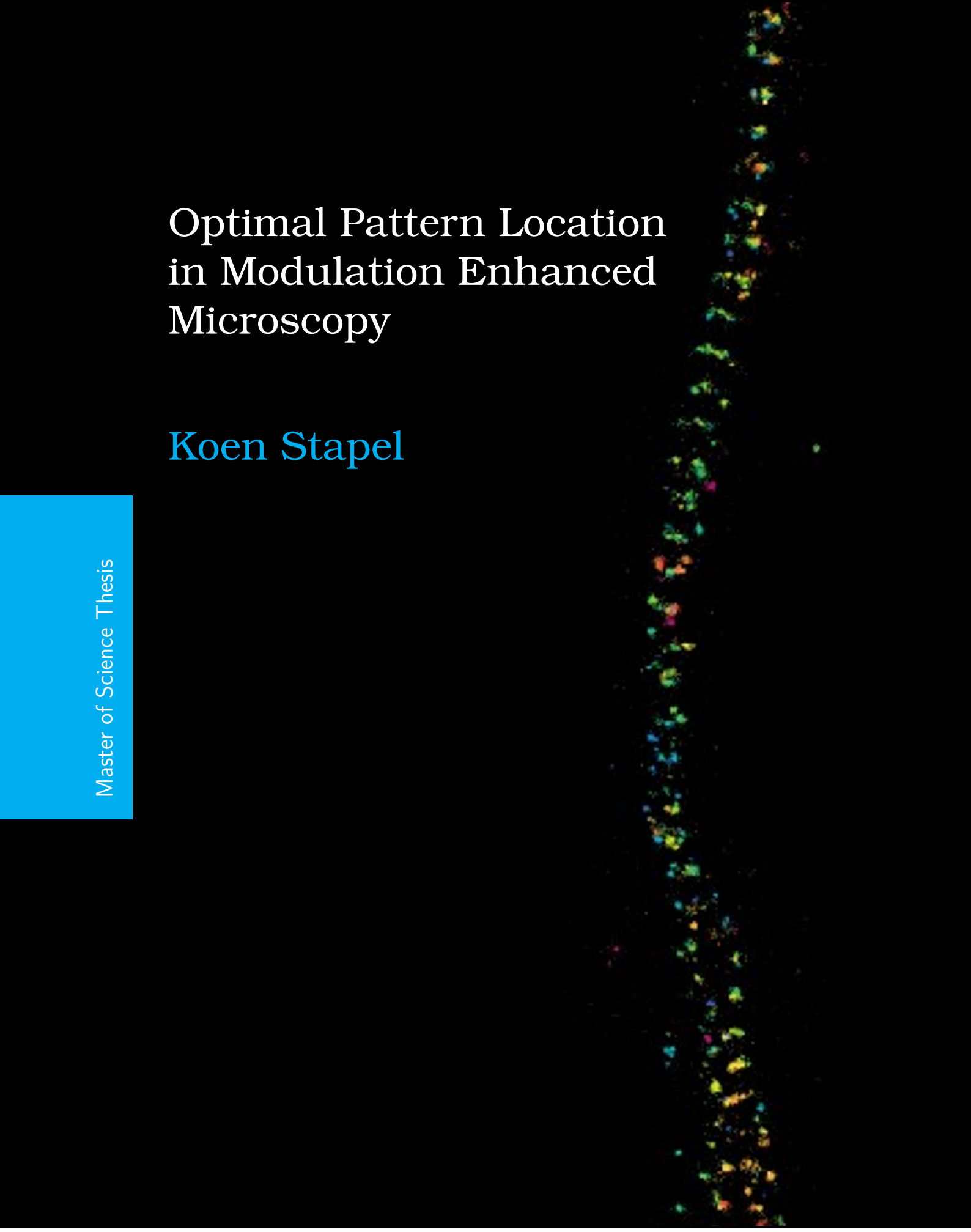
Optimal Pattern Location in Modulation Enhanced Microscopy

Koen Stapel

Master of Science Thesis



TU Delft Delft University of Technology

Delft Center for Systems and Control

Optimal Pattern Location in Modulation Enhanced Microscopy

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Koen Stapel

August 12, 2025

Faculty of Mechanical Engineering (ME) · Delft University of Technology

Cover image adapted from source: [1], licensed under CC BY-SA 4.0.

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical Engineering (ME) for acceptance a thesis entitled

Optimal Pattern Location in Modulation Enhanced Microscopy

by

Koen Stapel

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: August 12, 2025

Supervisor(s):

dr.ir. C.S. Smith

ir. D. Kalisvaart

Reader(s):

dr. M Kok

Abstract

Super-resolution fluorescence microscopy techniques enable researchers to observe biological processes at the nanometer scale, surpassing the diffraction limit of light. Among these, methods that combine illumination patterns with single-molecule localization offer high spatial resolution while maintaining photon efficiency. A leading example is MINFLUX microscopy, which achieves nanometer-scale localization by positioning a doughnut-shaped excitation beam, with a central intensity minimum, close to the emitter. This configuration allows for highly precise triangulation of the emitter's position using only a few detected photons. However, theoretical models often assume perfect modulation with a zero-intensity center, whereas in practical implementations, optical aberrations and alignment errors introduce residual intensity at the beam's center, degrading the localization precision.

To systematically investigate how deviations from ideal modulation affect localization performance, we introduce a modulation contrast parameter $m \in (0, 1]$, where $m = 1$ represents perfect modulation and values below one reflect increasing residual intensity at the excitation minimum. We extend the Cramér–Rao Lower Bound (CRLB) framework to incorporate this parameter, allowing us to quantify how imperfect modulation reduces the Fisher information and consequently increases the theoretical lower bound on localization precision.

We show that decreasing modulation contrast not only worsens achievable precision but also shifts the optimal illumination spacing L, challenging previously established scaling laws. We derive and validate a predictive formula for this optimal spacing, $L_{\text{opt}} \approx 1.30\,\sigma_{\text{illum}}\sqrt{1-m}$, which is experimentally accessible and maintains high precision. This relationship remains valid as long as the emitter lies within 40% of the pattern diameter. We further extend the framework to account for uncertainty in emitter position by incorporating prior information, showing that the optimal spacing increases with prior uncertainty σ_{prior}.

In addition, we evaluate iterative MINFLUX and show that under non-ideal conditions $m < 1$, the standard multi-step narrowing strategy becomes suboptimal. Instead, performing repeated measurements at the optimized spacing L_{opt} achieves significantly better precision over 50% improvement at $m = 0.95$.

Our results could provide a foundation for MINFLUX single-particle tracking, where selecting L_{opt} based on system contrast and prior uncertainty can maximize localization precision in

each frame while minimizing photon budget. Finally, experimental validation under non-ideal modulation conditions will be crucial to confirm the practical relevance of these predictions and to further refine theoretical models.

Table of Contents

Preface

When I started my thesis project, I thought: I've got this. Just another challenge to tackle head-on, like I always do. No big deal. But reality turned out to be a bit different. After a year, I ended up restarting with a completely new project, struggled endlessly with stubborn coding bugs, and found myself trying to make sense of papers that seemed impossible to understand. Throughout the project, I took on too many things at once. Saying "no" has never really been my strong suit, and that didn't exactly make writing a thesis any easier.

Luckily, I wasn't in it alone. ir. Sylke van der Kleij, my awesome girlfriend, stood by me through it all. She supported me every step of the way, lifted me up when I was down, and reminded me what I was capable of when I lost sight of it myself. I honestly don't know how I would've made it through without her.

Now, more than two years later, I'm finally wrapping up my master thesis. I'm proud of the final result, for sure. But even more proud of how I got here. It wasn't easy. There were some serious mental hurdles. But I pushed through, and I learned a lot about myself along the way.

Delft, University of Technology — Koen Stapel
August 12, 2025

Acknowledgements

I would like to express my gratitude to dr.ir. C.S. Smith for his role as supervisor. His support in the formal aspects of the project and his endorsement throughout the process were much appreciated.

A special thanks goes to ir. Dylan Kalisvaart for his invaluable daily supervision throughout this project. His guidance was instrumental not only in addressing technical and content-related challenges, but also in supporting my personal development during this process. I am particularly grateful for the safe and constructive environment he fostered, which made it easier to navigate difficult moments.

I also want to thank Mirthe Bronkhorst and Xinyue Zhang for their collaboration in developing the code base that underpinned this research. Their contributions, thoughtful mathematical discussions, and engaging spirit made the work both productive and enjoyable. Besides that I would like to thank all PhD candidates and students of the Quantitative Nanoscopy group for their feedback, interesting ideas, and the stimulating environment they helped create.

Finally, I want to give a special shout-out to the master's students up in the landscape in the Systems and Control tower. We have all been through the ups and downs together, but I especially want to thank my study buddies from Dodeka, the student athletics association. You were there during the toughest mental stretches, keeping things light and helping me push through. I am really grateful we went through this journey side by side.

Delft University of Technology
August 12, 2025

Koen Stapel

"Je hebt vaker spijt van wat je niet hebt gedaan, dan van wat je wel hebt gedaan."
— *Persoonlijke overtuiging*

"You more often regret the things you didn't do, than the things you did."
— *Personal belief*

Chapter 1

Introduction

Understanding the structure and function of biomolecules within living cells is essential for advancing our knowledge of health, disease, and drug development [2, 3]. High-resolution imaging of cellular components enables researchers to study dynamic biological processes and develop targeted therapies for various medical conditions [4, 5]. However, most biological structures are inherently transparent and exhibit little contrast under conventional light microscopy, making it challenging to visualize them with sufficient specificity or spatial resolution.

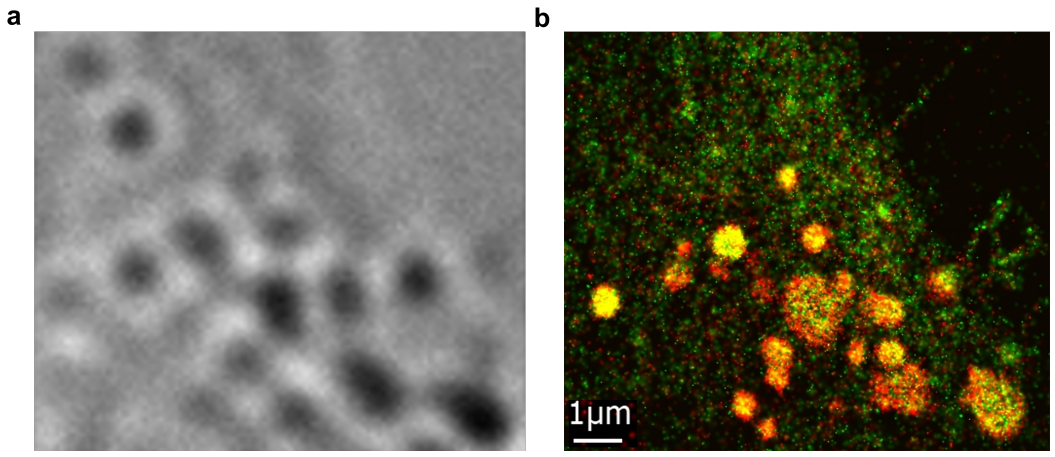


Figure 1-1: Comparison of imaging the same biological sample with conventional light microscopy **(a)** and fluorescence microscopy **(b)**. Adjusted from source: [6], under the CC BY-SA 3.0 license.

To address these limitations, fluorescence microscopy has become an indispensable tool in modern biology and medicine. By labeling specific molecules with fluorescent probes, it enables selective visualization of individual cellular components against an otherwise noisy background [7, 8]. As illustrated in Figure 1-1, this results in significantly enhanced contrast and molecular specificity. This allows researchers to identify, track, and quantify biomolecular processes with high spatial and temporal resolution in both fixed and living cells.

This chapter introduces the foundations of fluorescence microscopy and the motivation for this thesis. Section 1-1 explains the physical mechanisms of fluorescence, followed by Section 1-2, which describes the key components of a fluorescence microscope. Section 1-3 covers the point spread function and classical resolution limits. In Section 1-4, we provide an overview of super-resolution techniques that overcome the diffraction limit, and Section 1-4-1 focuses on single-molecule localization methods. Finally, Section 1-4-2 discusses modulation-enhanced strategies such as MINFLUX, and Section 1-5 presents the motivation and research objectives of this thesis.

1-1 Photon Emission in Fluorescence Microscopy

Fluorescence microscopy relies on the ability of fluorescent molecules, or fluorophores, to absorb light at one wavelength and emit it at a longer wavelength [7, 8]. This process is illustrated in Figure 1-2.

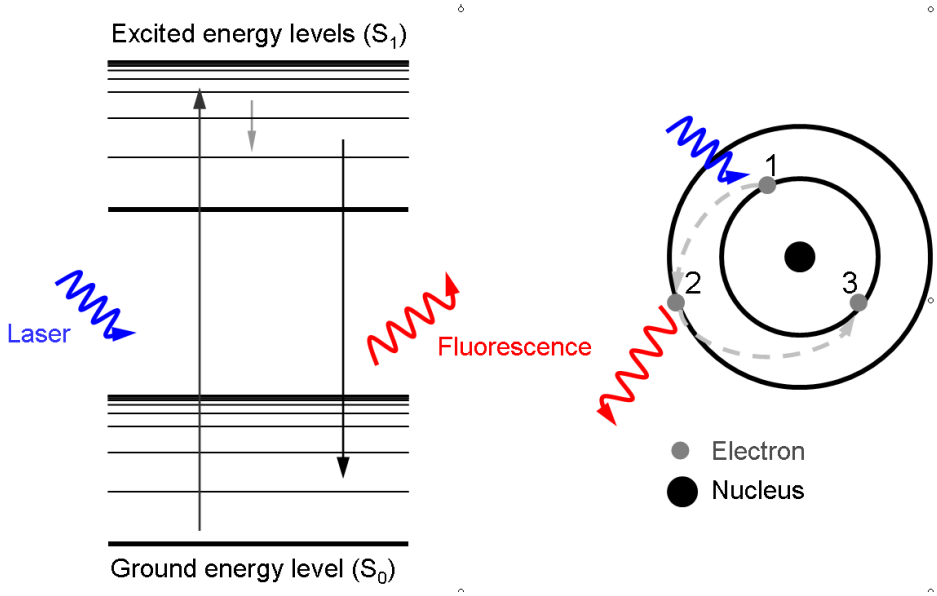


Figure 1-2: Schematic representation of the fluorescence process. The left panel shows a Jablonski diagram depicting electronic transitions. A molecule in its ground state (S_0) absorbs a photon from a laser (blue arrow), exciting it to a higher vibrational level of the first excited singlet state (S_1). The molecule rapidly relaxes non-radiatively (gray arrow) to the lowest vibrational level of S_1, followed by radiative relaxation (black arrow) back to a vibrational level of the ground state, emitting a photon of lower energy (red arrow). The right panel presents an orbital diagram, where an electron (gray dot) absorbs energy and moves from a lower orbital (1) to a higher orbital (2). It then returns to a lower orbital (3), releasing a photon. This visualizes fluorescence at the atomic level. Adjusted from source: [9], licensed under CC BY-SA 4.0.

As shown in the left panel of Figure 1-2, fluorescence begins when a fluorophore absorbs a photon of light, typically from a laser or another intense light source. The photon's energy must match the energy difference between the ground electronic state S_0 and one of the higher excited electronic states, such as S_1 [7, 8]. The energy E of the absorbed photon is given by:

$$E_{\text{absorption}} = \frac{hc}{\lambda_{\text{absorption}}} \tag{1-1}$$

where h is Planck's constant, c is the speed of light and $\lambda_{\text{absorption}}$ is the wavelength of the absorbed photon. Upon absorbing the photon, the fluorophore's electron transitions from the ground state S_0 to an excited state, typically S_1.

Once in the excited state S_1, the electron undergoes rapid vibrational relaxation, as indicated by the downward gray arrow. During this process, energy is dissipated through molecular vibrations without photon emission, reducing the electron's energy while remaining in the excited electronic state [7, 8]. This vibrational relaxation typically occurs within picoseconds, bringing the electron to the lowest vibrational level of S_1.

Following vibrational relaxation, the electron returns to the ground state S_0 by emitting a photon, a process known as fluorescence emission [7, 8]. This is represented by the black downward arrow in the Jablonski diagram, accompanied by the red squiggly arrow showing the emitted fluorescence. Because of the energy loss during the non-radiative relaxation, the emitted photon has lower energy (and thus a longer wavelength) than the absorbed photon.

The energy of the emitted photon, $\lambda_{\text{emission}}$, is determined by the energy difference between the lowest vibrational level of the excited state S_1 and a higher vibrational level of the ground state S_0. This relationship is expressed as:

$$E_{\text{emission}} = \frac{hc}{\lambda_{\text{emission}}} = E(S_1) - E(S_0) \tag{1-2}$$

where $\lambda_{\text{emission}}$ is the wavelength of the emitted photon, $E(S_1)$ represents the energy of the lowest vibrational level of the first excited electronic singlet state and $E(S_0)$ denotes the energy of the vibrational level in the ground electronic state to which the electron relaxes. The energy difference between these two levels determines the frequency of the fluorescence emission.

The right panel of Figure 1-2 presents a view using atomic orbitals. An electron initially occupies a lower energy orbital (position 1). Upon absorbing a photon (blue arrow), it is excited to a higher orbital (position 2). After a brief period, the electron returns to a lower-energy orbital (position 3), releasing a photon in the process (red arrow).

This difference in wavelength is the Stokes shift [7, 8]. The shift in wavelength can be expressed mathematically as:

$$\Delta\lambda_{\text{Stokes}} = \lambda_{\text{emission}} - \lambda_{\text{absorption}} \tag{1-3}$$

Finally, the cyclical nature of fluorescence is also visible in Figure 1-2, by showing that once the electron returns to the ground state, the molecule is ready to absorb another photon and repeat the process [7, 8].

1-2 Optical Setup of a Fluorescence Microscope

The optical setup of a fluorescence microscope consists of several components that work together to selectively excite and detect fluorescence from the specimen. Figure 1-3 illustrates the entire optical pathway.

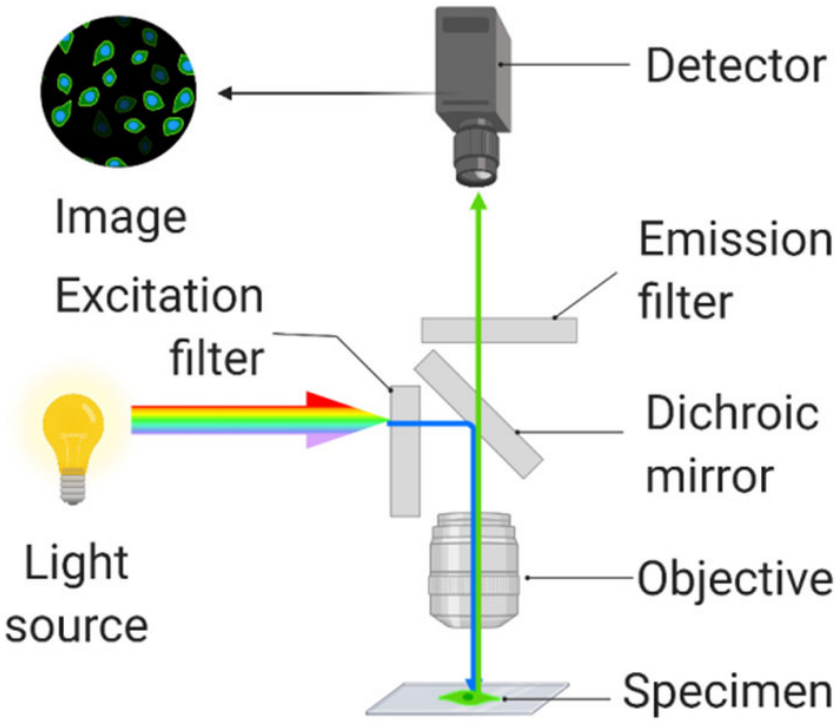


Figure 1-3: The fluorescence microscope optical setup. The diagram illustrates the components involved in fluorescence imaging: **Light source:** A high-intensity light source such as a mercury arc lamp, LED, or laser generates a broad spectrum of light. This light is used to excite the fluorophores in the specimen. **Excitation filter:** The excitation filter selects a narrowband of wavelengths that correspond to the excitation spectrum of the fluorophore, ensuring only the desired excitation light reaches the specimen. **Dichroic mirror:** The dichroic mirror reflects the excitation light toward the objective lens while transmitting the longer-wavelength emitted fluorescence back toward the detector. This wavelength-selective reflection is crucial for separating the excitation and emission paths. **Objective lens:** The objective lens focuses the excitation light onto a precise region of the specimen and collects the emitted fluorescence. The lens also plays a critical role in determining the resolution of the microscope. **Specimen:** The fluorophores within the specimen absorb the excitation light and re-emit light at a longer wavelength, a process known as fluorescence. **Emission filter:** The emission filter allows only the fluorescence emission to pass through, blocking any remaining excitation light. This step ensures that the final detected signal consists only of the emitted fluorescence. **Detector:** The fluorescence signal is captured by a detector, typically a camera or a PMT, which converts the light signal into a digital image. The detector's sensitivity and resolution are important for obtaining high-quality fluorescence images. This optical pathway ensures that the microscope selectively images the fluorescent light emitted by the sample while minimizing the interference from the excitation light. Source: [10], licensed under the CC BY 4.0 license.

The process begins with the light source, which provides the excitation light required to activate the fluorophores in the specimen. A common light source is a laser, capable of producing a broad spectrum of light [7, 8]. As shown in Figure 1-3, the light emitted by the source is directed toward the excitation filter.

The excitation filter selects a specific band of wavelengths from the light source. This ensures that only the wavelengths matching the excitation spectrum of the fluorophores reach the specimen. The filter blocks all other wavelengths, preventing unwanted excitation or background noise. In Figure 1-3, this step ensures precise excitation of the targeted fluorophores.

Next, the filtered light passes through the dichroic mirror. The dichroic mirror is a wavelength-selective component that reflects the excitation light toward the objective lens while allowing longer-wavelength emitted fluorescence to pass through. This separation of excitation and emission light ensures that only the desired fluorescence signal is detected [7, 8]. As illustrated in Figure 1-3, the dichroic mirror is responsible for the directing of the light path.

The reflected excitation light then enters the objective lens. The objective lens focuses the excitation light onto a precise region of the specimen. It also collects the emitted fluorescence from the specimen and directs it back through the optical pathway.

The specimen, located at the focal point of the objective lens, contains the fluorophores that are excited by the incoming light. Upon excitation, these fluorophores emit light at a longer wavelength. Figure 1-3 shows that the specimen is both the source of fluorescence and the focal point of excitation.

The emitted fluorescence then passes through the dichroic mirror and is directed toward the emission filter. The emission filter allows only the emitted fluorescence to pass through while blocking any remaining excitation light. This step ensures that the detected signal consists solely of the emitted fluorescence.

Finally, the filtered fluorescence reaches the detector. The detector, which is typically a camera or a PMT (photomultiplier tube), converts the light signal into a digital image [7, 8]. In Figure 1-3, the detector is the final component of the optical pathway, where the fluorescence signal is recorded.

1-3 Point Spread Function

The psf (point spread function) describes how a point source of light is spread by an optical system. It represents the system's response to a point source and determines the resolution limit of the system [7, 8]. The spreading of the point source occurs due to diffraction and optical aberrations. The primary reason for the spreading of the psf is diffraction, which is caused by the finite aperture of the optical system. The size of the aperture and the wavelength of the light affect the extent of the spreading.

The image of an object in an optical system is determined by the psf through a convolution process. In the Fourier domain, this relationship is expressed as:

$$I(f_x, f_y) = O(f_x, f_y)P(f_x, f_y) \tag{1-4}$$

Here, $I(f_x, f_y)$ is the image, $O(f_x, f_y)$ is the object, and $P(f_x, f_y)$ is the psf, all in the Fourier domain. The variables f_x and f_y represent spatial frequency coordinates, corresponding to the rate of intensity variation in the x and y directions. High spatial frequencies correspond to fine details, while low spatial frequencies represent smooth variations in intensity.

In the spatial domain, the same relationship is represented as:

$$I(x, y) = O(x, y) * P(x, y) \tag{1-5}$$

In this equation, $I(x, y)$ is the image, $O(x, y)$ is the object, $P(x, y)$ is the psf, and $*$ denotes the convolution operation. The variables x and y are spatial coordinates defining the position of points in the image plane. The convolution process determines how the psf spreads or blurs the intensity of the object in the resulting image, affecting resolution and quality.

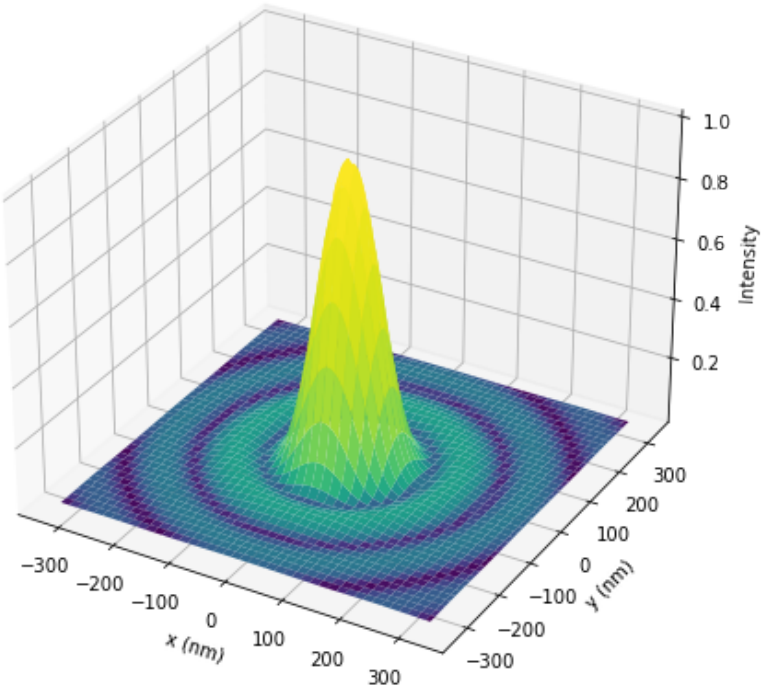


Figure 1-4: Visualization of the psf of a diffraction-limited optical system. The PSF represents the intensity distribution in the image plane resulting from a single point emitter, modeled here as an Airy pattern. The coordinates are scaled in nanometers based on a realistic Airy radius calculated for $\lambda = 650\,\text{nm}$ and numerical aperture NA = 1.4.

The Airy disk describes the smallest focused spot of light that a perfect lens with a circular aperture can produce [7, 8]. It represents the diffraction-limited psf in optical systems. As shown in Figure 1-4, the intensity is highest at the center and gradually diminishes through a series of concentric rings due to diffraction.

1-3-1 Diffraction Limit and Resolution

The diffraction limit represents the fundamental resolution limit of an optical system caused by the wave nature of light [7, 8]. It defines the smallest distance between two points that can still be resolved as separate entities. When light passes through an aperture, such as a lens or microscope objective, it diffracts, forming overlapping diffraction patterns that limit the system's resolving power.

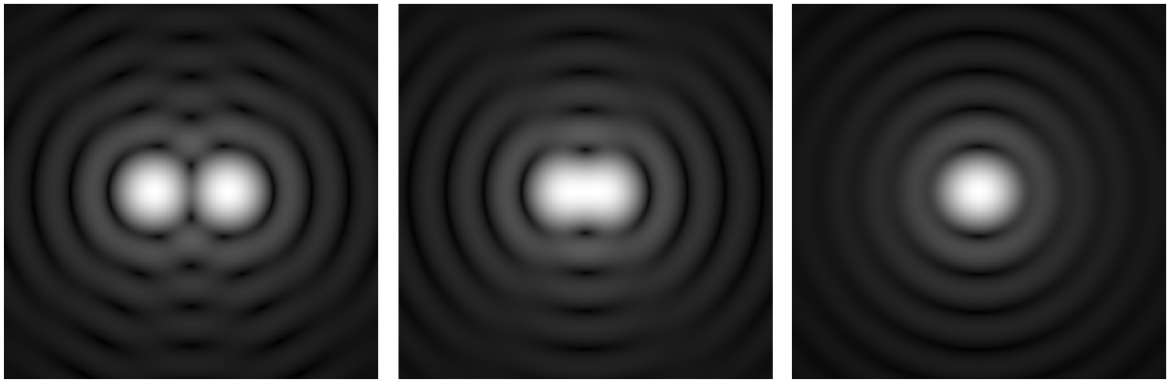


Figure 1-5: Illustration of the diffraction limit for resolving two closely spaced point sources. The left panel shows two points that are clearly resolvable, with the central maximum of one diffraction pattern separated from the other. The middle panel demonstrates the diffraction limit, where the central maximum of one overlaps with the first minimum of the other, making the points just resolvable. The right panel shows the points too close together to be resolved, with their diffraction patterns merging into a single peak.

The diffraction limit depends on two main factors: the wavelength of light (λ) and the numerical aperture (NA) of the imaging system. The wavelength determines how much the light spreads after diffraction, with shorter wavelengths providing better resolution. The numerical aperture (NA) is a measure of the light-gathering ability of the optical system and is calculated as:

$$\mathrm{NA} = n \sin \theta \tag{1-6}$$

Here, n is the refractive index of the medium between the lens and the sample, and θ is the half-angle of the maximum cone of light that enters the lens. A higher NA results in a smaller diffraction limit, allowing for better resolution. The diffraction limit can be defined by either the Rayleigh or Abbe limit.

Rayleigh Limit

The Rayleigh limit defines the minimum distance at which two point sources can be visually distinguished, given by:

$$d_R = 1.22 \frac{\lambda}{\mathrm{NA}} \tag{1-7}$$

This criterion assumes two Airy patterns are just resolvable when the central peak of one aligns with the first minimum of the other. It is commonly used in classical microscopy to describe the practical resolution limit [7, 8].

Abbe Limit

The Abbe limit defines the smallest resolvable feature based on spatial frequency, given by:

$$d_A = \frac{\lambda}{2\mathrm{NA}} \tag{1-8}$$

It reflects the maximum spatial frequency the system can capture and is used in theoretical modeling and advanced imaging techniques such as super-resolution microscopy [7, 8].

1-4 Super-Resolution Microscopy Techniques

To visualize the nanoscale organization of biological structures, imaging techniques must overcome the fundamental limitations imposed by diffraction. The diffraction limit is not simply a technical obstacle but a fundamental limit set by the wave nature of light [7]. It prevents the precise localization of structures that are closer together than approximately half the wavelength of the light used for imaging. For instance, individual proteins, organelles, and even densely packed molecular complexes cannot be clearly resolved because their dimensions are often below the diffraction limit. Therefore, there is a need for imaging techniques capable of resolving these structures beyond the diffraction limit, motivating the development of super-resolution microscopy methods. Super-resolution is a collective term for all techniques that overcome the diffraction limit [11, 12]. These methods push beyond the diffraction barrier to achieve a higher spatial resolution.

1-4-1 Single Molecule Localization Microscopy

Emission-based approaches in super-resolution microscopy focus on the precise localization of individual fluorescent molecules through SMLM (Single-Molecule Localization Microscopy) [13, 14, 15]. These techniques overcome the diffraction limit by temporally isolating fluorescence emissions from overlapping emitters. Figure 1-6 illustrates the process of SMLM in detail. The "ground truth" represents the actual positions of individual fluorescent emitters within a sample. However, due to the diffraction limit, these emitters appear as overlapping point spread functions in the "diffraction-limited image" in Figure 1-6, which blurs the individual emitters into indistinguishable regions.

To resolve this, SMLM activates a sparse subset of emitters at any given time. This is illustrated in the "single-molecule image stack" in Figure 1-6, where emitters are stochastically turned on and imaged across multiple frames. At each time point, the sparse activation ensures that the emitters are spatially separated, allowing their positions to be localized with sub-diffraction precision. This is achieved by fitting a Gaussian profile to the fluorescence

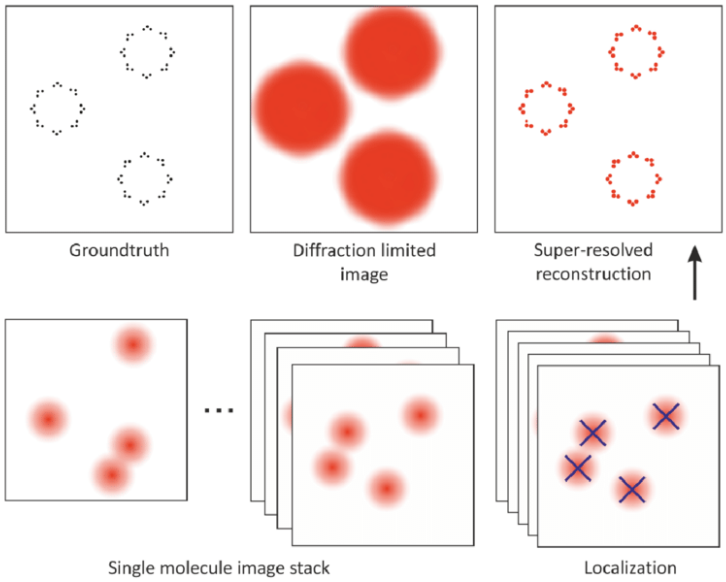


Figure 1-6: Principle of single-molecule localization microscopy (SMLM). The ground truth shows the true positions of individual fluorescent emitters. The diffraction-limited image demonstrates the overlap of point spread functions, limiting resolution due to the diffraction barrier. SMLM overcomes this limitation by activating a sparse subset of emitters in each frame, shown as the single-molecule image stack. The emitters are localized with sub-diffraction precision (indicated by blue crosses), and their positions are iteratively reconstructed across multiple frames to generate the super-resolved image. Source: [16], licensed under the CC BY 4.0 license.

signal from each emitter, with the localization marked by blue crosses, as shown in Figure 1-6. Once the positions of emitters are determined for each frame, these positions are combined iteratively over thousands of frames.

The localization precision σ_{SMLM} of an individual emitter follows a scaling law under the conditions that photon detection follows Poisson statistics and the psf is well-approximated by a Gaussian. Background noise (b) is treated as constant, and finite pixel size effects are incorporated. The lower bound of the precision can be estimated by the equation [17]:

$$\sigma_{\text{SMLM}} = \frac{\sigma_a}{\sqrt{N}} \sqrt{\frac{16}{9} + \frac{8\pi\sigma_a^2 b^2}{Na^2}}, \tag{1-9}$$

where $\sigma_a^2 = \sigma^2 + \frac{a^2}{12}$ accounts for the combined effects of the psf width σ and pixelation due to the detector with pixel size a. Here, N represents the number of detected photons, and b represents the background noise per pixel.

The emitter's fluorescence intensity is modeled by the psf, with pixelation effects accounted for by the term $\frac{a^2}{12}$. Localization precision improves with higher photon counts (N) due to reduced photon noise, as reflected by the term $\frac{\sigma_a}{\sqrt{N}}$. Background noise ($b > 0$) introduces additional uncertainty, described by $\frac{8\pi\sigma_a^2 b^2}{Na^2}$, which becomes significant under low photon counts or high noise levels, emphasizing the importance of minimizing b.

In the ideal case ($b = 0$), localization precision depends only on photon count, psf width and pixel size [17]:

$$\sigma_{\text{SMLM}} \propto \frac{\sigma_a}{\sqrt{N}}, \tag{1-10}$$

Several SMLM techniques achieve nanometer-scale resolution by sparsely activating fluorophores and localizing them over time. STORM and dSTORM use stochastic switching of fluorophores and achieve resolutions of 10–20 nm [14, 18]. PALM, which uses photoactivatable fluorescent proteins, reaches similar resolutions of 20–30 nm and is particularly well-suited for live-cell imaging [13, 19]. PAINT and DNA-PAINT rely on transient binding and achieve 5–10 nm resolution with high multiplexing capabilities [20, 15]. RESI builds on this with reversible switching and also reaches 5–10 nm resolution while reducing photodamage, making it ideal for long-term and live imaging applications [21].

1-4-2 Modulation-Enhanced Single Molecule Localization Microscopy

meSMLM (modulation enhanced SMLM) methods combine illumination-based and emission-based strategies to achieve unparalleled localization precision [22, 23]. These methods leverage excitation patterns, such as doughnut-shaped or sinusoidal illumination, and exploit the properties of emitted fluorescence to encode high-resolution spatial information. By integrating information from both the excitation and emission pathways.

Doughnut Pattern

In MINFLUX [22], the doughnut-shaped intensity minimum serves as the focal point for precise localization. Figure 1-7a shows how the intensity minimum is sequentially positioned at four locations (n_0, n_1, n_2, n_3) around a circular region (dashed line) that contains the fluorophore (red star). The spacing between the excitation positions is denoted by L, which determines the scale of the initial localization region. Figure 1-7b shows the corresponding photon counts measured for each excitation pattern.

The power of MINFLUX lies in its ability to combine two sources of information. First, the excitation pattern provides spatial resolution by creating a predictable relationship between the fluorophore's position and the recorded photon counts. Second, the photon counts themselves encode information about the fluorophore's location, as the fluorescence signal decreases when the fluorophore is closer to the intensity minimum. By analyzing these photon counts relative to the known positions of the intensity minimum, MINFLUX determines the location of the fluorophore with nanometer-scale precision, even with relatively low photon numbers [22].

The scaling law for the localization precision is given by [22]:

$$\sigma_{\text{MINFLUX}} \propto \frac{L}{\sqrt{N}} \tag{1-11}$$

where L is the spacing between excitation positions, and N is the number of detected photons. This precision surpasses traditional SMLM methods because the localization depends on the position of the intensity minimum rather than the centroid of a diffraction-limited spot.

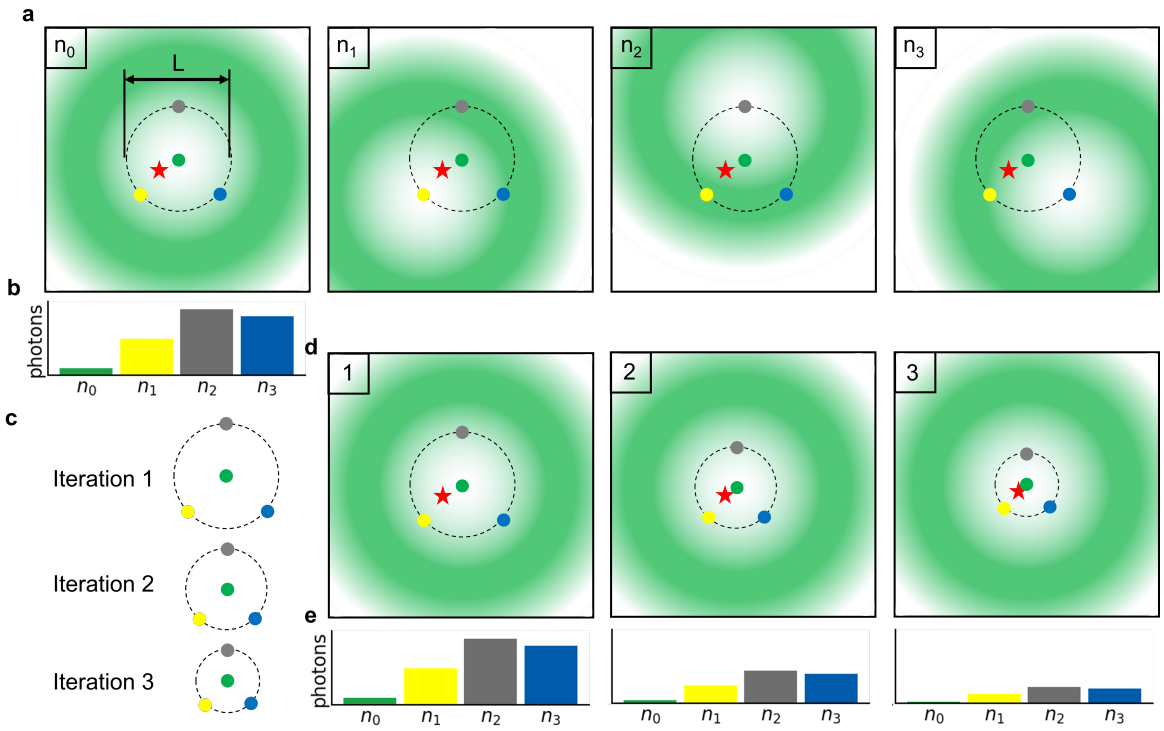


Figure 1-7: Explanation of MINFLUX and iterative MINFLUX; image heavily inspired by [22] and [24]. **(a)** Four excitation patterns n_0–n_3 around an unknown emitter (red star), with current estimate (green dot) and four colored null-centers (gray, blue, yellow, green) on a dashed ring of diameter L over a green-white gradient; only bottom/left spines are shown, top/right hidden. **(b)** Photon-count bar plot for n_0–n_3 (bars colored green, yellow, gray, blue), with x-axis labels n_0–n_3 and y-axis labeled "photons" , bottom/left spines visible. **(c)** Iterative refinement schematic: the dashed ring shrinks toward the red star, and the green-dot estimate converges . **(d)** Iterations 1–3, each showing a smaller dashed ring and updated green estimate relative to the red star; spines as in (a). **(e)** Photon-count bar plots for each iteration, with four colored bars (n_0–n_3).

Iterative MINFLUX [24] builds upon the original MINFLUX approach by introducing an iterative refinement process that progressively narrows the localization area. This method is shown in Figure 1-7.

Figure 1-7c shows the iterative refinement process in MINFLUX, beginning with an initial localization within a relatively large circular region. The doughnut-shaped excitation pattern is positioned at four locations to collect fluorescence data (Iteration 1). Based on the detected photon counts, the emitter's position is estimated and the region is narrowed. In the second and third iterations, the excitation points are placed within progressively smaller regions, using decreasing spacings L between the doughnut positions, thereby refining the emitter's estimated location. Figure 1-7d illustrates the excitation patterns applied at each iteration, showing how the localization region becomes smaller over time. Figure 1-7e presents the corresponding photon counts detected at each position for the three iterations, which are used to iteratively improve the emitter's localization.

This stepwise zooming technique achieves exponential improvements in localization precision with each iteration, as described by the scaling law [24]:

$$\sigma_{\text{iMINFLUX}} \propto \frac{1}{\left(\sqrt{N}\right)^k} \tag{1-12}$$

Here, σ_{iMINFLUX} is the localization precision, N is the number of detected photons, and k is the number of iterations. The iterative approach ensures that the precision improves exponentially with the number of refinement steps.

Sinusoidal Pattern

SIMFLUX [23] is a super-resolution technique that improves the precision of single-molecule localization by introducing sinusoidal excitation patterns that vary spatially across the sample. The intensity of the excitation field at the emitter's position determines how many photons are detected in each frame. Because the sinusoidal pattern is inherently one-dimensional, it is sequentially applied along both the x- and y-directions to enable full two-dimensional localization. By acquiring multiple images with the pattern shifted in phase, the method captures how the detected signal modulates with respect to the known illumination structure [23].

Because SIMFLUX combines both centroid estimation and pattern-based modulation, it can achieve higher localization precision than standard SMLM methods, often with fewer photons [23]. The technique is compatible with wide-field microscope setup and can be implemented using diffraction gratings or spatial light modulators to generate the required illumination patterns.

1-5 Thesis Motivation

Modern biology demands tools capable of visualizing structures and processes far below the diffraction limit of light. Fundamental mechanisms such as protein interactions, intracellular signaling, and molecular organization occur on nanometer scales [7]. Yet conventional fluorescence microscopy is fundamentally limited in resolution to approximately 200–250 nm due to the wave nature of light [7]. To overcome this, a diverse range of super-resolution techniques have been developed [11], enabling researchers to resolve subcellular structures with nanometer precision.

Among these, Single-Molecule Localization Microscopy (SMLM) has become a cornerstone by stochastically activating fluorescent emitters and localizing them with high precision [13, 14, 15]. However, its resolution is ultimately limited by the number of detected photons [17]. Modulation-enhanced approaches such as MINFLUX [22] and SIMFLUX [23] improve upon this by using illumination patterns to encode additional spatial information, achieving higher precision with fewer photons. The MINFLUX technique and its iterative variants represent some of the most photon-efficient approaches in super-resolution microscopy, enabling localization with nanometer precision using remarkably few photons [22, 24]. These methods use the Cramér–Rao Lower Bound (CRLB) to quantify the theoretical limits of localization precision based on the Fisher information content in photon detections [25, 26, 27].

Despite these advances, a crucial aspect of real-world implementation is often overlooked. Ideal MINFLUX assumes a perfect zero-intensity minimum in the center of the doughnut-shaped excitation beam, but such a condition is practically unachievable due to optical aberrations, imperfect interference, and system instabilities [28, 29, 30]. These imperfections result in residual intensity at the beam center and can be captured using a modulation contrast parameter $m \in (0, 1]$ [31]. Deviation from ideal modulation limits the spatial variation in excitation and reduces the Fisher information, thus impairing localization precision.

While the consequences of imperfect modulation have been studied in SIMFLUX [31], their impact on MINFLUX and iterative MINFLUX performance remains unexplored. In this thesis, we extend the CRLB framework to MINFLUX under realistic ($m < 1$) illumination conditions by incorporating modulation contrast as a variable. We aim to quantify how imperfect modulation affects localization precision, determine how this influences the optimal pattern spacing strategy, and evaluate the impact on iterative MINFLUX scaling. By integrating modulation contrast into theoretical analysis and simulation, we provide insight into the real-world performance limits of MINFLUX and propose guidelines for optimal system design under imperfect conditions.

This thesis is structured as follows. Chapter 2 introduces the estimation theory underpinning localization precision in single-molecule microscopy, including maximum likelihood estimation and the Cramér–Rao Lower Bound. Chapter 3 presents the main research contribution in the form of a manuscript titled *Optimal Pattern Location in Modulation Enhanced Microscopy*, which explores how imperfect modulation contrast affects localization performance in MINFLUX and iterative MINFLUX. A mathematical model incorporating arbitrary modulation contrast is developed, and the Fisher Information Matrix and CRLB are derived to investigate how these imperfections shift optimal pattern spacing and influence scaling behavior. Chapter 4 concludes the thesis with a discussion of key findings and possible extensions.

Chapter 2

Estimation of Parameters

Fluorescence microscopy detects individual photons emitted by fluorescent molecules and uses these measurements to infer the molecules positions over time. The randomness in photon emission and detection introduces variability in the measured positions of fluorescent molecules [17, 26]. Estimating molecular positions from photon counts requires statistical inference. Localization-based techniques, including SMLM and MINFLUX, infer emitter coordinates from photon detection events. The spatial resolution of these methods depends on the accuracy of parameter estimation from finite, noisy data [32, 33].

This chapter reviews estimation theory in the context of fluorescence microscopy. Section 2-1 defines estimands, estimators, and estimates. Section 2-2 formulates the likelihood function and the maximum likelihood estimator. Section 2-3 introduces the Cramér-Rao lower bound and its implications for localization precision.

2-1 Estimation

In statistical inference, the aim is to estimate an unknown parameter that characterizes the data-generating process. This parameter, called the estimand, is denoted by θ [25, 34]. It represents a fixed but generally unknown quantity, such as the mean, median, or variance of the data. To estimate θ, a random sample $X_1, X_2, \ldots, X_n$ is collected, where each X_i is a random variable representing an observation from the underlying distribution. Once the sample is observed, the realizations of these random variables are denoted as $x_1, x_2, \ldots, x_n$, and these observed values provide the basis for inferring θ.

Definition 1. *The estimand is the fixed but unknown parameter of a distribution, denoted by θ, that characterizes the data-generating process. It is the quantity of interest we aim to estimate using a random sample $X_1, X_2, \ldots, X_n$. The observed sample values $x_1, x_2, \ldots, x_n$ are realizations of these random variables and form the data used for estimation.*

To estimate an unknown parameter, we use an estimator, which is a rule or function that provides an estimate based on the observed sample [25, 34].

Definition 2. *An estimator, denoted by $\hat{\theta}(X_1, X_2, \ldots, X_n)$ (or simply $\hat{\theta}$), is a function of the random variables $X_1, X_2, \ldots, X_n$, expressed as:*

$$\hat{\theta} = g(X_1, X_2, \ldots, X_n)\,, \tag{2-1}$$

where g is a mapping from the sample space to the parameter space. The estimator is used to infer the value of an unknown parameter θ that characterizes the underlying distribution of the data. Since the estimator is defined in terms of random variables, it is itself a random variable.

Once we have an estimator, applying it to a given dataset produces an estimate, which is the specific value of the estimator for that dataset. An estimate is the realization of the estimator, obtained after observing the sample data [25, 34].

Definition 3. *An estimate is the numerical value produced by an estimator when applied to a particular sample. It is the actual result or outcome based on the observed data. The estimator, denoted $\hat{\theta}$, is a function of the random variables $X_1, X_2, \ldots, X_n$, and when evaluated at the observed sample values $x_1, x_2, \ldots, x_n$, it produces the estimate, denoted $\hat{\theta}(x_1, x_2, \ldots, x_n)$.*

An important property of estimators is *unbiasedness*. An estimator $\hat{\theta}$ is said to be unbiased if its expected value is equal to the true value of the estimand θ [25, 34]. In other words, an unbiased estimator does not systematically overestimate or underestimate the true parameter.

Definition 4. *An estimator $\hat{\theta}$ is said to be unbiased if its expected value is equal to the true value of the parameter being estimated. Formally, $\hat{\theta}$ is unbiased if*

$$\mathbb{E}[\hat{\theta}] = \theta\,, \tag{2-2}$$

where $\mathbb{E}[\hat{\theta}]$ denotes the expected value of the estimator $\hat{\theta}$, and θ represents the true value of the unknown parameter. Unbiasedness ensures that the estimator does not systematically overestimate or underestimate the parameter across repeated samples.

While unbiasedness ensures that an estimator is centered around the true parameter, it does not quantify the variability of the estimator across different samples. This variability is captured by the variance of the estimator [25, 34]. Variance indicates the consistency of an estimator: lower variance implies more stable estimates across different samples.

Definition 5. *The variance of an estimator $\hat{\theta}$ is the expected squared deviation of the estimator from its expected value. It is formally defined as:*

$$\mathrm{Var}(\hat{\theta}) = \mathbb{E}\Big[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2\Big]\,. \tag{2-3}$$

2-2 Likelihood Functions

The likelihood function represents the probability of observing the given sample data as a function of the unknown parameter. Suppose a random sample $X_1, X_2, \ldots, X_n$ is drawn from a distribution characterized by an unknown parameter θ. The likelihood function expresses how likely the observed data $x_1, x_2, \ldots, x_n$ are, given different possible values of θ [25, 34].

Definition 6. *The likelihood function, denoted by $L(\theta \mid X_1, X_2, \ldots, X_n)$, is the joint probability (or probability density) of the random sample $X_1, X_2, \ldots, X_n$ as a function of the unknown parameter θ. It represents how plausible different values of θ are, given the observed data.*

Lemma 1. *If the observations $X_1, X_2, \ldots, X_n$ are i.i.d., the likelihood function can be expressed as:*

$$L(\theta \mid X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} f(X_i \mid \theta)\,, \tag{2-4}$$

where $f(X_i \mid \theta)$ is the probability density function of X_i given θ.

Understanding the relationship between the likelihood function and the joint probability of the observed data is essential in parameter estimation [25, 34]. This connection allows the likelihood function to be derived directly from the data-generating process.

Lemma 2. *The likelihood function $L(\theta \mid X_1, X_2, \ldots, X_n)$ is equivalent to the joint probability of the observed data $X_1, X_2, \ldots, X_n$ given the parameter θ, provided that the probabilistic model is correctly specified. Formally:*

$$L(\theta \mid X_1, X_2, \ldots, X_n) = f(X_1, X_2, \ldots, X_n \mid \theta)\,, \tag{2-5}$$

where $f(X_1, X_2, \ldots, X_n \mid \theta)$ is the joint probability (or density) of the data given θ. This equivalence assumes that the model is correctly specified and that probabilities are well-defined for the observed data.

The likelihood function quantifies how plausible different parameter values θ are, given the observed data [25, 34]. Maximizing the likelihood function identifies the value of θ that makes the observed data most probable under the assumed statistical model. This value is known as the maximum likelihood estimate (MLE) of θ, and it represents the parameter value that best explains the observed data within the framework of the given model.

It is important to distinguish between the *maximum likelihood estimator* and the *maximum likelihood estimate*. The maximum likelihood estimator is a function of the random sample $X_1, X_2, \ldots, X_n$ that produces an estimate of θ. In contrast, the *maximum likelihood estimate* (MLE) is the specific numerical value of θ obtained when the estimator is applied to the observed data $x_1, x_2, \ldots, x_n$.

Definition 7. *The maximum likelihood estimate (MLE) of θ is the value of θ that maximizes the likelihood function $L(\theta \mid x_1, x_2, \ldots, x_n)$. In notation,*

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta}\ L(\theta \mid x_1, x_2, \ldots, x_n)\,. \tag{2-6}$$

Maximizing the likelihood function directly can sometimes be computationally challenging due to the product form of the likelihood, especially for large sample sizes [25, 34]. To simplify the optimization, the log-likelihood function is often used. Taking the logarithm of the likelihood (which does not change the location of the maximum) transforms the product of probabilities into a sum of log-probabilities, which reduces computational complexity. This transformation also improves numerical stability, as multiplying many small probabilities can lead to underflow, an issue mitigated by summing their logarithms. Additionally, the log-likelihood makes differentiation easier, aiding analytical or numerical optimization.

Definition 8. *The log-likelihood function, denoted by $\ell(\theta \mid x_1, x_2, \ldots, x_n)$, is the natural logarithm of the likelihood function $L(\theta \mid x_1, x_2, \ldots, x_n)$. Thus,*

$$\ell(\theta \mid x_1, x_2, \ldots, x_n) = \log L(\theta \mid x_1, x_2, \ldots, x_n)\,. \tag{2-7}$$

For i.i.d. observations, the log-likelihood is the sum of the log of the individual densities:

$$\ell(\theta \mid x_1, x_2, \ldots, x_n) = \log f(x_1 \mid \theta) + \log f(x_2 \mid \theta) + \cdots + \log f(x_n \mid \theta)\,. \tag{2-8}$$

2-3 Cramér–Rao Lower Bound

In estimation theory, a key objective is to minimize the variance of an estimator while keeping it unbiased. The Cramér–Rao lower bound (CRLB) provides a theoretical lower bound on the variance of any unbiased estimator for a given parameter [25, 34].

Definition 9. *The Fisher information, denoted by $I(\theta)$, quantifies the amount of information that the observed data carry about an unknown parameter θ. For a random variable X with probability density function $f(x \mid \theta)$, the Fisher information is defined as:*

$$I(\theta) = \mathbb{E}\left[\left(\frac{\partial \ln f(X \mid \theta)}{\partial \theta}\right)^2\right], \tag{2-9}$$

where the expectation is taken with respect to the distribution of X under the parameter value θ.

Lemma 3. *The Fisher information $I(\theta)$ can alternatively be expressed as the negative expectation of the second derivative of the log-likelihood:*

$$I(\theta) = -\,\mathbb{E}\left[\frac{\partial^2 \ln f(X \mid \theta)}{\partial \theta^2}\right]. \tag{2-10}$$

Theorem 1. *Let $\hat{\theta}$ be an unbiased estimator of a scalar parameter θ, based on a random sample $X_1, X_2, \ldots, X_n$ with probability density function $f(x \mid \theta)$. If certain regularity conditions are satisfied, the variance of $\hat{\theta}$ is bounded below by the inverse of the Fisher information:*

$$\mathrm{Var}(\hat{\theta}) \;\geq\; \frac{1}{I(\theta)}\,, \tag{2-11}$$

where $I(\theta)$ is the Fisher information.

The multivariate extension of the CRLB provides a lower bound on the covariance matrix of any unbiased estimator for a vector of parameters.

Theorem 2. *Consider $\hat{\boldsymbol{\theta}}$, an unbiased estimator for a parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^\top$, obtained from a random sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$. This sample is a realization of independent and identically distributed random variables $\mathbf{X} = (X_1, X_2, \ldots, X_n)^\top$ with joint probability density $p_{\boldsymbol{\theta}}(\mathbf{x})$ that depends on $\boldsymbol{\theta}$. Assume the log-likelihood $\ell(\boldsymbol{\theta} \mid \mathbf{x}) = \ln p_{\boldsymbol{\theta}}(\mathbf{x})$ is continuously differentiable with respect to $\boldsymbol{\theta}$.*

The entries of the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ *are given by*

$$[\mathbf{I}(\boldsymbol{\theta})]_{i,j} = -\,\mathbb{E}\left[\frac{\partial^2 \ell(\boldsymbol{\theta} \mid \mathbf{x})}{\partial\theta_i\,\partial\theta_j}\right], \tag{2-12}$$

where the derivatives are evaluated at the true parameter value $\boldsymbol{\theta}$*, and the expectation is taken with respect to the distribution of* $\mathbf{x}$ *at* $\boldsymbol{\theta}$*. If* $\mathbf{I}(\boldsymbol{\theta})$ *is invertible, then the covariance matrix* $\mathbf{C}_{\hat{\theta}}$ *of any unbiased estimator* $\hat{\boldsymbol{\theta}}$ *satisfies:*

$$\mathbf{C}_{\hat{\theta}} \succeq \mathbf{I}(\boldsymbol{\theta})^{-1}, \tag{2-13}$$

where $\mathbf{C}_{\hat{\theta}} = \mathbb{E}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\top]$*, and the notation* $\succeq$ *indicates that* $\mathbf{C}_{\hat{\theta}} - \mathbf{I}(\boldsymbol{\theta})^{-1}$ *is positive semi-definite.*

In fluorescence microscopy, estimation theory underpins the localization of single emitters. In single-molecule localization microscopy (SMLM) and MINFLUX, the goal is to estimate emitter positions—typically x and y coordinates—from photon counts measured across camera pixels or detector bins. Given a known point spread function or illumination profile, the likelihood function for emitter position can be constructed. The Cramér–Rao lower bound (CRLB) sets a theoretical lower limit on the variance of any unbiased estimator under a given noise model [26]. Under shot-noise-limited conditions, this bound predicts that localization precision scales with the inverse square root of the detected photon count [27]. In MINFLUX, which achieves high precision with few photons, the CRLB is used to quantify the trade-off between photon budget and spatial resolution [22].

Chapter 3

Optimal Pattern Location in MINFLUX

This chapter is written in the form of a standalone manuscript prepared for publication. It presents the theoretical modeling, simulation results, and analysis concerning the optimal placement of illumination patterns in MINFLUX and iterative MINFLUX microscopy, with a specific focus on the influence of imperfect modulation contrast.

Optimal Pattern Location in Modulation Enhanced Microscopy

Koen Stapel[1*], Dylan Kalisvaart[1*], Mirthe Bronkhorst[1], Xinyue Zhang[1], Carlas Smith[1*]

[1]Delft Center for Systems and Control, Delft University of Technology.

*Corresponding author(s)

Abstract

MINFLUX (minimal fluorescence photon fluxes) achieves nanometer-scale precision by probing single emitters with a donut excitation beam using the central intensity minimum. This principle enables localization that, for small displacements, becomes independent of the emission wavelength, further enhancing resolution. A critical yet underexamined parameter in this method is the pattern placement radius $\boldsymbol{L}$, which governs how spatial information is extracted from photon counts. Here, we show that optimizing $\boldsymbol{L}$ is essential for maximizing localization precision under realistic conditions where the illumination minimum is imperfect ($\boldsymbol{m < 1}$). Using Cramér–Rao lower bound simulations, we find that localization precision reaches a global optimum at a finite radius $\boldsymbol{L}_{\text{opt}}$, and that misplacing this radius leads to substantial performance losses. For modulation contrast $\boldsymbol{m = 0.95}$, tuning $\boldsymbol{L}$ to 34 nm improves localization precision by a factor of 1.6 relative to standard fixed-radius schemes. We derive a predictive relationship, $\boldsymbol{L}_{\text{opt}} \approx \mathbf{1.30}\boldsymbol{\sigma}_{\text{illum}}\sqrt{\mathbf{1} - \boldsymbol{m}}$, valid for $\mathbf{0.8} \leq \boldsymbol{m} \leq \mathbf{0.99}$. We incorporated a Gaussian prior on emitter position, to handle uncertainty in initial emitter position that is used to place the patterns on the emitter. We found the optimal spacing $\boldsymbol{L}_{\mathbf{opt}}$ for a given $\boldsymbol{\sigma}_{\mathbf{prior}}$. We further show that, in iterative MINFLUX, repeatedly applying this optimal $\boldsymbol{L}$ in each cycle outperforms conventional sequential shrinking strategies, improving precision by over 50% at $\boldsymbol{m = 0.95}$. These findings establish pattern placement as a decisive factor in modulation enhanced microscopy and provide a framework for extracting maximal information from minimal photon budgets. Experimental validation under biological imaging conditions will be critical to assess practical gains.

Keywords: MINFLUX, Single-Molecule Localization Microscopy (SMLM), Modulation, Cramér–Rao Lower Bound (CRLB), Imperfect Illumination, Modulation Contrast, Iterative Localization, Localization Precision

1 Introduction

Single-molecule localization microscopy (SMLM) has revolutionized optical imaging by enabling fluorescent molecules to be localized far beyond the diffraction limit of conventional microscopes [1]. Techniques such as photoactivated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) achieve nanometer-scale resolution by sequentially activating and localizing individual emitters, reconstructing super-resolved images molecule-by-molecule [1–3]. Further improvements in resolution have been realized through modulation enhancing strategies, notably exemplified by methods like MINFLUX (minimal emission flux), iterative MINFLUX, MINSTED, and SIMFLUX [4–7]. Figure S1 provides a schematic overview of (Iterative) MINFLUX. MINFLUX uses a targeted doughnut-shaped excitation pattern with a central intensity null to localize a single emitter, and iterative MINFLUX builds on this by repeatedly narrowing the illumination pattern around the refined position estimate [4, 5].

MINFLUX operates by centering a dark focal spot (intensity minimum) on the emitter's location, so that any slight displacement of the emitter from the null yields an imbalance in detected photon counts across multiple illumination positions [4]. By strategically positioning a set of doughnut illumination patterns around the molecule, MINFLUX extracts localization information more efficiently per photon than traditional Gaussian PSF localization approaching sub-nanometer precision [4, 8, 9]. This principle enables localization that, for small displacements, becomes independent of the emission wavelength, further enhancing resolution. Recent demonstrations using high-order vortex beams and two-photon excitation confirm the remarkable potential of MINFLUX, achieving localization precisions well below 1 nm under ideal conditions [8, 9]. The iterative MINFLUX technique extends the basic MINFLUX approach by successively zooming in on the emitter. After an initial coarse localization, the doughnut excitation pattern is re-centered and scaled to a smaller radius, and this process is repeated to further refine the position estimate [5]. Under perfect imaging conditions, each iteration of MINFLUX can improve localization precision by concentrating illumination where it yields the most information [5]. Beyond MINFLUX, related advances like MINSTED combine a coordinate-targeted STED depletion focus with MINFLUX-like readout, enabling real-time single-molecule tracking at nanometer spatial resolution and millisecond temporal resolution [6, 10].

In practice, optical aberrations mean that a real doughnut beam never features a perfectly zero-intensity center [11, 12]. Even with state-of-the-art optics, a residual light intensity remains at the nominal "dark" center of the focal spot, reducing the modulation contrast m of the illumination pattern from its ideal value of 1 [13, 14]. m (ranging from 0 to 1) quantifies how well the excitation is modulated: $m = 1$ corresponds to a perfect doughnut with zero central intensity, whereas $m < 1$ indicates a partially filled minimum [15]. An ideal doughnut pattern exhibits a deep null at the center, while an imperfect doughnut (low m) has a shallower intensity dip. Reduced modulation contrast directly diminishes the photon count differences that MINFLUX relies on, thus degrading the Fisher information available for localization [11, 13].

Under imperfect modulation ($m < 1$), the optimal illumination pattern spacing L in SIMFLUX is not the smallest possible; instead, localization precision is maximized at a finite, non-zero L [15]. Similar contrast-aware optimization remains largely unexplored for MINFLUX and its iterative implementations. Previous studies have optimized MINFLUX parameters under experimental imperfections [16] and improved iterative MINFLUX tracking by tuning dwell time, photon thresholds, and laser power [17]. However, neither of these studies investigated how imperfect modulation contrast ($m < 1$) impacts the choice of illumination pattern geometry in MINFLUX or iterative MINFLUX. In contrast, this work explicitly incorporates imperfect modulation conditions into the theoretical framework and optimizes pattern placement to maximize localization precision under m.

Motivated by this gap, this work optimizes the pattern locations in MINFLUX and iterative MINFLUX with a non-ideal modulation conditions. We incorporate a modulation contrast parameter into the MINFLUX theoretical framework and use Fisher-information-based analysis to calculate the Cramér–Rao lower bound (CRLB) for localization precision as a function of pattern placement. By optimizing the pattern locations for a given measured contrast m, we identify the strategies that maximize localization information. We determine how the optimal pattern spacing evolves as m deviates from 1, and we examine whether and when iterative MINFLUX provides additional benefit under these optimized conditions.

2 Method

2.1 Image Formation Model

We calculate the lower bound on localization precision by modeling image formation and deriving the Fisher information matrix and CRLB [18, 19]. The parameter vector is $\boldsymbol{\theta} = [\theta_x, \theta_y, \theta_I, \theta_b]^{\mathrm{T}}$. Here, θ_x and θ_y are emitter coordinates. θ_I is the expected signal photon count. θ_b is the expected background photon count per pattern [20]. The image formation model combines previous models [15, 20, 21].

The expected photon count at pixel (x_i, y_i) is

$$\mu_i(x_i, y_i, \boldsymbol{\theta}) = \theta_I P(\theta_x, \theta_y) E_x(x_i, \theta_x) E_y(y_i, \theta_y) + \theta_b, \tag{1}$$

where P is the illumination pattern at the emitter position. E_x and E_y are marginal PSFs [20]. Details are in Supplementary Note 1 [15, 21].

The Fisher information matrix is

$$I_{p,q}(\boldsymbol{\theta}) = \sum_{i=1}^{N} \frac{1}{\mu_i} \frac{\partial \mu_i}{\partial \theta_p} \frac{\partial \mu_i}{\partial \theta_q}, \tag{2}$$

where N is the number of pixels summed over patterns. Partial derivatives are in Supplementary Note 2 [21].

Photon detection in MINFLUX can be modeled either as a set of independent Poisson processes—one for each illumination pattern—or using a multinomial model in which the total number of detected photons θ_I is fixed [4]. The Poisson model

allows for variability in both the total photon count and the pattern-specific counts, whereas the multinomial model assumes a known θ_I and focuses only on the relative distribution of photons across patterns. In our simulations, we use a Poisson detection model but normalize to a constant total photon count.

We chose the Cramér–Rao lower bound (CRLB) over the Van Trees inequality (VTI) because the CRLB yields a deterministic, smooth objective function that is computable and suits efficient optimization [18, 22]. In contrast, the VTI (the Bayesian Cramér–Rao bound) entails integrating over a prior distribution of the emitter position [23, 24], often lacking closed-form expressions and requiring computationally intensive methods such as Monte Carlo integration [22, 25]. This introduces stochastic variability that can hinder optimization convergence. While the VTI provides a global uncertainty measure by averaging Fisher information over a prior [23, 26]. It is only applicable in scenarios where such a prior is explicitly defined, such as in iterative localization with prior knowledge of the emitter position. Since our focus lies on optimizing the precision for a fixed emitter position, the local CRLB is more appropriate. It quantifies the best-case variance at a specific parameter estimate and aligns with our goal of refining a single localization [18, 25]. Furthermore, literature demonstrates that Fisher-information-based objectives are effective in driving successive localization refinements [27].

2.1.1 Modulation Contrast

MINFLUX localizes emitters by sequentially illuminating with four doughnut patterns n_0 through n_3 ad in Figure S1, placed on a circle of diameter L [28]. The center pattern n_0 is at $(0, 0)$; the others are spaced 120° apart on this circle. Adjusting L controls the spatial sampling region and is key to iterative localization refinement. The excitation pattern in MINFLUX is modeled as a two-dimensional doughnut beam centered at (x_p, y_p) with width parameter σ_{illum} [21]. The doughnut pattern is

$$P_{\text{donut}}(x, y) = \frac{(x - x_p)^2 + (y - y_p)^2}{2\sigma_{\text{illum}}^2} \exp\left(-\frac{(x - x_p)^2 + (y - y_p)^2}{2\sigma_{\text{illum}}^2}\right), \tag{3}$$

Pattern position	Coordinates (x_p, y_p)
n_0 (center)	$(0, 0)$
n_1	$\left(\frac{L}{2}, 0\right)$
n_2	$\left(-\frac{L}{4}, +\frac{\sqrt{3}L}{4}\right)$
n_3	$\left(-\frac{L}{4}, -\frac{\sqrt{3}L}{4}\right)$

Table 1 Coordinates (x_p, y_p) of patterns n_0–n_3 with pattern separation L.

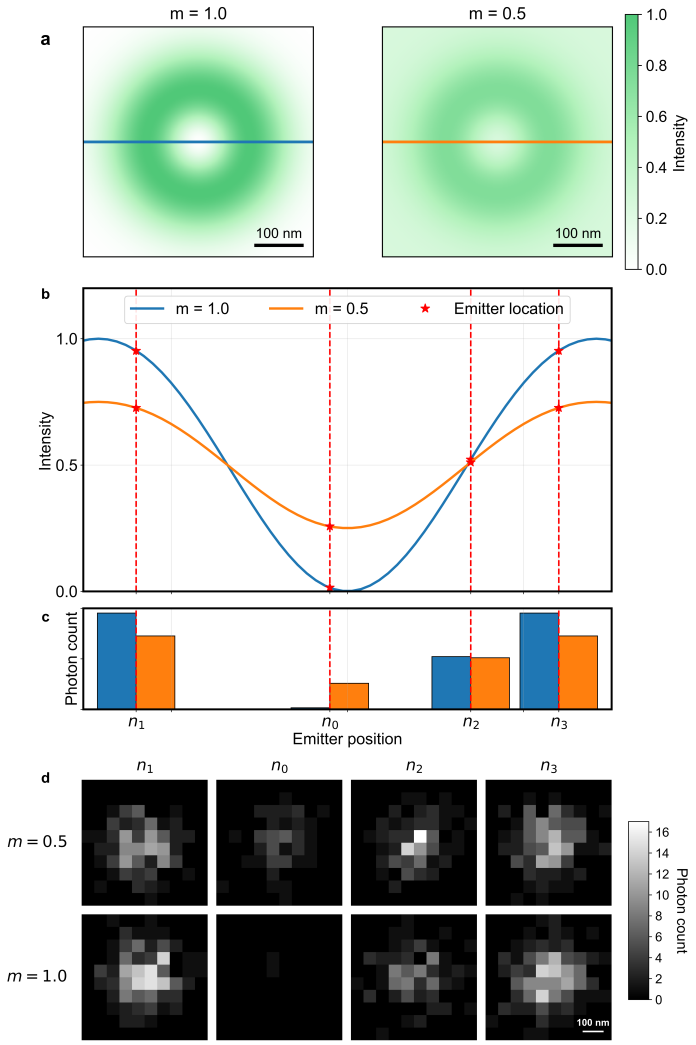


Fig. 1 **(a)** Two-dimensional simulated doughnut-shaped excitation patterns under perfect modulation ($m = 1.00$, left) and imperfect modulation ($m = 0.50$, right). Green indicates highest excitation intensity; white indicates zero. Insets show a single horizontal cross-section (colored line) at the vertical center of each donut. The perfect-modulation pattern ($m = 1.00$) has a deep central null, whereas the imperfect-modulation pattern ($m = 0.50$) retains a residual central intensity, reducing the contrast of the zero-node. **(b)** Horizontal cross-sections through these two doughnut patterns (same colored lines as in panel a), replotted on a 1D axis in units of σ_{illum}. The blue curve ($m = 1.00$) dips to zero at the center (emitter position), while the orange curve ($m = 0.50$) remains elevated. Overlayed red stars mark four predefined "pattern positions" n_0, n_1, n_2, n_3 (vertical dashed lines) at which the emitter is interrogated. Here, n_0 is the donut with its central minimum directly over the emitter (center pattern), and n_1–n_3 are three donut patterns shifted radially outward (outside patterns). Because the emitter's absolute coordinate changes relative to each donut pattern, the four red stars lie at different points on each cross-section. **(c)** Corresponding photon counts expected at each pattern position n_i, for both $m = 1.00$ (blue bars) and $m = 0.50$ (orange bars). These bar graphs are the four red-star corresponding photoncount from panel (b). **(d)** Eight simulated single-molecule detection images reflecting the summed photon counts under each combination of modulation contrast ($m = 1.00$ top row, $m = 0.50$ bottom row) and pattern position n_0–n_3.

Modulation contrast m quantifies how strongly an illumination pattern oscillates between its maximum and minimum intensity values and is defined as

$$m = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \tag{4}$$

In an ideal MINFLUX doughnut with a perfectly dark center, $I_{\min} = 0$ and thus $m = 1$. However, experiments in biological tissue show that optical scattering and aberrations lead to a nonzero doughnut minimum. For instance, residual intensity at the doughnut center of 3%, indicating degradation in doughnut contrast [13]. Optical aberrations fill in the doughnut's dark center [11]. Similar effects are seen in structured illumination microscopy (SIM), where aberrations reduce pattern contrast [29, 30]. Rather than modeling each aberration or scattering mechanism individually, using a finite $m < 1$ effectively aggregates these imperfections into a single, experimentally measurable parameter [15]. It should be noted, however, that modulation contrast only reflects illumination-forming imperfections. It does not account for photon shot noise, detector readout noise, point spread function mismatches [11].

Figure 1a shows that reducing the modulation contrast from $m = 1.0$ to $m = 0.5$ results in a partially filled doughnut minimum. In Figure 1b, this reduced contrast directly translates to shallower modulation of intensity across the pattern positions. The differences between maxima and minima flatten, meaning emitter displacement causes smaller changes in detected signal. Consequently, Figure 1c demonstrates how the photon count differences between pattern positions decrease with lower m, which reduces the Fisher information available for localization. Finally, Figure 1d shows that with imperfect modulation $m = 0.5$, the images become more similar across pattern positions.

2.2 Optimization

The objective is to optimize the illumination spacing parameter $L > 0$ to minimize the average localization uncertainty of the emitter. To obtain a single measure of localization uncertainty, we chose the root-mean-square average [4]. The cost function is defined as

$$\min_{L>0} \sigma_{x,y}(L), \quad \text{with} \quad \sigma_{x,y}(L) = \sqrt{\frac{\sigma_x(L)^2 + \sigma_y(L)^2}{2}}. \tag{5}$$

The values $\sigma_x(L)$ and $\sigma_y(L)$ are obtained from the diagonal entries of the Cramér–Rao lower bound (CRLB) covariance matrix. This matrix is computed by inverting the Fisher Information Matrix (FIM), which is numerically constructed for each candidate value of L [18].

The function $\sigma_{x,y}(L)$ is evaluated only for $L > 0$. While negative values of L would yield valid Fisher information and precision values, they correspond to a geometric reflection of the illumination pattern around the emitter. Because the physical setup assumes a fixed relative arrangement of illumination positions, only the magnitude of the spacing is optimized. The sign is not considered relevant, and the domain is restricted.

Figure S2 shows the CRLB-derived localization precision $\sigma_{x,y}(L)$ as a function of spacing L for modulation contrast values $m \in \{0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.99\}$. Each curve corresponds to a fixed m. The results show that for each m, $\sigma_{x,y}(L)$ reaches a unique global minimum at a finite $L = L_{\text{opt}}(m)$.

Derivatives of the CRLB with respect to L require differentiating through the Fisher Information Matrix (FIM), denoted by I, which is a symmetric matrix-valued function of L. Since the CRLB involves the inverse of I, the gradient includes terms of the form $-I^{-1}(\partial I/\partial L)I^{-1}$ [31]. These terms are computationally expensive to evaluate and not readily available in closed form.

The optimization of $\sigma_{x,y}(L)$ is performed using the Nelder–Mead simplex method [32]. This derivative-free algorithm is appropriate because the objective function involves matrix inversion and numerical integration, making analytical derivatives with respect to L computationally costly to obtain [31]. The Nelder–Mead method requires only pointwise evaluations of $\sigma_{x,y}(L)$, making it well-suited to scenarios where gradients are unavailable. Additionally, numerical evaluations show that $\sigma_{x,y}(L)$ is unimodal for fixed modulation contrast m, allowing local search methods like Nelder–Mead to reliably converge to the global minimum without requiring gradient information. The complete Nelder–Mead algorithm, including all design choices, is specified and justified in Supplementary Note 3. All parameter values appear in Table S1.

2.2.1 Prior Regularization

To determine the optimal illumination radius L in practical conditions where the emitter's position is not known exactly, we incorporate prior uncertainty into the optimization. We define the emitter offset vector as $\boldsymbol{\delta} = (x_{\text{offset}}, y_{\text{offset}})$, representing the displacement from the center of the illumination pattern. This uncertainty is modeled as a two-dimensional Gaussian distribution centered at the origin: $p_{\sigma_{\text{prior}}}(\boldsymbol{\delta}) = \mathcal{N}(\mathbf{0}, \sigma^2_{\text{prior}}\,\mathbf{I})$, where σ_{prior} denotes the standard deviation of the prior estimate. This formulation introduces a form of distributional regularization [33], in which the optimization accounts for variability in the emitter location rather than minimizing precision at a single fixed position.

To compute the expected localization precision under this prior, we define:

$$\mathbb{E}[\sigma_{x,y}(L)] = \iint_{\mathbb{R}^2} \sigma_{x,y}(L, \boldsymbol{\delta})\, p_{\sigma_{\text{prior}}}(\boldsymbol{\delta})\, d\boldsymbol{\delta} \approx \frac{1}{N} \sum_{i=1}^{N} \sigma_{x,y}(L, \boldsymbol{\delta}_i), \tag{6}$$

where $\sigma_{x,y}(L, \boldsymbol{\delta})$ is the localization uncertainty for a specific offset, and $\boldsymbol{\delta}_i \sim p_{\sigma_{\text{prior}}}$ are sampled offsets. The integral is approximated by a Monte Carlo average over N stratified samples.

The goal is then to find the illumination radius L that minimizes the expected localization uncertainty:

$$\min_{L>0} \mathbb{E}[\sigma_{x,y}(L)]. \tag{7}$$

Figure S6 shows that the resulting objective landscape is nearly convex and chas a single global minimum. However, non-convexity may appear for small values of L with

insufficient sampling. In our implementation, we found that using 500 stratified 2D samples yielded a global minimum optimization surface. Due to the unimodal structure and the lack of gradient information, we again employ the Nelder–Mead algorithm for minimizing the expected uncertainty. The complete optimization routine is described in Supplementary Note 4.

2.2.2 Iterative MINFLUX Optimization

The objective in iterative MINFLUX is to optimize a sequence of illumination spacings $(L_1, L_2, \ldots, L_k) \in \mathbb{R}^k_{>0}$ across k iterations to minimize the average localization uncertainty of the emitter. The cost function is defined as

$$\min_{L_1,\ldots,L_k>0} \sigma_{x,y}(L_1,\ldots,L_k), \quad \text{with} \quad \sigma_{x,y}(L_1,\ldots,L_k) = \sqrt{\frac{\sigma_x(L_1,\ldots,L_k)^2 + \sigma_y(L_1,\ldots,L_k)^2}{2}}. \tag{8}$$

The total Fisher Information Matrix is computed by summing the Fisher Information Matrices from each iteration [18]. Each term in the sum corresponds to one illumination pattern spacing L_i, such that the cumulative information reflects all k steps in the sequence.

Figure S3 and S4 visualize the optimization landscape of $\sigma_{x,y}(L_1, L_2)$ and $\sigma_{x,y}(L_1, L_2, L_3)$, respectively. These plots show smooth and structured objective surfaces with a clear minimum, suggesting that the function is amenable to local optimization [31]. However, global optimality cannot be guaranteed in higher dimensions, and the possibility of local minima increases with the number of parameters [34].

Although the optimization landscape appears amenable to local optimization, verifying whether a local solver can reliably reach the global minimum is essential. To assess this, we compared the results of a local optimization algorithm (Nelder–Mead) with a global method (Differential Evolution) [35]. As shown in Figure S5, the maximum observed difference in L_k between the two methods is only 0.08 nm. This small discrepancy suggests that both methods consistently converge to the same solution, indicating that the optimization landscape likely features a unique global minimum [31].

Given that the Nelder–Mead algorithm converged approximately 20 times faster in our simulations, it was selected for use in all subsequent optimization tasks. The full algorithm and all parameter settings are described in Supplementary Note 5.

3 Results

The expected localization precision in MINFLUX scales as $\sigma_{x,y} \propto L/\sqrt{N}$ [4]. Figure S7 confirms this: precision improves with $\sqrt{N}$ and grows linearly with L with $m = 1$.

Figure 2a illustrates how the optimal choice of illumination spacing L depends on the modulation contrast m. For perfect modulation ($m = 1$), the localization precision $\sigma_{x,y}$ continues to improve as L approaches zero, in line with the ideal theoretical prediction [4]. However, for imperfect modulation ($m < 1$), this scaling no longer holds. Instead, $\sigma_{x,y}$ reaches a minimum at a finite, nonzero L_{optimal}, and this optimum shifts to larger values as m decreases. For example, at $m = 0.95$, the optimal L is

approximately 34.38 nm. The optimal L is no longer as small as possible, but instead depends on m. These findings are consistent with those observed in SIMFLUX [15].

This behavior can be explained by considering the photon emission profile near the center of the illumination pattern. For perfect modulation ($m = 1$), the central region of the donut-shaped pattern emits no photons. As a result, even a small displacement of the emitter leads to a noticeable change in detected photon counts, since the emitter transitions from a dark region into a brighter region. This sharp contrast produces a high localization sensitivity. However, when $m < 1$, the central intensity is no longer zero. Although the spatial derivative of the illumination remains shallow near the center, the emitter now receives a nonzero photon count even without displacement. Consequently, the relative change in detected photons due to small displacements is reduced, lowering the available localization information. This effect favors larger illumination spacings, where the emitter samples regions with higher intensity gradients $L_{\text{opt}} > 0$.

The MINFLUX localization precision to the choice of L is quantified in Figure 2b. This panel shows the relative increase in localization uncertainty when L deviates from L_{optimal}, for various m values. For all $m < 1$, selecting a suboptimal L can substantially degrade localization precision. For instance, using a fixed small spacing such as $L = 5$ nm results in a localization precision that is at most 60% of the optimum for all values of m shown in the figure. For a modulation contrast of $m = 0.95$, using $L = 50\,\text{nm}$ and $L = 70\,\text{nm}$ (values originally suggested in the MINFLUX implementation [4]) results in localization precisions of 95.7% and 79.3% of the minimum achievable $\sigma_{x,y}$, respectively. For $m = 0.99$, the corresponding values are 56.6% and 42.8%.

Figure 2c shows that σ_x and σ_y exhibit symmetry about the y-axes. This follows from the pattern placement being symmetrical in the y-axis. The combined precision $\sigma_{x,y}$ displays a threefold rotational symmetry, consistent with three off-center illumination positions spaced 120° apart. The highest localization precision is achieved at the center, where information from all patterns overlaps maximally.

Figure 2a suggests a relation between the optimal L and the modulation contrast m, which is quantified in Figure 2d. The plot displays L_{optimal} as a function of $1-m$ in log–log scale. The data approximately follow a power law with slope 0.5, corresponding to

$$L_{\text{optimal}} \propto \sqrt{1-m}. \tag{9}$$

From Figure 2d, the goal is to directly express L_{opt} as a function of m. A dimensional analysis was performed. The quantity L_{opt} has units of nanometers (nm), while $\sqrt{1-m}$ is dimensionless. Therefore, the proportionality constant must also have units of nm. The initial hypothesis was that this constant depends on the illumination width σ_{illum}, which has units of nm. Figure 2e confirms this hypothesis. The data supports the following relation:

$$L_{\text{opt}} = A \cdot \sigma_{\text{illum}} \sqrt{1-m} \tag{10}$$

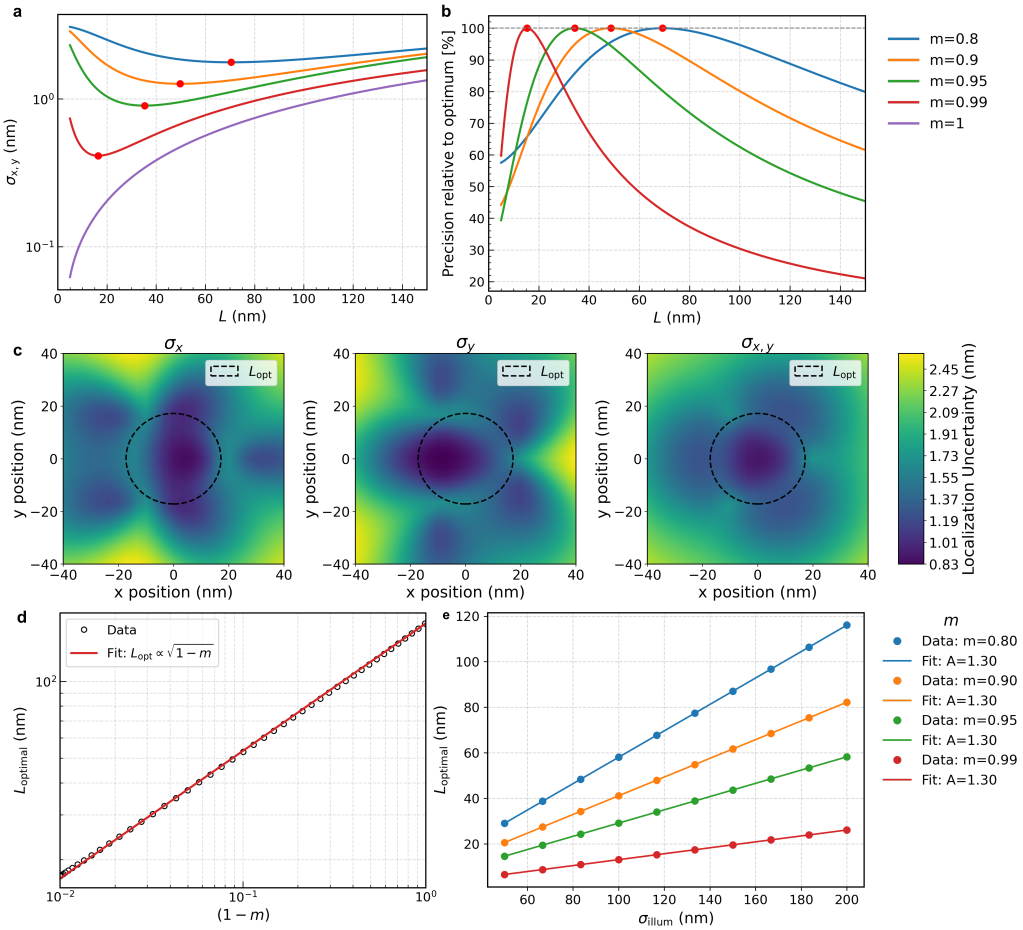


Fig. 2 **(a)** Localization precision $\sigma_{x,y}$ as a function of pattern separation L for various modulation contrasts m. Red dots indicate the optimal values L_{opt} that minimize the localization uncertainty. Imperfect modulation significantly shifts the optimal L. All simulations use an emitter photon count of $\theta_I = 1000$ and background photon count $\theta_b = 0$. **(b)** Relative increase in localization uncertainty $\sigma_{x,y}$ as a function of pattern separation L, normalized to the optimal value L_{opt} for each modulation contrast m. The y-axis shows the ratio $\sigma_{x,y}(L)/\sigma_{x,y}(L_{\text{opt}})$, expressed as a percentage. A value of 100% corresponds to the minimum localization uncertainty, i.e., when $L = L_{\text{opt}}$. Each curve corresponds to a different m from (a), except $m = 1$, which is excluded because it has no optimum. **(c)** Spatial maps of localization precision with $m = 0.95$ and $L = L_{\text{opt}}$, visualizing precision over a grid of emitter positions. The three subpanels show σ_x, σ_y, and $\sigma_{x,y}$. **(d)** Optimal pattern spacing L_{opt} as a function of $1 - m$ in log–log scale. The fitted line confirms the relationship $L_{\text{opt}} \propto \sqrt{1-m}$, with a slope of 0.50. **(e)** Fitted scaling relationship $L_{\text{opt}} = A \cdot \sigma_{\text{illum}} \sqrt{1-m}$ for multiple modulation contrasts m. Each pair of data and fit lines corresponds to a different m value.

This relation is dimensionally consistent. The proportionality constant A was determined to be 1.30. This value was obtained by fitting all data points corresponding to

four values of σ_{illum}, using only modulation contrasts $m \in [0.8, 0.99]$. A linear regression was performed on the transformed data. The fit yielded an R^2 score of 0.99998. Lower values of m were excluded due to optical setups being able to achieve these values m [36]. This results in the following predictive expression:

$$L_{\text{opt}} = 1.30 \cdot \sigma_{\text{illum}} \sqrt{1 - m} \tag{11}$$

The parameters σ_{illum} and m in Equation 11 can be determined experimentally by scanning the illumination pattern over a single emitter and recording the detected intensity, converted to photon counts. The modulation contrast m is then calculated as$m = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$, where $I_{\max}$ and $I_{\min}$ are the maximum and minimum detected intensities, respectively. The illumination width σ_{illum} can be estimated from the distance between the two intensity maxima in the scan profile. Alternatively, using the conversion in [14], the relation can be written as:

$$L_{\text{opt}} = 0.36 \cdot \frac{\lambda}{NA} \sqrt{1 - m} \tag{12}$$

We evaluate the domain in which Equation 11 remains accurate. Figures S8a and S8b show the absolute and relative difference between predicted and optimized L_{opt} over a domain of $m \in [0.8, 0.99]$ and $\sigma_{\text{illum}} \in [80, 150]$ nm. Within this range, the maximum absolute error in L_{opt} was 0.166 nm, and the maximum relative error was 0.239%. The impact of this deviation on localization precision σ_{xy} is shown in Figure S8c and S8d. The maximum deviation in σ_{xy} within this domain was 0.0000026 nm. A broader parameter range was also evaluated, with $m \in [0.1, 0.99]$ and $\sigma_{\text{illum}} \in [20, 500]$ nm. Here, the largest error in L_{opt} was 4.810 nm or 1.267%, and the largest deviation in σ_{xy} was 0.00016 nm. Outside this domain, particularly for low m and small σ_{illum}, the prediction error increased rapidly. Therefore, the validity domain of Equation 11 is:

$$0.1 \leq m \leq 0.99, \quad 20 \text{ nm} \leq \sigma_{\text{illum}} \leq 10000 \text{ nm} \tag{13}$$

The predictive relation in Equation 11 is derived under the assumption that the illumination pattern is centered on the emitter. This condition ensures that the optimization yields the smallest achievable localization uncertainty. In practice, this assumption does not always hold. During measurements, the emitter position is not known, and the illumination pattern may be offset from the true emitter location [4, 17]. Figure S9 shows how lateral emitter offsets affect the optimal illumination spacing. As the displacement increases, the optimal L grows and deviates from the prediction in Equation 11, consistent with prior observations [4, 17]. Figure S10 highlights that this deviation scales with the relative emitter–pattern offset. Figure S11 shows that to limit localization precision degradation to under 10%, the radial offset must satisfy $r_{\text{offset}} \leq 0.2 \cdot L$. Since L is the pattern diameter, this means the emitter must stay within 40% of the pattern radius to maintain high precision. However, Figure S11 also shows that localization precision degrades sharply once the emitter lies outside this diameter. This indicates that the relation in Equation 11 no longer holds

when the emitter is offset from the center. It is therefore more appropriate to interpret this relation as a lower bound (minimum), rather than a true optimum, which depends on the emitter's position relative to the pattern.

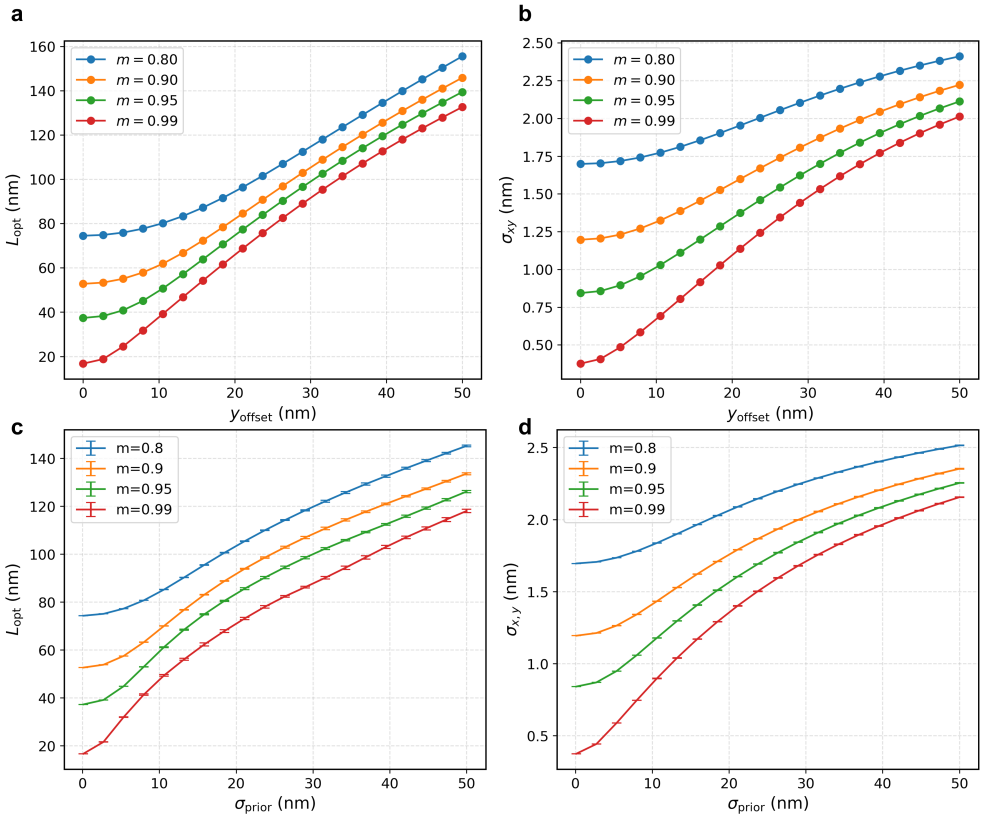


Fig. 3 **(a)** Optimal illumination spacing L_{opt} as a function of emitter offset along the y-axis, computed by minimizing the localization uncertainty for each fixed offset. **(b)** Corresponding localization precision $\sigma_{x,y}$ for each optimized spacing in panel (a). **(c)** Optimal illumination spacing L_{opt} as a function of prior uncertainty σ_{prior}, where the emitter position is assumed to follow a 2D Gaussian distribution. **(d)** Mean localization precision $\sigma_{x,y}$ evaluateqd at the optimal spacing from panel (c).

To find the pattern placement when Equation 11 does not hold, the effect of emitter displacement relative to the center of the illumination pattern on both the optimal spacing L_{opt} and the resulting localization precision is shown in Figures 3a and 3b. The value of L_{opt} increases monotonically with the emitter offset for all tested modulation contrasts. Equivalent results for offsets along the x-axis and along the diagonal direction $r = -\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}}$ are provided in Figure S12 and exhibit the same trend.

In practice, pattern placement is based on a prior estimate of the emitter position [4]. Therefore, Figures 3a and 3b, which assume a known offset, are not directly

applicable for determining the optimal L. If the emitter's offset is known precisely, no further localization would be required. To address this, we introduce the prior uncertainty σ_{prior}, defined such that the emitter position is modeled as $\begin{bmatrix} x_{\text{offset}} \\ y_{\text{offset}} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\text{prior}}\,\mathbf{I})$. Here, σ_{prior} represents the standard deviation of the initial estimate, which may be derived from the measurement variance [18], an CRLB [20], or existing analytical approximations [37–39].

Figures 3c and 3d show the optimal spacing L_{opt} and the resulting average localization precision $\sigma_{x,y}$ as a function of σ_{prior}, respectively. For all values of m, the optimal spacing increases with increasing prior uncertainty, and the resulting localization precision degrades accordingly. These plots can be used to select L_{opt} in a practical setting by evaluating σ_{prior} from an initial localization and then choosing the corresponding spacing. At $m = 0.99$ in Figure 3a, the optimal spacing L_{opt} increases from approximately 40 nm to over 100 nm when the offset increases from 0 to 50 nm, while under a Gaussian prior in Figure 3c with $\sigma_{\text{prior}} = 50$ nm, L_{opt} reaches only about 90 nm. These differences arise because the prior-based optimization averages over emitter positions, yielding more stable and conservative spacing values suitable for practical use.

To enable a rapid estimation of the optimal spacing L_{opt} based on a known prior uncertainty σ_{prior}, we fitted a fourth-order polynomial model to the data in Figure 3c. The resulting model is visualized in Figure S13. The fitted model is defined as follows:

$$\begin{aligned} L_{\text{opt}}(\sigma_{\text{prior}}, m, \sigma_{\text{illum}}) = {} & 1.30\,\sigma_{\text{illum}}\sqrt{1-m} \\ & + m^{1.991}\big(1.852\,\sigma_{\text{prior}} + 0.08362\,\sigma^2_{\text{prior}} \\ & \quad - 0.003163\,\sigma^3_{\text{prior}} + 3.124\times10^{-5}\,\sigma^4_{\text{prior}}\big). \end{aligned} \tag{14}$$

with a coefficient of determination of $R^2 = 0.98853$. This expression offers a convenient and computationally inexpensive way to approximate L_{opt}, which is useful for real-time decision-making in experimental workflows. An alternative approach is to use the look-up tables (Tables S2-S8). However, the optimal spacing also depends on additional system parameters such as pixel size, wavelength, and numerical aperture, which are not explicitly included in either the polynomial model or the tables. Therefore, the most accurate values of L_{opt} should ideally be obtained by re-running the full optimization procedure for the specific experimental configuration. This is not feasible in real-time due to the computational cost of the Monte Carlo integration, making the polynomial fit and the tables useful proxies for fast estimation. However, by running the code with the specific parameters of a given microscope setup, a custom table of accurate optimal L values can be generated and used for real-time decision-making.

Figure S14 shows that the localization precision $\sigma_{x,y}$ increases as the total photon count θ_I increases, for all values of L. The figure also demonstrates that the effect of photon count on the optimal L is minimal, with the maximum change in L_{opt} being less than 0.0025 nm across all tested modulation contrasts. Furthermore, Figure S14 confirms that the scaling relation $\sigma_{x,y} \propto \theta_I^{-1/2}$ holds even for imperfect modulation contrast. This scaling behavior is consistent with the result shown in Figure S7a for perfect illumination [4]. This shows the photon count will only affect the absolute size

of the information landscape and does not change the shape of the landscape, because the Fisher information in x and y scales with θ_I.

Figure S15a shows that the localization precision $\sigma_{x,y}$ increases as the total background-photon count σ_b increases for all values of L. Figure S15b demonstrates that the effect of background photons on the optimal L is minimal, with the maximum change in L_{opt} being less than 0.0025 nm across all tested modulation contrasts and signal photon counts. Figure S15c confirms that increasing background leads to an increase in localization uncertainty, regardless of the modulation contrast, and that the scaling relation $\sigma_{x,y} \propto \theta_I^{-1/2}$ remains valid even in the presence of background. In the Cramér–Rao lower bound (CRLB) expression, background photons act as an additive noise term, raising the uncertainty but not affecting the spatial derivatives in the Fisher information. In contrast, decreasing modulation contrast directly reduces the derivatives of the intensity profile with respect to position, which lowers the Fisher information, especially at the zero-intensity center where precise localization depends on steep gradients. As a result, while background photons degrade overall precision, they do not diminish the localization advantage provided by the central minimum in MINFLUX, whereas reduced modulation broadens and flattens the intensity pattern, weakening the signal gradient at the center and directly diminishing localization performance.

Figure S16 shows that adding more off-center patterns improves localization precision when the modulation contrast is imperfect ($m < 1$), but not when $m = 1$. As shown in Figure S17a, this is because for $m = 1$ the central pattern dominates the Fisher information, whereas for $m < 1$ it contributes little due to the loss of the zero-intensity point. Figure S17b further shows that at large L, the central pattern is the sole contributor for $m = 1$, while for $m < 1$, all patterns contribute equally but weakly. Figure S16c shows that increasing the number of illumination patterns changes the symmetry of the spatial precision map. For example, using five patterns introduces a fourfold symmetry, while six patterns yield a fivefold symmetry, and so on. This reflects the rotational symmetry imposed by the pattern geometry around the emitter. Figure S17c demonstrates that removing the central pattern reduces precision for $m = 1$, but improves it for $m < 1$, as expected from Figure S16.

Figure S18 compares two photon allocation strategies: pattern normalization (fixed photons per pattern) and global normalization (fixed total photons). The latter, used in MINFLUX [4] and in all other results in this study, allows photon counts to vary based on excitation intensity. As shown in Figure S18a, both approaches yield similar localization precision $\sigma_{x,y}$ as a function of L, and Figures S18b–c show that optimal L values differ by less than 5 nm across modulation contrasts. Figure S18d confirms that the resulting precision maps differ slightly but share the same threefold symmetry. In Figure S19, we compare two metrics for localization precision: the average-based CRLB $\sigma_{x,y}$ and the determinant-based CRLB σ_{det}. While Figure S19a shows that σ_{det} can attain lower minima for high m, Figures S19b and S19c show that both metrics produce nearly identical optimal L values and similar spatial precision maps.

3.1 Iterative MINFLUX

Figure S20 shows the localization precision $\sigma_{x,y}$ as a function of iteration number k for different values of α, using perfect modulation contrast ($m = 1$). The data reproduce the scaling law reported in [5], where the localization precision scales as $\sigma_{x,y} \propto 1/\sqrt{\theta_I}^k$. This relation predicts that, under ideal conditions, localization precision improves exponentially with the number of iterations.

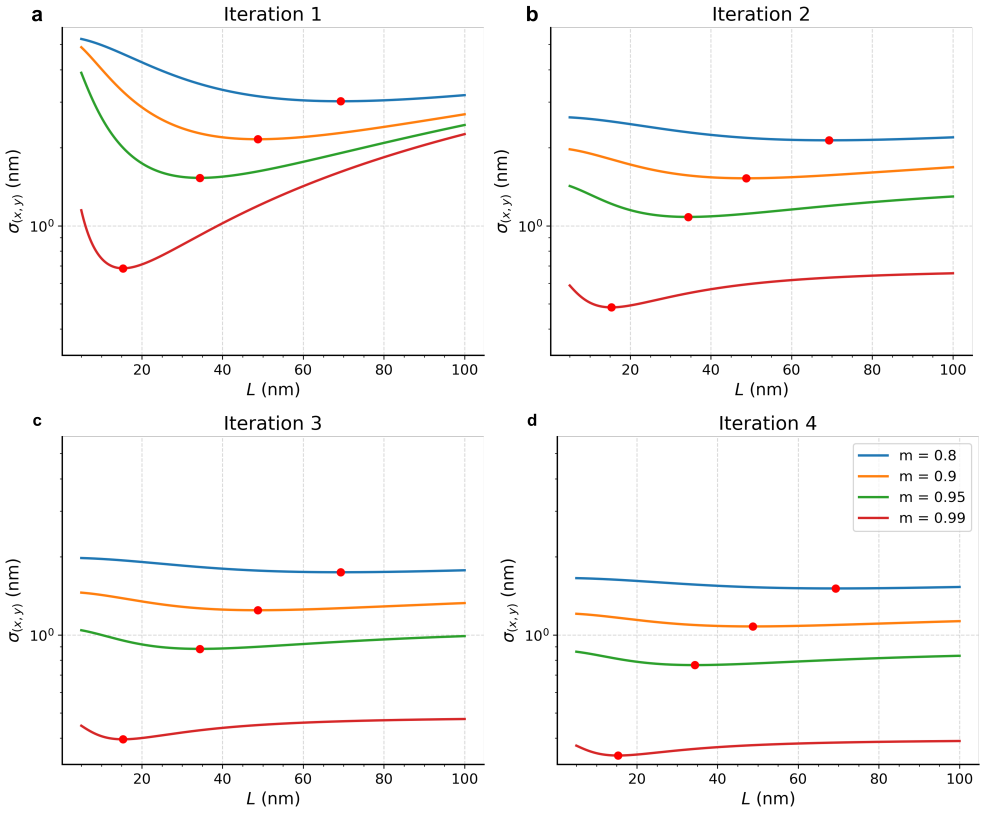


Fig. 4 Optimization of the pattern spacing L over four iterations of the iterative MINFLUX procedure. Each panel (Iteration 1–4) shows $\sigma_{(x,y)}$ versus L for four different modulation contrasts ($m = 0.8, 0.9, 0.95, 0.99$), with red markers indicating the optimal L that minimizes $\sigma_{(x,y)}$. The identical optimal L observed in each iteration demonstrates convergence to the same spacing across successive iterations.

After observing in Figure 2 that the optimal L for imperfect modulation contrast ($m < 1$) is no longer the smallest possible value, we questioned how this would affect localization precision in iterative MINFLUX. If decreasing L no longer always yields better precision, the typical iterative approach of successively reducing L may not be optimal. To investigate this, we optimized L at each iteration, as illustrated in

Figure 4. In this figure, each panel (Iterations 1–4) shows the localization precision $\sigma_{x,y}$ as a function of L. The optimal L remains the same at each iteration, regardless of k. This means that for imperfect modulation, the best result is obtained by repeating the same measurement multiple times with L fixed at its optimal value, rather than by shrinking L over successive iterations. In other words, the optimal iterative MINFLUX procedure converges to performing k identical MINFLUX measurements at the optimal L for that m.

We now compare against the iterative MINFLUX approach [5], where the illumination radius L_k at each iteration k is set according to the estimated localization precision from the previous step using $L_k = \alpha \cdot \sigma_{x,y}^{(k-1)}$, with α defining how aggressively the illumination contracts around the estimated emitter position.

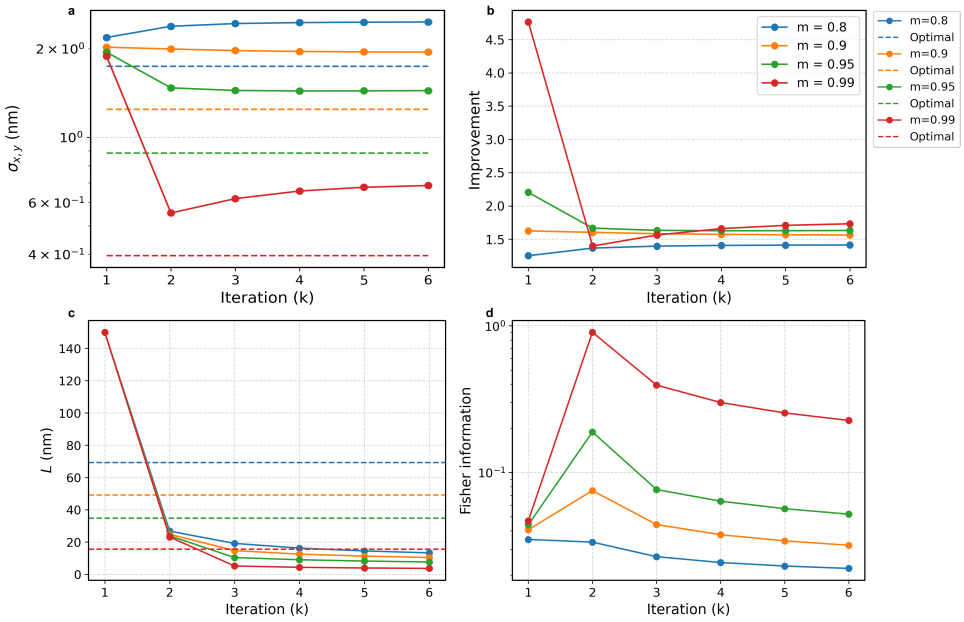


Fig. 5 Iterative MINFLUX simulation with $L_1 = 150$ and $\alpha = 5$ [5]. (a) The localization precision achieved in the simulation for different modulation contrasts m. For reference, the dashed lines indicate the precision attainable with optimal MINFLUX for each m. At iteration 1, the full photon budget is used in a single step; at iteration 2, the photon budget is divided evenly over two iterations, and so forth. (b) The ratio of the localization precision of optimal MINFLUX to that of iterative MINFLUX, quantifying the improvement achieved by MINFLUX. (c) The evolution of the illumination parameter L throughout the iterative simulation, starting from L_1 and updating according to $L_k = \alpha \cdot \sigma_{x,y}^{(k-1)}$ at each iteration. (d) The contribution of each pattern to the total Fisher information $\sqrt{\frac{I_{F_{x,x}}^2 + I_{F_{y,y}}^2}{2}}$ at each iteration. The Fisher information contribution peaks when L is closest to the optimal value found in (c).

Figure 5 presents a direct comparison with iterative MINFLUX [5]. Figure 5a shows that, for varying modulation contrasts, iterative MINFLUX yields worse localization precision compared to the optimal single-step strategy. Figure 5b quantifies this difference by plotting the ratio between the optimal and iterative precision: notably, for $m = 0.95$, the gain exceeds 1.5 (over 50% improvement) ($L_1 = 150$ nm, $\alpha = 5$). Figure 5c reveals how L adapts over iterations, while Figure 5d shows the Fisher information contributions from each pattern at every iteration. The Fisher information peaks when L approaches its optimal value, highlighting that spending more iterations near this optimal L maximizes localization precision. These results emphasize that iterative MINFLUX is intrinsically limited by its sequential adaptation process, especially under imperfect modulation conditions.

The same experiment is done with ($L_1 = 100$ nm and $\alpha = 15$) in Figure S21. This is to see the effect of how the precision will differ when getting closer or farther away from the optimal L, since in Figure 5 it will already past the optimal L after one update.

4 Conclusion

Our analysis demonstrates that optimizing the MINFLUX pattern radius L for a given modulation contrast m can improve localization performance under imperfect modulation. In particular, we find a finite optimal spacing $L_{\text{opt}}(m)$ exists for $m < 1$, in contrast to the ideal $m = 1$ case where decreasing L indefinitely improves precision [4]. At high contrast ($m = 0.95$), the optimal radius is 34 nm, which yields a localization uncertainty 50% lower than that achieved with a fixed small-radius illumination scheme of 5 nm. For example, using a suboptimal fixed spacing of $L = 70$ nm (as in the original MINFLUX implementation [4]) at $m = 0.95$ attains only 79% of the achievable precision, whereas tuning to L_{opt} recovers the full performance. We also established a predictive formula relating the optimal pattern radius to modulation contrast. Fitting the simulation data yields a scaling law for the optimal illumination spacing $L_{\text{opt}} \propto \sqrt{1-m}$, refined to $L_{\text{opt}} \approx 1.30, \sigma_{\text{illum}} \sqrt{, 1-m,}$. Both σ_{illum} and m can be determined experimentally by scanning the excitation pattern over an emitter and recording the resulting intensity. The modulation contrast is then computed as $m = (I_{\max} - I_{\min})/(I_{\max} + I_{\min})$, and σ_{illum} can be estimated from the spacing between the two maxima in the intensity profile. This expression for L_{opt} is valid over the practical range ($0.8 \leq m \leq 0.99$) with a relative error below 0.3%. However, this relation assumes the emitter lies under the central pattern and retains over 90% of the optimal localization precision when the emitter is within 40% of the pattern diameter. For larger emitter offsets, the localization uncertainty increases rapidly, and the analytical prediction no longer holds.

To extend this framework to practical applications where the emitter location is uncertain, we incorporated prior information in the form of a Gaussian position distribution. By modeling the emitter offset as a random variable with standard deviation

σ_{prior}. The resulting optimal spacing $L_{\text{opt}}(\sigma_{\text{prior}})$ increases with increasing prior uncertainty. These results provide a practical strategy for pattern placement, with an estimation σ_{prior} from an initial localization. σ_{prior} can be derived from the measurement variance [18], an CRLB [20], or existing analytical approximations [37–39].

When optimizing Iterative MINFLUX a single iteration of optimized MINFLUX already realizes the best possible precision, outperforming the standard multi-iteration procedure. The optimal illumination radius identified in the first measurement remains the same for subsequent iterations (for a given $m < 1$). In other words, the optimal iterative strategy simply converges to performing k identical exposures at L_{opt} rather than progressively contracting the pattern. Consequently, the conventional iterative MINFLUX approach [5] cannot achieve the same localization precision as our optimized single-step method. At $m = 0.95$, a single optimized MINFLUX measurement provides over a 1.5-fold improvement in precision, $> 50\%$ reduction in localization uncertainty, relative to the iterative scheme. However, iterative MINFLUX remains useful when the prior uncertainty σ_{prior} is large, provided that the illumination radius is never reduced below the optimal L predicted by our model.

These findings are relevant for single-particle tracking, where emitters are localized repeatedly across time. By adapting the illumination radius based on the evolving prior uncertainty, optimal precision can be maintained throughout the trajectory.

Experimental validation under realistic imaging conditions $m < 1$ will be essential to evaluate the practical accuracy of these predictions. Direct comparison between simulations and systematic measurements will determine whether the theoretical trends persist in practice.

Code and Data Availability

The code and Data used for the optimization in this work is available by contacting the authors.

References

[1] Betzig, E., Patterson, G.H., Sougrat, R., *et al.*: Imaging intracellular fluorescent proteins at nanometer resolution. Science **313**(5793), 1642–1645 (2006) https://doi.org/10.1126/science.1127344

[2] Rust, M.J., Bates, M., Zhuang, X.: Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). Nature methods **3**(10), 793–795 (2006) https://doi.org/10.1038/nmeth929

[3] Huang, B., Bates, M., Zhuang, X.: Super-resolution fluorescence microscopy. Annual review of biochemistry **78**, 993–1016 (2009) https://doi.org/10.1146/annurev.biochem.77.061906.092014

[4] Balzarotti, F., Eilers, Y., Gwosch, K.C., Gynna, A.H., Westphal, V., Stefani, F.D., Elf, J., Hell, S.W.: Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. Science **355**(6325), 606–612 (2017) https://doi.org/10.1126/science.aak9913

[5] Gwosch, K.C., Pape, J.K., Balzarotti, F., Hoess, P., Ellenberg, J., Ries, J., Hell, S.W.: Minflux nanoscopy delivers 3d multicolor nanometer resolution in cells. Nature Methods **17**, 217–224 (2019) https://doi.org/10.1038/s41592-019-0688-0

[6] Weber, M., Leutenegger, M., Stoldt, S., Jakobs, S., Mihaila, T.S., Butkevich, A.N., Hell, S.W.: Minsted fluorescence localization and nanoscopy. Nature photonics **15**(5), 361–366 (2021)

[7] Cnossen, J., Hinsdale, T., Thorsen, R., *et al.*: Localization microscopy at doubled precision with patterned illumination. Nature Methods **17**(1), 59–63 (2019) https://doi.org/10.1038/s41592-019-0657-7

[8] Tan, X.-J., Huang, Z.: Minflux nanoscopy enhanced with high-order vortex beams. Light: Science & Applications **14**(1), 1–8 (2025)

[9] Zhao, K., Xu, X., Ren, W., Jin, D., Xi, P.: Two-photon minflux with doubled localization precision. elight **2**(1), 5 (2022)

[10] Scheiderer, L., Emde, H., Hesselink, M., Weber, M., Hell, S.W.: Minsted tracking of single biomolecules. Nature Methods **21**(4), 569–573 (2024)

[11] Jahr, W., Velicky, P., Danzl, J.G.: Strategies to maximize performance in stimulated emission depletion (sted) nanoscopy of biological specimens. Methods **174**, 27–41 (2020) https://doi.org/10.1016/j.ymeth.2019.07.019

[12] Marin, Z., Scheiderer, L., Ries, J.: Simuflux: A realistic simulation framework for minflux microscopy. bioRxiv (2025) https://doi.org/10.1101/2025.04.08.647786 . Preprint

[13] Moosmayer, T., Kiszka, K.A., Westphal, V., Pape, J.K., Leutenegger, M., Steffens, H., Grant, S.G.N., Sahl, S.J., Hell, S.W.: Minflux fluorescence nanoscopy in biological tissue. Proceedings of the National Academy of Sciences **121**(52), 2422020121 (2024) https://doi.org/10.1073/pnas.2422020121

[14] Neupane, B., Chen, F., Sun, W., Chiu, D.T., Wang, G.: Tuning donut profile for spatial resolution in stimulated emission depletion microscopy. Review of Scientific Instruments **84**(4) (2013)

[15] Kalisvaart, D., Cnossen, J., Hung, S.-T., Stallinga, S., Verhaegen, M., Smith, C.S.: Precision in iterative modulation enhanced single-molecule localization microscopy. Biophysical Journal **121**(12), 2279–2289 (2022)

[16] Marin, Z., Ries, J.: Evaluating minflux experimental performance in silico. bioRxiv, 2025–04 (2025)

[17] Vogler, B.T., De Angelis, G., Zhao, Z., Eggeling, C., Reina, F.: Parameter optimization for iterative minflux microscopy enabled single particle tracking. bioRxiv, 2025–06 (2025)

[18] Kay, S.M.: Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, (1993)

[19] Ober, R.J., Ram, S., Ward, E.S.: Localization accuracy in single-molecule microscopy. Biophysical journal **86**(2), 1185–1200 (2004) https://doi.org/10.1016/S0006-3495(04)74193-4

[20] Smith, C.S., Joseph, N., Rieger, B., Lidke, K.A.: Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nature methods **7**(5), 373–375 (2010) https://doi.org/10.1038/nmeth.1449

[21] Kalisvaart, D., Hung, S.-T., Smith, C.S.: Quantifying the minimum localization uncertainty of image scanning localization microscopy. Biophysical Reports **4**(1) (2024)

[22] Hincks, I.: Utility of the Bayesian Cramér–Rao Bound (van Trees inequality). StackExchange Q&A, Jul 25 2017 (2017). https://stats.stackexchange.com/q/294392

[23] Van Trees, H.L.: Detection, Estimation, and Modulation Theory, Part I. Wiley, New York (2001)

[24] Gill, R.D., Levit, B.Y.: Applications of the van trees inequality: a bayesian cramér-rao bound. Bernoulli **1**(1/2), 59–79 (1995) https://doi.org/10.2307/3318681

[25] Tichavský, P., Muravchik, C.H., Nehorai, A.: Posterior cramér–rao bounds for discrete-time nonlinear filtering. In: IEEE Trans. Signal Processing (2010). Describes recursive computation of PCRLB vs. cost of Bayesian bounds

[26] Chen, W.-N., Özgür, A.: L_q Lower Bounds on Distributed Estimation via Fisher Information. arXiv preprint (2024). Notes that VTI yields minimax bounds by integrating over parameter priors

[27] Oh, M.-J., Love, D.J., Krogmeier, J.V., Brinton, C.G.: Dynamic and robust sensor selection strategies for wireless positioning. IEEE Trans. Signal Processing (2023). Uses CRLB-driven selection to optimize localization at known position

[28] Gustafsson, M.G.L.: Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. Journal of Microscopy **198**(1), 82–87 (2000)

[29] Li, Q., Jiang, X., Zhao, T., *et al.*: Polarization-controlled structured illumination microscopy. Light: Science Applications **12**, 29 (2023)

[30] Yang, B., Preibisch, S., Shroff, H.: Sub-diffraction-resolution live-cell imaging with standing-wave total internal reflection fluorescence microscopy. Nature Methods **11**, 579–582 (2014)

[31] Conn, A.R., Scheinberg, K., Vicente, L.N.: Introduction to Derivative-Free Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA (2009)

[32] Nelder, J.A., Mead, R.: A simplex method for function minimization. Computer Journal **7**(4), 308–313 (1965) https://doi.org/10.1093/comjnl/7.4.308

[33] Derman, E., Mannor, S.: Distributional robustness and regularization in reinforcement learning. arXiv preprint arXiv:2003.02894 (2020)

[34] Audet, C., Hare, W., Le Digabel, S., Tribes, C.: Surveys in derivative-free optimization. SIAM Review **63**(1), 1–45 (2021)

[35] Storn, R., Price, K.: Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization **11**(4), 341–359 (1997)

[36] Draper, N.R., Smith, H.: Applied Regression Analysis, 3rd edn. John Wiley & Sons, New York (1998)

[37] Thompson, R.E., Larson, D.R., Webb, W.W.: Precise nanometer localization analysis for individual fluorescent probes. Biophysical journal **82**(5), 2775–2783 (2002)

[38] 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**(5), 377–384 (2010)

[39] Stallinga, S., Rieger, B.: The effect of background on localization uncertainty in single emitter imaging. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 988–991 (2012). https://doi.org/10.1109/ISBI.2012.6235768 . IEEE

[40] Axelrod, D.: Total internal reflection fluorescence microscopy in cell biology. Traffic **2**(11), 764–774 (2001) https://doi.org/10.1034/j.1600-0854.2001.21104.x

[41] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes: The Art of Scientific Computing. Cambridge University Press (2007)

[42] Kroese, D.P., Taimre, T., Botev, Z.I.: Handbook of Monte Carlo Methods. John Wiley & Sons (2011)

[43] MacKay, D.J.C.: Information Theory, Inference, and Learning Algorithms. Cambridge University Press (2003)

[44] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press (2007)

[45] Das, S., Suganthan, P.N.: Differential evolution: A survey of the state-of-the-art. IEEE Transactions on Evolutionary Computation **15**(1), 4–31 (2011)

[46] Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., *et al.*: Scipy 1.0: fundamental algorithms for scientific computing in python. Nature methods **17**(3), 261–272 (2020)

Supplementary Note 1 Image Formation Model

To quantify localization precision in (iterative) MINFLUX, we model the photon detection process on the basis of a parametric image formation model. This framework enables the derivation of the Cramér-Rao Lower Bound (CRLB) precision bound. The parameter vector of the image formation model includes both the x and y coordinates of the emitter:

$$\theta = \begin{bmatrix} \theta_x & \theta_y & \theta_I & \theta_b \end{bmatrix}^T, \tag{S1}$$

where θ_x and θ_y are the position of the emitter in the x- and y-directions, θ_I is the expected photon count emitted by the emitter, and θ_b is the expected background count per unit area.

The image formation model is governed by a 2D psf, $h(x, y)$, which represents the spread of light from a point source in both the x and y directions. The image $g(x, y)$ of the object $f(x, y)$ is formed through a 2D convolution with the psf:

$$g(x, y) = h(x, y) * f(x, y), \tag{S2}$$

where $*$ denotes the two-dimensional convolution operator. The conservation of energy condition requires that the total area under the psf equals one:

$$\int\int_{\mathbb{R}^2} h(x, y)dxdy = 1. \tag{S3}$$

The object function $f(x, y)$ is then given by:

$$f(x, y) = P(x, y)\left(\theta_I \delta(x - \theta_x, y - \theta_y) + \theta_b\right), \tag{S4}$$

where $P(x, y)$ is the illumination pattern, and $\delta(x - \theta_x, y - \theta_y)$ represents a point emitter located at (θ_x, θ_y). This function incorporates both the signal from the emitter and the background noise. For the model to be valid, the illumination pattern $P(x, y)$ must be non-negative, accurately measurable, and constant during the imaging process to ensure reliable parameterization and reconstruction of the emitter's position.

To find an expression for the expected photon count on a pixel i, we integrate the image function $g(x, y)$ over the pixel area A_i, which has center coordinates (x_i, y_i). The result is:

$$\mu_i(x_i, y_i) = \theta_I P(\theta_x, \theta_y) \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} \int_{y_i - \frac{\Delta y}{2}}^{y_i + \frac{\Delta y}{2}} h(x - \theta_x, y - \theta_y)dxdy + \theta_b B_i, \tag{S5}$$

where B_i is given by:

$$B_i = \int_{x_i - \frac{\Delta x}{2}}^{x_i + \frac{\Delta x}{2}} \int_{y_i - \frac{\Delta y}{2}}^{y_i + \frac{\Delta y}{2}} \left(\int\int_{\mathbb{R}^2} h(\tau, \gamma) P(x - \tau, y - \gamma) d\tau d\gamma \right) dxdy. \tag{S6}$$

Since the experiments are assumed to be performed under total internal reflection fluorescence (TIRF) illumination, the background is considered spatially and pattern

independent. Therefore, we model the background contribution as a constant term across all patterns, i.e., $B_i = 1$ for all i [40].

Assuming a Gaussian psf model:

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{S7}$$

the expected photon count μ_i can be decomposed into the product of two separate integrals, E_x and E_y, which represent the contributions of the psf in the x and y directions, respectively:

$$\mu_i(x_i, y_i) = \theta_I P(\theta_x, \theta_y) E_x(x_i, \theta_x) E_y(y_i, \theta_y) + \theta_b B_i. \tag{S8}$$

The integrals $E_x(x_i, \theta_x)$ and $E_y(y_i, \theta_y)$ can be expressed in terms of the error function erf:

$$E_x(x_i, \theta_x) = \frac{1}{2}\left[\text{erf}\left(\frac{x_i - \theta_x + \frac{\Delta x}{2}}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{x_i - \theta_x - \frac{\Delta x}{2}}{\sigma\sqrt{2}}\right)\right], \tag{S9}$$

$$E_y(y_i, \theta_y) = \frac{1}{2}\left[\text{erf}\left(\frac{y_i - \theta_y + \frac{\Delta y}{2}}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{y_i - \theta_y - \frac{\Delta y}{2}}{\sigma\sqrt{2}}\right)\right]. \tag{S10}$$

The error function $\text{erf}(z)$ is defined as:

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt, \tag{S11}$$

where $\text{erf}(z)$ represents the probability of a random variable with a normal distribution and zero mean falling within the range $[-z, z]$.

These integrals take into account the finite pixel size and provide a more accurate model for the distribution of photon counts over the pixel area.

Substituting the expressions for E_x and E_y into the equation for μ_i:

$$\mu_i(x_i, y_i) = \theta_I P(\theta_x, \theta_y) E_x(x_i, \theta_x) E_y(y_i, \theta_y) + \theta_b B_i. \tag{S12}$$

Donut-shaped illumination introduces spatial modulation with a zero-intensity minimum at the center. The illumination pattern in 2D is defined as:

$$P_{\text{donut}}(x, y) = e\frac{(x - x_p)^2 + (y - y_p)^2}{2\sigma_{\text{illum}}^2} \exp\left(-\frac{(x - x_p)^2 + (y - y_p)^2}{2\sigma_{\text{illum}}^2}\right), \tag{S13}$$

where σ_{illum} is the standard deviation describing the spatial spread of the Gaussian envelope, and (x_p, y_p) is the center of the donut pattern. The quadratic term $\left[(x - x_p)^2 + (y - y_p)^2\right]$ creates the characteristic zero-intensity minimum at the center of the illumination.

The locations of x_p and y_p for iterative MINFLUX are dependent on the location estimate of the previous iteration. For our simulation with the MINFLUX and iterative MINFLUX case $\theta_{x,k-1}$ and $\theta_{y,k-1}$ are 0. They are defined in the table below:

Name	**Location on the circle with diameter L**	x_p	y_p
$P_{center}(x,y)$	center	$\theta_{x,k-1}$	$\theta_{y,k-1}$
$P_{\phi=0}(x,y)$	$\phi = 0$	$\theta_{x,k-1} + \frac{1}{2}L$	$\theta_{y,k-1}$
$P_{\phi=\frac{2}{3}\pi}(x,y)$	$\phi = \frac{2}{3}\pi$	$\theta_{x,k-1} - \frac{1}{4}L$	$\theta_{y,k-1} + \frac{\sqrt{3}}{4}L$
$P_{\phi=\frac{4}{3}\pi}(x,y)$	$\phi = \frac{4}{3}\pi$	$\theta_{x,k-1} - \frac{1}{4}L$	$\theta_{y,k-1} - \frac{\sqrt{3}}{4}L$

Supplementary Note 2 Cramér–Rao Lower Bound (CRLB)

To quantify the achievable localization precision for an unbiased estimator in MINFLUX and its iterative variants, we compute the Cramér–Rao Lower Bound (CRLB). The CRLB provides a lower bound on the variance of any unbiased estimator, given the statistical model of the imaging process. It is derived from the Fisher information matrix (FIM), which quantifies the sensitivity of the expected photon counts with respect to the parameters of interest.

The expected photon count at camera pixel (x_i, y_i) under an illumination pattern is given by:

$$\mu_i(x_i, y_i, \theta) = \theta_I P(\theta_x, \theta_y) E_x(x_i, \theta_x) E_y(y_i, \theta_y) + \theta_b, \tag{S14}$$

where $\theta = [\theta_x, \theta_y, \theta_I, \theta_b]^T$, $P(\theta_x, \theta_y)$ is the illumination intensity at the emitter location, and $E_x(x_i, \theta_x)$, $E_y(y_i, \theta_y)$ are the integrals of the Gaussian PSF over the pixel area in the x and y directions, respectively.

To compute the Fisher Information Matrix (FIM), we use the definition:

$$I_F(\theta)_{p,q} = \mathbb{E}\left[\frac{\partial \log L(\theta)}{\partial \theta_p} \cdot \frac{\partial \log L(\theta)}{\partial \theta_q}\right], \tag{S15}$$

where $p, q \in \{1, \ldots, d\}$ index the parameters in θ, and the expectation is taken over the distribution of the observed data.

In our imaging context, each camera pixel detects photons following a Poisson distribution. The probability of observing g_i photons at pixel i, given expected photon count $\mu_i(\theta)$, is:

$$P(g_i \mid \theta) = \frac{\mu_i(\theta)^{g_i} e^{-\mu_i(\theta)}}{g_i!}. \tag{S16}$$

The log-likelihood of observing a full image $g = (g_1, g_2, \ldots, g_N)$ is then:

$$\log L(\theta) = \sum_{i=1}^{N} \left[g_i \log \mu_i(\theta) - \mu_i(\theta) - \log(g_i!)\right]. \tag{S17}$$

Taking the partial derivative with respect to parameter θ_p, we get:

$$\frac{\partial \log L(\theta)}{\partial \theta_p} = \sum_{i=1}^{N} \left(\frac{g_i}{\mu_i(\theta)} - 1\right) \frac{\partial \mu_i(\theta)}{\partial \theta_p}. \tag{S18}$$

Now applying the expectation over all possible observations (i.e., over the Poisson distribution of g_i), we use the identity:

$$\mathbb{E}[g_i] = \mu_i(\theta),$$

so that:

$$\mathbb{E}\left[\left(\frac{g_i}{\mu_i(\theta)} - 1\right)\right] = 0.$$

However, when computing the expected product of derivatives for the FIM, we get:

$$I_F(\theta)_{p,q} = \mathbb{E}\left[\sum_{i=1}^{N}\left(\frac{g_i}{\mu_i(\theta)} - 1\right)\frac{\partial\mu_i(\theta)}{\partial\theta_p} \cdot \sum_{j=1}^{N}\left(\frac{g_j}{\mu_j(\theta)} - 1\right)\frac{\partial\mu_j(\theta)}{\partial\theta_q}\right] \tag{S19}$$

$$= \sum_{i=1}^{N}\mathbb{E}\left[\left(\frac{g_i}{\mu_i(\theta)} - 1\right)^2\right]\frac{\partial\mu_i(\theta)}{\partial\theta_p}\frac{\partial\mu_i(\theta)}{\partial\theta_q}, \tag{S20}$$

since g_i and g_j are independent for $i \neq j$, and the cross terms vanish. For a Poisson distribution:

$$\mathbb{E}\left[\left(\frac{g_i}{\mu_i(\theta)} - 1\right)^2\right] = \frac{1}{\mu_i(\theta)}.$$

This gives the final result [23]:

$$I_F(\theta)_{p,q} = \sum_{i=1}^{N}\frac{1}{\mu_i(\theta)}\frac{\partial\mu_i(\theta)}{\partial\theta_p}\frac{\partial\mu_i(\theta)}{\partial\theta_q}, \tag{S21}$$

which is the expression used to compute the Fisher Information Matrix in localization microscopy models such as MINFLUX.

The CRLB then provides a bound on the covariance matrix of any unbiased estimator $\hat{\theta}$:

$$\mathrm{Cov}(\hat{\theta}) \succeq I_F(\theta)^{-1}, \tag{S22}$$

from which the localization precision is computed as:

$$\sigma_x = \sqrt{[I_F(\theta)^{-1}]_{x,x}}, \quad \sigma_y = \sqrt{[I_F(\theta)^{-1}]_{y,y}}. \tag{S23}$$

We now provide the derivatives required for computing the FIM [21].

Derivatives of $\boldsymbol{\mu_i(x_i, y_i, \theta)}$

With respect to θ_x:

$$\frac{\partial\mu_i(x_i, y_i, \theta)}{\partial\theta_x} = \theta_I\frac{\partial P(\theta_x, \theta_y)}{\partial\theta_x}E_x(x_i, \theta_x)E_y(y_i, \theta_y) + \theta_I P(\theta_x, \theta_y)\frac{\partial E_x(x_i, \theta_x)}{\partial\theta_x}E_y(y_i, \theta_y) \tag{S24}$$

Derivative of $P(\theta_x, \theta_y)$ with respect to θ_x:

$$\frac{\partial P(\theta_x,\theta_y)}{\partial \theta_x} = e\left[\frac{\theta_x - x_p}{\sigma^2_{\text{illum}}} - \frac{(\theta_x - x_p)^2 + (\theta_y - y_p)^2}{2\sigma^4_{\text{illum}}}(\theta_x - x_p)\right]\exp\left(-\frac{(\theta_x - x_p)^2 + (\theta_y - y_p)^2}{2\sigma^2_{\text{illum}}}\right) \tag{S25}$$

Derivative of $E_x(x_i, \theta_x)$:

$$\frac{\partial E_x(x_i,\theta_x)}{\partial \theta_x} = \frac{1}{\sqrt{\pi}\sigma\sqrt{2}}\left[\exp\left(-\left(\frac{x_i - \theta_x - \frac{\Delta x}{2}}{\sigma\sqrt{2}}\right)^2\right) - \exp\left(-\left(\frac{x_i - \theta_x + \frac{\Delta x}{2}}{\sigma\sqrt{2}}\right)^2\right)\right] \tag{S26}$$

With respect to θ_y:

$$\frac{\partial \mu_i(x_i,y_i,\theta)}{\partial \theta_y} = \theta_I \frac{\partial P(\theta_x,\theta_y)}{\partial \theta_y} E_x(x_i,\theta_x)E_y(y_i,\theta_y) + \theta_I P(\theta_x,\theta_y)E_x(x_i,\theta_x)\frac{\partial E_y(y_i,\theta_y)}{\partial \theta_y} \tag{S27}$$

Derivative of $P(\theta_x, \theta_y)$ with respect to θ_y:

$$\frac{\partial P(\theta_x,\theta_y)}{\partial \theta_y} = e\left[\frac{\theta_y - y_p}{\sigma^2_{\text{illum}}} - \frac{(\theta_x - x_p)^2 + (\theta_y - y_p)^2}{2\sigma^4_{\text{illum}}}(\theta_y - y_p)\right]\exp\left(-\frac{(\theta_x - x_p)^2 + (\theta_y - y_p)^2}{2\sigma^2_{\text{illum}}}\right) \tag{S28}$$

Derivative of $E_y(y_i, \theta_y)$:

$$\frac{\partial E_y(y_i,\theta_y)}{\partial \theta_y} = \frac{1}{\sqrt{\pi}\sigma\sqrt{2}}\left[\exp\left(-\left(\frac{y_i - \theta_y - \frac{\Delta y}{2}}{\sigma\sqrt{2}}\right)^2\right) - \exp\left(-\left(\frac{y_i - \theta_y + \frac{\Delta y}{2}}{\sigma\sqrt{2}}\right)^2\right)\right] \tag{S29}$$

With respect to θ_I:

$$\frac{\partial \mu_i(x_i,y_i,\theta)}{\partial \theta_I} = P(\theta_x,\theta_y)E_x(x_i,\theta_x)E_y(y_i,\theta_y) \tag{S30}$$

With respect to θ_b:

$$\frac{\partial \mu_i(x_i,y_i,\theta)}{\partial \theta_b} = B_i = 1 \tag{S31}$$

These expressions can be substituted into Equation S21 to evaluate the Fisher information matrix. Inverting this matrix yields the Cramér–Rao lower bounds on the localization precision for each parameter.

Supplementary Note 3 MINFLUX Optimization

To enforce the constraint $L > 0$, a penalty term M is added to the objective function. This ensures that the optimizer avoids unphysical solutions where the illumination pattern would be reflected around the emitter due to negative spacing. The penalty is applied only outside the feasible region and does not affect the optimization landscape within the valid domain. The resulting cost function is defined as:

$$\min_{L\in\mathbb{R}} \sigma_{x,y}(L), \quad \text{with} \quad \sigma_{x,y}(L) = \sqrt{\frac{\sigma_x(L)^2 + \sigma_y(L)^2}{2}} + M. \tag{S32}$$

The penalty term M is given by:

$$M = \begin{cases} 0 & \text{if } L > 0, \\ 10^6 & \text{if } L \leq 0. \end{cases} \tag{S33}$$

This approach guarantees that the optimization remains numerically stable and does not interfere with the optimal solution when L remains in the valid physical range.

The Nelder–Mead algorithm is a direct search method used for unconstrained optimization of nonlinear functions [32]. It is derivative-free and operates solely based on function evaluations [41]. The algorithm maintains a simplex, which in one dimension is simply a line segment defined by two points. At each iteration, the algorithm evaluates the objective function at the vertices of the simplex and replaces one of the points based on a set of geometric operations: reflection, expansion, contraction, or shrinkage [41]. These operations are used to explore the domain and move the simplex toward a region with lower function values [41].

The algorithm begins from an initial guess x_0, which determines the starting position of the simplex. In one dimension, this simplex consists of x_0 and a second point offset from it. The choice of $x_0 = 50\,\text{nm}$ is based on the approximate location of the minimum as observed in Supplementary Figure S2.

Algorithm 1 Nelder–Mead Algorithm (1D Case)

Input: Initial point x_0, objective function $f(x)$, tolerance `xatol`
Initialize second point $x_1 = x_0 + \delta$, for small $\delta > 0$
Let x_l, x_h be the points with $f(x_l) \leq f(x_h)$
while $|x_h - x_l| >$ `xatol` **do**
    Reflection: $x_r = x_l + \alpha(x_l - x_h)$
    if $f(x_r) < f(x_l)$ **then**
        Expansion: $x_e = x_l + \gamma(x_r - x_l)$
        Replace $x_h \leftarrow x_e$ if $f(x_e) < f(x_r)$, else $x_h \leftarrow x_r$
    else if $f(x_r) < f(x_h)$ **then**
        Replace $x_h \leftarrow x_r$
    else
        Contraction: $x_c = x_l + \rho(x_h - x_l)$
        if $f(x_c) < f(x_h)$ **then**
            Replace $x_h \leftarrow x_c$
        else
            Shrink: $x_h \leftarrow x_l + \sigma(x_h - x_l)$
        end if
    end if
    Reorder: update x_l, x_h so that $f(x_l) \leq f(x_h)$
end while
Output: Approximate minimum x_l

The Nelder–Mead algorithm is used with fixed parameters: reflection coefficient $\alpha = 1$, expansion coefficient $\gamma = 2$, contraction coefficient $\rho = 0.5$, and shrinkage coefficient $\sigma = 0.5$ [41]. The stopping criterion is based solely on the simplex width, which must fall below a threshold `xatol` $= 10^{-6}$ nm [41].

Supplementary Note 4 Prior Location MINFLUX Optimization

The offset $\boldsymbol{\delta} = (x_{\text{offset}}, y_{\text{offset}})$ between the emitter and the illumination pattern is unknown. If it were known, the experimenter would not need to measure the position of the parameter. We now model the emitter position as a random variable drawn from a 2D Gaussian distribution centered at the illumination pattern origin.

$$p_{\sigma_{\text{prior}}}(\boldsymbol{\delta}) = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \sigma^2_{\text{prior}} & 0 \\ 0 & \sigma^2_{\text{prior}} \end{bmatrix}\right), \tag{S34}$$

Given this probabilistic model, we aim to select the pattern spacing L that minimizes the expected localization uncertainty over the distribution of emitter positions. Let $\sigma_{xy}(L, \boldsymbol{\delta})$ denote the achievable localization precision for a given spacing L and emitter offset $\boldsymbol{\delta}$. The expected localization precision is then defined as:

$$\mathbb{E}[\sigma_{x,y}(L)] = \iint_{\mathbb{R}^2} \sigma_{x,y}(L, \boldsymbol{\delta})\, p_{\sigma_{\text{prior}}}(\boldsymbol{\delta})\, d\boldsymbol{\delta}, \tag{S35}$$

where $p_{\sigma_{\text{prior}}}(\boldsymbol{\delta})$ is a 2D Gaussian distribution as defined in Equation (S34). This expression represents a Bayesian average over the prior uncertainty in emitter location, where $\sigma_{x,y}(L, \boldsymbol{\delta})$ denotes the CRLB-based localization uncertainty for a given offset and pattern spacing.

To compute this expectation numerically, we use a Monte Carlo approximation:

$$\mathbb{E}[\sigma_{x,y}(L)] \approx \frac{1}{N} \sum_{i=1}^{N} \sigma_{x,y}(L, \boldsymbol{\delta}_i), \tag{S36}$$

where $\boldsymbol{\delta}_i$ are random samples drawn from the prior $p_{\sigma_{\text{prior}}}(\boldsymbol{\delta})$. This method is justified by the law of large numbers, and its error converges as $\mathcal{O}(1/\sqrt{N})$, independent of the problem's dimensionality [42, 43].

To improve accuracy and reduce variance, we use stratified sampling rather than random sampling. In this approach, each axis is divided into N equal-probability intervals, and one sample is drawn from each, leading to a more uniform and representative coverage of the 2D Gaussian distribution [44]. This makes the resulting estimate smoother and more suitable for optimization. Specifically, we generate fixed stratified samples once per σ_{prior}, and reuse them across all evaluations of L. This allows the optimizer to reliably minimize the expected precision without introducing sampling noise into the objective function. This technique is known as sample average approximation.

The optimal pattern spacing is then given by:

$$L_{\text{opt}} = \arg\min_{L} \; \mathbb{E}[\sigma_{x,y}(L)] \tag{S37}$$

which defines the value of L that minimizes the expected localization error across the prior distribution of emitter positions.

Figure S6 shows that the objective landscape is nearly convex, except for small values of L, where some non-convexity appears. Nonetheless, the function exhibits a well-defined single optimum. A smooth landscape with a single well-defined minimum only emerges when the number of prior samples is sufficiently large. In our implementation, using 500 stratified 2D samples was necessary. Owing to the unimodal shape and lack of derivative information, we again use the Nelder–Mead algorithm to perform the optimization. The same algorithm from Supplementary Note 3 is used.

Supplementary Note 5 Iterative MINFLUX Optimization

As in the non-iterative case, a penalty term M is added to enforce the constraint $L_i > 0$ for all illumination spacings L_i in the sequence. The same justification applies, and the penalty does not affect the optimization within the feasible domain. The cost function becomes:

$$\min_{L_1,\ldots,L_k\in\mathbb{R}} \sigma_{x,y}(L_1,\ldots,L_k), \quad \text{with} \quad \sigma_{x,y}(L_1,\ldots,L_k) = \sqrt{\frac{\sigma_x(L_1,\ldots,L_k)^2 + \sigma_y(L_1,\ldots,L_k)^2}{2}} + M. \tag{S38}$$

where

$$M = \begin{cases} 0 & \text{if } L_i > 0 \text{ for all } i, \\ 10^6 & \text{otherwise.} \end{cases} \tag{S39}$$

In the multidimensional case, the Nelder–Mead algorithm generalizes the simplex from a line segment to a $(d+1)$-vertex polytope in d dimensions. For example, in k-dimensional iterative MINFLUX optimization, the simplex consists of $k+1$ distinct vectors $\mathbf{L}_i \in \mathbb{R}^k$. Each vector encodes a complete set of illumination spacings $(L_1,\ldots,L_k)$, and all geometric operations (reflection, expansion, contraction, and shrinkage) are applied to these vectors as a whole [41].

The same update rules apply component-wise, and the algorithm evaluates the cost function $\sigma_{x,y}(L_1,\ldots,L_k)$ at each vertex. The stopping criterion is again based on the diameter of the simplex, i.e., the maximum distance between any pair of vertices, which must fall below the threshold `xatol`. All parameter values and logic are preserved from the 1D case.

Supplementary Note 5.1 Differential Evolution

To verify that the local optimization converges to the global optimum, we compare it against the Differential Evolution (DE) algorithm [35]. DE is a population-based global optimization method suitable for nonlinear and non-differentiable functions. It maintains a population of candidate solutions and iteratively improves them using mutation, crossover, and selection operations.

At each generation, a new candidate is generated by adding the scaled difference between two population vectors to a third, and potentially combining it with the current candidate through crossover. If the resulting trial solution yields a lower objective function value, it replaces the current candidate. This mechanism promotes global exploration and avoids premature convergence [45].

In this work, the DE algorithm is implemented via the `differential_evolution` function from the SciPy library [46]. The `strategy` was set to `'best1bin'`, where the best solution in the population serves as the base vector for mutation. A population size of `popsize = 10` was used to ensure adequate exploration. The convergence tolerance `tol` was set to 10^{-6}, and the mutation factor was sampled uniformly from the interval $(0.5, 1.0)$. The crossover probability, given by `recombination = 0.7`, controlled the proportion of parameters inherited from the mutant vector. These settings provide a

balance between global exploration and precise convergence, making DE a suitable baseline for validating local optimization results.

Supplementary Note 6 Simulation Parameters

The simulations in this work were performed using a consistent set of imaging and system parameters, as listed in Table S1. These parameters were selected to match typical conditions in high-resolution MINFLUX experiments and to isolate the effects of modulation contrast and illumination pattern placement. All simulations assumed a total photon budget of 1000 photons per measurement, with no background signal. The emitter was placed at the origin, and the central illumination pattern was offset slightly to avoid singularities in the CRLB calculation.

Table S1 Simulation parameters used in all results.

Parameter	**Value**
Wavelength λ	640 nm
Numerical Aperture (NA)	1.4
PSF width σ_{psf}	$0.21 \cdot \lambda/\text{NA} = 96$ nm [37]
Donut peak-to-peak spacing	$0.78 \cdot \lambda/\text{NA} \approx 357$ nm [14]
Central pattern offset (x, y)	0.001 nm, 0.001 nm
Pixel size (object space)	65 nm
Pixel size (image space)	6.5 μm
Magnification	100
Image size	16×16 pixels
Emitter position (x, y, z)	0 nm, 0 nm, 0 nm
Initial ring radius L_0	150 nm
Total photon budget	1000 photons
Background intensity	0 photons/pixel
Initial guess (L_1)	50 nm
L-tolerance	10^{-6}
Penalty M for $L < 0$	10^{6}

Supplementary Note 7 Figures

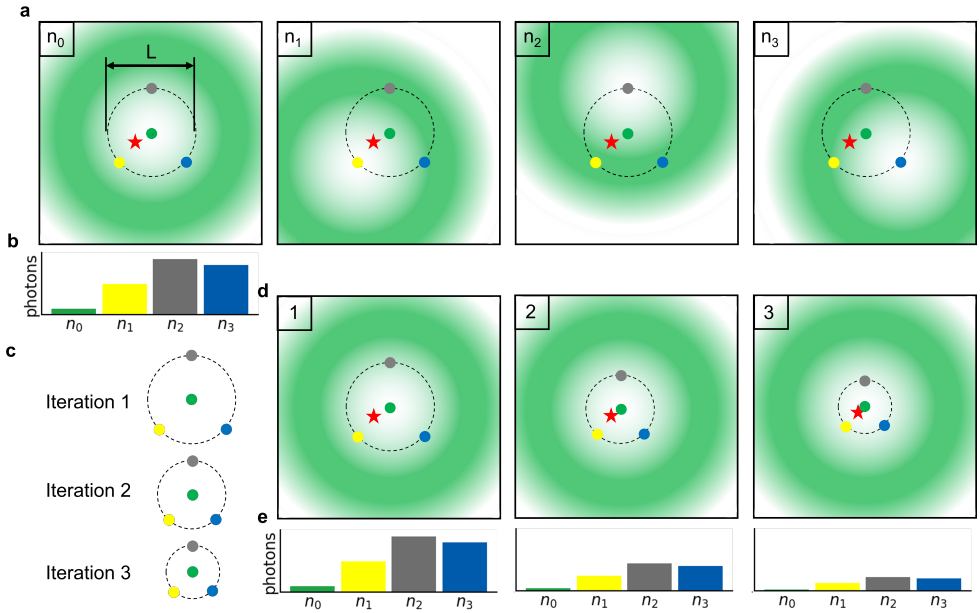


Fig. S1 Explanation of MINFLUX and iterative MINFLUX; image heavily inspired by [4] and [5]. **(a)** Four excitation patterns n_0–n_3 around an unknown emitter (red star), with current estimate (green dot) and four colored null-centers (gray, blue, yellow, green) on a dashed ring of diameter L over a green-white gradient; only bottom/left spines are shown, top/right hidden. **(b)** Photon-count bar plot for n_0–n_3 (bars colored green, yellow, gray, blue), with x-axis labels n_0–n_3 and y-axis labeled "photons" , bottom/left spines visible. **(c)** Iterative refinement schematic: the dashed ring shrinks toward the red star, and the green-dot estimate converges . **(d)** Iterations 1–3, each showing a smaller dashed ring and updated green estimate relative to the red star; spines as in (a). **(e)** Photon-count bar plots for each iteration, with four colored bars (n_0–n_3).

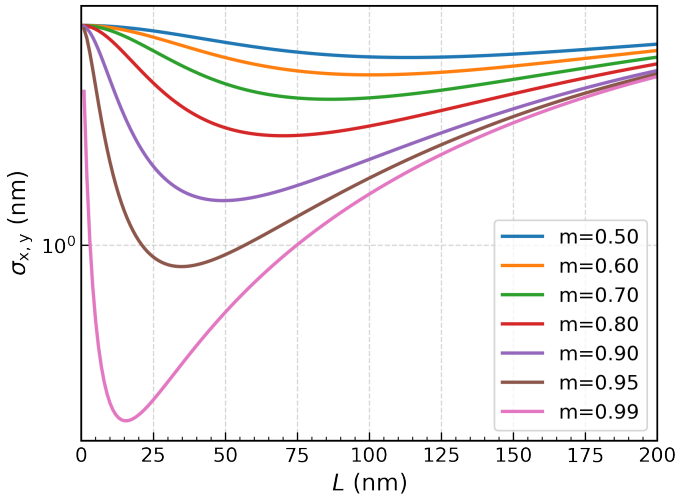


Fig. S2 Localization precision $\sigma_{x,y}$ as a function L for various modulation contrasts m. Each curve represents the optimization landscape for a specific m, showing how precision improves with increasing L up to an optimal value, beyond which performance degrades.

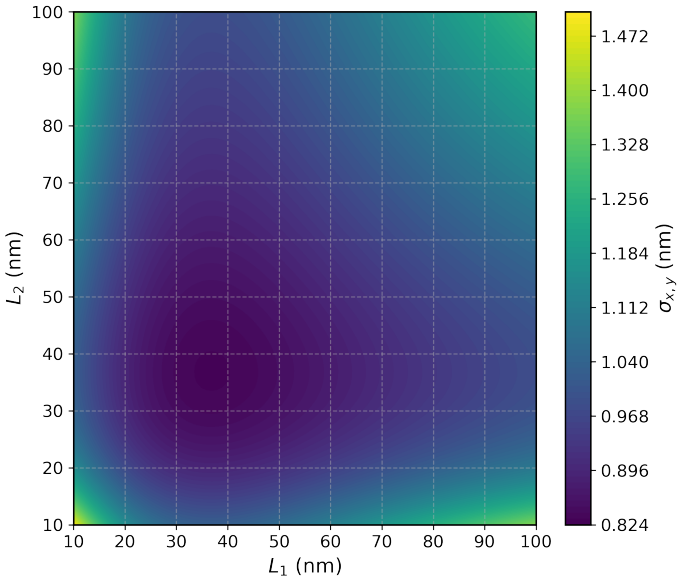


Fig. S3 Localization precision $\sigma_{x,y}$ as a function of two illumination pattern spacings L_1 and L_2. The colormap indicates the value of $\sigma_{x,y}$, with lower values shown in darker shades. The grid covers the range $L_1, L_2 \in [10, 100]$ nm.

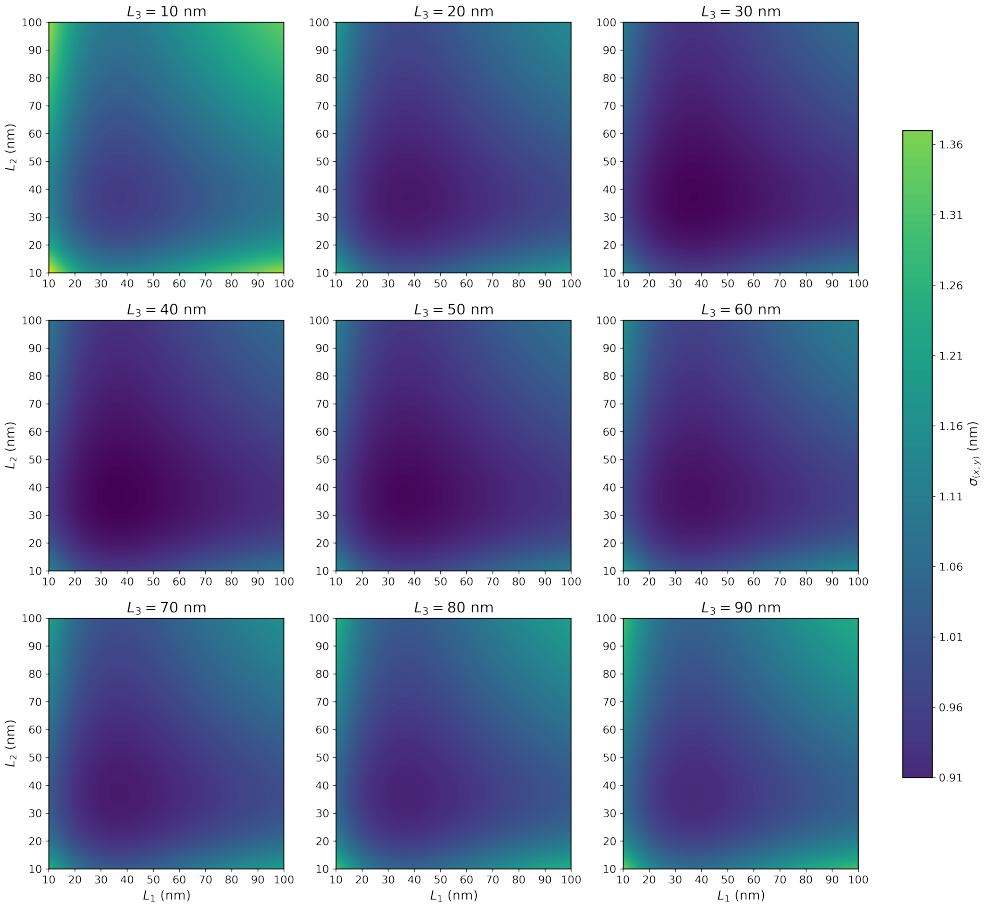


Fig. S4 Localization precision $\sigma_{x,y}$ as a function of two spacing parameters L_1 and L_2 across varying values of a third parameter L_3. Each panel corresponds to a different fixed value of L_3, showing the structure of the objective function in a 2D subspace for each setting.

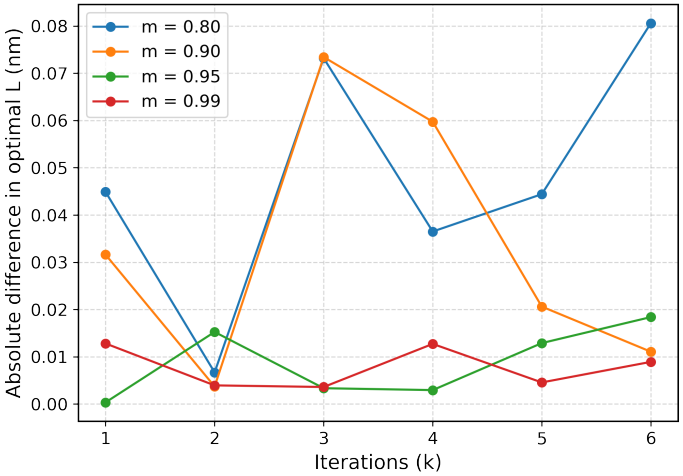


Fig. S5 Absolute difference in optimal illumination spacings L_k between local (Nelder–Mead) and global (Differential Evolution) optimization methods, plotted across iteration index k for various modulation contrasts $m \in \{0.80, 0.90, 0.95, 0.99\}$. The small differences across all k suggest that the optimization landscape is relatively well-behaved, and that local optimization is sufficient to reach near-global solutions in practice.

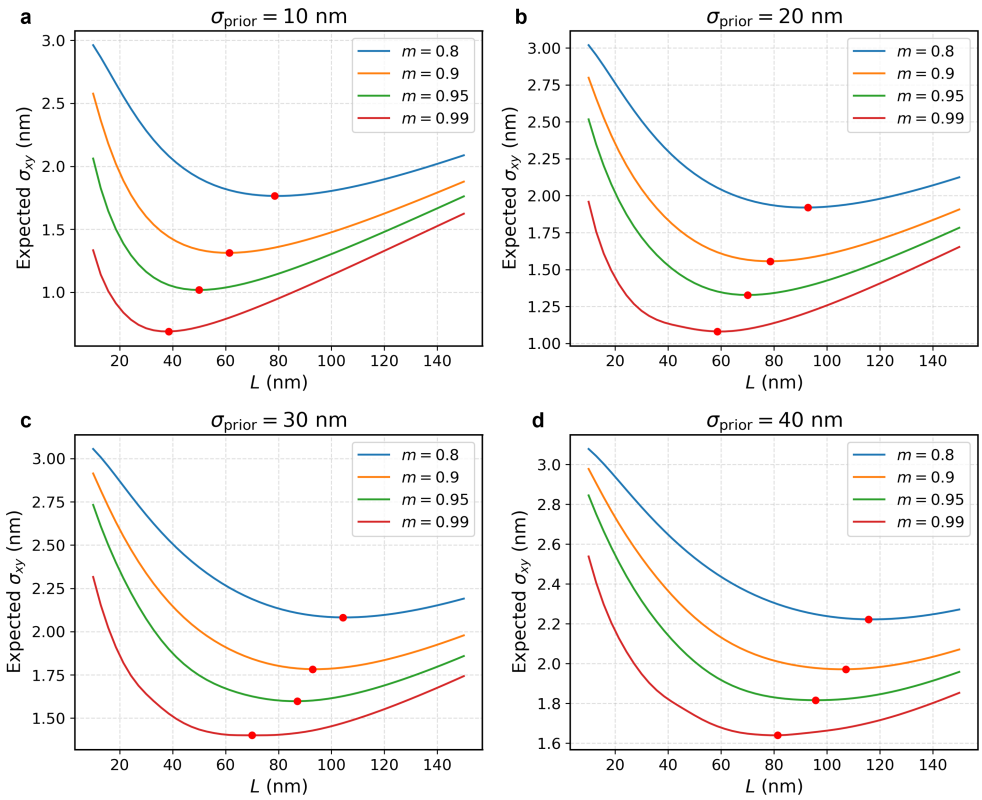


Fig. S6 Expected localization precision $\mathbb{E}[\sigma_{xy}]$ as a function of illumination spacing L for different modulation contrasts $m \in \{0.8, 0.9, 0.95, 0.99\}$ and prior standard deviations $\sigma_{\text{prior}} \in \{10, 20, 30, 40\}$ nm (top-left to bottom-right). Each curve corresponds to a different m, and the red dots indicate the optimal L that minimizes the expected precision. The convex shape of the curves illustrates that the objective function is well-behaved and amenable to numerical optimization.

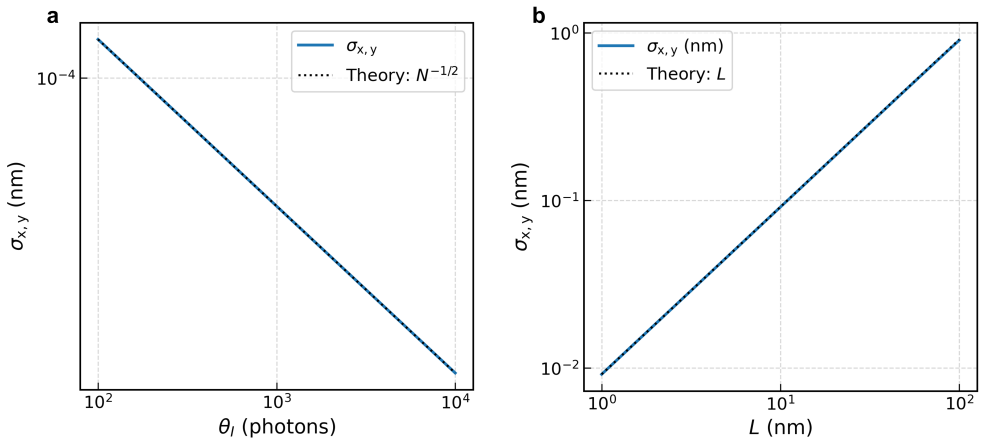


Fig. S7 **(a)** Theoretical scaling of localization precision $\sigma_{x,y}$ as a function of detected photon number N, demonstrating the expected $N^{-1/2}$ behavior for perfect modulation contrast ($m = 1$) at fixed $L = 200\,\text{nm}$. **(b)** Theoretical scaling of localization precision $\sigma_{x,y}$ as a function of L, illustrating the linear dependence $\sigma_{x,y} \propto L$ for perfect modulation contrast ($m = 1$) at fixed photon budget $N = 1000$, as predicted by [4].

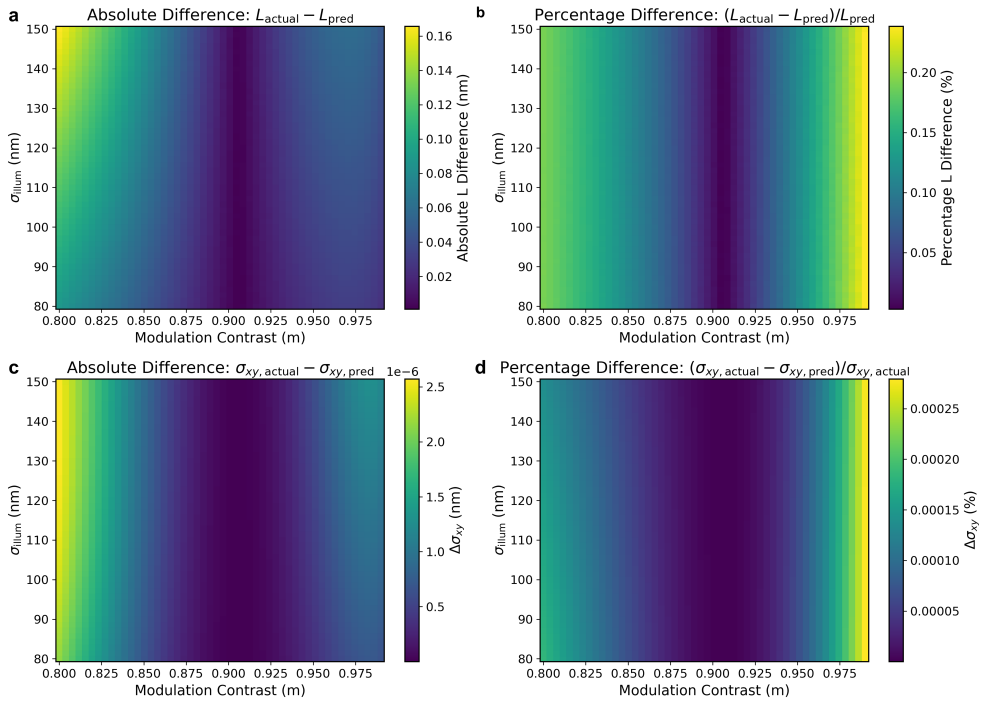


Fig. S8 Comparison between predicted and actual optimal parameters across the illumination domain defined by modulation contrast m and illumination width σ_{illum}. The predicted optimal spacing is given by $L_{\mathrm{opt}} = 1.30\,\sigma_{\mathrm{illum}}\sqrt{1-m}$. **(a)** Absolute difference in optimal spacing: $L_{\mathrm{actual}} - L_{\mathrm{pred}}$. **(b)** Percentage difference in optimal spacing: $(L_{\mathrm{actual}} - L_{\mathrm{pred}})/L_{\mathrm{pred}}$. **(c)** Absolute difference in localization precision: $\sigma_{xy,\mathrm{actual}} - \sigma_{xy,\mathrm{pred}}$. **(d)** Percentage difference in localization precision: $(\sigma_{xy,\mathrm{actual}} - \sigma_{xy,\mathrm{pred}})/\sigma_{xy,\mathrm{actual}}$. Color indicates the magnitude of each difference, highlighting systematic prediction errors over the parameter space.

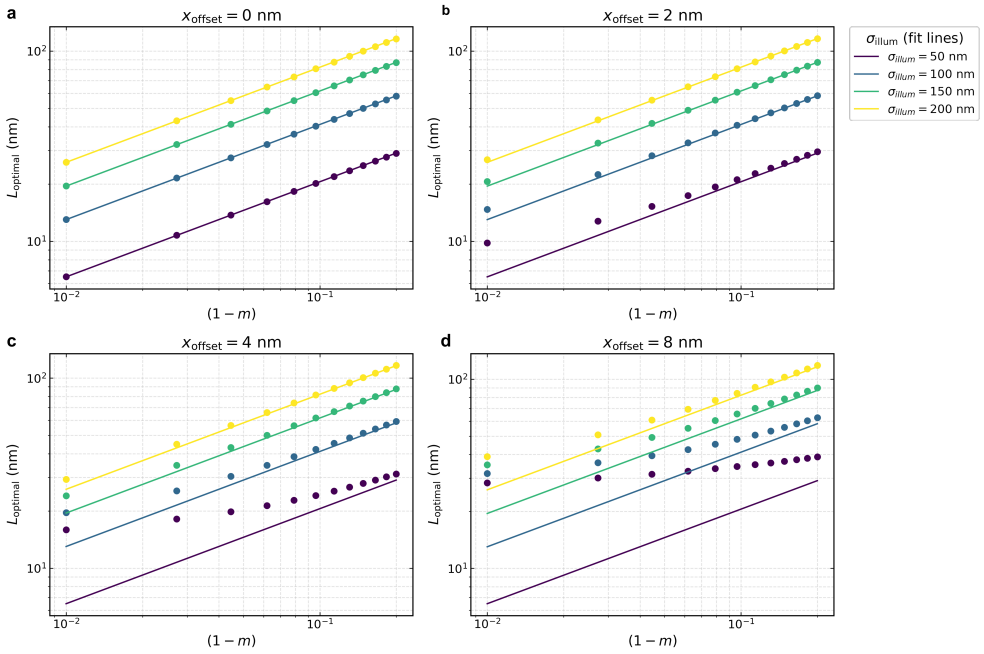


Fig. S9 The predictive model $L_{\text{opt}} = 1.30\,\sigma_{\text{illum}}\sqrt{1-m}$ across different lateral offsets of the emitter relative to the illumination pattern. The coefficient A is fitted using data at $x_{\text{offset}} = 0\,\text{nm}$ and held fixed in all subplots. **(a)**–**(d)** show the optimal pattern spacing L_{opt} as a function of $(1-m)$, for emitter offsets $x_{\text{offset}} = 0, 2, 4, 8\,\text{nm}$, respectively. Data points represent CRLB-optimized values, while solid lines show the fitted scaling law for each illumination width $\sigma_{\text{illum}} \in \{50, 100, 150, 200\}\,\text{nm}$. Despite increasing misalignment, the model remains predictive, particularly for larger σ_{illum}.

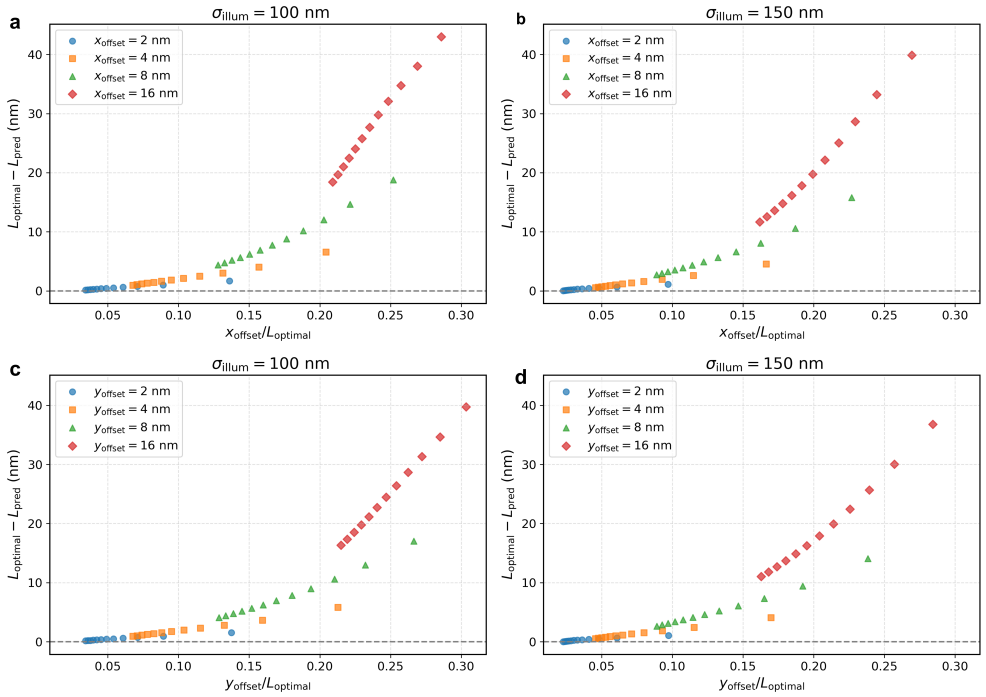


Fig. S10 Deviation between the numerically optimized and predicted illumination pattern spacing L_{optimal} as a function of relative emitter displacement. The predicted value is given by $L_{\text{pred}} = 1.30\,\sigma_{\text{illum}}\sqrt{1-m}$, and deviations are shown as $L_{\text{optimal}} - L_{\text{pred}}$. **(a)**–**(b)**: Influence of lateral offsets along the x-axis for $\sigma_{\text{illum}} = 100\,\text{nm}$ and $150\,\text{nm}$, respectively. **(c)**–**(d)**: Same as above, but for lateral offsets along the y-axis. Each point corresponds to a different modulation contrast m, and the horizontal axis represents the emitter offset normalized by the corresponding L_{optimal}. For small offsets, the prediction remains accurate; larger misalignments cause a systematic increase in optimal spacing beyond the predicted value.

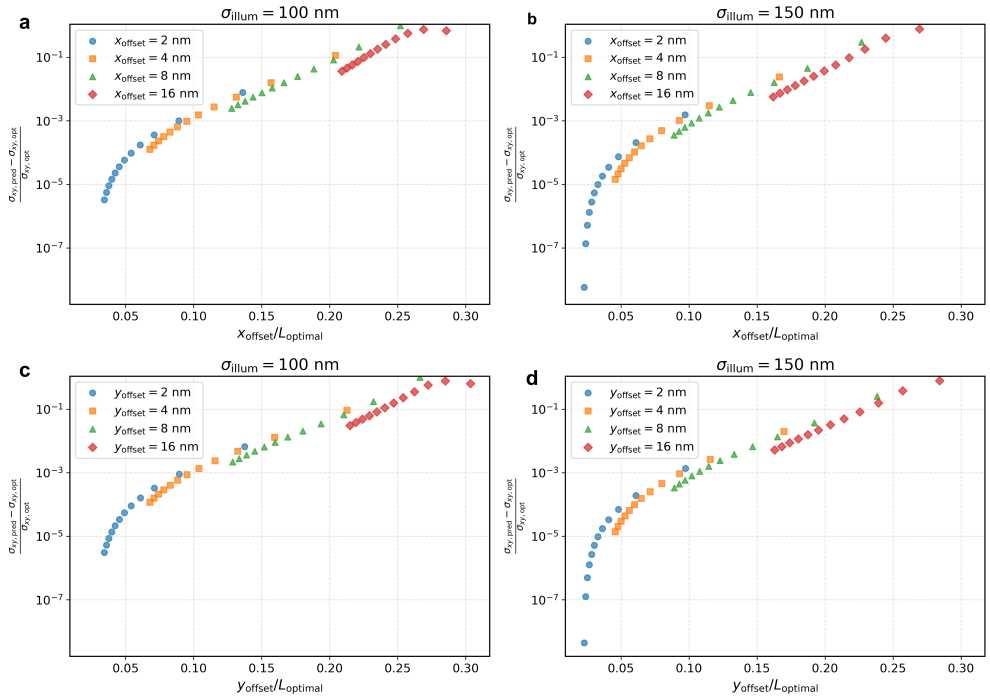


Fig. S11 Relative error in predicted localization precision due to lateral emitter misalignment. The predicted illumination spacing is given by $L_{\mathrm{pred}} = 1.3\,\sigma_{\mathrm{illum}}\sqrt{1-m}$. Each point compares the localization precision obtained using L_{pred} with the true optimum L_{opt}, computed via CRLB minimization. The vertical axis shows the relative precision error $(\sigma_{xy,\mathrm{pred}} - \sigma_{xy,\mathrm{opt}})/\sigma_{xy,\mathrm{opt}}$, while the horizontal axis expresses the normalized offset $x_{\mathrm{offset}}/L_{\mathrm{opt}}$ or $y_{\mathrm{offset}}/L_{\mathrm{opt}}$. **(a)**–**(b)** show errors for increasing x-offsets, and **(c)**–**(d)** for increasing y-offsets, at $\sigma_{\mathrm{illum}} = 100\,\mathrm{nm}$ and $150\,\mathrm{nm}$, respectively. The results indicate that the analytical prediction remains accurate (error $< 10\%$) for moderate misalignments relative to the pattern spacing.

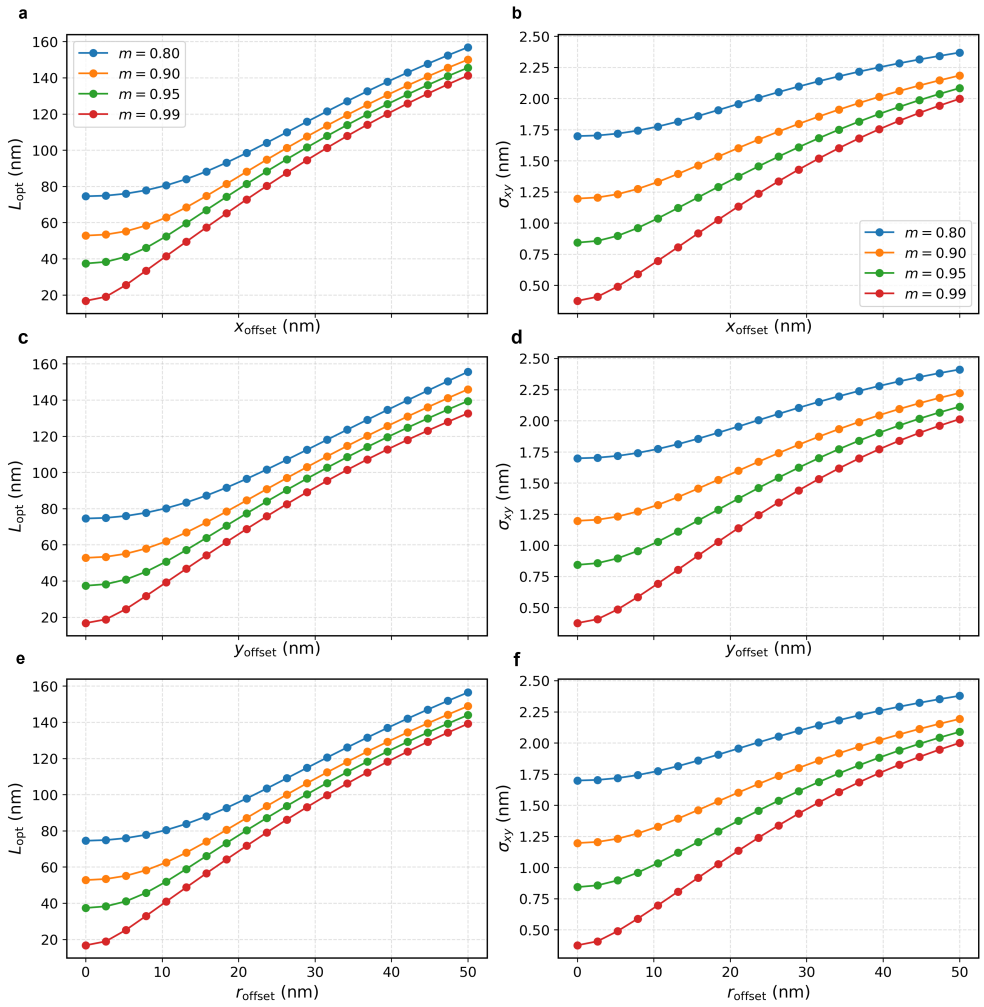


Fig. S12 **(a)** Optimal illumination spacing L_{opt} as a function of emitter offset along the x-axis, computed by minimizing the localization uncertainty for each fixed offset. **(b)** Corresponding localization precision σ_{xy} evaluated at the optimized spacing in panel (a). **(c)** Same as (a), but for emitter offsets along the y-axis. **(d)** Same as (b), but for offsets along the y-axis. **(e)** Same as (a), but for emitter offsets along the diagonal direction $r = -\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}}$. **(f)** Same as (b), but for diagonal emitter offsets as in (e).

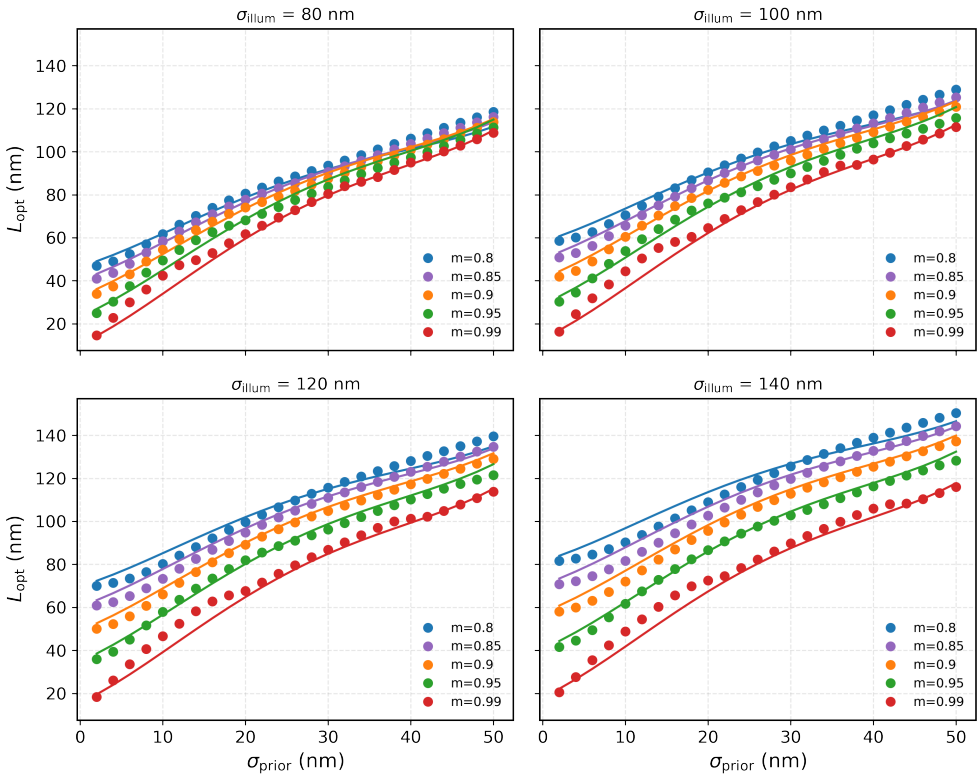


Fig. S13 Optimal illumination spacing L_{opt} as a function of prior localization uncertainty σ_{prior} for four illumination widths $\sigma_{\text{illum}} \in \{80, 100, 120, 140\}$ nm. Data points (circles) are shown alongside an analytical fit (solid lines) for modulation contrasts $m \in \{0.8, 0.85, 0.9, 0.95, 0.99\}$. The fit is based on a fourth-order polynomial model of the form: $L_{\text{opt}} = 1.3\,\sigma_{\text{illum}}\sqrt{1-m} + m^f \sum_{i=1}^{4} a_i\,\sigma_{\text{prior}}^i$ where the fitted coefficients are $a_1 = 1.852$, $a_2 = 8.362 \times 10^{-2}$, $a_3 = -3.163 \times 10^{-3}$, $a_4 = 3.124 \times 10^{-5}$, and exponent $f = 1.991$. The model achieves a coefficient of determination of $R^2 = 0.98853$, indicating excellent agreement between the model and simulation data.

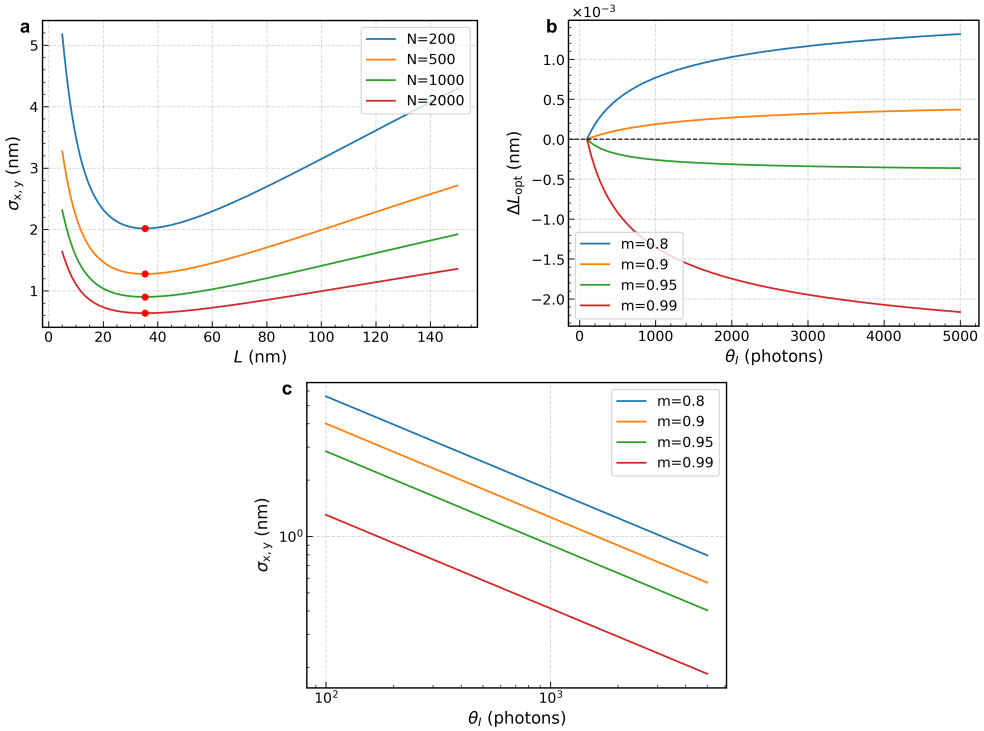


Fig. S14 **(a)** Precision curves $\sigma_{x,y}(L)$ for varying total signal-photon counts. Each curve shows the root-mean-square localization precision $\sigma_{x,y}$ (in nm) as a function of the pattern side length L (nm) for four total-photon budgets: $N = 200$ (blue), $N = 500$ (orange), $N = 1000$ (green), and $N = 2000$ (red). The curves are generated by sampling L uniformly between 5nm and 150nm, computing the two-dimensional Cramér–Rao lower bound (CRLB) at each L, and plotting $\sigma_{x,y}$ versus L. Solid circles mark the optimal side length $L_{\text{opt}}(N)$ that minimizes $\sigma_{x,y}$ for each photon budget. As expected, $\sigma_{x,y}$ decreases with increasing N, and L_{opt} shifts slightly toward smaller values as N grows. **(b)** Shift in optimal pattern side length $\Delta L = L_{\text{opt}}(N) - L_{\text{opt}}(N = 100)$ versus total signal photons N. For four modulation contrasts $m = 0.80$ (blue), $m = 0.90$ (orange), $m = 0.95$ (green), and $m = 0.99$ (red), we plot ΔL (in nm) against N (photons) on a logarithmic scale. It can be seen that the difference in L_{opt} is smaller than 0.0025 nm. **(c)** Log–log plot of $\sigma_{x,y}$ versus total signal photons N for multiple modulation contrasts. For $m = 0.80$ (blue), $m = 0.90$ (orange), $m = 0.95$ (green), and $m = 0.99$ (red), the optimal $L_{\text{opt}}(N)$ is first determined via the procedure in panel (a). Using those L_{opt} values, we compute $\sigma_{x,y}$ at each N and plot it on a double-logarithmic axis. Each nearly straight line confirms the asymptotic scaling $\sigma_{x,y} \propto N^{-1/2}$. Higher modulation contrast yields lower $\sigma_{x,y}$ at fixed N, as evidenced by the vertical separation of the lines.

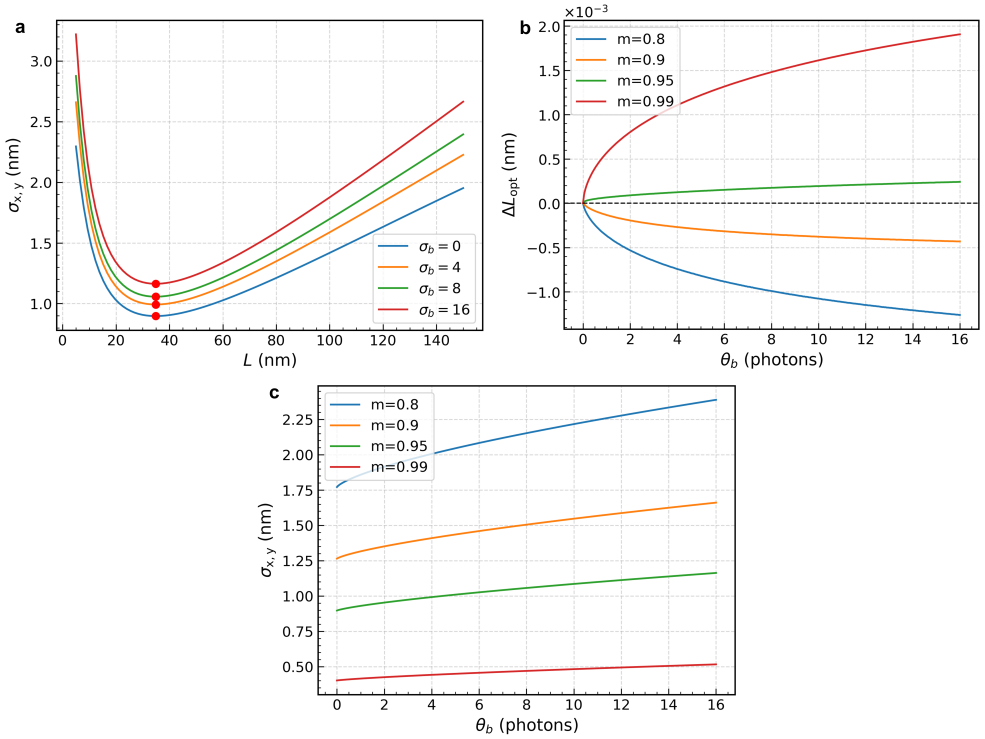


Fig. S15 **(a)** Precision curves $\sigma_{x,y}(L)$ for varying total background-photon counts. Each curve shows the root-mean-square localization precision $\sigma_{x,y}$ (in nm) as a function of the length L (nm) for four backgrounds: $\sigma_b = 0$ (blue), $\sigma_b = 4$ (orange), $\sigma_b = 8$ (green), and $\sigma_b = 16$ (red). The curves are generated by sampling L uniformly between 5 nm and 150 nm, computing the two-dimensional Cramér–Rao lower bound (CRLB) at each L, and plotting $\sigma_{x,y}$ versus L. Solid circles mark the optimal side length $L_{\mathrm{opt}}(\sigma_b)$ that minimizes $\sigma_{x,y}$ for each background level. As expected, $\sigma_{x,y}$ increases with larger σ_b, and L_{opt} shifts slightly toward larger values as background grows. **(b)** Shift in optimal pattern side length $\Delta L = L_{\mathrm{opt}}(\theta_b) - L_{\mathrm{opt}}(\theta_b = 0)$ versus total signal photons N for different modulation contrasts. For $m = 0.80$ (blue), $m = 0.90$ (orange), $m = 0.95$ (green), and $m = 0.99$ (red), we plot ΔL (in nm) against N (photons) on a logarithmic scale, using σ_b values from panel (a) to compute each $L_{\mathrm{opt}}(\sigma_b)$. It can be seen that the difference in L_{opt} is smaller than 0.0025 nm. **(c)** Localization precision $\sigma_{x,y}$ versus total background photons θ_b for multiple modulation contrasts at fixed $N = 1000$. For $m = 0.80$ (blue), $m = 0.90$ (orange), $m = 0.95$ (green), and $m = 0.99$ (red), we first determine the optimal side length $L_{\mathrm{opt}}(\sigma_b)$ as in panel (a), then compute $\sigma_{x,y}$ at each σ_b and plot it on a linear axis. Each curve rises monotonically, confirming that increasing background degrades precision.

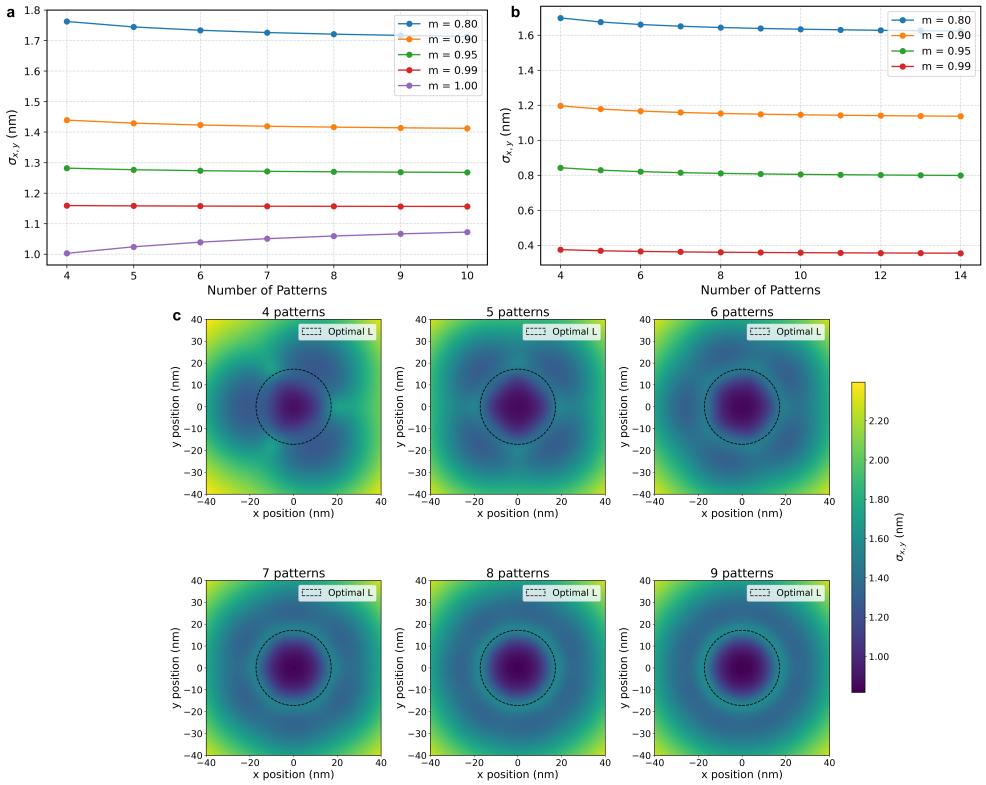


Fig. S16 **(a)** Precision versus number of patterns for a fixed length $L = 50\,\text{nm}$. When $m = 1$ (perfect modulation), adding off-center patterns decreases precision, whereas for $m < 1$ (imperfect modulation) precision increases as more patterns are added. Note that the optimal L for $m = 1$ is 0, so only the central pattern remains, so would not be a fair comparison. **(b)** Same as (a) but using the optimized length L for each modulation m. Again, adding patterns improves precision for $m < 1$. **(c)** Two-dimensional precision maps for different numbers of patterns, using $m = 0.95$ and the optimal L. The location of each pattern corresponds to regions of altered precision, illustrating how the pattern count influences spatial precision.

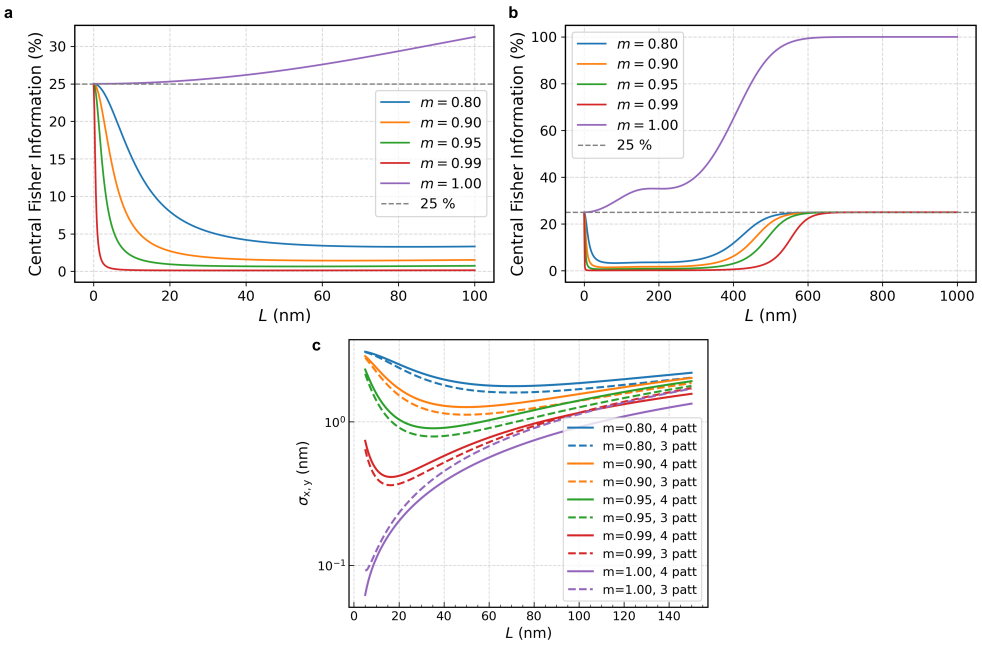


Fig. S17 **(a)** Sum of the Fisher information entries $F_{xx} + F_{yy}$ contributed by the central (middle) pattern, normalized by the total contribution of the non-center patterns, as a function of L. At $L = 0$, all patterns coincide and each contributes 25% of the total information. For perfect illumination ($m = 1$), the middle pattern dominates, whereas for imperfect illumination ($m < 1$) it contributes almost nothing for larger L. **(b)** Same quantity as in (a), but plotted for larger values of L. As large L , the relative contribution of the middle pattern becomes 100% for $m = 1$, due to the other patterns not being able to illuminate the emitter at this distance. For $m < 1$it becomes 25% due to each pattern having very small derivatives and becoming just wide field illumination. **(c)** Localization precision $\sigma_{\mathrm{x,y}}$ versus L with all four patterns (solid lines) and with the central pattern removed (dashed lines). Removing the middle pattern lowers the precision for perfect illumination (since the central pattern would otherwise provide infinite information), but for imperfect illumination the precision actually improves when the center is omitted.

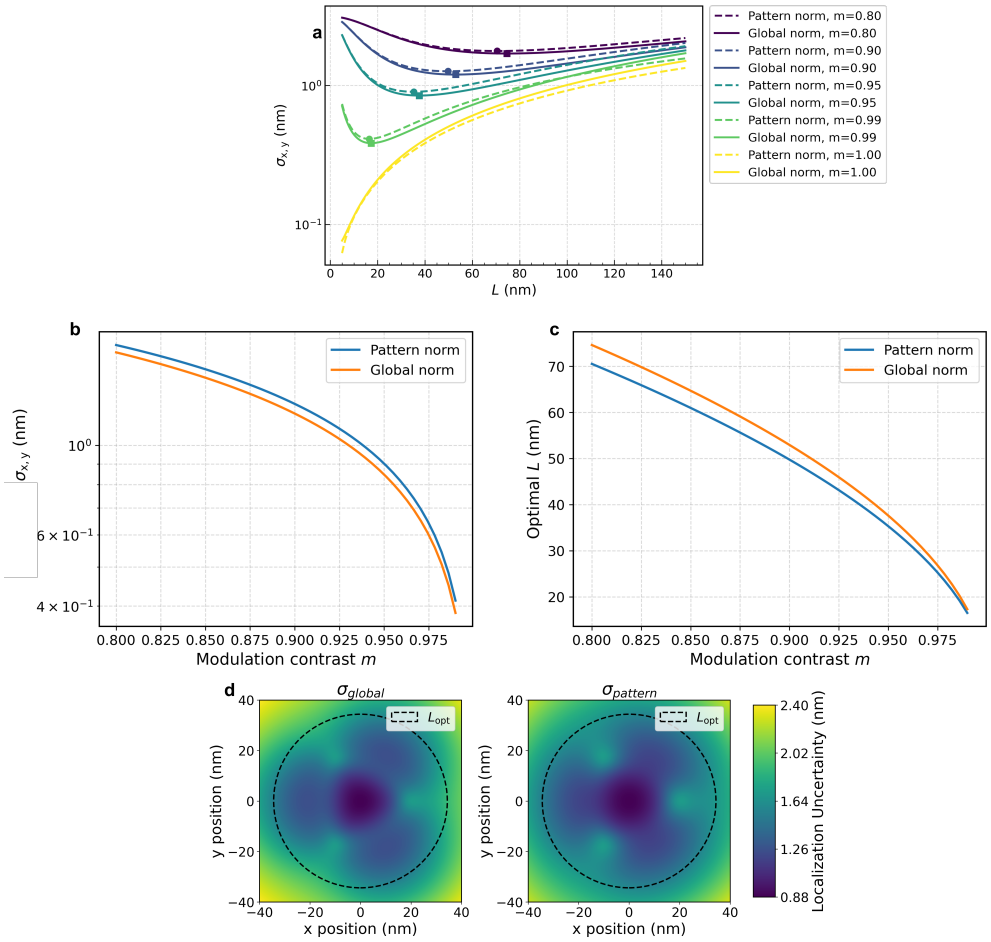


Fig. S18 **(a)** Localization precision $\sigma_{x,y}$ as a function of L for pattern normalization (dashed lines) and global normalization (solid lines) at modulation contrasts $m = 0.80,\ 0.90,\ 0.95,\ 0.99,\ 1.00$. Colored markers indicate the optimal L (excluding $m = 1$). **(b)** Optimal $\sigma_{x,y}$ versus m for both pattern-norm and global-norm methods, illustrating how precision degrades as modulation decreases. **(c)** Optimal L versus $1m$ for pattern-norm and global-norm methods. **(d)** Two-dimensional maps of $\sigma_{x,y}$ across emitter positions (x, y) at fixed $m = 0.95$, using the optimal L_{opt} determined at $(0, 0)$. Left panel shows the pattern-norm map, and right panel shows the global-norm map. In both maps, the dashed circle marks the radius L_{opt} at the center.

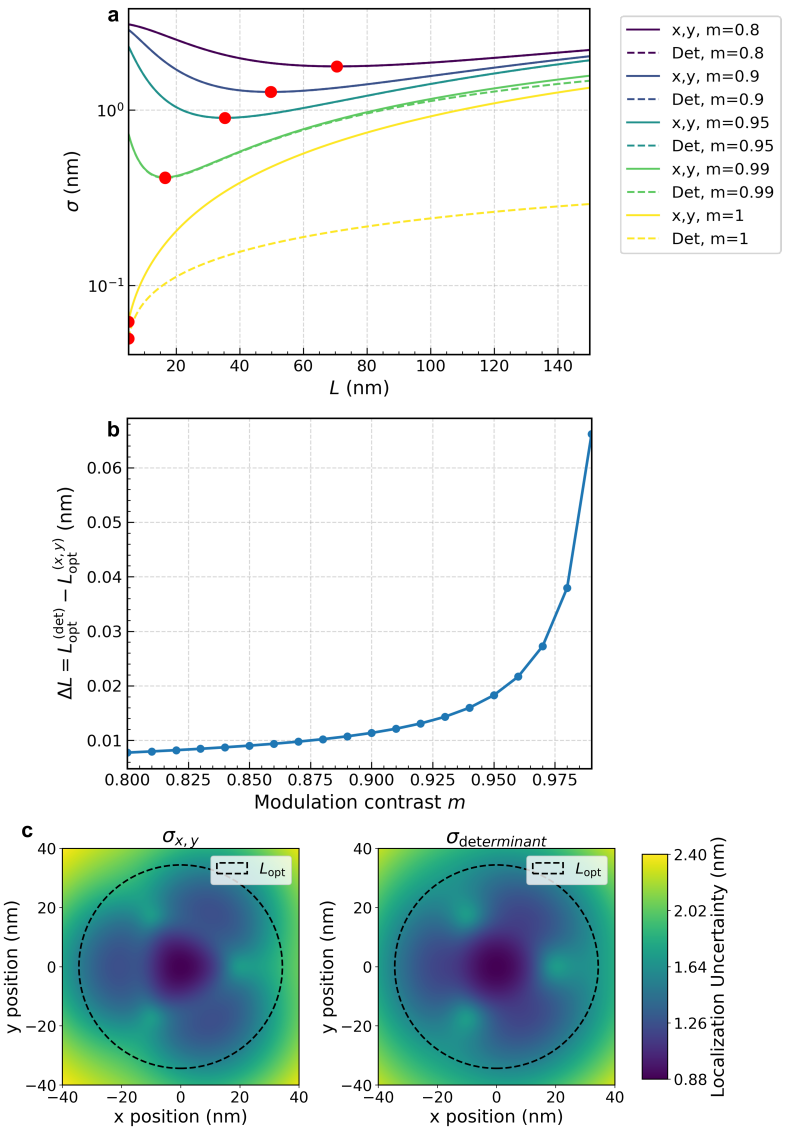


Fig. S19 **(a)** Comparison of average- and determinant-based CRLB curves $\sigma_{x,y}(L)$ (solid) and $\sigma_{\det}(L)$ (dashed) as a function of the length L for modulation contrasts $m = 0.80$ (purple), 0.90 (blue), 0.95 (teal), 0.99 (green), and 1.00 (yellow). Red circles mark the minima $L_{\text{opt}}^{(x,y)}$ where $\sigma_{x,y}(L)$ is smallest. **(b)** Difference $\Delta L = L_{\text{opt}}^{(x,y)} - L_{\text{opt}}^{(\text{det})}$ versus modulation contrast m. A positive ΔL indicates that the average-based optimum exceeds the determinant-based optimum. The curve rises monotonically with m, showing greater divergence of the two optima at high contrast. However, the difference is smaller than 0.1 nm. **(c)** Two-dimensional precision maps over the camera field of view (positions $x_{\text{emit}}, y_{\text{emit}}$). Left: $\sigma_{x,y}$ evaluated at $L_{\text{opt}}^{(x,y)}$, with a dashed circle indicating the locus of emitter positions at that optimum side length. Right: $\sigma_{\det}$ evaluated at $L_{\text{opt}}^{(\text{det})}$, with its corresponding dashed circle.

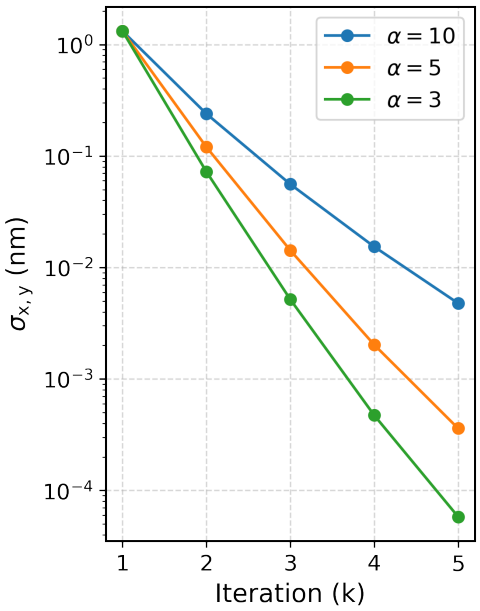


Fig. S20 Localization precision $\sigma_{x,y}$ of Iterative MINFLUX as a function of iteration number k for different aggression parameters α, where each illumination radius is set by $L_k = \alpha \cdot \sigma_{x,y}^{(k-1)}$. This result assumes perfect modulation contrast ($m = 1$) and reproduces the scaling behavior described in [5], where $\sigma_{x,y} \propto 1/\sqrt{N}^k$.

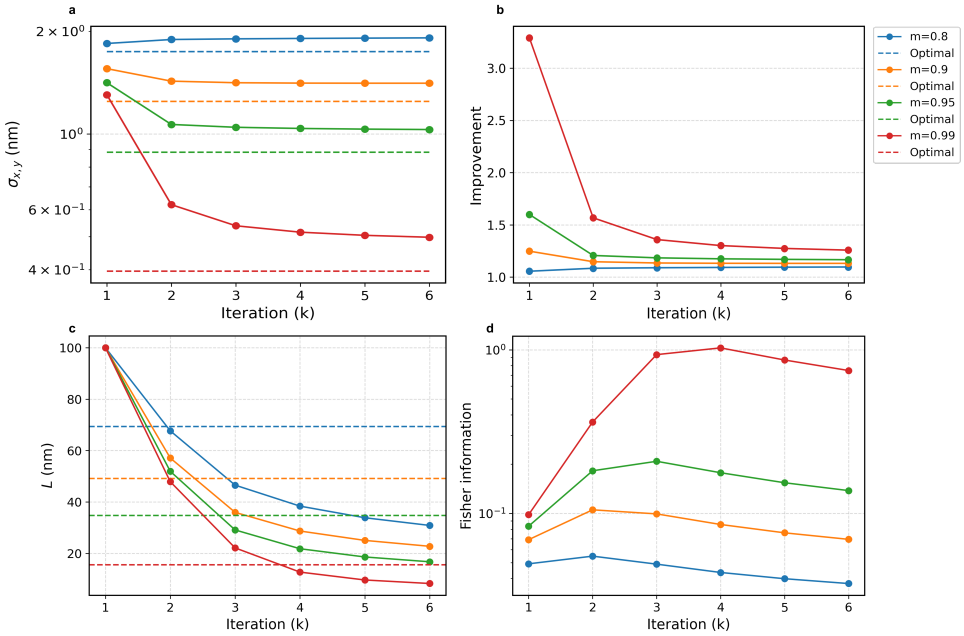


Fig. S21 Iterative MINFLUX simulation with $L_1 = 100$ and $\alpha = 15$. (a) The localization precision achieved in the simulation for different modulation contrasts m. For reference, the dashed lines indicate the precision attainable with optimal MINFLUX for each m. At iteration 1, the full photon budget is used in a single step; at iteration 2, the photon budget is divided evenly over two iterations, and so forth. (b) The ratio of the localization precision of optimal MINFLUX to that of iterative MINFLUX, quantifying the improvement achieved by MINFLUX. (c) The evolution of the illumination parameter L throughout the iterative simulation, starting from L_1 and updating according to $L_k = \alpha \cdot \sigma_{x,y}^{(k-1)}$ at each iteration. (d) The contribution of each pattern to the total Fisher information $\sqrt{\frac{I_{F_{x,x}}^2 + I_{F_{y,y}}^2}{2}}$ at each iteration. The Fisher information contribution peaks when L is closest to the optimal value found in (c).

Supplementary Note 8 Look Up Tables

Table S2 Optimale L (nm) for $m = 0.80$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	47.0	52.8	58.5	64.3	70.1	75.9	81.6	87.4
4.0	49.1	54.5	60.2	65.7	71.4	77.0	82.7	88.4
6.0	52.5	57.6	62.7	68.1	73.5	79.1	84.6	90.2
8.0	57.0	61.6	66.4	71.4	76.5	81.8	87.1	92.5
10.0	61.6	66.0	70.5	75.2	80.2	85.1	90.2	95.5
12.0	66.1	70.5	74.9	79.5	84.1	88.9	93.8	98.8
14.0	70.2	74.7	79.1	83.6	88.1	92.8	97.6	102.4
16.0	73.9	78.6	83.1	87.7	92.2	96.7	101.4	106.1
18.0	77.3	82.2	86.9	91.5	96.0	100.5	105.2	109.8
20.0	80.5	85.6	90.4	95.1	99.7	104.3	108.9	113.5
22.0	83.4	88.8	93.8	98.5	103.2	107.9	112.5	117.1
24.0	86.2	91.7	96.9	101.8	106.6	111.2	115.9	120.6
26.0	88.6	94.4	99.8	104.8	109.8	114.5	119.3	124.0
28.0	91.0	97.0	102.5	107.8	112.8	117.7	122.5	127.3
30.0	93.5	99.3	105.1	110.5	115.7	120.7	125.5	130.4
32.0	96.0	101.7	107.5	113.0	118.4	123.5	128.5	133.4
34.0	98.5	104.1	109.8	115.5	120.9	126.2	131.3	136.3
36.0	101.1	106.6	112.2	117.8	123.4	128.8	134.0	139.1
38.0	103.7	109.1	114.6	120.2	125.7	131.2	136.6	141.7
40.0	106.2	111.6	117.0	122.5	128.0	133.6	139.0	144.3
42.0	108.7	114.0	119.4	124.8	130.4	135.9	141.3	146.7
44.0	111.2	116.4	121.8	127.2	132.7	138.1	143.7	149.1
46.0	113.6	118.8	124.1	129.5	134.9	140.4	145.9	151.3
48.0	116.1	121.2	126.6	131.9	137.3	142.7	148.1	153.6
50.0	118.5	123.7	128.9	134.2	139.5	144.9	150.4	155.9

Table S3 Optimale L (nm) for $m = 0.85$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	41.0	45.9	50.9	55.9	60.9	65.9	70.9	75.9
4.0	43.6	48.2	52.9	57.7	62.5	67.3	72.3	77.2
6.0	47.9	52.0	56.2	60.7	65.2	69.8	74.5	79.3
8.0	53.1	56.9	60.7	64.8	68.9	73.3	77.7	82.2
10.0	58.4	62.0	65.6	69.5	73.4	77.3	81.6	85.8
12.0	63.0	66.8	70.5	74.2	78.0	81.8	85.8	89.8
14.0	67.3	71.2	75.0	78.8	82.5	86.3	90.2	94.1
16.0	71.0	75.2	79.2	83.0	86.9	90.6	94.5	98.3
18.0	74.5	78.9	83.1	87.1	90.9	94.8	98.6	102.4
20.0	77.7	82.3	86.7	90.9	94.8	98.8	102.6	106.4
22.0	80.5	85.5	90.1	94.4	98.4	102.4	106.4	110.2
24.0	83.0	88.4	93.2	97.7	101.9	105.9	110.0	113.9
26.0	85.5	90.9	96.0	100.7	105.1	109.3	113.4	117.4
28.0	88.0	93.4	98.6	103.5	108.1	112.5	116.7	120.8
30.0	90.6	95.8	101.0	106.1	110.9	115.5	119.8	124.0
32.0	93.2	98.3	103.4	108.5	113.5	118.2	122.7	127.0
34.0	95.9	100.8	105.9	110.9	115.9	120.8	125.4	129.9
36.0	98.5	103.4	108.3	113.3	118.3	123.2	128.0	132.6
38.0	101.1	105.9	110.8	115.7	120.7	125.6	130.5	135.2
40.0	103.7	108.4	113.3	118.1	123.0	128.0	132.8	137.6
42.0	106.2	110.9	115.7	120.5	125.4	130.2	135.2	139.9
44.0	108.7	113.4	118.1	122.9	127.7	132.6	137.4	142.3
46.0	111.2	115.8	120.5	125.3	130.1	134.9	139.7	144.5
48.0	113.7	118.2	122.9	127.7	132.4	137.2	142.0	146.8
50.0	116.2	120.5	125.2	129.9	134.7	139.5	144.2	149.1

Table S4 Optimale L (nm) for $m = 0.90$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	34.0	38.0	42.0	46.0	50.0	54.1	58.1	62.2
4.0	37.3	40.9	44.6	48.4	52.2	56.1	60.0	63.9
6.0	43.0	45.9	49.1	52.4	55.9	59.4	63.0	66.7
8.0	49.1	51.9	54.7	57.7	60.7	63.9	67.2	70.6
10.0	54.5	57.5	60.4	63.2	66.1	69.1	72.1	75.3
12.0	59.4	62.6	65.6	68.5	71.4	74.3	77.3	80.2
14.0	63.7	67.1	70.4	73.4	76.4	79.3	82.2	85.2
16.0	67.5	71.2	74.7	77.9	81.0	84.0	87.0	89.9
18.0	71.1	75.0	78.6	82.0	85.2	88.4	91.4	94.4
20.0	74.1	78.4	82.3	85.9	89.2	92.4	95.5	98.7
22.0	76.6	81.4	85.6	89.4	92.9	96.2	99.5	102.7
24.0	79.1	84.0	88.5	92.7	96.4	99.8	103.2	106.5
26.0	81.8	86.4	91.1	95.5	99.5	103.2	106.7	110.1
28.0	84.6	89.0	93.5	98.0	102.3	106.2	109.9	113.4
30.0	87.4	91.6	96.0	100.5	104.9	109.1	112.9	116.6
32.0	90.2	94.4	98.7	103.0	107.3	111.6	115.6	119.5
34.0	93.0	97.1	101.2	105.5	109.8	114.1	118.2	122.2
36.0	95.7	99.8	103.9	108.0	112.3	116.5	120.6	124.7
38.0	98.4	102.4	106.5	110.6	114.8	118.9	123.0	127.1
40.0	100.9	104.9	109.1	113.2	117.3	121.3	125.5	129.5
42.0	103.4	107.4	111.6	115.6	119.8	123.8	127.8	131.9
44.0	105.9	109.8	114.0	118.0	122.2	126.2	130.2	134.3
46.0	108.4	112.2	116.2	120.4	124.5	128.6	132.7	136.6
48.0	111.1	114.6	118.6	122.7	126.8	130.9	134.9	139.0
50.0	113.8	117.1	120.9	124.9	129.1	133.2	137.3	141.2

Table S5 Optimale L (nm) for $m = 0.925$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	29.8	33.2	36.6	40.1	43.6	47.1	50.5	54.1
4.0	34.0	36.9	39.8	43.0	46.2	49.5	52.8	56.2
6.0	40.4	42.7	45.2	47.9	50.6	53.5	56.5	59.6
8.0	46.7	49.1	51.5	53.9	56.3	58.9	61.5	64.3
10.0	52.3	54.9	57.4	59.8	62.3	64.7	67.1	69.6
12.0	57.1	60.1	62.7	65.3	67.7	70.2	72.7	75.1
14.0	61.4	64.6	67.6	70.3	72.9	75.4	77.9	80.3
16.0	65.4	68.8	71.9	74.8	77.6	80.2	82.7	85.2
18.0	68.8	72.6	75.9	79.0	81.9	84.6	87.3	89.9
20.0	71.6	75.9	79.6	82.8	85.9	88.8	91.6	94.2
22.0	74.1	78.6	82.7	86.3	89.6	92.7	95.5	98.3
24.0	76.8	81.1	85.4	89.4	93.0	96.2	99.2	102.1
26.0	79.7	83.7	87.9	92.0	95.9	99.5	102.7	105.7
28.0	82.7	86.5	90.5	94.5	98.5	102.3	105.8	109.1
30.0	85.6	89.4	93.1	97.0	101.0	104.9	108.6	112.0
32.0	88.5	92.3	95.9	99.7	103.5	107.3	111.2	114.8
34.0	91.4	95.0	98.8	102.4	106.1	109.8	113.6	117.3
36.0	94.1	97.8	101.5	105.1	108.8	112.4	116.1	119.8
38.0	96.8	100.4	104.1	107.7	111.4	115.0	118.6	122.2
40.0	99.2	103.0	106.6	110.3	114.0	117.6	121.1	124.7
42.0	101.7	105.4	109.1	112.8	116.5	120.1	123.7	127.2
44.0	104.3	107.7	111.5	115.2	118.9	122.6	126.1	129.6
46.0	107.0	110.1	113.8	117.5	121.2	124.9	128.5	132.1
48.0	109.8	112.7	116.0	119.8	123.5	127.3	130.9	134.5
50.0	112.6	115.2	118.4	122.0	125.7	129.5	133.1	136.7

Table S6 Optimale L (nm) for $m = 0.95$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	24.9	27.7	30.4	33.2	36.0	38.8	41.6	44.5
4.0	30.3	32.4	34.6	37.0	39.4	42.0	44.5	47.1
6.0	37.5	39.4	41.2	43.0	45.1	47.2	49.4	51.6
8.0	43.8	45.9	47.9	49.8	51.6	53.5	55.5	57.4
10.0	49.5	51.8	53.9	55.9	57.9	59.8	61.7	63.6
12.0	54.3	56.9	59.3	61.5	63.5	65.5	67.5	69.5
14.0	58.8	61.5	64.1	66.5	68.8	70.9	72.9	74.8
16.0	62.7	65.9	68.5	71.0	73.4	75.7	77.9	79.9
18.0	65.7	69.5	72.7	75.3	77.8	80.2	82.4	84.6
20.0	68.2	72.3	76.0	79.1	81.9	84.4	86.7	89.0
22.0	71.1	74.8	78.8	82.3	85.5	88.3	90.8	93.1
24.0	74.2	77.6	81.2	85.0	88.5	91.6	94.4	97.0
26.0	77.5	80.7	84.0	87.4	91.0	94.5	97.6	100.4
28.0	80.6	83.8	87.0	90.2	93.6	97.0	100.3	103.4
30.0	83.8	86.9	90.0	93.1	96.2	99.5	102.8	106.1
32.0	86.7	89.8	93.0	96.0	99.1	102.2	105.4	108.6
34.0	89.7	92.7	95.9	99.0	102.0	105.0	108.0	111.2
36.0	92.4	95.5	98.7	101.8	104.8	107.8	110.8	113.8
38.0	94.9	98.2	101.4	104.5	107.6	110.6	113.5	116.5
40.0	97.3	100.5	103.9	107.1	110.2	113.3	116.2	119.1
42.0	99.9	102.8	106.2	109.6	112.7	115.9	118.8	121.8
44.0	102.7	105.2	108.4	111.8	115.2	118.3	121.4	124.4
46.0	105.5	107.8	110.7	114.0	117.3	120.6	123.8	126.8
48.0	108.4	110.5	113.1	116.1	119.5	122.8	126.0	129.1
50.0	111.3	113.4	115.7	118.4	121.6	124.8	128.2	131.3

Table S7 Optimale L (nm) for $m = 0.975$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	19.0	20.8	22.6	24.5	26.3	28.3	30.2	32.2
4.0	26.4	27.5	28.8	30.2	31.6	33.0	34.6	36.2
6.0	33.6	35.1	36.4	37.7	39.0	40.2	41.4	42.7
8.0	40.1	41.6	43.0	44.5	45.9	47.3	48.5	49.8
10.0	45.4	47.3	49.1	50.8	52.2	53.6	55.0	56.3
12.0	50.7	52.6	54.4	56.2	57.9	59.5	60.9	62.3
14.0	55.2	57.6	59.5	61.2	63.0	64.6	66.2	67.8
16.0	58.2	61.6	64.1	66.2	68.0	69.6	71.2	72.8
18.0	60.5	64.1	67.5	70.2	72.4	74.3	76.0	77.7
20.0	64.1	66.7	69.8	73.2	76.0	78.4	80.4	82.2
22.0	67.9	70.2	72.7	75.5	78.6	81.5	84.0	86.1
24.0	71.5	73.8	76.1	78.4	81.1	83.9	86.8	89.3
26.0	74.9	77.3	79.6	81.9	84.1	86.6	89.2	91.9
28.0	78.3	80.7	83.0	85.3	87.5	89.7	92.0	94.5
30.0	81.6	83.9	86.3	88.6	90.9	93.0	95.2	97.3
32.0	84.9	87.2	89.5	91.8	94.1	96.2	98.4	100.5
34.0	87.8	90.3	92.6	94.9	97.2	99.4	101.6	103.6
36.0	90.1	93.0	95.5	97.9	100.2	102.4	104.6	106.7
38.0	92.5	95.1	98.1	100.7	103.0	105.3	107.5	109.7
40.0	95.3	97.2	100.0	103.1	105.7	108.0	110.3	112.5
42.0	98.3	99.8	102.0	104.9	108.0	110.6	113.0	115.2
44.0	101.2	102.7	104.5	106.9	109.8	112.7	115.4	117.7
46.0	104.1	105.6	107.3	109.2	111.6	114.5	117.4	120.0
48.0	107.0	108.6	110.2	112.0	114.0	116.4	119.1	122.0
50.0	109.9	111.5	113.1	114.8	116.6	118.6	121.0	123.8

Table S8 Optimale L (nm) for $m = 0.99$

σ_{prior} (nm) \ σ_{illum} (nm)	80	90	100	110	120	130	140	150
2.0	14.7	15.5	16.4	17.3	18.4	19.5	20.5	21.6
4.0	22.7	23.8	24.6	25.4	26.1	26.8	27.6	28.4
6.0	30.0	31.0	31.9	32.7	33.6	34.5	35.4	36.2
8.0	36.0	37.1	38.4	39.6	40.6	41.6	42.4	43.2
10.0	42.3	43.4	44.5	45.5	46.6	47.7	48.8	49.8
12.0	47.3	49.1	50.5	51.6	52.5	53.5	54.5	55.5
14.0	49.7	53.1	55.3	57.0	58.2	59.3	60.3	61.2
16.0	53.0	54.8	58.1	60.9	62.8	64.4	65.6	66.7
18.0	57.5	58.9	60.4	62.6	65.6	68.0	69.8	71.4
20.0	61.6	63.2	64.6	66.0	67.7	69.8	72.5	74.7
22.0	65.6	67.3	68.8	70.2	71.6	73.0	74.6	76.7
24.0	69.4	71.2	72.8	74.3	75.6	77.0	78.3	79.7
26.0	72.9	74.8	76.6	78.1	79.6	80.9	82.3	83.5
28.0	76.6	78.3	80.1	81.8	83.4	84.8	86.1	87.3
30.0	80.5	82.0	83.5	85.2	86.9	88.4	89.8	91.1
32.0	84.0	85.6	87.1	88.6	90.2	91.8	93.3	94.7
34.0	86.1	88.9	90.6	92.2	93.6	95.1	96.6	98.0
36.0	88.2	89.9	93.5	95.5	97.0	98.4	99.8	101.3
38.0	91.5	92.3	93.9	97.7	100.0	101.6	103.0	104.5
40.0	94.9	95.5	96.5	98.0	101.2	104.2	106.0	107.6
42.0	97.9	98.8	99.6	100.7	102.1	104.6	108.0	110.2
44.0	100.2	101.6	102.7	103.8	104.8	106.2	108.3	111.2
46.0	102.8	104.3	105.6	106.8	107.9	109.0	110.4	112.1
48.0	105.8	107.1	108.4	109.7	110.9	112.0	113.1	114.5
50.0	108.8	110.1	111.4	112.6	113.8	114.9	116.0	117.2

Chapter 4

Conclusion

This chapter concludes the thesis by summarizing the main findings in Section 4-1 and outlining future directions in Section 4-2. The focus lies on understanding how imperfect modulation contrast affects the theoretical limits of localization precision in MINFLUX-based techniques and on proposing further steps to refine these insights and validate them experimentally.

4-1 Summary

In this thesis, we investigated the impact of imperfect modulation contrast on the localization precision of two super-resolution microscopy techniques: MINFLUX [22] and iterative MINFLUX [24]. MINFLUX uses a doughnut-shaped excitation beam with a zero-intensity center to triangulate the position of fluorophores with nanometer precision. Iterative MINFLUX extends this approach by progressively narrowing the illumination pattern in successive steps, further improving localization precision.

Conventional theoretical models assume ideal conditions where the doughnut-shaped beam exhibits a perfect zero-intensity minimum at its center. However, in practical implementations, this assumption does not hold. Optical aberrations, imperfect interference, and system misalignment often lead to residual intensity at the beam center [28, 29, 30]. To address this discrepancy, we introduced a modulation contrast parameter $m \in (0, 1]$, where $m = 1$ corresponds to perfect modulation and lower values reflect increasing degrees of imperfection.

We extended the Cramér–Rao Lower Bound (CRLB) model to incorporate modulation contrast. This allowed us to quantify how imperfections in the excitation pattern degrade the Fisher information and, in turn, the theoretical lower bound on localization precision. This generalization enables the CRLB to reflect the reduced spatial information content in the fluorescence signal caused by residual intensity at the center of the excitation pattern. Our study demonstrates that optimizing the illumination spacing L in MINFLUX microscopy significantly improves localization performance under non-ideal modulation conditions ($m < 1$). While in the ideal case ($m = 1$) localization precision improves indefinitely with smaller L [22], we show that for realistic values of m, there exists a finite optimal spacing $L_{\text{opt}}(m)$ that

minimizes localization uncertainty. For example, at $m = 0.95$, setting $L = 34$ nm improves precision by over 50% compared to using a small fixed radius of 5 nm. Using suboptimal spacings, such as the originally implemented $L = 70$ nm [22], recovers only 79% of the achievable precision.

We derived and validated a predictive formula, $L_{\text{opt}} \approx 1.30\,\sigma_{\text{illum}}\sqrt{1-m}$, which enables experimental determination of L_{opt} through measurements of the illumination profile and modulation contrast. This relation holds with 90% precision across practical values of m and as the emitter lies within 40% of the pattern diameter.

To extend this optimization framework to more realistic scenarios where the emitter location is uncertain, we modeled prior information as a Gaussian distribution with standard deviation σ_{prior}. Our results show that the optimal spacing $L_{\text{opt}}(\sigma_{\text{prior}})$ increases with the uncertainty of the prior, offering a practical method for adaptive pattern placement. σ_{prior} can be estimated using the measurement variance [25], the Cramér–Rao bound [27], or analytical approximations [17].

In iterative MINFLUX, we found that using the same optimized spacing L_{opt} across iterations yields superior performance compared to the standard approach that progressively shrinks the pattern radius [24]. This optimized single-step strategy outperforms the conventional iterative scheme by over 50% in localization precision at $m = 0.95$, and converges to simply performing k identical exposures at the optimal L. Although iterative refinement remains beneficial when the prior uncertainty σ_{prior} is large, it is crucial not to reduce L below the optimal value, as doing so propagates information loss across iterations.

4-2 Future Outlook

While this thesis has advanced our theoretical understanding of how imperfect modulation contrast affects localization precision in (iterative) MINFLUX, several important directions remain open for further investigation.

These results directly inform the design of optimized illumination strategies for single particle tracking using MINFLUX. By selecting the optimal pattern spacing L_{opt} based on modulation contrast and prior uncertainty, one can maximize localization precision at each frame. Furthermore, avoiding suboptimal spacing prevents error accumulation over time, making real time tracking of dynamic processes more reliable, even under realistic imaging conditions $m < 1$ or variable emitter positions. This could enable higher precision tracking with fewer photons, which is especially valuable in live cell imaging where phototoxicity and speed are limiting factors.

Furthermore, all results presented here were derived using the Cramér–Rao lower bound (CRLB), which assumes no prior knowledge about the emitter's position. In iterative MINFLUX, however, each localization step benefits from previous estimates. The Van Trees inequality (VTI) provides a framework that incorporates prior information and could offer a more accurate theoretical lower bound in this iterative context. Applying the VTI to iterative MINFLUX would clarify how information accumulates over steps and whether current scaling laws remain valid when prior knowledge is accounted for.

Experimental validation under realistic imaging conditions $m < 1$ will be essential to evaluate the practical accuracy of these predictions. Direct comparison between simulations and systematic measurements will determine whether the theoretical trends persist in practice.

Appendix A

Proof VTI with a Gaussian prior

This result was derived during my master's thesis but was not utilized in the current project.

Proof. We aim to derive the expression for the prior information matrix using a Gaussian prior used in the Van Trees Inequality (VTI), which provides a Bayesian lower bound on the variance of any estimator. Specifically, we are interested in evaluating the achievable localization precision at a given iteration k, denoted by $\sigma_{x,k}$ and $\sigma_{y,k}$, based on both the prior knowledge and the observed data. The VTI is given by:

$$\mathrm{Var}(\hat{\theta}_k) \geq (I_{P,k} + I_{D,k})^{-1}\,, \tag{A-1}$$

where $\hat{\theta}_k$ is the vector of parameters estimated at iteration k, $I_{P,k}$ is the prior information matrix, and $I_{D,k}$ is the Fisher information matrix associated with the data likelihood at iteration k.

The prior information matrix can be found using:

$$[I_{P,k}]_{i,j} = \int_{\mathbb{R}^n} \frac{\partial \log(\lambda(\theta_k))}{\partial \theta_i} \frac{\partial \log(\lambda(\theta_k))}{\partial \theta_j} \lambda(\theta_k)\, d\theta_k \tag{A-2}$$

Here i, j is the i-th and j-th entry of the matrix or vector. So $[I_{P,k}]_{i,j}$ is just one scalar entry of the prior information matrix.

We assume a Gaussian distribution as the prior:

$$\lambda(\boldsymbol{\theta}_k) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})\right) \tag{A-3}$$

With $\boldsymbol{\theta}_{k-1}$ being the mean of our prior, chosen as the estimation in the previous iteration $k-1$. The covariance matrix $\boldsymbol{\Sigma}$ can be any matrix that is symmetric and positive definite.

Now we are taking the logarithm of the prior:

$$\log(\lambda(\boldsymbol{\theta}_k)) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1}) \tag{A-4}$$

Since the covariance matrix $\boldsymbol{\Sigma}$ is symmetric and positive definite, we can compute the derivatives. The derivative of logarithm of the prior will result into:

$$\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{i,k}} = -\frac{1}{2}\frac{\partial}{\partial \theta_{i,k}}(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1}) \tag{A-5}$$

Since we are working with matrixes, we can deconstruct this into summations of derivatives. We remove all the terms where the derivative will become zero, so the terms without $\theta_{i,k}$. This will result in the following three terms:

$$\begin{aligned}\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{i,k}} = &-\frac{1}{2}\frac{\partial}{\partial \theta_{i,k}}(\theta_{i,k} - \theta_{i,k-1})\Sigma_{i,i}^{-1}(\theta_{i,k} - \theta_{i,k-1}) \\ &- \frac{1}{2}\sum_{\substack{m=1 \\ m\neq i}}^{\tilde{n}}\frac{\partial}{\partial \theta_{i,k}}(\theta_{i,k} - \theta_{i,k-1})\Sigma_{i,m}^{-1}(\theta_{m,k} - \theta_{m,k-1}) \\ &- \frac{1}{2}\sum_{\substack{m=1 \\ m\neq i}}^{\tilde{n}}\frac{\partial}{\partial \theta_{i,k}}(\theta_{m,k} - \theta_{m,k-1})\Sigma_{i,m}^{-1}(\theta_{i,k} - \theta_{i,k-1})\end{aligned} \tag{A-6}$$

We will result with three terms, the term of both the i-th row and i-th column, the term for the whole i-th row, and the term for the whole i-th column. In the summation, $\tilde{n}$ is the total length of the vector $\theta_{i,k}$.

We will now calculate the derivatives of all the three separate terms:

$$\begin{aligned}\frac{\partial}{\partial \theta_{i,k}}(\theta_{i,k} - \theta_{i,k-1})\Sigma_{i,i}^{-1}(\theta_{i,k} - \theta_{i,k-1}) &= \frac{\partial}{\partial \theta_{i,k}}\Sigma_{i,i}^{-1}(\theta_{i,k} - \theta_{i,k-1})^2 \\ &= 2\Sigma_{i,i}^{-1}(\theta_{i,k} - \theta_{i,k-1})\end{aligned} \tag{A-7}$$

$$\frac{\partial}{\partial \theta_{i,k}}(\theta_{i,k} - \theta_{i,k-1})\Sigma_{i,m}^{-1}(\theta_{m,k} - \theta_{m,k-1}) = \Sigma_{i,m}^{-1}(\theta_{m,k} - \theta_{m,k-1}) \tag{A-8}$$

$$\frac{\partial}{\partial \theta_{i,k}}(\theta_{m,k} - \theta_{m,k-1})\Sigma_{i,m}^{-1}(\theta_{i,k} - \theta_{i,k-1}) = \Sigma_{i,m}^{-1}(\theta_{m,k} - \theta_{m,k-1}) \tag{A-9}$$

Now that we have the derivative of all the three terms, we can add them back up together.

$$\begin{aligned}\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{i,k}} &= -\frac{1}{2}\cdot 2\Sigma^{-1}_{i,i}(\theta_{i,k}-\theta_{i,k-1}) \\ &\quad -\frac{1}{2}\sum_{\substack{m=1\\ m\neq i}}^{\tilde{n}}\Sigma^{-1}_{i,m}(\theta_{m,k}-\theta_{m,k-1}) \\ &\quad -\frac{1}{2}\sum_{\substack{m=1\\ m\neq i}}^{\tilde{n}}\Sigma^{-1}_{i,m}(\theta_{m,k}-\theta_{m,k-1}) \\ &= -\sum_{m=1}^{\tilde{n}}\Sigma^{-1}_{i,m}(\theta_{m,k}-\theta_{m,k-1}) \end{aligned} \tag{A-10}$$

The same will be done for the derivative with respect to $\theta_{j,k}$

$$\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{j,k}} = -\sum_{n=1}^{\tilde{n}}\Sigma^{-1}_{j,n}(\theta_{n,k}-\theta_{n,k-1}) \tag{A-11}$$

Now that we have both derivatives, we can fill them into the Prior information integral.

$$\begin{aligned}[I_{P,k}]_{i,j} &= \int_{\mathbb{R}^n}\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{i,k}}\frac{\partial \log(\lambda(\boldsymbol{\theta}_k))}{\partial \theta_{j,k}}\lambda(\boldsymbol{\theta}_k)\,d\boldsymbol{\theta}_k \\ &= \sum_{m=1}^{\tilde{n}}\sum_{n=1}^{\tilde{n}}\Sigma^{-1}_{i,m}\Sigma^{-1}_{j,n}\int_{\mathbb{R}^n}(\theta_{m,k}-\theta_{m,k-1})(\theta_{n,k}-\theta_{n,k-1})\lambda(\boldsymbol{\theta}_k)\,d\boldsymbol{\theta}_k \end{aligned} \tag{A-12}$$

The summations can be taken out of the integral. The integral will now represent the expectation property of Gaussian distributions for the covariance matrix. So we can rewrite the integral into the following:

$$\begin{aligned}\int_{\mathbb{R}^n}(\theta_{m,k}-\theta_{m,k-1})(\theta_{n,k}-\theta_{n,k-1})\lambda(\boldsymbol{\theta}_k)\,d\boldsymbol{\theta}_k &= \mathbb{E}\left[(\theta_{m,k}-\theta_{m,k-1})(\theta_{n,k}-\theta_{n,k-1})\right] \\ &= \Sigma_{m,n}\end{aligned} \tag{A-13}$$

This will then result into:

$$[I_{P,k}]_{i,j} = \sum_{m=1}^{\tilde{n}}\sum_{n=1}^{\tilde{n}}\Sigma^{-1}_{im}\Sigma^{-1}_{jn}\Sigma_{mn} \tag{A-14}$$

This can be interpreted as the element-wise definition for a matrix product involving three matrices:

- $\boldsymbol{\Sigma}^{-1}$ with elements Σ^{-1}_{im} and Σ^{-1}_{jn},

- $\mathbf{\Sigma}$ with elements Σ_{mn}.

Since the matrixes are symetric we can first look at the inner Sum Over n First, consider the inner sum:

$$\sum_{n=1}^{n} \Sigma_{jn}^{-1}\Sigma_{mn}. \tag{A-15}$$

This part computes the row-column multiplication between the row of $\mathbf{\Sigma}$ indexed by m and the column of $\mathbf{\Sigma}^{-1}$ indexed by j. By definition of matrix inversion, this sum equals the Kronecker delta:

$$\sum_{n=1}^{n} \Sigma_{jn}^{-1}\Sigma_{mn} = \delta_{mj} \tag{A-16}$$

where δ_{mj} is the Kronecker delta:

$$\delta_{mj} = \begin{cases} 1 & \text{if } m = j, \\ 0 & \text{if } m \neq j \end{cases} \tag{A-17}$$

Now we will take the outer sum over m with the substituted result back into the outer sum:

$$[I_{P,k}]_{i,j} = \sum_{m=1}^{n} \Sigma_{im}^{-1}\delta_{mj} \tag{A-18}$$

Since $\delta_{mj} = 1$ only when $m = j$, the sum collapses to a single term:

$$[I_{P,k}]_{i,j} = \Sigma_{ij}^{-1} \tag{A-19}$$

This is because the covariance matrix and its inverse combine to form the identity matrix, and the identity matrix leaves only the elements of the inverse matrix $\mathbf{\Sigma}^{-1}$.

Since this is the case for every entry of the prior information matrix the full prior information amtrix will result into:

$$\mathbf{I}_{P,k} = \mathbf{\Sigma}^{-1} \tag{A-20}$$

We will be able to choose any symmetric positive definitive covariance matrix as a prior. For our case the optimal prior would be the Bayesian information matrix from the previous iteration. This would result into:

$$\mathbf{I}_{P,k} = I_{P,k-1} + I_{D,k-1} \tag{A-21}$$

With $I_{D,k-1}$ being the data information matrix in iteration $k-1$. Since $I_{P,k-1}$ will be determined the same way as $I_{P,k}$, we can rewrite this into:

$$\mathbf{I}_{P,k} = I_{D,k-1} + I_{D,k-2} + ... + I_{D,1} + I_{P,1} \tag{A-22}$$

So this results into the prior information only being dependent on the the data information from previous iterations and the first prior. The first prior will come from the CRLB data information. So the prior information matrix successfully incorporates the all the prior data information available. □

Bibliography

[1] abberior Instruments GmbH. MINFLUX image of a hippocampal neuron (Beta II spectrin). Wikimedia Commons, https://commons.wikimedia.org/wiki/File:MINFLUX_image_of_a_hippocampal_neuron.jpg, 2023. Licensed under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0); accessed 2025-07-23.

[2] Zeiss Microscopy. The fundamentals of fluorescence microscopy. https://zeiss-campus.magnet.fsu.edu/articles/basics/fluorescence.html, 2023. Accessed: 2025-05-11.

[3] Fluorescence microscopy - an overview. https://www.sciencedirect.com/topics/medicine-and-dentistry/fluorescence-microscopy, 2023. Accessed: 2025-05-11.

[4] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell*. Garland Science, 4th edition, 2002.

[5] Ralph Weissleder. Molecular imaging in cancer. *Science*, 312(5777):1168–1171, 2006.

[6] Christoph Cremer. Label-free localisation microscopy spdm - super resolution microscopy christoph cremer. https://commons.wikimedia.org/wiki/File:Label-free_Localisation_Microscopy_SPDM_-_Super_Resolution_Microscopy_Christoph_Cremer.jpg, 2020. Accessed: 2025-07-23.

[7] Ulrich Kubitscheck. *Fluorescence Microscopy: From Principles to Biological Applications*. John Wiley & Sons, 2017.

[8] Joseph R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Springer, Boston, MA, 3rd edition, 2006.

[9] Wikimedia Commons contributor. Laser-induced fluorescence. https://commons.wikimedia.org/wiki/File:Laser-induced_fluorescence.png, 2020. Accessed: 2025-07-23, licensed under CC BY-SA 4.0.

[10] Xinyi Wang and Yunyan Lai. Three basic types of fluorescence microscopy and recent improvement. *E3S Web of Conferences*, 290:01031, 01 2021.

[11] Lothar Schermelleh, Rainer Heintzmann, and Heinrich Leonhardt. A guide to super-resolution fluorescence microscopy. *Journal of Cell Biology*, 190(2):165–175, 07 2010.

[12] Xi Chen, Yu Wang, Xuewei Zhang, and Changsheng Liu. Advances in super-resolution fluorescence microscopy for the study of nano–cell interactions. *Biomaterials science*, 9(16):5484–5496, 2021.

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

[14] M. Bates M.J. Rust and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3(10):793–795, 2006.

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

[16] Alexa L. Mattheyses, Petya Atanasova, Daniel Graupe, and Ruby May Sullan. Sharpening super-resolution by single molecule localization microscopy in front of a tuned mirror. *bioRxiv*, January 2020. Licensed under CC BY 4.0.

[17] Reid E Thompson, Daniel R Larson, and Watt W Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal*, 82(5):2775–2783, 2002.

[18] Sebastian Van de Linde, Anna Löschberger, Teresa Klein, Meike Heidbreder, Steve Wolter, Mike Heilemann, and Markus Sauer. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature protocols*, 6(7):991–1009, 2011.

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

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

[21] Susanne CM Reinhardt, Luciano A Masullo, Isabelle Baudrexel, Philipp R Steen, Rafal Kowalewski, Alexandra S Eklund, Sebastian Strauss, Eduard M Unterauer, Thomas Schlichthaerle, Maximilian T Strauss, et al. Ångström-resolution fluorescence microscopy. *Nature*, 617(7962):711–716, 2023.

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

[23] Jelmer Cnossen, Taylor Hinsdale, Carlas S. Smith, Bernd Rieger, and Sjoerd Stallinga. Localization microscopy at doubled precision with patterned illumination. *Nature Methods*, 16:1040–1043, 2019.

[24] Klaus C. Gwosch, Jasmin K. Pape, Francisco Balzarotti, Philipp Hoess, Jan Ellenberg, Jonas Ries, and Stefan W. Hell. Minflux nanoscopy delivers 3d multicolor nanometer resolution in cells. *Nature Methods*, 17:217–224, 2019.

[25] Steven M. Kay. *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall, Upper Saddle River, NJ, USA, 1993.

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

[27] Carlas S Smith, Nilesh Joseph, Bernd Rieger, and Keith A Lidke. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature methods*, 7(5):373–375, 2010.

[28] Bhanu Neupane, Fang Chen, Wei Sun, Daniel T Chiu, and Gufeng Wang. Tuning donut profile for spatial resolution in stimulated emission depletion microscopy. *Review of Scientific Instruments*, 84(4), 2013.

[29] L. Scheiderer, Z. Marin, and J. Ries. Minflux fluorescence nanoscopy in biological tissue. *Proceedings of the National Academy of Sciences*, 122(12):e2422020121, 2025.

[30] Z. Marin, L. Scheiderer, and J. Ries. Simuflux: A realistic simulation framework for minflux microscopy. *bioRxiv*, 2025. Preprint.

[31] Dylan Kalisvaart, Jelmer Cnossen, Shih-Te Hung, Sjoerd Stallinga, Michel Verhaegen, and Carlas S Smith. Precision in iterative modulation enhanced single-molecule localization microscopy. *Biophysical Journal*, 121(12):2279–2289, 2022.

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

[33] Anja Balić, Philipp Berberich, Laura A. Masullo, Patrizia D'Este, and Aleksandra Radenovic. Information-optimal point spread function engineering for deep learning-enabled localization microscopy. *Nature Communications*, 13(1):1592, 2022.

[34] Frederik Dekking, Cornelis Kraaikamp, Hendrik Lopuhaä, and Ludolf Meester. *A Modern Introduction to Probability and Statistics: Understanding Why and How*. Springer, London, 2006.