Improved Localisation and Reconstruction in a Single Molecule Localisation Microscope to Enable Super-Resolution Imaging with Lowered Resource Pre-requisites

Pranjal Choudhury

Roll No: 206121051

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

of the Indian Institute of Technology Guwahati.

Supervisor: Prof. Bosanta Ranjan Boruah

Department of Physics

Indian Institute of Technology Guwahati

Guwahati, Assam, India

January, 2026

Indian Institute of Technology Guwahati

Doctor of Philosophy

Improved Localisation and Reconstruction in a Single Molecule Localisation Microscope to Enable Super-Resolution Imaging with Lowered Resource Pre-requisites

Pranjal Choudhury

Roll No: 206121051

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

of the Indian Institute of Technology Guwahati.

Supervisor: Prof. Bosanta Ranjan Boruah

Department of Physics

Indian Institute of Technology Guwahati

Guwahati, Assam, India, Pin- 781039

January, 2026

Dedicated to my parents, my sister and all the readers …

Declaration

Pranjal Choudhury
Roll No: 206121051
Department of Physics
Indian Institute of Technology Guwahati

I hereby declare that the results embodied in this thesis is the result of theory and experiment carried out by me at the Department of Physics, Indian Institute of Technology Guwahati, India, under the supervision of **Prof. Bosanta Ranjan Boruah**. This thesis has not been submitted to any university/ institute or elsewhere for the award of any degree, diploma or associateship.

Pranjal Choudhury
January, 2026

Certificate

Prof. Bosanta Ranjan Boruah
Department of Physics
Indian Institute of Technology Guwahati
Emails: brboruah@iitg.ac.in

This is to certify that the work contained in the thesis entitled **'Improved Localisation and Reconstruction in a Single Molecule Localisation Microscope to Enable Super-Resolution Imaging with Lowered Resource Pre-requisites'** by **Pranjal Choudhury** (Roll No. 206121051), a student of Department of Physics, Indian Institute of Technology Guwahati, for the award of degree of Doctor of Philosophy, has been carried out under my supervision.

The present thesis or any part thereof has not been submitted elsewhere for award of any other degree, diploma or associateship.

Prof. Bosanta Ranjan Boruah
January, 2026

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisor, Prof. Bosanta Ranjan Boruah, for introducing me to the field of optical microscopy and for his constant guidance, encouragement, and support throughout the course of this work. I am also grateful to the members of my PhD Doctoral Committee Prof. Ashwini Kumar Sharma, Prof. Dilip Pal, Dr. Rashmi Dutta Baruah, and Dr. Rishi Maiti for their valuable feedback and constructive suggestions during the course of my PhD.

I would like to thank the present and former Heads of the Department, as well as all the faculty members of the Department of Physics, IIT Guwahati, for their valuable insights, encouragement, and support during academic interactions and teaching assistantship duties. I also gratefully acknowledge the non-teaching staff of the department for their assistance with various administrative and laboratory related activities, and the Central Library for providing access to academic resources essential for my research. I am thankful to the medical staff of the institute for their care and support during my time here.

I extend my appreciation to my current lab mates Nedup, Anupam, Shilpa, Nimish, Dr. Amalesh, Neeraj, Pawan, Bishnu, and Dr. Deepjyoti, as well as my seniors Dr. Biswajit, Dr. Santanu, Dr. Ranjan, Dr. Goutam, Dr. Nagendra, Dr. Karuna, and Dr. Akanshu, for their help, discussions, and camaraderie. I also thank my old friends, including Snehashis, Prachurjya, Jyotirup, Priyanka, Indrajit, Raktim, Malaya, Gariyoshi, Pankaj, Bhaswata and Hiya. I am equally grateful to the batch mates and friends I made during my time here, including Nayan, Lwithwsa, Ankan, Samten, Sanu, Dipendu, Bhagwat, Himanshu, Dinesh, Shalini, Manish, Anterdipan and others whose presence made this journey more manageable and helped me through the demanding phases of this work. I also thank the members of the SPIE IIT Guwahati Student Chapter for giving me the opportunity to take part in and organise various educational and outreach events.

I am grateful to Prof. Paul French and the members of the Biophotonics Group at the Department of Physics, Imperial College London, including Prof. Mark, Prof. Chris, Dr. Edwin, Dr. Sunil, Dr. Jonathan, HuiHui, Miguel and Sara, for hosting me as a visiting student and providing me with the opportunity to learn and grow in a stimulating research environment.

I would like to acknowledge the Council of Scientific and Industrial Research (CSIR) for funding my PhD, as well as Imperial College London, UK for providing me the Global Development Hub Fellowship for my visit to Imperial. I would also thank ANRF, Optica, and IIT Guwahati for providing travel grants that enabled me to present my work at various international conferences.

Finally, I am deeply grateful to my parents and my sister for their unwavering love, patience, and encouragement throughout this journey.

Abstract

Super-resolution microscopy techniques overcome the diffraction limit of conventional optical microscopes and allow researchers, biologists, and pathologists to visualise nanometre-scale structures using visible light. Single Molecule Localisation Microscopy (SMLM), which includes techniques such as Photoactivated Localisation Microscopy (PALM), Stochastic Optical Reconstruction Microscopy (STORM), and Point Accumulation for Imaging in Nanoscale Topography (PAINT), enables researchers to adapt traditional fluorescence microscopes for super-resolution imaging with only minor additions to the imaging hardware. SMLM achieves this by temporally isolating neighbouring fluorophores through their stochastic fluorescence emission, imaging them in separate frames, and computationally localising their positions to map all labelled molecules with nanometre-level accuracy, ultimately reconstructing the super-resolved image.

Although highly effective, SMLM has several limitations. Capturing all fluorophore emitters often requires thousands of image frames, which can extend acquisition times to hours. The resulting datasets demand substantial storage capacity, and processing them requires efficient and fast computational tools. Emissions from sparse emitters are typically weak, necessitating the use of highly sensitive cameras for detection. However, this detection process inevitably introduces noise, making it difficult to distinguish true signals from background fluctuations. Accurate localisation also depends on fluorescent signals being well separated; signal overlap, particularly in densely labelled regions, reduces localisation accuracy and introduces artifacts into the reconstructed image. Prolonged imaging sessions further increase the risk of sample shifts caused by mechanical or thermal drift, which can result in significant artifacts in the final reconstruction.

This thesis addresses these challenges with the aim to make SMLM more accessible, particularly in low-resource settings. It presents a lightweight, modular, Python-based framework for emitter localisation and SMLM image reconstruction, implemented as a user-friendly graphical user interface (GUI). It also introduces a localisation routine that operates concurrently with image acquisition, enabling real-time image reconstruction. To improve signal detection in low signal-to-noise ratio (SNR) single-molecule fluorescence images, it proposes an adaptive thresholding method that effectively separates fluorescent

signals from noise and uneven backgrounds. To address high emitter densities, it develops an efficient convolutional neural network (CNN) for accurately localising overlapping signals. Finally, it addresses the problem of sample drift by implementing a cross-correlation-based drift correction scheme directly during the image acquisition stage.

Table of contents

List of figures

List of tables

Symbols, Acronyms and Abbreviations

Greek Symbols

π	$\simeq 3.14\ldots$
$\boldsymbol{i}$	Unit imaginary number ($\sqrt{-1}$)
λ	Wavelength
ν	Frequency
k	Propagation constant
μ	Refractive index

Mathematical Symbols

J_1	Bessel function of the first kind
$\otimes$	Convolution operation
μm	Micrometer
nm	Nanometer

Acronyms and Abbreviations

ANN	Artificial Neural Network

ASH	Averaged Shifted Histograms
BS	Beam Splitter
CAPL	CNN Assisted PSF Localisation
CCD	Charge-Coupled Device
CDF	Cumulative Distribution Function
CMOS	Complementary Metal-Oxide Semiconductor
CNN	Convolutional Neural Network
COM	Centre of Mass
CPU	Central Processing Unit
CRLB	Cramér–Rao Lower Bound
DH-PSF	Double Helix Point Spread Function
DoG	Difference of Gaussians
dSTORM	Direct Stochastic Optical Reconstruction Microscopy
FFT	Fast Fourier Transform
FPS	Frames Per Second
FRC	Fourier Ring Correlation
FWHM	Full Width at Half Maximum
GT	Ground Truth
HILO	Highly Inclined and Laminated Optical Sheet
HWP	Half Waveplate
ID	Iris Diaphragm
IR	Infrared
JI	Jaccard Index
LED	Light-Emitting Diode
LoG	Laplacian of Gaussian
LS	Least Squares
MLE	Maximum Likelihood Estimation

MSE	Mean Square Error
NA	Numerical Aperture
NDF	Neutral Density Filter
PALM	Photoactivated Localisation Microscopy
PAINT	Point Accumulation Imaging in Nanoscale Topography
PC	Personal Computer
PH	Pinhole
PLM	Peak Local Maximum
PSF	Point Spread Function
PSNR	Peak Signal-to-Noise Ratio
RAM	Random Access Memory
RMSE	Root Mean Square Error
SIM	Structured Illumination Microscopy
SLM	Spatial Light Modulator
SMLM	Single-Molecule Localisation Microscopy
SNR	Signal-to-Noise Ratio
STED	Stimulated Emission Depletion
STORM	Stochastic Optical Reconstruction Microscopy
TIRF	Total Internal Reflection Fluorescence
TP, FP, FN	True Positive, False Positive, False Negative

CHAPTER 1
Basics of Single-Molecule Localisation Microscopy and Thesis Overview

1.1 Introduction

To visualise and understand life processes has been a longstanding aim for researchers studying natural sciences. Effort to realise this goal has been advanced by the development of new microscopy methods, which have made it possible to see biological structures with unprecedented detail [1, 2]. Among these methods, light microscopy or optical microscopy holds a special place due to its unique advantages. It can image biological samples in a non-destructive manner, and can therefore be used to observe live cells and capture dynamic biological processes in real time. While light microscopy has played an important role in the establishment of cell biology, its use has been restricted by the fundamental limit of diffraction of light. For over a century, the diffraction limit of light has restricted the resolution achievable by optical microscopy to approximately 200-300 nanometers laterally and 500-700 nanometers axially [3]. This restriction implies that many intracellular structures and processes requiring sub diffraction limit resolution are beyond the reach of the optical microscopes.

The development of super-resolution microscopy techniques has broken this restriction and overcome the diffraction limit [4]. Among the various super resolution microscopy techniques, two of the most impactful ones are the Stimulated Emission Depletion (STED) microscopy [5, 6] and Single-Molecule Localisation Microscopy (SMLM) [7–9]. SMLM and STED can achieve resolution as small as tens of nanometres. STED confines fluorescence emission to a sub-diffraction volume by using a doughnut-shaped depletion laser beam. SMLM, on the other hand, utilises the stochastic photo-switching of some fluorescent dyes to separate emissions temporally, which enables the localisation of each individual fluorophore with accuracy surpassing the diffraction limit. These methods have provided the means for technology to gain a window into the molecular world as well as to visualise and quantify molecular processes in a highly detailed manner.

While both techniques are powerful, SMLM has the advantage of being relatively simpler to implement. It enables researchers to upgrade a standard fluorescence microscope to a super-resolution microscope. This ease has resulted in SMLM being widely adopted in research facilities to obtain data about the organisation of proteins, the dynamics of membranes, the processes of gene expression, and the details of other cellular mechanisms that were previously only accessible indirectly. Studies focused on the clustering of proteins and chromatin, the function of synapses, and the pathways of cellular signalling have significantly benefited from this technique [10–12].

This chapter provides a comprehensive introduction to the principles and methodologies of SMLM. It begins by establishing the foundations of light and fluorescence microscopy and then details the fundamental challenge posed by the diffraction limit. Following this, the chapter explores the super-resolution techniques that circumvent this resolution barrier and introduces SMLM in detail. We then present the theoretical framework and essential requirements for SMLM, with an emphasis on its image reconstruction process. The chapter concludes by outlining the key challenges that motivate this research and an overview of the thesis chapters.

1.2 Imaging and Light Microscopy

Light microscopy operates on the principle of using optical components to manipulate light that has interacted with a sample in order to produce a magnified image of its microstructures. It consist of an illumination system and an imaging system. The illumination system provides controlled light to the sample, while the imaging system magnifies and captures the light after it has interacted with the specimen.

The illumination system typically includes a lens assembly that collects and directs light from the source onto the specimen. Most modern microscopes use advanced light sources such as light-emitting diodes (LEDs) and lasers, which offer superior control over its intensity, wavelength, and temporal characteristics. This illumination system plays a crucial role in determining the quality and uniformity of illumination reaching the sample. The imaging system, on the other hand, consists of several key components, with the objective lens serving as the most critical element. The objective lens is responsible for collecting light from the specimen and forming the primary magnified image. Modern objectives are complex optical assemblies incorporating multiple lens elements, designed to minimise aberrations and maximise numerical aperture. The numerical aperture (NA), defined as $\mu\, sin(\alpha)$, where μ is the refractive index of the medium and α is the half-angle of the light cone accepted by the objective, is a fundamental parameter that determines both

the resolution and light-gathering power of the microscope. In epifluorescence microscopes, both excitation and emission light are routed through the same objective lens.

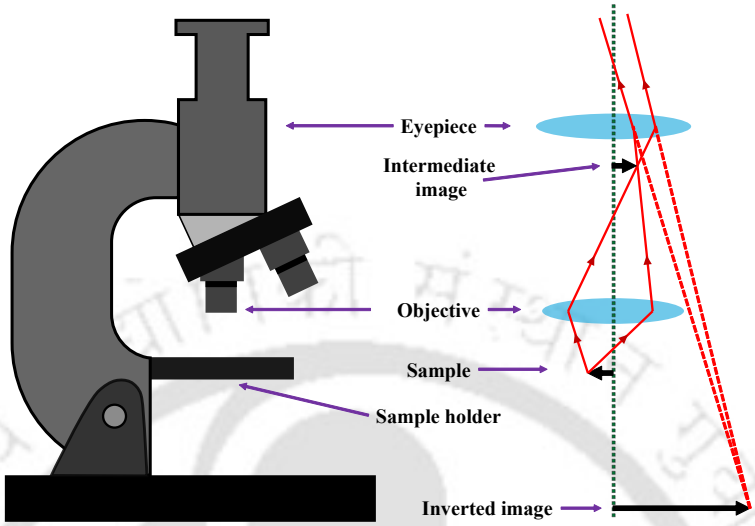

Figure 1.1. A compound microscope

Figure 1.1 shows a compound microscope with its various components illustrated on the left and its ray diagram on the right. The eyepiece provides additional magnification and relays the image formed by the objective to the observer. The total magnification of a microscope is the product of the objective and eyepiece magnifications, typically ranging from 10× to 1000× for conventional light microscopes. However, magnification alone does not determine the utility of a microscope; resolution, defined as the ability to distinguish two closely spaced objects as separate entities, is a more critical parameter. We will discuss resolution in greater depth in a later section.

Modern light microscopy has been significantly enhanced by technological advances in camera systems, particularly the development of complementary metal-oxide-semiconductor (CMOS) and charge-coupled device (CCD) cameras. These digital imaging systems offer superior sensitivity, dynamic range, quantitative capabilities, and faster acquisition speeds compared to traditional photographic methods. The integration of digital cameras with computer-controlled microscope systems has enabled automated image acquisition, real-time image processing, and sophisticated quantitative analysis.

1.3 Fluorescence Microscopy

Fluorescence microscopy is one of the most important light microscopy techniques in biological imaging, as it enables the visualisation of specific structures within the cell. Unlike bright-field microscopy, which relies on contrast generated by scattering or reflection of light from the sample, fluorescence microscopy exploits the phenomenon of fluorescence

to detect light emitted by specific molecules, known as fluorophores, that label structures of interest. The resulting image directly represents the spatial distribution of the labelled molecules.

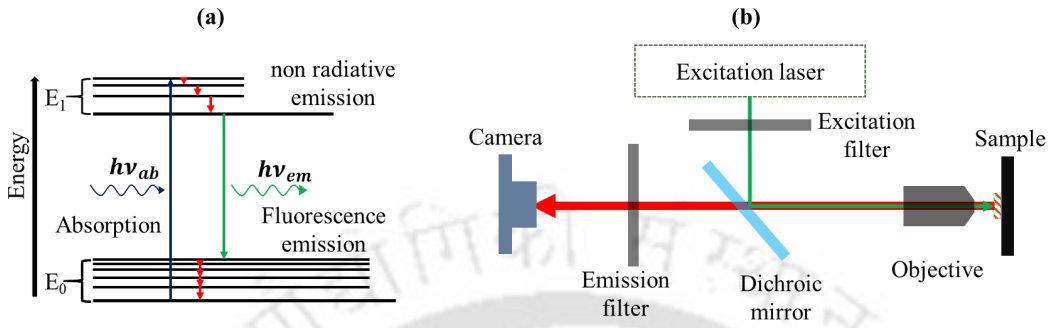

Figure 1.2. (a) A simplified Jablonski diagram depicting fluorescence emission. (b) A typical epifluorescence microscope with laser illumination.

The phenomenon of fluorescence arises from the quantum mechanical properties of fluorophore molecules. Upon absorption of a photon of sufficient energy, typically in the ultraviolet or visible range, an electron transitions from the ground state to a higher excited state. This excited state is short-lived, and within a few nanoseconds the electron relaxes back to the ground state, releasing excess energy in the form of a photon. Due to non-radiative energy losses such as vibrational relaxation in the excited state, the emitted photon has lower energy and hence a longer wavelength than the absorbed photon. This wavelength difference, known as the Stokes shift, enables the selective detection of fluorescence while effectively rejecting the excitation light. A schematic energy-level representation of this process using a simplified Jablonski diagram is shown in Figure 1.2 (a).

The overall architecture of a fluorescence microscope is similar to that of a standard light microscope but incorporates specialised optical components. These include a dedicated excitation light source, excitation filters, dichroic mirrors, emission filters, and a highly sensitive camera. As shown in Figure 1.2 (b), the excitation source, often comprising one or more high-power lasers, provides intense light at specific wavelengths to excite the fluorophores. An excitation filter ensures that only the desired spectral band reaches the sample. The dichroic mirror reflects this excitation light through the objective lens onto the specimen, while the same objective collects the fluorescence emission. Because the emitted light is red-shifted, it passes through the dichroic mirror and is further isolated by the emission filter before reaching the camera, resulting in a high-contrast image where fluorescent structures appear bright against a dark background.

The strength of fluorescence microscopy lies in its specificity, which is achieved by selectively labelling cellular components with fluorophores. While some biological molecules exhibit natural fluorescence, known as autofluorescence [13], the widespread applicability

of fluorescence microscopy emerged with the development of targeted fluorescent probes. These include small-molecule dyes such as DAPI for DNA [14], fluorescently labelled antibodies used in immunofluorescence, and genetically encoded fluorescent proteins such as GFP [15]. These probes can be directed to specific biomolecules with high precision, enabling high-contrast imaging and, under optimal conditions, the detection of single molecules [16]. Furthermore, the use of multiple fluorophores with distinct spectral properties allows multicolour imaging, facilitating the simultaneous visualisation of different cellular components [17].

Despite its capabilities, conventional fluorescence microscopy remains fundamentally limited by the diffraction of light. In the following sections, we examine how advanced super-resolution techniques have been developed to overcome this limitation and achieve nanoscale imaging resolution.

1.4 Diffraction Limit of Resolution

The fundamental limitation of optical microscopy arises from the wave nature of light. As light from a point source passes through the finite aperture of a lens, it undergoes diffraction, causing the image of an ideal point to spread into a three-dimensional intensity distribution of finite extent, known as the Point Spread Function (PSF) [18]. This diffraction-induced spreading gives rise to an intrinsic limit on spatial resolution that cannot be overcome by conventional optical design or increased magnification alone. This limit therefore represents a fundamental constraint on the ability of an optical system to distinguish closely spaced objects.

For circular apertures, such as those used in microscope objectives, the PSF takes the form of an Airy pattern in the focal plane, characterised by a bright central disk surrounded by concentric rings of decreasing intensity [19]. The finite spatial extent of the Airy disk defines the minimum separation at which two point sources can be resolved. The above diffraction limit can be expressed through Abbe's resolution criterion [20], which relates the achievable resolution to the wavelength of light and the numerical aperture of the objective. If λ is the imaging wavelength and NA is the numerical aperture, the lateral resolution limit as per the Abbe's criterion is given by,

$$d_{Abbe} = \frac{\lambda}{2\,NA} \tag{1.1}$$

In addition to Abbe's formulation, two other widely used criteria for defining optical resolution are the Rayleigh and Sparrow criteria. Rayleigh's criterion states that two point sources are just resolvable when the principal diffraction maximum of one Airy disk coincides with the first minimum of the other [21], yielding a resolution of $d_{\text{Rayleigh}} = \frac{0.61\,\lambda}{NA}$.

Sparrow's criterion defines an even stricter limit, corresponding to the smallest separation at which the combined intensity profile no longer exhibits a discernible minimum between two peaks [22], given by $d_{\text{Sparrow}} = \frac{0.47\,\lambda}{NA}$. To illustrate and compare these resolution

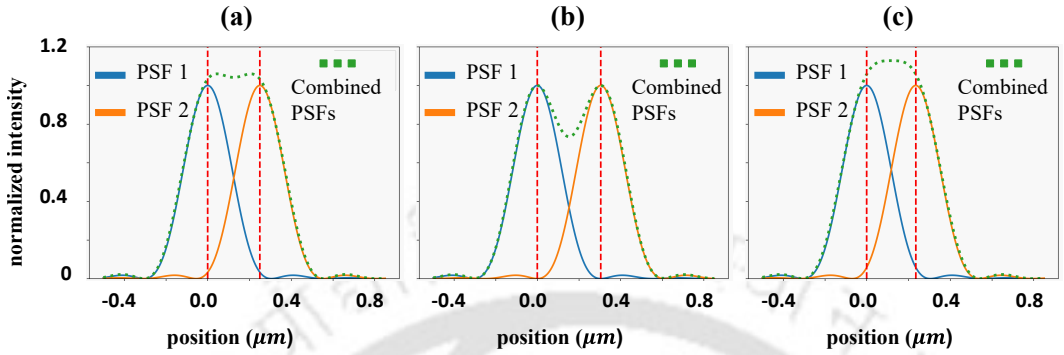

Figure 1.3. Simulated intensity profiles for two point emitters placed at lateral separations corresponding to (a) Abbe's, (b) Rayleigh's, and (c) Sparrow's resolution limits. Simulations assume $\lambda = 500$ nm and numerical aperture $NA = 1$. Plots show both individual PSFs and the resulting summed intensity distribution.

criteria, we simulate the image formation of two point emitters emitting at a wavelength of 500 nm and separated by distances corresponding to the Abbe, Rayleigh, and Sparrow limits. The simulations assume a diffraction-limited objective with unit numerical aperture ($NA = 1$). The resulting intensity distributions are shown in Figures 1.3(a–c), where each subplot displays the individual PSFs and their combined intensity profiles for separations corresponding to (a) Abbe's, (b) Rayleigh's, and (c) Sparrow's resolution criteria.

1.5 Super-Resolution Microscopy

As discussed, conventional optical microscopy is constrained by the diffraction limit, which restricts lateral resolution to approximately 200–300 nm and axial resolution to 500–700 nm. This resolution is insufficient to resolve most subcellular and molecular-scale structures. Super-resolution microscopy refers to a class of advanced imaging techniques that overcome this fundamental limitation. These techniques include: (i) frequency-space engineering in Structured Illumination Microscopy (SIM), (ii) spatial control of fluorescence emission in STED microscopy, (iii) temporal separation and localisation of individual emitters in SMLM, and (iv) spatially modulated excitation and photon-efficient triangulation in MINFLUX microscopy. Although based on distinct physical principles and optical implementations, all these approaches enable imaging beyond the diffraction limit, often reaching nanometre-scale resolution. Below, we outline the key super-resolution techniques.

1.5.1 Structured Illumination Microscopy (SIM)

Structured Illumination Microscopy is a super-resolution technique that can approximately double the spatial resolution of a conventional widefield microscope by illuminating the sample with a sequence of precisely engineered light patterns [23–26]. The core principle involves illuminating the specimen with a high-frequency sinusoidal intensity pattern, typically generated using a diffraction grating or a spatial light modulator (SLM), rather than uniform illumination. This patterned excitation interacts with sub-diffraction spatial features in the sample, producing lower-frequency Moiré fringes that encode otherwise unresolvable high-frequency information into the observable passband of the microscope.

To recover the information, a series of raw images is acquired, typically nine images for 2D SIM (three pattern orientations with three phase shifts each). A computational reconstruction algorithm then separates and reassigns the spatial frequency components to their correct locations, yielding a super-resolved image. Under linear illumination conditions, the theoretical lateral resolution achievable with SIM is given by,

$$d_{\mathrm{SIM}} \approx \frac{\lambda}{4NA} \tag{1.2}$$

1.5.2 Stimulated Emission Depletion (STED) Microscopy

Stimulated Emission Depletion microscopy was the first super-resolution technique to demonstrate that the diffraction barrier could be surpassed by actively manipulating the fluorescence emission process [5, 6]. The key idea is to engineer an effective PSF that is smaller than the diffraction limit. This is achieved using two spatially overlapped laser beams: a diffraction-limited excitation beam and a high-intensity depletion beam shaped into a doughnut profile with a central intensity minimum. The depletion beam, tuned to a longer wavelength, arrives shortly after excitation and forces fluorophores at the periphery of the excitation spot back to the ground state via stimulated emission, thereby suppressing their fluorescence. As a result, fluorescence is confined to a sub-diffraction region near the centre of the doughnut.

By scanning this reduced effective PSF across the sample, a super-resolved image is constructed. The achievable resolution depends on the intensity of the depletion beam and is given by,

$$d_{\mathrm{STED}} = \frac{\lambda}{2\,NA\,\sqrt{1 + I_{\mathrm{STED}}/I_{\mathrm{sat}}}} \tag{1.3}$$

where I_{STED} is the depletion beam intensity and I_{sat} is the fluorophore saturation intensity. Increasing I_{STED} progressively reduces the effective PSF size, enabling typical lateral resolutions of 30–50 nm and axial resolutions of approximately 100 nm [4].

1.5.3 MINFLUX Microscopy

MINFLUX (minimal photon fluxes) microscopy [27, 28] is a next-generation super-resolution technique that combines coordinate-targeted excitation with the photon efficiency of single-molecule localisation methods. It employs a doughnut-shaped excitation beam with a central intensity minimum, similar to STED, which is scanned in a defined pattern around an emitting fluorophore. By analysing the detected fluorescence intensity as the beam minimum approaches the emitter position, MINFLUX determines the fluorophore location using a triangulation-based approach with extremely high precision. The theoretical localisation precision is given by

$$\sigma_{\text{MINFLUX}} = \frac{L}{\gamma\sqrt{N}} \tag{1.4}$$

where L is the scan radius, N is the number of detected photons, and γ is a geometric factor. Experimental implementations have demonstrated localisation precisions of 1–3 nm in both two and three dimensions, with temporal resolutions on the order of tens of microseconds [28].

1.5.4 Single-Molecule Localisation Microscopy (SMLM)

Single-Molecule Localisation Microscopy techniques, including Photoactivated Localisation Microscopy (PALM) [7, 29], Stochastic Optical Reconstruction Microscopy (STORM) [8, 30], and Point Accumulation for Imaging in Nanoscale Topography (PAINT) [31, 32], achieve nanometre-scale resolution by temporally isolating the emission of individual fluorophores. This is accomplished through stochastic activation or photoswitching, ensuring that only a sparse subset of emitters is active in each image frame. Each active fluorophore produces a diffraction-limited PSF, which is identified and fitted, most commonly using a two-dimensional Gaussian model to determine its centroid with sub-pixel accuracy. The final super-resolved image is reconstructed by accumulating the localised centroids across all frames and rendering them as point-based histograms or Gaussian-convolved density maps.

The localisation precision for a 2D Gaussian-approximated PSF is given by [33–35],

$$\sigma_{\text{SMLM}} = \sqrt{\left(\frac{\sigma^2 + a^2/12}{N}\right)\left(\frac{16}{9} + \frac{8\pi(\sigma^2 + a^2/12)b^2}{a^2 N}\right)} \tag{1.5}$$

where σ is the PSF standard deviation, a is the pixel size, N is the number of detected photons, and b is the background noise per pixel. For practical interpretation, this expression is often approximated as $\sigma_{\text{SMLM}} \approx \sigma/\sqrt{N}$.

SMLM systems typically employ high-power activation and excitation lasers, high-NA oil-immersion objective lenses, highly sensitive EMCCD or sCMOS cameras, and imaging

buffers optimised for photoswitching chemistry [9]. From a practical perspective, SMLM is particularly attractive due to its relative simplicity and accessibility, as many conventional fluorescence microscopes can be upgraded to support SMLM with minimal hardware modifications. Beyond image reconstruction, SMLM yields precise molecular coordinate maps that enable advanced quantitative analyses such as molecule counting, clustering analysis, and single-particle tracking. Using these capabilities, SMLM has revealed nanoscale organisation in systems such as nuclear pore complexes, cytoskeletal networks, and membrane receptor assemblies, with typical lateral resolutions of 20–30 nm and axial resolutions of approximately 50 nm [8, 7, 36].

1.6 SMLM Modalities and Photoswitching Strategies

As discussed earlier, the principle of SMLM is to achieve sub-diffraction resolution by temporally isolating individual fluorescence emissions. This is realised through several distinct techniques, all of which share the common goal of capturing emission from a sparse subset of fluorophores in each imaging frame. To do so, they employ different photophysical or chemical mechanisms. One of the first methods developed is PALM, which typically employs genetically encoded photoactivatable fluorescent proteins [7]. A conceptually similar technique, STORM, was introduced originally utilising pairs of synthetic fluorescent dyes, an "activator" and a "reporter", to achieve reversible photoswitching [8]. A different approach is taken by PAINT, which relies on the transient binding and unbinding of fluorescent probes from a solution to the target structure, where the emission is governed by binding kinetics rather than intramolecular photoswitching [31]. The development of DNA-PAINT, using transient hybridisation of short, dye-labelled DNA strands, has made this a particularly powerful and versatile approach [32]. Figure 1.4 illustrates the stochastic fluorescence emission from a small subset of emitters in an image frame, their localisation, and the reconstruction of the final super-resolved image.

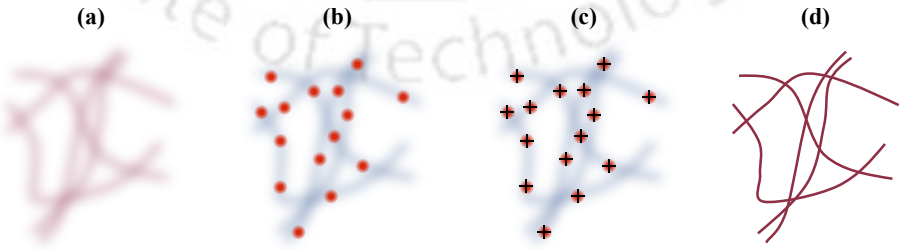

Figure 1.4. (a) A conventional widefield fluorescence image where many fluorophores emit simultaneously, resulting in overlapping PSFs and a blurred image. (b) A sparse, random subset of active fluorophores in an SMLM image frame, (c) localisation of fluorophores, and (d) the final super-resolved image reconstructed by accumulating localisations over many frames.

Although the specifics of fluorophore activation and deactivation differ, all these methods culminate in the same data acquisition process; which is acquisition of a long sequence of images, each containing a sparse set of diffraction-limited PSF spots. The following sections will focus in greater detail on the principles and practical implementation of one widely adopted SMLM method, STORM and its variant, dSTORM.

1.6.1 STORM and dSTORM

The key to STORM lies in the use of photoswitchable fluorophores that can be cycled between a fluorescent "on" state and a dark "off" state in a stochastic manner, a phenomenon commonly referred to as “blinking”. This blinking behaviour allows only a sparse, random subset of fluorophores to be active at any given time, ensuring that their individual emission patterns do not spatially overlap on the detector. These isolated emission events are captured over successive image frames and fitted with a PSF model to determine the precise sub-diffraction localisation of each molecule. After a few cycles, the fluorophores return to their dark state, allowing a new subset to be activated. By iterating this process over thousands of frames, a composite image with nanometre-scale resolution is reconstructed from the accumulated localisations.

As already mentioned, classical implementations of STORM use pairs of fluorescent dyes that include an activator and a reporter, typically conjugated to antibodies. The reporter dye (e.g., Cy5) is first switched to a stable dark state by the imaging laser. The activator dye (e.g., Cy3) is then excited with a separate, lower-wavelength laser, which helps to return the nearby reporter dye to its fluorescent state. The reporter dye then emits fluorescence under its specific excitation wavelength until it transitions back to a dark state. This cycle is repeated, allowing a new subset of fluorophores to be activated. While this approach provides high localisation precision, it requires complex dye chemistry and specific photophysical properties, which can limit accessibility and flexibility for labelling diverse biological samples.

To overcome these limitations, a simplified implementation known as direct STORM (dSTORM) was developed [30]. dSTORM eliminates the need for specialised activator-reporter dye pairs by exploiting the photoswitching behaviour of single conventional organic fluorophores like Alexa Fluor 647 or Cy5. As illustrated in Figure 1.5, the mechanism relies on driving fluorophores into a long-lived dark state. Under high-intensity laser excitation, a fluorophore is promoted from its ground state (E_0) to an excited singlet state (E_1), from which it can fluoresce. However, it can also transition via inter system crossing to a transient triplet state (E_T). In a reducing chemical environment, this triplet state molecule is then efficiently shelved into a stable, non-fluorescent dark "off" state (E_{dark}). The fluorophores stochastically return from this dark state to the ground state, creating

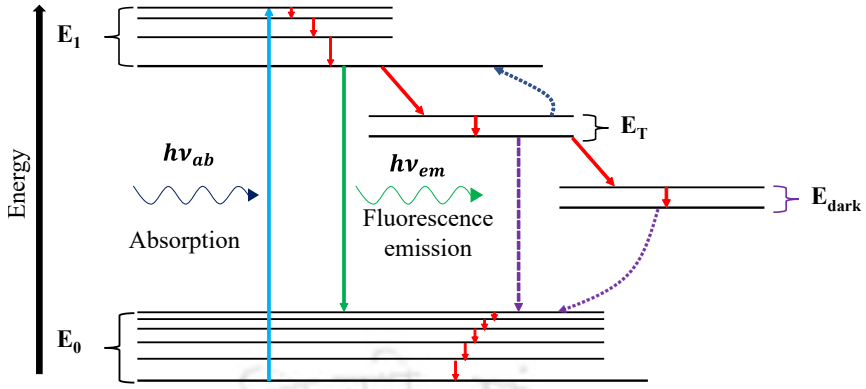

Figure 1.5. Jablonski diagram illustrating the principle of photoswitching in dSTORM. Upon excitation with a high-intensity laser (blue arrow), the fluorophore is promoted from the ground state (E_0) to the excited singlet state (E_1) and can then fluoresce back to E_0 (green arrow), defining the "on" state. Alternatively, the fluorophore can undergo inter system crossing (ISC) to the triplet state (E_T), from which it is shelved into a long-lived, non-fluorescent dark state (E_{dark}), often via reaction with a thiol in the imaging buffer. This "off" state persists until the fluorophore is reactivated (purple arrow), returning it to the ground state to re-enter the cycle.

the sparse and reversible blinking behaviour suitable for single-molecule localisation. This approach allows the use of standard immunofluorescence labelling techniques, significantly broadening the applicability of STORM imaging. Owing to its flexibility, simpler chemistry, and compatibility with standard labelling methods, dSTORM is taken as the basic system for the experimental work discussed in this thesis.

1.6.1.1 Experimental Setup Used in the Thesis

For our work a typical dSTORM microscope is built upon an inverted widefield fluorescence microscope (referred to as openFrame modular microscope), specifically modified for the demands of single-molecule imaging. The above openFrame microscope [37, 38] was built in a collaborative research between Department of Physics, IIT Guwahati and the Photonics group of Imperial College London. As shown in Figure 1.6, in the microscope illumination is provided by one or more high-power CW lasers, with common wavelengths being 462 nm, 532 nm, and 638 nm for exciting different fluorophores. An additional laser, at 405 nm, is used to reactivate fluorophores from the dark state in a controlled manner, allowing for consistent emitter density throughout the acquisition.

The excitation beam is focused at the back focal plane using lens L_1 and then passed through a high numerical aperture, oil-immersion objective lens to achieve uniform illumination at the sample plane. For improved optical sectioning and signal-to-background ratio, the system often operates in Total Internal Reflection Fluorescence (TIRF) or Highly Inclined and Laminated Optical (HILO) illumination mode, restricting illumination to a

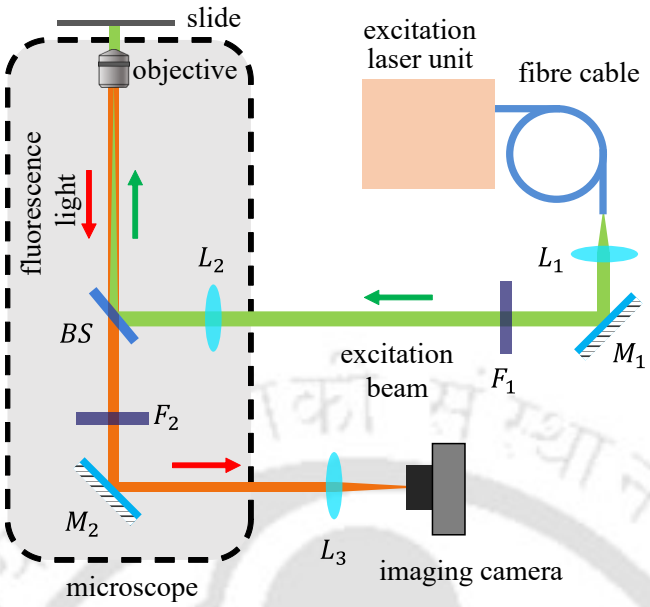

Figure 1.6. Schematic diagram of a typical dSTORM optical setup.

thin section of the sample near the coverslip. The emitted fluorescence light is gathered by the objective, passed through a dichroic mirror (BS) and an emission filter (F_2) to remove scattered excitation light, and is finally focused onto a highly sensitive camera by lens L_2.

To ensure positional stability during the extended acquisition period, the microscope is mounted on a vibration-isolated platform, and thermal drift is minimised using an environmental enclosure. The sample is mounted on a motorised or piezoelectric stage capable of nanometre-scale movement along the x-, y-, and z-axes, allowing for precise positioning and drift compensation. In many systems, fiducial markers such as fluorescent beads are co-immobilised with the sample and serve as stable references for post-acquisition drift correction.

1.6.1.2 Sample Preparation and Imaging

Sample preparation for dSTORM follows standard protocols for fixed-cell fluorescence microscopy but requires specific attention to dye selection and buffer conditions. Cells or tissue sections are chemically fixed, commonly using paraformaldehyde, to preserve ultrastructural detail. Following fixation, permeabilisation is performed using detergents like Triton X-100 to allow antibody access to intracellular epitopes. Fluorescent labelling is typically achieved through immunostaining, using primary antibodies targeting the protein of interest and secondary antibodies conjugated to photoswitchable dyes like Alexa Fluor 647 or Cy5 [3].

To achieve optimal blinking behaviour, the sample is mounted in a photoswitching buffer containing a reducing agent, such as β-mercaptoethylamine (MEA), and an oxygen-

scavenging system composed of glucose, glucose oxidase, and catalase [30]. This chemical environment stabilises the dark state of the fluorophores and enables repeated cycling between on and off states.

Once prepared, the sample is placed on the microscope stage and imaged over thousands of frames, with careful regulation of laser powers to maintain a sparse and stable blinking regime. Although image reconstruction and localisation analysis form the final stage of STORM and other SMLM techniques. These processes are discussed in the next section.

1.7 Image Formation and PSF Localisation in SMLM

As has been discussed already, image formation in SMLM involves light emitted from individual fluorophores, producing a characteristic PSF for each emitter in the camera plane. The raw data consist of a temporal sequence of image frames, with each frame containing emission from a stochastically different subset of fluorophores. Each image frame can be represented as a two-dimensional array of pixels, where the pixel intensities correspond to the number of photons captured by the detector during the specified exposure time [39].

The detected photon signal in each frame arises from multiple sources, including fluorescence emitted by in-focus fluorophores bound to the target structure, as well as contributions from out-of-focus emitters and autofluorescent background [9]. Additional spurious signals originating from the sample volume or optical path may further contribute to the measured intensity. The recorded image can therefore be interpreted as a convolution of the true fluorophore distribution with the microscope PSF, superimposed with measurement noise. Mathematically, the image formation process can be described as,

$$I(x, y) = N_P\left(F(x, y) \otimes \mathrm{PSF} + B\right) + N_G \tag{1.6}$$

where $I(x, y)$ denotes the observed intensity at pixel (x, y), $F(x, y)$ represents the distribution of active fluorophores, and B accounts for the mean background intensity. The Poisson noise operator $N_P()$ models signal-dependent photon shot noise, while N_G represents additive, signal-independent Gaussian readout noise [40, 39, 41].

Figure 1.7 schematically illustrates this process where panel (a) shows the spatial locations of fluorescent emitters; (b) depicts the convolution of these locations with the PSF; and (c) shows a representative SMLM frame recorded by the camera. Below we describe how individual PSFs are localised to recover emitter positions.

1.7.1 2D localisation

In an ideal aberration-free imaging system operating under far-field conditions, the PSF produced by a single point emitter is described by the Airy diffraction pattern. The intensity

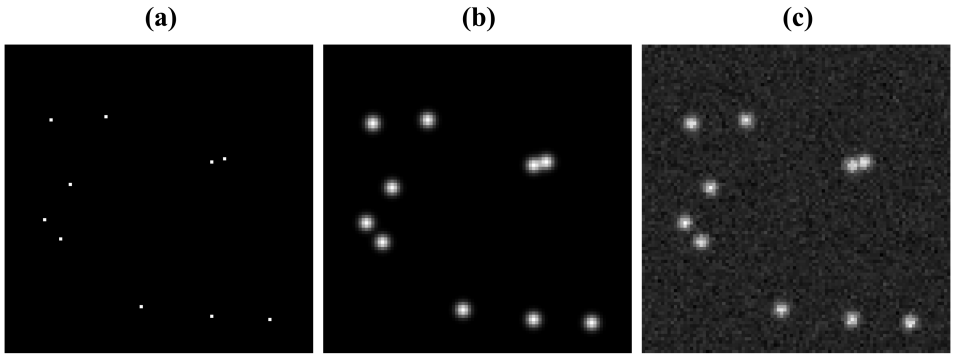

Figure 1.7. Image formation in SMLM. (a) Fluorophore locations. (b) Convolution of fluorophore locations with the PSF. (c) The resulting SMLM image recorded by the detector.

profile of the Airy PSF in the image plane is given by,

$$PSF_{\text{Airy}}(r) = I_0 \left[\frac{2J_1\left(\frac{ka}{R}r\right)}{\frac{ka}{R}r} \right]^2 \tag{1.7}$$

where I_0 is the peak intensity, J_1 is the first-order Bessel function, $k = 2\pi/\lambda$ is the wave number, a is the aperture radius, R is the effective focal length, and r is the radial coordinate in the image plane. This expression captures the oscillatory nature of diffraction arising from coherent wave interference.

While the Airy PSF provides an accurate physical description of the PSF, its analytical complexity and secondary ring structure pose challenges for its efficient use. As a result, the central lobe of the Airy pattern is commonly approximated by a two-dimensional Gaussian function [42]. This approximation enables efficient fitting using non-linear least squares or maximum likelihood estimation, while still achieving sub-diffraction localisation precision [43, 44].

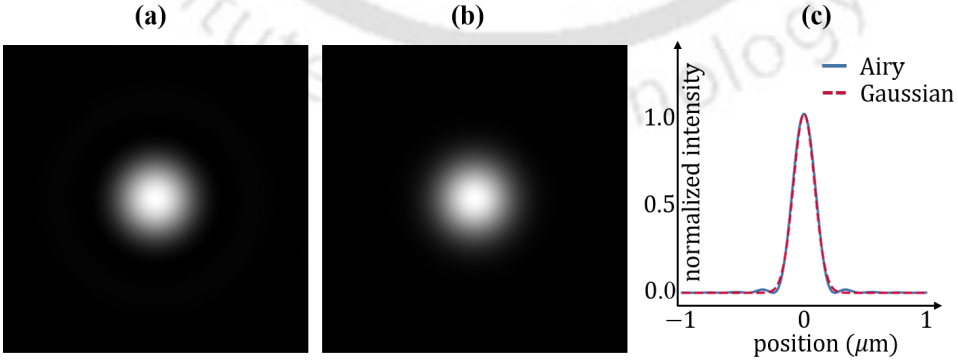

Figure 1.8. Two-dimensional PSF models: (a) Simulated Airy diffraction pattern for a 1.4 NA objective at 580 nm emission, (b) corresponding Gaussian approximation, and (c) radial intensity profiles through the centre of each PSF.

Figure 1.8 shows a comparison between the Airy PSF and its Gaussian approximation, demonstrating close agreement in the central region that dominates localisation accuracy. The analytical form of the Gaussian PSF is expressed as,

$$PSF_{\text{Gaussian}}(x, y) = I_0 \exp\left(-\frac{(x - x_0)^2}{2\sigma_x^2} - \frac{(y - y_0)^2}{2\sigma_y^2}\right) + b \tag{1.8}$$

where (x_0, y_0) denote the emitter coordinates, σ_x and σ_y are the PSF widths, I_0 is the peak amplitude, and b represents the background level. The fitting procedures used for localisation are discussed in detail in Chapter 2.

1.7.2 3D localisation

Although this thesis focuses on 2D localisation, several established methods enable axial (z) localisation in three-dimensional SMLM by encoding depth information into measurable changes in the PSF. These approaches rely on controlled optical modifications that introduce a calibrated z-dependence in the PSF shape or structure. The principal techniques include astigmatism, double-helix PSF (DH-PSF), and tetrapod PSF engineering.

Astigmatic PSF imaging is the most widely adopted approach due to its simplicity and robustness. It involves inserting a weak cylindrical lens into the detection path, typically between the tube lens and the camera [36]. This induces an asymmetric elongation of the PSF along mutually orthogonal axes above and below the focal plane. By fitting the resulting elliptical PSF with a Gaussian model, the axial position of the emitter can be inferred. Astigmatic localisation typically provides axial precisions of 50–70 nm over a range of approximately ± 0.5 μm [36].

The **double-helix PSF** employs a phase mask placed at the Fourier plane to reshape the PSF into two lobes that rotate about the optical axis as a function of axial position [45, 46]. The relative orientation of the lobes provides a monotonic mapping to the z-coordinate, enabling extended axial localisation ranges of up to ± 1.5 μm with near-uniform precision.

The **tetrapod PSF** is another phase-engineered design that produces a multi-lobed intensity pattern whose geometry varies systematically with axial position [47, 48]. Although offering long axial ranges and high photon efficiency, its increased structural complexity necessitates model-based fitting, typically using maximum likelihood estimation.

1.7.2.1 Alternative and Emerging 3D Localisation Methods

Additional strategies for 3D localisation include biplane imaging, light field microscopy, and machine learning-based approaches.

Biplane imaging splits the emission signal towards two detection planes separated along the optical axis, allowing axial localisation through comparative analysis of PSF shapes [49].

Light field microscopy captures both spatial and angular information using a microlens array, enabling volumetric localisation through computational reconstruction, albeit with some loss in lateral resolution [50–52].

More recently, **deep learning-based** methods have been introduced for 3D localisation, using convolutional neural networks trained on simulated or experimental PSFs to directly infer emitter positions [53–55]. These approaches have demonstrated strong robustness to noise, aberrations, and emitter overlap, and are particularly promising for dense or low-SNR imaging conditions.

1.8 Localisation Precision and Image Quality Assessment in SMLM

The theoretical limit on localisation precision in SMLM is set by the Cramér–Rao Lower Bound (CRLB) [56], which specifies the lowest possible variance achievable by any unbiased estimator of a parameter. In the context of emitter localisation, it defines the best achievable precision for a given imaging configuration. Assuming independent Poisson statistics for photon detection in each pixel, and a PSF model $I(x, y; \theta)$ where $\theta = (x_0, y_0, z_0)$ denotes the emitter coordinates, the Fisher Information Matrix (FIM) is given as,

$$F_{ij} = \sum_{x,y \in \mathrm{ROI}} \frac{1}{I(x, y; \theta)} \frac{\partial I(x, y; \theta)}{\partial \theta_i} \frac{\partial I(x, y; \theta)}{\partial \theta_j} \tag{1.9}$$

where $I(x, y; \theta)$ is the expected photon count at pixel (x, y) for a given θ. The photon count includes contributions from the signal and the background. The sum runs over all pixels in the region of interest (ROI) used for fitting. The CRLB for the i-th coordinate is then given by the corresponding diagonal element of the inverse FIM as,

$$\sigma^2_{\theta_i} \geq [F^{-1}]_{ii} \tag{1.10}$$

The CRLB depends on several experimental factors, including the number of detected photons N, background noise level b, pixel size a, and the spatial characteristics of the PSF. Engineered PSFs such as the double-helix or tetrapod designs are often optimised to minimise the axial CRLB across an extended depth of field, though this may increase the PSF footprint and computational complexity. Comparing CRLB profiles for different PSF

designs can guide the selection of the optimal imaging strategy under given experimental constraints.

1.8.1 Quantitative Evaluation of SMLM Localisation and Reconstructions

Quantitative evaluation of SMLM reconstructed images is critical for validating the performance of both the imaging setup and the downstream analysis pipeline. Unlike conventional microscopy, where resolution is determined by optical properties alone, SMLM performance is intricately linked to the precision and accuracy of emitter localisation. This necessitates objective metrics to quantify how well the localisation algorithm detects molecules and reconstructs spatial features.

The evaluation can be broadly categorised into two domains: (i) localisation accuracy and precision at the single-molecule level, and (ii) resolution of the final super-resolved image, which is typically assessed through spatial frequency domain analysis. Below we outline the key metrics used to assess these properties, some of which require known ground truth positions (e.g., from simulations), while others can be applied directly to experimental data.

1.8.2 Localisation Performance Metrics

Localisation performance can be analysed using concepts adapted from detection theory and statistical error analysis. Given a set of ground truth molecule positions $\mathcal{T} = \{\mathbf{r}_i\}$ and a corresponding set of estimated localisations $\mathcal{L} = \{\hat{\mathbf{r}}_j\}$, we can define True Positive (TP), False Negative (FN), and False Positive (FP) localisations based on whether they fall within a matching tolerance radius r_{tol} (typically 20-50 nm) of the ground truth. A TP is a correct detection corresponding to a ground truth emitter; a FN is a ground truth emitter that the algorithm failed to detect; while a FP is an erroneous localisation that does not correspond to a real emitter, often arising from background noise or artifacts. From these classifications, we can define the other performance parameters, namely, Recall, Precision, and Jaccard Index (JI) and the Root Mean Squared Error (RMSE)[57].

1.8.2.1 Recall, Precision, and Jaccard Index

Recall, also known as sensitivity, quantifies the algorithm's detection efficiency by measuring the fraction of actual emitters that were successfully found. A high recall score indicates that the algorithm is effective at identifying most of the true PSFs in the images. Conversely, Precision measures the reliability of the detected localisations, reflecting the fraction of reported localisations that are correct. High precision is critical for ensuring the final reconstructed image is not corrupted by false localisations. There is often a direct trade-off

between these two metrics; a more lenient detection threshold may increase recall at the expense of precision. To provide a balanced assessment of performance, the Jaccard Index is used. It measures the overall similarity between the ground truth and the detected sets by relating the number of true positives to the union of all ground truth and detected points. A high Jaccard Index signifies both high recall and high precision, indicating a robust and reliable localisation result. If $n()$ is the number operator, then recall, precision and JI can be expressed as,

$$\text{Recall} = \frac{n(\text{TP})}{n(\text{TP}) + n(\text{FN})} \tag{1.11}$$

$$\text{Precision} = \frac{n(\text{TP})}{n(\text{TP}) + n(\text{FP})} \tag{1.12}$$

$$\text{JI} = \frac{n(\text{TP})}{n(\text{TP}) + n(\text{FP}) + \text{n(FN)}} \tag{1.13}$$

1.8.2.2 Root Mean Square Error (RMSE)

For each correctly matched true positive pair, the localisation error is the Euclidean distance between the ground-truth position $\mathbf{r}_i$ and estimated positions $\hat{\mathbf{r}}_i$. The Root Mean Square Error (RMSE) is then calculated over all true positive pairs to provide the average deviation from the true position [58], which is an aggregate measure of localisation accuracy. The RMSE in this case is given as,

$$\text{RMSE} = \sqrt{\frac{1}{n(\text{TP})} \sum_{i=1}^{n(\text{TP})} \|\hat{\mathbf{r}}_i - \mathbf{r}_i\|^2} \tag{1.14}$$

Separate RMSE values may be found for the lateral (x, y) and axial (z) dimensions in 3D SMLM, providing insight into anisotropy in localisation accuracy. Figure 1.9 shows 3 ground truth emitter positions in blue dots, false negatives as red '+', and true positive, and false positive localisations in green and violet '+', respectively. Distance of a ground truth position from its true positive localisation is shown with a black arrow.

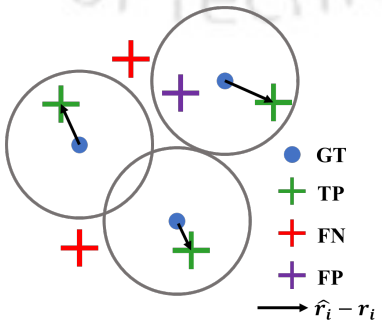

Figure 1.9. Pictorial depiction of localisation performance metrics. Here GT is ground truth, TP, FN and FP represent true positive, false negative and false positive localisation.

1.8.3 Fourier Ring Correlation (FRC)

While metrics like RMSE evaluate individual localisations, they do not directly assess the effective resolution of the final super-resolved image. Fourier Ring Correlation (FRC) [59, 60] can provide an objective, data-driven measure of resolution based on the reproducibility of spatial frequency content. FRC involves splitting the localisation dataset into two statistically independent subsets, reconstructing two super-resolved images and computing the normalised cross-correlation of their spatial frequency components. The FRC curve is defined as,

$$FRC(f) = \frac{\sum_{f\in \text{ring}} \tilde{I}_1(f) \cdot \tilde{I}_2^*(f)}{\sqrt{\sum_{f\in \text{ring}} |\tilde{I}_1(f)|^2 \cdot \sum_{f\in \text{ring}} |\tilde{I}_2(f)|^2}}, \tag{1.15}$$

where $\tilde{I}_1(f)$ and $\tilde{I}_2(f)$ are the Fourier transforms of the two images, evaluated over concentric frequency rings. The resolution is defined as the spatial frequency f_{res} at which the FRC curve falls below a threshold number, typically $1/7$ for localisation microscopy [60], and is given by,

$$f_{\text{res}} = \min\left\{ f \,:\, FRC(f) < \frac{1}{7} \right\}, \qquad \text{FIRE} = \frac{1}{f_{\text{res}}}. \tag{1.16}$$

Fourier image resolution (FIRE) number provides a quantitative measure of the image resolution. It is advantageous in the fact that it does not require knowledge of ground truth, making it well suited for real experimental data. However, it is influenced by localisation density and rendering parameters. When implemented rigorously, FRC provides an unbiased estimate of the spatial resolution achieved in a given dataset.

1.9 Issues with SMLM and Motivation of this Thesis

As discussed in previous sections, SMLM offers unprecedented resolution in biological imaging beyond the diffraction limit by localising individual fluorescent emitters. However, despite its potential, several limitations continue to hinder its widespread adoption in routine laboratory environments.

One such major barrier is the fragmented and non-adaptive nature of current SMLM workflows. In most existing setups, image acquisition, data processing, and super-resolution reconstruction are handled by separate tools, often with significant manual intervention. Given that SMLM typically produces thousands to tens of thousands of image frames, with each containing information of stochastically emitting fluorophores, this disjointed approach sometimes becomes inefficient. It prevents real-time feedback during acquisition, which hampers adaptive imaging strategies. Post-processing tools such as ThunderSTORM [61] and Picasso [32] are powerful but are generally limited to offline analysis and may not be easily adapted for real-time operation or modular customisation. Storing all raw frames,

which often are of 20–30 gigabytes in size, for later analysis also becomes impractical in high throughput or resource constrained settings, such as hospital imaging facilities or long-term live-cell experiments.

Another challenge arises from noise, particularly spatially varying backgrounds and low signal-to- noise ratios, which impair detection and localisation accuracy. Conventional thresholding methods frequently struggle on fluorescence microscopy images [62]. In SMLM, image pixel-intensity distributions are typically unimodal with a pronounced right skew and a heavy tail due to sparse bright PSFs superimposed on a non uniform background. Standard thresholding algorithms which are often designed for simpler unimodal or multimodal distributions do not handle SMLM images well. The wide dynamic range across fluorescence images can also cause incorrect segmentation of both bright and dim regions. Moreover, background variability coupled with random noise can obscure true emitter signals and make detection unreliable [63]. As a result, current global thresholding methods are ill-suited to local variations and often yield high false-positive rates or miss dim fluorophores [63].

A third bottleneck arises when imaging the sample at high emitter densities to accelerate data acquisition. Allowing multiple fluorophores to emit simultaneously reduces the number of image frames required to capture emission from all the fluorophores, thereby reducing total imaging time. But this practice often produces overlapping PSFs in the captured images. Standard single emitter fitting algorithms often fail under these conditions, leading to incorrect localisations and thus causing loss of resolution. While several multi emitter fitting and deconvolution approaches exist for such cases [64–68], they often perform poorly in low signal-to-noise ratio conditions and can be computationally expensive for large, dense datasets.

Finally, sample drift during long acquisitions poses a serious challenge. Imaging sessions lasting minutes to hours are vulnerable to mechanical instabilities and thermal fluctuations that cause relative motion between the sample and the imaging system [33, 9]. Even nanometre-scale shifts during acquisition reduces the localisation accuracy of the fluorophores, which blurs the final reconstructed image and negates the benefits of high-precision localisation [33].

1.10 Thesis overview

This thesis directly confronts the challenges mentioned above with the overarching goal of making SMLM a more robust, efficient, and accessible imaging modality for high-throughput applications. It proposes an integrated framework of computational and hardware solutions designed to streamline the SMLM workflow, from acquisition to reconstruction, making it

viable for the demanding realities of a clinical environment.The key contributions of this thesis in the form of chapter wise overview is provided below,

- **Chapter 1** provides a comprehensive introduction to the field of optical super-resolution microscopy, establishing the theoretical foundations of light microscopy, the challenge of the diffraction limit, and details of SMLM. It culminates in identifying the critical challenges faced in SMLM, especially under low resource constraints, which leads to the motivation behind this thesis work.
- **Chapter 2** addresses the need for a unified workflow by introducing a modular, Python-based 2D localisation software. This chapter details the design of a flexible platform that integrates multiple localisation algorithms and rendering strategies. A key feature presented in this thesis is a novel scheme for real-time data processing, enabling immediate feedback on image quality and experimental progress, which is a crucial step towards adaptive imaging.
- **Chapter 3** focuses on the challenge of robustly detecting fluorescent emitters in the presence of noise and non-uniform backgrounds. It introduces a novel adaptive thresholding scheme, which moves beyond conventional static thresholding by dynamically adjusting to image statistics, ensuring reliable PSF detection and minimising false positives and negatives, particularly under challenging signal-to-noise conditions.
- **Chapter 4** tackles the significant bottleneck of localising overlapping PSFs in high-density imaging scenarios. It presents the implementation and validation of a deep learning framework named CAPL (CNN-Assisted PSF Localisation). The chapter demonstrates how the convolutional neural network accurately detects and localises emitters even in crowded fields where traditional fitting algorithms fail, thereby enabling faster acquisitions and improved structural sampling.
- **Chapter 5** transitions to system-level enhancements for drift management. It primarily focuses on an image based real time lateral drift correction using cross-correlation. This method operates directly on the acquired images and corrects them for drift before processing for localisation.
- **Chapter 6** concludes the works presented in this thesis and provides some future directions.

Collectively, the work presented in this thesis aims to democratise high-resolution fluorescence imaging by combining algorithmic innovation with practical engineering. While the implementation is centred on STORM/dSTORM modalities, the tools and methodologies developed are generalisable to a broader class of single molecule localisation based

super-resolution techniques, thereby serving as a foundation for robust, high-performance imaging workflows in a variety of biological applications.

CHAPTER 2

A simple and efficient computational framework for localisation and image reconstruction

2.1 Introduction

As discussed in Chapter 1, SMLM techniques such as PALM, STORM, and PAINT have revolutionised biological imaging by enabling resolution beyond the diffraction limit. However, one of the key limitations identified earlier is the computational pipeline required to process the SMLM datasets. Processing these datasets often relies on resource intensive implementations or opaque, black-box software. In most existing workflows, the images are processed only after acquisition using tools such as ThunderSTORM [61], Picasso [32] or SMAP [69]. This post-acquisition processing approach increases storage demands, as all raw SMLM frames must be retained. In clinical or high-throughput environments, where hundreds of samples may be analysed daily, storing such large datasets is impractical.

This chapter presents the development of a lightweight, modular, and interpretable Python- based software framework [70] that supports both offline and real-time localisation and reconstruction in SMLM. The framework is designed to balance performance with simplicity, enabling robust localisation even on modest hardware. Its modular design allows researchers to replace or extend specific components according to experimental requirements. For real-time operation, the framework is adapted for the open-source microscope control software Micro-Manager [71] through a custom plugin, enabling immediate processing of images during acquisition and offering the option to discard raw frames after localisation, thereby greatly reducing storage requirements without compromising reconstruction quality.

The following sections describe the framework in detail, including its design, implementation, and evaluation. Each module is presented with its underlying methodology, and the system's performance is demonstrated on both simulated and experimental datasets.

2.2 Software Architecture and Workflow

The computational pipeline to generate a super-resolved image from raw microscope data follows a well-defined sequence of steps, as illustrated in Figure 2.1. As discussed in the previous chapter, each frame of an SMLM dataset contains only a sparse subset of active fluorophore PSFs, so the reconstruction pipeline begins by filtering the raw image to reduce background and enhance the PSFs. Candidate emitters are then detected and assigned approximate positions, which are refined to nanometre precision through sub-pixel localisation. Repeating this process for all frames yields a list of fluorophore coordinates that is subsequently rendered to produce the final super-resolved image. The software framework presented here is designed to execute this pipeline in a flexible, modular form.

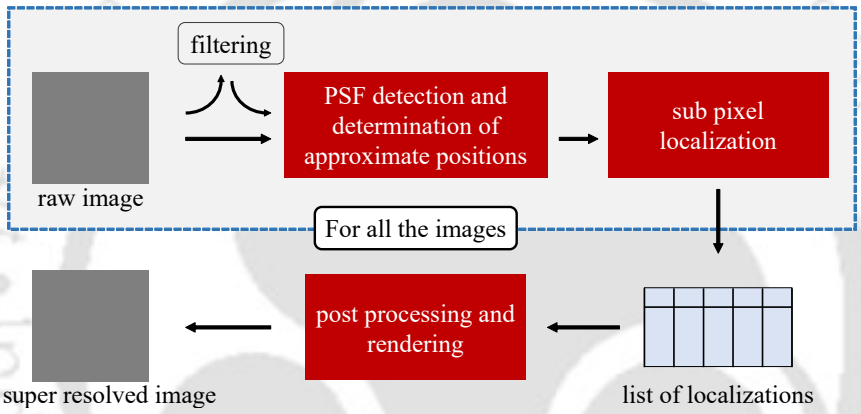

Figure 2.1. Localisation and reconstruction process in SMLM.

The software has been implemented using Python, an open-source programming language. In addition to the custom-developed codebase, it relies on widely used libraries including NumPy [72], SciPy [73], scikit-image [74], openCV [75], and matplotlib [76]. The framework supports two working modes: (i) an offline mode where image acquisition is followed by separate processing, and (ii) a real-time mode in which each frame is processed as soon as it is captured. For offline usage, we developed a Python script that was later adapted into a basic graphical user interface (GUI) as shown in Figure 2.2, enabling post-acquisition analysis. For real-time usage, we developed a plugin for the widely used open-source microscope control software Micro-Manager [71]. While both modes share a common core for image processing, the real-time setup includes an additional feature to delete each raw image after localisation data is extracted and appended to a master list, thereby optimising memory and storage use. We will discuss the real time reconstruction plugin in section 2.8.

The offline analysis is managed through the GUI, as shown in Figure 2.2, which guides the user through the modular processing steps. The workflow begins with the (a) `Load Image`

`Stack` module, where the user selects the input image stack. Subsequently, (b) `Filtering` step can be applied to enhance the signal-to-noise ratio (SNR) of fluorophore emissions using algorithms such as Gaussian blurring or mean filtering. Following this, (c) `Thresholding` algorithm segments the image, generating a mask (binary image) to isolate candidate PSFs from the background. As this is a critical step, these methods will be elaborated upon in Chapter 3. The mask is then passed to the (d) `Crude Localization` module to identify the approximate integer-pixel coordinates of each event. These estimates are subsequently refined to sub-pixel precision by the (e) `Sub-Pixel Localization` module, using methods like iterative least-squares fitting of a 2D Gaussian model. Critical parameters for this stage are also specified, notably the `Window Size`. This parameter defines the square pixel region around each candidate PSF used for fitting (e.g., a size of 7 corresponds to a PSF radius of 3 pixels). From `Window Size`, the program calculates `PSF Radius`, which is used internally. Finally, the resulting list of high-precision localisations is processed by the `Rendering` module, which uses the user-defined `Magnification` factor to construct the final super-resolved image from the localisation data. The localisation data as well as the rendered super resolved image can be saved to disk by the `Save Localizations (CSV)` and `Save Rendered Image Data` buttons.

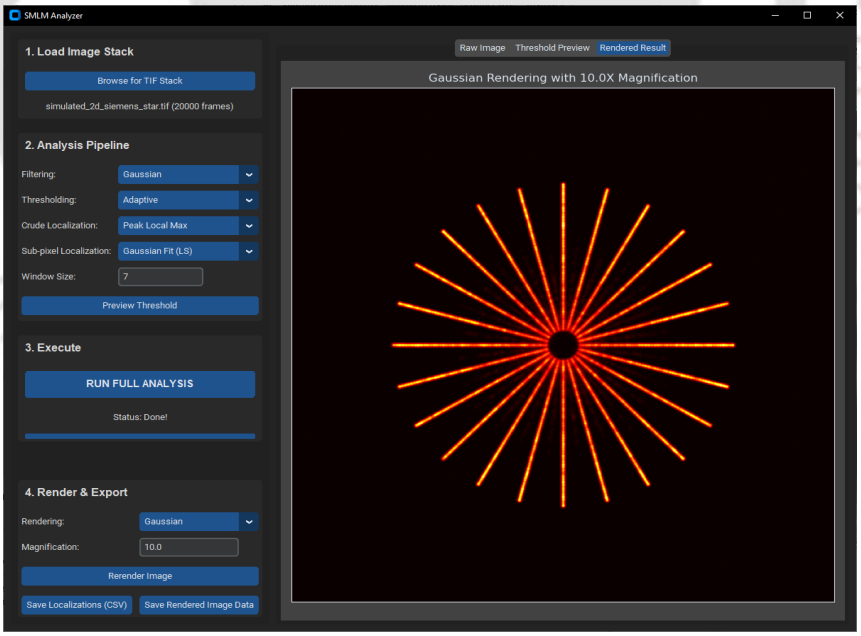

Figure 2.2. Screenshot of the GUI designed for SMLM image reconstruction.

The subsequent sections offer concise descriptions of each of the steps outlined above, along with the corresponding modular code elements and algorithms used.

2.3 Image Preprocessing and Noise Filtering

In fluorescence microscopy, the recorded intensity in each pixel arises from the stochastic arrival of photons emitted by fluorophores, superimposed with background fluorescence and detector readout noise. In imaging, noise refers to unwanted random variations in pixel intensity. The photon-limited nature of SMLM leads to Poisson-distributed shot noise [41], while the camera contributes an approximately Gaussian electronic noise component [77]. Because shot noise scales with the square root of the detected photon number, weak single-molecule emissions can be masked by statistical fluctuations, especially in frames with low activation density or elevated background. Additionally, photons from out-of-focus fluorophores and autofluorescence can introduce spatially varying background levels [78]. These fluctuations can obscure weak signals from individual fluorophores, introduce false positives, or degrade localisation accuracy. Thus, the accurate identification and localisation of single-molecule PSFs in SMLM depends heavily on effective noise suppression to reduce background fluctuations and enhance actual signal components. A detailed discussion on noise relevant to SMLM imaging is provided in Appendix C.

Preprocessing plays a crucial role in this regard; it conditions the raw image data by minimising random variations, compensating for uneven backgrounds, and boosting the overall SNR. These enhancements help isolate PSFs from the background more effectively, improving the reliability of emitter localisation.

In this section, we describe a few commonly used filtering techniques, namely mean (or box) filtering, Gaussian filtering, Laplacian of Gaussian (LoG), and Difference of Gaussian (DoG) filtering. It is worth noting that these filters are used solely for rough PSF detection. Once candidate regions are identified, the final sub-pixel localisation is always performed using the original, unfiltered images. This ensures that the precise intensity distribution of the PSF remains intact during the localisation process.

2.3.1 Mean Filter

Also referred to as the box filter, the mean filter replaces each pixel's value with the average of its neighbouring pixels within a specified window. This approach smooths out high-frequency noise and proves useful in situations with significant noise and sparse emitter distributions. For a kernel of size $n \times n$, the filtered value $I'(x, y)$ is given as,

$$I'(x, y) = \frac{1}{n^2} \sum_{i=-k}^{k} \sum_{j=-k}^{k} I(x + i, y + j), \quad \text{where } k = \frac{n-1}{2} \tag{2.1}$$

Numerical Implementation

The numerical implementation of mean filter relies on the `uniform_filter` function from

the `scipy.ndimage` library. The full, commented Python code is provided in Appendix D, Code Listing D.1.

2.3.2 Gaussian Filter

The Gaussian filter refines the box filter by using a weighted average, where weights diminish with distance from the central pixel. The corresponding 2D Gaussian kernel is expressed as,

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2.2}$$

Gaussian filtering helps preserve the shape of PSFs while reducing background noise and is particularly useful under low-SNR conditions. Both the kernel size and the standard deviation σ can be adjusted based on the nature of the acquired dataset.

Numerical Implementation

The numerical implementation of this filter relies on the `gaussian_filter` function from the `scipy.ndimage` library, which convolves the image with a Gaussian kernel. The Python implementation code is provided in Appendix D, Code Listing D.2.

2.3.3 Laplacian of Gaussian (LoG)

The Laplacian of Gaussian filter detects rapid intensity transitions, highlighting edges and blob-like regions in the image. It performs Gaussian smoothing followed by application of the Laplacian operator, which is defined as,

$$\nabla^2 G(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6}\right) \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2.3}$$

This filter is particularly suited to locating circular features such as PSFs and excels at eliminating both uniform and spatially structured backgrounds. It is frequently employed in blob detection applications and is effective in dense or complex imaging environments.

Numerical Implementation

This filter is implemented using the highly efficient `gaussian_laplace` function from the `scipy.ndimage` library, which computes the Laplacian of a Gaussian-smoothed image. The full Python code is provided in Appendix D, Code Listing D.3.

2.3.4 Difference of Gaussian (DoG)

The DoG filter serves as an efficient approximation of the LoG and is implemented by subtracting two Gaussian-blurred versions of the same image, each using a different standard

deviation. It is defined as,

$$\mathrm{DoG}(x, y) = G(x, y; \sigma_1) - G(x, y; \sigma_2) \tag{2.4}$$

where $\sigma_2 \gg \sigma_1$ This approach emphasises spatial frequencies within a defined range and is computationally less expensive than LoG, while still offering reliable PSF detection performance. It can be tuned to highlight features of a particular scale and is relatively insensitive to absolute signal levels.

Numerical Implementation

The implementation of this filter involves two applications of the `gaussian_filter` function from `scipy.ndimage` with two different standard deviations (σ_1 and σ_2). The final filtered image is obtained by subtracting the result of the wider blur from the narrower one. The complete function is available in Appendix D, Code Listing D.4.

Each of these filters offers specific benefits depending on the imaging scenario. For images with sparse emitters or low noise levels, simple box or Gaussian filtering may suffice. In contrast, LoG and DoG filters are more robust in noisy or crowded conditions, as they accentuate PSF-like structures while suppressing variations in background intensity.

2.4 PSF Detection and Crude Localisation

The next step in the SMLM workflow is to estimate the approximate positions of fluorescent emitters, denoted as $(\tilde{x_{cr}}, \tilde{y_{cr}})$, which we refer to as crude localisation. This step identifies likely regions containing single-molecule PSFs, which are then subjected to precise sub-pixel refinement. It is thus very important to detect the PSFs accurately and with high precision, such that all the fluorescence emission is detected, but at the same time, noise is excluded. Figure 2.3 (a) shows an image with a valid PSF, while (b) shows an image dominated by noise.

In our approach, the framework provides three distinct strategies for initial PSF detection and crude localisation, which are: Peak Local Maximum (PLM), Centre of Mass (COM), and Blob Detection. Each method presents a trade-off among speed, precision, and robustness to noise. We will discuss them in detail in the next few sections.

2.4.1 Peak Local Maximum (PLM)

This method detects PSFs by identifying pixels that are local maxima within a defined neighbourhood window. PLM is computationally efficient and best suited for scenarios with sparsely distributed emitters. A pixel is considered a valid local maximum if its intensity is higher than that of all pixels within a window (defined by `PSF Radius`) and also exceeds

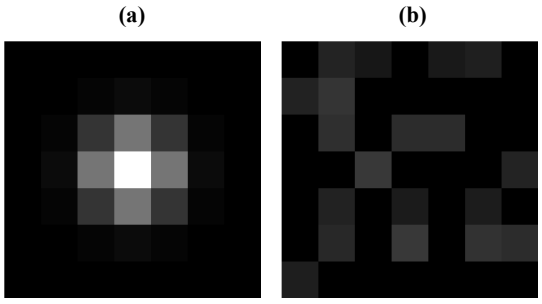

Figure 2.3. (a) A PSF array with symmetric intensity fall-off from the central peak. (b) An example image with noise, illustrating asymmetric intensity variation.

the specified `Threshold`. To prevent duplicate detections, peaks that are too close to each other are merged. While fast, PLM is vulnerable to detecting spurious peaks in high-noise conditions and may fail to detect faint PSFs. The process is summarised in Algorithm 1.

Algorithm 1: Crude localisation using Peak Local Maximum

Input: {image, Threshold, PSF Radius}
Output: $\{\tilde{x_{cr}}, \tilde{y_{cr}}\}$; coordinates of crude PSF centers
1. Set a neighbourhood defined by `PSF Radius`
2. Find local maxima above the specified `Threshold`
3. Merge peaks that lie within the minimum allowable distance
4. Output the coordinates of the remaining peaks

Numerical Implementation

This method is implemented using the `peak_local_max` function from the `skimage.feature` library, which efficiently identifies pixel intensities that are maxima within a defined neighbourhood. The full Python function is provided in Appendix D, Code Snippet D.5.

2.4.2 Centre of Mass (COM)

The COM technique estimates the centroid of intensity patterns indicative of PSFs. Unlike PLM, it incorporates the pixel intensity distribution, making it more precise in low-noise settings. Initially, local maxima and minima are identified. A peak is retained only if its contrast with the nearest minimum exceeds the `Threshold`. For each such region, the centroid is calculated as an intensity-weighted average within a square window defined by `PSF Radius` as,

$$[x_{centre}, y_{centre}] = \left[\frac{\sum x\, I(x,y)}{\sum I(x,y)}, \frac{\sum y\, I(x,y)}{\sum I(x,y)}\right] \tag{2.5}$$

While more accurate than PLM in low-noise cases, it is computationally heavier and may falter under severe noise. The process is summarised in Algorithm 2.

Numerical Implementation

The implementation follows a two-step process. First, valid peaks are identified in regions

Algorithm 2: Crude localisation using Centre of Mass (COM)

Input: {image, Threshold, PSF Radius}
Output: $\{\tilde{x_{cr}}, \tilde{y_{cr}}\}$; coordinates of centroid positions
1. Detect local maxima and minima using filtering operations
2. Retain maxima with sufficient intensity contrast
3. For each valid peak, define a window of `PSF Radius`
4. Calculate the centroid using an intensity-weighted average
5. Output the centroid coordinates

where the local intensity contrast, calculated using `maximum_filter`, exceeds the set threshold. The centroids of these valid regions are then precisely determined using the `centre_of_mass` function. All routines are from the `scipy.ndimage` library. The full implementation is provided in Appendix D, Code Snippet D.6.

2.4.3 Blob Detection using Laplacian of Gaussian

As previously discussed, the Laplacian of Gaussian (LoG) operator is highly effective for highlighting blob-like features of a specific size. In the context of PSF detection, this approach is particularly robust as it suppresses both uniform and structured background while enhancing emitter signals. The method operates by creating a multi-scale representation of the image, applying LoG filters at various scales to find candidate PSFs, which appear as local maxima in this scale-space [79]. Detections that overlap or fall outside the expected size limits of a PSF are then discarded. As it will be demonstrated in Section 2.4.4, this approach is the most reliable of the three crude localisation methods, especially in noisy conditions. The general process is summarised in Algorithm 3.

Algorithm 3: Crude localisation using LoG-based blob detection

Input: {image, threshold, PSF Radius}
Output: $\{\tilde{x_{cr}}, \tilde{y_{cr}}\}$; coordinates of detected PSFs
1. Apply a multi-scale LoG filter to the image
2. Identify local maxima in the resulting scale-space
3. Use thresholding to filter weak detections
4. Discard blobs smaller than a minimum size or larger than `PSF Radius`
5. Return the coordinates of valid detections

Numerical Implementation

This method is implemented using the `blob_log` function from the `skimage.feature` library, which provides a robust and efficient routine for detecting features across multiple scales. The full Python function can be found in Appendix D, Code Snippet D.7.

2.4.4 Quantitative Comparison of Crude Localisation Methods

To objectively evaluate the performance of the three crude localisation methods, a quantitative analysis was performed on a large synthetic dataset. A stack of 10,000 images, each containing a random number of PSFs with known ground-truth positions, was generated. Each image was corrupted with Mixed Poisson-Gaussian (MPG) noise [40, 39, 41], to simulate realistic experimental conditions. The performance of PLM, COM, and blob detection was assessed based on two key metrics: computational speed and detection accuracy.

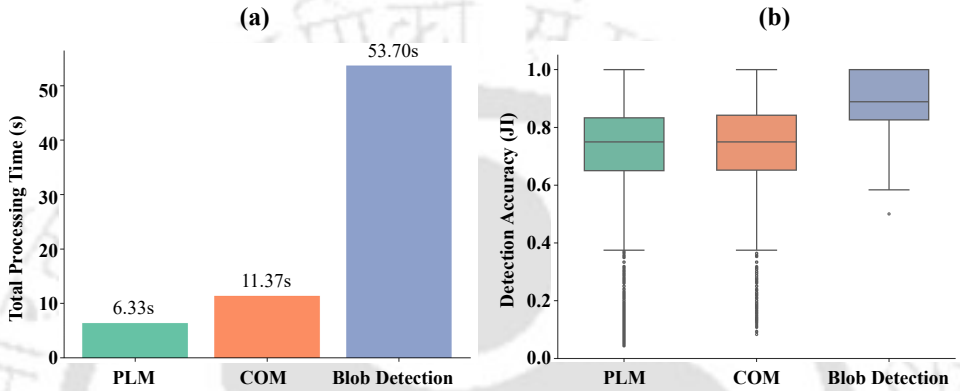

Figure 2.4. Comparison of Crude Localisation Methods. (a) Computational speed of the three crude localisation methods, measured as the total time required to process a dataset. (b) Detection accuracy of the crude localisation methods, evaluated using the Jaccard Index (JI). The box plot shows the distribution of accuracy scores across the dataset. The horizontal line indicates the median, the box represents the interquartile range (IQR), and the whiskers extend to 1.5 times the IQR.

Computational speed in terms of time taken in seconds, shown in Figure 2.4 (a), reveals the expected trade-off. PLM is the fastest method, as it involves only a simple search for local intensity peaks. COM is slightly slower due to the additional calculations required to find local minima for contrast evaluation and to compute the intensity-weighted centroid. Blob Detection is the most computationally intensive, as it involves convolving the image with LoG filters at multiple scales to build the scale-space representation. The analysis of detection accuracy, presented in Figure 2.4 (b), demonstrates the practical benefit of the more computationally intensive methods. Accuracy is quantified using the Jaccard Index, as described in Chapter 1, which is a robust metric that penalises both false positives (spurious detections) and false negatives (missed PSFs). The corresponding mean JI scores are 0.693, 0.721 and 0.889 respectively for PLM, COM and blob detection respectively. The results show that PLM has the lowest median accuracy and the widest distribution of scores, confirming its high sensitivity to noise. COM performs better, achieving a higher median accuracy, but its performance is still inconsistent, as indicated by the large interquartile range. Blob Detection clearly emerges as the superior method, exhibiting the highest median Jaccard

Index and the narrowest distribution of scores. This indicates not only higher accuracy but also significantly more reliable and consistent performance across the entire dataset. The inherent background subtraction and scale-specific filtering of the LoG operator make it highly resilient to noise, allowing it to robustly distinguish true PSFs from background fluctuations. While PLM and COM may offer advantages in speed for very clean data, this quantitative analysis confirms that blob detection provides the best balance of accuracy and reliability for typical SMLM data. Figure 2.5 illustrates a visual comparison of detection made by PLM, COM, and Blob detection on a random simulated noisy SMLM image from the stack.

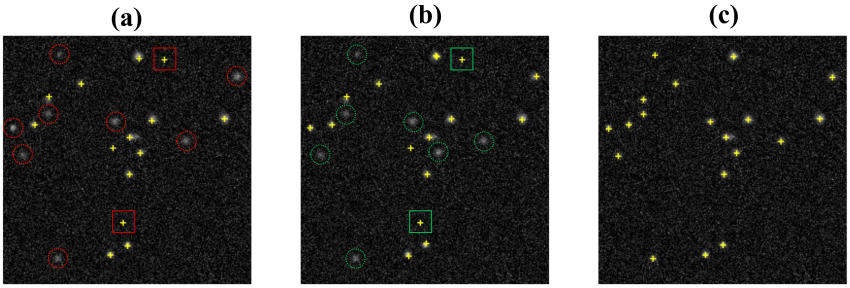

Figure 2.5. Crude localisation of PSF using (a) peak local maximum, (b) centre of mass and (c) blob detection methods on a simulated noisy image. True positive detections are marked with yellow '+'. Missed detections (false negatives) are shown in circles, while false positive detections are marked with squares.

2.5 Sub-Pixel Localisation

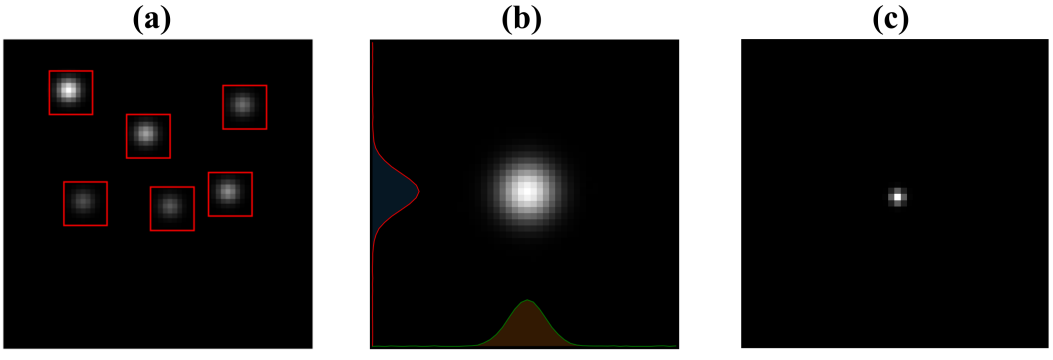

Figure 2.6. (a) A representative SMLM image with the PSFs being detected are highlighted in red squares. (b) An extracted PSF sub-array using crude localisation, (c) Sub-pixel localisation of the extracted PSF being represented as a Gaussian blob with its width corresponding to its localisation uncertainty.

Sub-pixel localisation refers to the process of refining the positions of detected PSFs corresponding to individual fluorescent molecules such that their coordinates are estimated

with greater precision than the native resolution of the camera pixels. This method allows localisation accuracy well below the pixel size by fitting a theoretical PSF model to the observed image data, while accounting for diffraction and optical aberrations introduced by the imaging system.

Several techniques are available for sub-pixel localisation, including least squares fitting (LS), maximum likelihood estimation (MLE), and phasor-based localisation algorithm (PLA). Each applies distinct mathematical principles and optimisation frameworks to match a predefined PSF model with the image of the emitter and determine its coordinates with high precision.

The procedure begins by extracting a sub-array from the raw image corresponding to each detected PSF, using the approximate coordinates $(\tilde{x_{cr}}, \tilde{y_{cr}})$ obtained through crude localisation. The area selected for fitting, referred to as the fitting window, D, is determined by the parameter `Window Size`, and is mathematically described by,

$$D = \{(x, y); x \in \{\tilde{x_{cr}} - \texttt{Window Size}/2, \dots, \tilde{x_{cr}} + \texttt{Window Size}/2\}, \\ y \in \{\tilde{y_{cr}} - \texttt{Window Size}/2, \dots, \tilde{y_{cr}} + \texttt{Window Size}/2\}\} \tag{2.6}$$

These cropped PSF regions, as seen in Figure 2.6 (b), are then matched to a model PSF function to extract the exact centre, illustrated in Figure 2.6 (c).

Although the Airy disk pattern is the theoretical PSF derived from diffraction theory, as introduced in Chapter 1, its direct implementation in fitting routines proves computationally intensive. Moreover, due to the limited spatial resolution of imaging sensors and the influence of noise, actual PSF measurements deviate from the Airy model. Consequently, the use of a 2D Gaussian function offers a practical and sufficiently accurate approximation of the PSF in SMLM [42], while remaining computationally tractable [43, 44]. The Gaussian approximation is justified because, for typical numerical apertures and pixel sizes, the central lobe of the Airy pattern dominates the recorded signal, and the side lobes fall below the noise floor. This results in a near-Gaussian intensity distribution over the region used for localisation. Moreover, the Gaussian model admits closed-form derivatives, enabling efficient optimisation and yielding localisation uncertainties that agree well with physically derived Cramér–Rao lower bound (CRLB), which specifies the lowest achievable localisation uncertainty permitted by the statistics of photon detection for a given photon budget. [33, 34, 43]. The Gaussian model is given as,

$$F(x, y; \Theta) = I_0 \, exp\left(-\frac{(x - x_0)^2}{2\sigma_x^2} - \frac{(y - y_0)^2}{2\sigma_y^2}\right) + b \tag{2.7}$$

where $\Theta = (I_0, x_0, y_0, \sigma_x, \sigma_y, b)$, with I_0 denoting the peak intensity, (x_0, y_0) as the centre of the Gaussian, σ_x and σ_y representing the spread along each axis, and b being the constant background level.

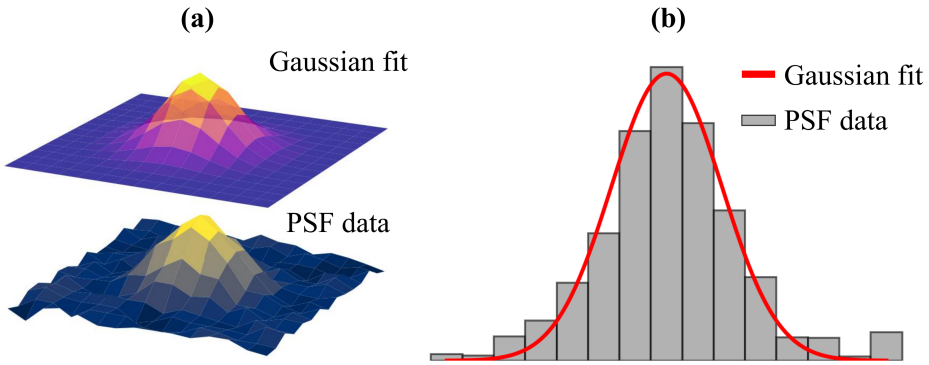

Figure 2.7. Surface plot of the (a) raw image data and the fitted Gaussian. (b) Line plot through the raw image and the fitted Gaussian.

Figure 2.7 (a) shows a surface plot of the observed PSF data with its Gaussian fit, while Figure 2.7 (b) provides a line plot comparison through the centre pixel.

2.5.1 Least Squares Fitting

Least squares (LS) fitting is a widely used method to extract model parameters by minimising the squared differences between observed pixel intensities and those predicted by a model. The cost function minimised during the fitting is given as,

$$S = \sum_{all\ pixels} [I(x, y) - F(x, y; \Theta)]^2 \tag{2.8}$$

where the model parameters Θ are iteratively refined to best fit the data [35]. The general workflow is summarised in Algorithm 4.

Algorithm 4: Sub-pixel localisation using 2D least squares fitting

Input: $\{(\tilde{x_{cr}}, \tilde{y_{cr}})$, image, window_size, init_params$\}$
Output: $\{amplitude, x_{sp}, y_{sp}, \sigma_x, \sigma_y\}$
1. For each crude localisation, extract a PSF sub-image of size `window_size`
2. Use `scipy.optimize.curve_fit()` to fit the 2D Gaussian PSF model
3. Adjust fitted coordinates to the global image frame
4. Return the complete set of fitted parameters

Numerical Implementation

Our implementation of LS fitting is centred around the `curve_fit` function from the `scipy.optimize` library. While multiple optimization strategies for LS fitting exist, such

as gradient descent [80] or the Nelder-Mead method [81], our implementation uses the robust and fast Levenberg-Marquardt (LM) algorithm [82]. A helper function is used to safely extract the PSF sub-image, even for molecules near the frame boundaries. The complete code is provided in Appendix D, Code Snippet D.9.

2.5.2 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) [83, 84] is a statistical fitting approach that incorporates noise characteristics in its model. This method determines the most likely parameter values by maximising the likelihood that the observed data originated from the proposed model. MLE is particularly well suited to localisation microscopy because the detection of fluorescence photons follows Poisson counting statistics. The likelihood formalism directly encodes the expected behaviour of a photon-limited detection process, ensuring that the fitted parameters correspond to the maximum-probability physical emitter position given the observed photon distribution [35].

Given the parametric model $F(x, y; \Theta)$, the likelihood function $L(I; \Theta)$ for a measured PSF intensity $I(x, y)$ under Poisson noise is given as,

$$L(I;\Theta) = \prod_{x,y} \frac{F(x,y;\Theta)^{I(x,y)}\, e^{-F(x,y;\Theta)}}{I(x,y)!} \tag{2.9}$$

The exponents can be linearised by implementing logarithms which gives the log-likelihood function $LL(I; \Theta)$. Neglecting constant terms, $LL(I; \Theta)$ becomes

$$LL(I;\Theta) = \sum_{x,y} \left[I(x,y) \cdot \log(F(x,y;\Theta)) - F(x,y;\Theta) \right] \tag{2.10}$$

Optimising the log-likelihood is equivalent to minimising its negative, which gives us the required parameters Θ^* as

$$\Theta^* = argmin(\Theta) \left[-\sum_{x,y} \left[I(x,y) \cdot \log(F(x,y;\Theta)) - F(x,y;\Theta) \right] \right] \tag{2.11}$$

The general process is summarised in Algorithm 5.

Numerical Implementation

The core of our MLE implementation is a custom cost function that calculates the negative log-likelihood for a Poisson noise model, as described by equation 2.10. This cost function is then minimized using the `minimize` routine from the `scipy.optimize` library, with the memory efficient 'L-BFGS-B' algorithm. This approach statistically determines the most probable set of PSF parameters. The full Python implementation is provided in Appendix D, Code Listing D.10.

Algorithm 5: PSF localisation using MLE

Input: { $(\tilde{x_{cr}}, \tilde{y_{cr}})$, image, window_size, init_params}
Output: $\{amplitude, x_{sp}, y_{sp}, \sigma_x, \sigma_y\}$
1. Extract a window of dimensions `window_size` around each crude PSF coordinate
2. Evaluate the log-likelihood using the model Gaussian
3. Minimise the negative log-likelihood via `scipy.optimize.minimize()`
4. Return the optimal parameters

2.5.3 Phasor-Based Localisation

In addition to the iterative techniques of LS and MLE, a fast, non-iterative phasor-based approach is also incorporated in the framework [85, 86]. It operates in the frequency domain, offering rapid estimation of PSF centres. This method calculates the 2D Fast Fourier Transform (FFT) of the PSF region and extracts the first-order Fourier coefficients in the x- and y-directions as,

$$F_x = r_x e^{i\theta_x}, \quad F_y = r_y e^{i\theta_y} \tag{2.12}$$

The phase angles θ_x and θ_y directly encode the emitter's position relative to the centre of the window. If the PSF sub-image has dimensions $D \times D$, the sub-pixel location within that window is given by,

$$x_0 = \frac{\theta_x}{2\pi} D, \quad y_0 = \frac{\theta_y}{2\pi} D \tag{2.13}$$

As a non-iterative method, the phasor approach is computationally very efficient, though it can be less precise than MLE under low signal-to-noise conditions. The process is summarised in Algorithm 6.

Algorithm 6: PSF localisation using the phasor-based approach

Input: $\{(\tilde{x_{cr}}, \tilde{y_{cr}})$, image, window_size }
Output: $\{x_{sp}, y_{sp}\}$; sub-pixel coordinates
1. Extract a PSF sub-image of size `window_size` centered at $(\tilde{x_{cr}}, \tilde{y_{cr}})$
2. Compute the 2D FFT of the sub-image
3. Determine the phase angles θ_x and θ_y from the first-order Fourier coefficients
4. Calculate the sub-pixel coordinates relative to the window using the phase angles
5. Adjust coordinates to the global image frame
6. Return the final sub-pixel coordinates

Numerical Implementation

This computations implemented in this method is mostly in the frequency domain. The core routine computes the 2D Fast Fourier Transform of the PSF sub-image using the `fft2` function from NumPy. The sub-pixel coordinates are then calculated directly from the

phase angles of the first-order Fourier coefficients, providing a rapid position estimate. The complete function is available in Appendix D, Code Snippet D.11.

2.5.4 Comparison of Sub Pixel Localisation Methods

The selection of an appropriate sub pixel localisation method requires balancing computational efficiency with localisation accuracy. The three methods considered in this study, namely LS fitting, MLE, and the phasor based approach, represent different points along this spectrum. To evaluate their performance in a systematic way, a quantitative analysis was carried out using a synthetic dataset of 500 PSFs of size (7 × 7) pixels with known ground truth positions that were corrupted with MPG noise.

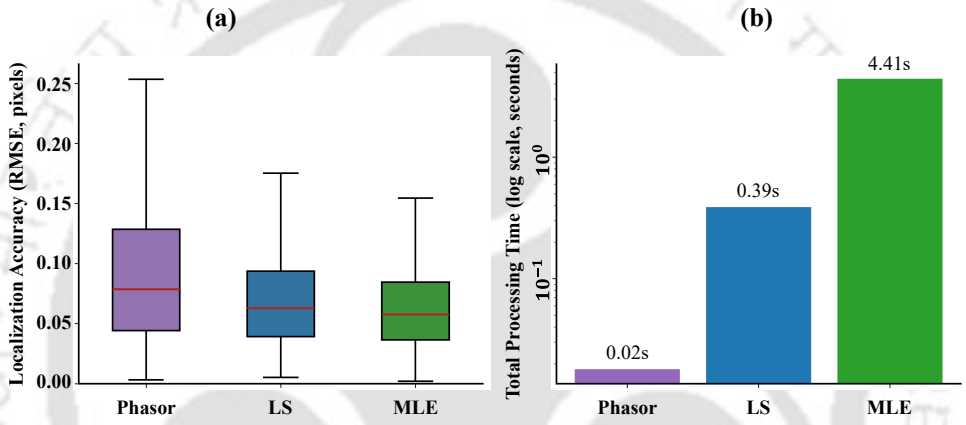

Figure 2.8. Quantitative comparison of the speed and accuracy of the sub pixel localisation methods on a dataset of 500 simulated PSFs. (a) Localisation accuracy, shown as a box plot of the localisation errors. The box indicates the median and interquartile range of the errors. (b) Computational speed, measured as the total time required to process the dataset (presented in logarithmic units).

From a theoretical perspective, MLE is expected to be the most accurate and reliable method. By explicitly modelling Poisson noise, it provides a statistically rigorous estimate of the emitter position and, under suitable conditions, and can achieve the CRLB for precision [33, 78]. Least squares fitting performs almost as well, particularly in conditions of high signal to noise ratio, but its precision decreases when the photon count is low [34]. The phasor based method, which avoids iterative fitting, is generally the least accurate because of its higher sensitivity to noise. This hierarchy is consistent with the quantitative results shown in Figure 2.8 (a). The phasor method produced the largest RMSE of 0.1174 pixels, whereas least squares fitting reduced this error to 0.0829 pixels. Maximum likelihood estimation achieved the best performance with a RMSE of 0.0755 pixels, thereby demonstrating its superior accuracy. A qualitative example of this behavior is shown in Figure 2.9 (b), which illustrates typical localisation results for the three methods on a single point spread function.

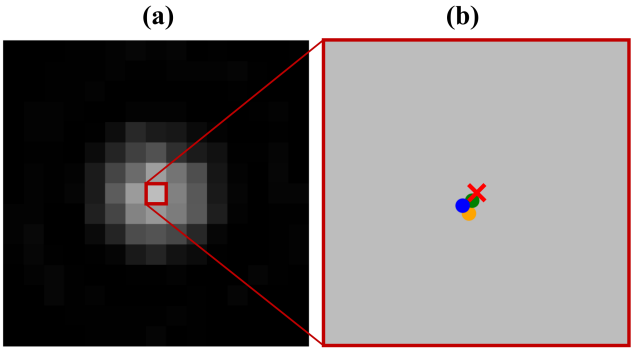

Figure 2.9. Qualitative localisation of a single PSF. **(a)** A simulated 7x7 pixel PSF. **(b)** The localised positions from LS (orange), MLE (green), and Phasor (blue) methods, with the true centre indicated as red cross.

When computational speed is considered, the trade off is reversed. The phasor based method is the most efficient, since it requires only a single fast Fourier transform for each point spread function [86]. Both least squares fitting and maximum likelihood estimation rely on iterative optimisation, which is inherently slower. As shown in Figure 2.8 (b), this difference is evident. The phasor method completed the analysis of the entire dataset in only 0.02 seconds. LS fitting required 0.39 seconds, while MLE was the slowest, taking 4.41 seconds to process the same data. These results confirm the expected trade off. While MLE remains the method of choice when the highest localisation accuracy is required, the phasor method is clearly preferable in scenarios where computational throughput is critical, such as real time acquisition feedback [87].

2.6 Rendering of Super-Resolved Image

The sub-pixel localisation module generates a list of coordinates and the corresponding standard deviations for each PSF identified in the raw images. Since the camera sensor is indexed using integer pixel coordinates, the obtained PSF positions are usually represented with decimal precision. This coordinate list serves as the basis for constructing super-resolved images using different visualisation techniques.

Visualisation plays a key role in interpreting super-resolution microscopy data, as it reveals the spatial organisation of molecules within the sample. Each localisation corresponds to an estimate of the true fluorophore position derived from the spatial distribution of detected photons. Rendering methods are therefore not merely visual tools but representations of the underlying molecular density. Depending on the rendering scheme, the output image may reflect either the raw localisation positions (as in scatter plots) or an estimate of the continuous emitter distribution (as in Gaussian rendering), which more closely represents the physical precision with which each molecule was localised.

In our implementation, we provide four commonly adopted rendering strategies, which are scatter plot, histogram, averaged shifted histogram, and Gaussian rendering. Figure 2.10 offers a visual comparison of the different methods supported by our software.

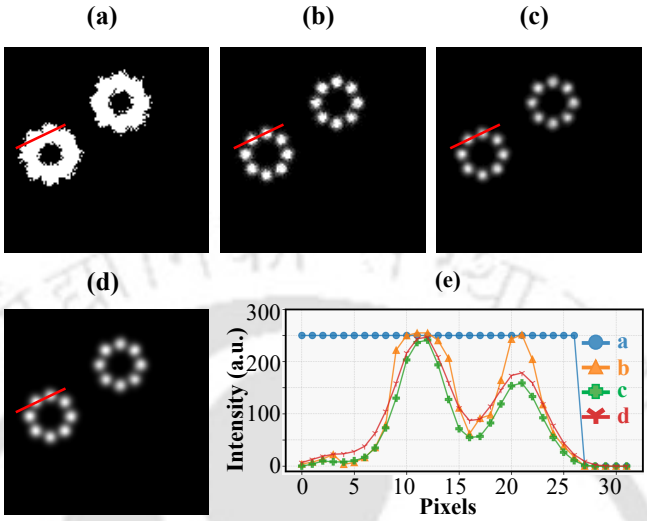

Figure 2.10. Visualisation of the reconstructed image using (a) scatter plot, (b) 2D histogram, (c) averaged shifted histogram. (d) Gaussian rendering. (e) show line plots along the line shown in (a), (b), (c) and (d), respectively.

2.6.1 Scatter Plot

A scatter plot is a straightforward method to display the spatial distribution of PSFs in two dimensions. Each localised emitter is represented as a single point, with its x and y coordinates denoting the sub-pixel position. In our implementation, the scatter plot is rendered by generating a binary image where pixels corresponding to localisation coordinates are assigned a value of 1, while all other pixels retain a value of 0. This process is repeated over all frames, thereby accumulating the localisations into a binary image that reflects emitter positions.

2.6.2 Histogram

Histograms offer another way to visualise molecular localisations by counting the frequency of x and y coordinates across discrete bins. Using NumPy's `histogram2d()` function, our code constructs a 2D histogram where each bin accumulates the number of localisation events falling within its bounds. Unlike the binary output of the scatter plot, this method results in a grayscale image in which pixel intensities are proportional to the number of events, thereby offering an intensity-weighted view of molecular clustering.

2.6.3 Averaged Shifted Histograms

The averaged shifted histogram (ASH) method [88] is a non-parametric density estimation technique well suited for high-resolution data visualisation. It estimates the spatial density of localisations by averaging multiple slightly offset histograms. To create the ASH, the localisation coordinates are first used to construct a 2D histogram. This histogram is then shifted in both the x and y directions using small offsets (e.g., one pixel) via the `numpy.roll()` function. These shifted histograms are recomputed for several offset values and averaged together to produce a smooth approximation of the underlying density distribution. Compared to conventional histograms, the ASH method reduces sensitivity to bin alignment and choice of bin size, thereby yielding a smoother and more robust visualisation of molecular density across the sample.

2.6.4 Gaussian Rendering

Gaussian rendering is widely utilised for visualising localisation data. In this approach, each molecular localisation is represented not as a discrete point but as a 2D Gaussian distribution. The Gaussian is centred on the emitter's sub-pixel position, and its standard deviation can be defined based on the estimated localisation precision [33, 78]. The final image is obtained by summing these Gaussian kernels over a fine-resolution pixel grid, which is often upsampled relative to the original camera resolution to improve visual clarity. This method produces smooth, continuous representations of molecular distributions that better reflect the actual spatial organisation of fluorophores within the sample. Additionally, Gaussian rendering facilitates quantitative analysis, such as spatial correlation and structural feature comparison. Although this method is computationally more demanding than scatter or histogram-based approaches, it remains popular due to its ability to faithfully preserve localisation precision and present data in an intuitive visual format.

2.7 Performance Evaluation of the Proposed Python Code on Standard Datasets

We validate the effectiveness of our code by localising PSFs in raw image stacks and reconstructing the corresponding super-resolved images using both simulated and experimental datasets. We first use experimental raw images available from the SRM-hub repository [87, 57]. The first dataset contains 3200 frames of microtubule structures, with each frame sized at (64×64) pixels and containing a high density of fluorophores. Figure 2.11 (a) shows the widefield image of the sample. To process this data, we apply blob detection for crude localisation and 2D least square fitting for sub-pixel localisation. The final super-resolved

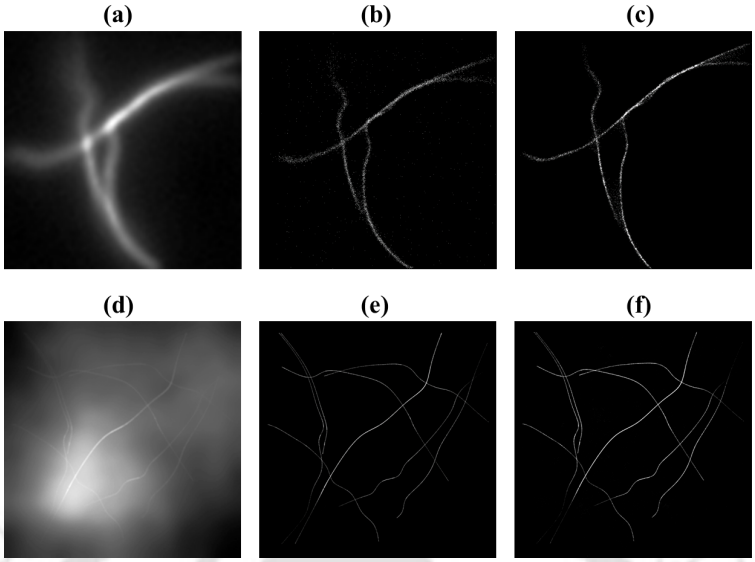

Figure 2.11. (a) and (d) are widefield images of tubulin, (b) and (e) are reconstructions using ThunderSTORM, while (c) and (f) are reconstructions using the proposed scheme.

image is rendered using the Averaged Shifted Histogram technique, as shown in Figure 2.11 (c).

The same dataset is also processed using ThunderSTORM, applying B-spline filtering and local maxima detection for coarse localisation, followed by maximum likelihood estimation for sub-pixel fitting. The resulting super-resolved image, rendered using Averaged Shifted Histograms, is shown in Figure 2.11 (b).

The second experimental dataset consists of 2400 image frames, each of dimension (256 × 256) pixels. These images capture tubulin structures also featuring dense fluorophore distributions. The widefield image for this sample is shown in Figure 2.11 (d), while Figures 2.11 (e) and (f) depict the corresponding super-resolved reconstructions using ThunderSTORM and our code, respectively, again using Averaged Shifted Histogram rendering.

By comparing Figures 2.11 (b) and (c), as well as (e) and (f), it is evident that our Python-based implementation produces super-resolved images of quality comparable to those obtained using ThunderSTORM.

2.8 A Real-Time Image Reconstruction Scheme for SMLM

SMLM datasets are typically very large, often surpassing 20 gigabytes per field of view, which makes both storage and processing highly demanding. This poses significant challenges in clinical environments, such as hospitals, where multiple samples may need to be imaged and storage capacity is limited. Additionally, the time taken for data acquisition and post-processing can delay results, limiting the feasibility of SMLM for time-sensitive tasks

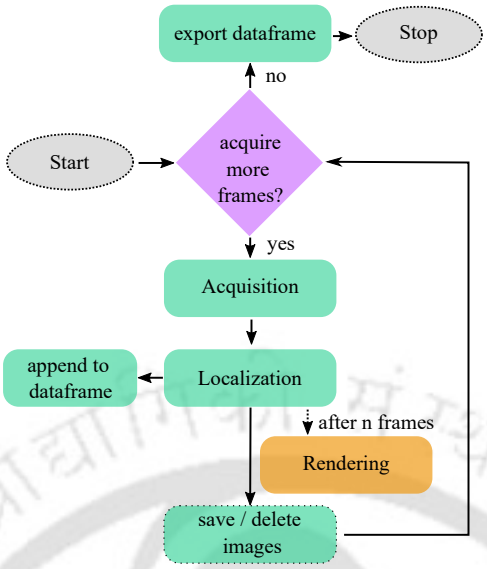

Figure 2.12. Flowchart for real-time SMLM image reconstruction and rendering pipeline. Acquisition proceeds frame by frame, followed by localisation and periodic rendering after every n frames. Localisation results are accumulated in a dataframe, which is exported at the end of the process as a CSV file. Optional image saving / deletion ensures memory efficiency.

like live-cell imaging or rapid diagnostics. Furthermore, real-time processing also offers several scientific advantages. Continuous reconstruction allows immediate assessment of fluorophore density, illumination stability, and photobleaching dynamics, which are fundamentally governed by photon emission statistics and dye photophysics. Monitoring these factors during acquisition ensures that the experimental conditions remain within the regime required for accurate single-molecule localisation.

To address these issues, we have developed a Micro-Manager plugin that interfaces directly with our Python code [89]. It integrates image acquisition and reconstruction within the same environment which enables real-time image reconstruction and visualisation during data acquisition. By incorporating image reconstruction into the acquisition workflow, this approach minimises both data storage demands and processing time, making SMLM more suitable for clinical and research use cases that require dynamic imaging and quick feedback. This advancement brings real-time SMLM closer to practical usage, offering a useful platform for high-throughput imaging and real-time monitoring of biological activity. Figure 2.12 shows a flow chart of the real time SMLM image processing implementation.

2.8.1 Image Acquisition and Parallel Analysis

Before analysis begins, the microscope is configured for STORM imaging and a multi-dimensional acquisition (MDA) session is initiated in Micro-Manager, with settings adjusted to save each frame as an individual file. Once acquisition starts, the custom plugin named

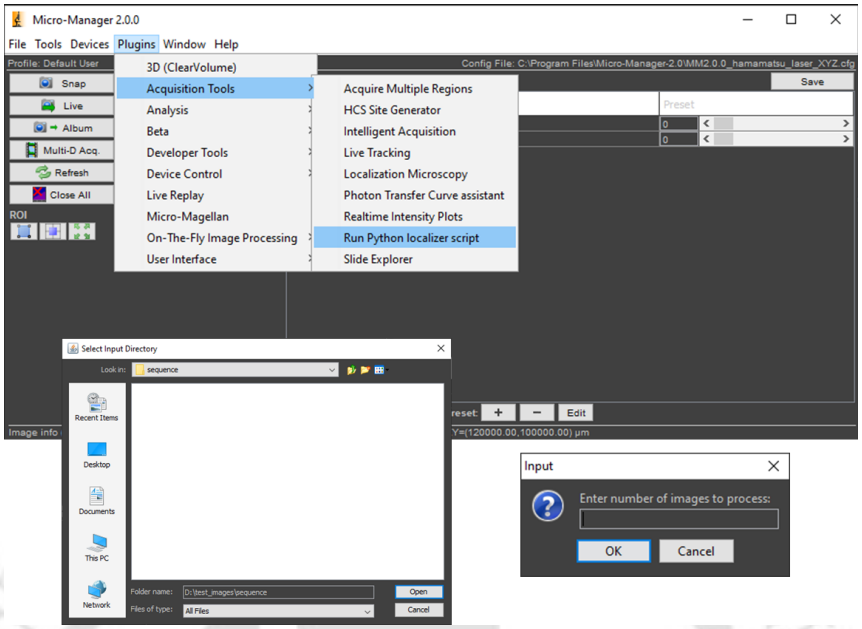

Figure 2.13. Implementation of the plugin *Run Python localisation script* in Micro-manager.

Run Python localizer script is launched within Micro-Manager. This plugin's user interface requires two inputs, the directory where the image frames are being saved and the total number of frames to be acquired. Upon receiving this input, the plugin activates a Python script that runs concurrently with image acquisition. The script continuously monitors the specified directory, processes each new image as it appears, and optionally deletes the processed frames to conserve storage space. We have utilised the fast peak local maximum for crude localisation and phasor based algorithm for near-instantaneous sub pixel localisation. This setup enables parallel acquisition and processing, allowing for real-time image reconstruction and visualisation. A snapshot of the plugin in operation is shown in Figure 2.13.

2.8.2 Application of the Scheme

To validate our plugin, we performed a proof-of-concept SMLM experiment using quantum dots (Qdot™ 705 ITK™ Carboxyl Quantum Dots, Invitrogen, 15–20 nm) in the openFrame modular microscope. Data were acquired at 33 frames per second using a Hamamatsu Orca Flash 4.0 v1 monochrome camera with excitation at 405 nm. A central region of interest (ROI) of dimensions (88 × 88) pixels was selected to ensure uniform illumination. The widefield view of the ROI is shown in Figure 2.14 (a). A total of 10,000 frames were recorded using MDA in Micro-Manager, while our plugin was simultaneously active. The live reconstructed image generated during acquisition is presented in Figure 2.14 (b). For reference, the same dataset was also processed after acquisition using ThunderSTORM, and the reconstructed

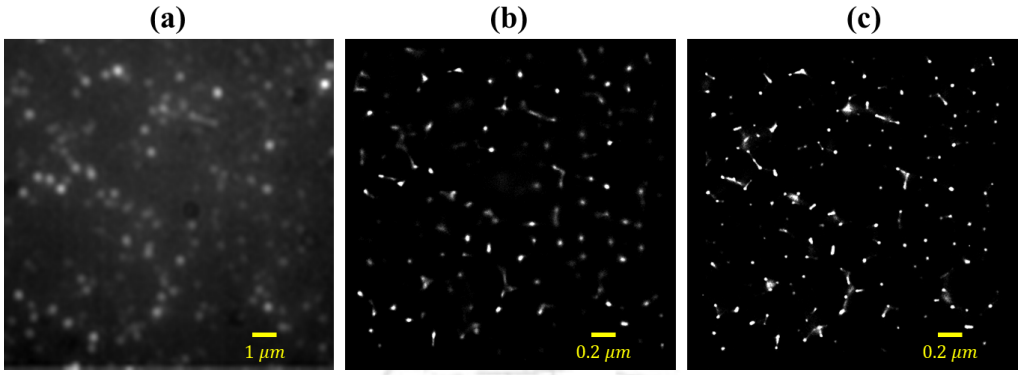

Figure 2.14. Reconstruction performed on a dataset of 10,000 images of blinking quantum dots. (a) shows the widefield image (88 × 88) pixels while (b) shows the real-time reconstructed image (440 × 440) pixels using our approach. (c) shows post-acquisition reconstructed image (440 × 440) pixels using ThunderSTORM.

image is shown in Figure 2.14 (c). Our plugin was able to deliver super-resolved results in real time, achieving reconstruction quality comparable to ThunderSTORM without the delays associated with post-processing.

2.9 Positioning Within the SMLM Software Ecosystem

The software framework presented in this chapter is intended to complement established SMLM analysis platforms such as ThunderSTORM, SMAP and Picasso, which offer extensive functionality and optimised implementations for a wide range of experimental conditions. In contrast, the proposed framework emphasises modularity, simplicity, and ease of extension. Implemented using standard Python-based tools, it allows individual components of the localisation pipeline to be modified or replaced with minimal overhead. This makes it particularly suitable for rapid prototyping and integration of custom methods, including those introduced in subsequent chapters.

Additionally, the lightweight design enables deployment in resource-constrained environments and facilitates real-time processing through integration with Micro-Manager. As such, the framework is best suited for specialised workflows requiring flexibility, customisation, or real-time capability, rather than as a direct replacement for comprehensive existing packages.

2.10 Chapter Summary

This chapter detailed a comprehensive Python-based software developed for the localisation and reconstruction of super-resolved images using single-molecule localisation microscopy. The program is organized in a modular fashion and structured simply to ensure ease of

understanding for general readers and straightforward adaptation to various applications. It supports both offline operation after image acquisition and real-time execution during acquisition. Demonstration examples confirm that the performance of the proposed software is comparable to that of established platforms like ThunderSTORM, particularly in terms of the quality of the super-resolved images it produces.

CHAPTER 3
An adaptive image thresholding method for SMLM images

3.1 Introduction

Single Molecule Localisation Microscopy (SMLM) achieves nanometre-scale resolution by precisely detecting and localising individual fluorescent emitters. As discussed previously, the foundation of this process lies in, what we call crude localisation, which is the accurate identification of fluorophore emitter point spread functions (PSFs) from noisy SMLM images. This initial detection step is crucial because only when PSFs are correctly identified can sub-pixel localisation yield high-quality, high-resolution reconstructions [33]. However, accurate detection remains challenging as PSF shapes vary due to optical distortions. From a physical perspective, the difficulty of detecting PSFs in SMLM originates from the photon-limited nature of fluorescence emission. Aberrations across the field of view, and heterogeneous illumination distort the PSF shape and shift its photon distribution [90, 91]. Additionally, background fluorescence and detector readout noise introduce fluctuations that follow mixed Poisson–Gaussian statistics [41]. These factors reduce the separability of signal and background in the images, making threshold selection a non-trivial problem [92]. A thresholding strategy that incorporates these statistical and optical considerations is therefore essential for reliable PSF detection in realistic imaging conditions.

In this chapter, we propose a novel adaptive thresholding strategy based on entropy analysis, specifically designed to separate PSFs from noise in SMLM images under diverse imaging conditions [93]. This method increases the robustness of PSF detection, particularly in cases involving inhomogeneous backgrounds or densely packed emitters. Once thresholding is complete, a crude localisation technique is applied to identify PSF candidates. We demonstrate that combining this entropy-based thresholding with blob detection significantly enhances the crude localisation step and leads to improved reconstruction of super-resolved images. We also compare the performance of our approach with existing techniques, including the widely used ThunderSTORM software.

3.2 Image Thresholding

Thresholding is a fundamental operation in digital image processing, widely employed to isolate significant features from the background [94]. By converting grayscale images into binary form using a chosen intensity threshold, this technique enables the segmentation of regions of interest such as patterns, objects, or signals from non-relevant background regions. Its simplicity and effectiveness make it indispensable across numerous fields, including medical diagnostics, document processing, industrial inspection, and advanced microscopy.

In SMLM, thresholding serves as a critical preprocessing step that allows for the identification of PSFs generated by fluorescent emitters [87]. The ability to distinguish these PSFs from background noise directly impacts the accuracy of subsequent localisation procedures and the quality of the final super-resolved image. In this context, reliable thresholding is essential for maintaining both spatial resolution and structural fidelity in reconstructed images.

3.2.1 Current Global and Local Thresholding Methods

A broad range of thresholding algorithms have been developed to address the diverse challenges associated with image segmentation, particularly in biomedical applications [95, 63, 96, 97]. One of the most straightforward is mean thresholding [94], where pixels are classified based on whether their intensity is above or below the image's global mean. This method presumes uniform illumination which is very uncommon in fluorescence microscopy. A more sophisticated global technique is Otsu's method [98], which selects a threshold by maximising the variance between the foreground and background pixel classes. Although effective in many cases, Otsu's algorithm assumes a bimodal histogram, which often fails to represent the intensity distributions in SMLM data where the foreground signal is too sparse to form a distinct second peak. Entropy-based approaches, such as Kapur's method [99], overcome this by maximizing the entropy of the segmented histogram. Extensions like Renyi's entropy method [100] and the minimum cross-entropy technique [101] further refine threshold selection, offering improved performance for multimodal histograms.

Local thresholding techniques have also gained traction, particularly for handling spatially varying illumination. Niblack's method [102] computes a local threshold for each pixel using neighbourhood statistics such as the local mean and standard deviation. Sauvola's technique [103], which builds on Niblack's algorithm, introduces contrast-dependent adjustments that make it more robust to intensity variations. Another notable method, the

T-point algorithm [104], calculates the threshold by fitting two linear segments to the tail of the histogram and identifying their intersection.

While each of these techniques offers certain advantages, they are not very effective when applied to SMLM images, which, as discussed, are characterised by sparse, low-intensity fluorophore signals superimposed on non-uniform and noisy backgrounds. In the next sections, we describe a new entropy-based adaptive thresholding method tailored to these challenges, designed to improve detection performance specifically in the context of SMLM datasets.

3.3 The Proposed Thresholding Method

As discussed in the previous chapters, the process of generating a super-resolved image from an SMLM dataset involves three prominent steps. First, each individual frame in the dataset undergoes pre-processing to differentiate potential PSFs from background noise. The resulting pre-processed image is then processed using a crude localisation algorithm to determine the approximate positions of the detected PSFs. Finally, these coordinates are refined through sub-pixel localisation techniques, which provide nanometre-scale accurate positions of the fluorophore emitters, which are then used for image reconstruction.

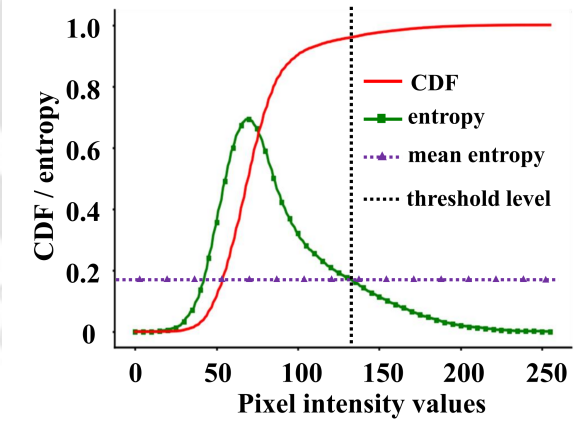

Figure 3.1. Plots showing the cumulative distribution function (CDF), $F(x)$, the entropy function $S(x)$, and the mean entropy as a function of pixel intensity x.

Here we propose an adaptive thresholding technique designed to dynamically distinguish fluorophore PSFs from noise during the detection and crude localisation step. The method begins with computing the histogram of pixel intensities from the input image, denoted by $N(x)$, where x is the pixel value and $N(x)$ represents the number of pixels having that

intensity. This histogram is normalised to generate a probability density function as

$$f(x)\,dx = \frac{N(x)}{\sum_{0}^{x_{\max}} N(x)}\,dx \tag{3.1}$$

This expression describes the likelihood of observing a pixel with intensity in the range x to $x + dx$ [105], with the maximum pixel value as x_{max}. The cumulative distribution function (CDF) [105] is then computed as

$$F(x) = \int_{0}^{x} f(y)\,dy \tag{3.2}$$

Using this CDF, we define the entropy function $S(x)$ [106], which measures the information content associated with the image and divides the pixel intensities into two classes at a given threshold x as

$$S(x) = -F(x)\log\left(F(x)\right) - [1 - F(x)]\log\left[1 - F(x)\right] \tag{3.3}$$

where $x \in (0, x_{\max})$. This entropy formulation effectively quantifies how well a threshold value partitions the image into two informative regions, i.e. foreground and background.

The optimal threshold is determined by identifying the intensity level at which the entropy value most closely matches its mean. This corresponds to the coordinate point on the higher side of the pixel intensity values where the entropy curve intersects the average entropy level. This can be mathematically expressed as,

$$\text{optimum\_threshold} = \max\left(x \mid |S(x) - \text{mean}(S(x))| < \alpha,\ x > 0\right) \tag{3.4}$$

where α is a small tolerance value, typically set around 10^{-2}. Figure 3.1 illustrates the CDF and entropy functions alongside the mean entropy line used in threshold selection.

In fluorescence microscopy, the pixel-intensity histogram typically consists of a broad background distribution superimposed with a sparse tail of higher-intensity pixels corresponding to emitters. The entropy curve reflects how informative a particular division of this histogram is. Low thresholds overemphasise background variability, while very high thresholds suppress genuine emitters. The mean entropy represents a balance point between the background and the PSF information. Selecting the threshold at the last intersection between the entropy curve and its mean value ensures that the chosen threshold lies at the transition where signal-bearing pixels begin to dominate the information content. This criterion provides a physically interpretable boundary between noise and fluorophore emissions, and performs robustly even when the histogram lacks clear bimodality.

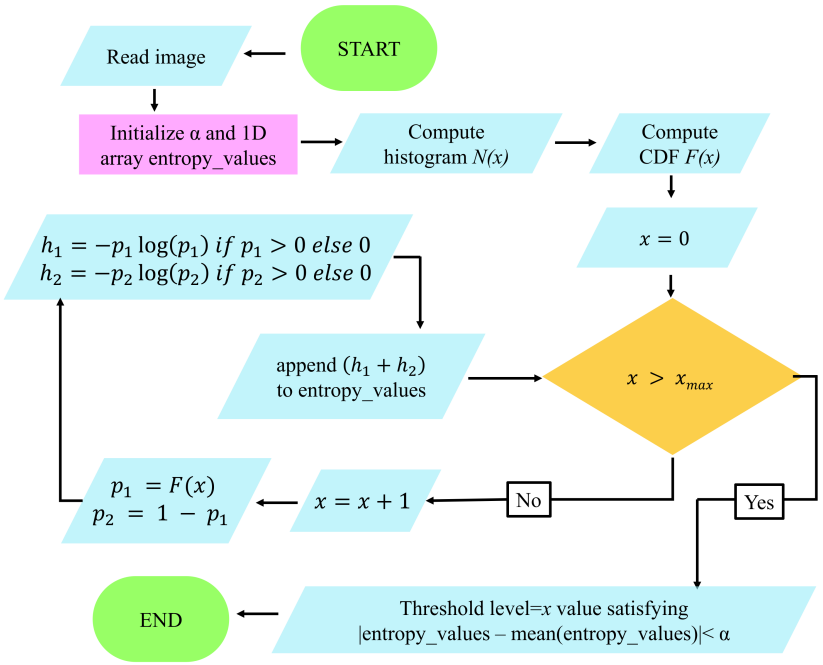

Figure 3.2. Flowchart showing the steps involved in the proposed adaptive entropy-based thresholding technique.

Unlike conventional entropy-based methods such as Kapur's algorithm [99], which may suffer from over-segmentation and include unnecessary background regions, our method selectively excludes such regions. Other entropy strategies, like cross-entropy thresholding [107] and fuzzy-entropic thresholding [108], tend to be computationally intensive. In contrast, the method we present balances accuracy and efficiency by using entropy values directly derived from the image histogram.

The implementation of the entropy based thresholding method in Python is presented in Appendix D, Code Snippet D.12. Once the threshold is determined, the image is binarized by retaining pixels whose intensity values correspond to entropy values above the mean. This step effectively suppresses low-information pixels likely to be noise while preserving candidate PSFs. The overall algorithm is summarised in the flowchart shown in Figure 3.2.

3.4 Comparative Performance Analysis

To assess the effectiveness of the proposed adaptive thresholding technique, we present a detailed comparative analysis against several commonly used thresholding methods. This includes both simulated and real datasets to evaluate signal-background separation and the preservation of structural details critical for accurate localisation.

3.4.1 Separation of Signal from Noise

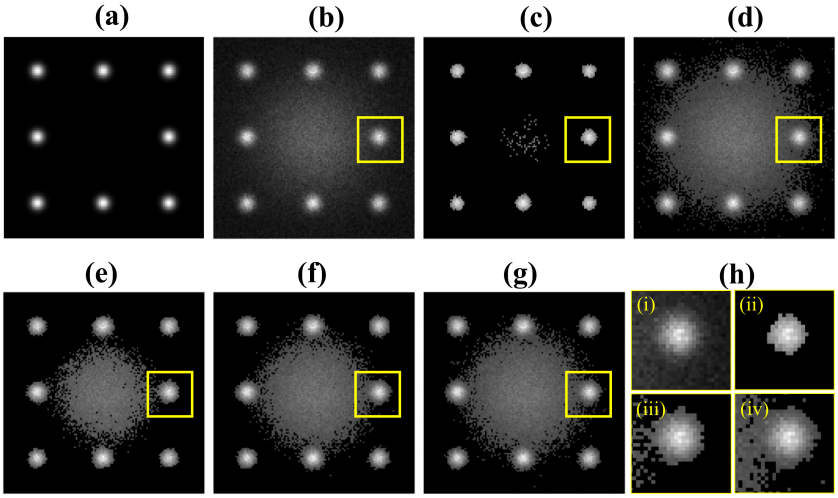

Figure 3.3. Comparison between existing thresholding methods and the proposed adaptive thresholding method on a simulated (128 × 128) image. (a) Ground truth with 8 Gaussian PSFs of intensities I_0 between 800–1000 photons and σ_x, σ_y between 3–3.5. (b) Image with added MPG noise. Thresholded images using (c) adaptive thresholding, (d) Kapur's method, (e) Otsu's method, (f) T-point method, and (g) mean thresholding. (h) Magnified views of yellow-boxed regions from (i) the noisy image and the thresholded images in (c), (e), and (f).

We created a synthetic grayscale image consisting of eight well-separated Gaussian PSFs using equation 1.8 given in Chapter 1. The original PSFs, with slightly varying intensities and FWHMs, are shown in Figure 3.3 (a). To emulate realistic imaging noise, we added mixed Poisson-Gaussian (MPG) noise [39–41], along with a low-frequency Gaussian disk background, which mimics uneven laser illumination [109]. The resulting noisy image, shown in Figure 3.3 (b), is modelled using equation 1.6 of Chapter 1, which is given as

$$I(x, y) = N_P\,(F(x, y) \otimes \mathrm{PSF} + B) + N_G \tag{3.5}$$

[as already stated in chapter 1, Np and Ng]

A more detailed explanation of noise models relevant to SMLM is included Appendix C.

We applied our adaptive thresholding method to the noisy image and compared it with Kapur's entropy method [99], Otsu's method [98], the T-point method [104], and mean-based thresholding. Figures 3.3 (c)–(g) show the thresholded images at threshold values, with the images being normalised within (0,255), as, 117 (adaptive), 57 (Kapur), 80.74 (Otsu), 71 (T-point), and 67.37 (mean). All pixels below the selected threshold were set to zero, with higher intensities preserved.

The yellow boxes highlight specific image regions consisting a single PSF for closer examination, with their magnified views displayed in Figure 3.3 (h). These zoomed-in

views reveal that the proposed method more reliably isolates PSFs from the background, particularly in noisy or overlapping regions.

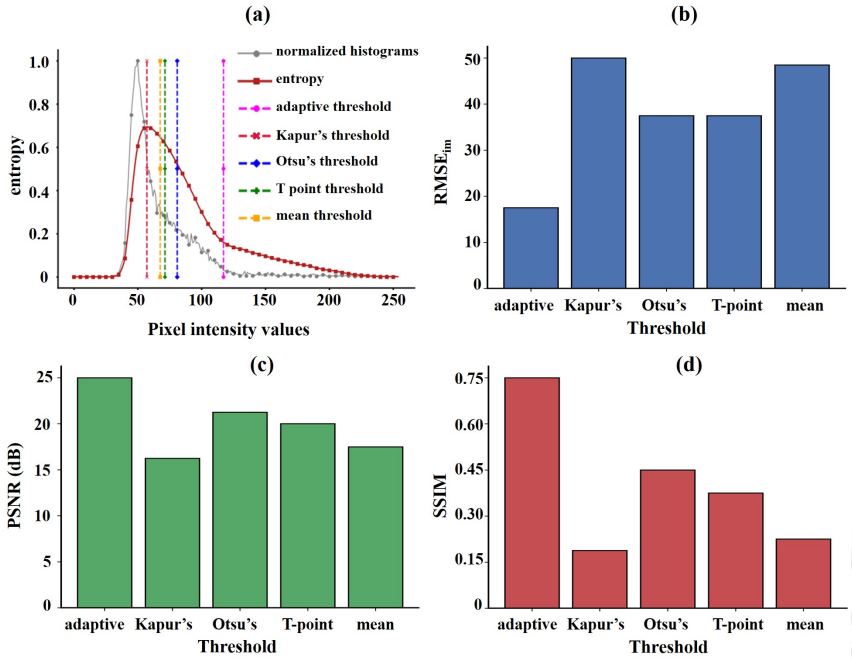

Figure 3.4. (a) Pixel intensity histogram and entropy plot showing thresholds obtained by various methods. Bar plots comparing (b) RMSE, (c) PSNR, and (d) SSIM between thresholded images and ground truth, averaged over 100 simulations.

Figure 3.4 (a) shows the pixel intensity histogram and associated entropy plot. To quantitatively assess performance, we calculated the image wise root mean squared error (RMSE_{im}), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) [110] between the thresholded and ground truth images. RMSE_{im} quantifies absolute pixel-wise deviation, which relates directly to signal loss or over-segmentation of PSFs. PSNR measures the ratio of true signal strength to noise-induced distortion, making it suitable for evaluating how well thresholding preserves photon-rich regions. SSIM captures structural fidelity, which is important for maintaining the morphology of PSFs and preventing artefacts that could bias localisation.

$$\text{RMSE}_{\text{im}} = \sqrt{\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I(i,j) - I_{th}(i,j))^2} \tag{3.6}$$

$$\text{PSNR} = 20\ \log_{10}\left(\frac{\max(I)}{\text{RMSE}_{\text{im}}}\right) \tag{3.7}$$

where I and I_{th} represent the ground truth and thresholded images, and (m, n) are their dimensions. These metrics were averaged over 1000 independent simulations. As shown in

Figures 3.4 (b)–(d), our method consistently achieved the lowest $RMSE_{im}$ and the highest PSNR and SSIM, demonstrating superior accuracy in signal preservation and noise rejection.

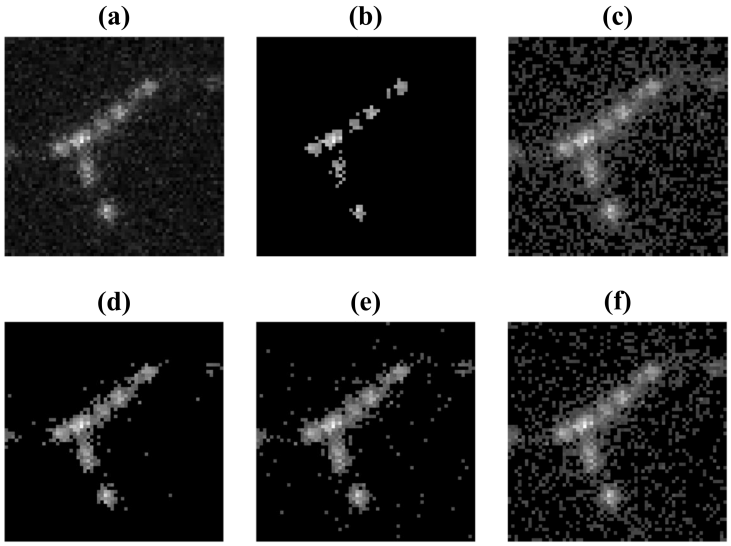

Figure 3.5. (a) Realistic 16-bit SMLM image frame (64 × 64) pixels. Thresholded results using (b) adaptive thresholding, (c) Kapur's method, (d) Otsu's method, (e) T-point method, and (f) mean thresholding.

To further assess practical performance, we tested the methods on a real 16-bit SMLM image from an open-source dataset [57]. As shown in Figure 3.5, the adaptive thresholding method isolates PSFs more effectively than other techniques. Among the traditional methods, Otsu's technique performs reasonably well but still falls short compared to the adaptive approach. Based on this evaluation, we focus on comparing our method against Otsu's in the subsequent sections.

3.4.2 Crude Localisation of PSFs in Thresholded Images

We next evaluated the effectiveness of three commonly used crude localisation techniques, peak local maximum, centre of mass, and blob detection, as described in Chapter 2, section 2.4, being applied after thresholding via either Otsu's method or the proposed adaptive method. To do this, we again created a dataset of 1000 synthetic SMLM images of dimensions (32×32) pixels by placing five Gaussian PSFs at random locations within the image. The images were then corrupted with noise as defined in equation (3.5) as $I_{\text{noisy}} = N_P(GT + 50) + 30\,N_G$. This noise level intentionally exceeds typical experimental conditions and is used to stress-test each localisation pipeline. The Poisson component reflects the low-photon regime characteristic of SMLM, while the added Gaussian term simulates amplified readout noise or background fluorescence.

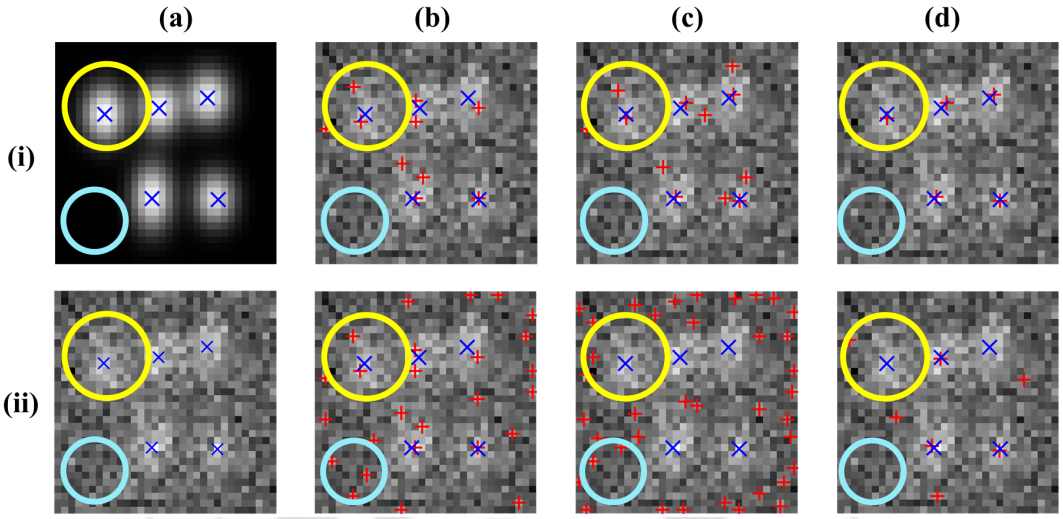

Figure 3.6. PSF detection accuracy. Column (a): ground truth (row i) and noisy image (row ii). Crude localisation using (b) peak local maximum, (c) centre of mass, and (d) blob detection after thresholding by adaptive method (row i) and Otsu's method (row ii).

In figure 3.6 we show a specific image from the above mentioned dataset being thresholded and processed for PSF detection. Column (a), displays the ground truth and corresponding noisy image of a specific image in rows (i) and (ii), respectively. In these images, red '+' symbols represent the PSF locations identified by the respective crude localisation technique, while blue '×' marks indicate the actual ground truth coordinates. Figures 3.6 (i)(b)–(d) show results after adaptive thresholding, and Figures 3.6 (ii)(b)–(d) correspond to those obtained following Otsu's thresholding. Yellow circles highlight areas with true PSFs, while blue circles denote regions containing only noise.

It is evident from the visual comparison that all localisation methods exhibit degraded performance when used after Otsu's thresholding. Among them, blob detection shows superior localisation accuracy, particularly when preceded by adaptive thresholding, with minimal false detections. For instance, in the yellow-highlighted region of Figure 3.6 (i)(d), blob detection successfully identifies a true PSF, while avoiding false positives in the noise-only area marked by the blue circle.

To quantitatively assess the localisation accuracy, we computed the Jaccard Index (JI) which measures how well the algorithm finds the correct number of emitters and lateral RMSE (which measures how accurately the detected positions match the ground truth), as introduced in Chapter 1, Section 1.8.2. As shown in Figure 3.7, the combination of blob detection with adaptive thresholding yields the highest JI and the lowest RMSE, indicating improved accuracy and robustness.

We further analysed a publicly available dataset consisting of 20,000 SMLM image frames representing microtubule structures, obtained from the references [57, 111]. Each

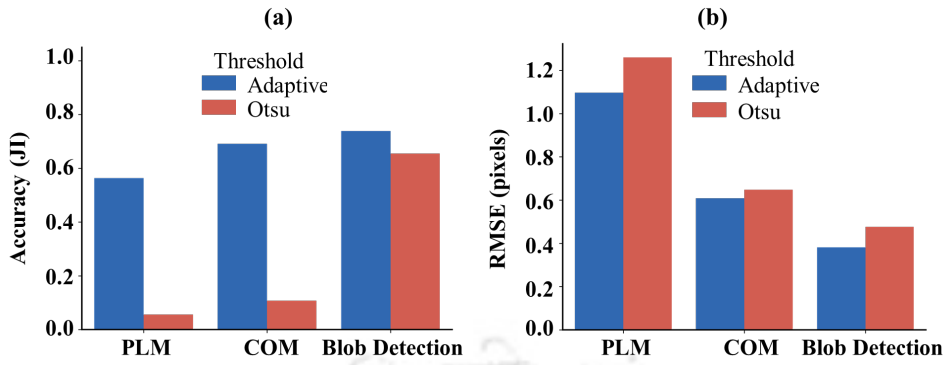

Figure 3.7. Bar plots of Jaccard Index and $\text{RMSE}_{x,y}$ for different localisation methods following adaptive or Otsu's thresholding.

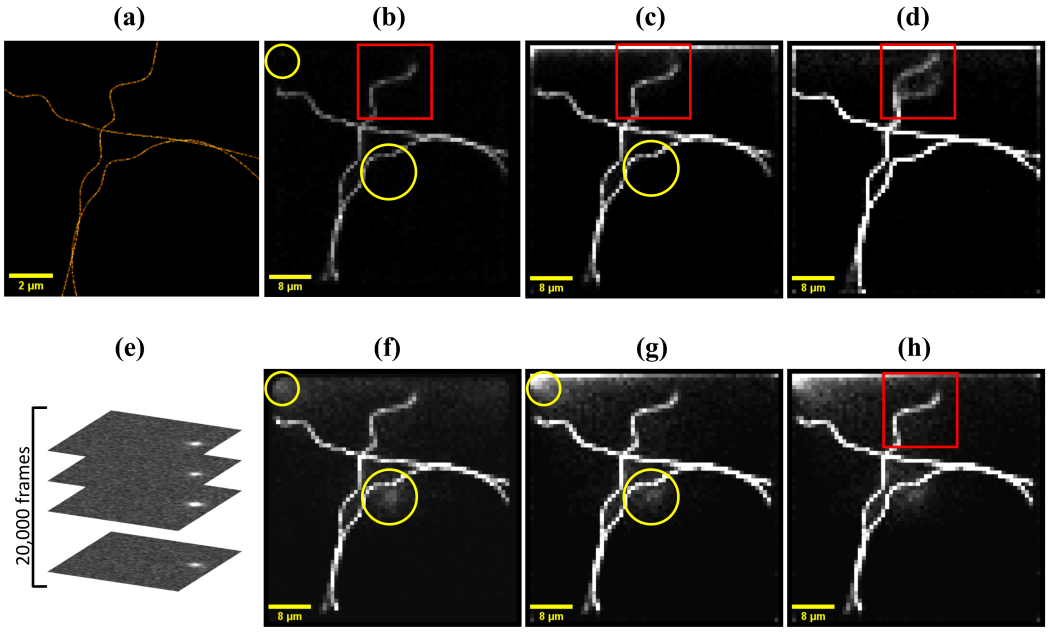

Figure 3.8. (a) Ground truth image of microtubules used to simulate a 20,000-frame SMLM dataset. Reconstructed images using (b)–(d) adaptive thresholding and (f)–(h) Otsu's thresholding, followed by (b, f) blob detection, (c, g) center of mass, and (d, h) peak local maximum.

frame was thresholded using both the proposed adaptive thresholding technique and Otsu's method. Crude localisation was then performed using blob detection, centre of mass, and peak local maximum. For each frame, pixels corresponding to detected PSF centres were activated to create binary masks. These were summed and normalised to generate the final reconstructed image.

Figures 3.8 (b)–(d) display reconstructions using blob detection, centre of mass, and peak local maximum after adaptive thresholding, while Figures 3.8 (f)–(h) show results following Otsu's thresholding. The reconstructions indicate that blob detection and centre of mass outperform peak local maximum, particularly in regions marked with red boxes.

This is because both methods take into account spatial connectivity of pixels, mitigating the problem of spurious detections from isolated high-intensity pixels, a limitation in peak local maximum.

Moreover, reconstructions based on Otsu's thresholding contain visible clusters of false detections, especially in the regions marked with yellow circles. Such artifacts are notably absent when using the adaptive thresholding approach. This further supports the superiority of our method for enhancing PSF detection and improving the quality of super-resolved reconstructions in noisy environments.

3.4.3 Reconstruction of Super-Resolved Image with Adaptive Thresholding

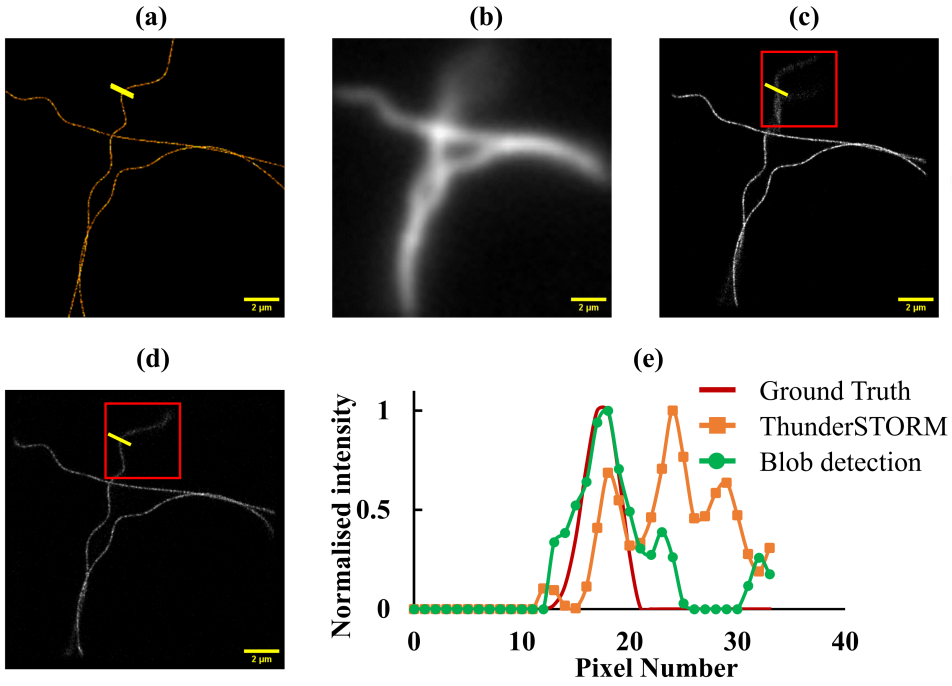

Figure 3.9. (a) Ground truth image of simulated microtubule structures. (b) Widefield equivalent image of the corresponding image stack. Super-resolved images reconstructed using (c) ThunderSTORM and (d) blob detection following adaptive thresholding. (e) Intensity plots along the yellow lines drawn in (a), (c), and (d). Images in (a), (c), and (d) are (256×256) pixels, while (b) is (64×64) pixels.

To further evaluate the efficacy of our method, we compared the quality of super-resolved images reconstructed using blob detection after adaptive thresholding with those generated by the ThunderSTORM software. ThunderSTORM was configured with its default settings, such as, a third order B-spline wavelet filter for preprocessing, local maximum for crude localisation, and weighted least-squares fitting using an integrated Gaussian PSF model for sub-pixel localisation.

Figure 3.9 (a) shows the ground truth image used in the simulation, and Figure 3.9 (b) presents the corresponding conventional widefield image. The super-resolved reconstructions obtained from ThunderSTORM and our proposed approach are shown in Figures 3.9 (c) and (d), respectively.

We observe that the blob detection method yields a reconstruction that is visually closer to the ground truth. In contrast, ThunderSTORM exhibits inaccuracies, particularly in noisy regions (highlighted by red boxes), where it introduces artifacts. To assess reconstruction accuracy quantitatively, we examined intensity profiles along a chosen region of interest, indicated by the yellow lines. The corresponding intensity plots are shown in Figure 3.9 (e). The plot from the adaptive thresholding and blob detection method closely follows the ground truth, whereas the ThunderSTORM result shows a false peak, indicating a misidentified structure.

3.4.4 Timing Performance

Crude Localisation Method	Thresholding Method	Time Taken (s)
Peak Local Maximum	Proposed	9.91
Peak Local Maximum	Otsu's	9.70
Centre of Mass	Proposed	82.45
Centre of Mass	Otsu's	81.43
Blob Detection	Proposed	105.53
Blob Detection	Otsu's	104.95

Table 3.1. Processing times for different combinations of thresholding and crude localisation methods applied to the SMLM dataset of 20000 image frames.

We also compared the execution times of the proposed adaptive thresholding method with those of Otsu's method, across all three crude localisation techniques on a set of 20,000 images. The computations were carried out on a standard desktop computer (Intel(R) Core(TM) i5-10400 CPU with 16 GB DDR4 RAM at 2667 MHz). Table 3.1 summaries the time taken by each method combination. The data reveals that the proposed method takes approximately same time as Otsu's method across all three localisation pipelines. This consistent performance advantage, while modest, demonstrates the computational efficiency of the entropy-based approach.

Notably, the results also highlight that the choice of the crude localisation technique has a significant impact on the overall processing time. The pipeline combining the proposed threshold with Peak Local Maximum was the most time-efficient combination, whereas blob detection was the most computationally intensive method overall. It is noteworthy

that while Blob Detection is the most accurate of the three localisation techniques, the high precision offered by it comes at a cost of enhanced computational time.

3.5 Chapter Summary

In this chapter, we introduced a dynamic and adaptive thresholding technique for distinguishing between signal and background noise in single-molecule localisation microscopy images. The method determines the optimal threshold for each SMLM image frame by identifying the intersection between the entropy curve and its mean value. This adaptive thresholding strategy has been applied to both simulated PSF images with added noise and realistic SMLM datasets. Across all tests, the method consistently outperformed several established thresholding techniques in accurately separating true fluorophore emissions from background noise. We then applied crude localisation methods to these thresholded images and found that blob detection, when combined with the proposed thresholding method, delivered the most accurate PSF localisation. Finally, we demonstrated that this combination enabled the reconstruction of super-resolved images that were visually and quantitatively closer to the ground truth compared to results produced by the widely used ThunderSTORM software. The method offers a significant improvement in image quality, and further optimisation could make it suitable for real-time or high-throughput applications.

CHAPTER 4
A CNN based method for dense PSF detection

4.1 Introduction

Single molecule localisation microscopy (SMLM) while highly effective for sparse activation conditions, encounters difficulties when fluorophores are activated simultaneously, resulting in overlapping point spread functions (PSFs) that compromise localisation accuracy.

In this chapter, we propose a convolutional neural network (CNN)-based method for accurately localising clustered PSFs in high-density SMLM images. The network is trained to predict fluorophore positions by generating Gaussian-like blobs centred at likely PSF locations. We further incorporate a fast blob detection routine as a post-processing step to achieve sub-pixel precision. Our approach offers improved localisation performance, particularly in challenging scenarios where traditional algorithms are limited, while maintaining computational simplicity and ease of implementation.

4.2 Physical and Computational Challenges in High-Density PSF Localisation

Localisation accuracy in SMLM is fundamentally governed by the physics of image formation. Each detected emission from a molecule can be expressed as the convolution of an idealised point emitter with the optical system's PSF. When two or more emitters lie within the width of the central lobe of the Airy pattern, their PSFs merge to form a composite intensity distribution. In such cases the information available for distinguishing individual emitters diminishes sharply, and traditional single-emitter fitting becomes biased. This behaviour is well captured by the Fisher information framework, which shows that the Cramér–Rao lower bound (CRLB) for localisation increases steeply as emitter separation decreases, reflecting the fundamental photon-limited resolvability of closely spaced fluorophores [34, 112].

Photon statistics further constrain localisation. As discussed in the previous chapters, fluorescence emission is a stochastic process governed by Poisson counting statistics, so the variance of the recorded photon number scales with the square root of the signal itself [33]. At typical photon budgets in SMLM, this shot noise often dominates, while detector readout noise contributes an additional approximately Gaussian component. Moreover, out-of-focus fluorescence, autofluorescence, and background scattering introduce spatially varying backgrounds that distort the apparent PSF shape. These effects collectively reduce the separability of overlapping PSFs and make the inverse problem of identifying individual emitters highly ill-posed, particularly at low signal-to-noise ratios or elevated fluorophore densities [113].

Classical multi-emitter algorithms attempt to address these limitations through iterative PSF decomposition or sparse recovery. Methods such as DAOSTORM [64] and 3D-DAOSTORM [65] iteratively subtract fitted PSFs from the residual image, while sparsity-based techniques like CSSTORM [66] and FALCON [68] formulate localisation as a convex optimisation problem that favours sparse emitter distributions. Although these methods perform well in moderate-density conditions, their accuracy degrades noticeably when emitter overlap becomes severe or when background statistics diverge from model assumptions. In addition, they are computationally expensive because the complexity of the fitting or optimisation procedure grows rapidly with emitter density.

Deep learning provides an alternative platform for overcoming these limitations by learning the highly nonlinear mapping from raw diffraction-limited images to emitter locations directly from data. Convolutional neural networks are especially well suited for this task because they can capture multi-scale spatial correlations and are robust to the diverse noise sources present in fluorescence microscopy. This has led to several successful frameworks such as ANNA-PALM [114], Deep-STORM [115], DeepSTORM3D [53], DECODE [54], and DeepCEL0 [116], all of which demonstrate substantial improvements over traditional fitting approaches in dense regimes. However, many of these networks involve deep architectures with millions of parameters, making them computationally demanding to train and deploy.

Here we introduce a lightweight CNN architecture [117] inspired by Deep-STORM and residual network principles. The model is trained on simulated SMLM images that incorporate realistic optical effects such as diffraction, pixelation, and Poisson-Gaussian noise, allowing it to learn emitter-specific spatial features directly from the physics of the imaging process. To maintain physical interpretability and obtain precise emitter coordinates, the CNN output is followed by a blob-detection routine that provides sub-pixel localisation without relying on heavy iterative optimisation. This hybrid strategy combines

the strengths of data-driven learning with the stability and transparency of model-based localisation, offering an efficient and effective solution for high-density SMLM.

4.3 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models inspired by the structure and function of biological neural networks. They are designed to learn complex relationships between input data and corresponding outputs through training. An ANN consists of layers of interconnected nodes or neurons, organised into an input layer, one or more hidden layers, and an output layer. Each neuron receives inputs from the previous layer, processes them using a set of learnable parameters (weights and biases), and passes the result through a non-linear activation function. These parameters are updated during training to minimise the error between the predicted and actual outputs. The introduction of non-linear activation functions allows neural networks to model complex, non-linear relationships. Common activation functions include the linear function, sigmoid (logistic), hyperbolic tangent (tanh), and rectified linear unit (ReLU) [118], along with their various extensions and hybrids [119, 120].

Considering f as the activation function, if w_{ij} denotes the weight connecting the j^{th} neuron in the previous layer to the i^{th} neuron in the current layer, and b is the bias term, then the output y_i of the i^{th} neuron is given by [121]

$$y_i = f\left(\sum_{j=1}^{n} w_{ij}x_j + b\right) \tag{4.1}$$

Here, x_j is the output of the j^{th} neuron in the previous layer, and n is the total number of neurons in that layer.

Training a neural network involves several component or parameters, such as, a loss function, an optimizer, and a learning rate. The loss function quantifies the difference between the predicted and actual values, guiding the network during training. Optimisers adjust the weights and biases to minimize the loss, while the learning rate controls the size of these adjustments in each training step. Each full pass through the training dataset is called an epoch. During training, callbacks can be employed to monitor progress and perform specific actions, such as saving model checkpoints or adjusting the learning rate dynamically. Together, these components enable the network to learn from data and generalize to unseen inputs.

4.3.1 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a specialised type of neural network designed for processing data with a grid-like structure, such as images. CNNs are particularly effective in image processing tasks due to their ability to extract spatial hierarchies of features.

At the heart of a CNN are convolutional layers, which apply a set of learnable filters (or kernels) to the input image. Each filter scans the image spatially, computing the weighted sum of the pixel intensities within its receptive field. This process generates a feature map that captures specific patterns such as edges, textures, or more complex structures. Mathematically, a 2D convolution operation is expressed as,

$$g(x, y) = k \otimes f(x, y) = \sum_{p=-a}^{a} \sum_{q=-b}^{b} k(p, q) \cdot f(x - p, y - q) \tag{4.2}$$

where $f(x, y)$ is the input image, $k(p, q)$ is the convolution kernel, and $g(x, y)$ is the resulting feature map. The parameters $a = \text{floor}(M/2)$ and $b = \text{floor}(N/2)$ correspond to half the kernel dimensions for a kernel of size $M \times N$.

In addition to convolutional layers, CNNs commonly use pooling layers to reduce the spatial dimensions of feature maps while preserving the most relevant information. Max pooling and average pooling are two standard methods, where the maximum or average value within a small window is selected. Pooling reduces the number of parameters, lowers computational cost, and mitigates overfitting.

For tasks requiring output with high spatial resolution, such as super-resolution microscopy where emitter coordinates must be preserved with sub-pixel accuracy or semantic segmentation, upsampling layers are employed. These layers increase the size of the feature maps through methods like nearest-neighbour interpolation or transposed convolutions, allowing the network to recover fine details in the output image.

4.4 The proposed CNN model for localisation of clustered PSFs

Our model is motivated by the strength of existing deep learning architectures in extracting spatial features and resolving fine-scale structures in complex image data. In particular, we take inspiration from the Deep-STORM architecture [115], which has shown impressive performance in the reconstruction of super-resolved images, and residual networks (ResNets) [122], known for their efficient training and ability to generalise well to complex tasks. By integrating elements from both, we design a hybrid network that balances accuracy, computational efficiency, and stability and ease of training. The following subsections outline

the structure, loss function, training methodology, and post-processing steps involved in the proposed approach.

4.4.1 Network architecture

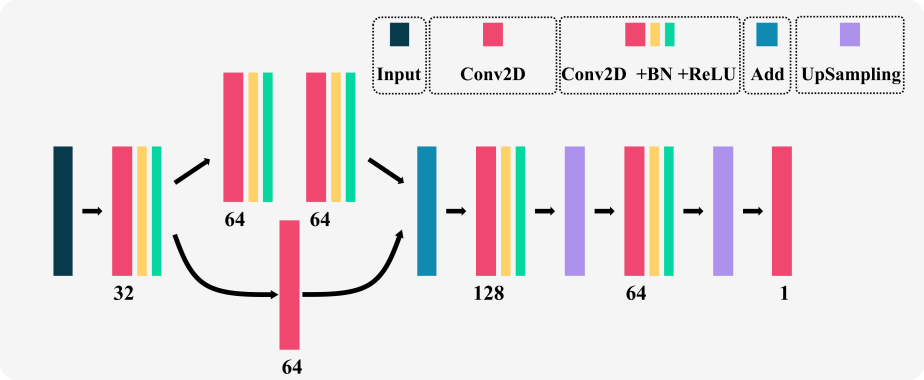

Figure 4.1. The architecture of the proposed CNN model for PSF localisation. Conv2D, BN and ReLU represent 2D convolution layer, batch normalisation layer and rectified linear unit activation function, respectively. Add and UpSampling2D represent the addition operation of layers and upsampling of layers, respectively. Each convolutional layer is specified with a number which indicates the number of associated filters.

Deep-STORM is a CNN based tool that has proven itself as highly effective for reconstructing super-resolved images in STORM setups. Consequently it has been gaining widespread usage and adaptation within the community. Meanwhile, another CNN-based scheme called residual networks [122, 123] have also gained significant interest, particularly in image enhancement in fluorescence imaging [124–127]. Besides, residual networks have showcased remarkable efficiency in optimising training processes and reducing the risk of overfitting, which is especially crucial in complex tasks like image enhancement and segmentation. Excited by the potential of both Deep-STORM and residual networks, we aim to harness their strengths and merge them into a novel hybrid model. Our proposed CNN architecture is meticulously designed to effectively capture and extract features from SMLM images, enabling accurate localisation of PSFs. The architecture comprises multiple layers, each serving a specific purpose in the feature extraction process. To achieve the same, we utilise Keras [128], a high-level neural network API, which is built on top of TensorFlow [129], an open-source machine learning framework developed by Google. The code for the project was written in Python.

Figure 4.1 illustrates the proposed CNN architecture, with each block representing a layer. At the core of the architecture are convolutional blocks which play a crucial role in learning hierarchical representations of the input images. These blocks consist of a (3×3) 2D convolutional layer followed by batch normalisation [130] and ReLU activation functions,

defined as $\mathrm{ReLU}(x) = max(0, x)$, which introduce nonlinearity by suppressing negative activations. Convolutional layers apply learnable filters to the input images, extracting various features such as edges, textures, and patterns enabling the network to identify local spatial correlations associated with overlapping PSFs and background noise. Batch normalisation ensures stable and efficient training by normalising the activations of each layer. In figure 4.1, each convolutional layer is specified with a number which indicates the number of associated filters. To facilitate the learning of deeper representations and improve gradient flow during training, a residual network block is incorporated into the architecture. The residual block comprises two convolutional blocks with a shortcut connection. The shortcut connection enables the network to learn residual mappings, capturing the difference between the input and output of each block. The hierarchical feature extraction process, coupled with residual connections and spatial dimension manipulation, empowers the network to learn intricate patterns and relationships within the input images, resulting in improved localisation performance. We also implement two (2×2) upsampling layers, yielding a fourfold increase in spatial sampling of the output image relative to the input. The final prediction layer is a (1×1) 2D convolutional layer with ReLU activation. Hence, the output of the network is a single-channel grayscale image.

Compared to the other CNN-based methods for localisation microscopy, our CNN architecture employs a simpler design with fewer layers. This translates to a smaller number of trainable parameters, leading to reduced training time while maintaining competitive performance. For instance, our network with five convolutional blocks when trained on an input of size (32,32,1) has 206,624 total parameters, whereas Deep-STORM and DeepCEL0 with six convolutional blocks have 389,392 parameters, which is a difference of 88.44%.

4.4.2 The loss function

A critical aspect of the accuracy of the designed CNN lies in the choice of an appropriate loss function. For regression tasks, normally, the MSE (mean square error) loss function [131, 132] is used. However, in our proposed scheme we employ a modified form of the MSE. We convolve the predicted image of the network with a 2D Gaussian to enhance its features. This modification reflects the physical extent of the PSF and reduces sensitivity to small spatial misalignments between predicted and true emitter locations. The convolved image is then used to calculate the loss against the true output image. The loss function in our case is thus given as,

$$\mathrm{MSE}(\mathrm{heatmap\_true}, \mathrm{heatmap\_pred}) = \frac{1}{N}\sum_{i=1}^{N}(\mathrm{heatmap\_true}_i - \mathrm{heatmap\_pred}_i)^2 \quad (4.3)$$

Here N is the total number of elements in the output image, heatmap_pred is the output of convolution of the CNN predicted image with the 2D Gaussian, and heatmap_true is the ground truth. In this work, the term heatmap refers to the two-dimensional output of the CNN which in this case represents the spatial representation of the emitter density. The above loss function quantifies the discrepancy between the predicted and true PSF density distributions and encourages agreement at the level of emitter density rather than exact pixel-wise correspondence, which is consistent with the stochastic nature of photon detection, thus guiding the model to learn accurate PSF localisation in the input images.

4.4.3 Training data generation and preprocessing

The proposed CNN model is trained on simulated data representing blinking SMLM images. The training data is generated using ThunderSTORM which implements a forward-model for fluorophore emission and detection, incorporating photon shot noise, background fluorescence, quantum efficiency, and pixelation. We use Gaussian PSFs of full width at half maximum (FWHM) ranging from 200 nm to 600 nm and intensity ranging from 1000-2000 photons. This range accounts for the fluctuations in detected photons, emission wavelength, and field dependent defocus aberration commonly encountered in experimental SMLM data. The density of fluorophores is kept at $8\mu m^{-2}$. We create 10,000 image frames, each of size (32×32) pixels, comprising 40-60 active fluorophores at random locations. These images (i.e. input images) are used as the independent variables in training. We also obtain a CSV file having the locations of each fluorophore in each frame, which is used to create the dependent variables (i.e. output images) for training.

Before training the CNN model, we preprocess the input images to ensure consistency and compatibility. This preprocessing involves normalising the pixel values of the images to the range [0, 1]. Furthermore, we calculate the mean and standard deviation of each image to monitor data distribution during training. These are then used to scale each image. Finally, the input and output images are reshaped to match the input dimensions expected by the CNN model. For training and validation the processed data is split in the ratio of 80:20.

4.4.4 Training of the CNN model

The model is trained for 30 epochs with a batch size of 16 using stochastic gradient descent-based Adam optimiser [133], with an initial learning rate of 0.001. During the training of our CNN model, we employ two essential callbacks of the above-stated neural network API, Keras, to enhance model performance and convergence. Firstly, the ModelCheckpoint callback saves the model weights after each epoch if the validation loss decreases, ensuring the retention of the best-performing model. This prevents overfitting and facilitates the

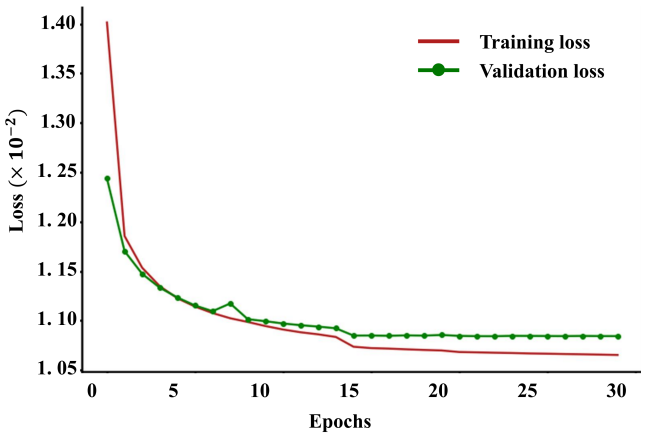

Figure 4.2. Plots of the training loss and validation loss for each epoch during training.

retrieval of optimal model weights for further evaluation or deployment. Secondly, we utilise the ReduceLROnPlateau callback to dynamically adjust the learning rate of the optimiser when the validation loss reaches a plateau. This adaptive learning rate adjustment strategy helps prevent the model from getting stuck in local minima and facilitates more efficient training. Figure 4.2 compares the training loss and validation loss for each epoch of training. To be noted here that in our implementation the total training time for 30 epochs on a PC

Hyperparameter	Value	Description
Number of Epochs	30	Total number of training epochs.
Batch Size	16	Number of samples processed before the model is updated.
Validation Split	0.2	Proportion of data used for validation during training.
Initial Learning Rate	0.001	Starting learning rate for the optimizer.
Learning Rate Scheduler	Reduce LR by 0.1 on plateau, minimum LR of 0.00005	Reduces the learning rate if the validation loss does not improve after 5 epochs.
L2 Weight	100	Regularization parameter to prevent overfitting by penalising large weights.

Table 4.1. Summary of hyperparameters used for model training

with an Intel(R) Core(TM) i5-10400 CPU with 16 GB of 2667 Mhz DDR4 RAM was 152.6 minutes. The particulars regarding various hyperparameters are summarised in Table 4.1.

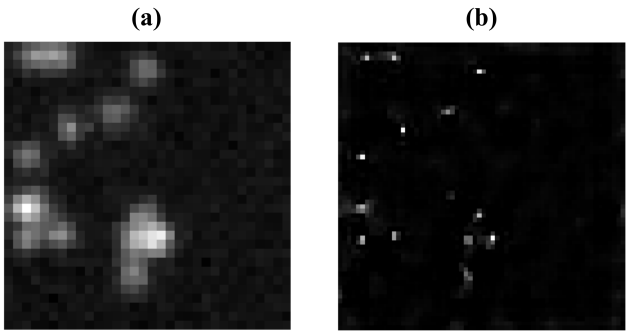

Figure 4.3. (a) A diffraction limited image comprising dense fluorophores and (b) the corresponding CNN output. (a) and (b) are of size (32×32) pixels and (128×128) pixels, respectively.

4.4.5 Incorporation of sub-pixel localisation

Once the model is trained, it can be used to predict the localisation of PSFs in unseen SMLM images. Figure 4.3 shows a representative scenario comprising a diffraction limited image of dense fluorophores as the input and the corresponding output of our CNN model. The output image has spikes at probable centres of the emission PSFs. As we can see, the network can segregate overlapped PSFs present in the diffraction limited image. However

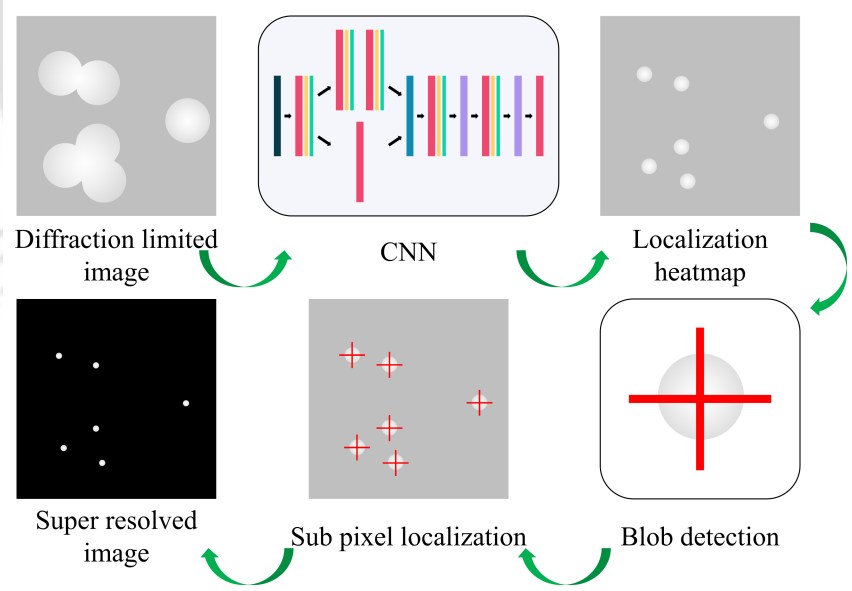

Figure 4.4. Schematic illustrating the steps involved towards sub-pixel localisation of densely packed PSFs in SMLM images.

the resolution of this image is limited by the pixel dimensions of the CNN output. To tackle this issue, we perform a sub-pixel localisation [134, 135] to find the centre of the PSF beyond the existing pixel resolution of the output image using a known fitting based routine. In our

case, though, we have used a sub-pixel localisation method called blob detection [136]. Blob detection is faster than the iterative fitting methods and much more accurate than methods that determine the centres based on peak pixel intensity, as shown in a previous work [93], and provides sufficiently accurate centre estimation for the heatmap peaks produced by the CNN. Figure 4.4 depicts the computational steps to be performed on an SMLM image for sub-pixel localisation of the overlapped PSFs.

The above scheme of identification of overlapped PSFs using our CNN model followed by sub-pixel localisation using blob detection will be henceforth referred to as proposed CNN assisted PSF localisation (CAPL) scheme.

4.5 Evaluation of the CAPL scheme

In this section, we will evaluate the performance of the CAPL scheme on both simulated and experimental SMLM images. Evaluation is carried out in terms of standard metrics, which include recall, precision, Jaccard index (JI), and $RMSE_{x,y}$, whose definitions are already provided in Chapter 1. Additionally, for experimental images where the ground truth is unknown, the quality of the output image can be evaluated using Fourier Ring Correlation (FRC) method [59, 60].

4.5.1 PSF Detection Accuracy

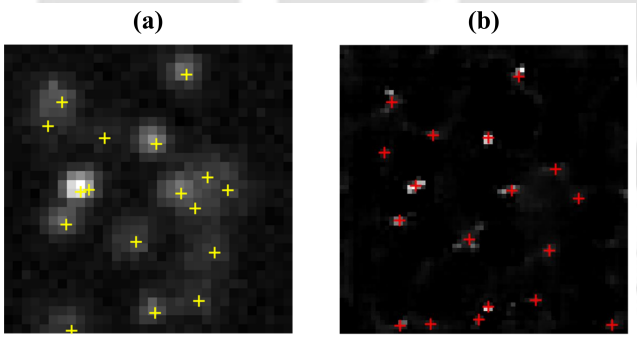

Figure 4.5. Evaluating the performance of the CAPL scheme for the detection of overlapped PSFs in simulated SMLM images. (a) shows the diffraction-limited images with the ground truths overlaid in yellow '+' marks, and (b) shows the detections of our CNN network overlaid on the CNN output image.

First, we investigate the performance of our CAPL scheme in accurately identifying overlapped PSFs. For this, We generate a stack of 50 SMLM images of dimensions (32 × 32) pixels in ThunderSTORM with a density of around two emitters per μm^2. The intensities of these PSFs range between 1000 to 2000 photons, with full width at half maximum (FWHM) between 200 nm and 600 nm. The various other simulation parameters considered are, pixel

size [nm]: 100nm, photoelectrons per A/D count: 3.6, quantum efficiency: 1.0, base level [A/D counts]: 30.0, and readout noise [electrons per pixel]: 0. This arrangement ensures a significant overlap of PSFs while also preserving some true background to enable an unbiased evaluation. Subsequently, these images are processed using CAPL scheme, and the outputs are compared with their ground truths.

Metric	recall	precision	JI	$RMSE_{x,y}$
Value	0.7799	0.9264	0.7345	0.5481

Table 4.2. Table showing recall, precision, JI and $RMSE_{x,y}$ calculated on 50 simulated SMLM images.

We depict a representative case among the 50 images in Figure 4.5. The true locations of the PSF are shown as yellow '+' while the detections using CAPL scheme are shown as red '+'. As seen in the figure, CAPL scheme can separate out overlapping PSFs in most cases, which are otherwise either left out or incorrectly localised by conventional reconstruction algorithms. It should be noted that we are showing the detections overlaid on top of the output of the CNN only and no background removal is done. The super-resolved image is reconstructed using the detected centre coordinates. The recall, precision, JI, and $RMSE_{x,y}$ parameters calculated for this dataset are presented in Table 4.2. It is noted that for the evaluation of these parameters, we consider a tolerance of 0.5 pixel for a detection to be true positive.

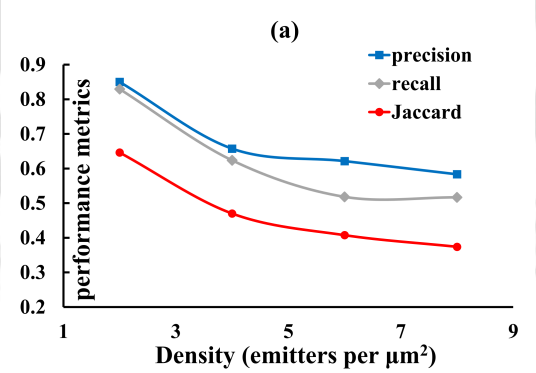

Figure 4.6. Plot of various performance indices when the density of fluorophore emitters is increased from two emitters per μm^2 to eight emitters per μm^2.

From the parameters, it is evident that our scheme performs well in the dense PSF dataset. The algorithm demonstrates high precision and recall, with minimal false positives and negatives. It achieves a good overlap between predicted and ground truth regions, as indicated by the $RMSE_{x,y}$ value. It should be emphasised that in the dense imaging regime considered here, strong overlap between neighbouring PSFs leads to significant coupling

between localisation parameters, rendering the classical single-emitter Cramér–Rao lower bound inapplicable. In such cases, the effective information content of the measurement is substantially reduced, and localisation errors on the order of tens of nanometres are expected even for optimal estimators. The observed $RMSE_{x,y}$ value is therefore consistent with the fundamental limitations imposed by overlapping diffraction-limited PSFs.

We further evaluate the performance of our CAPL scheme on four datasets, each with 50 images created using ThunderSTORM in the previous settings, but the density of fluorophore emitters is increased from two emitters per μm^2 to eight emitters per μm^2. To be noted here that emitter number 2.5 per μm^2 and above can be considered as higher emitter density [57]. Recall, precision and JI are evaluated and plotted as shown in Figure 4.6. We see that as the density increases, the performance indices show a gradual dip however the overall performance remains appreciable.

4.6 Implementation of the proposed CAPL scheme on standard datasets

To evaluate the effectiveness and robustness of the proposed CAPL scheme, we benchmark its performance on some standard simulated and realistic SMLM datasets with known ground truth. These datasets allow a systematic comparison against established localisation algorithms under controlled variations in emitter density and noise conditions. Performance is assessed using both qualitative reconstruction fidelity and quantitative localisation metrics commonly adopted in the SMLM literature.

4.6.1 Super-resolved Image Reconstruction in Low-Density Fluorophore Emitters in simulated and realistic SMLM images

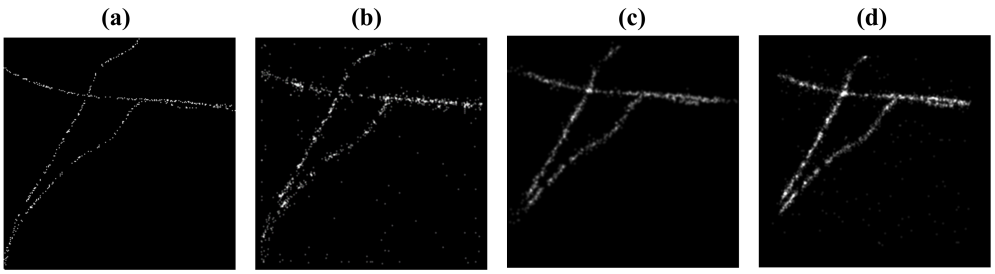

Figure 4.7. Comparison between ground truth image (a) and reconstructed images using (b) proposed CAPL scheme, (c) 3D-DAOSTORM and (d) ThunderSTORM.

We first investigate the performance of the proposed CAPL scheme to identify PSFs in images comprising low-density fluorophores by considering a simulated SMLM dataset

Method	recall	precision	JI	$RMSE_{x,y}$	SSIM
CAPL scheme	0.7657	0.6478	0.5406	0.9469	0.7753
ThunderSTORM	0.5245	0.9094	0.4985	1.8024	0.7942
3D-DAOSTORM	0.5106	0.6656	0.4064	0.5786	0.8293

Table 4.3. Table showing various indexes of assessment of SMLM localisation corresponding to Figure 4.7.

obtained from Super-Resolution Microscopy Hub [111]. We extract a patch of size (32 × 32) pixels from the first 2000 frames of the stack and process it using the CAPL scheme as well as 3D-DAOSTORM and ThunderSTORM. Figure 4.7 (a) shows the ground truth image, while (b), (c) and (d) show the reconstructions done using the three methods mentioned above. We see that the reconstruction done by all the methods are comparable visually. To compare them quantitatively, we estimate the precision, recall, JI and $RMSE_{x,y}$. We also estimate the structural similarity index (SSIM) [110] of the reconstructed images with the ground truth to assess how close are the super-resolved images reconstructed using the above-mentioned methods to the ground truth. The results are presented in Table 4.3. We observe that CAPL has better recall and JI scores as well as appreciable SSIM compared to the other methods. ThunderSTORM, on the other hand, which has a better precision has a much larger $RMSE_{x,y}$.

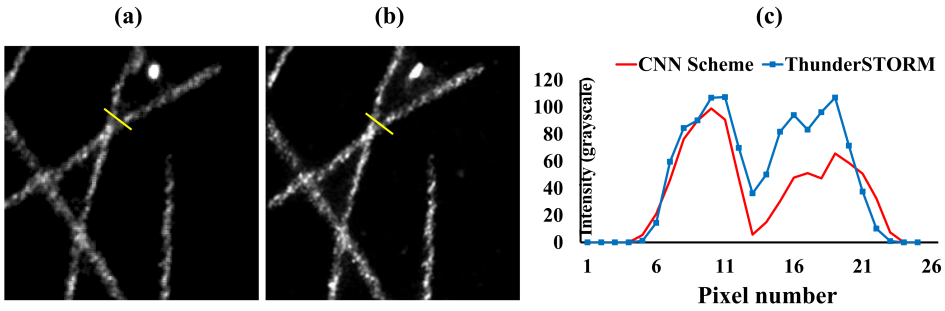

Figure 4.8. Comparison between reconstructed images using (a) CAPL scheme and (b) ThunderSTORM. (c) The line plots of pixel intensity values in the selected region of interest shown as yellow lines in (a) and (b).

We then consider a realistic SMLM dataset obtained from [87], which has 15000 image frames of size (32 × 32) pixels. The stack is then processed using both CAPL scheme and ThunderSTORM, and the results are shown in Figure 4.8. We also evaluate the FRC FIRE (Fourier Image REsolution) [60] number in both cases, which gives a resolution (fixed 1/7 criterion) of 5.345 and 4.958, for the CAPL scheme and ThunderSTORM, respectively. The corresponding FRC plots are shown in Figure 4.9. We therefore notice that for images

containing low-density emitters, the proposed scheme works almost with the same accuracy as ThunderSTORM.

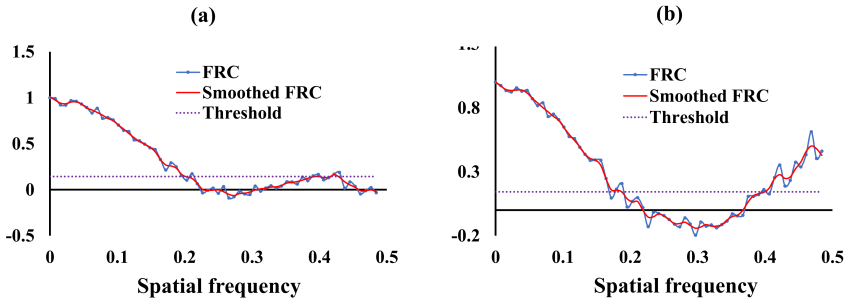

Figure 4.9. FRC plots corresponding to reconstructed images using (a) CAPL scheme and (b) ThunderSTORM.

4.6.2 Super-resolved Image Reconstruction in High-Density Fluorophore Emitters

The advantages of the CAPL scheme become more evident when we consider SMLM images formed by a high density of fluorophore emitters. In this scenario, we employ the three localisation methods, viz,; CAPL, 3D-DAOSTORM and ThunderSTORM on a high density dataset obtained from the same source as above [111]. To highlight the efficacy of each of the methods, we extract first 200 frames from the stack and process it using three methods. Row (i) of Figure 4.10 shows a representative diffraction limited image frame from the stack overlaid with ground truth locations marked in red '+' in (a), CAPL detections with blue '+' in (b), 3D-DAOSTORM detections in green circles in (c) and ThunderSTORM detections in yellow '+' in (d). On the other hand, row (ii) of the same figure shows the ground truth image in (a), and (b), (c) and (d) show the CAPL, 3D-DAOSTORM and ThunderSTORM reconstructed images, respectively. The assessment indexes of accuracy and precision are summarised in Table 4.4.

Method	Recall	Precision	JI	SSIM	Time (s)
CAPL scheme	0.6290	0.7321	0.5113	0.9052	0.78
ThunderSTORM	0.6412	0.3667	0.3043	0.8343	0.32
3D-DAOSTORM	0.7041	0.3232	0.2846	0.8968	1.40

Table 4.4. Quantitative performance metrics for SMLM localisation methods corresponding to Figure 4.10.

It is evident from Figure 4.10 and Table 4.4 that our CAPL scheme demonstrates superior accuracy, precision, tolerance against spatial overlap, and image quality compared to ThunderSTORM and 3D-DAOSTORM.

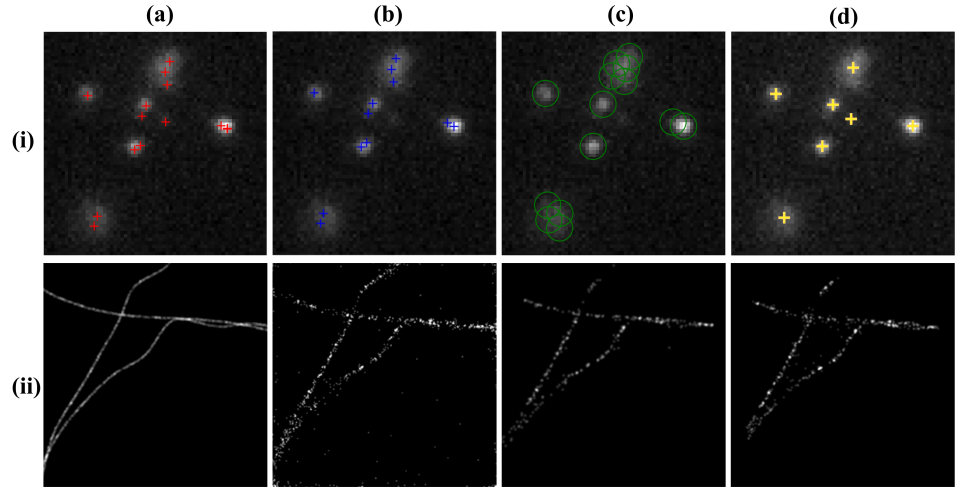

Figure 4.10. A comparison between the reconstructions of CAPL scheme, ThunderSTORM, and 3D-DAOSTORM from SMLM images of high-density emitters. Row (i) (a) shows a representative diffraction limited image frame overlaid with the ground truth locations marked in red '+', (b) shows CAPL detections with blue '+', (c) shows 3D-DAOSTORM detections in green circles and (d) shows ThunderSTORM detections in yellow '+'. Row (ii) (a) shows the Ground truth image, and the CAPL, 3D-DAOSTORM and ThunderSTORM reconstructed images in (b), (c) and (d), respectively.

4.6.3 Super-resolved Image Reconstruction in Experimental SMLM Data

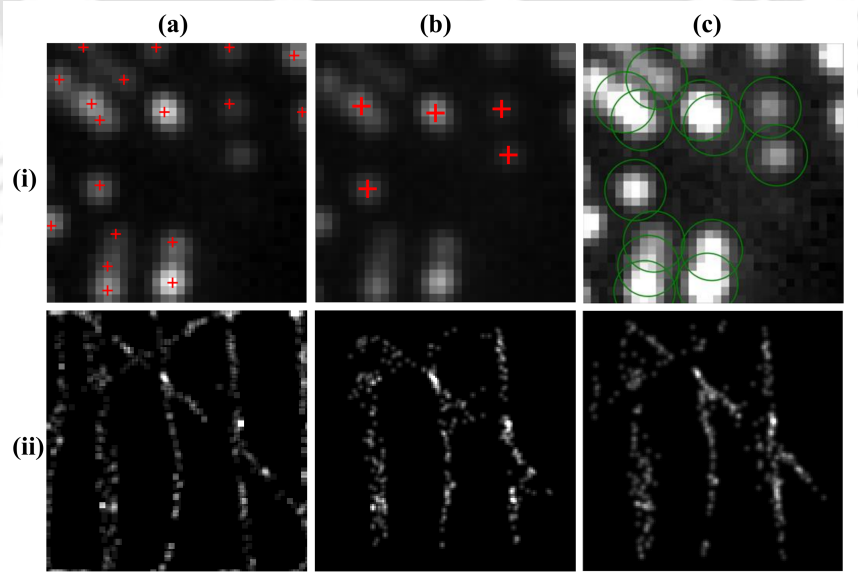

Figure 4.11. Implementation of CAPL scheme on an experimental dataset of 50 frames of size (32 × 32) pixels. Row (i) shows the diffraction-limited images overlayed by the PSF centres marked in red '+' identified using (a) CAPL scheme, (b) ThunderSTORM and (c) 3D-DAOSTORM on a particular image frame, while row (ii) shows the respective super-resolved images generated, each of size (128 × 128) pixels.

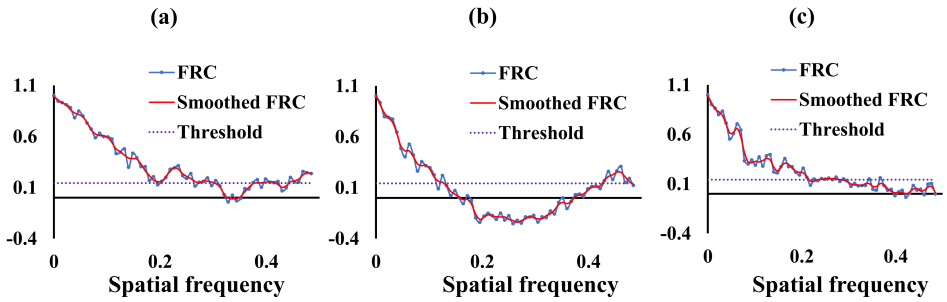

Figure 4.12. FRC plots corresponding to the reconstructions made using (a) CAPL scheme, (b) ThunderSTORM and (c) 3D-DAOSTORM .

We then implement our CAPL scheme on an experimental dataset [87, 137]. The stack has 500 image frames of size (128 ×128) pixels. We extract a patch of size (32 × 32) pixels from the first 50 frames and process the same using the proposed scheme, ThunderSTORM as well as 3D-DAOSTORM. Figure 4.11 (i) shows the detected PSFs in red '+' by each method for the 16th frame of the stack. We clearly see that our scheme can accurately detect most of the PSFs, in contrast with ThunderSTORM and 3D-DAOSTORM. CAPL could detect a total of 880 PSFs in 50 frames, while ThunderSTORM could detect 377 and 3D-DAOSTORM detected 721 PSFs. The effect of incorrect detections is reflected in the super-resolved images, shown in Figure 4.11 (ii). It is seen that our scheme creates a better quality super-resolved image compared to the other two methods. To evaluate the above outcomes quantitatively, we compute the FRC (fixed 1/7 criterion) for all above three cases and obtain FIRE numbers of 3.306, 7.851 and 4.633 for the CAPL scheme, ThunderSTORM and 3D-DAOSTORM, respectively. The FIRE number once again confirms our claim that our CAPL scheme provides superior localisation of densely packed PSFs in SMLM images. The corresponding FRC plots are shown in Figure 4.12.

4.7 Chapter Summary

This chapter introduces a novel approach for accurate PSF detection in SMLM under challenging high-density conditions, addressing the limitations of traditional methods when PSFs overlap. We developed a lightweight Convolutional Neural Network architecture to predict fluorophore positions as Gaussian heatmaps. This is followed by a blob detection routine as a post-processing step to achieve sub-pixel localisation, forming what we term the CNN Assisted PSF Localisation (CAPL) scheme. The proposed CAPL scheme was rigorously evaluated using both simulated and experimental SMLM datasets. Our results demonstrate that CAPL effectively separates overlapping PSFs and maintains high localisation accuracy across varying emitter densities. Quantitative comparisons using metrics such as recall, precision, Jaccard Index, RMSE, and SSIM consistently showed CAPL's superior performance,

especially in high-density imaging scenarios where traditional localisation algorithms like ThunderSTORM and 3D-DAOSTORM struggle with accuracy and artifact generation. Furthermore, assessments using FRC confirmed the enhanced resolution achieved by CAPL on both simulated and experimental data. By leveraging the power of deep learning while maintaining computational efficiency, the CAPL method thus offers a robust solution for advancing high-throughput and high-density super-resolution imaging.

CHAPTER 5
Real-Time Drift Correction for Single-Molecule Localization Microscopy

5.1 Introduction

Single Molecule Localisation Microscopy (SMLM), as discussed in the previous chapters, achieves super-resolution imaging by stochastically activating sparse subsets of fluorophores and localising them over thousands of sequential image frames. The accumulated localisations are then combined to reconstruct a super-resolved image [138, 9]. High-quality SMLM reconstructions therefore rely on accurate point spread function (PSF) detection, robust localisation, and effective handling of densely overlapping emitters. However, even when optimal detection and localisation algorithms are employed, the quality of the final reconstructed image can be severely degraded by sample drift occurring during the long acquisition times required for SMLM experiments.

Drift refers to the relative motion between the sample and the imaging system during data acquisition and is typically caused by thermal fluctuations, mechanical instabilities, or incomplete immobilisation of the specimen. Given that localisation precision in SMLM is often on the order of tens of nanometres, even sub-pixel drift can introduce significant localisation errors that accumulate over time. This accumulated error manifests as motion blur, loss of spatial fidelity, and structural distortion in the reconstructed image, ultimately limiting the achievable resolution.

Existing drift correction strategies can broadly be divided into hardware-based stabilisation approaches and computational post-processing methods. Hardware-based methods employ active feedback mechanisms, often interferometric, to physically stabilise the sample or objective in real time [47, 139]. While highly effective, such systems substantially increase the complexity and cost of the microscope setup and require careful calibration. Computational drift correction methods are more commonly used and are typically applied offline after data acquisition. These approaches rely either on fiducial markers embedded in the sample [140] or on correlation-based analysis of localisation data or reconstructed

images [141, 142]. However, a fundamental limitation of offline correction is the absence of feedback during acquisition, meaning that datasets compromised by excessive drift may only be identified after the experiment has concluded.

To address these limitations, this chapter presents a real-time, image-based drift correction framework that operates during data acquisition. The proposed method estimates lateral drift directly from the raw fluorescence images using frequency-domain cross-correlation and applies corrective pixel re-registration immediately to each acquired frame. The approach does not require fiducial markers, mechanical stabilisation hardware, or modification of the microscope stage. By integrating drift correction directly into the acquisition pipeline, the method preserves spatial consistency throughout long SMLM measurements and ensures that downstream localisation and reconstruction algorithms operate on temporally stable data.

5.2 Proposed Drift Correction Scheme

The proposed drift correction system is designed as a modular pipeline that processes the stream of raw images acquired by the camera. The fundamental steps involved in drift estimation and correction are illustrated in Figure 5.1. Each acquired image frame is denoted by $I_i(x, y)$, where i is the frame index. Drift is estimated by comparing each incoming frame with a reference frame, typically the first frame $I_1(x, y)$ of the acquisition.

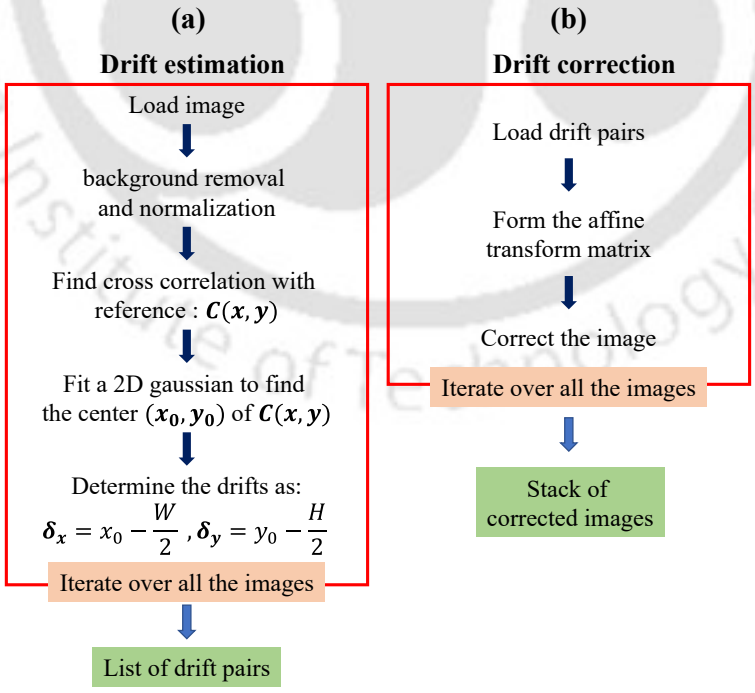

Figure 5.1. Illustration of the basic processes involved in (a) drift estimation and (b) generation of drift-corrected images.

Prior to drift estimation, each image undergoes background suppression and normalisation to improve robustness against uneven illumination and intensity fluctuations. In wide-field fluorescence microscopy, background illumination may arise from out-of-focus fluorescence, scattering within the optical system, and autofluorescence [40, 41]. To isolate diffraction-limited PSFs from these slowly varying background components, we apply a morphological top-hat operation for background removal [94]. This operation suppresses large-scale background features while preserving small, bright structures corresponding to individual emitters. Each frame is subsequently normalised to a common intensity range between 0 and 1 to ensure consistency across the image sequence. The estimated drift vector is then used to re-register pixel coordinates, restoring a consistent imaging geometry across the image frames. This framework can operate either in real time during acquisition or in offline modes for post-processing of pre-acquired datasets.

5.2.1 Drift Estimation and Correction

To determine the lateral displacement between the reference frame $I_1(x, y)$, which in this case corresponds to the first acquired image, and a subsequent frame $I_2(x, y)$, we compute their two-dimensional cross-correlation. This measure quantifies the relative translation required to align the subsequent image with the reference frame by maximising the overlap of their spatial features. Direct computation of the cross-correlation in the spatial domain is computationally expensive and therefore unsuitable for real-time operation. Instead, we exploit the convolution theorem and perform the computation efficiently in the Fourier domain as

$$C(x, y) = \mathcal{F}^{-1}\left(\mathcal{F}(I_1(x, y)) \cdot \overline{\mathcal{F}(I_2(x, y))}\right), \tag{5.1}$$

where $\mathcal{F}$ denotes the Fast Fourier Transform (FFT) and $\overline{\mathcal{F}(I)}$ denotes the complex conjugate of the Fourier transform of image I.

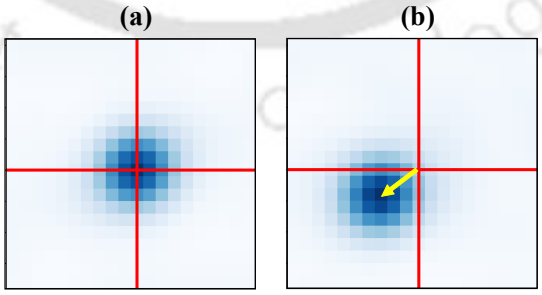

Figure 5.2. Cross-correlation peak of the reference frame (a) with itself and (b) with a different frame in which drift is present. The displacement of the peak, indicated by the yellow arrow, corresponds to the measured drift.

The resulting correlation matrix $C(x, y)$ exhibits a prominent peak whose location corresponds to the relative translation between the two images. To achieve sub-pixel accuracy, the peak position is refined by fitting a localised two-dimensional Gaussian model,

$$G(x, y) = A \exp\left(-\left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2}\right)\right) + B, \tag{5.2}$$

where the fitted centre (x_0, y_0) provides a high-precision estimate of the drift. Figure 5.2 illustrates the cross-correlation peak for both self-correlation and drifted image pairs. The drift vector $(\delta x, \delta y)$ is calculated relative to the centre of the correlation matrix with width W and height H as

$$\delta x = x_0 - \frac{W}{2} \quad \text{and} \quad \delta y = y_0 - \frac{H}{2}. \tag{5.3}$$

Once the drift vector is determined, the system applies a corrective translation to the image frame through pixel re-registration using an affine transformation,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\delta x \\ 0 & 1 & -\delta y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{5.4}$$

where (x', y') denote the corrected pixel coordinates and the matrix represents the affine transformation used to compensate for the measured drift.

5.2.2 Operating Modes

Our drift correction system is primarily designed to operate in real time during image acquisition. However, the same framework can also be deployed in two offline operating modes, where the image stack is first acquired in its entirety and subsequently processed for drift correction.

1. **Direct (real-time) mode:** Drift is estimated and corrected for every incoming frame during acquisition. In this mode, the drift estimation and correction steps are integrated directly into the image acquisition pipeline, allowing each frame to be corrected immediately upon capture.

2. **Interpolation method:** This is an accelerated offline mode where only a subset of frames (e.g., every N^{th} frame) is processed. The drift for the intermediate frames is then estimated using cubic spline interpolation [143]. By reducing the number of frames on which cross-correlation is performed, this approach can offer a substantial gain in computational efficiency, making it well suited for large, pre-acquired datasets.

3. **Projection-interpolation method:** This is another offline mode, particularly useful for sparse datasets. It bins small groups of consecutive frames, whose number is

denoted with as **B** via maximum intensity projection, thereby enhancing spatially consistent features while suppressing frame-to-frame stochastic fluctuations in emitter intensity. This improves the effective signal-to-noise ratio (SNR) prior to drift estimation, after which the interpolation method is applied.

Optionally, a Savitzky-Golay filter [144] can be applied to the calculated drift vectors in the offline modes to reduce high-frequency fluctuations and further smooth the estimated drift trajectory before correction.

5.3 Validation and Performance of the Drift Correction Scheme

The drift correction framework is first validated using simulated SMLM datasets with known ground-truth drift, as done in the previous chapters. The controlled setting in the simulated dataset enables a systematic assessment of the accuracy, performance, and operational limits of the proposed scheme under idealised conditions. Synthetic datasets consisting of 5000 frames are generated, as described in the previous chapters, with fluorophores randomly distributed on a two-dimensional plane and PSFs modelled as two-dimensional Gaussians corrupted with realistic Poisson shot noise and Gaussian readout noise. The simulations emulate the stochastic photoswitching behaviour of fluorophores and their image acquisition by a camera system.

In SMLM, any uncontrolled motion between the specimen plane and the imaging system introduces a time-dependent lateral displacement of the fluorophore PSFs. To replicate this effect in a controlled manner, a predefined drift trajectory is applied to the fluorophore positions. We then systematically evaluate the performance of the drift correction framework under varying conditions, including image size and emitter density,thereby enabling direct quantitative comparison between the imposed ground-truth drift and the estimated drift.

5.3.1 Accuracy and Impact on Localisation Quality

Figure 5.3 demonstrates the ability of the proposed scheme to re-centre a drifting PSF. We see that the real-time, interpolation, as well as the projection-interpolation modes can correct the drift, thus validating the effective operation of all the modes.

The influence of drift correction on localisation accuracy is illustrated in Figure 5.4, which shows that applying the correction transforms a blurred reconstruction into a sharp and well-localised image. This qualitative improvement is reflected in a marked enhancement in standard localisation performance metrics, as introduced in Chapter 1, including precision (the fraction of detections corresponding to true emitters), recall (the fraction of true emitters correctly identified), and the Jaccard index, which measures the overall similarity between the ground truth and detected emitter sets by relating the number of true positives to the

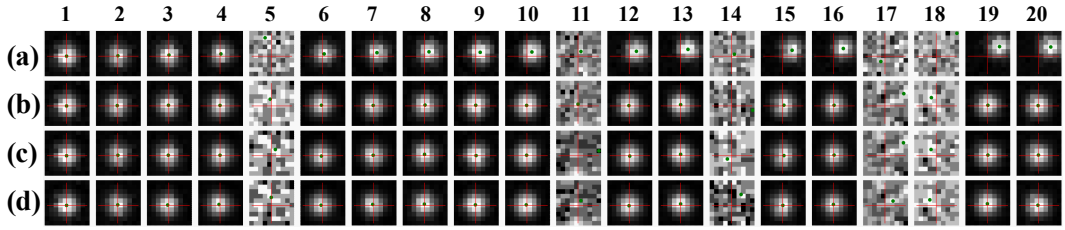

Figure 5.3. Visual validation of drift correction on a simulated PSF. (a) A single PSF drifting over 20 consecutive frames. (b–d) The same PSF after correction using the real-time, interpolation, and projection–interpolation modes, respectively. The system successfully re-centres the PSF to the reference position (red crosshairs), even in frames where the PSF is absent.

union of all ground truth and detected points [57]. Taken together, these results highlight the importance of effective drift correction in maintaining localisation accuracy during long acquisitions.

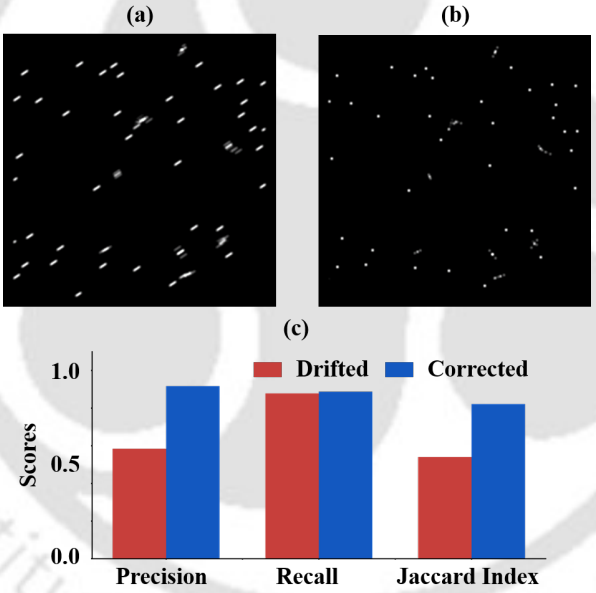

Figure 5.4. Impact of drift correction on localisation quality for a simulated SMLM stack of 5000 images with dimensions (128 × 128) pixels. (a) Reconstructed image in the presence of drift. (b) Reconstructed image after drift correction, showing sharp and well-localised emitters. (c) Corresponding improvement in localisation performance metrics.

5.3.2 Validation on High-Density DNA-PAINT Nanostructures

The method is further validated using publicly available DNA-PAINT datasets comprising DNA origami nanostructures with known ground-truth geometries [145]. Each structure has a rectangular geometry and consists of four binding sites located at its four vertices, where imaging strands carrying fluorophore molecules can transiently bind. The complete

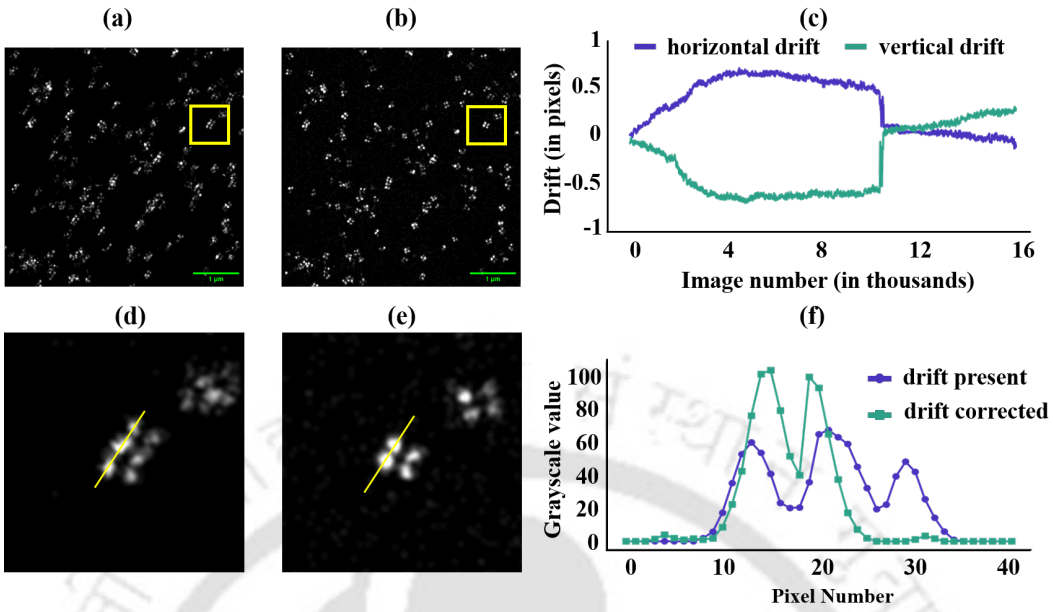

Figure 5.5. Offline drift correction on a DNA-PAINT dataset of DNA origami nanostructures. (a, d) Reconstructed image and a magnified region from the original, uncorrected data, where drift artifacts merge the nanostructures. (b, e) The same views after correction with our method, showing clearly resolved individual structures. (c) The estimated drift trace, showing significant displacement. (f) A line profile across two structures (indicated in c and d) shows two unresolved humps in the drifted data (blue) become two distinct peaks after correction (green).

dataset consists of 15,258 frames of size (512 × 256) pixels. Figures 5.5 (a) and (d) show that uncorrected drift leads to severe blurring and merging of nanostructures in the reconstructed image, whereas the application of drift correction effectively suppresses these artifacts. The corrected reconstructions, shown in Figures 5.5 (b) and (e), clearly resolve individual binding sites and restore the expected structural separation.

The accuracy of the correction is quantitatively confirmed in two ways. First, the line profile across two adjacent structures, shown in Figure 5.5 (f), demonstrates that the three broad, unresolved peaks observed in the drifted data are transformed into two sharp and distinct peaks after correction. Second, the measured average peak-to-peak separations within each structure in the corrected image are 63.25 ± 2.37 nm and 82.13 ± 3.61 nm, which are in excellent agreement with the known 60 nm and 80 nm design spacings of the DNA origami structures.

It is also noteworthy that the entire stack of 15,258 frames is processed in just 49 seconds using the interpolation method (using $N = 3$) on our hardware, corresponding to a processing speed of approximately 311 frames per second, which demonstrates the computational efficiency of the proposed offline correction strategy.

5.3.3 Performance Comparison of Operating Modes

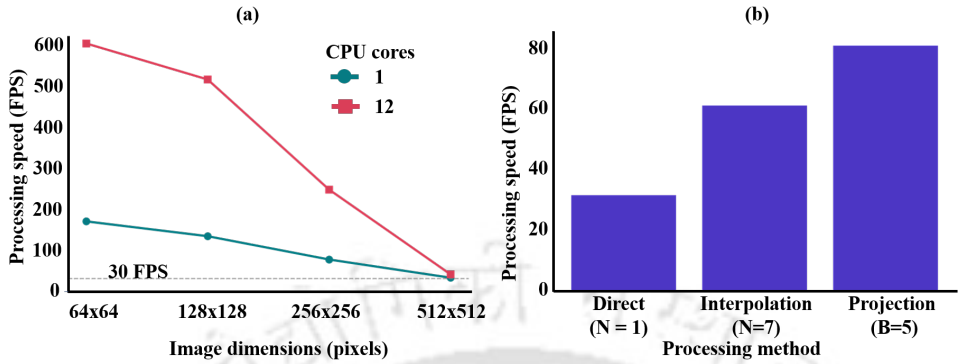

Figure 5.6. Performance scaling with image size and operating mode. (a) Processing speed (FPS) for different image dimensions using a single CPU core (1 worker) and all available CPU cores (12 workers). The grey dashed line at 30 FPS indicates a typical threshold for real-time, video-rate processing. (b) Speedup achieved when the interpolation and projection–interpolation methods are applied to an image stack of dimensions (512 × 512) pixels.

We benchmark the maximum processing speed, measured in frames per second (FPS), as a function of image size for the direct (i.e. real-time) drift correction method. This analysis is intended to assess whether the proposed framework can meet the timing constraints imposed by typical SMLM acquisition rates. Figure 5.6 (a) shows the performance scaling obtained using a single CPU core compared with a multi-core (12-worker) implementation. The use of multi-core processing provides a significant performance boost, enabling real-time drift correction at speeds exceeding 30 FPS. Even for large (512 × 512) pixel frames, the system maintains a processing speed of nearly 31 FPS on our CPU for both single-core and multi-core implementations, which is sufficient for sustained real-time operation in most SMLM experiments. We further observe that for smaller fields of view, as the camera frame rate is increased, the processing speed of the scheme scales approximately linearly.

We next process an image stack of dimensions (512 × 512) pixels using the interpolation method by analysing every seventh frame (using $N = 7$), and subsequently apply the projection-interpolation method with a group size of $B = 5$. The corresponding results are shown in Figure 5.6 (b), where a clear and consistent increase in processing speed is observed for both offline strategies.

We further assess the accuracy of the offline methods relative to the direct method by investigating the trade-off between processing speed and accuracy for the interpolation method as the interpolation interval is varied. The results, shown in Figure 5.7 (a), demonstrate that accuracy is only minimally impacted by this strategy. For a (512 × 512) dataset, increasing the interpolation interval from 1 (direct method) to 15 (processing every 15th

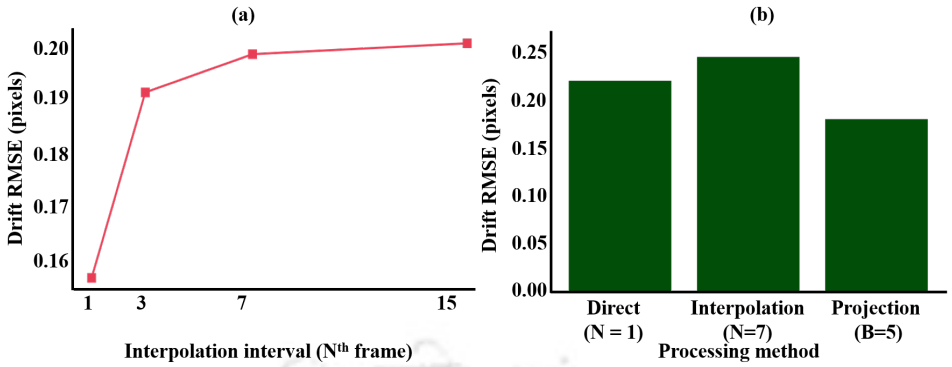

Figure 5.7. Accuracy trade-offs across different operating modes for a (512 × 512) dataset. (a) Evolution of total drift RMSE (in pixels) as the interpolation interval (N) increases from 1 to 15 frames. (b) Comparison of total drift RMSE for the direct, interpolation ($N = 7$), and projection-interpolation ($B = 5$) methods, demonstrating the improved accuracy of the projection-interpolation method under low-signal conditions.

frame) increases the total drift RMSE by only 0.043 pixels (i.e. from 0.162 to 0.205 pixels), which indicates that substantial computational savings can be achieved with only a modest reduction in accuracy. This confirms that interpolation is an effective and reliable strategy for post-processing large datasets.

The accuracy of the drift correction scheme also depends on the characteristics of the dataset and the underlying signal level. To account for this, the projection–interpolation method, which averages a small block of consecutive frames (e.g. $B = 5$) prior to cross-correlation, is designed to improve the effective SNR in scenarios with sparse emitter activation. We compare its performance against the standard interpolation method (using $N = 7$) for a (512 × 512) dataset containing sparse fluorophore activation (160 emitters per frame). As shown in Figure 5.7 (b), the projection-interpolation method yields a notable improvement in accuracy, reducing the RMSE by approximately 28% (from 0.246 to 0.177 pixels) and resulting in a lower error than that obtained using the direct method, which has an RMSE of 0.215 pixels.

The offline methods therefore provide a practical and efficient alternative to the direct method for post-acquisition processing of SMLM datasets. In this context, for sparse datasets, the projection–interpolation method offers a favourable trade-off between processing speed and drift estimation accuracy. A summary of processing speed and accuracy for the real-time and offline operating modes is provided in Table 5.1.

Image Size	Emitters	Direct (real-time)		Interpolation (N=7)		Projection (B=5)	
		FPS	RMSE	FPS	RMSE	FPS	RMSE
64×64	10	163.8	0.133	1039.4	0.018	645.5	0.180
128×128	40	129.9	0.183	757.6	0.222	617.5	0.147
256×256	160	76.6	0.156	384.6	0.196	262.4	0.155
	640	30.8	0.162	129.9	0.205	78.1	0.143
512×512	320	31.1	0.119	129.8	0.172	78.5	0.145
	160	30.9	0.215	129.5	0.246	79.3	0.177

Table 5.1. Summary of processing speed and accuracy for the real-time and offline drift correction modes using a single CPU core. The interpolation method uses an interval of $N = 7$, and the projection–interpolation method uses a block size of $B = 5$. RMSE denotes the total root mean square error relative to the ground-truth drift.

5.4 Experimental Implementation of Real-Time Drift Correction

The workflow for the experimental implementation of the direct, or real-time, drift correction mode is illustrated in Figure 5.8. At the start of image acquisition, the first frame (I_1) captured by the microscope camera is designated as the reference frame. For each subsequent frame (I_i), the system executes a cross-correlation algorithm relative to (I_1) to estimate the lateral drift vector. The corresponding corrective affine transformation is then applied immediately to generate a drift-corrected frame (I_i^C), which is subsequently added to a growing stack (S) for localisation and reconstruction.

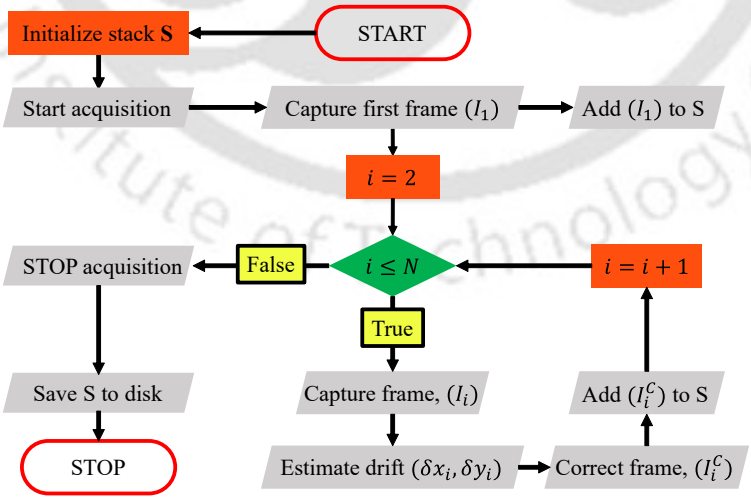

Figure 5.8. Steps involved in drift correction during image acquisition. This workflow corresponds to the direct method, or real-time operating mode.

To achieve true real-time operation, this entire processing cycle must be completed within the camera frame acquisition interval (e.g., < 40 ms). In cases involving larger image sizes, where the processing time exceeds the interval between consecutive frame acquisitions, the system can optionally operate the camera in triggered mode by issuing a trigger signal prior to the acquisition of each subsequent frame (I_i). Notably, the experimental implementation does not require any additional hardware, and the drift correction scheme can be readily integrated into existing acquisition software frameworks, such as Micro-Manager or custom Python or LabVIEW-based programs, and thus providing an accessible and flexible solution for live drift stabilisation in SMLM experiments.

5.4.1 Live Drift Correction on Quantum Dots

To demonstrate the real-time capability of the proposed system, we perform a live drift correction experiment on a sample of blinking quantum dots (QDs; Qdot 705 carboxyl quantum dots). The experiment is conducted using a modular openFrame-based single-molecule localisation microscope [37, 38], equipped with a 462 nm multimode excitation laser. Images are acquired using an IDS GigE industrial CMOS camera (model GV-79F0WP-M-GL, 2.74 μm pixel size) operating at full sensor acquisition with 4×4 on-chip binning. The system is controlled by a custom Python-based acquisition program that integrates the proposed drift correction module. Under these imaging conditions, the effective pixel size is 99.63 nm, calculated as $\left(\frac{2.74\times4}{110}\ \mu\text{m}\right)$.

Image acquisition is continued for 11 minutes and 06 seconds, during which a total of 9000 frames are recorded over the full field of view. This extended acquisition duration is deliberately chosen to induce and observe significant thermal and mechanical drift. It is important to note that the drift correction is applied instantaneously, and the processing introduces no additional latency beyond the standard camera acquisition time. During acquisition, the system computes the drift for each incoming frame and simultaneously stores both the original raw image stack and the live drift-corrected image stack. The estimated horizontal and vertical drift traces are shown in Figure 5.9 (a). Both image stacks are subsequently processed for STORM image reconstruction using the ThunderSTORM [61] plugin in Fiji. A representative localised QD from the original (uncorrected) stack is shown in Figure 5.9 (b), while the same localised QD in the drift corrected stack is shown in Figure 5.9 (c).

The results, shown in Figure 5.10, clearly highlight the effectiveness of the live drift correction. The reconstruction obtained from the uncorrected data as seen in Figure 5.10 (a) is severely compromised, with each quantum dot which is expected to appear as a sharp point, is smeared into an elongated streak due to cumulative sample drift over the acquisition period. Such reconstructions are unsuitable for reliable quantitative analysis. In contrast,

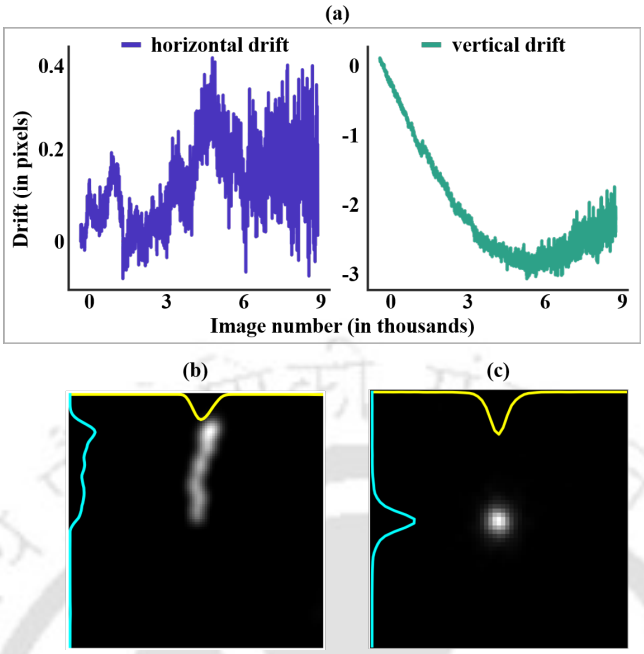

Figure 5.9. (a) Estimated drift in the horizontal and vertical directions for the acquired quantum dot (QD) image sequence. (b) & (c) A representative localised QD with the corresponding x and y intensity projections overlaid when drift is present and when it is corrected.

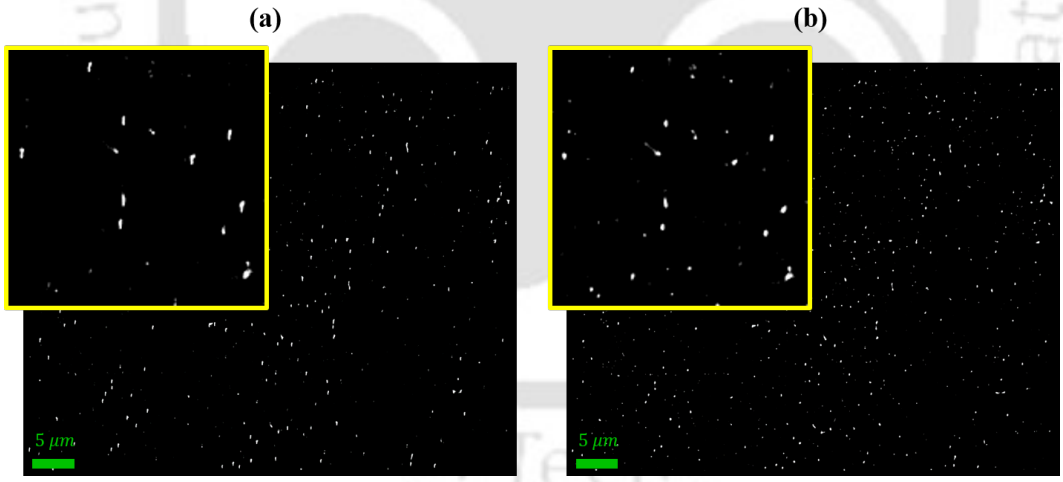

Figure 5.10. Real-time drift correction of 20 nm quantum dots over an 11-minute acquisition. (a) Reconstructed image from the uncorrected raw data, where significant drift causes severe motion blur, smearing each quantum dot into an elongated streak. (b) Reconstructed image from the data stack, corrected live during acquisition, showing sharp and well-resolved quantum dots.

the reconstruction obtained from the live drift-corrected stack, as seen in Figure 5.10 (b) exhibits sharp, distinct, and well-resolved quantum dots. The proposed system successfully tracks and compensates for drift in real time, thereby preserving the resolution of the final

image. Quantitative analysis of the localised quantum dots further reveals a substantial reduction in the full width at half maximum (FWHM), from an average of approximately 75 nm in the uncorrected data to approximately 34 nm after drift correction.

5.5 Chapter Summary

In this chapter, a real-time, image-based drift correction framework for Single Molecule Localisation Microscopy has been developed and validated. By estimating lateral drift through frequency-domain cross-correlation and applying corrective pixel re-registration during acquisition, the method preserves spatial consistency throughout long SMLM measurements. The approach requires no fiducial markers or additional hardware and integrates seamlessly into existing microscope setups. Extensive validation using simulated datasets, DNA-PAINT nanostructures, and live experimental measurements demonstrates that real-time drift correction substantially improves localisation accuracy and reconstructed image quality.

CHAPTER 6
Conclusions and Future Directions

6.1 Conclusions of the work

Single molecule localisation microscopy (SMLM) provides an easy yet efficient way to break the diffraction barrier and reconstruct a super-resolved image by gathering information from thousands of fluorescent images. SMLM based imaging techniques such as Stochastic Optical Reconstruction Microscopy (STORM) have contributed immensely to further research in biological and biomedical sciences. However, despite the advantages, SMLM suffers from some major limitations. In this thesis we attempted to address a few such limitations, especially pertaining to the image reconstruction process.The contributions of this thesis span from algorithmic development to system-level integration, culminating in a robust, efficient, and accessible SMLM workflow.

Chapter 1 established the theoretical and practical foundations of the work by introducing the principles of light and fluorescence microscopy, the diffraction limit, and an overview of super-resolution imaging techniques, with particular emphasis on SMLM and its dSTORM implementation. This chapter provided the necessary context for the methodological and computational developments presented in the subsequent chapters.

Chapter 2 detailed the development of a lightweight, modular, Python-based two-dimensional localisation software framework. Designed to support both offline and real-time processing through a Micro-Manager plugin, the framework integrates image preprocessing, PSF detection, sub-pixel localisation, and multiple rendering strategies for super-resolved image reconstruction. Benchmarking against established tools such as ThunderSTORM demonstrated comparable performance in standard localisation scenarios, while offering greater flexibility and computational efficiency.

A key contribution towards reconstruction of the super-resolved image was introduced in Chapter 3 through the development of an adaptive entropy-based image thresholding method for SMLM data. By dynamically determining optimal thresholds from pixel intensity histograms, the proposed approach demonstrated improved robustness to spatially varying background noise compared to conventional global and local thresholding techniques. When

combined with blob detection-based crude localisation, this method yielded superior PSF detection accuracy and improved reconstruction quality, particularly under low signal-to-noise and heterogeneous background conditions.

Chapter 4 presented a convolutional neural network–based localisation framework, termed the CNN-Assisted PSF Localisation (CAPL) scheme, designed specifically to address dense-emitter SMLM datasets. The lightweight CNN architecture was trained to predict fluorophore positions as Gaussian-like intensity distributions, with sub-pixel precision achieved through subsequent blob-based refinement. Extensive evaluation on simulated and experimental datasets demonstrated that CAPL outperforms conventional methods in terms of recall, precision, Jaccard Index, and reconstructed image quality, particularly in high-density regimes where PSF overlap is significant. Fourier ring correlation (FRC) analysis further confirmed the resolution improvements achieved using this approach.

Chapter 5 addressed the critical challenge of sample drift in long-duration SMLM acquisitions by introducing an image-based lateral drift correction framework. Using a Fast Fourier Transform (FFT)-based cross-correlation approach applied directly to raw image frames, the proposed method enabled accurate estimation and compensation of lateral drift in both offline and real-time operating modes. Validation on simulated and experimental datasets demonstrated significant improvements in reconstructed image fidelity, including enhanced resolution and reduced full-width at half-maximum (FWHM) of localised emitters. The computational efficiency and hardware-independent nature of this method make it particularly suitable for real-time drift stabilisation in practical SMLM experiments.

Finally, Chapter 6 consolidated the major contributions and findings of this thesis, besides highlighting the advances achieved in SMLM image processing through integrated software development, adaptive signal detection, deep-learning-based localisation, and real-time drift correction. In addition to summarising the core outcomes, the chapter reflected on the practical impact and limitations of the proposed methods, thereby motivating future research directions in this rapidly evolving field.

Overall, the work presented in this thesis contributes meaningfully to the broader field of super-resolution microscopy by addressing several fundamental challenges in SMLM image reconstruction. By improving localisation accuracy in noisy and high-density conditions and introducing practical solutions for system stability, the proposed methods collectively enhance the robustness, performance, and accessibility of SMLM workflows. The modular design and computational efficiency of the developed tools make them well suited for both offline analysis and real-time integration into microscopy systems, particularly in low-resource environments where access to specialised hardware or GPU acceleration and storage capacity may be limited.

Despite the promising results, several limitations warrant consideration. The adaptive thresholding approach, while offering improved accuracy, incurs a modest increase in computational cost relative to simpler global methods. The CNN-based localisation framework, although lightweight and efficient, may require retraining or fine-tuning when applied to datasets acquired under substantially different imaging conditions, such as altered noise characteristics or PSF shapes. Additionally, the image-based drift correction strategy can perform optimally when the field of view is large, containing sufficient structural features.

In reflecting on the overall objectives of this thesis, the primary goal of developing computationally efficient and accessible methods for SMLM analysis, particularly in low-resource settings, has been largely achieved. The work presented demonstrates that it is possible to balance localisation accuracy, computational efficiency, and practical usability within a unified framework. While there remain areas for further refinement, the contributions of this thesis represent a meaningful step towards more accessible and adaptable SMLM workflows. On a personal level, this research has provided valuable insight into both the theoretical and practical aspects of super-resolution microscopy, and has contributed, albeit modestly, to the ongoing development of the field.

6.2 Directions for Future Research

Addressing the limitations stated in the previous section can form a natural basis for future refinement and extension of the proposed approaches proposed in the thesis. The field of Single-Molecule Localisation Microscopy continues to evolve rapidly, driven by advances in both experimental techniques and computational analysis. Building on the contributions of this thesis, several other promising avenues for future exploration can be envisioned to further enhance the capabilities, accessibility, and utility of SMLM.

An important direction yet to be completed in this thesis is the development of a hardware-based autofocus and axial stabilisation system using an astigmatic vortex beam [146]. In this approach, an infrared laser beam (830 nm) is structured into a vortex profile using a phase mask or spatial light modulator and passed through two astigmatic lenses with different focal lengths oriented orthogonally. When focused onto the glass–oil interface of the sample and collected by a secondary detection camera, the reflected beam produces a characteristic two-lobed intensity distribution whose peak-to-peak separations along the orthogonal axes vary systematically with axial displacement. Preliminary investigations, conducted through numerical simulations demonstrated that this response can be calibrated prior to imaging to generate a lookup table relating axial position to lobe separation. During acquisition, the lookup table could be used to estimate and correct focus drift in real time. Integrating this axial focus-lock mechanism with the image-based lateral drift correc-

tion framework presented in this thesis would enable full three-dimensional stabilisation, offering a compact and hardware-efficient solution for long-term SMLM experiments.

One exciting new direction is the development of point-cloud based analysis pipelines. Traditional super-resolved images are usually rendered by binning localised fluorophore coordinates onto a 2D grid, which may obscure structural information at fine scales. Treating SMLM data as native point clouds opens the door to techniques drawn from computer graphics and 3D vision, such as spatial clustering, topology-aware segmentation, and geometric deep learning. These approaches could allow more accurate quantification of nanoscale architecture and dynamic molecular interactions directly from localisation coordinates.

Equally important is the advancement of visualisation tools that go beyond static reconstructions. Real-time, interactive visualisation environments leveraging game engines or virtual reality platforms could enable researchers to explore complex spatial structures in 3D, annotate regions of interest, and trace molecular trajectories. Such systems could also integrate confidence metrics from localisation software, helping users assess the reliability of different features in reconstructed datasets.

Another promising line of research is the design of modular, "plug-and-play" hardware-software ecosystems. With the increasing availability of open-source microscopy hardware (e.g., OpenFrame) and real-time control platforms, there is a need for standardised interfaces that allow easy integration of localisation algorithms, drift correction, and adaptive feedback mechanisms with the imaging system. This could lead to intelligent microscopes that dynamically adjust imaging parameters based on real-time localisation density, drift estimates, or fluorophore photophysics.

On the data processing front, cloud-based SMLM pipelines could democratise access to super-resolution imaging by enabling users to upload raw data and receive analysed results without requiring high-performance local hardware. Combining this with pre-trained deep learning models that adaptively refine themselves on user data would further streamline analysis and encourage broader adoption.

Finally, the integration of SMLM with emerging modalities such as correlative light and electron microscopy (CLEM), light-sheet SMLM, or optogenetic manipulation, will require innovative strategies for data fusion and synchronisation. Developing tools to co-register and interpret multimodal datasets could uncover entirely new dimensions of subcellular organisation and dynamics.

References

[1] Frank De Lange, Alessandra Cambi, Richard Huijbens, Bärbel de Bakker, Wouter Rensen, Maria Garcia-Parajo, Niek van Hulst, and Carl G Figdor. Cell biology beyond the diffraction limit: near-field scanning optical microscopy. *Journal of cell science*, 114(23):4153–4160, 2001. 1

[2] Brian Herman and John J Lemasters. *Optical microscopy: emerging methods and applications*. Elsevier, 2012. 1

[3] Bo Huang, Mark Bates, and Xiaowei Zhuang. Super-resolution fluorescence microscopy. *Annual review of biochemistry*, 78:993–1016, 2009. 1, 12

[4] Stefan W Hell. Far-field optical nanoscopy. *science*, 316(5828):1153–1158, 2007. 1, 7

[5] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.*, 19(11):780–782, Jun 1994. 1, 7

[6] Thomas A. Klar and Stefan W. Hell. Subdiffraction resolution in far-field fluorescence microscopy. *Opt. Lett.*, 24(14):954–956, Jul 1999. 1, 7

[7] Eric Betzig, George H Patterson, Rachid Sougrat, O Wolf Lindwasser, Scott Olenych, Juan S Bonifacino, Michael W Davidson, Jennifer Lippincott-Schwartz, and Harald F Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *science*, 313(5793):1642–1645, 2006. 1, 8, 9

[8] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature methods*, 3(10):793–796, 2006. 8, 9

[9] Mickaël Lelek, Melina T Gyparaki, Gerti Beliu, Florian Schueder, Juliette Griffié, Suliana Manley, Ralf Jungmann, Markus Sauer, Melike Lakadamyali, and Christophe Zimmer. Single-molecule localization microscopy. *Nature Reviews Methods Primers*, 1(1):39, 2021. 1, 9, 13, 20, 79

[10] Bonnie O Leung and Keng C Chou. Review of super-resolution fluorescence microscopy for biology. *Applied spectroscopy*, 65(9):967–980, 2011. 2

[11] Masahito Yamanaka, Nicholas I Smith, and Katsumasa Fujita. Introduction to super-resolution microscopy. *Microscopy*, 63(3):177–192, 2014.

[12] Jeroen Vangindertael, Rafael Camacho, Wouter Sempels, Hideaki Mizuno, Peter Dedecker, and KPF Janssen. An introduction to optical super-resolution microscopy for the adventurous biologist. *Methods and applications in fluorescence*, 6(2):022003, 2018. 2

[13] Monica Monici. Cell and tissue autofluorescence research and diagnostic applications. *Biotechnology annual review*, 11:227–256, 2005. 4

[14] Jan Kapuscinski. Dapi: a dna-specific fluorescent probe. *Biotechnic & histochemistry*, 70(5):220–233, 1995. 5

[15] Martin Chalfie, Yuan Tu, Ghia Euskirchen, William W Ward, and Douglas C Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, 1994. 5

[16] Takashi Funatsu, Yoshie Harada, Makio Tokunaga, Kiwamu Saito, and Toshio Yanagida. Imaging of single fluorescent molecules and individual atp turnovers by single myosin molecules in aqueous solution. *Nature*, 374(6522):555–559, 1995. 5

[17] Alex M Valm, Jessica L Mark Welch, Christopher W Rieken, Yuko Hasegawa, Mitchell L Sogin, Rudolf Oldenbourg, Floyd E Dewhirst, and Gary G Borisy. Systems-level analysis of microbial community organization through combinatorial labeling and spectral imaging. *Proceedings of the National Academy of Sciences*, 108(10):4152–4157, 2011. 5

[18] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005. 5

[19] George Biddell Airy. On the diffraction of an object-glass with circular aperture. *Transactions of the Cambridge Philosophical Society*, 5:283, 1835. 5

[20] Ernst Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für mikroskopische Anatomie*, 9(1):413–468, 1873. 5

[21] Rayleigh. Xxxi. investigations in optics, with special reference to the spectroscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 8(49):261–274, 1879. 5

[22] Carroll Mason Sparrow. On spectroscopic resolving power. *Astrophysical Journal, vol. 44, p. 76*, 44:76, 1916. 6

[23] Mark AA Neil, Rimas Juškaitis, and Tony Wilson. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Optics letters*, 22(24):1905–1907, 1997. 7

[24] Rainer Heintzmann and Christoph G Cremer. Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. In *Optical biopsies and microscopic techniques III*, volume 3568, pages 185–196. SPIE, 1999.

[25] Mats GL Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy*, 198(2):82–87, 2000.

[26] Rainer Heintzmann, Thomas M Jovin, and Christoph Cremer. Saturated patterned excitation microscopy—a concept for optical resolution improvement. *Journal of the Optical Society of America A*, 19(8):1599–1609, 2002. 7

[27] Francisco Balzarotti, Yvan Eilers, Klaus C Gwosch, Arvid H Gynnå, Volker Westphal, Fernando D Stefani, Johan Elf, and Stefan W Hell. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science*, 355(6325):606–612, 2017. 8

[28] Klaus C Gwosch, Jasmin K Pape, Francisco Balzarotti, Philipp Hoess, Jan Ellenberg, Jonas Ries, and Stefan W Hell. Minflux nanoscopy delivers 3d multicolor nanometer resolution in cells. *Nature methods*, 17(2):217–224, 2020. 8

[29] Samuel T Hess, Thanu PK Girirajan, and Michael D Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–4272, 2006. 8

[30] Mike Heilemann, Sebastian Van De Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie International Edition*, 47(33):6172–6176, 2008. 8, 10, 13

[31] Alexey Sharonov and Robin M Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, 103(50):18911–18916, 2006. 8, 9

[32] Joerg Schnitzbauer, Maximilian T Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. Super-resolution microscopy with dna-paint. *Nature protocols*, 12(6):1198–1228, 2017. 8, 9, 19, 23

[33] Russell E Thompson, Daniel R Larson, and Watt W Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical journal*, 82(5):2775–2783, 2002. 8, 20, 33, 37, 40, 47, 62

[34] Raimund J Ober, Sripad Ram, and E Sally Ward. Localization accuracy in single-molecule microscopy. *Biophysical journal*, 86(2):1185–1200, 2004. 33, 37, 61

[35] Kim I Mortensen, L Stirling Churchman, James A Spudich, and Henrik Flyvbjerg. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature methods*, 7(5):377–381, 2010. 8, 34, 35

[36] Bo Huang, Wenqin Wang, Mark Bates, and Xiaowei Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813, 2008. 9, 15

[37] J Lightley, S Kumar, MQ Lim, E Garcia, F Görlitz, Y Alexandrov, T Parrado, C Hollick, E Steele, K Roßmann, et al. openframe: a modular, sustainable, open microscopy platform with single-shot, dual-axis optical autofocus module providing high precision and long range of operation. *Journal of Microscopy*, 292(2):64–77, 2023. 11, 89

[38] Anupam Bharadwaj, Ranjan Kalita, Amalesh Kumar, Anupam Sarma, Bithiah G Jaganathan, Sunil Kumar, Frederik Gorlitz, Jonathan Lightley, Chris Dunsby, Mark Neil, et al. A cost-effective, modular, research-grade optical microscope. *Curr. Sci.*, 126:244–254, 2024. 11, 89

[39] Jennifer C Waters. Accuracy and precision in quantitative fluorescence microscopy, 2009. 13, 31, 52

[40] Y Garini, A Gil, I Bar-Am, D Cabib, and N Katzir. Signal to noise analysis of multiple color fluorescence imaging microscopy. *Cytometry: The Journal of the International Society for Analytical Cytology*, 35(3):214–226, 1999. 13, 31, 81

[41] Anna Jezierska, Hugues Talbot, Caroline Chaux, Jean-Christophe Pesquet, and Gilbert Engler. Poisson-gaussian noise parameter estimation in fluorescence microscopy imaging. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1663–1666. IEEE, 2012. 13, 26, 31, 47, 52, 81

[42] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied optics*, 46(10):1819–1829, 2007. 14, 33

[43] Sjoerd Stallinga and Bernd Rieger. Accuracy of the gaussian point spread function model in 2d localization microscopy. *Optics express*, 18(24):24461–24476, 2010. 14, 33

[44] Olivier Haeberlé, Mehdi Ammar, Hiromitsu Furukawa, Koji Tenjimbayashi, and Peter Török. Point spread function of optical microscopes imaging through stratified media. *Optics Express*, 11(22):2964–2969, 2003. 14, 33

[45] Sri Rama Prasanna Pavani and Rafael Piestun. Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system. *Optics express*, 16(26):22048–22057, 2008. 15

[46] Matthew D Lew, Steven F Lee, Majid Badieirostami, and WE Moerner. Corkscrew point spread function for far-field three-dimensional nanoscale localization of pointlike objects. *Optics letters*, 36(2):202–204, 2011. 15

[47] Ginni Grover, Wyatt Mohrman, and Rafael Piestun. Real-time adaptive drift correction for super-resolution localization microscopy. *Optics express*, 23(18):23887–23898, 2015. 15, 79

[48] Yoav Shechtman, Lucien E Weiss, Adam S Backer, Steffen J Sahl, and WE Moerner. Precise three-dimensional scan-free multiple-particle tracking over large axial ranges with tetrapod point spread functions. *Nano letters*, 15(6):4194–4199, 2015. 15

[49] Manuel F Juette, Travis J Gould, Mark D Lessard, Michael J Mlodzianoski, Bhupendra S Nagpure, Brian T Bennett, Samuel T Hess, and Joerg Bewersdorf. Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nature methods*, 5(6):527–529, 2008. 16

[50] Robert Prevedel, Young-Gyu Yoon, Maximilian Hoffmann, Nikita Pak, Gordon Wetzstein, Saul Kato, Tina Schrödel, Ramesh Raskar, Manuel Zimmer, Edward S Boyden, et al. Simultaneous whole-animal 3d imaging of neuronal activity using light-field microscopy. *Nature methods*, 11(7):727–730, 2014. 16

[51] Michael Broxton, Logan Grosenick, Samuel Yang, Noy Cohen, Aaron Andalman, Karl Deisseroth, and Marc Levoy. Wave optics theory and 3-d deconvolution for the light field microscope. *Optics express*, 21(21):25418–25439, 2013.

[52] Ruth R Sims, Sohaib Abdul Rehman, Martin O Lenz, Sarah I Benaissa, Ezra Bruggeman, Adam Clark, Edward W Sanders, Aleks Ponjavic, Leila Muresan, Steven F Lee, et al. Single molecule light field microscopy. *Optica*, 7(9):1065–1072, 2020. 16

[53] Elias Nehme, Daniel Freedman, Racheli Gordon, Boris Ferdman, Lucien E Weiss, Onit Alalouf, Tal Naor, Reut Orange, Tomer Michaeli, and Yoav Shechtman. Deepstorm3d: dense 3d localization microscopy and psf design by deep learning. *Nature methods*, 17(7):734–740, 2020. 16, 62

[54] Artur Speiser, Lucas-Raphael Müller, Philipp Hoess, Ulf Matti, Christopher J Obara, Wesley R Legant, Anna Kreshuk, Jakob H Macke, Jonas Ries, and Srinivas C Turaga. Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature methods*, 18(9):1082–1090, 2021. 62

[55] Miguel A Boland, Jonathan PE Lightley, Edwin Garcia, Sunil Kumar, Chris Dunsby, Seth Flaxman, Mark AA Neil, Paul MW French, and Edward AK Cohen. Model-free machine learning-based 3d single molecule localisation microscopy. *Journal of Microscopy*, 299(1):77–87, 2025. 16

[56] Kim A Winick. Cramér–rao lower bounds on the performance of charge-coupled-device optical position estimators. *JOSA A*, 3(11):1809–1815, 1986. 16, 117

[57] Daniel Sage, Thanh-An Pham, Hazen Babcock, Tomas Lukes, Thomas Pengo, Jerry Chao, Ramraj Velmurugan, Alex Herbert, Anurag Agrawal, Silvia Colabrese, et al. Super-resolution fight club: assessment of 2d and 3d single-molecule localization microscopy software. *Nature methods*, 16(5):387–395, 2019. 17, 40, 54, 55, 72, 84

[58] Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15:1–28, 2015. 18

[59] Niccolo Banterle, Khanh Huy Bui, Edward A Lemke, and Martin Beck. Fourier ring correlation as a resolution criterion for super-resolution microscopy. *Journal of structural biology*, 183(3):363–367, 2013. 19, 70

[60] Robert PJ Nieuwenhuizen, Keith A Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature methods*, 10(6):557–562, 2013. 19, 70, 73

[61] Martin Ovesnỳ, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M Hagen. Thunderstorm: a comprehensive imagej plug-in for palm and storm data analysis and super-resolution imaging. *Bioinformatics*, 30(16):2389–2390, 2014. 19, 23, 89

[62] I Izeddin, J Boulanger, V Racine, CG Specht, A Kechkar, D Nair, A Triller, D Choquet, M Dahan, and JB Sibarita. Wavelet analysis for single molecule localization microscopy. *Optics express*, 20(3):2081–2095, 2012. 20

[63] Mehmet Sezgin and Bu¨ lent Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1):146–168, 2004. 20, 48

[64] Seamus J Holden, Stephan Uphoff, and Achillefs N Kapanidis. Daostorm: an algorithm for high-density super-resolution microscopy. *Nature methods*, 8(4):279–280, 2011. 20, 62

[65] Hazen Babcock, Yaron M Sigal, and Xiaowei Zhuang. A high-density 3d localization algorithm for stochastic optical reconstruction microscopy. *Optical nanoscopy*, 1:1–10, 2012. 62

[66] Lei Zhu, Wei Zhang, Daniel Elnatan, and Bo Huang. Faster storm using compressed sensing. *Nature methods*, 9(7):721–723, 2012. 62

[67] Eran A Mukamel, Hazen Babcock, and Xiaowei Zhuang. Statistical deconvolution for superresolution fluorescence microscopy. *Biophysical journal*, 102(10):2391–2400, 2012.

[68] Junhong Min, Cédric Vonesch, Hagai Kirshner, Lina Carlini, Nicolas Olivier, Seamus Holden, Suliana Manley, Jong Chul Ye, and Michael Unser. Falcon: fast and unbiased reconstruction of high-density super-resolution microscopy data. *Scientific reports*, 4(1):4577, 2014. 20, 62

[69] Jonas Ries. Smap: a modular super-resolution microscopy analysis platform for smlm data. *Nature methods*, 17(9):870–872, 2020. 23

[70] Pranjal Choudhury and Bosanta Ranjan Boruah. Localization and image reconstruction in a storm based super-resolution microscope. *Image Processing On Line*, 14:64–85, 2024. 23

[71] Arthur D Edelstein, Mark A Tsuchida, Nenad Amodaj, Henry Pinkard, Ronald D Vale, and Nico Stuurman. Advanced methods of microscope control using μmanager software. *Journal of biological methods*, 1(2):e10, 2014. 23, 24

[72] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. 24

[73] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. 24

[74] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in Python. *PeerJ*, 2:e453, 6 2014. 24

[75] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000. 24

[76] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007. 24

[77] Fang Huang, Tobias MP Hartwich, Felix E Rivera-Molina, Yu Lin, Whitney C Duim, Jane J Long, Pradeep D Uchil, Jordan R Myers, Michelle A Baird, Walther Mothes,

et al. Video-rate nanoscopy using scmos camera–specific single-molecule localization algorithms. *Nature methods*, 10(7):653–658, 2013. 26

[78] Alex Small and Shane Stahlheber. Fluorophore localization algorithms for super-resolution microscopy. *Nature methods*, 11(3):267–279, 2014. 26, 37, 40

[79] Tony Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318, 1993. 30

[80] Haskell B Curry. The method of steepest descent for non-linear minimization problems. *Quarterly of Applied Mathematics*, 2(3):258–261, 1944. 35

[81] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965. 35

[82] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978. 35

[83] Steven M Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993. 35

[84] Athanasios Papoulis and S Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002. 35

[85] Bin Yu, Danni Chen, Junle Qu, and Hanben Niu. Fast fourier domain localization algorithm of a single molecule with nanometer precision. *Optics letters*, 36(22):4317–4319, 2011. 36

[86] Koen JA Martens, Arjen N Bader, Sander Baas, Bernd Rieger, and Johannes Hohlbein. Phasor based single-molecule localization microscopy in 3d (psmlm-3d): An algorithm for mhz localization rates using standard cpus. *The Journal of chemical physics*, 148(12):123311, 2018. 36, 38

[87] Daniel Sage, Hagai Kirshner, Thomas Pengo, Nico Stuurman, Junhong Min, Suliana Manley, and Michael Unser. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature methods*, 12(8):717–724, 2015. 38, 40, 48, 73, 76

[88] David W Scott. Averaged shifted histograms: effective nonparametric density estimators in several dimensions. *The Annals of Statistics*, pages 1024–1040, 1985. 40

[89] Pranjal Choudhury, Amalesh Kumar, and Bosanta Ranjan Boruah. Implementation of real-time image reconstruction and visualization in single molecule localization microscopy. In *Advances in Microscopic Imaging V*, volume 13937, pages 164–166. SPIE, 2025. 42

[90] Daniel Burke, Brian Patton, Fang Huang, Joerg Bewersdorf, and Martin J Booth. Adaptive optics correction of specimen-induced aberrations in single-molecule switching microscopy. *Optica*, 2(2):177–185, 2015. 47

[91] Kayvan F Tehrani, Jianquan Xu, Yiwen Zhang, Ping Shen, and Peter Kner. Adaptive optics stochastic optical reconstruction microscopy (ao-storm) using a genetic algorithm. *Optics express*, 23(10):13677–13692, 2015. 47

[92] Leonhard Möckl, Anish R Roy, and WE Moerner. Deep learning in single-molecule microscopy: fundamentals, caveats, and recent developments. *Biomedical optics express*, 11(3):1633–1661, 2020. 47

[93] Pranjal Choudhury and Bosanta R Boruah. Adaptive image thresholding and localization of point spread functions with enhanced precision for single molecule localization based super-resolution microscopy. *Optics and Lasers in Engineering*, 178:108234, 2024. 47, 70

[94] Rafael C Gonzalez. *Digital image processing*. Pearson education india, 2009. 48, 81

[95] Prasanna K Sahoo, SAKC Soltani, and Andrew KC Wong. A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2):233–260, 1988. 48

[96] Linda Mahmoudi and Ali El Zaart. A survey of entropy image thresholding techniques. In *2012 2nd international conference on advances in computational tools for engineering applications (ACTEA)*, pages 204–209. IEEE, 2012. 48

[97] Julien Ghaye, Madhura Avinash Kamat, Linda Corbino-Giunta, Paolo Silacci, Guy Vergeres, Giovanni De Micheli, and Sandro Carrara. Image thresholding techniques for localization of sub-resolution fluorescent biomarkers. *Cytometry part a*, 83(11):1001–1016, 2013. 48

[98] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979. 48, 52

[99] Jagat Narain Kapur, Prasanna K Sahoo, and Andrew KC Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer vision, graphics, and image processing*, 29(3):273–285, 1985. 48, 51, 52

[100] Prasanna Sahoo, Carrye Wilkins, and Jerry Yeager. Threshold selection using renyi's entropy. *Pattern recognition*, 30(1):71–84, 1997. 48

[101] Chun Hung Li and CK Lee. Minimum cross entropy thresholding. *Pattern recognition*, 26(4):617–625, 1993. 48

[102] Wayne Niblack. *An introduction to digital image processing*. Strandberg Publishing Company, 1985. 48

[103] Jaakko Sauvola and Matti Pietikäinen. Adaptive document image binarization. *Pattern recognition*, 33(2):225–236, 2000. 48

[104] Nicolas Coudray, Jean-Luc Buessler, and Jean-Philippe Urban. Robust threshold estimation for images with unimodal histograms. *Pattern Recognition Letters*, 31(9):1010–1019, 2010. 49, 52

[105] Daniel Zwillinger and Stephen Kokoska. *CRC standard probability and statistics tables and formulae*. Crc Press, 1999. 50

[106] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948. 50

[107] Samy Sadek and Ayoub Al-Hamadi. Entropic image segmentation: a fuzzy approach based on tsallis entropy. *Int J Comput Vis Signal Process*, 5(1):1–7, 2015. 51

[108] Haniza Yazid, Shafriza Nisha Basah, Saufiah Abdul Rahim, Muhammad Juhairi Aziz Safar, and Khairul Salleh Basaruddin. Performance analysis of entropy thresholding for successful image segmentation. *Multimedia Tools and Applications*, 81(5):6433–6450, 2022. 51

[109] Ian Khaw, Benjamin Croop, Jialei Tang, Anna Möhl, Ulrike Fuchs, and Kyu Young Han. Flat-field illumination for quantitative fluorescence imaging. *Optics express*, 26(12):15276–15288, 2018. 52

[110] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. Ieee, 2003. 53, 73

[111] Super-resolution microscopy hub | dataset. https://srm.epfl.ch/srm/dataset/challenge-3D-simulation/MT1.N1.LD/index.html, 2016. Accessed: 2024-03-12. 55, 73, 74

[112] Sripad Ram, E Sally Ward, and Raimund J Ober. Beyond rayleigh's criterion: a resolution measure with application to single-molecule microscopy. *Proceedings of the National Academy of Sciences*, 103(12):4457–4462, 2006. 61

[113] Hendrik Deschout, Francesca Cella Zanacchi, Michael Mlodzianoski, Alberto Diaspro, Joerg Bewersdorf, Samuel T Hess, and Kevin Braeckmans. Precisely and accurately localizing single emitters in fluorescence microscopy. *Nature methods*, 11(3):253–266, 2014. 62

[114] Wei Ouyang, Andrey Aristov, Mickaël Lelek, Xian Hao, and Christophe Zimmer. Deep learning massively accelerates super-resolution localization microscopy. *Nature biotechnology*, 36(5):460–468, 2018. 62

[115] Elias Nehme, Lucien E Weiss, Tomer Michaeli, and Yoav Shechtman. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica*, 5(4):458–464, 2018. 62, 64

[116] Pasquale Cascarano, Maria Colomba Comes, Andrea Sebastiani, Arianna Mencattini, Elena Loli Piccolomini, and Eugenio Martinelli. Deepcel0 for 2d single-molecule localization in fluorescence microscopy. *Bioinformatics*, 38(5):1411–1419, 2022. 62

[117] Pranjal Choudhury and Bosanta R Boruah. Neural network-assisted localization of clustered point spread functions in single-molecule localization microscopy. *Journal of Microscopy*, 297(2):153–164, 2025. 62

[118] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings, 2011. 63

[119] Andrea Apicella, Francesco Donnarumma, Francesco Isgrò, and Roberto Prevete. A survey on modern trainable activation functions. *Neural Networks*, 138:14–32, 2021. 63

[120] Shiv Ram Dubey, Satish Kumar Singh, and Bidyut Baran Chaudhuri. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503:92–108, 2022. 63

[121] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016. 63

[122] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009. 64, 65

[123] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 65

[124] Florian Kromp, Lukas Fischer, Eva Bozsaky, Inge M Ambros, Wolfgang Dörr, Klaus Beiske, Peter F Ambros, Allan Hanbury, and Sabine Taschner-Mandl. Evaluation of deep learning architectures for complex immunofluorescence nuclear image segmentation. *IEEE Transactions on Medical Imaging*, 40(7):1934–1949, 2021. 65

[125] Hassan Abdallah, Brent Formosa, Asiri Liyanaarachchi, Maranda Saigh, Samantha Silvers, Suzan Arslanturk, Douglas J Taatjes, Lars Larsson, Bhanu P Jena, and Domenico L Gatti. Res-cr-net, a residual network with a novel architecture optimized for the semantic segmentation of microscopy images. *Machine Learning: Science and Technology*, 1(4):045004, 2020.

[126] Jiji Chen, Hideki Sasaki, Hoyin Lai, Yijun Su, Jiamin Liu, Yicong Wu, Alexander Zhovmer, Christian A Combs, Ivan Rey-Suarez, Hung-Yu Chang, et al. Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nature methods*, 18(6):678–687, 2021.

[127] Zhiwei Zhou, Weibing Kuang, Zhengxia Wang, and Zhen-Li Huang. Resnet-based image inpainting method for enhancing the imaging speed of single molecule localization microscopy. *Optics Express*, 30(18):31766–31784, 2022. 65

[128] François Chollet et al. Keras. https://keras.io, 2015. 65

[129] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pages 265–283, 2016. 65

[130] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. pmlr, 2015. 65

[131] Erich L Lehmann and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006. 66

[132] Donald F Specht et al. A general regression neural network. *IEEE transactions on neural networks*, 2(6):568–576, 1991. 66

[133] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 67

[134] Robert J Valkenburg, Alan M McIvor, and P Wayne Power. Evaluation of subpixel feature localization methods for precision measurement. In *Videometrics III*, volume 2350, pages 229–238. SPIE, 1994. 69

[135] A Patwardhan. Subpixel position measurement using 1d, 2d and 3d centroid algorithms with emphasis on applications in confocal microscopy. *Journal of Microscopy*, 186(3):246–257, 1997. 69

[136] Kay Thwe Min Han and Bunyarit Uyyanonvara. A survey of blob detection algorithms for biomedical images. In *2016 7th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*, pages 57–60. IEEE, 2016. 70

[137] Debora Keller Nicolas Olivier and Suliana Manley. Dataset real_high_density | 2d, 2013. https://srm.epfl.ch/srm/dataset/challenge-2D-real/Real_High_Density/index.html [Accessed: 19 February 2024]. 76

[138] Lothar Schermelleh, Alexia Ferrand, Thomas Huser, Christian Eggeling, Markus Sauer, Oliver Biehlmaier, and Gregor PC Drummen. Super-resolution microscopy demystified. *Nature cell biology*, 21(1):72–84, 2019. 79

[139] Simao Coelho, Jongho Baek, Matthew S Graus, James M Halstead, Philip R Nicovich, Kristen Feher, Hetvi Gandhi, J Justin Gooding, and Katharina Gaus. Ultraprecise single-molecule localization microscopy enables in situ distance measurements in intact cells. *Science Advances*, 6(16):eaay8271, 2020. 79

[140] Alexander Balinovic, David Albrecht, and Ulrike Endesfelder. Spectrally red-shifted fluorescent fiducial markers for optimal drift correction in localization microscopy. *Journal of Physics D: Applied Physics*, 52(20):204002, 2019. 79

[141] Yina Wang, Joerg Schnitzbauer, Zhe Hu, Xueming Li, Yifan Cheng, Zhen-Li Huang, and Bo Huang. Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Optics express*, 22(13):15982–15991, 2014. 80

[142] Michael J Wester, David J Schodt, Hanieh Mazloom-Farsibaf, Mohamadreza Fazel, Sandeep Pallikkuth, and Keith A Lidke. Robust, fiducial-free drift correction for super-resolution imaging. *Scientific reports*, 11(1):23672, 2021. 80

[143] Hsieh Hou and H Andrews. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on acoustics, speech, and signal processing*, 26(6):508–517, 1978. 82

[144] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964. 83

[145] Jelmer Cnossen, Tao Ju Cui, Chirlmin Joo, and Carlas Smith. Drift correction in localization microscopy using entropy minimization. *Optics Express*, 29(18):27961–27974, 2021. 84

[146] Pranjal Choudhury and Bosanta Ranjan Boruah. Astigmatic vortex-beam-based autofocus system for super resolution microscopy. In *Advances in Microscopic Imaging V*, volume 13937, pages 129–131. SPIE, 2025. 95

[147] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013. 111

APPENDIX A
Theoretical Derivation of the Point Spread Function

Introduction

In high-resolution fluorescence microscopy, particularly in single-molecule localisation techniques such as STORM, the image of a single fluorophore is blurred by diffraction, forming a characteristic point spread function (PSF). The PSF governs the resolution of the imaging system and serves as the theoretical foundation for localisation and deconvolution algorithms.

While a full vectorial diffraction theory is required for a completely accurate description of the electromagnetic field for high numerical aperture (NA) objectives, the fundamental properties of the PSF can be effectively described using a scalar approximation. The Debye-Wolf integral [147], in its scalar form, provides an excellent and widely-used model for the field near the focal region. This appendix presents a full derivation of the PSF in both the lateral and axial directions using this scalar Debye-Wolf formalism.

The Debye-Wolf Integral

We begin by considering the Huygens-Fresnel principle as applied to a converging spherical wavefront W. The complex electric field $U(P)$ at a point P near the geometric focus is described by the integral,

$$U(P) = -\frac{i}{\lambda}\frac{Ae^{-ikf}}{f}\iint_W \frac{e^{iks}}{s}\,dS \tag{A.1}$$

In this expression, A is the constant amplitude of the wave on the spherical wavefront W, and f is the focal length, which corresponds to the radius of curvature of the wavefront. The term s represents the distance from a point Q on the wavefront to the observation point P, while $k = 2\pi/\lambda$ is the wave number corresponding to the wavelength λ of the light. The

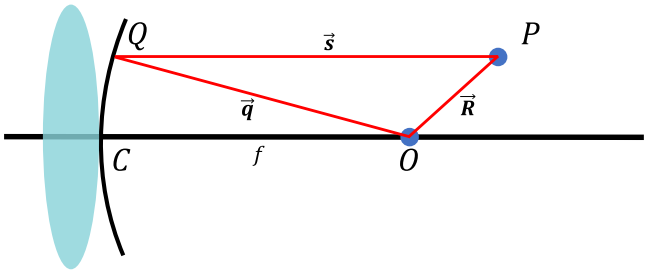

Figure A.1. Schematic illustrating the geometry used in deriving the Debye-Wolf integral. A converging spherical wavefront from a high-NA objective focuses at point O, with the observation point P located near the focal region. The vector $\vec{q}$ denotes the direction from the focus to a point Q on the wavefront, and $\vec{R}$ is the position vector of point P with respect to the focal point. The path length $s = QP$ is approximated using the dot product $\vec{q} \cdot \vec{R}$ for calculating the phase term in the integral.

exponential factor e^{-ikf}/f accounts for the phase and amplitude of the spherical wave at the reference sphere.

To simplify this integral for the region near the geometric focus, we consider approximations valid for observation points close to the origin. First, the denominator $1/s$ can be approximated by $1/f$ since $s \approx f$ near the focus. However, the phase term e^{iks} requires a more accurate treatment due to its sensitivity to small variations in path length. For this, we adopt the approximation

$$s \approx f - \vec{q} \cdot \vec{R} \tag{A.2}$$

Here, $\vec{q}$ is the unit vector directed from the focus O to a point Q on the wavefront, and $\vec{R}$ is the position vector of the observation point P relative to the focus. Substituting this into the expression for $U(P)$, we obtain

$$U(P) \approx -\frac{iAf}{\lambda} e^{-ikf} \iint_W e^{ik(f-\vec{q}\cdot\vec{R})}\, dS \tag{A.3}$$

The exponential can be separated into two factors: $e^{ik(f-\vec{q}\cdot\vec{R})} = e^{ikf} e^{-ik(\vec{q}\cdot\vec{R})}$. Substituting this back into the integral, the terms e^{-ikf} and e^{ikf} cancel, simplifying the expression to

$$U(P) = -\frac{iAf}{\lambda} \iint_W e^{-ik(\vec{q}\cdot\vec{R})}\, dS \tag{A.4}$$

At this point, the surface integral over the physical wavefront area W is transformed into an integral over the solid angle Ω that the wavefront subtends as seen from the focus. Since the surface area element on a sphere of radius f is $dS = f^2 d\Omega$, we substitute and obtain

$$U(P) = -\frac{iAf}{\lambda} \iint_\Omega e^{-ik(\vec{q}\cdot\vec{R})} f^2\, d\Omega = -\frac{iAf^3}{\lambda} \iint_\Omega e^{-ik(\vec{q}\cdot\vec{R})}\, d\Omega \tag{A.5}$$

To simplify notation, we define a constant $C = -\frac{iAf^3}{\lambda}$, and thus express the Debye-Wolf integral in its compact form, which is given as,

$$U(P) = C \iint_{\Omega} e^{-ik(\vec{q}\cdot\vec{R})}\, d\Omega \tag{A.6}$$

To obtain an explicit form of this integral, we represent the observation point P in cylindrical coordinates as $\vec{R} = (r_p \cos\psi_p, r_p \sin\psi_p, z_p)$, where r_p is the radial distance from the optical axis, ψ_p is the azimuthal angle, and z_p is the axial coordinate. A point Q on the spherical wavefront is described using spherical coordinates (θ, ϕ), where θ is the polar angle (measured from the optical axis) and ϕ is the azimuthal angle. The unit vector $\vec{q}$ in the direction from the focus to point Q is given by $\vec{q} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$.

The dot product $\vec{q}\cdot\vec{R}$ is then evaluated as,

$$\begin{aligned}\vec{q}\cdot\vec{R} &= r_p\cos\psi_p\sin\theta\cos\phi + r_p\sin\psi_p\sin\theta\sin\phi + z_p\cos\theta \\ &= r_p\sin\theta\left(\cos\psi_p\cos\phi + \sin\psi_p\sin\phi\right) + z_p\cos\theta \\ &= r_p\sin\theta\cos(\phi-\psi_p) + z_p\cos\theta\end{aligned} \tag{A.7}$$

Substituting this into the Debye integral, and using $d\Omega = \sin\theta\, d\theta\, d\phi$, the expression for the field becomes

$$U(r_p, z_p) = C\int_0^{\alpha}\int_0^{2\pi} e^{-ik[r_p\sin\theta\cos(\phi-\psi_p)+z_p\cos\theta]}\sin\theta\, d\theta\, d\phi \tag{A.8}$$

This final expression represents the electric field at any point (r_p, z_p) near the focus of a high numerical aperture system, and forms the basis for further derivation of lateral and axial PSF models.

Lateral PSF in the Focal Plane

To derive the lateral PSF, we set $z_p = 0$, corresponding to the focal plane. The integral simplifies to

$$U(r_p, 0) = C\int_0^{\alpha}\int_0^{2\pi} e^{-ikr_p\sin\theta\cos(\phi-\psi_p)}\sin\theta\, d\theta\, d\phi \tag{A.9}$$

The inner integral over ϕ yields the zeroth-order Bessel function of the first kind using the identity

$$\int_0^{2\pi} e^{-ix\cos(\phi-\psi_p)}\, d\phi = 2\pi J_0(x) \tag{A.10}$$

Substituting $x = kr_p \sin\theta$, we get

$$U(r_p, 0) = 2\pi C \int_0^{\alpha} J_0(kr_p \sin\theta) \sin\theta \, d\theta \tag{A.11}$$

In the paraxial limit, or assuming a uniform wavefront, this yields the first-order Bessel function. Defining the dimensionless lateral coordinate $v = kr_p \sin\alpha = \frac{2\pi NA}{\lambda} r_p$, the normalised intensity is the Airy pattern

$$I(v) = I_0 \left[\frac{2J_1(v)}{v}\right]^2 \tag{A.12}$$

Axial PSF Along the Optical Axis

To accurately model the axial PSF for a high-NA system (e.g., $NA = 1.4$), the scalar Debye-Wolf integral must include the apodization factor $\sqrt{\cos\theta}$, representing the Abbe sine condition. Setting $r_p = 0$, the field along the optical axis is

$$U(0, z_p) = 2\pi C \int_0^{\alpha} \sqrt{\cos\theta} e^{-ikz_p \cos\theta} \sin\theta \, d\theta \tag{A.13}$$

We change variables using $w = \cos\theta$, with $dw = -\sin\theta d\theta$. The limits transform from $\theta \in [0, \alpha]$ to $w \in [1, \cos\alpha]$,

$$U(0, z_p) = 2\pi C \int_{\cos\alpha}^{1} \sqrt{w} e^{-ikz_p w} \, dw \tag{A.14}$$

Unlike the paraxial case, this integral involving $\sqrt{w}$ does not result in a simple sinc function. However, the paraxial approximation (ignoring apodization where $\sqrt{w} \approx 1$) serves as a useful analytical baseline. In that limit, the integration yields

$$U(0, z_p) \approx \frac{2\pi C}{ikz_p} \left[e^{-ikz_p \cos\alpha} - e^{-ikz_p}\right] \tag{A.15}$$

The resulting axial intensity distribution is,

$$I(0, z_p) = I_0 \left[\frac{\sin\left(\frac{kz_p(1-\cos\alpha)}{2}\right)}{\frac{kz_p(1-\cos\alpha)}{2}}\right]^2 \tag{A.16}$$

Introducing axial coordinate $u = \frac{2\pi n z_p}{\lambda} \sin^2(\alpha/2)$, we obtain

$$I(z_p) = I_0 \mathrm{sinc}^2(u) \tag{A.17}$$

In experimental widefield imaging with $NA = 1.4$, the axial profile often deviates from this ideal symmetric form. Refractive index mismatches between the immersion oil ($n \approx 1.515$)

and the sample medium introduce spherical aberration, which manifests as an additional phase term $\Phi(\theta)$ in the exponent of the Debye-Wolf integral. This leads to the characteristic axial elongation and asymmetry frequently observed in single-molecule localisation microscopy.

The results of the scalar diffraction theory are visualised in Figure A.2, which shows the distinct structure of the PSF in both the lateral and axial planes.

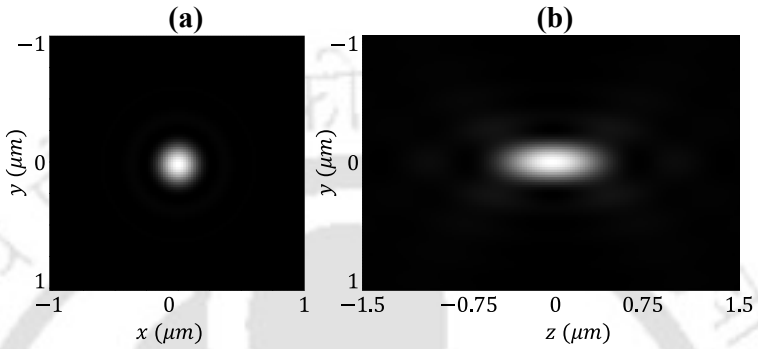

Figure A.2. The theoretical PSF as predicted by scalar diffraction theory. (a) The lateral (x-y) plane, showing the characteristic Airy pattern. (b) The axial (z-y) plane, showing the elongation along the optical axis.

APPENDIX B
Derivation of Localization Precision in SMLM

Cramér–Rao Lower Bound

The Cramér–Rao Lower Bound (CRLB) provides a fundamental theoretical limit on the variance of any unbiased estimator of a parameter [56]. In the context of SMLM, the CRLB defines the best possible precision with which the position of an isolated fluorophore can be determined, given the photon statistics, detector properties, and background noise. It serves as a benchmark for evaluating the performance of localisation algorithms and for optimising imaging system design.

Model for a Pixelated Gaussian PSF

Consider a fluorescent emitter located at position (x_0, y_0). Due to diffraction, the emitted photons form a PSF on the detector that can be modeled by a two-dimensional Gaussian distribution. The expected number of photons detected in a pixel with coordinates (j, k) and pixel size a is given by integrating the Gaussian function over the area of the pixel as,

$$E_{j,k}(\theta) = \frac{N}{2\pi\sigma^2} \int_{ja}^{(j+1)a} \int_{ka}^{(k+1)a} \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}\right) dx\, dy + b \tag{B.1}$$

Here, N denotes the total number of detected signal photons, σ is the standard deviation of the PSF, and a represents the pixel size. The term b corresponds to the mean number of background photons per pixel. The parameter vector is $\theta = (N, x_0, y_0, \sigma, b)$.

When the PSF is significantly larger than the pixel size ($\sigma \gg a$), this integral can be accurately approximated by evaluating the Gaussian at the pixel's center and scaling by the pixel area as,

$$E_{j,k}(\theta) \approx \frac{Na^2}{2\pi\sigma^2} \exp\left(-\frac{((j+0.5)a - x_0)^2 + ((k+0.5)a - y_0)^2}{2\sigma^2}\right) + b \tag{B.2}$$

Fisher Information Matrix (FIM)

The Fisher Information Matrix (FIM) quantifies the information that the measured pixel values carry about the unknown parameter vector θ. For photon detection, which follows Poisson statistics, the elements of the FIM are given by

$$F_{ij} = \sum_{j,k} \frac{1}{E_{j,k}(\theta)} \frac{\partial E_{j,k}(\theta)}{\partial \theta_i} \frac{\partial E_{j,k}(\theta)}{\partial \theta_j} \tag{B.3}$$

The CRLB states that the variance of any unbiased estimator $\hat{\theta}_i$ is bounded by the inverse of the FIM:

$$\text{var}(\hat{\theta}_i) \geq [F^{-1}]_{ii} \tag{B.4}$$

We are interested in the precision of estimating the position x_0. To find the corresponding FIM element F_{x_0,x_0}, we first compute the partial derivative of Equation B.2 with respect to x_0 is

$$\frac{\partial E_{j,k}}{\partial x_0} = \underbrace{\frac{Na^2}{2\pi\sigma^2} \exp\left(-\frac{((j+0.5)a - x_0)^2 + ((k+0.5)a - y_0)^2}{2\sigma^2}\right)}_{G_{j,k}} \cdot \frac{((j+0.5)a) - x_0}{\sigma^2} \tag{B.5}$$

Letting $E_{j,k} = G_{j,k} + b$, the FIM element for x_0 becomes

$$F_{x_0,x_0} = \sum_{j,k} \frac{G_{j,k}^2}{G_{j,k} + b} \cdot \frac{((j+0.5)a - x_0)^2}{\sigma^4} \tag{B.6}$$

Analytical Approximation of the FIM

Equation B.6 is the exact form of the FIM element and is typically evaluated numerically. However, an analytical approximation can be derived by converting the discrete sum over pixels into a continuous integral over space, under the assumption that the pixel size is small ($\sigma \gg a$). The summation is replaced by an integral scaled by the pixel area a^2 as,

$$F_{x_0,x_0} \approx \frac{1}{a^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{G(x,y)^2}{G(x,y) + b} \frac{(x - x_0)^2}{\sigma^4} \, dx \, dy \tag{B.7}$$

where $G(x, y)$ is the continuous form of the Gaussian signal. Solving this integral is non-trivial, but its solution shows that the total information is a combination of information from different noise sources. The inverse of this FIM element, which gives the variance, is

therefore a sum of variance contributions from each source. The detailed derivation leads to three primary terms that constitute the final expression for localisation variance.

Final Expression for Localisation Precision

By performing the integration in Equation B.7 and adding a term for pixelation noise, the CRLB for the variance in the x_0 position is approximated as,

$$\text{var}(\hat{x}_0) \geq \frac{\sigma^2 + a^2/12}{N} + \frac{8\pi\sigma^4 b^2}{a^2 N^2} \tag{B.8}$$

The expression for the variance in y_0 is identical. The localisation precision is the standard deviation, obtained by taking the square root as,

$$\sigma_{\text{SMLM}} \geq \sqrt{\frac{\sigma^2 + a^2/12}{N} + \frac{8\pi\sigma^4 b^2}{a^2 N^2}} \tag{B.9}$$

Each term in this expression corresponds to a specific source of uncertainty:

- $\frac{\sigma^2}{N}$: This term represents the uncertainty due to photon shot noise. Precision improves as the number of collected photons (N) increases.
- $\frac{a^2/12}{N}$: This term arises from the pixelation of the detector. It represents the quantisation error from assigning a continuous signal to discrete pixels and is minimised by using smaller pixels.
- $\frac{8\pi\sigma^4 b^2}{a^2 N^2}$: This term represents the uncertainty contribution from background photons. Its impact is most significant in low signal-to-noise ratio (SNR) conditions.

Simplified Case

In an ideal, background-free ($b = 0$) scenario with infinitely small pixels ($a = 0$), the expression for localisation precision simplifies to its most fundamental form as,

$$\sigma_{\text{SMLM}} = \frac{\sigma}{\sqrt{N}} \tag{B.10}$$

This highlights the direct relationship between the size of the PSF (σ) and the number of collected photons (N) in determining the ultimate limit of localisation precision.

APPENDIX C
Noise Models Relevant to SMLM

Introduction

Single molecule localisation microscopy (SMLM) operates under extremely low-light conditions and relies on accurately detecting faint PSFs emitted by individual fluorophores. The localisation accuracy and precision are strongly influenced by various sources of noise in the imaging system. This appendix outlines the primary noise models relevant to SMLM image formation, discussing their statistical nature and visual impact through illustrative simulations.

Poisson (Photon Shot) Noise

Photon shot noise is a fundamental type of noise that arises from the discrete and stochastic nature of photon emission and detection. It follows a Poisson distribution, meaning that the variance of the detected signal is equal to its mean. As a result, brighter regions in an image (those with higher photon counts) exhibit more absolute noise, while darker regions show less. However, the signal-to-noise ratio (SNR) improves with increasing signal because it scales with the square root of the photon count. The discrete nature of photon arrivals causes intensity fluctuations in each pixel, and for a pixel with an expected value λ, the probability of detecting k photons is given by the Poisson distribution as,

$$P(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{C.1}$$

In SMLM, photon counts from individual fluorophores often range from a few hundred to a few thousand, making shot noise a dominant factor that limits localisation precision. Figure C.1(b) demonstrates the effect of Poisson noise on a simulated SMLM image, revealing how noise levels vary with signal intensity.

Gaussian (Electronic Readout) Noise

Gaussian noise, in contrast, originates from the electronic processes involved in image acquisition and is largely independent of the light signal. This type of noise is additive and normally distributed, typically modeled with zero mean and constant variance. It includes contributions from readout noise which is introduced during the conversion of charge to a digital signal by the camera electronics, and thermal noise or dark current, which results from thermally generated electrons in the detector even in the absence of light. Although thermal noise is fundamentally Poissonian, for the short exposure times typical in SMLM, its effect is often well approximated by Gaussian noise. Gaussian noise affects all parts of the image uniformly, regardless of local intensity, and its presence is particularly significant in low-signal scenarios. Figure C.1(c) shows an example of an SMLM image affected by Gaussian noise, illustrating its spatially uniform appearance.

Mixed Poisson-Gaussian Noise

In practice, SMLM images are simultaneously affected by both Poisson and Gaussian noise components. The resulting noise is therefore a combination of signal-dependent and signal-independent elements. A realistic model of SMLM image formation incorporates this mixture, where the total signal is first corrupted by Poisson noise reflecting photon statistics, and then further perturbed by additive Gaussian readout noise. This can be mathematically represented as,

$$I(x, y) = N_P(S(x, y) + B_{\text{photon}}) + N_G(0, \sigma^2_{\text{readout}}) \tag{C.2}$$

Here, $S(x, y)$ denotes the expected fluorescence signal, while B_{photon} is the average background photon count, which also contributes to Poisson noise. The operator $N_P(\cdot)$ applies the Poisson process to simulate the stochastic nature of photon arrival, and $N_G(0, \sigma^2_{\text{readout}})$ represents zero-mean Gaussian noise introduced during the readout stage. This combination closely resembles the actual noise observed in experimental SMLM images. Algorithms that aim to localize fluorophores with high precision must be robust against this mixed noise model. Figure C.1(d) shows a simulated example of an SMLM image corrupted by both types of noise, while Figure C.1(e) depicts only background noise under the same mixed model for comparison.

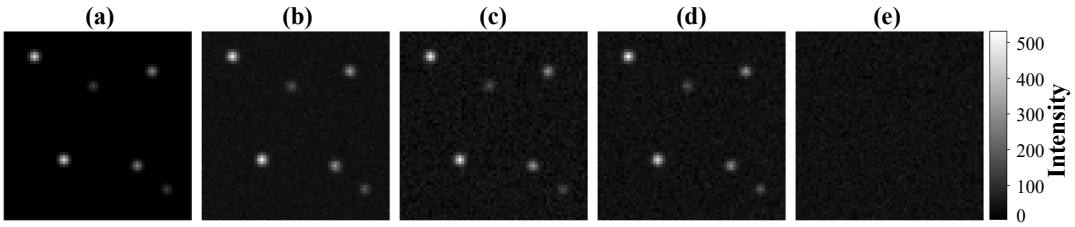

Figure C.1. Simulated SMLM images illustrating different noise models: (a) Clean PSFs without noise. (b) PSFs with Poisson noise (signal-dependent). (c) PSFs with Gaussian noise (uniform across the image). (d) Mixed Poisson-Gaussian noise showing realistic conditions. (e) Mixed noise applied to background only.

Effect of Photon Count on Signal-to-Noise Ratio

SNR is a fundamental metric that quantifies the clarity of a signal in the presence of noise. In the context of SMLM, it determines how well the signal from a fluorescent molecule's PSF can be distinguished from various noise sources, such as photon shot noise and camera readout noise. A higher SNR implies more precise and reliable localisation.

The SNR for an imaging system affected by both Poisson (photon) noise and Gaussian (electronic) noise is typically defined as,

$$\text{SNR} = \frac{N}{\sqrt{N + \sigma^2_{\text{readout}}}} \tag{C.3}$$

where N denotes the number of detected signal photons, and $\sigma^2_{\text{readout}}$ is the variance of the Gaussian (readout) noise.

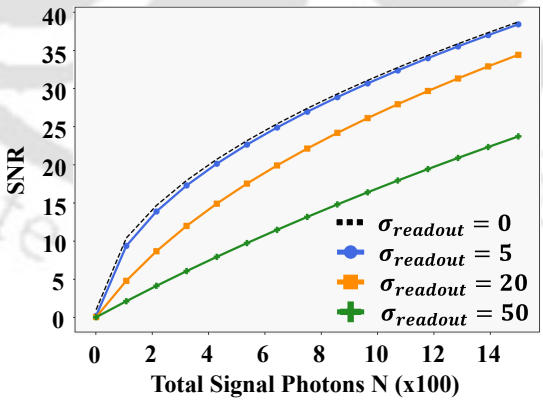

Figure C.2. Plot of Signal-to-Noise Ratio (SNR) as a function of the total number of signal photons (N), under varying Gaussian noise levels. The dashed line shows the ideal Poisson-limited case. At low photon counts, higher readout noise significantly lowers the SNR, impacting localisation precision in SMLM.

In the ideal case of a Poisson-limited system with negligible readout noise ($\sigma_{\text{readout}} \approx 0$), the SNR simplifies to $\sqrt{N}$, reflecting that the noise is solely governed by photon statistics.

However, in real-world imaging systems, especially under low-light conditions common in SMLM, electronic noise can significantly degrade the SNR. As the number of detected photons increases, the effect of Gaussian noise becomes less significant, and the system approaches the Poisson-limited regime.

Figure C.2 illustrates how SNR varies with increasing signal photons under different levels of Gaussian noise. It is evident that higher readout noise leads to lower SNR, especially at low photon counts, thereby affecting the accuracy of PSF localisation. At high photon counts, the differences between curves diminish, and SNR approaches the theoretical maximum.

APPENDIX D
Numerical Implementation of Key Algorithms

This appendix contains the full Python implementations of the key algorithms discussed in this thesis. The code is organised to correspond with the methods presented in Chapter 2 and Chapter 3. All implementations rely on the core scientific Python libraries including NumPy for array manipulation, SciPy for optimisation and image processing, and Scikit-image for feature detection.

All the codes developed in this thesis are hosted on Github at https://github.com/Pranjal264.

Image Preprocessing and Noise Filtering

This section provides the Python code for the image preprocessing filters discussed in Chapter 2, section 2.3. Each function takes a 2D NumPy array as input and returns the filtered image.

Mean Filtering

Code Snippet D.1. Python implementation of the Mean filter using SciPy.

```
import numpy as np
from scipy.ndimage import uniform_filter

def mean_filter(image: np.ndarray, kernel_size: int) -> np.ndarray:
    """
    Applies a mean (or box) filter to an image.
    Args:
        image:       The input image as a NumPy array.
        kernel_size: The size of the square kernel window (e.g., 3 for a 3x3
                     window).
    Returns:
```

```
        The filtered image as a NumPy array.
    """
    #  scipy.ndimage.uniform_filter efficiently calculates the average of
    elements within a window of a given size. mode=`constant' handles
    boundaries.
    return uniform_filter(image, size=kernel_size, mode="constant", cval=0)
```

Gaussian Filtering

Code Snippet D.2. Python implementation of the Gaussian filter using SciPy.

```
import numpy as np
from scipy.ndimage import gaussian_filter

def apply_gaussian_filter(image: np.ndarray, sigma: float) -> np.ndarray:
    """
    Applies a Gaussian filter to an image.
    Args:
        image: The input image as a NumPy array.
        sigma: The standard deviation of the Gaussian kernel.
    Returns:
        The filtered image as a NumPy array.
    """
    #  scipy.ndimage.gaussian_filter convolves the image with a Gaussian
    kernel of the specified standard deviation.
    return gaussian_filter(image, sigma=sigma)
```

Laplacian of Gaussian (LoG) Filtering

Code Snippet D.3. Python implementation of the Laplacian of Gaussian (LoG) filter.

```
import numpy as np
from scipy.ndimage import gaussian_laplace

def apply_log_filter(image: np.ndarray, sigma: float) -> np.ndarray:
    """
    Applies a Laplacian of Gaussian (LoG) filter to an image.
    Args:
        image: The input image as a NumPy array.
        sigma: The standard deviation of the Gaussian component.
    Returns:
        The filtered image as a NumPy array.
    """
```

```
    #  scipy.ndimage.gaussian_laplace efficiently computes the LoG by
    convolving the image with the second derivative of a Gaussian.
    #  The result is often multiplied by -1 to make blobs appear as peaks.

    return -1 * gaussian_laplace(image, sigma=sigma)
```

Difference of Gaussian (DoG) Filtering

Code Snippet D.4. Python implementation of the Difference of Gaussian (DoG) filter.

```
import numpy as np
from scipy.ndimage import gaussian_filter

def apply_dog_filter(image: np.ndarray, sigma1: float, sigma2: float) ->
    np.ndarray:
    """
    Applies a Difference of Gaussian (DoG) filter to an image.
    Args:
        image:  The input image as a NumPy array.
        sigma1: The standard deviation of the narrower Gaussian.
        sigma2: The standard deviation of the wider Gaussian (sigma2 >
    sigma1).
    Returns:
        The filtered image as a NumPy array.
    """
    # Create the two blurred versions of the image
    gaussian1 = gaussian_filter(image, sigma=sigma1)
    gaussian2 = gaussian_filter(image, sigma=sigma2)
    # Subtract the wider blur from the narrower blur
    return (gaussian1 - gaussian2)
```

PSF Detection and Crude Localization

This section contains the implementations for the three crude PSF detection methods discussed in Chapter 2, Section 2.4.

Peak Local Maximum

Code Snippet D.5. Python implementation for Peak Local Maximum (PLM) detection.

```
import numpy as np
from skimage.feature import peak_local_max
```

```
def detect_plm(image: np.ndarray, threshold: float, psf_radius: int) ->
    tuple:
    """
    Detects probable PSF locations using skimage.feature.peak_local_max.
    This function implements the PLM algorithm.
    Args:
        image:      The input image as a NumPy array.
        threshold:  The absolute intensity threshold for a peak to be
                    considered.
        psf_radius: The minimum allowable distance between peaks, in pixels.
    Returns:
        A tuple (x_coords, y_coords) of the detected PSF coordinates.
    """
    #  Find coordinates of local maxima that are above a threshold and
    separated by at least psf_radius.
    coordinates = peak_local_max(image, min_distance=psf_radius,
                  threshold_abs=threshold)

    # peak_local_max returns (row, col) which corresponds to (y, x)
    y_coords, x_coords = coordinates.T
    return x_coords, y_coords
```

Centre of Mass

Code Snippet D.6. Python implementation for Centre of Mass (COM) detection.

```
import numpy as np
import scipy.ndimage as ndimage
from scipy.ndimage import maximum_filter, minimum_filter

def _find_peaks_with_contrast(data: np.ndarray, neighborhood_size: int,
    threshold: float):
    """Helper function to find peaks with sufficient contrast."""
    data_max = maximum_filter(data, neighborhood_size)
    maxima   = (data == data_max)
    data_min = minimum_filter(data, neighborhood_size)

    # A peak is valid if its intensity contrast exceeds the threshold
    diff = (data_max - data_min) > threshold
    maxima[diff == 0] = 0
    return ndimage.label(maxima)
```

```
def detect_com(image: np.ndarray, threshold: float, neighborhood_size: int) -
    > tuple:
    """
    Estimates PSF positions using the Centre of Mass (COM) method.
    Args:
        image:      The input image as a NumPy array.
        threshold:  The minimum intensity contrast for a peak to be
                    considered.
        neighborhood_size: The size of the window for finding local extrema.
    Returns:
        A tuple (x_coords, y_coords) of the detected PSF centroids.
    """
    # 1. Find regions corresponding to valid peaks
    labeled_regions, num_objects = _find_peaks_with_contrast(
                                    image, neighborhood_size,
                                    threshold)
    if num_objects == 0:
        return np.array([]), np.array([])
    # 2. Calculate the center of mass for each labeled region
    coordinates =   np.array(ndimage.center_of_mass(image,
                    labeled_regions,
                    range(1, num_objects + 1)))
    # Coordinates are returned as (row, col) which is (y, x)
    y_coords, x_coords = coordinates.T
    return x_coords, y_coords
```

Blob Detection

Code Snippet D.7. Python implementation for LoG-based blob detection.

```
import numpy as np
from skimage.feature import blob_log

def detect_log_blob(image: np.ndarray, threshold: float, psf_radius: int) -
    > tuple:
    """
    Detects PSFs as blobs using the Laplacian of Gaussian (LoG) method.
    Args:
        image:      The input image as a NumPy array.
        threshold:  The detection threshold. Lower values find more, fainter
                    blobs.
        psf_radius: The maximum expected radius of a PSF in pixels. Used as
                    max_sigma.
    Returns:
```

```
        A tuple (x_coords, y_coords, sigmas) for the detected blobs.
    """
    # blob_log finds blobs of different sizes by applying LoG at multiple
    scales (sigmas).
    # min_sigma is set to 1 to avoid detecting single noisy pixels.
    blobs = blob_log(image, min_sigma=1, max_sigma=psf_radius,
            num_sigma=10, threshold=threshold)
    if blobs.size == 0:
        return np.array([]), np.array([]), np.array([])
    # Blobs are returned as (row, col, sigma) which is (y, x, radius)
    y_coords, x_coords, sigmas = blobs.T
    return x_coords, y_coords, sigmas
```

Sub Pixel Localization

This section details the code for the sub-pixel refinement algorithms, including the necessary helper functions and PSF models used for fitting.

2D Gaussian PSF model

Code Snippet D.8. Helper functions and symmetric 2D Gaussian model.

```
import numpy as np
from scipy.optimize import curve_fit, minimize
from scipy.fft import fft2
#2D gaussian PSF model
def _gaussian_2d(coords, A, x0, y0, sigma, C):
"""A symmetric 2D Gaussian function."""
        y, x = coords
        return A * np.exp(-((x - x0)**2 + (y - y0)**2) / (2 * sigma**2)) + C

#function to safely extract PSFs from widefield image
def _get_roi(image, r, c, half_box):
        """Extracts the ROI and checks if it is within image boundaries."""
        if (r - half_box < 0 or r + half_box + 1 > image.shape[0] or
            c - half_box < 0 or c + half_box + 1 > image.shape[1]):
                return None
        return image[r - half_box : r + half_box + 1,
                     c - half_box : c + half_box + 1].astype(np.float64)
```

Least squares fitting

Code Snippet D.9. Sub-pixel localization using Least Squares (LS) fitting.

```
def fit_ls(image, crude_coords, box_size=7):
"""Refines crude localizations using Least Squares fitting."""
        half_box = box_size // 2
        y_grid, x_grid = np.mgrid[:box_size, :box_size]
        refined_coords = []
        for r, c in crude_coords:
            roi = _get_roi(image, r, c, half_box)
            if roi is None: continue
            try:
            # Initial guess for the fit parameters
                initial_guess = [roi.max() - roi.min(), half_box, half_box,
                                 1.0, roi.min()]
                popt, _ = curve_fit(_gaussian_2d, (y_grid.ravel(),
                           x_grid.ravel()), roi.ravel(), p0=initial_guess)
                amp, fit_y, fit_x, fwhm, _ = popt
                # Quality control: check if the fit is near the center
                if abs(fit_x - half_box) < 1.5 and abs(fit_y - half_box)
                   < 1.5:
                    # Convert from ROI coordinates to global image
    coordinates
                        global_x = c - half_box + fit_x
                        global_y = r - half_box + fit_y
                        refined_coords.append((global_y, global_x))
            except (RuntimeError, ValueError):
                continue
        return np.array(refined_coords)
```

Maximum Likelihood Estimation

Code Snippet D.10. Sub-pixel localization using Maximum Likelihood Estimation (MLE).

```
def fit_mle(image, crude_coords, box_size=7):
"""Refines crude localizations using Maximum Likelihood Estimation."""
        half_box = box_size // 2
        y_grid, x_grid = np.mgrid[:box_size, :box_size]
        refined_coords = []

        def neg_log_likelihood(params, data):
            model = _gaussian_2d((y_grid.ravel(), x_grid.ravel()),
                       *params)
            model[model <= 0] = 1e-9 # Avoid log(0)
            return -np.sum(data * np.log(model) - model)
        for r, c in crude_coords:
```

```
            roi = _get_roi(image, r, c, half_box)
            if roi is None: continue
            initial_guess = [roi.max() - roi.min(), half_box, half_box, 1.0,
                             roi.min()]
            # Add bounds for stability
            bounds = [(0, None), (0, box_size), (0, box_size), (0.1,
                      half_box), (0, None)]
            result = minimize(neg_log_likelihood, initial_guess,
                     args=(roi.ravel(),), method='L-BFGS-B', bounds=bounds)

            if result.success:
                    amp, fit_y, fit_x, fwhm, _ = result.x
                    if abs(fit_x - half_box) < 1.5 and
                       abs(fit_y - half_box) < 1.5:
                        refined_coords.append((r - half_box + fit_y,
                        c - half_box + fit_x))
        return np.array(refined_coords)
```

Phasor Based Localization

Code Snippet D.11. Sub-pixel localization using Phasor-based phase analysis.

```
def fit_phasor(image, crude_coords, box_size=7):
"""Refines crude localizations using a fast, non-iterative Phasor method."""
    half_box = box_size // 2
    refined_coords = []
    for r, c in crude_coords:
       roi = _get_roi(image, r, c, half_box)
       if roi is None: continue
       try:
         fft_values = fft2(roi)
                window_pixel_size = roi.shape[0]
        # Extract and correct phase for X position
         angX = np.angle(fft_values[0, 1])
         if (angX > 0): angX = angX - 2 * np.pi
         fit_x = (abs(angX) / (2 * np.pi / window_pixel_size))
         # Extract and correct phase for Y position
         angY = np.angle(fft_values[1, 0])
         if (angY > 0): angY = angY - 2 * np.pi
         fit_y = (abs(angY) / (2 * np.pi / window_pixel_size))
         # Quality control and global coordinate conversion
         if abs(fit_x - half_box) < 1.5 and abs(fit_y - half_box)
            < 1.5:
             refined_coords.append((r - half_box + fit_y,
```

```
                c - half_box + fit_x))
            except:
                continue
    return np.array(refined_coords)
```

Adaptive entropy thresholding

Code Snippet D.12. Python implementation of the proposed adaptive entropy-based thresholding algorithm.

```
import numpy as np
from scipy import ndimage

def custom_threshold_optimized(image, alpha=0.01):
  """
  Calculate the optimal threshold using a vectorized custom entropy-based
    method.
        Args:
            image (numpy.ndarray): Input grayscale image. Must be of an
    integer type.
            alpha (float): Tolerance parameter for finding the
    threshold.
        Returns:
            optimal_threshold (int): Optimal threshold value for image
    segmentation.
  """
        # Ensure the image is not empty
        if image.size == 0:
            return 0
        # Use np.bincount for fast histogram calculation.
        image = image.astype(np.int32)  # Ensure integer type for bincount
        max_val = int(np.max(image))
        hist = np.bincount(image.ravel(), minlength=max_val + 1)
        # Normalize histogram to get probabilities
        hist_norm = hist / hist.sum()
        # Calculate the cumulative distribution function (CDF)
        P1 = np.cumsum(hist_norm)
        # Calculate entropy for thresholds from 0 to max_val-1
        threshold_values = np.arange(max_val)
        P1 = P1[threshold_values]
        # Probability of the second class (foreground)
        # Using np.maximum to avoid floating point inaccuracies close to 1.0
```

```
        P2 = np.maximum(1.0 - P1, 0)
        #compute all thresholds at once.
        log_arg1 = np.where(P1 > 0, P1, 1)
        log_arg2 = np.where(P2 > 0, P2, 1)
        entropy_values = -P1 * np.log(log_arg1) - P2 * np.log(log_arg2)
        if entropy_values.size == 0:
            return 0 # Handles images with max value of 0
        # Find the optimal threshold from the entropy values.
        mean_entropy = np.mean(entropy_values)
        candidate_indices = np.flatnonzero(np.abs(entropy_values -
                            mean_entropy) < alpha)
        # Add a fallback for cases where no threshold meets the criteria.
        if candidate_indices.size == 0:
            return np.argmin(np.abs(entropy_values - mean_entropy))
        # Return the last threshold that satisfies the condition.
        return candidate_indices[-1]
```

List of publications and other scientific activities

Publications leading to this thesis

1. **Pranjal Choudhury**, Anupam Bharadwaj and Bosanta Ranjan Boruah. "Real-Time Lateral Drift Stabilization in Single Molecule Localization Microscopy Using Live Image Re-Registration." (Manuscript under communication).

2. **Pranjal Choudhury** and Bosanta Ranjan Boruah. "Neural network-assisted localization of clustered point spread functions in single-molecule localization microscopy." Journal of Microscopy 297.2 (2025): 153-164.

3. **Pranjal Choudhury**, and Bosanta Ranjan Boruah "Adaptive image thresholding and localization of point spread functions with enhanced precision for single molecule localization based super-resolution microscopy." Optics and Lasers in Engineering 178 (2024): 108234.

4. **Pranjal Choudhury**, and Bosanta Ranjan Boruah. "Localization and Image Reconstruction in a STORM Based Super-resolution Microscope." Image Processing On Line 14 (2024): 64-85.

Other Publications

1. Nagendra Kumar, **Pranjal Choudhury**, and SS Goutam Buddha. "Deep learning-driven wavefront sensing for grating-array based wavefront sensor." IEEE Sensors Journal vol. 25, no. 3 (2025): 4769-4776.

2. Amalesh Kumar, Anupam Bharadwaj, **Pranjal Choudhury**, Sam P. Mathew, Bithiah Grace Jaganathan, and Bosanta Ranjan Boruah. "Tuning the excitation laser power in a stochastic optical reconstruction microscope for Alexa Fluor 647 dye in Vectashield mounting media." Review of Scientific Instruments 95, no. 8 (2024).

Conference Proceedings

1. **Pranjal Choudhury** and Bosanta Ranjan Boruah. "Structure Estimation in Single Molecule Localization Microscopy Using Parametric and Deep Learning Methods." In Frontiers in Optics + Laser Science 2025 (pp. JTu5A.56-1). Optica Publishing Group.

2. **Pranjal Choudhury** and Bosanta Ranjan Boruah. "Astigmatic vortex beam based autofocus system for super resolution microscopy." In European Conference on Biomedical Optics 2025 (pp. 129–131). SPIE and Optica Publishing Group.

3. **Pranjal Choudhury**, Amalesh Kumar and Bosanta Ranjan Boruah. "Implementation of real time image reconstruction and visualization in single molecule localization microscopy." In European Conference on Biomedical Optics 2025 (pp. 164–166). SPIE and Optica Publishing Group.

4. **Pranjal Choudhury**, Nagendra Kumar, Anupam Bharadwaj, and Bosanta Ranjan Boruah. "Advancing Super-Resolution Imaging: A Comprehensive Analysis of Gaussian Fitting in STORM Microscopy." In 2023 IEEE Workshop on Recent Advances in Photonics, pp. 1-3. IEEE, 2023.

5. Anupam Bharadwaj, **Pranjal Choudhury**, Ranjan Kalita, Amit Sharma, Amalesh Kumar, Bithiah Grace Jaganathan, Sunil Kumar, Jonathan Lightley, Edwin Garcia, P. M. W. French and Bosanta Ranjan Boruah . "A Cost-Effective Implementation of Optical Super-Resolution Microscopy with STORM." In International Conference on Photonics, (pp. 13-22). Springer Nature Singapore, 2023.

6. Nagendra Kumar, SS Goutam Buddha, **Pranjal Choudhury**, Karuna Sindhu Malik, and Santanu Konwar. "Fast and Efficient Wavefront Estimation Algorithm Using Slope Values in Zonal Wavefront Sensor." In 2023 IEEE Workshop on Recent Advances in Photonics, (pp. 1-3). IEEE, 2023.

7. Nagendra Kumar, **Pranjal Choudhury**, Nedup Sherpa, Akanshu Chauhan, SS Goutam Buddha, Karuna S. Malik, Santanu Konwar, and Bosanta Ranjan Boruah. "Improvement in wavefront measurement accuracy in grating array-based zonal wavefront sensors by modulating the intensity profile of the incident beam." In Practical Holography XXXVII: Displays, Materials, and Applications, vol. 12445, (pp. 173-180). SPIE, 2023.

8. Akanshu Chauhan, Nedup Sherpa, Nagendra Kumar, **Pranjal Choudhury**, SS Goutam Buddha, and Bosanta Ranjan Boruah. "Orthogonality of Zernike modes in phase profiles estimated using Zonal wavefront sensor and Transport of Intensity phase retrieval method." In Quantitative Phase Imaging IX, vol. 12389, (pp. 89-94). SPIE, 2023.

Conference Presentations

1. Structure Estimation in Single Molecule Localization Microscopy Using Parametric and Deep Learning Methods - Optica Fio+LS 2025, Denver, USA.
2. Astigmatic Vortex Beam-Based Autofocus System for Super Resolution Microscopy - Optica-SPIE ECBO 2025, Munich, Germany.
3. Implementation of Real-Time Image Reconstruction and Visualization in Single Molecule Localization Microscopy - Optica-SPIE ECBO 2025, Munich, Germany.
4. Convolutional Neural Network in Localization based Super-Resolution Optical Microscopy - PANE 2024, Tezpur University, India.
5. Image reconstruction modalities in a single molecule localization microscope - 2024, Imperial College London, UK.
6. Role of Gaussian Fitting in the image reconstruction in a STORM based super resolution microscope - PHOTONICS 2023, IISC Bangalore, India.
7. A python program for image reconstruction in a STORM based super resolution microscope - RIC 2023, IIT Guwahati, India.
8. Image reconstruction in a STORM based super resolution microscope - NCETP 2021, Tezpur University, India.

School/ Workshop attended

1. Participated in "IEEE Sensors Winter School" organised by Department of Electronics and Electrical Engineering Department, Dec 5-6, 2023 Indian Institute of Technology Guwahati, India.
2. Participated in "IEEE Frontiers in Optics and Terahertz Photonics" organised by IEEE Photonics Society Student Chapter, May 12, 2023, Indian Institute of Technology Guwahati, India.
3. Participated in the "5-day Online Workshop on Machine Learning & Deep Learning Techniques with Applications" conducted by the Centre for Intelligent Cyber Physical Systems and TIH Division IIT Guwahati Technology Innovation & Development Foundation, Jan 7-11,2023, Indian Institute of Technology Guwahati, India.
4. Participated in "Summer School on Optics & Photonics: Fundamentals and Applications" organised by Indian Institute of Technology Delhi Optics and Photonics Centre

and Delhi Chapter, National Academy of Sciences, India (NASI) June 6-24, 2022, Indian Institute of Technology Delhi.

Achievements

1. Awarded Optica Student Leadership 2025.
2. Awarded ANRF ITS to participate in ECBO 2025, Munich Germany.
3. Awarded Imperial College London Global Development Hub Fellowship 2024 to conduct research at Imperial College London as a visiting PhD student.
4. Received CSIR Junior Research Fellowship to pursue PhD.

* * *