The evaluation of ERK1 dimerization models using fluorescence correlation Spectroscopy and the development of analysis algorithms for single-molecule super resolution

Fang Huang

Follow this and additional works at: http://digitalrepository.unm.edu/phyc_etds

Part of the Astrophysics and Astronomy Commons, and the Physics Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Physics & Astronomy ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
Fang Huang
Candidate
Department of Physics and Astronomy

This dissertation is approved, and it is acceptable in quality and form for publication on microfilm:

Approved by the Dissertation Committee:

__________________________, Chairperson

__________________________

Accepted:

__________________________
Dean, Graduate School

Date
The Evaluation of ERK1 Dimerization Models
using Fluorescence Correlation Spectroscopy and the Development of
Analysis Algorithms for Single-Molecule Super Resolution

by

Fang Huang

M.S., Physics, University of New Mexico, 2009
B.S., University of Science and Technology of China, 2004

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico

Albuquerque, New Mexico

December, 2011
©2011, Fang Huang
Dedication

To my grandfather, Huang, Zhongfeng

my parents

and

my dearest wife
Acknowledgements

First I want to express my deepest gratitude to my Ph.D. advisor, Dr. Keith Lidke. Dr. Keith Lidke has been a perfect advisor for me during my time in LidkeLab in University of New Mexico. His wisdom, devotion to science, humbleness and kindness have inspired me both in my life and research.

I also want to thank my co-advisors, Dr. James Thomas and Dr. Diane Lidke who have brought me into the field of biophysics and provided tremendous help during my Ph.D. studies.

I want to thank Michael Malik for all the helps and important suggestions on my project.

I want to thank Samantha Schwartz for teaching me cell culture techniques and helping so much on my project.

I want to thank Jason Byars for helping me advance my knowledge in C++ programming and helping re-organizing the code.
I want to thank W. Duncan Wadsworth for helpful discussion about statistics.

I want to thank Patrick Cutler for tons of helpful discussions on almost every topics in my project.

I want to thank Fenfei Liu, Kathrin Spendier and Andy Maloney for all the helps down in B10W.

I want to thank Kristopher Marjon, Niki Marjon and especially Lilli for being our good friends.

I want to thank Peter Relich for his helps in the lab.

I want to thank everyone who have helped me these years.

Thank You Very Much!
The Evaluation of ERK1 Dimerization Models
using Fluorescence Correlation Spectroscopy and the Development of
Analysis Algorithms for Single-Molecule Super Resolution

by

Fang Huang

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Physics

The University of New Mexico
Albuquerque, New Mexico

December, 2011
The Evaluation of ERK1 Dimerization Models
using Fluorescence Correlation Spectroscopy and the Development of
Analysis Algorithms for Single-Molecule Super Resolution

by

Fang Huang

M.S., Physics, University of New Mexico, 2009
B.S., University of Science and Technology of China, 2004
Ph.D., Physics, University of New Mexico, 2011

Abstract

In this dissertation, we describe three advancements in the area of fluorescence spectroscopy and microscopy. First, we have implemented Fluorescence Correlation Spectroscopy on a home-built confocal microscope. We proposed F/N analysis which stands for the ratio between average intensity and average number of moving objects in FCS as a method to detect dimerization of fluorescently tagged proteins diffusing in three dimensions. We used the technique to evaluate the dimerization hypothesis of ERK1 protein during nucleus translocation and it shows that ERK1 nucleus translocation is dimerization independent. Second, we have developed a multi-emitter analysis algorithm for single-molecule super resolution techniques that is designed to localize multiple overlapping emitters within a small sub-region. We have demonstrated its advantage compared with single emitter fitting using both simulated data sets and cell data sets. Fitting processes are implemented in GPU
using CUDA to further improve the speed of analysis. Third, we proposed a robust, consistent and reliable rejection algorithm using the log-likelihood ratio as a test statistic and have demonstrated its advantage by comparing to two commonly used test statistics in the literature. To predict the correct distribution of LLR under Poisson noise when expectation values approaches zero, we proposed a hypothesis which is used to obtain the correct distribution of LLR and together with the hypothesis we proposed a empirical function which could be used to obtain the distribution.
Contents

List of Figures xvii

List of Tables xxix

Glossary xxx

1 Introduction to Fluorescence Microscopy 1

References .................................................... 7

References 7

2 The Evaluation of ERK1 Dimerization Models using FCS 9

2.1 Introduction to FCS ............................................. 11

2.2 Experimental Setup and Data Analysis .......................... 14

2.2.1 Experimental Setup ........................................ 14

2.2.2 Data Acquisition ........................................... 15

2.2.3 FCS Diffusion Coefficient Analysis ........................ 15
Contents

2.2.4 F/N Dimerization Analysis ............................... 18
2.3 Complications Encountered during FCS Measurement ............... 20
   2.3.1 Laser intensity fluctuations .............................. 20
   2.3.2 Photobleaching and Blinking ............................. 22
2.4 Control Experiments ......................................... 23
   2.4.1 Diffusion Coefficient Measurement Control ................. 23
   2.4.2 F/N dimerization control ................................ 24
2.5 Dimerization Test on Erk1 Protein after Stimulation ............... 24
   2.5.1 Experimental Setup ...................................... 25
   2.5.2 Cell preparation (Performed by Dr. Diane S. Lidke) ......... 26
   2.5.3 Results .................................................. 27
2.6 Conclusion ................................................... 30
References ....................................................... 30

3 Simultaneous Multiple-emitter Fitting for SM-SR Imaging .............. 34
   3.1 Introduction to Single-Molecule Super Resolution ................. 35
   3.2 Theoretical Basis for the Multi-emitter Fitting Algorithm ....... 38
      3.2.1 Multiple Emitter Extension to the Pixelized Single Emitter
            Model .................................................. 38
      3.2.2 Maximum Likelihood Estimator ............................. 39
3.3  The Analysis Procedure .............................................. 40
3.3.1  Image Pre-processing and Segmentation ...................... 41
3.3.2  Multi-emitter Fitting Analysis (MFA) ............................ 42
3.3.3  Model Selection .................................................. 45
3.3.4  Precision of the Estimated Parameters ....................... 46
3.3.5  Filtering and SR Image Reconstruction ....................... 46
3.4  Computational and Experimental Methods ....................... 47
3.4.1  Hardware and Software Implementation of Analysis Routines . 47
3.4.2  Estimator Precision and Algorithm Performance Testing .... 48
3.4.3  Synthetic Data Generation .................................... 49
3.4.4  SM-SR Imaging ................................................ 50
3.5  Results and Discussion ............................................. 51
3.5.1  Optimal Sub-region Size and N_{max} .......................... 51
3.5.2  Uncertainty Estimator Performance ........................... 53
3.5.3  Algorithm Performance Versus Density and Intensity Distribu-
       tion ........................................................ 54
3.5.4  Pattern Simulation Results .................................... 57
3.5.5  Algorithm Speed ............................................... 61
3.5.6  Imaging of Actin Structures .................................. 62
3.6  GPU Based MATLAB-MEX Programming ......................... 64
Contents

3.7 Conclusion .................................................................................. 66

3.8 Acknowledgements ...................................................................... 67

References ......................................................................................... 67

References 67

4 Rejection Algorithms for Single Emitter Fitting in SR Analysis 71

4.1 Introduction .................................................................................. 72

4.1.1 Why a Rejection Algorithm is Needed? ............................... 72

4.1.2 The Rejection Algorithm at Optimal Active Density in SM-SR
Experiments ....................................................................................... 75

4.1.3 Rejection Algorithm for Some Possible Complications ........ 78

4.2 Rejection Algorithm Review .......................................................... 79

4.2.1 Criteria of a Good Rejection Algorithm ............................... 79

4.2.2 Brief Review of Current Rejection Algorithms ................. 81

4.2.3 Elliptical Test Review ............................................................... 82

4.2.4 SSE Test Review ..................................................................... 84

4.2.5 Poisson Distributed Noise Model in PALM ....................... 85

4.3 Rejection Algorithms Demonstrations .......................................... 86

4.3.1 Acceptance Demonstration for Elliptical Test and SSE Test .. 86

4.3.2 Rejection Demonstration of Elliptical Test and SSE Test .... 90

4.3.3 Discussion on Demonstrations of Acceptance and Rejection .. 93
Contents

F Fisher Information Matrix for Poisson Distributed Data 141

G List of Software Recommended to Compile CUDA-MEX Files 143

H Included Publication 145
List of Figures

2.1 Experimental setup of our home built FCS system on inverted microscope (Eclipse TE200, Nikon). Core components consist of an argon ion laser (Ion Laser Technology 5425A-00C-2), the microscope, a photon-counting avalanche photodiode (APD) (SPCM, PerkinElmer Life Sciences), a 60 × , 1.4 NA oil immersion objective (Zeiss) and a auto/cross-correlator (Flex02-03D). . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .11

2.2 Demonstration of the principle of effective sample volume where fluorophore diffuses into the sample volume during FCS experiment. Sample volume, defined by the focus of the Gaussian beam and the pinhole before the APD, provide an extremely small excitation volume within the sample. Fluorescent proteins that move within the sample could potentially enter the sample volume and get excited into a fluorescent state during the dwell time within the sample volume. . . . 13

2.3 Demonstration of obtained auto-correlation function and its fitting to FCS function using free GFP in PBS buffer. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .17
2.4 F/N and Diffusion coefficient results obtained using Fluorescence Correlation Spectroscopy. FCS was performed on free GFP, WT-ERK-GFP and Δ4-ERK-GFP constructs respectively in both the nucleus and the cytoplasm of the live cells. In the table \textit{starved} indicates measurements before serum addition, and \textit{+serum} data were acquired in the cytoplasm from 1 to 10 min after serum addition and in the nucleus from 10 to 15 min. Values are given as average ± S.E. 27

2.5 Dimerization of GFP-ERK1-WT was not detected by FCS in either the nucleus and the cytoplasm. FCS was used to extract the fluorescence intensity per mobile object (F/N). No significant change in F/N was observed after stimulation for GFP-ERK1-WT, indicating that ERK1 remained monomeric after activation. Each symbol after activation indicates a single measurement. Connected symbols are measurements from the same cell over time. Values are normalized to the F/N value before stimulation and absolute values are found in previous table 2.4. The gray band represents the single cell standard deviation of pre stimulation measurements. 28

2.6 Dimerization of GFP-ERK1-Δ4 was not detected by FCS in either the nucleus or the cytoplasm. This was expected given the well-proved inability of dimerization of GFP-ERK1-Δ4. This experiment served as a negative control for F/N measurement on live cell. Connected symbols are measurements from the same cell over time. Values are normalized to the F/N value before stimulation and absolute values are found in Table 2.4. The gray band represents the single cell standard deviation of pre stimulation measurements. 29
List of Figures

3.1 Proximity of emitters as a function of emitter density. The probabilities of finding N=1-5 emitters within a $8\sigma_{PSF} \times 8\sigma_{PSF}$ square sub-region ($\sigma_{PSF} = 127$ nm) at different densities were calculated for a uniformly distributed population of emitters and plotted as a function of density. As the emitter density increases beyond $1 \mu m^{-2}$, the fraction of sub-regions containing single emitters reduces dramatically (red line), emphasizing the need for fitting algorithms that can accommodate multiple emitters within a single sub-region. 36

3.2 Illustration of execution steps in the multi-emitter estimation task. (a) Fitting algorithm flowchart. (b) through (e): Demonstration of the results from each estimation task from the 1 emitter model through the 4 emitter model. The 5 emitter model fitting is not performed by the algorithm, because of the low photon counts in the deflated image. 44

3.3 Single fluorophore intensity distribution of the organic fluorophore Alexa Fluor 647 obtained from the data set described in section 4.4.1 taken in TIRF condition. The distribution is modeled as a normal distribution with $\mu = 800, \sigma = 100$. 50
List of Figures

3.4 Performance of the precision estimate. (a) A comparison between the precision predicted from the CRLB and from the modified Fisher information matrix. A series of simulated images of two emitters at various separations between their centers were generated. MFA was performed on these images and the precision estimates calculated by the modified Fisher information matrices (F(\theta) Estimated Std. Dev.) were compared with that obtained from the CRLB (Estimated Uncertainty CRLB), precisions obtained from the CRLB generated by emitter’s true position (Theoretical Uncertainty CRLB), and the observed standard deviation of the estimates (Observed Std. Dev.). (b) The CDF (integral of histogram) of the uncertainty estimator accuracy obtained using the modified Fisher information matrices for random placements of multiple emitters. . . . . . . . . . . . . . . 53

3.5 Performance versus active emitter density and intensity distribution in low intensity case. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of 300 ± 30 (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes \( N_{\text{max}} = 5 \). The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position. 55
List of Figures

3.6 Performance versus active emitter density and intensity distribution in cases that the intensity distribution is similar as what we have observed from Alexa Fluor 647. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of 800 ± 100 (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes $N_{\text{max}} = 5$. The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position. ................................................................. 56

3.7 Performance versus active emitter density and intensity distribution for extremely high emitter intensity case. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of 5000 ± 30 (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes $N_{\text{max}} = 5$. The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position. ................................................................. 57

3.8 Simulated data set and its sum projection. (a) The emitter position histogram used in generating synthetic data. (b) Sum projection of the generated image which demonstrated the diffraction limited pattern observed under conventional microscope. ................................................. 58
List of Figures

3.9 Results from single emitter fit and multi emitter fit at the density of 1 \( \mu m^{-2} \). (a) Single emitter fitting result at a density of 1 \( \mu m^{-2} \) with \( N_{\text{max}} = 1 \). (b) Multiple emitter fitting result at a density of 1 \( \mu m^{-2} \) with \( N_{\text{max}} = 5 \). At 1 \( \mu m^{-2} \) case, \( N_{\text{max}} = 1 \) resulted in 12848 emitters localized while \( N_{\text{max}} = 5 \) localized 30354 emitters. The contrast of images (a) to (b) were globally adjusted across all images for optimal display. 59

3.10 Results from single emitter fit and multi emitter fit at the density of 6 \( \mu m^{-2} \). (a) Single emitter fitting result at a density of 6 \( \mu m^{-2} \) with \( N_{\text{max}} = 1 \). (b) Multiple emitter fitting result at a density of 6 \( \mu m^{-2} \) with \( N_{\text{max}} = 5 \). In 6 \( \mu m^{-2} \) case, \( N_{\text{max}} = 1 \) resulted in 519 emitters localized while \( N_{\text{max}} = 5 \) localized 33580 emitters. The contrast of images (a) to (b) were globally adjusted across all images for optimal display. 60

3.11 Comparison of SM-SR fitting routines for imaging the actin meshwork within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) images were obtained using conventional TIRF microscopy. Scales bars represent 5 \( \mu m \) in (a) and 1 \( \mu m \) in (b) 62

3.12 Comparison of SM-SR fitting routines for imaging the actin meshwork within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) SM-SR images were generated using \( N_{\text{max}} = 1 \) as a demonstration of single emitter fitting result. Actin rich regions, seen in top right of 3.11a are missing using single emitter routines (\( N_{\text{max}} = 1 \)) in (b). Scales bars represent 5 \( \mu m \) in (a) and 1 \( \mu m \) in (b) 63
List of Figures

3.13  Comparison of SM-SR fitting routines for imaging the actin meshwork within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) SM-SR images were generated using $N_{\text{max}} = 5$ as a demonstration of multi emitter fitting result. Actin rich regions, seen in top right of 3.11b, 3.12b, 3.13b are missing using single emitter routines ($N_{\text{max}} = 1$) 3.12b, but successfully fit using the MFA ($N_{\text{max}} = 5$) 3.13b. The increase in molecular density found using the MFA ($N_{\text{max}} = 5$) routine also reveals a more complete depiction of the underlying actin structure, outlining possible actin corrals seen in the center of 3.13a. Scales bars represent 5 µm in (a) and 1 µm in (b) . . . . . . . . . . 64

4.1  Histogram of ellipticity for isolated emitter sub-region. 1000 single emitter sub-regions corrupted individually by Poisson noise are fitted using bivariate Gaussian method and then ellipticity values are calculated after obtaining the eigenvalues from the resulting covariance matrices $\Sigma$. Histogram of ellipticity value for these 1000 sub-region is shown below where ellipticity threshold 15% is also shown in the same plot as a red dashed vertical line which indicates any sub-region that would result in a ellipticity value bigger than 15% (to the right of the threshold line) will be excluded from further analysis. . . . . 88

4.2  Histogram of SSE and a proposed threshold for isolated emitter sub-region. 1000 single emitter sub-regions corrupted individually by Poisson noise are fitted using LS fitting and then SSE value defined in Eq. 4.19 are plotted as histogram. A proposed threshold which associate with a possible failure rate 1600 for this specific case are shown as a red dashed line. SSE rejection algorithm will reject any SSE result from fitting that are larger than this threshold value (to the right of the red line) from further reconstruction process. . . . . 89
List of Figures

4.3 Histogram of ellipticity. 1000 sub-regions that are the same configuration which has 2 emitter with a distance of 2 pixel away from each other are analyzed using elliptical test method. Ellipticity obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the elliptical test. Threshold of 0.15 ellipticity are also shown in the figure as a red dashed line. Sub-regions that result in a larger than threshold ellipticity value would then be excluded from future analysis. 91

4.4 Histogram of SSE values. 1000 sub-regions that are the same configuration which has 2 emitter with a distance of 2 pixel away from each other are analyzed using SSE test method. SSE values obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the SSE test. Threshold SSE of 1600 are also shown in the figure as a red dashed line. Sub-regions that result in a larger than threshold SSE value are excluded from future analysis. 92

4.5 Histograms of ellipticity for simulated sub-regions. For every configurations ($d = 1, d = 2, d = 3, d = 4$ and $d = 5$), 1000 Poisson noised sub-regions that are the same configuration which has 2 emitter with a distance of $d$ pixel away from each other are analyzed using elliptical test method. Ellipticity obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the elliptical test. Threshold of 0.15 ellipticity are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold ellipticity value would then be excluded from further analysis. 94
List of Figures

4.6  Histogram of $SSE$ for simulated sub-region with single emitters series of various intensities and threshold the of 1600. For every configurations ($N = 300, N = 800, N = 1200, N = 2000$ and $N = 5000$), 1000 Poisson noised sub-regions that are the same configuration are analyzed using $SSE$ test. $SSE$ values obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the acceptance process of the $SSE$ test at various intensities. Threshold of $SSE = 1600$ are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold $SSE$ value would then be excluded from future analysis.

4.7  Histogram of LLR for simulated sub-regions with single emitters series of various intensities and a p-value threshold of 5%. For every configurations ($I = 300, I = 800, I = 1200, I = 2000$ and $I = 5000$), 1000 Poisson noised sub-regions are analyzed using LS fitting. LLR values obtained for each individually Poisson noise corrupted sub-region and its fitting model are plotted as histograms to illustrate the acceptance processes of the LLR test at various intensities. A p-value threshold of 5% are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than p-value threshold LLR value would then be excluded from future analysis.
List of Figures

4.8 Histogram of LLR obtained from fitting result and a 5% p-value threshold. For every configurations \((d = 1, d = 2, d = 3, d = 4\) and \(d = 5\)), 1000 Poisson noised sub-regions that are the same configuration which has 2 emitter with a distance of \(d\) pixel away from each other are analyzed using LS fitting to a 2-D Gaussian. LLR values obtained from the results for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the LLR goodness of fitting test. A 5% p-value threshold are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold LLR value would then be excluded from future analysis. . . . . . . . . . . . . . . . . . . . . 106

4.9 Numerical evaluation of \(\langle LLR \rangle\) as a function of \(\lambda\) using Eq. 4.30. \(\langle LLR \rangle\) starts from 0 as \(\lambda \rightarrow 0\) and overshoots to about 1.2 at \(\lambda = 1.33\) and then falls back down and approaches 1 as \(\lambda > 10\) . . . . . . . . 110

4.10 Histograms of LLR values obtained 5 different configuration with different sub-regions sizes. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as \(4 \times 4\), \(8 \times 8\), \(12 \times 16\), \(12 \times 16\) and \(20 \times 20\). Each sub-region contain only one emitter with \(\sigma_{PSF} = 1.2\) pixels and background=0.01. Thus with increasing sub-region size, the number of pixels that are low-count pixels increases in each configuration with sub-region size. Also shown on the plot is the \(\chi^2\) distribution with degrees of freedom \(K\) for each configuration which is expected to be the distribution of LLR values. . . . . . . . . . . . . . . . . . . . . 112
4.11 Histograms of LLR values obtained 5 different configuration with different sub-regions sizes with addition of the Poisson noised background and its corresponding $\chi^2_K$ distribution. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. Each sub-region are then added an additional Poisson noised background with expected count of 10 for each pixel. LLR are calculated using 4.22 and shown using histogram with the corresponding $\chi^2_K$ for each configuration.

4.12 Histograms of LLR values obtained 5 different configuration with different sub-regions sizes and its corresponding $\chi^2$ distribution with degrees of freedom predicted using Lidke-Huang hypothesis. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. LLR is then obtained using the method described this section, and plotted as histogram for each configuration. $\chi^2$ distribution with degrees of freedom predicted by LH hypothesis are also shown for each configuration.

4.13 Plot of numerically calculated $\langle LLR(m) \rangle$ and $LHA(m)$. LHA provided a functional form $LHA(m)$ for calculating $\langle LLR(m) \rangle$ at various $m$ values.
List of Figures

4.14 Histograms of LLR values obtained 5 different configuration with different sub-regions sizes and its corresponding $\chi^2$ distribution with degrees of freedom predicted using Lidke-Huang hypothesis and Lidke-Huang approximation. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{\text{PSF}} = 1.2$ pixels and background=0.01. LLR is then obtained using the method described in section 4.6.6, and plotted as histogram for each configuration. $\chi^2$ distribution with degrees of freedom predicted by LH hypothesis calculated by LH approximation are also shown for each configuration. 121

B.1 Focused laser spot on grating pattern 132
List of Tables

2.1 Effects of laser intensity on measured diffusion coefficient. . . . . . . 22

3.1 Time consumption and performance. Images (128×128×1000) of emitters that were randomly placed with uniform distribution are generated using two active emitter densities. Both were fitted by \( N_{\text{max}} = 1 \) and \( N_{\text{max}} = 5 \) and the speed for each part of the algorithm was recorded. Total numbers of fitting sub-regions are 146,597 and 336,959 in low and high density case respectively. . . . . . . . . . . . 61
<table>
<thead>
<tr>
<th>Glossary</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-D</td>
<td>Two dimensional</td>
</tr>
<tr>
<td>3-D</td>
<td>Three dimensional</td>
</tr>
<tr>
<td>APD</td>
<td>Avalanche Photo Diode</td>
</tr>
<tr>
<td>ACF</td>
<td>Auto-Correlation Function</td>
</tr>
<tr>
<td>ALU</td>
<td>Arithmetic Logic Unit</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Density Function</td>
</tr>
<tr>
<td>CRLB</td>
<td>Cramer-Rao Lower Bound</td>
</tr>
<tr>
<td>CCD camera</td>
<td>Charge-Coupled Device camera</td>
</tr>
<tr>
<td>CUDA</td>
<td>Compute Unified Device Architecture</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>DIC</td>
<td>Differential Interference Contrast microscopy</td>
</tr>
<tr>
<td>DSDNA</td>
<td>Double Stranded DNA</td>
</tr>
<tr>
<td>DI</td>
<td>Deionized Water</td>
</tr>
<tr>
<td>dSTORM</td>
<td>Direct Stochastic Optical Reconstruction Microscopy</td>
</tr>
</tbody>
</table>
Glossary

EM-CCD camera  Electron Multiplier Charge-Coupled Device camera
fluorophores  Fluorescent molecules
FIONA  Fluorescence Imaging with One Nanometer Accuracy
fPALM  Fluorescence Photo Activation Localization Microscopy
FRAP  Fluorescence Recovery After Photobleach
FLIP  Fluorescence Loss In Photobleaching
FCS  Fluorescence Correlation Spectroscopy
GFP  Green Fluorescent Protein
GSDIM  Ground State Depletion microscopy followed by Individual Molecule return
GPU  Graphics Processing Unit
HeLa cells  a cell line that was derived from cervical cancer cells taken from Henrietta Lacks, a patient who eventually died of her cancer on October 4, 1951.
LS  Least Square
LLR  Log-Likelihood Ratio
MLE  Maximum Likelihood Estimator
MFA  Multi-emitter Fitting Analysis
MEX  MATLAB Executable
NIH  National Institutes of Health
### Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSF</td>
<td>National Science Foundation</td>
</tr>
<tr>
<td>NR method</td>
<td>Newton-Raphson method</td>
</tr>
<tr>
<td>N.A.</td>
<td>Numerical Aperture</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PALM</td>
<td>Photo-Activated Localization Microscopy</td>
</tr>
<tr>
<td>STED</td>
<td>Stimulated Emission Depletion Microscopy</td>
</tr>
<tr>
<td>SSIM</td>
<td>Saturated Structured Illumination Microscopy</td>
</tr>
<tr>
<td>SR</td>
<td>Super Resolution</td>
</tr>
<tr>
<td>SM-SR</td>
<td>Single-Molecule Super Resolution</td>
</tr>
<tr>
<td>STORM</td>
<td>Stochastic Optical Reconstruction Microscopy</td>
</tr>
<tr>
<td>SSDNA</td>
<td>Single Stranded DNA</td>
</tr>
<tr>
<td>SSE test</td>
<td>Sum Squared Error test</td>
</tr>
<tr>
<td>TIRF(M)</td>
<td>Total Internal Reflection Fluorescence (Microscopy)</td>
</tr>
<tr>
<td>YFP</td>
<td>Yellow Fluorescent Protein</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction to Fluorescence Microscopy

Ever since the Greek and Roman period around 1000 A.D., humans used specially shaped spherical surfaced glasses to magnify structures which would otherwise be too small to be distinguished by the human vision system by itself [1]. However it was not until the 1700s that the first compound microscope (multiple lenses) was developed by Zacharias Jansen and John Lipperhey in the Netherlands [1]. Based on the compound microscope, developments have continued to improve its usability and magnification by perfecting lens craftsmanship from the 18th century until now [2]. The most common type of optical microscopes nowadays usually consists of the illumination component, a sample stage, an objective and an eye piece. Samples to be examined under microscope are usually illuminated by the transmitted white light generated by the illumination component and the transmitted light/reflected light is then collected and magnified by the objective and then passed into eye piece for viewing purpose. Current microscope setup usually includes a side port together with an eye piece system so that users could either view the magnified sample by their eyes or collect the image using a camera attached on the side port.
Chapter 1. Introduction to Fluorescence Microscopy

One of the major drawbacks of illumination based, bright field microscopy is the inability to distinguish the target from the surrounding environment. In other words, bright field microscopy lacks signal to background ratio or contrast especially for thin samples that are around microns thick. Developments have been made to increase the contrast in illumination based microscopy. Popular modifications of illumination based microscopy to enhance contrast include Dark Field microscopy, where optics are designed such that only the scattered light from the sample is collected, Phase Contrast microscopy where the phase differences of the transmitted light introduced by differences of sample thickness are converted into intensity of light and thus gained the ability to differentiate index of refraction differences in very thin samples, and Differential Interference Contrast microscopy (DIC) where interferometry principle is used to gain additional contrast for various index of refraction differences across the sample [3, 4].

Together with all these beautiful discoveries in development of illumination based, bright field microscopy, fluorescence microscopy was born. British mathematician and physicist Sir George G. Stokes first described fluorescence in 1852, where he described the phenomenon of fluorescence, as exhibited by fluorspar and uranium glass [5]. He found these materials had the ability to absorb the invisible ultra-violet light (shorter wavelength) and emit a longer wavelength light, which is now called ‘Stocks Shift’ in honor of his discovery in fluorescence. Fluorescence microscopy takes advantages of the fluorescence property for certain materials, which we called ‘fluorescent dyes’ or ‘fluorohores’, in order to further enhance the contrast between sample and the surrounding background. By tagging a biological sample with fluorescent dyes, early scientists have used fluorescence microscopy to observe the emitted fluorescence from the sample using illumination of ultra-violet light source which is invisible to human eyes. Thus fluorescence which generated by the fluorescently tagged sample provides unsurmountable signal to noise ratio compared with techniques in illumination based bright field microscopy. With the development of dye
Chapter 1. Introduction to Fluorescence Microscopy

and filter systems, excitation sources and dye categories that could be used in fluorescence microscopy have increased dramatically. Currently fluorescence microscopy has permeated all cell and molecular biology fields. Together with the illumination based microscopy, electron microscopy and other microscopy methods, fluorescence microscopy has played an important role in the advancement of cell biology studies.

Among all fluorescence microscopy varieties, the most popular ones are wide field fluorescence microscopy, confocal microscopy, two-photon microscopy and total internal reflection microscopy. Wide field fluorescence microscopy is implemented using the simplest fluorescence microscopy method where the sample is illuminated using an excitation lamp or laser and then the emitted fluorescence from the sample are passed through the filter which are used to block transmitted or scattered excitation light and collected by an eye piece or a mounted camera. Since a large area of the sample in both lateral and axial direction are illuminated by the illumination lamp, out of focus fluorescence blur greatly decreased the resolution and the signal to noise ratio wide field fluorescence microscopy could achieve. At the same time, as the sample plane is essentially a 2-D plane, it’s very hard for wide field fluorescence microscopy to obtain information along axial axis. Confocal microscopy, which uses a focused illumination beam and a pinhole-detector setup together with scanning optics provides a solution for 3-D fluorescence microscopy. The focused laser excitation beam together with a pin-hole setup creates an effective detection volume which is about hundreds of nanometers in the lateral direction and couple microns in axial direction. Thus by scanning this sample volume across the biological sample in 3-D, scanning confocal microscopy enables scientist to obtain a 3-D fluorescence images of the sample with high resolution. Two-photon microscopy, by using the non-linearity of two photon excitation, employs a similar idea as confocal microscopy by limiting the excitation and detection volume to a tiny volume in order to increase resolution. This small effective excitation/detection volume is then scanned across the sample to create 3-D sample images. However, scanning process is highly speed limited in
both confocal and two photon microscopy and thus the time resolution is sacrificed in order to gain the spatial resolution in 3-D. Total Internal Reflection Fluorescence microscopy (TIRF) was introduced by Daniel Axelrod as an alternation to scanning microscopy [6]. TIRF uses the evanescent wave to selectively illuminate and excite fluorophores in a surface region of the sample immediately adjacent to the glass-water interface by directing the excitation beam through the coverslip-water surface with an incidence angle greater than the critical angle. This thin layer of excitation from the evanescent field (usually cited as around 200 nm, however it depends on the angle of incidence) provides a very restricted excitation volume and thus fluorophores within the sample solution but away from the thin excitation volume will not be excited. The contrast and signal to noise ratio provided by TIRFM together with its scanning-free system are appreciated by cell biologist especially in membrane protein studies. Although signal to noise ratio is maximized using fluorescence microscopy, the fundamental resolution barrier of optical microscopy that is caused by the wave property of light still prevented studying individual protein at the single molecule level.

Due to the diffraction limit experimentally described by Ernst Karl Abbe in 1874, the resolution limit for the optical microscope is defined by its numerical aperture, which is defined as $N.A. = n \sin \theta$, by equation.

$$d = \frac{\lambda}{2N.A.} \quad (1.1)$$

While this equation is believed to be first described by Joseph Louis Lagrange in the 18th century, Abbe is the first scientist who experimentally demonstrated and explained the diffraction limit. Due to the diffraction limit, small object such as organic dyes or fluorescent proteins that are around tens of nanometers in diameter would appear as a relatively large Airy disc pattern of around 200 nm when imaged
through the objective in the microscope system. The size of this pattern varies with the Numerical Aperture (N.A.) and wavelength of light. For example, Green Fluorescent Protein (GFP) whose emission peak wavelength is around 520 nm passing through a high numerical aperture objective with $N.A. = 1.45$ will result in a diffraction limited spots with a diameter around 350 nm, which is much larger than the size of the GFP (around 4 nm hydrodynamic size).

This barrier was believed to be an insurmountable limitation of optical microscopy until the recent development of various super resolution microscopy techniques. Two general types of techniques that have been developed to overcome this barrier are illumination based super resolution technique and single molecule based super resolution technique. Stimulated emission depletion microscopy (STED), ground state depletion microscopy (GSD), and saturated structured illumination microscopy (SSIM) are the super resolution techniques based on alternating the effective excitation light pattern by using the nonlinear response of fluorophores to excitation laser in order to achieve an effective illumination spot that are much smaller than the diffraction limit [7]. The other types of super resolution technique are based on single molecule localization and is usually referred to as single-molecule super resolution techniques (SM-SR).

The concept of single molecule localization was introduced as early as 1981 by Barak and Webb in single particle tracking experiments [8]. It was not until 2002 when Thompson et. al. proposed a functional estimation on localization uncertainty that quantitative localization analysis of single molecule started to attract the attention of scientists from various fields [9]. Fluorescence imaging with one nanometer accuracy (FIONA), which is introduced by Selvin et.al.in 2003, demonstrated the idea that by localizing single emission spots, one could localize the center of the diffraction limited spot with a certain precision of estimation, and thus provide an estimation of the emitter center with a high localization precision [10]. Localization
Chapter 1. Introduction to Fluorescence Microscopy

techniques can only be done on isolated fluorophores, which would benefit single molecule studies where the labeling density are sparse, however continuous structures of interests would not be revealed due to a low sampling density. Lidke et.al.in 2005, introduced an idea by using blinking behavior of quantum dots to distinguish single emitter events from highly labeled structures [11].

Methods with similar ideas that were more easily applied to cellular imaging were then published in the literature such as stochastic optical reconstruction microscopy (STORM), photo-activated localization microscopy (PALM), fluorescence photo activation localization microscopy (fPALM), direct stochastic optical reconstruction microscopy (dSTORM) and ground state depletion microscopy followed by individual molecule return (GSDIM). As the above methods all related to isolate and localize single molecule, it is in general called ‘Single-Molecule Super Resolution’ (SM-SR) techniques [12].

These techniques have revolutionized fluorescence microscopy with the ability to achieve a resolution as low as 20 nm, which is an order of magnitude improvement from conventional optical microscopy that is limited by diffraction to $\lambda/2NA$ or approximately 300 nm.

During the years of my study in UNM, I have used confocal fluorescence microscopy techniques described above in order to investigate specific biology questions and also developed and implemented new analysis methods that advance the localization based SM-SR techniques. In this thesis, first, I describe the work I have done on using Fluorescence Correlation Spectroscopy (FCS) with the confocal microscope to investigate the dimerization model for ERK1 in nucleus translocation. I then described the multi-emitter fitting algorithm that we have developed for SM-SR technique that relaxes several important restrictions applied on this technique. Finally, I proposed a log-likelihood ratio statistic to be used in a robust and consistent rejection algorithm that is implemented in SM-SR techniques and I showed
Chapter 1. Introduction to Fluorescence Microscopy

using simulated data set that our proposed rejection algorithm is consistent, robust
and statistically explainable compared with current rejection algorithms that are
commonly used in SM-SR literature.

The following three chapters of my thesis are based on two published peer-reviewed
papers and one manuscript in preparation at LidkeLab in UNM. These included pa-
pers can be found in Appendix H.

References

[1] B. O. Payne, Microscope design and construction (Cooke, Troughton, Simms,
1954).


macher, “Light microscopy,” Techniques in microscopy for biomedical applica-


fluorescence,” Annual Review of Biophysics and Bioengineering 13, 247–268


calization analysis for individual fluorescent probes,” Biophysical Journal 82,
2775–2783 (2002).
Chapter 1. Introduction to Fluorescence Microscopy


Chapter 2

The Evaluation of ERK1 Dimerization Models using FCS

Fluorescence microscopy has provided biologists an important tool to study cell structure and protein movements. Compared with transmitted light based microscopy, fluorescence microscopy provides a low background and high signal to noise ratio approach to observe cell structures under the microscope. With recent developments of fluorescent proteins such as, GFP [1], YFP [2] and mCherry [3]...etc, fluorescence microscopy can be performed in an almost non-invasive way using transfection where genetic modification can be done to the DNA of a cell in order to make it express the fluorescent protein on its own. By using genetically expressed fluorescent proteins, fluorescent imaging can be done in live cells with minimum disturbance toward the target [4]. In fluorescence microscopy, fluorescence emitted from the fluorescent objects –fluorophores (emitters), is collected by either a camera such as EM-CCD camera which is set on the image plane of the microscope for wide field setup, or a single photon detector such as an avalanche photo diode (APD)-pinhole combination for confocal setup [5]. Quantitative analysis can be done on the data collected from the camera or APD to help understand the underlying biological behaviors.
Diffusion coefficients [6] of specific biologically interesting proteins are considered an important parameter for understanding biological behaviors within the cell. Fluorescence recovery after photobleaching (FRAP) [7, 8] and fluorescence loss in photobleaching (FLIP) [9, 10], provide two methods to measure diffusion coefficients of an ensemble [10]. In FRAP, a bright laser beam is focused by the objective and scanned across a certain region of a cell, fluorescent proteins or organic dyes within that region will get photobleached after the scanning process and appear as a dark region immediately after scanning, when viewed with a dim excitation laser. Due to the random diffusion behavior of the fluorescently tagged protein, the fluorescence intensity of photobleached area will recover after a certain period of time and the recovery curve can be extracted and analyzed to obtain an ensemble diffusion coefficient of the tagged protein [7]. In contrast to FRAP, FLIP detects the fluorescent loss in the unbleached area next to the bleached area caused by interchange between fluorescent protein and photobleached protein between photobleached area and unphotobleached areas [11].

While these measurements provide methods for estimating diffusion coefficients for proteins of interest, the measured diffusion coefficient can only be considered as a measure in an ensemble level [10]. To measure the diffusion coefficient at the molecular or the single protein level, one could use Fluorescence Correlation Spectroscopy (FCS) [12, 13].

We have implemented a FCS setup on a home-built confocal microscope and proposed a F/N analysis method which stands for the ratio between average fluorescence intensity and average number of moving object that could be used to detect dimerization. We then investigated ERK1 diffusion behaviors after stimulation and also verified the dimerization model with F/N analysis using FCS.
2.1 Introduction to FCS

Fluorescence correlation spectroscopy (FCS) introduced by Magde, Elson and Webb in 1972 [14] has evolved for many years and the current FCS measurement is usually performed in a confocal setup [15, 16].

![Experimental setup of our home built FCS system on inverted microscope (Eclipse TE200, Nikon). Core components consist of an argon ion laser (Ion Laser Technology 5425A-00C-2), the microscope, a photon-counting avalanche photodiode (APD) (SPCM, PerkinElmer Life Sciences), a 60 ×, 1.4 NA oil immersion objective (Zeiss) and a auto/cross-correlator (Flex02-03D).](image)

The excitation laser beam is guided through the side port of the microscope and aligned parallel to the optical axis of the microscope before being guided into the
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

objective. The excitation laser is focused onto the sample plane of the microscope by the objective and thus provides a excitation profile on the sample plane. Fluorescence emitted from the fluorescent proteins or dyes are then collected using the same objective and guided back out through the side port of the microscope. A dichroic mirror is used to split the excitation and emission light where excitation is reflected into the side port and emission are allowed to pass back through and be focused into pinhole-APD setup. The pinhole placed before the detection devices is used for the exact purpose as used in confocal setup—for exclusion of out of focus light. The effective detection volume after the pinhole is smaller than the excitation profile and modeled as an elongated 3D Gaussian centered at the focal plane with z axis being its longer axis and x,y being its shorter axis [17].
Figure 2.2: Demonstration of the principle of effective sample volume where fluorophore diffuses into the sample volume during FCS experiment. Sample volume, defined by the focus of the Gaussian beam and the pinhole before the APD, provide an extremely small excitation volume within the sample. Fluorescent proteins that move within the sample could potentially enter the sample volume and get excited into a fluorescent state during the dwell time within the sample volume.

Fluorescent objects are usually considered infinitely small compare to the sample volume [13]. While they diffuse within the sample, a very small number of them will diffuse into or out of the sample volume in a fairly quick manner, on the time scales of microsecond or millisecond depending on their sizes and viscosity of the solution. When they pass through the sample volume, they are excited by the focused laser and therefore start to emit fluorescence. The emitted fluorescence can be detected by the APD continuously until the fluorophore exits the sample volume. During
this dwell time of the fluorophore, emitted fluorescence increases the intensity and this increase in total intensity is detected by the APD. The auto-correlation function is then extracted from the intensity history and analyzed to obtain the dwell time of the fluorophore within the sample volume. The extracted dwell time provides reliable information for diffusion coefficient of the specific tagged protein on a single molecule level [18].

2.2 Experimental Setup and Data Analysis

2.2.1 Experimental Setup

As shown in Fig. 2.1, our homebuilt FCS setup consists of an excitation component, a microscope component, and a detection component.

The excitation component is comprised of a Argon ion laser which emits at the wavelength of 488 nm, a narrow pass interference filter (CVI-Melles Griot) to better select the wavelength output from the laser, and a laser amplitude stabilization system (370 Conoptics). The output laser beam from the excitation component is a collimated laser beam whose wavelength is sharply peaked around 488 nm.

The excitation laser beam is then focused onto the back focal plane (conjugate image plane) of the microscope by a concave lens and a dichroic mirror is used to guide the laser beam into the side port of the microscope. The excitation laser beam again becomes collimated after the tube lens located on the bottom of the microscope, and then passes into the objective. The objective then focuses the excitation beam onto the sample plane.

The fluorescence emitted by fluorophores is then transmitted back into the objective, through the tube lens and then out of the microscope. A long pass dichroic
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

mirror which allows a 97% transmission for the fluorescence around 525 nm was used to reflect the excitation laser and transmit the longer wavelength fluorescence. Another lens was used to focus the emission beam on to the pinhole-APD setup. In our setup, we used a pinhole-APD (APD with a small detection area) instead of pinhole and APD setup. The APD has a detection area around 50 um, which could be considered as equivalent setup compared to the pinhole and APD setup. Small detection area on the APD behaves in a similar way as the pinhole where the out of focus emission fluorescence is rejected and in focus emission fluorescence is passed through the small area and detected by the APD.

2.2.2 Data Acquisition

Single photon events that are acquired by the APD are then collected by an autocorrelator which collects the photon history and converts it into the autocorrelation function with user defined time resolution. An autocorrelator card then outputs the autocorrelation function calculated from the signal and writes as a *.sin file on the computer connected to it. The autocorrelation function is defined as

\[ G(\tau) = \frac{\langle F(t) \cdot F(t + \tau) \rangle}{\langle F(t) \rangle^2} \]  

(2.1)

where \( G(\tau) \) is the autocorrelation function, \( \tau \) is the correlation time, and \( F(t) \) is the intensity of fluorescence at time \( t \).

2.2.3 FCS Diffusion Coefficient Analysis

Given the autocorrelation function described in Eq. 2.1. An analytical solution can be obtained by following the derivation in [19] which gives:
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

\[ G(\tau) = \frac{1}{\langle N \rangle} \cdot \frac{1}{1 + \frac{\tau}{\tau_D}} \cdot \frac{1}{\sqrt{1 + k^2 \cdot \frac{\tau}{\tau_D}}} \]  

(2.2)

Where \( \langle N \rangle \) is the average number of particles within the sample volume, \( k \) is a fixed parameter that accounts for the differing axial and lateral dimensions of the sample volume, \( \tau_D \) is the lateral diffusion time that a molecule stays in the focal volume. We could then derive the diffusion coefficient from \( \tau_D \) using:

\[ D = \frac{r_0^2}{4 \cdot \tau_D} \]  

(2.3)

where \( r_0 \) is the beam waist of the sample volume.

Given the analytical form for the autocorrelation function in Eq. 2.2, one then fits this function to the autocorrelation curve obtained from the experiment in order to obtain the diffusion coefficient and average number of molecules within the sample volume.
Fig. 2.3 showed a typical FCS experimental data and its functional fit to Eq. 2.2. ACF stands for autocorrelation function which is as defined in Eq. 2.1. The FCS measurement was performed on free GFP in PBS buffer for the demonstration. By fitting the ACF, the diffusion coefficient D and the average number of particle \( \langle N \rangle \) can be obtained. Local concentration can also be calculated given known \( r_0 \) and k.
2.2.4 F/N Dimerization Analysis

As shown in previous section, the average number of fluorescent objects within the sample volume $\langle N \rangle$ is obtained at the same time by fitting the autocorrelation function 2.2. In the same data set, the average intensity $\langle F \rangle$ could be obtained as well. One could then calculate the F/N value which would then provide a measurement of average brightness of each moving object.

$$F/N = \frac{\langle F \rangle}{\langle N \rangle}$$ (2.4)

This brightness per moving object (F/N) are then used later for monitoring dimerization events in live cells. During the dimerization event, two proteins both tagged with fluorescent proteins on each of them will form a dimer and then move as an entire entity. This entity is recognized in our FCS experiment as a single moving fluorescent object which has a brightness two times of of a single fluorescent protein as being a dimer state with 2 fluorescent proteins. The conformation changes can be detected by monitoring the changes in F/N value and thus used as a method to detect dimerization events.

Although F/N analysis provides a factor of two difference between a dimer system and a monomer system, the behavior of F/N in systems with mixture of dimers and monomers are not defined yet. One can describe a multi-component FCS where FCS is performed on a system that has multiple non-interacting species as:

$$G(\tau) = \frac{1}{(\sum_k Q_k N_k)^2} \sum_j Q_j^2 N_j \frac{1}{\left(1 + \frac{\tau}{\tau_{D_j}}\right)} \cdot \frac{1}{\sqrt{1 + k^2 \cdot \frac{\tau}{\tau_{D_j}}}}$$ (2.5)

Where $Q_k$ is the quantum yield for species $k$, $N_k$ is the particle number of species $k$ and $\tau_{D_i}$ is the lateral diffusion time for species $i$ [20].
It is assumed that the dimer, which is formed by 2 monomers can still be approximated as a sphere. Thus the radius differences between monomer and dimer is relatively small and is around 1:1.2 for monomer:dimer. We could then assume that lateral diffusion times $\tau_D$ are almost the same for dimer and monomer. Thus after comparing with Eq. 2.5 the apparent number of molecules detected using FCS in this system by fitting Eq. 2.2 given that the ratio of quantum yield between monomer and dimer are 1:2, can be described as

$$\frac{1}{N_{\text{app}}} = \frac{N_1 + 4N_2}{(N_1 + 2N_2)^2}$$

(2.6)

Where $N_{\text{app}}$ stands for apparent number of molecules within the sample volume obtained from fitting the single component FCS formula 2.2 to the data obtained the FCS experiment performed on system with a mixture of dimers and monomers. $N_1$ is the number of monomers in sample volume and $N_2$ is the number of dimers in the sample volume in the mixture system.

We then consider a system that initially only contains monomers and then, due to a certain system change, a fraction of the monomers in the system dimerized. Assuming the fraction of dimerization is $\alpha$, then $N_1 = (1-\alpha)N$ and $N_2 = \frac{\alpha}{2}N$, where $N$ is the average number of monomers in the sample volume in the system initially. Substitute the above calculation into Eq. 2.6, we could then find out,

$$\frac{1}{N_{\text{app}}} = \frac{1 + \alpha}{N}$$

(2.7)

where $\alpha$ is the fraction of dimerization and $N_{\text{app}}$ stands for apparent number of emitters obtained directly from fitting ACF obtained from FCS measurement from the system. $N$ stands for total number of monomers before dimerization. During
dimerization the total number of individual fluorophore doesn’t change, thus the F/N value would increase proportionally with increasing dimerization portion $\alpha$. Therefore in a system that only contains monomers and starts to dimerize after a certain system change, we are expecting a linear increase in its F/N value corresponding to its dimerization fraction.

### 2.3 Complications Encountered during FCS Measurement

#### 2.3.1 Laser intensity fluctuations

The detected fluorescence intensity is linearly dependent on the excitation laser intensity focused on the sample plane. Variations in excitation laser intensity will directly influence the autocorrelation curve obtained. We have observed a sinusoidal variation pattern for our excitation laser intensity in our setup (data not shown) before we installed a laser stabilizer.

To obtain the laser intensity variation curve, a mirror is put on the sample stage of the microscope while the excitation laser after the objective should focus exactly on to the sample plane which is also the reflection surface of the mirror. The power of the excitation laser is usually pretty high and the reflected beam contains almost 100% of the laser power resulting in a high laser power reflected back through the eye piece or side port of the microscope. A Neutral density filter of $ND = 5$ reduces the laser power to $10^{-5}$ of the input intensity. After setting up the neutral density filter, by looking at the focus spot in eye piece, we can see that the focused laser spot from the objective and the reflection surface of the mirror are overlapping with each other which is achieved in the alignment process of the confocal setup. The
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

excitation laser is then reflected back by the mirror and detected by the APD that is on the conjugate image plane of the confocal setup. The intensity history detected by the APD could then be recorded and plotted over time to extract the fluctuation information about the excitation laser intensity.

Laser fluctuations observed in this manner could result from several sources. The fluctuation might directly come from the variation of the laser source or it could come from the spatial instability of the laser beam which will also cause this detected fluctuation in our confocal setup. Another source that is also important is the room temperature. Fluctuations in room temperature in the specific location could also result in changes of behaviors of our optical system and confocal setup is considerably sensitive to this type of changes as focused excitation beam are usually around the size of microns. Changes due to this manner could be avoided by using a advanced room temperature control system (suggested by Travis Gould, Yale University, 2011) to maintain the room temperature within $\pm 1^\circ C$, however it’s not possible to reach this requirement at the time of this experiment due to realistic issues.

Tests were performed after installing the laser stabilizer (370 Conoptics) (data not shown) and the detected intensity fluctuations from laser are gone after warming up the laser for at least 1 hour. This suggested that the fluctuation observed for the laser intensity mostly comes from the variation of laser intensity from the laser source and also the temperature variation when the laser source is warming up. This brought us an idea that the temperature changes might make the measurement invalid when we perform FCS on live cells. Therefore when we performed FCS experiments on live cells, extra caution was executed to minimize the temperature change during experiment by turning off the AC system within the darkroom. To give a stabilized laser intensity, the laser is turned on at least 1 hour before starting the cell experiments and laser intensity fluctuation history is recorded before cell experiment.
2.3.2 Photobleaching and Blinking

Fluorescent dyes or fluorescent proteins in their excited state have a certain probability to either undergo an irreversible destruction or enter a long lifetime dark state. The fluorescence of the fluorophore turns off immediately after either of these events and would result in mistakenly shorter diffusion time from FCS measurement. Thus FCS measurement in this situation would provide a larger diffusion coefficient than the true diffusion coefficient. Although these processes are yet to be understood, the rate of the fluorophore to be photobleached or to enter the dark state is nearly proportional to the excitation intensity.

To ensure that photobleaching did not cause misinterpretation of FCS measurements [21], the effect of laser intensity was studied, using GFP in phosphate buffered saline (PBS) solutions. The laser intensity at the sample was controlled using neutral density filters.

<table>
<thead>
<tr>
<th>Laser Intensity (kW/cm²)</th>
<th>40</th>
<th>10</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>D obtained from FCS (µm²/s)</td>
<td>80.02±1.71</td>
<td>45.24±2.63</td>
<td>44.64±3.00</td>
</tr>
</tbody>
</table>

Table 2.1: Effects of laser intensity on measured diffusion coefficient.

Where D stands for diffusion coefficient. As shown in Table 2.1, although 40 kW/cm² gave an artifactually fast diffusion estimate, at intensities below 10 kW/cm² the diffusion estimate was stable and reproducible. The diffusion coefficient of 45 µm²/s gives a hydrodynamic GFP diameter of 4 nm, consistent with literature estimates. 4 kW/cm² was chosen for later experiments which are performed on cells.
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

2.4 Control Experiments

2.4.1 Diffusion Coefficient Measurement Control

Control experiments were done for our FCS setup using fluorescein tagged (5’ end) single stranded DNA (SSDNA)(Alpha DNA, Montreal, Sequence: fluorescein-5 ’A G G A C C C A G A A T C C G A T A G G A C C G 3’) in different solution environments. Notice that glycerol-water solution can provide viscosities from 0.001 Pa·s to 1.5 Pa·s. A mixture of glycerol and DI water is used. Two different buffers are made from glycerol and water, with ratios of water to glycerol as 0.33:1 and 0:1 (pure water) respectively. According to literature, the viscosity for these two solution systems have a ratio of 3:1 at room temperature [22]. Thus the diffusion coefficient for our fluorophores in these two buffer systems are expected to have a ratio around 1:3 [6]. FCS measurement were performed on these two systems and $\tau_D$ are obtained by fitting the autocorrelation function with theoretical FCS function. For the system that contains both water and glycerol with a volume ratio of 0.33:1, the measure $\tau_D$ is 1.4 ms, while for pure water, the measured $\tau_D$ is 0.43 ms. This observation is consistent with what we have expected due the viscosity difference for these two buffer system. However, noticing that Einstein relationship assumes a spherical object [6], the SSDNA isn’t necessarily spherical in both buffer system. In fact due to the difference of hydrophobic properties in these systems, the (average) shape of our SSDNA might vary in these two buffer systems which would introduce slight changes in diffusion coefficient based on the shape of the object [23]. In spite of these possible complications, this control demonstrated that FCS is reliable in measuring diffusion coefficient on single molecule level.
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

2.4.2 F/N dimerization control

In the previous section, we have proposed a measure of average brightness per moving object (F/N) as a dimerization indicator in FCS measurements. F/N controls were done using single stranded DNA (SSDNA) (Alpha DNA, Montreal) tagged with fluorescein at 5end and the double stranded DNA with fluorescein at each end which is annealed by heating SSDNA with its complementary strand at 86°C and slowly cooling to room temperature. Sequence of the two complementary SSDNA used in this control experiment are: fluorescein- 5’A G G A C C C A G A A T C C G A T A G G A C C G 3’ and fluorescein- 5’ C G G T C C T A T C G G A T T C T G G G T C C T 3’.

<table>
<thead>
<tr>
<th>Fluorescein concentrations</th>
<th>0.7 µM</th>
<th>0.4 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA type</td>
<td>DSDNA</td>
<td>SSDNA</td>
</tr>
<tr>
<td>F/N value</td>
<td>500±14</td>
<td>250±2</td>
</tr>
<tr>
<td></td>
<td>DSDNA</td>
<td>SSDNA</td>
</tr>
<tr>
<td></td>
<td>432±6</td>
<td>205±12</td>
</tr>
</tbody>
</table>

FCS Experiments were performed in PBS solution with laser intensity of 10 kW/cm². As shown in table above, F/N values showed a 2.05 ± 0.06 fold difference between DSDNA and SSDNA which is exactly what we have expected for the F/N ratio between single-tagged object (SSDNA) and double-tagged object (DSDNA).

2.5 Dimerization Test on Erk1 Protein after Stimulation

Stimulation of numerous cell surface receptors leads to activation of Raf/MEK/ERK signaling pathway. In this kinase cascade, Raf phosphorylates MEK, and MEK phosphorylates ERK which is primarily located in the cytoplasm of resting cells [24]. Upon phosphorylation of ERK, ERK accumulates in the nucleus of the cell which
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

is required for cell cycle entry [25, 26]. Dimerization of ERK has been proposed as a requirement for this nuclear translocation [27]. In order to evaluate the ERK1 dimerization model, FCS experiments were performed on both wild type (WT) GFP-ERK1-WT and mutant type (∆4) GFP-ERK1-∆4 in timed stimulation experiments. Its mutant GFP-ERK1-∆4 which has been shown consistently dimerization-deficient in vitro [27] is also tested using FCS as a comparison.

2.5.1 Experimental Setup

Fluorescence correlation spectroscopy (FCS) was performed with an inverted microscope (Eclipse TE200, Nikon) using the 488-nm line (narrow pass interference filter, CVI-Melles Griot) of an Argon ion laser (Ion Laser Technology, 5425A-00C-2) as the excitation source. The beam intensity was stabilized with a laser amplitude stabilization system (370 Conoptics) as described above. Fluorescence emission was detected with a photon-counting avalanche photodiode (SPCM-AQR, PerkinElmer Life Sciences). Digital signal correlation was performed using a Flex02-03D auto/cross-correlator (Correlator.com). A 60×, 1.4 NA oil immersion objective (Zeiss) was used. Correlation data at times shorter than that corresponding to the laser stabilization feedback frequency (200 kHz) were discarded. The beam waist, $r_0$, measured using a variation of the knife edge scan method [28], was 0.266 μm. The axial acceptance was determined to be 4 μm, by scanning a thin fluorescent film in z. FCS correlation curves were fit with MATLAB using the Gaussian acceptance approximation [19] shown in Eq. 2.2.

All measurements were performed with a laser intensity estimated as 4 kW/cm² in the sample plane obtained from measurements with an optical power meter (model S110, Thorlabs) and estimations of image field. Brightness (F/N), was used to monitor the aggregation state of ERK. The number of molecules (N) per volume was
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

used to calculate GFP-ERK concentration. Because the axial acceptance of FCS measurements was larger than the height of the cells. We performed electron microscopy measurements of starved MEF\textsuperscript{ERK1−/−} cells, leading to estimates of nuclear and cytoplasmic thickness, 2.2 ± 0.4 and 1.7 ± 0.6 \(\mu\)m, respectively. The calculation of concentration were corrected accordingly.

2.5.2 Cell preparation (Performed by Dr. Diane S. Lidke)

Full-length cDNA of human ERK1 subcloned into pcDNA3 vector (Invitrogen) was subjected to site directed mutagenesis with QuikChange site-directed mutagenesis kit (Stratagene) to remove the codons of amino acids Pro\textsuperscript{193} − Asp\textsuperscript{196} and generate ER1-Δ4. GFP-ERK1-WT and GFP-ERK1-Δ4 were generated by PCR cloning of human ERK1 or human ERK1-δ4 into PCR2.1 TOPO (Invitrogen) followed by ligation into XhoI and BamHI sites of pEGFP-C1 vector from Clonetech.

Mouse embryo fibroblast ERK1 knockout cells (MEF\textsuperscript{ERK1−/−} [29]) were cultured in DMEM +10% fetal calf serum and were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer’s protocol. Transient transfections were used within 48 h. For microscopy, cell monolayers were cultured in 8-well Lab-Tek Chamber slides (Nunc, Rochester, NY). Cells were starved (0 or 0.1 % serum) for 4-12 h. Live cell imaging was carried out in Tyrode’s buffer with 20 mM glucose and 0.1% bovine serum albumin, and samples were maintained at 34-36 °C by an objective heater (Bioscience Tools).
Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS

2.5.3 Results

ERK1 Diffusion Behavior Measured By FCS

<table>
<thead>
<tr>
<th></th>
<th>( F/N )</th>
<th>( D )</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>counts/object/s</td>
<td>( \mu m^2/s )</td>
<td></td>
</tr>
<tr>
<td>GFP cytoplasm</td>
<td>319 ± 10</td>
<td>17.5 ± 1.3</td>
<td>7</td>
</tr>
<tr>
<td>GFP nucleus</td>
<td>308 ± 8</td>
<td>16.1 ± 0.7</td>
<td>18</td>
</tr>
<tr>
<td>WT cytoplasm starved</td>
<td>316 ± 8</td>
<td>7.7 ± 0.4</td>
<td>31</td>
</tr>
<tr>
<td>WT cytoplasm + serum</td>
<td>308 ± 5</td>
<td>6.6 ± 0.3</td>
<td>49</td>
</tr>
<tr>
<td>WT nucleus starved</td>
<td>283 ± 14</td>
<td>7.1 ± 0.8</td>
<td>13</td>
</tr>
<tr>
<td>WT nucleus + serum</td>
<td>299 ± 11</td>
<td>6.8 ± 0.6</td>
<td>19</td>
</tr>
<tr>
<td>( \Delta 4 ) cytoplasm starved</td>
<td>346 ± 8</td>
<td>6.4 ± 0.4</td>
<td>16</td>
</tr>
<tr>
<td>( \Delta 4 ) cytoplasm + serum</td>
<td>309 ± 9</td>
<td>6.7 ± 0.4</td>
<td>9</td>
</tr>
<tr>
<td>( \Delta 4 ) nucleus starved</td>
<td>296 ± 13</td>
<td>5.2 ± 0.4</td>
<td>17</td>
</tr>
<tr>
<td>( \Delta 4 ) nucleus + serum</td>
<td>278 ± 10</td>
<td>5.7 ± 0.4</td>
<td>20</td>
</tr>
</tbody>
</table>

Figure 2.4: \( F/N \) and Diffusion coefficient results obtained using Fluorescence Correlation Spectroscopy. FCS was performed on free GFP, WT-ERK-GFP and \( \Delta 4 \)-ERK-GFP constructs respectively in both the nucleus and the cytoplasm of the live cells. In the table \( \text{starved} \) indicates measurements before serum addition, and \(+ \text{serum}\) data were acquired in the cytoplasm from 1 to 10 min after serum addition and in the nucleus from 10 to 15 min. Values are given as average \( \pm \) S.E.

Using FCS, we determined diffusion coefficients of the ERK constructs (WT and \( \Delta 4 \)) in both the cytoplasm and nucleus, before and after activation. No evidence for systematic changes in diffusion was seen after stimulation, and no significant differences between WT and mutant were detected (Table 2.4). This result indicates that the acceleration of the nucleocytoplasmic rate of shuttling upon stimulation was not a result of increased mobility in a given compartment. Mobile (free) GFP-ERK1 diffused at a rate one-third that of free GFP. From hydrodynamic considerations, GFP-ERK should have a diffusion constants 60-70\% that of free GFP, on the basis of their molecular masses and assuming ellipsoidal molecules with axial ratios not exceeding 4\[23\]. Thus The slower diffusion of GFP-ERK likely reflects interactions
with nuclear and cytoplasmic partner proteins.

**ERK1 Dimerization Is Not Detected in Living Cells**

![Image](image-url)  

Figure 2.5: Dimerization of GFP-ERK1-WT was not detected by FCS in either the nucleus and the cytoplasm. FCS was used to extract the fluorescence intensity per mobile object (F/N). No significant change in F/N was observed after stimulation for GFP-ERK1-WT, indicating that ERK1 remained monomeric after activation. Each symbol after activation indicates a single measurement. Connected symbols are measurements from the same cell over time. Values are normalized to the F/N value before stimulation and absolute values are found in previous table 2.4. The gray band represents the single cell standard deviation of pre stimulation measurements.
Figure 2.6: Dimerization of GFP-ERK1-Δ4 was not detected by FCS in either the nucleus or the cytoplasm. This was expected given the well-proved inability of dimerization of GFP-ERK1-Δ4. This experiment served as a negative control for F/N measurement on live cell. Connected symbols are measurements from the same cell over time. Values are normalized to the F/N value before stimulation and absolute values are found in Table 2.4. The gray band represents the single cell standard deviation of pre stimulation measurements.

FCS experiments were performed in live cell with both GFP-ERK1-WT and GFP-ERK1-Δ4. To search for evidence of dimerization, the number of independent mobile objects (N) in the focal volume was determined by FCS, and the brightness per moving object (F/N), was calculated. Dimerization of ERK would cause an increase in F/N, because the number of diffusing objects would be reduced upon dimerization, whereas the total fluorescence intensity would remain the same. In live cells, no changes in dimerization state (invariant F/N values) of GFP-ERK1 was
detected after stimulation with serum as shown in Fig. 2.5. GFP-ERK1-WT and GFP-ERK1-Δ4 shown in Fig. 2.5 and 2.6 showed the same behavior. In addition, the F/N values for ERK constructs did not differ significantly from the value for free monomeric GFP expressed in MEF$^{ERK1-/-}$ cells as shown in Table 2.4.

2.6 Conclusion

We have built a FCS system based on a confocal microscope setup. Due to issues with fluctuating laser intensity and room temperature, we have implemented a protocol to give reliable FCS measurements on the system. We have proposed F/N method to detect dimerization. Using SSDNA and DSDNA combination system, we successfully detected a two fold increase of object intensity (F/N) on DSDNA compared with SSDNA. We performed FCS measurement on ERK-1 in live cells to address the dimerization hypothesis that has been proposed as a requirement for ERK-1 nuclear translocation. Timed stimulation experiments are performed to monitor F/N decrease upon dimerization and no significant change in F/N was observed. Together with other measurement in [30], the result indicates ERK-1 nuclear translocation is dimerization independent.

References


Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS


Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS


Chapter 2. The Evaluation of ERK1 Dimerization Models using FCS


Chapter 3

Simultaneous Multiple-emitter Fitting for SM-SR Imaging

In this chapter, we describe an analysis algorithm that localizes multiple emitters simultaneously within an isolated sub-region using a multi-emitter model and the Maximum Likelihood Estimator. We have developed several techniques that are essential for the convergence of the fitting, and proposed a phenomenological modified Fisher information matrix that is designed to solve the singularity problem in Gaussian mixture models. We demonstrated the advantages of multi-emitter fitting analysis in both simulated and experimental data sets and also demonstrated that by implementing most of the analysis procedure in GPU using Compute Unified Device Architecture (CUDA, Nvidia) technology, the overall time consumption of the multi-emitter fitting analysis is on the scales of minutes.
3.1 Introduction to Single-Molecule Super Resolution

Single molecule based super resolution (SM-SR) techniques have revolutionized fluorescence microscopy, achieving spatial resolution of approximately 20 nm, an order of magnitude improvement from conventional fluorescence microscopy that is limited by diffraction to $\lambda/2NA$ or approximately 250 nm [1, 2, 3, 4, 5]. The SM-SR concept relies on making precise and accurate estimations of the positions of individual emitters that label the structure of interest. Resolution is then a function of both the position uncertainty and the sampling density. This concept is realized by exploiting some properties of the fluorescent probes that result in a small subset of emitters being in a fluorescent state during the acquisition of any single image. Acquired images that contain different subsets of active emitters can then be analyzed and used to generate a SR image, providing sufficient sampling density and localization precision. Initial demonstrations of SM-SR used a variety of probes including quantum dots [6, 7], photo-activatable proteins [8, 9] and organic dyes [10] and the number of probes that have been demonstrated for use in SM-SR continues to grow [4, 11].

In the case of 2D imaging, which is the focus of this work, an advantage of SM-SR over other SR techniques such as STED[12], 4Pi[13], and SSIM[14] is that it can be implemented using a relatively simple and conventional microscope such as an objective based Total Internal Reflectance Fluorescence (TIRF) microscope setup. However, the technique relies heavily on the analysis of the acquired data, primarily in making estimates of the position of on the order of $10^6$ emitters. To simplify and speed analysis, conventional analysis approaches only attempt to localize well separated, single emitter events and data that does not fit this model is rejected. Experimental conditions must then be optimized to give a single-frame active emitter density that makes best use of the data and yet minimizes acquisition time [15].
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

Figure 3.1: Proximity of emitters as a function of emitter density. The probabilities of finding $N=1$-$5$ emitters within a $8\sigma_{\text{PSF}} \times 8\sigma_{\text{PSF}}$ square sub-region ($\sigma_{\text{PSF}} = 127$ nm) at different densities were calculated for a uniformly distributed population of emitters and plotted as a function of density. As the emitter density increases beyond $1 \mu m^{-2}$, the fraction of sub-regions containing single emitters reduces dramatically (red line), emphasizing the need for fitting algorithms that can accommodate multiple emitters within a single sub-region.

SM-SR fitting routines that disregard events that cannot be fit to a single emitter profile result in some fraction of data being discarded. The potential loss of information is demonstrated in Fig 3.1, which shows that at an active emitter density of $1 \mu m^{-2}$, more than $55\%$ of $8\sigma_{\text{PSF}} \times 8\sigma_{\text{PSF}}$ ($\sigma_{\text{PSF}} = 127$ nm) sub-regions contain 2 or more active emitters. Such nearby or overlapping emission patterns could result in a failure of the single emitter model and the data not being used in the SR image reconstruction. The distribution of the number of emitters found within these $8\sigma_{\text{PSF}} \times 8\sigma_{\text{PSF}}$ sub-regions ($\sigma_{\text{PSF}} = 127$ nm) as a function of density is also
shown in Fig. 3.1 and illustrates that with increasing active emitter density, isolated single-emitter events become rare and therefore a majority of the position estimates will get discarded due to an unacceptable fit to a single emitter model. It is clear that a multiple-emitter fitting approach would enable the analysis of images containing higher single-frame density of active emitters. Analysis of multiple emitters simultaneously in one sub-region does not necessarily impact the position uncertainties as visually overlapping emitters (around 100 nm between emitter centers) can be localized with similar uncertainties [16, 17]. In practice, a multi-emitter fitting model would allow one or more of several important quantities to be improved, which would result in a much better localization in cases where single frame active emitter densities are relatively high [15].

Here, we describe an analysis method that uses the Maximum Likelihood Estimator (MLE) in order to perform simultaneous position estimates of multiple emitters within a small sub-region. In contrast to other techniques that use deflation methods, whereby the best single fluorophore fit is made to the image and the analysis proceeds with the residuum image that is calculated by subtracting the single fluorophore fit model [18, 19, 20, 21], all emitter positions within the sub-region are estimated simultaneously. The sub-region data is fit to models assuming \( N \) emitters, where \( N \) is varied from \( N=1 \), to \( N = N_{\text{max}} \) using a process that we will subsequently refer to as Multi-emitter Fitting Analysis (MFA). Based on the log-likelihood, a chi square distributed test statistic is used to either choose one model, or reject all fitting models. In this manuscript, we describe a procedure that allows robust application of the MFA, including model selection criteria, uncertainty calculations, and other procedures for analyzing a SM-SR data set and image reconstruction.
3.2 Theoretical Basis for the Multi-emitter Fitting Algorithm

3.2.1 Multiple Emitter Extension to the Pixelized Single Emitter Model

The impulse response of a microscope to a point source of light is defined as the point spread function (PSF) and in the 2D case, can be well approximated by the Gaussian function [22, 23]:

$$\text{PSF}(x,y) = \frac{1}{2\pi\sigma_0^2} e^{-\frac{(x^2+y^2)}{2\sigma_0^2}}$$ (3.1)

where $\sigma_0$ represents the standard deviation of the Gaussian.

Given the pixelization that occurs from a CCD based detector system, this continuous distribution can be modified to represent the expected photon count in pixels on the camera. For an individual pixel $k$ located at a position $\{x, y\}$ and assumed to rectangular, the expected number of photons in that pixel, which are emitted from a point object in focus, can be calculated by integrating Eq. 3.1 across the pixel assuming a square shaped pixel. This pixelized single emitter profile is given as:

$$\mu_k(x, y) = I_0 \Delta E_x(x, y) \Delta E_y(x, y) + b_0$$ (3.2)

where $\mu_k(x, y)$ is the expected photon count for a given pixel $k$, $I_0$ is the total emitted photon counts expected, $b_0$ is the background and $\Delta E_x(x, y)$ and $\Delta E_y(x, y)$ are:

$$\Delta E_x(x, y) = \frac{1}{2} \left( \text{erf}\left( \frac{x - x_0 + \frac{1}{2}}{\sqrt{2}\sigma_0} \right) - \frac{1}{\sqrt{2}\sigma_0} \right)$$ (3.3a)

$$\Delta E_y(x, y) = \frac{1}{2} \left( \text{erf}\left( \frac{y - y_0 + \frac{1}{2}}{\sqrt{2}\sigma_0} \right) - \frac{1}{\sqrt{2}\sigma_0} \right)$$ (3.3b)

Where $x_0$ and $y_0$ are emitter positions.
This model can be extended to account for emission from multiple emitters by assuming each emitter contributes independently to the expected photon counts at a given pixel $k$. The expected photon count for pixel $k$, $\mu_k(x, y)$ generated by $N$ emitters can then be calculated by summing over the total number of emitters $N$ and is defined as:

$$
\mu_k(x, y) = \sum_{i} I_0 \Delta E_{x_i}(x, y) \Delta E_{y_i}(x, y) + b_0 \quad (3.4)
$$

### 3.2.2 Maximum Likelihood Estimator

To estimate the emitter positions, we maximize the likelihood function [24]:

$$
L(\theta|D) = \prod_k L(u_k|d_k) = \prod_k \frac{\mu_k(x, y)^{d_k} e^{-\mu_k(x, y)}}{d_k!} \quad (3.5)
$$

where the likelihood of the parameters $\theta$ given the data $D$ is modeled as a photon counting process for each pixel, with the expected counts given by the multi-emitter model $\mu_k$ defined in Eq. 3.4 and the observed counts $d_k$. The maximum likelihood estimator (MLE) is used to estimate the emitter positions $\{x_1, y_1\}, \ldots, \{x_N, y_N\}$ and the background fluorescence rate $b_0$, giving $\hat{\theta} = \{b_0, x_1, y_1, \ldots, x_N, y_N\}^T$. To ensure robust estimation, we find that it is necessary to confine the intensity parameter $I_i = I_0$ in Eq. 3.4, where $I_0$ is obtained from independent measurements.

Maximization of Eq. F.2 can be performed using the Newton-Raphson method (NR) to iteratively maximize the log-likelihood. The iterative step for parameter $\theta_i$
can be written for a Poisson noise model as follows [25]:

\[
\theta_i \rightarrow \theta_i - \left[ \sum_k \frac{\partial \mu_k(\theta_i)}{\partial \theta_i} \left( \frac{d_k}{\mu_k(\theta_i)} - 1 \right) \right]^{-1} \left[ \sum_k \frac{\partial^2 \mu_k(\theta_i)}{\partial \theta_i^2} \left( \frac{d_k}{\mu_k(\theta_i)} - 1 \right) - \frac{\partial \mu_k(\theta_i)}{\partial \theta_i} \frac{d_k}{\mu_k(\theta_i)^2} \right]^{-1} (3.6)
\]

All derivatives of \( \mu(\theta) \) are identical in form to those from the single-emitter model and are given in Appendix E as well as in [25].

3.3 The Analysis Procedure

Our fitting routine operates independently on each image of a time series. First, a series of image filters are applied to each frame to find points of interests and then each frame is partitioned into an array of sub-regions around these points. In each sub-region, the positions of \( N \) proposed emitters in a model of \( N = 1 \) to \( N_{\text{max}} \) are found sequentially where the \( N \) emitter model uses position information from the \( N - 1 \) emitter model. We generate the p-value from a test statistic based on the log-likelihood ratio (LLR) to compare fits for each model. The model with the highest p-value is selected and the associated uncertainties and fits are determined based on a modified Fisher information matrix. The process is repeated for all frames and a reconstructed image is generated from the estimates by placing bivariate Gaussian shapes at the estimated locations using estimator uncertainties to build the bi-variate covariance matrix. Below we outline these steps in further detail.
3.3.1 Image Pre-processing and Segmentation

For each data set, all frames are analyzed independently. Experimentally acquired images are first offset and gain corrected to convert pixel intensity values to photon counts. To aid sub-region selection, a two step image filtering process is carried out to reduce Poisson noise and background and to identify potential emitter locations. The first filtering step is calculated from the original image $I$, as follows:

$$A_1 = \text{uniform}[I, (2\sigma_{PSF} + 1)] - \text{uniform}[I, (2 \times (2\sigma_{PSF} + 1))]$$  \hspace{1cm} (3.7)

where $\text{uniform}[image, q]$ represents a uniform filter process with a square kernel size $q$ operating on the 2-D matrix $image$. The uniform filter acts as a smoothing filter by reassigning the value of each pixel to the average pixel value within the square kernel centered at the pixel position. The analysis is not strongly dependent on the smoothing filter so the uniform filter is chosen for speed. Subtraction means a pixel-wise subtraction between results obtained for each filter process. The second filtering step is performed on the first filtered image $A_1$ as follows:

$$A_2 = \text{max}[A_1, (5\sigma_{PSF})]$$  \hspace{1cm} (3.8)

where $\text{max}[image, q]$ represents a maximum filter process used to obtain local maximum values within a square kernel size $q$. Through this process, all pixels within a kernel take the maximum value within the kernel. These two filtered images $A_1$ and $A_2$ are then compared pixel-wise to identify regions of interest:

$$A_3 = \begin{cases} 
\emptyset & \text{if } A_1 \neq A_2 \\
1 & \text{if } A_1 = A_2
\end{cases}$$  \hspace{1cm} (3.9)

Through this process, pixels with local maximum intensities in the uniformly
filtered image $A_1$ are identified in $A_3$. Sub-regions of size $6\sigma_{\text{PSF}} \times 6\sigma_{\text{PSF}}$ that are centered at pixels where $A_3 = 1$ are selected for further analysis.

### 3.3.2 Multi-emitter Fitting Analysis (MFA)

Each sub-region is analyzed using a Multi-emitter Fitting Analysis as depicted in Fig. 3.2. The analysis proceeds sequentially from a $N = 1$ model to a $N = N_{\text{max}}$ model. For the $N = 1$ model, the center of mass of the sub-region is used as the initial position estimate. For the $N \neq 1$, multi-emitter models, the $N - 1$ position estimates found in the previous step are used as $N - 1$ of the initial position estimates. The remaining initial position estimate is found by calculating the residuum image generated by a subtraction of the $N - 1$ model (Eq. 3.4) from the data in the sub-region. If the value of the maximum intensity pixel in the residuum image is low enough to assume that all emitters in the sub-region have been found, the analysis does not proceed further. Otherwise, from the residuum image, the last initial estimate is calculated from the position of the pixel with the maximum count value, giving $\{x_{\text{def}}, y_{\text{def}}\}$ and then is adjusted in a ”Push&Pull” process to $\{x_{\text{adj}}, y_{\text{adj}}\} = \{x_{\text{def}} \pm \sigma_{\text{PSF}}/2, y_{\text{def}} \pm \sigma_{\text{PSF}}/2\}$. If $\{x_{\text{def}}, y_{\text{def}}\}$ is within $\sigma_{\text{PSF}}$ of the edge of the sub-region, that position is likely to correspond an emitter outside of the region, and the sign of the adjustment is such to move the adjusted position further away from the center of the sub-region. Otherwise, the sign of the adjustment is such to move the adjusted position towards the center of mass of the $N - 1$ position estimates. This compensates for the effect that in a $N - 1$ model of an underlying $N$ emitter system, the estimated positions of N-1 emitters are displaced such that after deflation, the position of the maximum value pixel is biased away from the actual position of that emitter. This effect is illustrated in Fig. 3.2b. We found that the ”Push&Pull” adjustment of only one of the initial position estimates is sufficient to allow robust convergence. The initial estimates are then updated using a fixed
number of iterations of Eq. 3.6. After obtaining estimates for each model, models with location estimates outside the fitting boundary, which is a $8\sigma_{\text{PSF}} \times 8\sigma_{\text{PSF}}$ square region concentric with image sub-region (red box Fig. 3.2(b)-(e)), are discarded. Models with positions estimates within the fitting boundary but outside the data sub-region (black region between red and yellow box in Fig. 3.2(b)-(e)) are allowed since emitters located in this region will affect the data sub-region. The position and background estimates, along with their log-likelihood, are saved for each remaining model for a further model selection process.
Figure 3.2: Illustration of execution steps in the multi-emitter estimation task. (a) Fitting algorithm flowchart. (b) through (e): Demonstration of the results from each estimation task from the 1 emitter model through the 4 emitter model. The 5 emitter model fitting is not performed by the algorithm, because of the low photon counts in the deflated image.
3.3.3 Model Selection

To compare between models, we used a test statistic based on the log-likelihood ratio (LLR) as an indicator for the quality of fit. The LLR is shown in Eq. 4.22 and approximates a chi square distribution with $K - (2N + 1)$ degrees of freedom, where $K$ is the number of pixels in the sub-region and $N$ is the number of emitters in the model.

$$LLR = -2 \ln \left[ \frac{L(\hat{\theta}|D)}{L(D|D)} \right]$$  

$$= -2 \sum_{i=1}^{K} \left[ -\mu_i + d_i + d_i \ln(\mu_i) - d_i \ln(d_i) \right] \quad (3.10)$$

where $K$ is the total number of pixels within the sub-region, $\mu_i$ and $d_i$ are the expected count by proposed model and experimental obtained count value for $i^{th}$ pixel respectively, $D$ represents the sub-region data, $\hat{\theta}$ are the MLE estimates and $L(D|D)$ gives the upper limit of likelihood of the data set with Poisson noise (when $\mu_k = d_k$). The model is accepted if it has the maximum chi-square p-value of all models and passes the p-value threshold set by user. Considering that the variance of intensities in real or realistically simulated data would broaden the LLR distribution and thus result in a smaller p-value, typically a small p-value of $10^{-3}$ to $10^{-6}$ is used as the threshold in our analysis and is still sufficient to reject incorrect models and the un-converged fit. After obtaining the uncertainty for the position estimates, emitters with estimated positions near (within $\sigma_{PSF}/2$) or outside of the sub-region boundary are discarded. The parameters describing the remaining emitters are passed to the image reconstruction process.
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

3.3.4 Precision of the Estimated Parameters

For unbiased estimators, the Cramer-Rao Lower Bound (CRLB), given as $\text{var}(\hat{\theta}) \geq I_{\theta}^{-1}$ where $I(\theta)_{i,j} = E\left[\frac{\partial \ln L(\mu(\theta)|D)}{\partial \theta_i} \frac{\partial \ln L(\mu(\theta)|D)}{\partial \theta_j}\right]$ is the Fisher information matrix, is often used to calculate the precision of estimated parameters [26, 16, 25]. However, as known from the analysis of Gaussian mixture models [27], the Fisher information matrix is singular at $\{x_i, y_i\} = \{x_j, y_j\}$, and near this singular point, cannot be used to correctly calculate estimator precision. We implemented a phenomenological correction to the Fisher information matrix by modifying the off diagonal terms that give rise to the singularity. Given our parameter set $\theta = \{b_0, x_1, y_1, ... x_N, y_N\}^T$, the corrections are given by:

$$F(\theta)_{i,j} = \begin{cases} \frac{A}{A+1} I(\theta)_{i,j} & \text{(i,odd) & (j,odd) & (i \neq 1, j \neq 1) & (i \neq j)} \\ \frac{A}{A+1} I(\theta)_{i,j} & \text{(i,even) & (j,even) & (i \neq j)} \\ I(\theta)_{i,j} & \text{other} \end{cases}$$

(3.11)

$A$ is given by $A = \frac{|(\theta_i - \theta_j)^2|}{\sigma_i \times \sigma_j}$, where $\sigma_i$ and $\sigma_j$ are the intermediate precision calculations obtained from $F(\theta)$ assuming $A = 0$. $F(\theta)$, which we designate the modified Fisher Information matrix, replaces the original Fisher Information matrix in our precision calculation process, is non-singular at $\{x_i, y_i\} = \{x_j, y_j\}$ and quickly converges to $I(\theta)$ once far from the point of singularity. Thus it provides reasonable precision estimates in the regions both near and far from the point of singularity.

3.3.5 Filtering and SR Image Reconstruction

After obtaining estimates and their uncertainties, a rejection process is performed to remove repeated localizations that can occur due to overlapped sub-regions. An emitter estimate is removed if there is another estimate with a smaller uncertainty.
within a distance of the previous emitter’s localization uncertainty coming from
the the same image frame but a different sub-region. Another filtering process is
performed to remove the estimates with position uncertainties greater than the res-
olution threshold and would therefore not contribute to the desired resolution in
the reconstructed image. The SR image is reconstructed by adding bivariate 2-D
Gaussian shapes to the SR image at the location of the position estimates. The
covariance of the bivariate Gaussian is constructed using the appropriate elements
of $F(\theta)^{-1}$ and indicate the asymmetry of the position uncertainties that arise from
the multi-emitter localization process.

3.4 Computational and Experimental Methods

3.4.1 Hardware and Software Implementation of Analysis
Routines

Numerical analysis are performed using MATLAB (The Mathworks, USA), the imag-
ing processing toolbox, DipImage [28] and c-language codes that are compiled to
MATLAB mex files and initiated from within the MATLAB environment. GPU
code (Nvidia CUDA [29]) are managed through c-language codes that are also com-
piled to MATLAB mex files and runs within the MATLAB environment. All CPU
based code runs on a single thread.

Image pre-processing and segmentation are implemented in c-code. The array of
isolated sub-regions are passed into the GPU global device memory for the MFA.
The MFA for each sub-region is independently carried out by a single thread on the
GPU similar to that is described in previous publication [25] using 50 iteration steps
in NR iteration process. The model selection is performed in the same thread as
part of the MFA. The fitted parameters for successful models are passed back to
the CPU. The generation of $F(\theta)$ and its inversion, by LU decomposition with back substitute method [30], are implemented on the GPU executing with one thread per sub-region. The resulting uncertainties for each parameter are passed back to the CPU. The filtering of position estimates by sub-region position and their uncertainties is performed on the CPU. Reconstruction of the SR image is performed in a manner inverse to the sub-region selection. First, in the GPU, an up-sampled sub-region is generated that corresponds to each position estimate and its uncertainties. The bivariate Gaussian shapes for the position estimates are added to the sub-region. All generated up-sampled sub-regions are passed back to the CPU and assembled into a single up-sampled SR image.

3.4.2 Estimator Precision and Algorithm Performance Testing

In order to demonstrate the performance of the modified Fisher information matrix in calculating the estimator precision, two types of data sets were generated and analyzed. First, a series of simulated images of two emitters that had increasing separations between their centers were generated. For each separation, 1000 identical two-emitter images were generated, corrupted by Poisson noise and fitted by MFA. Second, images of 1000 configurations of random placements of 1, 2, 3, 4 and 5 emitters were replicated 1000 times, corrupted by Poisson noise and fitted by MFA. The performance of the modified Fisher information matrix in providing the correct precision estimates was demonstrated by comparing the observed standard deviation of estimates and the precision of the estimator calculated using the modified Fisher information matrix. Estimator accuracies for each emitter distribution were calculated by taking the ratio between the mean of the uncertainty estimates and the observed uncertainty.
Algorithm performance was also tested on simulation data where 2D Gaussian shapes were randomly placed with uniform distribution through the image with their actual position registered for later calculation. The total expected photon count per emitter was selected from a normal distribution with $\mu = 800$, $\sigma = 100$. A background count rate of 5 count/pixel was added to the image, and then the image was corrupted with Poisson noise. After fitting these images using MFA with a target resolution of 20 nm or 50 nm, the localization fraction was calculated by taking the ratio between the number of correctly localized emitters which is defined as having a registered emitter position near the localized emitter within the target resolution and the total number of emitters in simulation. The error rate of the algorithm was obtained by calculating the ratio between the number of mis-localized emitters which is defined as having no actual emitter position near the localized emitter within the target resolution and the total number of emitters obtained from fitting.

### 3.4.3 Synthetic Data Generation

Synthetic image series in a Siemens star pattern with 50 non-empty slice regions were generated such that the maximum width of each slice (on the outer diameter) equals 213 nm. A fixed labeling density $\rho_0 = 5000 \, \mu m^{-2}$ and off rate $k_{\text{off}} = 0.8 \, \text{frame}^{-1}$ were used, with varied $k_{\text{on}}$ to generate variations in active densities ($\rho_{\text{active}}$) according to:

$$\rho_{\text{active}} = \rho_0 \times \frac{k_{\text{on}}}{k_{\text{off}} + k_{\text{on}}} \tag{3.12}$$

A blinking trace was generated for each emitter using the transition rates $k_{\text{on}}$ and $k_{\text{off}}$ for dark to active, and active to dark transitions respectively and were designed to emulate realistic photophysical properties. As in all of our simulations, the active emitters were represented as a 2D Gaussian shapes, with $\sigma = \sigma_{\text{PSF}} = 1.2$ pixels (127 nm). To represent the experimentally observed variation in emitter brightness, for each emitter, the total expected photon count per frame was selected from a
normal distribution with $\mu = 800, \sigma = 100$. Shown in Fig. 3.3 is the single frame intensity distribution of Alexa Fluor 647. A background count rate of 5 count/pixel was added to the image, and then the image was corrupted with Poisson noise. Calculation of the density of active emitters assumes a pixel size of 106 nm, which is the back-projected pixel size in the experimental system.

![Image](a)

Figure 3.3: Single fluorophore intensity distribution of the organic fluorophore Alexa Fluor 647 obtained from the data set described in section 4.4.1 taken in TIRF condition. The distribution is modeled as a normal distribution with $\mu = 800, \sigma = 100$.

### 3.4.4 SM-SR Imaging

**Cell Culture – Prepared with Samantha Schwartz, UNM, 2010**

Human epithelial cervical cancer (HeLa) cells were cultured in Minimum Essential Media (Gibco) supplemented with Fetal Bovine Serum (HyClone), L-Glutamine and Penn-Strep. Five hours after plating at low confluence onto 8-well Labtek chambers (Nunc), cells were serum starved approximately 10 hours and fixed using 4% paraformaldehyde in phosphate buffered saline (PBS). After three washes in PBS, cells were permeabilized (0.5%v/v Triton X-100) at room temperature for 15 minutes.
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

in the presence of 3% BSA to reduce non-specific binding. Cells were again washed three times in PBS before addition of Alexa Fluor 647 Phalloidin (Invitrogen). Phalloidin was added at four times the recommended concentration (approximately 660 nM) to ensure dense labeling. Cells were washed five times and imaged in the presence of an oxygen scavenging system including 50 mM MEA [31] as a reducing agent.

Microscopy and Data Acquisition

Single molecule imaging experiments were performed in an epi-fluorescence microscope setup consisting of an inverted microscope (IX71, Olympus America Inc.), 1.45 NA TIRF objective (U-APO 150x NA 1.45, Olympus America Inc.), 635 nm diode laser (Radius 635, Coherent Inc.), and an electron multiplying CCD camera Ixon (897, Andor Technologies PLC.) with EM gain set to $\approx 200$. The epi-fluorescence filter setup consisted of a dichroic mirror (650 nm, Semrock) and an emission filter (692/40, Semrock). The sample chamber was mounted in a 3D piezostage (Nano-LPS, Mad City Labs). 5000 images were taken in a TIRF configuration at 20 frames/second. Drift correction was not implemented, but from independent measurements we estimate a drift of less than 25 nm over the acquisition time. Frames were $256 \times 256$ pixels with a pixel size of 0.106 $\mu$m.

3.5 Results and Discussion

3.5.1 Optimal Sub-region Size and $N_{\text{max}}$

Various sub-region sizes ranging from $4\sigma_{PSF}$ to $8\sigma_{PSF}$ were evaluated in the aspects of both localization fraction and error rate that are defined in section 3.4.2. Small sub-regions tend to isolate individual emitters from one another better compared to
larger sub-regions and thus results in sub-regions containing fewer emitters. However, the smaller area decreased the amount of information that could be used in fitting and thus the error rate increases compared to larger sub-regions. Large sub-regions provide more accurate estimates compared to a smaller sub-region but the probability of introducing more emitters within or near the sub-region increases quadratically with the width of the square sub-region. We have tested our algorithm performance under different sub-region sizes, such as $4\sigma_{\text{PSF}}$, $5\sigma_{\text{PSF}}$, $6\sigma_{\text{PSF}}$, $7\sigma_{\text{PSF}}$, $8\sigma_{\text{PSF}}$, various active emitter densities from $0.1 \, \mu\text{m}^{-2}$ to $10 \, \mu\text{m}^{-2}$, various emitter intensities from 200 to 5000 and various intensity variance. After comparing these plots (data not shown), we found that sub-region size of $6\sigma_{\text{PSF}}$ shows the best compromise of error rates and localization fraction. $N_{\text{max}}$ values ranging from 1 to 8 were tested. Large $N_{\text{max}}$ tend to generate a more complex likelihood surface and thus the possibility for the estimator being stuck at a local minimum increases with $N_{\text{max}}$. The complexity introduced by multi-emitter model results in higher error rates and thus $N_{\text{max}}$ was restricted to 5 in our analysis.
3.5.2 Uncertainty Estimator Performance

Figure 3.4: Performance of the precision estimate. (a) A comparison between the precision predicted from the CRLB and from the modified Fisher information matrix. A series of simulated images of two emitters at various separations between their centers were generated. MFA was performed on these images and the precision estimates calculated by the modified Fisher information matrices $F(\theta)$ Estimated Std. Dev. were compared with that obtained from the CRLB (Estimated Uncertainty CRLB), precisions obtained from the CRLB generated by emitter’s true position (Theoretical Uncertainty CRLB), and the observed standard deviation of the estimates (Observed Std. Dev.). (b) The CDF (integral of histogram) of the uncertainty estimator accuracy obtained using the modified Fisher information matrices for random placements of multiple emitters.

Using simulations, estimator precision calculations for various emitter configurations were calculated from our modified Fisher information matrix $F(\theta)$ of Eq. 3.11 and compared with observed standard deviations. Singularity of the Fisher Information matrix for the multi-component Gaussian model when 2 (or more) emitter centers that have a separation less than 100 nm results in a failure of the CRLB to correctly estimate the precision of the position estimation. This effect is demonstrated in Fig. 3.4a. Fig. 3.4a also shows that calculations based on $F(\theta)$ gave a correct estimator precision (compared to the observed standard deviation of the
estimates) in the regions both near and far from the point of singularity of the two emitter model, with only a small but conservative deviation below 50 nm. We also show the performance of $F(\theta)$ based precision calculations for random configurations of multiple emitters by looking at the estimator accuracy, defined as the ratio of the precision calculated using $F(\theta)$ to the observed standard deviation of the estimates. The cumulative distribution (the normalized integral of the histogram) of the estimator accuracy is shown in Fig. 3.4b and demonstrates that the estimator accuracy distribution (corresponding to the derivative the CDF) of is narrowly peaked around 1 for the 1-3 emitter models (ideal) and is conservative (reported precision is larger than observed standard deviation) on the 4 and 5 emitter models where the estimator accuracy distribution is peaked around 1.1 and 1.2 respectively.

3.5.3 Algorithm Performance Versus Density and Intensity Distribution

We have tested our algorithm on simulated data sets where emitters were randomly placed with uniform distribution in a $64 \times 64$ image representing an area of $46 \ \mu m^2$ in our microscope camera setup. By increasing the number of active emitters within the image, density increased from 0.01 $\mu m^{-2}$ to 10 $\mu m^{-2}$. Both single ($N_{\text{max}} = 1$) to multi ($N_{\text{max}} = 5$) emitter fitting algorithm were performed on these data sets and localization fraction (defined in 4.2) were calculated.
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

Figure 3.5: Performance versus active emitter density and intensity distribution in low intensity case. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of $300 \pm 30$ (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes $N_{\text{max}} = 5$. The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position.
Figure 3.6: Performance versus active emitter density and intensity distribution in cases that the intensity distribution is similar as what we have observed from Alexa Fluor 647. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of 800 ± 100 (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes $N_{\text{max}} = 5$. The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position.
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

Figure 3.7: Performance versus active emitter density and intensity distribution for extremely high emitter intensity case. Shown are the results of MFA analysis of images with spatially random distributed emitters with normally distributed intensities of 5000 ± 30 (a),(b). Localization error is calculated as the distance from the estimated position to the found position and in all cases assumes $N_{\text{max}} = 5$. The median localization error is where the cumulative distribution reaches 0.5. Localization fraction is the fraction of emitters that are correctly localized as determined by being found within either 20 nm or 50 nm from the known position.

Fig. 3.5,3.6 and 3.7 shows the performance of the MFA analysis for various densities and intensity distributions. The simulations show that the localization fraction improvement from $N_{\text{max}} = 1$ to $N_{\text{max}} = 5$ is most significant at a densities higher than 1 $\mu \text{m}^{-2}$. We note that at high intensities with narrow intensity distribution (Fig. 3.7a,3.7b) the localization error improves, but the localization fraction does not. This is attributed to high sensitivity to model mismatches created by the fixed intensity assumption and emitters outside the fitting window.

3.5.4 Pattern Simulation Results

Simulated Siemens star pattern images were generated such that the labeled region active emitter density is 1.0 $\mu \text{m}^{-2}$ and 6 $\mu \text{m}^{-2}$. These two sets of data were analyzed using $N_{\text{max}} = 1$ and $N_{\text{max}} = 5$. Results of the analysis are shown in Fig. 3.9a through
Figure 3.8: Simulated data set and its sum projection. (a) The emitter position histogram used in generating synthetic data. (b) Sum projection of the generated image which demonstrated the diffraction limited pattern observed under conventional microscope.
Figure 3.9: Results from single emitter fit and multi emitter fit at the density of 1 $\mu$m$^{-2}$. (a) Single emitter fitting result at a density of 1 $\mu$m$^{-2}$ with $N_{\text{max}} = 1$. (b) Multiple emitter fitting result at a density of 1 $\mu$m$^{-2}$ with $N_{\text{max}} = 5$. At 1 $\mu$m$^{-2}$ case, $N_{\text{max}} = 1$ resulted in 12848 emitters localized while $N_{\text{max}} = 5$ localized 30354 emitters. The contrast of images (a) to (b) were globally adjusted across all images for optimal display.
Figure 3.10: Results from single emitter fit and multi emitter fit at the density of 6 \( \mu \text{m}^{-2} \). (a) Single emitter fitting result at a density of 6 \( \mu \text{m}^{-2} \) with \( N_{\text{max}} = 1 \). (b) Multiple emitter fitting result at a density of 6 \( \mu \text{m}^{-2} \) with \( N_{\text{max}} = 5 \). In 6 \( \mu \text{m}^{-2} \) case, \( N_{\text{max}} = 1 \) resulted in 519 emitters localized while \( N_{\text{max}} = 5 \) localized 33580 emitters. The contrast of images (a) to (b) were globally adjusted across all images for optimal display.

At relatively low densities, results from \( N_{\text{max}} = 1 \) and \( N_{\text{max}} = 5 \) are similar. For \( N_{\text{max}} = 1 \) shown in Fig. 3.9a, 12848 emitters were localized and accepted for use in the SR reconstruction, and for \( N_{\text{max}} = 5 \) shown in Fig. 3.9b, 30354 emitters were accepted and used. In the high density case, shown in Fig. 3.10a and Fig. 3.10b, there was nearly two orders of magnitude (519 versus 33580) more position estimations accepted and used in the reconstruction. As shown in the projections of the SR images, the \( N_{\text{max}} = 1 \) fitting performs better near the edges of the structures where the local active emitter density is smaller. It is interesting to note that at the low density, \( N_{\text{max}} = 5 \) fits almost 3 times more emitters than \( N_{\text{max}} = 1 \) case, and thus the pattern result shows a better resolved structure near the center and provides better resolution compared to \( N_{\text{max}} = 1 \) fitting result.
Algorithm speed was tested under conditions including various active densities and $N_{\text{max}}$. Tests were performed on two set of data (data size: 128×128×1000) with densities $1 \, \mu m^{-2}$ and $5 \, \mu m^{-2}$. Algorithm execution was divided into several major sections and the run times for each section were recorded. As shown in Table 3.1, the operation time for MFA $N_{\text{max}} = 5$ was 176 s for the $1 \, \mu m^{-2}$ case and 408 s for the $5 \, \mu m^{-2}$ case. When performing single emitter operation ($N_{\text{max}} = 1$), this run time decreased dramatically to 17 s and 30 s respectively. This dramatic difference is caused by the complexity introduced by fitting multiple emitters, such as fitting to multiple models, the deflation process, NR iteration on more parameters, the Fisher information modification and also a larger Fisher information matrix. However, the fraction of localization also dramatically increased when comparing single fitting results to multi fitting results as over 100 times more emitters were localized at a density of $5 \, \mu m^{-2}$ and almost 3 times more at a density of $1 \, \mu m^{-2}$.

<table>
<thead>
<tr>
<th>$N_{\text{max}}$</th>
<th>Processing Time (s)</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fitting Routine</td>
<td>Preprocessing</td>
</tr>
<tr>
<td>$N_{\text{max}} = 1$</td>
<td>CPU</td>
<td>GPU</td>
</tr>
<tr>
<td>Low Density: $1 , \mu m^{-2}$, Total Simulated Emitters: 181222</td>
<td>5.69</td>
<td>8.41</td>
</tr>
<tr>
<td>High Density: $5 , \mu m^{-2}$, Total Simulated Emitters: 906127</td>
<td>5.72</td>
<td>162.25</td>
</tr>
</tbody>
</table>

Table 3.1: Time consumption and performance. Images (128×128×1000) of emitters that were randomly placed with uniform distribution are generated using two active emitter densities. Both were fitted by $N_{\text{max}} = 1$ and $N_{\text{max}} = 5$ and the speed for each part of the algorithm was recorded. Total numbers of fitting sub-regions are 146,597 and 336,959 in low and high density case respectively.
3.5.6 Imaging of Actin Structures

![Figure 3.11: Comparison of SM-SR fitting routines for imaging the actin mesh-work within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) images were obtained using conventional TIRF microscopy. Scales bars represent 5 µm in (a) and 1 µm in (b)](image)

Imaging the actin mesh-work within HeLa cells demonstrates the improvements in SM-SR fitting made possible by the MFA’s multi emitter analysis ($N_{\text{max}} = 5$) compared with single emitter analysis ($N_{\text{max}} = 1$). For samples with high labeling densities, such as those possible when using small molecule fluorescent probes such as Alexa Fluor 647 phalloidin, regions of interest that could be seen using conventional microscopy (Fig. 3.11b), may appear to be discontinuous when analyzed using $N_{\text{max}} = 1$ that can not process high active densities (Fig. 3.12b). By analyzing these data sets using MFA ($N_{\text{max}} = 5$), less events were discarded. The reconstructed SR image from $N_{\text{max}} = 5$ showed more continuous structures and ultimately, enabled finer detail of the underlying protein distributions to be revealed (Fig. 3.13a). It is shown in Fig. 3.13 that although the branching structures were resolved nicely using $N_{\text{max}} = 1$, structures toward the middle backbone can’t be resolved, because the backbone structure are composed of intercrossing actin fibers and thus possessed
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

a higher local emitter density than isolated line structures. As shown in Fig. 3.13a, MFA ($N_{\text{max}} = 5$) achieved to resolve the backbone structure better than single emitter fitting algorithm ($N_{\text{max}} = 1$).

Figure 3.12: Comparison of SM-SR fitting routines for imaging the actin mesh-work within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) SM-SR images were generated using $N_{\text{max}} = 1$ as a demonstration of single emitter fitting result. Actin rich regions, seen in top right of 3.11a are missing using single emitter routines ($N_{\text{max}} = 1$) in (b). Scales bars represent 5 $\mu$m in (a) and 1 $\mu$m in (b)
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

Figure 3.13: Comparison of SM-SR fitting routines for imaging the actin mesh-work within a HeLa cell labeled with Alexa 647 phalloidin. (a) and (b) SM-SR images were generated using $N_{\text{max}} = 5$ as a demonstration of multi emitter fitting result. Actin rich regions, seen in top right of 3.11b, 3.12b, 3.13b are missing using single emitter routines ($N_{\text{max}} = 1$) 3.12b, but successfully fit using the MFA ($N_{\text{max}} = 5$) 3.13b. The increase in molecular density found using the MFA ($N_{\text{max}} = 5$) routine also reveals a more complete depiction of the underlying actin structure, outlining possible actin corrals seen in the center of 3.13a. Scales bars represent 5 µm in (a) and 1 µm in (b).

3.6 GPU Based MATLAB-MEX Programming

As required by the Nyquist–Shannon sampling theorem [32], in order to achieve a resolution around 10s of nanometers, data sets should at least contain tens of thousands of frames. Each frame are then isolated into sub-regions and then analyzed using localization algorithm. Such a huge mount of computation load would result in hours or even days of processing time if performed on single threaded CPU.

CUDA is a parallel computing architecture developed by Nvidia that allows users to access the virtual instruction set and memory of the parallel computational elements in CUDA enabled GPUs. While CPU development has focused on
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

multi-tasking processing, GPUs emphasize fast parallel operation across hundreds of Arithmetic Logic Units (ALUs) that are located on each graphic card. SM-SR analysis in this manner is perfectly suitable for GPU processing since hundreds of thousands of sub-regions are to be analyzed, and further more the analysis process among sub-regions are completely independent at the localization stage, facilitating parallel computing.

Programming in CUDA can be achieved using CUDA supported C Language [29] supported by Nvidia. Usually 4 significant steps are followed in addition to normal C Language programming:

1. Memory allocation and transfer from host memory (RAM) to device memory (GPU-RAM).
2. Definition of block size (number of threads) and grid size (number of blocks) per multi-processor.
3. Design of kernel functions that target individual thread and shared memory usage.
4. Memory transfer from device memory to host memory.

I have setup a small collection of progressive sophisticated core samples code and a tutorial in the Lidkelab—wiki which could be used as a learning material for CUDA programming.

Since we have implemented most of our instrumentation control and analysis routines in MATLAB, we also implemented all of our GPU code using MATLAB-MEX environment. While MATLAB allows users to import C Language code and to compile into a ‘MATLAB executable’, as so called ‘MEX’ file, we have combined MATLAB-C and C-CUDA together to create MATLAB-C-CUDA MEX (CUDA-MEX) files. These CUDA-MEX files are programmed using both syntaxes from
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

MATLAB-C and C-CUDA and thus was able to be executed in MATLAB environment and perform massive parallel computing assignment using CUDA technology on GPU which introduced CUDA technology directly in the MATLAB environment. Programming in the MATLAB-C-CUDA environment follows almost the same steps described in previous list except that user now has to comply with syntax conventions in MEX programming.

Although compiling programmed C-CUDA code in Visual Studio 2008 (VS2008) should be straight forward by following the online tutorial on Lidkelab−wiki (which I have used as a general resource page to provide programming information for several graduate students from our collaborators), trying to compile CUDA-MEX files can be quite confusing. Although by following the tutorial step by step should get MEX-CUDA file compiled in VS2008, a easier way is to use the pre-made CUDA-MEX project template available also on our wiki page. Appendix G gives a complete list of software to be installed in order to make CUDA-MEX compiling successful for a newly installed windows system. This list is also included on the our wiki.

3.7 Conclusion

The MFA method we have developed allows the analysis of images with average active emitter densities up to 10 \( \mu \text{m}^{-2} \). This capability relaxes an important constraint in SM-SR, allowing an order of magnitude improvement in the speed of acquisition and/or the maximum supported duty cycle of the emitters. Although our approach is based on the maximum likelihood estimator, robust estimation of multiple emitter positions also requires strategies such as making good initial estimates, making accurate uncertainty estimates, and the model selection and rejection algorithm. Higher density imaging allows shorter acquisition times, but results in more computational complexity in analysis. By implementing key analysis steps in GPU hardware, we
Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging

show the MFA method can complete the analysis of real data sets on the time scale of minutes[33].

3.8 Acknowledgements

We want to thank Marcel Bruchez and Yan Qi from Carnegie Mellon University for providing the HeLa cell line for this work. Bernd Rieger and Robert Nieuwenhuizen are thanked for the helpful suggestions and discussions for the manuscript. We also thank W. Duncan Wadsworth for insightful discussions about test statistics. This work was supported by NIH grant #R21RR024438, NIH grant #1P50GM085273, NIH grant #2U54RR022241 and NSF grant #0954836.

References


Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging


Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging


Chapter 3. Simultaneous Multiple-emitter Fitting for SM-SR Imaging


Chapter 4

Rejection Algorithms for Single Emitter Fitting in SR Analysis

In this chapter, we first demonstrate the importance of rejection algorithms in single-molecule super resolution (SM-SR) analysis. We propose a method to calculate the optimal active emitter density to maximize SM-SR efficiency by maximizing the localization density. We then conclude that rejection algorithms are essential for SM-SR analysis. Second, we reviewed the most common two rejection algorithms that are used in literature: elliptical test and sum squared error (SSE) test. We demonstrated their behaviors of acceptance and rejection, and their unexpected and inconsistent behavior in various configurations. Next we proposed the log-likelihood ratio (LLR) that is used in a robust and consistent rejection algorithm and demonstrated its advantages over the other rejection algorithms. Third, we investigated the property of LLR at small expectation values and demonstrated its discrepancy from $\chi^2$ distribution in this region. We then proposed three methods which can be used to solve this discrepancy issue by introducing Lidke–Huang (LH) hypothesis. We also introduced a empirical function to reduce the computational complexity. Finally, we conclude that with our proposed correction method, performing significance test on
LLR provided a robust, consistent and statistically rigorous rejection algorithm for SM-SR analysis.

4.1 Introduction

4.1.1 Why a Rejection Algorithm is Needed?

As discussed in previous chapter, single molecule super resolution relies on the separation of single emitter events so that in each acquired frame from the EM-CCD camera, only a small portion of emitter population is in the fluorescence state during that exposure time. In order to resolve a structure with a resolution of tens of nanometers, a minimum labeling density required is required which is given by the Nyquist – Shannon sampling theorem [1, 2],

$$\Delta_{\text{Nyquist}} = \frac{2}{N^{1/D}} \tag{4.1}$$

Where $\Delta_{\text{Nyquist}}$ stands for the smallest resolvable size, $N$ stands for density of localized emitter or localization density and $D$ is the dimension of the structure.

For example, to achieve a 10 nm resolution ($\Delta_{\text{Nyquist}} = 10$ nm), the minimum localization density can be calculated as $(2/\Delta_{\text{Nyquist}})^{1/D}$ which gives a localization density of $4 \times 10^4 / \mu \text{m}^2$ in two dimensions. In order to accumulate such a localization density as fast as possible, number of localized particle per frame should be maximized by adjusting active density per frame, $\rho_{\text{act}}$. $\rho_{\text{act}}$ is can be calculated based on the labeling density $\rho_0$ and the activation rate including both transition rates $k_{\text{on}}$ and $k_{\text{off}}$ for dark to active, and active to dark transitions respectively which would result in a active emitter density per frame described below,
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

\[ \rho_{\text{act}} = \rho_0 \times \frac{k_{\text{on}}}{k_{\text{off}} + k_{\text{on}}} \] (4.2)

We can optimize the active emitter density \( \rho_{\text{act}} \) in a SM-SR imaging experiment in order to maximize the isolated emitter events per frame to achieve maximum localization density. However, a high \( \rho_{\text{act}} \) result in a higher probability of overlapping for emitters and a decrease on number of localizable emitters per frame since only isolated emitters can be localized accurately.

For most commonly used single molecule analysis routines operated on around 400-700 nm emission wavelengths [3] with a \( \sigma_{\text{PSF}} \) around 140 nm, reliably localizing 1 particle in a 1.3 \( \mu \text{m} \times 1.3 \mu \text{m} \) isolate sub-region which does not contain any other emitter center within this sub-region could well be achieved and attempting to perform localization analysis on sub-regions that contains more than a single emitter center would often result in biased position estimates or localization artifacts [4]. So theoretically, if every particle could be activated in a manner that each of them could be isolated in each frame using a 1.3 \( \mu \text{m} \times 1.3 \mu \text{m} \) sub-region, single molecule analysis routines could provide reliable localization at a density of 0.6 \( /\mu\text{m}^2 \) which is \( \frac{1}{1.3\mu\text{m} \times 1.3\mu\text{m}} \). Then to achieve a resolution of 10 nm, \( 4 \times 10^4 \) particles have to be localized per square micron. Thus, in this perfect situation, one needs 67,000 frames to achieve such a resolution.

However, in a super resolution experiment, emitters can not be activated in such a uniform way to provide exactly one emitter per region but are activated randomly across the biological structure. Assuming, a random uniform distributed active population of emitters in a 2-D surface, the average number of emitters that could be isolated with a sub-region size \( a \) without any other emitter within the sub-region, \( N_{\text{loc}} \) could be calculated using binomial distribution,
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

\[ N_{\text{loc}} = \binom{n}{1} (1 - p)^{n-1} \]  

(4.3)

where \( n \) is the total number of active emitters, \( p = a/A \) which stands for the probability of placing an emitter within the sub-region. We then make two assumptions for further evaluation of the problem.

B1 Single emitter sub-regions that contain only 1 emitter center near the center of the regions and doesn’t have any other emitters center in any of their pixels are accepted by rejection algorithm for localization analysis and result in a localized emitter. Other sub-regions will be excluded from the fitting.

B2 Boundary effects can be ignored by assuming the total area \( A \) of the image is significantly larger than the sub-region size \( a \).

Given the above assumptions, we can then try to calculate an expected localization density per frame \( \rho_{\text{loc}} \) for a given active emitter density per frame \( \rho_{\text{active}} \), sub-region size \( a \) and total area of the image, \( A \).

For example, given the active emitter density \( \rho_{\text{act}} = 1/\mu m^2 \), the total area of image \( A = 50 \mu m \times 50\mu m = 2500\mu m^2 \) and sub-region size \( a = 1.69 \mu m^2 \), \( N_{\text{loc}} \) can be then calculated using Eq. 4.3 with \( n = \rho_{\text{act}} \times A \) and \( p = a/A \). Given previous assumption, B1 and B2 and conditions specified, we then obtain the number of localized emitter per frame \( N_{\text{loc}} \) is 461. Noticing that the number of total active emitters per frame is \( n = \rho_{\text{act}} \times A = 2500 \), we can conclude that at a density of \( 1/\mu m^2 \), only 18.4% of the population \( \left( \frac{461}{2500} \right) \) could be isolated within a 1.3 \( \mu m \times 1.3 \mu m \) sub-region and thus leads to a successful localization. For each individual frames, the localization density can be then calculated as \( 18.4\% \times \rho_{\text{act}} \) or \( \rho_{\text{loc}} = \frac{\rho_{\text{loc}}}{A} \) which gives us 0.18/\( \mu m^2 \). The corresponding rejection percentage can be then calculated as 1-18% which gives 82%.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

This calculation demonstrates that out of all sub-regions selected for a potential single emitter fitting, 82% of them should be rejected by a rejection algorithm. Thus it is essential for any rejection algorithms in SM-SR analysis to be consistent and robust so that it will perform such a rejection process reliably which then leads to reliable localization estimates.

4.1.2 The Rejection Algorithm at Optimal Active Density in SM-SR Experiments

We now expand our idea described in the above section to calculate the localization density for any given active emitter density $\rho_{\text{act}}$ and sub-region size $a$. Noticing that the localization density is a product of active emitter density and localization fraction, increasing the active emitter density could result in a decreasing localization fraction and decreasing active emitter density will in turn result in a increasing localization fraction. A trade-off has to be made between the 2 multiplying factors to maximize the localization density. We could extend this idea to any possible sub-region size $a$ in order to determine the optimal active emitter density that could maximize the localization density $\rho_{\text{loc}}$. The following assumptions are made:

C1 Single emitter sub-regions that contain only 1 emitter center near the center of the regions and doesn’t have any other emitters center in any of their pixels are accepted by rejection algorithm for localization analysis and result in a localized emitter. Other sub-regions will be excluded from the fitting.

C2 Boundary effect are ignored by treating the image as a infinite plane: $A \rightarrow \infty$

From Eq. 4.3, we divide each side of the equations by the total image size $A$ on both right hand side and left hand side of the equation. We then obtain:
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

\[
\frac{N_{\text{loc}}}{A} = \frac{\binom{n}{1}(1 - p)^{n-1}}{A} \quad (4.4)
\]

Noticing that \( (\binom{n}{1}) = n \) and \( \frac{N_{\text{loc}}}{A} = \rho_{\text{loc}} \) which stands for localization density, one could write Eq. 4.5 as

\[
\rho_{\text{loc}} = \frac{n(1 - p)^{n-1}}{A} \quad (4.5)
\]

given \( n = \rho_{\text{act}} \times A \) and \( p = a/A \) we obtain

\[
\rho_{\text{loc}} = \rho_{\text{act}}(1 - a/A)^{A\rho_{\text{act}}-1} \quad (4.6)
\]

With the assumption of [C2], where total area of the image is infinite: \( A \to \infty \) with

\[
\lim_{x \to \infty} (1 + a/x)^{bx} = e^{ab} \quad (4.7)
\]

and combining Eq. 4.6,[C2] and Eq. 4.7, one then obtain

\[
\rho_{\text{loc}} = \rho_{\text{act}}e^{-a\rho_{\text{act}}} \quad (4.8)
\]

Thus Eq. 4.8 provide a theoretical calculation for localization density with given active emitter density and sub-region size. We could substitute the previous defined a and \( \rho_{\text{act}} \) which we used in previous section and obtain that \( \rho_{\text{loc}} = 0.18 \, \mu\text{m}^2 \) which is consistent with the result calculated above.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

In order to calculate the optimal active emitter density $\rho_{\text{opt}}$ where the localization density $\rho_{\text{loc}}$ is at its maximum point for a certain sub-region size $a$, the derivative of Eq. 4.8 respect to $\rho_{\text{act}}$ should vanish at this stationary point. Thus we have

$$e^{-a\rho_{\text{opt}}} - a\rho_{\text{opt}}e^{-a\rho_{\text{opt}}} = 0$$

(4.9)

which gives us $\rho_{\text{opt}} = \frac{1}{a}$. So when $\rho_{\text{act}} = \rho_{\text{opt}}$, we will obtain the maximum localization density which means that we would be able to localize the maximum number of emitters using our localization algorithm and thus performing SM-SR imaging at its maximized efficiency. To obtain the localization density at this maximum point, we substitute $\rho_{\text{opt}} = \frac{1}{a}$ into Eq. 4.8 and it gives us

$$\rho_{\text{loc}}^{\text{max}} = \frac{1}{ae}$$

(4.10)

So at the optimal density configuration where $\rho_{\text{act}} = \rho_{\text{opt}} = \frac{1}{a}$, the maximum localization density we could obtain is $\frac{1}{ae}$ where $a$ is the area size of the sub-region and $e$ is the base of the natural logarithm.

Eq. 4.8 provides a way to calculate the theoretical localization density per frame to be expected given a experimental specified active emitter density per frame. Eq. 4.10 and Eq. 4.9 together provides a method to achieve the maximum efficiency providing a specific fitting sub-region size by varying active emitter density which could be achieved by controlling the laser power or buffer condition.

The optimal active density for a sub-region size $1.3\mu m \times 1.3\mu m$ is $0.6 \mu m^2$ with maximum localization density of $0.21/\mu m^2$. The fraction of non-localized emitter can be calculated by $(0.6 /\mu m^2 - 0.21 /\mu m^2)/0.6 \mu m^2$ which equals to 65%. This result indicates that during super resolution analysis, the majority of obtained result
from the fittings that are performed on isolated sub-regions are to be discarded, the rejection algorithm, such as a threshold according previous knowledge of intensity or shape, or a test statistics used as an indicator of goodness of fitting for single emitter model.

4.1.3 Rejection Algorithm for Some Possible Complications

We can now achieve a maximum localization density by adjusting the experimental setup such as the laser power and the buffer condition which affect \( k_{\text{on}} \) and \( k_{\text{off}} \) rates. Let’s now considering achieving a resolution target of 10 nm in the optimal situation. Assuming our sub-region size equals to 1.3 \( \mu m \times 1.3 \mu m \), in order to maximize the localization density, the active density is adjusted to 0.6/\( \mu m^2 \), and the resulting localization density is 0.21/\( \mu m^2 \). To achieve the 10 nm resolution, which requires 40,000 localization per square micron, one will need at least 200,000 frames working at optimal active density to achieve a uniform resolution in this single molecule super resolution experiment. Acquisition of these images could result a long experimental time. During this long acquisition time, situations that would not be recognized in experiments with short duration, could arise as a problem during this extended period of time. During this long period of time, the microscope stage could drift in both lateral and axial directions, this can be solved by placing a fiducial beads within the sample and shifting sample plane using a piezo stage based on the position feedback from localization analysis of the fiducial beads. The object might move during the acquisition period and thus fluorophores that are originally in focus might move in and out of the focal plane or the surface plane of a TIRF microscope [5]. Fluorophores could lose their affinity of the structure and leave the structure, forming a free emitter diffusing within the sample. These free fluorophores could thereafter bind back on to the structure at various location and provide extra difficulty in analysis. Mobility of the fluorophore could also change resulting in a
change in emission pattern during the experiments. Thus when one analyzes images obtained from SM-SR experiments, usually by isolating sub-regions and localizing single emitter center, rejection algorithms should be carefully designed so that the data we obtained from the experiments won’t be misinterpreted.

4.2 Rejection Algorithm Review

4.2.1 Criteria of a Good Rejection Algorithm

As discussed in the introduction section, in order to achieve maximized efficiency in achieving target resolution in SM-SR experiments, which is achieved when maximizing the localization density, one should control the active emitter density so that it is optimal with respect to the sub-region size. When operating at this optimal density, the analysis that fits sub-regions which are isolated for each emitter has to be able to reject incorrect fittings such as biased fitting caused by influences of a nearby emitter, out of focus emitters and regions that are heavily corrupted by high autofluorescence. From the calculation made in the previous section, on average, over 70% of fitting attempts have to be rejected to avoid biased fitting result. Autofluorescence and out of focused emitters could make this percentage well above 70% for some experimental conditions.

The rejection algorithm, which not only rejects the inappropriate fittings caused by various sources but also accepts correct fitting at its maximum efficiency, is thus essential for the successes of super resolution analysis. A well defined rejection algorithm should include following properties:

1. Behavior of the rejection algorithm should be consistent for various experimental conditions, such as intensity, emission wavelength, pixel size of the EM-CCD
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

camera and etc.

2. The rejection algorithm should have a well-defined statistical meaning whose statistical properties could be included for any quantitative analysis performed based on obtained position estimates such as clustering analysis.

3. The noise model should be tested to see whether the obtained data can be generated from the proposed model under a certain noise model which could be derived from the setup of the experiments.

Given desired properties for a rejection algorithm, we propose the log likelihood ratio (LLR) as the test statistics that is used in rejection algorithm. LLR is described as

\[
LLR = -2 \ln \left[ \frac{L(\hat{\theta}|D)}{L(D|D)} \right]
\] (4.11)

Where \( L(\hat{\theta}|D) \) stands for the likelihood for a given parameter set \( \hat{\theta} \) to generate the data set \( D \), and \( L(D|D) \) stands for the likelihood for the data set to generate the data itself under a specific noise model.

We demonstrated our proposed statistics to asymptotically approaches \( \chi^2 \) distribution [6, 7] and its statistical properties using simulated data set. A description and comparison between the commonly used rejection algorithm including elliptical test [8, 3, 9] and SSE test [10, 11] are presented and comparison on performances of these algorithms on simulated data are performed and described as well.

In general, as demonstrated in the following section, the LLR is \( \chi^2 \) distributed with degrees of freedom K-N [6, 12]. It can be reliably applied as a rejection algorithm at various emitter intensities, emission wavelength and pixel sizes without
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

inconsistent rejection behavior. It not only a test on the shape similarity or absolute value similarity between model and the data, but also takes the noise model of the system under consideration. Statistical meaning of LLR based significance test can also assist further quantitative analysis.

However, in sub-regions with near zero background values, the LLR follows a \( \chi^2 \) distribution but with a different degrees of freedom other than K-N. We have proposed a Lidke-Huang hypothesis together with Lidke-Huang approximation in order to correctly calculate the degrees of freedom in low pixel count sub-regions. We have demonstrated our theory using simulated data set for both LH hypothesis and LH approximation. Both hypothesis and approximation are suitable not only for calculating LLR in SM-SR analysis but also for any application to calculate LLR test statistics under a Poisson noise assumption.

4.2.2 Brief Review of Current Rejection Algorithms

Usual methods for rejecting un-converged fits, fits that are performed on noisy peaks or fits that are biased due to influences from a nearby emitter, include thresholding on a certain fitting parameter with prior knowledge such as intensity, \( \sigma_{PSF} \), relative position of emitter centers within the sub-region, and the most importantly thresholding on a proposed test statistics. In STORM and dSTORM [8, 3, 9], Zhuang and Bates use an elliptical test where isolated sub-regions were fitted to elliptical Gaussians in order to obtain the longest axes and shortest axes of the elliptical shapes. According to the obtained axis values, ellipticity values defined as \( |(a - b)/a + b| \), where a and b are the longest and shortest axes of the elliptical shapes respectively, are calculated and thus sub-regions that leads to high ellipticity values or having large difference between the value \( a/b \) and 1 are excluded from further analysis. In PALM [8], Betzig et al, used a sum squared error test, where after obtaining the fitting result model,
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

a subtraction is performed between generated model sub-image and data sub-image. The pixel values of residual image are then squared and summed to obtain this sum squared error (SSE). Therefore larger SSE values indicate larger pixel values in the residual image and thus indicate higher discrepancy between the model sub-image and the data sub-image. Estimates resulting in a larger than threshold SSE value are rejected while estimates resulting in a small SSE values are accepted for further reconstruction of super resolution image.

4.2.3 Elliptical Test Review

As described in [8], sub-regions isolated as potential candidates of single emitters in the obtained frame, were fit using nonlinear least square regression to a continuous ellipsoidal Gaussian:

\[
I(x, y) = A + I_0 e^{[-(x'/a)^2-(y'/b)^2]/2} \tag{4.12}
\]

where

\[
x' = (x - x_0) \cos \theta - (y - y_0) \sin \theta \quad y' = (x - x_0) \sin \theta - (y - y_0) \cos \theta \tag{4.13}
\]

Where A is the background fluorescence level, \(I_0\) is the intensity of the emitter, a and b define the widths of the Gaussian distribution along x’,y’ directions and \(x_0\) and \(y_0\) define the center position of the Gaussian distribution within the sub-region and \(\theta\) is the tilt angle of the elliptical Gaussian relative to the axes along the sub-region square pixels.

Ellipticity could be then calculated as:
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

\[ \text{ellipticity} = \left| \frac{a - b}{a + b} \right| \]  

(4.14)

Since elliptical Gaussian isn’t a common terminology used to describe such a Gaussian shape, we are going to illustrate this method in another equivalent way using a more commonly used concept ‘Bivariate Gaussian’. In short the elliptical test fits the sub-region with a transformation of a bivariate Gaussian where axes have been rotated such that the long axis and short axis of the elliptical shape aligns with either x’ or y’ axis after rotation and thus covariance coefficients vanishes. Variance terms in Eq. 4.12 could then represent the longest axis and shortest axis of the ellipsoidal shape without the need of a covariance term.

To calculate the ellipticity without rotating the axis, one can use a bivariate Gaussian instead of the elliptical Gaussian form shown below,

\[
f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp \left( -\frac{1}{2(1 - \rho^2)} \left[ \frac{(x - \mu_x)^2}{\sigma_x^2} + \frac{(y - \mu_y)^2}{\sigma_y^2} - \frac{2\rho(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} \right] \right)
\]

(4.15)

where \(\sigma_x^2\) and \(\sigma_y^2\) are the variances in x and y directions respectively and \(\rho\) stands for the correlation coefficient between x and y. One could calculate the covariance matrix which is defined as

\[
\Sigma = \begin{pmatrix}
\sigma_x^2 & \rho \sigma_x^2 \sigma_y^2 \\
\rho \sigma_x^2 \sigma_y^2 & \sigma_y^2
\end{pmatrix}
\]

(4.16)

Given covariance matrix \(\Sigma\), eigenvalues of \(\Sigma\) could then be obtained, as \(\lambda_1\) and \(\lambda_2\). The square root of these two eigenvalues then represent the standard deviation of
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

longest and shortest axes of the ellipsoidal shape and could then be used to calculate the ellipticity of this bivariate Gaussian as

$$\text{ellipticity} = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_2}}{\sqrt{\lambda_1} + \sqrt{\lambda_2}}$$  \hspace{2em} (4.17)$$

The obtained ellipticities are used as a threshold rejection statistics where any sub-regions that result in larger than 15% (0.15) ellipticities are rejected and thus excluded from further analysis.

4.2.4 SSE Test Review

In PALM [10], obtained frames are first converted into photon counts by subtraction of a constant dark state offset (and divided by the camera gain). A differential image is obtained by performing a subtraction in current image from the last adjacent image. Thus any molecule disappearing or appearing events could be shown on the differential image as a intensity peak either negative or positive. A small sub-region is isolated centered around the center of the intensity peak in the differential image and areas where the sub-region occupies in image $image_n$ is filled with background from $image_{(n+1)}$. So that if the molecule also exists in frame n-1, it will also appear as a peak in the next differential image. This process is repeated frame by frame backward to the beginning to identify and isolate the data from all significant peaks in the image stack.

If isolated sub-region contains peaks that are closer than 1 pixel in consecutive frames they will be recognized as a single molecule and sub-regions will be summed together for analysis. After fitting processed (summing over same individual molecule) sub-regions using equal variate two-dimensional Gaussian,
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

\[ f(x, y) = I_0 \exp \left[ -\frac{(x - x_0)^2}{2\sigma^2} - \frac{(y - y_0)^2}{2\sigma^2} \right]. \]  

(4.18)

Each pixel is then weighted by the inverse of the squared residual and then fit with weighted least square fitting. Parameter estimates from 2nd weighted least square fitting trial is chosen as the result for localization analysis for the sub-region. Fitting results are stored as well as the sum of the squared error (SSE) which can be described as

\[ sse = \sum_{i=1}^{K} (d_i - u_i)^2 \]  

(4.19)

where \( i \) is individual pixel number, \( K \) is the total number of pixels in the sub-region, \( d_i \) is the counts for each pixel from the data and \( u_i \) is the expected count value calculated from the fitting model.

This SSE value is then to be used in a threshold rejection statistics where any fitting with SSE values bigger than \( SSE_{\text{min}} \) is excluded from further reconstruction processes.

### 4.2.5 Poisson Distributed Noise Model in PALM

Noticing that in localization analysis described in PALM [10], sub-regions are isolated from a residue image which are the result of a subtraction of the current frame from the previous frame. Subtraction process operated on images with Poisson noised photon counts could be problematic since the result from subtraction of two Poisson noised image are no-longer Poisson distributed any more. In general the subtraction Process introduces extra spreading of the noise distribution around the mean.
and the results in a distribution which has a variance that equals to the summation of variances from the previous Poisson and is described by Skellam distribution. However, assuming single emitter events using subtraction method often comes from subtraction of a single emitter emission pattern and low count background, Skellam distribution approaches to a Poisson distribution in this case and thus we are going to assume sub-regions obtained from this process still follow Poisson distribution model.

4.3 Rejection Algorithms Demonstrations

4.3.1 Acceptance Demonstration for Elliptical Test and SSE Test

In order to demonstrate how the above rejection algorithm works, simulated sub-regions are generated by placing 2-D equal variate Gaussian blob in the center of image where pixel intensities are calculated by integrating the 2-D Gaussian over the pixel range for each pixel as described in previous chapter.

\[ \mu_k(x, y) = I_0 \Delta E_x(x, y) \Delta E_y(x, y) + b_0 \]  

(4.20)

where \( \mu_k(x, y) \) is the expected photon count for a given pixel \( k \), \( I_0 \) is the total emitted photon counts expected, \( b_0 \) is the background and \( \Delta E_x(x, y) \) and \( \Delta E_y(x, y) \) are:

\[ \Delta E_x(x, y) = \frac{1}{2} \left( \text{erf} \left( \frac{x - x_0 + \frac{1}{2}}{\sqrt{2}\sigma_0} \right) - \text{erf} \left( \frac{x - x_0 - \frac{1}{2}}{\sqrt{2}\sigma_0} \right) \right) \]  

(4.21a)

\[ \Delta E_y(x, y) = \frac{1}{2} \left( \text{erf} \left( \frac{y - y_0 + \frac{1}{2}}{\sqrt{2}\sigma_0} \right) - \text{erf} \left( \frac{y - y_0 - \frac{1}{2}}{\sqrt{2}\sigma_0} \right) \right) \]  

(4.21b)
Where $x_0$ and $y_0$ are emitter positions and $\sigma_0^2$ is the variance of the 2-D Gaussian.

The simulated sub-regions are then corrupted by Poisson noise to simulated noise model coming from EM-CCD camera. Then rejection procedures of elliptical test and the SSE test are performed as described in previous section.

For the elliptical test, 1000 single emitter sub-regions were corrupted individually by Poisson noise and were fitted using bivariate Gaussian method and then elipticity values are calculated after obtaining the eigenvalues from the resulting covariance matrices $\Sigma$. Histogram of ellipticity values for these 1000 sub-region is shown below where the elipticity threshold 15% shown in the same plot as a dashed vertical line. Any sub-regions that would result in elipticity values bigger than 15% will be excluded from further analysis.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.1: Histogram of ellipticity for isolated emitter sub-region. 1000 single emitter sub-regions corrupted individually by Poisson noise are fitted using bivariate Gaussian method and then ellipticity values are calculated after obtaining the eigenvalues from the resulting covariance matrices $\Sigma$. Histogram of ellipticity value for these 1000 sub-region is shown below where ellipticity threshold 15% is also shown in the same plot as a red dashed vertical line which indicates any sub-region that would result in a ellipticity value bigger than 15% (to the right of the threshold line) will be excluded from further analysis.

Fig. 4.1 demonstrates that the ellipticity test used in [8, 3] performed pretty well when accepting isolated emitter sub-region as valid sub-region for further analysis.

For SSE test, subregions are fitted to the same 2-D Gaussian using non-Linear Least Square (LS) regression in order to find position estimates. Based on the esti-
mate result from LS fitting, a model image is then generated using the the same 2-D Gaussian method used in fitting process and the sum of the squared residue (SSE) are calculated and histogram plot is generated from the SSE values obtained for the 1000 individually Poisson noise sub-regions.

![Histogram of SSE and a proposed threshold for isolated emitter sub-region.](image)

Figure 4.2: Histogram of SSE and a proposed threshold for isolated emitter sub-region. 1000 single emitter sub-regions corrupted individually by Poisson noise are fitted using LS fitting and then SSE value defined in Eq. 4.19 are plotted as histogram. A proposed threshold which associate with a possible failure rate 1600 for this specific case are shown as a red dashed line. SSE rejection algorithm will reject any SSE result from fitting that are larger than this threshold value (to the right of the red line) from further reconstruction process.

Fig. 4.2 demonstrates the acceptance case for SEE test used in PALM and fPALM
type of analysis. However, there’s no information about which specific threshold for SSE value that is used in [8] other than a brief mention of $\chi^2 < (\chi^2)_{\text{min}}$. The inability to assign a specific value for the SEE test is caused by the intrinsic property of the SEE test: The center of distribution of SSE test statistics shifts dramatically with various emitter intensity level. Thus any preset thresholds behave differently in rejection behavior under various emitter intensities.

\subsection{4.3.2 Rejection Demonstration of Elliptical Test and SSE Test}

In order to demonstrate the rejection processes from both test, based on the single emitter image that is generated in the previous section based on the integral of the 2-d Gaussian, another Gaussian shaped emitter is added 2 pixels away from the first emitter. This 2-emitter image should then be rejected by rejection algorithm as having multiple emitters within the sub-region which would lead to biased or incorrect localization if accepted. Thus it is used to demonstrate the rejection process for both of the rejection algorithms.

For the ellipticity test, sub-regions are again fitted using bivariate Gaussian method and then elipticity values are calculated after obtaining the eigenvalues from the resulting covariance matrices. Histogram of ellipticity values is plotted where elipticity threshold 15\% is shown in the same plot as a dashed vertical line.
Figure 4.3: Histogram of ellipticity. 1000 sub-regions that are the same configuration which has 2 emitter with a distance of 2 pixel away from each other are analyzed using elliptical test method. Ellipticity obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the elliptical test. Threshold of 0.15 ellipticity are also shown in the figure as a red dashed line. Sub-regions that result in a larger than threshold ellipticity value would then be excluded from future analysis.

Fig. 4.3 demonstrated that the ellipticity test used in [8, 3] type of analysis rejects properly when 2 emitters are both within the same sub-region with a distance of 2 pixels between them.

For SSE test, sub-regions are fitted to the same 2-D Gaussian using MLE. Based on the obtained estimate from MLE, a single emitter model image is then generated.
and the sum of the squared residue (SSE) is calculated and shown in histogram.

![Histogram of SSE values](image)

**Figure 4.4:** Histogram of SSE values. 1000 sub-regions that are the same configuration which has 2 emitter with a distance of 2 pixel away from each other are analyzed using SSE test method. SSE values obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the SSE test. Threshold SSE of 1600 are also shown in the figure as a red dashed line. Sub-regions that result in a larger than threshold SSE value are excluded from future analysis.

Fig. 4.4 demonstrated that SSE test properly when 2 emitter are both within the same sub-region with a distance of 2 pixels between them. Most of the SSE values obtained from fitting 2 emitter image using single emitter model are bigger than the threshold which are rejected by the SSE test correctly.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

4.3.3 Discussion on Demonstrations of Acceptance and Rejection

Here we have demonstrated both acceptance and rejection processes relating to both elliptical test and SSE test. In the acceptance demonstration, we have assumed fixed intensities throughout the test in order to make demonstration simple. However, in realistic situations, emitter intensities vary dramatically in different experimental-dye setup and thus fixed intensity won’t be necessarily a completely test for the behavior of the rejection algorithms. When we include intensity fluctuations of the simulated emitters, we can then start to notice the problem introduced by these rejection algorithms. For rejection process demonstration where we introduced another emitter which locates 2 pixels away from the center emitter to be rejected by described algorithm, we could also introduce emitters that has a various distance from the center emitter to better simulate the realistic situation which is done in the following section. However, doing so also leads to a unexpected behavior for these rejection algorithms and thus demonstrated below as the limitation of the pre-described rejection algorithms.

4.3.4 Demonstration of Rejection Error in Elliptical Test

2 emitter sub-region are simulated to demonstrate the limitation of the elliptical test on rejecting non-single emitter sub-regions.

First, a single emitter is generated in the center of sub-region using Eq. 4.20, and another emitter is then placed at a random position which has a distance \( d \) from first emitter. The distance \( d \) is varied from 1-5 pixel where at each specific configuration, 1000 sub-regions in this configuration are generated by placing the 2nd emitter on any random positions that have a distance \( d \) from the 1st emitter center.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Ellipticity values are calculated as described in Section 4.2.3, and then are shown using histograms in each configuration. The 0.15 ellipticity threshold is also shown in the figure as a 2-D vertical plane.

![Histograms of ellipticity for simulated sub-regions.](image)

Figure 4.5: Histograms of ellipticity for simulated sub-regions. For every configurations \((d = 1, d = 2, d = 3, d = 4 \text{ and } d = 5)\), 1000 Poisson noised sub-regions that are the same configuration which has 2 emitter with a distance of \(d\) pixel away from each other are analyzed using elliptical test method. Ellipticity obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the elliptical test. Threshold of 0.15 ellipticity are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold ellipticity value would then be excluded from further analysis.

As shown in Fig. 4.5, configurations when distance between the 2 emitters \(d\) equals 2,3, and 4 pixels are properly rejected by the ellipticity test while configurations with \(d = 1\) and 5 are mostly accepted where they should be rejected as having more
than 1 emitter in the sub-regions which introduce biased fits. In the $d = 1$ case, sub-regions that contain 2 fluorophores that are very close with each other (1 pixel=106 nm) result in closely overlapping Gaussian distribution, fitting elliptical Gaussians on these sub-regions result in semi circular Gaussian shapes instead of the elongated Gaussian shapes. Thus, the ellipticity values obtained from the elliptical test are close to 0 and then are accepted as having ellipticity values smaller than the cut off value of 0.15. For $d = 5$ case, sub-regions contains 2 fluorophores that are relatively far from each other, and thus since one of the emitter locates in the middle of the image and the other emitter locates outside of the sub-region, the other emitter is still close to the boundary of the sub-region so that some part of its emission pattern is still within the sub-region. Since the sub-region contains part of the emission pattern from outside emitters, it affects the result obtained from the localization analysis and leads to a biased position estimates. Sub-regions of this type should be rejected to avoid biased estimates and misinterpreting the data set. However, elliptical test failed to reject these regions since such non-overlapping emission pattern isn’t producing a big enough ellipticity value.

4.3.5 Demonstration of Acceptance Error in SSE Test

To demonstrate the dependence of $SSE$ with intensities of the emitters, another set of single molecule sub-regions is generated by placing 2-D bivariate Gaussian emitter in the center of sub-region. Intensity $I_0$ of the generated emitter varied from 300-5000, and for each intensity configuration, 1000 identical sub-regions are generated and then corrupted using poison noise individually. $SSE$ values are calculated using algorithm described in the above Section 4.2.4 based on position estimates obtained from fitting 2-D Gaussian as in Eq. 4.18. $SSE$ values for each configurations are shown using histogram and a $SSE$ threshold is shown as a 2-D vertical plane.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.6: Histogram of $SSE$ for simulated sub-region with single emitters series of various intensities and threshold the of 1600. For every configurations ($N = 300, N = 800, N = 1200, N = 2000$ and $N = 5000$), 1000 Poisson noised sub-regions that are the same configuration are analyzed using $SSE$ test. $SSE$ values obtained for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the acceptance process of the $SSE$ test at various intensities. Threshold of $SSE = 1600$ are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold $SSE$ value would then be excluded from future analysis.

Fig. 4.6 shown above demonstrates the inconsistent acceptance behavior for emitters with various intensities. Low intensity emitters ($N = 300$) could easily pass the threshold at a 100% rate as shown in figure while isolated emitters with higher intensities ($N = 1200$) have a more than 70% chance to be rejected regardless that they are still isolated emitter sub-regions. Emitters with extremely high intensities ($N = 5000$), although being isolated emitters within each sub-region and providing
reliable position estimates, will be rejected at almost 100% rate. Since a higher intensity would provide a higher count for each pixel, under Poisson noise, fluctuation is also higher.

### 4.3.6 Discussion

As shown in the previous two sections, the unexpected behaviors for both elliptical test and SSE test make them unreliable and inconsistent. Noticing that threshold values used in both test are also an arbitrary number obtained purely from user experiences and objective assumptions, it would be very hard to predict the behavior of the rejection process in much more complicated realistic experimental situations with the presence of cell autofluorescence and binding and unbinding of the dyes,...etc. As mentioned in Section 4.1, over 60% of sub-regions are subjected to be rejected in order to produce reliable estimates of emitters, a reliable and predictable rejection algorithm is in great need for SM-SR analysis in order to unbiasedly interpret obtained data set. A potential property that is favorable for rejection algorithm is the ability to be rigorously explained in statistical analysis which could then be combined with other analysis method in order to potentially develop new analysis method to extract biologically interesting information.

### 4.4 Log-likelihood Ratio as a Rejection Algorithm

#### 4.4.1 Introduction to the LLR test

In order to obtain a rejection algorithm that maintains good statistical properties such as consistent rejection behavior at different intensity levels, being able to predict a goodness of fit based on specific noise model and possessing a statistical ex-
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Planation, we implemented a rejection algorithm based on using Log-likelihood ratio statistics. Log-likelihood ratio (LLR), sometimes mentioned as “deviance” in statistical literature, has been a well known goodness of fit test for Poisson noised data from mid twentieth century. LLR are defined as the logarithmic of the ratio between the likelihood for the model to generate data set and the likelihood for the data to generate itself under the proposed noise model.

\[
LLR = -2 \ln \left[ \frac{L(\hat{\theta}|D)}{L(D|D)} \right] \quad (4.22)
\]

where \( L(\hat{\theta}|D) \) denotes the likelihood for the proposed model \( \hat{\theta} \) to generate data set \( D \), and \( L(D|D) \) is the likelihood for the data set to generate itself under the proposed noise model.

For example, under Poisson distributed noise assumption, for a single pixel on a CCD camera which is setup to detect photons. From the theoretical calculation, we are expecting to get 100 photons \( (m = 100) \) in this pixel during a certain exposure time. In a certain photon counting experiment, the photon we obtained from that pixel is 105 \( (d = 105) \) which is not exactly the expectation value \( m \) due to Poisson noise (shot noise). Thus the probability for obtaining 105 photons experimentally given our expecting \( m = 100 \), or in another word, the likelihood of \( m = 100 \) given \( d = 105 \) is calculated by Poisson distribution:

\[
L(m|d) = P(d|m) = \frac{e^m m^d}{d} \quad (4.23)
\]

Where \( L(m|d) \) stands for the likelihood function and \( P(d|m) \) is the probability density function. \( m \) stands for the expectation value and \( d \) is the observed count in the detector.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

One should notice that this probability is a relatively small which would be varying depending on the expectation value \( m \) and the random variable \( d \). The likelihood for the model \( m \) to generate the data set, in our case a single data value \( d \), is then 0.034. The likelihood for the data to generate itself under Poisson noise is then \( L(d|d) \) which equals to 0.039. LLR could be then calculated by calculating the logarithm of the ratio from these 2 likelihood and multiplying it with \(-2\), as in Eq. 4.22. The LLR for our expectation of 100 photons and data of 105 photons are then 0.27. Noticing that providing this LLR value of 0.27 here doesn’t provide any information for helping to decide whether our model is correct for the data or not. We should then proceed to investigate the statistical properties of LLR in order to make any conclusion based on this observed LLR in rigorous way.

We notice that LLR decreases with increasing model to data likelihood \( L(m|d) \) and thus a smaller LLR value indicates higher likelihood, \( L(m|d) \) values and a high LLR value would mean a low likelihood values for \( L(m|d) \). In order to decide whether the result of our theoretical calculation, 100 photons, is a correct model for our experimental data set, in our case a single data, 105 photons under Poisson noise, a decision mechanism such as a threshold, is needed. For example if the LLR is higher than a threshold, we could reject this model and if the LLR is lower than a threshold, we could accept this model.

In model testing, the p-value is often generated to provide such a mechanism which stands for the probability for the model to obtain a test statistics (in our case LLR) from a data set which is at least as extreme as the observed data or worse given the proposed model is correct. This p-value is then compared with the significance level \( \alpha \) which is often set to 0.05 or 0.1. In order to calculate the p-value, the test statistic have to follow a certain well defined distribution so that one could obtain the cumulative distribution function (CDF) for its distribution. Thus if the distribution of our test statistics are known or derivable, we could then calculate the p-value for
the model and compare it with significance level $\alpha$ in order to reject or accept the model.

As mentioned in the next section that the proposed LLR follows a $\chi^2$ distribution with a degrees of freedom, $K-N$ where $K$ is the total number of pixels in the sub-region and $N$ is the total number of parameters we are trying to estimate in the model.

### 4.4.2 LLR in Literature

Scientists have made great efforts in finding a test statistic as a goodness of fit parameter ever since the beginning era of the statistics subject. Pearson (known as the "Father of the statistics") have proposed ‘Pearson’s test’ as the goodness of fitting which applies perfectly on normal distributed random variable. Loglikelihood ratio test have been brought up as a goodness of fitting statistics immediate following the recent development of Maximum Likelihood Estimator (MLE). Unlike Pearson’s test that applies on normally distributed data sets, LLR could be applied on various distributions (such as Poisson distribution, Gaussian distribution) as long as it comply the regulatory condition which is stated in [13]. What is also provided in [13] is the derivation of the $\chi^2$ distributed LLR for identical independent distributed (i.i.d.) random variable with a degree of freedom of 1. The case of independent identically distributed (i.i.d.) random variable’s goodness of fitting test is considered as standard case in modern literature. However in model testing for parameter fitting model, random variables are not all necessarily i.i.d.s because parameters in the model could introduce reduction of freedom among data bins. It has been proved that in the non-standard case, LLR follows $\chi^2$ distribution with degrees of freedom,$K - N$ where $K$ is the total number of data bins and $K$ is total number of free parameters in the model [6]. In SM-SR analysis, $K$ will be the total number of
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

pixels in the sub-region and N is the total number of free parameters in the single emitter model such as background, intensity, position estimates \((x_0, y_0, z_0)\) for 3-D SM-SR analysis.

### 4.4.3 LLR Demonstration on a Single Pixel

Given the theorem proved in [13] that LLR for i.i.d should follow \(\chi^2\) distribution with degrees of freedom of 1, one can calculate the p-value for our LLR of 0.27 by evaluating CDF of \(\chi^2_{df=1}\) distribution integrating the \(\chi^2\) PDF from 0.27 to infinity.

\[
p = \int_{x=0.27}^{\infty} \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}} dx \tag{4.24}
\]

As the result, our p-value for our single pixel example is 0.6 which is above the significance level 0.05 and it means the probability for our model to generate a data set that is at least as extreme as the observed data set or worse is 0.6 (which means it’s quite often), and thus we accept this model.

### 4.4.4 LLR as a Rejection Algorithm in SM-SR Analysis

In super resolution analysis, a single emitter within a sub-region is fit with 2-D Gaussian to extract its localization information. The result obtained from the localization fitting routine would then result in a 2-D Gaussian model which is proposed to be the model for the single emitter within the sub-region. To determine whether or not to reject this proposed model, we calculate the LLR based on the model and sub-region data.

First, one new sub-region which is the same size as the fitting sub-region is generated. Based on the position obtained from the fitting, a single emitter model is gener-
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

ated using 2-D Gaussian and placed on the new sub-region at the resulting estimated position. LLR are then calculated from the data sub-region \((D = d_1, ..., d_i, ..., d_K)\) and newly generated model sub-region \((M = m_1, ..., m_i, ..., m_K)\) as follows:

\[
LLR_{SR} = \sum_{i=1}^{K} [-m_i + d_i + d_i \ln(m_i) - d_i \ln(d_i)]
\]  

(4.25)

where \(i\) is the individual pixel number and \(K\) is the total number of pixels in the sub-region.

LLR obtained from the previous calculation will be distributed as a \(\chi^2\) distribution with a degrees of freedom \(K-N\) where we denote as \(\chi^2_{K-N}\). Thus, one could obtain the p-value for the model-data set from the distribution \(\chi^2_{K-N}\) and then compare it with the significance level \(\alpha\) which could be defined by the analysts based on their practical interest.

As discussed above, LLR follows a \(K-N\) distribution which doesn’t depend on the intensity of the single emitter or the relative pixel size of the system. The distribution of our test statistics provides a consistent rejection behavior for all emitters with various intensity or \(\sigma_{PSF}\). Using LLR as the test statistic for model rejection will provide a consistent and statistically reasonable rejection behavior whose properties could also be included and analyzed quantitatively in further analysis, such as clustering analysis. As the noise model (in our case, Poisson noise) and fitting parameters are all included in the calculation, LLR essentially provide a a measure of goodness of fit for both localization parameters and the noise model.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

4.5 LLR Implementation and Demonstrations

4.5.1 Demonstration of Acceptance and Rejection Behavior of LLR in SM-SR

In previous section, we have proved that LLR follows $\chi^2_{K-N}$ independent of emitter intensity, relative PSF sizes and pixel size on the camera and we also provided a p-value threshold (significance level) method in order to provide a reliable, statistically explainable rejection behavior. In this section, we demonstrated LLR acceptance and rejection behavior using simulated data set which ellipticity test and SSE test have shown their limitation on in Section 4.3. Our proposed LLR rejection algorithm is implemented for both simulated data set and either completely or partially solved the problem which other rejection algorithms have encountered. More importantly, the rejection behavior of LLR can now be explained using statistics and we no longer have to depend on an arbitrary threshold which is objectively set based on experience. Instead, we now use a significance level of 5% as the rejection threshold to decide for acceptance or rejection.

4.5.2 LLR Acceptance Versus SSE Test

Using LLR as our rejection algorithm, we tested our proposed rejection algorithm on the similar simulations mentioned in testing sse test. Tests were performed on the sub-region series containing single emitter with various intensities. As previously mentioned, a set of single molecule sub-regions were generated by placing 2-D bivariate Gaussian emitter in the center of sub-region. Intensities (I) of the generated emitter were varied from 300-5000, and for each intensity configuration, 1000 identical sub-regions were generated and corrupted using poison noise and then fitted using LS fitting routine. LLR was calculated for each fitting results as discussed in
previous section and plotted as a histogram each for the individual configuration. A p-value threshold of 5% is also shown in the figure as vertical plane.

![Histogram of LLR for simulated sub-regions with single emitters series of various intensities and a p-value threshold of 5%.](image.png)

**Figure 4.7**: Histogram of LLR for simulated sub-regions with single emitters series of various intensities and a p-value threshold of 5%. For every configurations ($I = 300, I = 800, I = 1200, I = 2000$ and $I = 5000$), 1000 Poisson noised sub-regions are analyzed using LS fitting. LLR values obtained for each individually Poisson noise corrupted sub-region and its fitting model are plotted as histograms to illustrate the acceptance processes of the LLR test at various intensities. A p-value threshold of 5% are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than p-value threshold LLR value would then be excluded from future analysis.

As shown in Fig. 4.7 and compared with Fig. 4.6, using LLR as the rejection statistics has solved the problem of inconsistent rejection behavior when using SSE as the rejection algorithm. All configurations at various intensities generate a $\chi^2$
distributed LLR with a degrees of freedom 45. Given such a consistent rejection behavior, we could use the p-value threshold (significance level) method to provide a consistent rejection behavior for all various intensities of emitters or even different types of emitters. Although the degrees of freedom of $\chi^2$ distribution changes with different sub-region sizes, as long as the p-value threshold is fixed, we could also obtain consistent rejection behavior among various size of the sub-regions, various fitting parameter numbers and it is also consistent through different models such as equal variate model or bi-variate model in either 2D or 3D cases.

4.5.3 LLR Rejection Versus Elliptical Test

We then tested our proposed rejection algorithm on the similar simulations mentioned in testing elliptical test. Similarly as in section 4.3.4, a single emitter was generated in the center of sub-region using Eq. 4.20, and another emitter was then placed at a random position which is distance $d$ away from first emitter. The distance $d$ was varied from 1-5 pixel, where at each specific distance value, 1000 sub-region in this configuration were generated by placing the 2nd emitter on any random position that has a distance $d$ from the 1st emitter center with a polar angle $\theta$ ranged from $0^\circ$ to $360^\circ$. Fittings were then performed using LS fitting routine and after obtaining the fitting result, LLR values were calculated for each fitting results and plotted as histogram for an individual configuration for $d = 1 - 5$. A p-value threshold of 5% is also shown in the figure as vertical plane.
Figure 4.8: Histogram of LLR obtained from fitting result and a 5% p-value threshold. For every configurations \( (d = 1, d = 2, d = 3, d = 4 \text{ and } d = 5) \), 1000 Poisson noised sub-regions that are the same configuration which has 2 emitter with a distance of \( d \) pixel away from each other are analyzed using LS fitting to a 2-D Gaussian. LLR values obtained from the results for each individually Poisson noise corrupted sub-region are plotted as a histogram to illustrate the rejection process of the LLR goodness of fitting test. A 5% p-value threshold are also shown in the figure as a light blue semi-transparent surface. Sub-regions that result in a larger than threshold LLR value would then be excluded from future analysis.

From Fig. 4.8 and compare with Fig. 4.5, we still have the same problem as we had when using elliptical test with closely overlapping emitter. However, sub-regions where emitters are separated with a distance of 5 pixels, are properly rejected using LLR. Although, closely overlapping emitters were not correctly rejected using both
method, we can use existing knowledge for the target emitter to setup an intensity threshold so that the high intensity result from fitting these overlapping emitters with single emitter model would be identified as a non-single emitter sub-region and then the sub-region is rejected by the intensity threshold.

### 4.5.4 Discussion

As previously demonstrated in this section, although SSE test provides a inconsistent rejection behavior across various intensities of the emitters, LLR provides a reliable rejection behavior in various intensities. In rejection behavior demonstration, elliptical test fails to reject sub-regions with multiple emitter within the sub-region and LLR provides a partial solution for this problem. LLR is reliable on rejection of multi-emitter sub-region when emitters are more than 1 pixels distance with each other. However, when emitter are close to each other and fitting routine could mistakenly fit the 2 emitter using single emitter model and thus provide a biased localization, LLR alone isn’t enough to reject this kind of sub-regions. But by combining with intensity threshold method based on prior knowledge, we could expect to reliably reject these sub-regions which both LLR and elliptical test failed to reject. In general, using LLR as the test statistic for model rejection provides a consistent and statistically reasonable rejection behavior. Its properties could also be included and analyzed quantitatively in further analysis, such as clustering analysis. As the noise model and fitting parameters are all included in the calculation, LLR essentially provides a measure of goodness of fit for both localization parameters and the assumed noise model.

However, as Poisson distribution has a finite parameter space where $m \in (0, \infty)$, the property of LLR when as $m \to 0$ is yet to be investigate. Although in usual case, obtained images in SM-SR experiment would rarely have a background count near
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

0, there’s still cases we met that falls into this region. The following section provides a theory to calculate the correct distribution of LLR when there’s low count pixels within the sub-region.

4.6 LLR at small values

4.6.1 Introduction

It is shown in section 4.4, that log likelihood ratio for a single bin data set would follow a chi-square distribution when statistical distribution of the noise model comply certain regulatory conditions mentioned in [13, 7]. However, although LLR under Poisson distributed noise asymptotically approaches a \( \chi^2 \) distribution with degree of freedom of \( K - N \), LLR for data set with bins which have low expectation values (low expected bin count), does not follow \( \chi^2_{K-N} \) \[12\]. This situation could also happen in our single-molecule super resolution analysis where sub-regions have a lot low count pixels such that assuming LLR for these sub-region following a \( \chi^2_{K-N} \) distribution would give incorrect rejection behaviors.

4.6.2 Mean of LLR

To potentially correct this problem, we took the mean of our LLR for a single bin with Poisson distributed noise data. As our LLR is asymptotically approaching \( \chi^2 \) with \( df = 1 \) and \( \chi^2 = df \) is a property of \( \chi^2 \) distribution, we expect the mean of the LLR asymptotically approaches 1. The mean value of LLR for Poisson noise case is given by

\[
\langle LLR \rangle = 2 < m - d + d \ln d - d \ln m >
\] (4.26)
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

with definition of mean operation for Poisson noise as

\[ < x > = \sum_{0}^{\infty} x P(d|m) \]  \hspace{1cm} (4.27)

where \( P(d|m) \) is the pdf of Poisson distribution

\[ P(d|m) = \frac{e^{m} m^d}{d!} \]  \hspace{1cm} (4.28)

Noticing that in equation 4.26, \(< m > = m, < d > = m\), one could write Eq. 4.26 as

\[ \langle LLR \rangle = 2(-m \ln m + \sum_{0}^{\infty} d \ln dP(d|m)) \]  \hspace{1cm} (4.29)

thus,

\[ \langle LLR \rangle = 2(-m \ln m + \sum_{0}^{\infty} d \ln \frac{e^{m} m^d}{d!}) \]  \hspace{1cm} (4.30)

Eq. 4.30 provides us a theoretical way to calculate the mean of LLR under Poisson noise model for a single bin and we then evaluated it in a range of \( m \) values numerically using Eq. 4.30.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.9: Numerical evaluation of $\langle LLR \rangle$ as a function of $\lambda$ using Eq. 4.30. $\langle LLR \rangle$ starts from 0 as $\lambda \to 0$ and overshoots to about 1.2 at $\lambda = 1.33$ and then falls back down and approaches 1 as $\lambda > 10$.

The above figure shows the behavior of $\langle LLR \rangle$ as a function of $m$. $\langle LLR \rangle$ rises quickly from 0 at $m \to 0$ to about $m = 0.5$ when it reaches 1. After it overshoots to about 1.2, it falls back to asymptotically approaches 1 as $m > 10$. 

110
4.6.3 The LLR is not $\chi^2$ Distributed at Small Expectation Values

The effect of the $\langle LLR \rangle$’s divergence from 1 when $\lambda$ is small affects our rejection algorithm when at least 1 pixel has a low expectation counts. The following simulated sub-regions that consist of various numbers of low count pixels have been made to illustrate this effect. To isolate the problem, we now no longer perform fittings to the sub-regions but directly compare the model sub-region with its Poisson noised sub-region which would then follow $\chi^2_K$ since $N = 0$ without fitting performed. Series of sub-regions are generated with 1 emitter in the center which have a intensity of 800 and $\sigma_{PSF}$ of 1.2 and background for these sub-region are set to 0.01 to introduce low count pixels. In the five different configurations, sub-region sizes are set at 5 different values such as, $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. For each configuration, 1000 identical sub-regions were generated and corrupted using Poisson noise individually. The LLR values obtained from the model sub-region and Poisson noised sub-region was calculated using Eq. 4.22 and histogram of the LLR is plotted for each configuration. Along with each histogram, $\chi^2$ distributions with degrees of freedom of $K$ which is the number of pixel of the sub-region, are also plotted to illustrate the discrepancy from LLR distribution and $\chi^2_K$. 

111
Figure 4.10: Histograms of LLR values obtained 5 different configuration with different sub-regions sizes. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. Thus with increasing sub-region size, the number of pixels that are low-count pixels increases in each configuration with sub-region size. Also shown on the plot is the $\chi^2$ distribution with degrees of freedom $K$ for each configuration which is expected to be the distribution of LLR values.

As shown in Fig. 4.10, in the case of $4 \times 4$ sub-regions, single emitters in the center with a $\sigma_{PSF} = 1.2$ pixel resulted in sub-regions that are filled with the single emission patterns. Thus no low value pixels exists in these sub-region configuration and LLR distribution follows $\chi^2_K$ distribution where $K = 16$. On the other hand, when sub-region size equals $12 \times 12$, a significant number of pixels have intensity contributions
from the emitters and with a low background setup (0.01), the expected count values for these pixels are around 0. Thus in the configuration of $12 \times 12$ sub-regions, LLR distribution does not follow $\chi^2_{12 \times 12}$. For configurations with sub-region sizes of $16 \times 16$ and $20 \times 20$, a even large number of low count pixels introduces even more significant discrepancy between LLR distribution and $\chi^2_K$.

4.6.4 Proposed Correction Method I - Add Background

We here proposed 3 methods to solve this issue concerning the discrepancy of LLR from $\chi^2$ distribution when bins with low expectation values are present.

Noticing that this problem is caused only by bins with low expectation values, one of the ways to fix this issue is to add a random Poisson distributed background noise in the sub-region. In this way, the background is increased and no longer being around 0 where the discrepancy occurs. Since the summation of 2 Poisson random variables is still a Poisson random variable, we didn’t alter the noise model at all by adding the Poisson noised background. A demonstration of this method is shown below where we add a poison noised background whose expectation value is 10 in the sub-region series which result in all pixel counts in the sub-regions series are above the discrepancy region.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.11: Histograms of LLR values obtained 5 different configuration with different sub-regions sizes with addition of the Poisson noised background and its corresponding $\chi^2_K$ distribution. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. Each sub-region are then added an additional Poisson noised background with expected count of 10 for each pixel. LLR are calculated using 4.22 and shown using histogram with the corresponding $\chi^2_K$ for each configuration.

Adding additional Poisson noised background on the original sub-region will increase the expected background level from 0.01 to 10 and thus no-pixel in the sub-region could then be categorized as low count pixel. As shown in Fig. 4.11, in all size configurations, LLR follows a $\chi^2_K$ exactly and thus demonstrated the success in solving the discrepancy issues.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

However, this method could only be used if the signal to noise ratio is high in the original sub-region. When the signal to noise ratio is low in the sub-region, adding additional Poisson noised background would result in a dramatic increase in uncertainty in the position estimation and might result in a total inability for the localization routine to localize the emitter centers.

4.6.5 Proposed Correction Method II - LH Hypothesis With Numerical Evaluation

We here proposed two methods to solve this problem without adding background to the image. These two method are all based on the following Lidke-Huang hypothesis, which is yet to be proved at the time of publishing of this manuscript. Given the
definition of LLR that the log likelihood ratio test statistics for Poisson distributed data set is defined as:

\[ LLR = -2 \sum_{i=1}^{K} \left[ -m_i + d_i + d_i \ln(m_i) - d_i \ln(d_i) \right] \] (4.31)

where \( K \) is the total data bin number, \( m_i \) is the expected count for the \( i^{th} \) bin and \( d_i \) is the \( i^{th} \) bin count from the trial, Lidke-Huang hypothesis can then be described as:

**Lidke-Huang Hypothesis:** The test statistics distribution of Loglikelihood ratio (LLR) for Poisson distributed data set, as defined in Eq. 4.31, follows a \( \chi^2 \) distribution with degrees of freedom \( \sum_{i=1}^{K} \langle LLR(m_i) \rangle \) which could be non-integer and \( \langle LLR(m_i) \rangle \) are defined as:

\[ \langle LLR(m_i) \rangle = 2(-m_i \ln m_i + \sum_{k=0}^{\infty} k \ln k \frac{e^{m_i} m_i^k}{k!}) \] (4.32)
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Lidke-Huang (LH) hypothesis provided a way to calculate the distribution of LLR with an extended $\chi^2$ distribution with non-integer degrees of freedom for Poisson noised data set. One could notice that the degrees of freedom defined in LH hypothesis, $\sum_{i=1}^{K} \langle LLR(m_i) \rangle$, together with Eq. 4.32 asymptotically approaches to the total number of bins in the data set as every $m_i \to \infty$. However, LH hypothesis also provide the correct distribution when low count bins exists in the Poisson noised data set and original log likelihood ratio theory is broken in these situations.

Based on LH hypothesis, we propose another method that by using LH hypothesis to obtain the degrees of freedom of the expected $\chi^2$ distribution to provide a correction of the $\chi^2$ distribution so that it is used to describe the LLR distribution.

To evaluate Eq. 4.32, numerical calculation could be used in order to calculate $\langle k \ln k \rangle$ by summing from 0 to 100 instead of 0 to $\infty$. As $\langle LLR(m_i) \rangle$ asymptotically approaches 1 when $m_i$ gets bigger, we only need to make corrections up to $m_i = 25$ and thus the summation range of $k$ from 0 to 100 is fairly enough to provide a reliable accuracy and we use $\langle LLR(m_i) \rangle = 1$ when $m_i > 25$ which is sufficient since the asymptotical behavior of the $\langle LLR(x) \rangle$
Figure 4.12: Histograms of LLR values obtained 5 different configuration with different sub-regions sizes and its corresponding $\chi^2$ distribution with degrees of freedom predicted using Lidke-Huang hypothesis. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. LLR is then obtained using the method described this section, and plotted as histogram for each configuration. $\chi^2$ distribution with degrees of freedom predicted by LH hypothesis are also shown for each configuration.

Fig. 4.12 demonstrated, using LH hypothesis and numerically evaluated $\langle LLR \rangle$ provided a reliable $\chi^2$ distribution to describe the data set with various number of low count bins while original LLR theory breaks. In all 5 configuration, from sub-regions size of $4 \times 4$ to sub-region size of $20 \times 20$, $\chi^2$ provided by LH hypothesis gives a nice fit for the actually LLR distribution. When degrees of freedom of $\chi^2$ is
non-integer, we use corresponding Γ distribution to describe the statistical behavior.

4.6.6 Proposed Correction Method III - LH Approximation

Although, LH hypothesis provided a good description of LLR distribution, the evaluation of \(<LLR(m_i)>\) for data bin remain a tedious process since a summation from 0 to 100 has to be calculated as well as a calculation of \(k!\) in every loop. In case of super resolution analysis, thousands of frames are collected and for each frame, hundreds of sub-region are isolated and each of the sub-region would be then fitted and pass through the rejection algorithm. As a result performing numerical calculation for each sub-region individually by looping \(k\) from 0 to 100 will cause in a major decrease in analysis speed and would possible result in a huge increasing on total analysis time. To reduce the computational complexity, we proposed our "Lidke-Huang Approximation" (LHA):

\[
LHA(m) = \frac{1}{(1 - ce^{ab})^{2d} - 1} \left[ (1 - ce^{-a(m-b)})^{2d} + 1 \right] + 1 \quad (4.33)
\]

where \(a=0.7188\), \(b=0.9795\), \(c=1.36\) and \(d=1.84\).

\(LHA(m)\) is an empirical function that could be used as an approximation to the \(\langle LLR(m)\rangle\) value at various \(m\) especially for small \(m\) values where the original theory breaks.

As shown below are the result from LH approximation overlayed with the numerical calculation result of \(\langle LLR(m)\rangle\) for \(m\) ranging from 0.001 to 25.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.13: Plot of numerically calculated $\langle LLR(m) \rangle$ and $LHA(m)$. LHA provided a functional form $LHA(m)$ for calculating $\langle LLR(m) \rangle$ at various $m$ values.

Given that LH approximation could provide us a reliable approximation for $\langle LLR \rangle$, we could then use the following formula to calculate the degrees of freedom for the LLR in both large and small $m$ cases. We should notice that using Eq. LHAfitplot shown below would work with all ranges of $m$ in predicting the degrees of freedom and it converges to K when $\lambda \to \infty$ in all bins of data set.
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

Figure 4.14: Histograms of LLR values obtained 5 different configuration with different sub-regions sizes and its corresponding $\chi^2$ distribution with degrees of freedom predicted using Lidke-Huang hypothesis and Lidke-Huang approximation. Series of sub-regions were generated in 5 different configurations of sub-region sizes such as $4 \times 4$, $8 \times 8$, $12 \times 12$, $16 \times 16$ and $20 \times 20$. Each sub-region contain only one emitter with $\sigma_{PSF} = 1.2$ pixels and background=0.01. LLR is then obtained using the method described in section 4.6.6, and plotted as histogram for each configuration. $\chi^2$ distribution with degrees of freedom predicted by LH hypothesis calculated by LH approximation are also shown for each configuration.

As shown in Fig. 4.14, using LH approximation provided a easy way to calculate $\langle LLR \rangle$ proposed in LH hypothesis. In each configuration, distribution of LLR follows the predicted $\chi^2$ distribution calculated using LH approximation. However, we notice that for sub-regions with larger sizes, the peak of LLR distribution isn’t overlapping
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

perfectly with the $\chi^2$ from LHA. This is caused by the LHA approximation although being a very close approximation for $\langle LLR \rangle$ as shown in figure 4.13, still have non-zero error especially when $m > 4$. Although the maximum error for each estimates are around the level of 0.2 and decreases with increasing $m$, in the case of $20 \times 20$ sub-regions with large number pixels with counts around 10, the error effect is magnified such that the inconsistency between LHA and numeric calculation are summed across all pixel within the sub-region and introduced a shift of $\chi^2$ distribution which is predicted by LHA from the true distribution.

4.6.7 Discussion

Since the $\chi^2_K$ distribution of LLR statistics breaks down when expected value approaches 0, using LLR as the rejection algorithm faces obstacle when sub-region contains near 0 count pixels. We have solved this problem by 3 method, adding background, Lidke-Huang hypothesis with numerical calculation and Lidke-Huang hypothesis with Lidke-Huang approximation. Any of these proposed method provided a reliable way to predict the distribution for the LLR and thus make LLR reliable in all possible ranges of SM-SR data sets.

4.7 Conclusion

In this chapter, we first introduced the rejection algorithm used in SM-SR analysis and demonstrated the importance of its role in obtaining unbiased and reliable position estimates for target emitters. We reviewed two rejection algorithm that are commonly used in literature, elliptical test and SSE test and demonstrated their acceptance and rejection process together with their limitations. We then proposed our LLR test which complies to our criteria for a good test statistics as being a consistent,
Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis

statistically explainable and robust test statistics to be used in rejection algorithm. Demonstration are then shown to illustrate the acceptance and rejection behavior and then LLR are compared with the other two rejection algorithms in order to show its advantages. In general, LLR as the test statistic for rejection algorithm provides a robust and statistically reasonable rejection behavior whose properties could also be included and analyzed quantitatively in further analysis, such as clustering analysis. As the noise model and fitting parameters are all included in the calculation, LLR essentially provides a measure of goodness of fit for both position estimates and the assumed noise model.

When sub-region have near 0 count pixels, discrepancy raises between the proposed $\chi^2$ distribution and actually distribution of LLR. In order to correctly predict the distribution of LLR for data sets that have low count bins, we proposed Lidke-Huang hypothesis which provide a way to obtain the correct distribution of LLR in such case. Lidke-Huang approximation is proposed as a method to reduce the computational complexity introduced by using LHH. In the same time, We have provided 3 methods to predict the LLR distribution in presence of near 0 pixels within the sub-region: adding background, Lidke-Huang hypothesis with numerical calculation, Lidke-Huang hypothesis with Lidke-Huang approximation. We have demonstrated using simulation that all these 3 methods are proved to be reliable in obtaining the correct LLR distribution which is used in rejection algorithm.

References


Chapter 4. Rejection Algorithms for Single Emitter Fitting in SR Analysis


Chapter 5

Conclusion

We have implemented Fluorescence Correlation Spectroscopy in our home-built confocal setup and performed F/N analysis using FCS on live cells in order to evaluate the hypothesis of dimerization of ERK1 protein during nucleus translocation. We have provided evidence toward the conclusion that ERK1 nucleus translocation is dimerization independent.

We have developed a multi-emitter analysis for single-molecule super-resolution techniques that is designed to localize multiple overlapping emitters within a small sub-region. Several methods have been implemented in the algorithm which are essential for fitting convergence. To solve the singularity problem in the Fisher information matrix when using Gaussian mixture models, we have proposed a phenomenological fix by using a modified Fisher information matrix in order to provide reliable uncertainty estimates. We demonstrated the advantages of multi-emitter fitting analysis in both simulated and experimental data sets. We demonstrated that by implementing most of the analysis procedure in GPU using Compute Unified Device Architecture (CUDA, Nvidia) technology, the overall time consumption of the multi-emitter fitting analysis is on the scales of minutes.
Chapter 5. Conclusion

We have proposed the log-likelihood ratio as a test statistic that could be used in a reliable, robust and consistent rejection algorithm. We have demonstrated both its improved acceptance and rejection behavior as compared to elliptical test and SSE test that are commonly used in the literature. However, when sub-regions contain pixels with low counts, LLR diverges from the original proposed $\chi^2$ distribution. We proposed a hypothesis which predicts the true distribution of LLR at low expectation values and also provided an empirical function that is used to obtain the true distribution while avoiding intensive numerical calculations.
Appendices

A  FCS Alignment Protocol

B  Measurements of FCS sample volume waist

C  Tips on FCS Measurement

D  Calculation of Newton – Raphson Iteration Formula

E  Derivatives used in Iterative Formula and Fisher Information Matrix

F  Derivation of Fisher Information Matrix for Multi-Emitter Model

G  List of Software Recommended to Compile CUDA-MEX Files

H  Included Publications
Appendix A

FCS Alignment Protocol
-implemented with Dr. James Thomas

1. Find the Back Focal Plane. Remove the objective, set the Edmund grating on the nosepiece. By using the stage insert with a large hole, shine a HeNe back through the optics. With the side port (sp) lens in position, find where the grating focuses (i.e. where the diffracted orders cross.) Position the dichroic mirror here.

2. Next: rear-project to find the optical axis of the microscope. For this step, any objective lens may be used.

- A pinhole is centered in the field of view. ("FCS pinhole" recommended)
- The threaded iris is placed on the side port.
- The iris is centered with the centering screws.
- The iris is removed and the side port (sp) lens is threaded on.
Appendix A. FCS Alignment Protocol - implemented with Dr. James Thomas

- The focused image of the pinhole is marked with a second iris (field iris).

*For this last step, it may be useful to illuminate the pinhole with a HeNe to get more intensity. It may also be useful to use a sample with features (letters, numbers) to determine the best focus more easily.*

3. Remove the sp lens and the objective lens. Align the argon laser through the pinhole (field) iris, and in the back focal plane. Use the iris with the attached plastic sheeting to center in the back focal plane of the objective lens.

Use the downstream mirror to center the back focal plane, and the upstream mirror to center in the pinhole (field) iris. The beam should automatically center through the sp iris.

4. Remove the sp iris and screw in the sp lens. Add the f=250 mm focusing lens about 10” (250mm) upstream of the pinhole (field) iris. Adjust the position of this lens to keep the laser centered in the back focal plane.

5. Add the objective lens. Using ND 4, image the laser spot off a scratched mirror or off the Edmund grating. It should be close to the center of the crosshairs. You may use the dichroic adjustments to do some fine positioning of the beam right on the crosshairs.

6. Remove the ND, coarsely align the reflected laser spot onto the APD. Do fine alignment by maximizing count rate using the Correlator Program.
Appendix B

Measurement of FCS Effective Sample Volume Waist

Beam waist measurement for Gaussian beam. Before we take measurement, it’s better to have a knowledge base for Gaussian beam if we don’t. The following website might be helpful.

http://en.wikipedia.org/wiki/Gaussian_beam

http://www.mellesgriot.com/products/optics/gb_1.htm

To measure radial waist of a Gaussian beam, a grating mirror (in FCS drawer) is used for its internal calibration by the width of its stripes. Usually because of the limited distance which piazo is able to move, we use the narrowest 200LPMM part.

Here’s a list of four methods to achieve the goal of beam waist measurement on x-y direction. Two of them involve Fourier Transform of two convoluted signal and the other two is quite fundamental which might be better choice when the accuracy is less important.
Appendix B. Measurement of FCS Effective Sample Volume Waist

B.1 CCD Camera method

Place the grating mirror on the sample plane and focus the laser on to the mirror. Move it until you can clearly see the entire spot without being blocked by any reflection metal stripes on the grating. Record the width information (e.g. 200LPMM) where you put your focus on. Take a image with the CCD camera. The width can be calculated by counting pixels of the focus spot in the TIF picture.

B.2 Differentiation method

Mount grating mirror onto translation stage and use our piezo to precisely control the translation. (We have this already built up in B10W.) Use a slow frequency function generator (0.1Hz is preferred to generate more data per circle) and amplifier to deliver the piezo a voltage from. As observed by myself, the piezo can move within 5um steadily with a voltage rage 0Vs- 55Vs. After placing the piezo stage, function generator and the amplifier in place and all connected, start the function generator and do a little alignment so that the focus spot can move between an area of Dark-Bright-Dark shown as Figure 1. Use APD to track the reflected focus intensity, and save it as a *.trc file using Flex-02 software provided with the autocorrelator.

Now we have a trace looks like a square wave but the edge is not as sharp. This is because the laser focus is not an infinitesimal point but a Gaussian spot which is probably 0.5-1.0m. Select one of a perfect edge from the square wave and differentiate them in Igor. The Gaussian like curve after differentiation can be fit to a Gaussian and then we can derive the beam waist from the we got from fitting. Then we can calculate the real beam size by using the width of the grating stripe to determine the size of 1 point in the Igor fitting.
Appendix B. Measurement of FCS Effective Sample Volume Waist

Figure B.1: Focused laser spot on grating pattern

B.3 Fourier transform method 1

As mentioned in section B.2, we can obtain a intensity trace of the focus spot moving between the stripes of the grating glass and get a curve like a square-wave. The squarewave like curve is a result of convolution between a Gaussian function and a square-wave function. One can extract the perfect square-wave by applying a threshold on the recorded intensity trace and the MATLAB code for generate perfect square-wave from signal is also provided as supplementary in this protocol.

We know that from the properties of Fourier transform that

\[ F[f(x) \ast g(x)] = F[f(x)] \cdot F[g(x)] \]  \hspace{1cm} (B.1)

where \( F[f(x)] \) stands for the fourier transform of signal \( f(x) \), \( \ast \) stands for the convolution operation between 2 signals. Given that our detected pseudo square-wave (psw(x)) is the result of convolution between a true square-wave (sw(x)) and
Appendix B. Measurement of FCS Effective Sample Volume Waist

the Gaussian beam \((g(x))\), where \(psw(x) = sw(x) \otimes g(x)\), we could then use Eq.B.1 and solve for our Gaussian beam function \(g(x)\),

\[
g(x) = F^{-1} \left[ \frac{F[psw(x)]}{F[sw(x)]} \right]
\]  

(B.2)

where \(F^{-1}\) stands for inverse fourier transform operation. Eq.B.2 provides us a deconvolution method in order to obtain the gaussian beam profile.

B.4 Fourier transform method 2

Since the Fourier transform of a Gaussian function is still a Gaussian but with an inverse variance, the ratio fourier transforms between \(psw(x)\) and \(sw(x)\) can be directly fit to a Gaussian. Then the real variance of the Gaussian beam profile can be calculated by take the inverse of the variance result obtained from fitting process.
Appendix C

Important tips on FCS Measurement

Here I summarized items that require attention when performing FCS experiment on live cells in the FCS setup in B10W, PandA, University of New Mexico. When we perform the cell experiment, it is suggested to try to pick up the cells which are dim but not broken. A few test experiments should be carried out before taking the real data.

1. make sure that APD is still aligned by adjusting the X,Y knob of APD translator to maximize the count rate in the data collection program.

2. Use 1/3 - 1/4 drop of oil and try to make the oil cover only flat part on top of the objective lens.

3. Try to make average count rate around 20K-100K per second. This will give the best result of FCS on GFPs in our setup. (ND=2, Noisy eater is set to auto, 5.0)
Appendix C. Important tips on FCS Measurement

4. Wait for 20-30 mins for the sample (oil or slides) to be heated up before proceed cell experiment.

5. To avoid photobleach on the cell, use mercury lamp as few as possible. Illuminate the sample only when you try to pick up the cells and shut it down right after the cell is founded.

6. Turn the APD on 15s before click ”start” button in data collection software.

7. Clear the objective lens every time when change into another well for experiment.

8. Temperature vibration could be critical, AC should be shut down 1 hour before cell experiment.

9. 488 nm laser should be turned on at least 1 hour before experiment.

10. To reduce air flow which would make serious effects on FCS data, avoid any large movements within the darkroom during acquisition period.
Appendix D

Calculation of *Newton – Raphson* Iterative Update Formula

Maximum Likelihood Estimator (MLE) in chapter 3 is achieved by maximizing the likelihood function:

\[
L(\theta|D) = \prod_k L(u_k|d_k) = \prod_k \mu_k(x,y)^{d_k} e^{-\mu_k(x,y)} \frac{d_k}{d_k!} \tag{D.1}
\]

Instead of maximizing the likelihood function described in Eq. D.1, we maximized the log-likelihood function which is equivalent to directly maximize the likelihood function. Loglikelihood function are given by taking logarithmic on both side of Eq. D.1, which is given by,

\[
l(\theta|D) = \sum_k [d_k \ln \mu_k - \mu_k - d_k \ln(d_k) + d_k]
\]
Appendix D. Calculation of Newton – Raphson Iterative Update Formula

where $l(\theta|D)$ is the loglikelihood for $\theta$ given data set $\{D\}$.

Since at the maximum point of loglikelihood function, the derivative of it must be zero. Newton – Raphson method is used to iteratively update the parameter in order to find this stationary point. Iteration of Newton – Raphson method are given below for function $f(x)$, root $x_0$ could be found by an iteration process by the following equation

$$x^{n+1} = x^n - \frac{f(x^n)}{f'(x^n)}.$$

(D.2)

Apply Newton – Raphson iteration process on $l'(\theta|D)$ in order to find the stationary point, we would then obtain,

$$\theta_i^{n+1} = \theta_i^n - \frac{l'(\theta_i^n)}{l''(\theta_i^n)}$$

(D.3)

where $\theta_i^{n+1}$ stands the value of parameter $\theta_i$ at $n+1$ iteration step, and $l'(\theta_i^n)$ and $l''(\theta_i^n)$ stands for the first derivative and second derivative respectively. $l'(\theta_i^n)$ is obtained by taking the derivative of loglikelihood function $l(\theta|D)$ where we obtain,

$$l'(\theta_i) = \sum_k \frac{\partial \mu_k(\theta_i)}{\partial \theta_i} \left( \frac{d_k}{\mu_k(\theta_i)} - 1 \right)$$

(D.4)

and by taking the derivative of the above equation, we obtain the second derivative form for loglikelihood function as

$$l''(\theta_i) = \sum_k \left[ \frac{\partial^2 \mu_k(\theta_i)}{\partial \theta_i^2} \left( \frac{d_k}{\mu_k(\theta_i)} - 1 \right) - \frac{\partial \mu_k(\theta_i)}{\partial \theta_i} \frac{d_k}{\mu_k(\theta_i)^2} \right]$$

(D.5)
Appendix D. Calculation of Newton – Raphson Iterative Update Formula

We then combine Eq. D.3, D.4 and D.5 together and we obtain Newton – Raphson parameter update equation for $\theta_i$ at any step $n$ as

$$\theta_i^{n+1} \rightarrow \theta_i^n - \left[ \sum_k \frac{\partial \mu_k(\theta_i^n)}{\partial \theta_i} \left( \frac{d_k}{\mu_k(\theta_i^n)} - 1 \right) \right]$$

$$\left[ \sum_k \frac{\partial^2 \mu_k(\theta_i^n)}{\partial \theta_i^2} \left( \frac{d_k}{\mu_k(\theta_i^n)} - 1 \right) - \frac{\partial \mu_k(\theta_i^n)}{\partial \theta_i} \frac{d_k}{\mu_k(\theta_i^n)^2} \right]^{-1}$$

(D.6)

This finished the proof of Eq. D.6 in Chapter 2.

In order to calculate the update step for each parameter to be estimated such as position estimates $(x,y)$, intensity $(I)$ and background $(b)$, derivatives which is calculated in Appendix E.
Appendix E

Derivatives of Multi-Emitter Model

Given multi emitter model as

\[ \mu_k(x, y) = \sum_{i=1}^{N} I_0 \Delta E_{x_i}(x, y) \Delta E_{y_i}(x, y) + b_0 \]  

(E.1)

where \( \mu_k(x, y) \) is the expected model value at pixel \( k \) locating at position \( (x,y) \), \( I_0 \) is the intensity of the fluorophore, \( b_0 \) stands for the estimated background in the model and \( \Delta E_{x_i}, \Delta E_{y_i} \) can be described as

\[ \Delta E_x(x, y) = \frac{1}{2} \left( \text{erf} \left( \frac{x - x_0 + \frac{1}{2}}{\sqrt{2} \sigma_0} \right) - \text{erf} \left( \frac{x - x_0 - \frac{1}{2}}{\sqrt{2} \sigma_0} \right) \right) \]  

(E.2a)

\[ \Delta E_y(x, y) = \frac{1}{2} \left( \text{erf} \left( \frac{y - y_0 + \frac{1}{2}}{\sqrt{2} \sigma_0} \right) - \text{erf} \left( \frac{y - y_0 - \frac{1}{2}}{\sqrt{2} \sigma_0} \right) \right) \]  

(E.2b)

where \( x_0 \) and \( y_0 \) is the parameter of the emitter center location which are to be estimated using Newton – Raphson iteration method.
Appendix E. Derivatives of Multi-Emitter Model

First derivatives of $\mu_k(\theta)$ with respect to each parameter $\theta_i$ could then be calculated as

\[
\frac{\partial \mu_k^N(\theta)}{\partial \theta_{x_i}} = \frac{\theta_{l_i}}{2\sigma\sqrt{2\pi}} (e^{-\frac{(x_k - \theta_{x_i} - \frac{1}{2})^2}{2\sigma^2}} - e^{-\frac{(x_k - \theta_{x_i} + \frac{1}{2})^2}{2\sigma^2}}) \Delta E_y(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \tag{E.3a}
\]
\[
\frac{\partial \mu_k^N(\theta)}{\partial \theta_{y_i}} = \frac{\theta_{l_i}}{2\sigma\sqrt{2\pi}} (e^{-\frac{(y_k - \theta_{y_i} - \frac{1}{2})^2}{2\sigma^2}} - e^{-\frac{(y_k - \theta_{y_i} + \frac{1}{2})^2}{2\sigma^2}}) \Delta E_x(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \tag{E.3b}
\]
\[
\frac{\partial \mu_k^N(\theta)}{\partial \theta_{l_i}} = \Delta E_x(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \Delta E_y(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \tag{E.3c}
\]
\[
\frac{\partial \mu_k^N(\theta)}{\partial \theta_b} = 1 \tag{E.3d}
\]

Thus 2nd derivatives could also be obtained by taking the derivative of Eq. E.3, as

\[
\frac{\partial^2 \mu_k(\theta)}{\partial \theta_{x_i}^2} = \frac{\theta_{l_i}}{2\sigma^3\sqrt{2\pi}} \left( (x_k - \theta_{x_i} - \frac{1}{2}) e^{-\frac{(x_k - \theta_{x_i} - \frac{1}{2})^2}{2\sigma^2}} - (x_k - \theta_{x_i} + \frac{1}{2}) e^{-\frac{(x_k - \theta_{x_i} + \frac{1}{2})^2}{2\sigma^2}} \right) \Delta E_y(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \tag{E.4a}
\]
\[
\frac{\partial^2 \mu_k(\theta)}{\partial \theta_{y_i}^2} = \frac{\theta_{l_i}}{2\sigma^3\sqrt{2\pi}} \left( (y_k - \theta_{y_i} - \frac{1}{2}) e^{-\frac{(y_k - \theta_{y_i} - \frac{1}{2})^2}{2\sigma^2}} - (y_k - \theta_{y_i} + \frac{1}{2}) e^{-\frac{(y_k - \theta_{y_i} + \frac{1}{2})^2}{2\sigma^2}} \right) \Delta E_x(\theta_{x_i}, \theta_{y_i}, \theta_{l_i}) \tag{E.4b}
\]
\[
\frac{\partial^2 \mu_k(\theta)}{\partial \theta_{l_i}^2} = 0 \tag{E.4c}
\]
\[
\frac{\partial^2 \mu_k(\theta)}{\partial \theta_b^2} = 0 \tag{E.4d}
\]
Appendix F

Fisher Information Matrix for For Poisson Distributed Data

The $i, j$ element of the Fisher information matrix is given by the expectation of the product of two log-likelihood:

$$I(\theta)_{i,j} = E\left[\frac{\partial \ln(L(\theta|D))}{\partial \theta_i} \frac{\partial \ln(L(\theta|D))}{\partial \theta_j}\right]$$  \hspace{1cm} (F.1)

where $\theta$ is the set of parameters being estimated $\theta = [\theta_1...\theta_N]$ and $L(\theta|D)$ is the likelihood of the data set $\{D\}$ given the model generated by $\theta$. Here the parameters are $\theta = [\theta_x, \theta_y, ..., \theta_{I_0}, \theta_{bg}]$. For a Poisson process the likelihood is given by

$$L(\theta|D) = \prod_k L(u_k|d_k)$$

$$= \prod_k \frac{\mu_k(x,y)^{d_k} e^{-\mu_k(x,y)}}{d_k!}$$  \hspace{1cm} (F.2)

Using Eq. F.1, Eq. F.2 and the Stirling approximation ($\ln n! \approx n \ln n - n$ for large $n$), we could obtain
Appendix F. Fisher Information Matrix for Poisson Distributed Data

\[ I(\theta)_{i,j} = E\left[ \sum_k (d_k - \mu_k(x, y))^2 \frac{1}{\mu_k(x, y)^2} \frac{\partial \mu_k(x, y)}{\partial \theta_i} \frac{\partial \mu_k(x, y)}{\partial \theta_j} \right] \] (F.3)

Using the fact that \( E[(d_k - \mu_k(x, y))^2] \) is the variance and equal to the expected value for a Poisson process,

\[ I(\theta)_{i,j} = \sum_k \frac{1}{\mu_k(x, y)} \frac{\partial \mu_k(x, y)}{\partial \theta_i} \frac{\partial \mu_k(x, y)}{\partial \theta_j} \] (F.4)
Appendix G

List of Software Recommended to Compile CUDA-MEX Files

1. CUDA 4.0 package:
   - Graphic card driver
   - Toolkit
   - SDK
   - Samples

2. Visual Studio 2008 with Service Pack 1, including:
   - Microsoft Visual Studio 2008
   - Microsoft Developer Network
   - Microsoft SQL Server 2005
   - Microsoft SDK v6.0A

3. CUDA VS Wizard
Appendix G. List of Software Recommended to Compile CUDA-MEX Files

4. Parallel Nsight 2.0
Appendix H

Included Publication

