UNIVERSITY OF GENOVA

PHD PROGRAM IN BIOENGINEERING AND ROBOTICS

Array detection enables large localization range for simple and robust MINFLUX

by

Sanket Patil

Thesis submitted for the degree of *Doctor of Philosophy* (38° cycle)

January 2026

Dr. Giuseppe Vicidomini	Supervisor
Dr. Eli Slenders	Supervisor
Prof. Paolo Massobrio	Head of the PhD program

Thesis Jury:

Dr. Luciano Masullo	External examiner
Dr. Takahiro Deguchi	External examiner

Dibris

Department of Informatics, Bioengineering, Robotics and Systems Engineering

"Wisdom is what I hate most, it is absolute conformist stupidity."

Slavoj Žižek

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgments. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Sanket Patil
March 2026

Acknowledgements

S P

Abstract

Fluorescence microscopy enables visualization of cellular structures with molecular specificity, yet conventional methods remain limited by diffraction to resolutions of $\sim 250\,\text{nm}$. Super-resolution microscopy overcomes this resolution barrier, with minimal fluorescence photon fluxes microscopy (MINFLUX) representing the current state-of-the-art in point scanning microscopy. MINFLUX probes single emitters with an intensity minimum in a targeted coordinate pattern (TCP) and typically achieves 1 nm to 3 nm localization uncertainty. However, this performance is coupled to a small localization range as the localization uncertainty grows rapidly when the emitter lies outside the TCP. Therefore MINFLUX requires a pre-localization step and iterative recentering to keep the emitter inside the TCP, combined with increasing laser power to ensure signal from the emitter. Here we present ***ISM-FLUX***, a technique that combines the MINFLUX concept with image scanning microscopy (ISM). ISM-FLUX combines orbital scanning of a donut-shaped excitation with a 5×5 single-photon avalanche diode (SPAD) array detector, extending the localization range to $\sim 600 \times 600\,\text{nm}^2$ while achieving 2 nm to 10 nm localization uncertainty with $\sim 10^3$ photons. Unlike single-element detection, the SPAD array detector provides spatial information of photons, resulting in camera-like information that prevents Cramér-Rao bound (CRB) divergence across the extended field. The orbital scanning geometry operates on standard galvanometric mirrors, requiring only a vortex phase plate and SPAD array to convert a confocal microscope into an ISM-FLUX system. We validated ISM-FLUX experimentally using DNA-origami nanorulers with 20 nm and 40 nm binding-site separations, resolving these distances even when the structures are positioned outside the TCP orbit. To support long acquisitions, we built a custom 3D active stabilization module that maintains sub-nanometer sample drift over hours in closed loop. The same platform integrates a co-registered wide-field single-molecule localization microscopy (SMLM) path for large field-of-view (FoV) context and correlative workflows, achieving $\sigma_{\text{NeNA}} \approx 7.83\,\text{nm}$. Together, these results show that ISM-FLUX delivers MINFLUX-style localization over an extended range within a point-scanning architecture. By simplifying MINFLUX microscopy and extending its capabilities within existing confocal infrastructure, ISM-FLUX provides an accessible path toward molecular-scale imaging that could accelerate broader adoption of single-molecule localization techniques in biological research.

Table of contents

List of figures

Acronyms

AOD	acousto-optic deflector
AOM	acousto-optic modulator
APD	avalanche photodiode
BFP	back focal plane
CCD	charge-coupled device
CMOS	complementary metal-oxide semiconductor
CRB	Cramér-Rao bound
DAC	digital-to-analog converter
DNA-PAINT	DNA-based points accumulation for imaging in nanoscale topography
FLIM	fluorescence-lifetime imaging microscopy
FoV	field-of-view
fPALM	fluorescence photoactivation localization microscopy
FPGA	field-programmable gate array
FWHM	full width at half maximum
GNP	gold nanoparticles
HILO	highly inclined laminated optical sheet
ISM	image scanning microscopy
MDF	molecule detection function
MINFLUX	minimal fluorescence photon fluxes microscopy
MINSTED	minimal STED
MLE	maximum likelihood estimator
NeNA	nearest-neighbor based analysis
OTF	optical transfer function
PAINT	points accumulation for imaging in nanoscale topography
PALM	photoactivated localization microscopy
PDE	photon detection efficiency

PMT	photomultiplier tube
PSF	point spread function
RCC	redundant cross-correlation
RESI	resolution enhancement by sequential imaging
RESOLFT	reversible saturable optical linear fluorescence transitions
ROSE	repetitive optical selective exposure
SBR	signal-to-background ratio
SIM	structured illumination microscopy
SMLM	single-molecule localization microscopy
SPAD	single-photon avalanche diode
SRM	super-resolution microscopy
STED	stimulated emission depletion microscopy
STORM	stochastic optical reconstruction microscopy
TCP	targeted coordinate pattern
TIRF	total internal reflection fluorescence

Chapter 1

Introduction

Cells are the basic units of biological organization that carry out complex molecular processes required for all lifeforms. These processes are comprised of cellular-scale interactions that influence fields such as molecular medicine, biotechnology, and astrobiology. Understanding these interactions aids therapeutic development, biomaterial design, and our understanding of biological mechanisms in different environments. However, understanding cellular architecture and dynamics has proven challenging due to the complexity of these systems. Cellular systems are complex because numerous components such as proteins, lipids, nucleic acids, etc., interact dynamically to sustain functional balance [1]. While researchers have developed multiple tools and techniques to probe molecular interactions, many cellular processes remain incompletely understood because they are susceptible to internal and external perturbations and because of limitations in available imaging technologies.

Modern analytical approaches have broadened the scope of cellular research. Modern omics technologies, including genomics, transcriptomics, and proteomics, can jointly profile DNA sequence, RNA abundance, protein levels, and metabolites from the same biological system, enabling integrated, cross-layer analysis [2, 3]. One such development is fluorescence microscopy combined with specific labeling strategies, which allows visualization and tracking of individual cells, organelles, and bio-molecules to understand their interactions and functions [4, 5]. The non-invasive nature of light, coupled with the target-specific fluorescent labeling of biological entities, offers higher contrast and greater molecular specificity than many other light-based methods. However, conventional optical microscopy has historically been limited by the diffraction barrier, restricting spatial resolution to approximately $\lambda/2$ ($\sim$250 nm laterally, $\sim$500 nm axially for visible light). This fundamental constraint, first described by Ernst Abbe in 1873 [6], restricted its utility primarily to either the bulk quantifi-

cation of molecules residing within this diffraction-limited volume or the analysis of sparsely expressed or labeled molecules that did not spatially overlap within such a confined region.

Recognition of this limitation drove the development of super-resolution microscopy techniques [7–9]. These advancements allowed subcellular studies at the single-molecule level using imaging techniques capable of achieving single-digit nanometer spatial resolution. In particular, single-molecule localization microscopy (SMLM) determines the positions of individual fluorescently labeled biomolecules with localization precision well below the diffraction limit. By sequentially localizing many such molecules, SMLM reconstructs images of cellular structures with molecular-scale resolution. A major recent advancement within the field of single-molecule localization microscopy is minimal fluorescence photon fluxes microscopy (MINFLUX) [10, 11], which attains nanoscale precision with high photon efficiency, representing the current state-of-the-art in super-resolution imaging.

This chapter focuses on the progression of microscopy from diffraction-limited fluorescence microscopy to MINFLUX-like nanoscopy by following the improvements in the optical resolution. Confocal microscopy provides the point-scanning reference architecture with optical sectioning and a well-defined effective (confocal) system point spread function (PSF), determined by the combination of excitation and detection PSFs [12–14]. Image scanning microscopy (ISM) then shows how the same scanning geometry can yield improved spatial detail by using detector arrays and photon reassignment, without requiring specialized fluorophore switching [15, 16]. Building on these concepts, super-resolution microscopy introduces approaches that surpass the diffraction limit by shaping the illumination or exploiting fluorophore state control and localization-based inference [7–9]. This framework introduces MINFLUX as a confocal-type localization method that uses structured excitation with intensity minima to localize single emitters with high photon efficiency [10, 11]. Overall, this work will concentrate on the single-molecule localization using MINFLUX-like microscopy techniques.

1.1 Confocal microscopy

Confocal microscopy arose from the idea that out-of-focus light can be rejected by coupling point illumination with a conjugate detection pinhole, thereby detecting only in focus fluorescence from the sample. Confocal microscopy achieves modest resolution improvement over widefield techniques by combining point illumination with pinhole-based detection, producing an effective PSF that is the product of illumination and detection PSFs [13]. The PSF is the intensity distribution in the image plane produced by an ideal point emitter; it is the system's

impulse response. For incoherent fluorescence imaging under uniform imaging conditions, the recorded image is the object convolved with the PSF [13, 14, 17, 18]. This multiplicative effect of the excitation and detection PSFs narrows the effective system PSF by approximately $\sqrt{2}$, yielding a theoretical lateral resolution improvement from ∼250 nm to ∼180 nm, though practical implementations typically achieve 200 nm to 250 nm [14, 19]. The gap between theoretical and practical resolution arises primarily from the trade-off between resolution and signal-to-noise ratio: while smaller pinholes (< 0.5 Airy units) maximize resolution, they also reject a substantial fraction of in-focus fluorescence, leading to unacceptably low photon counts and poor image quality [14]. Consequently, practical confocal imaging uses pinhole sizes around 1 A.U. (Airy unit), prioritizing adequate signal collection over maximal resolution gain. While this represents improvement over conventional widefield microscopy, confocal microscopy remains fundamentally diffraction-limited.

These confocal-like point scanning techniques use a single-point detector (avalanche photodiode (APD), photomultiplier tube (PMT), etc.) to detect photon signals as intensities and associate them with the positions defined by the scanning system. Marvin Minsky's 1957 patent introduced this double-focusing, stage-scanned concept, defining an early architecture for confocal imaging [12]. In a confocal laser-scanning microscope, a focused laser beam illuminates a diffraction-limited spot in the specimen. The beam is scanned point-by-point across the sample using galvanometric mirrors or other beam-steering optics, and fluorescence emission from each position is collected through the same objective lens (descanned configuration) or through a separate detection path. A conjugate pinhole placed at the image plane blocks out-of-focus light, passing only fluorescence originating from the focal volume, enabling optical sectioning. A schematic representation of a confocal laser scanning microscope is shown in Figure 1.1 (a). The detected photon signal at each scan position is registered and assembled into an image where pixel value reflects the fluorescence intensity at the corresponding spatial location in the specimen. This yields a lateral full width at half maximum (FWHM) value around 200 nm to 250 nm for visible light depending on the excitation wavelength and the numerical aperture ($\mathrm{FWHM}_{\perp} \approx \lambda/2\mathrm{NA}$)[14].

Practical high-speed scanning followed by Petrán and Egger's tandem-scanning (Nipkow disk) microscope in the late 1960s, enabling parallel confocal operation for reflected-light and fluorescence imaging [20, 21]. With stable lasers and galvanometric scanners, beam-scanning confocal instruments matured in the 1980s; Brakenhoff and colleagues quantified resolution gains (effective PSF narrowing of about 1.4×), and White, Amos, and Fordham demonstrated optical sectioning with higher contrast in thick fluorescent specimens [19, 22]. These developments were systematized by Wilson, Sheppard, and later Pawley, whose

theoretical and practical treatments summarized confocal performance and trade-offs for biological imaging [13, 14, 17].

Contemporary confocal implementations span spinning-disk and beam-scanning designs. Spinning-disk (Nipkow) systems with microlens-enhanced pinhole arrays (e.g., Yokogawa CSU series) provide high frame rates and low phototoxicity for live-cell imaging, and improved disk designs and optics reduce pinhole crosstalk and increase field-of-view (FoV) [23–25]. Beam-scanning confocals routinely use resonant scanners for video-rate sectioning, high-quantum-efficiency GaAsP/HyD detectors for sensitivity, spectral detection and unmixing for multicolor imaging, and computational deconvolution algorithms to enhance both lateral and axial resolution. However, the deconvolution effectiveness depends on adequate photon counts, reinforcing the need to balance pinhole size against signal collection [14, 22]. ISM addresses this limitation by replacing the single pinhole with a pixelated detector array, *i.e.*, camera. Each element here acts as a small pinhole, recovering photons that would otherwise be rejected.

1.2 Image scanning microscopy

Image scanning microscopy (ISM), a term coined by Mueller and Enderlein in 2010 [16], describes an imaging system that combines the camera-based detection of widefield techniques with the point-scanning architecture of laser-scanning microscopes. Early attempts to improve confocal resolution without sacrificing signal began with theoretical analyses of detector displacement and "pixel reassignment" by Sheppard in the late 1980s [15], followed by subsequent work examining how detector size affects resolution [26]. [16] showed that replacing the single point detector with a multi-element detector yields, at each scan position, a small detector-array image (*micro-image*) that can be shifted according to the reassignment rule and summed to form a higher-resolution image. Subsequent work by Sheppard and colleagues later provided experimental demonstrations and practical reassignment procedures consistent with this framework [27]. A schematic representation of an Image Scanning Microscope is provided in Figure 1.1 (b).

During an ISM acquisition the excitation focus is raster-scanned as in a classical confocal microscope, but the descanned fluorescence is recorded on a pixelated detector (charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) camera [28], GaAsP PMT [29], or, increasingly, single-photon avalanche diode (SPAD) arrays [30, 31]) that is positioned in the image (pinhole) plane. Because every detector element functions as an independent pinhole, the effective PSF of each micro-image is offset from the optical axis

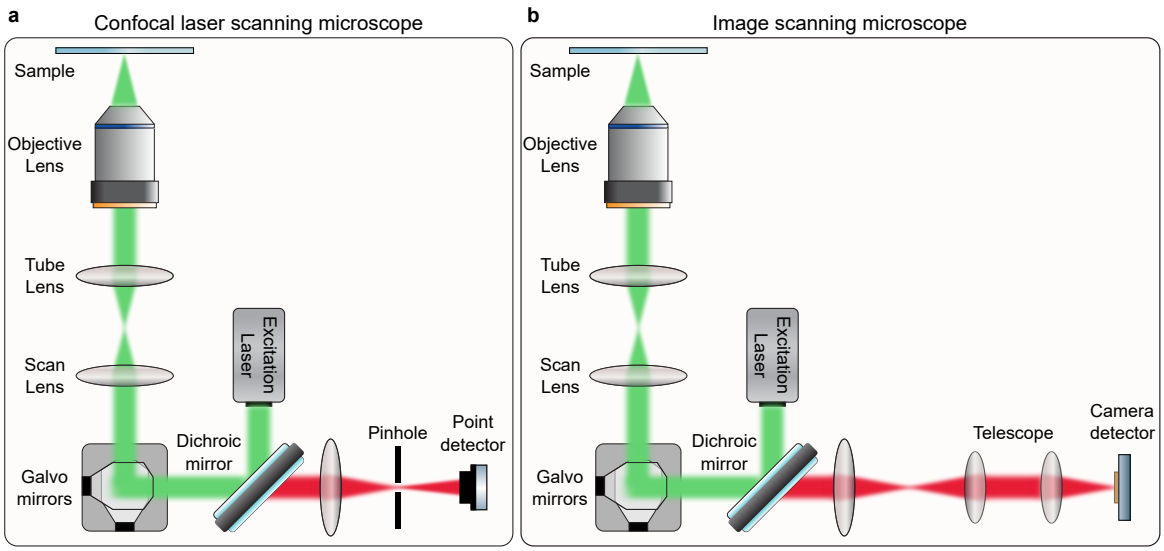

Figure 1.1 Schematic representation of Confocal laser-scanning microscope and image-scanning microscope. a) In a conventional confocal laser-scanning microscope, the excitation beam is raster-scanned across the sample and the emitted fluorescence is descanned and focused onto a conjugate pinhole, which rejects out-of-focus light before detection on a single-point detector. **b)** In an image-scanning microscope, the same scanning and descanning geometry is retained, but the pinhole is replaced by a detector array placed in a plane conjugate to the pinhole plane. A relay ($4f$) telescope images and magnifies the intermediate image onto the array, enabling the acquisition of a detector-element-resolved signal at each scan position for subsequent pixel reassignment and reconstruction.

in a known way. For each scan coordinate $\mathbf{r}_{\text{scan}}$ the signal arriving on detector pixel i carries an additional lateral offset $\mathbf{r}_i$ set by the imaging optics.

To understand the resolution gain, it is instructive to begin from the PSF of a conventional confocal microscope. In its general form, the detected signal depends on three contributions: the excitation PSF, the emission PSF, and the pinhole transmission function $p(\mathbf{r})$:

$$h_{\text{conf}}(\mathbf{r}) = h_{\text{ex}}(\mathbf{r}) \cdot \left[(h_{\text{em}} \otimes p)(\mathbf{r})\right] \tag{1.1}$$

where $h_{\text{ex}}(\mathbf{r})$ and $h_{\text{em}}(\mathbf{r})$ are the excitation and emission PSFs, $p(\mathbf{r})$ is the pinhole transmission function, and $\mathbf{r}$ is the lateral position vector in the sample plane. In the limit of a vanishingly small (closed) pinhole, $p(\mathbf{r}) \to \delta(\mathbf{r})$, the convolution reduces to the identity and Equation 1.1 simplifies to the product:

$$h_{\text{conf}}(\mathbf{r}) = h_{\text{ex}}(\mathbf{r}) \cdot h_{\text{em}}(\mathbf{r}) \tag{1.2}$$

Taking the Fourier transform of Equation 1.2, and noting that multiplication in real space corresponds to convolution in Fourier space, the corresponding optical transfer function (OTF) is:

$$H_{\text{conf}}(\mathbf{k}) = H_{\text{ex}}(\mathbf{k}) \otimes H_{\text{em}}(\mathbf{k}) \tag{1.3}$$

where H_{conf}, H_{ex}, and H_{em} are the OTFs corresponding to h_{conf}, h_{ex}, and h_{em} respectively, and $\mathbf{k}$ is the spatial frequency vector. Each individual OTF has a finite support with cutoff frequency $k_{\text{cut}} = 2\,\text{NA}/\lambda$, where NA is the numerical aperture of the objective and λ is the relevant wavelength. Since the convolution of two functions extends the support to the sum of their individual radii, and under the approximation $\lambda_{\text{ex}} \approx \lambda_{\text{em}} = \lambda$ (i.e., $h_{\text{ex}} \approx h_{\text{em}}$), the cutoff of H_{conf} extends to:

$$k_{\text{conf}}^{\max} = k_{\text{ex}} + k_{\text{em}} \approx \frac{4\,\text{NA}}{\lambda} = 2\,k_{\text{WF}} \tag{1.4}$$

where $k_{\text{ex}} = 2\,\text{NA}/\lambda_{\text{ex}}$ and $k_{\text{em}} = 2\,\text{NA}/\lambda_{\text{em}}$ are the cutoff frequencies of the excitation and emission OTFs respectively, and $k_{\text{WF}} = 2\,\text{NA}/\lambda$ is the cutoff frequency of a conventional widefield microscope. This doubling of the OTF support represents the hard physical limit on resolution imposed by diffraction.

This factor-of-2 frequency extension also has a direct counterpart in real space. Approximating both h_{ex} and h_{em} as Gaussians of equal standard deviation σ:

$$h_{\text{ex}}(\mathbf{r}) = \exp\left(-\frac{|\mathbf{r}|^2}{2\sigma^2}\right), \qquad h_{\text{em}}(\mathbf{r}) = \exp\left(-\frac{|\mathbf{r}|^2}{2\sigma^2}\right) \tag{1.5}$$

their product (Equation 1.2) is itself a Gaussian with a narrowed standard deviation σ_{conf}:

$$h_{\text{conf}}(\mathbf{r}) = \exp\left(-\frac{|\mathbf{r}|^2}{\sigma^2}\right) = \exp\left(-\frac{|\mathbf{r}|^2}{2\sigma_{\text{conf}}^2}\right), \qquad \sigma_{\text{conf}} = \frac{\sigma}{\sqrt{2}} \tag{1.6}$$

ISM recovers this same $\sqrt{2}$ lateral resolution improvement, but without the signal penalty of a physically closed pinhole. In ISM, each pixel of the array detector acts as an independent infinitesimal pinhole (approximated by a Dirac delta $\delta(\mathbf{r} - \mathbf{r}_i)$) so that the closed-pinhole condition is satisfied locally at every detector element simultaneously. Summing the reassigned images across all pixels therefore combines the resolution benefit of a closed pinhole with the photon collection efficiency of a fully open aperture, effectively eliminating the resolution–signal trade-off that limits classical confocal operation. Pixel reassignment exploits this directly: each detector pixel i records a sub-image that is shifted by half the known

detector offset $\mathbf{r}_i$ before summation [15]:

$$I_{\mathrm{ISM}}(\mathbf{r}) = \sum_i I_i\left(\mathbf{r} + \tfrac{1}{2}\mathbf{r}_i\right) \tag{1.7}$$

where I_i is the image recorded on detector pixel i, $\mathbf{r}_i$ is its lateral offset from the optical axis in sample-space units, and I_{ISM} is the final reconstructed ISM image. This reassignment and coherent summation contracts the effective PSF according, yielding a typical lateral resolution of 120 nm to 140 nm with improved signal recovery [16, 32]. When compared instead to practical 1-A.U. confocal operation ($\sim$200 nm to 250 nm), the same $\sqrt{2}$ scaling corresponds to $\sim$140 nm to 180 nm lateral FWHM under comparable optical conditions.

Compared with classic confocal microscopy operating at a conventional 1-Airy-unit pinhole, ISM offers: (i) $\sim 1.4\times$ lateral resolution enhancement for reassignment-based ISM; when the detector-array data are additionally processed with linear deconvolution/model-based reconstruction, this increases to $\sim 1.7\times$ (Airyscan-type processing) [33–35], with both values representing increasing approaches toward the $2\times$ diffraction-imposed ceiling derived above; (ii) up to 4–8$\times$ higher signal-to-background ratio (SBR) because the pinhole is effectively fixed to ~ 1 A.U.; (iii) intrinsic optical sectioning equivalent to confocal; and (iv) full compatibility with standard fluorophores, multi-colour imaging, two-photon excitation, fluorescence-lifetime imaging microscopy (FLIM) [31], and stimulated emission depletion microscopy (STED)-ISM hybrids for live-cell nanoscopy [36]. The technique also reduces photodose, which benefits long-term live imaging.

Early ISM prototypes relied on low-noise EM-CCDs, but the field has shifted to custom CMOS SPAD arrays that provide nanosecond timing, low dark counts, and multi-kilohertz read-out rates. Vicidomini and colleagues introduced a 5×5 SPAD-array ISM platform and demonstrated photon-resolved ISM (PR-ISM) with simultaneous fluorescence-lifetime imaging (FLISM) [31, 37]. Replacing cameras with multi-anode PMTs (Airyscan) or single-photon avalanche diode (SPAD) arrays has increased detection rates into the multi-MHz regime. Modern 1-MP CMOS SPAD arrays achieve ≥ 100 MHz aggregate count rates with on-chip gating and time-correlated read-out [38]. Subsequent work extended SPAD-ISM to two-photon excitation, background-rejection strategies, and time-correlated single-photon counting for dynamic studies of biomolecular condensates [39, 40]. Because every photon can be time-tagged with a few hundred picoseconds precision, SPAD-based ISM can multiplex structural imaging with functional readouts of lifetime, anisotropy, or fluctuation analysis while keeping the optical path unchanged.

ISM is now available in several commercial instruments and is an active research topic. In the commercial space, Zeiss introduced its Airyscan (2014) and Airyscan 2 (2019) systems, which uses a 32-channel hexagonal GaAsP array with proprietary reassignment to resolve structures ∼120 nm apart in the lateral dimension and provide a 4–8× signal-to-noise improvement [34, 35]. Hardware acceleration, such as the ZEISS Airyscan Fast mode, can quadruple scan speed through an elongated excitation line while retaining super-resolution capabilities [33]. Nikon introduced its N-SPARC detector [41] on the AX/AX R platforms, featuring a 25-element silicon-photomultiplier array with claimed resolutions of 100 nm XY and 300 nm Z resolution; in parallel, its NSPARC system (NIS-Elements 2023) integrates a 25-channel detector and a reconstruction pipeline to offer a 1.6× resolution improvement for large fields of view [42, 43].

Alongside these commercial developments, open-hardware projects have made ISM more accessible to researchers. Projects such as simpleISM provide 3D-printed adapter modules and FPGA-based processing solutions that enable retrofitting of existing confocal platforms [44], while BrightEyes-ISM provides hardware designs and reconstruction algorithms for academic implementation [45]. Advanced reconstruction pipelines based on inverse-problem solvers have further expanded the computational toolkit available to the ISM community [46]. Current work explores ISM variants aimed at higher resolution or deeper imaging. Focus-ISM combines multi-element detector arrays with computational background rejection algorithms to enable high-resolution imaging deep within scattering tissues [36, 47], while hybrid STED-ISM protocols achieve sub-70 nm resolution using depletion illumination doses compatible with live-cell imaging of photosensitive specimens. Together, these advances place ISM between confocal microscopy and higher-resolution nanoscopy methods.

1.3 Super-resolution Microscopy

Over the past two decades, advanced microscopy techniques that break the diffraction limit have emerged. These techniques allow us to visualize biological structures such as nuclear pores, viruses, chromatin complexes, and cytoskeletal filaments with resolutions approaching the nanometer scale [7, 48]. These techniques are collectively called super-resolution microscopy. Foundational work in the field was recognized by the 2014 Nobel Prize in Chemistry, awarded to Eric Betzig, Stefan W. Hell, and William E. Moerner. [49]. Development of STED microscopy marked the starting point for modern super-resolution methods [7, 48], followed by stochastic optical reconstruction microscopy (STORM) [9], photoactivated localization microscopy (PALM) and fluorescence photoactivation localization

microscopy (fPALM) [8, 50], and DNA-based points accumulation for imaging in nanoscale topography (DNA-PAINT), which expanded upon the original points accumulation for imaging in nanoscale topography (PAINT) method by Sharonov & Hochstrasser [51, 52]. These latter three methods are collectively termed SMLM, as they achieve super-resolution through precise localization of individual fluorescent molecules [53]. In parallel, Structured illumination microscopy (SIM) emerged as an imaging-based strategy that shifts high spatial frequencies into the passband via patterned excitation, enabling roughly twofold lateral-resolution improvement and, with 3D-SIM, axial gains as well [54, 55]. A simplified overview of the light paths and basic working principles of these super-resolution microscopy (SRM) techniques is shown in Figure 1.2.

Early super-resolution microscopy development focused primarily on proof-of-concept demonstrations and technological advances, including extensions to three-dimensional and multi-color imaging capabilities. Current SRM methods can be classified into two major categories based on their excitation and detection mechanisms: Imaging-based techniques and localization-based techniques [56, 57]. This distinction determines the spatial and temporal resolution, and implementation characteristics of each method. Within SRM, the term *nanoscopy* denotes super-resolution approaches capable of theoretically unlimited spatial resolution; namely those that achieve confinement through deterministic state control (e.g., STED, reversible saturable optical linear fluorescence transitions (RESOLFT)) or sequential localization (e.g., PALM, STORM, MINFLUX), in contrast to linear reconstruction or reassignment strategies such as SIM and ISM that yield modest, fixed gains [7, 8, 10, 56]. The distinction reflects the scaling of resolution: state-control and localization methods can, in principle, reach single-nanometer precision given sufficient photons and probe stability. Conversely, SIM and ISM provide a determined improvement in resolution relative to the diffraction limit regardless of illumination intensity [16, 54, 57].

1.3.1 Imaging-based super-resolution microscopy

Imaging-based SRM techniques achieve sub-diffraction spatial resolution through deterministic optical or computational enhancement of microscopy images without requiring sequential localization of individual fluorophores. These primarily include STED, RESOLFT, and SIM (Figure 1.2 (b) & (c)). STED nanoscopy achieves resolution enhancement by co-aligning a diffraction-limited excitation focus with a phase-structured depletion beam, typically a vortex or "donut" pattern with helical phase featuring a central intensity zero. Through stimulated emission, excited molecules at the periphery are forced into the ground state,

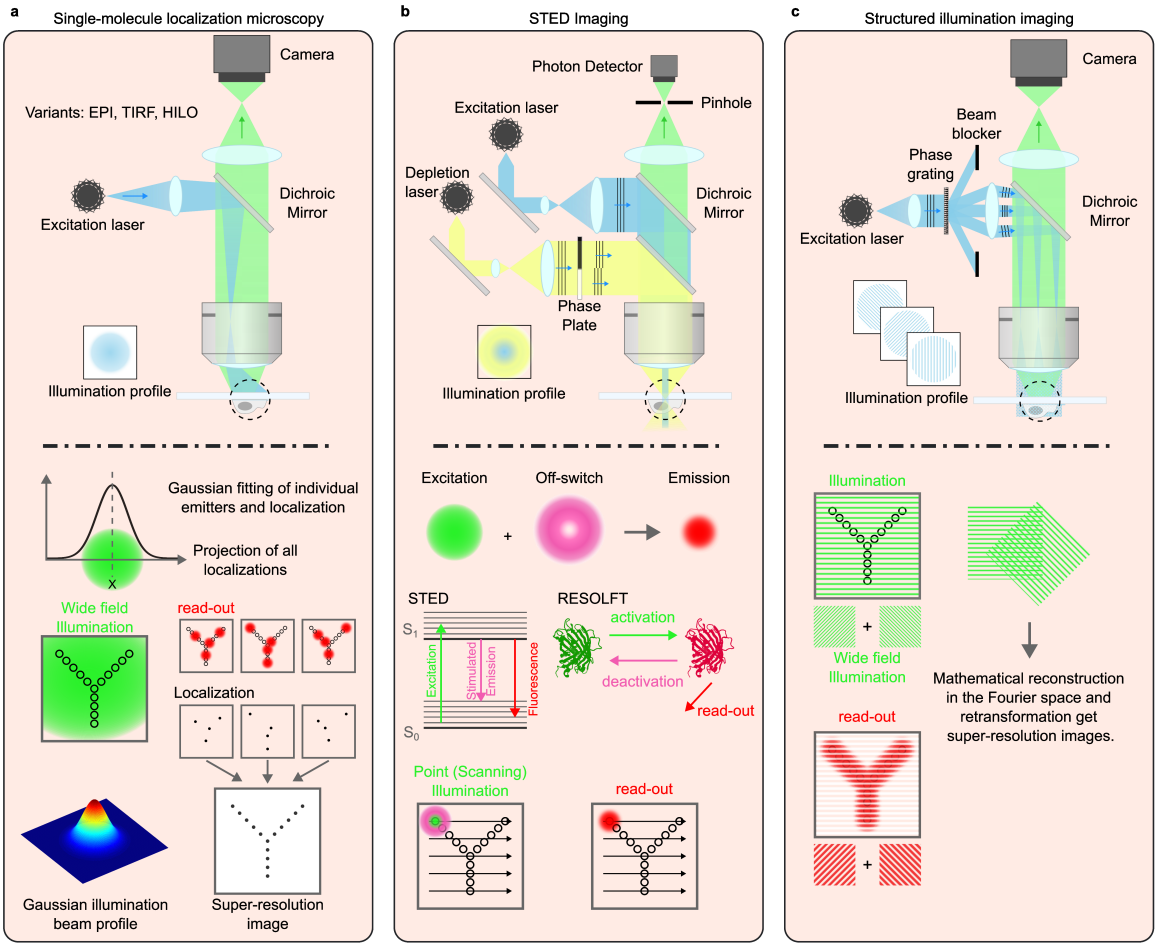

Figure 1.2 Simplified light-paths and basic working principles of common super-resolution microscopy techniques. a) SMLM using widefield imaging. Top: excitation laser with a Gaussian beam profile reflected by the dichroic mirror into the objective to uniformly illuminate the sample; fluorescence is collected through the same objective and imaged onto a camera. Variants include epi, total internal reflection fluorescence (TIRF), and highly inclined laminated optical sheet (HILO). Bottom: under widefield illumination, only a sparse subset of fluorophores is stochastically activated per frame. Each isolated emitter is fitted with a 2D Gaussian function to determine its precise coordinates. This process is repeated over thousands of frames, and all localizations are projected onto a single coordinate map to reconstruct a super-resolution image. **b)** STED microscopy. Top: a confocal-based setup where the excitation laser is co-aligned with a depletion laser that passes through a phase plate to generate a donut-shaped beam with zero intensity at its center. Both beams are focused through the objective, and emission is collected via a pinhole onto a photon detector. Bottom: the donut-shaped depletion beam forces peripheral fluorophores into the ground state via stimulated emission, confining fluorescence to a sub-diffraction central spot. The sample is scanned point-by-point, and the resulting image is built from sequential readouts, achieving resolution beyond the diffraction limit. **c)** SIM imaging. Top: excitation light passes through a phase grating to generate a sinusoidal interference pattern and the structured pattern is projected onto the sample via a dichroic mirror. Emission is captured on a widefield camera. Bottom: the sample is illuminated with stripe patterns at multiple phases and orientations. High-frequency sample information, normally undetectable, becomes encoded through frequency mixing (moiré effect). Mathematical reconstruction in Fourier space separates and recombines these components, extending resolution approximately twofold beyond the diffraction limit.

shrinking the effective fluorescence spot to a sub-diffraction area determined by the depletion intensity [7, 48]. Three-dimensional nanoscopy extends this principle by combining the lateral vortex pattern for XY confinement with an axial intensity minimum created by a bottle-beam or similar phase mask, thereby providing isotropic sub-diffraction resolution in all three dimensions [48]. Various implementation approaches exist, including pulsed and continuous-wave (CW) depletion schemes. In CW-STED configurations, time-gated detection selectively captures late fluorescence photons that arrive after depletion events, which sharpens the effective PSF while operating at a lower depletion power [58–60]. These approaches can achieve sub-100 nm lateral resolution while preserving confocal sectioning. They also remain compatible with multicolor imaging through a selection of dye–depletion wavelength combinations [58].

RESOLFT microscopy builds upon the same "zero-intensity confinement" principle of STED microscopy but uses reversibly switchable fluorophores combined with patterned light fields containing intensity zeros. This approach toggles fluorophores between bright and dark states using intensities that are orders of magnitude lower than those needed for stimulated emission, thereby reducing photobleaching and phototoxicity to levels more suitable for live-cell imaging [61–63]. Both STED and RESOLFT show how deterministic PSF engineering coupled with optical state transitions can break the diffraction limit in scanning microscopes while preserving confocal-like optical sectioning via pinhole detection or, in three-dimensional variants, via axial PSF engineering.

SIM achieves $\sim 2\times$ lateral-resolution improvement by frequency mixing (Moiré shifting) between the specimen and a sinusoidal illumination pattern, with 3D-SIM extending axial resolution and enabling optical sectioning [54, 55]. In principle, nonlinear (e.g., saturated) SIM generates higher harmonics and can therefore extend the accessible spatial frequencies without a fixed upper bound, although in practice the resolution is limited by SBR, photobleaching, and achievable pattern contrast [64–66].

1.3.2 Localization-based super-resolution microscopy

Single-molecule localization microscopy (SMLM) refers to fluorescence imaging techniques that, regardless of their widefield or confocal implementation, reconstruct a super-resolved image from the sequential detection and localization of sparsely emitting fluorophores (Figure 1.2 (a)). This is achieved through photoswitching [67], a stochastic process by which small subsets of individual emitters are randomly activated or switched in consecutive acquisitions, creating a fluorophore 'blinking' (ON/OFF) behavior. Photoswitching is a stochastic

process for individual molecules; however, switching probabilities can be controlled through laser excitation or modulation of the chemical environment [68]. In each acquisition frame, only a small subset of molecules are in the active (emissive) state, ensuring that each molecule appears as a spatially isolated, diffraction-limited spot and that the images of individual molecules do not overlap. Sequential acquisition of thousands of image frames from the same FoV ensures that most fluorescent molecules are active in at least one frame. After imaging, the emitter positions are then estimated with sub-diffraction precision (commonly through Gaussian fitting of the images of individual molecules) and aggregated across thousands of frames to form the final image [8, 9, 50]. This combined localization and computational reconstruction yields a super-resolved image with resolution well below the diffraction limit [69].

SMLM methods are commonly grouped by detection and excitation modality into *wide-field* type (e.g., PALM, STORM, DNA-PAINT) and *confocal*-type (e.g., MINFLUX and Minimal STED (MINSTED) [70, 71]). These SMLM techniques achieves lateral resolutions on the order of ∼20 nm to 30 nm and axial resolutions of ∼50 nm to 80 nm in typical biological samples, with better values attainable under optimized conditions and with DNA-PAINT, which can reach localization precisions of ∼2 nm or better under optimized imaging conditions [69, 72–74].

In SMLM, the achievable precision for localizing an isolated emitter depends on photon number N, the width parameter s of the PSF (i.e., the standard deviation σ of its Gaussian approximation), pixel size a, and background noise b (photons/pixel). A widely used approximation for the lower bound of the standard error in one lateral dimension is [75]:

$$\sigma_{\text{loc}}^2 \approx \frac{s^2 + \frac{a^2}{12}}{N} + \frac{8\pi s^4 b^2}{a^2 N^2} \tag{1.8}$$

This scaling makes explicit the central SMLM tradeoffs: increasing N (brighter dyes, longer exposures) improves precision as $1/\sqrt{N}$, whereas increasing background or motion blur degrades precision [76]. In practice, lateral localization precisions of ∼5 nm to 15 nm are routine for bright organic dyes in fixed cells. DNA-PAINT can push this precision to the few-nanometer regime because its blinking mechanism is decoupled from the fluorophore: transient binding of imager strands drives the on/off switching, enabling the use of exceptionally bright, photostable organic dyes (e.g., Cy3b) that yield high photon counts per localization event, while continuous replenishment of imager strands from solution effectively reducing photobleaching [77, 78]. Three-dimensional SMLM is achieved by PSF engineering such as astigmatism, double-helix PSF, 4Pi detection, or interferometric approaches yielding

axial precisions of ∼20 nm to 50 nm under favorable conditions [69, 79–82]. Algorithmic advances since 2010 include multi-emitter fitting for higher-density imaging [83], sCMOS noise modeling, pointwise error estimation beyond simple Cramér-Rao bound (CRB) proxies [76] , and open-source pipelines (e.g., ThunderSTORM [84]) for reproducible analysis.

Widefield-based single-molecule localization microscopy

Historically, the widefield-based SMLM techniques progressed in the mid-2000s with three notable research contributions. PALM demonstrated super-resolution by photoactivating sparse subsets of photoactivatable or photoconvertible fluorescent proteins and fitting their PSFs to estimate molecular positions [8, 50]. STORM exploited the photoswitching of organic dyes in specialized buffers to stochastically cycle fluorophores between dark and bright states, resulting in higher photon yields and thus higher localization precision than typical fluorescent proteins [9, 69]. In parallel, PAINT introduced transient probe–target interactions to generate stochastic blinking without requiring photoswitchable chemistry [52]. DNA-PAINT subsequently used programmable, transient DNA hybridization to produce blinking kinetics with high localization precision, and Exchange-PAINT enabled highly multiplexed 2D/3D imaging via sequential buffer exchanges [73, 85]. Quantitative extensions such as qPAINT convert binding kinetics into absolute or relative target copy numbers, decoupling counting from dye photophysics [86].

Widefield SRM techniques are camera-based modalities that overcome the diffraction limit by using structured illumination and/or fluorescence spatio-temporal statistics [54, 55], while retaining parallel acquisition across the FoV. In the localization-based class, these techniques rely on the principle that isolated fluorescent molecules can be localized with high precision, provided their PSFs do not spatially overlap [87]. This capability is achieved by labeling biological targets with specialized probes that cycle between emissive (ON) and non-emissive (OFF) states, or exhibit spectral alterations (e.g., switching from green to red emission) [8, 9, 50]. Widefield SRM techniques include SMLM methods such as STORM, PALM, fPALM, DNA-PAINT, and their derivatives.

Recent work has focused on increasing speed, multiplexing, and quantitative accuracy. DNA-PAINT variants and analysis toolkits (e.g., Exchange-PAINT, qPAINT) have matured into robust, high-plex imaging with sub-10 nm resolution in vitro and ∼10 nm to 20 nm in cells [73, 85, 86]. Repetitive optical selective exposure (ROSE) uses interferometric excitation patterns to enhance localization precision achieving down to ∼5 nm feature resolution laterally and sub-2 nm axially in ROSE-Z at comparable photon budgets [88, 89]. For multiplexing and counting, Exchange-PAINT and qPAINT have broadened applications in fixed

cells and complex tissues [85, 86]. Software ecosystems for SMLM (e.g., ThunderSTORM, SMAP, Picasso) and experimental strategies (e.g., active stabilization, axial modulation) have improved precision, 3D fidelity, and throughput [82, 84].

SMLM techniques have been applied to mapping macromolecular assemblies (e.g., nuclear pore complexes), quantifying stoichiometries in situ (qPAINT), resolve cytoskeletal architectures, and study receptor nanoscale organization in membranes under pharmacological perturbations [82, 85, 86, 90]. There are also multiple commercial products available that are based on SMLM such as: i) Nikon N-STORM integrates high-power TIRF/epifluorescence with optimized buffer/detection for STORM; ii) Zeiss ELYRA supports PALM/STORM and DNA-PAINT workflows with 2D/3D options; iii) Leica GSDIM implements ground-state depletion imaging; these systems include synchronization, drift correction, and reconstruction pipelines documented in application notes and user literature.

Recently, the Jungmann group introduced a new widefield-based SMLM technique called resolution enhancement by sequential imaging (RESI) [90], which represents the current state of the art of camera-based super-resolution methods. RESI extends DNA-PAINT's barcoding to subdivide the original docking-strand single-target population into M mutually exclusive, orthogonal barcodes that are imaged sequentially. For each physical target j, the technique acquires K_j independent localizations across multiple imaging rounds. The resulting K_j localizations are averaged and the final localization precision for target j follows the statistical relation:

$$\sigma_{\mathrm{RESI},j} = \frac{\sigma_0}{\sqrt{K_j}} \tag{1.9}$$

where $\sigma_0 \approx 4\,\mathrm{nm}$ to $8\,\mathrm{nm}$ represents the single-round DNA-PAINT precision. Thus, the error decreases as $1/\sqrt{K}$ and can reach $< 1\,\mathrm{nm}$ (6–9 Å) in vitro and $< 2\,\mathrm{nm}$ to $3\,\mathrm{nm}$ in cellular environments, where probe size limits effective resolution [90]. This way averaging multiple sampling events yields sub-nanometer precision without requiring brighter fluorophores or more complex optics.

Masullo *et al.* combined RESI with bio-orthogonal metabolic sugar labeling to achieve ångström lateral precision imaging of individual monosaccharides within cellular glycocalyx structures [91]. Their approach uncovered sub-protein-scale organizational patterns of cell surface glycans that had remained inaccessible to previous super-resolution techniques, revealing discrete sugar clustering with spatial resolution approaching $\approx 0.8\,\mathrm{nm}$. Similarly, Pachmayr *et al.* used RESI to map the nanoscale epitope geometry of clinically relevant therapeutic antibodies on cancer cell surfaces, directly correlating sub-3 nm binding con-

figurations with therapeutic efficacy in cancer immunotherapy applications [92]. These studies illustrate how RESI's spatial resolution supports structure–function analysis at the single-protein level within intact cellular environments.

However, mathematically, the achievable RESI resolution is limited by photon number. The CRB for an averaged position is:

$$\sigma_{\mathrm{CRB}} = \frac{\sigma_0}{\sqrt{K}} \propto \frac{1}{\sqrt{N_{\mathrm{tot}}}} \tag{1.10}$$

but the sequential barcoding ensures that photons from individual emitters are collected without crosstalk from neighboring molecules, so K can, in principle, be increased until stage or sample drift and dye-linkage errors dominate.

In practical terms, classical SMLM delivers $\sim$20 nm to 30 nm resolution; DNA-PAINT often reaches < 10 nm; ROSE pushes lateral to $\sim$5 nm and axial to < 2 nm; RESI demonstrates sub-nanometer resolution in vitro and single-protein resolution in cells, contingent on barcoding, sequential acquisition, and sample stability [88, 90, 93].

Confocal-based single-molecule localization microscopy

Confocal-type SMLM acquires localization information with point excitation and single-pixel (or effectively single-aperture) detection. Beyond photon-count-based readout, Enderlein and colleagues combined a SMLM-like localization scheme with confocal laser scanning into a fluorescence-lifetime SMLM (FL-SMLM), emphasizing the feasibility of confocal SMLM modalities and their complementarity to camera-based approaches [94, 95]. In this approach, the sample is scanned with a confocal beam at high frame rates and the resulting images are analyzed using a widefield-like localization procedure. Radmacher *et al.* subsequently extended this concept by incorporating ISM-style array detection and pixel reassignment into FL-SMLM, demonstrating a twofold resolution gain relative to conventional confocal-based localization while retaining nanosecond temporal resolution for lifetime discrimination [96].

Confocal-type localization methods such as MINFLUX and its derivatives (Figure 1.3) localize single emitters by probing fluorescence with excitation patterns that contain intensity zeros (donut-shaped beam), by iteratively positioning the minimum relative to the emitter. This allows for nanometer and sub-nanometer precision with lower photon counts than camera-based SMLM techniques [10]. In practice, MINFLUX has demonstrated molecular-scale localization in imaging and tracking: lateral/axial precisions of approximately $\sim$1 nm to 3 nm in 3D imaging of cellular protein complexes and organelles, with typical photon budgets on the order of $10^2 - 10^3$ photons per localization [11, 97, 98]. Three-dimensional and multi-

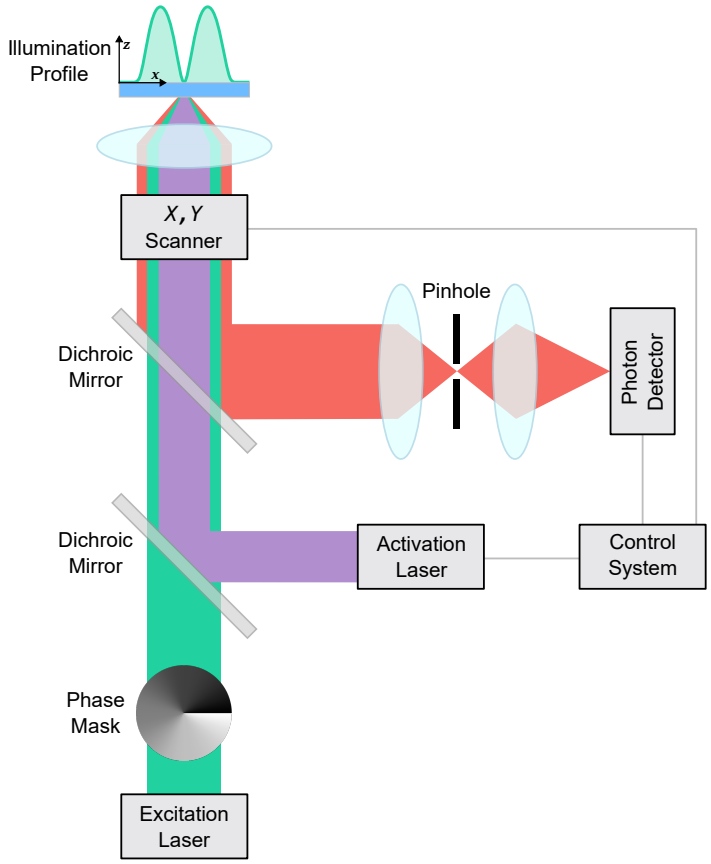

Figure 1.3 Simplified MINFLUX setup. An excitation laser beam (green) is shaped by a vortex-phase mask forming a donut intensity spot in the focal plane of the objective lens. The beam is overlapped with a photoactivation beam (purple) and focused into the sample after passing through an xy scanning system. Fluorescence is descanned, deflected by a dichroic mirror (DC), filtered by a confocal pinhole and detected by photon-counting APDs. Image adapted from [11].

color implementations have been established in cells, enabling isotropic few-nanometer precision on targets such as the nuclear pore complex and MICOS subunits in mitochondria [11, 98]. A related MINFLUX-based strategy, MINSTED, uses STED-generated intensity minima as a movable coordinate for emitter localization while operating at comparatively low light doses [70].

Chapter 2

State of the Art

Since the inception of super-resolution microscopy, many advances have improved the performance of these microscopes. Recent developments in confocal-based single-molecule localization microscopy have led to minimal fluorescence photon fluxes microscopy (MINFLUX) and its derivaties, which represents the current state of the art of scanning-based super-resolution methods [10, 11]. MINFLUX and its derivatives use a donut-shaped excitation beam with a central intensity zero. The beam minimum is iteratively positioned in a targeted coordinate pattern (TCP) relative to the fluorophore, and the residual fluorescence is collected (Figure 2.1). The achievable localization precision in 2D, under a quadratic approximation of the donut excitation profiles, follows the Cramér-Rao bound (CRB):

$$\sigma_{\text{MINFLUX}} \simeq \frac{L}{2\sqrt{2N}} \tag{2.1}$$

where L represents the TCP radius and N is the number of detected photons. This expression is the limiting form of the full result (Supplementary Eq. S27 in [10]), which uses four donut-beam exposures in a 2D geometry. In conventional single-molecule localization microscopy (SMLM), the prefactor s—the standard deviation of the diffraction-limited PSF—scales with PSF the width ($\sim \lambda/(2\text{NA})$). In MINFLUX, the prefactor $L/(2\sqrt{2})$ depends on user-controlled scan parameters, often yielding an order-of-magnitude improvements in photon efficiency. The method attains $\sim 1\,\text{nm}$ to $3\,\text{nm}$ three-dimensional precision with $10^2 - 10^3$ detected photons, compared with $10^4 - 10^5$ photons for similar precision in camera-based methods.

Soon after MINFLUX was introduced in 2017, iterative and three-dimensional implementations appeared. Gwosch *et al.* reported isotropic sub-3 nm precision on cellular structures, including nuclear pore complexes and mitochondrial MICOS assemblies [11, 98]. The

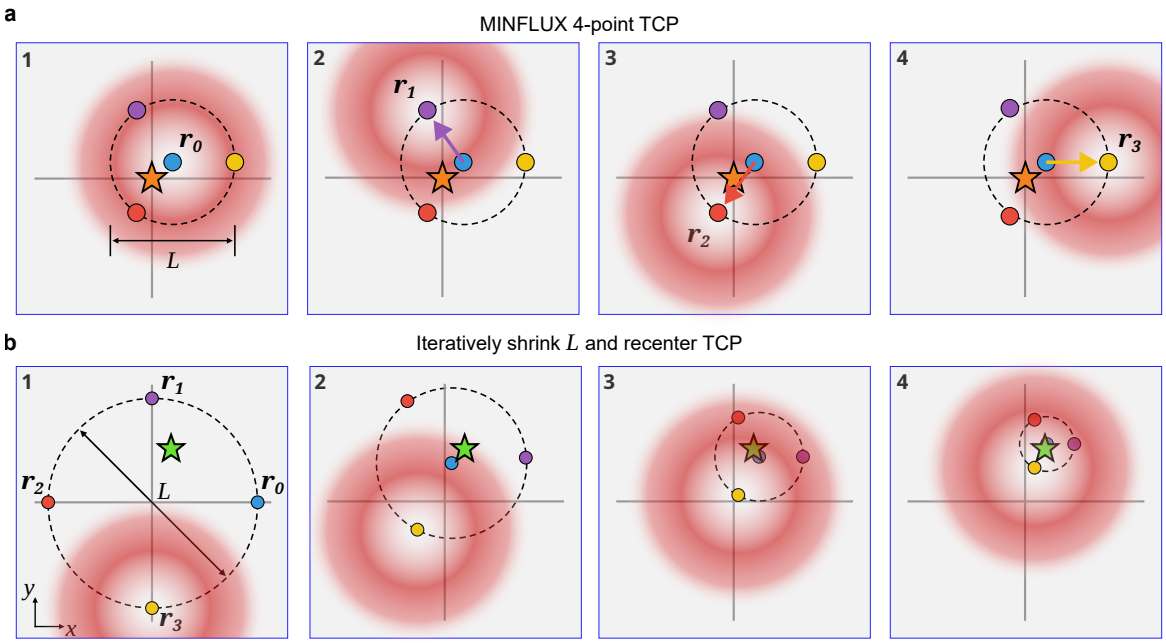

Figure 2.1 MINFLUX excitation beam targeted coordinate pattern. a) Classical MINFLUX 4-point TCP; targeting the beam to four designated coordinates constituting the TCP (blue, purple, red, yellow and beam on orange target in red) based on [10]. **b)** Modified TCP based on [11] where the diameter L is iteratively shrunk and the pattern in recentered on the emitter.

photon efficiency of MINFLUX and related methods is useful for live-cell tracking, where photobleaching and photostability limits the tracking and imaging times.

In 2021, Minimal STED (MINSTED) was introduced, combining a stimulated emission depletion microscopy (STED) beam with a MINFLUX-like implementation [70]. MINSTED replaces the intensity zero of the donut excitation pattern with the minimum of a co-aligned STED beam. This approach exploits STED's capacity to create an effectively sub-diffraction detection volume around the beam minimum, with the localization precision bound given by:

$$\sigma_{\text{MINSTED}} \simeq \frac{r}{\sqrt{N}} \tag{2.2}$$

where r is the distance between fluorophore and STED beam minimum. The distance r can be reduced below the diffraction limit by increasing STED power, and the stimulated emission process provides intrinsic background suppression through temporal filtering of spurious fluorescence events. MINSTED achieves sub-nanometer precision with $\sim 200-1000$ detected photons per localization, while maintaining low background levels that can be challenging for the excitation-based approaches.

Recent developments in MINFLUX increased spatial precision and expanded measurement capabilities. Sahl *et al.* introduced a planar-projection MINFLUX geometry with

photoactivatable fluorophores to directly measure intramolecular distances down to 1 Å laterally and 1 nm axially [99]. This approach enables distance measurements spanning 1 nm to 10 nm which is the domain of single-molecule Förster resonance energy transfer (FRET) while providing linear distance readouts without the orientation-dependent complications of energy-transfer methods. Using this technique, the authors mapped distance distributions within individual proteins and complexes in live cells, enabling new studies of conformational dynamics. Rickert *et al.* developed a 4π MINFLUX configuration with opposing objectives to generate interference-enhanced axial gradients by interfering their respective excitation minima [100]. This doubles the axial intensity gradient near the focal minimum, yielding isotropic $\sigma_{x,y,z} \lesssim 1\,\text{nm}$ with sub-millisecond temporal resolution. This increases the spatiotemporal information rate for applications requiring high spatial and temporal resolution, such as molecular-motor tracking and confined-diffusion analysis.

Advancements in MINFLUX have also spawned a plethora of derivative techniques that utilize the core concepts, but deviate from the original design in order to make the technique more approachable and feasible for a broader research audience. Notable examples include RASTMIN, which replaces iterative TCPs with a raster scan of an intensity minimum to approximate MINFLUX performance on a standard confocal setup [101, 102]. Pulsed-interleaved MINFLUX (pMINFLUX) introduces pulsed, temporally interleaved, spatially shifted vortex beams and time-correlated single-photon detection, enabling access to fluorescence lifetime alongside position estimates and simplifying aspects of the control scheme [103, 104]. Lifetime-enabled variants extend pMINFLUX to co-tracking, FRET-assisted colocalization, and lifetime multiplexing of two spectrally similar dyes, enabling simultaneous multi-emitter measurements at the nanometer scale [105]. To achieve precise 3D localization with low photon counts, GET-pMINFLUX combines pMINFLUX with graphene energy transfer, extracting axial position from fluorescence lifetime while using pMINFLUX for lateral coordinates, reaching sub-2 nm lateral and sub-nanometer axial precision near the graphene interface [104].

Beyond these experimental implementations, several groups proposed and analyzed excitation and patterning strategies that retain the MINFLUX principle of probing fluorophores with intensity minima. These approaches seek to reduce instrumentation complexity or expand the performance envelope of the technique. Groups have theoretically explored high-order vortex beams to increase localization sensitivity [106]. Multiphoton and two-photon MINFLUX variants have also been proposed theoretically with predictions of improved precision or larger usable field-of-view (FoV) without requiring fundamental changes to detection hardware. Theoretical analyses of multiphoton RASTMIN have suggested that

sub-nanometer precision can be preserved across larger scan areas [102]. These external approaches extend MINFLUX applicability and aim to make coordinate-targeted SMLM easier to implement. Importantly, they aim to preserve the key characteristics: high photon efficiency and nanometer-scale precision.

Soon after the conception of MINFLUX, Abberior Instruments released commercial MINFLUX systems, including standalone MINFLUX and integrated MINFLUX-Flex, providing multicolor 3D localization (2 nm) and single-particle tracking up to 100 kHz over fields of about 20 µm × 20 µm [97]. These systems incorporate factory-aligned beam scanning, closed-loop stage stabilization with sub-nanometer precision, and real-time reconstruction algorithms, making molecular-scale fluorescence microscopy accessible to researchers without extensive optical expertise.

Together, wide-field resolution enhancement by sequential imaging (RESI) and confocal-type MINFLUX set the present state-of-the-art of single-molecule super-resolution microscopy. RESI attains Ångström-scale precision via temporal multiplexing and averaging, while MINFLUX reaches comparable single-localization precision with fewer photons but higher instrumental complexity. These parallel developments provide high spatial resolution sufficient to access molecular-scale structural information in cellular systems.

2.1 Current Challenges

Although MINFLUX offers nanometer-level precision for single-molecule imaging and tracking, it has some challenges and shortcomings. Major limitations include restricted localization range, a mandatory iterative pre-localization step, a small FoV, and technical complexity. The restricted localization range arises from TCPs, which can cause localization uncertainty to grow rapidly with displacement. As a result, the CRB diverges rapidly when the emitter lies outside the TCP. Consequently, MINFLUX-like methods perform an iterative pre-localization step to find molecules before imaging or tracking. This reduces the photon budget and increases photobleaching, longer acquisition times and sometimes lowering achievable precision. The FoV of a single MINFLUX is dictated by the L parameter, which is the diameter of the TCP. However, in MINFLUX, high precision is obtained by iteratively decreasing L (Figure 2.1 (b)). This, in turn, reduces the FoV per iteration, lowering practical throughput and increasing imaging time.

Beyond throughput, the point-scanning nature of MINFLUX limits contextual imaging. Because only a single TCP region is probed at any time, acquiring a cell-wide overview requires sequential tiling or mosaic acquisition, which is time consuming and incompatible

with dynamic samples. This makes it difficult to correlate nanoscale structural details with their broader cellular environment. In contrast, widefield SMLM methods acquire entire fields of view simultaneously, enabling rapid contextual imaging alongside high-density localization. A hybrid approach, combining MINFLUX precision with widefield SMLM context, would therefore leverage the strengths of both modalities.

Furthermore, MINFLUX, in its iterative step, requires increasing the beam power to detect signal from the donut minima [11]. This increases intensity at the donut crest, which can saturate or photobleach dyes and, in some cases, damage the sample, complicating live-cell imaging. Because of the TCP design, MINFLUX struggles with simultaneous multi-emitter localization within one TCP scan, requiring high-density samples to be sparsified chemically or by photophysics. Moreover, MINFLUX uses complex, expensive hardware and typically requires substantial technical expertise to build and operate such systems. Lastly, MINFLUX being a closed-source system, most implementations remain proprietary or lab-specific. Open, user-friendly software is limited, which hinders the reproducibility and mass adoption in a broader community.

2.2 Aim of the work

This thesis aims to address some key limitations of MINFLUX microscopy by extending the method with concepts from image-scanning microscopy. We present ***ISM-FLUX*** (image scanning microscopy (ISM)-FLUX), a technique that uses structured illumination through spatially displaced doughnut-shaped beams combined with structured detection via an asynchronous 5×5 single-photon avalanche diode (SPAD) array detector. This detector records the signal from the probed region of interest at each exposure, enabling direct analysis of the spatial distribution of detected photons. The approach extends the localization range relative to MINFLUX-based methods while maintaining similar precision across a larger FoV.

Conventional MINFLUX performs best when molecules lie within a small TCP, typically with a radius L on the order of tens of nanometers. However, localization becomes challenging when molecules fall outside this region. Our approach uses the array detector to simultaneously capture information across a larger area, extending the effective localization range to several hundred nanometers. Inside the TCP, our approach achieves localization uncertainty similar to conventional MINFLUX, and for emitters outside the TCP, it reaches precision comparable to wide-field SMLM. This extends the usable range of ISM-FLUX to several hundred nanometers.

The array-detector architecture supports alternative TCP patterns (e.g., orbital scanning) with similar performance, simplifying optical implementation on a standard confocal microscope. This simplification reduces instrumental complexity compared to iterative MINFLUX implementations.

The objective of the work is to establish the experimental foundations of ISM-FLUX and to define practical operating limits. We believe this approach will emerge as a significant asset for studying cellular dynamics. Its design aims to provide nanometer-scale information while remaining straightforward to use; making it approachable for a broader audience.

Chapter 3

ISM-FLUX: Working Principle and Localization Framework

This chapter provides a comprehensive overview of the working principles of Image scanning microscopy (ISM)-FLUX and the associated single-molecule localization framework. We discuss the key design choices underlying ISM-FLUX, including the orbital scanning of a focused donut-shaped excitation beam combined with a single-photon avalanche diode (SPAD) array detector in a descanned detection path. Furthermore, we establish the forward model and the corresponding molecule detection functions. Finally, this formalism is used to derive the maximum-likelihood estimator for precise molecular localization.

3.1 Single-Photon Avalanche Diode (SPAD) array

A typical confocal laser scanning microscope combined with a single-photon avalanche diode (SPAD) array detector, with minimal modifications is shown in Figure 3.1 (a). The detector comprises a 5×5 or a 7×7 single-photon avalanche diode array with an integrated active quenching circuit fabricated on the same silicon die. In the SPAD array, upon photon absorption, an avalanche current is triggered across the reverse-biased junction. This avalanche is sensed and actively quenched by the on-chip circuitry, and the bias is subsequently restored after a user-selectable hold-off interval that is common to all pixels. After each detection, the pixel is quenched and then kept inactive for a programmable hold-off (dead time; T_D) between $T_D = 25\,\text{ns}$ to $200\,\text{ns}$, and the system achieves a typical timing resolution of $\sim 150\,\text{ps}$ (jitter) relative to the actual detection instant. Jitter is defined as the standard deviation of the temporal error between the photon's actual detection instant and the timestamp reported by the system. This can cause spurious detections within µs after an event due to the charge

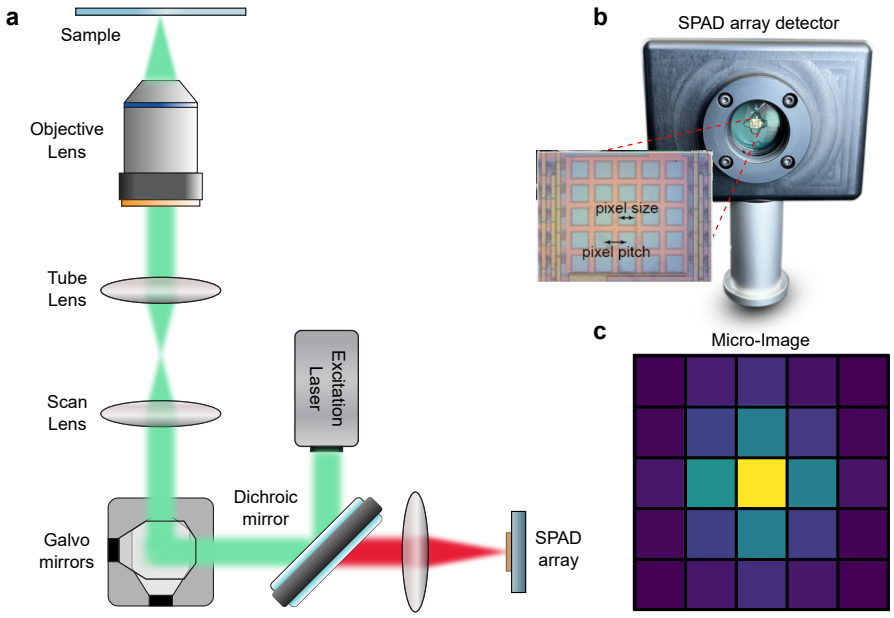

Figure 3.1 Laser scanning microscope with SPAD array detector. **a)** Sketch of the laser scanning microscope equipped with a SPAD array detector. **b)** Example image of a physical detector and with a 5×5 chip inside it. **c)** The photon count from the 25 detector elements is visualized as a matrix, called a *micro-image*.

carriers trapped during an avalanche. This probability is mitigated by selecting longer hold-off times at the expense of maximum photon count rate. Individual pixels are SPADs with a size of 50 µm, arranged on a square grid with a pitch of 75 µm [30] as shown in Figure 3.1 (b). Only a fraction of the die area is photosensitive; the inter-pixel spacing reduces optical crosstalk and provides room for circuit/layout constraints.

The spectral response is characterized by the photon detection efficiency (PDE) curve available across the 400 nm to 1000 nm range, as shown in Figure 3.2 (a). The PDE decreases with the wavelength, from ∼45 % at 450 nm to an average of ∼10 % in the $650 - 750$ nm range and approaches the noise floor near 1000 nm. For fluorescence emission in the red spectral region at 700 nm, the active sensing area is ∼1.4 A U, which allows the collection of ∼91 % of photons incident upon the imaging plane when the emitter is in focus. The detector sensor, however, does not exhibit uniform photosensitivity across its entire surface because of the dead-zones situated between pixel elements, resulting in photon losses. We quantify this using the fill factor (photosensitive area divided by total chip area). These dead-zones reduce optical and electrical crosstalk between the neighboring pixels due to the relatively large inter-pixel spacing. Earlier-generation SPAD arrays were limited to fill factor of approximately 44 % [30]. Newer SPAD arrays use microlenses to redirect light

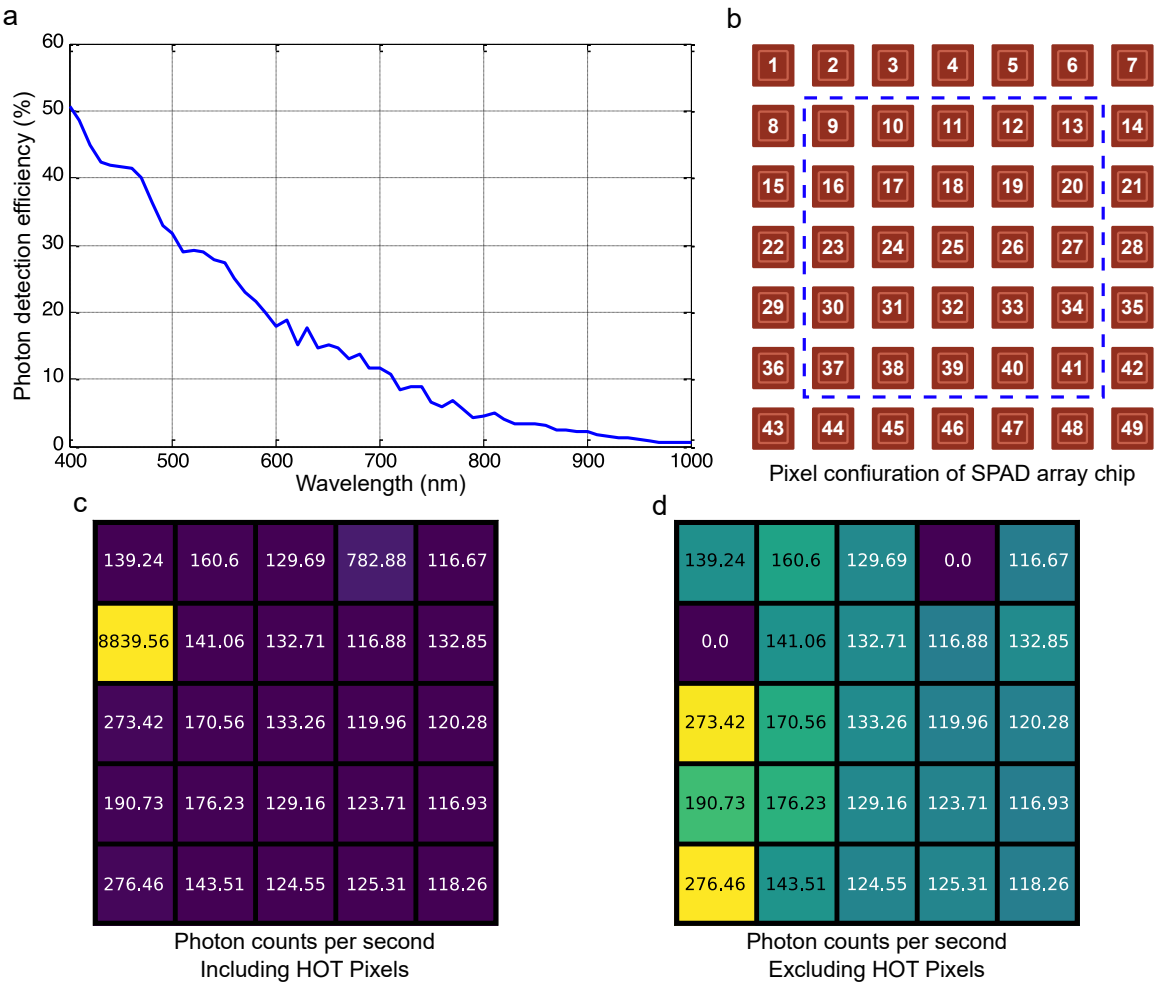

Figure 3.2 Details of the SPAD array detector. **a)** Physical pixel layout of the 7×7 detector chip and the selected 5×5 region. **b)** Average Photon Detection Efficiency for the SPAD array at different wavelengths. **c)** Dark counts map of the 5×5 detector region. The result is the average of 100 acquisitions of 1 s. **d)** Dark counts map of the detector region *excluding* the hot pixels.

from inter-pixel regions onto the active areas, increasing the effective fill factor to more than 80 % [107].

In our setup, the SPAD array is a 7×7 (49) pixel detector; however, for our measurement we operate the detector in a 5×5 readout mode (25 central pixels) as shown in the Figure 3.2 (b). From 25 pixels, we obtain a set of 25 photon count values $\mathbf{n} = (n_1, n_2, n_3, ..., n_K)$ which can be rearranged in a 5×5 image as shown in Figure 3.1 (c), which we call a *micro-image*. We use the central 5×5 (25) (Figure 3.2 (b)) pixels because the full 7×7 (49) pixels increases the collection diameter (or width) to $\sim$1.96 A U. This increases background with little additional signal from in-focus emitters. The pixel-wise map of the dark counts of the SPAD array detector at 25 °C is shown in Figure 3.2 (c). At 25 °C, we measure a total dark

count rate of $\phi_{\text{dark}} = 13.04\,\text{kHz}$ for the selected 5×5 region (Figure 3.2 (c), 1 s integration averaged over 100 acquisitions). Two hot pixels dominate the dark noise, contributing $\sim$74 % of the total; excluding them yields typical pixel dark count rates of 120 Hz to 250 Hz for the remaining 23 pixels (Figure 3.2 (d)).

Although the detector described above was used in our experiments, there are new "red-enhanced" detectors available that increase the detection efficiency in the red (650 nm to 750 nm) range by up to 5× [108].

3.2 Localization uncertainty and the Cramér-Rao bound

3.2.1 Localization Estimators

Localization is the process of inferring the unknown position coordinates, $\mathbf{r_E}$, of a fluorescent emitter based on a set of noisy measurements and a finite photon budget N. Although the hardware differs between widefield single-molecule localization microscopy (SMLM) and minimal fluorescence photon fluxes microscopy (MINFLUX)-like techniques, the underlying statistical problem is similar. In both cases, the information regarding the fluorophore position is encoded in the distribution of detected photons across a set of measurement channels. In camera-based SMLM, these channels correspond to physical pixels on a detector array illuminated by a widefield illumination. In MINFLUX, the channels correspond to temporal bins associated with specific excitation beam positions.

According to estimation theory, one way to quantify the theoretical lower limit of the localization uncertainty is to calculate the so-called Cramér-Rao bound [109]. The Cramér-Rao bound (CRB) (Section 3.2.2) is a statistical measure that provides a theoretical lower bound on the variance of any unbiased estimator of an unknown parameter. Several estimators are used in SMLM, with different trade-offs in bias and variance.

We can describe these modalities using a unified observation model. Let the dataset consist of N_c independent measurement channels, indexed by $k = 1, \ldots, N_c$. The number of photons detected in the k-th channel, denoted by n_k, follows a Poisson distribution with a mean expectation μ_k that depends on the emitter's position $\mathbf{r_E}$:

$$n_k \sim \text{Poisson}(\mu_k(\mathbf{r_E})), \quad k = 1, \ldots, N_c. \tag{3.1}$$

The expected count $\mu_k(\mathbf{r_E})$ is the sum of the signal from the fluorophore and a background term:

$$\mu_k(\mathbf{r_E}) = N p_k(\mathbf{r_E}) + \beta_k, \tag{3.2}$$

where N is the total number of signal photons detected from the emitter across all channels, $p_k(\mathbf{r_E})$ is the probability of a photon landing in channel k given the emitter is at $\mathbf{r_E}$, such that $\sum_k p_k(\mathbf{r_E}) = 1$, and β_k is the expected background count (e.g., from dark counts or out-of-focus fluorescence).

Assuming conditional independence of photon counts across channels, the full parameter vector $\boldsymbol{\theta} = (\mathbf{r_E}, N, \{\beta_k\}, \ldots)$ can be estimated via maximum likelihood. The log-likelihood function is given by

$$\ell(\boldsymbol{\theta}) = \sum_{k=1}^{N_c} \left[n_k \log \mu_k(\boldsymbol{\theta}) - \mu_k(\boldsymbol{\theta}) \right] + \text{const}, \tag{3.3}$$

This framework covers common estimators (centroid estimation, least-squares fitting, and maximum-likelihood estimation) under specific assumptions on $p_k(\mathbf{r_E})$ and β_k.

In widefield SMLM, $p_k(\mathbf{r_E})$ represents the integration of the emission point spread function (PSF) over the k-th pixel area. In MINFLUX-like (ISM-FLUX) localization techniques, $p_k(\mathbf{r_E})$ is referred to as the Molecule Detection Function (molecule detection function (MDF)), which describes the convolution of the excitation intensity profile and the detection efficiency. The goal of any localization estimator is to recover $\mathbf{r_E}$ from the observation vector $\mathbf{n} = \{n_1, \ldots, n_K\}$. Several estimators exist, offering different trade-offs between computational speed and localization accuracy. The most common models used in SMLM are Center-of-Mass (centroid) estimation, Gaussian fitting, and maximum likelihood estimator (MLE).

Center-of-Mass (Centroid) Estimation The simplest estimator calculates the first moment of the intensity distribution. For a 2D image, the centroid coordinate $\hat{\mathbf{r}}_{\text{CoM}}$ is given by:

$$\hat{\mathbf{r}}_{\text{CoM}} = \frac{\sum_k \mathbf{r}_k n_k}{\sum_k n_k}, \tag{3.4}$$

where $\mathbf{r}_k$ is the spatial coordinate of the k-th pixel center. The centroid estimator is fast but typically suboptimal. It ignores the Poissonian nature of the noise and the specific shape of the PSF or MDF. Furthermore, it is highly susceptible to background noise; a uniform background pulls the estimate towards the center of the field-of-view (FoV), introducing a systematic bias that decreases with the increase in signal-to-background ratio (SBR). In practice, centroid estimation is mainly used for initialization rather than final localization

in high-precision workflows [110]. Practical real-world use of the centroid estimation has already been shown in the single-particle tracking literature [111].

Nonlinear Least Squares (Gaussian Fitting) A common approach in camera-based SMLM is to approximate the PSF as a 2D Gaussian function and fit it to the data using nonlinear least squares (NLLS). The estimator minimizes the sum of squared differences between the observed counts n_k and the model $\mu_k(\mathbf{r})$:

$$\hat{\mathbf{r}}_{\mathrm{NLLS}} = \underset{\mathbf{r}}{\operatorname{argmin}} \sum_k w_k \left(n_k - \mu_k(\mathbf{r})\right)^2 . \tag{3.5}$$

The weights w_k are typically chosen to be the inverse of the expected variance, $1/\sigma_k^2$. This method assumes Gaussian noise statistics. For modern scientific-complementary metal-oxide semiconductor (CMOS) cameras, where read noise varies pixel-by-pixel, this weighting allows the estimator to account for detector-specific noise maps. However, the performance degrades when the PSF deviates from a Gaussian or at lower photon counts typical of SMLM ($N < 500$), where the Gaussian noise approximation for photon shot noise becomes inaccurate [112].

Maximum Likelihood Estimation (MLE) For localizing emitters in MINFLUX-like localization techniques, Maximum Likelihood Estimation is commonly used by formulating a Poisson likelihood for the detected photon counts [10, 87]. For Poisson-distributed data, the probability of observing a specific set of counts $\mathbf{n}$ given the parameters $\boldsymbol{\theta} = (\mathbf{r}, N, \dots)$ is given by the product of individual Poisson probabilities. The log-likelihood function, $\mathscr{L}(\boldsymbol{\theta}|\mathbf{n})$, is:

$$\mathscr{L}(\boldsymbol{\theta}|\mathbf{n}) = \sum_{k=1}^{N_c} \left(n_k \ln(\mu_k(\boldsymbol{\theta})) - \mu_k(\boldsymbol{\theta}) - \ln(n_k!)\right). \tag{3.6}$$

Since the term $\ln(n_k!)$ depends only on the data, maximizing $\mathscr{L}$ is equivalent to maximizing:

$$\mathscr{L}'(\mathbf{r}) = \sum_{k=1}^{N_c} \left(n_k \ln(N p_k(\mathbf{r}) + \beta_k) - (N p_k(\mathbf{r}) + \beta_k)\right). \tag{3.7}$$

The MLE, $\hat{\mathbf{r}}_{\mathrm{MLE}}$, is the coordinate vector that maximizes Eq. 3.7. Unlike the analytical centroid solution, the MLE requires numerical optimization. For widefield SMLM, $p_k(\mathbf{r})$ is typically modeled as a Gaussian or a cubic spline [75, 113]. For MINFLUX and ISM-

FLUX, $p_k(\mathbf{r})$ corresponds to the MDFs generated by the structured excitation beams. MLE is asymptotically efficient and unbiased, meaning that as N increases, its variance approaches the CRB and the bias approaches zero [109, 114]. In the ISM-FLUX analysis we use a modified version of the maximum likelihood algorithm, which is detailed in the Appendix A.

High-Density (multi-emitter) Localization The estimators described above assume a single isolated emitter within the analysis window. When the density of active fluorophores is high, their PSFs or MDFs overlap, and single-emitter models fail. High-density algorithms, multi-PSF fitting (DAOSTORM [115]), sparsity-based deconvolution (WindSTORM [116]), or multi-emitter MLE, extend the likelihood model to account for M overlapping emitters by defining the mean count as a superposition: $\mu_k = \sum_{j=1}^{M} N_j p_k(\mathbf{r}_j) + \beta_k$. Solving this requires determining both the number of emitters M and their respective parameters, often utilizing iterative deflation or sparse deconvolution techniques.

3.2.2 The Cramér-Rao Bound (CRB)

The precision of any unbiased estimator is limited by the information content of the data and it is important to characterize this performance for all the localization methods. This limit is quantified by the Cramér-Rao bound (CRB), which states that the covariance matrix of the estimated parameters, $\mathrm{cov}(\hat{\boldsymbol{\theta}})$, is bounded from below by the inverse of the Fisher Information Matrix (FIM), $\mathbf{F}$:

$$\mathrm{cov}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{F}^{-1}. \tag{3.8}$$

For the Poissonian observation model described in Eq. 3.7, the elements of the Fisher Information Matrix are given by:

$$F_{ij} = \sum_{k=1}^{N_c} \frac{1}{\mu_k(\boldsymbol{\theta})} \left(\frac{\partial \mu_k(\boldsymbol{\theta})}{\partial \theta_i} \right) \left(\frac{\partial \mu_k(\boldsymbol{\theta})}{\partial \theta_j} \right). \tag{3.9}$$

This links precision to photon count and to the spatial gradients of the measurement model. Precision depends on two factors: the photon count (since $\mu_k \propto N$, $F \propto N$, and precision $\sigma \propto 1/\sqrt{N}$) and the sensitivity of the measurement to changes in position (the gradient $\partial \mu_k / \partial \mathbf{r}$).

In conventional camera-based SMLM, the gradient is determined by the slope of the emission PSF. Since the PSF width is diffraction-limited, the information content per photon is fixed by the numerical aperture. In contrast, MINFLUX-like approaches manipulate the

gradient term via structured excitation. Placing an intensity zero near the emitter makes the detected rate strongly sensitive to small displacements ([10] - Supplementary information). This sharp gradient increases the Fisher information significantly, allowing these methods to achieve nanometer precision with orders of magnitude fewer photons than required in widefield SMLM. The localization uncertainty generally scales as:

$$\sigma_{\text{CRB}} \propto \frac{L}{\sqrt{N}}, \tag{3.10}$$

where L characterizes the steepness of the confinement (e.g., the diameter of the targeted coordinate pattern). MINFLUX reduces L iteratively, while ISM-FLUX uses array detection to retain information when the emitter is offset from the intensity minimum.

3.3 ISM-FLUX Principle and Mathematical Framework

ISM-FLUX estimates the emitter position using both the scan position and the photon distribution over the SPAD array. While standard MINFLUX encodes position solely based on the photons collected by a single-point detector, and widefield SMLM encodes position solely through the diffraction-limited spot on a camera, ISM-FLUX utilizes a four-dimensional data set: the photons are detected as a function of both the excitation beam position and the detector pixel coordinate.

In a typical implementation, a donut-shaped excitation beam with a central intensity minimum is focused into the sample. This beam is steered to a sequence of N_c discrete positions, $\mathbf{r}_1, \ldots, \mathbf{r}_{N_c}$, forming the targeted coordinate pattern (TCP). In the standard confocal configuration, the detector is placed in the descanned path; consequently, when the excitation beam moves, the effective observation volume moves with it. The system uses an asynchronous SPAD array detector consisting of $N_d = 25$ independent elements (pixels). Because the detector is in a descanned configuration (i.e., the emitted fluorescence is routed back through the galvanometric scan mirrors so that beam deflection is undone), the fluorescence emission from a point source located at $\mathbf{r}_E$ is imaged onto the array. However, unlike a static widefield microscope, the relative position of the emitter image on the array changes as the scanner moves the excitation beam.

We can mathematically describe the expected signal from an emitter using the convolution of the optical response functions (excitation PSF and detection PSF) of the system. Let $h(\mathbf{r})$ denote the intensity distribution of the donut-shaped excitation beam focused at the origin of the reference system. When the beam is steered to a scan position $\mathbf{r}_i$, the illumination

intensity at the emitter position $\mathbf{r}_E$ is given by $h(\mathbf{r}_E - \mathbf{r}_i)$. The fluorescence emission is incoherent and follows the system's PSF. We define the emission PSF $h_2(\mathbf{r})$ as the probability density that a photon emitted from an emitter at the origin is detected at position $\mathbf{r}$ in the image plane, expressed in sample coordinates. The array detector is composed of discrete elements, where the j-th element (for $j = 1, \ldots, N_d$) integrates the photon flux over its active area defined by a window function $w_j(\mathbf{r})$. In a descanned configuration, the scanners add a displacement to the optical path that effectively shifts the detector's FoV in the sample plane by $\mathbf{r}_i$. Therefore, the probability of a photon emitted from $\mathbf{r}_E$ hitting the j-th pixel while the beam is at position $\mathbf{r}_i$ depends on the relative coordinate $\mathbf{r}_E - \mathbf{r}_i$.

The expected number of photons μ_{ij} detected in element j during illumination step i is proportional to the product of the local excitation intensity and the collection probability of that specific pixel for the given emitter position:

$$\mu_{ij}(\mathbf{r}_E) = C \cdot h(\mathbf{r}_E - \mathbf{r}_i) \cdot \iint h_2(\mathbf{r} - \mathbf{r}_E)\, w_j(\mathbf{r} - \mathbf{r}_i)\, d^2\mathbf{r} \tag{3.11}$$

where C is a constant accounting for the molecular brightness, exposure time, and quantum efficiency. The integral term represents the "detection PSF" of the j-th pixel. Since the detector array is small and centered on the optical axis, this integral describes a standard confocal detection volume, but slightly offset for each pixel j.

This model leads to two localization regimes inherent to the technique. When the emitter is located near the center of the TCP (where $\mathbf{r}_E \approx \mathbf{r}_{\text{center}}$), the photon flux is reduced as well as the shot noise is minimal, making this minimum a uniquely informative operating point, providing the high precision characteristic of MINFLUX. Conversely, when the emitter is located far from the TCP center, the term $h(\mathbf{r}_E - \mathbf{r}_i)$ becomes relatively flat. In this regime, the localization information is carried primarily by the integral term: the spatial distribution of photons across the detector array indices j. In this regime, the detector spatial distribution provides camera-like localization information, and the likelihood typically remains unimodal even outside the TCP center.

This defines the MDF as the combination of these two terms, denoted as $p_{ij}(\mathbf{r}_E)$ in the subsequent statistical analysis. This function maps the probability of observing a photon in a specific channel $\{i, j\}$ for any given emitter position in the FoV.

3.3.1 Practical implementation: The orbital scanning targeted coordinate pattern (TCP)

While the theoretical framework described above is valid for any arbitrary set of beam positions $\mathbf{r}_i$, in practice, the scan trajectory is chosen to trade off precision, complexity, and temporal resolution. Standard MINFLUX implementations typically utilize acousto-optic deflector (AOD) to steer the beam. AODs are inertia-free, allowing the beam to jump quickly between the discrete points of a standard 4-point TCP. AODs can respond in 100 ns to 200 ns for small-signal modulation (10 % to 90 %) [117, 118] and random-access switching time of 0.5 µs to 2 µs for full-bandwidth jumps [119]. However, AODs have a limited range and add significant complexity to the optical path and are not standard components in commercial laser-scanning microscopes.

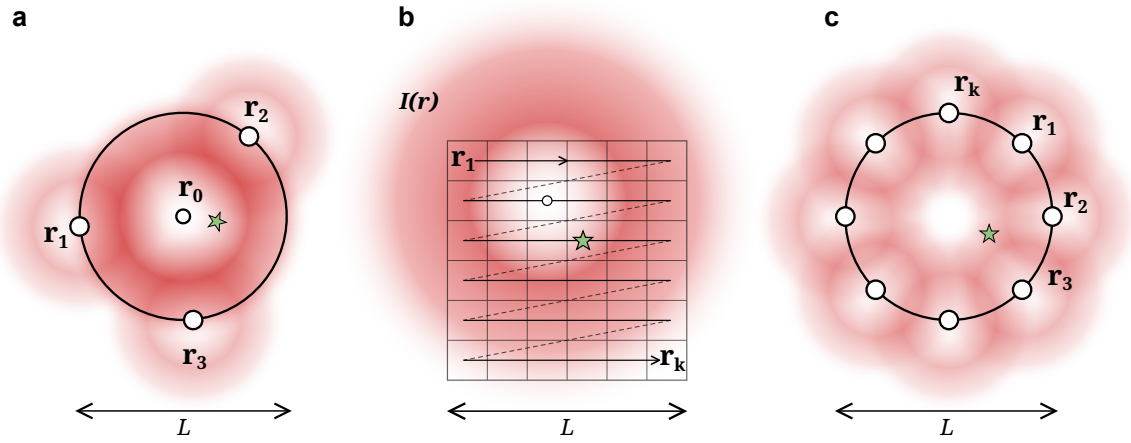

Figure 3.3 Comparison of different targeted coordinate pattern implementations. Schematic comparison of **a)** a discrete TCP (e.g., four-point MINFLUX), **b)** a raster-scanned TCP (RASTMIN-like), and **c)** the continuous orbital TCP used in ISM-FLUX (Donut beam diameter and TCP diameter not to scale).

To remain compatible with standard confocal laser-scanning microscopes, we use galvanometric mirrors. Galvanometer scanners are mechanical systems with mass and inertia, and thus, behave as harmonic oscillators. Attempting to drive them with the Heaviside step functions required for a discrete 4-point TCP induces mechanical oscillations around the target position. Reducing these mechanical oscillations requires a settling dead time at each point, which reduces the duty cycle and thus the total collected photons.

We considered a raster-scanning approach (Figure 3.3 (b)) similar to the RASTMIN technique, which was performed using standard galvanometric mirrors [101]. While RASTMIN reports 50 Hz frame rate under their conditions [101], in our microscope the galvo inertia

limited stable small-field raster scanning to 20 Hz for the scan parameters required here. At this temporal resolution, blinking events are often shorter than the frame duration and are therefore incompletely captured, leading to a degradation in localization precision.

To overcome these limitations, ISM-FLUX uses a continuous orbital scanning TCP as shown in Figure 3.3 (c). By driving the x and y galvanometers with phase-shifted sine waves, the excitation beam traces a circular trajectory around the optical axis. This motion allows for high-speed operation (typically ≥ 500 Hz) with high repeatability and low mechanical noise. This follows the sequential structured-illumination scheme of Masullo et al. (2022), where the beam position varies continuously [120]. This orbit scan is discretized and is divided into N_c positions, effectively creating a TCP consisting of N_c arc segments arranged in a circle. The beam moves continuously and photons are assigned to one of N_c angular bins per orbit in post-processing. In this geometry, the MINFLUX parameter L is defined as the diameter of the circular trajectory.

Beam Shaping and Detection Orbital scanning in ISM-FLUX can be performed using both Gaussian and donut-shaped excitation beam. The donut shaping of the excitation beam is achieved using a static vortex phase plate (VPP) with a topological charge of $m = 1$ (details in Section 4.2.2). Compared to Spatial Light Modulators (SLMs), a vortex phase plate is a passive component and can be inserted in the collimated beam path to generate the donut profile used to minimize μ_{ij} (expected number of photons detected in element j during illumination step i) at the TCP center. This simplifies the setup while still allowing switching between donut and gaussian excitation by inserting/removing the vortex phase plate. Finally, as mentioned before, the detection path replaces the single-point detector of a standard confocal system with a 5×5 asynchronous SPAD array. The array detector provides structured detection to be combined with the structured illumination. The SPAD array provides higher temporal resolution than a camera (jitter $\sim$150 ps, per-pixel dead time $\sim$100 ns) and supports fluorescence lifetime analysis.

Similar to RASTMIN [101], the TCP is generated by using the galvanometric scan mirrors to steer the donut beam. As we scan in an orbit, for each position, an image of the probed sample region is collected forming the microimage as shown in Figure 3.4. When a single emitter is present, these images vary in two aspects: the total number of detected photons per image and the spatial distribution of photons across the detector, which traces a circular path due to the descanned detection mode. While MINFLUX localizes solely based on photon counts, ISM-FLUX additionally considers the spatial coordinate of the photon detection events shown by the middle and last row of Figure 3.4.

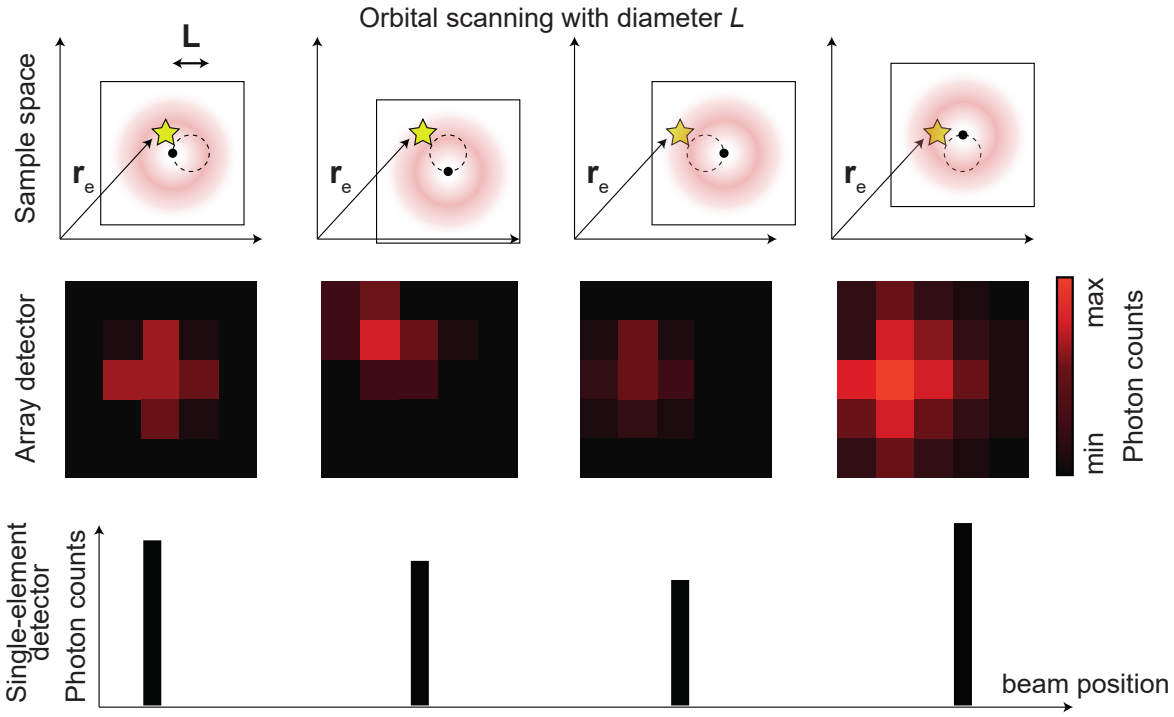

Figure 3.4 Concept of single-molecule localization with ISM-FLUX. The donut-shaped excitation beam moves in a circular pattern around a single emitter **positioned at** $\mathbf{r}_E$ (yellow star). For each position on the circle, the array detector registers a micro-image of the emitted fluorescence. The final photon count data set is four-dimensional: $I(c,\alpha,x_d,y_d)$, with c the orbit number, α the angular coordinate within the orbit, and (x_d,y_d) the detector element coordinates of the array detector. For simplicity, we describe each detector element with a number $I(c,\alpha,d)$. In conventional MINFLUX-based techniques, a single-element detector counts the number of fluorescence photons, discarding their position in the image plane, $I(c,\alpha)$.

We represent the measurement as an $N_c \times N_d \times N_t$ photon-count tensor, with N_c the number of positions in the circle, N_d the number of detector elements ($N_d = 25$), and N_t the number of consecutive orbits for which photons of the emitter were detected. Summing over the N_t orbits, we have a reduced event data set n_{ij}, with $i \in \{1,2,...,N_c\}$ and $j \in \{1,2,...,N_d\}$.

3.3.2 Optimization of Scanning Parameters

The performance of the orbital scanning strategy is governed by two fundamental parameters: the number of sampling points per orbit, denoted as N_c (or `nPoints`), and the diameter of the circular orbit, L. In the orbital approach, N_c and L are chosen to trade off CRB-limited precision against galvo dynamics.

Note: Unless stated otherwise, all CRB simulations in Figures 3.5–3.9 assume a position-dependent photon budget $N(\mathbf{r})$ proportional to the cumulative expected detected signal, i.e. $N(\mathbf{r}) \propto \sum_{i,j} p_{ij}(\mathbf{r})$, reflecting the higher excitation rate experienced by emitters near the donut crest.

Discretization of the Orbit (N_c)

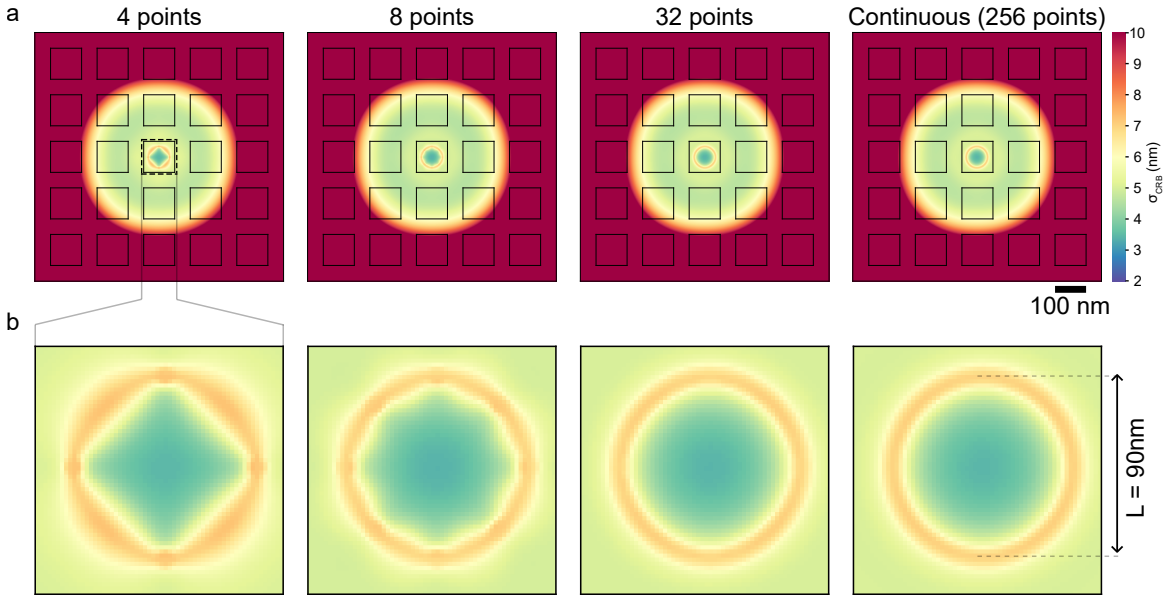

Figure 3.5 Number of orbit poins versus Cramér-Rao bound. a) CRB for a molecule emitting $N = 100$ photons in the TCP center for different number of points during the orbital scanning. N is rescaled for the other emitter positions. L =90 nm. **b)** Zoomed-in section of the center of the CRB map.

To determine the minimum sampling density required for uniform lateral localization precision, we performed CRB simulations for various discrete TCP geometries consisting of N_c points distributed uniformly on a circle. As shown in Figure 3.5, for low values such as $N_c = 4$ or $N_c = 8$, the CRB map exhibits significant anisotropy. In these configurations, the precision is high along specific axes aligned with the beam positions but degrades along the intermediate angles, mirroring the behavior of a static multipoint TCP.

As N_c increases, the discretized TCP approximates a continuous circular trajectory, and the resulting localization precision becomes radially symmetric. For $N_c \geq 32$, the CRB map is approximately isotropic and increasing to ($N_c = 64$ or higher) gives little additional

improvement. Larger N_c (> 32) mainly increases data size and computational time, with minimal gain in spatial information. We therefore set $N_c = 32$ for subsequent experiments. This value provides a sufficiently dense discretization to capture the continuous orbital motion of the galvanometer scanners while maintaining a manageable data size.

Trajectory Diameter (L) and Scanning Speed

The diameter L determines the steepness of the excitation intensity gradient near the optical axis. According to MINFLUX theory, a smaller L yields higher photon efficiency and better precision at the TCP center ($\sigma_{\mathrm{CRB}} \propto L/\sqrt{N}$), but causes the uncertainty to diverge rapidly for molecules located slightly away from the center. In order to figure out the best TCP diameter, we simulated the CRB for different L values as shown in Figure 3.6. From the simulations we can see that the CRB increases outside the TCP for all the different L values. Furthermore, smaller L gives best precision near the TCP center, but the CRB grows steeply with the FoV. Intermediate values like $L \approx 90\,\mathrm{nm}$ keeps CRB low over a larger FoV compared to smaller L value. Based on CRB simulations (Figure 3.6), we targeted an intermediate diameter of $L \approx 90\,\mathrm{nm}$. It yields single-digit-nanometer TCP near the TCP center while preventing the CRB from growing too steeply within the target region.

Experimentally achieving this specific diameter requires accounting for the physical response of the galvanometer mirrors. We operated the galvo mirrors at a scan frequency of approximately 521 Hz (orbit time $\approx$1.92 ms). On a side note: $L = 90\,\mathrm{nm}$ was the smallest repeatable value that we were able to achieve with our galvo mirrors at the given operating speed. At this speed the mechanical inertia of the mirrors acts as a low-pass filter, causing a magnitude attenuation in the scan amplitude. We characterized this behavior by driving the mirrors with different input voltages (representing the "Imposed L") and measuring the actual resulting diameter ("Measured L") using the descanned image formed on the SPAD array.

As illustrated in Figure 3.7, the measured diameter is consistently smaller than the imposed diameter due to this inertial damping, a frequency-dependent attenuation of scan amplitude due to galvo dynamics. This effect is frequency-dependent and increasing the scan speed further attenuates the amplitude of the orbital scan. For instance, increasing the scan rate by a factor of 10 results in a significantly smaller measured L for the same input voltage. Therefore, to obtain $L = 90\,\mathrm{nm}$ at 521 Hz, we drive the galvos with a larger commanded amplitude to compensate for attenuation. This mismatch means the orbit parameters must be measured experimentally rather than relying on the software settings. However, the

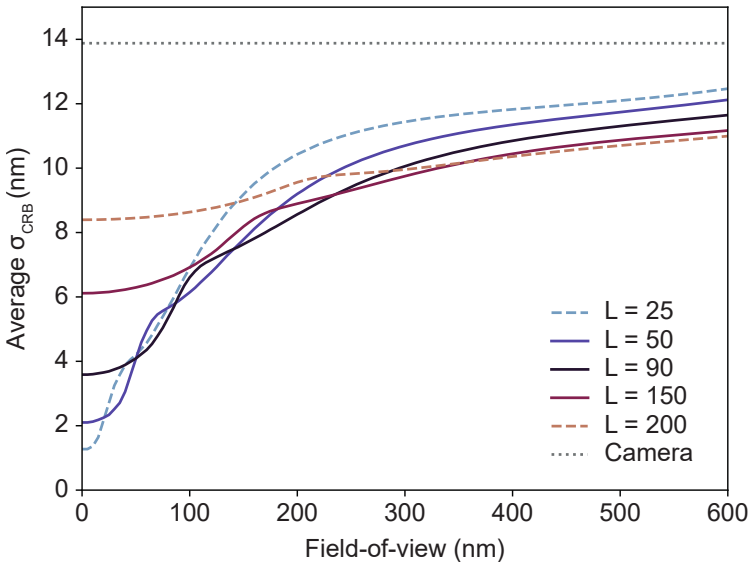

Figure 3.6 Average Cramér-Rao bound scaling along field-of-view for different targeted coordinate pattern diameters. Average CRB within a circular FoV as a function of the FoV diameter. FoV=0 corresponds to the TCP center, and FoV=L to the TCP circle. Comparison of the CRB for different L values. The dotted line shows the CRB for camera-based detection [75] for a pixel size of 50 nm projected onto the sample plane, assuming no background, and other settings similar to those used for ISM-FLUX.

image-based nature of ISM-FLUX allows us to retrieve these parameters directly from the data via a self-calibration procedure, which is detailed in Section 4.5.2.

To evaluate the performance of the different scanning geometries, we compared the theoretical localization uncertainty as a function of the FoV for MINFLUX, RASTMIN, and ISM-FLUX, considering both single-element and SPAD array detection implementations (Figure 3.8). The simulations assumed a photon budget of $N = 100$, a background of $N_{bg} = 25$ (resulting in a Signal-to-Background Ratio, $SBR = 4$), and the TCP diameter $L = 100\,\text{nm}$ (gray vertical dashed line) for all configurations.

Using an array detector extends the usable FoV for all scan geometries and reduces the CRB divergence outside the TCP compared with single-element detection. ISM-FLUX with a SPAD array shows a more homogeneous uncertainty across the FoV (Figure 3.8 blue dashed line) and does not diverge outside the TCP.

Figure 3.9 shows how L changes the localization uncertainty: In our implementation, we kept the laser power, the TCP position, and the L value constant during the localization process. As a result, the emission rate depends on the emitter position relative to the TCP (Figure 3.9). With position-dependent N, the expected uncertainty is $\sim$3.5 nm to 7 nm for $L = 90\,\text{nm}$ and $N \approx 100 - 650$. We use $L = 90\,\text{nm}$ to balance central precision

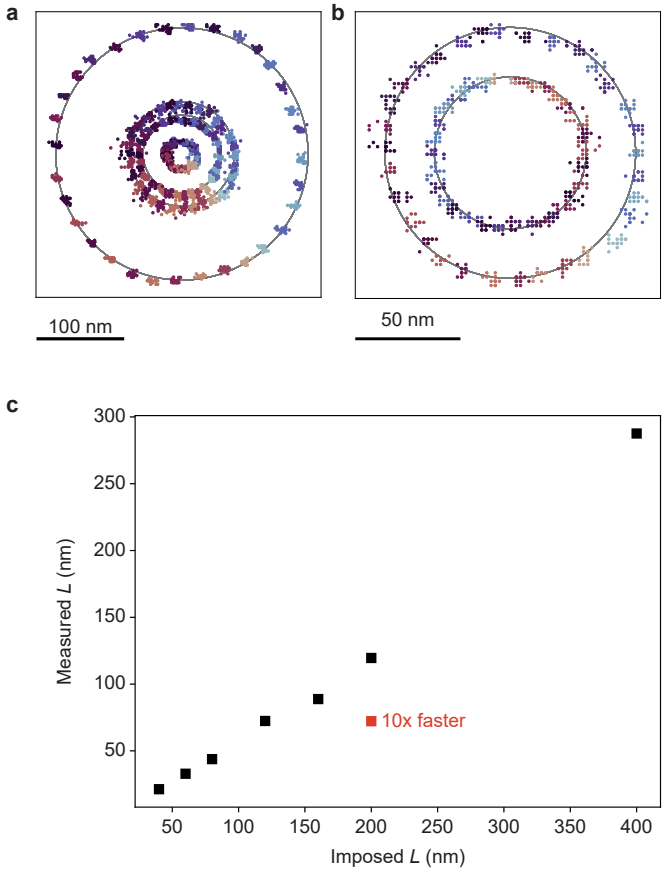

Figure 3.7 Imposed versus measured targeted coordinate pattern diameter for different scan speeds. a) 32 positions of the TCP measured for different L values. From smallest to largest TCP: $L = 60\,\text{nm}$, $160\,\text{nm}$, $200\,\text{nm}$, $400\,\text{nm}$. Orbit direction counterclockwise, from blue to red. **b)** The same imposed L of $100\,\text{nm}$ results in a different observed L and starting angle, depending on the scan speed. The smaller circle corresponds to a $10\times$ faster scan ($192\,\mu\text{s}$ per circle) than the larger circle **c)** Imposed L vs. measured L. Except for the indicated data point, all orbit times were $1.92\,\text{ms}$ per circle.

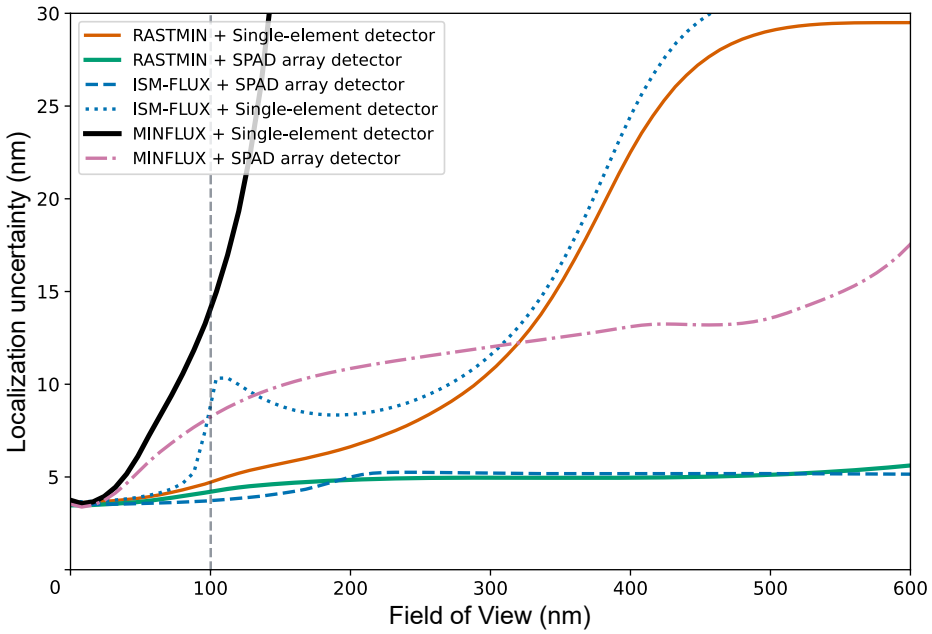

Figure 3.8 Localization uncertainty vs field-of-view for different targeted coordinate pattern patterns. Localization uncertainty as a function of the emitter displacement from the TCP center for six configurations - {MINFLUX, RASTMIN, ISM-FLUX} $\times$ {Single element detector, SPAD array detector}. The signal and background are normalized to $N = 100$ signal photons and $N_{\mathrm{bg}} = 25$ background photons at the TCP center (SBR $= 4$), with $L = 100\,\mathrm{nm}$ (gray vertical dashed line) for all scanning geometries.

against uniformity of σ_{CRB} across the FoV. Note that keeping L large enough also keeps the difference between the expected photon counts for an emitter in the TCP center and an emitter near the donut maximum below one order of magnitude, Figure 3.9. Since we do not adjust the laser power depending on the emitter's position, the radiant flux received by a molecule in the TCP center should be sufficiently high to detect its fluorescence signal above the background. Simultaneously, for a molecule positioned near the peak of the donut-shaped illumination, the radiant flux should be low enough not to induce blinking, bleaching, or saturation of the fluorescence.

The array detector plays an important role here. For large L, where the donut gradient provides little information (the "flat" regions of the excitation profile), the spatial information from the detector array $w_j(\mathbf{r})$ compensates for the lack of structured illumination contrast. This helps keep the uncertainty more uniform over a larger FoV and reduces the divergence seen with single-element detection in the Figure 3.6 & Figure 3.8.

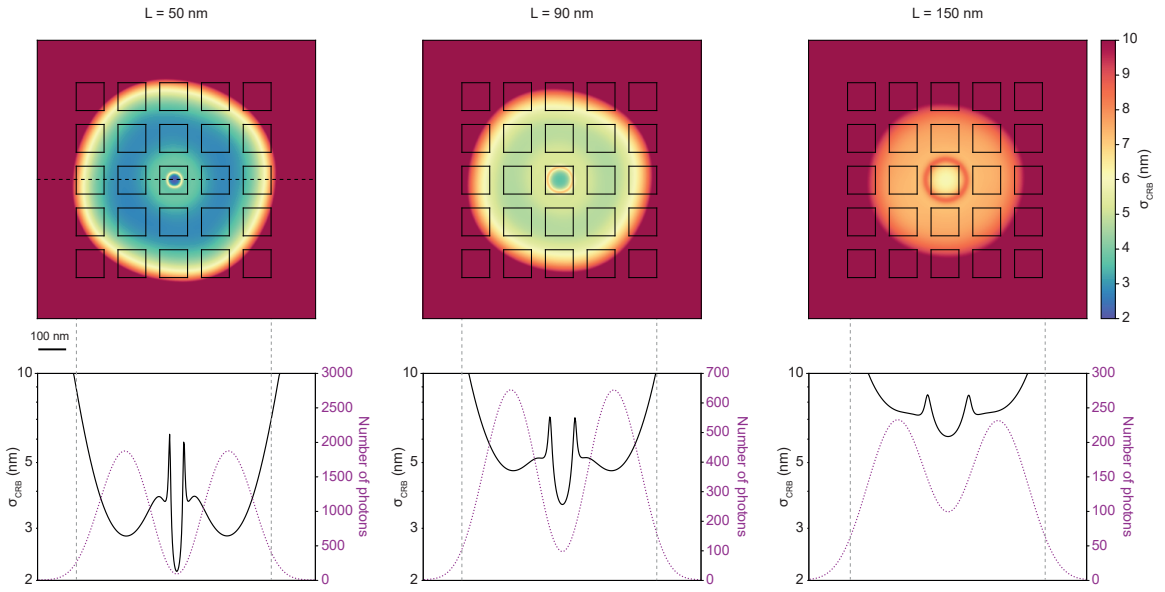

Figure 3.9 Cramér-Rao bound maps for different targeted coordinate pattern diameters. CRB for a molecule emitting $N = 100$ photons in the TCP center. N is rescaled for the other emitter positions, assuming $N(\mathbf{r})$ is proportional with $\sum_{i=1}^{32}\sum_{j=1}^{25} p_{i,j}(\mathbf{r})$.

3.3.3 The Molecule Detection Function (MDF) and Trajectory Size

The statistical framework of ISM-FLUX relies on the Molecule Detection Function (MDF), denoted as $p_{ij}(\mathbf{r}_E)$. This function represents the probability of detecting a photon in detector pixel j at scan position i for an emitter located at $\mathbf{r}_E$. Figure 3.10 shows how p_{ij} is constructed from the excitation pattern and detection PSF. The donut-shaped excitation PSF, shown in Figure 3.10 (a), serves as the basis. During the measurement, this pattern is translated to discrete positions $\mathbf{r}_i$ along the orbital trajectory. Figure 3.10 (b) displays the resulting MDFs for selected detector elements (columns) at different points along the scan orbit (rows). Each tile in this matrix represents the effective observation volume for a specific spatio-temporal channel $\{i, j\}$, defined by the overlap between the shifted excitation beam intensity and the FoV of that specific detector pixel. A stationary fluorophore produces a characteristic spatio-temporal pattern distributed across the N_d (25) detector elements and N_c (32) positions.

The shape and gradient of these MDFs depend on the TCP diameter, L. We evaluated the effect of L using CRB simulations for multiple trajectory diameters (Figure 3.6). In the center of the TCP, the localization uncertainty σ_{CRB} scales linearly with L. For example, a small trajectory ($L = 30\,\text{nm}$) yields an extremely steep intensity gradient at the donut minimum, theoretically allowing for sub 2 nm precision with $\sim$100 photons. However, this high central

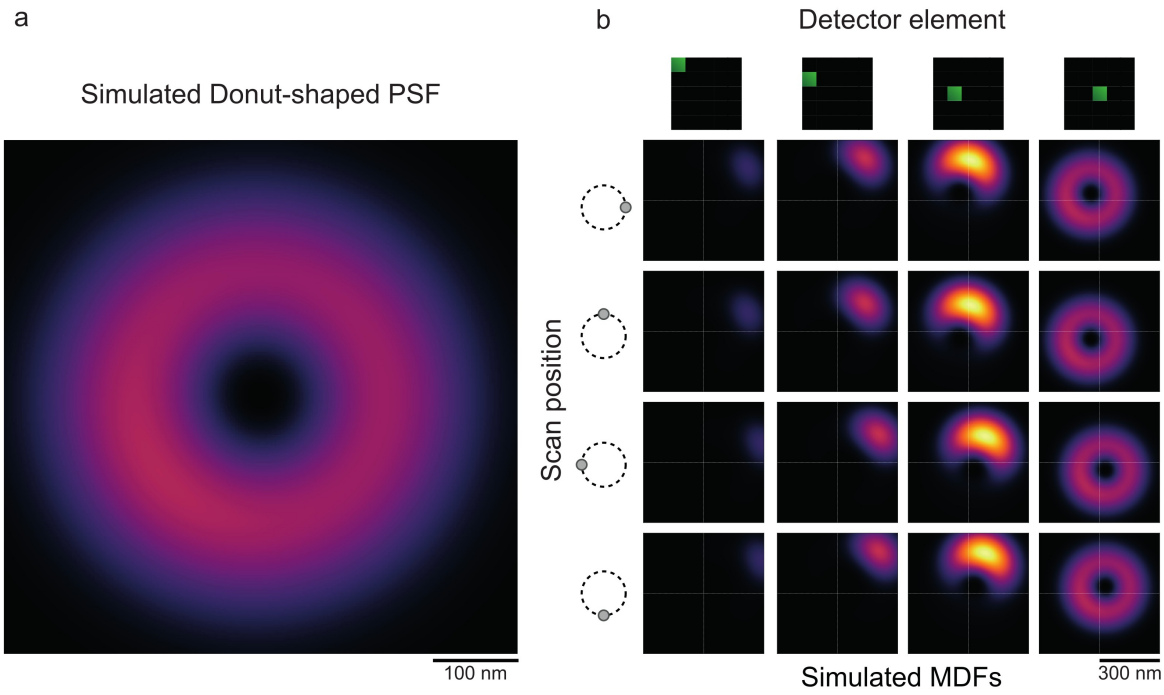

Figure 3.10 Construction of molecule detection functions for ISM-FLUX a) Simulated donut-shaped excitation intensity distribution used for orbital scanning. **b)** Simulated MDFs $p_{ij}(\mathbf{r})$ illustrating the combined effect of structured illumination (scan position index i) and structured detection (detector element index j) on the expected photon distribution across the FoV. Each MDF corresponds to one spatio-temporal channel $\{i, j\}$ and is obtained by combining the shifted excitation pattern with the pixel-integrated detection PSF of element j in descanned detection. The full MDF set comprises $N_c \times N_d$ channels (here $N_c = 32$, $N_d = 25$). For simplicity, we show 4 positions and 4 detector elements.

precision comes at the cost of the relatively reduced FoV for a given localization uncertainty value.

3.3.4 Regimes of localization precision

The achievable localization precision is governed by the number of detected photons N. In widefield SMLM, precision scales as $\sigma \propto \sigma_{em}/\sqrt{N}$, where σ_{em} represents the diffraction-limited width of the emission PSF. In MINFLUX-based approaches, this diffraction limit is replaced by the trajectory parameter L, leading to $\sigma \propto L/\sqrt{N}$. ISM-FLUX transitions between a MINFLUX-like regime and a camera-like regime depending on emitter position. When the molecule is near the optical axis (TCP center), the system operates in the "MINFLUX regime," where the structured illumination dominates the information content. Near the TCP

center, the structured-illumination gradient yields nanometer-scale uncertainty with fewer photons compared to widefield SMLM.

While, when the molecule is located outside the TCP, the excitation intensity increases, and the system transitions into the "Camera-based regime." In this region, the localization uncertainty is governed by the diffraction-limited detection term ($\sigma \propto \sigma_{em}/\sqrt{N}$). However, as the molecule moves away from the TCP center, it enters the higher-intensity crest of the donut beam. As the emitter approaches the donut crest, the detected photon number (N_{AV}) increases and partly compensates for the reduced gradient. The higher photon count partly offsets the loss of L-scaling and reduces uncertainty growth over the extended FoV.

Having established the physical mechanisms governing signal detection and defined the resulting Molecule Detection Functions (MDF, p_{ij}), we now turn to the statistical framework required to inverse this problem of estimating the fluorophore position $\mathbf{r}$ from the noisy photon count measurements.

Chapter 4

Methods and Implementation

This chapter describes the experimental implementation of the instrument used throughout this thesis and the analysis workflow applied to the resulting data. The system combines three coupled components: (i) the $3D$ active sample stabilization module, (ii) an image scanning microscopy (ISM)-FLUX microscope for targeted high-precision localization, and (iii) an integrated camera-based widefield single-molecule localization microscopy (SMLM) pathway intended for correlative, survey-scale imaging. These modules share core optics and a common sample coordinate frame, enabling sequential measurements on the same region of interest without mechanical realignment.

4.1 System overview and rationale

Beyond the core ISM-FLUX microscope, we implemented two additional optical subsystems. First, we added a $3D$ active stabilization subsystem to compensate sample drift in real time during long acquisitions. Second, we integrated a widefield SMLM detection path to obtain large-area context images for correlative workflows with ISM-FLUX. Here we explain the motivation for each subsystem and summarize the constraints that shaped the final design.

4.1.1 Active sample stabilization

The performance of super-resolution microscopy, particularly of SMLM, depends on the inherent stability of the whole imaging system, including the sample. System drift, a pervasive challenge affecting all imaging platforms, significantly hinders the achievable resolution in super-resolution microscopy (SRM). This drift arises from a combination of thermal fluctuations and mechanical vibrations within both the imaging system itself and the

surrounding environment. The negative impact of system drift becomes evident in the form of a declining image quality ranging from minor imperfections in low-resolution imaging to severe spatial resolution degradation, potential misinterpretation of the data, and even rendering the data unusable in super-resolution applications.

Approaches to address sample drift can be broadly categorized into two methods: active stabilization [121–123] and post-processing corrections [124–126]. Active stabilization focuses on real-time drift correction by continuously monitoring the sample's position, for example, by tracking fiducial markers, and by using mechanical or piezoelectric actuators to adjust the sample's position to counteract any drift. This approach can be implemented for all three axes or for axial direction only in scenarios where lateral drift is negligible or addressed through separate corrections [127]. Post-processing corrections address drift after image acquisition. Drift estimation can be achieved either through mathematical model estimation [128] or by imaging fiducial markers alongside the sample [121, 123]. Notably, without acquiring a full $3D$ image stack, axial drift cannot be effectively compensated during post-processing because of information loss due to de-focus. Indeed, axial drift leads to out-of-focus images, with inherent information loss as a consequence. Active stabilization is preferred in low-throughput SMLM techniques, such as minimal fluorescence photon fluxes microscopy (MINFLUX), RASTMIN, pMINFLUX, as well as ISM-FLUX [129, 130], where drift cannot be estimated robustly from sparse fluorescence events alone, albeit at the cost of relying on fiducials as a proxy for emitter motion. Recently Sun et al. introduced NanoDriftGuard, an open-source real-time stabilization pipeline that tracks non-fluorescent gold nanoparticles in a dedicated near-IR channel and drives an XYZ piezo stage for continuous 3D drift correction. Using correlation against a small reference stack (with intentionally defocused fiducial images for improved axial sensitivity), they report isotropic ångström-scale stability at $> 50\,\mathrm{Hz}$ and quantify residual inter-channel drift in long acquisitions [131].

The use of active stabilization techniques to enhance sample stability has become more widespread in recent years [127, 131–135]. Additionally, multiple commercial systems with axial stabilization capabilities have emerged [136–138]. However, the adaptability of these commercial solutions is often hindered by their limited compatibility with custom-built microscope systems. This has led to the development of numerous custom approaches for sample stabilization, broadly categorized into fiducial marker-based stabilization and reflection-based stabilization. Fiducial marker-based stabilization tracks the position of pre-introduced fiducial markers, such as gold nanoparticles (GNP), gold nanorods, or fluorescent beads, within the sample. Drift is continuously monitored through a dedicated imaging

pathway that tracks the position of these markers in real-time. Reflection-based stabilization makes use of the reflection of a focused beam off the coverslip. An axial movement of the sample leads to a change in the position or shape of the image of the reflected beam. In both cases, a feedback loop adjusts the sample position to actively compensate for any detected displacement. Each approach presents distinct advantages and limitations. Fiducial-based techniques offer the benefit of three-dimensional stabilization but necessitate the introduction of fiducial markers during sample preparation. Conversely, reflection-based techniques, while requiring no sample alteration, are restricted to axial correction. Consequently, fiducial-based stabilization strategies are often the preferred choice for low throughput SMLM methods, such as MINFLUX, which require real-time 3*D* stabilization.

A general limitation of fiducial-marker-based stabilization is that it implicitly assumes that the motion inferred from the tracked fiducials is representative of the motion of the fluorescent emitters of interest. In practice, this equivalence can be imperfect. For example, fiducials and fluorophores can reside at different axial positions, can be attached to different layers (glass surface versus cellular structures), and are measured in different optical channels with distinct wavelengths and filtering, such that channel-specific distortions or non-common-path effects can contribute to relative motion. In contrast, drift-correction methods that estimate drift directly from the fluorescence localizations (high-throughput post-processing) avoid this specific assumption because drift is inferred from the same signal used for reconstruction.

Several active 3*D* stabilization systems for custom-built microscopes currently exist [11, 101, 139]. While these approaches yield stabilization results around 1 nm for all axes, their implementation is either hindered by the need for a costly setup, including a PC with a suitable graphics card for GPU computing, a more complex sample preparation, such as the use of functionalized polystyrene beads, a limited feedback update rate ($\leq 20\,\mathrm{Hz}$), or the lack of open-source availability. So for our needs, we built a fast, open-source 3*D* active sample stabilization system that delivers similar performance but at a low-cost, requiring neither complex sample preparation nor extensive hardware installation. The setup is designed as a modular, standalone add-on that is compact and can be easily integrated into a custom microscope. We use this setup to keep the sample stable during ISM-FLUX measurements.

4.1.2 Widefield single-molecule localization microscopy add-on for correlative imaging

ISM-FLUX is designed for targeted, high-precision localization within a relatively small effective field-of-view (FoV). While this operating mode is well suited for interrogating

selected structures with nanometer-level precision, it provides limited contextual information about the surrounding sample architecture. For the next stage of this work, we therefore aim to extend the instrument from targeted precision toward multi-scale imaging within a single platform by adding a widefield SMLM modality. We integrate a camera-based widefield detection path which when combined with stochastic, time-separated blinking enables sequential localization of many closely spaced fluorophores and thereby supports super-resolved imaging over large, uniformly illuminated areas [9, 69, 140, 141]. In this combined workflow, widefield SMLM provides survey-scale context and a cross-comparison layer for reconstructed geometry and analysis choices, whereas ISM-FLUX refines selected regions of interest with higher photon efficiency and improved localization precision.

A key design goal of this integration is to enable correlative experiments without mechanical realignment. To that end, the widefield path is implemented on a shared optical train with ISM-FLUX; importantly using the same objective and tube (i.e., final imaging) lens. The excitation path is switched between modalities using a removable mirror. Sharing these core imaging optics helps in establishing a common sample coordinate frame and reduces the complexity of correlative measurements. ISM-FLUX, integrated with a single-photon avalanche diode (SPAD) array, allows for much faster temporal sampling because of the high detector readout speed relative to a standard scientific-complementary metal-oxide semiconductor (CMOS) camera (milliseconds vs nanoseconds). The SPAD array also supports asynchronous readout, which means that we can, in the future, use the same setup for fluorescence lifetime imaging. The widefield SMLM system will also allow for establishing multicolor protocols in future that can complement scanning-restricted FoV constraints, and it will provide a super-resolved reconstruction that can be used as a cross-comparison layer between both modalities for the sample's shape and structural geometry over a larger scale, drift control, and analysis choices.

The combined system of widefield SMLM and ISM-FLUX on a single, co-registered platform, is fully assembled and operational (Section 4.2.3), and its performance has been characterized as described in Section 5.3. Building on this foundation, we are planning correlative experimental workflows in which a widefield SMLM data set acquired with stochastic optical reconstruction microscopy (STORM) or DNA-based points accumulation for imaging in nanoscale topography (DNA-PAINT) is first used to survey a large FoV and identify regions of interest, followed by ISM-FLUX imaging of selected sub-regions at higher spatial resolution under active stabilization. While the individual modalities have been validated independently, the full correlative pipeline of combining survey-scale mapping with nanometer-scale interrogation in a single acquisition session has not yet been demonstrated

and represents the next experimental step. This will provide a unified workflow for contextual imaging, precision localization, and cross-modal cross-checks.

4.2 Hardware Implementations

4.2.1 Sample stabilization Setup

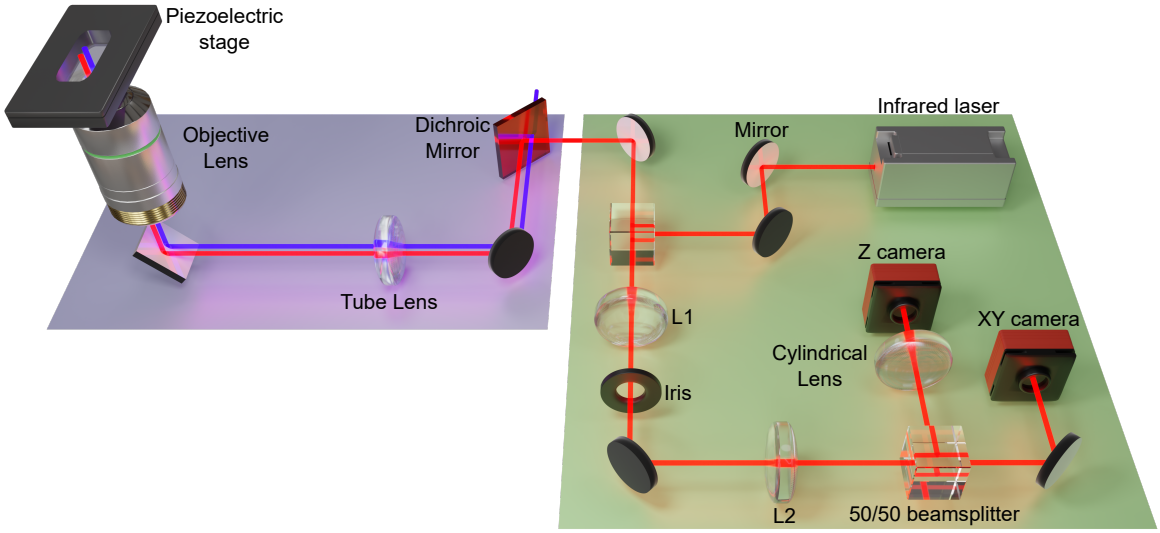

Figure 4.1 Stabilization system optical setup Optical setup of the stabilization system. The green part shows the stabilization setup, which is coupled to a microscope (partially shown in purple). The stabilization beam is coupled into the imaging path using a short-pass dichroic mirror, depicted in red in the imaging part.

The optical configuration is shown in Figure 4.1. An infrared laser (L780P010 - 780 nm, 10 mW, Thorlabs, New Jersey, USA) serves as the illumination source for the fiducial markers. The collimated beam is split by a 50/50 cube beam splitter (BS014, 700 nm to 1100 nm, Thorlabs) and coupled into the microscope with a near-infrared (NIR) short-pass dichroic mirror (AHF 745/SP BrightLine HC Shortpass Filter, F39-745). The short-pass filter is positioned between the scan lens and tube lens in the ISM-FLUX system (Section 4.2.2). An alternative configuration with the filter between the objective and tube lens is investigated in Section 4.2.3. By adjusting the mirrors positioned before the beam splitter, the beam is laterally displaced from the optical axis at the back aperture of the objective, leading to highly inclined laminated optical sheet (HILO) illumination after passing through the tube lens and objective lens (Nikon Plan Apo VC 100×/1.40 Oil OFN25 DIC N2). The HILO illumination enables imaging over a large FoV ($\sim 36 \times 36\,\mu m^2$), which allows tracking of

multiple particles simultaneously to improve robustness and precision (in our experiments, we use a concentration providing a number of particles between 5 and 15). The backscattered signal from the fiducial markers is collected by the same objective and reflected by the NIR dichroic mirror to enter the stabilization part again.

Table 4.1 List of essential components required for building the stabilization system

Part name	**Product code**
Infrared laser	***[ThorLabs]*** **L780P010** - 780 nm, 10 mW, Ø5.6 mm, A pin code, laser diode
Mirror (x5)	***[ThorLabs]*** **PF10-03-P01** - Ø1" Protected silver mirror
Beam splitter (x2)	***[ThorLabs]*** **BS014** - 50:50 Non-polarizing beam splitter cube, 700 - 1100 nm
Lens (x2)	***[ThorLabs]*** **AC254-200-B-ML** - f=200 mm, Ø1" Achromatic doublet, SM1-threaded mount, ARC: 650-1050 nm
Iris	***[ThorLabs]*** **ID15/M** - Mounted standard iris, Ø15.0 mm max aperture
Cylindrical lens	***[ThorLabs]*** **LJ1516RM-B** - f = 1000.0 mm, Ø1", N-BK7 mounted plano-convex round cyl Lens, ARC: 650 - 1050 nm
Camera (x2)	***[ThorLabs]*** **CS165MU1/M** - Zelux® 1.6 MP monochrome CMOS camera
Shortpass dichroic mirror	***[AHF]*** **FF01-745/SP-25** - 745/SP BrightLine HC Shortpass Filter

Here, the beam passes through the 50/50 cube beamsplitter and enters a 4f system composed of two lenses, L_1 and L_2 ($f = 200\,\text{mm}$ each). Both the tube lens and L_1, and L_1 and L_2 form a 4f system. In the conjugate plane between L_1 and L_2, an image of the back focal plane (BFP) is formed, and the light reflected by the cover slip is focused off-axis. An iris diaphragm (IDA25/M, Thorlabs) blocks the light reflected by the coverslip while transmitting most of the backscattered light. Finally, the signal is split with another 50/50 beam splitter and projected onto two CMOS cameras (CS165MU1/M - Zelux, Thorlabs), used to independently stabilize the lateral and axial direction, respectively. A cylindrical lens ($f = 1000\,\text{mm}$) induces astigmatism into the latter beam path to enable axial position estimation. Using two cameras instead of one effectively prevents aberrations caused by the cylindrical lens from introducing errors in lateral position calculations. The captured images from both cameras are transmitted to a central computer for further processing and

calculations. Drift corrections are sent to a three-axis nanopositioning piezoelectric stage (P-545.3R8S PInano XY(Z), E-727 Controller, Physik Instrumente).

To stabilize the sample, fiducial markers such as GNP can be used. Here, we use GNPs with a diameter ranging from 100 nm to 150 nm. The choice of GNPs stems from their several advantageous characteristics: readily available and cost-effective at various sizes, and capable of scattering laser light across a wide range of wavelengths [142]. The components required for building the stabilization are listed in Table 4.1.

4.2.2 ISM-FLUX Setup

We designed the ISM-FLUX setup from the ground up to remain as close as possible to a conventional ISM microscope, introducing only minor modifications. For excitation we used a 635 nm diode laser (LDH-D-C-635M; PicoQuant GmbH, Berlin, Germany) driven by a PDL 800-D controller for power control and operated in continuous mode. The laser output was passed through a Glan–Thompson polarizer followed by a half-wave plate to set the linear polarization. The beam was demagnified with a 2:1 telescope and then spatially cleaned by coupling into a single-mode, polarization-maintaining fiber (PM fiber). Because the source exhibits a transverse multimode profile, coupling efficiency into a single-mode fiber was limited to 7 % to 10 %, yielding ∼7 mW to 10 mW at the output of the fiber, which was sufficient for our single-molecule measurements.

After collimation, the beam traversed an achromatic half-wave plate and an achromatic quarter-wave plate specified for 500 nm to 900 nm (B. Halle, Berlin, Germany) to generate circular polarization. The beam then passed through a vortex phase plate (V-633-10-1; Vortex Photonics, Planegg, Germany) with topological charge $m = 1$, designed for 633 nm, to produce a donut-shaped excitation. For optimal performance, the input was circularly polarized and the phase plate was held perpendicular to the beam such that the back-reflection ($\sim 6.5\,\%$ of incident power) overlapped the incident beam. A better overlap of these two beams yields a better donut minimum.

A 1:1 relay telescope imaged the vortex plate to the galvanometric scanners via a 643 nm dichroic beamsplitter (F48-643; AHF, Germany). The vortex plate was placed in the BFP of the relay so that it was optically conjugate to the galvanometric scanners (see Figure 4.2), preserving the donut profile and preventing distortions during scanning. Scanning was performed with a pair of galvanometric mirrors (Saturn 1B 56S; ScannerMAX, Sanford, FL, USA) with small-step response $< 90\,\mu s$. The scanner relay was coupled to a Nikon scan lens ($f = 60\,mm$) and a Nikon tube lens ($f = 200\,mm$). A near-infrared dichroic between the scan

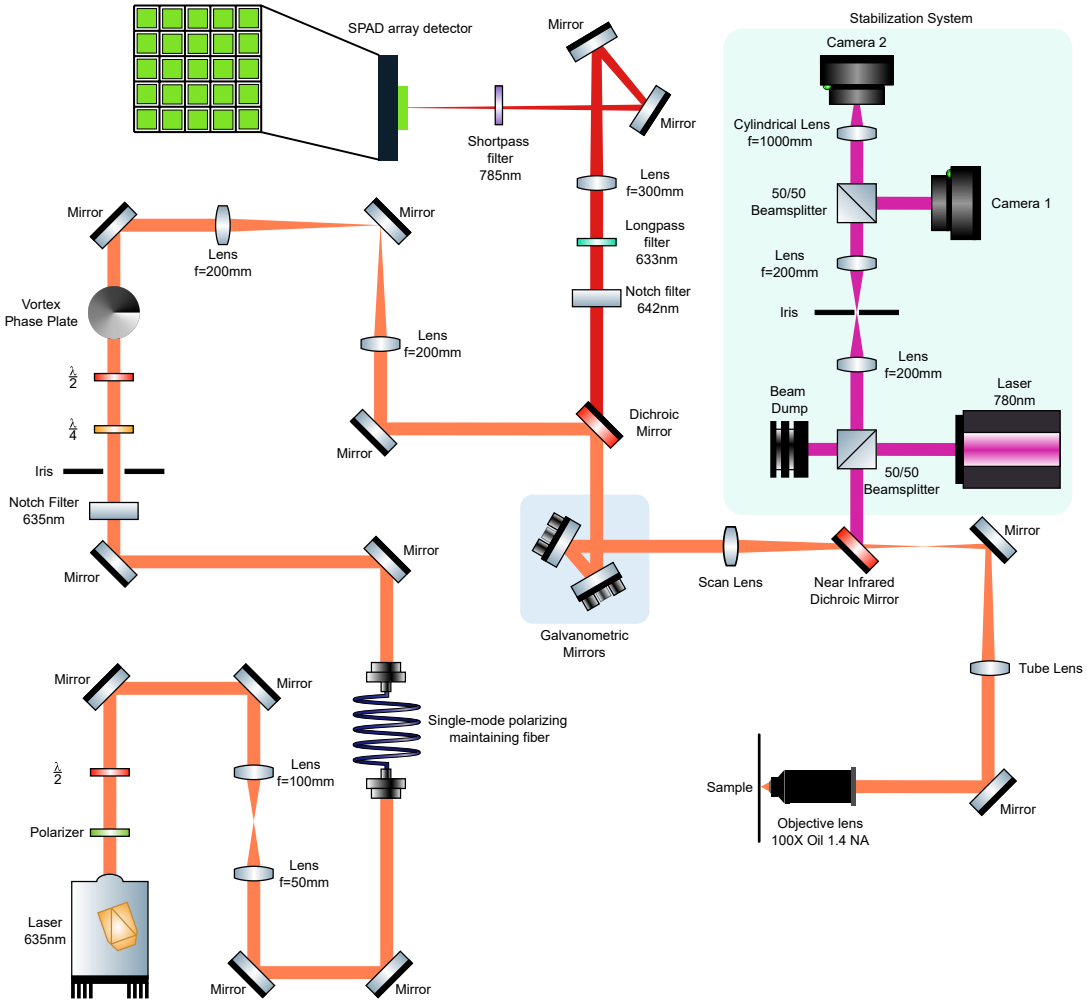

Figure 4.2 Detailed ISM-FLUX optical setup $\frac{\lambda}{2}$ and $\frac{\lambda}{4}$ denote the half-wave and quarter-wave plate. Near infrared dichroic mirror used in this system is Shortpass 750nm. The optical setup highlighted in green is the stabilization system used to keep the sample in focus during the imaging session.

and tube lenses coupled the sample-stabilization channel into the ISM-FLUX path (detailed in section 4.2.1 and 4.3.1). All measurements used a Nikon Plan Apo VC $100\times/1.40$ oil-immersion objective (OFN25 DIC N2).

Fluorescence from the sample was de-scanned by the galvanometric mirrors and transmitted by the dichroic into a clean-up filter set: BSP01-785R-25 short-pass ($\lambda_{\text{cut}} = 785\,\text{nm}$), a 642 nm notch filter, and LP01-633R-25 long-pass ($\lambda_{\text{cuton}} = 633\,\text{nm}$). This train transmitted fluorescence while rejecting residual reflected excitation and stabilization laser beams. The emission was then focused with a 300 mm lens (AC254-300-A-ML; Thorlabs, Germany),

placed conjugate to the scan lens, onto a 5×5 SPAD array detector (PRISM Light; Genoa Instruments, Genoa, Italy). The resulting detection aperture was ~ 1.4 Airy units at the detector, so the array served as an effective confocal pinhole to suppress out-of-focus background. The sample was stabilized throughout the entire data acquisition using the aforementioned sample stabilization system.

4.2.3 Widefield single-molecule localization microscopy optical setup

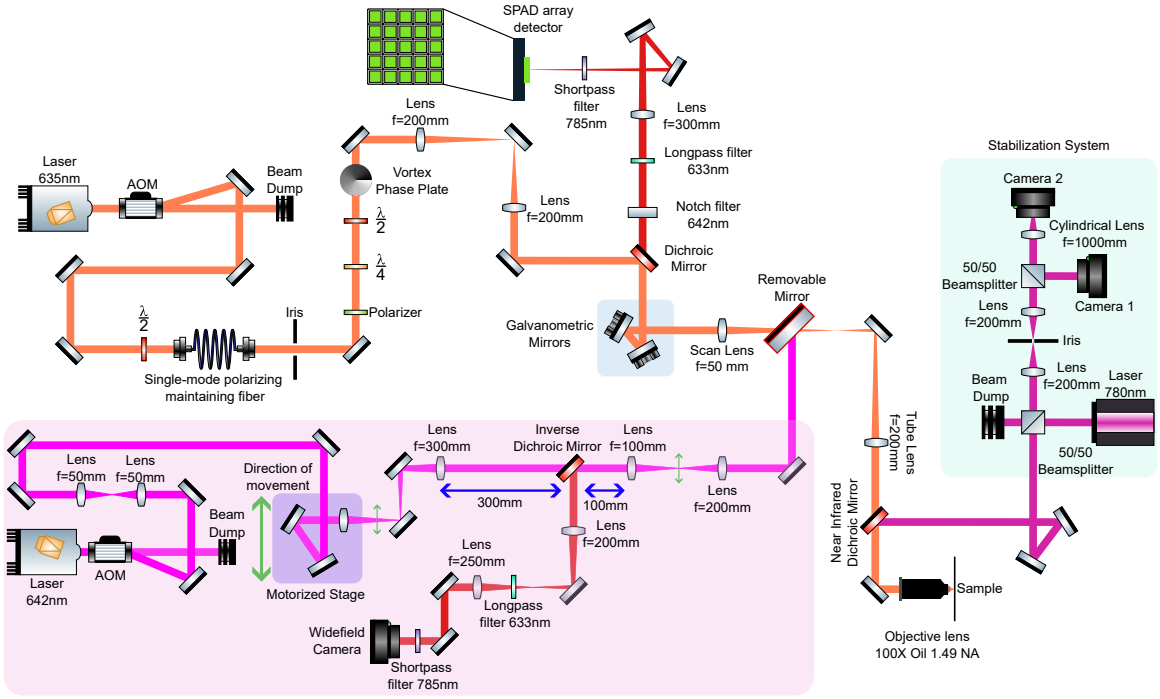

Figure 4.3 Detailed Widefield SMLM optical setup integrated into ISM-FLUX. The components highlighted by the pink box and the optical path denoted by the pink beam shows the Widefield SMLM setup. Widefield SMLM path intersects the ISM-FLUX path at the removable mirror.

To enable correlative super-resolution imaging, a widefield SMLM path was integrated with the existing ISM-FLUX setup. Alongside the building of the widefield SMLM setup, we also rebuilt the ISM-FLUX setup to further improve its performance. We added an acousto-optic modulator (AOM) to the ISM-FLUX setup as well to have a granular control over the output power. The hybrid design facilitates sequential imaging of identical regions of interest without mechanical realignment, thereby minimizing experimental artifacts and enabling direct comparison of localization precision between modalities. Furthermore, as described in Section 4.2.1, we repositioned the near-infrared short-pass dichroic to the space between

the objective and the tube lens. In this configuration, the dichroic introduces less distortion into the ISM-FLUX excitation path, thereby reducing aberration-induced degradation of the excitation focus. The widefield illumination source is a continuous-wave (CW) visible fiber laser (VFL-P-1000-624-OEM1, MPB Communications Inc., Quebec, Canada) operating at 642 nm. The laser delivers up to 1 W output power with a beam quality factor $M^2 < 1.1$ and maintains long-term wavelength stability of ± 0.1 nm. Laser power control is achieved via a graphical user interface (GUI-VFL V4.0.1) through laser driver. The laser exhibits a thermal transient during power adjustments, requiring up to 30 min to stabilize. To enable real-time imaging and sample scouting, an AOM (MTS110-A3-VIS, AA Opto-Electronic, Orsay, France) was inserted into the beam path. The AOM provides sub-microsecond modulation bandwidth (nominal rise time $< 1\,\mu$s) and achieves an extinction ratio of > 45 dB, allowing both rapid power gating and fine intensity control without perturbing the laser's thermal equilibrium.

The laser output first passes through a 1:1 Keplerian telescope constructed from two 50 mm lenses. This telescope collimates the divergent fiber output, and establishes a defined beam diameter of 2 mm. The telescope relays the beam waist to the subsequent translation stage. A motorized linear stage (Z925B, Thorlabs, New Jersey, USA) with 25 mm travel range translates the beam parallel to the optical table (Figure 4.3). Translating the excitation beam in a plane conjugate to the objective BFP shifts the incidence angle at the coverslip, enabling continuous adjustment between epifluorescence, HILO, and total internal reflection fluorescence (TIRF) illumination. The stage incorporates a 60 mm focal length lens and steering mirrors to ensure the beam remains centered on all subsequent optical elements when used in the epi imaging mode. Following translation, the beam undergoes two-stage expansion. A 300 mm focal length lens provides $5\times$ magnification, increasing the beam diameter to ~ 10 mm. This expanded beam is incident upon an inverse dichroic beamsplitter (T650 SPXR, AHF, Germany) that transmits the 642 nm excitation while reflecting fluorescence ≥ 662 nm. The inverse configuration was deliberately chosen because of our mechanical constraints of maintaining optical compatibility with the ISM-FLUX setup. The dichroic is positioned in a conjugate plane to prevent beam clipping during translation. A subsequent 1:2 telescope provides final $10\times$ beam magnification before injection into the ISM-FLUX path via a removable magnetic mirror.

A tube lens ($f = 200$ mm) focuses the beam at the BFP of a $100\times$ oil-immersion TIRF objective (CFI Apochromat, NA 1.49, Nikon). We changed the objective lens for this setup in order to be compatible with high-power widefield illumination. Fluorescence emission is collected through the same objective, reflected by the dichroic beamsplitter, and magnified

1.25× before spectral filtering. The filter set comprises: (i) a shortpass edge filter (F76-789, $\lambda_{\text{cut}} = 785\,\text{nm}$), (ii) a laser beamsplitter (F48-640, 640 nm), and (iii) a longpass filter (F76-635, $\lambda_{\text{cut-on}} = 635\,\text{nm}$). This combination suppresses excitation laser scatter and residual 780 nm infrared stabilization light. The total system magnification is 67.5×, projecting $\sim 96\,\text{nm}$ per pixel onto the Hamamatsu Orca Flash4.0 V3 camera sensor (native pixel size 6.5 µm). The modular optical design permits switching between ISM-FLUX and widefield SMLM modalities without realignment. This capability has been verified experimentally and will be exploited in future correlative measurements.

4.3 Control and Data Acquisition Software

For sample-stabilization control, we use custom Python software that coordinates hardware control and data acquisition for the entire stabilization subsystem. For imaging and data acquisition, we use two software systems: BrightEyes-MCS and custom-written LabVIEW software. The galvanometric scan mirror drivers and SPAD array detector were connected to a field-programmable gate array (FPGA) module (NI PXIe-1071 with an NI PXIe-7856R board, National Instruments, USA) that was controlled with the BrightEyes-MCS software to acquire the system point spread function (PSF) and to scout and image samples for general testing and with a custom-written LabVIEW software to perform orbital scanning for ISM-FLUX. The same PC (Dell Precision 5820 Tower X-Series) was used for data acquisition and controlling the sample stabilization. For widefield/STORM scouting and data acquisition, we used a separate PC dedicated to the camera readout. Both scouting and acquisition were performed using Hamamatsu TokuPic software (Hamamatsu Photonics, Japan).

4.3.1 Stabilization Control Software

The stabilization system integrates a central computer (same as ISM-FLUX) interfacing with the CMOS cameras and the nanopositioning piezoelectric stage. The computer stabilizes the sample by imaging the fiducial markers, performing real-time $3D$ localization of the particles, and repositioning the stage accordingly. The software is developed in Python and runs entirely on the CPU. Figure 4.4 illustrates the pseudocode for the stabilization software. We use multiprocessing to parallelize fiducial marker localization in the lateral direction, accelerating drift calculations. The decision to forgo GPU utilization ensures broad compatibility across a diverse range of systems, including systems equipped solely with integrated GPUs or non-NVIDIA dedicated options, where GPU utilization might not

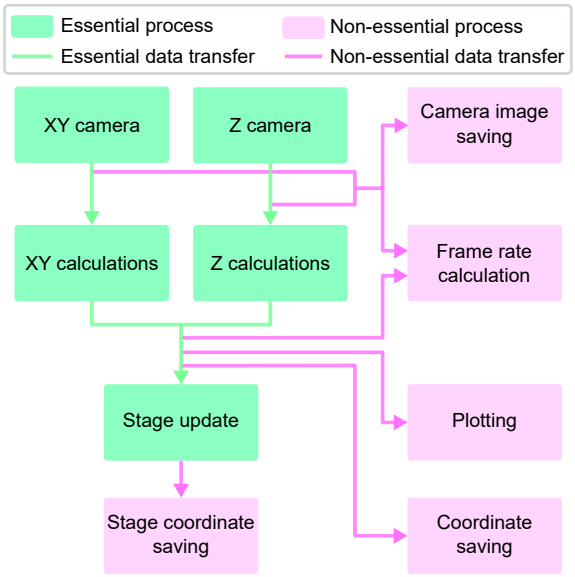

Figure 4.4 Sample stabilization code workflow. Flowchart with pseudocode describing the stabilization code that controls the hardware. The code is split into 10 different processes that are categorized into two types. The first type (green) represents the essential processes required for the stabilization code. The second type (pink) represents the non-essential part of the stabilization code. Similarly, the green arrows show the essential data transfer, and the pink arrows show the non-essential data transfer.

offer a performance advantage and could even impede processing speed due to data transfer overhead.

Before initiating the stabilization process, the software connects to the two CMOS cameras and the nanopositioning piezoelectric stage. All three pieces of hardware require dynamic-link library (.dll) files to establish a connection using the USB 3.0 protocol. These .dll files, provided by the proprietary control software of the cameras (Thorlabs) and the piezoelectric stage (Physik Instrumente), can be installed by the user via a straightforward Windows installation wizard. After the hardware connection, the user can crop the FoV for each camera independently, Figure 4.5 (a), balancing a larger FoV to capture more particles with a smaller FoV to achieve a higher frame rate. Subsequently, the user selects a chosen number of particles within this FoV for lateral stabilization. The software then performs a z-scan using the same piezoelectric stage employed for stabilization to measure the PSF shape at various axial positions, Figure 4.5 (a).

Then, depending on the selected settings, up to 10 processes are initiated, 5 are essential for actively stabilizing the sample (one process per camera for image capture, one process per camera for image analysis, one process to control the stage, and one process for real-

time plotting), while the remaining 5 are optional (for miscellaneous tasks, such as data saving, statistical calculations, and data logging) and can be toggled on or off depending on experimental requirements. This implementation accommodates processes running at different speeds, allowing capturing images at a high frame rate (*e.g.*, $\sim 120\,\mathrm{Hz}$), even when stage updates occur at a lower frequency ($\sim 20\,\mathrm{Hz}$). Disabling the non-essential processes can enhance acquisition and calculation speed when the CPU is fully utilized.

All 10 processes operate in parallel, utilizing a shared memory space allocated at the start of the program to facilitate simultaneous data access. In this way, the data generated by each process is shared with the other processes. Within this shared memory, data is classified into two types: new data (indicated by green arrows) generated with each process iteration and duplicate data (indicated by pink arrows), which is used for data saving and logging purposes. A potential challenge arises when multiple processes attempt to access the shared memory concurrently for reading or writing, which can lead to memory corruption and disrupt the stabilization program. To address this issue, we have implemented a memory lock mechanism. The memory lock is implemented in such a way that whenever a process needs to read from or write to the shared memory buffer, the memory is locked by that particular process handler so that no other process can access that memory space. While the memory lock introduces a minor time penalty of less than 100 µs per iteration, it enables efficient data exchange between processes without the need for explicit inter-process communication (IPC) mechanisms, ensuring stable and uninterrupted program operation.

Positional deviations in $3D$ are assessed by comparing measured particle locations with the reference points established at the start of the measurement. Detected deviations trigger corresponding negative feedback commands sent to the piezoelectric stage. This closed-loop control system actively returns the sample to its relative position with respect to the stabilization beam path. In more detail, the software reads the cameras and stores the last 25 frames in memory. A separate process takes, for each camera, the weighted average of these frames, with the highest weight for the latest frame and an exponentially decreasing weight for earlier frames. For lateral drift correction, the particles are localized, and their average drift from the reference points is stored in memory. A separate process checks with a fixed rate ($20\times$ per second) whether or not the measured average drift exceeds a threshold of 1 nm and, if necessary, updates the lateral position of the stage. Axial drift is assessed by comparing the PSF shape in the averaged frame with the reference curve obtained during the initialization procedure. Similarly to the lateral drift, a separate process updates the axial position of the stage with a fixed rate of 20 Hz.

The software is engineered to optimize the balance between computational speed and memory usage, ensuring peak performance during background operation. To evaluate its capabilities, we tested the stabilization software on a laptop equipped with an Intel i7-9750H CPU. The laptop achieved comparable results, albeit with increased CPU utilization. This heightened consumption is attributable to the laptop CPU's lower clock speed and instructions per cycle. The system's response time for a given axis exhibits variability depending upon the precise code execution stage within the process runtime. In a worst-case scenario, a response may necessitate a duration of two "if check cycles" if the fluctuation is detected immediately following a completed check for that specific axis. Conversely, an immediate response can be achieved under optimal conditions if the fluctuation is detected concurrently with the axis check. Parallelization of the "if checks" is infeasible due to the inherent serial nature of the input command, restricting the piezoelectric stage to the execution of a single instruction at any given time.

During corrective movements, the piezoelectric stage exhibits inertia, leading to an overshoot of the target sample position. The piezoelectric stage incorporates an internal feedback loop with a proportional–integral–derivative (PID) controller, which corrects the position error using a weighted combination of the instantaneous error (P), the accumulated error over time (I), and the rate of change of the error (D), thereby reducing overshoot caused by the stage's inertia. However, in scenarios characterized by elevated environmental vibrations, we have observed persistent stage movement oscillations despite this corrective mechanism. To prevent oscillatory behavior caused by overshooting, we multiply the measured drift for each dimension with a factor α ($0 < \alpha \leq 1$). This gain reduction allows approaching the set point gradually, *i.e.*, in multiple stage movement cycles, which improves the stability. This calculated reduction accounts for the inertial effect, ensuring the final stage position aligns with the desired value. Notably, the magnitude of this signal reduction remains an adjustable parameter, enabling users to fine-tune it based on specific hardware configurations and experimental parameters. The software also logs all the data that is generated while the stabilization is running. The user can use this data in postprocessing for further corrections or simple experimental data verification (Section 4.5.3).

Finally, the stabilization software additionally implements a session-resume capability. At the start of each experiment, it stores the complete set of configuration parameters required to re-establish the previous operating state in a dedicated configuration file (e.g., camera exposure times, lateral and axial fields of view/ROIs, selected tracking particle ROIs and their positions, and the axial FFT-based calibration factor). When stabilization is terminated and later restarted, the user is prompted to either reload the previously saved configuration or

repeat the full calibration workflow from scratch. Reloading the configuration automatically restores the camera ROI settings and the calibration parameters, enabling rapid continuation of stabilization at the same sample location with minimal interruption, thereby reducing both downtime and drift-induced displacement.

4.3.2 BrightEyes-MCS (Microscope Control Software)

BrightEyes-MCS is an open-source software suite developed in our lab for multichannel scanning microscopy and data acquisition [143]. The architecture comprises two layers: low-level firmware on a National Instruments FPGA board and high-level Python software on a host computer. The firmware, written in NI LabVIEW for FPGA, handles hardware interfacing and data streaming through FIFO buffers. The high-level software automatically loads the firmware onto the FPGA, configures operational registers (e.g., pixel count, dwell time), and processes streamed data. The system supports 25 primary digital detector channels; for the experimental configuration, an NI USB-7856R FPGA board was paired with an Analog Devices EVAL-AD5764 external digital-to-analog converter (DAC) evaluation board. Minimum dwell time per pixel is 1.0 µs, optionally subdivided to 0.5 µs per time bin. Scanning control is implemented through analog outputs that can be connected to linear galvanometer mirrors and piezoelectric scanners. Each of the eight analog outputs can be configured as X, Y, or Z scanning signals, or as constant-voltage sources for peripherals such as AOM. Two analog inputs accommodate the photomultiplier tube (PMT) detection.

In our experiments, BrightEyes-MCS was used for routine confocal/ISM-style acquisitions, calibrations and, day-to-day operation of the microscope. We used it to acquire raster-scanned images for sample scouting and stabilization-control measurements. In addition, we recorded SPAD-array datasets to characterize the system PSF and to generate the molecule detection function (MDF) calibration data required for ISM-FLUX localization. The software was also used to image the samples using a PMT during the setup prototyping and building process. Although BrightEyes-MCS also supports orbital scanning, circular scans for ISM-FLUX were implemented with dedicated custom control to meet the specific synchronization requirements of the targeted coordinate pattern and data readout.

4.3.3 Custom LabVIEW software for ISM-FLUX

A custom LabVIEW program developed in our lab is used to configure and execute the ISM-FLUX orbital scan at a high level. The program allows the user to specify the orbit geometry and acquisition parameters directly in sample units (µm), including the targeted

coordinate pattern (TCP) diameter (typically $L = 90\,\text{nm}$), the number of discrete orbit points, the per-point photon binning configuration (time resolution and number of time bins per point), and the orbit center. Time resolution (Δt) refers to the bin width used to discretize photon detections for each orbit point (i.e., the temporal sampling of the detector signal during the dwell time), while the number of time bins (N_{bins}) per point specifies how many such temporal bins are recorded per orbit point (See Section 4.5.3). The orbit center is set through independent X/Y offsets, implemented as DC voltage offsets to the galvo driver inputs that translate the scan pattern to the desired measurement location. In addition, the user selects the number of orbits to acquire (typically 300 000), which defines the total measurement duration together with the dwell settings.

During acquisition, the galvos are driven using a sinusoidal input voltage and photon counts are binned into $N_c = 32$ angular positions per orbit for analysis and synchronization. Each discrete position corresponds to one voltage pair (V_x, V_y) sent to the galvo drivers. The voltages are computed from the user-defined orbit parameters in µm using a pre-measured galvo calibration factor (conversion from displacement to voltage). Photon timing is recorded as a fixed-length time-trace per scan position with N_{bins} bins. The effective dwell time per scan position is given by $t_{\text{dwell}} = \Delta t \times N_{\text{bins}}$, where Δt is the time resolution (in µs) and N_{bins} is the number of time bins stored per orbit position.

4.4 Sample Preparation Protocols

4.4.1 Sample preparation for active sample stabilization

Gold nanoparticles samples for stabilization experiments were prepared using 100 nm diameter, OD 1 GNP (Sigma-Aldrich, Product No. 742031). The GNP suspension was sonicated for 15 min to reduce aggregation prior to deposition. Glass coverslips were cleaned by rinsing three times with 70 % ethanol, then coated with Poly-L-Lysine (Sigma-Aldrich, Product No. P4832) by incubation for 30 min at room temperature, followed by rinsing with ddH_2O. The sonicated GNP solution was drop-cast onto the Poly-L-Lysine-coated coverslips and allowed to settle for 10 min to 15 min before being sealed and left to dry under room temperature.

4.4.2 Sample preparation for ISM-FLUX calibration and imaging

The same GNP sample was used for ISM-FLUX imaging as well. The sample for single fluorophore measurements contained passively adsorbed antibodies on an untreated glass

surface in an 80 % glycerol concentration. The stock solution of anti-mouse IgG-Atto647N (50185-1ML-F, Sigma Aldrich, Steinheim, Germany), was initially diluted in PBS, and for the final single-molecule concentration in a home-made 80 % (w/w) glycerol solution (G5516-500ML, Sigma Aldrich, Steinheim, Germany). A coverslip of the dimension of a microscope slide (24 x 60 mm, No. 1.5H, Paul Marienfeld GmbH, Lauda-Königshofen, Germany) was cleaned with mQ water and mounted into the microscope. Right before the start of the measurement, 20ul of the antibody-glycerol solution was dropped onto the coverslip. Immediately afterward, antibodies that were adsorbed or moving on the coverslip were visible, as well as in a solution. Measurements were performed only on non-moving antibodies. To evaluate the single fluorophore signal, only the signal from the last intensity step before complete photobleaching was evaluated.

The DNA origami, with DNA-PAINT labeling, was commercially purchased from GATTAquant (GATTAquant GmbH, Munich, Germany). Both nanorulers used for imaging experiments consisted of three collinear binding sites with 40 nm spacings ((GATTA-PAINT HiRes NP 40) or 20 nm spacings (GATTA-PAINT HiRes NP 20R), and imaged using Atto655-labeled imager strands. The average binding time (ON-time) of the origami was $\sim$ 100 ms.

4.4.3 Sample preparation for widefield single-molecule localization microscopy

For single-molecule localization microscopy, we prepared two samples: one for DNA-PAINT and one for STORM. Both samples were prepared using HeLa cells. HeLa cells were fixed 24 h after plating. Cells were washed once with sterile 1$\times$ Dulbecco's phosphate-buffered saline (DPBS; Sigma Life Science) and incubated in a fixation solution containing 0.2 % glutaraldehyde and 0.2 % formaldehyde in DPBS for 15 min at room temperature. After fixation, cells were washed three times with 1$\times$ PBS and stored in 1$\times$ PBS supplemented with 0.1 % sodium azide at 4 °C until immunolabeling.

Immunolabeling for dSTORM For blocking and permeabilization, samples were incubated for 30 min at room temperature in PBS containing 3 % bovine serum albumin (BSA) and 0.1 % Triton X-100. Primary antibodies (Anti-TOMM20 antibody [EPR15581-54], Abcam Limited, Cambridge, United Kingdom) were diluted in the same buffer (typical dilution 1:500) and incubated with the cells for 60 min at room temperature; for multi-target imaging, primary antibodies were applied simultaneously. Samples were washed three times

in PBS containing 3 % BSA and 0.1 % Triton X-100 and then incubated for 60 min at room temperature with secondary antibodies conjugated to dSTORM fluorophores (typical dilution 1:500; secondary antibodies applied simultaneously for multi-target imaging). Following immunolabeling, samples were post-fixed in 3.7 % formaldehyde in PBS, washed three times with 1× PBS, incubated in ammonium chloride solution (2.7 g L^{-1} NH_4Cl in PBS), and washed three additional times with 1× PBS. Coverslips were stored in 1× PBS containing 0.1 % sodium azide at 4 °C until imaging.

Mounting for dSTORM imaging Immediately before imaging, coverslips were mounted in EverSpark 2.0 dSTORM imaging buffer (Idylle Labs, Paris, France). An adhesive square spacer was placed around the central region of a curved microscope slide to define the imaging chamber. A 200 µL droplet of imaging buffer was placed on the slide, and the coverslip was inverted onto the droplet. The buffer vial was resealed immediately after opening to minimize oxygen ingress. Trapped air bubbles were removed prior to sealing. The chamber was sealed using a two-component sealant.

Immunolabeling for DNA-PAINT DNA-PAINT immunolabeling followed the dSTORM protocol described above, with the following modifications. Secondary antibodies were conjugated to DNA-PAINT docking strands and were diluted in DNA-PAINT antibody buffer (Massive Photonics; typical dilution 1:500). After labeling, coverslips were stored in DNA-PAINT washing buffer (Massive Photonics) at 4 °C until imaging.

Before imaging, DNA-PAINT imager strands were diluted in DNA-PAINT imaging buffer (Massive Photonics) to a final concentration in the sample chamber in the range 0.1 nmol to 0.7 nmol. For drift correction, 100 nm GNPs (Sigma-Aldrich, 742031) were drop-cast onto both samples at the stock concentration (no dilution) until a final density of 5 to 10 GNPs per 30 × 30 μm^2 was reached.

4.5 Data Analysis

4.5.1 Stabilization Calibration and Position estimation

Position Estimation

Inclined (HILO) illumination plays a crucial role in enabling the simultaneous illumination and subsequent selection of numerous particles for stabilization purposes. Selecting multiple

particles minimizes the influence of suboptimal particles, potentially inadequately immobilized on the surface and thus hindering optimal sample stabilization. Our stabilization software allows the users to select particles prior to stabilizing the sample (as illustrated in Figure 4.5 (a)). Notably, averaging the positions of multiple particles leads to a reduction in errors associated with pinpointing the exact stage position.

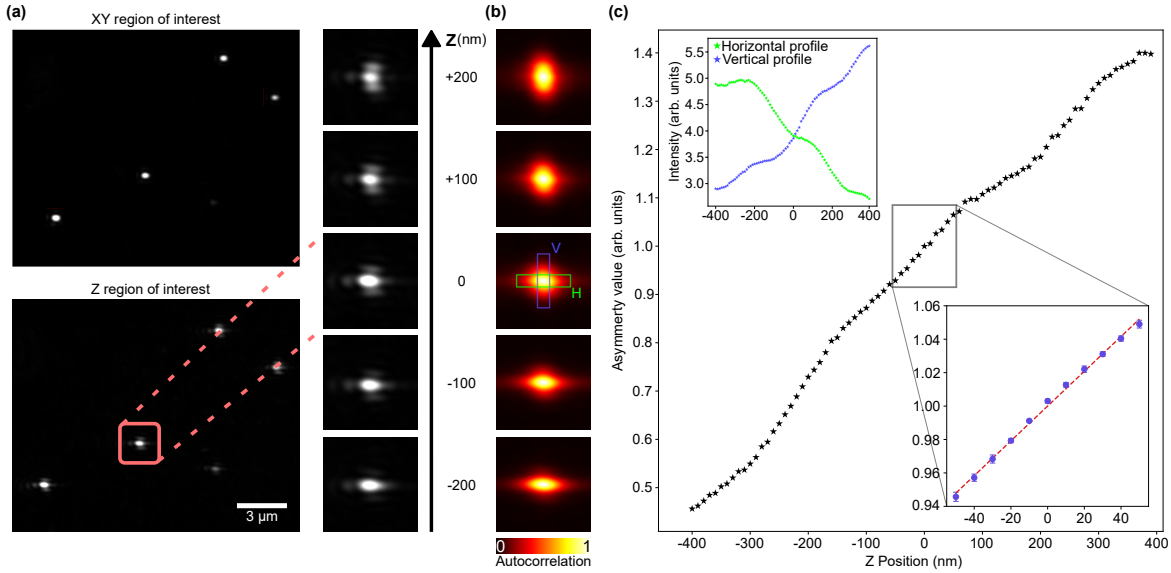

Figure 4.5 Stabilization Axial Calibration. a) Exemplary camera images of the XY and Z camera. The column on the right shows the astigmatism induced by the cylindrical lens, as imaged by the Z camera at different axial planes. **b)** Autocorrelation of the images captured by the Z camera for the same z positions as in **a)**. The axial position of the sample is derived from the asymmetry of the autocorrelation function, calculated as the ratio of the summed intensities of the pixel values in H (green) and V (blue). **c)** Measured autocorrelation asymmetry as a function of the z position. The calibration measurement performed at the start of each experiment covers a smaller range (from $-50\,\mathrm{nm}$ to $50\,\mathrm{nm}$, see inset), where the asymmetry can be well approximated as a linear curve (dotted line). The calibration curve and the error bars are the mean and standard deviation over ten iterations. The inset plot in the top left shows the horizontal and vertical components (shown in **b)**) of the summed intensities over the green and blue boxes

In the stabilization loop, we estimate drift from the apparent motion of the selected fiducial markers and use it as a proxy for the sample motion experienced by the emitters during the measurement. This approach is effective when fiducials and emitters are mechanically coupled to the same substrate or structure. We estimate the lateral drift in the camera frame of reference $\mathbf{d}(t)$ at time point t by averaging the measured changes in the $2D$ position of N fiducial markers compared to the start of the measurement: $\mathbf{d}(t) = \frac{1}{M}\sum_{m=1}^{M}(\mathbf{x}_m(t) - \mathbf{x}_m(0))$, with $\mathbf{x}_m(t)$ position of marker m at time t. We use the local gradients method [144] to measure $\mathbf{x}_m(t)$ because of its computational efficiency compared to fitting a Gaussian function.

Conversion of drift from pixels to nanometer is done based on a calibration measurement with a reflective grid sample (R1L3S3PR, ThorLabs).

Axial drift estimation is based on astigmatism introduced by the cylindrical lens positioned in front of the camera and comparing the shape of the PSF with respect to the start of the measurement. To estimate the relative z-position, we calculate the autocorrelation G in a user-selected region of the image, ideally containing multiple particles for enhanced robustness and precision (Figure 4.5 (b, c)). We define a measure for the PSF shape A as the ratio $A = \sum_i G_i / \sum_j G_j$ with $i \in V$ and $j \in H$, with H and V perpendicular rectangular regions centered around the autocorrelation peak. Conversion from A to axial displacement in nanometers is based on a reference curve measured at the start of every measurement (Figure 4.5 (c)).

Axial Calibration

The axial calibration involves sampling multiple planes in both the positive (+z) and negative (-z) directions with respect to the reference plane. At each sampled plane, we calculate the corresponding A value. These values are then fitted using linear regression (see the inset of Figure 4.5 (c)). The slope of the fit represents the conversion factor $\Delta A/\Delta z$, expressed in units per nanometer. We typically perform the calibration procedure for a small range (-50 nm to 50 nm) with respect to the initial z position with multiple (eg. 9 or 11) sampling places involved in between. This allows us to approximate the calibration curve by a linear curve. Appendix B contains more details on the expected axial localization uncertainty using this method.

The astigmatism-based approach offers distinct advantages for axial stability by facilitating high-precision, long-range sample stabilization. By leveraging auto-correlation, we can establish a linear fit to astigmatism values across a substantial z-range of ~ 800 nm (-400 nm to 400 nm), as demonstrated in Figure 4.5 (c). This extensive range allows the system to achieve stable Z-plane stabilization throughout this Z-depth, granting users the flexibility to conduct experiments at any desired Z-plane with assured stabilization.

4.5.2 ISM-FLUX molecule detection function and orbit calibrations

Measuring the molecule detection functions

As mentioned in section 3.3.1, ISM-FLUX data analysis requires a set of $N_c \times N_d$ MDFs, with N_c the number of illumination positions and N_d the number of detector elements in the array. The MDF describes the position-dependent probability to excite and detect a fluorescence

photon from a molecule at a given position within the sample for a given position in the TCP and a given detector element [145]. We measured the MDFs in two steps. First, we measured $I_{0j}(\mathbf{r}_E)$ by scanning over a single GNP with the galvo scan mirrors and detecting the signal with the SPAD array detector, Figure 4.6 (a). $I_{0j}(\mathbf{r}_E)$ denotes the (reference) detector image: the expected photon counts recorded on SPAD array element j for an emitter at position $\mathbf{r}_E$.

The scan settings were FoV $2\times2\,\mu m^2$, pixel size $2\,nm^2$, and a scan speed corresponding to 15 s to 30 s per image. We averaged over 5-6 images while the stabilization software kept the particle in place. The laser power was adjusted to keep the total detected photon count rate (sum 5×5) around 1 MHz in the doughnut maximum. The 2 nm pixel size was chosen as a trade-off between better localization uncertainty (with a small pixel size) and smaller MDF data size (with a bigger pixel size). Due to small day-to-day changes in the experimental setup, we measured $I_{0j}(\mathbf{r}_E)$ before each series of experiments. $I_{0j}(\mathbf{r}_E)$ takes the form of 1000 pixel $\times$ 1000 pixel $\times$ 25 detector elements. We noted that for small FoVs, the images were stretched in the x-direction (the fast scan axis) due to the limited precision of the galvo scanner, Figure 4.6 (a). Therefore, we squeezed the images by 20 % in the x-direction in post-processing using the resize function from the scikit-image package in Python. The second step is using $I_{0j}(\mathbf{r}_E)$ to generate $I_{ij}(\mathbf{r}_E)$ by using the orbital scan calibration. Knowing the radius and starting angle or the orbit and the direction of rotation, we shifted the 25 images $I_{0j}(\mathbf{r}_E)$ in post-processing to each of the 32 scan positions, resulting in a full MDF data set, containing 1000 pixel $\times$ 1000 pixel $\times$ 800 images, Figure 4.6 (b). For measurements in which the two hot detector elements were excluded, the MDF data set contained 736 (23×32) images. All images were smoothed with a combination of a 2D uniform filter from SciPy and a GaussianBlur from OpenCV-Python until visually smooth.

Simulation of the MDFs and calculation of the CRB

To simulate $I_{0j}(\mathbf{r}_E)$, we used the BrightEyes-ISM [46] and s^{2}ISM [47] Python packages. First, we acquired a conventional ISM dataset of a fixed HeLa cell stained for α-tubulin (details in [146]). The experimental shift vectors were obtained through phase correlation of the images in (a) with the image from the central detector element (12), Figure 4.7. We assumed no circular symmetry-breaking aberrations and a detector pitch smaller than one airy unit to ensure evenly-spaced shift vectors. From the shift vectors of the images generated by the inner 3×3 detector elements, we fitted the orientation, rotation, and magnification of the system. Using the fitted parameters, we simulated a set of 25 MDFs assuming 635 nm excitation wavelength, 660 nm emission wavelength, NA 1.4, refractive index 1.5, vortex phase plate as a mask (angular momentum 1), pixel pitch and pixel size of the 5×5 array

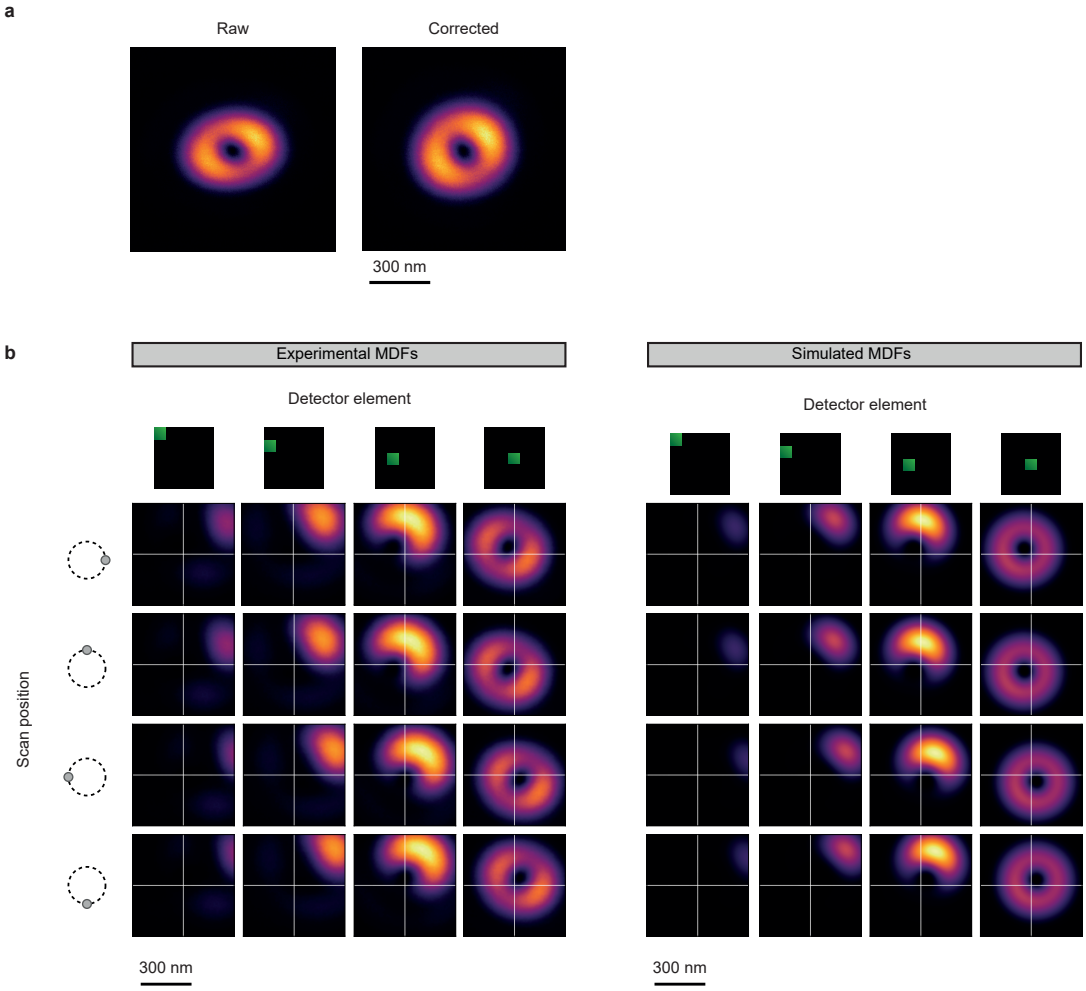

Figure 4.6 Molecule detection function for ISM-FLUX localization acquired using Gold nanoparticles. a) Experimental MDF for the central detector element, obtained by scanning a GNP. Scan settings: 2000 × 2000 pixel, 2 nm/pixel, pixel dwell time 30 μs, 10 repetitions. Galvanometric limitations lead to an apparent astigmatic image, corrected in post-processing. **b)** Subset of **a)** experimental and **b)** simulated MDFs. The full MDF data set is a collection 25 x 32 images, a combination of 25 detector elements and 32 scan positions. For simplicity, we show 4 positions and 4 detector elements. The experimental MDFs are stretch-corrected and smoothed in post-processing.

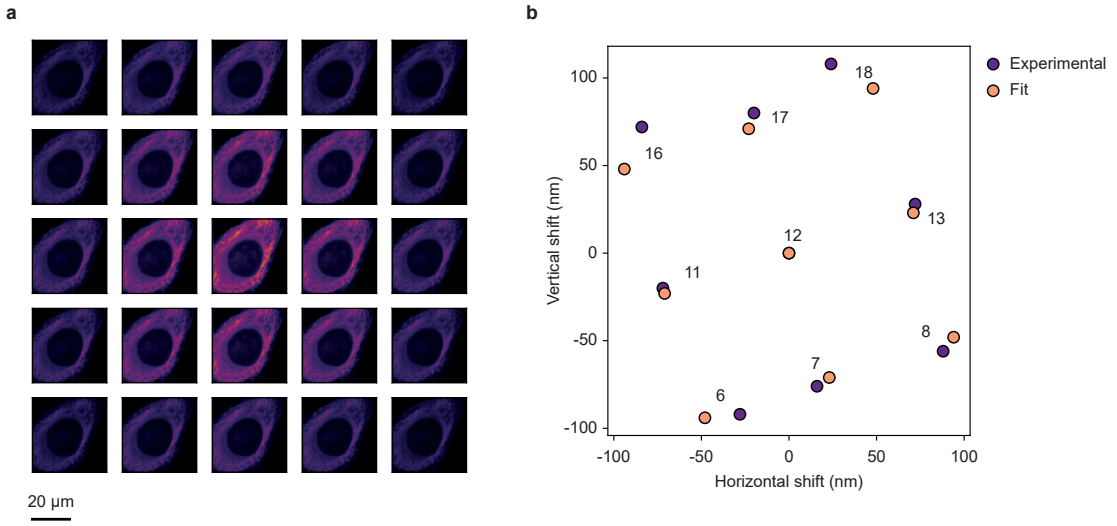

Figure 4.7 Extracting the detector orientation, rotation, and system magnification from the shift vectors of an ISM data set. a) Image of a fixed HeLa cell with α-tubulin staining. The images of the 5x5 detector elements are shown. **b)** From the shift vectors of the 3x3 detector elements closest to the detector center, the detector orientation, rotation, and system magnification were estimated.

detector projected into the sample space 150 nm and 100 nm, respectively. The simulation space pixel size was 2.5 nm, 1000×1000 pixels. To obtain $I_{ij}(\mathbf{r}_E)$, Figure 4.6 (b), we followed the same protocol as before, without the smoothing.

For the calculation of σ_{CRB}, we used the code described in [129], using as input parameters the simulated MDFs, assuming a total of 100 detected photons at each position and signal-to-background ratio (SBR) set to 10 to simulate optimal conditions. For the simulations of Figure 3.9, we scaled σ_{CRB} with $\sqrt{N}$, with N the expected number of photons at each position. The map of $N(x,y)$ was found by summing all MDFs, rescaled to have $N = 100$ in the center.

Measuring the orbit parameters and self-calibration procedure

The three parameters describing the spatial characteristics of the orbital scan are the diameter L, the starting phase α_0, and the direction of rotation. All parameters may be derived from a reference measurement or from the single-molecule measurement itself.

For the reference measurement, we used a doughnut-shaped illumination profile but the calibration of the orbital scan can be done with a Gaussian profile as well. We focused on a single GNP and zoomed in, ultimately leading to a situation in which the galvo is not scanning anymore, but an image of the GNP can be seen on the array detector. Then,

using the LabVIEW software (Section 4.3.3), we started a circular scan around that position, keeping the orbit period (1.92 ms) and induced radius (100 nm) the same as for the ISM-FLUX measurements. We measured for about 19 seconds and skipped the first couple of seconds in the analysis to remove galvo artifacts at the start of the scan. The laser power was chosen to have a photon count rate of about 7 MHz overall detector elements combined, orders of magnitude higher than the dark count rate but low enough not to induce saturation effects. As a result of the descanned detection mode, the movement of the galvo is reflected in the position of the particle on the array detector. To find the orbit parameters, we first summed, for each position in the orbit, the photon counts of 100 orbits. Then, we localized the particle for each position in the orbit with the maximum likelihood estimator, resulting in a circle of 32 points. We repeated this procedure for the whole data set, leading to a set of circles, from which the radius, starting angle, and orientation can be measured as explained in the protocol of Figure 4.8. Note that, due to inertia, the actual orbit radius and starting angle may differ from the imposed values, Figure 3.7.

ISM-FLUX self-calibration

All orbit parameters can also be directly retrieved from the Single-molecule data, Figure 4.9. Since the detector is placed in descanned mode, moving the illumination beam leads to an identical apparent movement of the detector in the sample space and thus leads to a corresponding opposite shift in the image of a molecule. In other words, localizing each on-event for every angle of the TCP individually results in 32 distinct localizations, each exhibiting a relative shift with respect to the other 31 localizations. From the resulting pattern, one can derive all orbit parameters, *i.e.*, the diameter, the starting angle, and the direction of rotation. Given the low number of photon counts per position in the TCP, we modified this concept by binning the photon counts in 4 quadrants of 8 consecutive points of the TCP, resulting in 4 sets of localizations rather than 32. We performed the localization analysis for different L values and calculated for which L value the relative shift between the four sets is the smallest. The resulting value was in good agreement with the true L value that we measured in a reference measurement, Figure 4.8.

To show how the orbit diameter can be recovered from the measurement itself, we consider the apparent displacement of a single emitter in the descanned detector frame (*i.e.* the micro-image) for two opposite excitation positions on the TCP. Specifically, we use two opposite points on the orbit to obtain an estimator for the orbit diameter. Consider a molecule at position $\mathbf{r}_E$ and two opposite points on the TCP, located at $\mathbf{r}_1$ and $\mathbf{r}_{16}$. When the laser beam is at position $\mathbf{r}_1$ or $\mathbf{r}_{16}$, the detector center is also at $\mathbf{r}_1$ or $\mathbf{r}_{16}$, respectively. Thus, in the

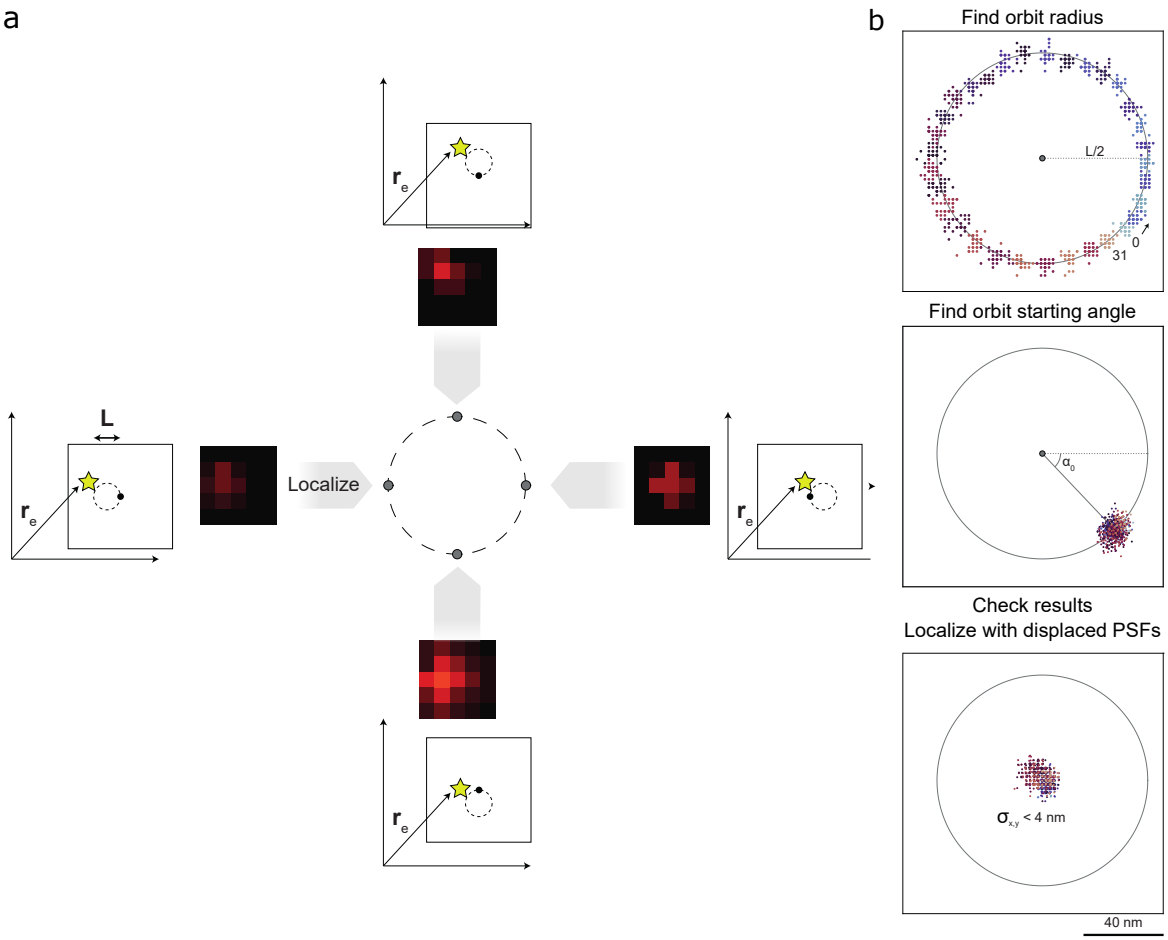

Figure 4.8 Measurement of the circular orbit parameters from a reference measurement. **a)** A fiducial marker is imaged while scanning the sample with a circular motion. Each position on the circle leads to a different micro-image, both in terms of the center of mass as well as the number of photons. The marker is localized in each micro-image individually using the 5x5 PSFs. Note that in this case only the distribution of the scattered photons on the detector is used, not the intensity differences between different images. Hence, also a Gaussian beam can be used for calibrating the galvanometric scanners. **b)** Analyzing the 32 micro-images using an MLE leads to a distribution of localizations on a circle from which the center is estimated by taking the mean coordinates and the radius is estimated as the mean distance from the center. Mean and standard deviation over 30 circles yields $L = (105.2 \pm 0.9)$ nm. Next, all points are rotated back an integer number of $1/32 * 2\pi$ and from the mean resulting coordinates, the starting angle of the orbit is estimated. Having both the orbit radius and starting angle, a new series of PSFs can be calculated: one set of 25 PSFs for each of the 32 positions on the circle, calculated by displacing each set of 25 PSFs according to its position in the orbit. The resulting set of 32x25 PSFs is used to analyze ISM-FLUX data, i.e. data from a full orbit or a sum over several consecutive orbits will result in a single localization, in which both the distributions of the photons on the detector as well as the change in photon flux during the orbit is taken into account.

detector frame-of-reference and assuming no (shot) noise, the image of the molecule moves from being centered around $\mathbf{r}_E - \mathbf{r}_1$ to $\mathbf{r}_E - \mathbf{r}_{16}$. The shift of the molecule $\mathbf{L}$ as seen by the detector is:

$$\mathbf{L} = \mathbf{r}_{16} - \mathbf{r}_1, \tag{4.1}$$

with $|\mathbf{L}| = L$, the diameter of the TCP. Thus, L can be directly derived from two micro-images taken at opposite angles of the TCP.

However, $\mathbf{r}_1$ and $\mathbf{r}_{16}$ are experimental localizations subject to a non-zero localization uncertainty. As a more robust approach, we propose to localize events for different points on the TCP, assuming a TCP diameter ℓ. In other words, we shift the two emitter positions in the detector reference frame with $\pm\boldsymbol{\ell}/2$, with $\boldsymbol{\ell}$ a vector on the line connecting the two points on the TCP and $|\boldsymbol{\ell}| = \ell$:

$$\boldsymbol{\ell} = \ell \frac{\mathbf{r}_{16} - \mathbf{r}_1}{|\mathbf{r}_{16} - \mathbf{r}_1|} \tag{4.2}$$

The two emitter positions are then

$$\mathbf{r}_E - \mathbf{r}_1 - \boldsymbol{\ell}/2, \tag{4.3}$$

$$\mathbf{r}_E - \mathbf{r}_{16} + \boldsymbol{\ell}/2. \tag{4.4}$$

The difference $\Delta\mathbf{r}$ between the two positions is

$$\Delta\mathbf{r} = -\mathbf{r}_1 + \mathbf{r}_{16} - \boldsymbol{\ell} = \mathbf{L} - \boldsymbol{\ell} \tag{4.5}$$

For $\boldsymbol{\ell} = \mathbf{L}$, $\Delta\mathbf{r} = 0$.

In reality, $\mathbf{r}_1$ and $\mathbf{r}_{16}$ are localizations with uncertainty σ^2. Assuming the localizations can be drawn from a Gaussian distribution with standard deviation σ, the mean squared distance $\langle d^2 \rangle$ between the two positions and their mean is

$$\langle d^2 \rangle = \left(\frac{|\Delta\mathbf{r}|}{2} \right)^2 + \sigma^2 = \frac{(L-\ell)^2}{4} + \sigma^2 \tag{4.6}$$

Thus, $\langle d^2 \rangle$ reaches a minimum where $\ell = L$, and the function value at this point corresponds to the experimental localization uncertainty.

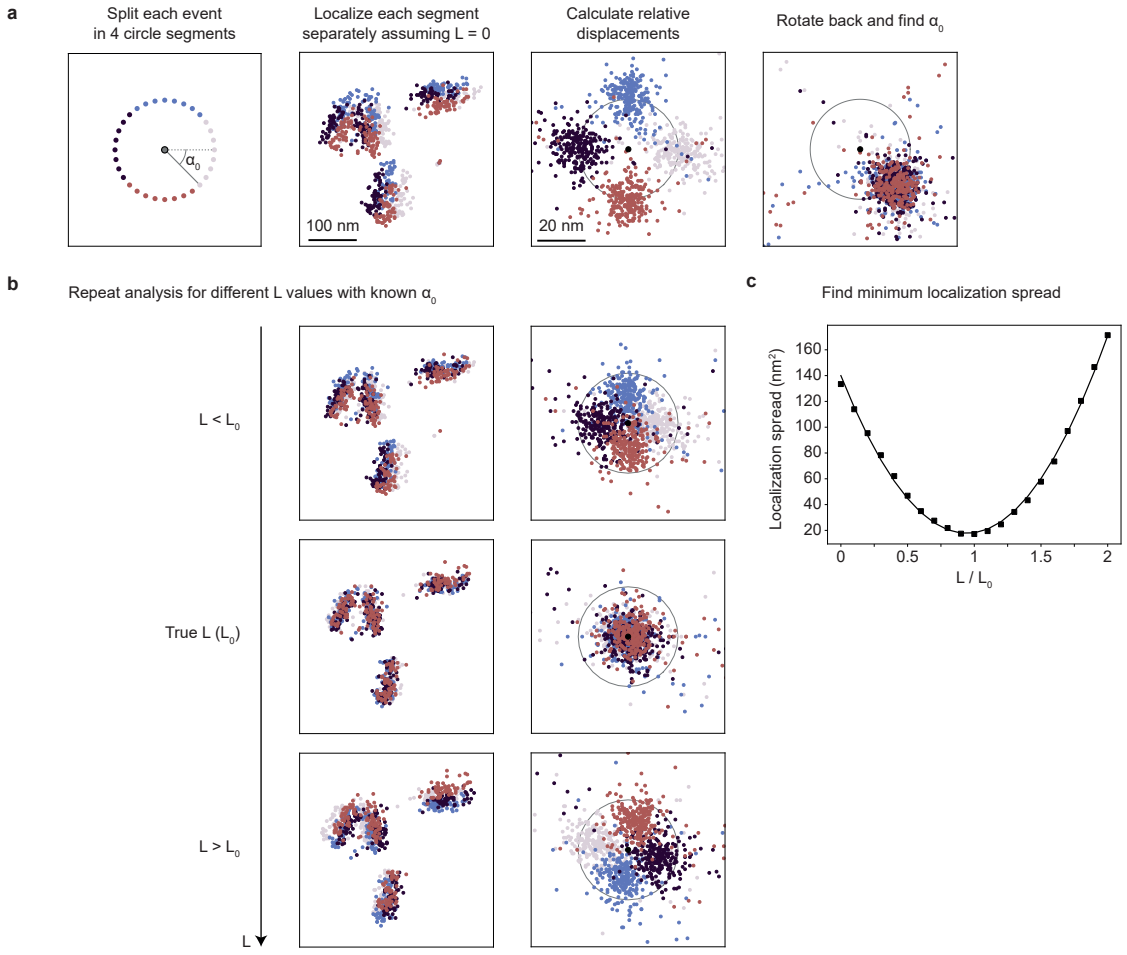

Figure 4.9 ISM-FLUX self-calibration using single-molecule data. a) Extracting the starting angle of the orbit α_0. Each blinking event is split into four segments, indicated by the different colors. The MLE is applied to each segment separately assuming $L = 0$. The result is four sets of segment-localizations that are displaced with respect to each other. Calculating for each event the relative displacement between the segments (*i.e.,* the displacement between the localization of each segment and the mean of the four segments) results in four clusters on a circle. By rotating each segment s back by an amount equal to $-\pi/4 - s\pi/2$ with $s \in \{0,1,2,3\}$, all clusters overlap and the starting angle α_0 can be measured. We calculated the median (x,y) position of each cluster and estimated α_0 as the mean $\pm$ standard deviation, which is $\alpha_0 = -0.81 \pm 0.02$, in good agreement with the value of $\alpha_0 = -0.812 \pm 0.008$ found in a reference measurement, Figure 4.8. **b)** The orbit diameter is found by repeating the analysis assuming different L values. For each L value, the localization spread, as defined in Section 4.5.2, is calculated from the relative displacements. **c)** The localization spread as a function of L (scatter plot) is fitted with a second-order polynomial curve (line plot). The fit, $y = 137x^2 - 259x + 140$ has a minimum around $L/L_0 = 0.95$, with $L_0 = 61.4\,\mathrm{nm}$, the diameter found in the reference measurement.

Since the photon counts in a single-molecule on-event are rather low, we split the data from each event into 4 circle segments of 8 consecutive points in the orbit. Then, we approximate $\langle d^2 \rangle$ as the mean squared displacement between the four localizations and their mean. As a metric for the overall localization spread of all events combined, we take the median value of all $\langle d^2 \rangle$ values. We repeat this calculation for a set of ℓ values and we fit the resulting curve with a second-order polynomial. The ℓ value that minimizes the localization spread corresponds to the true L value. The protocol is illustrated in Figure 4.9.

4.5.3 Data formats, logging, and analysis workflows

Stabilization logs and file formats

During each stabilization run, the software writes the experiment metadata and time series outputs to a dedicated session folder. For reproducibility and later analysis, the logging is organized into *ten* files (including the configuration file): `config_dict.pkl` at the session root directory and nine files in the `CSV/` subdirectory (Table 4.2). The recorded outputs fall into three practical categories: (i) commanded **stage position** updates (per axis), (ii) **individual particle position** time traces used for drift estimation (lateral per particle; axial as a single value per frame), and (iii) **axial calibration and diagnostic parameters**, which store the quantities required to convert the axial observable into nanometers and to document acquisition performance.

Table 4.2 Files saved during stabilization.

Category	**Filename**
Configuration	`config_dict.pkl`

Particle positions	`X_coordinates.csv`
	`Y_coordinates.csv`
	`Z_coordinates.csv`

Stage positions	`X_position.csv`
	`Y_position.csv`
	`Z_position.csv`

Axial calibration / diagnostics	`astigmatism_values.txt`
	`Z_axis_FFT_calibration.txt`
	`Framerate_information.txt`

Both particle- and stage-position logs include a timestamp with microsecond resolution (formatted as `HH_MM_SS_ffffff`), generated at the time the corresponding value is queued for saving. Lateral particle coordinates are stored *per particle* in the same row, enabling

post-experiment inspection of individual fiducials as well as combined statistics (Table 4.3). Stage positions are stored whenever a move command is issued (i.e., when the controller is updated), providing a direct record of the correction trajectory and allowing drift to be verified or re-estimated during post-processing (Table 4.4). Along the axial channel, both the particle-derived quantity and the stage update are saved as a single scalar per timestamp: the axial estimator produces one value for the selected ROI, and the stage log records the corresponding commanded z position.

Table 4.3 Example format of the lateral particle-position file `X_coordinates.csv`. The first column is the timestamp t, followed by per-particle position p_i (in nm).

timestamp **(HH_MM_SS_ffffff)**	p_1	p_2	p_3	p_4	p_5
09_06_28_511870	158.29	−38.85	133.42	130.67	58.62
09_06_28_561872	156.92	−41.06	132.09	129.18	56.96
09_06_28_598872	158.02	−40.85	132.14	129.83	57.48

Table 4.4 Example format of the stage-position file `X_position.csv`. A new entry is written when an x-axis move command is executed.

t **(HH_MM_SS_ffffff)**	X (µm)
09_06_29_935066	66.04168
09_06_29_999098	66.08621
09_06_30_063098	66.13075

In addition to these ten files, the software stores image outputs to support validation and alternative post-processing workflows. At the start of the experiment, an overview image of the selected fiducial ROIs is saved as `Images/XY_rois.png`. Furthermore, a dedicated process can periodically save raw camera frames from both channels as uncompressed TIFF files (e.g., `xy_image_{t}.tiff` and `z_image_{t}.tiff`), where the saved image corresponds to the most recent acquired frame available at the time of saving. Because both cameras can run at > 100 fps, saving every frame is typically impractical; instead, the user specifies a time interval Δt (in seconds) that sets how often images are written to disk.

After termination of the stabilization run, the software writes a short acquisition summary to `Framerate_information.txt`. This log consolidates the total number of acquired frames for the lateral (XY) and axial (Z) cameras, the corresponding average frame rates, and the total elapsed acquisition time (Table 4.5). The summary provides a compact record of the effective throughput of the tracking pipeline and is useful for quality control and for correlating downstream analyses with the actual sampling density.

Table 4.5 Example contents of `Framerate_information.txt` written at the end of an experiment.

Quantity	Value
XY_Frames	265840
Z_Frames	173271
Frame_Rate_XY (Hz)	140.146
Frame_Rate_Z (Hz)	91.345
Total_Time (s)	1896.876

Widefield SMLM localization pipeline

All SMLM (STORM and DNA-PAINT) data is acquired on a widefield fluorescence microscope equipped with an sCMOS camera (Hamamatsu ORCA-Flash4.0 V3) and recorded using the manufacturer software (Hamamatsu TokuPic). Data were collected as single-color time series and saved as 16 bit TIFF stacks. For STORM measurements, the exposure time was set to 15 ms to 25 ms; for DNA-PAINT, exposures were typically 100 ms. Depending on the sample and labeling density, $\sim 60\,000$ frames to $80\,000$ frames were recorded per acquisition. The pixel size was set to 96 nm per pixel for subsequent analysis as explained in Section 4.2.3.

Localizations were obtained using either ThunderSTORM (Fiji/ImageJ) [84] or Picasso [147]. Picasso was used as the default workflow; when file size or memory constraints prevented reliable processing, localization was performed in ThunderSTORM using fixed settings across datasets. In ThunderSTORM, images were pre-filtered with a lowered Gaussian filter ($\sigma = 0.9$ pixel). Candidate emitters were identified by local-maximum detection using a threshold of $1.4 \times \mathrm{std}(\mathrm{Wave.F1})$ and a 4-neighbourhood connectivity criterion. Sub-pixel localization was then performed by fitting an integrated Gaussian point-spread function within a fitting radius of 3 pixel, using least-squares optimization with an initial PSF width of $\sigma = 1.6$ pixel. No explicit camera calibration (gain/offset correction or sCMOS variance map) was applied; instead, only the effective pixel size was specified.

Post-localization filtering was applied to remove implausible or low-quality fits. Specifically, localizations were retained when the fitted Gaussian width satisfied $\sigma \in [90\,\mathrm{nm}, 200\,\mathrm{nm}]$ and the lateral localization uncertainty fulfilled $\sigma_{xy} < 25$ nm to 30 nm (dataset-dependent threshold within this range). The effective localization precision was additionally quantified using nearest-neighbor based analysis (NeNA) following the approach introduced by Endesfelder *et al.* [148], implemented in a custom Python script. No Fourier ring correlation (FRC) analysis was performed, as the quantitative results presented here do not rely on a single scalar image-resolution estimate.

During all measurements, the sample stabilization system was active to reduce drift. Under these conditions, residual sample drift was maintained below $\sigma < 1.5\,\text{nm}$ along all axes, and therefore no additional drift correction was applied during post-processing. This choice reduced the number of tunable post-processing steps and also avoids introducing drift-model artifacts into the reconstructed localization datasets.

ISM-FLUX data collection formats and conversions

The ISM-FLUX acquisition produces two complementary data products: a PSF dataset and the experimental photon time-trace data. The PSF is acquired using the Brighteyes-MCS software [143] in `.h5` format and is subsequently converted to `.tiff` for the downstream analysis. This conversion is performed using the open-source python package `BrightEyes-FFS` [149]. The TIFF representation is convenient for inspection and for interfacing with standard image-processing routines, while retaining the original sampling and intensity values.

Table 4.6 Parameters stored in `{binfile_name}_info.txt`. Values shown correspond to the example provided.

Field	**Example value**	**Description (as used here)**
Date / time	10/5/2025, 5:36 PM	Acquisition start timestamp
TIME RESOLUTION [µs]	6.000000E+0	Time-bin width Δt
NUMBER OF TIME BINS PER PIXEL	10	Number of time bins N_{bins} stored per pattern position
NUMBER OF PIXELS	32	Number of spatial positions per orbit (N_c)
NUMBER OF LINES	300000	Number of orbits (N_{orbits})
NUMBER OF FRAMES	1	Number of frames stored in this file
TOTAL NUMBER OF DATA POINTS	512000000	Total scalar samples written to the `.bin` stream
RANGE X [µm]	1.000061E-1	Scan range in x
RANGE Y [µm]	0.000000E+0	Scan range in y
RANGE Z [µm]	0.000000E+0	Scan range in z
OFFSET X [V]	9.082001E+0	Galvo driver offset in x
OFFSET Y [V]	-1.076300E+1	Galvo driver offset in y
OFFSET Z [V]	0.000000E+0	Galvo driver offset in z

The experimental measurement is collected using the LabVIEW software (Section 4.3.3) as a binary `.bin` file, accompanied by a text metadata file `{binfile_name}_info.txt`

(Table 4.6). The binary stream encodes time-binned photon counts with a microsecond time base for each detector channel and for each point in the repeated excitation pattern. In the configuration used in our analysis, each data point corresponds to a time bin of width $\Delta t = 6\,\mu s$ and $N_{bins} = 10$ time bins are stored per position. Each circular orbit of $N_c = 32$ points and detector of $N_d = 25$ pixels, yields a per-orbit data block of $N_{bins} \times N_c \times N_d$ samples. This means that the final data dimension is:

$$N_{bins} \times N_c \times N_d \times N_{orbits}$$

With $N_{orbits} = 300{,}000$ user-defined repetitions, typically the total acquisition time is

$$T = \Delta t \times N_{bins} \times N_c \times N_d \times N_{orbits} = 6\,\mu s \times 10 \times 32 \times 25 \times 300{,}000 = 576\,s.$$

For localization analysis, the `.bin` file is parsed into a NumPy array with dimensions matching the acquisition settings and is analyzed together with the measured PSF (MDF).

ISM-FLUX data analysis and localization pipeline

Once all the data is preprocessed and converted into necessary file formats, the data analysis is performed in Python. A scheme of the analysis protocol is shown in Figure 4.10 and Figure 4.11. For the origami measurements, we started by calculating a time trace, Figure 4.10 (a), obtained by summing all photons in all detector channels (excluding the two that have high dark-noise, Figure 3.2), for all positions on the orbit. Then, we calculated a histogram, which was fitted with a sum of two Gaussian distributions as a fit function. The count value corresponding to the first peak was used as a threshold to distinguish the signal from the background. Each time the photon counts exceeded the threshold, a new event started. The event ended when the intensity trace fell again below the threshold. For all orbits within an event, photon counts observed in the same detector element for the same position in the orbit were summed, resulting in a reduced event data set n_{ij}, with $i \in \{1,2,...,N_c\}$ and $j \in \{1,2,...,N_d\}$, with $N_c = 32$ and $N_d = 23$. The events were further filtered, Figure 4.10 (b), based on the total number of photons, the duration, and the standard deviation $\sigma_{measured}$ of the count fluctuations within an event. Given a high background count rate stemming from a combination of out-of-focus fluorescence and detector dark counts, a threshold of at least 2000 photons per event was enforced. Furthermore, only events with a duration of at least 5 orbits were considered. Lastly, we calculated for each event the standard deviation of the photon counts over the orbits and discarded events for which this standard deviation was more than 2.5 times the expected value for a Poisson process, *i.e.*, $\sqrt{N}$. This filter reduces the chance of including events characterized by pronounced blinking or the presence of a secondary

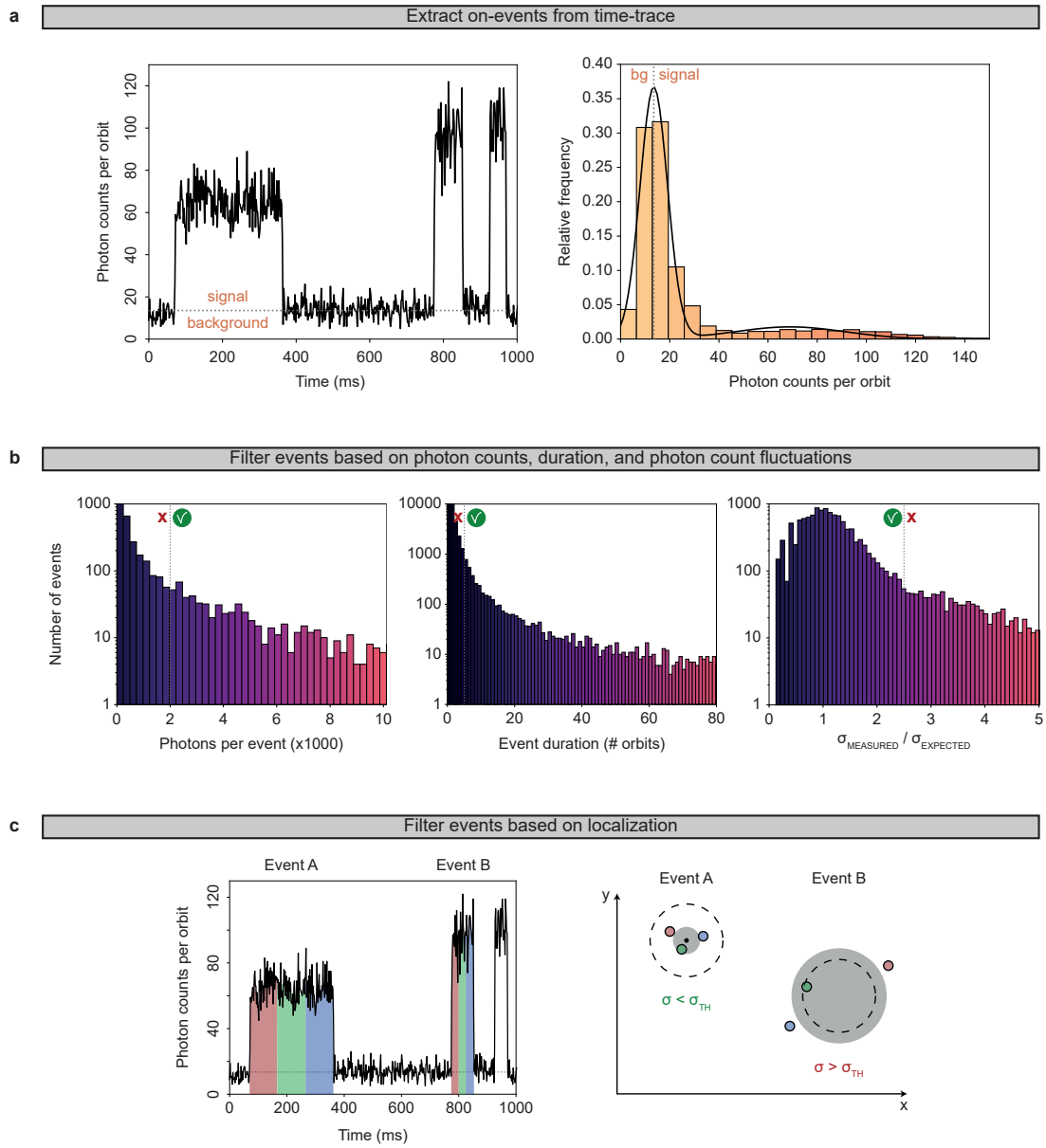

Figure 4.10 Analysis protocol for ISM-FLUX on DNA-origami measured with DNA-PAINT. a) Part of the time trace, obtained by summing all photons in all detector channels for all positions on an orbit, and histogram and fit of the full time trace (bg = background). **b)** All events pass through three filters, based on the total number of photons (left), the duration (center), and the standard deviation σ_{measured} of the count fluctuations within an event (right). **c)** Each event that passes all filters is split into three equally long chunks, resulting in three independent localizations, colored in red, green, and blue. If the three localizations are closer to each other than a user-chosen threshold, i.e. $\sigma < \sigma_{TH}$, the event is accepted, otherwise, the event is discarded. The grey circle has a radius of σ, the dotted line indicates the threshold. Here, event A is accepted, and the three localizations are merged into a single one, obtained by taking the mean coordinates (black dot). Event B is discarded.

fluorophore bound to a nearby docking site. Each event that passed all filters was split into three equally long segments, Figure 4.10 (c), resulting in three independent localizations. If the three localizations were closer to each other than a user-chosen threshold, *i.e.*, $\sigma < \sigma_{TH}$, the event was accepted; otherwise, the event was discarded. Here, $\sigma = \sqrt{\frac{(\sigma_x^2+\sigma_y^2)}{2}}$, with σ_x and σ_y the standard deviation of the x and y locations, respectively.

For the measurements with the fiducial markers, no event detection was needed. Instead, we moved the marker every second, Figure 5.4 and Figure 5.5 (a-c), with the piezo-electric stage to a pre-programmed position while continuously performing the orbital scan and detecting the signal with the array detector. The resulting trace was binned several times with increasing bin lengths to increase the photon counts. For Figure 5.4 and Figure 5.5 (a), each of the 17 positions was probed 50-100 times, and the standard deviation in x and y or the three particle positions closest to the TCP center, indicated with the rectangles in panels (a, b), were used to calculate the localization uncertainty. For the point detector data in Figure 5.5 (a), outlier localizations were removed by considering only the retrieved positions within the TCP. For Figure 5.5 (b), we probed each location at least 39 times, and we calculated the covariance matrix of the localizations and used the eigenvectors and the square root of the eigenvalues of the matrix to find the orientation and size of the principal components of the spread, respectively.

For the measurements with the fixed fluorophores, the intensity trace was binned several times with increasing bin lengths to increase the photon counts. For a given bin length, each set of three consecutive bins was analyzed as follows: first, the fluorophore position was calculated for each of the three events. The standard deviation in (x,y) was used as an estimate for the localization uncertainty for that set. The final localization uncertainty for a given N was calculated as the mean $\pm$ standard error over all sets. Similarly, N was calculated as the mean $\pm$ standard error over all sets. Depending on the bin length, between 22 and 366 events of 75-1300 photons (fluorophore position 1) and between 50 and 2500 events of 43-2200 photons (fluorophore position 2) were analyzed.

Localizations were calculated with a brute-force maximum likelihood approach, similar to [129]. Each event consisted of either 32×25 or 32×23 photon count values, with and without hot detector elements, respectively, and an equally large set of MDFs. For the GNP and the 20 nm origami, we use the experimental MDFs; for the 40 nm origami the theoretical MDFs. We analyzed all events in a data set assuming the same SBR, set to 3 for the fiducial marker, 3.13 for the 40 nm origami, and 3.3 for the 20 nm origami, in all cases assuming the background to be uniformly distributed over all positions in the orbit and all detector elements.

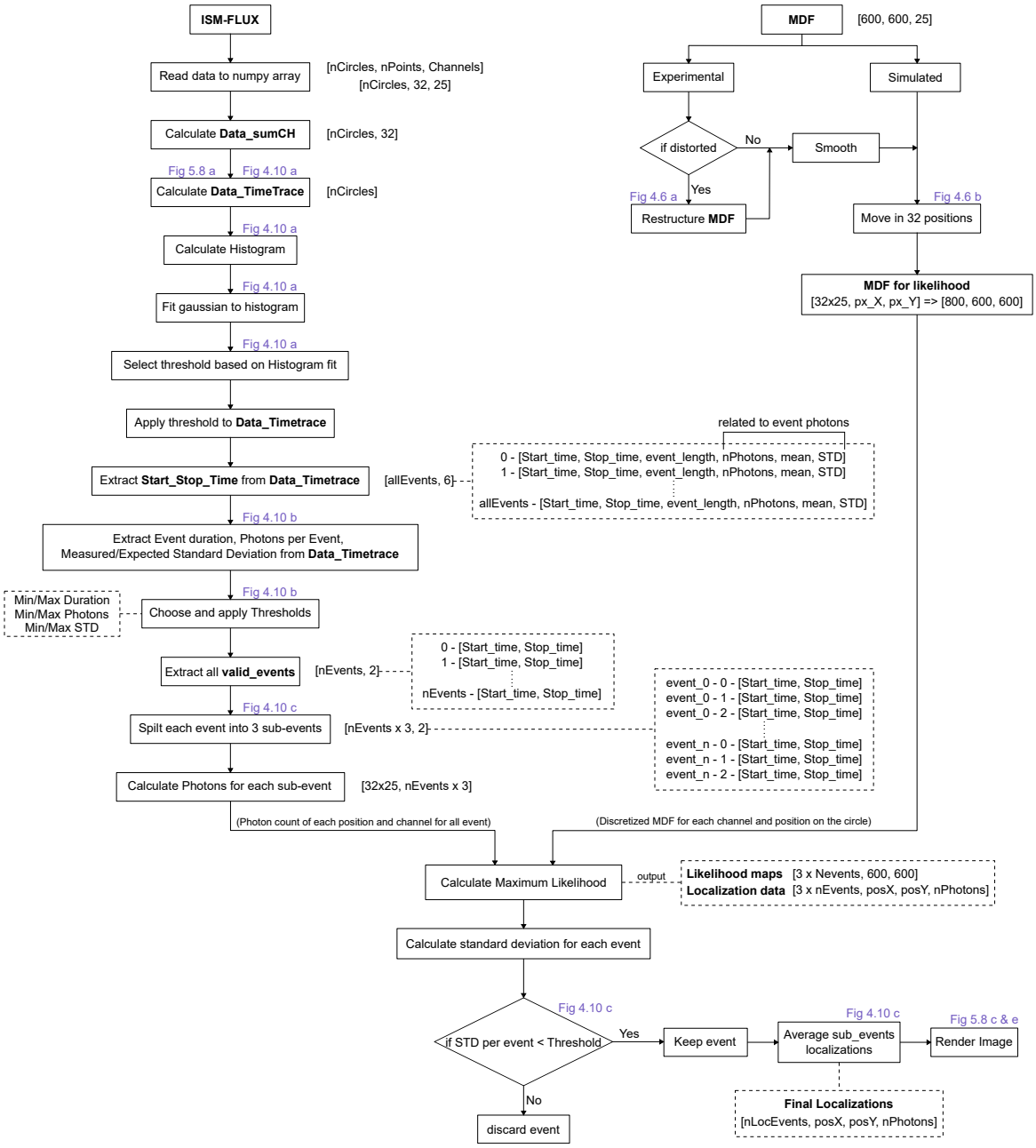

Figure 4.11 Analysis pipeline for ISM-FLUX data analysis.

Chapter 5

Performance Characterization and Results

This chapter focuses on the experimental results obtained with the stabilization, image scanning microscopy (ISM)-FLUX and the widefield single-molecule localization microscopy (SMLM) setups. First, we quantify the performance of a 3D active sample-stabilization system and validate its impact on long-duration imaging. We then evaluate ISM-FLUX using gold nanoparticles (GNP), fixed fluorophores, and DNA-origami nanorulers to evaluate the localization performance, emphasizing localization range and localization uncertainty relative to single-element detection. Finally, we benchmark the system in a widefield SMLM modality and assess how active stabilization affects reconstruction quality and localization repeatability over extended acquisitions.

5.1 Sample Stabilization

First, we measured the performance of the stabilization system using the calculated particle positions with respect to the start of the measurement in the cameras' frames of reference (lateral and axial), as measured by the stabilization system itself during closed-loop stabilization (Figure 5.1 (a)). For all axes, we obtained approximately Gaussian distributions centered near zero, where zero is defined relative to the first frame, with a standard deviation below 1 nm for a 1000 s measurement. Since the nanopositioning piezo system has a closed-loop resolution of 1 nm, and our stabilization software uses a 1 nm threshold for stage updates, we cannot expect a better precision. Note that, due to differences in the frame rate of the cameras and the update rate of the stage, the measured drift may be higher than the 1 nm threshold for multiple consecutive frames, as shown in the insets in Figure 5.1 (a). The

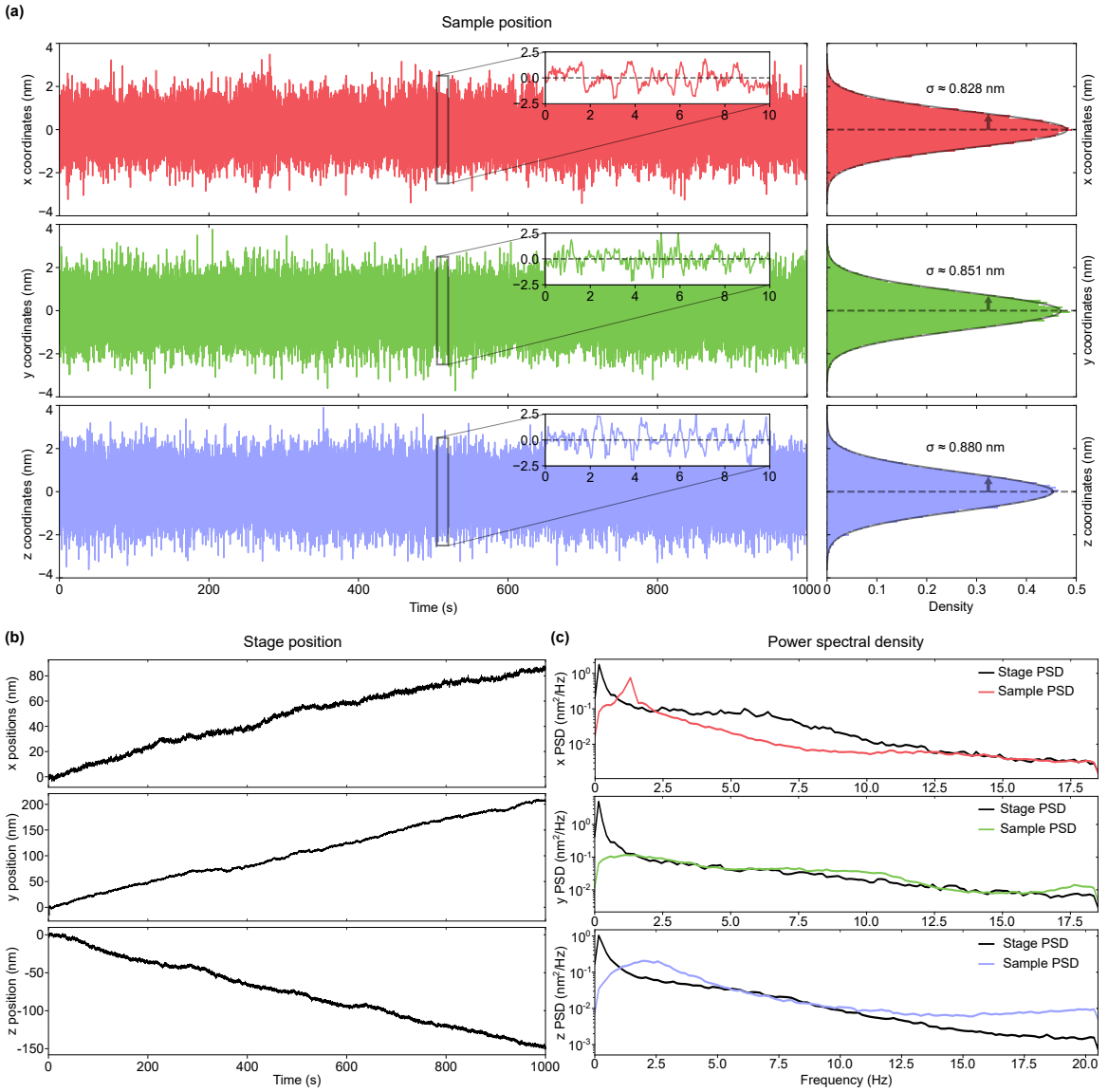

Figure 5.1 Sample stabilization results of data collected from piezoelectric stage. a) Sample position as monitored by the stabilization system and the corresponding histograms showing the distribution of the sample position across time. The parameter σ is the standard deviation of the Gaussian fit to the histogram. **b)** Corresponding position of the piezoelectric stage for each axis. **c)** Power spectral density (PSD) of the stabilized sample position (colored) and the stage position (black).

corresponding stage position (Figure 5.1 (b)), shows about 100 nm to 200 nm drift in the x, y direction, respectively, and about 150 nm drift in the axial direction over 1000 s. Comparison of the power spectral density curves of the stage and sample position (Figure 5.1 (c)) shows that drift and low-frequency vibrations are suppressed by the stabilization software by more than one order of magnitude in all dimensions. The repositioning of the stage amplifies some vibrations, as can be seen around 1 Hz for the x-direction and 2.5 Hz for the z-direction. However, the amplitudes of these vibrations are low enough to keep the overall sample movement below 1 nm standard deviation in each direction.

As a control measurement, we performed long-term confocal imaging of the GNP sample with the stabilization system running. Specifically, we removed the band-pass filter from our custom image scanning microscope (ISM-FLUX setup) to record reflection images in descanned mode using the single-photon avalanche diode (SPAD) array detector [31] (Figure 5.2 (a)). Subsequently, we conducted time-series experiments utilizing our microscope control software, BrightEyes-MCS [143]. We calculated the lateral drift in the confocal images by phase correlating each image with a reference image taken at the start of the measurement. With the stabilization software running, we observed a drift of (-0.8 ± 1.2)nm and (-0.6 ± 0.9)nm ($mean \pm SD$) in the x and y direction, respectively, over the course of 4 hours (Figure 5.2 (c)). Although these values exceed the sub-nm precision calculated by the stabilization system, these results may be expected given the additional uncertainty introduced by the galvanometric scan system or other drift components in sections of the beam path that are not common with the stabilization path. Because the stabilization readout and the imaging readout use different optical paths and different estimators, the reported residual motion depends on the measurement modality. With the stabilization system turned off, we measured a lateral drift of more than 500 nm on the x-axis and more than 900 nm on the y-axis, with additional drift in the axial direction, moving the sample out of focus (Figure 5.2 (b) & (d)).

For both lateral and axial drift calculations, we prioritized lightweight algorithms over slower, more precise alternatives [144]. Consequently, the update speed of our system is not limited by the software but by the rate at which the piezo stage controller can process movement commands, here ~ 20 Hz for each axis. For applications where speed and accuracy are less critical, such as time-lapse measurements in confocal or widefield microscopy, alternative algorithms may be preferable. For instance, the local gradients method is well-suited to sparse particles whose images are radially symmetric. However, for denser particle distributions or larger, non-radially symmetric structures, other methods, such as phase correlation, will perform better. Similarly, when only axial stabilization is required, analyzing

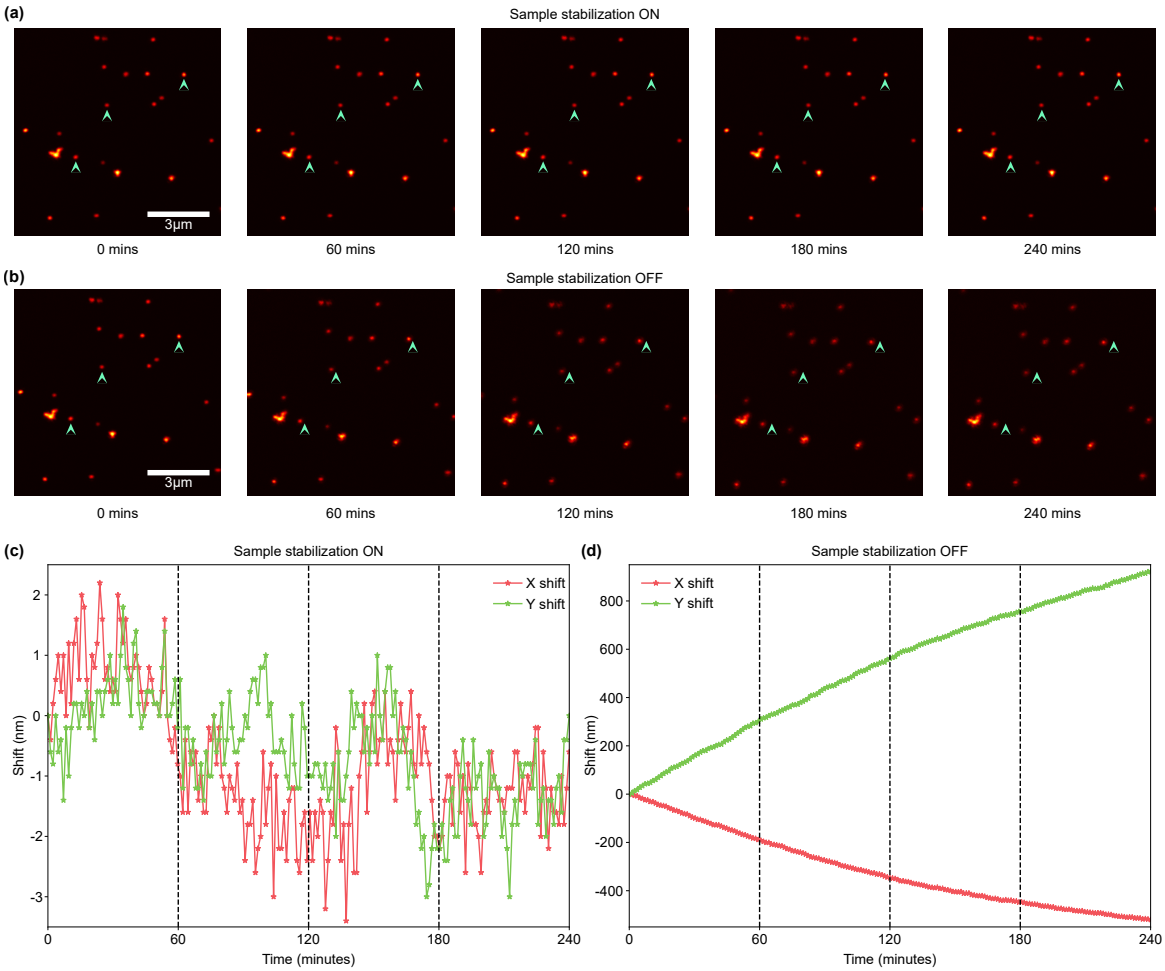

Figure 5.2 Confocal imaging of gold nanoparticless with and without active stabilization. **a)** Subset of a time series taken with the active stabilization on. The sample consisted of GNPs attached to the coverslip surface. Images were acquired every minute for four hours. With the stabilization system on, there is no visible drift. **b)** The same measurement with the stabilization system off shows a clear drift in both the lateral and axial directions over the course of four hours. **c)**, **d)** Measured drift in each image with respect to the start of the time-lapse, calculated by phase-correlation, with the stabilization system on **c)** and off **d)**.

the reflection of a beam off a cover slip can be advantageous, as it eliminates the need for fiducial markers in the sample. Although we have not implemented these alternatives here, the open-source code can be easily adapted to accommodate these custom scenarios.

5.2 ISM-FLUX

Cramér-Rao bound analysis with SPAD array detector

As detailed in Appendix A, we extend the maximum likelihood approach from minimal fluorescence photon fluxes microscopy (MINFLUX) to the use of an array detector [129]. The likelihood function for the emitter at position $\mathbf{r}_E$ can be written as

$$\mathscr{L}(\mathbf{r}_E|\{n_{ij}\}) = \frac{N!}{\prod_{i=1}^{N_c}\prod_{j=1}^{N_d} n_{ij}!}\prod_{i=1}^{N_c}\prod_{j=1}^{N_d} p_{ij}(\mathbf{r}_E)^{n_{ij}} \tag{5.1}$$

where, $p_{ij}(\mathbf{r}_E)$ denotes the molecule detection function (MDF), *i.e.*, the emitter position-dependent probability to detect a photon during the i-th position in the circle and in the j-th element of the detector. All MDFs need to be known for the localization and can either be simulated using Eq. A.15 or experimentally measured by raster-scanning an emitter.

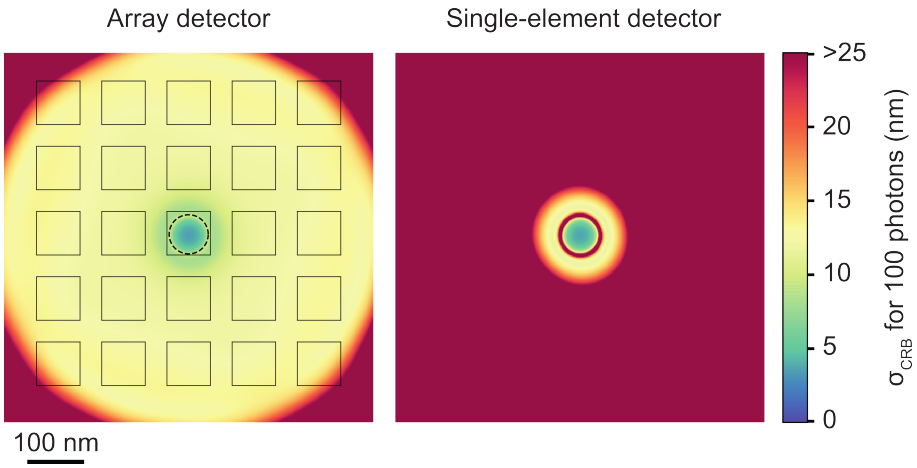

Figure 5.3 Cramér-Rao bound as a function of the emitter position for 100 detected photons. Comparison between an array detector and a single-element detector for $L = 90\,\text{nm}$.

The minimal uncertainty σ_{CRB} in localizing an emitter with an unbiased estimator, *i.e.*, the Cramér-Rao bound (CRB), depends on various factors, such as L, N, and the position of the emitter relative to the center of the targeted coordinate pattern (TCP). For an emitter at the center of the scanned orbit, $\sigma_{\text{CRB}} \propto L/\sqrt{N}$ [129], Figure 3.6. Thus, the smaller the diameter of the orbital scan, *i.e.*, the lower the L value, the more precisely the emitter can be localized for a given N. For a fixed L value, σ_{CRB} scales with $1/\sqrt{N}$. These properties are shared by both MINFLUX and ISM-FLUX. Similar properties hold for positions in close proximity to the TCP center. In this scenario, the emitter's position is primarily inferred

from the total number of photons detected at each position within the TCP, with the photon distribution on the array detector contributing minimal additional information.

However, a key difference between both techniques emerges when the emitter is located outside the circle described by the TCP. With a single-element detector, the localization range is restricted to a circle with a diameter below the TCP diameter. Figure 5.3 compares the CRB across emitter positions for array detection versus a virtual single-element detector at $L = 90\,\text{nm}$ (100 detected photons). Outside of this range, σ_{CRB} diverges, and the maximum likelihood estimator (MLE) is highly biased. In contrast, the MLE with array detectors shows a non-diverging σ_{CRB} and an unbiased position estimation, at least for high photon counts, for a much larger range, Figure 5.3. Farther from the TCP center, the dominant source of information shifts to the image of the emitter, resulting in an increased σ_{CRB} that approaches the value of camera-based techniques. However, thanks to the structured illumination, σ_{CRB} remains lower (*i.e.,* better) for ISM-FLUX than for conventional camera-based techniques. As a result, the localization range is not set by L but by the lateral size of the illumination profile and the field-of-view (FoV) of the detector – which should be small enough to block out-of-focus fluorescence. *E.g.*, for $L = 90\,\text{nm}$ and 100 detected photons, σ_{CRB} increases from about 3.5 nm in the center to about 12 nm near the edge of the detector. The larger L, the more homogeneous σ_{CRB} becomes. Large L values essentially represent camera-based SMLM with structured illumination [150, 151], offering a $\sqrt{2}$-fold improvement in localization uncertainty compared to conventional camera-based methods.

In our implementation, we kept the laser power, the TCP position, and the L value constant during the localization process. As a consequence, also the emission rate of the emitter depends on its position with respect to the TCP, Figure 3.9. Due to the varying N, the effective localization uncertainty within the localization range is between 3.5-7 nm for L =90 nm and a photon count between 100 and 650. We chose a constant L value of 90 nm as a trade-off between minimum localization uncertainty in the center and maximum σ_{CRB} homogeneity. Note that keeping L large enough also keeps the difference between the expected photon counts for an emitter in the TCP center and an emitter near the doughnut maximum below one order of magnitude, Figure 3.9. Since we do not adjust the laser power depending on the emitter's position, the radiant flux received by a molecule in the TCP center should be sufficiently high to detect its fluorescence signal above the background. Simultaneously, for a molecule positioned near the peak of the doughnut-shaped illumination, the radiant flux should be low enough not to induce blinking, bleaching, or saturation of the fluorescence.

Note that we discretized the TCP orbit into 32 points. The calculation of σ_{CRB} shows that very similar results can be obtained with fewer probed points, Figure 3.5. However, to ensure a smooth orbital scan in which potential inertia effects caused by the galvo's motion and stopping is minimal, we opted to use 32 points.

Enhanced localization range with ISM-FLUX

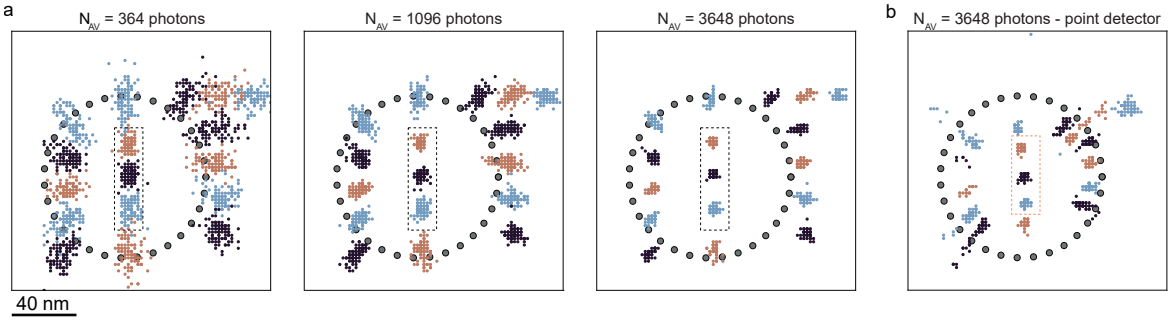

Figure 5.4 ISM-FLUX localization of gold nanoparticles moved in IIT shaped pattern. **a)** Retrieved localizations for different particle positions, obtained by moving the piezo-stage to 17 different positions, and different photon numbers, obtained by changing the integration time. The rectangles indicate the three particle locations used to calculate the average experimental localization uncertainty within the TCP. **b)** Treating the data as if they were recorded with a single-element detector.

We validated the localization performance of ISM-FLUX by repeatedly localizing a GNP translated to different positions with a nanometer-precise piezo-electric stage. During the entire temporal series, the particle moved to different positions, the TCP remained constant, and the SPAD array detected photons in synchronization with the galvanometer mirror. We began by checking the setup's performance in localizing a particle positioned inside or near the TCP circle, Figure 5.4 (a). We moved the particle to 17 different positions, forming the logo of the Italian Institute of Technology (IIT). We calculated the localization uncertainty using $\sigma = \sqrt{(\sigma_x^2 + \sigma_y^2)/2}$, with σ_x and σ_y being the standard deviations of the estimated positions in x and y, respectively. For the three positions closest to the TCP center, σ is better than 4 nm for an average of 364 detected photons, Figure 5.5 (a). To evaluate position estimation with different photon counts, we binned the temporal series in post-processing for increasing numbers of TCP orbits, resulting in a continuously decreasing localization uncertainty, which approached a limiting value of about 2 nm for 3000 photons. In addition to the uncertainty arising from photon statistics, other factors, such as residual vibrations and system drift, contribute to the overall uncertainty limit.

For comparison, we mimicked the scenario of a single-element detector by summing all photon counts across the 25 elements of the array detector. Estimating the position based on only 32 photon counts leads to a similar localization uncertainty for positions close to the TCP center, Figure 5.4 (b). However, close to the border of the TCP and outside the TCP, we get an imprecise and biased estimate of the particle, highlighting the limited localization range with a single-element detector. Note that the virtual single-element detector and the array detector are the same size (1.4 Airy Units) and, therefore, have the same FoV. The increased localization range of ISM-FLUX is entirely due to the subdiffractional spatial sampling provided by the array detector.

To verify whether the localization range closely matches the FoV of the array detector, we moved the particle in a grid pattern across the entire array detector FoV (about 600 $\times$ 600 nm^2) with a step size of 100 nm. At each grid position, we localized the particle after a predetermined number of orbits, leading to a different total number of detected photons (up to a factor of 4) per imposed position. Despite a varying number of photon counts, we observed a uniform σ across the FoV of the detector, Figure 5.5 (b). This uniformity is maintained for longer integration times, resulting in a lower (*i.e.,* better) σ for all positions, according to $1/\sqrt{N}$, Figure 5.5 (b) & (c).

Building on these results, it remains to be shown that we can use the single-digit localization uncertainty outside the TCP to resolve single-digit spatial increments. Therefore, we repeated the experiment of GNP localization with a step size of 6 nm. For the GNP placed outside the TCP, we observed a σ value below 2 nm, and consequently, we could resolve the 6 nm steps, Figure 5.5 (d). The localization results of Figure 5.5 (d) are shown in Figure 5.6. The ability to resolve these steps stems mainly from the subdiffractional sampling of the detected photons. Instead, a localization scheme based on a single-element detector would require additional prelocalization.

Resolving DNA origami nanorulers with ISM-FLUX

For high enough photon counts, we achieved single-digit localization uncertainty for bright scatterers. Here, we demonstrate the ability of ISM-FLUX to localize single fluorophores.

First, we used ISM-FLUX to localize fixed ATTO647N fluorophores. The sample was raster-scanned to find single fluorophores, Figure 5.7 (a), followed by ISM-FLUX measurements at different positions close to the fluorophore positions, Figure 5.7 (b) & (c). For a fluorophore in the TCP center, we found localization uncertainty values close to σ_{CRB} for all N values. Analyzing the same data for a single-element detector gives almost identical

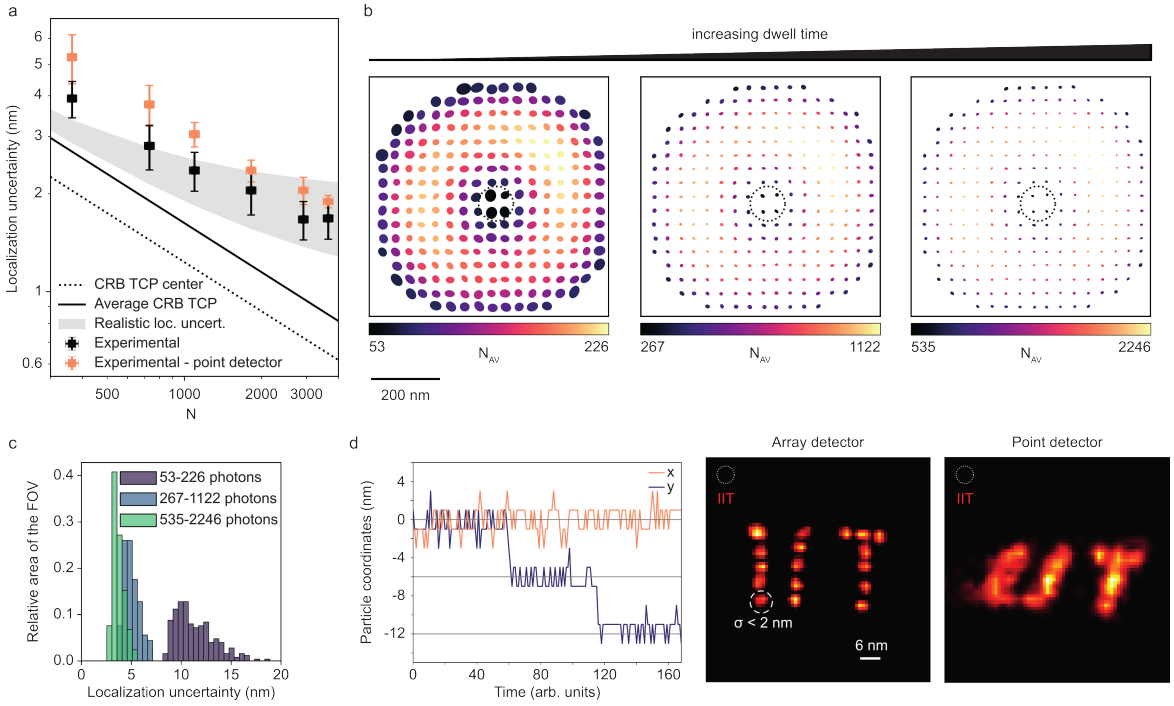

Figure 5.5 Localization uncertainty of ISM-FLUX over the field-of-view at different photon budget. a) Average localization uncertainty of ISM-FLUX as a function of N, calculated as the average and standard deviation for the 3 positions, indicated by the dotted rectangles in Figure 5.4 **a)** and **b)**. The comparison with σ_{CRB} for the TCP center (dotted line) and the average within the TCP (solid line) are plotted. The 'realistic localization uncertainty' is calculated as $\sqrt{\sigma_{\text{CRB}}^2 + \sigma_{\text{DC}}^2}$, with $\sigma_{\text{DC}} = 1-2$ nm. **b)** Calculated localization uncertainty over a large FoV, obtained by moving the orbital scan position with respect to the GNP to 250 positions. The color indicates the number of detected photon counts for a constant dwell time. From left to right, the measurement times per localization were 1, 5, and 10 orbits, respectively, with a period of 1.92 ms per orbit. A similar measurement was performed by moving the sample and keeping the TCP steady. **c)** Histograms of the data from **b)**, in which the localization uncertainty is defined as the average of half of the width and height of the ellipses. The relative area of the FoV is approximated as the number of positions in **b)** divided by the total number of positions to which we moved the nanoparticle. **d)** Our ISM-FLUX system has a minimum localization uncertainty of about 2 nm, limited by the stabilization performance and computational power needed for the MLE. For high enough photon counts, we reach this limit of $\sigma < 2$ nm, and, as a result, we can resolve a pattern with 6 nm steps. The time trace shows a portion of the calculated (x, y) coordinates as a function of time. The images show the reconstruction for all particle positions and a comparison obtained by treating the data as if they were recorded with a single-element detector. The GNP is outside the TCP for all positions. The experimental MDFs were used for analysis. See Figure 5.6 for localization results of **d)** and Figure 5.4 **a)** with the simulated MDFs.

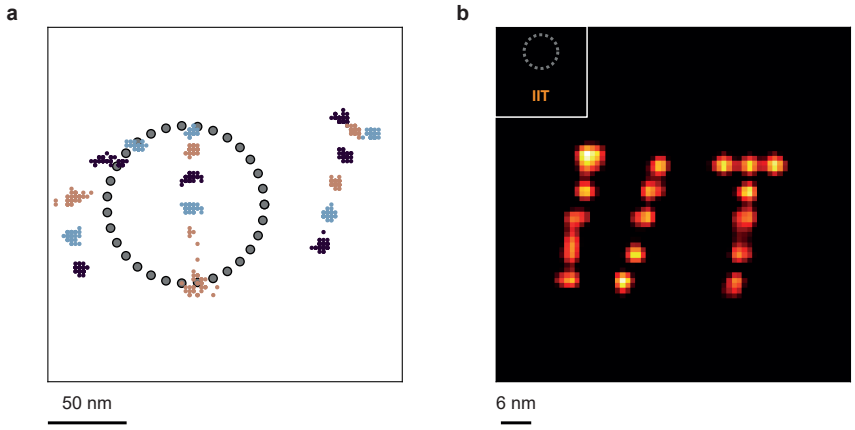

Figure 5.6 Gold nanoparticle localization results for simulated molecule detection functions. a) Shows the localization of GNP close to the TCP using a point detector. For **b)**, the GNP movement was outside the TCP, as indicated in the top left inset.

results. For a molecule outside the TCP, ISM-FLUX is less photon-efficient, as illustrated by the higher localization uncertainty. However, the spatial information from the array detector still results in an unbiased position estimation. Instead, with a single-element detector, we cannot retrieve the fluorophore position.

Having demonstrated the increased localization range of ISM-FLUX on single molecules, we now demonstrate the resolving capability of ISM-FLUX with two types of DNA origami nanorulers with labeling distances of 20 nm and 40 nm. As a switching mechanism, we use DNA-based points accumulation for imaging in nanoscale topography (DNA-PAINT), but ISM-FLUX is also compatible with other mechanisms. We first took advantage of the confocal architecture by performing conventional raster-scanned imaging over a large FoV ($> 20 \times 20\,\mu m$) to scout the sample and select origami positions. The large localization range of ISM-FLUX (about $600 \times 600\,nm^2$) eliminates the requirement for a more precise pre-localization step. For ISM-FLUX, we used an orbital scan frequency of 1.92 ms/orbit, which ensured that we could sum over many orbits for the duration of an average event time of 100 ms.

Given the ISM-FLUX measurement, ON-events were identified by thresholding the intensity trace, Figure 5.8, and the following filters were applied to the on-events: (i) a lower threshold on the total number of photons of the event, (ii) an upper threshold on the intensity fluctuations during the on-event, and (iii) a lower threshold on the duration, see Section 4.5.2.

Each remaining event was split into three equal-duration segments and localized using the maximum likelihood approach. We calculated the experimental localization uncertainty

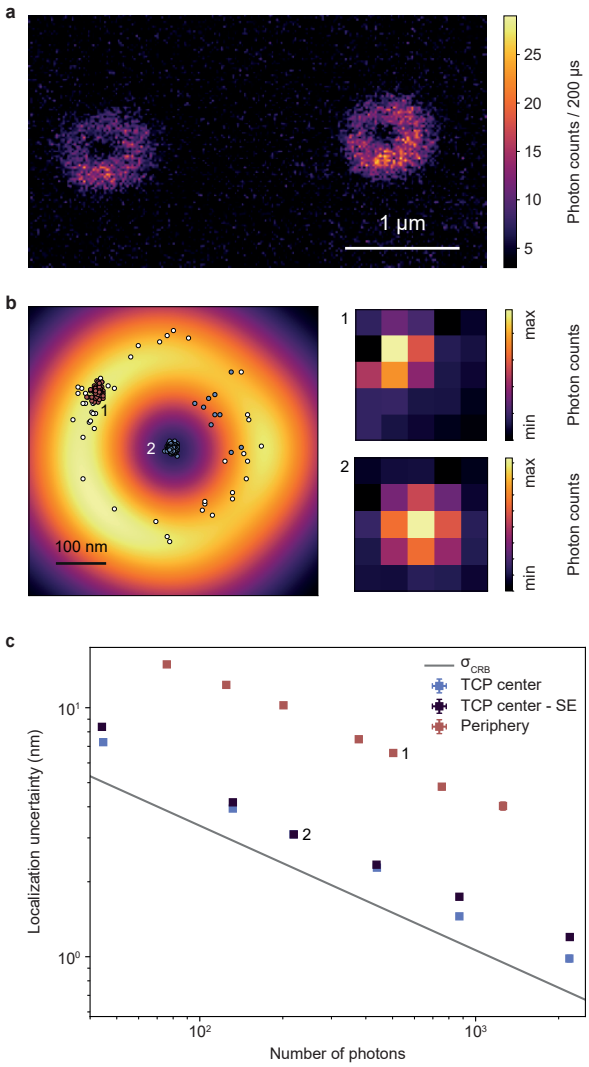

Figure 5.7 ISM-FLUX imaging and localization experiment on fixed single ATTO647N molecules. a) Raster-scanned image of individual fluorescent molecules, obtained by summing the signals of all detector elements for each scan position. **b)** ISM-FLUX on two emitters, one in the periphery (1) and another in the TCP center (2), are sequentially localized until bleached. The white scatter plot represents the same data for the peripheral molecule, analyzed as if using a single-element detector. The fingerprints show the summed photon counts over all TCP positions and orbits. **c)** Corresponding ISM-FLUX localization uncertainty for different photon counts, obtained by chopping the time trace in segments of different sizes. The two numbered data points correspond to the scatter plots in **a)**. SE = analysis as if a single-element detector was used. Error bars represent standard errors but are mostly too small to be visible. σ_{CRB} shows the CRB for the TCP center.

σ for each event, resulting in a histogram with a median value of 4.9 nm for the 40 nm origami, Figure 5.8 (b). Note that since each event is split into 3 subevents, the value of 4.9 nm corresponds to the localization uncertainty obtained for minimum $2000/3 \approx 667$ photons, of which only about 80 % (534 photons) are signal and the remaining counts are either background or dark counts. We excluded events where σ exceeded a threshold of 8.25 nm, resulting in the removal of 69 out of 284 total events (approximately 24 %). For the remaining events, each set of three localizations was averaged, leading to a final result of 215 events. Summing the micro-images of all events, Figure 5.8 (b), the so-called "fingerprint" reveals the presence of at least one origami outside the TCP. However, plotting the individual localizations and the reconstructed image, Figure 5.8 (c), shows four origami, all located outside of the TCP. Thanks to the large localization range of ISM-FLUX, all origami can be viewed in a single measurement. Conversely, when analyzing the same dataset for a single-element detector, we see a highly biased and imprecise localization map, as predicted by the estimation theory.

We repeated the imaging and analysis procedure for the origami structure with 20 nm spacing, Figure 5.8 (d, e). As shown by the fingerprint, the origami is located closer to the TCP but still outside the TCP. Because of small differences between the experimental and simulated MDFs close to the doughnut minimum caused by small aberrations in the imaging system, we used the experimental MDFs for the analysis of this origami. As a result, the localization uncertainty histogram shows a wider distribution, Figure 5.8 (e), compared to Figure 5.8 (c), for which we used the simulated MDFs. We set the threshold for accepting events to $\sigma < 5$ nm, resulting in 36 localizations for the three spots combined. The scatter plot and the reconstruction show that the 20 nm spacing between the docking sites can be resolved, Figure 5.8 (e).

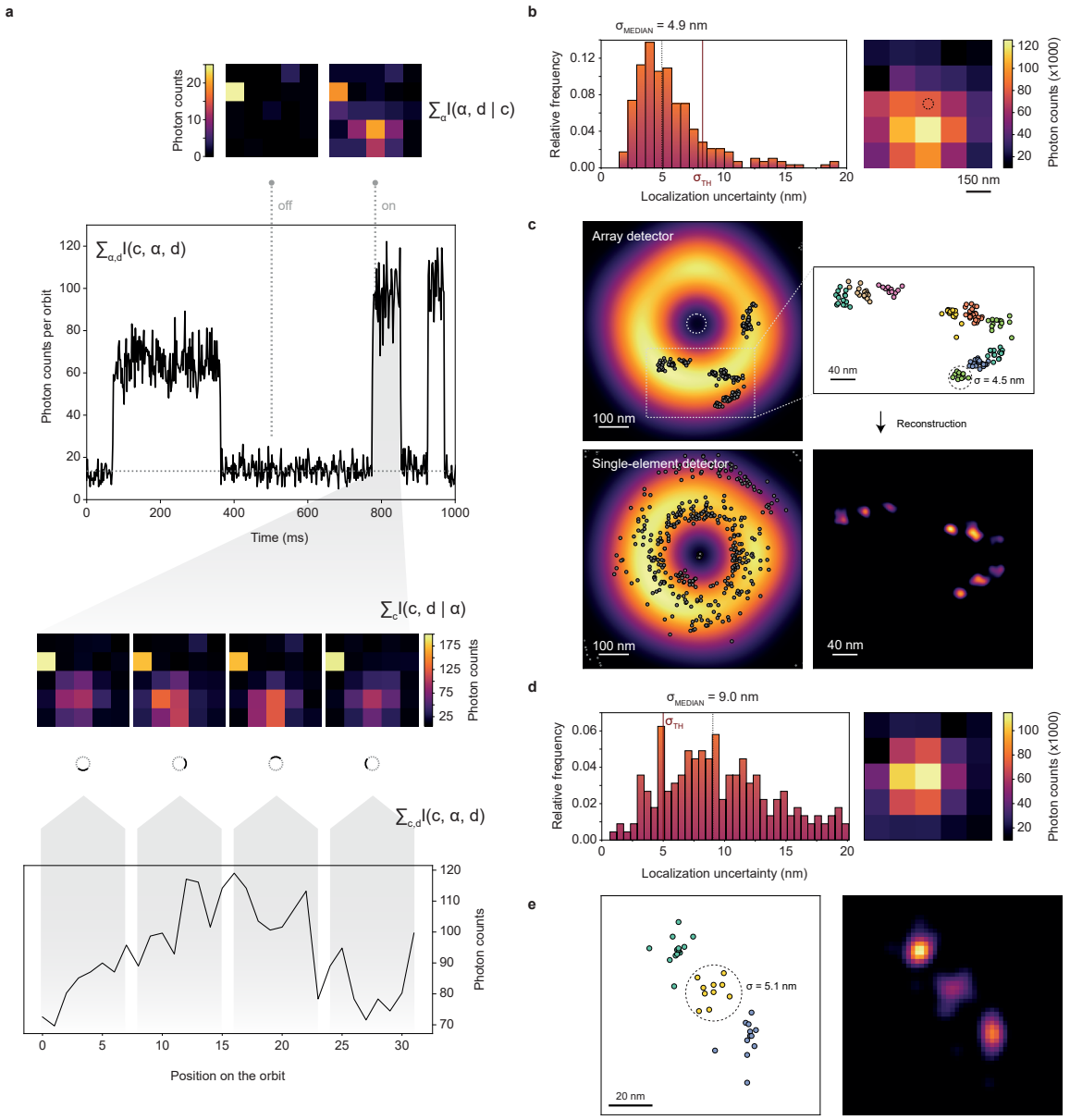

Figure 5.8 ISM-FLUX localization experiment on DNA origami nanorulers. a-c) ISM-FLUX on DNA origami with 40 nm spacing between nearest neighbors. **a)** Different representations of the same data set show the on-off blinking and the signal on the detector for a representative on-event. **b)** Histogram of the experimental localization uncertainty of all blinking events and the sum of all micro-images during the on-events. It shows a segment of a 9.6-minute ISM-FLUX measurement. **c)** Scatter plot of the localizations (top-left). Only localizations that pass through all filters are shown. The background image shows the sums of all MDFs. The colored scatter plot (right) shows a zoomed-in region, with the three docking sites of each origami shown in a different color. Classification based on k-means clustering with 9 clusters. Reconstructed images made with a Gaussian kernel with a fixed standard deviation of 6 nm. The single-element detector plot (bottom-left) shows the localizations obtained by summing the signal over all detector elements, thus ignoring the spatial information on the detection side. **d-e)** Data are similar to **b-c)** for a DNA origami sample with 20 nm spacing between nearest neighbors. The final localization uncertainty, calculated as the square root of the average variance in x and y for all localizations within a cluster, are (7.1 $\pm$ 1.1) nm and (4.9 $\pm$ 0.4) nm for the nanorulers with 40 nm spacing and 20 nm spacing, respectively.

5.3 Widefield single-molecule localization microscopy

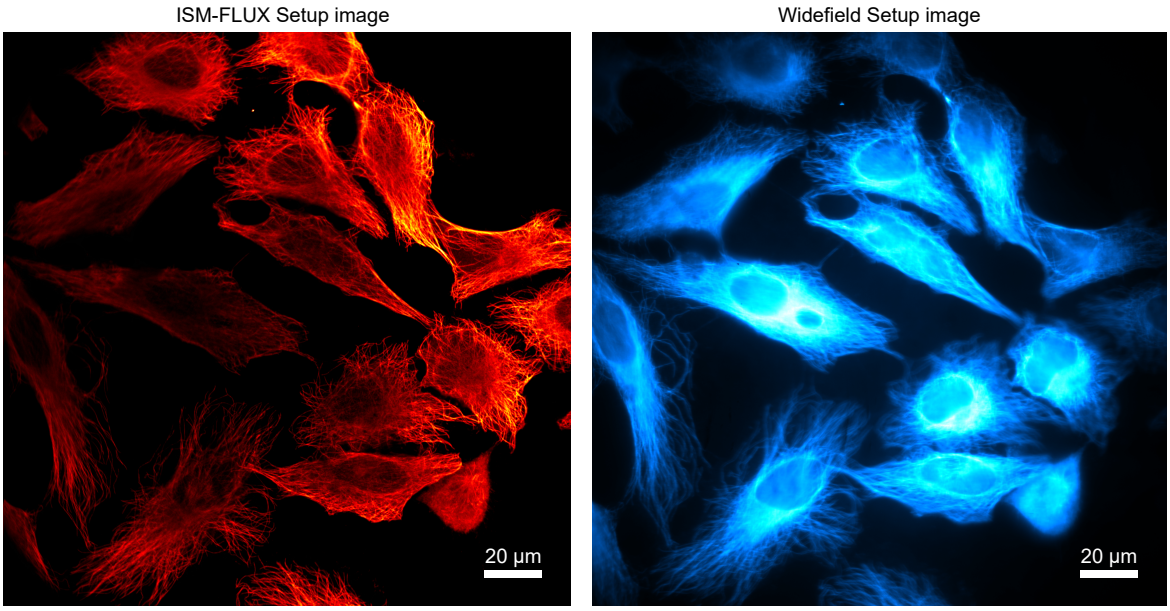

Figure 5.9 field-of-view comparison. field-of-view for the ISM and widefield system using the same sample. Left side shows the complete scanning FoV of the ISM system and right side shows the complete camera FoV of the widefield system.

As both imaging modalities share the relay optics (Widefield and ISM), the widefield SMLM path does not use a dedicated detection arm with separate lenses after the dichroic beamsplitter, as is common in traditional widefield SMLM setups. Instead, an inverse dichroic mirror transmits the excitation and reflects sample fluorescence to a scientific-complementary metal-oxide semiconductor (CMOS) camera (ORCA-Flash4.0 V3). The detection path provides 67.5× total magnification, yielding a pixel size of ∼96 nm per pixel and a widefield FoV of approximately $198 \times 198\,\mu m^2$. We verified that the ISM-FLUX scanning FoV and the widefield camera FoV coincide by imaging the sample using both modalities as shown in Figure 5.9.

Although cellular samples are commonly treated as quasi-2D structures, they exhibit significant axial thickness of several micrometers. Conventional widefield excitation (epi-illumination) generates fluorescence throughout this entire depth, producing substantial background signal from regions outside the focal plane. Total internal reflection fluorescence (TIRF) addresses this by restricting excitation to an evanescent field of approximately 100 nm to 200 nm, thereby reducing out-of-focus contribution and improving the image contrast and offering a moderate sectioning effect near the coverslip.

To characterize the TIRF imaging capabilities of our system, we inserted a Bertrand lens (phase telescope) before the camera. This configuration enables direct imaging of the objective's back focal plane (BFP) and its conjugate planes onto the camera sensor. We selected an 60 mm Bertrand lens and positioned it between the final relay lens (250 mm) and the camera to ensure proper focusing of the BFP onto the sensor.

For these tests, we used HeLa cells immunostained for α-tubulin with Alexa Fluor 647. After placing the sample on the objective and focusing the image on the camera sensor, we inserted the Bertrand lens as described above. This arrangement simultaneously imaged both the undercritical angle fluorescence (UAF) ring and the excitation beam's focal spot at the BFP, as shown in Figure 5.10 (a). In this panel, the inner part of the circular ring corresponds to UAF emission, while the central spot (marked with a blue pointer) indicates the excitation beam's position at the BFP. To map the excitation beam's trajectory across the BFP, we translated the motorized platform (containing steering mirrors and lenses) from one extreme position to the other, capturing its position at multiple points shown by images in Figure 5.10 (a). The complete BFP imaging sequence is shown in the movie here - BFP-imaging-full-sweep.

The middle image in Figure 5.10 (a) shows the beam at the BFP center, corresponding to conventional epi-illumination, while the leftmost and rightmost images in Figure 5.10 (a) captures the beam at the peripheral position approaching the critical angle, enabling TIRF illumination. Initially, visualizing fluorescence beyond the critical angle, including the supercritical angle fluorescence (SAF) ring, proved challenging due to the minimal refractive index mismatch between the mounting medium of our choice and coverslip glass. When the refractive indices of the glass and the sample are very similar, SAF and anything beyond the critical angle become harder to visualize, since the critical angle is given by $\sin\theta_c = n_{\text{sample}}/n_{\text{glass}}$. If $n_{\text{sample}} \approx n_{\text{glass}}$, then $\sin\theta_c \approx 1$, and therefore $\theta_c \approx 90^\circ$, compressing the SAF region into a very narrow annulus at the edge of the pupil. We used Invitrogen ProLong Diamond Antifade Mountant (cured refractive index ≈ 1.47) with coverslip glass ($n_{\text{glass}} \approx 1.5$). This small index difference ($\Delta n \approx 0.03$) yields a thin SAF ring which the accessible angular capabilities of the present illumination path did not permit stable excitation beyond the critical angle under these mounting conditions.

We later repeated a similar sample (α-tubulin in HeLa cells, imaged with Atto647N) using water as the medium rather than our standard mounting media. Water provided better index mismatch $\Delta n \approx 0.17$ ($n_{\text{water}} \approx 1.33$; $n_{\text{glass}} \approx 1.5$), allowing clearer visualization of region beyond the critical angle and the SAF ring through the Bertrand lens. In the leftmost and rightmost panels of Figure 5.10, acquired at TIRF incidence, a second excitation spot

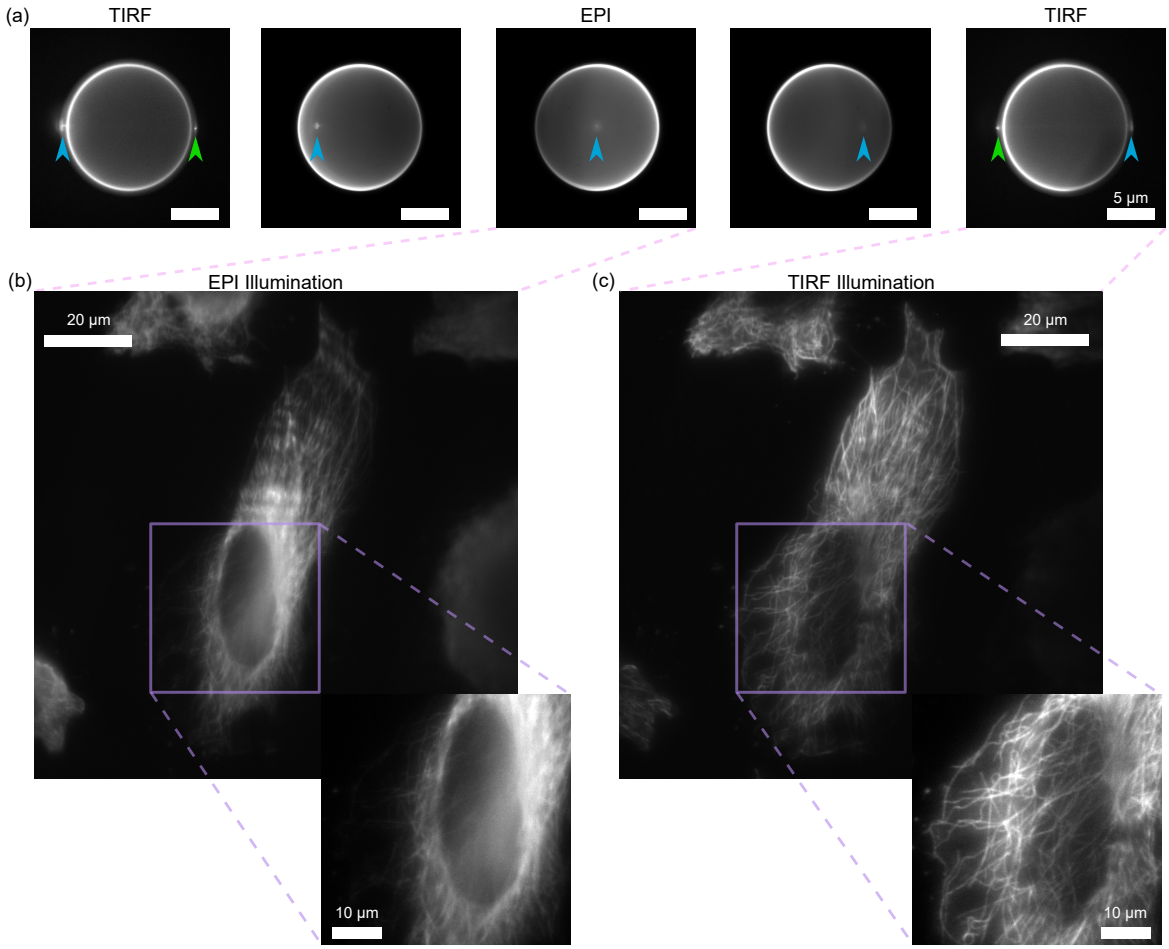

Figure 5.10 Characterization of TIRF illumination and optical sectioning: a) BFP imaging sequence showing the translation of the excitation beam from the center (epi-illumination) toward the periphery (TIRF). The central spot in the epi-illumination image (blue arrow) indicates the position of excitation beam. The spot on the opposite side in the TIRF illumination image (green arrow) shows the beam collected by the objective after reflecting from the sample. **b)** HeLa cells immunostained for α-tubulin imaged under Epi-illumination. **c)** The same sample imaged under TIRF illumination, demonstrating superior signal-to-background ratio and resolution of sub-nuclear structures (insets).

appears at the diametrically opposite edge of the ring. This spot forms when excitation light reflects from the glass–sample interface beyond the critical angle and is collected back through the objective lens. Its presence provides an independent confirmation that the illumination angle lies beyond the critical angle and that the system is operating in the TIRF regime.

After removing the Bertrand lens, we imaged the same cells under both epi-illumination (Figure 5.10 (a)-middle) and TIRF (Figure 5.10 (a)-right). Microtubule structures appeared substantially sharper in the TIRF image (Figure 5.10 (b)) than under epi-illumination (Figure 5.10 (c)). Insets focusing on the nuclear region shows that epi-illumination blurs intranuclear details due to out-of-focus fluorescence, whereas TIRF clearly resolves subnuclear structures, demonstrating the contrast enhancement provided by the evanescent field.

However, we noticed two limitations of the TIRF sectioning in our setup (Figure 5.10): (i) microtubule structures close to the cell center (near the nucleus) are not clearly visualized or resolved, and (ii) the images exhibit fringe-like intensity modulations indicative of non-uniform illumination. In our opinion these two factors may contribute:

- First, optical sectioning with TIRF is insufficient in our case: although evanescent-wave illumination is used, the high axial structural density of HeLa cells still generate out-of-focus fluorescence, preventing reliable visualization and resolution of microtubules near the center of the nucleus (Figure 5.10 (c)).

- Second, coherent laser illumination interferes with scattered light within the sample. Our spatially and temporally coherent laser beam interferes with light scattered from imperfections, surface roughness, and refractive index variations in the sample, coverslip, immersion medium, and other optical components. This creates uneven illumination with bright and dark regions formed by interference fringes, reducing image quality. Figure 5.10 (b) shows this effect at the top edge of the cell, where the illumination appears non-homogeneous with fringe-like patterns. This non-uniform excitation degrades both image homogeneity and quantitative interpretability, as has been noted previously in coherent widefield and TIRF illumination systems [152, 153].

The potential solutions to mitigate these problems are discussed in Chapter 6.

To evaluate widefield SMLM performance under DNA-PAINT conditions, we imaged α-tubulin in fixed HeLa cells and reconstructed the data using ThunderSTORM. GNPs (100 nm diameter) were drop-cast onto the coverslip to provide fiducials for the active stabilization system (5–10 particles per $30\times30\,\mu m^2$ FoV of the stabilization cameras). A total of ~47,000 frames were acquired with an exposure time of 50 ms per frame. Over the full acquisition, the

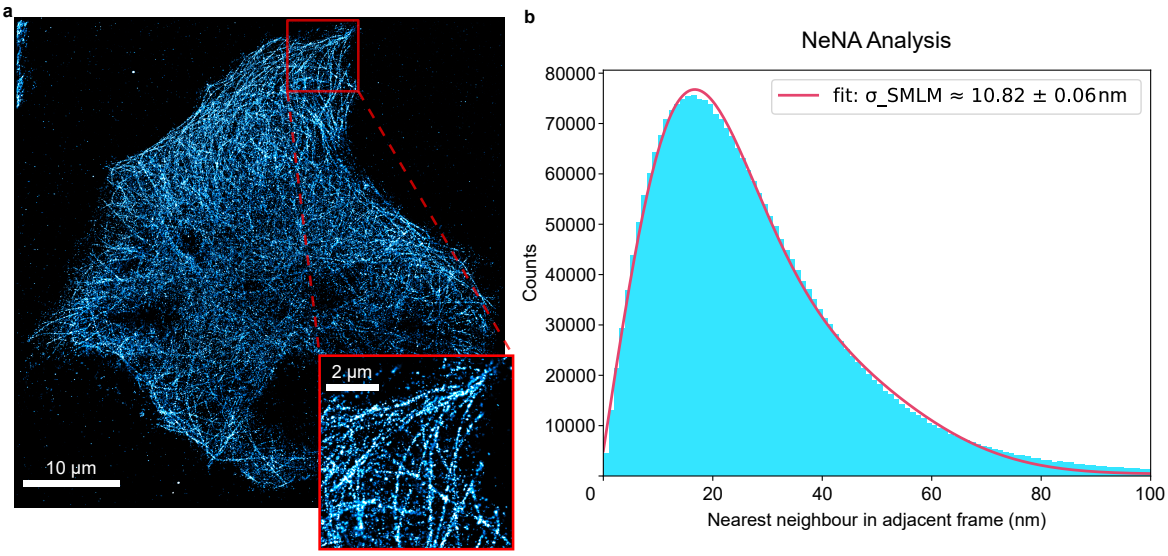

Figure 5.11 DNA-PAINT Super-resolution reconstruction of α-tubulin in HeLa cells: a) Widefield DNA-PAINT imaging of α-tubulin in HeLa cells. The image was reconstructed from 47,000 frames (50 ms exposure) using the ThunderSTORM. Inset shows the zoomed-in image of the top section of the DNA-PAINT localization result. **b)** Nearest neighbor analysis of the localization data based on [148].

measured sample position traces showed residual motion below $\sigma \approx 1.5\,\text{nm}$ along all axes while stabilization was active and therefore, no additional post-acquisition drift correction was applied. The reconstructed cell-scale microtubule network is shown in Figure 5.11 (a) as a normalized Gaussian rendering using ThunderSTORM. To quantify the performance of our system and the localizations, we computed nearest-neighbor based analysis (NeNA) using a separate Python implementation following Endesfelder et al. [148], using nearest neighbors between adjacent frames. Localizations were filtered only by fitted Sigma ($90 < \sigma < 200\text{nm}$, in ThunderSTORM units) and reported uncertainty ($< 100\,\text{nm}$). The NeNA fit yielded $\sigma_{\text{SMLM}} = (10.82 \pm 0.06)\,\text{nm}$ as shown in Figure 5.11 (b).

The measured localization precision of $\sigma_{\text{SMLM}} = (10.82 \pm 0.06)\,\text{nm}$ is consistent with values reported in the literature for DNA-PAINT imaging of microtubules in fixed cells. Endesfelder et al., who introduced the NeNA method, reported typical values of $\sim 10\,\text{nm}$ to $15\,\text{nm}$ for cellular targets under widefield illumination [148]. More recent studies using optimized configurations have achieved slightly better precisions: Stehr et al. demonstrated sub-10nm precision using flat-top TIRF illumination [154], while Zaza et al. reported 6–13nm for microtubules using spinning disk confocal with optical photon reassignment [155]. Under standard TIRF conditions comparable to our setup, Steen et al. reported cellular localization

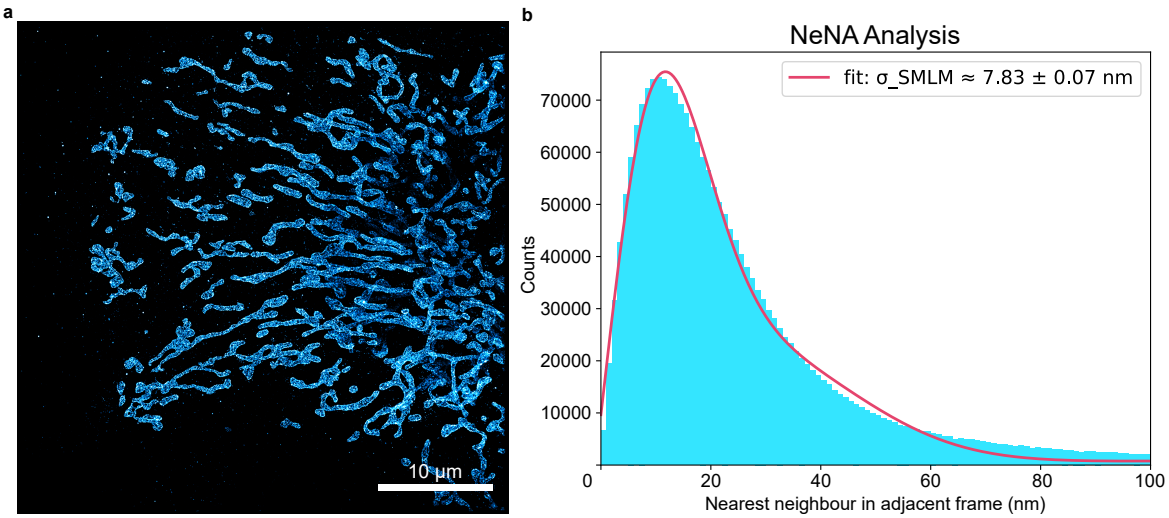

Figure 5.12 STORM Super-resolution reconstruction of TOM20 in HeLa cells: a) Widefield stochastic optical reconstruction microscopy (STORM) imaging of *TOM*20 in HeLa cells (labeled with AF647). The image was reconstructed from 60,000 frames (25 ms exposure) using the ThunderSTORM. **b)** Nearest neighbor analysis of the localization data based on [148].

precisions of 8–16nm depending on the dye and target [74]. Our result thus represents competitive performance for a widefield TIRF system.

We next imaged the mitochondria-associated target (TOM20, AF647) under STORM conditions to assess performance on a second, commonly used SMLM modality. The acquisition comprised ∼60,000 frames at 25 ms exposure per frame over an area of $\sim 43 \times 43\,\mu m^2$. The sample was pre-bleached using high excitation power to reach a sparse blinking regime, and no 405 nm activation was used during the acquisition. Active stabilization remained enabled throughout, and the measured sample position over the full acquisition showed residual motion below $\sigma \approx 1.5$ nm along all axes. Figure 5.12 (a) shows the normalized Gaussian rendering (ThunderSTORM) of the resulting localization reconstruction. NeNA was computed from the same localization list using the same adjacent-frame definition [148] and the same minimal filtering ($90 < \sigma < 200$, uncertainty < 100 nm). The fitted precision was $\sigma_{\mathrm{SMLM}} = (7.83 \pm 0.07)$ nm as shown in Figure 5.12 (b).

The measured localization precision of $\sigma_{\mathrm{SMLM}} = (7.83 \pm 0.07)$ nm is consistent with the higher photon yield typically achieved in dSTORM with Alexa Fluor 647 [156]. This result can be attributed to the fact that AF647 in thiol-based switching buffer emits several thousand photons per switching cycle. The achieved precision is also consistent with literature values for dSTORM on cellular targets; Dempsey et al. reported ∼ 8 nm to 17 nm localization precision for AF647 under optimized switching conditions [156], while Provost et al. demonstrated stable ∼ 10 nm precision for cellular dSTORM imaging [157]. Our

result of $\sigma_{\mathrm{SMLM}} = (7.83 \pm 0.07)\,\mathrm{nm}$ thus represents high-end performance for widefield dSTORM, enabled in part by the active stabilization system maintaining $\sigma \approx 1.5\,\mathrm{nm}$ residual drift throughout the acquisition.

Finally, to show the impact of stabilization on widefield SMLM over a long acquisition, we recorded TOM20-STORM datasets with stabilization enabled and disabled. The data was acquired on the same sample and acquisition parameters, but on different cells. The top row of Figure 5.13 visualizes the raw localizations as a scatter plot, with localization color mapped to frame index (0 min to 25 min across $\sim 60,000$ frames). With stabilization enabled (Figure 5.13 (a)), localizations from early and late times remain spatially co-registered at the level expected from the rendered structural width. With stabilization disabled (Figure 5.13 (b)), we observed a systematic temporal lateral displacement is apparent across multiple structures alongside a gradual axial shift that led to a loss of focus over time. Without active stabilization, the accumulated drift causes localizations from different time points to be mapped to different coordinate frames, so that repeated observations of the same structure no longer register. As a result, fine features are broadened or split into multiple displaced copies, and the reconstructed morphology is visibly degraded as shown in Figure 5.13 (b).

The NeNA analysis (bottom row, Figure 5.13) reflects this change in effective localization precision. For the stabilized dataset, NeNA results in $\sigma_{\mathrm{SMLM}} = (9.02 \pm 0.08)\,\mathrm{nm}$, whereas for the unstabilized dataset it increased to $\sigma_{\mathrm{SMLM}} = (11.97 \pm 0.11)\,\mathrm{nm}$. NeNA was computed from nearest-neighbor distances between adjacent frames and therefore reflects localization repeatability at short time separations, which is affected by photon statistics, background, point spread function (PSF) model mismatch (including defocus), and short-timescale motion [148]. In this comparison, the concurrent appearance of time-dependent spatial drift in the raw localization plot and the increase in σ_{SMLM} provide consistent evidence that suppressing sample motion is necessary to maintain localization precision over long, large-field acquisitions, particularly when axial drift introduces additional PSF variation.

To assess whether post-acquisition drift correction is beneficial when active stabilization is already engaged, we applied drift correction to the stabilized TOM20-STORM dataset (Figure 5.13 (a)) using SMAP [158]. Drift correction was performed using the RCC approach, in which the acquisition is partitioned into temporal blocks and pairwise displacements are estimated by FFT-based cross-correlation of the corresponding super-resolution images, followed by spline interpolation of the inferred drift trajectory [125]. Importantly, the RCC correction was applied to the same ThunderSTORM localization list (no re-localization in SMAP), and NeNA was computed with the same adjacent-frame definition and analysis pipeline as above.

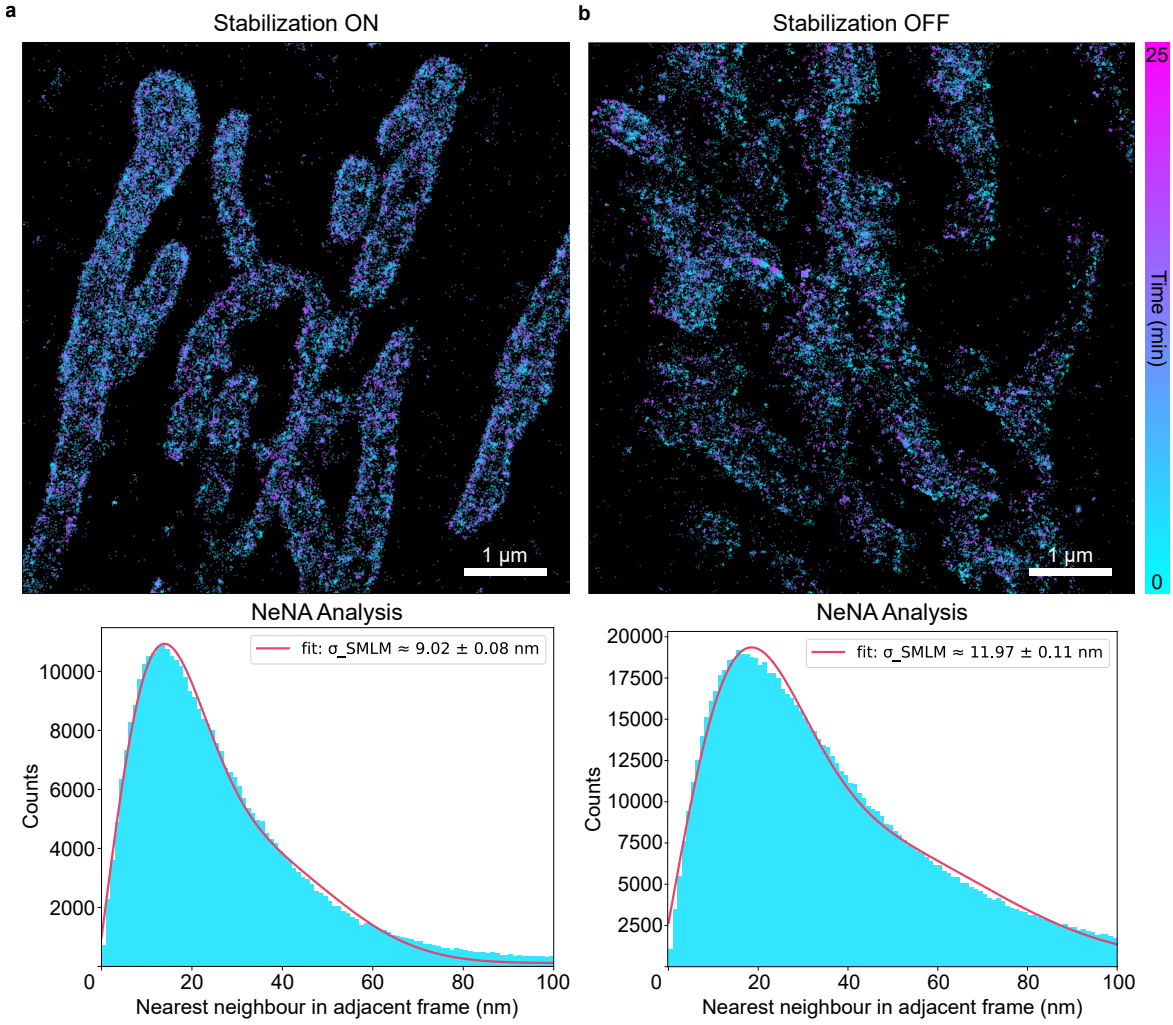

Figure 5.13 Super-resolution reconstruction with and without active stabilization. a) Widefield STORM imaging of $TOM20$ in HeLa cells with active sample stabilization. Nearest neighbor analysis results in $\sigma_{\mathrm{SMLM}} = (9.02 \pm 0.08)\,\mathrm{nm}$. **b)** Widefield STORM imaging of the same sample (different cell) with sample stabilization turned off. As a result, the localizations drift in lateral direction while also losing focus over time. This leads to a worse nearest neighbor analysis value of $\sigma_{\mathrm{SMLM}} = (11.97 \pm 0.11)\,\mathrm{nm}$.

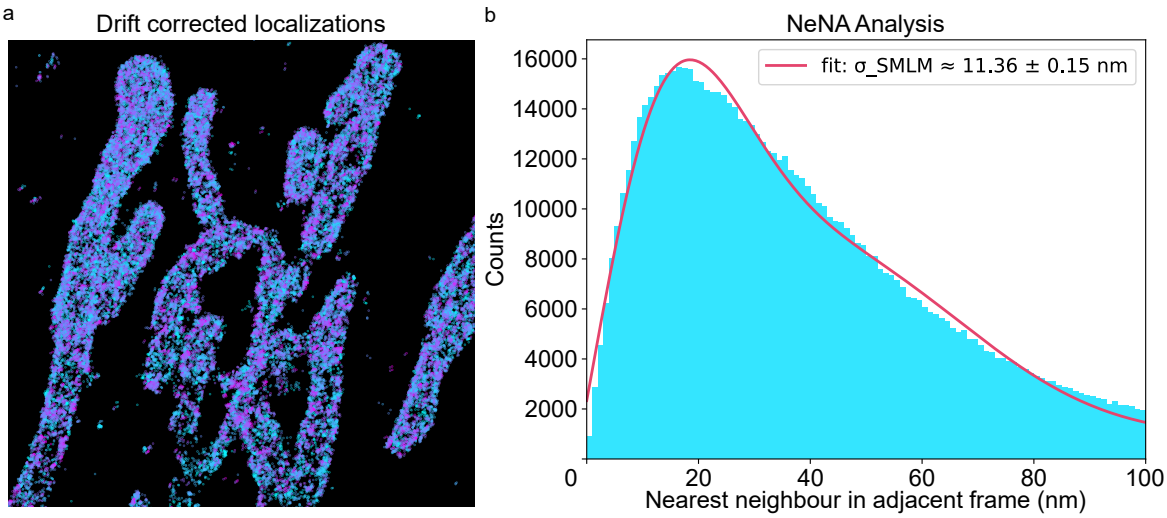

Figure 5.14 Nearest-neighbor based analysis of actively stabilized acquisition after drift correction. **a)** Shows the reconstruction of drift corrected localizations of the dataset from Figure 5.13 **a)**. **b)** Nearest-neighbor based analysis of the localization data from Figure 5.13 **a)** after post-processing drift correction using redundant cross-correlation (RCC).

After RCC correction, NeNA yielded $\sigma_{\mathrm{SMLM}} = (11.36 \pm 0.15)\,\mathrm{nm}$, compared with $\sigma_{\mathrm{SMLM}} = (9.02 \pm 0.08)\,\mathrm{nm}$ for the stabilized dataset without post-processing as shown in Figure 5.14. In this regime where the measured residual motion during acquisition is already at the nanometer scale, RCC did not improve the effective localization precision and instead introduced additional broadening to the fit. Therefore, under the present conditions, drift correction did not improve the results and can be omitted in this workflow, thus eliminating an additional post processing step in the analysis pipeline.

Chapter 6

Conclusions and Outlook

This chapter outlines a brief summary of the thesis and its main conclusions. The chapter closes with planned improvements to the instrument and experimental workflow. These developments are aimed at improving performance and simplifying day-to-day operation.

Discussion and Conclusions

This PhD thesis explored the concept of utilizing a single-photon avalanche diode (SPAD) array detector for performing single-molecule localization microscopy (SMLM) in a minimal fluorescence photon fluxes microscopy (MINFLUX)-like fashion using a standard laser scanning microscope. Taken together, this thesis establishes (i) an experimentally validated image scanning microscopy (ISM)-FLUX implementation based on orbital scanning and array detection, (ii) an analysis and calibration workflow that supports localization across an extended field-of-view (FoV) without the divergence characteristic of single-element detection outside the targeted coordinate pattern (TCP), and (iii) an integrated stabilization strategy that enables long acquisitions without relying on post-processing drift correction.

Reliable nanometer-scale localization over minutes-to-hours required active control of sample drift. For this thesis, we implemented a three-dimensional stabilization module based on real-time tracking of fiducial markers and closed-loop correction with a piezoelectric stage. We present an active 3D stabilization system for microscopy imaging, utilizing real-time monitoring of infrared back-scattering from fiducial markers within the sample. We capture and analyze widefield images at a rate of about 60 Hz and update the stage position at a rate of 20 Hz (Section 4.2.1). In contrast to some earlier methods, our system benefits from parallel processing, enabling significantly faster image acquisition and drift correction. This allows for a more rapid response to sample movement and enhances the overall system stability. The

20 Hz limit in our system is not an inherent restriction of the software but rather a limitation imposed by the digital communication protocol of the piezoelectric stage used in this setup. Our approach yields sub-nanometer precision across all three spatial dimensions for several hours. The optical system is compact and portable, facilitating easy installation and transfer between different microscopes. The software is open-source and runs entirely on the CPU, removing the need for specialized hardware, such as a high-end GPU-equipped computer (Section 4.3.1). We anticipate that these advantages will make super-resolution microscopy, as well as time-lapse confocal and widefield microscopy experiments, more accessible.

In closed loop, the residual sample motion measured by the stabilization readout remained below $\sim 1\,\text{nm}$ standard deviation over 1000 s, while an independent validation in the microscope imaging channel yielded sub-nanometer to nanometer-scale drift over 4 h and prevented the multi-hundred-nanometer displacements observed when stabilization was disabled (Section 5.1). These results set the practical floor for high-precision ISM-FLUX measurements and reduce the need for drift correction steps that would otherwise be required in post-processing.

We demonstrated that ISM-FLUX can localize a single fluorescent molecule with an uncertainty between 2 nm to 10 nm for ~ 1000 photons and an orbital scanning diameter $L = 90\,\text{nm}$. Thanks to the unique spatial information leveraged by the array detector, we achieved this localization uncertainty within a $600 \times 600\,\text{nm}^2$ region. Specifically, the localization uncertainty depends on the position of the fluorophore relative to the TCP, which is best at the TCP center. In contrast to MINFLUX, the localization uncertainty in ISM-FLUX does not diverge for fluorophores outside of the TCP, resulting in greater robustness and an extended localization range (Section 5.2).

As a result, the typical iterative procedure of MINFLUX in which the position and diameter of the TCP are changed is not crucial. Probing the central TCP point is not required to have a single local maxima in the maximum likelihood estimator (MLE) map for the array detector [10, 129, 130]. The ability to have a pure circular TCP is advantageous because a galvanometric scanning system of a confocal microscope can perform orbital scanning at high speeds [159, 160]. Due to various factors such as inertia, the actual scan diameter performed by the galvanometric scan system is lower than the diameter as imposed by the signal delivered to the galvanometer drivers by the microscope control software (Section 4.3). Therefore, a calibration step to derive the exact orbit parameters is crucial for the ISM-FLUX data analysis. However, having an array detector in de-scanned mode greatly simplifies this procedure, as this configuration allows self-calibration (Section 4.5.2).

A limitation of ISM-FLUX and MINFLUX is their lower localization throughput compared to camera-based SMLM due to the sequential single-molecule localization inherent in these approaches. Shorter fluorophore on-times and higher rates of on-events could partially speed up acquisition. In DNA-based points accumulation for imaging in nanoscale topography (DNA-PAINT), a higher rate of on-events can be achieved by increasing the imager concentration, but this would significantly raise background noise. Alternatively, other labeling strategies can be explored, ranging from common stochastically photoswitchable dyes and fluorogenic dyes to newly developed photoactivatable dyes [161]. In this context, ISM-FLUX, by using an array detector, opens up the possibility of localizing multiple fluorophores simultaneously [129], which can further increase throughput. In practice, this class of scanning-based localization methods benefits from active stabilization because the low number of localizations per unit time limits the reliability of drift estimation from the signal channel alone. A current technical drawback of ISM-FLUX is the lower photon-collection efficiency of SPAD array detectors compared to the single-element SPAD detectors typically used in MINFLUX, particularly when using far-red fluorophores. However, we believe that this limitation is expected to improve with the introduction of red-enhanced SPAD array detectors [108, 160].

In this ISM-FLUX implementation, single-molecule localization is limited to two dimensions, but an extension to 3D is conceivable in several ways: encoding the axial information (i) on the detection side, (ii) on the excitation side, or (iii) in the fluorescence lifetime. In the first approach, the axial position can be encoded in the molecule image by placing a cylindrical lens in front of the array detector, as was shown in [111] for tracking and commonly used in camera-based localization. This approach is technically simple but does not leverage the benefits of the MINFLUX principle along the axial dimension. In the second approach, one can generate an intensity minimum in the axial direction on the excitation side and translate this along the third dimension, similar to 3D MINFLUX [11]. Lastly, the axial position can be encoded in the fluorescence lifetime when combined with metal-induced energy transfer (MIET) [162] or graphene energy transfer (GET) [104]: the reduction of the molecule's lifetime when it is in proximity to a metal or graphene film can be measured to determine its axial position. However, this method restricts localization to a thin layer near the film. Furthermore, lifetime imaging using MINFLUX and its derivatives is a challenging task due to the limited number of signal photons from the sample.

While ISM-FLUX does not require an iterative procedure like MINFLUX, it could still benefit from one. Specifically, a two-step ISM-FLUX implementation can be envisioned: the first step coarsely localizes the molecule within a $600 \times 600\,\mathrm{nm}^2$ region, and in the

second step, the TCP is centered on the molecule to fully leverage the MINFLUX concept. This implementation would also guarantee consistent localization uncertainty across the entire localization range of ISM-FLUX. An advantage of this implementation is its natural extension to real-time single-molecule tracking. In this scenario, another benefit of the large localization range of ISM-FLUX is that an incorrect real-time estimate of the emitter position, and consequently an incorrect updated TCP position, will not result in a failed localization. Instead, it will only lead to a suboptimal localization in terms of photon efficiency.

Essentially, from an optical architecture perspective, converting a confocal laser scanning microscope into an ISM-FLUX system requires only the addition of a phase plate or a spatial light modulator, and an array detector. From the microscope control perspective, circular scanning is necessary instead of conventional raster scanning. Combined with the greater penetration depth compared to total internal reflection-based techniques, the self-calibration procedure, and automated analysis, we foresee a future where ISM-FLUX enables the easy integration of SMLM with single-digit nanometer localization uncertainty into existing confocal microscopes. On one hand, we demonstrated that asynchronous readout SPAD arrays enable a straightforward confocal-based implementation of the MINFLUX principle, with only a relatively small compromise in localization performance. On the other hand, this work highlights the significant advantages that also the original iterative MINFLUX implementation may achieve by replacing the conventional single-element detector with a SPAD array.

To benchmark performance in a complementary modality and to support correlative measurements on the same optical platform, we integrated a widefield SMLM detection path sharing the objective and tube lens with ISM-FLUX (Section 4.2.3). Using this subsystem, we reconstructed widefield SMLM datasets acquired under DNA-PAINT and direct-stochastic optical reconstruction microscopy (STORM) conditions, and quantified localization repeatability using nearest-neighbor based analysis (NeNA). Across these datasets, NeNA yielded localization repeatability in the single-digit to low-double-digit nanometer range, with the best case reaching $\sigma_{\mathrm{NeNA}} = 7.83\,\mathrm{nm}$. In a representative long acquisition, NeNA increased from 9.02 nm with active stabilization to 11.97 nm without stabilization (Section 5.3, consistent with the time-dependent displacement and focus degradation observed in the unstabilized reconstruction.

Within this thesis, the widefield and scanning modalities have been implemented and characterized independently, serving complementary roles: widefield SMLM provides structural context over a larger FoV, whereas ISM-FLUX targets selected regions with improved photon efficiency and a larger localization range than single-element MINFLUX-like im-

plementations. The shared stabilization and shared coordinate frame between modalities are designed to reduce ambiguity when comparing reconstructions and selecting regions of interest for high-precision measurements; a correlative workflow that is planned as the next experimental step.

The combined results define practical operating regimes and limitations that are directly relevant for applying MINFLUX-like localization in confocal architectures and for integrating scanning and camera-based nanoscopy within a common, stabilized coordinate frame. We believe that the combination of asynchronous readout SPAD arrays and MINFLUX will form a powerful and transformative union. This potential will be further enhanced by integrating photon time-tagging capabilities to fully realize the possibilities of the emerging photon-resolved microscopy paradigm [37, 146].

Future plans

The next development step is to reduce manual intervention during experiments and to improve illumination uniformity in widefield SMLM. In the current implementation, switching between ISM-FLUX and widefield SMLM relies on manual intervention. We plan to automate this mode switching by motorizing the relevant fold mirror using a software-controlled linear translation stage (Thorlabs LTS150C/M; 150 mm Translation Stage with Stepper Motor), which should improve repeatability and reduce the risk of small alignment changes introduced by handling. We further plan to use COS7 cells for our SMLM experiments, which prove more suitable for total internal reflection fluorescence (TIRF) applications due to their flatter morphology and superior substrate adhesion, reducing axial thickness [163]. In addition, we plan to add a speckle reducer in the STORM excitation path (before the removable mirror shared between modalities). This will help reduce the coherence-related interference fringes caused by stray scattering from the sample and optical components, which leads to spatially non-uniform excitation. Temporally averaging these patterns is expected to homogenize the illumination profile and, in practice, to reduce frame-to-frame and field-to-field variations in localization density [164, 165]. A second hardware extension is the addition of further excitation wavelengths to broaden the range of compatible labels and enable multicolour experiments. We plan to integrate 405 nm laser as an activation/reactivation channel for STORM and 488 nm laser to support multicolour SMLM workflows, including correlative measurements in which a second spectral channel provides complementary context.

For the sample stabilization setup, we plan for an alternative marker-free stabilization approach to expand the system's applicability in scenarios unsuitable for fiducial marker utilization. We plan to do this using the lens-gated switching mechanism. This approach would involve a single, switchable lens, enabling the system to transition between fiducial-based and reflection-based stabilization modes dynamically. Such a capability would be particularly beneficial in experimental workflows demanding the execution of diverse imaging modalities on a single imaging setup. The lens-gated system would facilitate rapid on-the-fly adaptation between these stabilization methods, potentially accommodating distinct imaging techniques during imaging experiments.

On the analysis and application side, we plan to streamline the widefield localization workflow beyond the current ThunderSTORM-based processing by adopting high-density and/or deep learning based fitting approaches where appropriate. We plan to use tools including DAOSTORM for multi-emitter fitting [115] and DECODE for dense localization with uncertainty estimation [166]. The goal is a more standardized pipeline with explicit quality control and fewer manual parameter choices, so that comparisons across datasets and modalities are better controlled. Finally, we plan to implement a correlative workflow in which the same sample is first imaged with STORM to provide large-field structural context and then imaged with ISM-FLUX for nanometer-scale precision of selected regions of interest.

In conclusion, incorporating these new improvements and adding new features to the system will significantly broaden the range of applications that it can be used for. Our technique, which combines simplicity and informativeness, has the potential to set a new standard. With the rapid development of SPAD array technologies, we believe that our approach represents a promising avenue for single-molecule localization microscopy.

References

1. Alberts, B. *Molecular Biology of the Cell* 6th ed. 1464 pp. ISBN: 978-1-315-73536-8 (W.W. Norton & Company, New York, Aug. 7, 2017).
2. Dai, X. & Shen, L. Advances and Trends in Omics Technology Development. *Frontiers in Medicine* **9.** ISSN: 2296-858X (July 1, 2022).
3. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. *Genome Biology* **18,** 83. ISSN: 1474-760X (May 5, 2017).
4. Lichtman, J. W. & Conchello, J.-A. Fluorescence microscopy. *Nature Methods* **2,** 910–919. ISSN: 1548-7105 (Dec. 2005).
5. Tsien, R. Y. The green fluorescent protein. *Annual Review of Biochemistry* **67,** 509–544. ISSN: 0066-4154 (1998).
6. Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie* **9,** 413–468. ISSN: 0176-7364 (Dec. 1, 1873).
7. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters* **19,** 780–782. ISSN: 0146-9592 (June 1, 1994).
8. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science (New York, N.Y.)* **313,** 1642–1645. ISSN: 1095-9203 (Sept. 15, 2006).
9. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3,** 793–796. ISSN: 1548-7105 (Oct. 2006).
10. Balzarotti, F. *et al.* Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355,** 606–612 (Feb. 10, 2017).
11. Gwosch, K. C. *et al.* MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells. *Nature Methods* **17,** 217–224. ISSN: 1548-7105 (Feb. 2020).
12. Marvin, M. *US Patent* 3013467A (1961).
13. Wilson, T. *Theory and practice of scanning optical microscopy* in collab. with Internet Archive. 234 pp. ISBN: 978-0-12-757760-9 (London ; Orlando : Academic Press, 1984).
14. *Handbook Of Biological Confocal Microscopy* (ed Pawley, J. B.) (Springer US, Boston, MA, 2006).
15. SHEPPARD, C. J. R. Super-resolution in confocal imaging. *Super-resolution in confocal imaging* **80,** 53–54. ISSN: 0030-4026 (1988).

16. Müller, C. B. & Enderlein, J. Image scanning microscopy. *Physical Review Letters* **104,** 198101. ISSN: 1079-7114 (May 14, 2010).
17. Zaleckas, V. J. Theory and Practice of Scanning Optical Microscopy By Tony Wilson and Colin J. R. Sheppard. *MRS Bulletin* **10,** 2–2. ISSN: 1938-1425 (May 1, 1985).
18. Goodman, J. W. *Introduction to Fourier Optics* 520 pp. ISBN: 978-0-9747077-2-3 (Roberts and Company Publishers, 2005).
19. Brakenhoff, G. J., Blom, P. & Barends, P. Confocal scanning light microscopy with high aperture immersion lenses. ISSN: 1365-2818.
20. Egger, M. D. & Petráň, M. New Reflected-Light Microscope for Viewing Unstained Brain and Ganglion Cells. *Science* (July 21, 1967).
21. Petráň, M., Hadravský, M., Egger, M. D. & Galambos, R. Tandem-Scanning Reflected-Light Microscope*. *JOSA* **58,** 661–664 (May 1, 1968).
22. White, J. G., Amos, W. B. & Fordham, M. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *Journal of Cell Biology* **105,** 41–48. ISSN: 0021-9525 (July 1, 1987).
23. *CSU-W1 - Confocal Scanner Unit | Yokogawa America* https://www.yokogawa.com/us/solutions/products-and-services/life-science/spinning-disk-confocal/csu-w1-confocal-scanner-unit/ (2025).
24. Tiziani, H. J. & Uhde, H. M. Three-dimensional analysis by a microlens-array confocal arrangement. *Applied Optics* **33,** 567–572. ISSN: 1559-128X (Feb. 1, 1994).
25. Azuma, T. & Kei, T. Development of a High-speed Super-resolution Confocal Scanner. **60** (2017).
26. Sheppard, C. J. R. & Cogswell, C. J. Confocal Microscopy with Detector Arrays. *Journal of Modern Optics* **37,** 267–279 (1990).
27. Sheppard, C. J. R., Mehta, S. B. & Heintzmann, R. Superresolution by image scanning microscopy using pixel reassignment. *Optics Letters* **38,** 2889–2892. ISSN: 1539-4794 (Aug. 1, 2013).
28. De Luca, G. M. R. *et al.* Re-scan confocal microscopy: scanning twice for better resolution. *Biomedical Optics Express* **4,** 2644–2656 (Nov. 1, 2013).
29. Scipioni, L., Lanzanò, L., Diaspro, A. & Gratton, E. Comprehensive correlation analysis for super-resolution dynamic fingerprinting of cellular compartments using the Zeiss Airyscan detector. *Nature Communications* **9** (Nov. 30, 2018).
30. Buttafava, M. *et al.* SPAD-based asynchronous-readout array detectors for image-scanning microscopy. *Optica* **7,** 755–765. ISSN: 2334-2536 (July 20, 2020).
31. Castello, M. *et al.* A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nature Methods* **16,** 175–178. ISSN: 1548-7105 (Feb. 2019).
32. Sheppard, C. J. R. The Development of Microscopy for Super-Resolution: Confocal Microscopy, and Image Scanning Microscopy. *Applied Sciences* **11,** 8981. ISSN: 2076-3417 (Jan. 2021).
33. Huff, J. The Fast mode for ZEISS LSM 880 with Airyscan: high-speed confocal imaging with super-resolution and improved signal-to-noise ratio. *Nature Methods* **13,** i–ii. ISSN: 1548-7091, 1548-7105 (Nov. 2016).

34. Huff, J. in *Advanced Optical Methods for Brain Imaging* (eds Kao, F.-J., Keiser, G. & Gogoi, A.) 83–102 (Springer, Singapore, 2019). ISBN: 978-981-10-9020-2.
35. Wu, X. & Hammer, J. A. in *Confocal Microscopy: Methods and Protocols* (eds Brzostowski, J. & Sohn, H.) 111–130 (Springer US, New York, NY, 2021). ISBN: 978-1-0716-1402-0.
36. Tortarolo, G. *et al.* Focus image scanning microscopy for sharp and gentle super-resolved microscopy. *Nature Communications* **13,** 7723. ISSN: 2041-1723 (Dec. 13, 2022).
37. Tortarolo, G. *et al.* Compact and effective photon-resolved image scanning microscope. *Advanced Photonics* **6,** 016003. ISSN: 2577-5421, 2577-5421 (Jan. 2024).
38. Scott, R., Jiang, W., Qian, X. & Deen, M. J. A Multi-Time-Gated SPAD Array with Integrated Coarse TDCs. *Electronics* **11,** 2015. ISSN: 2079-9292 (Jan. 2022).
39. Koho, S. *et al.* Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nature Communications* **10,** 3103. ISSN: 2041-1723 (July 15, 2019).
40. Perego, E. *et al.* Single-photon microscopy to study biomolecular condensates. *Nature Communications* **14,** 8224. ISSN: 2041-1723 (Dec. 12, 2023).
41. Nikon Corporation. *Nikon introduces the AX R MP with NSPARC Super-resolution Multiphoton Confocal Microscope* https://www.nikon.com/company/news/2023/0727_ax_r_mp_01.html (2025).
42. Nikon Corporation. *AX / AX R with NSPARC Brochure* (2024).
43. Nikon Instruments Inc. *AX / AX R with NSPARC | Specifications* https://www.microscope.healthcare.nikon.com/products/confocal-microscopes/ax/specifications (2025).
44. Goswami, M., Lachmann, R., Kretschmer, R. & Heintzmann, R. simpleISM—A straight forward guide to upgrade from confocal to ISM. *PLOS ONE* **17,** e0279378. ISSN: 1932-6203 (Dec. 27, 2022).
45. Zunino, A., Castello, M. & Vicidomini, G. Reconstructing the image scanning microscopy dataset: an inverse problem. *Inverse Problems* **39,** 064004. ISSN: 0266-5611 (Apr. 2023).
46. Zunino, A. *et al.* Open-source tools enable accessible and advanced image scanning microscopy data analysis. *Nature Photonics* **17,** 457–458. ISSN: 1749-4893 (June 2023).
47. Zunino, A. *et al.* Structured detection for simultaneous super-resolution and optical sectioning in laser scanning microscopy. *Nature Photonics* **19,** 888–897. ISSN: 1749-4893 (Aug. 2025).
48. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences* **97,** 8206–8210 (July 18, 2000).
49. *Press release: The Nobel Prize in Chemistry 2014* NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2014/press-release/ (2025).
50. Hess, S. T., Girirajan, T. P. K. & Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal* **91,** 4258–4272. ISSN: 0006-3495 (Dec. 1, 2006).

51. Jungmann, R. *et al.* Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami. *Nano Letters* **10,** 4756–4761. ISSN: 1530-6984 (Nov. 10, 2010).
52. Sharonov, A. & Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences* **103,** 18911–18916 (Dec. 12, 2006).
53. Huang, B., Bates, M. & Zhuang, X. Super-resolution fluorescence microscopy. *Annual Review of Biochemistry* **78,** 993–1016. ISSN: 1545-4509 (2009).
54. Gustafsson, M. G. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* **198,** 82–87. ISSN: 0022-2720 (Pt 2 May 2000).
55. Gustafsson, M. G. L. *et al.* Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination. *Biophysical Journal* **94,** 4957–4970. ISSN: 0006-3495 (June 15, 2008).
56. Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature Reviews. Molecular Cell Biology* **18,** 685–701. ISSN: 1471-0080 (Nov. 2017).
57. Schermelleh, L. *et al.* Super-resolution microscopy demystified. *Nature Cell Biology* **21,** 72–84. ISSN: 1476-4679 (Jan. 2019).
58. Willig, K. I., Harke, B., Medda, R. & Hell, S. W. STED microscopy with continuous wave beams. *Nature Methods* **4,** 915–918. ISSN: 1548-7105 (Nov. 2007).
59. Vicidomini, G. *et al.* Sharper low-power STED nanoscopy by time gating. *Nature Methods* **8,** 571–573. ISSN: 1548-7105 (June 5, 2011).
60. Moffitt, J. R., Osseforth, C. & Michaelis, J. Time-gating improves the spatial resolution of STED microscopy. *Optics Express* **19,** 4242–4254. ISSN: 1094-4087 (Feb. 28, 2011).
61. Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proceedings of the National Academy of Sciences* **102,** 17565–17569 (Dec. 6, 2005).
62. Andresen, M. *et al.* Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proceedings of the National Academy of Sciences* **102,** 13070–13074 (Sept. 13, 2005).
63. Habuchi, S. *et al.* Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. *Proceedings of the National Academy of Sciences* **102,** 9511–9516 (July 5, 2005).
64. Gustafsson, M. G. L. Nonlinear Structured-Illumination Microscopy: Wide-Field Fluorescence Imaging with Theoretically Unlimited Resolution. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 13081–13086 (Sept. 13, 2005).
65. Heintzmann, R., Jovin, T. M. & Cremer, C. Saturated Patterned Excitation Microscopy—A Concept for Optical Resolution Improvement. *Journal of the Optical Society of America A* **19,** 1599–1609 (Aug. 2002).

66. Ingerman, E. A., London, R. A., Heintzmann, R. & Gustafsson, M. G. L. Signal, Noise and Resolution in Linear and Nonlinear Structured-Illumination Microscopy. *Journal of Microscopy* **273,** 3–25 (Jan. 2019).

67. Heilemann, M. *et al.* Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie (International Ed. in English)* **47,** 6172–6176. ISSN: 1521-3773 (2008).

68. Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods* **8,** 1027–1036. ISSN: 1548-7105 (Dec. 2011).

69. Huang, B., Wang, W., Bates, M. & Zhuang, X. Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy. *Science* **319,** 810–813 (Feb. 8, 2008).

70. Weber, M. *et al.* MINSTED fluorescence localization and nanoscopy. *Nature Photonics* **15,** 361–366. ISSN: 1749-4893 (May 2021).

71. Weber, M. *et al.* MINSTED nanoscopy enters the Ångström localization range. *Nature Biotechnology* **41,** 569–576. ISSN: 1546-1696 (Apr. 2023).

72. Juette, M. F. *et al.* Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nature Methods* **5,** 527–529. ISSN: 1548-7105 (June 2008).

73. Nieves, D. J., Gaus, K. & Baker, M. A. B. DNA-Based Super-Resolution Microscopy: DNA-PAINT. *Genes* **9,** 621. ISSN: 2073-4425 (Dec. 11, 2018).

74. Steen, P. R. *et al.* The DNA-PAINT palette: a comprehensive performance analysis of fluorescent dyes. *Nature Methods* **21,** 1755–1762. ISSN: 1548-7091 (Aug. 2024).

75. Mortensen, K. I., Churchman, L. S., Spudich, J. A. & Flyvbjerg, H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods* **7,** 377–381. ISSN: 1548-7105 (May 2010).

76. Lindén, M., Ćurić, V., Amselem, E. & Elf, J. Pointwise error estimates in localization microscopy. *Nature Communications* **8,** 15115. ISSN: 2041-1723 (May 3, 2017).

77. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nature Protocols* **12,** 1198–1228. ISSN: 1750-2799 (June 2017).

78. Strauss, S. & Jungmann, R. Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT. *Nature Methods* **17,** 789–791. ISSN: 1548-7091 (Aug. 2020).

79. Pavani, S. R. P. *et al.* Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences* **106,** 2995–2999 (Mar. 3, 2009).

80. Shtengel, G. *et al.* Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proceedings of the National Academy of Sciences* **106,** 3125–3130 (Mar. 3, 2009).

81. Aquino, D. *et al.* Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. *Nature Methods* **8,** 353–359. ISSN: 1548-7105 (Apr. 2011).

82. Huang, F. *et al.* Ultra-High Resolution 3D Imaging of Whole Cells. *Cell* **166,** 1028–1040. ISSN: 0092-8674, 1097-4172 (Aug. 11, 2016).

83. Huang, B. An in-depth view. *Nature Methods* **8,** 304–305. ISSN: 1548-7105 (Apr. 2011).
84. Ovesný, M., Křížek, P., Borkovec, J., Svindrych, Z. & Hagen, G. M. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics (Oxford, England)* **30,** 2389–2390. ISSN: 1367-4811 (Aug. 15, 2014).
85. Jungmann, R. *et al.* Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nature Methods* **11,** 313–318. ISSN: 1548-7105 (Mar. 2014).
86. Jungmann, R. *et al.* Quantitative super-resolution imaging with qPAINT. *Nature Methods* **13,** 439–442. ISSN: 1548-7105 (May 2016).
87. Thompson, R. E., Larson, D. R. & Webb, W. W. Precise Nanometer Localization Analysis for Individual Fluorescent Probes. *Biophysical Journal* **82,** 2775–2783. ISSN: 0006-3495 (May 1, 2002).
88. Gu, L. *et al.* Molecular resolution imaging by repetitive optical selective exposure. *Nature Methods* **16,** 1114–1118. ISSN: 1548-7105 (Nov. 2019).
89. Gu, L. & Ji, W. Recent progress on single-molecule localization microscopy. *Biophysics Reports* **7,** 365–376. ISSN: 2364-3439 (Oct. 31, 2021).
90. Reinhardt, S. C. M. *et al.* Ångström-resolution fluorescence microscopy. *Nature* **617,** 711–716. ISSN: 1476-4687 (May 2023).
91. Masullo, L. A. *et al.* Ångström-resolution imaging of cell-surface glycans. *Nature Nanotechnology,* 1–7. ISSN: 1748-3395 (July 28, 2025).
92. Pachmayr, I. *et al.* Resolving the structural basis of therapeutic antibody function in cancer immunotherapy with RESI. *Nature Communications* **16,** 6768. ISSN: 2041-1723 (July 23, 2025).
93. Gu, L. *et al.* Molecular-scale axial localization by repetitive optical selective exposure. *Nature Methods* **18,** 369–373. ISSN: 1548-7105 (Apr. 2021).
94. Thiele, J. C. *et al.* Confocal Fluorescence-Lifetime Single-Molecule Localization Microscopy. *ACS Nano* **14,** 14190–14200. ISSN: 1936-0851 (Oct. 27, 2020).
95. Oleksiievets, N. *et al.* Single-Molecule Fluorescence Lifetime Imaging Using Wide-Field and Confocal-Laser Scanning Microscopy: A Comparative Analysis. *Nano Letters* (July 6, 2022).
96. Radmacher, N. *et al.* Doubling the resolution of fluorescence-lifetime single-molecule localization microscopy with image scanning microscopy. *Nature Photonics* **18,** 1059–1066. ISSN: 1749-4885. arXiv: 2404.12186 [physics.optics] (11 Nov. 2024).
97. Schmidt, R. *et al.* MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. *Nature Communications* **12,** 1478. ISSN: 2041-1723 (Mar. 5, 2021).
98. Pape, J. K. *et al.* Multicolor 3D MINFLUX nanoscopy of mitochondrial MICOS proteins. *Proceedings of the National Academy of Sciences of the United States of America* **117,** 20607–20614. ISSN: 1091-6490 (Aug. 25, 2020).
99. Sahl, S. J. *et al.* Direct optical measurement of intramolecular distances with angstrom precision. *Science* (Oct. 11, 2024).

100. Rickert, J. D., Held, M. O., Engelhardt, J. & Hell, S. W. 4Pi MINFLUX arrangement maximizes spatio-temporal localization precision of fluorescence emitter. *Proceedings of the National Academy of Sciences* **121,** e2318870121 (Mar. 12, 2024).

101. Masullo, L. A. *et al.* An alternative to MINFLUX that enables nanometer resolution in a confocal microscope. *Light: Science & Applications* **11,** 199. ISSN: 2047-7538 (June 30, 2022).

102. Xu, X., Jia, S. & Xi, P. Raster-scanning Donut simplifies MINFLUX and provides alternative implement on other scanning-based microscopes. *Light: Science & Applications* **11,** 293. ISSN: 2047-7538 (Oct. 10, 2022).

103. Masullo, L. A. *et al.* Pulsed Interleaved MINFLUX. *Nano Letters* (Dec. 18, 2020).

104. Zähringer, J. *et al.* Combining pMINFLUX, graphene energy transfer and DNA-PAINT for nanometer precise 3D super-resolution microscopy. *Light: Science & Applications* **12,** 70. ISSN: 2047-7538 (Mar. 10, 2023).

105. Cole, F. *et al.* Super-resolved FRET and co-tracking in pMINFLUX. *Nature Photonics* **18,** 478–484. ISSN: 1749-4893 (May 2024).

106. Tan, X.-J. & Huang, Z. MINFLUX nanoscopy enhanced with high-order vortex beams. *Light: Science & Applications* **14,** 184. ISSN: 2047-7538 (May 6, 2025).

107. *Genoa Instruments* https://www.genoainstruments.com/prism.html#two (2025).

108. Gulinatti, A., Ceccarelli, F., Ghioni, M. & Rech, I. Custom silicon technology for SPAD-arrays with red-enhanced sensitivity and low timing jitter. *Optics Express* **29,** 4559–4581. ISSN: 1094-4087 (Feb. 1, 2021).

109. Kay, S. M. *Fundamentals of Statistical Signal Processing: Estimation theory* book. ISBN: 978-0-13-487840-9 (Prentice-Hall PTR, 2013).

110. Cheezum, M. K., Walker, W. F. & Guilford, W. H. Quantitative Comparison of Algorithms for Tracking Single Fluorescent Particles. *Biophysical Journal* **81,** 2378–2388. ISSN: 0006-3495 (Oct. 1, 2001).

111. Bucci, A. *et al.* 4D Single-particle tracking with asynchronous read-out single-photon avalanche diode array detector. *Nature Communications* **15,** 6188. ISSN: 2041-1723 (July 23, 2024).

112. Abraham, A. V., Ram, S., Chao, J., Ward, E. S. & Ober, R. J. Quantitative study of single molecule location estimation techniques. *Optics Express* **17,** 23352–23373. ISSN: 1094-4087 (Dec. 21, 2009).

113. Ober, R. J., Ram, S. & Ward, E. S. Localization Accuracy in Single-Molecule Microscopy. *Biophysical Journal* **86,** 1185–1200. ISSN: 0006-3495 (Feb. 1, 2004).

114. Van. Trees, H. L. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory* 2nd, 1–768. ISBN: 978-0471226970 (Wiley, New York, USA, 2001).

115. Holden, S. J., Uphoff, S. & Kapanidis, A. N. DAOSTORM: an algorithm for high-density super-resolution microscopy. *Nature Methods* **8,** 279–280. ISSN: 1548-7105 (Apr. 2011).

116. Ma, H., Xu, J. & Liu, Y. WindSTORM: Robust online image processing for high-throughput nanoscopy. *Science Advances* **5,** eaaw0683 (Apr. 26, 2019).

117. Goutzoulis, A. & Pape, D. R. *Design and Fabrication of Acousto-Optic Devices* 1–528 (Marcel Dekker, New York, USA, 1994).

118. Roy, P. Acousto-Optic Deflector: Principles, Working and Applications. *International Journal of Advanced Networking and Applications* **15,** 6126–6130. ISSN: 0975-0290 (2023).

119. Franz, D. *et al.* Characterization of a hybrid scanning system comprising acousto-optical deflectors and galvanometer scanners. *Applied Physics B* **128,** 55 (2022).

120. Masullo, L. A., Lopez, L. F. & Stefani, F. D. A common framework for single-molecule localization using sequential structured illumination. *Biophysical Reports* **2,** 100036. ISSN: 2667-0747 (Mar. 9, 2022).

121. Lee, S. H. *et al.* Using fixed fiduciary markers for stage drift correction. *Optics Express* **20,** 12177–12183. ISSN: 1094-4087 (May 21, 2012).

122. Mlodzianoski, M. J. *et al.* Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Optics Express* **19,** 15009–15019. ISSN: 1094-4087 (Aug. 1, 2011).

123. McGorty, R., Kamiyama, D. & Huang, B. Active microscope stabilization in three dimensions using image correlation. *Optical Nanoscopy* **2,** 3. ISSN: 2192-2853 (Apr. 18, 2013).

124. Han, R. *et al.* Drift correction for single-molecule imaging by molecular constraint field, a distance minimum metric. *BMC Biophysics* **8,** 1. ISSN: 2046-1682 (Jan. 13, 2015).

125. Wang, Y. *et al.* Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Optics Express* **22,** 15982–15991. ISSN: 1094-4087 (June 30, 2014).

126. Geisler, C. *et al.* Drift estimation for single marker switching based imaging schemes. *Optics Express* **20,** 7274–7289. ISSN: 1094-4087 (Mar. 26, 2012).

127. Silvestri, L. *et al.* Universal autofocus for quantitative volumetric microscopy of whole mouse brains. *Nature Methods* **18,** 953–958. ISSN: 1548-7105 (Aug. 2021).

128. Przybylski, A., Thiel, B., Keller-Findeisen, J., Stock, B. & Bates, M. Gpufit: An open-source toolkit for GPU-accelerated curve fitting. *Scientific Reports* **7,** 15722. ISSN: 2045-2322 (Nov. 16, 2017).

129. Slenders, E. & Vicidomini, G. ISM-FLUX: MINFLUX with an array detector. *Physical Review Research* **5,** 023033. ISSN: 2643-1564 (Apr. 17, 2023).

130. Slenders, E., Patil, ., Held, M. O., Zunino, A. & Vicidomini, G. Array detection enables large localization range for simple and robust MINFLUX. *Light: Science & Applications* **14,** 234. ISSN: 2047-7538 (July 3, 2025).

131. Sun, X. *et al.* NanoDriftGuard: Open-source isotropic ångström-scale active stabilization for super-resolution microscopy. *Optics and Lasers in Engineering* **190,** 108957. ISSN: 0143-8166 (2025).

132. Rahmani, A., Cox, T., Achary, A. T. A. & Ponjavic, A. Astigmatism-based active focus stabilisation with universal objective lens compatibility, extended operating range and nanometer precision. *Optics Express* **32,** 13331–13341. ISSN: 1094-4087 (Apr. 8, 2024).

133. Berg, B. v. d. *et al. A Modular Approach to Active Focus Stabilization for Fluorescence Microscopy* Jan. 6, 2023.

134. Lin, S., He, Y., Feng, D., Piliarik, M. & Chen, X.-W. Optical Fingerprint of Flat Substrate Surface and Marker-Free Lateral Displacement Detection with Angstrom-Level Precision. *Physical Review Letters* **129,** 213201 (Nov. 15, 2022).

135. Chen, S.-Y., Heintzmann, R. & Cremer, C. Sample drift estimation method based on speckle patterns formed by backscattered laser light. *Biomedical Optics Express* **10,** 6462–6475. ISSN: 2156-7085 (Dec. 1, 2019).

136. *The Nikon Perfect Focus System (PFS)* Nikon's MicroscopyU. www.microscopyu.com/tutorials/the-nikon-perfect-focus-system-pfs (2024).

137. *Leica DMI6000 B with Adaptive Focus Control Leica DMI6000 B with Adaptive Focus Control* https://www.leica-microsystems.com/products/light-microscopes/p/leica-dmi6000-with-adaptive-focus-control/ (2024).

138. *TruFocus | Z-Drift Compensation System | Olympus LS* https://www.olympus-lifescience.com/en/microscopes/inverted/ix83/trufocus/ (2024).

139. Coelho, S., Baek, J., Walsh, J., Gooding, J. J. & Gaus, K. 3D active stabilization for single-molecule imaging. *Nature Protocols* **16,** 497–515. ISSN: 1750-2799 (Jan. 2021).

140. Almada, P., Culley, S. & Henriques, R. PALM and STORM: Into Large Fields and High-Throughput Microscopy with sCMOS Detectors. *Methods* **88,** 109–121 (2015).

141. Klevanski, M. *et al.* Automated Highly Multiplexed Super-Resolution Imaging of Protein Nano-Architecture in Cells and Tissues. *Nature Communications* **11,** 1552 (2020).

142. Jain, P. K., Lee, K. S., El-Sayed, I. H. & El-Sayed, M. A. Calculated Absorption and Scattering Properties of Gold Nanoparticles of Different Size, Shape, and Composition: Applications in Biological Imaging and Biomedicine. *The Journal of Physical Chemistry B* **110,** 7238–7248. ISSN: 1520-6106 (Apr. 1, 2006).

143. Donato, M., Slenders, E., Zunino, A., Bega, L. & Vicidomini, G. BrightEyes-MCS: a control software for multichannel scanning microscopy. *Journal of Open Source Software* **9,** 7125. ISSN: 2475-9066 (Nov. 24, 2024).

144. Kashchuk, A. V. *et al.* Particle Localization Using Local Gradients and Its Application to Nanometer Stabilization of a Microscope. *ACS Nano* **17,** 1344–1354. ISSN: 1936-0851 (Jan. 24, 2023).

145. Dertinger, T. *et al.* Two-Focus Fluorescence Correlation Spectroscopy: A New Tool for Accurate and Absolute Diffusion Measurements. *ChemPhysChem* **8,** 433–443. ISSN: 1439-7641 (2007).

146. Rossetta, A. *et al.* The BrightEyes-TTM as an open-source time-tagging module for democratising single-photon microscopy. *Nature Communications* **13,** 7406. ISSN: 2041-1723 (Dec. 1, 2022).

147. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-Resolution Microscopy with DNA-PAINT. *Nature Protocols* **12,** 1198–1228 (2017).

148. Endesfelder, U., Malkusch, S., Fricke, F. & Heilemann, M. A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochemistry and Cell Biology* **141,** 629–638 (2014).

149. Slenders, E. *VicidominiLab/BrightEyes-FFS* https://www.github.com/VicidominiLab/BrightEyes-FFS (2025).

150. Reymond, L., Huser, T., Ruprecht, V. & Wieser, S. Modulation-enhanced localization microscopy. *Journal of Physics-Photonics* **2,** 041001 (2020).

151. Cnossen, J. *et al.* Localization microscopy at doubled precision with patterned illumination. en. *Nature Methods* **17,** 59–63. ISSN: 1548-7105 (Jan. 2020).

152. Gregor, I. & Enderlein, J. Image scanning microscopy. *Current Opinion in Chemical Biology. Chemical Genetics and Epigenetics • Molecular Imaging* **51,** 74–83. ISSN: 1367-5931 (Aug. 1, 2019).

153. Ellefsen, K. L., Dynes, J. L. & Parker, I. Spinning-Spot Shadowless TIRF Microscopy. *PLOS ONE* **10,** e0136055. ISSN: 1932-6203 (Aug. 26, 2015).

154. Stehr, F. *et al.* Flat-top TIRF illumination boosts DNA-PAINT imaging and quantification. *Nature Communications* **10** (Mar. 2019).

155. Zaza, C., Joseph, M. D., Dalby, O. P. L. & Baddeley, D. Super-resolution imaging in whole cells and tissues via DNA-PAINT on a spinning disk confocal with optical photon reassignment. *Nature Communications* **16** (Mar. 2025).

156. Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nature Methods* **8,** 1027–1036 (2011).

157. Provost, A., Rousset, C., *et al.* Innovative particle standards and long-lived imaging for 2D and 3D dSTORM. *Scientific Reports* **9,** 17967 (2019).

158. Ries, J. SMAP: a modular super-resolution microscopy analysis platform for SMLM data. *Nature Methods* **17,** 870–872 (2020).

159. Petrášek, Z. & Schwille, P. Precise Measurement of Diffusion Coefficients using Scanning Fluorescence Correlation Spectroscopy. *Biophysical Journal* **94,** 1437–1448 (2008).

160. Slenders, E. *et al.* Cooled SPAD array detector for low light-dose fluorescence laser scanning microscopy. *Biophysical Reports* **1,** 100025 (2021).

161. Lincoln, R. *et al.* A general design of caging-group-free photoactivatable fluorophores for live-cell nanoscopy. *Nature Chemistry* **14,** 1013–1020 (2022).

162. Hauke, L. *et al.* Metal-Induced Energy Transfer (MIET) for Live-Cell Imaging with Fluorescent Proteins. *ACS Nano* **17,** 8242–8251 (2023).

163. Fish, K. N. Total Internal Reflection Fluorescence (TIRF) Microscopy. *Current Protocols in Cytometry* **50,** 12.18.1–12.18.13. ISSN: 1934-9300 (2009).

164. Saini, D., Westbrook, L., Mehl, R., Dorey, M. & Ramm, J. *Optical fiber based despeckler for fluorescence microscopy and imaging* in *Optical Fibers and Sensors for Medical Diagnostics, Treatment and Environmental Applications XXIII* Optical Fibers and Sensors for Medical Diagnostics, Treatment and Environmental Applications XXIII. **12372** (SPIE, Mar. 6, 2023), 49–54.

165. Affannoukoué, K. *et al.* Super-resolved total internal reflection fluorescence microscopy using random illuminations. *Optica* **10,** 1009. ISSN: 2334-2536 (Aug. 20, 2023).

166. Speiser, A. *et al.* Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods* **18,** 1082–1090. ISSN: 1548-7091 (Sept. 2021).

Appendix A

Maximum likelihood estimator

Having defined the photon-counting observation model and the Cramér-Rao bound (CRB) as a benchmark for estimator performance (Section 3.2), we now specify the estimator used throughout this work. We adopt a maximum-likelihood formulation because it matches the Poisson statistics of photon detection and, in the high-photon limit, is efficient in the sense that its variance approaches the CRB. For image scanning microscopy (ISM)-FLUX, the likelihood is most conveniently expressed in terms of the molecule detection function (MDF), which encode the probability $p_{ij}(\mathbf{r}_E)$ that a detected photon is registered in detector element j while the excitation minimum is at scan position i.

Consider a system in which a fluorophore, i.e., a point-emitter, is sequentially illuminated with N_c illumination patterns, e.g. by moving a donut beam to different positions. The number of photons detected with a single-element detector can be assumed to be Poisson distributed with λ_i the expected number of photons under illumination i ($i \in \{1,2,...,N_c\}$). Thus, the probability of observing $\mathbf{n} = (n_1, n_2, ..., n_{N_c})$ photons is

$$P(\mathbf{n}|\{\lambda_i\}) = \prod_{i=1}^{N_c} \frac{e^{-\lambda_i} \lambda_i^{n_i}}{n_i!}. \tag{A.1}$$

However, this is the probability of having $\mathbf{n}$ photons. What we are interested in is the probability $P(\mathbf{n}|N)$, with $N = \sum_i n_i$. Using Bayes' theorem, we have

$$P(\mathbf{n}|N,\{\lambda_i\}) = \frac{P(\mathbf{n}|\{\lambda_i\})}{P(N|\{\lambda_i\})} = \frac{\prod_{i=1}^{N_c} \frac{e^{-\lambda_i}\lambda_i^{n_i}}{n_i!}}{\frac{e^{-\lambda_{tot}}\lambda_{tot}^{N}}{N!}} = \frac{N! \prod_{i=1}^{N_c} \frac{e^{-\lambda_i}\lambda_i^{n_i}}{n_i!}}{e^{-\lambda_{tot}}\lambda_{tot}^{N}} \tag{A.2}$$

with $\lambda_{tot} = \sum \lambda_i$.

We can write $\lambda_i = N p_i$, with N the total number of detected photons, p_i the probability that if a photon is detected, it is detected under illumination i, and $\sum_i p_i = 1$. Then,

$$P(\mathbf{n}|N,\{p_i\}) = \frac{N!\prod_{i=1}^{N_c}\left(\frac{e^{-Np_i}N^{n_i}p_i^{n_i}}{n_i!}\right)}{e^{-N}N^N} = \frac{N!}{\prod_{i=1}^{N_c}n_i!}\prod_{i=1}^{N_c}p_i^{n_i} \tag{A.3}$$

which is a multinomial distribution.

We now turn to the ISM-FLUX case with an array detector. For ISM-FLUX, we can define a probability p_{ij} that if a photon detection event happens, it happens under illumination i in detector-element j. We can write the dependence of $p_{ij}(\mathbf{r}_E)$ on the emitter position $\mathbf{r}_E$. Assume a (donut) illumination intensity profile $h(\mathbf{r}-\mathbf{r}_i)$ that moves to N_c different positions $\mathbf{r}_i$. Assume further that the expected number of emitted photons is proportional to the illumination intensity. Then, the fraction of signal photons emitted under illumination i is:

$$f_i(\mathbf{r}_E) = \frac{h(\mathbf{r}_E-\mathbf{r}_i)}{\sum_{p=1}^{N_c}h(\mathbf{r}_E-\mathbf{r}_p)} \tag{A.4}$$

The fluorescence emitted by the single emitter follows a spatial distribution $h_2(\mathbf{r}-\mathbf{r}_E)$ equal to the emission point spread function (PSF) centered at position $\mathbf{r}_E$. We have,

$$\int_{\mathbb{R}^2} h_2(\mathbf{r}-\mathbf{r}_E)\,dxdy = 1 \tag{A.5}$$

The probability $g_{ij}(\mathbf{r}_E)$ that a photon emitted under illumination i is detected by detector element j is

$$g_{ij}(\mathbf{r}_E) = \int_{\mathbb{R}^2} h_2(\mathbf{r}-\mathbf{r}_E)w_{ij}(\mathbf{r})\,dxdy \tag{A.6}$$

Here, $w(\mathbf{r})$ is a 2D window function describing the detector element. Note that, since the detector is placed in descanned mode, the detector *moves* when the excitation beam moves, hence the dependence on both the indices i and j:

$$w_{ij}(\mathbf{r}) = w(\mathbf{r}-\mathbf{r}_i-\mathbf{s}_j) \tag{A.7}$$

with $\mathbf{s}_j$ the position of detector element j within the array. Note that the detector has a finite number of elements and thus a finite size hence, not all emitted photons will be detected:

$$\begin{aligned}\sum_j g_{ij}(\mathbf{r}_E) &= \sum_j \int_{\mathbb{R}^2} h_2(\mathbf{r}-\mathbf{r}_E) w_{ij}(\mathbf{r})\, dx\, dy \\ &= \int_{\mathbb{R}^2} h_2(\mathbf{r}-\mathbf{r}_E) \sum_j w_{ij}(\mathbf{r})\, dx\, dy = \int_A h_2(\mathbf{r}-\mathbf{r}_E)\, dx\, dy < 1\end{aligned} \tag{A.8}$$

with A the overall area of all detector elements.

If a photon is observed, the probability it happened under illumination i in detector element j is:

$$p_{ij} = \frac{P(\text{detection event in i,j})}{P(\text{detection event})} = \frac{f_i(\mathbf{r}_E) g_{ij}(\mathbf{r}_E)}{\sum_{p=1}^{N_c} \sum_{q=1}^{N_d} f_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)} \tag{A.9}$$

Plugging in A.4 yields

$$p_{ij} = \frac{\frac{h(\mathbf{r}_E-\mathbf{r}_i)}{\sum_{p=1}^{N_c} h(\mathbf{r}_E-\mathbf{r}_p)} g_{ij}(\mathbf{r}_E)}{\sum_{p=1}^{N_c} \left(\frac{h(\mathbf{r}_E-\mathbf{r}_p)}{\sum_{q=1}^{N_c} h(\mathbf{r}_E-\mathbf{r}_q)} \sum_{q=1}^{N_d} g_{pq}(\mathbf{r}_E) \right)} = \frac{h(\mathbf{r}_E-\mathbf{r}_i) g_{ij}(\mathbf{r}_E)}{\sum_{p=1}^{N_c} \sum_{q=1}^{N_d} h(\mathbf{r}_E-\mathbf{r}_p) g_{pq}(\mathbf{r}_E)} \tag{A.10}$$

where, using Eq. A.6, the numerator can be written as

$$\begin{aligned}I_{ij}(\mathbf{r}_E) &= h(\mathbf{r}_E-\mathbf{r}_i) \int_{\mathbb{R}^2} h_2(\mathbf{r}-\mathbf{r}_E) w(\mathbf{r}-\mathbf{r}_i-\mathbf{s}_j)\, dx\, dy \\ &= h(\mathbf{r}_E-\mathbf{r}_i) \int_{\mathbb{R}^2} h_2(\mathbf{r}) w(\mathbf{r}-\mathbf{s}_j+\mathbf{r}_E-\mathbf{r}_i)\, dx\, dy\end{aligned} \tag{A.11}$$

For $\mathbf{r}_i = 0$, we find the equation for the signal in detector element j for a point-source at position $\mathbf{r}_E$ and detector element at position $\mathbf{s}_j$. Experimentally, one can measure $I_{0j}(\mathbf{r}_E)$ by moving a point source in 2D with a piezoelectric stage and measuring the resulting signal for each position and in each detector element or, equivalently, by scanning the point source with the galvanometric scanners.

For $\mathbf{r}_i \neq 0$, we have $I_{ij}(\mathbf{r}_E) = I_{0j}(\mathbf{r}_E - \mathbf{r}_i)$. Thus, when the targeted coordinate pattern (TCP) is known, one can simply shift the resulting point-source images to the right positions to get the full set of $I_{ij}(\mathbf{r}_E)$. One can find the denominator in Eq. A.10 by summing all images. For the ISM-FLUX setup in this work, we have $N_d = 25$ detector elements and $N_c = 32$ positions.

In the presence of background, Eq. A.10 becomes:

$$p_{ij} = \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E) + b_{ij}}{\sum_{p=1}^{N_c} \sum_{q=1}^{N_d} \left(h(\mathbf{r}_E - \mathbf{r}_p) g_{pq}(\mathbf{r}_E) + b_{pq} \right)} \tag{A.12}$$

with b_{ij} is the expected background counts per channel j under illumination i. With a SPAD array detector, the main contributions to the background counts are the out-of-focus fluorescence and the dark counts. We assume that both contributions are equal for each detector element and each illumination, thus $b_{ij} = b$:

$$p_{ij} = \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E) + b}{N_c N_d b + \sum_{p=1}^{N_c} \sum_{q=1}^{N_d} h(\mathbf{r}_E - \mathbf{r}_p) g_{pq}(\mathbf{r}_E)} \tag{A.13}$$

In the remainder we use a uniform-background approximation for clarity; in the experimental analysis we exclude hot pixels and incorporate measured background rates per pixel when needed.

Define the signal-to-background ratio as:

$$\mathrm{SBR}(\mathbf{r}_E) = \frac{\sum_{p=1}^{N_c} \sum_{q=1}^{N_d} h(\mathbf{r}_E - \mathbf{r}_p) g_{pq}(\mathbf{r}_E)}{N_c N_d b} = \frac{\sum_p \sum_q h_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)}{N_c N_d b} \tag{A.14}$$

Where we introduced a shorter notation for the double sum. Then,

$$\begin{aligned} p_{ij}(\mathbf{r}_E) &= \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E) + b}{N_c N_d b + N_c N_d b\, \mathrm{SBR}(\mathbf{r}_E)} \\ &= \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E) \sum_p \sum_q h_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)}{N_c N_d b (\mathrm{SBR}(\mathbf{r}_E) + 1) \sum_p \sum_q h_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)} + \frac{b}{N_c N_d b (\mathrm{SBR}(\mathbf{r}_E) + 1)} \\ &= \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E) \mathrm{SBR}(\mathbf{r}_E)}{(\mathrm{SBR}(\mathbf{r}_E) + 1) \sum_p \sum_q h_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)} + \frac{b}{N_c N_d b (\mathrm{SBR}(\mathbf{r}_E) + 1)} \\ &= \frac{\mathrm{SBR}(\mathbf{r}_E)}{\mathrm{SBR}(\mathbf{r}_E) + 1} \frac{h(\mathbf{r}_E - \mathbf{r}_i) g_{ij}(\mathbf{r}_E)}{\sum_p \sum_q h_p(\mathbf{r}_E) g_{pq}(\mathbf{r}_E)} + \frac{1}{\mathrm{SBR}(\mathbf{r}_E) + 1} \frac{1}{N_c N_d} \end{aligned} \tag{A.15}$$

Assuming no prior information, we can express a likelihood function $\mathscr{L}$ as

$$\mathscr{L}(\mathbf{r}_E | \{n_{ij}\}) = \frac{N!}{\prod_{i=1}^{N_c} \prod_{j=1}^{N_d} n_{ij}!} \prod_{i=1}^{N_c} \prod_{j=1}^{N_d} p_{ij}(\mathbf{r}_E)^{n_{ij}} \tag{A.16}$$

with p_{ij} from Eq. A.15.

The log-likelihood function, dropping constant terms, is

$$\ell(\mathbf{r}_E|\{n_{ij}\}) = \sum_{i=1}^{N_c}\sum_{j=1}^{N_d} n_{ij} \ln p_{ij}(\mathbf{r}_E) \tag{A.17}$$

The maximum likelihood estimation for the emitter position is

$$\mathbf{r}_E^{\mathrm{MLE}} = \arg\max(\ell(\mathbf{r}_E|\{n_{ij}\})) \tag{A.18}$$

A.1 Bias of Maximum likelihood estimator in ISM-FLUX

The maximum likelihood estimator (MLE) is not inherently unbiased. At low photon counts, the expected estimate can deviate from the ground truth.We quantified MLE bias by simulation across emitter positions and photon counts, comparing MLE estimates to ground truth. We consecutively positioned the emitter at various points along a line passing through the TCP center. We used the simulated MDFs, Figure 3.10, to draw photon counts from a multinomial distribution with relative probabilities equal to the MDF values at the emitter position. We assumed 0 background. Then, we used the MLE to retrieve the emitter position and stored the difference in (x,y) between the ground truth and the retrieved position. We repeated this simulation 1000$\times$ for each emitter position and each photon count number.

The result in Fig. A.1, shows that the MLE is always unbiased in the TCP center and on the TCP circle, regardless of the number of photon counts. For other positions within the TCP, as well as outside the TCP, we find a position- and photon-count-dependent bias. For high photon counts, i.e. $N > 1000$, the bias is far below 1 nm in both directions for all positions within the detector field-of-view (FoV) and the MLE can be considered unbiased. For lower photon counts, the bias increases in both directions but remains below 10 nm for the whole FoV for $N \geq 30$. Note that for most positions on the horizontal line, the bias in the x (radial) direction is higher than in the y (tangential) direction.

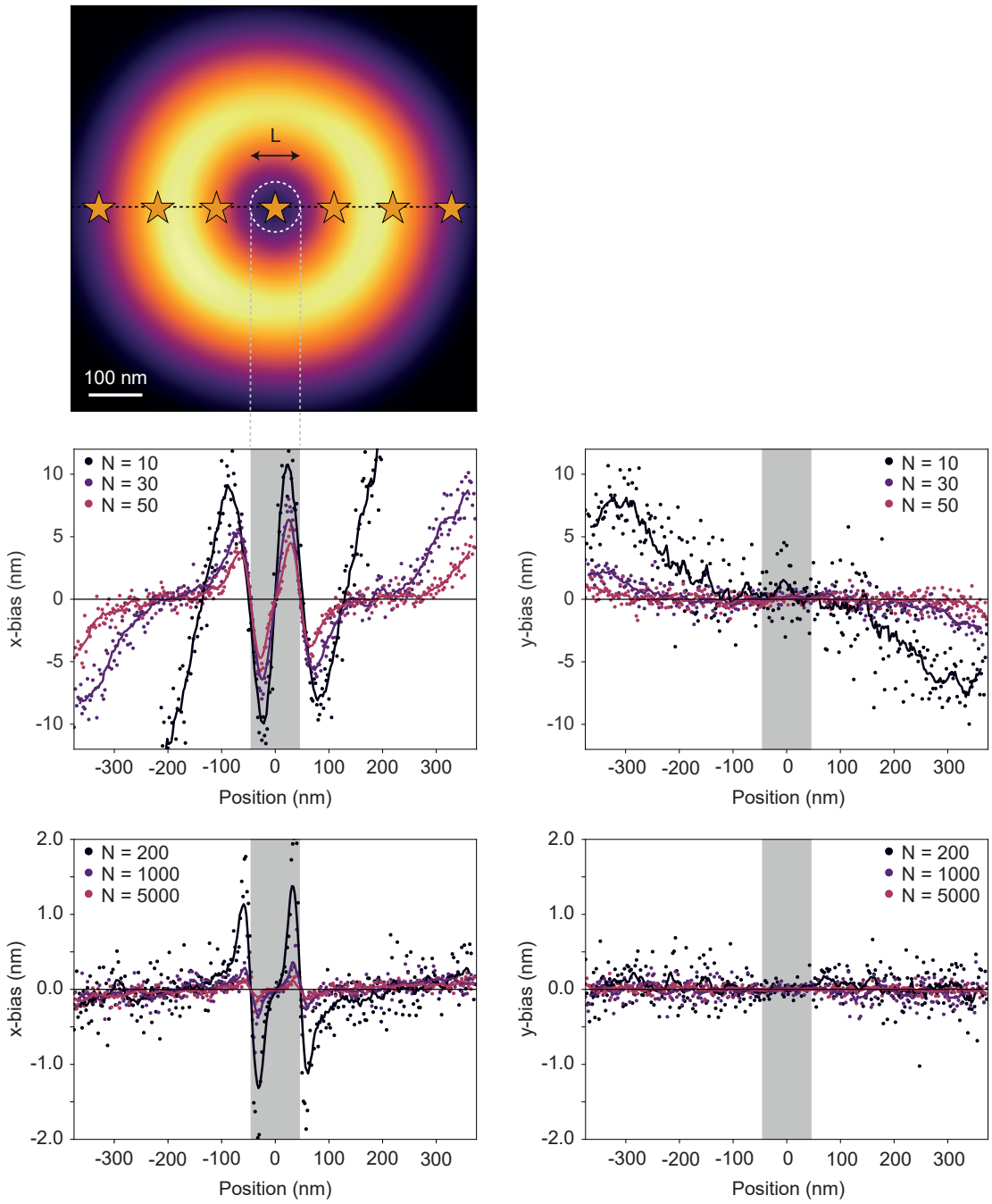

Figure A.1 Bias of the MLE in the x and y directions for different photon counts (N) and various emitter positions. $L = 90\,\text{nm}$. The emitter is moved to 300 positions in 2.5 nm steps along a horizontal line passing through the TCP center (indicated by the dotted black line in the intensity plot). For each position, the bias in both directions is plotted. The scatter plots display the raw data, while the line plots show the corresponding moving averages with a window size of 10.

Appendix B

Sample stabilization axial localization precision

We approximate the point spread function (PSF) shape as a 2D Gaussian function with beam waists $\sigma_x(z)$, $\sigma_y(z)$ in the x,y direction:

$$\sigma_x(z) = \sigma_0(1 - \alpha z) \tag{B.1}$$

$$\sigma_y(z) = \sigma_0(1 + \alpha z), \tag{B.2}$$

with σ_0 the PSF width in the focal plane ($z = 0$) and α a parameter with units m^{-1} that describes the strength of the astigmatism.

Without noise, the image of a single particle centered around the origin is:

$$I(x,y) \propto \exp\left(-\frac{x^2}{\sigma_x^2}\right) \exp\left(-\frac{y^2}{\sigma_y^2}\right). \tag{B.3}$$

We check for the asymmetry in the PSF by summing the intensities $I(x,y)$ along the x- (for $y = 0$) and y- (for $x = 0$) direction. For simplicity, we take the integral from $-\infty$ to ∞:

$$\int_{-\infty}^{\infty} I(x, y = 0)dx \propto \sqrt{\pi}\sigma_0(1 - \alpha z) \tag{B.4}$$

$$\int_{-\infty}^{\infty} I(x = 0, y)dx \propto \sqrt{\pi}\sigma_0(1 + \alpha z) \tag{B.5}$$

We estimate the z-position of the particle by comparing the total number of photons collected along the x-axis $(x, y = 0)$ and the y-axis $(x = 0, y)$, *i.e.,* we have two photon count numbers, n_x and n_y, describing the total number of photons collected along each axis. We assume that n_x and n_y are Poisson distributed with the expected number of photons λ_x and λ_y proportional to p_x and p_y, respectively.

$$p_x = \frac{\sqrt{\pi}\sigma_0(1-\alpha z)}{\sqrt{\pi}\sigma_0(1-\alpha z)+\sqrt{\pi}\sigma_0(1+\alpha z)} = \frac{1}{2} - \frac{\alpha z}{2} \tag{B.6}$$

and, similarly,

$$p_y = \frac{1}{2} + \frac{\alpha z}{2} \tag{B.7}$$

Without prior information, the likelihood for observing n_x, n_y photons, given a total of $N = n_x + n_y$ photons, is given by a multinomial distribution with event probabilities p_x and p_y.

$$\mathscr{L} = \frac{N!}{n_x! n_y!} p_x^{n_x} p_y^{n_y} \tag{B.8}$$

Then,

$$\frac{d^2 \log \mathscr{L}}{dz^2} = \frac{-\alpha^2 n_x}{(1-\alpha z)^2} - \frac{\alpha^2 n_y}{(1+\alpha z)^2} \tag{B.9}$$

In the focal plane (z = 0), we find

$$\frac{d^2 \log \mathscr{L}}{dz^2} = -\alpha^2 n_x - \alpha^2 n_y = -\alpha^2 N \tag{B.10}$$

A lower bound on the standard deviation of any unbiased estimator, *i.e.,* the Cramér-Rao bound (CRB), is given by square root of the reciprocal of the Fisher information (with E the expected value):

$$\sigma_{\mathrm{CRLB}} = \sqrt{-\mathrm{E}\left[\frac{d^2}{dz^2}\log \mathscr{L}\right]} = \frac{1}{\sqrt{\alpha^2 N}} = \frac{1}{\alpha\sqrt{N}} \tag{B.11}$$

In order to estimate α for our system, we define the PSF anisotropy factor F

$$F = \frac{\sigma_y}{\sigma_x} = \frac{1+\alpha z}{1-\alpha z} \approx 1 + 2\alpha z \tag{B.12}$$

Experimentally, we use the PSF asymmetry A defined using the horizontal and vertical waists of the autocorrelation function (see Section 4.5.1). However, since the autocorrelation function broadens a 2D Gaussian function by the same factor of $\sqrt{2}$ in the x- and y-direction, the asymmetry ratio remains unchanged. Thus, A is a good approximation of the ratio σ_y/σ_x. From Figure 4.5 (c), we can estimate α to be about $\alpha \approx 5 \times 10^{-4}\,\mathrm{m}^{-1}$. Thus, at least a million photons per frame are needed for all particles combined to reach a localization uncertainty below 2 nm.