# Super-Resolution Image Reconstruction for Single-Molecule Localization Microscopy

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Jiaqing Huang, B.E.

Graduate Program in Department of Electrical and Computer Engineering

The Ohio State University

2016

Dissertation Committee:

Dr. Yuejie Chi, Advisor

Dr. Jianjie Ma, Co-Advisor

Dr. Philip Schniter

Dr. Mingzhai Sun

© Copyright by

Jiaqing Huang

2016

# Abstract

Single-molecule localization microscopy achieves sub-diffraction-limit resolution by localizing a sparse subset of stochastically-activated emitters in each frame. The temporal resolution, determined by the number of frames to acquire a super-resolution image, is limited by the maximal emitter density that can be handled by image reconstruction algorithms. Hence to enhance the temporal resolution, we develop signal processing algorithms for super-resolution image reconstruction.

In two-dimensional (2D) super-resolution image reconstruction, compressive sensing based reconstruction algorithms utilize the sparsity prior of the signal in each frame and can greatly increase the emitter density. We apply spectrum estimation methods to achieve comparable performance with a much lower computational cost. By transforming the images into the frequency domain, the localization of emitters is equivalent to estimation of 2D frequencies. The Matrix Enhancement and Matrix Pencil (MEMP) algorithm is applied to extract the emitter locations, resulting in fast computation and comparable performance with compressive sensing based methods.

In three-dimensional (3D) super-resolution image reconstruction, the point spread function (PSF) is engineered to encode the axial information, e.g., by introducing a cylindrical lens. However, the large effective volume of 3D PSF increases the possibility of overlapping emitters in each frame and makes image reconstruction challenging. To address this problem, we developed a 3D high-density super-resolution imaging

system which includes a multi-focus optical platform with astigmatism and a computational framework using compressive sensing. A de-biasing step is introduced to reduce the intrinsic bias caused by the discretization in compressive sensing. This framework improves performance over a single-focus platform due to better differentiability along the axial direction.

However, the discrete formulation of compressive sensing for super-resolution image reconstruction suffers from high computational complexity and gridding error due to model mismatch. We propose a novel super-resolution algorithm for both 2D and 3D image reconstruction, dubbed TVSTORM, which promotes the sparsity of activated emitters without discretizing their locations. Several strategies are pursued to improve the reconstruction quality under the Poisson noise model, where the computational time is reduced by an order-of-magnitude. Numerical results on both simulated and cell imaging data are provided to validate the favorable performance of TVSTORM.

Dedicated to Wenjia and to my parents.

# Acknowledgments

First of all, I would like to thank my advisors Prof. Yuejie Chi and Prof. Jianjie Ma for their guidance, support, patience and encouragement over the past several years. I'm grateful for the great environment they provide for my research. I have benefitted a lot from Prof. Chi's feedback on my research, as well as her valuable advise and help. I sincerely appreciate Prof. Ma for the trust and freedom he gives me when I pursue and conduct research of my interest. I would also like to thank Dr. Mingzhai Sun whom I have been working with very closely. He has been a great mentor and friend to me and I have learned so much from him both in research and life experience. I want to thank Prof. Philip Schniter for being on my candidacy and dissertation committee.

I thank all my labmates in both Chi Lab and Ma Lab for their help. Especially, I want to thank Yuanxin Li, for inspiring and entertaining discussions, Kristyn Gumpper and Matthew Sermersheim, for devoting hundreds of hours of back-and-forth sample preparation that makes the real imaging in this dissertation possible. I thank all my friends in Columbus for making my PhD life here such a memorable one. I also want to thank the Ohio Supercomputing Center for the generous computational resources they provide for my research, without which some calculations in this dissertation could take months.

Lastly, I thank my parents for their unconditional love and support. I thank Wenjia for being with me during my PhD life. Words cannot describe how lucky I am to have her in my life.

# Vita

Jun, 2012 ..... B.E., Electrical Engineering,
Shandong University, China

2012-present ..... Graduate Research Associate,
The Ohio State University

# Publications

## Research Publications

M. Sun, J. Huang, F. Bunyak, K. Gumpper, G. De, M. Sermersheim, G. Liu, P. Lin, K. Palaniappan and J. Ma. "Superresolution Microscope Image Reconstruction by Spatiotemporal Object Decomposition and Association: Application in Resolving T-tubule Structure in Skeletal Muscle," *Optics Express*, vol. 22, no. 10, pp. 12160-12176, May. 2014.

J. Huang, M. Sun, K. Gumpper, Y. Chi and J. Ma. "3D Multifocus Astigmatism and Compressed Sensing (3D MACS) based Superresolution Reconstruction," *Biomedical Optics Express*, vol. 6, no. 3, pp. 902-917, Feb. 2015.

J. Huang, K. Gumpper, Y. Chi, M. Sun and J. Ma. "Fast Two-Dimensional Super-Resolution Image Reconstruction Algorithm for Ultra-High Emitter Density," *Optics Letters*, vol. 40, no. 13, pp. 2989-2992, Jun. 2015.

J. Huang, M. Sun and Y. Chi. "Super-Resolution Image Reconstruction for High-Density 3D Single-Molecule Microscopy," *International Symposium on Biomedical Imaging*, (Prague, Czech Republic), Apr. 2016.

J. Huang, M. Sun, J. Ma and Y. Chi. "Super-Resolution Image Reconstruction for High-Density 3D Single-Molecule Microscopy," *submitted to IEEE Trans. on Computational Imaging*, Jan. 2016.

## Fields of Study

Major Field: Electrical and Computer Engineering

# Table of Contents

**Page**

# List of Tables

**Table** **Page**

# List of Figures

# Chapter 1: Introduction

Fluorescence microscopy has been widely used in molecular and cell biological research and proven to be a powerful tool for a variety of studies. It allows direct observations of dynamic cellular processes owing to its minimum invasiveness and high biochemical specificity. However, due to optical diffraction, the resolution of conventional fluorescence light microscopes is limited to ∼200 nm in the lateral direction and ∼500 nm in the axial direction, preventing them from imaging many subcellular structures.

In recent years, a number of super-resolution microscopy techniques have been developed to break the diffraction limit and improve the lateral resolution to ∼20 nm and the axial resolution to ∼50 nm. These super resolution microscopy techniques include saturated structured illumination microscopy (SSIM) [1], stimulated emission depletion (STED) microscopy [2], and single-molecule localization microscopy (SMLM) such as photoactivated localization microscopy (PALM) [3] and stochastic optical reconstruction microscopy (STORM) [4]. These super resolution microscopy techniques, although are still under active development, will revolutionize the biological imaging field as they provide the unprecedented optical resolution smaller than the wavelength of light, thus enabling the observations of previously unresolvable subcellular structures and biological processes. The Nobel Prize in Chemistry was awarded

jointly to Eric Betzig, William Moerner and Stefan Hell in 2014 for "the development of super-resolved fluorescence microscopy", which brings "optical microscopy into the nanodimension" [5]. Among these techniques, SMLM utilizes the switching dynamics of fluorophores and achieves sub-diffraction-limit spatial resolution by localizing a sparse subset of stochastically activated emitters in each frame. This process is repeated for multiple frames, with each frame having a stochastically different subset of active fluorophores, and the final super-resolution image is reconstructed by aggregating the localized fluorophore positions. Theoretically, the spatial resolution of SMLM is only limited by the localization accuracy of the fluorophores.

## 1.1 Biological Applications of SMLM

SMLM has been used to image numerous cellular structures and provide the unprecedented resolution that allows observations of fine structures beyond the diffraction limit. Here, we list a few examples to demonstrate the application of SMLM in the biological field.

In skeletal muscles, the transverse (T)-tubular invagination of the sarcolemmal membrane touches the terminal cisternae of the junctional sarcoplasmic reticulum to form a triad junction structure that provides the structural framework for excitation-contraction (EC) coupling [6]. Most of the ultrastructural analyses of the triad junction involve electron microscopy [6]. In [7], we use the antibody mitsugumin29 (MG29), a skeletal muscle-specific protein, to label the T-tubule structure in the rat FDB fiber and image the MG29-labelled T-tubule network using SMLM in combination with a spatio-temporal object decomposition and association algorithm. As shown in Figure 1.1 (a), the traditional epifluorescence microscopy produces striated

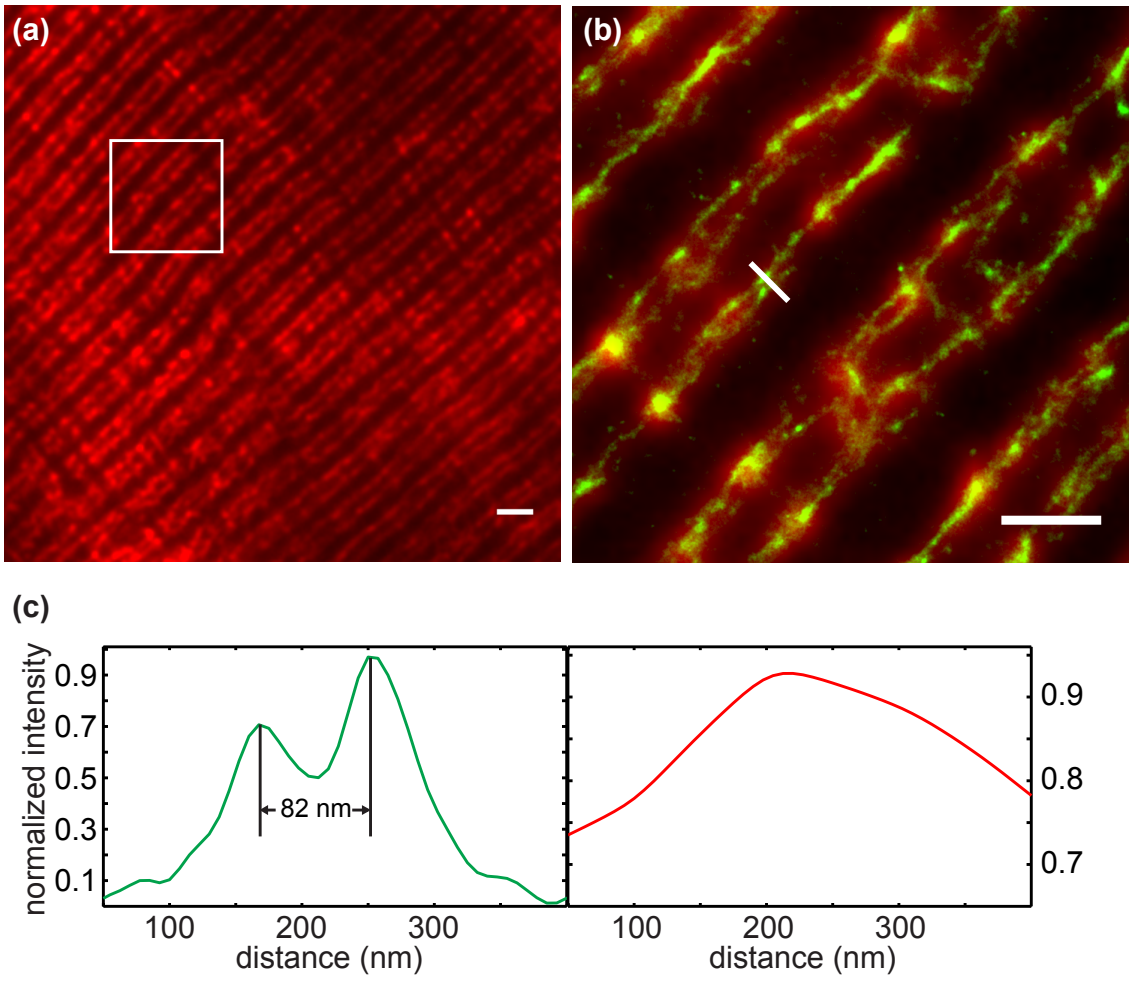



Figure 1.1: SMLM imaging of T-tubule structures in rat FDB fiber. (a) A conventional fluorescence image of the MG29-labelled T-tubule network in rat skeletal muscle. Scale bar: 2 $\mu$m. (b) Overlay of the epifluorescence image (red) of the T-tubule and the corresponding SMLM image (green). Scale bar: 1 $\mu$m. (c). The line profiles of the selected region in (b) with SMLM (green) and epifluorescence image (red).

pattern of MG29-labelled T-tubule network, with a doublet pattern that resembles the typical membrane structure in rat skeletal muscle. Within each of the doublets, there are two T-tubule membranes running in parallel that could not be resolved in traditional optical imaging. With SMLM, we are able to reconstruct the images, clearly showing the two sides of the T-tubule membrane within a single doublet. Figure 1.1 (b) shows the overlay of the SMLM image (green) and the conventional epifluorescence image (red). Figure 1.1 (c) shows the line profiles of the selected region, which is the intensity distribution along the selected line. The green curve corresponds to the line profile from the SMLM image, which shows clearly the distance between the two T-tubule membranes is about 82 nm. The conventional epifluorescence image completely misses the fine structure (red curve). To our knowledge, this is the first time that the T-tubule structure in skeletal muscle cells is resolved optically.

In another study [8], Huang et al. investigate the interactions between mitochondria and microtubules through multicolor three-dimensional (3D) SMLM imaging. The mitochondria are double membrane-bound organelles found in most cells that are responsible for biochemical processes of respiration and energy production. The microtubules are microscopic tubular structures in cytoplasm. The constant transportation and reorganization of mitochondria are driven by motor proteins on cytoskeleton, of which the microtubules are components. Their spatial relations, therefore, are of great interest to biologists who want to study the movement of mitochondria. In [8], two-color SMLM imaging of mitochondria and microtubules in mammalian cells are performed and the comparison with conventional fluorescence image is shown in Figure 1.2 (a) and (b). The SMLM image of mitochondria and microtubules clearly demonstrates greater resolution than conventional fluorescence

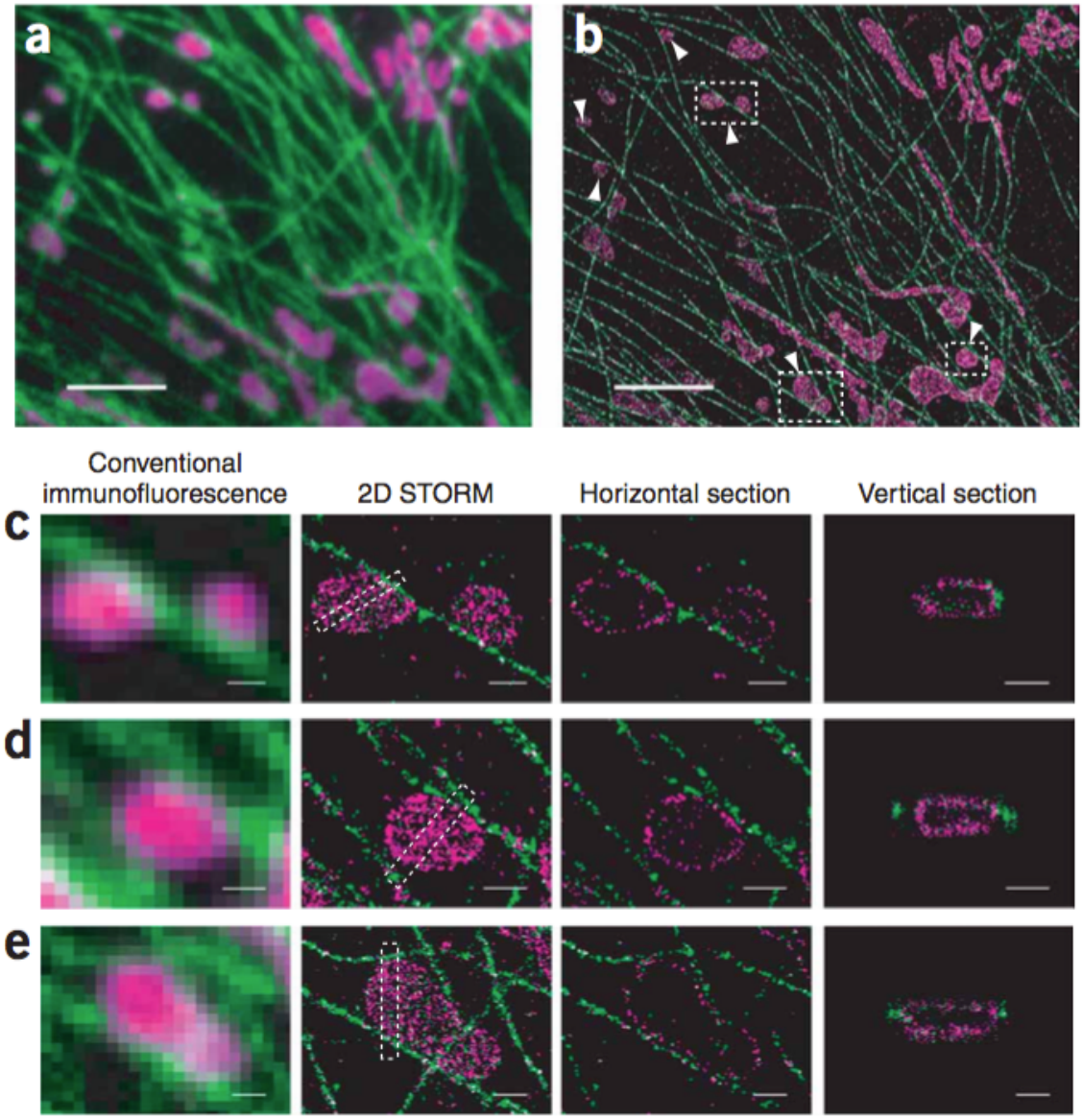



Figure 1.2: Two-color 3D SMLM images of mitochondria (magenta) and microtubules (green). (a) A conventional fluorescence image. (b) A SMLM image of the same area. (c-e) Magnified views of the boxed regions in (b). Adapted from [8].

image. Figure 1.2 (c-e) show the magnified views of the boxed regions in Figure 1.2 (b). In Figure 1.2 (c), the contact between mitochondria and microtubules are observed in both conventional fluorescence microscopy and SMLM. However, many false contacts between mitochondria and microtubules are also observed in the conventional fluorescence microscopy due to the diffraction limit as shown in Figure 1.2 (d). Mitochondria are also often observed to be surrounded by microtubules (Figure 1.2 (e)). The spatial relations between mitochondria and microtubules are further studied in [8] and a noncontiguous inchworm-like interaction mode is observed, which may facilitate the movement of mitochondria to maintain a dynamic morphology.

## 1.2 Motivation: Improving the Temporal Resolution of SMLM

The temporal resolution of SMLM, defined as the time required to generate a super-resolution image, becomes the limiting factor of this technique due to the iterative nature of the imaging procedure. To achieve a super-resolution image with a spatial resolution of 20 to 50 nm, the image acquisition can take tens of thousands of frames, requiring lengthy acquisition time, on the order of several minutes. This greatly limits the application of super-resolution techniques from imaging fast dynamic processes. For example, Mitsugumin 53 (MG53) is a muscle-specific protein which has been reported to facilitate the repair of membrane damage in muscle cells [9]. However, the movement of MG53 under cell damage is a fast process and happens on the order of seconds. Super-resolution techniques with higher temporal resolution are thus needed in order to resolve and study the movement and clustering process of MG53 under cell damage in both two-dimensional (2D) and 3D imaging.

To improve the temporal resolution of SMLM, two different strategies have been proposed. One is by controlling the photophysical properties of the fluorescence probes. In a seminal work by Jones et al. [10], a temporal resolution as high as 0.5 second was achieved in live cell imaging by using engineered bright emitters with a strong activation laser and a sensitive low light detector. However, a strong excitation laser can induce phototoxicity to cells. The other approach involves increasing the density of activated emitters in each frame so that the number of frames required to localize the same number of emitters can be significantly reduced. However, a high density of activated emitters causes the point spread functions (PSF) of adjacent emitters to overlap, which results in poor performance of the widely used single emitter localization methods [4]. In 3D super-resolution imaging where the 3D PSF is extended along the axial direction, the situation is further exacerbated since the PSFs of adjacent emitters have more overlapping regions and similarities.

## 1.3 Contributions and Outline of the Dissertation

In this dissertation, we aim to present fast and high-performance super-resolution image reconstruction algorithms that can analyze images of high emitter density reliably in both 2D and 3D cases, as well as new imaging modalities that can improve the super-resolution imaging performance. These innovations not only improve the temporal resolution and reconstruction quality of super-resolution microscopy, but also provide a better tool for biological researchers.

We start with 2D image reconstruction. The state-of-the-art reconstruction algorithm for SMLM is based on compressive sensing. However, it requires discretization

over a continuous 2D space and solving a large-scale optimization problem, thus suffering from high computational complexity. Such discretization also produces intrinsic gridding error in the reconstruction. To address this, we transform the imaging process of SMLM into the frequency domain. The problem of emitter localization then can be formulated as a 2D frequency estimation problem. We develop MempSTORM [11] based on 2D spectrum analysis which is able to achieve similar performance with compressive sensing based reconstruction but is over two orders of magnitude faster.

In 3D super-resolution imaging, due to the extended PSF shape, nearby emitters have more overlapping regions and similarities, increasing the difficulty to distinguish emitters along the axial direction. We propose a multi-focus astigmatism compressive sensing based 3D imaging platform which enhances the asymmetry of system PSF along the axial direction by combining astigmatic optics and multi-camera techniques. An algorithm based on compressive sensing is developed which reduces the bias introduced by gridding and increases the emitter density.

Motivated by the need to reduce the extreme long analysis time required in 3D image reconstruction due to discretization, we consider the notion of total-variation norm, which can be seen as a generalization of the $\ell_1$ norm for finite-dimensional vectors to the continuous space without imposing a discrete grid for the locations of the emitters. We propose TVSTORM, a penalized maximum likelihood estimator under the Poisson noise regularized by the total variation norm of the activated emitters. TVSTORM demonstrates an order-of-magnitude improvement on the computational cost over compressive sensing based reconstruction due to the elimination of optimizing over a fine-grained grid. It also shows a significant improvement on the

localization accuracy in terms of detection rate, false discovery rate and precision, without adding post-processing steps.

The outline of the dissertation is as follows. Chapter 2 reviews the background information of SMLM as well as other super-resolution imaging techniques. In Chapter 3, we describe the signal model of SMLM used across this dissertation as well as existing reconstruction algorithms based on single emitter localization and compressive sensing. Chapter 4 presents a fast 2D image reconstruction algorithm - MempSTORM (based on the publication [11]). In Chapter 5, we propose a multi-focus astigmatism and compressive sensing based 3D imaging platform (based on the publication [12]). Next, we present TVSTORM for both 2D and 3D image reconstruction in Chapter 6 (based on the publication [13] and [14]). Finally, we conclude in Chapter 7.

# Chapter 2: Background

Image contrast, the difference in the color or brightness, is essential for humans to distinguish objects. Scientists have been pursuing to increase the contrast of microscopic images for a long time by various means. These techniques include dark field microscopy, phase contrast microscopy [15], polarized light microscopy [16], differential interference contrast (DIC) microscopy [17], etc. However, the most popular and probably the best technique by far to create contrast is through fluorescence.

## 2.1 Fluorescence Excitation and Emission

Illuminated by excitation light of specific wavelength, fluorophores can emit light with lower energy (longer wavelength). This phenomenon, called Stokes shift, is named after Sir George Stokes who first discovered that fluorspar and uranium glass could convert invisible ultra-violet light into visible radiation that is of longer wavelength in 1852 [18]. In 1945, Aleksander Jabłoński formulated the physical and molecular mechanism [19] of fluorescence excitation and emission; the corresponding diagram showing the transitions of electronic states is therefore called Jabłoński diagram.

The principle of fluorescence excitation (absorption) and emission is shown in Figure 2.1. The upper part of Figure 2.1 is a Jabłoński diagram where $S_0$ is the ground

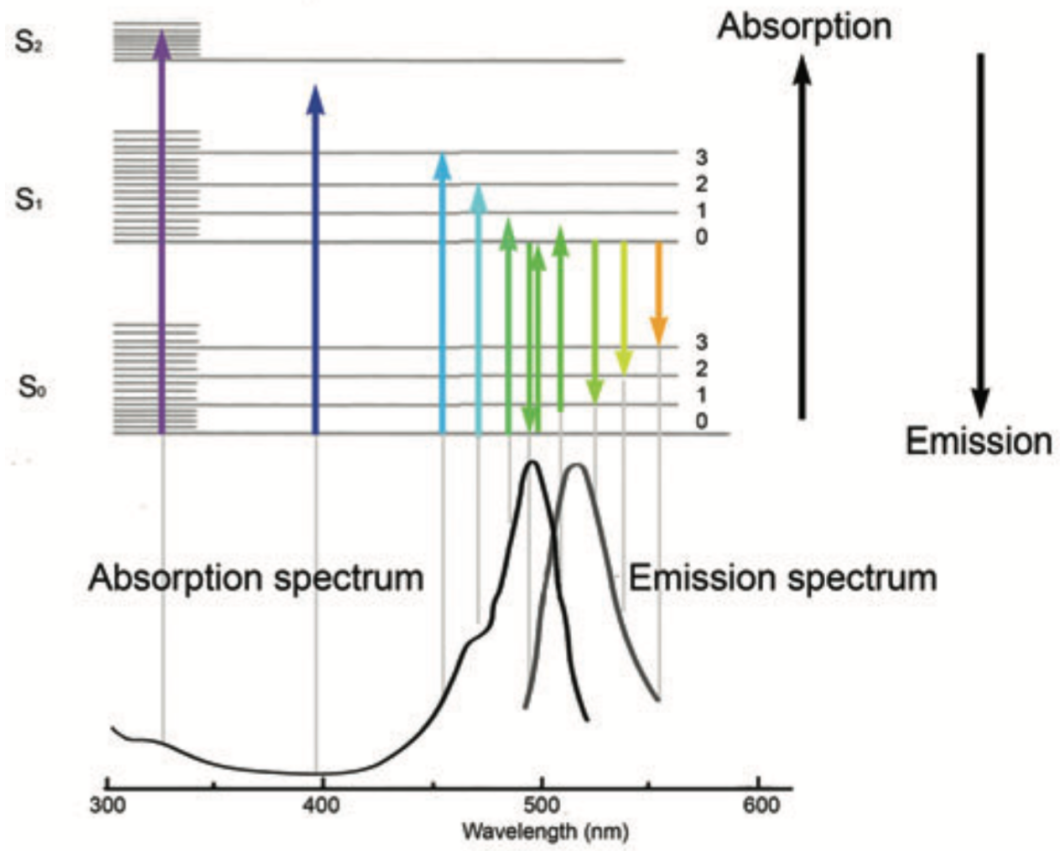



Figure 2.1: Spectral properties of fluorophore excitation and emission. Adapted from [20].

state and represents the energy of a fluorescent molecule without the excitation of light. $S_1$ and $S_2$ are excited states and they occur when an outer electron is boosted to a different orbital which is farther away from the nucleus. When a fluorophore absorbs the energy of a photon that is greater than the difference between $S_0$ and the lowest energy level of $S_1$, it will transfer from the ground state $S_0$ to an excited state of either $S_1$ or $S_2$, depending on the absorbed energy. Therefore, a range of spectrum can excite a fluorophore, as shown in the bottom half of Figure 2.1 as

the absorption spectrum. Due to the minimum energy required for excitation, the absorption spectrum typically has a sharp cutoff on the long-wavelength side.

There are several pathways that an excited fluorophore can return to the ground state. First, a process called internal conversion typically transfers the excited fluorophore to the lowest energy level of $S_1$. The energy difference is released through vibrational relaxation which does not involve the emission of photons. The internal conversion from the excited state to $S_0$ is also possible but for most fluophores this conversion is not preferred due to the large energy gap. From the lowest energy level of $S_1$, the typical energy pathway to the ground state is through the emission of a photon. Since $S_0$ has multiple energy levels, there is a range of the spectrum for the emitted photons, as shown in the bottom half of Figure 2.1 as the emission spectrum. The time for both excitation and emission are very short, with excitation time on the order of femtoseconds and emission time on the order of nanoseconds.

Due to the lost energy during the process of internal conversion and vibrational relaxation, the emission wavelength typically is higher than the excitation wavelength, creating a difference in the peak spectrums, i.e., the Stokes shift. The overlap of absorption spectrum and emission spectrum depends on the magnitude of the Stoke shift. A larger Stokes shift results in a smaller overlap and vice versa.

Another pathway for excited fluorophores to return to the ground state is by phosphorescence, which involves intersystem crossing between singlet states and triplet states. This process will not be described here and the interested readers may refer to [20] for details.

## 2.2 Fluorescence Microscopy

The use of fluorophores in the biological field was first initiated in the 1930s and has gained much popularity since then. Thousands of fluorescence probes have been created to label almost every biological structure. Furthermore, the development of genetic engineering allows genetic modification of proteins to carry a fluorescent reporter, bringing enormous growth to cell biology as well as optical microscopy. In 2008, the Nobel Prize in Chemistry was awarded jointly to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien "for the discovery and development of the green fluorescent protein, GFP" [21]. Because fluorophores are highly sensitive, specific and provide very high contrast, fluorescence microscopy has found numerous applications in the field of cell biology, including but not limited to, discovering patterns of certain proteins, tracking the movement of molecules, studying intracellular signaling, etc. We describe the principle of fluorescence microscopy in this section.

There are a few key components of fluorescence microscopy: light source of specific wavelength, objective, emission filter and detector. To begin with, samples are first stained with fluorophores and illuminated with light of certain wavelength. Then the Stokes-shifted emission light of fluorophores are separated from illumination light through the use of a spectral emission filter and collected by an advanced electronic light detection device such as a charge-coupled device (CCD) camera. Typically epifluorescence illumination is used where the illumination light and emission light share the objective and part of light path. The emission filter should filter out most of the illumination light while keeping as much energy from emission as possible. Contemporary light detectors enable the detection of even a single molecule's behavior such as blinking, providing extremely high sensitivity.

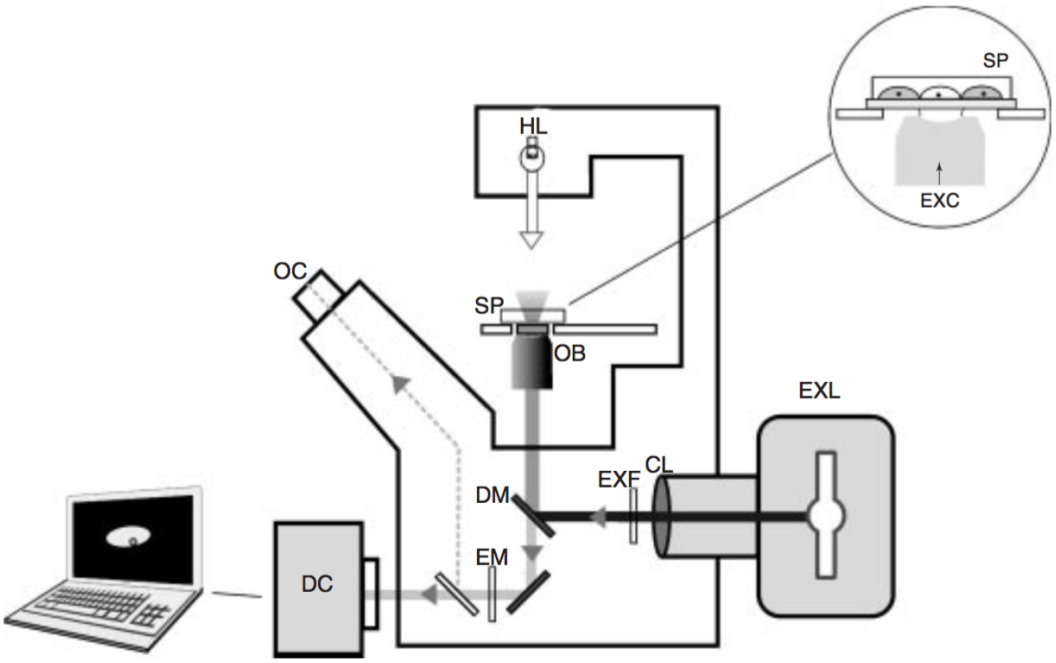



Figure 2.2: Schematic diagram of a typical inverted epifluorescence microscope. HL: halogen lamp; SP: specimen; OB: objective lens; OC: ocular (eyepiece); DC: digital camera; EXL: excitation light source; CL: collector lens; EXF: excitation filter; DM: dichroic mirror; EM: emission filter; EXC: exciting light incident on the specimen. Adapted from [22].

Figure 2.2 shows the schematic diagram of a typical inverted epifluorescence microscope, where the objective is below the sample and pointing upwards. Such configuration is useful for imaging live cells maintained in standard growth medium. The excitation light source, often in the region of short wavelength, first passes through excitation filter to allow only selected wavelength corresponding to the absorption spectrum of fluorophores. The filtered excitation light then reflects from a dichroic mirror, via the objective to irradiate the samples. After the sample fluoresces through

the process of excitation and emission, the emission light is gathered by the objective, passed through the dichroic mirror and filtered by the emission filter. A digital camera collects the filtered emission light and produces a fluorescence image.

## 2.3 Diffraction Limit

Due to optical diffraction, the image of an infinitesimal object, after any optical system, will become a finite-sized spot. The PSF is used to describe the response of a point source to an optical imaging system. Mathematically, let $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^2$ index the spatial coordinates in the object plane and the detector plane, respectively, $\chi(\boldsymbol{w})$ represent the light intensity emitted by the object, and $s(\boldsymbol{v})$ represent the light intensity received at the detector plane, then the input-output relation of a fluorescence microscope can be characterized as [23]:

$$s(\boldsymbol{v}) = \kappa \int f(\boldsymbol{v} - \boldsymbol{w})\chi(\boldsymbol{w})d\boldsymbol{w}, \tag{2.1}$$

where $\kappa$ is a normalization constant. Notice that (2.1) is a linear convolution of $\chi(\boldsymbol{w})$ with the low-frequency PSF $f(\boldsymbol{v})$ induced by diffraction [23]:

$$f(\boldsymbol{v}) = \left[2\frac{J_1(k_{\text{em}}\text{NA}\|\boldsymbol{v}\|_2)}{k_{\text{em}}\text{NA}\|\boldsymbol{v}\|_2}\right]^2, \tag{2.2}$$

where $k_{\text{em}} = 2\pi/\lambda$ is the emission wavenumber, $\lambda$ is the wavelength of emission light, NA is the numerical aperture of the objective lens and $J_1(\cdot)$ is the first-order Bessel function of the first kind. Numerical aperture is a dimension-less number that characterizes the range of angles that can be accepted or emitted by the optical system, defined as $\text{NA} = 2n \cdot \sin\alpha$, where $n$ is the refractive index of surrounding medium and $\alpha$ is the maximum half-angle under which cone of light can enter or exit

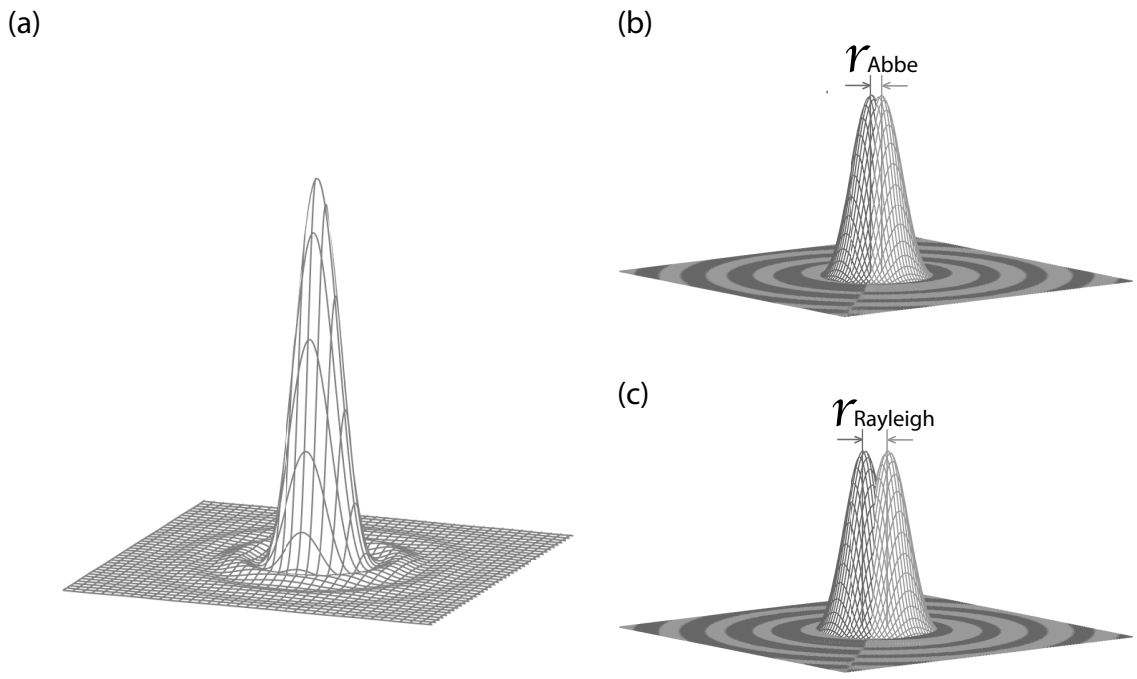



Figure 2.3: (a) Surface plot of diffraction-limited imaging system PSF (Airy disk). Minimum separation according to (b) Abbe criterion and (c) Rayleigh criterion.

the lens. The pattern of a diffraction-limited system PSF, characterized in (2.2), is also called Airy disk, the surface plot of which is shown in Figure 2.3 (a).

To characterize the resolution of a diffraction-limited system, two different criteria, i.e., the Abbe and Rayleigh criteria [24, 25], have been proposed. They both define the resolution as the minimum physical distance in space that two point sources need to be apart in order for them to be distinguished as individuals. The distinction between the two is their slightly different definitions of resolvability of two objects after diffraction. According to the Abbe criterion, the resolution is defined as

$$r_{\text{Abbe}} = \frac{\lambda}{2\text{NA}}, \tag{2.3}$$

while Rayleigh defines the resolution as

$$r_{\text{Rayleigh}} = \frac{0.61\lambda}{\text{NA}}. \tag{2.4}$$

The surface plots of the intensity distribution of two point sources placed with a separation of $r_{\text{Abbe}}$ and $r_{\text{Rayleigh}}$ are shown in Figure 2.3 (b) and (c).

For the axial direction, the width of PSF is 2-3 times the width in the lateral direction [26]. Therefore, with visible light and a high-end fluorescence microscope equipped with an oil immersion objective of NA $\approx 1.4$, the lateral resolution is $\sim 200$ nm and the axial resolution is $\sim 500$ nm. Such resolution is sufficient for resolving structures on the tissue or cell level where the desired object of interest is of the scale above 200 nm. However, for subcellular imaging where the scale of many structures of interest is well below 100 nm, traditional fluorescence microscopy is no longer a feasible technique. By using electrons instead of photons for imaging, electron microscopy (EM) achieves nanometer-scale or even higher resolution. However, it loses the labelling specificity and requires a very sophisticated sample preparation procedure which damages the sample permanently, prohibiting live cell or tissue imaging. Therefore, it is desirable to develop imaging techniques which still possess the advantage of fluorescence microscopy, i.e., minimum invasiveness, sensitivity, selectivity, specificity and high contrast, but has a higher resolution beyond the diffraction limit.

## 2.4 Sub-Wavelength Imaging Techniques

To improve the diffraction-limited resolution of conventional microscopy, a number of sub-wavelength imaging techniques have been proposed. They include total internal reflection fluorescence (TIRF) microscopy [27], confocal microscopy [28], multiphoton

microscopy [29], 4Pi microscopy [30], and structured illumination microscopy (SIM) [31]. Their mechanism and realization are summarized here:

- TIRF microscopy: When the incidence angle of excitation light is under the critical angle, the excitation light is totally internally reflected, creating a very thin layer of exponentially decaying wave (evanescent wave). This evanescent wave can excite only the fluorophores near the glass-water (or glass-oil) surface, enhancing the resolution in the axial direction. Usually a layer of less than 200 nm depth can be observed. This technique is especially useful for observing molecular events close to the cell surface.

- Confocal microscopy: By placing a spatial pinhole at the conjugate focal (confocal) plane of the sample, out-of-focus light excited by the sample is reduced, sharpening the system PSF. After a 2D plane is scanned, an optical slice of the sample is detected, enabling reconstruction of 3D structures by scanning the axial direction. Confocal microscopy has become one of the most widely used imaging techniques in the biological field.

- Multiphoton microscopy: Bu using more than one photon to excite a fluorophore, it not only minimizes the scattering in the tissue, but also suppresses the background signal. Multiphoton microscopy provides a much deeper penetration and reduces the effective size of excitation PSF, allowing optical sectioning in 3D imaging.

- 4Pi microscopy: As Abbe's diffraction limit (2.3) indicates, the PSF size of diffraction-limited imaging system decreases as the NA increases. By using two opposing objective lenses focused to the same spot, light emitted by the sample

can be collected coherently from both sides, leading to a theoretically maximal solid angle of $4\pi$. This effectively increases NA and substantially improves the axis resolution as well. A similar technique, $I^5M$ microscopy [32], which features a high-speed data acquisition has been proposed later. Both 4Pi and $I^5M$ microscopes achieve an axial resolution of $\sim$100 nm.

- SIM: A periodic high-frequency sinusoidal illumination pattern, created in the lateral direction, the axial direction, or both by interference, is used in SIM to modulate the spectral frequencies of the image. Information at higher frequency, which previously is unobservable, is shifted to lower frequency which can be observed by the microscope. Images modulated at different orientations are acquired and demodulated to its original high frequency to synthesize the high-resolution image. A theoretical resolution improvement by a factor of 2 can be achieved in both lateral and axial directions.

The above introduced techniques modify the optical imaging setup and achieve significant resolution improvement especially in the axial direction. However, their resolution is still limited by the diffraction of light. The best resolution achieved by these techniques is 100 nm in both lateral and axial directions as reported by $I^5S$ [33], a combination of $I^5M$ and SIM. Such resolution improvement, although very significant, is still not enough for nanometer-scale structures visualization.

## 2.5 Spatially Pattern Excited Super-Resolution Microscopy

In the past few years, a number of breakthroughs have been made in the far-field microscopy field to bypass the diffraction limit in both lateral and axial directions [1–4, 34]. These techniques, which modulate the excitation or emission nonlinearly

either in the spatial domain or in the temporal domain, can be divided into two categories: spatially pattern excited super-resolution microscopy and SMLM. In this section, we introduce spatially pattern excited super-resolution microscopy.

### 2.5.1 Saturated Structured Illumination Microscopy

As discussed earlier, SIM makes use of a spatially varying illumination to render high frequency information. However, the resolution limit induced by diffraction also applies to the patterned illumination, resulting in a mere factor of two in the improvement of highest detectable frequency if the sample responds linearly to the illumination intensity. Therefore, the key to further improve the resolution is to break the linearity between illumination and excitation. Such nonlinearity can be achieved when fluorophores are illuminated with a strong illumination and reach saturation. As shown in Figure 2.4 (a), the illumination light distribution follows a sinusoidal signal with a strong peak intensity. The excitation responds nonlinearly with the illumination and reaches to saturation level when illumination is above certain threshold. Such saturated excitation pattern can modulate the sample with higher order of frequencies which can be recovered with frequency analysis. Figure 2.4 (b) shows the generated pattern after regular and saturated excitation pattern.

In SSIM, the fundamental resolution limit no longer comes from the diffraction of light, rather, it is due to the saturation level of excitation. A resolution of 50 nm in the lateral direction has been demonstrated in [1]. In principle, it can also be applied to 3D imaging where the saturated excitation pattern is applied along the axial direction.

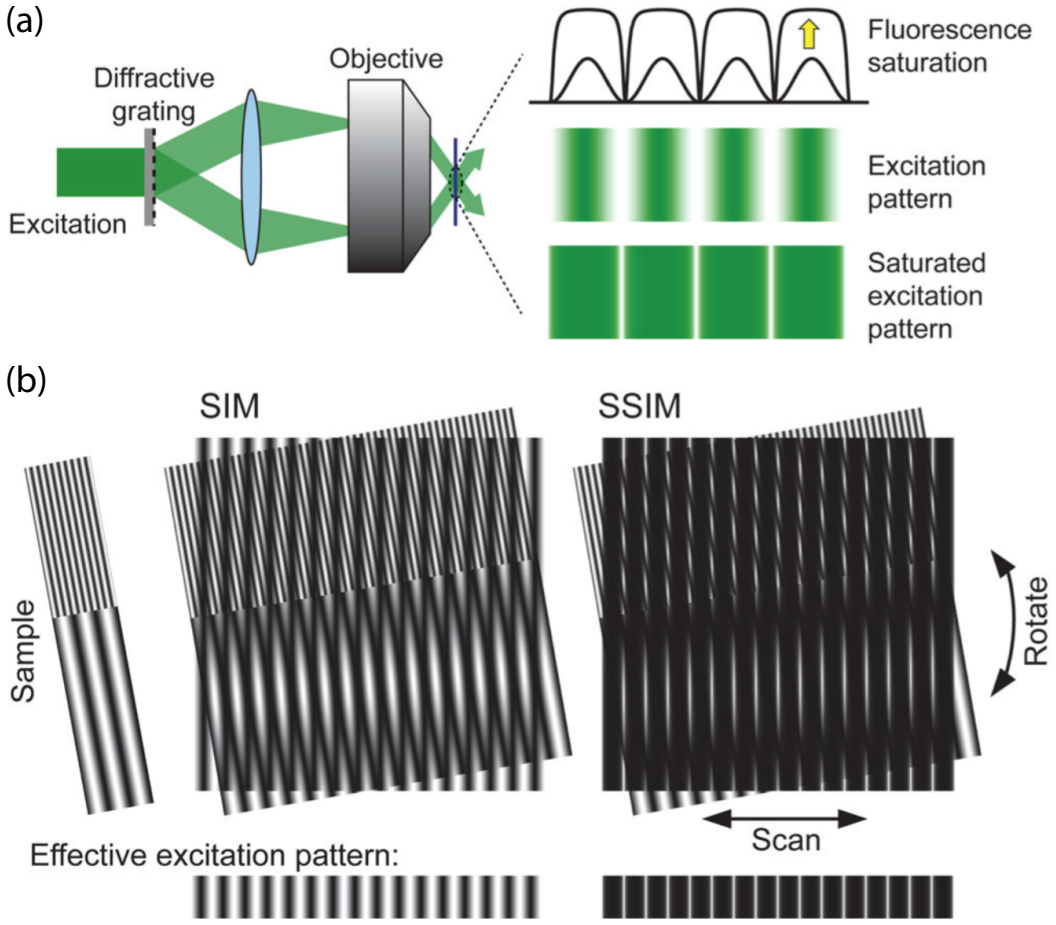



Figure 2.4: The principle of SSIM. (a) The generation of illumination pattern and nonlinear response due to saturation. (b) Generated moiré pattern after regular and saturated excitation pattern. Adapted from [26].

### 2.5.2 Stimulated Emission Depletion Microscopy

STED microscopy shares similar concepts with SSIM as they both make use of the nonlinear interaction between sample and illumination. In STED microscopy, the key effect is stimulated emission which was discovered by Einstein in 1916. Stimulated emission describes the phenomenon of an excited molecule (fluorophore) releasing a photon upon illumination with light of energy that matches the difference between the excitation state and the ground state. Such effect can be used to deplete targeted excited fluorophores. Figure 2.5 (a) shows the process of stimulated emission.

Unlike SSIM where the nonlinear saturation is used for excitation, STED microscopy uses nonlinear saturation for de-excitation (depletion). Such saturated depletion is achieved with a laser of appropriate wavelength which is called STED laser. An excitation laser first excites the fluorophores to the excited state $S_1$. The STED laser, after passing through a specially designed phase mask, generates a doughnut-shaped pattern with a zero intensity in the center and deplete the fluorophores off the center to the ground state $S_0$. Figure 2.5 (b) shows the schematic drawing of STED microscopy. Note that the STED laser pattern is still subject to the diffraction limit. However, similar to SSIM, the nonlinear relationship between the depletion and STED pattern can be achieved when the STED laser intensity reaches saturation level of depletion. Fluorophores illuminated with STED laser intensity higher than a threshold will be depleted. Therefore, the saturated depletion pattern has a nonzero intensity area that is smaller than the diffraction limit, shrinking the effective PSF width. The shrunk PSF is then scanned over the sample to generate a super-resolution image. Figure 2.5 (c) shows the process of saturated depletion shrinking the effective PSF width.

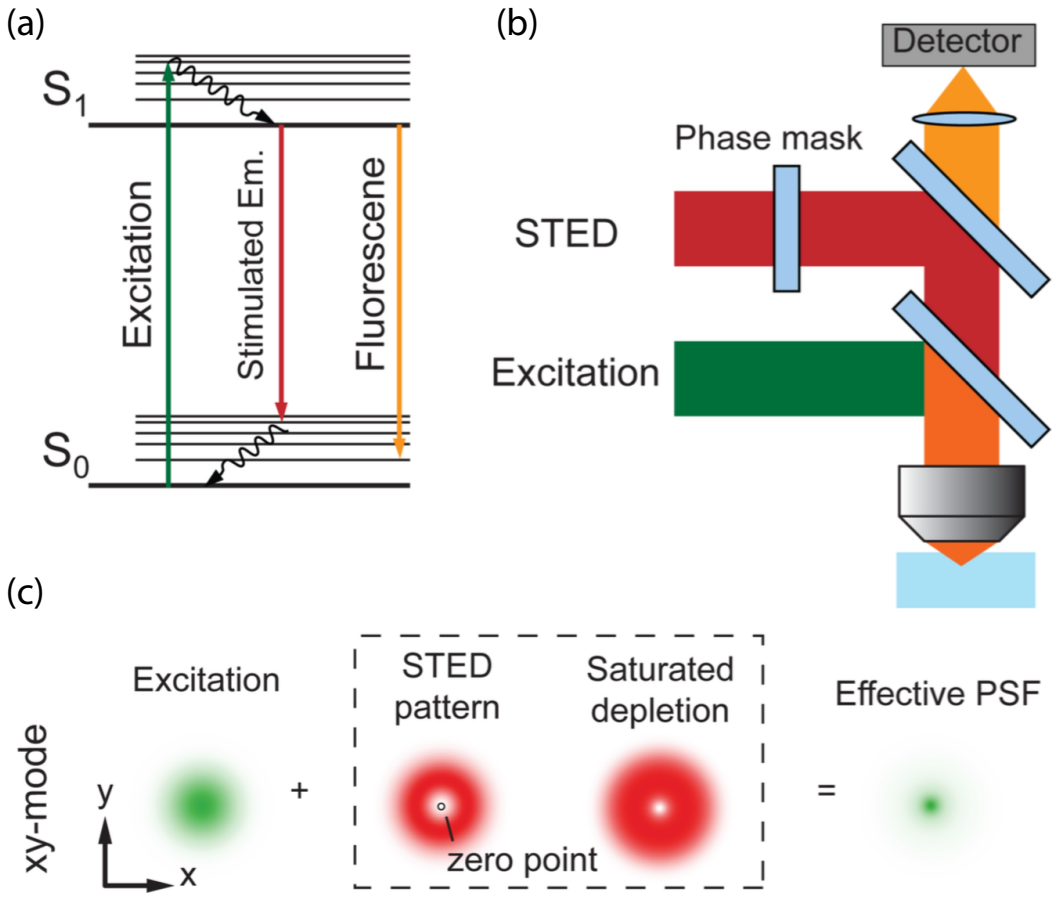



Figure 2.5: The principle of STED microscopy. (a) Illustration of stimulated emission. (b) System setup of STED microscopy. (c) Saturated depletion shrinking the effective PSF width. Adapted from [26].

In STED microscopy, the resolution is determined by the effective PSF width which is limited by the saturation level of depletion, i.e., the achievable power of STED laser, instead of the diffraction. In practice, a lateral resolution of ∼30 nm has been reported in [35].

## 2.6 Single-Molecule Localization based Super-Resolution Microscopy

Single-molecule localization microscopy relies on the precise localization of the activated single emitters. Stochastic optical reconstruction microscopy (STORM) [4], direct STORM (dSTORM) [36,37], photo-activated localization microscopy (PALM) [3], and fluorescence photoactivation localization microscopy (fPALM) [38] all belong to this category. The concept was proposed around the same time in 2006 by three independent groups of Samuel Hess [38], Xiaowei Zhuang [4] and Eric Betzig [3], respectively, and such super-resolution method holds superior capability of achieving a typical lateral resolution of ∼20 nm [39–41]. In this dissertation, emitters and fluorophores are used interchangeably.

### 2.6.1 Principle

Single molecule localization based super-resolution microscopy improves the spatial resolution significantly by combining two techniques: single-molecule localization and photoswitchable fluorescence. Although the use of single-molecule techniques to precisely localize and track molecules to a much higher precision than the diffraction limit has been known for a long time [42–44], they were only applied to imaging a few separated fluorophores. This is because when multiple fluorophores are located within the diffraction limit, their PSFs become overlapping, preventing precise localization: for general fluorescently labelled samples, hundreds of fluorophores may be located within a diffraction-limited region. The key to overcome this barrier is to separate the fluorophores in time instead of space. This can be realized by photoswitchable fluorophores, which are fluorophores that can be switched between activated

and inactivated states upon illumination of light. PALM and fPALM use photoactivable fluorescent proteins which irreversibly photobleach after emission; STORM and dSTORM use photoswitchable fluorescent dyes which have multiple switching cycles. The emission of all these fluorescent dyes or proteins can be controlled to allow only a random sparse subset of fluorophores to be activated at a time. Mathematically, in each frame the signal at the object plane $\chi(\boldsymbol{w})$ consists only a few Dirac spikes:

$$\chi(\boldsymbol{w}) = \sum_i c_i \delta(\boldsymbol{w} - \boldsymbol{w_i}), \tag{2.5}$$

where $\boldsymbol{w}_i$ represents the location of emitter $i$ and $c_i \geq 0$ is the corresponding brightness. Super-resolution is achieved by the precise localization of sparsely activated emitters and repeated activation to allow different subsets of emitters to be activated for localization. The final super-resolution image is reconstructed as a superposition of the localized emitter positions from thousands of frames. This principle is illustrated in Figure 2.6.

### 2.6.2 Data Analysis

The images acquired by SMLM are very different from those by the traditional fluorescence microscopy and require sophisticated reconstruction algorithms to generate the final super-resolution image. A number of different reconstruction algorithms have been proposed and we will summarize two most popular choices, i.e., single emitter algorithm and compressive sensing based reconstruction, in Chapter 3. These algorithms take raw images of blinking fluorophores as input and output a list of localized emitter positions. Some algorithms may also produce the associated localization accuracy for each emitter.

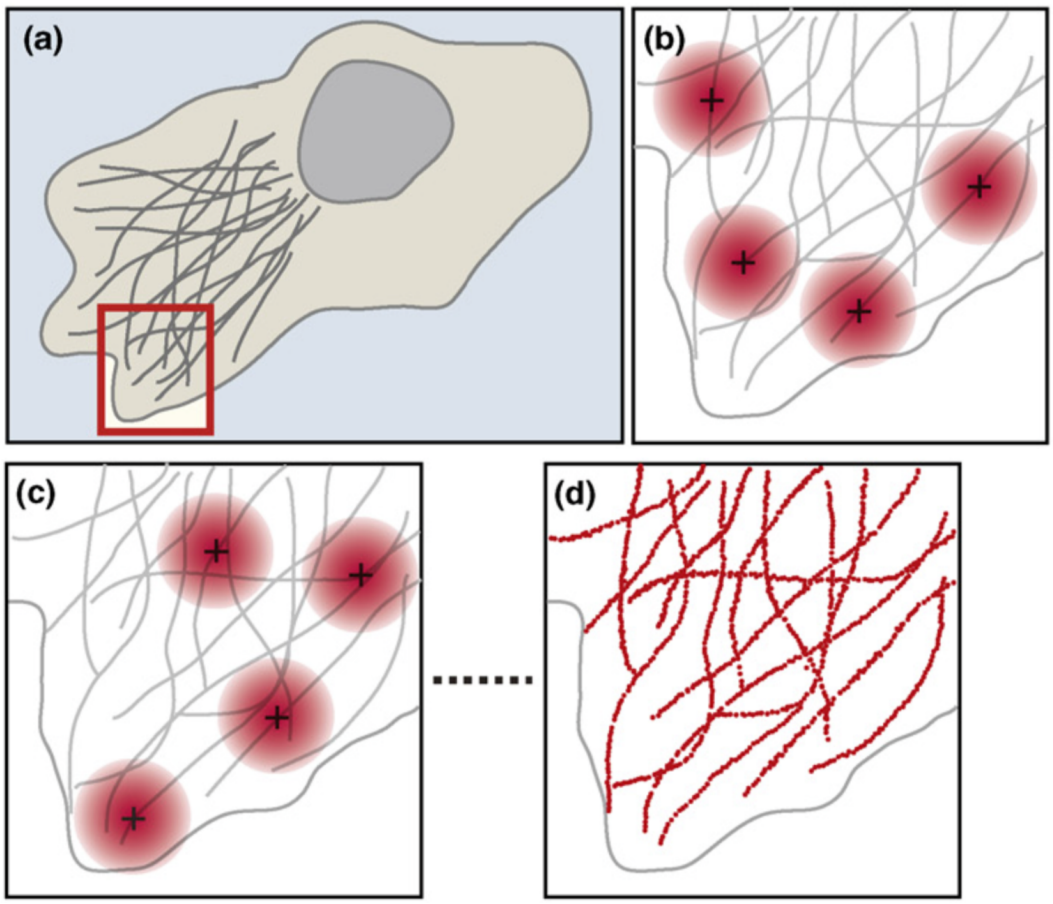



Figure 2.6: SMLM imaging principle. (a) Schematic drawing of a cell within which the gray filaments are labelled with photoswitchable fluorophores. (b) One typical acquired frame where only a sparse subset of fluorophores are activated. (c) A subsequent frame where a different random subset of fluorophores are activated. (d) The high-resolution image reconstructed from the localized fluorophore positions after sufficient number of frames are acquired. Adapted from [45].

### 2.6.3 Drift Correction

Typically SMLM requires a relatively long acquisition time due to its iterative nature. Therefore, considering the high localization accuracy, any drift in the lateral or axial direction will deteriorate the final reconstruction quality significantly. In practice, there are many factors that could cause the sample to drift during the acquisition, including but not limited to, thermal variance of the microscope, vibration of the optical table, air flow in the room, etc. To reduce the effect of sample drift to the minimum, several drift correction methods have been proposed.

**Lateral Drift**

The drift in the lateral direction is typically corrected offline. One method involves adding fiducial markers such as fluorescent beads, quantum dots, or gold particles in the sample and they are visible during the acquisition and stay stationary relative to the sample. These markers serve as reference points and their positions over time can be determined and used to offset the movement of the sample in a post-processing step.

The second method is to calculate the sample movement as a function of time using the acquired SMLM images themselves. The acquired SMLM images are divided into small chunks and reconstruction of each chunk can be analyzed by image alignment algorithms to determine the sample drift of each chunk. The calculated drift can then be subtracted from the localization of each chunk. Another similar approach is to acquire a non-fluorescence image such as DIC image every few seconds. The sample drift can be calculated using these non-fluorescence images.

#### Axial Drift

The drift in the axial direction is more difficult to correct in the post-processing step than the lateral drift due to the lower axial resolution. Therefore, active correction mechanism is needed during the acquisition. Typically, a weak infrared laser beam is directed to pass the objective and reflect back at the glass-oil (or glass-air) interface. The reflection is focused on a 2D position sensitive detector which can detect the centroid of the beam. The movement of laser beam on QPD represents axial movement of sample and can be used to correct the focus drift in real time [46].

### 2.6.4 Visualization

The output of SMLM is a list of localized fluorophore positions and cannot be directly visualized. The most common approach of visualization is to create an oversampled grid with the value of each grid pixel representing the number of emitters within. One can also take into account the localization accuracy and render the reconstruction as a sum of 2D Gaussian kernels where each kernel represents a localization with its width representing the corresponding localization uncertainty. Other approaches include quadtree-based adaptive histogram and Delaunay triangulation [47].

### 2.6.5 Resolution

The resolution of SMLM is determined by many factors such as localization precision, drift correction accuracy, labelling density, etc, and is not easily derived. Here, for simplicity, we only consider the localization accuracy in the case of non-overlapping fluorophores. The localization problem can be viewed as determining the position of a fluorophore given $I$ independent measurements, where $I$ is the number of collected photons. Each measurement has an uncertainty $\sigma$ due to diffraction, where $\sigma$ is the

standard deviation of the system PSF. Therefore, if the positions of $I$ photons can be determined exactly, according to the central limit theorem, when $I$ is large enough, the mean of $I$ photons' positions approximately follows a normal distribution with mean of the fluorophore position and standard deviation of $\sigma/\sqrt{I}$. So roughly speaking, the localization accuracy of the fluorophore is limited by $\sigma/\sqrt{I}$. This precision, although still depending on the PSF width, is scaled by the number of collected photons, resulting in a resolution not limited by the diffraction of light.

Thompson et al. [44] further take into account the pixelation of camera image and noise effect to propose a resolution limit for localization of a single fluorophore as

$$\sigma_{\text{loc}} = \sqrt{\frac{\sigma^2 + G/12}{I} + \frac{8\pi\sigma^4 b^2}{GI^2}}, \tag{2.6}$$

where $G$ is the area of a camera pixel and $b$ is the background noise level. Another estimation of the resolution limit given by Mortensen et al. [48] is represented as

$$\sigma_{\text{loc}} = \sqrt{\frac{\sigma_a^2}{I}\left(\frac{16}{9} + \frac{8\pi\sigma_a^2 b^2}{IG}\right)}, \tag{2.7}$$

where $\sigma_a^2 = \sigma^2 + G/12$.

### 2.6.6 Three-Dimensional Imaging

Most structures of interest to biologists have a 3D morphology in nature while most 3D microscopy imaging techniques have the lowest resolution in the axial direction. Therefore, there has been a growing interest to extend SMLM to 3D. The first method that is capable of 3D single-molecule imaging is proposed by Huang et al. [49] based on optical astigmatism. By placing a weak cylindrical lens into the optical path, a difference of focus in the $x$ and $y$ direction is created. Thus, the ellipticity and orientation of the emitter's PSF varies along the $z$ direction, as shown in Fig. 2.7(a),

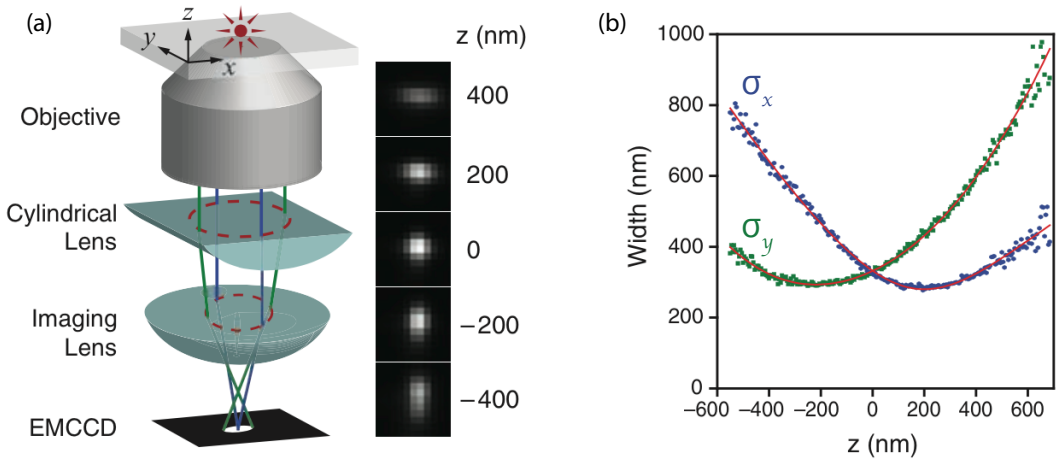



Figure 2.7: The scheme of 3D SMLM. (a) The optical diagram illustrates the principle of 3D localization microscope by introducing a cylindrical lens. The $z$ position of the emitter can be localized according to the ellipticity of the image. The right panel shows the image of an emitter in different $z$ locations. (b) Calibration curve of system PSF width $\sigma_x$ and $\sigma_y$ as a function of $z$ location. Adapted from [49].

making it possible to differentiate emitters in different $z$ locations. When the emitter is located on the average focal plane, the PSF is round shaped. When the emitter is located above the average focal plane, the PSF becomes elliptical and is elongated in the $y$ direction and vice versa. Fig. 2.7(b) shows the calibration curve of system PSF width $\sigma_x$ and $\sigma_y$ as a function of $z$ location. It is clear that the symmetry of system PSF in $z$ direction is violated. The dependence of $\sigma_x$ and $\sigma_y$ on the emitter axial position $z_0$ can be formulated as

$$\begin{aligned}\sigma_x(z_0) &= \sigma_{x,0}\sqrt{1+\sum_{i=2}^{4} A_{x,i}\left(\frac{z_0-c_x}{d_x}\right)^i},\\ \sigma_y(z_0) &= \sigma_{y,0}\sqrt{1+\sum_{i=2}^{4} A_{y,i}\left(\frac{z_0-c_y}{d_y}\right)^i},\end{aligned} \tag{2.8}$$

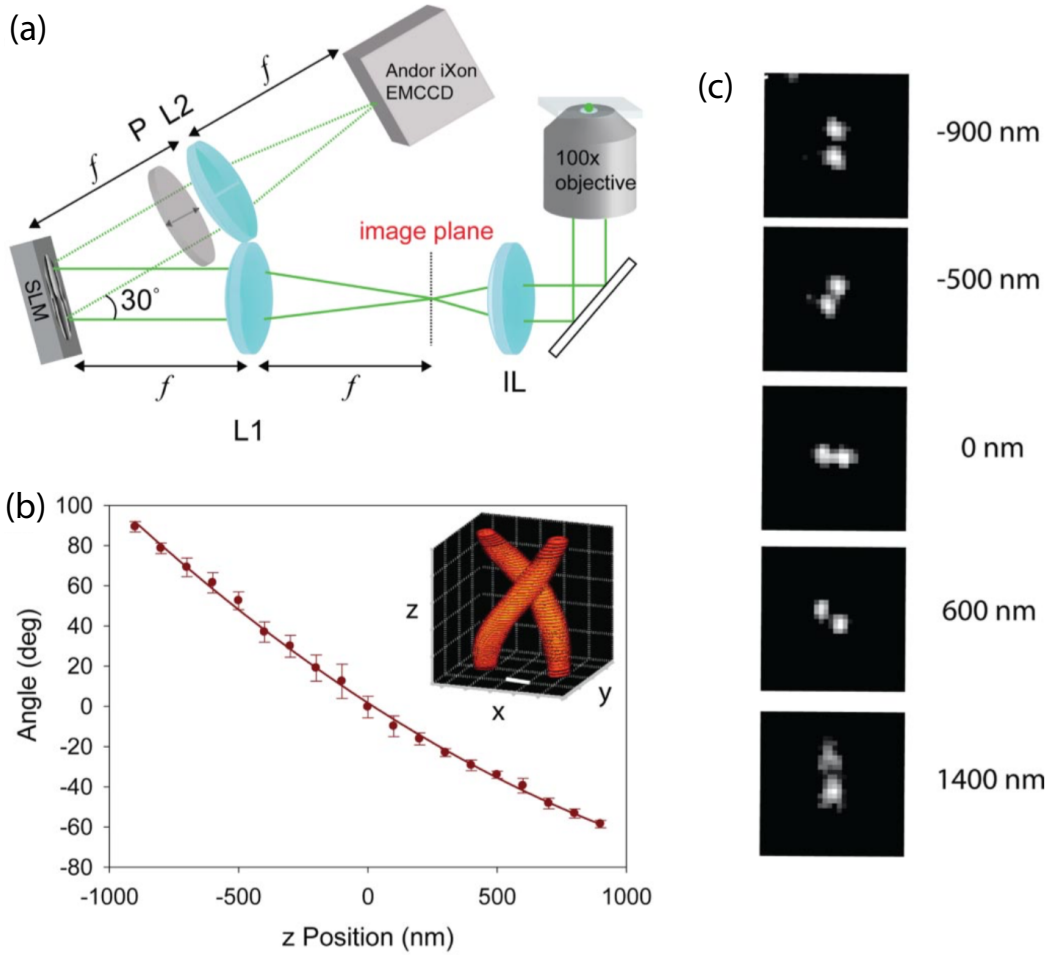



Figure 2.8: DH-PSF imaging principle. (a) The optical setup of single-molecule DH-PSF imaging. (b) The calibration curve of the angle between two lobes as a function of $z$ position. (c) Images of an emitter in different $z$ locations. Adapted from [50].

where $\sigma_{x,0}$ and $\sigma_{y,0}$ are the width of PSF in the $x$ and $y$ directions at focus, $c_x$ and $c_y$ are the offsets from the focal plane, $d_x$ and $d_y$ are the focus depths, $A_{x,i}$ and $A_{y,i}$, $i = 2, 3, 4$, are used to correct the non-linearity of the optical system. Astigmatism based 3D imaging is relatively easy to implement and achieves an axial resolution of $\sim 50$ nm [49].

Another method is proposed by Pavania et al. [50] where the system PSF is engineered to have two lobes using a spatial light modulator. The angle of two lobes changes at different axial locations. It is termed double-helix PSF (DH-PSF) as the PSF over $z$ positions resembles a double-helix. Therefore, the emitter lateral position can be determined as the center of two lobes while its axial location is determined by the angle between two lobes with a calibrated lookup table. The imaging principle of DH-PSF based 3D SMLM is shown in Figure 2.8. DH-PSF is able to image samples with a large field of depth (over 2 $\mu m$).

## 2.7 Home-Built SMLM System

We have set up a SMLM system (Fig. 2.9) based on an inverted microscope (IX71, Olympus America Inc.) with 1.49 NA 100× oil immersion TIRF objective. Four lasers, 405 nm, 488 nm, and 647 nm diode lasers (Vortran Laser Technology Inc.) and 561 nm diode-pumped solid-state laser (CrystaLaser), covering the spectrum of the most commonly used fluorophores, are combined in a single optical path using dichroic mirrors. These four laser beams are then expanded and focused onto the back focal plane of the objective lens, where the focus point is allowed to move through the use of a translational stage. Therefore, the illumination light can be controlled at a high incidence angle, narrowing the depth of illumination to improve the signal-to-noise ratio (SNR). A multi-band dichroic mirror (FF405/496/593/649, Semrock) is used to separate emission from excitation laser. Multiple emission filters are placed before cameras to further reduce the contamination from the excitation laser to a negligible level.

Three thermoelectrically cooled electron multiplying CCD (EMCCD) cameras (iXon Ultra 897, Andor Technologies PLC.) are used for imaging with the electron-multiplying (EM) gain set to 255 to provide multi-focus astigmatism based 3D imaging platform (Chapter 5). For regular 2D imaging, only one camera is used. EMCCD uses on-chip signal amplification technology to raise the signal level via a process called impact ionization in order to reduce the effect of camera readout noise and dark current for better SNR.

The sample holder is mounted on a 3D piezo stage (Nano-LPS, Mad City Lab). An infrared 980 nm laser is used in combination with the piezo stage for the axial drift correction [49]. A computer streams the continuously read data from cameras directly to a hard disk. All instruments and acquisition are controlled by home-written LabVIEW programs (National Instruments Corp.).

Several pictures of our home-built system setup are shown in Figure 2.10 (a - d). Figure 2.10 (a) shows a portion of the excitation section where four lasers are expanded combined to a single optical path. Figure 2.10 (b) shows the rest portion of excitation section as well as the axial drift correction system. The inverted microscopy and sample stage are shown in Figure 2.10 (c). Figure 2.10 (d) shows the detection section. A screenshot of the LabVIEW programs for laser, stage and camera control and image acquisition is shown in Figure 2.10 (e).

## 2.8 Cell Culture and Immuno-Fluorescence Staining

To test our system on experimental data, we prepare samples for our imaging system and the procedure is described here. Dulbeccos modified Eagles medium (DMEM, Invitrogen) is supplemented with 10% FBS, 100 U/mL penicillin and 100

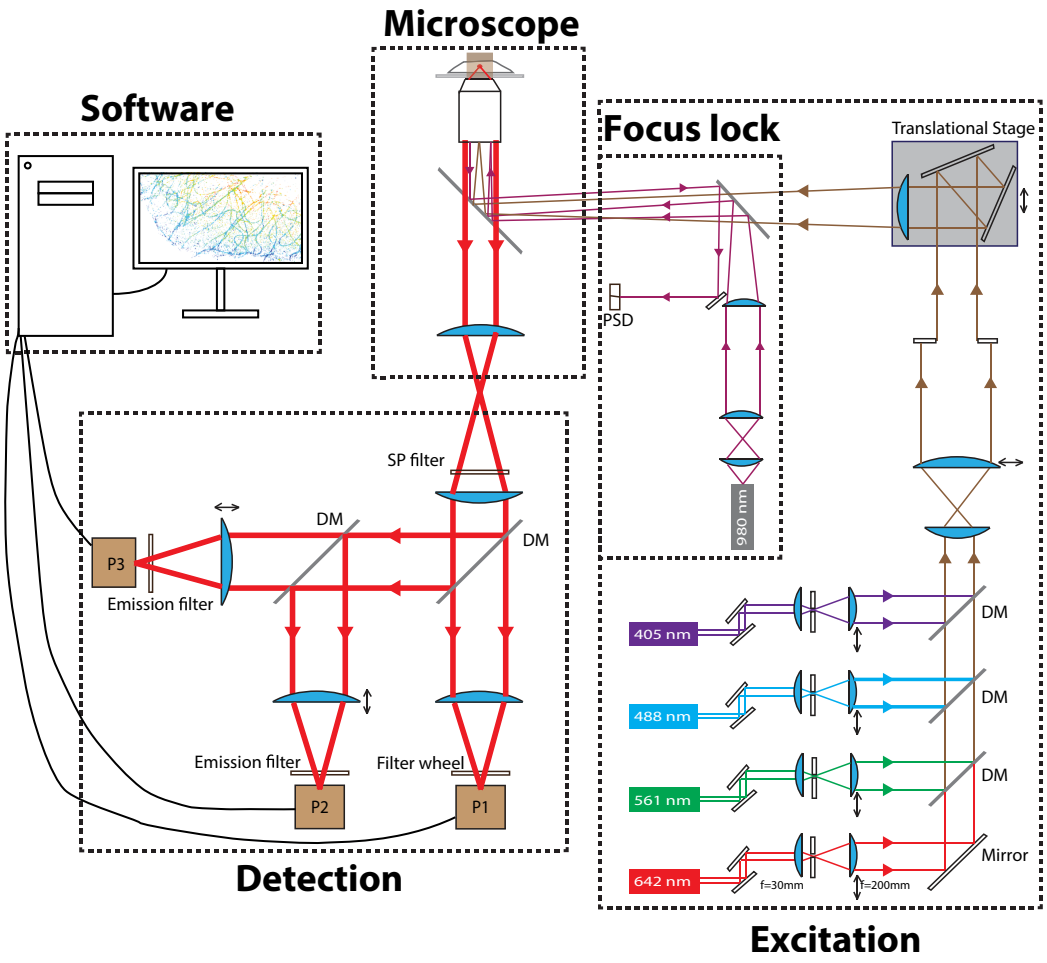



Figure 2.9: Diagram of the home-built SMLM setup composed of Excitation, Microscope, Focus lock, Detection and Software. PSD: position sensitive detector; DM: dichroic mirror; SP: short-pass.

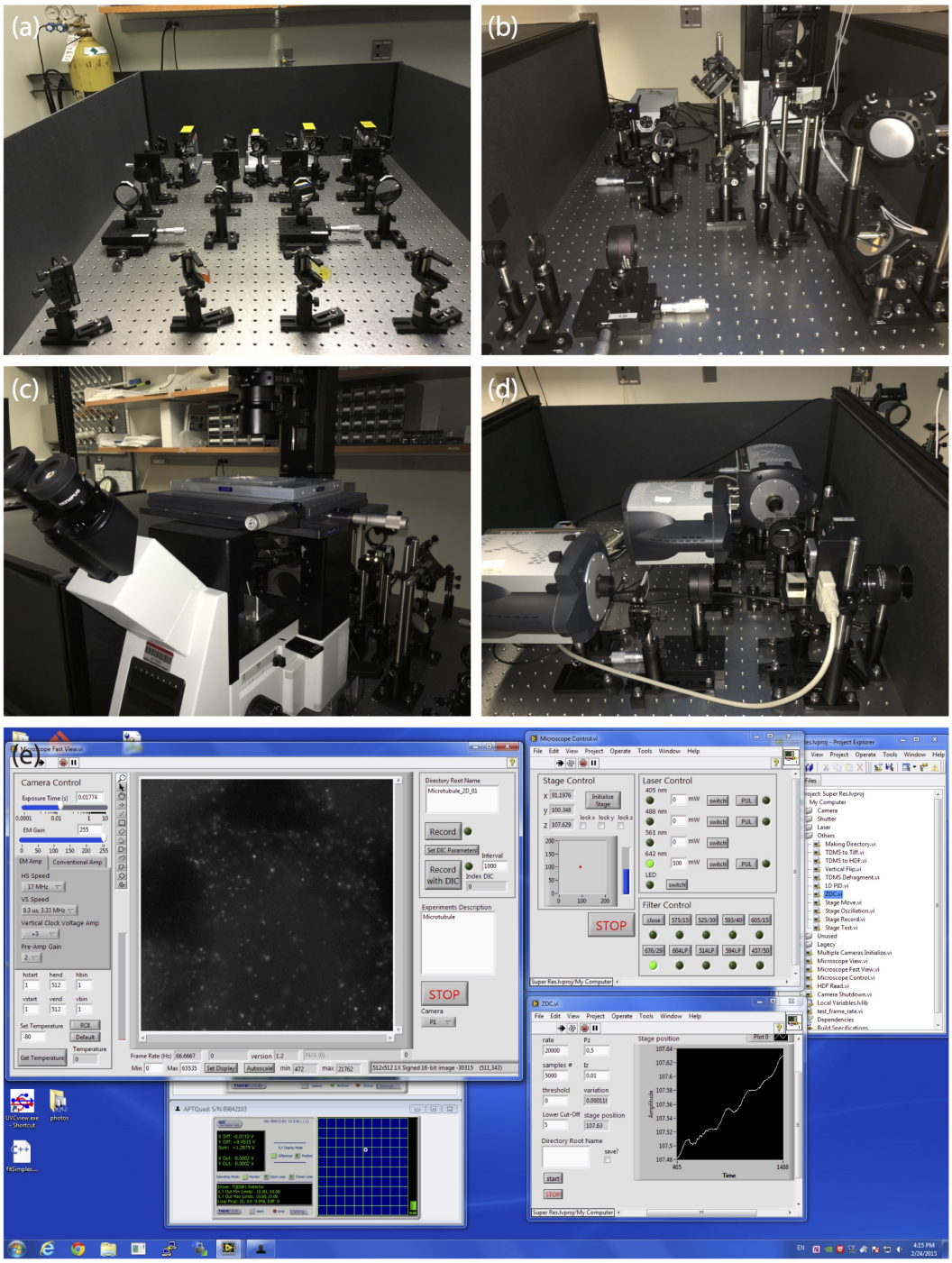



Figure 2.10: Pictures of the home-built SMLM system. (a) A portion of excitation. (b) Axial drift correction system and the rest portion of excitation. (c) The inverted microscope and sample stage. (d) Detection system. (e) A screenshot of the LabVIEW programs for laser, stage and camera control and image acquisition.

$\mu$g/ml streptomycin and warmed to 37 $^oC$ before adding to cells. 250,000 HeLa cells are seeded onto a 35 mm, No. 1.5 glass window dish (MatTek Corp) at 37 $^oC$ with 5% $CO_2$. After culturing cells for 24 hours, cells are fixed at room temperature in a fixation buffer (3% paraformaldehyde and 0.1% glutaraldehyde in PBS) for 10 minutes, and then rapidly washed 3 times with PBS. Cells are blocked in blocking buffer (3% bovine serum albumin (BSA) and 0.2% Triton X-100) for 1 hour at room temperature. After blocking, $\alpha$-tubulin (Sigma Aldrich) is diluted in blocking buffer and applied to cells for 1 hour at room temperature. Cells are washed 3 times 5 minutes each in PBS on a rocking platform and subsequently treated with appropriate secondary antibody, in this instance Mouse-IgG, conjugated to AlexaFluor647 for 30 minutes. Antibodies are secured through a post-fixation step using the fixation buffer for 5 minutes. Imaging buffer (10% glucose, 50 mM Tris pH 8.5, 10 mM NaCl, 14 mg Glucose Oxidase, 50 $\mu$L 20 mg/mL catalase, and 1X $\beta$-mercaptoethanol (Sigma Aldrich)) is used throughout experimentation.

# Chapter 3: Signal Model and Existing Algorithms

The success of SMLM depends on many factors such as sample labelling density and axial drift correction. Nonetheless, the most important components of SMLM are: (i) photoswitchable fluorophores which can switch between the activated state and inactivated state stochastically upon illumination with light of specific wavelength; (ii) high-performance reconstruction algorithms that can localize fluorophores' positions precisely. In the last chapter we have introduced the principle of photoswitchable fluorophores which make single-molecule localization possible. In this chapter we present existing reconstruction algorithms for SMLM. The growing use of SMLM in the biological imaging field has led to substantial development of localization algorithms for both 2D and 3D reconstruction, extending from isolated emitter PSFs to the analysis of overlapping PSFs.

## 3.1 Two-Dimensional Signal Model

We begin by introducing the signal model of SMLM in 2D imaging. In each frame, a sparse subset of emitters are activated. Let $\theta^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_I^{(i)}] = [\theta_L^{(i)}, \theta_I^{(i)}]$ be the parameters for the $i^{\text{th}}$ emitter, where $\theta_L^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}] \in \mathcal{S}$ are the coordinates in the $x$ and $y$ dimensions, respectively,

$$\mathcal{S} = \Big\{ \theta_L = [\theta_x, \theta_y] \Big| \theta_x \in [x_{\min}, x_{\max}],\ \theta_y \in [y_{\min}, y_{\max}] \Big\}, \tag{3.1}$$

and $\theta_I^{(i)} > 0$ denotes its intensity. Let $\Theta = \{\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(P)}\}$ be the set of parameters, where $P$ is the total number of emitters. We can write the set of activated emitters $\chi = \chi(x, y|\Theta)$ as a sparse superposition of point sources, given as

$$\chi = \chi(x, y|\Theta) = \sum_{i=1}^{P} \theta_I^{(i)} \delta(x - \theta_x^{(i)}, y - \theta_y^{(i)}), \tag{3.2}$$

where $\delta(x - x_0, y - y_0)$ is a Dirac measure located at $(x_0, y_0)$. For notational convenience, we also use $\chi(\Theta)$ to denote $\chi(x, y|\Theta)$.

Due to diffraction, the point source signal $\chi$ is low-pass filtered by the microscope PSF before forming the image, as illustrated by (2.1). Accurate PSF modelings such as Richards-Worf model [51] and Gibson-Lanni model [52] require taking account of many factors including the vector nature of light waves and interfaces between the sample and objective lens. However, the numerical evaluation of these models are computationally inefficient. Accurate approximation using Airy PSF is used by many researchers [53], as shown earlier in (2.2), which is the squared modulus of the Fourier transform of the circular aperture. To further reduce the evaluation complexity and simplify the calculations, an isometric Gaussian function is often used to approximate the system PSF [54]:

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{3.3}$$

where $\sigma$ is the standard deviation in the lateral direction. Such approximation can be very accurate when the emitter is in focus. However, when the emitter is out-of-focus, the PSF shape can become more complicated and such case will be discussed later.

Therefore the signal received at the detector plane can be written as

$$s(x, y|\chi) = (f * \chi)(x, y) = \sum_{i=1}^{P} \theta_I^{(i)} f(x - \theta_x^{(i)}, y - \theta_y^{(i)}). \tag{3.4}$$

The signal is captured by a camera, a pixelated light detection device that measures the number of photons striking each pixel. If the pixel size is sufficiently small, uniform intensity within each pixel can be assumed and the measurement at each pixel can be approximated as the multiplication between the PSF intensity at the pixel center and the pixel area. However, due to physical constraints and noise, the pixel size cannot be infinitesimal. Therefore, the expected number of photons received at the $(m,n)^{\text{th}}$ camera pixel, denoted as $\mu[m,n|\chi]$, can be written as an integration of $s(x,y)$ over the area of a pixel:

$$\begin{aligned}\mu[m,n|\chi] &= \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} s(x,y|\chi)\, dxdy, \\ &= \sum_{i=1}^{P} \frac{\theta_I^{(i)}}{4} \left[ Q\left( \frac{m-\theta_x^{(i)}+\frac{1}{2}}{\sqrt{2}\sigma} \right) - Q\left( \frac{m-\theta_x^{(i)}-\frac{1}{2}}{\sqrt{2}\sigma} \right) \right] \\ &\quad \times \left[ Q\left( \frac{n-\theta_y^{(i)}+\frac{1}{2}}{\sqrt{2}\sigma} \right) - Q\left( \frac{n-\theta_y^{(i)}-\frac{1}{2}}{\sqrt{2}\sigma} \right) \right], \end{aligned} \tag{3.5}$$

where $Q(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

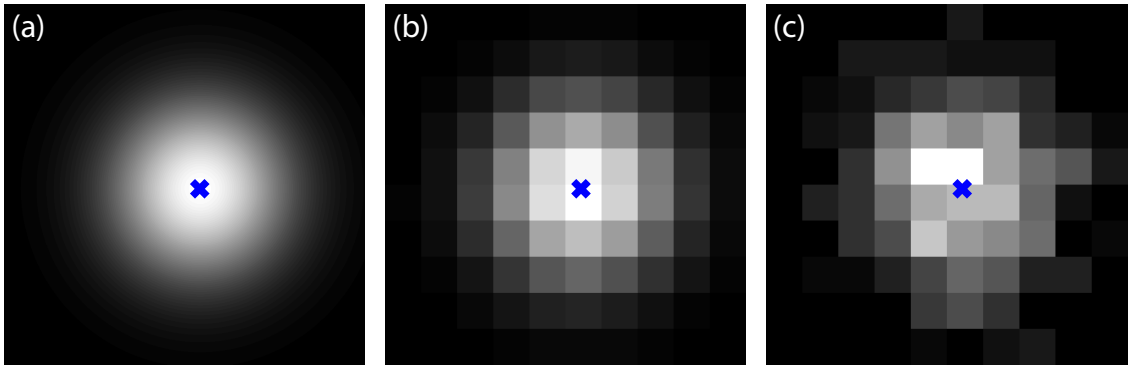



Figure 3.1: The effect of pixelation and noise. (a) Noise-free and pixel-free signal at detector plane. (b) Pixelated image without noise. (c) Pixelated image after noise. The blue cross represents the true emitter location.

The actual number of photons hitting the camera at the $(m,n)^{\text{th}}$ pixel, denoted as $y[m,n]$, follows an independent Poisson distribution with the parameter $\mu[m,n|\chi]$, given as

$$y[m,n|\chi] \sim \text{Poisson}(\mu[m,n|\chi]). \tag{3.6}$$

The effect of pixelation and noise is shown in Figure 3.1. Figure 3.1 shows the signal of a single emitter at the detector plane. The pixelated image is shown in Figure 3.1 (b) and Figure 3.1 (c) is the actual acquired image after noise.

## 3.2 Three-Dimensional Signal Model

In 3D imaging, the parameter for the $i^{\text{th}}$ emitter becomes $\theta^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_z^{(i)}, \theta_I^{(i)}] = [\theta_L^{(i)}, \theta_I^{(i)}]$, where $\theta_L^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_z^{(i)}] \in \mathcal{S}$ are the coordinates in $x$, $y$ and $z$ dimensions, respectively and

$$\mathcal{S} = \Big\{\theta_L = [\theta_x, \theta_y, \theta_z] \Big| \theta_x \in [x_{\min}, x_{\max}],\ \theta_y \in [y_{\min}, y_{\max}],\ \theta_z \in [z_{\min}, z_{\max}]\Big\}. \tag{3.7}$$

Then the set of activated emitters $\chi = \chi(x,y,z|\Theta)$ can also be written as a sparse superposition of point sources, given as

$$\chi = \chi(x,y,z|\Theta) = \sum_{i=1}^{P} \theta_I^{(i)} \delta(x - \theta_x^{(i)}, y - \theta_y^{(i)}, z - \theta_z^{(i)}), \tag{3.8}$$

where $\delta(x - x_0, y - y_0, z - z_0)$ is a Dirac measure located at $(x_0, y_0, z_0)$.

For different 3D imaging schemes the system PSF varies. In this dissertation we focus on astigmatism based 3D imaging configuration. Herein, the system PSF can be approximated as an asymmetric 2D Gaussian function with the ellipticity determined by the emitter location along the $z$ direction:

$$f(x,y|z_0) = \frac{1}{2\pi\sigma_x(z_0)\sigma_y(z_0)} e^{-\left(\frac{x^2}{2\sigma_x(z_0)^2} + \frac{y^2}{2\sigma_y(z_0)^2}\right)}, \tag{3.9}$$

where $\sigma_x(z_0)$ and $\sigma_y(z_0)$ are the standard deviations in the $x$ and $y$ directions and their dependence on the emitter axial location $z_0$ is formulated in (2.8).

Therefore the signal at the detector plane is given as

$$s(x, y|\chi) = \sum_{i=1}^{P} \theta_I^{(i)} f(x - \theta_x^{(i)}, y - \theta_y^{(i)}|\theta_z^{(i)}). \tag{3.10}$$

Integrating $s(x, y|\chi)$ over the pixel area, we have the expected number of photons received at the $(m, n)^{\text{th}}$ camera pixel written as

$$\begin{aligned}
\mu[m, n|\chi] &= \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} s(x, y|\chi)\, dxdy, \\
&= \sum_{i=1}^{P} \frac{\theta_I^{(i)}}{4} \left[ Q\left( \frac{m - \theta_x^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) - Q\left( \frac{m - \theta_x^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) \right] \\
&\quad \times \left[ Q\left( \frac{n - \theta_y^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) - Q\left( \frac{n - \theta_y^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) \right].
\end{aligned} \tag{3.11}$$

The actual camera signal at the $(m, n)^{\text{th}}$ camera pixel also suffers from the intrinsic shot noise and follows a Poisson distribution with parameter $\mu[m, n|\chi]$ as shown in (3.6).

Denote the camera image as $y = \{y[m, n]\}$. The objective of both 2D and 3D super-resolution image reconstruction, therefore, is to estimate the point source signal $\chi(\Theta)$ (or $\Theta$), given the observed image $y$.

## 3.3 Single Emitter Algorithm

A number of super-resolution image reconstruction algorithms have been proposed for SMLM since its invention. Though they vary in the underlying assumptions and optimization objectives, they can be roughly divided into two categories: single emitter fitting based and multiple emitters localization based methods. Single emitter fitting based reconstruction, also known as the Single Emitter Algorithm (SEA) [4],

first identifies candidate emitters that are well separated from others. Since the system's PSF can be well approximated by a Gaussian function, fitting to a Gaussian function is then performed on a small area surrounding the candidate emitters. For 2D image reconstruction, emitters not on focus are eliminated according to the fitted PSF width and intensity. In 3D case, the PSF widths are used to calculate emitters' axial positions. This procedure is repeated for all the frames and can be implemented in parallel with a multi-core central processing unit (CPU) or a graphics processing unit (GPU) efficiently [55].

### 3.3.1 Peak Identification

The identification of candidate emitters serves as the first step in SEA. A Gaussian filtering is applied to the image to remove high-frequency noise and low-frequency background variations. The Gaussian kernel should have the same width as the system PSF and integrate to zero. Let $\acute{y}$ represent the filtered image and we have

$$\acute{y}[u,v] = (f * y)[u,v], \tag{3.12}$$

where $f[u,v]$ is the discrete representation of system PSF. Another filter commonly used for candidate emitters detection is Laplacian of Gaussian (LoG) filter. After filtering, the image is then searched for local maxima above a given threshold $\zeta$, i.e.,

$$\begin{aligned}(\acute{\theta}_x^{(i)}, \acute{\theta}_y^{(i)}) = (u,v) \quad \text{if} \quad & \acute{y}[u,v] > \acute{y}[u+1,v] \quad \& \quad \acute{y}[u,v] > \acute{y}[u-1,v] \quad \& \\ & \acute{y}[u,v] > \acute{y}[u,v+1] \quad \& \quad \acute{y}[u,v] > \acute{y}[u,v-1] \quad \& \\ & \acute{y}[u,v] > \zeta. \end{aligned} \tag{3.13}$$

These local maxima $\{(\acute{\theta}_x^{(i)}, \acute{\theta}_y^{(i)})\}$ are considered to be peak positions and their surrounding boxes $\mathcal{S}^{(i)}$ are extracted as candidates:

$$\mathcal{S}^{(i)} = \{(u,v) \mid |u - \acute{\theta}_x^{(i)}| < \frac{\text{blockwidth}}{2} \;\&\; |v - \acute{\theta}_y^{(i)}| < \frac{\text{blockwidth}}{2}\}. \tag{3.14}$$

Figure 3.2 (a) and (b) show a typical SMLM image and identified candidate emitters in the red boxes. Figure 3.2 (c) shows the images of a few extracted emitters.

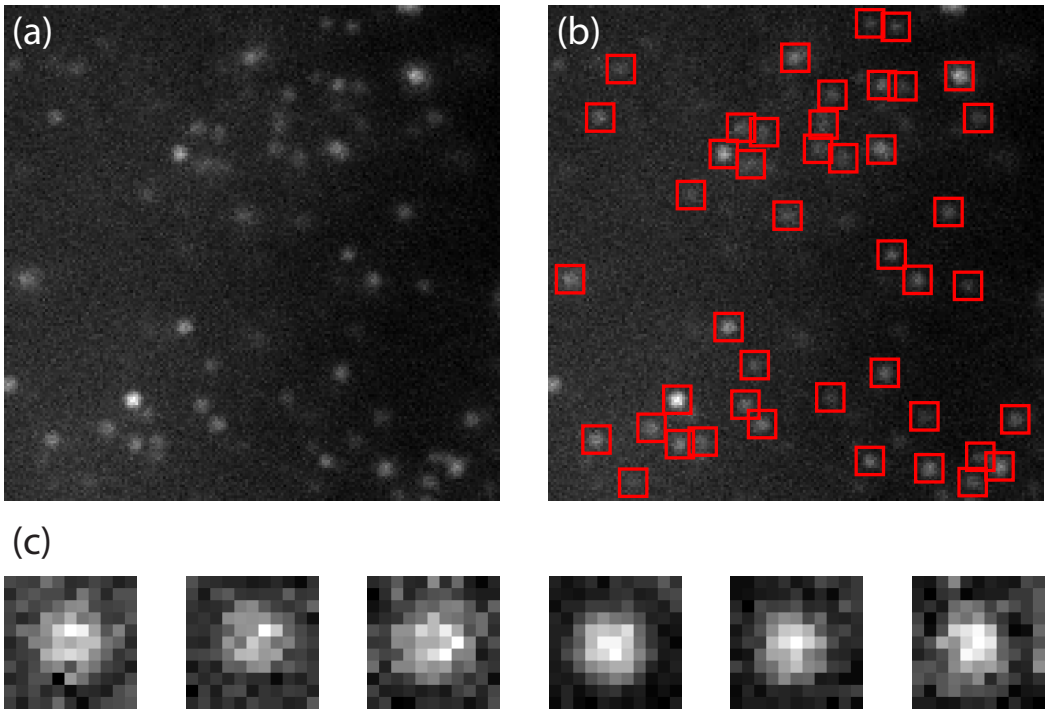



Figure 3.2: (a) A typical SMLM image. (b) Identified candidate emitters shown in red boxes. (c) Images of a few extracted emitters.

### 3.3.2 Localization

After images of candidate emitters are extracted, the position of these extracted emitters are further refined with sub-pixel resolution using localization algorithms.

There are a number of localization algorithms that have been proposed and we summarize them here.

**QuickPALM**

The most straightforward and intuitive localization algorithm is probably to calculate the weighted centroid of extracted image (center of mass) as the estimated position of emitters [56], i.e.,

$$\begin{aligned}\hat{\theta}_x^{(i)} &= \frac{\sum_{(m,n)\in\mathcal{S}^{(i)}} m \cdot y[m,n]}{\sum_{(m,n)\in\mathcal{S}^{(i)}} y[m,n]},\\ \hat{\theta}_y^{(i)} &= \frac{\sum_{(m,n)\in\mathcal{S}^{(i)}} n \cdot y[m,n]}{\sum_{(m,n)\in\mathcal{S}^{(i)}} y[m,n]},\end{aligned} \tag{3.15}$$

where $\mathcal{S}^{(i)}$ is the set of pixels corresponding to the extracted square region of $i^{\text{th}}$ emitter, as defined in (3.14). The spot shape can also be characterized using four parameters: left radius $\hat{\sigma}_l^{(i)}$, right radius $\hat{\sigma}_r^{(i)}$, top radius $\hat{\sigma}_t^{(i)}$ and bottom radius $\hat{\sigma}_b^{(i)}$, defined as:

$$\begin{aligned}\hat{\sigma}_l^{(i)} &= \frac{\sum_{(m,n)\in\mathcal{S}_l^{(i)}} y[m,n](\hat{\theta}_x^{(i)} - m)}{\sum_{(m,n)\in\mathcal{S}_l^{(i)}} y[m,n]}, \hat{\sigma}_r^{(i)} = \frac{\sum_{(m,n)\in\mathcal{S}_r^{(i)}} y[m,n](m - \hat{\theta}_x^{(i)})}{\sum_{(m,n)\in\mathcal{S}_r^{(i)}} y[m,n]},\\ \hat{\sigma}_t^{(i)} &= \frac{\sum_{(m,n)\in\mathcal{S}_t^{(i)}} y[m,n](\hat{\theta}_y^{(i)} - n)}{\sum_{(m,n)\in\mathcal{S}_t^{(i)}} y[m,n]}, \hat{\sigma}_b^{(i)} = \frac{\sum_{(m,n)\in\mathcal{S}_b^{(i)}} y[m,n](n - \hat{\theta}_y^{(i)})}{\sum_{(m,n)\in\mathcal{S}_b^{(i)}} y[m,n]},\end{aligned} \tag{3.16}$$

where

$$\begin{aligned}\mathcal{S}_l^{(i)} &= \{(m,n)|m < \hat{\theta}_x^{(i)} \;\&\; (m,n) \in \mathcal{S}^{(i)}\},\\ \mathcal{S}_r^{(i)} &= \{(m,n)|m \geq \hat{\theta}_x^{(i)} \;\&\; (m,n) \in \mathcal{S}^{(i)}\},\\ \mathcal{S}_t^{(i)} &= \{(m,n)|n < \hat{\theta}_y^{(i)} \;\&\; (m,n) \in \mathcal{S}^{(i)}\},\\ \mathcal{S}_b^{(i)} &= \{(m,n)|n \geq \hat{\theta}_y^{(i)} \;\&\; (m,n) \in \mathcal{S}^{(i)}\}.\end{aligned} \tag{3.17}$$

When the emitter PSF is isotropic (such as in 2D case), $\hat{\sigma}_l^{(i)} \approx \hat{\sigma}_r^{(i)} \approx \hat{\sigma}_t^{(i)} \approx \hat{\sigma}_b^{(i)}$. However, when there are multiple PSFs overlapping, violating the single emitter presumption, such symmetry is no longer valid and the symmetry level in the $x$ and $y$

direction, defined as $S_x$ and $S_y$, can be used to as a filtering criteria in a later step:

$$\begin{aligned} S_x &= 1 - \frac{\left|\hat{\sigma}_l^{(i)} - \hat{\sigma}_r^{(i)}\right|}{\hat{\sigma}_l^{(i)} + \hat{\sigma}_r^{(i)}}, \\ S_y &= 1 - \frac{\left|\hat{\sigma}_t^{(i)} - \hat{\sigma}_b^{(i)}\right|}{\hat{\sigma}_t^{(i)} + \hat{\sigma}_b^{(i)}}. \end{aligned} \tag{3.18}$$

Higher values of $S_x$ and $S_y$ represent higher symmetry along the corresponding axis.

**Least-Squares Fitting**

Another commonly used localization algorithm seeks to minimize the mismatch between the data and the emitter PSF by varying the parameters of the emitter [4]. We seek the parameters of the emitter that minimize the least-squares error:

$$\hat{\theta}^{(i)} = \underset{\theta}{\operatorname{argmin}} \sum_{(m,n)\in\mathcal{S}^{(i)}} (y[m,n] - \mu[m,n|\{\theta\}])^2, \tag{3.19}$$

where $\mathcal{S}^{(i)}$, as defined in (3.14), is the set of pixels in the extracted square region of $i^{\text{th}}$ emitter. This minimization problem can be solved by many optimization routines efficiently, such as the widely used Levenberg-Marquadt algorithm [48, 57]. However, ordinary least-squares fitting assumes a uniform variance in the errors (homoscedasticity). As described earlier, for photon count images, the noise for each pixel follows a Poisson distribution where the variance is the expected number of photons hitting each pixel, causing a nonuniform noise variance (heteroscedasticity). Therefore, a weighted least-squares minimization can be used:

$$\hat{\theta}^{(i)} = \underset{\theta}{\operatorname{argmin}} \sum_{(m,n)\in\mathcal{S}^{(i)}} \frac{(y[m,n] - \mu[m,n|\{\theta\}])^2}{\sigma^2_{(m,n)}}, \tag{3.20}$$

where $\sigma^2_{(m,n)}$ is the variance of the pixel $(m,n)$, typically estimated as $\max(y[m,n],1)$ to avoid division by zero. One can also substitute the denominator with $\mu[m,n|\{\theta\}]$.

If a pixel value is expected to have a large variance, then the squared error of that pixel should have less contribution to the total error than the error of a pixel whose variance is small. Therefore, more emphasis is put on the tail of PSF.

**Radial Symmetry**

In another study, Parthasarathy [58] utilizes the radially symmetric property of system PSF and determines the position of an emitter as the point of maximal radial symmetry. For each pixel of a radially symmetric image without noise, its gradient will point towards the center. For noisy images, the radial center is estimated to be the point that has the smallest total distance to the intensity gradients of all pixels. Parthasarathy provides an analytic solution to this problem without iterative fitting, so that this method is 100 times faster than fitting-based methods. However, it can not be readily applied to 3D emitter localization where both lateral position and the axial position of emitters need to be estimated.

### 3.3.3 Filtering

Since the sparse activation of emitters is a stochastic process, there may be multiple nearby fluorophores activated at the same time, resulting in overlapping PSFs. In such a scenario, their PSFs will be regarded as a single peak and only one position will be returned, resulting in a higher localization error. This can be reduced by adding a subsequent filtering step where the width of PSF returned by the localization algorithm can be used as a criteria to filter out overlapping emitters. Other criteria also exist such as the symmetry level $S_x$ and $S_y$ calculated in (3.18). Out-of-focus emitters can also be filtered out at this step by using the fitted width and background.

## 3.4 CSSTORM

Despite its simplicity and efficiency, SEA can only handle images of low emitter density ($< 1$ emitter/$\mu m^2$) and suffers from low temporal resolution. With low emitter density images, tens of thousands of frames are needed to achieve a super-resolution image of 20 to 50 nm resolution. This greatly limits its application in live cell imaging and other fast dynamic processes, where the temporal resolution required is often within one minute. In order for faster super-resolution image acquisition, algorithms that can handle higher density than SEA are in urgent need. Several algorithms have been proposed to analyze emitters of intermediate to high density [7, 59–61]. Among these algorithms, compressive sensing based reconstruction algorithm, also known as CSSTORM [59], utilizes the sparsity prior of the signal in each frame and achieves the state-of-the-art recall rate and localization accuracy when the density is high ($>5$ emitters/$\mu m^2$).

### 3.4.1 Background

The Nyquist-Shannon theorem states that the sampling rate should be at least twice the maximum bandwidth of the signal to avoid losing information. In 2004, Emmanuel Candès, Justin Romberg, Terence Tao [62] and David Donoho [63] proposed that given the signal is sparse in certain basis, it can be reconstructed with much less number of measurements than what the Nyquist-Shannon theorem requires. This sampling technique is known as compressive sensing (CS).

**Sparsity**

Let a real-valued, one-dimensional vector $x \in \mathbb{R}^{n_1}$ be the signal of interest and $\{\psi_i\}_{i=1}^{n_1}$ be a set of orthonormal basis vectors, where $\psi_i \in \mathbb{R}^{n_1}, 1 \leq i \leq n_1$. A basis matrix $\Psi \in \mathbb{R}^{n_1 \times n_1}$ is defined as $\Psi = [\psi_1, \psi_2, \cdots, \psi_{n_1}]$ and can be used to represent $x$ as

$$x = \Psi s, \tag{3.21}$$

where $s = [s_1, \cdots, s_{n_1}]^T \in \mathbb{R}^{n_1}$ with $s_i = \langle x, \psi_i \rangle$. Therefore, $s$ can be seen as a representation of $x$ in the $\Psi$ domain. If at most $K$ coefficients of $s$ are non-zero, $x$ is defined as $K$-sparse and can be represented as a linear combination of only $K$ basis vectors of $\{\psi_i\}_{i=1}^{n_1}$. We are interested in the case when $K \ll n_1$.

In practice, the signal of interest may not be strictly $K$-sparse. Instead, the coefficients of $s$ may decay rapidly, having only a few large coefficients and many small coefficients. In such a case, we define $s_K$ as the best $K$-term approximation of $s$:

$$s_K = \underset{z \in \Sigma_K}{\operatorname{argmin}} \|s - z\|_1 \tag{3.22}$$

and $x_K := \Psi s_K$, where $\Sigma_K = \{z \in \mathbb{R}^{n_1} : |\text{supp}(z)| \leq K\}$ and $\text{supp}(z)$ is the support of $z$. Since $\Psi$ is an orthonormal basis, we have $\|x_K - x\|_2 = \|s_K - s\|_2$, where $\|\cdot\|_2$ represents the $\ell_2$ norm. Therefore, the approximation error of $x_K$ is negligible as long as $\|s_K - s\|_2$ is small enough. If $x$ is $K$-sparse or the coefficients of $x$ in the $\Psi$ domain decay quickly (in other words, compressible), it can be represented without much information loss by preserving only the large coefficients and their locations of $s$. This yields a sparse approximation of the signal $x$, an idea underlying many compression algorithms such as JPEG, JPEG2000, MP3 and MPEG. Figure 3.3 shows

the comparison between an original image and its wavelet representation. Most of the wavelet coefficients are close to zero, demonstrating a high compressibility in the wavelet domain.

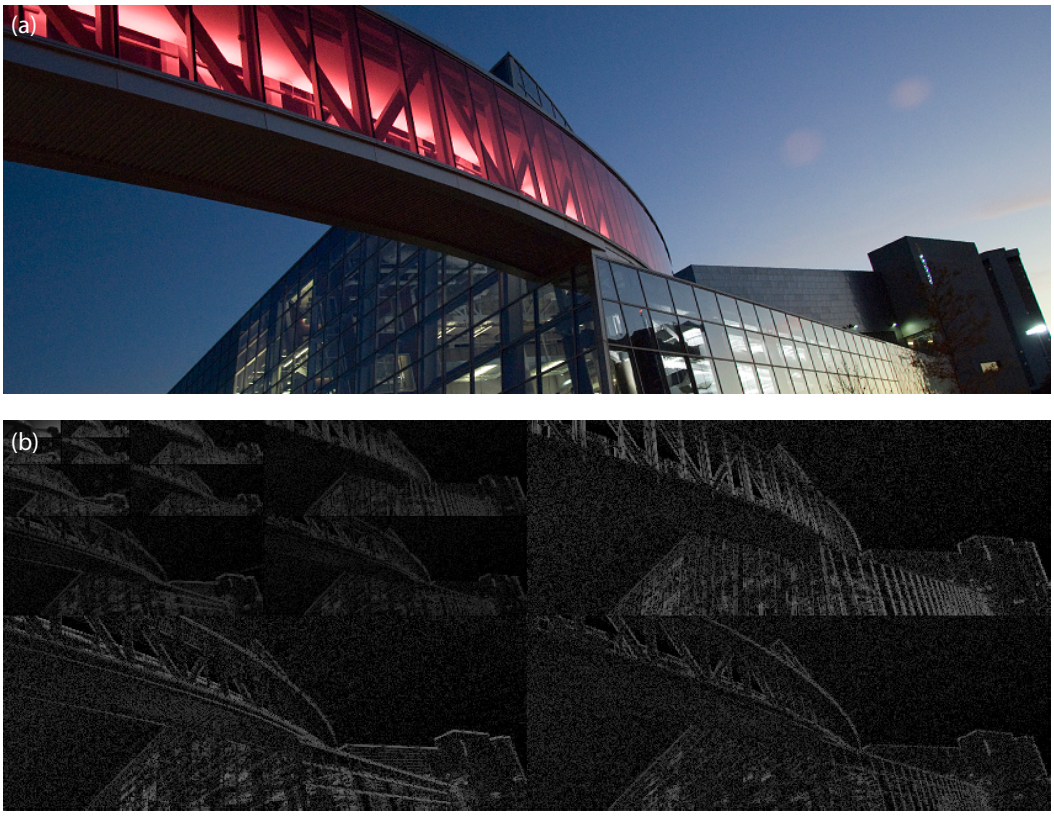



Figure 3.3: Illustration of the sparse representation. (a) Original image. (b) Wavelet transform.

### Noiseless Measurements

Consider a linear measurement system that collects $n_2 < n_1$ measurements $\{y_j\}_{j=1}^{n_2}$, where $y_j = \langle \phi_j, x \rangle, 1 \leq j \leq n_2$. Let $\Phi = [\phi_1, \phi_2, \cdots, \phi_{n_2}]^T \in \mathbb{R}^{n_2 \times n_1}$ and $y \in \mathbb{R}^{n_2}$

where $y = [y_1, \cdots, y_{n_2}]^T$, we have

$$y = \Phi x = \Phi \Psi s = As, \tag{3.23}$$

where $A = \Phi\Psi \in \mathbb{R}^{n_2 \times n_1}$. The goal of CS, therefore, is to infer $s$ given the knowledge of $y$ and $A$. This is very common in practice such as when the sensors are expensive and therefore limited, or when the sensing process is slow as in magnetic resonance imaging. Since $A$ is a matrix of size $n_2 \times n_1$, this problem is underdetermined: there are an infinite number of $\tilde{s}$ that satisfy $y = A\tilde{s}$.

### $\ell_1$ minimization

Given the prior knowledge that $s$ is sparse, probably the most intuitive solution to (3.23) is to find the sparsest $\tilde{s}$ that satisfies $y = A\tilde{s}$, i.e.,

$$\hat{s} = \underset{\tilde{s}}{\operatorname{argmin}} \|\tilde{s}\|_0 \quad \text{subject to } y = A\tilde{s}, \tag{3.24}$$

where $\|\cdot\|_0$ is the $\ell_0$ norm and is calculated as the number of nonzero elements. In principle, an impressively small number of measurements is required to achieve perfect reconstruction under mild condition on $\Phi$ and $\Psi$. In fact, only $K + 1$ random measurements are needed for reconstructing a $K$-sparse signal with $\ell_0$ minimization [64]. Unfortunately, solving $\ell_0$ minimization (3.24) is a combinatorial problem and computationally inefficient. It is known to be NP-complete [65], requiring enumerating $\binom{n_1}{K}$ possible locations of the nonzero entries of $s$.

On the other hand, minimization of the $\ell_1$ norm has been used as a sparsity-promoting function for decades [66, 67]. Substituting the $\ell_0$ norm with the $\ell_1$ norm yields the following minimization problem:

$$\hat{s} = \underset{\tilde{s}}{\operatorname{argmin}} \|\tilde{s}\|_1 \quad \text{subject to } y = A\tilde{s}. \tag{3.25}$$

This algorithm is also known as Basis Pursuit [68] and can be solved efficiently via linear programming. The geometric interpretation of using the $\ell_1$ norm versus the more commonly used $\ell_2$ norm as in least-squares reconstruction is shown in Figure 3.4. We define the solution space $\mathcal{H} := \mathcal{N}(A) + s$, where $\mathcal{N}(A)$ is the null space of $A$. Therefore, $y = A\tilde{s}$ for any $\tilde{s} \in \mathcal{H}$. It can be seen that the $\ell_1$ norm ball is pointing towards the coordinate axes. Therefore, when the value of it increases, it has a higher chance contacting the solution space at the axes, resulting in a sparser solution than $\ell_2$ norm.

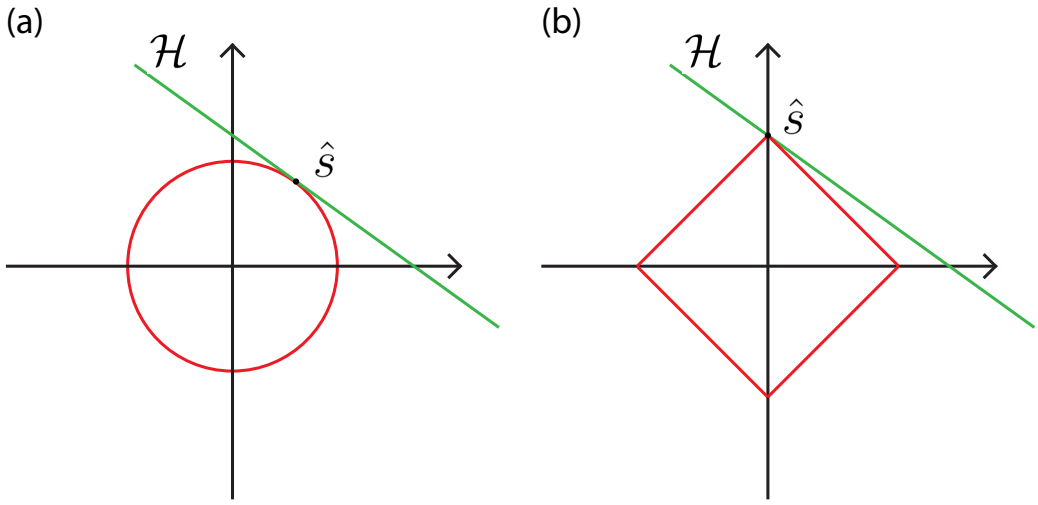



Figure 3.4: Comparison of (a) $\ell_2$ minimization and (b) $\ell_1$ minimization solutions. The green line $\mathcal{H}$ is the solution space and the red line is the $\ell_1$ or $\ell_2$ norm ball.

### Restricted Isometry Property

To evaluate the reconstruction performance using $\ell_1$ minimization (3.25), a notion called Restricted Isometry Property (RIP) [69] is introduced. For $K = 1, 2, \cdots, n_1$,

the isometry constant $\delta_K$ of a matrix $A$ is defined as the smallest number that satisfies

$$(1-\delta_K)\|s\|_2^2 \leq \|As\|_2^2 \leq (1+\delta_K)\|s\|_2^2 \tag{3.26}$$

for all $K$-sparse vectors $s$. The value of $\delta_{2K}$ is of special interest and if $\delta_{2K}$ is not close to 1, $A$ preserves the Euclidean length of any $2K$-sparse signals. This implies that the distance between any two $K$-sparse vectors is well preserved after measurements, i.e., for all $K$-sparse vectors $s_1$ and $s_2$, we have $(1-\delta_K)\|s_1-s_2\|_2^2 \leq \|As_1-As_2\|_2^2 \leq (1+\delta_K)\|s_1-s_2\|_2^2$. This is important for reliable reconstruction of (3.25) as described below.

Assume $\delta_{2K} < \sqrt{2}-1$, it can be shown that [69] the solution $\hat{s}$ to the $\ell_1$ minimization problem (3.25) obeys

$$\begin{aligned} \|\hat{s}-s\|_1 &\leq C_0\|s-s_K\|_1, \\ \|\hat{s}-s\|_2 &\leq C_0 K^{-1/2}\|s-s_K\|_1, \end{aligned} \tag{3.27}$$

for some constant $C_0$, where $s_K$ is the best $K$-sparse approximation of $s$ as described earlier. Therefore,

1. if $s$ is $K$-sparse, i.e., $s = s_K$, and in combination with (3.27), we have exact recovery because $\hat{s} = s$;

2. if $s$ is not $K$-sparse, the $\ell_1$ error of reconstruction is linearly bounded by the $\ell_1$ error of the best $K$-sparse approximation. For $\ell_2$ error, the bound is scaled down by $K^{1/2}$.

**Noisy Measurements**

In practice, the measurements are typically corrupted with noise such as shot noise or white noise. Consider the measurement system of the form:

$$y = As + z, \tag{3.28}$$

where $z$ is unknown but bounded noise, with $\|z\|_2 \leq \epsilon$.

The constraint in the $\ell_1$ minimization problem (3.25) can be relaxed to accommodate the noise term introduced in (3.28). Therefore, we have the reconstruction problem

$$\hat{s} = \underset{\tilde{s}}{\operatorname{argmin}} \|\tilde{s}\|_1 \quad \text{subject to } \|y - A\tilde{s}\|_2 \leq \epsilon, \tag{3.29}$$

where $\epsilon$ is an upper bound on the noise level. This problem can be solved efficiently with linear programming.

Again, assume $\delta_{2K} < \sqrt{2} - 1$, it has be shown [69] that the solution to the optimization problem (3.29) obeys

$$\|\hat{s} - s\|_2 \leq C_0 K^{-1/2} \|s - s_K\|_1 + C_1 \epsilon, \tag{3.30}$$

for some constants $C_0$ and $C_1$. Compared with the bound in the noiseless case (3.27), only a term proportional to the noise level is added. Besides, the constants $C_0$ and $C_1$ are typically very small. This result guarantees the robustness of $\ell_1$ minimization for all sparse and compressible signals when the measurement matrix obeys the RIP condition.

The natural question to raise is, how to design the measurement matrix so the RIP condition holds? One way is to design the measurement matrix randomly. One can choose the element of $\Phi$ by sampling independent and identically distributed random variables from a normal distribution with zero mean and variance $1/n_1$. Then with an overwhelming probability, the matrix $A$ obeys the RIP condition given that [63,70,71]

$$n_2 \geq C \cdot K \log(n_1/K), \tag{3.31}$$

where $C$ is a small constant.

In summary, the theory of compressive sensing states that a signal can be reconstructed robustly from only $\mathcal{O}(K\log(n_1/K))$ measurements by finding sparse solutions to underdetermined linear systems. The recovery is possible under two conditions: (a) the signal is sparse in some domain; and (b) the measurements are highly incoherent with respect to the representation domain, which can be enforced by the RIP condition.

### 3.4.2 Formulation in Super Resolution

In SMLM, only a sparse subset of emitters are activated and the image acquired at the camera is a linear combination of the sparsely activated emitters' PSFs. This can be formulated as inverting an underdetermined linear system where the sparsity of the signal is exploited. The goal of CSSTORM for SMLM image reconstruction, is to robustly recover the emitters' locations under high emitter density.

In CS, the notion of sparsity is essential and generally the signal is sparse with respect to a finite dictionary. In SMLM, the observed signal $y$ is a mixture of a few continuously parameterized PSFs. Thus $y$ has a sparse representation with respect to a continuously-parameterized dictionary, where the dictionary contains the PSFs located at different positions. A straightforward strategy to solve the sparse recovery problem with a continuously parameterized dictionary is to discretize the parameter space [72]. As first described in [59], a discrete grid with spacing $\triangle$ much smaller than the camera pixel size is introduced to represent the discrete parameter space. In 2D case, we define the grid locations as a finite set $\mathcal{S}_m$:

$$\mathcal{S}_m = \Big\{\theta_L = [\theta_x, \theta_y] \Big| \theta_x \in \mathcal{S}_x,\ \theta_y \in \mathcal{S}_y \Big\}, \tag{3.32}$$

where

$$\mathcal{S}_x = \{x_{\min}, x_{\min} + \triangle, x_{\min} + 2\triangle, \cdots, x_{\max}\}, \tag{3.33}$$

$$\mathcal{S}_y = \{y_{\min}, y_{\min} + \triangle, y_{\min} + 2\triangle, \cdots, y_{\max}\}. \tag{3.34}$$

Then we define $x \in \mathcal{R}^{|\mathcal{S}_x||\mathcal{S}_y| \times 1}$ as the emitter signal, where each entry of $x$ representing the brightness of an emitter located at the point. Since only a sparse number of emitters are active at a time, $x$ is sparse. The observed camera image $y$ is a linear measurement of the underlying sparse emitters that locate approximately on the discrete grid. Mathematically,

$$y \sim \text{Poisson}(Ax). \tag{3.35}$$

Poisson$(\cdot)$ represents the Poisson distribution with the variance determined to be its parameter. The matrix $A$ is the column-wise concatenation of shifted PSF at corresponding locations on the discrete grid. Define the location of an emitter corresponding to the $j^{\text{th}}$ entry of $x$ as $\theta_L^j$, we have the $j^{\text{th}}$ column of $A$ determined as $A_j = \{\mu[m, n|\{\theta_L^j, 1\}]\}$, where $\mu(\cdot)$ is defined in (3.5). Thus the signal $x$ can be reconstructed by solving the following $\ell_1$ minimization problem:

$$\begin{gathered} \text{minimize } \|x\|_1 \\ \text{subject to } x \geq 0 \\ \|y - Ax\|_2 \leq \xi \cdot \sqrt{\|y\|_1}, \end{gathered} \tag{3.36}$$

where $\xi$ is a regularization parameter that controls the trade-off between the fidelity to the observation and the sparsity of the signal. The quadratic constraints are introduced to accommodate the inherent Poisson noise. A large $\xi$ enhances the sparsity level of $x$ and allows larger deviation from the acquired data, while a small $\xi$ results

in higher fidelity to the acquired data and less sparse $x$. In practice, a value of $\xi$ between 1.5 and 2 is used. The optimization problem (3.36) can be solved using linear programming with quadratic constraints [59].

Unfortunately, the matrix $A$ does not satisfy the RIP due to the high correlations between its columns. In [59], a simplified case is considered to estimate the limit of emitter density for robust reconstruction. It is shown that the smallest distance of two emitters for them to be "perfectly" detected is $\sim$2.5 $\sigma$, resulting in a maximal emitter density of $\sim$7.8 $\mu m^{-2}$.

CSSTORM has been proven as a robust and high-performance algorithm for high-density super-resolution image reconstruction [59]. It allows an order of higher emitter density than what the SEA algorithm can handle and greatly increases its temporal resolution. It also achieves a significant improvement of the localization precision. The principle of CSSTORM can also be applied to 3D SMLM image reconstruction, where a 3D grid can be introduced to represent emitters locations.

# Chapter 4: 2D Fast Reconstruction Algorithm via Frequency Estimation

As described in the last chapter, CSSTORM achieves state-of-the-art recall rate and precision when it was first proposed. However, it requires solving a large-scale optimization problem and suffers from high computational complexity. It also suffers from the intrinsic gridding error as it discretizes the 2D space. We notice that the mathematical formulation of the imaging process of super-resolution microscopy can be transformed into the frequency domain and the problem of emitter localization can be turned into a 2D frequency estimation problem. We have developed a method, dubbed MempSTORM [11], based on a 2D spectrum estimation method called Matrix Enhancement and Matrix Pencil (MEMP) [73] and extract the number of emitters and positions by estimating the 2D frequencies. By verification through simulations and experiments, this method is proven to handle the same level of density as CSSTORM and at the same time achieves much faster computation speed. The most time-consuming computations for the proposed method are a truncated singular value decomposition (SVD) and two generalized eigenvalue decomposition operations.

## 4.1 Signal Model

We start with the model of 2D SMLM. Recall that $\Theta = \{\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(P)}\}$ is the set of parameters, where $P$ is the total number of emitters and $\theta^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_I^{(i)}] = [\theta_L^{(i)}, \theta_I^{(i)}]$ is the parameters for the $i^{\text{th}}$ emitter. We then rewrite the signal $s(x, y|\chi)$ received at camera position $(x, y)$ as

$$s(x, y|\chi) = \sum_{i=1}^{P} \theta_I^{(i)} f(x - \theta_x^{(i)}, y - \theta_x^{(i)}), \tag{4.1}$$

where $f(\cdot)$ is the system PSF. The 2D Continuous-Time Fourier Transform (CTFT) of the image can be written as:

$$\mathcal{S}(u, v|\chi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y|\chi) e^{-j2\pi(xu+yv)} dxdy \tag{4.2}$$

$$= \mathcal{F}(u, v) \sum_{i=1}^{P} \theta_I^{(i)} e^{-j2\pi(\theta_x^{(i)}u + \theta_y^{(i)}v)}, \tag{4.3}$$

where $\mathcal{F}(u, v)$ is the CTFT of $f(x, y)$.

The image $\mu = \{\mu[m, n|\chi]\}$ acquired by the camera is a matrix of size $M \times N$, where $\mu[m, n|\chi]$ represents the intensity value of the pixel at $(m, n)$ obtained by integrating $s(x, y|\chi)$ over the pixel area, approximated as:

$$\mu[m, n|\chi] = \int_{m-0.5}^{m+0.5} \int_{n-0.5}^{n+0.5} s(x, y|\chi) dxdy \approx G \cdot s(m, n|\chi), \tag{4.4}$$

where $1 \leq m \leq M$, $1 \leq n \leq N$, and $G$ is the area of a pixel. The Discrete Fourier Transform (DFT) $\Gamma = \{\Gamma[k, l]\}$ of $\mu$ can be approximated as

$$\Gamma[k, l] \approx F[k, l] \sum_{i=1}^{P} G\theta_I^{(i)} e^{-j2\pi(\frac{k\theta_x^{(i)}}{M} + \frac{l\theta_y^{(i)}}{N})}, \tag{4.5}$$

where $F[k, l] = \mathcal{F}(k/M, l/N)$, for $1 \leq k \leq M$ and $1 \leq l \leq N$. Denote $f_{1i} = \frac{\theta_x^{(i)}}{M}$, $f_{2i} = \frac{\theta_y^{(i)}}{N}$, and divide both sides of (4.5) by $F[k, l]$, we obtain

$$R[k, l] = \frac{\Gamma[k, l]}{F[k, l]} \approx \sum_{i=1}^{P} G\theta_I^{(i)} e^{-j2\pi(kf_{1i} + lf_{2i})}. \tag{4.6}$$

After these transformations, the problem of emitter localization becomes a 2D frequency estimation problem in (4.6). We apply the MEMP method [73] for the frequency estimation step, considering its low computational cost and capability to resolve high-density signals.

## 4.2 Matrix Enhancement and Matrix Pencil

We further rewrite (4.6) into

$$R[k,l] \approx \sum_{i=1}^{P} G\theta_I^{(i)} b_i^k d_i^l, \tag{4.7}$$

where $b_i = e^{-j2\pi f_{1i}}$, $d_i = e^{-j2\pi f_{2i}}$ are called the 2D poles. With this notation we can write $R = \{R[k,l]\}$ as an $M \times N$ matrix with the following factorization:

$$R \approx BCD, \tag{4.8}$$

where

$$B = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_P \\ \vdots & \vdots & \ddots & \vdots \\ b_1^{M-1} & b_2^{M-1} & \cdots & b_P^{M-1} \end{bmatrix} \in \mathbb{C}^{M\times P}, \tag{4.9}$$

$$C = G \cdot \begin{bmatrix} \theta_I^{(1)} & 0 & \cdots & 0 \\ 0 & \theta_I^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \theta_I^{(P)} \end{bmatrix} \in \mathbb{R}^{P\times P}, \tag{4.10}$$

$$D = \begin{bmatrix} 1 & d_1 & \cdots & d_1^{N-1} \\ 1 & d_2 & \cdots & d_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & d_P & \cdots & d_P^{N-1} \end{bmatrix} \in \mathbb{C}^{P\times N}. \tag{4.11}$$

From (4.8)-(4.11), in principle one can obtain $\{b_i; i = 1, \cdots, P\}$ and $\{d_i; i = 1, \cdots, P\}$ from the left and right principal singular vectors of $R$. However, $b_i$ and $d_i$ cannot be obtained from the SVD when either set of $\{b_i; i = 1, \cdots, P\}$ or $\{d_i; i =$

$1, \cdots, P\}$ does not contain distinct elements, due to the rank deficiency of $R$, i.e. rank($R$)$< P$.

To solve this problem, a partition-and-stacking process, referred to as matrix enhancement [73], is applied before SVD. The enhanced matrix $R_e$ is defined as a block Hankel matrix of size $K \times (M - K + 1)$:

$$R_e = \begin{bmatrix} R_0 & R_1 & \cdots & R_{M-K} \\ R_1 & R_2 & \cdots & R_{M-K+1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{K-1} & R_K & \cdots & R_{M-1} \end{bmatrix}, \tag{4.12}$$

where each block $R_m$, $0 \leq m \leq M-1$, is a Hankel matrix of size $L \times (N - L + 1)$ defined as:

$$R_m = \begin{bmatrix} R[m,0] & R[m,1] & \cdots & R[m,N-L] \\ R[m,1] & R[m,2] & \cdots & R[m,N-L+1] \\ \vdots & \vdots & \ddots & \vdots \\ R[m,L-1] & R[m,L] & \cdots & R[m,N-1] \end{bmatrix}.$$

It is shown in [73] that rank($R_e$)=$P$ as long as the two pencil parameters $K$ and $L$ satisfy

$$M - I + 1 \geq K \geq I, \text{ and } N - I + 1 \geq L \geq I, \tag{4.13}$$

eliminating the earlier described rank deficiency issue.

In the noise-free case, the SVD of $R_e$ can be given as:

$$R_e = \sum_{i=1}^{P} \sigma_i u_i v_i^H = U \Sigma V^H, \tag{4.14}$$

where $U = [u_1, \cdots, u_P] \in \mathbb{C}^{KL \times P}$, $\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_P) \in \mathbb{R}^{P \times P}$, $V = [v_1, \cdots, v_P] \in \mathbb{C}^{(M-K+1)(N-L+1) \times P}$, respectively. In the case when $R_e$ is noisy, we can similarly define $U$, $\Sigma$ and $V$ as the top $P$ left singular vectors, singular values, and right singular vectors of $R_e$.

Denote the $i^{\text{th}}$ column of $U^H$ as $w_i$, then $U^H = [w_1, w_2, \cdots, w_{KL}]$. Further define $U_p \in \mathbb{C}^{KL \times P}$ by permuting the rows of $U$ as:

$$\begin{aligned} U_p^H = [&w_1, w_{L+1}, \cdots, w_{(K-1)L+1}, \\ &w_2, w_{L+2}, \cdots, w_{(K-1)L+2}, \cdots, \\ &w_L, w_{2L}, \cdots, w_{KL}]. \end{aligned} \tag{4.15}$$

Define $U_1 \in \mathbb{C}^{(K-1)L \times P}$ as the submatrix of $U$ by deleting its last $L$ rows, $U_1^H = [w_1, \cdots, w_{(K-1)L}]$, and $U_2 \in \mathbb{C}^{(K-1)L \times P}$ as the submatrix of $U$ by deleting its first $L$ rows, $U_2 = [w_{L+1}, \cdots, w_{KL}]$, then the matrix pencil $U_2 - \lambda U_1$ can be written as [73]:

$$U_2 - \lambda U_1 = E(Y_d - \lambda I)T, \tag{4.16}$$

where $Y_d = \text{diag}(b_1, \cdots, b_P)$ is a diagonal matrix of the poles $\{b_i; i = 1, \cdots, P\}$, $E$ and $T$ are full-rank matrices. Thus the poles $\{b_i; i = 1, \cdots, P\}$ can be found as the rank-reducing numbers $\lambda$ such that the matrix pencil $U_2 - \lambda U_1$ is rank-deficient. This can be achieved by solving a generalized eigenvalue decomposition.

Similarly, $\{d_i; i = 1, \cdots, P\}$ can be found as the rank-reducing numbers of the matrix pencil $U_4 - \lambda U_3$, where $U_3 \in \mathbb{C}^{K(L-1) \times P}$ is the submatrix of $U_p$ by deleting its last $K$ rows, $U_3^H = [w_1, w_{L+1}, \cdots, w_{KL-1}]$, and $U_4 \in \mathbb{C}^{K(L-1) \times P}$ is the submatrix of $U_p$ by deleting its first $K$ rows, $U_4^H = [w_2, w_{L+2}, \cdots, w_{KL}]$. This again can be achieved by solving a generalized eigenvalue decomposition problem.

After extracting $\{b_i; i = 1, \cdots, P\}$ and $\{d_i; i = 1, \cdots, P\}$ separately, we need to further determine the one-to-one (or one-to-more, if there're multiplicity in the retrieved poles) correspondence between the poles to fully determine the set of 2D poles. First, we consider all the possible pairings between $\{b_i; i = 1, \cdots, P\}$ and $\{d_i; i = 1, \cdots, P\}$, with a total of no more than $P^2$ pairs. We then pose a non-negative least-squares problem, by minimizing the quadratic loss between the data

matrix $R$ and the fitting using all possible pairs, given in (4.8), under the constraint that the coefficient of each pair is non-negative. We then select $P$ pairs corresponding to the highest coefficients. From the paired 2D poles $\{(b_i, d_i); i = 1, \cdots, P\}$, the set of 2D frequencies $(f_{1i}, f_{2i})$ as well as the positions $(\theta_x^{(i)}, \theta_y^{(i)})$ of the $P$ emitters can be calculated.

In implementing the above MempSTORM method, a pair of pencil parameters $K$ and $L$ need to be chosen for matrix enhancement. (4.13) is a sufficient condition for the rank of enhanced matrix $R_e$ to be $P$. Under this condition, the rule of thumb is to choose $K$ and $L$ such that the enhanced matrix $R_e$ is as square as possible, i.e., choose $K$ to be close to $(M + 1)/2$ and $L$ to be close to $(N + 1)/2$. Moreover, since the number of emitters $I$ is not known a priori, we choose a threshold value and determine the number of emitters as the number of singular values of $R_e$ that is larger than the threshold. The threshold value is chosen such that the sum energy of the selected singular vectors is 80-90% of the total.

For super-resolution image reconstruction, the noise in the frequency domain has similar energy across different frequencies due to the Poisson noise in the spatial domain. However, the energy of the signal is not uniformly distributed in the frequency domain due to the shape of the PSF. Instead, most of the energy is concentrated in the low frequency with the high frequency components dominated by noise. In practice, we only use entries that correspond to the low frequency region of $R$ and apply MEMP on a truncated data matrix. For an image of size $30 \times 30$, we only use the submatrix of size $11 \times 11$ of (4.6) that corresponds to the low-frequency region of the image to generate the matrix enhancement.

## 4.3 Performance

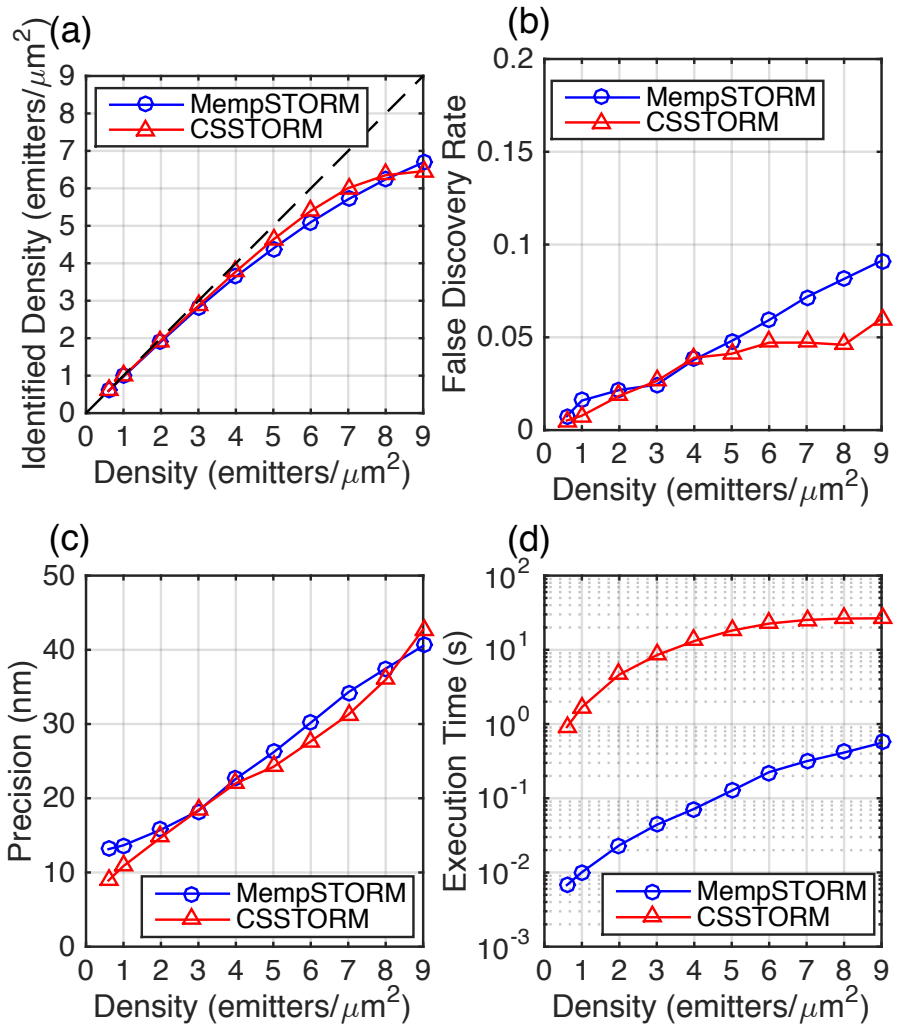



Figure 4.1: Performance comparisons of (a) the identified density, (b) false discovery rate, (c) precision and (d) execution time between MempSTORM and CSSTORM under different emitter densities.

To evaluate the performance of MempSTORM, we generate a series of simulated STORM movies of size $30 \times 30$ across a range of emitter densities (0.5 emitter/$\mu m^2$ to 9 emitters/$\mu m^2$). An average photon number of 1500 per emitter is used in

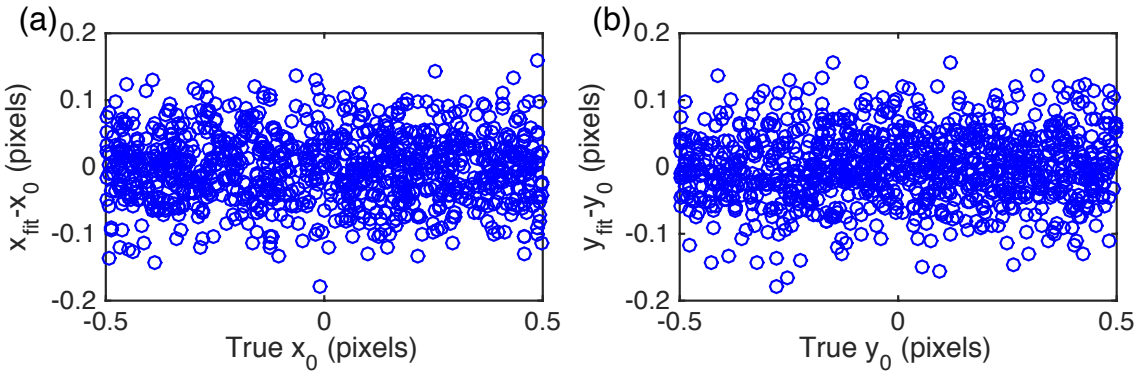



Figure 4.2: The bias of single emitter localization using MempSTORM in (a) $x$ direction and (b) $y$ direction.

the simulation. The images are comparatively analyzed using MempSTORM and CSSTORM [59]. We implement CSSTORM using $\ell_1$-homotopy, which has been proven to be two orders of magnitude faster than CVX implementation [74]. An up-sampling factor of 8 is used in CSSTORM. As shown in Fig. 4.1, MempSTORM achieves the same level of identified density and false discovery rate as CSSTORM, and maintains similar precision with CSSTORM. The false discovery rate is defined as the percentage of incorrectly detected emitters in all detected emitters. Most importantly, MempSTORM is more than 100 times faster than CSSTORM in most densities (Fig. 4.1 (d)).

To examine the bias of MempSTORM, we simulate a stack of images of size $30 \times 30$ with a single emitter positioned randomly. The photon number of the emitter is set to be 1500. The localization bias is plotted as a function of the emitter location in both $x$ and $y$ directions. As shown in Fig. 4.2, there is no obvious bias associated with the MempSTORM method.

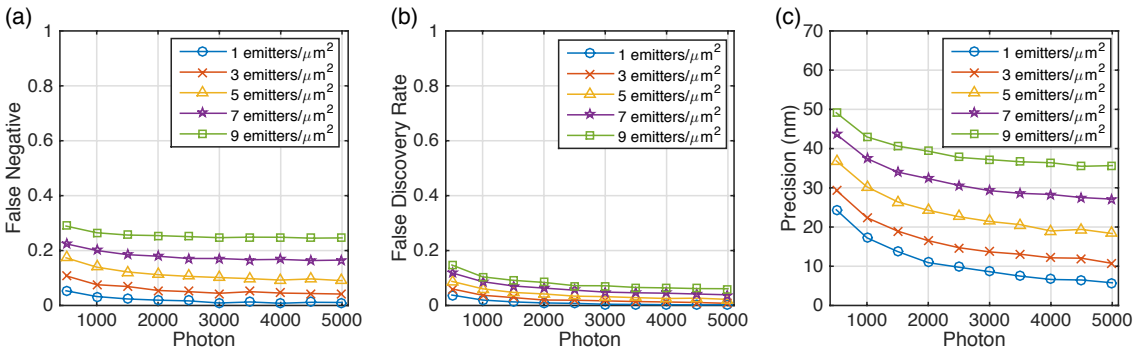



Figure 4.3: Performance comparisons of (a) false negative, (b) false discovery rate and (c) precision when the emitter photon varies under different densities.

We then evaluate the performance of MempSTORM when the number of photons varies. A series of STORM images with different emitter photon numbers are generated and analyzed. The performance of MempSTORM deteriorates when the number of photons decreases, but within a tolerable range (Fig. 4.3). Even when the photon number is as low as 500, MempSTORM can still detect a reasonable amount of emitters with high precision (<50 nm).

To further test the practical applicability of MempSTORM, we analyze STORM images of microtubules stained with Alexa 647 in HeLa cells. In total, 5000 frames are acquired at 56 frames per second. Figure. 4.4 shows the comparison of reconstructed microtubule images using the MempSTORM and CSSTORM methods, where MempSTORM achieves similar performance with CSSTORM. However, the reconstruction time using MempSTORM is only 20 minutes; while with the same computational configurations, CSSTORM requires more than 100 hours.

## 4.4 Summary

In summary, we have developed the MempSTORM method for super-resolution image reconstruction with high emitter densities. MempSTORM converts the problem of emitter localization into frequency estimation and is computational efficient. It achieves the same performance as the current state-of-the-art method, CSSTORM, but is more than 100 times faster.

The experiments in this chapter were performed on the compute nodes of the Ohio Supercomputing Center Oakley cluster. Each node has one Intel Xeon X5650 CPU clocked at 2.66GHz and 48 GB of memory. Currently, the algorithm is implemented in MATLAB. By porting the algorithm to C/C++ and using GPU parallelism, we expect to further accelerate the analysis and put online reconstruction of dense images into practice.

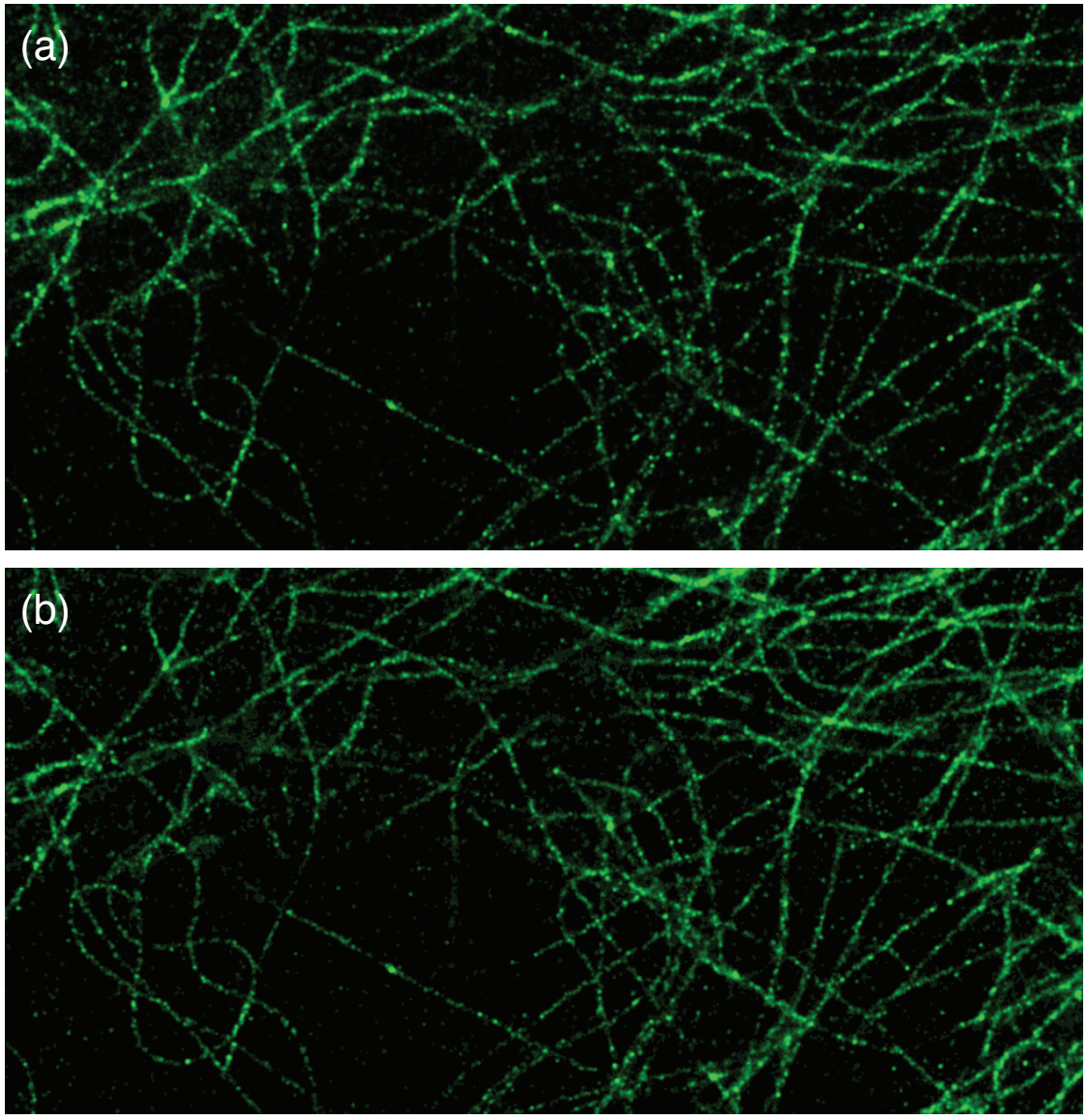



Figure 4.4: Reconstructed microtubule image using 5000 raw frames with (a) MempSTORM and (b) CSSTORM.

# Chapter 5: 3D Multi-Focus Astigmatism and Compressive Sensing based Super-Resolution Image Reconstruction

For 3D super-resolution imaging, the challenge arises from the fact that the 3D PSF is extended along the axial direction, and the PSFs of nearby emitters have more overlapping regions and similarities, causing ambiguity to distinguish the emitter $z$ locations. Traditional 3D image reconstruction algorithm follows a similar strategy as SEA algorithm. First, isolated emitters are extracted and fitted to a Gaussian function on a small area surrounding the candidate emitters. The axial location of the emitter is then determined according to the fitted width $\sigma_x$ and $\sigma_y$ of Gaussian function using the calibration curve. Such procedure is repeated for all the frames and is referred to as 3D-SEA. 3D-SEA suffers the same shortcoming as SEA as it can only handle images of low emitter density.

In this chapter, we propose a multi-focus astigmatism compressive sensing based 3D imaging platform (3D MACS) [12] that allows us to efficiently locate high-density emitters. A multi-focus technique, in combination with astigmatic optics is applied to enhance the asymmetry of the PSF along the axial direction. Accordingly, we develop a new algorithm which takes advantage of compressive sensing with a fast $\ell_1$-Homotopy optimization method. More importantly, a debiasing method is developed, which not only reduces the bias introduced by gridding, an intrinsic drawback of

compressive sensing based methods, but also significantly increases the computational speed. With 3D MACS, we are able to reconstruct super-resolution images with as high as 10 emitters/$\mu m^2$ density and reconstruct 3D microtubule images within a couple of seconds, which makes this method suitable for live cell imaging.

## 5.1 Imaging System

Our 3D multi-plane astigmatism STORM system (Fig. 5.1(a)) is based on the setup introduced in Section 2.7. Three identical EMCCD cameras (iXon Ultra 897, Andor Technologies PLC.) are used to image three different focus planes with all of the EM gains set to 255 and the distance between each focus plane is 180 nm. The three cameras are synchronized by an external hardware trigger. Three identical plano-convex cylindrical lenses (Thorlabs) with focus length of 1000 mm are placed on the emission paths to introduce asymmetry to the 3D PSF. Such system can be easily transformed into a single-focus astigmatism imaging system (3D SACS) without change of the reconstruction algorithm.

## 5.2 3D Astigmatic PSF and Multi-Plane Alignment

To acquire the calibration curves of all the three cameras, we imaged the 100 nm fluorescent beads at different $z$ positions by moving the piezo stage along the axial direction. Figure 5.1(b) shows images of the beads where each camera is focused on a different $z$ plane. For each focus plane, the bead images are fitted to an asymmetric 2D Gaussian function (3.9) to acquire $\sigma_x$ and $\sigma_y$.

To align the three cameras, we took images of fluorescent beads that are sparsely attached to a coverslip. For a typical 512 $\times$ 512 image, the average number of beads

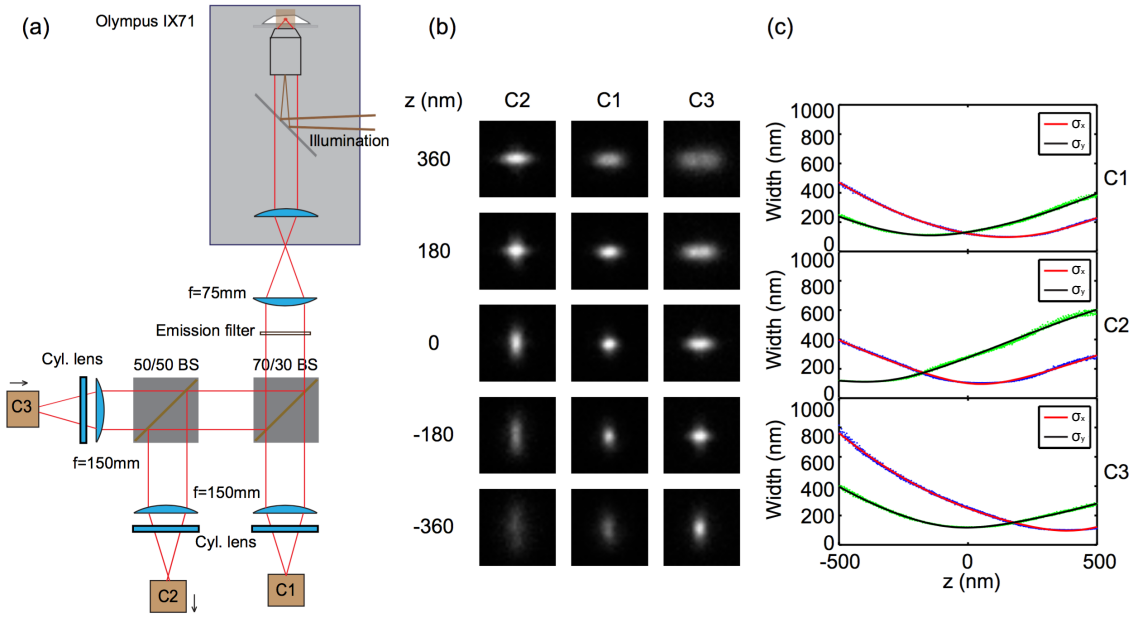



Figure 5.1: (a) Images of an emitter on the three cameras with various $z$ focuses (b) Calibration curves of the PSFs. $\sigma_x$ and $\sigma_y$ are functions of the focus position $z$ for different cameras.

is about 50. We moved the piezo stage along the $z$ direction 100 steps with a step size of 10 nm. The total travel distance (1000 nm) covers the whole depth of the image (800 nm). The centers, $(x_0, y_0)$, of the fluorescence bead images are determined by fitting to 3.9. All the bead centers from the second and third cameras are paired with those on the first camera. Local weighted mean transformation [75] is applied to register the three camera images. The registration error, $\sigma_{\text{loc}}$, for the second and third cameras using the first camera as the reference is within 6 nm (Fig. 5.2).

## 5.3 Signal Model and $\ell_1$-Homotopy

As mentioned in Section 3.4, the use of compressive sensing in 2D super-resolution image reconstruction has been well studied. In the case of 3D single-focus astigmatic

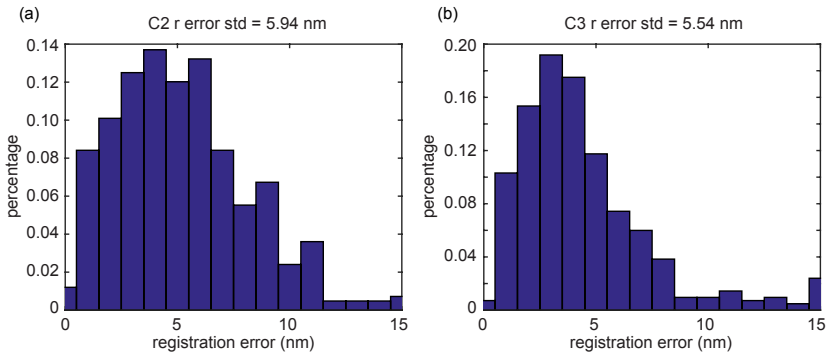



Figure 5.2: The image registration error distribution. The image from the first camera is used as the reference and the second and third camera images are registered to the first camera image. The standard deviation of the registration error for the second and third cameras are 5.94 nm and 5.54 nm respectively.

configuration, (3.9) is used to approximate the system PSF, with $\sigma_x$ and $\sigma_y$ as a function of the axial position $z$. Here, $y$ is vectorized from the measured image to form a one-dimensional vector, and the true emitter location $x$ is also a one-dimensional vector. In case of the 3D multi-focus astigmatic imaging, $y$ is concatenated from the three camera images. The true emitter location $x$ is also a one-dimensional vector that is concatenated from a 3D high-resolution grid, and the measurement matrix $A$ is of size $n_2 \times n_1$ where $n_2$ is the number of pixels in $y$ and $n_1$ is the number of pixels in $x$. The measurement matrix is determined by the system PSF. Typically $n_2$ is much smaller than $n_1$ which means the number of measurements are much smaller than the signal dimension.

Several methods have been developed to reconstruct the original signal $x$ through the $\ell_1$ minimization (3.36). The interior point method was used in [59,76] and utilized the implementation of the CVX package [77]. Another method is the $\ell_1$-Homotopy

algorithm that was used by Babcock et al. in [74]. The authors show that the $\ell_1$-Homotopy algorithm is roughly 10-250 fold faster than the CVX method without sacrificing the localization accuracy [74]. We apply the $\ell_1$-Homotopy algorithm in our 3D image reconstruction to make use of the rapid reconstruction rate compared to other methods. Here, the choice of the $\xi$ in (3.36) is critical. For simulated images, $\xi$ of value 1.5 is used to account for the Poisson noise. For real data acquired by the cameras, a factor of $\sqrt{2}$ is applied due to the noise introduced by the gain registers and hence $\xi$ of value 2.1 is used [48].

To reduce the computational complexity of the large optimization problem of 3D image reconstruction, which grows exponentially with the image size, we divide the image into small blocks and perform the optimization on these blocks individually. Since the cylindrical lens introduces de-focusing, the system PSF now spans a larger area than in the 2D case. Hence, we use block size of $20 \times 20$ and have 4 pixels overlapping with neighboring blocks on each side. For each pixel, we divide it into a $8 \times 8 \times 11$ voxel, which spans from -0.4 $\mu m$ to 0.4 $\mu m$ in the $z$ direction. Therefore, the measurement matrix $A$ is of size (20×20×3, 20×20×8×8×11), which is (1200, 281600). Using floating-point numbers, $A$ takes 1.35 GB of memory, which can easily fit into the memory of a modern computer. In our experiment, we used a Nvidia Tesla K40 GPU that has a global memory of 12 GB. The acquired $x$ represents a 3D up-sampled grid with the value at each entry representing the intensity of the emitter located at the corresponding grid. Algorithm 5.3 shows the detailed implementation of 3D MACS.

**Algorithm 1** 3D MACS

**Input**: $y$
**Output**: $x$ and $pos$
Select $\xi$ $(= 1.5)$
$A =$ GeneratePSF() /* according to (3.9) */
$\epsilon = \xi \cdot \|y\|_1$
$x =$ homotopy($y$, $A$, $\epsilon$) /* $\ell_1$-Homotopy */
$x =$ Debias($y$, $A$, $x$) /* 3D Local Debiasing */
$\mathcal{S}_0 = \{i| \, |x_i| > tol\}$ /* support set */
$pos =$ weightedcentroid($\mathcal{S}_0$, $x$) /* 3D weighted centroid of each */
/* connected components */

## 5.4 Debiasing

The finite discrete grid formulation in compressive sensing introduces bias to the estimation of the emitter locations and reduces the resolution of the final reconstructed super-resolution image [78]. Here we introduce a debiasing step followed by a 3D centroid calculation, which significantly reduces the bias and improves localization accuracy.

After the $\ell_1$ minimization step, we extract the support of the reconstructed $x$ (denote as $x_{\mathcal{S}_0}$) and the corresponding measurement matrix $A_{\mathcal{S}_0}$. The support of a vector is the set of all the non-zero elements of the vector. First, we dilate $x_{\mathcal{S}_0}$ and then minimize the least-squares objective $\|y - A_{\mathcal{S}_0} x_{\mathcal{S}_0}\|_2^2$ to obtain the best fit to the data over the dilated support of $x$ under non-negative constraint [79]. In the dilation step, we use a total of 75 voxels, 5 along the $x$ and $y$ directions and 3 along the $z$ direction. The physical size of the dilated volume is 45 nm $\times$ 45 nm $\times$ 240 nm. We used a larger physical length (240 nm) to compensate the lower resolution along the $z$ direction. We also tested other dilation sizes and found that the performance is

not sensitive to the dilation sizes. The 3D weighted centroids of the optimized $x_S$ are returned as the estimated positions of the emitters. Algorithm 5.4 shows the detailed implementation of the 3D local debiasing.

**Algorithm 2** 3D Local Debiasing

**Input**: $A$, $y$, $x$ from $\ell_1$ homotopy
**Output**: debiased $x$
$\mathcal{S}_0 = \{i| \, |x_i| > tol\}$ `/* support set */`
**if** $\mathcal{S}_0 \neq \emptyset$ **then**
  $B = ones(a_1, a_2, a_3)$ `/* 3D structuring element */`
  $\mathcal{S}_0 = \mathcal{S}_0 \oplus B$ `/* dilate support set */`
  $Z = \{1, \cdots, n\} \setminus \mathcal{S}_0$ `/* non-support set */`
  $x_Z = 0$ `/* set non-support set to zero */`
  $x_{\mathcal{S}_0} = \text{argmin}_{x \geq 0} \|A_{\mathcal{S}_0} x - b\|_2^2$ `/* optimize over dilated support set */`

## 5.5 Single-Focal SEA and Multi-Focal SEA

To compare 3D SACS with single emitter fitting based methods, we implement the SEA method that is developed for single focal plane reconstruction (SF-SEA) [4, 80]. Briefly, wavelet based de-noising followed by thresholding is used to find the candidate emitters. Then the candidate emitters are fitted with a 2D Gaussian function (3.9) to get $\sigma_x$ and $\sigma_y$, which are used to determine the $z$ location using the calibration curve defined in (2.8). SF-SEA is not designed to work for the multi-focal system. To compare with the 3D MACS, we develop and optimize a SEA method for the multi-focal system (MF-SEA). Specifically, the following optimization is performed.

$$\hat{\theta} = \underset{\tilde{\theta}}{\text{argmin}} \sum_t \sum_m \sum_n (y_t[m,n] - f_t[m,n|\{\tilde{\theta}\}])^2, \tag{5.1}$$

where $f_{\theta,m}(i,j)$ is defined by (3.9). $i$, $j$, and $m$ are the indices of the pixel and the focal plane, respectively. $y_t$ represents the image from the focal plane $t$. The essence of the MF-SEA is that we optimize the cost function (5.1) with all the focal planes taken into consideration simultaneously instead of optimizing the cost function for each focal plane individually.

Following the optimization is the filtering step. We use the ratio of the residual to the original image to characterize the quality of fitting. We define,

$$\beta = \frac{\sum_t \sum_m \sum_n (y_t[m,n] - f_t[m,n|\{\hat{\theta}\}])^2}{\sqrt{\sum_t \sum_m \sum_n |y_t[m,n]|}} \tag{5.2}$$

which is used to determine whether the candidate emitter would be kept. A low $\beta$ value represents higher fidelity to the data. We keep candidate emitters with $\beta$ less than 2.1 to be consistent with the criterion of $\xi$ used in 3D MACS and 3D SACS.

## 5.6 Results and Discussion

### 5.6.1 3D SACS and 3D MACS

To test our 3D SACS and 3D MACS methods, we simulate $P$ molecules that were randomly distributed in a volume of 0.9 $\mu m \times 0.9$ $\mu m \times 0.8$ $\mu m$. These molecules are imaged either on one camera (3D SACS) or on three cameras (3D MACS) with the experimentally measured PSF model ((3.9) and (2.8)). In all the simulations, the pixel size is 75 nm, which matches our experimental pixel size. In the case of 3D MACS, the distance between two consecutive focus planes is 180 nm. All the parameters are chosen to match the measured experimental values. An average photon number of 1500 per emitter is used in the simulation. In the case of 3D MACS, the total number of photons is evenly distributed on the three cameras. Both Poisson noise and Gaussian noise are added to model the photon noise and the camera readout noise

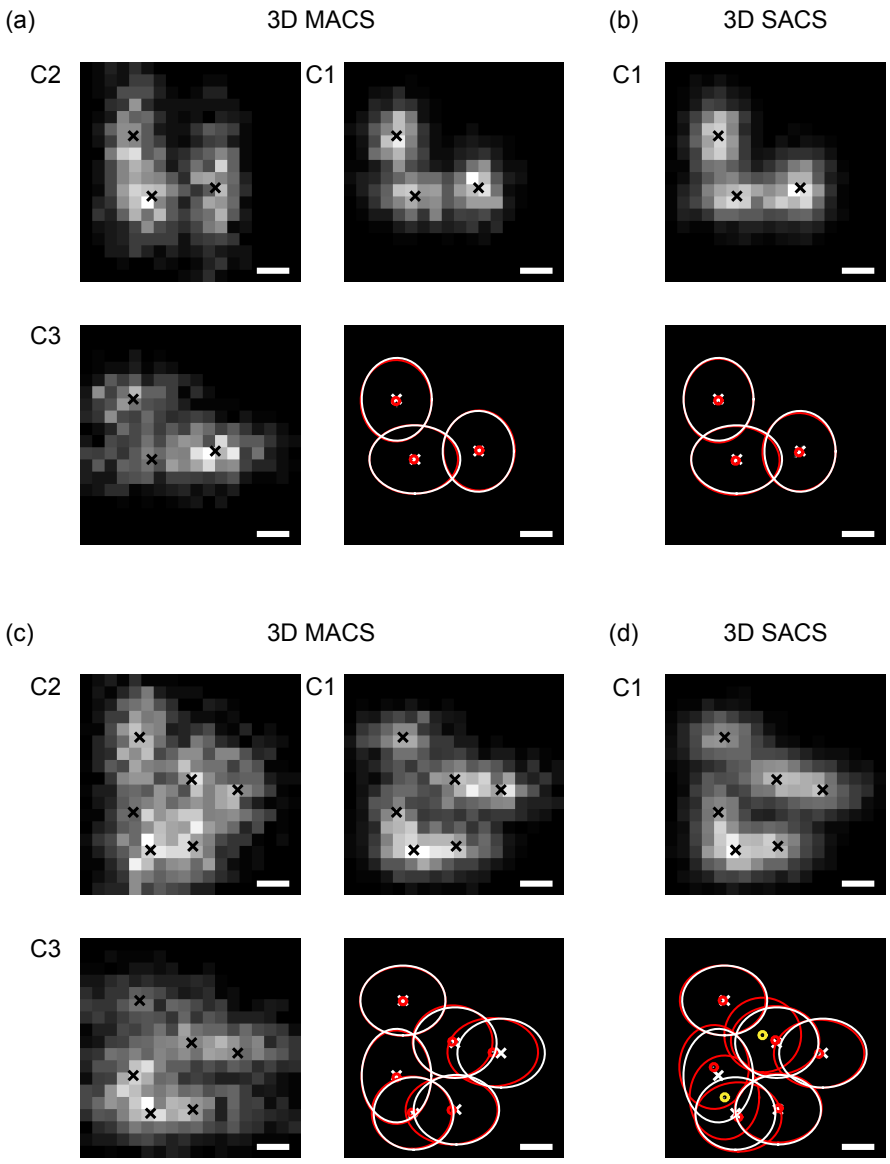



Figure 5.3: Comparison of the performance of 3D SACS and 3D MACS at low (a, b) and high density (c, d) of emitters. C1, C2, and C3 label the images formed on the three different cameras. At low emitter density, both 3D SACS and MACS can identify all the emitters with similar accuracy (lower right panel of (a) and lower panel of (b)). We use an ellipsoid to represent the $x$, $y$ and $z$ locations of an emitter. The center of ellipsoid is determined by the $x$ and $y$ positions and the shape is determined by the $z$ position. Red represents the results from our calculation and white represents the actual location from the simulation. When the emitter density increases (c, d), 3D MACS (lower right panel of (c)) correctly identified all the emitters, while SACS clearly show ambiguity (lower panel of (d)) with two wrongly detected emitters (colored yellow).

respectively. By adjusting the number of emitters $N$ per frame, we are able to control the emitter density. As shown in Fig. 5.3(a) and 5.3(b), at a low emitter density (3 emitters/$\mu m^2$), both 3D SACS and 3D MACS are able to correctly identify all the emitters with similar localization accuracy. However, when the density is increased (8 emitters/$\mu m^2$), 3D MACS performs markedly better. As shown in Fig. 5.3(c) and 5.3(d), 3D MACS is able to recover all the emitters without any false positive; while 3D SACS gives two false positive emitters. When the density is increased even more, the differential behavior of 3D MACS and 3D SACS becomes more obvious.

Astigmatism induced by the cylindrical lens breaks the symmetry of the PSF, providing depth information. However, as noticed in [76], the oval shape of the PSF caused by astigmatism makes it more difficult to distinguish overlapping emitters. For example, an off-focus emitter forms an elliptical spot on the camera, while two overlapping on-focus emitters can also form an elliptical image. In this case, it is impossible to differentiate whether the observed spot is formed by an off-focus emitter or two overlapping on-focus emitters, unless extra information is provided. Since we apply three different focuses in 3D MACS, the two on-focus emitters will look different from the off-focus emitter on the other two focus planes. Therefore, we are able to eliminate ambiguity and improve 3D MACS performance over 3D SACS, as shown in Fig. 5.3(c) and 5.3(d), by using additional information provided by extra cameras at a cost of lower photons per camera.

In our study, we use a Gaussian PSF model to generate the measurement matrix $A$. Due to the aberration induced by the refractive index mismatch, particularly at deeper site in the sample, the Gaussian PSF model can have limitations. The Gibson-Lanni [81, 82] PSF model could be applied when the astigmatism induced by

the cylindrical lens is taken into consideration. Adaptation of the Gibson-Lanni PSF model to our 3D MACS can potentially lead to an improvement in 3D resolution.

### 5.6.2 Debiasing

As discussed in Section 3.4, CS is based on a discrete formulation although the emitters locate over a continuum. This formulation intrinsically introduces bias and reduces the localization accuracy. In [59], the authors use geometry centroids from the reconstructed high resolution images as the coordinates of the emitters, which partially resolves the bias issue. The bias introduced by the discrete grid becomes dominant when the grid size gets bigger. In our 3D MACS, the up-sampling factors in the $x$, $y$, and $z$ direction are 8, 8 and 11 respectively, which gives rise to a 9 nm$\times$9 nm$\times$80 nm voxel. Due to the larger length along the $z$ direction, the bias is much bigger along the axial direction.

Figure 5.4(a) and 5.4(b) show localization error along the axial direction before and after the debiasing step respectively. Before debiasing, the periodic bias is evident. This is introduced by the finite size of a voxel. Emitters that are closer to the borders of a voxel tend to fall onto the edges of the voxel. After debiasing, the periodic bias is reduced significantly as shown in Fig. 5.4(b).

We also assess how debiasing improves localization accuracy along the $x$ and $y$ directions. As shown in Fig. 5.4(c), debiasing provides limited improvement along the lateral directions, consistent with the observation in [59], because of the much smaller length scale (9 nm). In contrast, the improvement along the axial direction is apparent due to the larger length of the voxel along the $z$ direction. As expected, decreasing the voxel length along $z$ by increasing the up-sampling factor improves

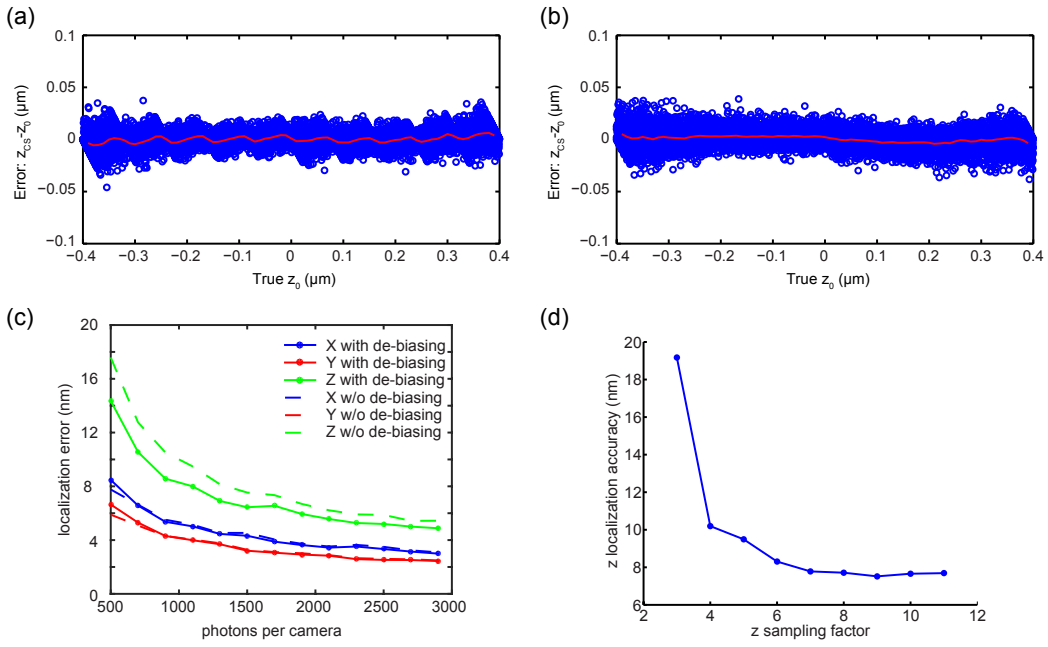



Figure 5.4: Bias at different $z$ locations (a) without and (b) with debiasing step. (c) Localization accuracy increases with debiasing step, particularly for the axial direction (green solid and dashed lines). For the $x$ and $y$ directions, debiasing barely improves the localization accuracy. (d) $z$ localization accuracy improves with the increase of the $z$ sampling factor.

the localization accuracy (Fig. 5.4(d)). However, the improvement curve levels off when the up-sampling factor gets to about 10, in which case the voxel size along $z$ direction is small enough to minimize the bias. Debiasing not only improves the

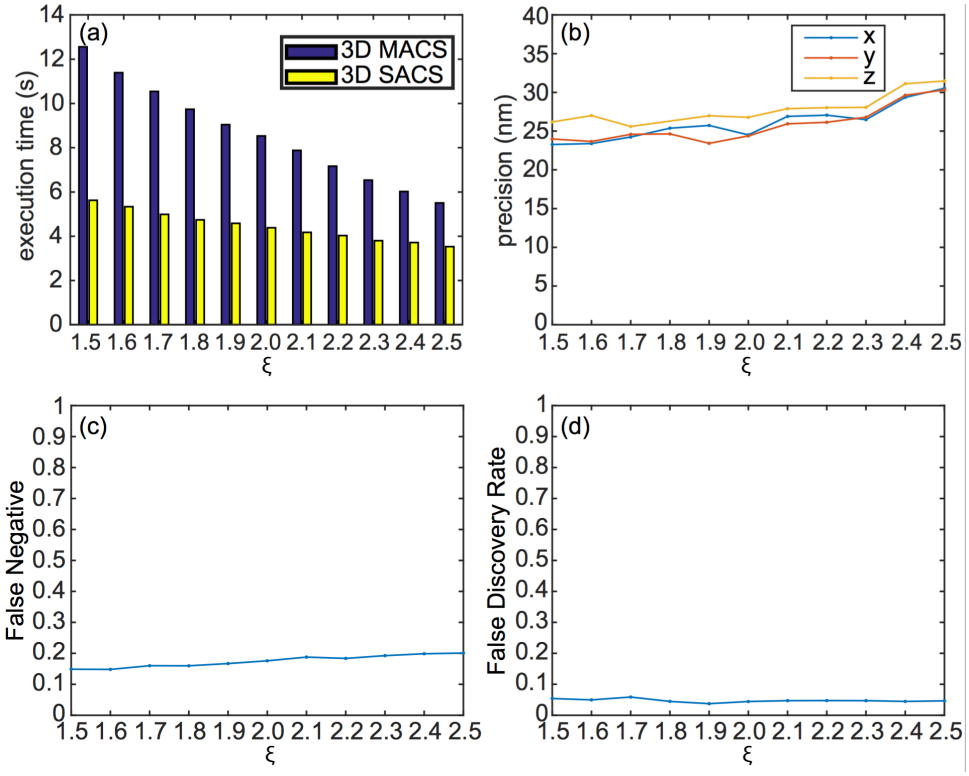



Figure 5.5: (a) The execution time of both 3D MACS and 3D SACS decreases with the increase of $\xi$. (b) The $x$, $y$, and $z$ direction localization error only slightly increase with the increase of $\xi$. The false negative (c) and false discovery rate (d) stay relatively constant as $\xi$ increases from 1.5 to 2.5.

localization accuracy but also speeds up the reconstruction process. The most time consuming part of the reconstruction process is the $\ell_1$ optimization step, which heavily depends on the choice of the value of $\xi$. Choosing a lower value of $\xi$ increases the optimization fidelity, minimizing the error between the estimated emitters and the

true locations of the emitters, but at the cost of a longer computation time. On the other hand, a higher value of $\xi$ accelerates the computational speed at the cost of a larger localization error. We test the effect of $\xi$ on the computation time and performance. As shown in Fig. 5.5(a), the computation time reduces significantly with the increase of $\xi$, while the localization errors (Fig. 5.5(b)) slightly increase but still within the acceptable range (<30 nm in $x$, $y$ and $z$ directions). In the simulation, the emitter density we use is 8 emitters/$\mu m^2$. We also test the false negative and false discovery rate (Fig. 5.5(c) and 5.5(d)) both of which stay below 20% when epsilon increases from 1.5 to 2.5. All these results suggest that we are able to gain computational speed without sacrificing performance.

### 5.6.3 Performance Evaluation of 3D SACS and 3D MACS

To demonstrate the performance of our 3D SACS and 3D MACS methods, we generate a series of simulated 20 $\times$ 20 camera pixel STORM movies across a range of emitter densities (0.1 emitter/$\mu m^2$ to 10 emitters/$\mu m^2$) and analyze them with SF-SEA, MF-SEA, 3D SACS, and 3D MACS. Figure 5.6(a) demonstrates that both 3D SACS and 3D MACS could be used to identify as high as 10 emitter/$\mu m^2$ density of emitters with acceptable lateral and axial localization accuracy. Both SF-SEA and MF-SEA could not be used when emitter density is higher than 1 emitter/$\mu m^2$. In addition, 3D MACS shows superior performance in identifying high density emitters and better $x$, $y$ and $z$ localization accuracy than 3D SACS. Figure 5.6(b) also shows that the false discovery rate of 3D MACS is much lower than 3D SACS. It is important to note that in Fig. 5.6(b), the false discovery rate of both SF-SEA and MF-SEA are lower than the 3D MACS. This lower false discovery rate is due to SEA

eliminating any ambiguous detection, resulting in a low identified emitter density as shown in Fig. 5.6(a). Figure 5.6(c) and 5.6(d) show the localization error in $x$, $y$ and $z$ directions for 3D MACS, 3D SACS, SF-SEA and MF-SEA. 3D MACS achieves the lowest localization error with 3D SACS performing slightly worse.

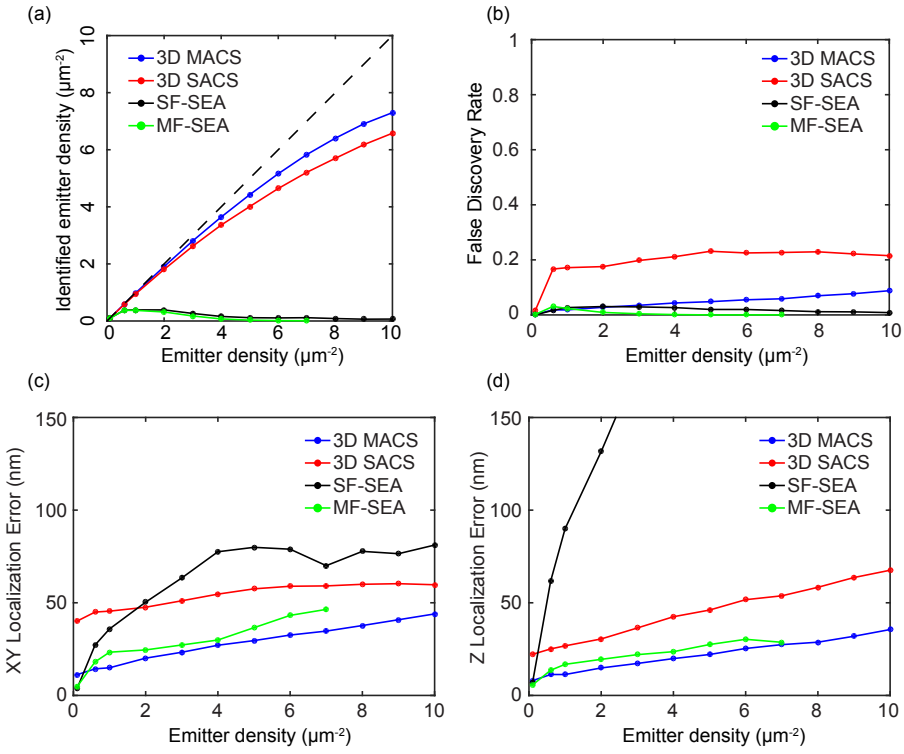



Figure 5.6: (a) A comparison of the identified emitter densities using 3D SACS, 3D MACS, SF-SEA and MF-SEA methods. (b) False discovery rates of the emitters using the three different methods. (c) XY and (d) Z localization precision at various emitter densities.

We also study the computation time as a function of emitter density using 3D MACS and 3D SACS methods implemented on a GPU (Nvidia Tesla K40). Table 5.1 shows that the computation time per block ($20 \times 20$) increases as the emitter density increases. Due to the increased computational complexity, 3D MACS takes about

| density ($\mu m^{-2}$) | 3D SACS (s) | 3D MACS (s) |
|---|---|---|
| 2 | 1.14 | 2.31 |
| 3 | 1.79 | 3.08 |
| 4 | 2.56 | 5.23 |
| 5 | 3.27 | 6.45 |
| 6 | 4.44 | 9.40 |
| 7 | 5.40 | 10.91 |
| 8 | 5.86 | 12.56 |

Table 5.1: Change of the execution time with 3D SACS and 3D MACS under different emitter densities.

double the time of 3D SACS. However, when the improved localization accuracy and detection rates are considered (Fig. 5.6), in general, we recommend using 3D MACS rather than 3D SACS wherever practical. In addition, the computation time can be further decreased with the use of multiple GPUs in parallel, such as the Nvidia Maximus multi-GPU system.

To evaluate the performance of 3D MACS on structures with variable emitter densities, we simulate three different types of structures: a sphere, three crossed lines, and a 3D stellate (Fig. 5.7). The radius of the sphere is 60 nm, the minimum distance in the $z$ direction between the two crossed lines is 460 nm and the distance between two consecutive branches of the stellate is 57 nm. In the simulation, the densities used are all 8 emitters/$\mu m^2$. The number of frames used is 3000. For MF-SEA, due to the low detection rate under high emitter density, the reconstruction of all the structures are not satisfactory. On the other hand, 3D MACS is able to resolve their structures clearly. For the sphere (Fig. 5.7(a)), the XY, XZ, and YZ cross sections clearly show that the spherical structure is resolved. The three crossed lines (Fig. 5.7(b)) simulate linear structures that are frequently observed in biological

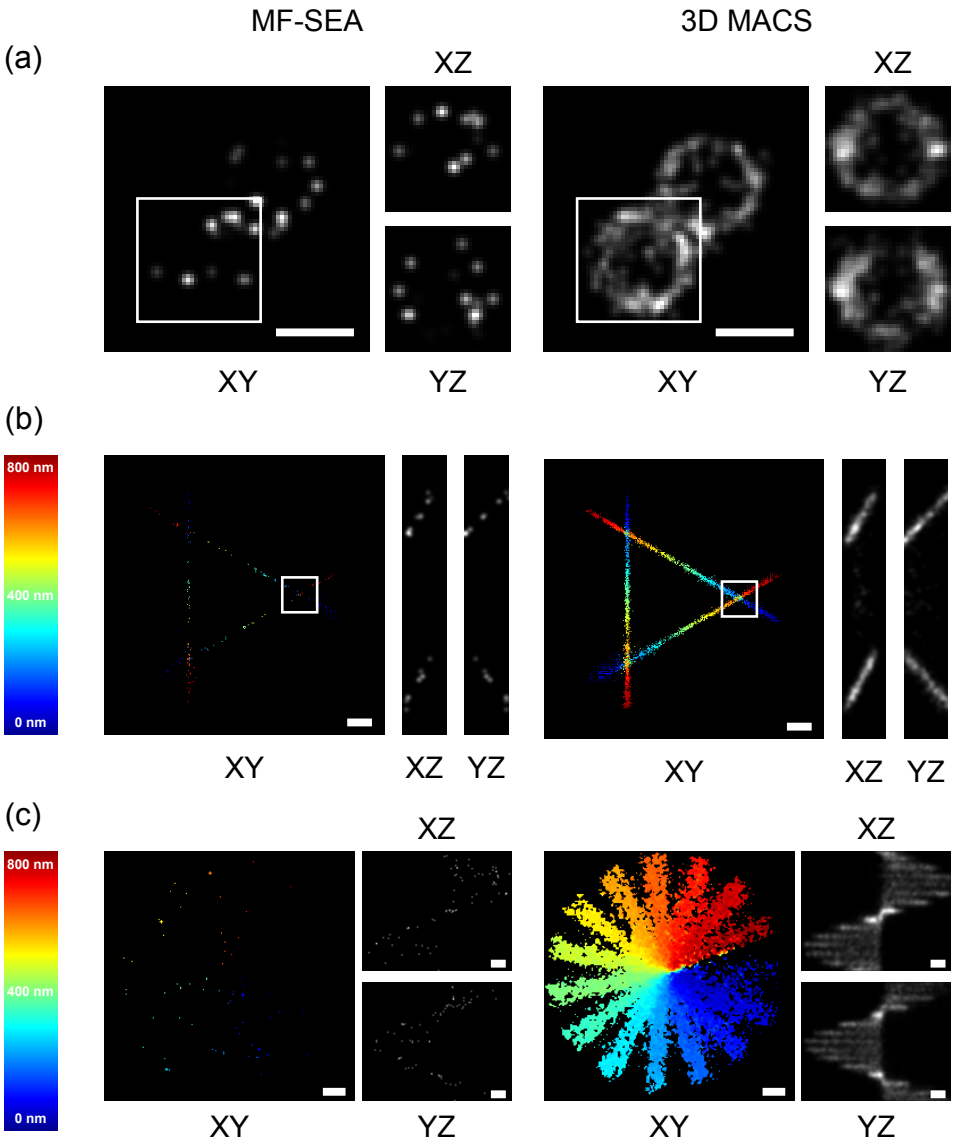



Figure 5.7: The XY, XZ, and YZ projection images of (a) simulated spheres, (b) 3-crossinged lines, and (c) 3D stellate with MF-SEA and 3D MACS using 3000 frames. Scale bar: 100 nm.

experiments, such as the microtubule structures. The artificial 3D stellate structure (Fig. 5.7(c)) demonstrates the ability of 3D MACS in resolving gradually changing structures.

In this study we use either one or three focal planes, however, our method can be applicable to other multifocal microscopy system [85, 86]. With the increase of the number of focal planes, the coverage of the sample depth increases. For example, Abrahamsson et al. [87] have developed a system using a 9-plane setup to track molecules with depth of 2.25-18 $\mu m$. Since the total number of photons of a fluorophore is limited, the number of photons collected at each focal plane decreases with the increase of number of focal planes, which will compromise the localization accuracy. Therefore, it is important to balance the number of focal planes and their spacing toward achieving an optimum spatial resolution [88].

### 5.6.4 Reconstruction of 3D microtubule STORM images using 3D MACS

Fig. 5.8(a) shows the reconstructed 3D microtubule image of a fixed HeLa cell using 1000 raw frames with our 3D MACS method. The frame rate used in our experiment is 56 frames per second (fps), which means that we are able to acquire the image in less than 20 seconds. Even with only 1000 frames, we are able to reconstruct very continuous 3D microtubule images. Figures 5.8(b) through 5.8(d) show the zoom-in region, averaged image, and one typical raw frame, respectively. Fig. 5.8(e) shows the YZ cross section of the region specified by the selected region in (b). The overlay of two microtubules is clearly seen in the cross section and Fig. 5.8(f) is the line profile along $z$ direction. We observe 35 nm and 36 nm of full width at

half maximum (FWHM) in $z$ direction with the distance between two microtubules as 112 nm.

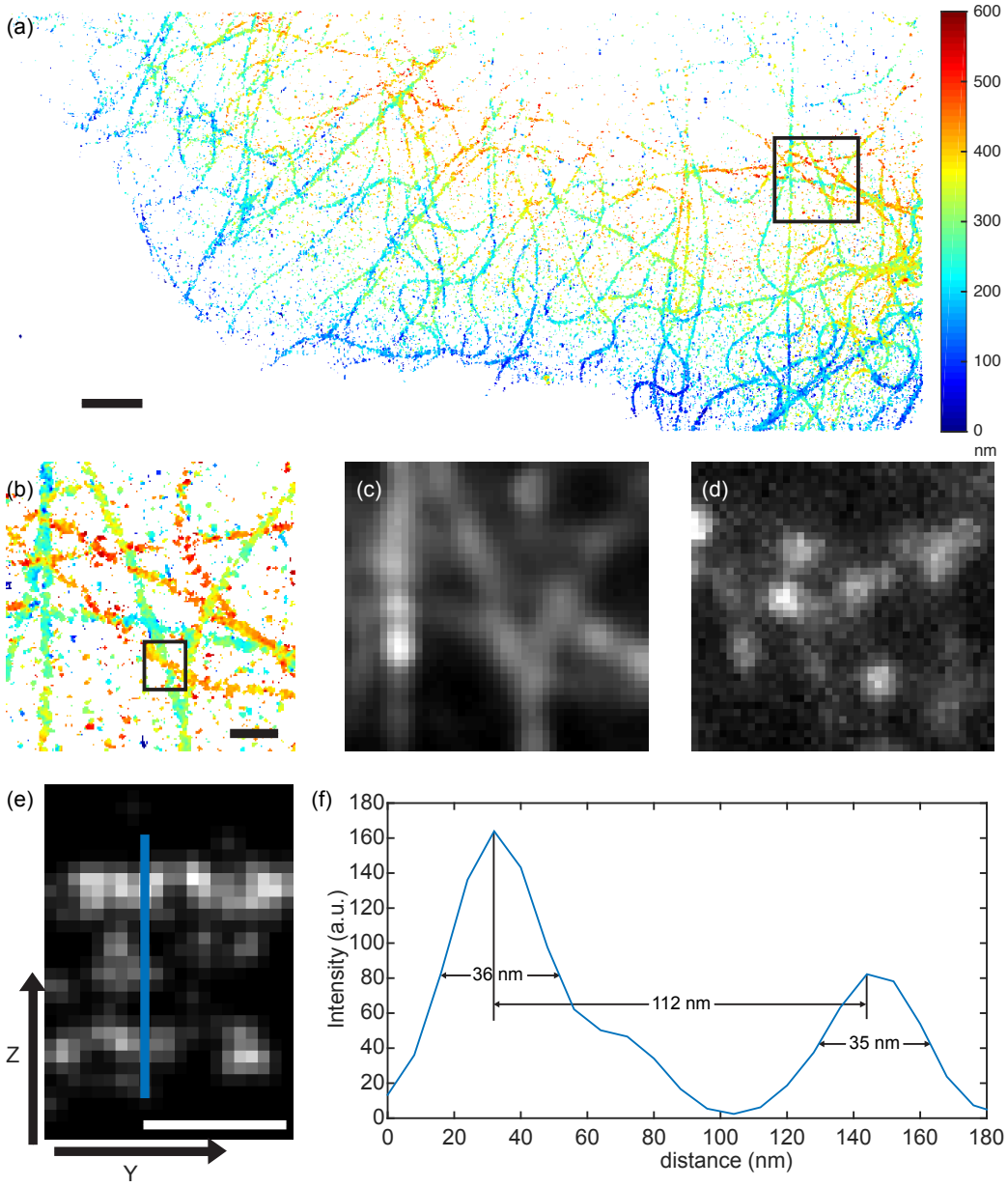



Figure 5.8: (a) Reconstructed Microtubule image (bar: 1000 nm). (b) Zoom in of the selected region in (a) (bar: 500 nm). (c) Averaged image. (d) One typical raw frame. (e) YZ projection view of the selected region in (b) (bar: 100 nm). (f) Line profile of (e).

## 5.7 Summary

In summary, we have developed a multi-focus astigmatism imaging system for 3D SMLM imaging (3D MACS and 3D SACS). The designed optical system, in combination with $\ell_1$-Homotopy procedure followed by $\ell_2$ debiasing, allows us to acquire 3D super-resolution images in a much shorter time span with high emitter densities, as high as 10 emitters/$\mu m^2$. The introduced $\ell_2$ debiasing step and the 3D weighted centroid refinement allow us to reduce the stringency of $\xi$ to achieve faster reconstruction speed. We also implement the algorithm on a GPU to reduce reconstruction time. Currently, we implement our algorithm mostly using Matlab. By porting the algorithm completely to C/C++, we expect to further accelerate the reconstruction.

# Chapter 6: Total-Variation Norm Regularized Reconstruction under Poisson Noise

In this chapter, we propose a novel super-resolution image reconstruction algorithm, dubbed TVSTORM, for 3D image reconstruction under the Poisson noise model, which is more appropriate for photon count data. The camera image is treated as an observation drawn from a Poisson distribution whose parameters are determined by the 3D PSF profile and the locations of activated emitters. The reconstruction is performed by minimizing a negative Poisson log-likelihood penalized by the total-variation norm [89] of the point source signal composed of the activated emitters. The total-variation norm can be viewed as a continuous analog of the $\ell_1$ norm for finite-dimensional vectors to promote emitter sparsity without discretizing their locations. We solve the proposed optimization problem efficiently following a variant of the alternating descent conditional gradient method in [90] and [91]. Specifically, in each iteration, TVSTORM first selects a new point source and adds it to the current estimate, whose location is determined by a first-order linearization of the Poisson log-likelihood function over a coarse grid, and then refines the estimate of all included point sources by gradient descent using backtracking line search. Through numerical experiments, TVSTORM demonstrates an order-of-magnitude improvement on the computational cost over CSSTORM due to the elimination of optimizing over a

fine-grained grid. It also shows a significant improvement on the localization accuracy over CSSTORM in terms of detection rate, false discovery rate and precision, without adding post-processing steps.

The rest of this chapter is organized as follows. Section 6.1 describes the signal model of super-resolution image reconstruction. We present the proposed TVSTORM algorithm in Section 6.2. Numerical experiments on both synthetic and real data are demonstrated in Section 6.3. Finally, we conclude in Section 6.4.

## 6.1 Signal Model

The image formulation process of 3D SMLM has been described in detail in Section 3.2. Here we rewrite a few key definitions for reference.

Let $\Theta = \{\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(P)}\}$ be the set of parameters, where $\theta^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_z^{(i)}, \theta_I^{(i)}] = [\theta_L^{(i)}, \theta_I^{(i)}]$ is the parameter for the $i^{\text{th}}$ emitter and $\theta_L^{(i)} = [\theta_x^{(i)}, \theta_y^{(i)}, \theta_z^{(i)}] \in \mathcal{S}$ are the coordinates in $x$, $y$ and $z$ dimensions, respectively. The set of activated emitters $\chi = \chi(x, y, z|\Theta)$ is written as a sparse superposition of point sources, given as

$$\chi = \chi(x, y, z|\Theta) = \sum_{i=1}^{P} \theta_I^{(i)} \delta(x - \theta_x^{(i)}, y - \theta_y^{(i)}, z - \theta_z^{(i)}), \tag{6.1}$$

We denote the admissible set of $\chi$ as

$$\mathcal{G} = \{\chi = \chi(\Theta)|\Theta = \{\theta^{(i)}\}_{i=1}^{P}, P \in \mathbb{Z}^+, \theta_L^{(i)} \in \mathcal{S}, \theta_I^{(i)} \geq 0, 1 \leq i \leq P\}. \tag{6.2}$$

The expected number of photons received at the $(m, n)^{\text{th}}$ camera pixel is written as

$$\begin{aligned}\mu[m, n|\chi] = &\sum_{i=1}^{P} \frac{\theta_I^{(i)}}{4} \left[ Q\left( \frac{m - \theta_x^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) - Q\left( \frac{m - \theta_x^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) \right] \\ &\times \left[ Q\left( \frac{n - \theta_y^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) - Q\left( \frac{n - \theta_y^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) \right].\end{aligned} \tag{6.3}$$

The number of photons hitting the camera at the $(u,v)^{\text{th}}$ pixel, denoted as $y[u,v]$, follows an independent Poisson distribution with the parameter $\mu[u,v|\chi]$, given as

$$\Pr(y[m,n]=\omega|\chi)=\frac{\mu[u,v|\chi]^{\omega}e^{-\mu[u,v|\chi]}}{\omega!},\quad \omega\in\mathbb{Z}^{+}. \tag{6.4}$$

Denote the camera image as $y=\{y[u,v]\}$. The objective of super-resolution image reconstruction is to estimate the point source signal $\chi(\Theta)$, given the observed image $y$.

## 6.2 Proposed Approach

In this section, we describe the proposed TVSTORM algorithm for high-density 3D super-resolution imaging under the Poisson noise model.

### 6.2.1 Theoretical Framework

We first define the loss function, $\ell(y|\chi)$, as the negative Poisson log-likelihood of observing $y$ given $\chi$. According to (6.4),

$$\begin{aligned}\ell(y|\chi)&=-\log\left(\prod_u\prod_v\Pr(y[u,v]|\chi)\right)\\&=\sum_u\sum_v\left(\mu[u,v|\chi]-y[u,v]\log(\mu[u,v|\chi])\right)+C,\end{aligned}$$

where $C$ is a constant that does not depend on $\chi$. As suggested in [92], for Poisson log-likelihood, it is advantageous to introduce a small offset $0<\eta\ll 1$ to improve stability, where we modify $\ell(y|\chi)$ as

$$\ell(y|\chi)=\sum_u\sum_v\left(\mu[u,v|\chi]-y[u,v]\log(\mu[u,v|\chi]+\eta)\right). \tag{6.5}$$

On the other hand, we wish to motivate the underlying sparsity of the activated emitters. To this end, we define the total-variation norm of the point source signal $\chi$,

$\|\chi\|_{\mathrm{TV}}$, whose measure-theoretic definition can be found in [89]. The total-variation norm can be seen as a generalization of the $\ell_1$ norm for finite-dimensional vectors to the continuous space without imposing a discrete grid for the location of the emitters. For the point source signal in (6.1), $\|\chi\|_{\mathrm{TV}} = \sum_{i=1}^{P} \theta_I^{(i)}$.

We propose to seek the point source signal $\chi$ that minimizes the loss function $\ell(y|\chi)$ penalized by its total-variation norm, given as

$$\hat{\chi} = \underset{\tilde{\chi}\in\mathcal{G}}{\operatorname{argmin}}\ \ell(y|\tilde{\chi}) + \xi\|\tilde{\chi}\|_{\mathrm{TV}}, \tag{6.6}$$

where $\xi$ is a regularization parameter that controls the trade-off between the fidelity to the observation and the sparsity of the emitters. We denote this approach as TVSTORM.

### 6.2.2 Implementation

Unfortunately, the algorithm (6.6) is in general non-convex for 3D imaging due to the likelihood term, and challenging to solve. Nonetheless, we develop an iterative algorithm that can be regarded as a variant of the alternating descent conditional gradient method in [90] and [91]. The description is given in Algorithm 3. TVSTORM is an iterative algorithm, where in each iteration, a new point source is first selected and added to the current estimate of $\chi$, and then the estimate of $\chi$ is refined by gradient descent using backtracking line search. The algorithm stops when the intensity of the most recently added point source falls below a given threshold.

Let $\hat{\chi}^{(t)} = \chi(x, y, z|\hat{\Theta}^{(t)})$ be the estimate of the emitter object at the $t^{\mathrm{th}}$ iteration, where $\hat{\Theta}^{(t)}$ represents the parameters of the point sources in $\hat{\chi}^{(t)}$. At the $(t+1)^{\mathrm{th}}$ iteration, the SELECT step aims to add one point source (with parameter $\hat{\theta}^{(t+1)}$) to the emitter object of the previous iteration, $\hat{\chi}^{(t)}$, by minimizing the first-order Taylor

**Algorithm 3** TVSTORM

**Input Parameter:** threshold $\gamma$
$t \leftarrow 0$
$\hat{\Theta}^{(0)} \leftarrow \varnothing$
$\hat{\chi}^{(0)} \leftarrow \chi(x, y, z|\hat{\Theta}^{(0)})$
**repeat**
  ▷ **SELECT**
  $\hat{\theta}_L^{(t+1)} \leftarrow \underset{\theta \in \mathcal{S}_{\text{coarse}}}{\operatorname{argmin}} \left\langle \frac{\partial \ell(y|\hat{\chi}^{(t)})}{\partial \mu(\hat{\chi}^{(t)})}, \mu(\chi(x, y, z|\theta)) \right\rangle.$
  $\hat{\Theta}^{(t+1)} \leftarrow \hat{\Theta}^{(t)} \cup \{\hat{\theta}_L^{(t+1)}\}$
  ▷ **REFINE**
  $\hat{\Theta}^{(t+1)} \leftarrow \text{REFINE}(y, \hat{\Theta}^{(t+1)})$
  $\hat{\chi}^{(t+1)} \leftarrow \chi(x, y, z|\hat{\Theta}^{(t+1)})$
  $t \leftarrow t + 1$
**until** $\hat{\theta}_I^{(t)} < \gamma$

series approximation of $\ell(y|\hat{\chi})$ around $\hat{\chi}^{(t)}$, which is

$$\ell\left(y\middle|\hat{\chi}^{(t)} + \chi\left(x, y, z|\theta\right)\right) \approx \ell(y|\hat{\chi}^{(t)}) + \left\langle \frac{\partial \ell(y|\hat{\chi}^{(t)})}{\partial \hat{\chi}^{(t)}}, \chi(x, y, z|\theta) \right\rangle. \tag{6.7}$$

Let $\mu(\chi) = \{\mu(u, v|\chi)\}$ denote the noise-free image generated from $\chi$. It can be shown that [90]

$$\left\langle \frac{\partial \ell(y|\hat{\chi}^{(t)})}{\partial \hat{\chi}^{(t)}}, \chi(x, y, z|\theta) \right\rangle = \left\langle \frac{\partial \ell(y|\hat{\chi}^{(t)})}{\partial \mu(\hat{\chi}^{(t)})}, \mu(\chi(x, y, z|\theta)) \right\rangle, \tag{6.8}$$

which does not depend on the intensity. Therefore, the location of the new point source is selected as

$$\hat{\theta}_L^{(t+1)} = \underset{\theta \in S_{\text{coarse}}}{\operatorname{argmin}} \left\langle \frac{\partial \ell(y|\hat{\chi}^{(t)})}{\partial \mu(\hat{\chi}^{(t)})}, \mu(\chi(x, y, z|\theta)) \right\rangle, \tag{6.9}$$

where $S_{\text{coarse}}$ is a coarse grid over $\mathcal{S}$. We only require a coarse grid since the locations will be refined afterwards.

The REFINE step aims to find the maximum likelihood estimate of $\chi^{(t+1)}$ with the number of point sources fixed by minimizing the loss function using iterative gradient descent over $\hat{\Theta}^{(t+1)}$. For each parameter $\theta \in \hat{\Theta}^{(t+1)}$, we first find the direction that

decreases the loss function by calculating the partial derivative of the loss function with respect to $\theta$, whose expressions are given in Appendix A. The step size is then determined using backtracking line search to speed up convergence. The details of this step are described in Algorithm 4.

**Algorithm 4** REFINE$(y, \Theta)$

1: **Input Parameters:** $\alpha_0$, $\tau \in (0,1), c \in (0,1)$
2: **repeat**
3:  **for** every $\theta^{(i)}$ in $\Theta$ **do**
4:  **for** every $\theta_j^{(i)}$ in $\theta^{(i)}$ **do**
5:  $\alpha \leftarrow \alpha_0$
6:  $\widetilde{\Theta} \leftarrow \Theta$
7:  **repeat**
8:  $\widetilde{\theta}_j^{(i)} \leftarrow \theta_j^{(i)} - \alpha \frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_j^{(i)}}$
9:  $\alpha \leftarrow \alpha \times \tau$
10:  **until** $\ell(y|\chi(\widetilde{\Theta})) \leq \ell(y|\chi(\Theta)) - \alpha c \| \frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_j^{(i)}} \|_2^2$
11:  $\Theta \leftarrow \widetilde{\Theta}$
12:  **end for**
13:  **end for**
14: **until** convergence

### 6.2.3 Discussions

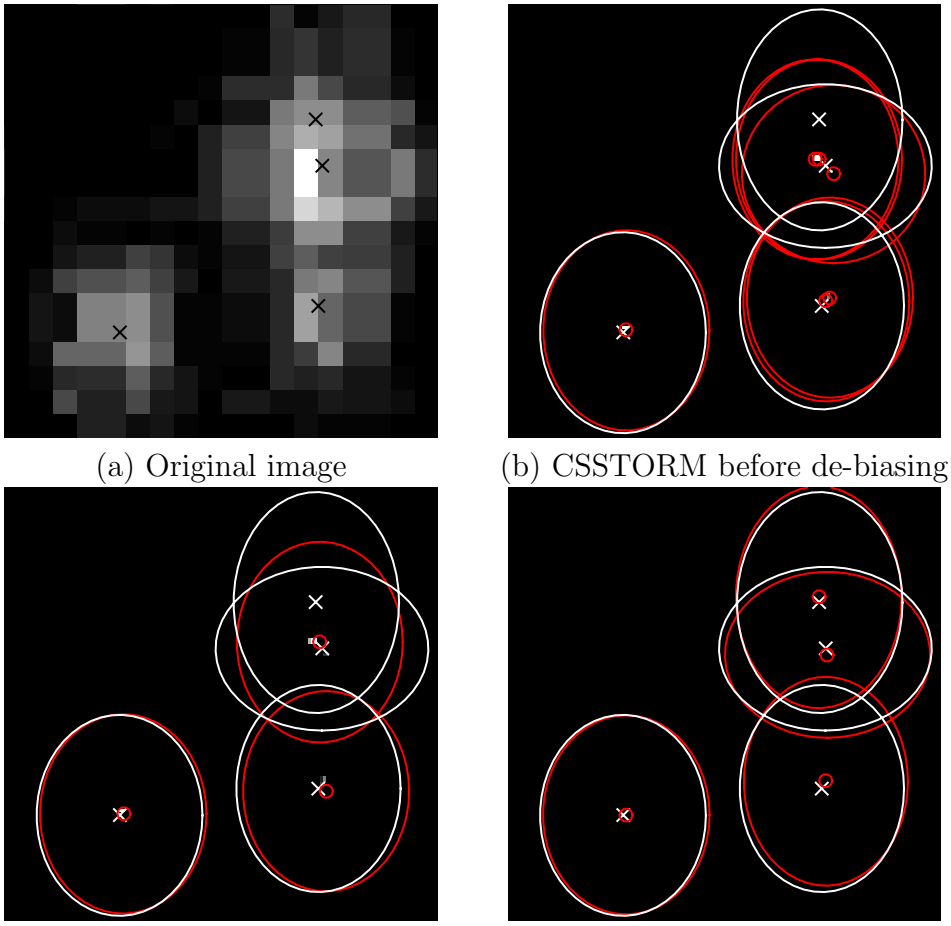
Since a coarse grid is used in the SELECT step to find the starting locations of emitters, the main computation is committed in the REFINE step, specifically, in the re-evaluation of $\ell(y|\chi(\widetilde{\Theta}))$ each time $\widetilde{\Theta}$ is altered (line 10 in Algorithm 4). Assume the image size is $n \times n$. According to (6.3), the re-evaluation of $\ell(y|\chi(\widetilde{\Theta}))$ requires $\mathcal{O}(|\tilde{\Theta}| \cdot n^2)$ calculations. Since each time the re-evaluation only requires changing the parameters of one emitter (line 8 in Algorithm 4), efficient implementation of this

step requires only $\mathcal{O}(n^2)$ calculations by maintaining $\chi(\widetilde{\Theta} \setminus \{\widetilde{\theta}^{(i)}\})$ when refining the parameters of the $i^{\text{th}}$ emitter.

The stopping criteria of TVSTORM is determined by threshold $\lambda$ in Algorithm 3. Fortunately, we find the performance of TVSTORM relatively insensitive to the selection of $\lambda$. In our implementation, we set $\lambda$ to half of the expected emitter intensity. In the REFINE step, the selection of $\alpha_0$ determines the maximum step size, while $\tau$ regulates the shrinking factor of the step size and $c$ involves the Armijo-Goldstein condition [93] (line 10 in Algorithm 4). To avoid evaluating $\chi(\widetilde{\Theta})$ at too many points, we set $\alpha_0 = 10^3$, $\tau = 10^{-1}$ and $c = 10^{-6}$.

## 6.3 Numerical Experiments

We conduct numerical simulations and real experiments to demonstrate the preferable performance of TVSTORM. The images used for numerical simulations are generated by randomly distributing certain amount of emitters in a 3D or 2D space. The number of emitters is determined by the required density of each simulation; the volume of 3D space is 0.8 $\mu m$ $\times$ 0.8 $\mu m$ $\times$ 0.8 $\mu m$ and the volume of 2D space is 0.8 $\mu m$ $\times$ 0.8 $\mu m$ unless otherwise stated. The pixel size of the generated images is 75 $nm$ $\times$ 75 $nm$, which matches our real system setting. Equations (3.9) and (2.8) are used to generate the PSF model for simulation whose parameters are calibrated in [12]. The simulated images are then corrupted with Poisson noise with the means of the detected photon number determined by the expected number of photons of each pixel, as described in Section 6.1.



(a) Original image (b) CSSTORM before de-biasing

(c) CSSTORM after de-biasing (d) TVSTORM

Figure 6.1: Emitter localization using CSSTORM and TVSTORM. (a) Original image; (b) CSSTORM before debiasing; (c) CSSTORM after debiasing; and (d) TVSTORM.

### 6.3.1 Comparisons with CSSTORM on 3D Image Reconstruction

We first examine the reconstruction quality using CSSTORM and TVSTORM on a single frame. We generate an image with four emitters that are randomly distributed in a 3D space, with intensity of 300 photons each, as shown in Fig. 6.1 (a). For CSSTORM, an up-sampling factor of 8 in the lateral direction and 9 in the axial

direction is used in the discretization. The output from CSSTORM typically requires post-processing such as de-biasing [12] in order to mitigate the gridding error, while TVSTORM does not include post-processing steps. Fig. 6.1 (b) and (c) show the image reconstruction from CSSTORM before and after de-biasing, where we use an ellipsoid to represent the spatial locations of an emitter with the center representing its lateral position and the shape representing its axial position. The reconstruction is shown in red, while the ground truth is shown in white. As seen from Fig. 6.1 (b), the reconstruction from CSSTORM before de-biasing contains many false positives. After de-biasing, nearby output emitters are clustered together but one emitter is missing, as shown in Fig. 6.1 (c). Contrarily, the reconstruction from TVSTORM, as shown in Fig. 6.1 (d), identifies all emitters correctly with high precision.

Next, to evaluate the average performance of TVSTORM, we generate a series of SMLM images under different densities (0.75 emitter/$\mu m^3$ to 11.25 emitters/$\mu m^3$). The emitters are randomly distributed in a 3D volume with the intensity set as 500. The images are then corrupted with Poisson noise. Fig. 6.2 compares the performance of TVSTORM with CSSTORM in terms of identified density, false discovery rate, precision and execution time with respect to the emitter density. The identified density is defined as the number of correctly detected emitters per area. The precision is defined as the standard deviation of localization errors. Indeed, TVSTORM is able to detect more emitters with an improved precision while maintaining a lower false discovery rate than CSSTORM. Additionally, the execution time of TVSTORM is much faster than that of CSSTORM due to the elimination of a fine-grained grid during optimization.

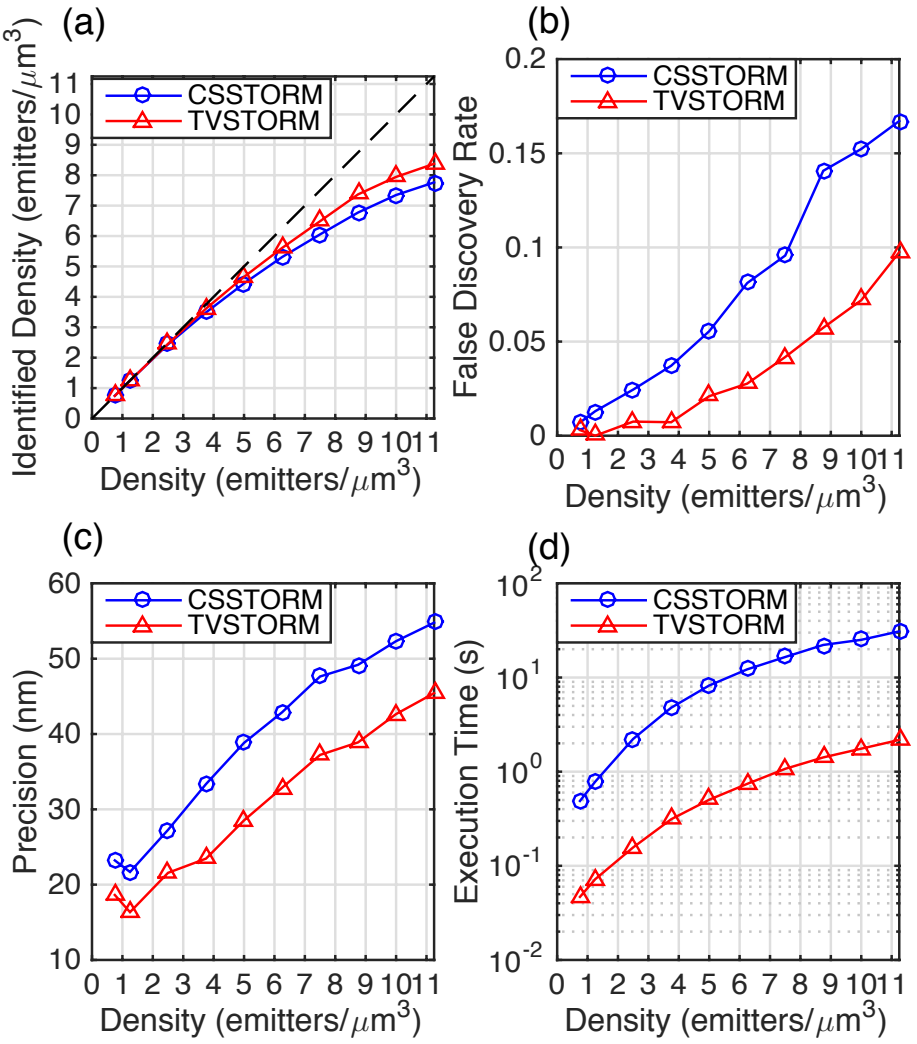



Figure 6.2: Performance comparisons of 3D image reconstruction between TVSTORM and CSSTORM: (a) identified density, (b) false discovery rate, (c) precision and (d) execution time with respect to the emitter density.

### 6.3.2 Performance under Different Photon Levels

We then evaluate the performance of TVSTORM as the SNR varies. We generate a series of SMLM images with emitters randomly distributed in a 3D volume under different SNR levels by varying the emitter photon numbers from 100 to 1000, covering the photon levels of normal fluorescent proteins and fluorescent dyes. Fig. 6.3 shows that the performance of TVSTORM deteriorates with the decrease of the SNR. However, even with an emitter photon number as low as 100, TVSTORM is still able to detect a large portion of emitters with small false discovery rate at an acceptable precision, as shown by the identified density, false discovery rate, and precision with respect to the emitter density at different photon numbers in Fig. 6.3 (a), (b) and (c), respectively. By explicitly considering Poisson noise model, TVSTORM is relatively insensitive to the SNR and allows the use of fluorophores with a lower photon yield.

### 6.3.3 Block Width Selection

The image acquired at the camera is very large and cannot be handled as a whole. It is typically divided into small overlapping blocks which are analyzed independently to improve the computation time. To avoid compromising the performance, blocks are overlapped and the block width is set to be larger than the PSF width. Here we study the effect of block width on the performance of TVSTORM by simulating an image of size 45 $\mu m$ $\times$ 45 $\mu m$ and analyzing the image under different block widths. The photon number is set to 500 and the density is 6 emitters/$\mu m^3$. In our simulation, we first divide the total image into non-overlapping blocks of size $w \times w$, where $w$ is the block width. For each block, a margin of 3 pixels in each direction is added, creating an extended block of size $(w + 6) \times (w + 6)$. Then we

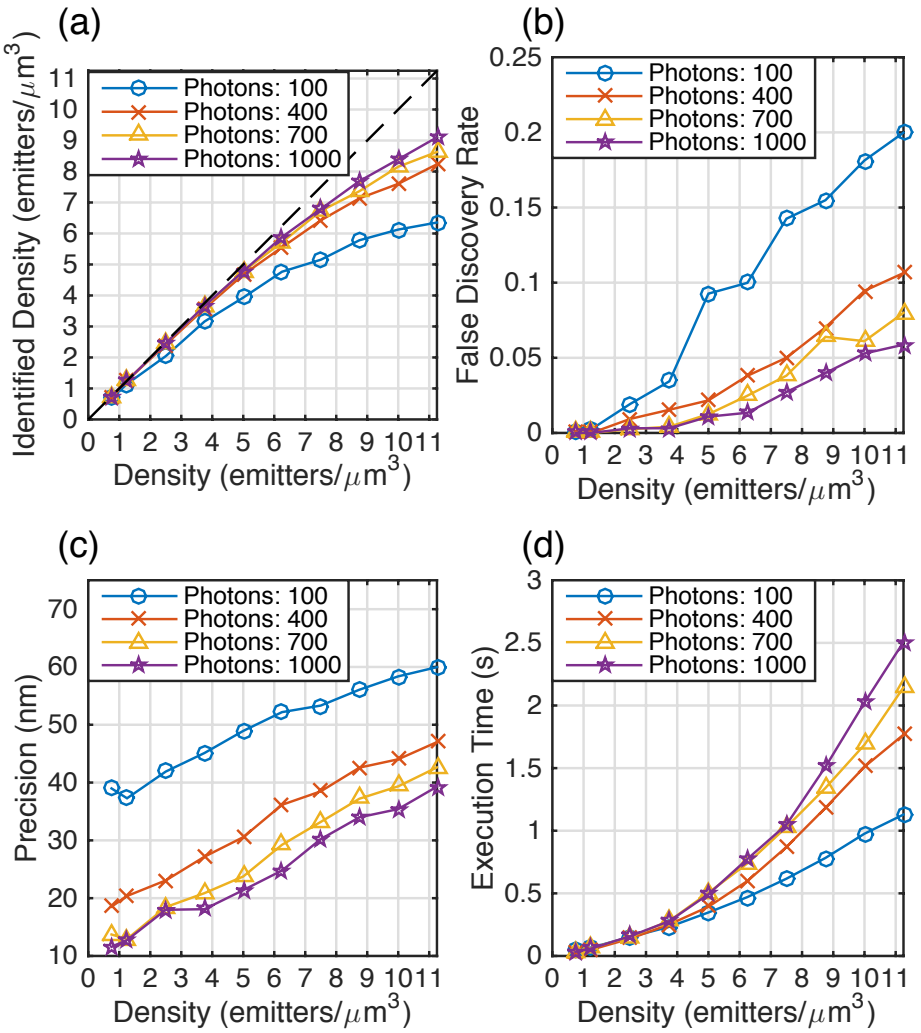



Figure 6.3: Performance of 3D image reconstruction across different photon levels: (a) identified density, (b) false discovery rate, (c) precision and (d) execution time with respect to the emitter density.

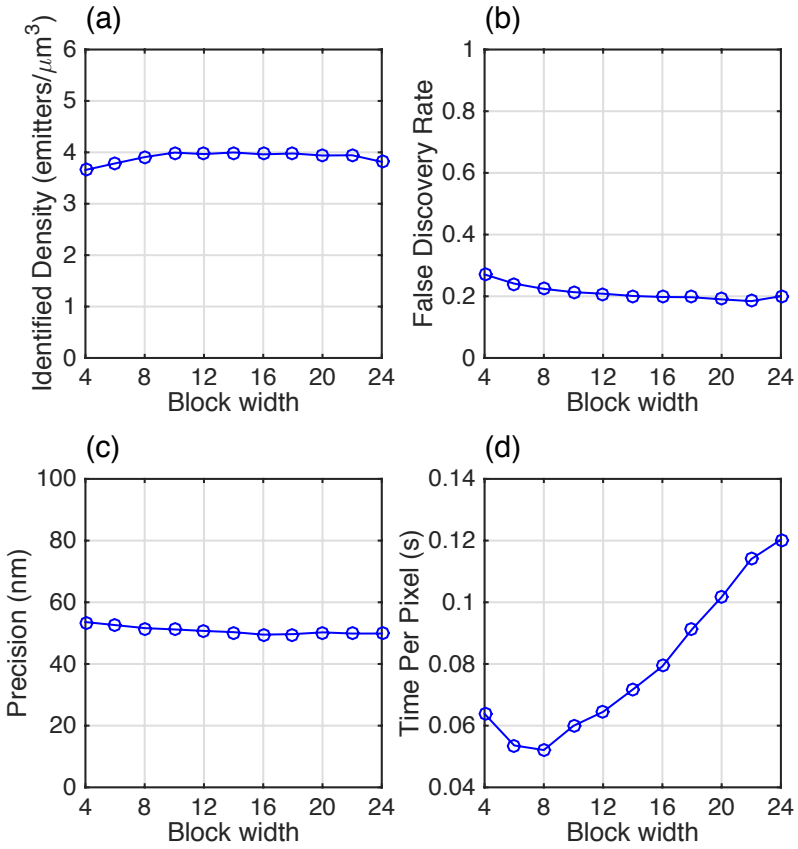



Figure 6.4: Performance of 3D image reconstruction across different block widths: (a) identified density, (b) false discovery rate, (c) precision and (d) execution time per pixel.

apply TVSTORM on the extended block and only emitters within the central $w \times w$ region are kept to avoid the duplicate emitter identification. Fig. 6.4 (a), (b) and (c) show that the identified density, false discovery rate and precision maintain stable as the block width changes. However, the execution time per pixel decreases first and then increases with the block width, as shown in Fig. 6.4 (d). The lowest execution time per pixel occurs when the block width is 8.

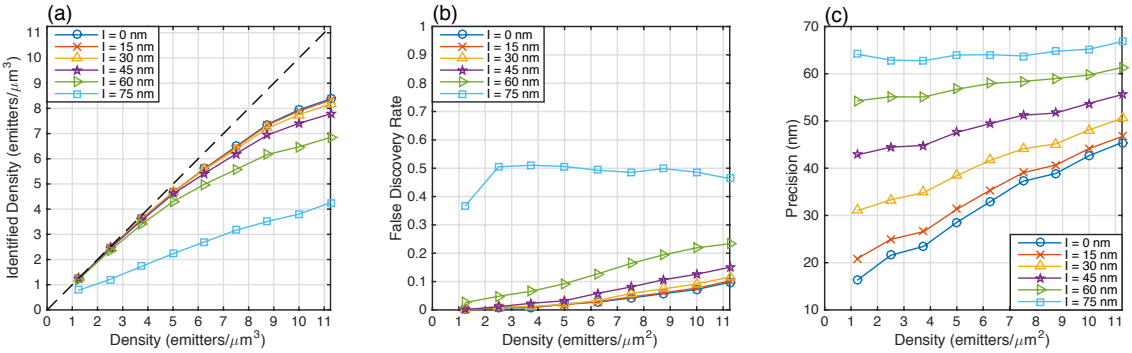



Figure 6.5: 3D image reconstruction performance comparisons across different mismatch level: (a) identified density, (b) false discovery rate, (c) precision.

### 6.3.4 Model Mismatch

In practice, due to the misalignment and lens imperfection, the center coordinate of the emitter's PSF suffers from small drifts as the emitter moves along the axial direction [94]. In specific, (6.3) becomes

$$\mu[u,v|\chi] = \sum_{i=1}^{P} \frac{\theta_I^{(i)}}{4} \left[ Q\left( \frac{u-\theta_x^{(i)} - \Delta_x(\theta_z^{(i)}) + \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) - Q\left( \frac{u-\theta_x^{(i)} - \Delta_x(\theta_z^{(i)}) - \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) \right] \times \left[ Q\left( \frac{v-\theta_y^{(i)} - \Delta_y(\theta_z^{(i)}) + \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) - Q\left( \frac{v-\theta_y^{(i)} - \Delta_y(\theta_z^{(i)}) - \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) \right],$$

where $\Delta_x(\theta_z^{(i)})$ and $\Delta_y(\theta_z^{(i)})$ are not known. To study the effect of this model mismatch on TVSTORM, we generate images under different mismatch level and analyze them using TVSTORM. The deviation of the emitter lateral coordinate is generated as $\Delta_x(z) = I\sin(5z)$ and $\Delta_y(z) = I\cos(5z)$, where we vary $I$ as different mismatch level. As shown in Fig. 6.5, as $I$ increases, the performance of TVSTORM decreases within a tolerable range until $I$ reaches about 75 $nm$. This indicates that calibration of the deviation along the axial direction is essential for accurate performance of

super-resolution algorithms, which can be performed using techniques for example in [94], and extra care should be taken when a large deviation is expected.

### 6.3.5 Performance Comparisons of 2D Image Reconstruction

Although TVSTORM is designed for 3D image reconstruction, it can be easily modified for the 2D setting. Here we study the performance of TVSTORM for 2D image reconstruction. We generate a series of simulated SMLM images across a range of emitter densities (1 emitter/$\mu m^2$ to 9 emitters/$\mu m^2$). The average number of photons for each emitter is 500. We also apply CSSTORM [59] and MempSTORM [11] on the same images for comparison. An up-sampling factor of 8 is used in CSSTORM.

Fig. 6.6 compares the performance of TVSTORM with CSSTORM and MempSTORM in terms of identified density, false discovery rate, precision and execution time with respect to the emitter density. TVSTORM demonstrates the best performance in terms of identified density, false discovery rate and precision. It is also an order-of-magnitude faster than CSSTORM but is slower than MempSTORM, which is known for its extreme fast execution time.

### 6.3.6 Real Experiments

To demonstrate the practical applicability of TVSTORM, both 2D and 3D SMLM images of microtubules stained with Alexa 647 in HeLa cells are acquired with the same system setup and imaging condition described earlier and analyzed with TVSTORM. Since typically the image is not homogenous in certain areas, e.g., center

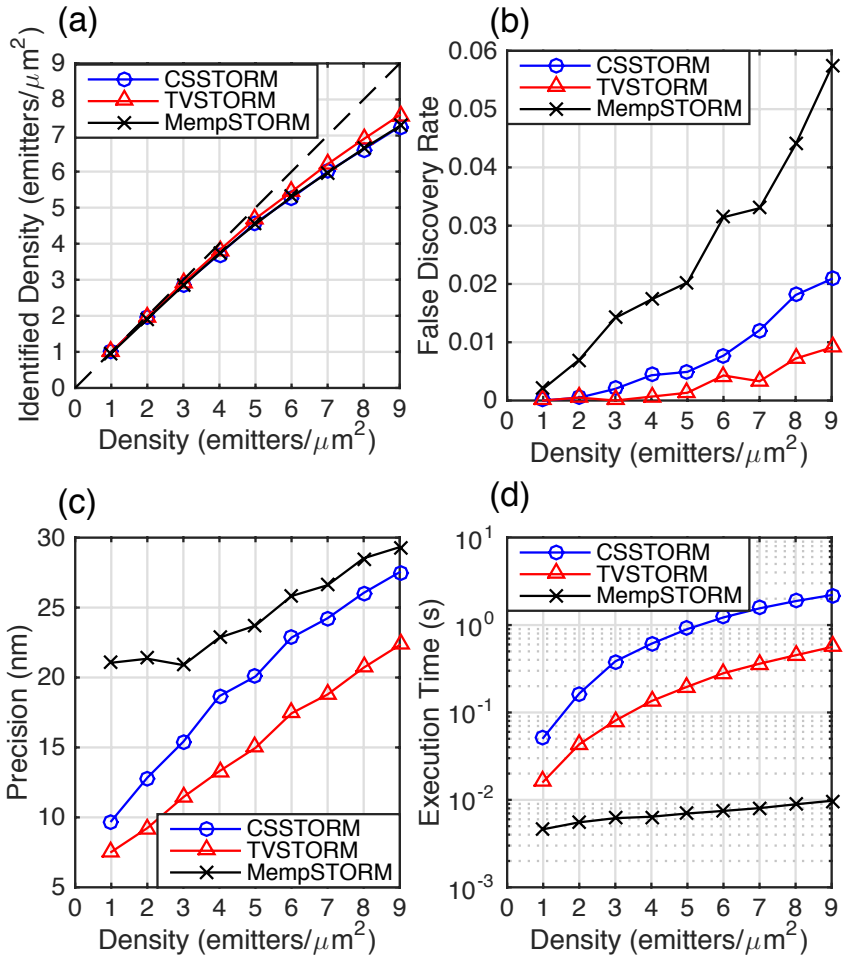



Figure 6.6: Performance comparisons of 2D image reconstruction between TVSTORM, CSSTORM and MempSTORM: (a) identified density, (b) false discovery rate, (c) precision and (d) execution time with respect to the emitter density.

region of the cell, a background term is added in our model (6.3) to take this inhomogeneity into account, i.e.,

$$\mu[u,v|\chi] = \sum_{i=1}^{P} \frac{\theta_I^{(i)}}{4} \left[ Q\left( \frac{u-\theta_x^{(i)}+\frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) - Q\left( \frac{u-\theta_x^{(i)}-\frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \right) \right] \times \left[ Q\left( \frac{v-\theta_y^{(i)}+\frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) - Q\left( \frac{v-\theta_y^{(i)}-\frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})} \right) \right] + \theta_b, \tag{6.10}$$

where $\theta_b$ is an extra parameter to be estimated. This change accounts for various background across different blocks. It also results in a straightforward modification in Algorithm 4, where we apply gradient descent using backtracking line search to refine $\theta_b$.

Fig. 6.7 (a) shows the averaged image from 10000 frames of SMLM images and Fig. 6.7 (b) is the 3D super-resolution image reconstructed using TVSTORM, where the structure of 3D microtubule can be well resolved with the axial coordinate represented in different colors. Similarly, the averaged image and reconstructed super-resolution image for 2D STORM imaging are shown in Fig. 6.8 (a) and (b), respectively, and demonstrate high quality reconstruction and fine details that cannot be resolved in the averaged low-resolution image.

## 6.4 Summary

In this chapter, TVSTORM is proposed for 3D super-resolution image reconstruction, which is a penalized maximum likelihood estimator under the Poisson noise by the total variation norm of the activated emitters. TVSTORM avoids the intrinsic bias of CSSTORM due to gridding, and is computationally more efficient, with better detection rate, false discovery rate, and precision. Furthermore, TVSTORM can be easily adapted to 2D super-resolution image reconstruction or other single-molecule

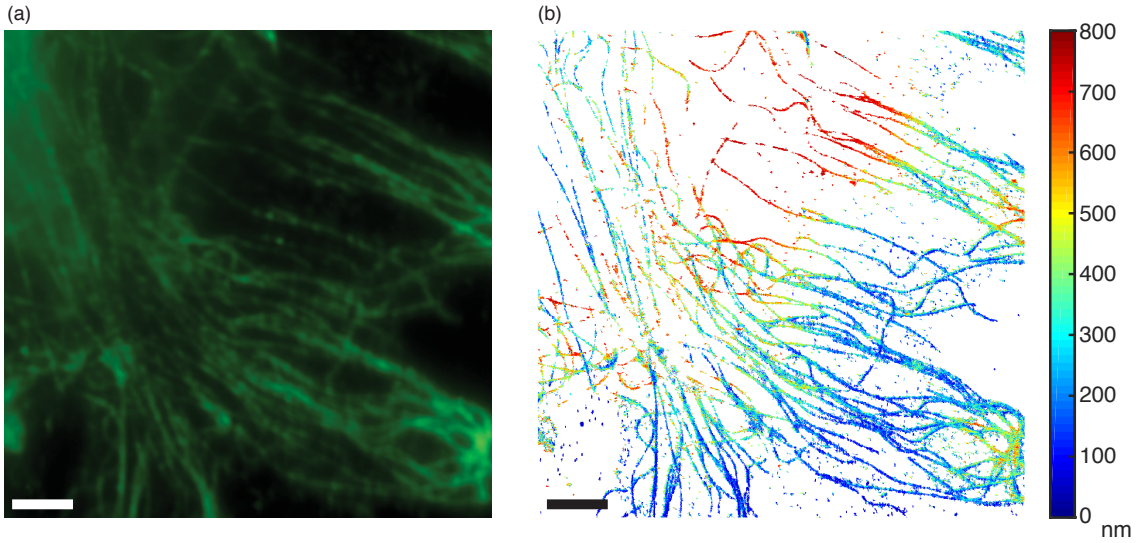



Figure 6.7: (a) Averaged image from 10000 frames and (b) 3D image reconstruction result of microtubules stained with Alexa 647. (bar: 3 $\mu m$)

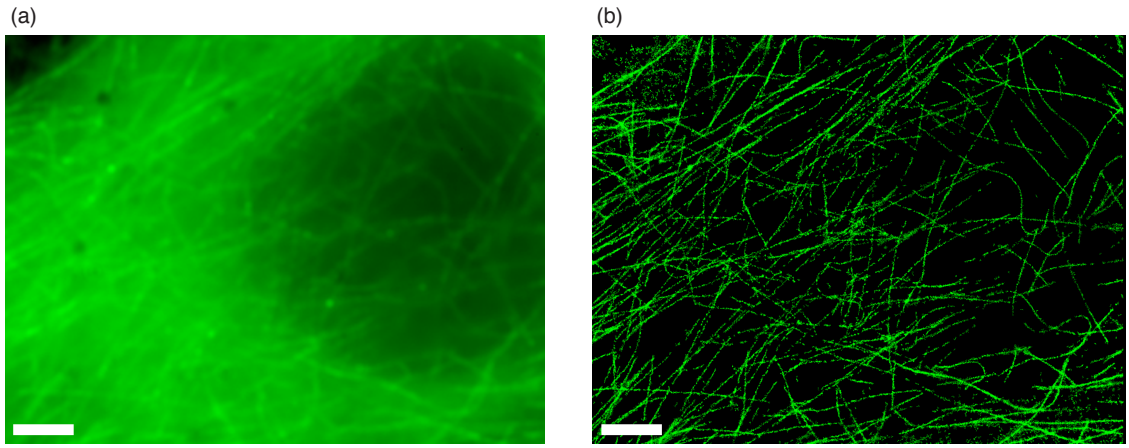



Figure 6.8: (a) Averaged image from 10000 frames and (b) 2D image reconstruction result of microtubules stained with Alexa 647. (bar: 3 $\mu m$)

microscopy with different PSF configurations. Extensive simulation results are provided to demonstrate the superior performance of TVSTORM. Experimental results are also present to show its practical applicability.

# Chapter 7: Conclusions

The ability to observe finer structures have been advancing the biological field for a long time and will continue to do so. The invention of single-molecule based super-resolution microscopy provides a new means to study cellular processes at the single molecule level. As this technique continues to be adopted by more and more labs, it becomes increasingly important to develop reconstruction algorithms and imaging platforms that can improve the super-resolution image quality.

## 7.1 Summary of the Work

In this dissertation, we presented fast and high-performance super-resolution image reconstruction algorithms that can analyze images of high emitter density reliably in both 2D and 3D cases, as well as new imaging modality that can improve the super-resolution imaging performance. These innovations not only improve the temporal resolution and reconstruction quality of super-resolution microscopy, but also provide a better tool for biological researchers.

For 2D image reconstruction, we transformed the super-resolution imaging model to the frequency domain and turned the problem of emitter localization into 2D spectrum estimation, a problem often encountered in signal processing. We developed an algorithm (MempSTORM) based on a 2D spectrum-estimation method to extract

the number of emitters and their positions by determining the 2D frequencies, which speeds up the execution time over two orders of magnitude.

For 3D imaging platform, we developed a multi-focus astigmatism compressive sensing based 3D imaging platform that allows us to efficiently locate high density emitters. We applied a multi-focus technique, in combination with astigmatic optics to enhance the asymmetry of the PSF along the axial direction. Accordingly, we developed a new algorithm that combines $\ell_1$ minimization and a debiasing method, which not only reduces the bias introduced by gridding but also significantly increases computation speed.

For 3D image reconstruction, we proposed TVSTORM, a penalized maximum likelihood estimator under the Poisson noise regularized by the total variation norm of the activated emitters. TVSTORM avoids the intrinsic bias caused by gridding, and is computationally more efficient, with better detection rate, false discovery rate, and precision. Furthermore, TVSTORM can be easily adapted to 2D super-resolution image reconstruction or other single-molecule microscopy with different PSF configurations.

## 7.2 Possible Future Research

Possible future research directions include incorporating the temporal information of SMLM images across neighboring frames into reconstruction algorithms. The time-varying signal in super-resolution microscopy has more structure than mere sparsity. The signals of two consecutive frames have a strong correlation due to the short exposure time. Most of the positions of emitters do not change across frames and

such structure has not been utilized by previous researches. The incorporation of such information could potentially further improve the image reconstruction quality.

Another direction is to study the reconstruction algorithms for adaptive optics based 3D SMLM imaging. In practice, due to optical imperfections and sample scattering, aberrations of the system PSF exist, deteriorating the image quality and reconstruction accuracy. Adaptive optics uses a deformable mirror to actively correct the wavefront distortion, allowing reshaping the system PSF and correcting aberrations. With adaptive optics, the ideal PSF shape for optimal localization accuracy could be achieved and the imaging depth could be further improved.

# Appendix A: Derivation in TVSTORM

In this appendix, we derive the partial derivatives of $\ell(y|\chi(\Theta))$ over every parameter $\theta_j^{(i)}$, which is required for the REFINE procedure as described in Algorithm 4. To begin with, we have

$$\begin{aligned}\frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_j^{(i)}} &= \sum_m \sum_n \left( -\frac{y[m,n]}{\mu[m,n|\Theta]} \frac{\partial \mu[m,n|\Theta]}{\partial \theta_j^{(i)}} + \frac{\partial \mu[m,n|\Theta]}{\partial \theta_j^{(i)}} \right) \\ &= \sum_u \sum_v \left( \frac{\mu[m,n|\Theta] - y[m,n]}{\mu[m,n|\Theta]} \frac{\partial \mu[m,n|\Theta]}{\partial \theta_j^{(i)}} \right). \end{aligned} \tag{A.1}$$

We then derive the partial derivative of $\mu[m,n|\Theta]$ over $\theta_x^{(i)}$, $\theta_y^{(i)}$, $\theta_z^{(i)}$, and $\theta_I^{(i)}$. Let

$$\begin{aligned}\tilde{m}_+^{(i)} &= \frac{m - \theta_x^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})}, \tilde{m}_-^{(i)} = \frac{m - \theta_x^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_x(\theta_z^{(i)})}, \\ \tilde{n}_+^{(i)} &= \frac{n - \theta_y^{(i)} + \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})}, \tilde{n}_-^{(i)} = \frac{n - \theta_y^{(i)} - \frac{1}{2}}{\sqrt{2}\sigma_y(\theta_z^{(i)})}.\end{aligned}$$

Then according to (6.3) and following simple calculus, we have

$$\begin{aligned}\frac{\partial \mu[m,n|\Theta]}{\partial \theta_x^{(i)}} &= \frac{\theta_I^{(i)}}{4} \left[ \frac{\partial \left( Q\left(\tilde{m}_+^{(i)}\right) - Q\left(\tilde{m}_-^{(i)}\right) \right)}{\partial \theta_x^{(i)}} \right] \left[ Q\left(\tilde{n}_+^{(i)}\right) - Q\left(\tilde{n}_-^{(i)}\right) \right] \\ &= \frac{\theta_I^{(i)}}{4} \left[ -\frac{2}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{2}\sigma_x(\theta_z^{(i)})} \left( e^{-\left(\tilde{m}_+^{(i)}\right)^2} - e^{-\left(\tilde{m}_-^{(i)}\right)^2} \right) \right] \left[ Q\left(\tilde{n}_+^{(i)}\right) - Q\left(\tilde{n}_-^{(i)}\right) \right] \\ &= \frac{\theta_I^{(i)}}{2\sqrt{2\pi}\sigma_x(\theta_z^{(i)})} \left[ e^{-\left(\tilde{m}_-^{(i)}\right)^2} - e^{-\left(\tilde{m}_+^{(i)}\right)^2} \right] \times \left[ Q\left(\tilde{n}_+^{(i)}\right) - Q\left(\tilde{n}_-^{(i)}\right) \right]. \end{aligned} \tag{A.2}$$

Similarly,

$$\frac{\partial \mu[m,n|\Theta]}{\partial \theta_y^{(i)}} = \frac{\theta_I^{(i)}}{2\sqrt{2\pi}\sigma_y(\theta_z^{(i)})} \left[ Q\left(\tilde{m}_+^{(i)}\right) - Q\left(\tilde{m}_-^{(i)}\right) \right] \times \left[ e^{-\left(\tilde{n}_-^{(i)}\right)^2} - e^{-\left(\tilde{n}_+^{(i)}\right)^2} \right]. \tag{A.3}$$

Additionally, using the chain rule, $\frac{\partial\mu[m,n|\Theta]}{\partial\theta_z^{(i)}}$ can be written as:

$$\frac{\partial\mu[m,n|\Theta]}{\partial\theta_z^{(i)}} = \frac{\partial\mu[m,n|\Theta]}{\partial\sigma_x(\theta_z^{(i)})}\frac{\partial\sigma_x(\theta_z^{(i)})}{\partial\theta_z^{(i)}} + \frac{\partial\mu[m,n|\Theta]}{\partial\sigma_y(\theta_z^{(i)})}\frac{\partial\sigma_y(\theta_z^{(i)})}{\partial\theta_z^{(i)}}, \tag{A.4}$$

where

$$\begin{aligned}
\frac{\partial\sigma_x(\theta_z^{(i)})}{\partial\theta_z^{(i)}} &= \frac{\sigma_{x,0}\left(\sum_{i=2}^{4} i\cdot A_{x,i}\left(\frac{z_0-c_x}{d_x}\right)^{i-1}\right)}{2\sqrt{1+\sum_{i=2}^{4} A_{x,i}\left(\frac{z_0-c_x}{d_x}\right)^{i}}}, \\
\frac{\partial\sigma_y(\theta_z^{(i)})}{\partial\theta_z^{(i)}} &= \frac{\sigma_{y,0}\left(\sum_{i=2}^{4} i\cdot A_{y,i}\left(\frac{z_0-c_y}{d_y}\right)^{i-1}\right)}{2\sqrt{1+\sum_{i=2}^{4} A_{y,i}\left(\frac{z_0-c_y}{d_y}\right)^{i}}}.
\end{aligned} \tag{A.5}$$

From (6.3), $\frac{\partial\mu[m,n|\Theta]}{\partial\sigma_x(\theta_z^{(i)})}$ can be calculated as:

$$\begin{aligned}
\frac{\partial\mu[m,n|\Theta]}{\partial\sigma_x(\theta_z^{(i)})} =& \frac{\theta_I^{(i)}}{4}\left(\frac{\partial\left[Q\left(\tilde{m}_+^{(i)}\right)-Q\left(\tilde{m}_-^{(i)}\right)\right]}{\partial\sigma_x(\theta_z^{(i)})}\right)\left[Q\left(\tilde{n}_+^{(i)}\right)-Q\left(\tilde{n}_-^{(i)}\right)\right] \\
=& \frac{\theta_I^{(i)}}{4}\left[\frac{2}{\sqrt{\pi}}\cdot\left(e^{-\left(\tilde{m}_+^{(i)}\right)^2}\cdot-\tilde{m}_+^{(i)}-e^{-\left(\tilde{m}_-^{(i)}\right)^2}\cdot-\tilde{m}_-^{(i)}\right)\right]\times\left[Q\left(\tilde{n}_+^{(i)}\right)-Q\left(\tilde{n}_-^{(i)}\right)\right] \\
=& \frac{\theta_I^{(i)}}{2\sqrt{\pi}\sigma_x(\theta_z^{(i)})}\left[e^{-(\tilde{m}_-^{(i)})^2}\cdot\tilde{m}_-^{(i)}-e^{-(\tilde{m}_+^{(i)})^2}\cdot\tilde{m}_+^{(i)}\right]\times\left[Q\left(\tilde{n}_+^{(i)}\right)-Q\left(\tilde{n}_-^{(i)}\right)\right].
\end{aligned} \tag{A.6}$$

Similarly,

$$\frac{\partial\mu[m,n|\Theta]}{\partial\sigma_y(\theta_z^{(i)})} = \frac{\theta_I^{(i)}}{2\sqrt{\pi}\sigma_x(\theta_z^{(i)})}\left[Q\left(\tilde{m}_+^{(i)}\right)-Q\left(\tilde{m}_-^{(i)}\right)\right]\times\left[e^{-(\tilde{n}_-^{(i)})^2}\cdot\tilde{n}_-^{(i)}-e^{-(\tilde{n}_+^{(i)})^2}\cdot\tilde{n}_+^{(i)}\right]. \tag{A.7}$$

Lastly, it is easy to derive $\frac{\partial\mu[m,n|\Theta]}{\partial\theta_I^{(i)}}$ as:

$$\frac{\partial\mu[m,n|\Theta]}{\partial\theta_I^{(i)}} = \frac{1}{4}\left[Q\left(\tilde{m}_+^{(i)}\right)-Q\left(\tilde{m}_-^{(i)}\right)\right]\left[Q\left(\tilde{n}_+^{(i)}\right)-Q\left(\tilde{n}_-^{(i)}\right)\right]. \tag{A.8}$$

Plugging (A.2), (A.3), (A.4) and (A.8) into (A.1), we obtain the partial derivative of the loss function $\ell(y|\chi(\Theta))$ over $\theta_x^{(i)}$, $\theta_y^{(i)}$, $\theta_z^{(i)}$, and $\theta_I^{(i)}$ as in (A.9), (A.10), (A.11) and (A.12), respectively.

$$\frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_x^{(i)}} = \frac{\theta_I^{(i)}}{2\sqrt{2\pi}\sigma_x(\theta_z^{(i)})} \sum_u \sum_v \left[\frac{\mu[m,n|\Theta] - y[m,n]}{\mu[m,n|\Theta]}\right] \times \left[e^{-\left(\tilde{m}_-^{(i)}\right)^2} - e^{-\left(\tilde{m}_+^{(i)}\right)^2}\right] \times \left[Q\left(\tilde{n}_+^{(i)}\right) - Q\left(\tilde{n}_-^{(i)}\right)\right], \tag{A.9}$$

$$\frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_y^{(i)}} = \frac{\theta_I^{(i)}}{2\sqrt{2\pi}\sigma_y(\theta_z^{(i)})} \sum_u \sum_v \left[\frac{\mu[m,n|\Theta] - y[m,n]}{\mu[m,n|\Theta]}\right] \times \left[Q\left(\tilde{m}_+^{(i)}\right) - Q\left(\tilde{m}_-^{(i)}\right)\right] \times \left[e^{-\left(\tilde{n}_-^{(i)}\right)^2} - e^{-\left(\tilde{n}_+^{(i)}\right)^2}\right], \tag{A.10}$$

$$\frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_I^{(i)}} = \frac{1}{4} \sum_u \sum_v \left[\frac{\mu[m,n|\Theta] - y[m,n]}{\mu[m,n|\Theta]}\right] \times \left[Q\left(\tilde{m}_+^{(i)}\right) - Q\left(\tilde{m}_-^{(i)}\right)\right] \times \left[Q\left(\tilde{n}_+^{(i)}\right) - Q\left(\tilde{n}_-^{(i)}\right)\right], \tag{A.11}$$

$$\frac{\partial \ell(y|\chi(\Theta))}{\partial \theta_z^{(i)}} = \sum_u \sum_v \left[\frac{\mu[m,n|\Theta] - y[m,n]}{\mu[m,n|\Theta]}\right] \times \left[\frac{\partial \mu[m,n|\Theta]}{\partial \sigma_x(\theta_z^{(i)})} \frac{\partial \sigma_x(\theta_z^{(i)})}{\partial \theta_z^{(i)}} + \frac{\partial \mu[m,n|\Theta]}{\partial \sigma_y(\theta_z^{(i)})} \frac{\partial \sigma_y(\theta_z^{(i)})}{\partial \theta_z^{(i)}}\right]. \tag{A.12}$$

# Bibliography


[1] M. G. Gustafsson, "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 37, pp. 13 081–13 086, 2005.

[2] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Optics letters*, vol. 19, no. 11, pp. 780–782, 1994.

[3] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.

[4] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature methods*, vol. 3, no. 10, pp. 793–796, 2006.

[5] E. Betzig, S. W. Hell, and W. E. Moerner, "The nobel prize in chemistry 2014," *Nobel Media AB*, 2014.

[6] C. Franzini-Armstrong, J. Heuser, T. Reese, A. Somlyo, and A. Somlyo, "T-tubule swelling in hypertonic solutions: a freeze substitution study." *The Journal of physiology*, vol. 283, p. 133, 1978.

[7] M. Sun, J. Huang, F. Bunyak, K. Gumpper, G. De, M. Sermersheim, G. Liu, P.-H. Lin, K. Palaniappan, and J. Ma, "Superresolution microscope image reconstruction by spatiotemporal object decomposition and association: application in resolving t-tubule structure in skeletal muscle," *Opt. Express*, vol. 22, no. 10, pp. 12 160–12 176, 2014.

[8] B. Huang, S. A. Jones, B. Brandenburg, and X. Zhuang, "Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution," *Nature methods*, vol. 5, no. 12, pp. 1047–1052, 2008.

[9] C. Cai, H. Masumiya, N. Weisleder, N. Matsuda, M. Nishi, M. Hwang, J.-K. Ko, P. Lin, A. Thornton, X. Zhao, Z. Pan, S. Komazaki, M. Brotto, H. Takeshima, and J. Ma, "MG53 nucleates assembly of cell membrane repair machinery," *Nature cell biology*, vol. 11, no. 1, pp. 56–64, 2009.

[10] S. A. Jones, S.-H. Shim, J. He, and X. Zhuang, "Fast, three-dimensional super-resolution imaging of live cells," *Nature methods*, vol. 8, no. 6, pp. 499–505, 2011.

[11] J. Huang, K. Gumpper, Y. Chi, M. Sun, and J. Ma, "Fast two-dimensional super-resolution image reconstruction algorithm for ultra-high emitter density," *Optics Letters*, vol. 40, no. 13, pp. 2989–2992, 2015.

[12] J. Huang, M. Sun, K. Gumpper, Y. Chi, and J. Ma, "3D multifocus astigmatism and compressed sensing (3D MACS) based superresolution reconstruction," *Biomedical optics express*, vol. 6, no. 3, pp. 902–917, 2015.

[13] J. Huang, M. Sun, and Y. Chi, "Super-resolution image reconstruction for high-density 3D single-molecule microscopy," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*. IEEE, 2016.

[14] J. Huang, M. Sun, J. Ma, and Y. Chi, "Super-resolution image reconstruction for high-density 3D single-molecule microscopy," *Computational Imaging, IEEE Transactions on*, submitted.

[15] F. Zernike, "How I discovered phase contrast," *Science*, vol. 121, no. 3141, pp. 345–349, 1955.

[16] W. C. McCrone, L. B. McCrone, and J. G. Delly, *Polarized light microscopy*. Ann Arbor Science Publishers Inc. and McCrone Research Institute, 1978.

[17] M. Pluta, "Nomarski's dic microscopy: a review," in *Phase Contrast and Differential Interference Contrast Imaging Techniques and Applications*. International Society for Optics and Photonics, 1994, pp. 10–25.

[18] G. G. Stokes, "On the change of refrangibility of light," *Philosophical Transactions of the Royal Society of London*, vol. 142, pp. 463–562, 1852.

[19] A. Jabloński, "General theory of pressure broadening of spectral lines," *Physical Review*, vol. 68, no. 3-4, p. 78, 1945.

[20] J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy," *Nature methods*, vol. 2, no. 12, pp. 910–919, 2005.

[21] T. R. Y. Shimomura, Osamu and M. Chalfie, "The nobel prize in chemistry 2008," *Nobel Media AB*, 2008.

[22] U. Kubitscheck, *Fluorescence microscopy: from principles to biological applications*. John Wiley & Sons, 2013.

[23] J. W. Goodman and S. C. Gustafson, "Introduction to fourier optics," *Optical Engineering*, vol. 35, no. 5, pp. 1513–1513, 1996.

[24] E. Abbe, "Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung," *Archiv für mikroskopische Anatomie*, vol. 9, no. 1, pp. 413–418, 1873.

[25] L. Rayleigh, "Investigations in optics, with special reference to the spectroscope," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 8, no. 49, pp. 261–274, 1879.

[26] B. Huang, M. Bates, and X. Zhuang, "Super resolution fluorescence microscopy," *Annual review of biochemistry*, vol. 78, p. 993, 2009.

[27] D. Axelrod, "Cell-substrate contacts illuminated by total internal reflection fluorescence." *The Journal of cell biology*, vol. 89, no. 1, pp. 141–145, 1981.

[28] M. Minsky, "Microscopy apparatus," Dec. 19 1961, uS Patent 3,013,467.

[29] W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: multiphoton microscopy in the biosciences," *Nature biotechnology*, vol. 21, no. 11, pp. 1369–1377, 2003.

[30] S. Hell and E. H. Stelzer, "Fundamental improvement of resolution with a 4pi-confocal fluorescence microscope using two-photon excitation," *Optics Communications*, vol. 93, no. 5-6, pp. 277–282, 1992.

[31] M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *Journal of microscopy*, vol. 198, no. 2, pp. 82–87, 2000.

[32] M. Gustafsson, D. Agard, and J. Sedat, "I5M: 3D widefield light microscopy with better than 100nm axial resolution," *Journal of microscopy*, vol. 195, no. 1, pp. 10–16, 1999.

[33] L. Shao, B. Isaac, S. Uzawa, D. A. Agard, J. W. Sedat, and M. G. Gustafsson, "I$^5$S: wide-field light microscopy with 100-nm-scale resolution in three dimensions," *Biophysical journal*, vol. 94, no. 12, pp. 4971–4983, 2008.

[34] J. Fölling, M. Bossi, H. Bock, R. Medda, C. A. Wurm, B. Hein, S. Jakobs, C. Eggeling, and S. W. Hell, "Fluorescence nanoscopy by ground-state depletion and single-molecule return," *Nature methods*, vol. 5, no. 11, pp. 943–945, 2008.

[35] V. Westphal and S. W. Hell, "Nanoscale resolution in the focal plane of an optical microscope," *Physical review letters*, vol. 94, no. 14, p. 143903, 2005.

[36] M. Heilemann, S. van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, "Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes," *Angewandte Chemie International Edition*, vol. 47, no. 33, pp. 6172–6176, 2008.

[37] M. Heilemann, S. van de Linde, A. Mukherjee, and M. Sauer, "Super-resolution imaging with small organic fluorophores," *Angewandte Chemie International Edition*, vol. 48, no. 37, pp. 6903–6908, 2009.

[38] S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical journal*, vol. 91, no. 11, pp. 4258–4272, 2006.

[39] J. C. Vaughan, S. Jia, and X. Zhuang, "Ultrabright photoactivatable fluorophores created by reductive caging," *Nature methods*, vol. 9, no. 12, pp. 1181–1184, 2012.

[40] A. Pertsinidis, K. Mukherjee, M. Sharma, Z. P. Pang, S. R. Park, Y. Zhang, A. T. Brunger, T. C. Südhof, and S. Chu, "Ultrahigh-resolution imaging reveals formation of neuronal snare/munc18 complexes in situ," *Proceedings of the National Academy of Sciences*, vol. 110, no. 30, pp. E2812–E2820, 2013.

[41] A. Pertsinidis, Y. Zhang, and S. Chu, "Subnanometre single-molecule localization, registration and distance measurements," *Nature*, vol. 466, no. 7306, pp. 647–651, 2010.

[42] L. S. Barak and W. W. Webb, "Fluorescent low density lipoprotein for observation of dynamics of individual receptor complexes on cultured human fibroblasts," *The Journal of cell biology*, vol. 90, no. 3, pp. 595–604, 1981.

[43] ——, "Diffusion of low density lipoprotein-receptor complex on human fibroblasts," *The Journal of cell biology*, vol. 95, no. 3, pp. 846–852, 1982.

[44] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophysical journal*, vol. 82, no. 5, pp. 2775–2783, 2002.

[45] M. Bates, B. Huang, and X. Zhuang, "Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes," *Current opinion in chemical biology*, vol. 12, no. 5, pp. 505–514, 2008.

[46] M. Bates, S. A. Jones, and X. Zhuang, "Stochastic optical reconstruction microscopy (STORM): a method for superresolution fluorescence imaging," *Cold Spring Harbor Protocols*, vol. 2013, no. 6, pp. pdb–top075 143, 2013.

[47] D. Baddeley, M. B. Cannell, and C. Soeller, "Visualization of localization microscopy data," *Microscopy and Microanalysis*, vol. 16, no. 01, pp. 64–72, 2010.

[48] K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nature methods*, vol. 7, no. 5, pp. 377–381, 2010.

[49] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[50] S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. Moerner, "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function," *Proceedings of the National Academy of Sciences*, vol. 106, no. 9, pp. 2995–2999, 2009.

[51] B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems. ii. structure of the image field in an aplanatic system," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 253, no. 1274. The Royal Society, 1959, pp. 358–379.

[52] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *JOSA A*, vol. 8, no. 10, pp. 1601–1613, 1991.

[53] A. V. Abraham, S. Ram, J. Chao, E. Ward, and R. J. Ober, "Quantitative study of single molecule location estimation techniques," *Optics express*, vol. 17, no. 26, pp. 23 352–23 373, 2009.

[54] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics*, vol. 46, no. 10, pp. 1819–1829, 2007.

[55] C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nature methods*, vol. 7, no. 5, pp. 373–375, 2010.

[56] R. Henriques, M. Lelek, E. F. Fornasiero, F. Valtorta, C. Zimmer, and M. M. Mhlanga, "QuickPALM: 3D real-time photoactivation nanoscopy image processing in imagej," *Nature methods*, vol. 7, no. 5, pp. 339–340, 2010.

[57] S. Wolter, A. Löschberger, T. Holm, S. Aufmkolk, M.-C. Dabauvalle, S. van de Linde, and M. Sauer, "rapidSTORM: accurate, fast open-source software for localization microscopy," *Nature methods*, vol. 9, no. 11, pp. 1040–1041, 2012.

[58] R. Parthasarathy, "Rapid, accurate particle tracking by calculation of radial symmetry centers," *Nature Methods*, vol. 9, no. 7, pp. 724–726, 2012.

[59] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nature methods*, vol. 9, no. 7, pp. 721–723, 2012.

[60] S. J. Holden, S. Uphoff, and A. N. Kapanidis, "DAOSTORM: an algorithm for high-density super-resolution microscopy," *Nat. Meth.*, vol. 8, no. 4, pp. 279–280, 2011.

[61] E. A. Mukamel, H. Babcock, and X. Zhuang, "Statistical deconvolution for super-resolution fluorescence microscopy," *Biophys. J.*, vol. 102, no. 10, pp. 2391–2400, 2012.

[62] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[63] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[64] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[65] E. J. Candès, M. Rudelson, T. Tao, and R. Vershynin, "Error correction via linear programming," in *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on.* IEEE, 2005, pp. 668–681.

[66] J. F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, vol. 38, no. 5, pp. 826–844, 1973.

[67] F. Santosa and W. W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 4, pp. 1307–1330, 1986.

[68] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[69] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[70] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[71] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[72] G. Tang, B. N. Bhaskar, and B. Recht, "Sparse recovery over continuous dictionaries-just discretize," in *Signals, Systems and Computers, 2013 Asilomar Conference on*. IEEE, 2013, pp. 1043–1047.

[73] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Process*, vol. 40, no. 9, pp. 2267–2280, 1992.

[74] H. P. Babcock, J. R. Moffitt, Y. Cao, and X. Zhuang, "Fast compressed sensing analysis for super-resolution imaging using L1-homotopy," *Optics express*, vol. 21, no. 23, pp. 28 583–28 596, 2013.

[75] A. Goshtasby, "Image registration by local approximation methods," *Image and Vision Computing*, vol. 6, no. 4, pp. 255–261, 1988.

[76] L. Gu, Y. Sheng, Y. Chen, H. Chang, Y. Zhang, P. Lv, W. Ji, and T. Xu, "High-density 3D single molecular analysis based on compressed sensing," *Biophysical journal*, vol. 106, no. 11, pp. 2443–2449, 2014.

[77] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008.

[78] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2182–2195, 2011.

[79] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 586–597, 2007.

[80] I. Izeddin, J. Boulanger, V. Racine, C. Specht, A. Kechkar, D. Nair, A. Triller, D. Choquet, M. Dahan, and J. Sibarita, "Wavelet analysis for single molecule localization microscopy," *Optics express*, vol. 20, no. 3, pp. 2081–2095, 2012.

[81] S. Frisken Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *JOSA A*, vol. 8, no. 10, pp. 1601–1613, 1991.

[82] H. Kirshner, F. Aguet, D. Sage, and M. Unser, "3-D PSF fitting for fluorescence microscopy: implementation and localization application," *Journal of microscopy*, vol. 249, no. 1, pp. 13–25, 2013.

[83] J. Min, C. Vonesch, H. Kirshner, L. Carlini, N. Olivier, S. Holden, S. Manley, J. C. Ye, and M. Unser, "FALCON: fast and unbiased reconstruction of high-density super-resolution microscopy data," *Scientific reports*, vol. 4, 2014.

[84] J. Min, S. J. Holden, L. Carlini, M. Unser, S. Manley, and J. C. Ye, "3D high-density localization microscopy using hybrid astigmatic/biplane imaging and sparse image reconstruction," *Biomedical optics express*, vol. 5, no. 11, pp. 3935–3948, 2014.

[85] A. Barsic, G. Grover, and R. Piestun, "Three-dimensional super-resolution and localization of dense clusters of single molecules," *Scientific reports*, vol. 4, 2014.

[86] S. Jia, J. C. Vaughan, and X. Zhuang, "Isotropic three-dimensional super-resolution imaging with a self-bending point spread function," *Nature photonics*, vol. 8, no. 4, pp. 302–306, 2014.

[87] S. Abrahamsson, J. Chen, B. Hajj, S. Stallinga, A. Y. Katsov, J. Wisniewski, G. Mizuguchi, P. Soule, F. Mueller, C. D. Darzacq, X. Darzacq, C. Wu, C. I. Bargmann, D. A. Agard, M. Dahan, and M. G. L. Gustafsson, "Fast multicolor 3D imaging using aberration-corrected multifocus microscopy," *Nature methods*, vol. 10, no. 1, pp. 60–63, 2013.

[88] A. Tahmasbi, S. Ram, J. Chao, A. V. Abraham, F. W. Tang, E. Sally Ward, and R. J. Ober, "Designing the focal plane spacing for multifocal plane microscopy," *Optics express*, vol. 22, no. 14, pp. 16 706–16 721, 2014.

[89] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.

[90] N. Boyd, G. Schiebinger, and B. Recht, "The alternating descent conditional gradient method for sparse inverse problems," *arXiv preprint arXiv:1507.01562*, 2015.

[91] N. Rao, P. Shah, and S. Wright, "Forward-backward greedy algorithms for atomic norm regularization," *Signal Processing, IEEE Transactions on*, vol. 63, no. 21, pp. 5798–5811, 2015.

[92] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: sparse poisson intensity reconstruction algorithms–theory and practice," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 1084–1096, 2012.

[93] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.

[94] M. D. Lew, M. A. Thompson, M. Badieirostami, and W. Moerner, "In vivo three-dimensional superresolution fluorescence tracking using a double-helix point spread function," in *BiOS*. International Society for Optics and Photonics, 2010, pp. 75 710Z–75 710Z.