{\pi((x-Mx_0)^2+(y-My_0)^2)} dxdy, \tag{5.14}$$

$$\mu_\theta^2(k,t) := \Lambda_0(t-t_0) \int_{C_k} \frac{J_1^2(a\sqrt{(x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2})}{\pi((x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2)} dxdy, \tag{5.15}$$

with $a = 2\pi n_a/(\lambda M)$, n_a denoting the numerical aperture of the objective lens and λ denoting the wavelength of the photons.

5.8.3.5 Analytical expression of $\partial\mu_\theta/\partial\theta$

For $\theta = (x_0, y_0, d, \phi) \in \Theta$, let μ_θ be given by Eq. **5.13**, $r_{01} := M(x_0, y_0)$, $r_{02} := M(x_0 + d\cos\phi, y_0 + d\sin\phi)$ and $a = 2\pi n_a/(\lambda M)$. For $M > 0$, define $||r - r_{01}|| := \sqrt{(x-Mx_0)^2+(y-My_0)^2}$ and $||r - r_{02}|| = \sqrt{(x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2}$, where $r = (x,y) \in \mathbb{R}^2$. Then

$$\frac{\partial\mu_\theta(k,t)}{\partial\theta} := \left[\frac{\partial\mu_\theta(k,t)}{\partial x_0} \quad \frac{\partial\mu_\theta(k,t)}{\partial y_0} \quad \frac{\partial\mu_\theta(k,t)}{\partial d} \quad \frac{\partial\mu_\theta(k,t)}{\partial\phi}\right], \quad k = 1, \ldots, N_p, \quad \theta \in \Theta,$$

where the entries of the row vector in the above equation are given below:

$$\begin{aligned}
\frac{\partial\mu_\theta(k,t)}{\partial x_0} &= 2aM\Lambda_0(t-t_0)\left(\int_{C_k}(x-Mx_0)\frac{J_1(a||r-r_{01}||)J_2(a||r-r_{01}||)}{\pi||r-r_{01}||^3}dr\right.\\
&\qquad\left.+\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial y_0} &= 2aM\Lambda_0(t-t_0)\left(\int_{C_k}(y-My_0)\frac{J_1(a||r-r_{01}||)J_2(a||r-r_{01}||)}{\pi||r-r_{01}||^3}dr\right.\\
&\qquad\left.+\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial d} &= 2aM\Lambda_0(t-t_0)\left(\cos\phi\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right.\\
&\qquad\left.+\sin\phi\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial\phi} &= 2aM\Lambda_0(t-t_0)\left(-d\sin\phi\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right.\\
&\qquad\left.+d\cos\phi\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),
\end{aligned}$$

with $r = (x,y) \in \mathbb{R}^2$, $dr := dxdy$, $\theta \in \Theta$ and $k = 1, \ldots, N_p$.

5.8.4 Calculation of the resolution measure when additional spatial information is present

This section discusses the calculation of the Fisher information matrix for the global estimation problem of determining the unknown parameter θ from data acquired before and after the first photobleaching event for a single-molecule pair that exhibits a double-step photobleaching behavior. Without loss of generality, the location coordinates of the single molecule that photobleaches first are set to be $(x_0 + d\cos\phi, y_0 + d\sin\phi)$ and the location coordinates of the single molecule that photobleaches last are set to be (x_0, y_0). The data acquired before and after the first photobleaching event are mutually independent and therefore the Fisher information matrix for the global estimation problem can be written as

$$\mathbf{I}_{tot}(\theta) := \mathbf{I}(\theta) + \mathbf{I}_a(\theta), \quad \theta \in \Theta.$$

In the above equation, $\mathbf{I}(\theta)$ and $\mathbf{I}_a(\theta)$ denote the Fisher information matrices that are calculated for the problem of estimating the unknown parameter θ from data acquired before and after the first photobleaching event, respectively. The expressions of $\mathbf{I}(\theta)$ for a pixelated detector in the presence and absence of noise sources are given in Section *Fisher information matrix for a pixelated detector*. The matrix $\mathbf{I}_a(\theta)$ is of the form $\mathbf{I}_a(\theta) := \begin{bmatrix} \mathbf{I}_l(\theta) & 0 \\ 0 & 0 \end{bmatrix}$, $\theta \in \Theta$, where $\mathbf{I}_l(\theta)$ denotes the Fisher information matrix for the problem of determining the location (x_0, y_0) of a single molecule. In the absence of Gaussian noise $\mathbf{I}_l(\theta)$ is given by

$$\mathbf{I}_l(\theta) = \sum_{k=1}^{N_p} \frac{1}{\mu_\theta^1(k,t) + \beta^1(k,t)} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \\ \\ \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} & \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix}, \quad \theta \in \Theta,$$

where $\mu_\theta^1(k,t)$ is given in Eq. 5.14, $\beta^1(k,t)$ denotes the mean of the additive Poisson noise in the kth pixel, and the partial derivatives $\partial\mu_\theta^1(k,t)/\partial x_0$ and $\partial\mu_\theta^1(k,t)/\partial y_0$ are given by

$$\frac{\partial \mu_\theta^1(k,t)}{\partial \zeta_0} = 2aM\Lambda_0(t-t_0) \int_{C_k} (\zeta - M\zeta_0) \frac{J_1(a||r - r_{01}||)J_2(a||r - r_{01}||)}{\pi||r - r_{01}||^3} dr, \tag{5.16}$$

for $\zeta \in \{x, y\}$, $\theta \in \Theta$, $k = 1, \ldots, N_p$ and $||r - r_{01}|| := \sqrt{(x - Mx_0)^2 + (y - My_0)^2}$. In the presence of Gaussian noise, $\mathbf{I}_l(\theta)$ is given by

$$\mathbf{I}_l(\theta) = \sum_{k=1}^{N_p} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \\ \\ \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} & \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix} \times$$

$$\left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta^1(k,t)]^{l-1} e^{-\nu_\theta^1(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2} \right)^2}{p_{\theta,k}^1(z)} dz - 1 \right), \quad \theta \in \Theta,$$

where $\nu_\theta^1(k,t) := \mu_\theta^1(k,t) + \beta^1(k,t)$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, μ_θ^1 is given in Eq. **5.14**, β^1 is given above, $\partial \mu_\theta^1(k,t)/\partial x_0$ and $\partial \mu_\theta^1(k,t)/\partial y_0$ are given in Eq. **5.16**, and the expression for $p_{\theta,k}^1$ is analogous to that given in Eq. **5.9**, but with the term $\nu_\theta(k,t)$ in Eq. **5.9** replaced by $\nu_\theta^1(k,t)$.

5.8.5 Maximum-likelihood estimation

The estimation of the unknown parameters is carried out on the data that is contained in a pixel array. We first consider the scenario when the pixel array is extracted from an individual image, which contains photons from both point sources. The log-likelihood function for the data in the pixel array is given by

$$\ln(\mathcal{L}(\theta \mid z_1, \ldots, z_{N_p})) := \ln\left(\prod_{k=1}^{N_p} p_{\theta,k}(z_k) \right) = \sum_{k=1}^{N_p} \ln(p_{\theta,k}(z_k)), \quad \theta \in \Theta, \tag{5.17}$$

where N_p denotes the total number of pixels in the pixel array, z_k denotes the detected photon count at the kth pixel in the pixel array, and $p_{\theta,k}$ denotes the probability density function of z_k that is given by Eq. 5.9, $k = 1, \ldots, N_p$. For the distance estimation problem with the data acquired by a pixelated detector in the presence of noise sources, the vector of unknown parameters is set to be $\theta = (x_0, y_0, d, \phi)$ (see Fig. 5) and the image function of the point source is assumed to be the Airy profile. The maximum-likelihood estimate of θ is obtained

by substituting the expression for μ_θ given by Eq. 5.13 in $p_{\theta,k}$ (Eq. 5.9) and determining the value of θ that maximizes the log-likelihood function $\ln(\mathcal{L}(\theta))$.

We next consider the scenario when the pixel array is obtained by adding N_1 pixel arrays, which are extracted from N_1 individual images that contain photons from both point sources. The log-likelihood function for the data in the summed pixel array is given by

$$\ln(\tilde{\mathcal{L}}(\theta \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p})) := \sum_{k=1}^{N_p} \ln(\tilde{p}_{\theta,k}(\tilde{z}_k)), \quad \theta \in \Theta, \tag{5.18}$$

where $\tilde{z}_k$ denotes the detected photon count at the kth pixel in the summed pixel array and $\tilde{p}_{\theta,k}$ denotes the density function of $\tilde{z}_k$, $k = 1, \ldots, N_p$, which is given by

$$\tilde{p}_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\tilde{\sigma}_{w,k}} \sum_{l=0}^{\infty} \frac{[\tilde{\nu}_\theta(k,t)]^l e^{-\tilde{\nu}_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\tilde{\eta}_k}{\tilde{\sigma}_{w,k}}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \tag{5.19}$$

In Eq. 5.19, $\tilde{\nu}_\theta(k,t) := N_1(\mu_\theta(k,t) + \beta(k,t))$, $\tilde{\eta}_k = N_1\eta_k$, and $\tilde{\sigma}_{w,k} = \sqrt{N_1}\sigma_{w,k}$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, where $\mu_\theta(k,t)$ ($\beta(k,t)$) denotes the mean photon count at the kth pixel from the point sources (scattering-noise sources) in the individual pixel array, η_k and $\sigma^2_{w,k}$ denote the mean and the variance of the readout noise at the kth pixel, respectively, in the individual pixel array. For the distance estimation problem, the maximum-likelihood estimate of $\theta = (x_0, y_0, d, \phi)$ is obtained by substituting the expression for μ_θ given by Eq. 5.13 in $\tilde{p}_{\theta,k}$ (Eq. 5.19) and determining the value of θ that maximizes the log-likelihood function $\ln(\tilde{\mathcal{L}}(\theta))$.

If the point sources exhibit a double step photobleaching behavior, the images acquired after the first photobleaching event can also be used to estimate θ. Here, the experimental data that is used to estimate θ consists of two summed pixel arrays. One of the summed pixel arrays is obtained by adding N_1 pixel arrays that are extracted from N_1 individual images acquired before the first photobleaching event (i.e., images that contain photons from both point sources). In this summed pixel array the detected photon count at the kth pixel is denoted as $\tilde{z}_k$, $k = 1, \ldots, N_p$, where N_p denotes the total number of pixels. The other summed pixel array is obtained by adding N_2 pixel arrays that are extracted from N_2 individual images acquired after the first photobleaching event and $\tilde{z}^1_k$ denotes the detected photon count at the kth pixel

in this summed pixel array. The log-likelihood function $\ln(\mathcal{L}_T(\theta))$ for the data contained in the two summed pixel arrays is given by

$$\ln(\mathcal{L}_T(\theta) \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p}; \tilde{z}^1_1, \ldots, \tilde{z}^1_{N_p})) := \ln(\tilde{\mathcal{L}}(\theta \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p})) + \ln(\tilde{\mathcal{L}}^1(\theta \mid \tilde{z}^1_1, \ldots, \tilde{z}^1_{N_p})), \quad (5.20)$$

where $\theta \in \Theta$ and $\ln(\tilde{\mathcal{L}})$ ($\ln(\tilde{\mathcal{L}}^1)$) denotes the log-likelihood function corresponding to $\{\tilde{z}_1, \ldots, \tilde{z}_{N_p}\}$ ($\{\tilde{z}^1_1, \ldots, \tilde{z}^1_{N_p})\}$). The expression for $\ln(\tilde{\mathcal{L}})$ is given in Eq. 5.18, and the expression for $\ln(\tilde{\mathcal{L}}^1)$ is given by $\ln(\tilde{\mathcal{L}}^1(\theta \mid \tilde{z}^1_1, \ldots, \tilde{z}^1_{N_p})) := \sum_{k=1}^{N_p} \ln(\tilde{p}^1_{\theta,k}(\tilde{z}^1_k))$, $\theta \in \Theta$, where $\tilde{p}^1_{\theta,k}$ denotes the density function of $\tilde{z}^1_k$ that is given by

$$\tilde{p}^1_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\tilde{\sigma}^1_{w,k}} \sum_{l=0}^{\infty} \frac{[\tilde{\nu}^1_\theta(k,t)]^l e^{-\tilde{\nu}^1_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\tilde{\eta}^1_k}{\tilde{\sigma}^1_{w,k}}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \quad (5.21)$$

In Eq. 5.21, $\tilde{\nu}^1_\theta(k,t) := N_2(\mu^1_\theta(k,t) + \beta^1(k,t))$, $\tilde{\eta}^1_k = N_2\eta^1_k$, and $\tilde{\sigma}^1_{w,k} = \sqrt{N_2}\sigma^1_{w,k}$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, where $\mu^1_\theta(k,t)$ ($\beta^1(k,t)$) denotes the mean photon count at the kth pixel from one of the point sources (scattering noise sources) in the pixel array that is extracted from the image acquired after the first photobleaching event, and η^1_k and $(\sigma^1_{w,k})^2$ denote the mean and the variance of the readout noise at the kth pixel, respectively in the pixel array that is extracted from the image acquired after the first photobleaching event. Thus, for the distance estimation problem, the maximum-likelihood estimate of $\theta = (x_0, y_0, d, \phi)$ from the two summed pixel arrays is obtained by substituting the expression for μ_θ given by Eq. 5.13 in $\tilde{p}_{\theta,k}$ (Eq. 5.19) and substituting the expression for μ^1_θ given by Eq. 5.14 in $\tilde{p}^1_{\theta,k}$ (Eq. **5.21**), and then determining the value of θ that maximizes the log-likelihood function $\ln(\mathcal{L}_T(\theta))$.

We note that Eqs. 5.17, 5.18 and 5.20 can be used to obtain the maximum-likelihood estimate of θ in a wide variety of imaging conditions. For instance, consider the scenario when the image function of the point source is described by a profile that is different from the Airy profile. In this case, we use Eq. 5.10 (Eq. 5.11) to obtain an expression for μ_θ (μ^1_θ) in terms of the desired image profile, and then maximize the corresponding log-likelihood function to obtain the maximum-likelihood estimate. In all of the above cases, the maximum-likelihood estimation is carried with the optimization toolbox of MATLAB in the MIATool software environment (software available on request).

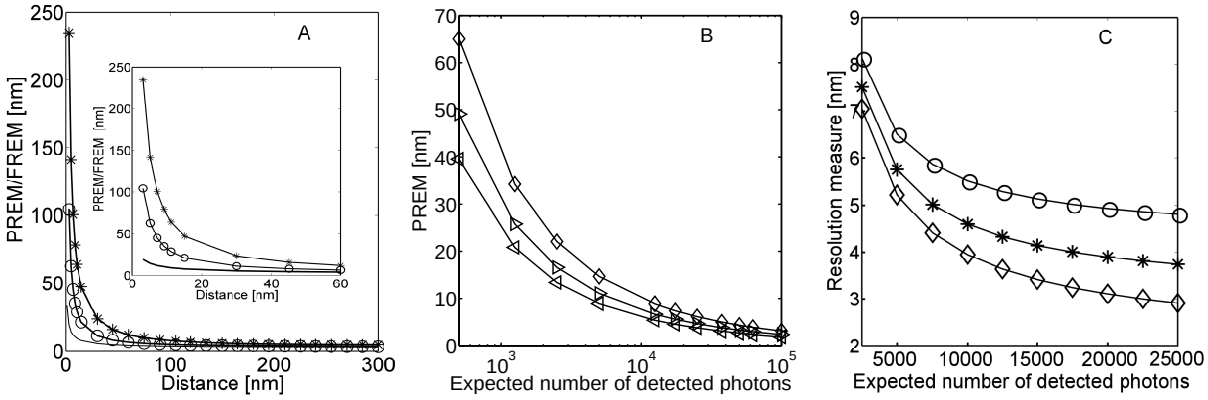

Figure 5.2. Behavior of the PREM as a function of distance and photon count. Panel A shows the PREM as a function of the distance of separation between two Cy5 molecules in the presence ($*$) and absence ($\circ$) of noise sources for a pixelated detector. The PREM given in eq. 5.1 is also shown for reference (—). The inset shows the same for distances in the range of $1 - 60$ nm. Panel B shows the PREM for a pixelated detector as a function of the expected number of detected photons per molecule in the presence of noise sources for different distances of separation; $d = 30$ nm ($\diamond$), $d = 40$ nm ($\triangleright$) and $d = 50$ nm ($\triangleleft$). Panel C shows the effect of using additional spatial information on the resolution measure for a pair of Cy5 molecules ($d = 10$ nm) that exhibits a double-step photobleaching behavior. The panel shows the resolution measure as a function of the expected photon count collected from the single molecule after the first photobleaching event for a pixelated detector in the presence of noise sources. The plots shown consider three scenarios, i.e., when the expected number of photons collected from each single molecule before the first photobleaching event are 2500 ($\circ$), 5000 ($*$) and 12500 ($\diamond$). In all panels, the photon detection rate Λ_0 of each Cy5 molecule is set to be $\Lambda_0 = 2500$ photons/s, the acquisition time is set to be 1 s, the pixel dimension is set to be $12.9\ \mu m \times 12.9\ \mu m$, the pixel array size is set to be 13×13, the mean of the additive Poisson noise is set to be 80 photons/pixel/s, the mean and standard deviation of the additive Gaussian noise is set to be 0 e^-/pixel and 8 e^-/pixel, respectively, the noise statistics is assumed to be the same for all pixels, and one of the single molecules is assumed to be at the center of the pixel array. All other numerical parameters are analogous to those used in Fig. 5.1.

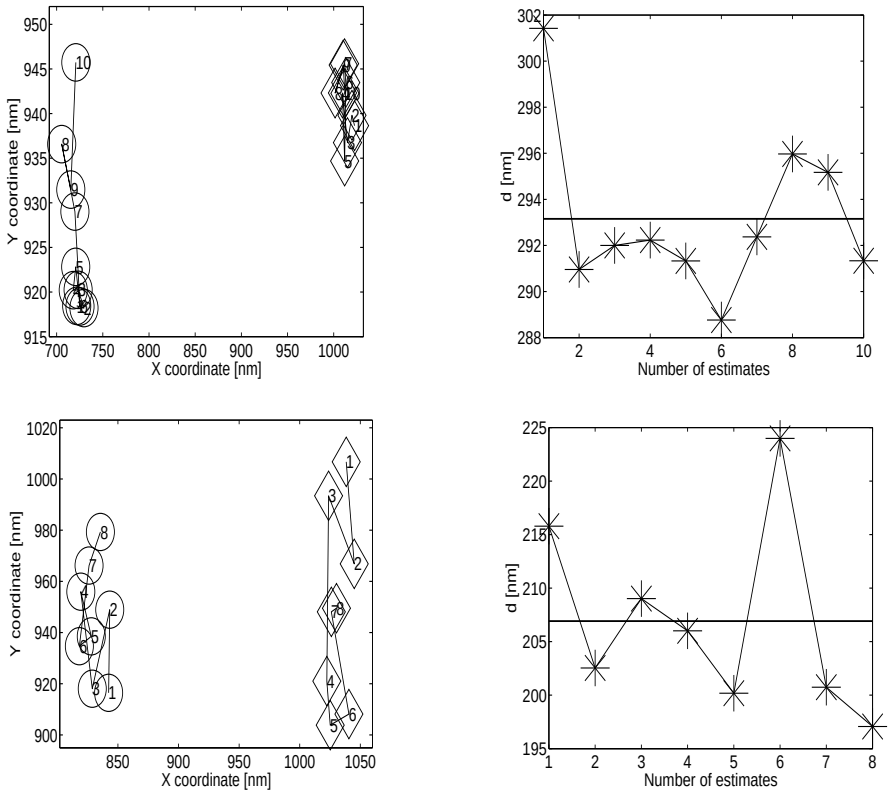

Figure 5.3. Results of distance estimation for closely spaced Cy5 molecules.
The figure shows the estimates of the X-Y coordinates and the distance of separation for two pairs of closely spaced Cy5 single molecules. Top row shows results for a single molecule pair with a mean distance of separation of 293 nm (Data Analysis 1 in Table 5.1). Bottom row shows results for a single molecule pair with a mean distance of separation of 207 nm (Data Analysis 2 in Table 5.1).

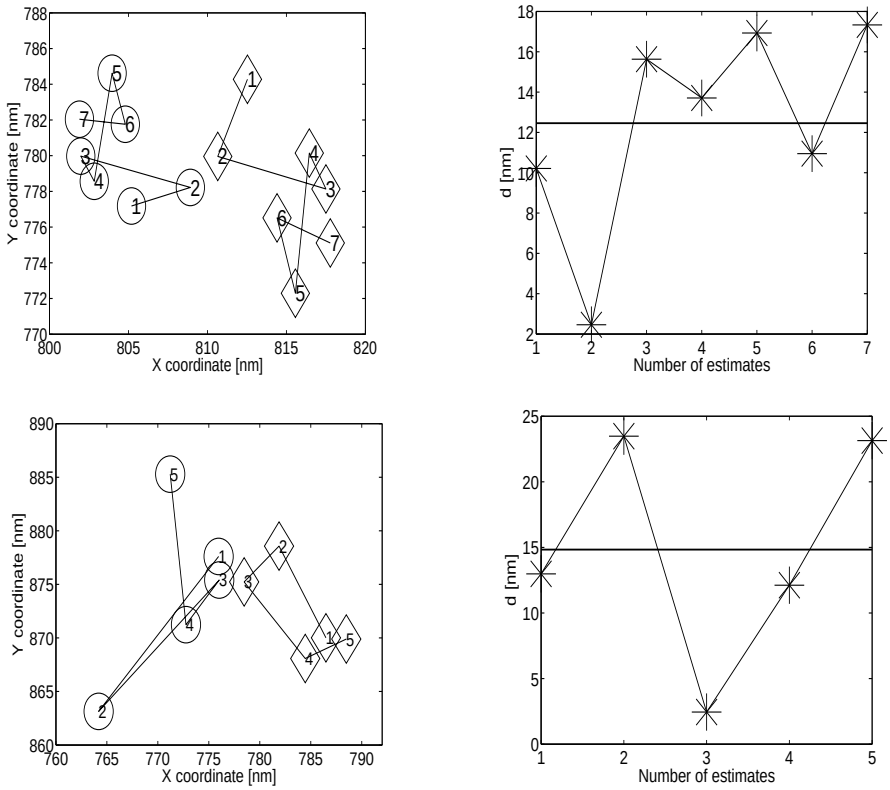

Figure 5.4. Results of distance estimation for fluorescently labeled DNA nanorulers. The figure shows the estimates of the X-Y coordinates and the distance of separation for two DNA nanorulers. Top row shows results for the DNA nanoruler with a mean distance of separation of 12.5 nm (Data Analysis 5 in Table 5.1). Bottom row shows results for the DNA nanoruler with a mean distance of separation of 14.8 nm (Data Analysis 6 in Table 5.1).

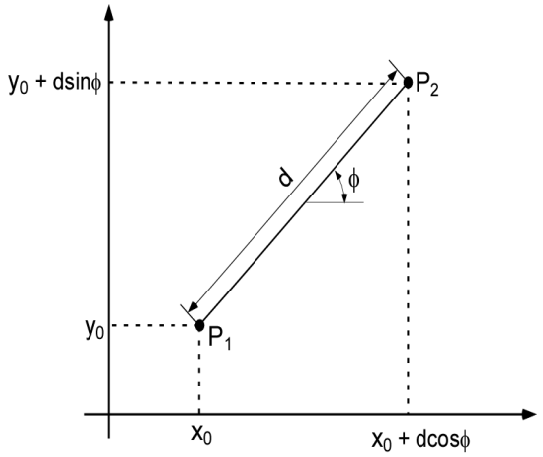

Figure 5.5. Coordinate system used for the derivation of the generalized PREM.

REFERENCES

[1] G. Seisenberger, M. U. Reid, T. Endress, H. Büning, M. Hallek, and C. Bräuchle, "Real-time single-molecule imaging of the infection pathway of an adeno-associated virus," *Science*, vol. 294, pp. 1929–1932, 2001.

[2] H. Taguchi, T. Ueno, H. Tadakuma, M. Yoshida, and T. Funatsu, "Single molecule observation of protein-protein interactions in the chaperonin system," *Nature Biotechnology*, vol. 19, pp. 861–865, 2001.

[3] M. Tomishige, Y. Sako and A. Kusumi, "Regulation mechanism of the lateral diffusion of Band 3 in erythrocyte membranes by the membrane skeleton," *Journal of Cell Biology*, vol. 142, no. 4, pp. 989–1000, 1998.

[4] R. Lino, I. Koyama, and A. Kusumi, "Single molecule imaging of greeen fluorescent proteins in living cells: E-Cadherin forms oligomers on the free cell surface," *Biophysical Journal*, vol. 80, pp. 2667–2677, 2001.

[5] P. R. Smith, I. E. G. Morrision, K. M. Wilson, N. Fernandez, and R. J. Cherry, "Anomalous diffusion of major histocompatibility complex class-I molecules on HeLa cells determined by single particle tracking," *Biophysical Journal*, vol. 76, pp. 3331–3344, 1999.

[6] J. Gelles, B. J. Schnapp, and M. P. Scheetz, "Tracking kinesin-driven movements with nanometric-scale precision," *Nature*, vol. 331, pp. 450 – 453, 1988.

[7] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, "Myosin walks hand over hand: single fluorophore imaging with 1.5 nm localization," *Science*, vol. 300, pp. 2061–2065, 2003.

[8] W. E. Moerner and L. Kador, "Optical detection and spectroscopy of single molecules in a solid," *Physical Review Letters*, vol. 62, pp. 2535–2538, 1989.

[9] W. E. Moerner and D. P. Fromm, "Methods of single-molecule fluorescence spectroscopy and microscopy," *Review of Scientific Instruments*, vol. 74, no. 8, pp. 3597–3619, 2003.

[10] S. Weiss, "Fluorescence spectroscopy of single biomolecules," *Science*, vol. 283, pp. 1676–1683, 1999.

[11] O. Shimomura, F. H. Johnson, and Y. Saiga, "Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea." *Journal of Cell Composition and Physiology*, vol. 59, pp. 223–239, 1962.

[12] J. Zlatanova and K. V. Holde, "Single-molecule biology: what is it and how does it work?" *Molecular Cell*, vol. 24, pp. 317–329, 2006.

[13] M. Bruchez, P. Gin, S. Weiss, and A. P. Alivisatos, "Semiconductor nanocrystals as fluorescent biological labels," *Science*, vol. 281, pp. 2013–2016, 1998.

[14] J. K. Jaiswal and S. M. Simon, "Potentials and pitfalls of fluorescent quantum dots for biological imaging," *Trends in Cell Biology*, vol. 14, pp. 497–504, 2004.

[15] R. Rigler and E. S. Elson, *Fluorescence Correlation Spectroscopy: Theory and Applications*. Springer Verlag, 2001.

[16] S. Inoue and K. R. Spring, *Video Microscopy: The fundamentals*, 2nd ed. Plenum Press, 1997.

[17] D. J. Denvir and C. G. Coates, "Electron-multiplying CCD technology: application to ultrasensitive detection of biomolecules," *Proceedings of the SPIE*, vol. 5626, pp. 502–512, 2002.

[18] D. J. Denvir and E. Conroy, "Electron multiplying CCDs," *Proceedings of the SPIE*, vol. 4877, pp. 55–68, 2003.

[19] D. Axelrod, *Light microscopy in biology*. Oxford University Press, Oxford, UK, 1999, ch. 11.

[20] G. I. Mashanov, D. Tacon, A. E. Knight, M. Peckham, and J. E. Molloy, "Visualizing single molecules inside living cells using total internal reflection fluorescence microscopy," *Methods*, vol. 29, pp. 142–152, 2003.

[21] R. M. Dickson, D. J. Norris, Y. L. Tzeng, and W. E. Moerner, "On/off blinking and switching behaviour of single molecules of green fluorescent protein," *Nature*, vol. 388, pp. 355–358, 1997.

[22] G. C. Holst, *CCD arrays, cameras and displays*, 2nd ed. SPIE Press, Bellingham, WA, 1998.

[23] D. S. Martin, M. B. Forstner, and J. A. Käs, "Apparent subdiffusion inherent to single particle tracking," *Biophysical Journal*, vol. 83, pp. 2109–2117, 2002.

[24] N. Bobroff, "Position measurement with a resolution and noise limited instrument," *Review of Scientific Instruments*, vol. 57, no. 6, pp. 1152–1157, 1986.

[25] G. J. Schütz, H. Schindler, and T. Schmidt, "Single-molecule microscopy on model membranes reveals anamalous diffusion," *Biophysical Journal*, vol. 73, pp. 1073–1080, 1997.

[26] U. Kubitscheck, O. Kückmann, T. Kues, and R. Peters, "Imaging and tracking single GFP molecules in solution," *Biophysical Journal*, vol. 78, pp. 2170–2179, 2000.

[27] R. N. Ghosh and W. W. Webb, "Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules,," *Biophysical Journal*, vol. 66, pp. 1301–1318, 1994.

[28] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophysical Journal*, vol. 82, pp. 2775–2783, 2002.

[29] M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophysical Journal*, vol. 81, pp. 2378–2388, 2001.

[30] D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with a charge coupled device camera," *Journal of the Optical Society of America A*, vol. 10, no. 5, pp. 1014–1023, 1993.

[31] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, and R. L. White, "Compensation for read out noise in charge coupled device images," *Journal of the Optical Society of America A*, vol. 12, no. 2, pp. 272–283, 1995.

[32] R. J. Ober, C. Martinez, X. Lai, J. Zhou, and E. S. Ward, "Exocytosis of IgG as mediated by the receptor, FcRn: An analysis at the single molecule level," *Proceedings of the National Academy of Sciences USA*, vol. 101, no. 30, pp. 11 076–11 081, 2004.

[33] S. Y. Kim, Z. Gitai, A. Kinkhabwala, L. Shapiro, and W. E. Moerner, "Single molecules of the bacterial actin MreB undergo directed treadmilling motion in Caulobacter crescentus," *Proceedings of the National Academy of Sciences USA*, vol. 103, pp. 10 929–34, 2006.

[34] S. Inoue, "Imaging of unresolved objects, superresolution, and precision of distance measurement with video microscopy," in *Methods in Cell Biology*, D. Taylor, Y. L. Wang, L. Wilson, and P. Matsudaira, Eds. Academic Press, NY, USA, 1989, vol. 30B, pp. 85–112.

[35] J. E. N. Jonkman, J. Swoget, H. Kress, A. Rohrbach, and E. H. K. Stelzer, "Resolution in optical microscopy," in *Methods in Enzymology*, G. Marriott and I. Parker, Eds. Academic Press, New York, USA, 2003, vol. 360A, pp. 416–446.

[36] S. Ram, E. S. Ward, and R. J. Ober, "How accurately can a single molecule be localized in three dimensions using a fluorescence microscope?" *Proceedings of the SPIE*, vol. 5699, pp. 426–435, 2005.

[37] F. Aguet, D. V. D. Ville, and M. Unser, "A maximum likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles," *Optics Express*, vol. 13, pp. 10 503–10 522, 2005.

[38] J. Enderlein, "Tracking of fluorescent molecules diffusing within membranes," *Applied Physics B*, vol. 71, pp. 773–777, 2000.

[39] ——, "Positional and temporal accuracy of single molecule tracking," *Single Molecule*, vol. 3, pp. 225–230, 2000.

[40] A. J. Berglund and H. Mabuchi, "Feedback controller design for tracking of a single fluorescent particle," *Applied Physics B*, vol. 78, pp. 653–659, 2004.

[41] V. Levi, Q. Ruan, and E. Gratton, "3-d particle tracking in a two-photon microscope: application to the study of molecular dynamics in cells," *Biophysical Journal*, vol. 88, 2005.

[42] T. E., J. Enderlein, S. Syed, S. A. Mckinney, R. G. Petschek, T. Ha, Y. E. Goldman, and P. R. Selvin, "Defocussed orientation and position imaging (DOPI) of myosin V," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 6495–6499, 2006.

[43] L. Holtzer, T. Meckel, and T. Schmidt, "Nanometric three-dimensional tracking of individual quantum dots in cells," *Applied Physics Letters*, vol. 90, p. 053902, 2007.

[44] P. V. Roessel and A. H. Brand, "Imaging into the future: visualizing gene expression and proteins interactions with fluorescent proteins," *Nature Cell Biology*, vol. 4, pp. E15–E20, 2002.

[45] A. J. den Dekker and A. van den Bos, "Resolution: a survey," *Journal of the Optical Society of America A*, vol. 14, pp. 547–557, 1997.

[46] P. C. D. Hobbs, *Building electro-optical systems*. John Wiley and Sons, 2000.

[47] Z. Zalevsky and D. Mendlovic, *Optical superresolution*. Springer Verlag, New York, USA, 2003.

[48] M. P. Gordon, T. Ha, and P. R. Selvin, "Single molecule high resolution imaging with photobleaching," *Proceedings of the National Academy of Sciences USA*, vol. 101, no. 17, pp. 6462–6465, 2004.

[49] X. Qu, D. Wu, L. Mets, and N. F. Scherer, "Nanometer-localized multiple single-molecule fluorescence microscopy," *Proceedings of the National Academy of Sciences USA*, vol. 101, no. 31, pp. 11 298–11 303, 2004.

[50] K. A. Lidke, B. Rieger, T. M. Jovin, and R. Heintzmann, "Superresolution by localization of quantum dots using blinking statistics," *Optics Express*, vol. 13, no. 18, pp. 7052–7062, 2005.

[51] R. J. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single molecule microscopy," *Biophysical Journal*, vol. 86, pp. 1185–1200, 2004.

[52] S. Ram, E. S. Ward, and R. J. Ober, "Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy," *Proceedings of the National Academy of Sciences USA*, vol. 103, pp. 4457–4462, 2006.

[53] S. Ram, J. Chao, P. Prabhat, E. S. Ward, and R. J. Ober, "A novel approach to determining the three-dimensional location of microscopic objects with applications to 3D particle tracking," *Proceedings of the SPIE*, vol. 6444, pp. D1–D9, 2007.

[54] C. R. Rao, *Linear statistical inference and its applications.* Wiley, New York, USA., 1965.

[55] S. Zacks, *The theory of statistical inference.* John Wiley and Sons, New York, USA, 1971.

[56] M. Born and E. Wolf, *Principles of Optics.* Cambridge University Press, Cambridge, UK, 1999.

[57] S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of performance limits for optical microscopes," *Multidimensional Systems and Signal Processing*, vol. 17, pp. 27–58, 2006.

[58] B. Saleh, *Photoelectron statistics.* Springer Verlag, Berlin, Germany, 1978.

[59] S. M. Kay, *Fundamentals of statistical signal processing.* Prentice Hall PTR, New Jersey, USA, 1993.

[60] D. L. Snyder and M. I. Miller, *Random point processes in time and space.*, 2nd ed. Springer Verlag, New York, USA, 1991.

[61] A. Papoulis and S. U. Pillai, *Probability, random variables and stochastic processes.*, 4th ed. McGraw Hill, New York, USA, 2002.

[62] K. A. Winick, "Cramer-Rao lower bounds on the performance of charge coupled device optical position estimators," *J. Opt. Soc. Am. A*, vol. 3, no. 11, pp. 1809–1815, 1986.

[63] S. Ross, *Introduction to probability models.* Academic Press, CA., 2000.

[64] T. M. Apostol, *Mathematical analysis.* Addison Wesley Publishing Company, Boston, USA, 1974.

[65] W. Rudin, *Real and Complex Analysis.* Mc Graw Hill, New York, USA, 1987.

[66] J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. Mc Graw Hill, USA, 1996.

[67] A. Santos and I. T. Young, "Model-based resolution: applying the theory in quantitative microscopy," *Applied Optics*, vol. 39, pp. 2948–2958, 2000.

[68] G. N. Watson, *A treatise on the theory of Bessel functions.* Cambridge University Press, Cambridge, UK, 1958.

[69] F. Zhang, *Matrix Theory.* Springer Verlag, New York, USA, 1999.

[70] F. Bowman, *Introduction to Bessel functions.* Dover, New York, USA, 1968.

[71] T. Schmidt, G. J. Schütz, W. Baumgartner, H. J. Gruber, and H. Schindler, "Characterization of photophysics and mobility of single molecules in a fluid lipid membrane," *Journal of Physical Chemistry*, vol. 99, pp. 17 662–17 668, 1995.

[72] T. Coleman, M. A. Branch, and A. Grace, *MATLAB programming reference manual, Version 6.* The Mathworks Inc, MA, 2000.

[73] T. Kues, R. Peters, and U. Kubitscheck, "Visualization and tracking of single protein molecules in the cell nucleus," *Biophysical Journal*, vol. 80, pp. 2954–2967, 2001.

[74] MATLAB, *Optimization Toolbox, for use with MATLAB - User guide ver 2.* Mathworks Inc, USA, 1999.

[75] M. A. and A. S. T. Kogure, "Lighting up cells: labelling proteins with fluorophores," *Nature Cell Biology Supplement*, no. S1-S6, 2003.

[76] P. Prabhat, S. Ram, E. S. Ward, and R. J. Ober, "Simultaneous imaging of different focal planes in fluorescence microscopy for the study of cellular dynamics in three dimensions," *IEEE Transactions on Nanobioscience*, vol. 3, no. 4, pp. 237–242, 2004.

[77] ——, "Simultaneous imaging of several focal planes in fluorescence microscopy for the study of cellular dynamics in 3D," *Proceedings of the SPIE*, vol. 6090, pp. 115–121, 2006.

[78] P. Prabhat, Z. Gan, J. Chao, S. Ram, C. Vaccaro, S. Gibbons, R. J. Ober, and E. S. Ward, "Elucidation of intracellular pathways leading to exocytosis of the Fc receptor, FcRm, using multifocal plane microscopy," *Proceedings of the National Academy of Sciences USA*, vol. 104, pp. 5889–5894, 2007.

[79] V. Ghetie and E. S. Ward, "Multiple roles for the major histocompatibility complex class I- related receptor FcRn," *Annual Review of Immunology*, vol. 18, pp. 739–766, 2000.

[80] E. W. Ades, F. J. Candal, R. A. Swerlick, V. G. George, S. Summers, D. C. Bosse, and T. J. Lawley, "HMEC-1: establishment of an immortalized human microvascular endothelial cell line," *Journal of Investigative Dermatology*, vol. 99, pp. 683–690, 1992.

[81] R. J. Ober, C. Martinez, C. Vaccaro, J. Zhou, and E. S. Ward, "the site and dynamics of IgG salvage by the MHC class I-related receptor, FcRn," *Journal of Immunology*, vol. 172, pp. 2021–2029, 2004.

[82] C. Vaccaro, J. Zhou, R. J. Ober, and E. S. Ward, "Engineering the Fc region of immunoglobulin G to modulate in vivo antibody levels," *Nature Biotechnology*, vol. 23, pp. 1283–1288, 2005.

[83] S. F. Gibson and F. Lanni, "Diffraction by a circular aperture as a model for three-dimensional optical microscopy," *Journal of the Optical Society of America A*, vol. 6, no. 9, pp. 1357–1367, 1989.

[84] S. F. F. Gibson, "Modelling the 3D imaging properties of the fluorescence light microscope," Ph.D. dissertation, Carnegie-Mellon University, 1990.

[85] C. J. R. Sheppard and P. Török, "An electromagnetic theory of imaging in fluorescence microscopy, and imaging in polarization fluorescence microscopy," *Bioimaging*, vol. 5, pp. 205–218, 1997.

[86] J. Enderlein, "Theoretical study of detection of a dipole emitter through an objective with high numerical aperture," *Optics Letters*, vol. 25, no. 9, pp. 634–636, 2000.

[87] P. Török, P. Varga, Z. Laczik, and G. R. Booker, "Electromagnetic diffraction of light focussed through a planar interface between materials of mismatched refractive indices: an integral representation," *Journal of the Optical Society of America A*, vol. 12, no. 10, pp. 2660–2671, 1995.

[88] B. Lounis and W. E. Moerner, "Single photons on demand from a single molecule at room temperature," *Nature*, vol. 407, pp. 491–493, 2000.

[89] L. Fleury, J. M. Segura, G. Zumofen, B. Hecht, and U. Wild, "Nonclassical photon statistics in single molecule fluorescence at room temperature," *Physical Review Letters*, vol. 84, no. 6, 2000.

[90] P. Tinnefeld, C. Müller, and M. Sauer, "Time varying photon probability distribution of individual molecules at room temperature," *Chemical Physics Letters*, vol. 345, pp. 252–258, 2001.

[91] G. I. Mashanov, D. Tacon, M. Peckham, and J. E. Molloy, "The spatial and temporal dynamics of pleckstrin homology domain binding at the plasma membrane measured by imaging single molecules in live mouse myoblasts," *Journal of Biological Chemistry*, vol. 279, no. 15, pp. 15 274–15 280, 2004.

[92] P. Tinnefeld, D. P. Herten, and M. Sauer, "Photophysical dynamics of single molecules studied by spectrally resolved fluorescence lifetime imaging microscopy," *Journal of Physical Chemistry A*, vol. 105, pp. 7989–8003, 2001.

[93] A. P. Alivisatos, W. Gu, and C. Larabell, "Quantum dots as cellular probes," *Annual Review in Biomedical Engineering*, vol. 7, pp. 55–76, 2005.

[94] H. Balci, T. Ha, H. L. Sweeney, and P. R. Selvin, "Inter-head distance measurements in myosin VI via SHRImP support a simplified hand-over-hand model," *Biophysical Journal*, vol. 89, pp. 413–417, 2005.

[95] S. Hell and E. H. K. Stelzer, "Properties of a 4Pi confocal fluorescence microscope," *Journal of the Optical Society of America A*, vol. 9, pp. 2156–2166, 1992.

[96] T. A. Klar, S. Kajobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academdy of Sciences USA*, vol. 97, no. 15, pp. 8206–8210, 2000.

[97] M. Gustafsson, D. A. Agard, and J. W. Sedat, "I^5m: widefield light microscopy with better than 100nm axial resolution," *Journal of Microscopy*, vol. 195, pp. 10–16, 1998.

[98] M. Gustafsson, "Extended resolution fluorescence microscopy," *Current Opinion in Structural Biology*, vol. 9, pp. 627–634, 1999.

[99] J. M. Paar, N. T. Harris, D. Holowka, and B. Baird, "Bivalent ligands with rigid double-stranded DNA spacers reveal structural constraints on signaling by FcεRI," *Journal of Immunology*, vol. 169, pp. 856–864, 2002.

BIOGRAPHICAL STATEMENT

Sripad Ram was born in Chennai (Madras), India, in 1979. He received the B.Sc. degree in applied sciences from PSG College of Technology, Coimbatore, India, in 1999 and the M.Sc. degree in physics from Indian Institute of Technology, Chennai, India, in 2001. He joined the UTA/UTSW Bioengineering Graduate Program in 2001 and has been working in the laboratory of Professors Sally Ward and Raimund Ober since then. His research interests include fluorescence microscopy, cell biology and statistical signal processing. He is a student member of the IEEE, the SPIE and the Biophysical Society.