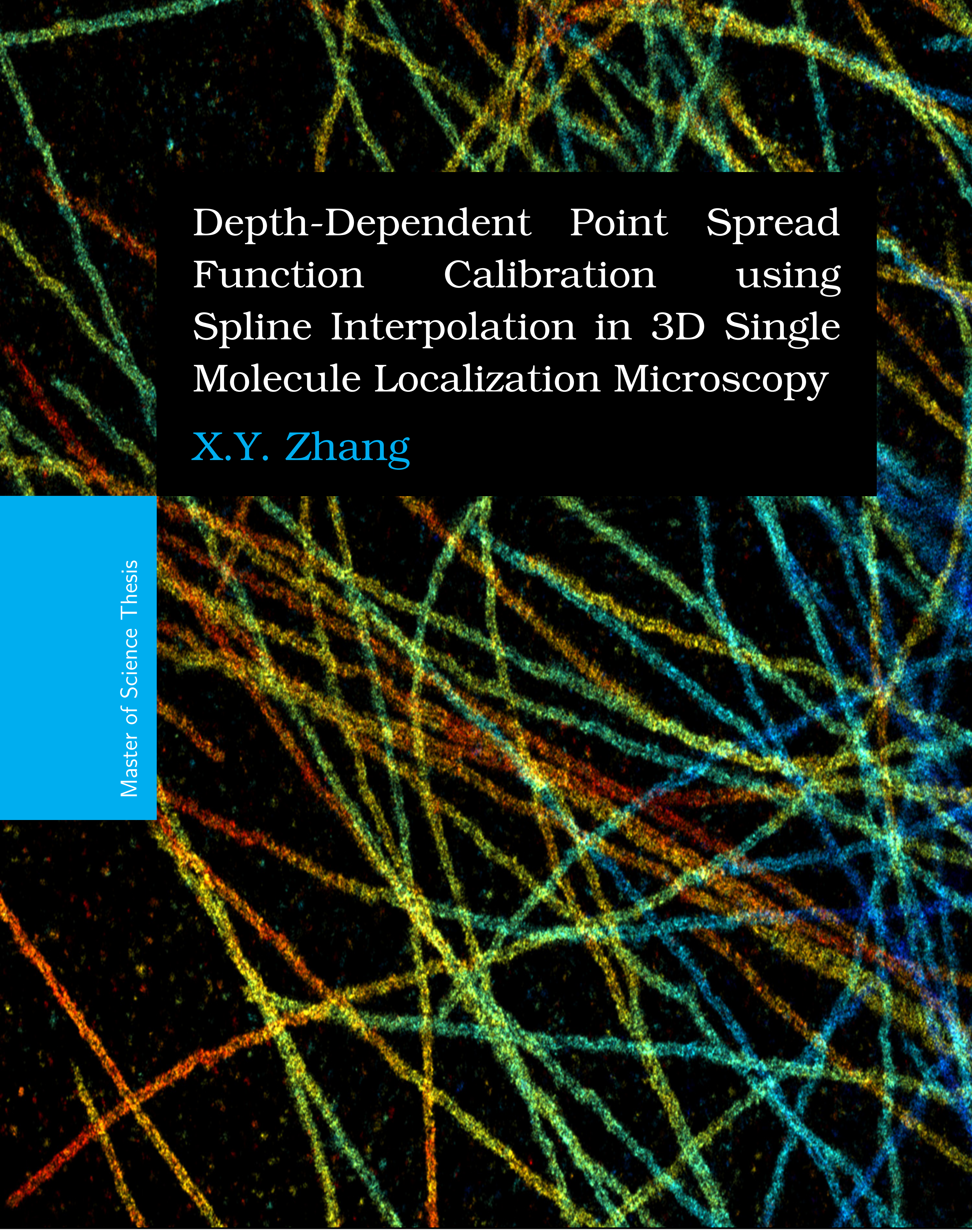


Depth-Dependent Point Spread Function Calibration using Spline Interpolation in 3D Single Molecule Localization Microscopy

X.Y. Zhang

Master of Science Thesis

Delft University of Technology

Delft Center for Systems and Control

Depth-Dependent Point Spread Function Calibration using Spline Interpolation in 3D Single Molecule Localization Microscopy

Master of Science Thesis

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

X.Y. Zhang

April 12, 2026

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

Cover page image from Ries Lab. Available at: https://rieslab.de/#gallery (accessed on 12/04/2026).

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

Depth-Dependent Point Spread Function Calibration using Spline Interpolation in 3D Single Molecule Localization Microscopy

by

X.Y. Zhang

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: April 12, 2026

Supervisor(s):

Dr.ir. C.S. Smith

Serafim Korovin

Reader(s):

Dr. N.J. Myers

Dr. O.A. Soloviev

Abstract

In three-dimensional single-molecule localization microscopy (SMLM), emitter positions are estimated by fitting a model of the microscope's Point Spread Function (PSF) to measured data. In practice, PSF models are typically calibrated using bead data acquired near the coverslip, and are assumed to remain valid representations at larger imaging depths. However, refractive index mismatch between the immersion medium, coverslip, and sample introduces depth-dependent spherical aberrations, causing the PSF shape to vary with imaging depth. As a result, a PSF model calibrated at the coverslip leads to degraded lateral localization precision and substantial axial bias when applied several micrometers deep into the sample. In this work, we introduce a depth-dependent PSF calibration approach that interpolates between calibration datasets acquired at multiple imaging depths. Calibration stacks are reconstructed at arbitrary depths using Catmull–Rom spline interpolation and used to calibrate cubic spline (cspline) models for localization. Simulations show that a conventional coverslip-calibrated model results in mean absolute axial biases exceeding 294 nm at an imaging depth of 5 μm. In contrast, the proposed approach reduces the axial bias up to 99%, consistently achieving axial bias below 5 nm. In addition, the lateral localization precision improves by 62% and 61% in x and y, respectively. Validation on experimentally acquired bead data demonstrates an axial bias reduction of 80% compared to coverslip calibration. These results show that interpolation of calibration data across depth compensates for depth-dependent PSF mismatch, enabling accurate and precise 3D localization over extended imaging depths without requiring additional optical hardware.

Table of Contents

Acknowledgements

First and foremost, I would like to thank my daily supervisor, Serafim Korovin, for his continuous weekly support and guidance throughout this master's thesis project. Additionally, without a doubt, I would like to thank my supervisor Dr.ir. Carlas Smith for his support during moments when I felt unsure about my confidence in the project, and for his willingness to help me with my graduation process. Lastly, I would like to thank all my friends and family, who have continuously supported me and instilled confidence in me.

Introduction

Many biological questions require visualizing the three-dimensional organization of cellular structures at the nanometer scale. For example, understanding how proteins assemble at synapses, how nuclear pores are distributed across the nuclear envelope, or how cytoskeletal filaments remodel in living cells contributes to our understanding of cellular function and to identifying treatments in the biology of advanced disease [36, 8].

Conventional light microscopy has been a valuable tool to visualize subcellular structures [8]. However, its fundamental constraint on spatial resolution due to diffraction means we cannot resolve structures beyond ~200–300 nm [1]. Single-molecule localization microscopy (SMLM) is a class of super-resolution imaging techniques that overcomes the diffraction limit [6, 54, 59]. In SMLM, fluorescent molecules are activated sparsely across many imaging frames so that their fluorescence spots remain spatially separated. This allows each molecule to be localized individually with high precision, and the accumulation of these localizations over time forms a high-resolution image. Accurate localization relies on fitting an appropriate model of the microscope's Point Spread Function (PSF). In SMLM, the achievable resolution is no longer primarily limited by the optical hardware, but strongly depends on the estimation precision of the localization algorithms.

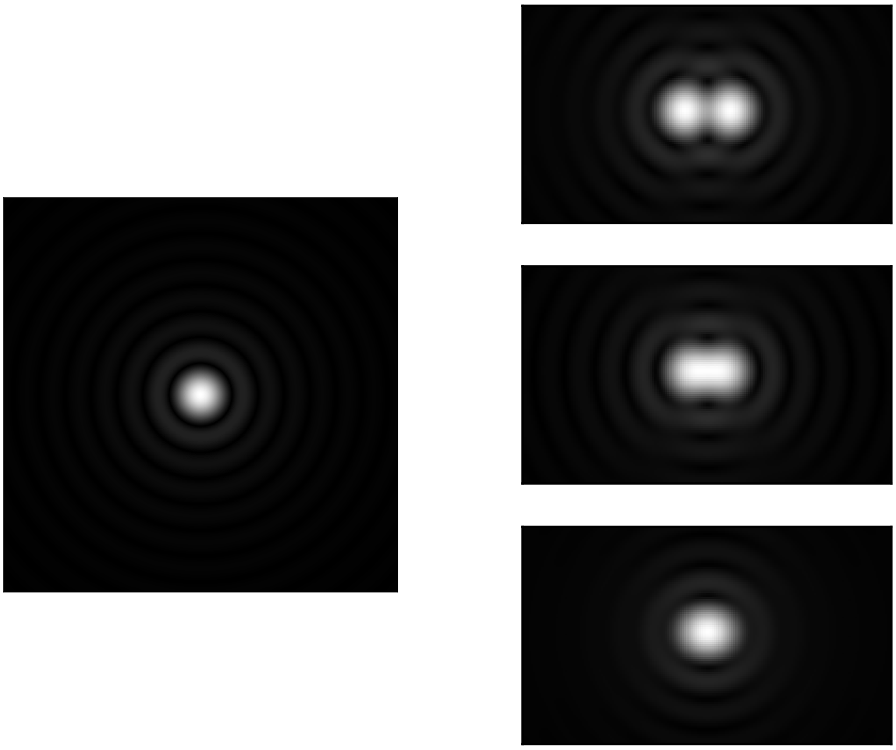
This chapter introduces the concepts of fluorescence microscopy and SMLM relevant to this thesis. Section 1.1 discusses the concept of resolution and its fundamental diffraction limit in optical microscopy. Section 1.2 explains the physical mechanisms underlying fluorescence microscopy. Section 1.3 describes the principles of SMLM, and Section 1.4 discusses how SMLM is typically extended to perform super-resolution imaging in three dimensions. Section 1.5 introduces the concept of refractive-index mismatch and explains how the resulting depth-dependent aberrations distort the PSF, leading to model mismatch between the calibrated model and the actual PSF in 3D SMLM, which represents the central problem addressed in this thesis. Lastly, Section 1.6 presents the motivation and outline of this thesis.

1.1 Diffraction-limited resolution

The spatial resolution of an optical microscope is defined as the smallest resolvable distance between two objects. Physicist Ernst Abbe discovered in 1873 that the resolution of an optical microscope is fundamentally limited by the diffraction limit of light [1]. The term diffraction has been defined by Sommerfeld and can be described as the bending of light waves at the edges of an opening or an obstacle, such as an aperture [38]. Abbe's discovery meant that ultimately, it is not the quality of the optical instrument used in an imaging system that determines the resolution, but a combination of the wavelength of light used and the aperture of the objective lens. Therefore, the smallest resolvable lateral distance d between two objects by a microscope is limited by

$$d = \frac{\lambda}{2\mathrm{NA}}, \tag{1.1}$$

where λ is the wavelength of the emitted light, $\mathrm{NA} = n \sin \alpha$ is the numerical aperture (NA) of the objective lens, n is the refractive index of the immersion medium, and α, the effective half-angle of light collected by the objective. As a consequence of the diffraction limit, an infinitesimally small light source imaged using a light microscope does not produce an image of a point, but a finite-sized spot called a PSF. This principle is visualized in Figure 1.1a.



(a) Simulated Airy disc pattern generated by a single ideal point source.

(b) Simulated Airy disc patterns generated by light from two point sources.

Figure 1.1: Simulated Airy disc patterns. (a) The PSF of an ideal point source imaged through a circular aperture forms a diffraction pattern known as the Airy disc, characterized by a bright central spot and concentric rings. (b) PSFs generated by two point sources at varying distances. From top to bottom, the sources get increasingly close together: well-separated, just barely resolvable, and too close to distinguish. This illustrates the fundamental resolution limit of optical microscopy, where diffraction causes overlapping patterns that make it difficult to resolve individual emitters.

In an ideal imaging system with a circular aperture, the PSF forms an Airy pattern. When two point sources are closer to each other than the resolution limit, their Airy patterns overlap and appear as a single spot, making it impossible to resolve the image as coming from two separate sources. This occurrence is illustrated in Figure 1.1b

As seen in Equation 1.1, the resolution limit depends on the wavelength of light and the NA of the objective. Taking into account that visible light ranges between ~$400-700$ nm and a typical NA is 1.4 [13], the diffraction limit lies roughly between ~$150-250$ nm.

This range restricts the ability of optical microscopes to resolve the subcellular organization of individual molecules or molecular complexes, which are often smaller than this limit. For example, the nuclear pore complex has a diameter of only ~120 nm [18], the viral envelope of HIV is ~100 nm [81], and biological structures such as microtubules (~25 nm), actin filaments (~7 nm), and synaptic vesicles (~40 nm) [80, 27] remain obscured by Abbe's diffraction limit in conventional microscopy.

1.1.1 Point Spread Function

The PSF describes the response of an imaging system to a point source [24]. An object can be considered as a collection of point emitters, each producing a shifted copy of the PSF. The recorded image is then formed by adding these contributions. Assuming a shift-invariant system, this can be written as a convolution of the object with the PSF [3, 55]:

$$i(x, y, z) = o(x, y, z) \otimes h(x, y, z), \tag{1.2}$$

where $\otimes$ denotes convolution, $i(x, y, z)$ is the continuous image intensity distribution, $o(x, y, z)$ the object being imaged, $h(x, y, z)$ the PSF, and (x, y, z) are the spatial coordinates.

In the frequency domain, the Fourier transform of the PSF is referred to as the optical transfer function (OTF), which describes the frequency response of the optical system. In the frequency domain, Equation 1.2 becomes a multiplication:

$$I(f_x, f_y, f_z) = O(f_x, f_y, f_z)\, H(f_x, f_y, f_z), \tag{1.3}$$

where capitals I, O, and H denote the respective Fourier transforms of the image, object, and PSF respectively, and (f_x, f_y, f_z) are spatial frequency coordinates.

1.2 Fluorescence microscopy

Over the past two decades, several super-resolution microscopy methods have been developed to overcome the diffraction limit of conventional optical microscopy and enable three-dimensional imaging of biological structures with nanometer-scale resolution [57]. Among these methods, fluorescence-based approaches play a central role, as they enable the detection of individual molecules with high contrast.

Fluorescence microscopy is widely used in the life and biomedical sciences, ranging from cell biology [21] to neuroscience [77], due to its ability to visualize biological structures and processes with high specificity, contrast, and sensitivity. It addresses the major challenge of poor contrast between the imaging sample and background, as biological samples are often optically transparent [63].

The term "fluorescence" was coined by George Stokes in 1852 [67]. In simple terms, fluorescence is the emission of light that occurs within nanoseconds after the absorption of a shorter wavelength of light. The difference between the excitation and emission wavelengths, called *Stokes shift*, is a key property that makes fluorescence a powerful contrasting tool. By filtering out the excitation light while transmitting the emitted fluorescence, it becomes possible to selectively image fluorescently labeled structures with high contrast. This contrasting mechanism is superior to that of traditional optical microscopy techniques that rely on absorption, scattering, or reflection, in which small objects often provide little optical contrast against the background [42].

A more detailed explanation of the excitation and emission processes can be illustrated using a Jablonski diagram, such as the one in Figure 1.2. Molecules with fluorescent properties are called fluorophores. When a fluorophore is illuminated with light of a suitable wavelength, it absorbs energy and transitions from the ground state S_0 to higher energy states, called excited singlet states $(S_1, S_2, \ldots, S_n)$.

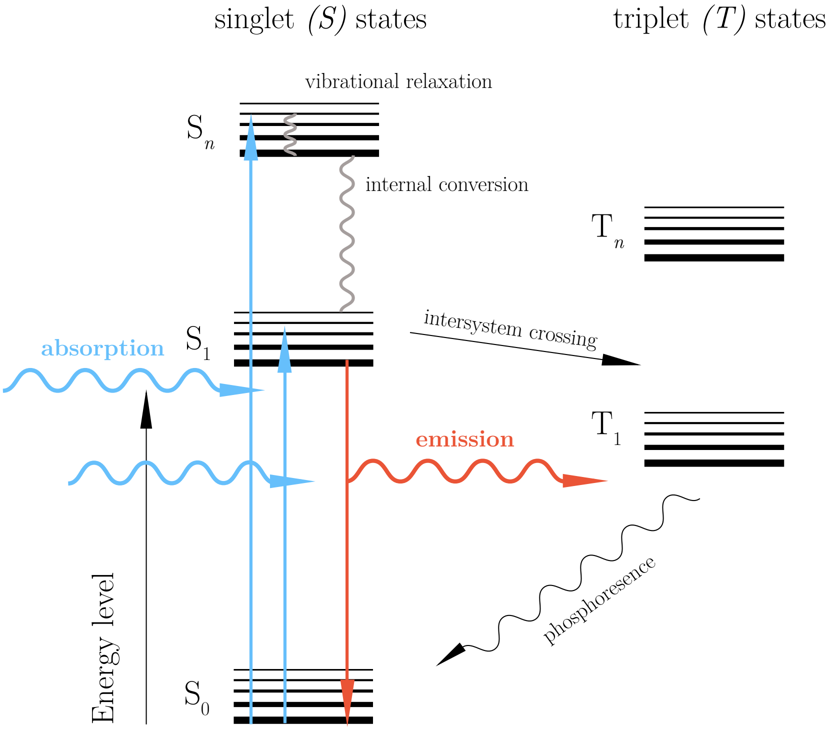


Figure 1.2: Example of a simple Jablonski diagram illustrating the states and transitions of a typical fluorophore. Upon absorption of photons, the fluorophore transitions from the ground state (S_0) to higher singlet states (S_n), followed by non-radiative relaxation to S_1. From there, it can return to S_0 via photon emission (fluorescence) or undergo intersystem crossing to a triplet state (T_1), potentially leading to phosphorescence or entering non-fluorescent dark states.

Within a few picoseconds after excitation to an excited singlet state, the fluorophore undergoes rapid vibrational relaxation to the lowest vibrational level of the first excited singlet state S_1 through non-radiative processes. From this state, it can return to the ground state S_0 either by releasing energy as emitted photons, the process known as fluorescence, or through non-radiative decay without fluorescence. This emission process forms the basis of fluorescence microscopy.

In addition to fluorescence, fluorophores can undergo transitions to non-fluorescent states.

In some cases, the fluorophore may undergo intersystem crossing to enter a triplet state $(T_1, T_2, \ldots, T_n)$. From the triplet state, it may relax back to S_0, potentially producing phosphorescence, or transition into reversible dark states (D) or long-lived dark states (LLD) through reactions with thiolate molecules [43]. Recovery to the ground state from these dark states can be influenced by adjusting the surrounding oxygen concentration.

These transitions between fluorescent and non-fluorescent states are commonly referred to as switching "ON" and "OFF". The period during which a fluorophore cycles between the ground state S_0 and the excited state S_1 while emitting photons is considered the "ON" state, whereas transition to dark states is considered the "OFF" state. If the fluorophore enters a long-lived dark state from which it cannot recover, it is said to be photobleached. Photobleaching results in permanent chemical damage to the fluorophore, preventing fluorescence even if the molecule continues to be excited.

1.3 Single-Molecule Localization Microscopy

SMLM is a class of super-resolution techniques that overcomes the diffraction limit by localizing individual fluorescent emitters. When two emitters are located close to each other, beyond the diffraction limit, their PSFs overlap, making it impossible to distinguish them as two individual emitters (as illustrated in Figure 1.1b). To address this, SMLM relies on sparsely activating or switching fluorophores such that only a subset of emitters is active at any given time. The goal is to image emitters in isolation, avoiding overlapping PSFs and enabling precise localization of single molecules. By repeating this process over many frames, the positions of many fluorescent molecules can be reconstructed to form a super-resolved image.

Most SMLM techniques sequentially image isolated subsets of fluorophores and compute their position by exploiting the photoswitching properties of fluorescent molecules. However, these techniques differ in the type of fluorophore used and in the mechanism by which "ON" and "OFF" switching is induced. The SMLM technique photoactivated localization microscopy (PALM) was introduced by Betzig et al. [6, 28, 7], and uses photoactivatable fluorescent proteins that can be switched on with UV illumination. In contrast, stochastic optical reconstruction microscopy (STORM) [54] and related techniques, such as direct stochastic optical reconstruction microscopy (dSTORM) [26], use synthetic fluorophores (dyes) that undergo photoswitching in the presence of specific buffers. A different approach is point accumulation in nanoscale topography (PAINT) [59], which achieves sparse emission not by photoswitching but through fluorophores that switch between free diffusion and immobilization by temporarily binding dyes to a target molecule. DNA-PAINT [58] is a prominent variant, where this transient binding is achieved by short dye-labeled DNA strands that hybridize with complementary target sequences.

Despite the differences in fluorophore type and switching mechanisms, all SMLM techniques rely on the common underlying principle of temporally separating emitters to enable precise localization. In practice, this requires an optical imaging system to detect fluorescence and a computational pipeline to extract molecule positions from the recorded data.

1.3.1 Widefield epifluorescence microscopy

To acquire SMLM data, an optical setup is required to excite fluorophores and detect their emission. This is typically achieved using widefield epifluorescence microscopy, a widely used configuration in fluorescence imaging and many SMLM techniques. In this setup, both the excitation and emission light paths pass through the same objective lens, as illustrated in Figure 1.3. The excitation light, typically produced by a high-intensity source such as a mercury lamp, LED lamp, or laser, is directed onto the sample using a dichroic mirror. This mirror reflects light at the excitation wavelength while transmitting light at longer wavelengths corresponding to the emitted fluorescence. Most fluorescence imaging is performed using epifluorescence microscopy due to its simplicity, wide accessibility, and compatibility with live-cell imaging [68].

The emitted fluorescence travels back through the same objective and passes through the dichroic mirror. An emission filter then removes any remaining excitation light before the fluorescence signal is recorded by a sensitive camera or photodetector. The recorded images obtained with such a setup serve as input for subsequent processing steps that detect and localize individual emitters.

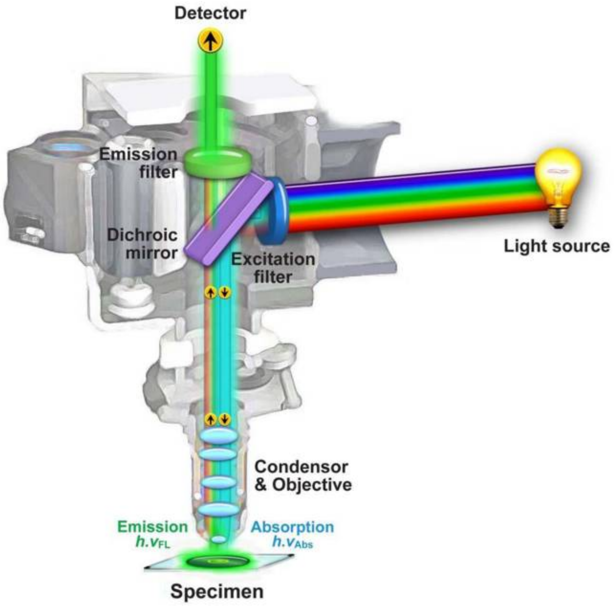


Figure 1.3: Schematic of an epifluorescence microscope. Excitation light from a broadband or laser light source passes through an excitation filter that selects the appropriate wavelength for fluorophore excitation. The filtered light is reflected by a dichroic mirror and focused onto the sample through the objective. Fluorescence from the sample travels back through the same objective and, due to the longer emission wavelength (Stokes shift), it now transmits through the dichroic mirror. An emission filter removes residual excitation light before the fluorescence signal is detected by a sensitive camera or photodetector. Source: [34]

1.3.2 Workflow of a single-molecule localization microscopy experiment

From these recorded image sequences, a typical workflow for an SMLM experiment consists of the following steps [63]: image acquisition, pre-processing, detection, localization, post-processing, and visualization. These steps are schematically shown in Figure 1.4.

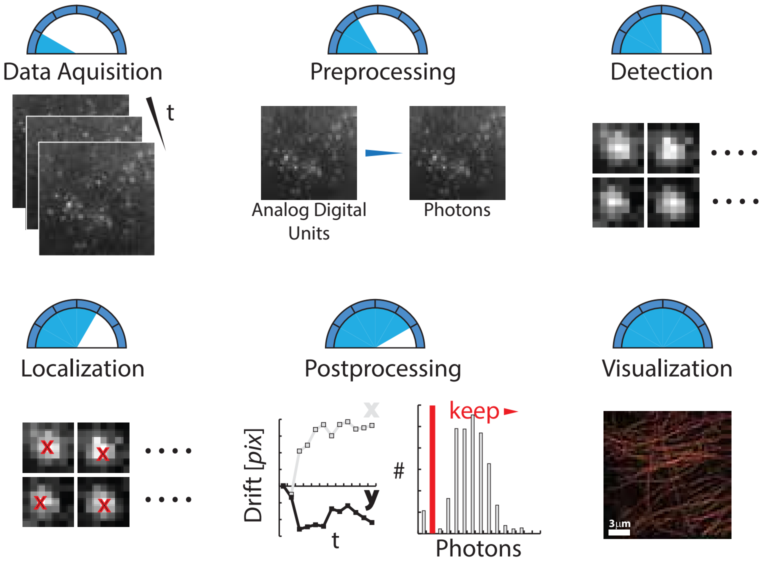


Figure 1.4: Overview of the SMLM image processing pipeline. In the data acquisition step, a sample labeled with fluorescent emitters is imaged over time, with only a subset of emitters active in each frame due to blinking or photobleaching. During pre-processing, measured intensities are converted into photon counts suitable for quantitative analysis. The detection step identifies regions of interest containing single emitters. In the localization step, the precise emitter positions are estimated for each detected region. Post-processing includes quality control, drift correction, and merging of multiple localizations. Finally, the results are visualized to reconstruct a super-resolution image. Source: [63].

Data acquisition

SMLM experiments typically require imaging millions of fluorophores over $10,000 - 100,000$ camera frames to reconstruct a super-resolution image [63]. The number of fluorophores active in each frame should be maximized to reduce acquisition time, while remaining low enough to ensure isolated PSFs suitable for localization.

The isolated subset of fluorophores in each frame is achieved through photoswitching or binding kinetics, depending on the specific method (e.g., PALM, STORM, PAINT). The acquisition speed depends on both the switching rate of the fluorophores and the frame rate of the camera, with frame rates up to approximately $3,200$ frames per second reported for modern systems [32].

The switching rate of fluorescent dyes such as Cy5 and Alexa Fluor 647 can be increased at higher laser powers [26]. Standard continuous-wave lasers provide excitation and activation, and the emitted photons are detected with high-sensitivity cameras, such as an electron-multiplying charge-coupled device (EM-CCD) or a scientific complementary metal-oxide semiconductor (sCOMS) camera.

Preprocessing

The recorded images contain noise and camera-specific artifacts that must be corrected before detection and localization can be performed. Unlike EM-CCDs, images acquired with sCOMS cameras exhibit non-negligible and pixel-dependent readout noise. As a result, SMLM algorithms, originally designed for Poisson-distributed and pixel-independent noise in EM-CCDs [47, 64], cannot provide reliable position estimates without per-pixel calibration. Per-pixel calibration of offset, gain, and variance is therefore required to avoid systematic localization errors [32]. Background signal is often removed to reduce heterogeneous background noise [38], for example, by subtracting the background averaged over the entire image sequence [14].

Detection

Each frame is analyzed to identify regions of interest (ROIs) that potentially contain single emitters [63]. ROIs can be selected by extracting local intensity maxima using a threshold relative to the background level. The resulting ROIs often correspond to small pixel regions that likely contain single molecules and serve as the input for the subsequent localization step.

The performance of the detection algorithm can be described by two quantities [63]: the detection efficiency (also called sensitivity or recall), defined as the ratio of correctly detected events to the total number of true events, and the false-positive rate, defined as the ratio of all false detections over the total number of detections. In practice, both false positives and false negatives occur due to noise [35].

Localization

For each region of interest (ROI), the position of a single molecule is estimated with subpixel accuracy by fitting an image formation model to the data [63]. Estimation algorithms in localization microscopy include least-squares fitting and maximum likelihood estimation (MLE). Under appropriate conditions and with an accurate noise and image formation model, MLE can approach the theoretical lower bound on the localization precision given by the Cramèr-Rao lower bound (CRLB) [47, 16, 62].

Accurate localization, therefore, depends on the choice of the image formation model and, in particular, on the PSF model used during fitting. Here, the term PSF refers to the physical image of a point emitter formed by the microscope, while the PSF model denotes the mathematical representation used to describe this image during localization. The accuracy of localization depends on how well this model captures the observed PSF shape. In practice, Gaussian approximations of the PSF are widely used due to their computational simplicity. Since this thesis focuses on the localization step, a detailed description of parameter estimation using an MLE algorithm is given in Chapter 2.

Post-processing

The post-processing step can be divided into three parts [63]. First, quality control is performed, and estimated parameters from the localization step are compared with those from the detection step to filter out unreliable detections. This ensures a high-quality reconstructed super-resolution image. Second, drift correction compensates for mechanical or thermal drifts of the microscope stage or moving emitters during the acquisition process. The final step combines the individual localizations to construct a super-resolution image. In many SMLM experiments, a single molecule remains “ON” across several camera frames and therefore can be localized multiple times. Merging multiple localizations of the same molecule can improve localization precision [63].

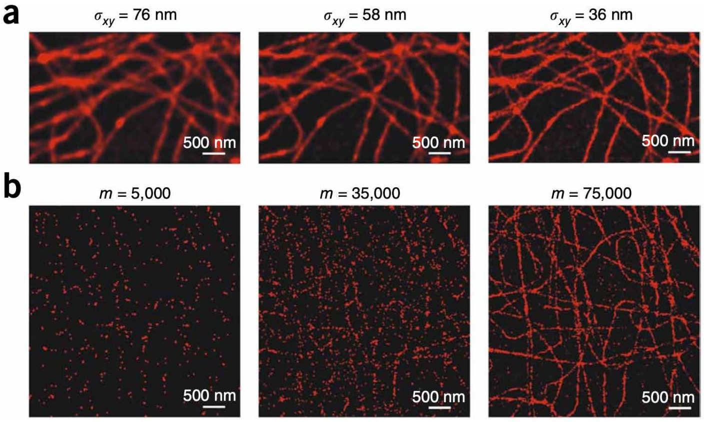


Figure 1.5: Experimental images of microtubules in HeLa cells are shown, illustrating the influence of localization precision and density of fluorescently labeled molecules on the resolution in localization microscopy. In (a), it can be seen that as the localization precision of the estimated $\sigma_{x,y}$ position of the emitter improves (decreasing values of $\sigma_{x,y}$), the image results in less blur. In (b), the images show that an increased density of fluorescent labels (increasing value of the number of localized emitters m) results in many more emitters being visualized. Source: [16]

Visualization

In the final step, the localized positions are combined to form a super-resolution image. One common approach is to define a grid with bin sizes similar to the localization precision, count the number of localizations in each bin, and convert these counts into pixel intensities [38].

Alternatively, each localization can be represented as a two-dimensional Gaussian profile, centered at the estimated position. The width of the Gaussian reflects the localization precision of that position and is often normalized as a probability density [6, 28, 63].

As a result of super-resolution microscopy techniques such as SMLM, the resolution of an SMLM image is no longer limited by Abbe's diffraction limit. Instead, it is determined by the precision of the localization algorithm and the density of fluorescent labels across the sample [48]. This relationship can be seen in Figure 1.5.

1.4 3D Single-molecule localization microscopy using PSF engineering

In conventional SMLM, the lateral (x, y) positions of single emitters are estimated from the recorded images, while information along the optical axis (z-dimension) is not directly encoded in the image. As a result, emitters located at different axial positions can produce images with similar PSF shapes, making it difficult to determine their axial position relative to the focal plane. To overcome this limitation, SMLM was extended early on to estimate the 3D positions of single-molecule emitters through various microscope modifications. This

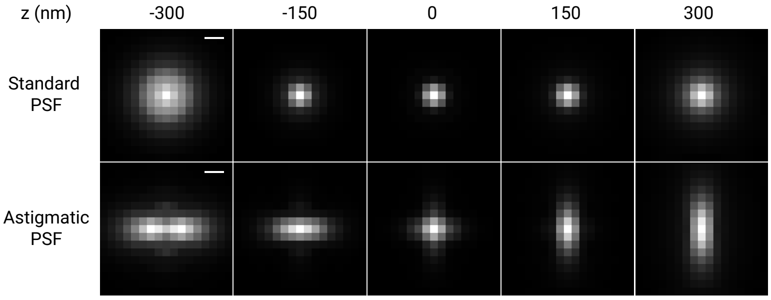


Figure 1.6: Simulated PSFs. A standard PSF across a z-range of -300 to 300 nm from the focal plane. The PSF shape is symmetric above and below the focal plane. Astigmatic PSF simulated with an astigmatism magnitude of 60 nm (rms wavefront amplitude) across the same z-range. Simulations are carried out with: $\lambda = 671$ nm, pixel size= 100 nm, NA= 1.43, $n_{med} = 1.43$, $n_{cov} = 1.51$, and $n_{imm} = 1.52$. Scale bar = 250 nm.

thesis focuses on PSF engineering, a common approach for encoding axial position because it is relatively easy to implement and robust [15]. The principle of PSF engineering relies on introducing suitable aberrations, such that the z-position of the single molecule is encoded into the shape of the PSF. Here, z denotes the axial coordinate (along the optical axis) measured relative to the focal plane (typically at $z = 0$), and positive z indicates positions above the focal plane toward the objective. The optical axis is the central axis of the imaging system, perpendicular to the image plane.

In a standard wide-field epifluorescence microscope, the PSF changes almost symmetrically above and below the focal plane. This symmetry introduces ambiguity about whether we are imaging above or below the focal plane [75] and is illustrated in Figure 1.6, which compares standard and astigmatic PSF shapes across an axial range. This ambiguity can be resolved by a technique known as PSF engineering, in which the axial position of the emitter is encoded in deliberate changes to the PSF shape. Various engineered PSF shapes have been proposed to encode axial information into the PSF. This thesis focuses on astigmatic PSFs. A simple way to generate an astigmatic PSF is by adding an optical component, such as a cylindrical lens or a glass wedge [5], to induce astigmatism [29, 30]. This generates PSFs that are circular in the focal plane, and become elliptical in either the x or y direction as a function of the axial position, either above or below the focal plane. This ellipticity is visible in the same figure.

For astigmatic PSFs, the axial position can be estimated by relating measurable PSF parameters to the emitter's z-position through a calibration curve [75], which can be derived by fitting a PSF model to an axial stack of PSF images. These images are obtained by imaging fluorescent beads embedded in the sample at different focal planes. Imaging is achieved by changing the distance between the objective and the sample. For astigmatic PSFs, the calibration curve is typically constructed using the width of the PSF in the x and y directions (w_x and w_y), which vary with the axial position. The z-position can be estimated by matching the measured pair (w_x, w_y) to the corresponding value in the calibration curve [31]. Figure 1.7 shows an example of such a calibration curve, illustrating how the PSF widths change as a function of the axial position and enable estimation of the emitter's z-position.

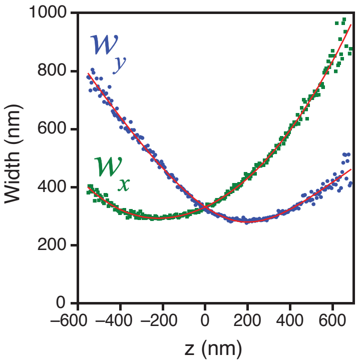


Figure 1.7: An example of a calibration curve for an astigmatic PSF. The x and y widths, w_x, w_y respectively as a function of z-position. Source: [30].

.

1.5 Refractive index mismatch induced aberrations

Refraction is the bending of light rays when they pass between media with different refractive indices, due to a change in the propagation speed of light [25, 23]. Most commonly, refraction occurs at an interface between two optical media. When these media have different refractive indices, this is referred to as a refractive index mismatch. As illustrated in Figure 1.8, incident light rays in a medium with refractive index n_1 are bent at the interface with a second medium of refractive index n_2. The angles of incidence and refraction, θ_1 and θ_2 respectively, are related by Snell's law: $n_1 \sin\theta_1 = n_2 \sin\theta_2$.

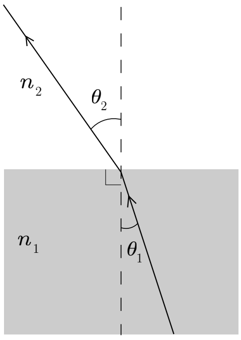
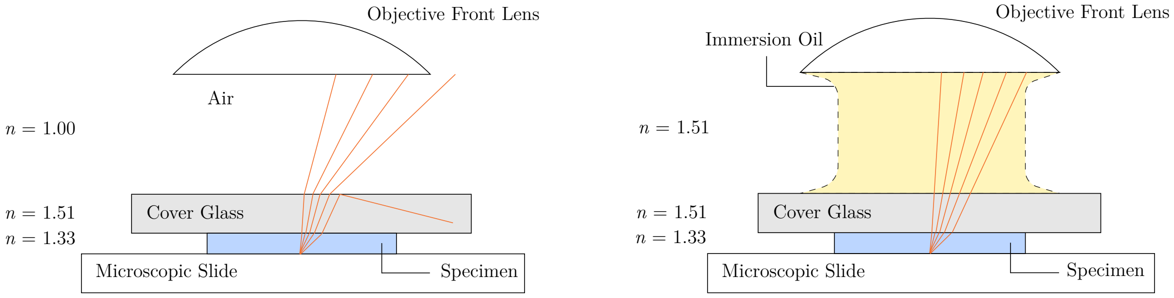


Figure 1.8: Illustration of refraction at the interface between two media with refractive indices n_1 and n_2. In this case, $n_2 < n_1$, and therefore the refracted ray bends away from the normal, resulting in $\theta_2 > \theta_1$.

High-resolution fluorescence microscopy, including SMLM, commonly relies on objectives with a high NA. A higher NA increases the light-gathering ability of the objective and improves the resolving distance according to Abbe's expression in Equation (1.1). One way to achieve high NA is to use immersion objectives, such as oil-immersion lenses, in which the immersion medium has a higher refractive index than air. However, when the refractive indices of the objective immersion and sample-mounting media differ, a refractive-index mismatch occurs within the imaging system. This mismatch introduces optical aberrations that distort the observed PSF. As a result, the measured PSF deviates from the PSF model used for localization, reducing localization accuracy in 3D imaging.

Figure 1.9 contains two schematic drawings of microscopes with different objective immersion media. In Figure 1.9a, an air objective lens with $n = 1.00$ collects a smaller fraction of the emitted light, as high-angle light rays may undergo total internal reflection at the interface between the sample and glass coverslip. For an oil-immersion objective with $n = 1.51$, light is collected more efficiently. However, this also introduces a refractive index mismatch between the immersion and sample-mounting media (typically aqueous-based). Such mismatches lead to spherical aberrations, whose effects become more pronounced with increasing imaging depth, distorting the PSF and introducing a mismatch between the observed PSF and the model used for localization [11].



(a) Air objective.

(b) Oil immersion objective.

Figure 1.9: Comparison of schematic microscope imaging configurations. (a) When an air objective is used, no immersion medium is present between the coverslip and the objective lens. This results in a lower numerical aperture (NA) and reduced light collection efficiency. (b) With an oil immersion objective, the immersion oil (with refractive index $n = 1.51$) approximately matches the refractive index of the coverslip and objective lens.

This effect can be understood by considering how variations in refractive index alter the optical path length of the emitted wavefront. Light emitted by a single molecule propagates through the sample and microscope optics before reaching the camera. Variations in refractive index within the optical system cause the emitted light to experience different optical path lengths (OPLs) [10, 22]. As light passes through different media, it travels at different speeds depending on the refractive-index distribution, resulting in phase delays across the wavefront and, therefore, a distorted wavefront [17].

1.5.1 Zernike polynomials

The phase aberration function, which describes the distortion to the wavefront, can be mathematically represented as a weighted sum of basis functions in polar coordinates as

$$\phi(\rho, \theta) = \sum_k a_k X_k(\rho, \theta), \tag{1.4}$$

where a_k are the coefficients that quantify the contribution of the basis functions $X_k(\rho, \theta)$, often referred to as modes. In principle, any suitable set of basis functions may be used, but in optics, these modes are often chosen for mathematical convenience to be Zernike polynomials introduced by Zernike [78, 23]. Zernike polynomials are defined in polar coordinates (ρ, θ), where $\rho \in [0, 1]$ is the normalized radial coordinate and $\theta \in [0, 2\pi)$ is the azimuthal angle. They are parametrized by two integers n and m, where $|m| \leq n$, and $n - |m|$ is even. The corresponding Zernike polynomial is defined as

$$Z_n^{\pm m}(\rho, \theta) = R_n^m(\rho) \begin{cases} \cos(m\theta) & \text{if } m > 0 \\ \sin(|m|\theta) & \text{if } m < 0 \\ 1 & \text{if } m = 0 \end{cases}. \tag{1.5}$$

Here, the radial polynomial $R_n^m(\rho)$ is given by

$$R_n^m(\rho) = \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} \rho^{n-2s}. \tag{1.6}$$

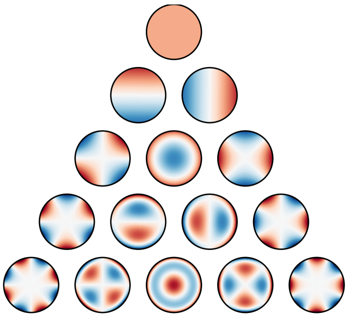
Zernike polynomials form an orthogonal basis on the unit disk and therefore:

$$\int_0^{2\pi} \int_0^1 Z_n^m(\rho, \theta)\, Z_{n'}^{m'}(\rho, \theta)\, \rho\, d\rho\, d\theta = \frac{\pi}{n+1}\, \delta_{nn'}\, \delta_{mm'}. \tag{1.7}$$

Though they are defined in polar coordinates, Zernike polynomials can be expressed as closed-form polynomials in Cartesian coordinates (see Table 1.1). Each polynomial in the basis corresponds closely to well-known aberrations such as defocus, astigmatism, coma, or spherical aberrations [10, 50]. In Figure 1.10a, the Zernike polynomials are depicted with two indices and arranged in a pyramid structure, traditionally done for presentation purposes [73].

$$\begin{matrix} \mathrm{Z}_0^0(x,y) \\ \mathrm{Z}_1^{-1}(x,y) \quad \mathrm{Z}_1^1(x,y) \\ \mathrm{Z}_2^{-2}(x,y) \quad \mathrm{Z}_2^0(x,y) \quad \mathrm{Z}_2^2(x,y) \\ \mathrm{Z}_3^{-3}(x,y) \quad \mathrm{Z}_3^{-1}(x,y) \quad \mathrm{Z}_3^1(x,y) \quad \mathrm{Z}_3^3(x,y) \\ \mathrm{Z}_4^{-4}(x,y) \quad \mathrm{Z}_4^{-2}(x,y) \quad \mathrm{Z}_4^0(x,y) \quad \mathrm{Z}_4^2(x,y) \quad \mathrm{Z}_4^4(x,y) \end{matrix}$$

(a) Zernike polynomials in a pyramid up to $m, n = 4$.



(b) Phase functions of Zernike polynomials up to radial degree $n, m = 4$. Adapted from an image by Tatoute[1].

Figure 1.10: Side-by-side representation of Zernike polynomials: symbolic vs. visual heatmap.

In this thesis, this representation is useful because refractive-index mismatch predominantly introduces spherical aberration, whose magnitude increases with imaging depth [9, 51].

---

[1] Source: Wikimedia Commons, licensed under CC BY-SA 4.0.

Noll's ordering (j)	n	m	Aberration name	Polar form $Z_j(r,\theta)$	Cartesian form $Z_j(x,y)$
1	0	0	Piston	1	1
2	1	1	Tilt	$2r\cos\theta$	$2x$
3	1	-1	Tilt	$2r\sin\theta$	$2y$
4	2	0	Defocus	$\sqrt{3}(2r^2-1)$	$\sqrt{3}(2x^2+2y^2-1)$
5	2	2	Astigmatism	$\sqrt{6}r^2\cos 2\theta$	$\sqrt{6}(x^2-y^2)$
6	2	-2	Astigmatism	$\sqrt{6}r^2\sin 2\theta$	$\sqrt{6}(2xy)$
7	3	-1	Coma	$\sqrt{8}(3r^3-2r)\sin\theta$	$\sqrt{8}(3x^2y+3y^3-2y)$
8	3	1	Coma	$\sqrt{8}(3r^3-2r)\cos\theta$	$\sqrt{8}(3x^3+3xy^2-2x)$
9	3	3	Trefoil	$\sqrt{8}r^3\cos 3\theta$	$\sqrt{8}(x^3-3xy^2)$
10	3	-3	Trefoil	$\sqrt{8}r^3\sin 3\theta$	$\sqrt{8}(3x^2y-y^3)$
11	4	0	Spherical	$\sqrt{5}(6r^4-6r^2+1)$	$\sqrt{5}(6(x^2+y^2)^2-6(x^2+y^2)+1)$

Table 1.1: Noll's single-index ordering of the Zernike polynomials up to $j = 11$, including radial degree n, azimuthal frequency m, standard aberration name, and expressions in polar and Cartesian coordinates.

1.5.2 Spherical aberrations

Spherical aberration occurs when light rays passing through a lens at different distances from the optical axis do not converge at the same focal point. The optical axis is the central axis of the imaging system, passing through the optical elements. The z-axis of object and image space coincides with this axis [56], as illustrated in Figure 1.11. In fluorescence microscopy, the magnitude of spherical aberration increases with imaging depth when there is a refractive-index mismatch among the objective immersion medium, coverslip, and sample mounting medium [51].

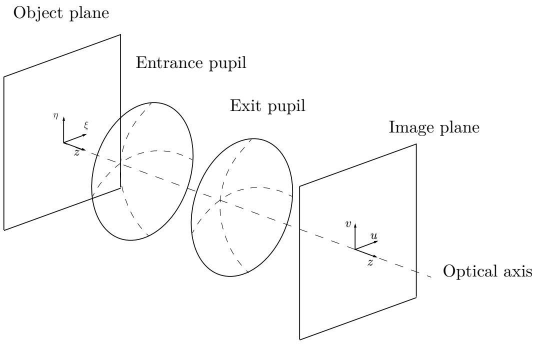


Figure 1.11: Schematic illustration of the geometric definitions used in optical systems. Light propagates from the object plane with coordinates (ξ, η, z), through the entrance pupil and exit pupil to the image plane with coordinates (u, v, z). The optical axis is the central line passing through the system, coinciding with the z-axis in both object and image coordinates.

1.5.3 Effects of refractive index mismatches

In microscopy, refractive index mismatches commonly occur in two scenarios [17]: either when an oil-immersion objective is used to image an aqueous live-cell sample, or when an air-immersion objective is used to image a higher-refractive-index mounting medium or tissue-clearing solution. In both cases, spherical aberrations arise. As a consequence, the detected PSF becomes distorted, leading to image blurring and signal loss as the three-dimensional PSF becomes axially elongated [10].

Additionally, refractive index mismatches introduce a focal shift [17], meaning that the physical position does not correspond directly to the nominal focal position set by objective displacement. As a result, the axial scaling of the image becomes distorted. If the refractive index of the objective immersion medium is higher than that of the sample mounting medium, the imaged structures appear axially elongated. Conversely, when the immersion medium has a lower refractive index than the sample, the structures appear axially compressed.

These depth-dependent distortions cause a mismatch between the PSF used during calibration (often measured near the coverslip) and the PSF observed at greater depths within the sample. Consequently, if not accounted for, this leads to systematic localization errors (bias) and compromised three-dimensional localization precision in 3D imaging [41, 51].

1.6 Thesis motivation

This section summarizes the challenge of refractive index mismatch-induced aberrations in 3D SMLM that motivates this thesis, with particular focus on their impact on localization bias and precision. Three-dimensional super-resolution imaging methods based on SMLM enable nanoscale imaging beyond Abbe's diffraction limit. As discussed in Subsection 1.3.2, the localization precision of single emitters is a key factor determining the final image quality in localization microscopy.

In 3D SMLM with PSF engineering, it is typically assumed that a PSF model calibrated on data acquired near the coverslip generalizes to data acquired deeper within the sample. However, biological imaging is largely conducted in aqueous mounting media with a refractive index of approximately 1.33, whereas the glass coverslip on which samples are mounted, and the oil in which the objective lens is immersed, typically have refractive indices of approximately 1.51 and 1.52, respectively. This refractive index mismatch causes emitted fluorescence to experience different optical path lengths as it propagates through the optical system, resulting in distorted wavefronts that manifest as depth-dependent spherical aberrations.

These depth-dependent spherical aberrations decrease the peak intensity of the PSF as emitted photons spread over a larger area, effectively elongating the 3D PSF. As a result, the axial cross-section of the PSF profile becomes increasingly asymmetric with respect to the focal position at imaging depths of only a few micrometers [41, 51]. Additionally, refractive index mismatch introduces a focal shift, where the physical distance between an emitter in the sample and the coverslip (the actual focus position) does not match the displacement required by the objective to bring the emitter into focus (the nominal focus position). Applying a PSF model calibrated near the coverslip to such aberrated data results in reduced localization precision in the lateral coordinates and increased bias in the estimated axial position.

Several approaches have been proposed to mitigate the effects of depth-dependent aberrations or to accurately model the PSF. These include active aberration correction using Adaptive Optics (AO) [12, 46], and physically derived [61] or phase-retrieved PSF models

[44, 24, 79]. While these methods can accurately model or compensate for aberrations, they often require specialized optical hardware, dedicated expertise to implement complex AO, or calibration of numerous optical system parameters, such as objective lens properties, refractive indices, aberration functions, and emission wavelengths. These requirements can limit their widespread adoption.

An alternative approach was proposed by Li *et al.* [41], which used a cubic spline (cspline) approach to create a look-up-table (LUT) to correct bias in estimated axial positions by interpolating between correction values across different imaging depths. This approach enables accurate axial localization without requiring hardware modifications. However, since the underlying PSF model remains mismatched to the experimentally observed PSF, the resulting degradation in localization precision remains uncorrected.

This thesis proposes a depth-dependent calibration approach to address this model mismatch. By interpolating between calibration data at multiple imaging depths using Catmull–Rom splines, and evaluating the interpolation at an intermediate depth, we can obtain an accurate PSF model at arbitrary intermediate depths within this interpolated range. Additionally, it investigates how the interpolated calibration performs compared to using the nearest available calibration to the evaluation depth. This thesis aims to provide a hardware-free solution that mitigates both localization bias and precision loss arising from a mismatched model due to refractive-mismatch-induced aberrations.

The remainder of this thesis is organized as follows. Chapter 2 introduces the estimation-theory concepts underlying localization in SMLM, including MLE and the theoretical lower bound on achievable variance, the CRLB. Chapter 3 provides an introduction to spline functions and establishes the notation used throughout the remainder of the thesis, as spline models play a central role in the proposed calibration approach. Chapter 4 presents the proposed depth-dependent calibration method in the form of a manuscript titled *Depth-Dependent Point Spread Function Calibration via Spline Interpolation in 3D Single-Molecule Localization Microscopy*, and evaluates the method using simulated and experimental data. The chapter also includes supplementary material with additional figures and a detailed explanation of the interpolation procedure. Finally, Chapter 5 summarizes the main findings of this thesis and discusses possible directions for future work.

Estimation preliminaries for single-molecule localization

This chapter introduces the estimation concepts required for the thesis. It became clear from Chapter 1 that the resolution of single-molecule localization microscopy (SMLM) images is no longer limited by Abbe's diffraction limit, but instead depends on the precision of the localization algorithm and the density of the fluorescently labeled molecules [16]. Achieving this requires a localization algorithm that can estimate the parameters of individual molecules, such as their positions and the number of photons they emit.

Localization in SMLM can be formulated as an inverse problem: given the recorded signal on the detector pixels, the task is to estimate the underlying parameters of the single molecule that generated the observed image [62, 75]. To do so, an image formation model that links single-molecule parameters to the signal measured by the detector is required. Moreover, an estimation algorithm is needed that recovers these parameters from noisy measurements.

This chapter begins by describing the image formation model that determines how a single emitter appears on the camera detector. In Section 2.2 we introduce several approaches for modeling the Point Spread Function (PSF). Sections 2.3 and 2.4 discuss how a maximum likelihood estimation (MLE) algorithm can be used in combination with a modified Levenberg-Marquardt (LM) algorithm to estimate the single-molecule parameters. Lastly, Section 2.5 discusses several metrics for evaluating the localization performance of the estimated single-molecule parameters.

2.1 Image Formation Model

The estimation of single-molecule parameters is commonly performed by fitting an appropriate image formation model to the observed signal recorded by the camera pixel [62]. A commonly used image formation model is given by

$$\mu_k = \theta_I \,\mathrm{PSF}(\vec{r}_k - \vec{r}_0) + \theta_b, \tag{2.1}$$

where μ_k is the expected photon count at the k^{th} pixel, $\vec{r}_0 = (\theta_x, \theta_y, \theta_z)$ denotes the spatial coordinates of the single-molecule emitter, $\vec{r}_k$ denotes the spatial coordinate of the center of pixel k, θ_I is the total number of photons emitted by the molecule, θ_b is the background photon count per pixel, and PSF is the point spread function of the imaging system. The emitter parameters to be estimated can be combined into the following parameter vector:

$$\vec{\theta} = \begin{pmatrix} \theta_x & \theta_y & \theta_z & \theta_I & \theta_b \end{pmatrix}^T . \tag{2.2}$$

The number of photons detected during a camera exposure is typically modeled as a random variable following a Poisson distribution [45]. Let c_k denote the observed photon count in the k^{th} pixel. The recorded image $\vec{c} = (c_1, \ldots, c_K)$ is therefore a noisy realization of the expectation predicted by the image formation model, where K denotes the total number of pixels in the region of interest (ROI), and each pixel value is distributed as $c_k \sim \text{Poisson}(\mu_k)$.

This formulation casts single-molecule localization as an inverse problem: given the observed pixelated image $\vec{c}$, the task is to recover the underlying emitter parameters $\vec{\theta}$ such that the expected signal $\mu_k(\vec{\theta})$ best resembles the observed measurement.

2.2 Point Spread Function Models

In the image formation model in Equation 2.1, $\text{PSF}(\vec{r}_k - \vec{r}_0)$ represents a model of how photons emitted by a molecule are distributed over the detector pixels. In practice, this model approximates the true PSF of the optical system. Differences between the observed photon counts c_k and the expected counts μ_k predicted by the model arise from measurement noise as well as mismatches between the assumed PSF model and the true PSF [2]. In this section, we introduce several approaches to approximate the PSF model.

2.2.1 2D Gaussian PSF Model

The PSF of an imaging system can be approximated by a two-dimensional Gaussian distribution

$$\text{PSF}(x, y) = \frac{1}{2\pi\sigma_0^2} \exp\left(-\frac{(x - \theta_x)^2}{2\sigma_0^2} - \frac{(y - \theta_y)^2}{2\sigma_0^2} \right), \tag{2.3}$$

where σ_0 denotes the standard deviation of the Gaussian distribution in 2D and (θ_x, θ_y) the spatial coordinates of the emitter. Gaussian PSF models are widely used because of their conceptual simplicity and computational efficiency, despite being only an approximation to the physical diffraction-limited PSF [66].

2.2.2 3D Gaussian Astigmatic PSF Model

The Gaussian approximation of the PSF in Equation 2.3 can be extended to enable 3D localization through PSF engineering. In astigmatic 3D imaging, the PSF becomes elliptical in the x and y directions depending on the axial position of the emitter. The resulting PSF model approximated with a Gaussian as follows [64]:

$$\text{PSF}(x, y) = \frac{1}{2\pi\sigma_x(\theta_z)\sigma_y(\theta_z)} \exp\left(-\frac{(x - \theta_x)^2}{2\sigma_x(\theta_z)^2} - \frac{(y - \theta_y)^2}{2\sigma_y(\theta_z)^2} \right). \tag{2.4}$$

Here $(\theta_x, \theta_y, \theta_z)$ denote the spatial coordinates of the emitter. The standard deviations $\sigma_x(\theta_z)$ and $\sigma_y(\theta_z)$ depend on the axial position z as follows:

$$\sigma_x(z) = \sigma_{0x}\sqrt{1 + \frac{(z-\gamma)^2}{d^2} + A_x\frac{(z-\gamma)^3}{d^2} + B_x\frac{(z-\gamma)^4}{d^2}}, \tag{2.5}$$

$$\sigma_y(z) = \sigma_{0y}\sqrt{1 + \frac{(z-\gamma)^2}{d^2} + A_y\frac{(z-\gamma)^3}{d^2} + B_y\frac{(z-\gamma)^4}{d^2}}, \tag{2.6}$$

where σ_{0x} and σ_{0y} denote the in-focus standard deviations in the x and y directions, d is the parameter related to the depth of focus that influences the axial dependence of the PSF widths and must be determined through system identification [64], and A_x, A_y, B_x, and B_y are empirical constants determined experimentally.

2.2.3 Vectorial PSF Model

For imaging systems equipped with high-numerical aperture (NA) objectives, the vector nature of light cannot be neglected, and a vectorial PSF model is required. The vectorial PSF model accounts for the dipole orientation of a single emitter, the effects of high NA objectives, interfaces between media, polarization, and optical aberrations [66].

The emitted light can be described by the electric field E_{dipole}. This electric field propagates from the dipole located in the sample mounting medium through the coverslip and immersion medium with refractive indices n_{med}, n_{cov}, and n_{imm}, respectively. The emitted field is collected by the objective and mapped to the electric field in the pupil plane, denoted by E_{pupil}.

The electric field component $l = x, y$ in the image plane is proportional to the emission dipole component $j = x, y, z$, and can be expressed as the Fourier transform of the electric field components in the pupil plane [61, 72]:

$$E_{\text{pupil},lj}(W, \vec{\rho}, d, z_{\text{stage}}) = A(\vec{\rho})q_{lj}(\vec{\rho})\exp\left(iW(\vec{\rho}) + i\left[dk_{z,0}(\vec{\rho}) - z_{\text{stage}}k_{z,2}(\vec{\rho})\right]\right), \tag{2.7}$$

where $\vec{\rho} = (\rho_x, \rho_y)$ are the normalized pupil coordinates. The term $A(\vec{\rho})$ denotes the amplitude factor [76, 69], which includes the well-known aplanatic correction factor [61]:

$$A(\rho) = \sqrt{\frac{n_{\text{imm}}\cos(\theta_{\text{imm}})}{n_{\text{med}}\cos(\theta_{\text{med}})}}\frac{1}{\sqrt{n_{\text{med}}\cos(\theta_{\text{med}})}} = \frac{\left(n_{\text{imm}}^2 - \text{NA}^2\rho^2\right)^{1/4}}{\left(n_{\text{med}}^2 - \text{NA}^2\rho^2\right)^{1/2}}, \tag{2.8}$$

where NA is the numerical aperture of the objective lens. The terms $q_{lj}(\vec{\rho})$ are polarization vector components defined in [65]. The function $W(\vec{\rho})$ represents the aberration function expressed as a sum of Zernike polynomials. The parameter d denotes the imaging depth, defined as the distance between the coverslip and the focal plane, and z_{stage} represents the position of the objective stage relative to the focal plane.

The wave vector $k(\vec{\rho})$ is a function of the normalized pupil coordinates defined as:

$$\vec{k}(\vec{\rho}) = (k_x(\vec{\rho}),\, k_y(\vec{\rho}),\, k_{z,i}(\vec{\rho})) = \frac{2\pi}{\lambda}\left(\text{NA}\rho_x,\, \text{NA}\rho_y,\, \sqrt{n_i^2 - \text{NA}^2\|\vec{\rho}\|_2^2}\right), \tag{2.9}$$

where n_i denotes the refractive index of medium i and λ is the wavelength of the emitted light. Lastly, to obtain the resulting incoherent PSF from emission of a freely-rotating dipole emitter, the six Fourier transforms of the electric field components in the pupil plane are quadratically summed:

$$H(\vec{r}_k-\vec{r}_0) = \frac{1}{3w_n} \sum_{l=x,y} \sum_{j=x,y,z} \int_{\xi_j} \int_{|\vec{\rho}|<1} E_{\text{pupil},lj}(W, \vec{\rho}, d, z_{\text{stage}}) \exp\Big[-i\vec{k}(\vec{\rho}) \cdot (\vec{r}_k - \vec{r}_0)\Big] \, d^2\rho \, d^2\xi_k, \tag{2.10}$$

where ξ_k is the k^{th} pixel, $\vec{k}(\vec{\rho}) = (k_x(\vec{\rho}),\, k_y(\vec{\rho}),\, k_{z,\text{med}}(\vec{\rho}))$, and w_n is a normalization factor defined in [69].

Vectorial PSF models provide a physically accurate description of the image formation process. In contrast, Gaussian PSF models often appear sufficient when emitters rotate freely, since the resulting PSF represents an average over all dipole orientations [66]. However, when the dipole orientation is fixed or when optical aberrations are present, the PSF can become asymmetric and is no longer well described by a Gaussian approximation [65]. Such mismatches can introduce localization biases that may reach several tens of nanometers [66, 19].

Vectorial PSF models capture these effects more accurately, but their evaluation requires computing Fourier transforms and is therefore computationally more expensive. In this thesis, the vectorial PSF model is therefore only used in the image formation model to generate simulated data, as it allows the refractive indices of the optical media and the resulting spherical aberrations to be explicitly controlled. However, for localization algorithms, the computational cost of evaluating vectorial PSFs can be high, motivating the use of alternative PSF representations, such as spline models.

2.2.4 Cubic spline PSF Model

Cubic splines are piecewise polynomial functions defined over a set of connected intervals. Their function value and first and second derivatives are continuous at the knots, which are the points at which the polynomial segments connect. Spline functions have also been used to model the 3D PSF for fitting images of single molecules [4, 40]. Cubic spline (cspline) models provide a representation of experimentally measured PSF, enabling the modeling of complex PSF shapes and aberrations that may occur at different imaging depths using calibration data.

The 3D PSF can be approximated by a cubic spline defined per voxel (i, j, k) as follows:

$$f_{i,j,k}(x, y, z) = \sum_{m=0}^{3} \sum_{n=0}^{3} \sum_{o=0}^{3} A_{i,j,k,m,n,o} \left(\frac{x - t_i}{\Delta t}\right)^m \left(\frac{y - t_j}{\Delta t}\right)^n \left(\frac{z - u_k}{\Delta u}\right)^o, \tag{2.11}$$

$$t_i \le x \le t_{i+1}, t_j \le y \le t_{j+1}, u_k \le z \le u_{k+1}, \Delta t = t_{i+1} - t_i = t_{j+1} - t_j, \Delta u = u_{k+1} - u_k$$

where Δt denotes the voxel size in the lateral (x, y) directions, and Δu denotes the voxel size along the axial (z) direction. The coordinates t_i, t_j, and u_k define the lower bounds of the voxel (i, j, k) along the x, y and z directions, respectively.

To uniquely specify the cspline $f_{i,j,k}(x, y, z)$, there are $(k+1)^d$ cspline coefficients $A_{i,j,k,m,n,o}$ required in each voxel, where k denotes the polynomial order and d the dimensionality of the spline. For a cubic spline $(k = 3)$ in three dimensions $(d = 3)$, this corresponds to $(k + 1)^d = 4^3 = 64$ coefficients.

2.3 Maximum Likelihood Estimation fitting

We have defined an image formation model that predicts the expected photon counts μ_k, and introduced various approaches to modeling the PSF. To estimate the parameters of a single-molecule emitter, the image formation model must be fitted to the observed image using an estimation algorithm.

An estimator is a function that maps the observed data to estimates of the underlying parameters. Consider the observed data vector $\vec{c} = (c_1, \ldots, c_N)$, where each c_k is a realization of an independent photon count in pixel k. Assuming photon detection follows a Poisson process, each pixel value is modeled as an independent random variable with mean $\mu_k(\vec{\theta})$. To simplify notation, we will continue writing $\mu_k(\vec{\theta})$ as μ_k. The likelihood of observing the data $\vec{c}$ given the parameters $\vec{\theta}$ is therefore:

$$L(\vec{\theta} \mid \vec{c}) = \prod_{k=1}^{N} p(c_k; \vec{\theta}) = \prod_{k=1}^{N} \frac{\mu_k^{c_k} e^{-\mu_k}}{c_k!}. \tag{2.12}$$

where $p(c_k; \vec{\theta})$ is the probability mass function of observing the measurement of $\vec{c}$ given the underlying parameters $\vec{\theta}$. The maximum likelihood estimation method determines the parameter values $\vec{\theta}$, such that the likelihood function is maximized. The parameter values $\hat{\theta}$ that maximize the likelihood define the maximum likelihood estimator:

$$\hat{\theta} = \arg\max_{\vec{\theta}} L\left(\vec{\theta} \mid \vec{c}\right) \tag{2.13}$$

In practice, it is advantageous to maximize the logarithm of the likelihood. Since the logarithm converts products into sums while preserving the location of the maximum, and the logarithm function is monotonous, maximizing the likelihood function over $\vec{\theta}$ leads to the same argument as maximizing $\log L(\vec{\theta} \mid \vec{c})$ over $\vec{\theta}$:

$$\begin{aligned}\log L(\vec{\theta} \mid \vec{c}) &= \log \prod_k \frac{\mu_k^{c_k} e^{-\mu_k}}{c_k!} \\ &= \sum_k \log \frac{\mu_k^{c_k} e^{-\mu_k}}{c_k!} \\ &= \sum_k c_k - \mu_k + c_k \log \frac{\mu_k}{c_k}.\end{aligned} \tag{2.14}$$

Stirling's approximation is sometimes used to simplify expressions: $\log c_k! \approx c_k \log c_k - c_k$ [64]. However, it is not required for the optimization since $\log(c_k!)$ does not depend on $\vec{\theta}$. Additionally, since many numerical optimizers in statistical packages usually work by minimization rather than maximization. Therefore, this expression is often rewritten as the negative log-likelihood is minimized instead:

$$-\log L\left(\vec{\theta} \mid \vec{c}\right) = \sum_k \mu_k - c_k - c_k \log \frac{\mu_k}{c_k} \tag{2.15}$$

It is also commonplace to express the objective function in terms of the log-likelihood ratio with respect to the maximum possible likelihood $L(\vec{c}|\vec{c})$. Taking the negative logarithm converts the product in the likelihood function into a sum, and multiplying by a factor of two yields the objective function:

$$\chi^2_{\mathrm{MLE}} = -2\log\frac{L\left(\vec{\theta}\,|\,\vec{c}\right)}{L\left(\vec{c}\,|\,\vec{c}\right)} = 2\left[\sum_k\left(\mu_k - c_k\right) - \sum_{k,c_k>0} c_k \log\left(\mu_k/c_k\right)\right], \tag{2.16}$$

where the factor of two is introduced by convention to obtain an objective function that is comparable to the classical χ^2 statistic used in least-squares fitting [37]. Minimizing the χ^2_{MLE} yields the maximum likelihood estimate of the underlying parameters $\vec{\theta}$ for the noisy observed image.

2.4 Modified Levenberg-Marquardt optimization algorithm

The objective function in Equation 2.16 was introduced in the context of MLE fitting for Poisson distributions by Laurence et al. [37]. To minimize this function, they employed a modified version of the LM algorithm introduced by Nishimura and Tamura [49].

As the LM algorithm can be viewed as a modification of the Newton–Raphson method, we first briefly recall this method. Newton–Raphson is an iterative multidimensional root-finding method that relies on the gradient and the Hessian of the objective function [52]. In this method, the search direction $\Delta\theta_i$ for the i^{th} free parameter θ_i is given by

$$\Delta\theta_i = H_{ij}^{-1} J_j, \tag{2.17}$$

where H_{ij} and J_j are the Hessian and Jacobian of the objective function, respectively. However, the inversion of H_{ij} may become numerically ill-conditioned if the Hessian matrix is almost singular. Instead, the Levenberg-Marquardt method solves this problem by using an additional dimensionless damping factor λ [52]. In the Levenberg-Marquardt algorithm, the search direction $\Delta\theta_i$ is now given by

$$\Delta\theta_i = \left(H_{ij} + \lambda I\right)^{-1} J_j. \tag{2.18}$$

If $\|H_{ij}\| >> \lambda$, the Levenberg-Marquardt algorithm starts to resemble the Newton-Raphson method, and if the norm is much smaller than λ, the algorithm takes a step in the steepest descent direction instead [70].

The resulting optimization problem is generally non-convex, and therefore, the algorithm does not always converge to the global minimum. There are multiple local minima that the localization scheme can converge to, and as such, the initial conditions θ_0 are important. This is especially the case under conditions of model mismatch. In this thesis, a multi-start optimization strategy was implemented.

In the Levenberg-Marquardt algorithm modified to find the maximum likelihood estimate for a Poisson process, the Hessian matrix is used without the second partial derivatives term, as computation of the second derivatives is often quite difficult and can be destabilizing when the model fits badly or is contaminated by outlier points [40, 52]:

$$H_{ij} = \sum_k \frac{\partial\mu_k}{\partial\theta_i}\frac{\partial\mu_k}{\partial\theta_j}\frac{c_k}{\mu_k^2}. \tag{2.19}$$

The Jacobian J_j is defined as

$$J_j = \sum_k \frac{\partial\mu_k}{\partial\theta_j}\frac{\left(c_k - \mu_k\right)}{\mu_k} \tag{2.20}$$

The parameter update $\Delta\vec{\theta}$ is obtained by solving the linear system defined by Equations (2.19) and (2.20). The parameters are then iteratively updated according to

$$\vec{\theta}_{i+1} = \vec{\theta}_i + \Delta\vec{\theta}_i. \tag{2.21}$$

In this thesis, the damping factor λ was set to 10, and the algorithm was run for 30 iterations.

2.5 Evaluating localization performance

Once the single-molecule parameters have been estimated using the localization algorithm described in the previous sections, it is important to quantify the quality of these estimates. In single-molecule localization microscopy, the performance of a localization algorithm is typically evaluated in terms of localization bias and localization precision.

2.5.1 Localization bias and precision

The localization bias of the estimated x position θ_x can be written as:

$$\mathrm{Bias}(\theta_x) = \langle\theta_x\rangle - x, \tag{2.22}$$

where the angled brackets $\langle\cdot\rangle$ denote statistical averaging, and x is the ground truth x-position of the single molecule, only known in simulation.

The localization precision is determined by the variance of the estimated position and can be calculated as follows:

$$\mathrm{Var}(\theta_x) = \langle(\theta_x - \langle\theta_x\rangle)^2\rangle. \tag{2.23}$$

In practice, the localization precision is often reported as the standard deviation of the estimated position rather than the variance, since it has the same physical units as the position estimate.

2.5.2 Cramér-Rao lower bound

For an unbiased estimator and assuming the likelihood of the data satisfies the regularity conditions, the Cramèr-Rao lower bound (CRLB) provides a theoretical lower bound on the variance of the estimator [71, 35].

The CRLB is strongly related to a mathematical measure called the Fisher information, which quantifies the sensitivity of an observable quantity (our noisy SMLM image) to changes in its underlying parameters (e.g., emitter position). The CRLB provides a lower bound on the variance for a scalar parameter θ and is computed as follows [35]

$$\mathrm{CRLB}(\theta) = \frac{1}{-\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\ln p(\vec{c};\theta)\right]}, \tag{2.24}$$

where $\ln p(\vec{c};\theta)$ denotes the log-probability density function of observing the data $\vec{c}$, given the parameter θ. If $\hat{\theta}$ is an unbiased estimate of θ, the lower bound of the variance cannot exceed the CRLB

$$\mathrm{var}(\hat{\theta}) \geq \mathrm{CRLB}(\theta). \tag{2.25}$$

In the case a vector of parameters $\vec{\theta} \in \mathbb{R}^p$ is estimated, the lower bound on the variance of each parameter θ_i is given by the corresponding diagonal element of the inverse Fischer information matrix (FIM) $I(\vec{\theta})$

$$\mathrm{Var}(\hat{\theta}_i) \geq [I^{-1}(\vec{\theta})]_{ii}, \tag{2.26}$$

where the FIM is a $p \times p$ matrix defined as

$$\left[I(\vec{\theta})\right]_{ij} = \mathbb{E}\left[\frac{\partial^2 \ln p(\vec{c}; \vec{\theta})}{\partial\theta_i \partial\theta_j}\right], \tag{2.27}$$

for $i, j = 1, 2, \ldots, p$. In the scalar case (2.24), $p = 1$ and therefore $I(\theta) = \mathbb{E}\left[\frac{\partial^2 \ln p(\vec{c};\theta)}{\partial\theta^2}\right]$ is a 1×1 FIM making it a scalar.

Combining the log of the expression for the likelihood in (2.12) and the expression for the FIM in (2.27), we obtain the following expression for the FIM applied to SMLM:

$$\left[I(\vec{\theta})\right]_{ij} = \mathbb{E}\left[\sum_k (c_k - \mu_k)^2 \frac{1}{\mu_k^2} \frac{\partial\mu_k}{\partial\theta_i} \frac{\partial\mu_k}{\partial\theta_j}\right]. \tag{2.28}$$

Recall that in (2.1) we defined the image formation model assuming a Poisson process of for the arrival of photons at pixels, where μ_k is the expected photon count at the k^{th} pixel and c_k is the observed photon count. For a Poisson distribution, the variance is equal to the mean, such that $\mathbb{E}\left[\sum_k (c_k - \mu_k)^2\right]$ simply becomes μ_k. Substituting this into (2.28) leads to the following expression for the FIM:

$$\left[I(\vec{\theta})\right]_{ij} = \sum_k \frac{1}{\mu_k} \frac{\partial\mu_k}{\partial\theta_i} \frac{\partial\mu_k}{\partial\theta_j}. \tag{2.29}$$

Given an image formation model, signal-level, background, and noise models, the CRLB provides the theoretical lower bound on the achievable localization precision. In practice, several works have shown that well-designed localization algorithms based on the MLE approach this bound under ideal conditions [2, 47, 64]. However, this optimal precision can only be achieved when the assumed image formation model accurately describes the measured data. Model mismatches, resulting from depth-dependent aberrations caused by refractive-index mismatches, may therefore lead to localization precision that deviates from the CRLB [62] as the regularity conditions imposed on the likelihood between the model and data are violated.

2.5.3 χ^2-test

The χ^2-test can be used to assess the goodness-of-fit between the observed experimental data and the image predicted by the fitted model parameters [61]. The χ^2 statistic is defined as:

$$\chi^2 = \sum_{k=1}^{K} \frac{(c_k - \mu_k(\vec{\theta}))^2}{\mu_k(\vec{\theta})}, \tag{2.30}$$

where c_k is the observed photon count in pixel k, and $\mu_k(\vec{\theta})$ is the expected photon count predicted by the image formation model evaluated at the estimated parameters $\hat{\theta}$. Here, K denotes the total number of pixels in the ROI.

Spline interpolation preliminaries

This chapter introduces the spline-interpolation principles required for this thesis. In Chapter 2, cubic spline models were introduced as a possible representation of the Point Spread Function (PSF) in localization algorithms [4, 40]. In this chapter, rather than focusing on splines as a means to represent the PSF, the focus shifts to spline interpolation itself, which forms the mathematical basis of the proposed depth-dependent method.

The idea of using splines to create smooth curves dates back centuries to the Middle Ages, when draftsmen used long, flexible strips of wood to draw smooth lines when designing ship hulls and bows [74]. By fixing the strip of wood at several points, a smooth curve passing through those points could be obtained. In modern applications, spline functions follow a similar principle, but instead of wooden strips, they are now constructed mathematically using piecewise polynomial functions. Today, splines are widely applied in fields such as automotive design, medical imaging, system identification, and computer-aided geometric design [33, 53].

Mathematically, a spline is defined as a piecewise polynomial function that maintains a degree of smoothness at the boundaries between adjacent polynomial segments (knots). Many spline functions are constructed by combining simple polynomial curves into a single smooth function. A well-known example is the Hermite spline, which defines the curve shape through endpoint positions and tangents. Before introducing such spline constructions in detail, the polynomial notation used in this thesis will be established, which forms the mathematical foundation for spline interpolation and is adopted from [60].

3.1 Polynomial representation

Piecewise polynomials are a fundamental tool for representing curves in computer graphics and geometric modeling. The general form of an n^{th}-degree polynomial is:

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n, \tag{3.1}$$

where t denotes the curve parameter, a_i are the coefficients of the polynomial, and $n \in \mathbb{N}$ denotes the degree of the polynomial, i.e., the highest power of t for which the corresponding coefficient $a_n \neq 0$. Equation 3.1 can also be written in the canonical form of a polynomial by using summation notation [60]:

$$f(t) = \sum_{i=0}^{n} a_i t^i. \tag{3.2}$$

A polynomial can therefore be seen as being either zero or a finite sum of non-zero terms, where each term consists of a coefficient a_i multiplied by a power of the parameter t. The coefficients determine the shape of the polynomial's curve.

In many applications, particularly in curve modeling and interpolation, cubic polynomials ($n = 3$) are commonly used because they provide sufficient flexibility to represent smooth curves while remaining computationally efficient. These cubic polynomials form the building blocks of many spline constructions, including Hermite splines introduced in the next section [20].

3.2 Cubic Hermite splines

A cubic Hermite spline is constructed by joining multiple cubic Hermite segments defined on consecutive intervals $[t_i, t_{i+1}]$. Each segment interpolates two points with predefined tangents at those points.

Consider a spline segment $f_i(x)$ defined on the interval $[t_i, t_{i+1}]$. The segment interpolates between p_i at t_i and p_{i+1} at t_{i+1} with prescribed tangents p'_i and p'_{i+1}, respectively. These interpolation conditions and a smooth transition between adjacent segments are imposed by the following constraints, written as

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} p_i \\ p'_i \\ p_{i+1} \\ p'_{i+1} \end{bmatrix}. \tag{3.3}$$

Hermite segments are typically defined on the unit interval $t \in [0, 1]$. Therefore, $f_i(0)$ and $f_i(1)$ correspond to the function values at the start and end of each segment, respectively. For a cubic Hermite spline, each segment is represented by a cubic polynomial:

$$f_i(r) = a_i r^3 + b_i r^2 + c_i r + d_i, \tag{3.4}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, and $t_i \leq x < t_{i+1}$. Substituting these expressions into the constraints yields the following system:

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} p_i \\ p'_i \\ p_{i+1} \\ p'_{i+1} \end{bmatrix} = \begin{bmatrix} d_i \\ c_i \\ a_i + b_i + c_i + d_i \\ 3a_i + 2b_i + c_i \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix}. \tag{3.5}$$

Solving for the polynomial coefficients yields:

$$\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \mathbf{C}^{-1} \begin{bmatrix} p_i \\ p'_i \\ p_{i+1} \\ p'_{i+1} \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i \\ p'_i \\ p_{i+1} \\ p'_{i+1} \end{bmatrix}. \tag{3.6}$$

Substituting these coefficients back into Equation 3.4 gives the matrix form of a cubic Hermite spline segment:

$$f_i(r) = \begin{bmatrix} r^3 & r^2 & r & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_i \\ p'_i \\ p_{i+1} \\ p'_{i+1} \end{bmatrix}, \tag{3.7}$$

with $r = (x - t_i)/(t_{i+1} - t_i)$ and $t_i \leq x < t_{i+1}$. It can be seen from the interpolation conditions in Equation 3.3, that adjacent spline segments share both the function value and the first derivative at their junctions. Therefore, cubic Hermite splines are C^1-continuous.

3.3 Catmull–Rom splines

A Catmull–Rom spline can be viewed as a special case of a cubic Hermite spline. Similar to a Hermite spline, a 1D Catmull–Rom spline $f_i(x)$ is defined for a segment $[t_i, t_{i+1}]$ by interpolation between p_i at t_i and p_{i+1} at t_{i+1}, with prescribed tangents p'_i and p'_{i+1} at those points. Unlike a Hermite spline, where the tangents are arbitrarily chosen values, a Catmull–Rom defines the tangents using points p_{i-1} and p_{i+2} consecutive to points p_i and p_{i+1} respectively. The tangent at p_i is determined by the vector $p_{i+1} - p_{i-1}$, and the tangent at p_{i+1} from $p_{i+2} - p_i$.

To ensure a smooth transition between segments f_i and f_{i+1}, the following constraints are imposed:

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} p_i \\ \frac{p_{i+1} - p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2} - p_i}{2} \end{bmatrix}, \tag{3.8}$$

for each segment i. $f_i(0)$ and $f_i(1)$ correspond to the function value at the start and end of each segment. For a cubic Catmull–Rom spline, each segment is represented by a cubic polynomial:

$$f_i(r) = a_i r^3 + b_i r^2 + c_i r + d_i, \tag{3.9}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, and $t_i \leq x < t_{i+1}$. Substituting this polynomial expression into the constraints from Equation 3.8 yields the following fully determined system of equations to obtain 4 coefficients:

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} d_i \\ c_i \\ a_i + b_i + c_i + d_i \\ 3a_i + 2b_i + c_i \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \begin{bmatrix} p_i \\ \frac{p_{i+1} - p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2} - p_i}{2} \end{bmatrix}. \tag{3.10}$$

To obtain the coefficients a_i, b_i, c_i and d_i, we solve the system using matrix inversion:

$$\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \mathbf{C}^{-1} \begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -0.5 & 0 & 0.5 \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{3.11}$$

where we expressed the tangents through the quadruplet of the points $p_{i-1}, p_i, p_{i+1}, p_{i+2}$ as follows:

$$\begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -0.5 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}. \tag{3.12}$$

Finally, substituting these coefficients back into Equation 3.9 gives the matrix form of a cubic Catmull–Rom spline segment:

$$f_i(r) = \begin{bmatrix} r^3 & r^2 & r & 1 \end{bmatrix} \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{3.13}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, with $t_i \leq x < t_{i+1}$. The spline segment defined on this interval interpolates the points p_i and p_{i+1}, such that

$$f(t_i) = p_i, \quad f(t_{i+1}) = p_{i+1}.$$

An example of a 1D cubic Catmull–Rom spline segment is shown in Figure 3.1. In contrast to Hermite splines, Catmull–Rom splines do not require the tangents at the endpoints to be predefined. Instead, the tangents are estimated directly from neighboring data points.

In this thesis, PSF calibration stacks are available only at discrete imaging depths. The derivatives of the PSF with respect to neighboring depth are not known beforehand. Catmull–Rom splines therefore provide an appropriate interpolation scheme, as they generate a smooth depth-dependent PSF model using only the neighboring calibration stacks. In practice, however, the available calibration depths are not guaranteed to be uniformly spaced. The standard Catmull–Rom formulation introduced in this chapter therefore serves as the base for the proposed interpolation method, but will be extended to handle non-uniform parameter spacing. The non-uniform Catmull–Rom formulation used in the proposed depth-dependent calibration method is presented in Chapter 4, and its derivation can be found in the appendix of that chapter.

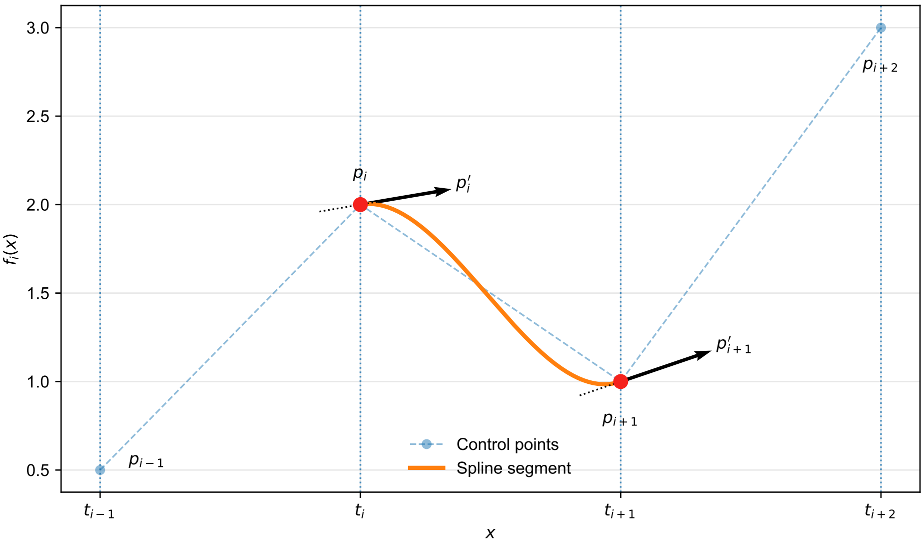


Figure 3.1: Example of a 1D Catmull–Rom spline segment defined by the four points $p_{i-1}, p_i, p_{i+1}, p_{i+2}$ with values 0.5, 2, 1, and 3, respectively. The corresponding parameter values are uniformly spaced at $t_{i-1} = 0.0$, $t_i = 1.0$, $t_{i+1} = 2.0$, and $t_{i+2} = 3.0$. The spline segment interpolates the points p_i and p_{i+1}, and the tangent vectors p'_i and p'_{i+1} are determined according to the Catmull–Rom spline formulation.

Manuscript

This chapter presents the main results of the thesis written as a paper manuscript. The manuscript is written using the Springer Nature LaTeX authoring template and includes the methods, results, supplementary notes, and figures.

Depth-Dependent Point Spread Function Calibration via Spline Interpolation in 3D Single-Molecule Localization Microscopy

Xin Yue Zhang[1*], Serafim Korovin[1*] and Carlas Smith[1*]

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

Abstract

In three-dimensional single-molecule localization microscopy (SMLM), emitter positions are estimated by fitting a model of the microscope's point spread function (PSF) to measured data. In practice, PSF models are typically calibrated using bead data acquired near the coverslip, and are assumed to remain valid representations at larger imaging depths. However, refractive index mismatch between the immersion medium, coverslip, and sample introduces depth-dependent spherical aberrations, causing the PSF shape to vary with imaging depth. As a result, a PSF model calibrated at the coverslip leads to degraded lateral localization precision and substantial axial bias when applied several micrometers deep into the sample. In this work, we introduce a depth-dependent PSF calibration approach that interpolates between calibration datasets acquired at multiple imaging depths. Calibration stacks are reconstrcuted at arbitrary depths using Catmull–Rom spline interpolation and used to calibrate cspline models for localization. Simulations show that a conventional coverslip-calibrated model results in mean absolute axial biases exceeding **294** nm at an imaging depth of **5** $\boldsymbol{\mu}$m. In contrast, the proposed approach reduces the axial bias up to **99%**, consistently achieving axial bias below **5** nm. In addition, the lateral localization precision improves by **62%** and **61%** in $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. Validation on experimentally acquired bead data demonstrates an axial bias reduction of **80%** compared to coverslip calibration. These results show that interpolation of calibration data across depth compensates for depth-dependent PSF mismatch, enabling accurate and precise 3D localization over extended imaging depths without requiring additional optical hardware.

1 Introduction

Single-molecule localization microscopy (SMLM) is a super-resolution imaging method that enables visualization of biological structures with nanometer-scale precision beyond the diffraction limit of conventional light microscopy [1–3]. In 3D SMLM, the axial position of a single molecule is typically inferred from the shape of the point spread function (PSF), using PSF engineering approaches such as astigmatic [4], double-helix [5], or tetrapod [6] PSFs.

To estimate single-molecule parameters such as its position, maximum likelihood estimation (MLE) [7] is commonly employed. Under appropriate conditions, the variance of the MLE approaches the theoretical lower bound given by the Cramér–Rao lower bound (CRLB), provided an accurate PSF model is used. PSF models are typically obtained by calibrating an analytical PSF model to a stack of experimental bead images acquired by scanning a fluorescent bead immobilized near the coverslip along the axial direction. The assumption is that the PSF measured near the coverslip resembles that obtained deeper in the sample.

However, biological imaging is typically conducted in aqueous mounting media with a refractive index of approximately 1.33, whereas the glass coverslip on which samples are mounted, and the oil in which the objective lens is immersed, have refractive indices of approximately 1.51 and 1.52, respectively. This refractive index mismatch causes light to propagate at different speeds across interfaces, leading to distorted wavefronts that can be characterized by depth-dependent spherical aberrations [8, 9].

These aberrations reduce peak intensity as emitted photons spread over a larger area, effectively elongating the 3D PSF. As a result, the axial cross-section of the PSF profile becomes increasingly asymmetric with respect to the focal position at depths of only a few micrometers. Additionally, refractive index mismatch introduces a focal shift, in which the physical distance between an emitter in the sample and the coverslip (the actual focus position) does not match the displacement required by the objective to bring the emitter into focus (the nominal focus position). Applying a PSF model calibrated near the coverslip to such aberrated data, therefore, results in reduced localization precision and increased bias in the estimated single-molecule position.

Several approaches have been proposed to mitigate the effects of depth-dependent aberrations or to accurately model the PSF. These include active aberration correction using adaptive optics (AO) [10, 11], and physically derived [12] or phase-retrieved PSF models [13–15]. While these methods can accurately model or compensate for aberrations, they often require specialized optical hardware, dedicated expertise for implementing the complex AO, or calibration of numerous optical system parameters, such as objective lens properties, refractive indices, aberration functions, and emission wavelengths. These requirements can limit their widespread adoption.

An alternative approach was proposed by Li *et al.* [16], who used a cubic spline (cspline) approach to create a look-up-table (LUT) to correct bias in estimated axial positions by interpolating between correction values across depth. Although this approach improves axial bias without hardware modifications, it does not correct the underlying mismatch between the PSF model and the observed PSF shape, and therefore does not recover localization precision.

In this work, we introduce a depth-dependent PSF calibration approach that interpolates between calibration datasets acquired at multiple imaging depths. Using Catmull–Rom spline interpolation, calibration stacks are reconstructed at arbitrary depths, and used to calibrate cspline PSF models throughout the imaging volume. Additionally, we compare this approach to calibration at the nearest available depth, providing insight into the effect of using interpolated versus locally calibrated PSF models.

Using both simulated and experimental data, we show that coverslip-calibrated PSF models lead to degraded lateral localization precision and biased axial localization at increasing imaging depths. In contrast, the proposed depth-dependent calibration approach provides an improved PSF representation across depth, reducing localization bias without degrading precision, while remaining practical to implement without additional optical hardware.

2 Methods

2.1 Image formation model

To simulate and estimate single-molecule images, we adopt a standard image formation model describing the expected photon counts detected on the camera sensor. The expected number of photons μ_k detected at the k^{th} camera pixel from a single molecule is given by [12]:

$$\mu_k = \theta_I \, H(\vec{r}_k - \vec{r}_0) + \theta_b, \tag{1}$$

where θ_I denotes the total photon count of the emitter, $H(\vec{r}_k - \vec{r}_0)$ represents the point spread function (PSF) of the imaging system, θ_b denotes the background photon count per pixel, $\vec{r}_0 = (\theta_x, \theta_y, \theta_z)$ are the emitter coordinates, and $\vec{r}_k$ denotes the spatial coordinate of the center of pixel k. The following sections describe the PSF model used to simulate the data and the algorithm used to estimate the parameter vector $\vec{\theta}$ by fitting the image formation model to observed images.

2.2 Vectorial point spread function simulation

To simulate realistic single-molecule images that include astigmatism from 3D SMLM with PSF engineering and depth-dependent spherical aberrations arising from refractive-index mismatch, the vectorial PSF model is employed. This model describes the PSF of a freely rotating dipole, consistent with previous studies [12, 17–20].

The vectorial PSF is obtained by propagating the dipole emission through the optical system and computing the resulting electric field distribution in the image plane. The vectorial PSF model accounts for the dipole orientation of a single emitter, the effects of high NA, media interfaces, polarization, and aberrations [17]. The emitted electric field E_{dipole} propagates from the dipole located in the mounting medium through the coverslip and immersion medium with refractive indices n_{med}, n_{cov}, and n_{imm}, respectively. The electric field E_{dipole} is collected by the objective and transferred to the electric field in the pupil plane E_{pupil}. The electric field component $l = x, y$ in the image plane is proportional to the emission dipole component $j = x, y, z$, and can be expressed as the Fourier transform of the electric field components in the pupil plane [12, 20]:

$$E_{\mathrm{pupil},lj}(W, \vec{\rho}, d, z_{\mathrm{stage}}) = A(\vec{\rho}) q_{lj}(\vec{\rho}) \exp\left(iW(\vec{\rho}) + i\left[dk_{z,0}(\vec{\rho}) - z_{\mathrm{stage}} k_{z,2}(\vec{\rho})\right]\right), \quad (2)$$

where $\vec{\rho} = (\rho_x, \rho_y)$ are the normalized pupil coordinates, $A(\vec{\rho})$ is the amplitude[21, 22], including the well-known aplanatic correction factor given by [12]:

$$A(\rho) = \sqrt{\frac{n_{\mathrm{imm}} \cos(\theta_{\mathrm{imm}})}{n_{\mathrm{med}} \cos(\theta_{\mathrm{med}})}} \frac{1}{\sqrt{n_{\mathrm{med}} \cos(\theta_{\mathrm{med}})}} = \frac{\left(n_{\mathrm{imm}}^2 - \mathrm{NA}^2 \rho^2\right)^{1/4}}{\left(n_{\mathrm{med}}^2 - \mathrm{NA}^2 \rho^2\right)^{1/2}}, \quad (3)$$

where NA is the numerical aperture of the objective lens. $q_{lj}(\vec{\rho})$ are polarization vector components defined elsewhere [23], $W(\vec{\rho})$ is the aberration function expressed as a sum of Zernike polynomials, d is the imaging depth, defined as the distance between the coverslip and the focal plane, z_{stage} is the position of the objective stage relative to the focal plane. The wave vector $k(\vec{\rho})$ is a function of the normalized pupil coordinates by:

$$\vec{k}(\vec{\rho}) = (k_x(\vec{\rho}),\, k_y(\vec{\rho}),\, k_{z,i}(\vec{\rho})) = \frac{2\pi}{\lambda} \left(\mathrm{NA}\rho_x,\, \mathrm{NA}\rho_y,\, \sqrt{n_i^2 - \mathrm{NA}^2 \|\vec{\rho}_2^2\|}\right), \quad (4)$$

where n_i and λ are the refractive indices for the mounting medium and the immersion medium, and the wavelength of emitted light. Lastly, to calculate the resulting incoherent PSF from emission of a freely-rotating dipole, the six Fourier transforms of the electric field components in the pupil plane are quadratically summed:

$$\begin{aligned} H(\vec{r}_k - \vec{r}_0) = \frac{1}{3w_n} \sum_{l=x,y} \sum_{j=x,y,z} \Bigg| \int_{\xi_j} \int_{|\vec{\rho}|<1} & E_{\mathrm{pupil},lj}(W, \vec{\rho}, d, z_{\mathrm{stage}}) \\ & \times \exp\left[-i\vec{k}(\vec{\rho}) \cdot (\vec{r}_k - \vec{r}_0)\right] \, d^2\rho \, d^2\xi_k \Bigg|^2 . \end{aligned} \quad (5)$$

where ξ_k is the k^{th} pixel, $\vec{k}(\vec{\rho}) = (k_x(\vec{\rho}),\, k_y(\vec{\rho}),\, k_{z,\mathrm{med}}(\vec{\rho}))$, and w_n is a normalization factor defined elsewhere [22]. The simulated vectorial point spread functions were computed using VectorialPSF by van Velde [24]. In this work, the vectorial PSF model is used to generate simulated calibration stacks that incorporate both astigmatism and depth-dependent spherical aberrations arising from refractive index mismatch. Compared to scalar PSF models, the vectorial PSF formulation is required to model high NA effects, polarization, and aberrations, for a realistic simulation of engineered PSFs in 3D SMLM.

2.3 Maximum likelihood estimation for SMLM

Maximum likelihood estimation (MLE) is employed to estimate the parameters of a single-molecule emitter from observed camera images. The vector containing the parameters to be estimated is

$$\vec{\theta} = \left(\theta_x \ \theta_y \ \theta_z \ \theta_I \ \theta_b\right)^T. \tag{6}$$

The observed photon counts recorded by state-of-the-art EMCCD and sCMOS cameras can be accurately modeled as realizations of a Poisson process with a mean equal to the expected photon count given in Equation 1 [7]. Under this model, maximizing the likelihood corresponds to finding the parameter set $\vec{\theta}$ for which the expected image μ_k best explains the observed photon counts.

The log-likelihood of observing an image $\vec{c} = (c_1, \ldots, c_K)$ under the Poisson noise model is given by:

$$\log \ell(\vec{\theta} \,|\, \vec{c}\,) = \sum_{k=1}^{K} c_k - \mu_k + c_k \log \frac{\mu_k}{c_k}, \tag{7}$$

where K denotes the total amount of pixels in the ROI $\vec{c}$. The maximum likelihood estimate $\hat{\theta}$ is found by minimizing the negative log-likelihood of Equation 7, through a modified Levenberg-Marquardt method. In the Levenberg-Marquardt algorithm, the search direction is given by:

$$\Delta\theta_i = \left(H_{ij} + \lambda I\right)^{-1} J_j, \tag{8}$$

where H_{ij} and J_j are the Hessian and Jacobian matrix, respectively, λ is the damping factor, and I is a diagonal matrix containing the diagonal elements of the Hessian matrix [25]. In the modified Levenberg-Marquardt algorithm, the Hessian matrix is approximated and used without the second partial derivatives term [25, 26]:

$$H_{ij} = \sum_k \frac{\partial \mu_k}{\partial \theta_i} \frac{\partial \mu_k}{\partial \theta_j} \frac{c_k}{\mu_k^2}. \tag{9}$$

The Jacobian J_j is defined as:

$$J_j = \sum_k \frac{\partial \mu_k}{\partial \theta_j} \frac{(c_k - \mu_k)}{\mu_k}. \tag{10}$$

The parameter update $\Delta\vec{\theta}$ is obtained by solving the linear system defined by Equations 9 and 10. The parameters are then iteratively updated according to:

$$\vec{\theta}_{i+1} = \vec{\theta}_i + \Delta\vec{\theta}_i. \tag{11}$$

The optimization is repeated until convergence or until a maximum number of iterations is reached. In the following section, we describe how the point spread function H used in the image formation model is constructed using a cubic spline representation.

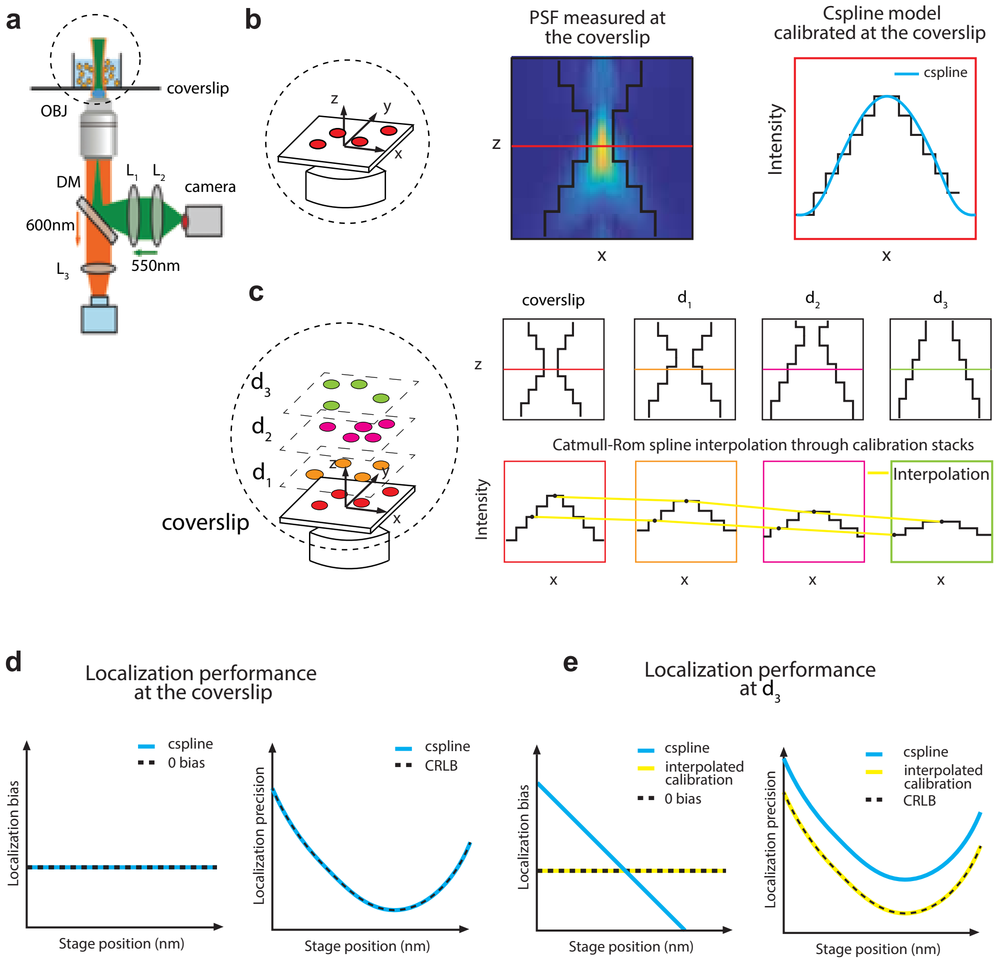


Fig. 1 Schematic of the proposed depth-dependent PSF calibration with interpolation. (a) Schematic of the optical setup. DM: dichroic mirror, OBJ: objective lens, L: lens. Drawing adapted from [27]. (b) Schematic overview of a conventional cspline calibration. A calibration stack is acquired at the coverslip and used to construct a cspline model. The PSF is illustrated as an intensity profile in x (colored), together with a discretized, pixel-sampled representation (black staircase) for visualization. (c) Schematic overview of the proposed depth-dependent calibration. PSF stacks are acquired at multiple depths d_i using beads embedded in agarose gel. The intensity profiles show that the PSF shape changes with imaging depth d. By interpolating between these PSF stacks using a Catmull–Rom spline (yellow curve), the interpolation can be evaluated at arbitrary depths. (d) Localization at the coverslip using a coverslip-calibrated cspline model yields no localization bias and precision approaching the Cramér-Rao lower bound (CRLB), the theoretical lower bound on achievable variance. (e) At greater depths, the coverslip-calibrated model (blue curves) yields biased estimates and compromised localization precision. The model obtained through the depth-dependent calibration approach (yellow curves) exhibits unbiased estimations with precision approaching the CRLB due to reduced model mismatch.

2.4 Cubic spline calibration at the coverslip

In this section, we describe the construction of a cubic spline (cspline) model for the 3D PSF, in which the PSF is represented as a piecewise third-order polynomial within each voxel (i, j, k). The PSF modeled as a cubic spline (cspline) can be written as [25, 28]:

$$\begin{aligned} f_{i,j,k}(x, y, z) &= \sum_{m=0}^{3}\sum_{n=0}^{3}\sum_{o=0}^{3} A_{i,j,k,m,n,o}\left(\frac{x-t_i}{\Delta t}\right)^m \left(\frac{y-t_j}{\Delta t}\right)^n \left(\frac{z-u_k}{\Delta u}\right)^o, \\ & t_i \leq x \leq t_{i+1}, t_j \leq y \leq t_{j+1}, u_k \leq z \leq u_{k+1}, \\ & \Delta t = t_{i+1} - t_i = t_{j+1} - t_j, \Delta u = u_{k+1} - u_k. \end{aligned} \tag{12}$$

Here, Δt and Δu denote the voxel size of the PSF along the x, y and z dimensions, respectively. The coordinates t_i, t_j, and u_k define the lower bounds of voxel (i, j, k). To uniquely specify the cspline $f_{i,j,k}(x, y, z)$ within each voxel, there are $(p+1)^d$ cspline coefficients $A_{i,j,k,m,n,o}$ needed, where p is the polynomial degree and d is the number of dimensions. For a cubic spline ($p = 3$) in three dimensions ($d = 3$), this results in $(p+1)^d = 4^3 = 64$ coefficients.

To calibrate a cspline model at the coverslip, simulated calibration data were first generated using the image formation model of Equation 1, with the vectorial PSF model described in Section 2.2. The spline coefficients were computed using the SMAP software package developed by Ries [29]. A stack of centered emitters was simulated over an axial range of ±800 nm around an imaging depth of $d = 0$ nm corresponding to the coverslip, with a z-spacing of 10 nm. Simulations used an emission wavelength of $\lambda = 671$ nm, an emitter photon count of 20000 photons, an astigmatic Zernike coefficient $Z_2^2 = 80$ mλ (rms), and a uniform background of 5 photons per pixel. Cspline representations of the PSF are a data-driven approach that allows the model to capture complex PSF shapes without relying on an explicit analytical model that depends on precise knowledge of optical parameters. The simulation parameters were chosen to reproduce realistic imaging conditions in 3D SMLM. The refractive indices, NA, and emission wavelength were selected to align with the oil-immersion microscopy configuration reported in the literature [30], ensuring that refractive-index mismatch-induced aberrations are realistically modeled. The photon count and background level were chosen to provide a sufficiently high signal-to-noise ratio, allowing the impact of PSF model mismatch to be isolated from noise-induced localization errors.

2.5 Depth-dependent PSF calibration construction

In this section, we describe the proposed depth-dependent PSF calibration approach. Conventional cspline model calibration approaches assume that a PSF measured at a single depth remains valid throughout the imaging volume, which leads to model mismatch at larger depths. A depth-dependent PSF calibration is constructed by generating PSF stacks at a discrete set of imaging depths $\{d_i\}$ using the image formation model described in Section 2.1.

To obtain a calibration stack at an intermediate evaluation depth d_{eval}, we propose an interpolation scheme between four neighboring calibration stacks using a Catmull–Rom spline. Specifically, for an evaluation depth $d_{\text{eval}} \in [d_i, d_{i+1}]$, the four neighboring stacks at depths $d_{i-1}, d_i, d_{i+1}, d_{i+2}$ are used. The Catmull–Rom formulation enables local interpolation directly on the sampled PSF data, producing an interpolated 3D PSF stack at depth d_{eval}.

A key property of the proposed approach is that the Catmull–Rom formulation is adapted to handle non-uniformly spaced calibration depths (see Appendix A). This is essential in practice, as experimentally acquired calibration stacks are typically not uniformly distributed in the imaging volume. Compared to global interpolation approaches, this formulation is computationally efficient due to its local nature and direct operation on sampled data.

In contrast, a full four-dimensional cubic spline representation over (x, y, z, d) requires storing polynomial coefficients per voxel that grow exponentially with dimensionality, i.e., $\mathcal{O}(4^d N)$. In the proposed approach, interpolation is performed only along the depth dimension for each voxel, using a constant number of neighboring calibration stacks. As a result, the computational cost and memory requirements scale linearly with the number of calibration depths, i.e., $\mathcal{O}(N)$, as no global coefficient model over (x, y, z, d) is constructed.

The interpolated PSF stack is subsequently converted into a calibration dataset compatible with the SMAP framework, and cspline coefficients are computed in the same procedure as described in Section 2.4. This yields a cspline PSF model at the evaluation depth d_{eval}, enabling PSF fitting at arbitrary depths within the calibrated range.

To evaluate the proposed depth-dependent calibration approach, localization is performed on simulated and experimental images using MLE. The negative log-likelihood is minimized using the modified Levenberg–Marquardt algorithm described in Section 2.3, using a damping factor of 10 and a maximum of 30 iterations. The overall workflow of the proposed approach, together with its impact on localization performance, is illustrated schematically in Figure 1.

2.6 Experimental data processing

In this section, we describe how the experimental data are processed before fitting and localization. To assess our proposed method on experimental data, we use a dataset provided by the Ries Group at the European Molecular Biology Laboratory (EMBL). This dataset was also used by Li *et al.* [16]. In this section, we describe how the data is preprocessed to obtain bead stacks ready for interpolation at various imaging depths.

2.6.1 Data preprocessing

Unlike the simulated single-molecule images, experimental images are not reported in photon counts but in raw camera units, typically expressed as analog-to-digital units (ADUs). For this dataset, offset correction can be applied directly using the metadata value of 400 ADU. However, a full conversion to photons cannot be performed because the ADU-to-electron conversion factor is not provided in the metadata, and the EM gain was reported as 1. Therefore, all processing is performed on offset-corrected ADU values.

After correcting the ADU values for the offset, candidate bead locations must be detected. The dataset contained large-field-of-view (FOV) stacks with a size of 320×279 pixels. To detect well-isolated individual beads within these FOVs, we employed a peak detection algorithm by Huang *et al.* [31]. This peak detection is described in detail in Appendix G.

After peaks are detected, a square region-of-interest (ROI) of size 27×27 pixels is cropped around the peak. Next, we align these stacks such that they are centered with respect to the ROI. To align each bead stack laterally, the z-plane with the highest peak intensity was selected as the reference plane. A 2D elliptical Gaussian with a constant background was fitted to this reference plane using nonlinear least-squares optimization. The fitted center coordinates were then used to compute the subpixel shift required to bring the bead center to the geometric center of the ROI. This shift was applied uniformly to all z-planes using third-order spline interpolation.

2.6.2 Depth assignment

In experimental data, the imaging depth of each processed bead stack must be estimated. To assign an imaging depth to each processed bead stack, the nominal focal plane of each bead is determined using an approach based on that of Li *et al.* [25]. Here, a cspline model calibrated on the bead stack from beads immobilized at the coverslip is fitted to the processed bead stacks from the deep axial scan after preprocessing. For each bead, this fitting procedure yields an estimated axial z_{fit} as a function of the objective stage position z_{stage}. The nominal focal position of the bead is defined as the stage position at which the fitted coordinate returns $z_{\text{fit}} = 0$. This position is determined with sub stepsize accuracy by linear interpolation between the two neighboring frames around the sign following Li *et al.* [16]. This condition corresponds to the symmetry plane of the engineered astigmatic PSF and provides a repeated axial reference through depth even in the presence of depth-dependent spherical aberrations [16].

In practice, the processed bead stacks spanned a large axial range and were fitted with a model calibrated at the coverslip. As such, the fitted curve may contain discontinuities or multiple sign changes, as illustrated in Figure I14 (see Appendix I). This behavior arises from using a coverslip-calibrated model across a deep axial scan, which leads to invalid fits outside the local focal region of a bead, causing the fitter to converge to different local minima at different stage positions. To account for the multiple frames at which there is a zero crossing, the nominal focal position is not taken from the first sign change in the full fitted curve. Instead, the continuous segments within the fitted axial positions are identified. The continuity of a segment is defined by subsequent fits not exceeding a predefined threshold. Segments whose axial span is smaller than a predefined minimum range are discarded. Among these remaining candidates, the longest segment that contains a sign change is used for depth assignment. The stage position at which $z_{\text{fit}} = 0$ within this selected segment is then estimated by interpolation between the two fitted points surrounding the sign change.

Lastly, the imaging depth of each bead relative to the coverslip is defined as

$$d = z_0 - z_{\mathrm{cs}},$$

where z_0 denotes the nominal focal position of the bead obtained from the deep scan and z_{cs} is the coverslip reference position. Following the approach of Li *et al.* [16], the coverslip reference position z_{cs} is defined as the lowest nominal focal position among all processed bead stacks in the dataset. This choice provides a consistent experimental reference corresponding to the shallowest detected focal plane, which is assumed to be closest to the physical coverslip. Subtracting z_{cs} expresses all fitted bead positions relative to a common experimental coverslip reference.

Lastly, each stack is cropped to a local axial interval of $[-700, 700]$ nm around the assigned depth, yielding a local calibration stack. These depth-assigned calibration stacks form the control points for the interpolation procedure described in Section 2.5. The processed bead stacks are subsequently used to construct calibration models and to generate synthetic datasets for performance evaluation, as described in the following section.

2.7 Evaluation of localization performance using experimental calibration data

To evaluate localization performance on experimental data, we perform a hybrid experimental-simulation approach. While the cspline PSF models are calibrated from the processed experimental bead stacks, the localization performance cannot be assessed directly from the experimental data alone. Each bead stack represents only a single realization of the measured image, which is insufficient to estimate statistical measures such as localization bias and precision.

To account for this, for each depth-assigned bead stack, a cspline model is first calibrated as described in Section 2.5. This calibrated model is then used to fit the same stack using MLE as described in Section 2.3, yielding estimates of the photon count and background level as functions of axial position.

These fitted photon and background values are subsequently used to generate synthetic single-molecule image stacks using the image-formation model in Equation 1. For each axial position, 60 independent realizations are generated by drawing from a Poisson distribution with a mean equal to the output of the image formation model, thereby reproducing the stochastic nature of photon detection and shot noise.

This procedure enables the statistical evaluation of localization performance, specifically the computation of localization bias and precision, which require multiple independent realizations at each axial position. While such statistics could be obtained from repeated experimental measurements, such as capturing repeated frames at each objective stage position, this was not available in the provided dataset. Therefore, the proposed approach provides a practical alternative by combining experimentally derived PSF models with simulated photon statistics to enable multiple realizations for performance evaluation.

2.8 Evaluation metrics

Localization performance is evaluated by the localization bias, localization precision, the Cramér–Rao lower bound (CRLB), and the χ^2 statistic.

The localization bias of an estimated parameter is defined as the mean error between the estimated and ground-truth value. For example, for the lateral coordinate θ_x, the bias is given by:

$$\mathrm{Bias}(\theta_x) = \langle \hat{\theta}_x - \theta_x \rangle, \tag{13}$$

where $\langle \cdot \rangle$ denotes averaging. The same definition is used for the y and z-coordinates.

Localization precision is defined as the standard deviation of the localization error, which is the square root of the variance of the estimated parameters, calculated as follows:

$$\sigma(\theta_x) = \sqrt{\langle (\hat{\theta}_x - \langle \hat{\theta}_x \rangle)^2 \rangle}. \tag{14}$$

The CRLB provides a theoretical lower bound on the achievable variance of an unbiased estimator, assuming the likelihood of the data satisfies the regularity conditions [32]. Using the log-likelihood function from Equation 7, the CRLB, which provides a lower bound on the variance of each estimated emitter parameter, is given by the diagonal elements of the inverse of the Fisher information matrix (FIM). The FIM is calculated as follows:

$$\left[I(\vec{\theta}) \right]_{ij} = \sum_k \frac{1}{\mu_k} \frac{\partial \mu_k}{\partial \theta_i} \frac{\partial \mu_k}{\partial \theta_j}. \tag{15}$$

The derivation of this expression is given in Appendix B. To calculate the CRLB for a given calibrated cspline model, the analytical expressions for the partial derivative terms are given in Appendix C.

Model quality is evaluated by the χ^2-statistic. This statistic is a goodness-of-fit measure that quantifies the statistical difference between the expected photon count produced by the model evaluated at the fitted parameters, and the measured photon count in the observed image [12]. The χ^2-statistic is calculated as follows:

$$\chi^2 = \sum_{k=1}^{K} \frac{(c_k - \mu_k)^2}{\mu_k}, \tag{16}$$

where c_k is the measured photon count of the k^{th} pixel, and μ_k is the expected photon count of that pixel as predicted by the fitted image formation model. Here, K denotes the total number of pixel observations, i.e., the number of pixels in the fitted ROI multiplied by the number of focal slices. Assuming c_k follows a Poisson distribution with mean equal to μ_k, the mean and variance of the χ^2-statistic can be calculated as:

$$\mathbb{E}\left[\chi^2\right] = K, \tag{17}$$

and

$$\sigma^2\left[\chi^2\right] = 2K + \sum_{k=1}^{K} \frac{1}{\mu_k}, \tag{18}$$

respectively. Therefore, in Section 3, the mean and variance of the observed χ^2 values, calculated according to Equation 16 from the fitted results, will be compared to the theoretical mean and variance.

3 Simulation results

In this section, we evaluate the localization performance of single-molecule localization with a conventional cspline PSF model calibrated at the coverslip under refractive-index-mismatch conditions, and our proposed method in simulation.

To provide a baseline for comparison, we first evaluate localization performance in the absence of model mismatch. This baseline is obtained by calibrating a cspline model on bead data acquired at the coverslip and subsequently using this model to fit data simulated at the same imaging depth. Model quality is assessed using pixel-wise residuals and the χ^2 statistic, while localization performance is quantified by bias and precision. We then compare this baseline to the case in which the same coverslip-calibrated model is applied at larger imaging depths, where refractive index mismatch induces depth-dependent aberrations and model mismatch.

3.1 Localization with a cspline model calibrated at the coverslip

To evaluate the depth-dependent calibration approach, we first establish a baseline by assessing the localization performance of a cspline model calibrated and evaluated at the coverslip. The calibration procedure is described in Section 2.4. Figure 2a shows the resulting calibration and the PSFs generated by this model at different axial positions. We evaluate the agreement between the calibrated model and the ground truth PSF at the evaluation depth using a pixel-wise difference, similar to Babcock *et al.* [28]. Figure 2a also shows the corresponding residuals with respect to the ground truth PSFs at the evaluation depth (coverslip). The maximum pixel-wise difference is 3×10^{-3}, indicating close agreement between the calibrated model and the ground-truth PSFs at the evaluation depth. This is expected as the cspline model is calibrated on data obtained at the same depth as the evaluation depth, as shown in Figure 2b.

Next, localization performance was evaluated using simulated single-molecule images generated with the image formation model in Equation 1 and the vectorial PSF model described in Section 2.2. The evaluation depth was set to the coverslip ($d = 0$ nm), and emitter positions were sampled over an axial range of $z \in [-600, 600]$ nm with a step size of 10 nm. The evaluation range was reduced to $z \in [-600, 600]$ nm to avoid convergence issues near the boundaries of the cspline calibration range. For each axial position, 100 independent noise realizations were generated by drawing Poisson-distributed photon counts with a mean equal to the expected photon count per pixel. The emitter photon count was set to 5000 with a uniform background of 5 photons per pixel. This photon count was chosen such that the localization precision approaches

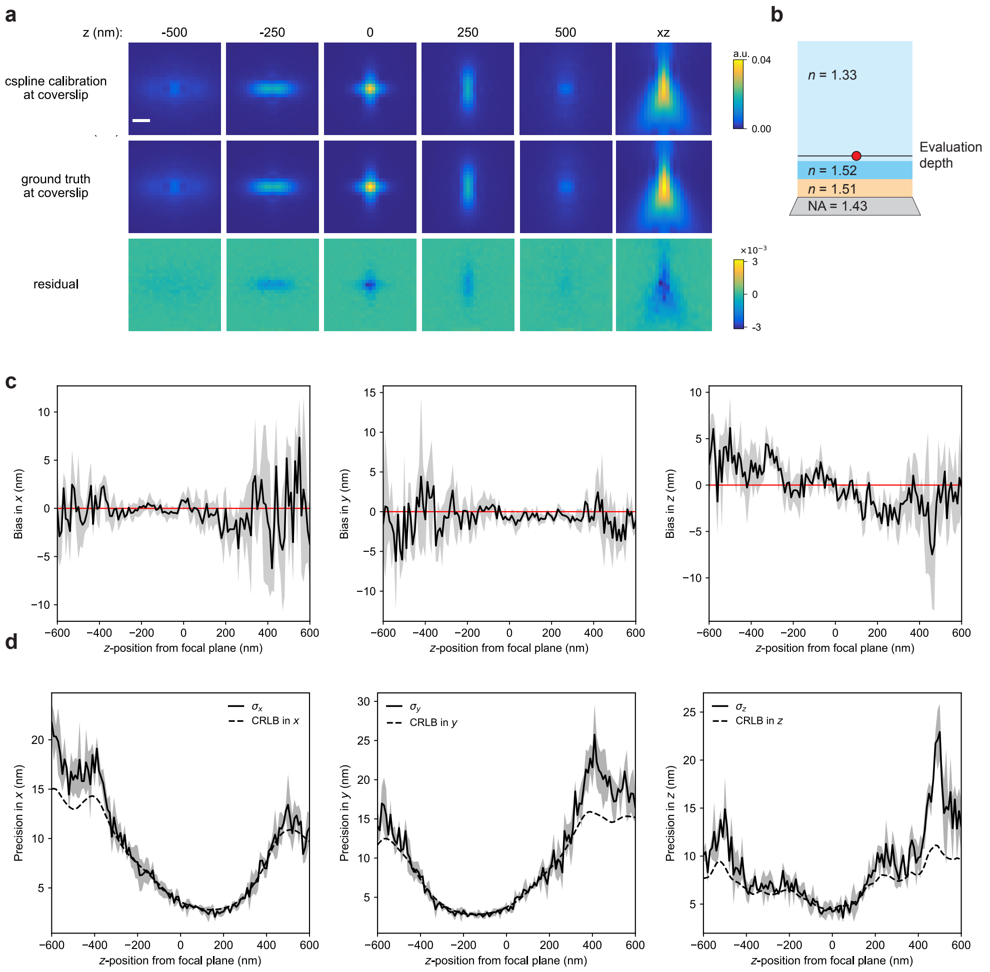


Fig. 2 Localization performance of a cspline model calibrated and evaluated at the coverslip. (a) PSFs of the cspline model calibrated at the coverslip at various axial positions (top). The ground truth PSFs at the evaluation depth $d = 0$ nm (middle). The residual between the calibrated model and data at the evaluation depth is computed as a pixel-wise difference (bottom). Scale bar: 500 nm; a.u. = arbitrary units. (b) The calibration is obtained from a bead (red circle) immobilized on the coverslip. The evaluation depth is at the same depth as the bead. (c) Localization bias in x, y, and z. Black lines and shaded areas in localization plots represent the mean and standard deviation of the biases. Red lines indicate zero bias. (d) Localization precision in x, y, and z alongside the Cramér-Rao lower bound (CRLB). The errors are unbiased, and precision reaches the CRLB for x, y, and z, respectively, indicating optimal performance in the absence of model mismatch.

the CRLB under the given imaging conditions (see Figure E2 in Appendix E). Each simulated stack consisted of 121 focal slices with an ROI size of 27 × 27 pixels. The optical parameters were the following: pixel size = 100 nm, NA = 1.43, wavelength of emitted light $\lambda = 671$ nm, astigmatic Zernike coefficient $Z_2^2 = 80$ mλ (rms), $n_{\mathrm{imm}} = 1.51$, $n_{\mathrm{cov}} = 1.52$, and $n_{\mathrm{med}} = 1.33$. These parameters were chosen to reflect typical experimental conditions in 3D SMLM using oil-immersion microscopy [30]. Emitter parameters were estimated by minimizing the negative log-likelihood using the modified Levenberg–Marquardt algorithm described in Section 2.3.

As shown in Figure 2c, the bias in the estimated x, y, and z positions remains centered around zero in the absence of model mismatch. The mean bias across the evaluated axial range was -0.675 nm, 0.307 nm, and -0.174 nm for x, y, and z, respectively. Figure 2d shows that the localization precision reaches the Cramér–Rao lower bound (CRLB) in all dimensions.

The χ^2 value is computed between the model evaluated at the fitted parameters $(\hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z, \hat{\theta}_N, \hat{\theta}_b)$ and the observed data across all pixels. The mean χ^2 was $\bar{\chi}^2 = 1.346 \times 10^5$ with a standard deviation of 5.5×10^3. For the full stack, the total number of statistically independent measurements is $K = 27 \times 27 \times 121 = 8.8 \times 10^4$. Thus, the mean observed χ^2 value is approximately 53% higher than the expected value. This deviation is attributed to the high photon count used in this simulation, which amplifies the effect of small differences between the fitted and observed photon counts. As shown in Figure E3 in Appendix E, the χ^2 statistic increasingly deviates from its expected value for photon counts above 2×10^3.

Having established the baseline performance, we now evaluate the effect of model mismatch by fitting the coverslip-calibrated cspline model to data simulated at an imaging depth of 5 μm above the coverslip. At this depth, several micrometers above the coverslip, refractive index mismatch induces significant spherical aberrations, resulting in pronounced changes in PSF shape [16]. The simulated single-molecule images are generated with the same optical parameters. At an imaging depth of 5 μm, we simulated single-molecule images with an increased emitter photon count of 20000, whilst maintaining a background count of 5 photons per pixel. At increased imaging depths, the PSF spreads over a larger spatial footprint, reducing the peak photon count per pixel and thereby decreasing the signal-to-noise ratio. To isolate the effect of model mismatch from this reduction in SNR, a higher total photon count was used.

As shown in Figure 3a, the cspline calibration remains the same, namely at the coverslip. However, it can be seen that the ground truth PSFs at an evaluation depth of 5 μm above the coverslip have changed compared to the PSFs at the coverslip in Figure 2. As a result, the pixel-wise residual between the coverslip-calibrated cspline model and the ground truth PSF at 5 μm increases substantially, with a maximum pixel-wise difference of 29×10^{-3} compared to 3×10^{-3} at the coverslip. The ground-truth PSF at 5 μm is axially elongated and exhibits a change in its orientation. In particular, for $0 \leq z \leq 500$ nm, the coverslip-calibrated model is elongated along the y direction, whereas the ground-truth PSF at 5 μm remains elongated along the x direction.

Next, we assess the localization performance of the coverslip-calibrated cspline model at an evaluation depth of 5 μm. As shown in Figure 3c, the localization bias of a coverslip-calibrated cspline model fitted to images at 5 μm worsens in all dimensions, compared to fitting at the coverslip case. The maximum value for the absolute bias in x and y increased from 6.2 nm and 7.4 nm to 15.5 nm and 16.2 nm in x and y. Moreover, the bias in z has deteriorated substantially, where the mean absolute axial bias across the evaluated axial range $z \in [-600, 600]$ has risen from 7.5 nm to 294 nm. We report

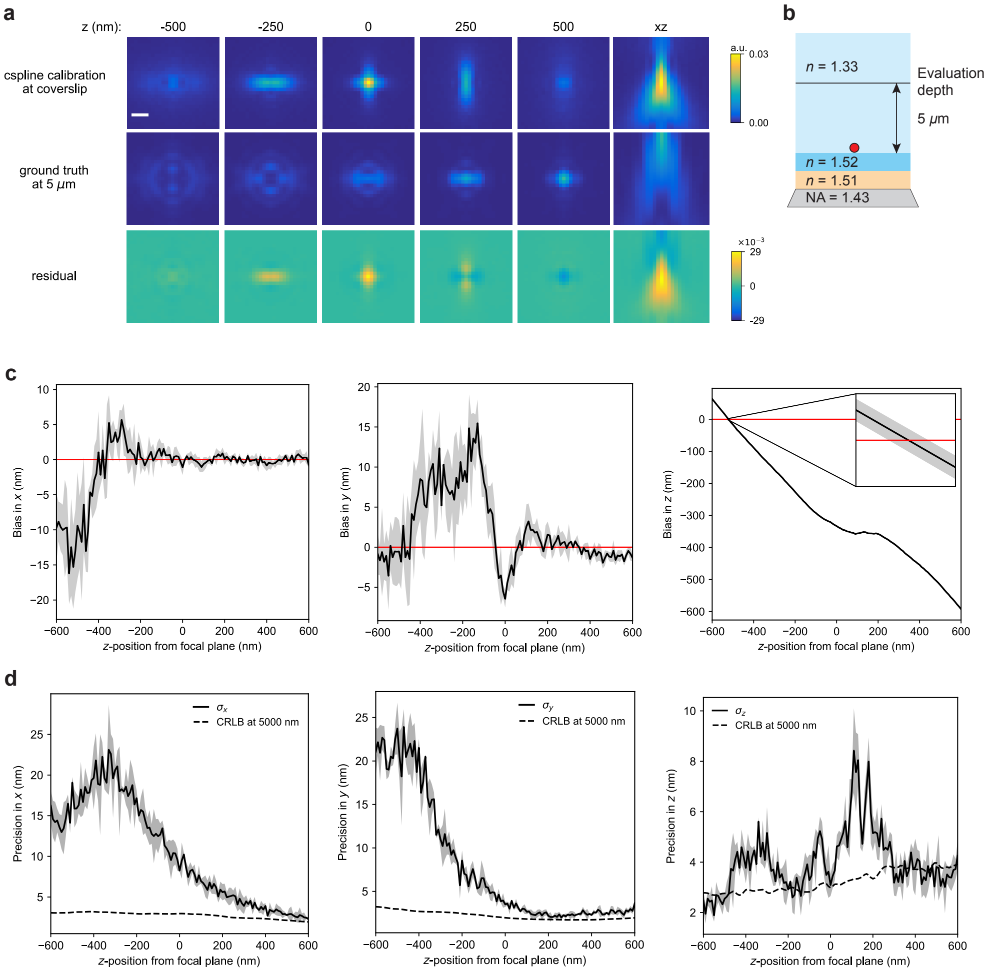


Fig. 3 Localization performance of a cspline model calibrated at the coverslip and evaluated at 5μm above the coverslip. (a) PSFs of the cspline model calibrated at the coverslip at various axial positions (top). The ground truth PSFs at the evaluation depth of $d = 5$ μm above the coverslip (middle). The residual between the coverslip-calibrated model and data at the evaluation depth is computed as a pixel-wise difference (bottom). Scale bar: 500 nm; a.u. = arbitrary units. (b) The calibration is obtained from a bead (red circle) immobilized on the coverslip. The evaluation depth is 5 μm above the coverslip. (c) Localization bias in x, y, and z. Black lines and shaded areas in localization plots represent the mean and standard deviation of the biases. Red lines indicate zero bias. (d) Localization precision in x, y, and z alongside the CRLB. The errors are significantly biased, especially in z. Localization precision was also compromised due to the model mismatch.

the mean absolute bias to avoid cancellation effects when positive and negative biases occur across the evaluated axial range.

At an evaluation depth of 5 μm, the localization precision also degrades. In Figure 3d, it can be seen that across the entire evaluated axial range, the mean localization precision in x has risen to a value of 10.5 nm compared to 7.9 nm at the coverslip. The localization precision in y changes to a mean precision of 11.6 nm compared to 10.0 nm at the coverslip. However, between the range $z \in [-600, -200]$, the mean localization precision increased from 8.2 nm to 16.8 nm in y.

Notably, the axial localization precision occasionally appears to surpass the CRLB. However, the CRLB is only a valid lower bound under the assumption of an unbiased estimator and a correctly specified likelihood model. In this case, the estimator is biased (see Figure 3c), and the fitted model does not match the data-generating process. As a result, the regularity conditions underlying the CRLB are violated, and it no longer represents a valid theoretical bound. Consequently, the comparison between the observed precision and the CRLB is not meaningful in this setting. Still, the CRLB is included for reference, illustrating the degradation in performance under model mismatch.

Lastly, we assess the model quality with the fitted parameters to the data. For the whole fitted stack, the expected χ^2-value remains at $K = 27 \times 27 \times 121 = 8.8 \times 10^4$, because this is solely dependent on the number of pixels in the fitted stack. The mean χ^2-value is 2.8×10^5 with a standard deviation of 9.8×10^3, which is 222% higher than the expected value. Since the data at an imaging depth of 5 μm differs from the calibration data at the coverslip, the resulting χ^2 value is significantly increased. This value will be used as a baseline to assess whether the proposed depth-dependent calibration method improves the goodness-of-fit.

3.2 Localization with a cspline model calibrated at the nearest and interpolated depth

In the previous section, we established the performance of the conventional approach of calibrating a cspline model on data from beads immobilized at the coverslip. We showed that while performance when fitting at the coverslip is optimal (unbiased and precision reaching the CRLB), at an imaging depth of 5 μm above the coverslip, the results become substantially biased and lateral localization precision deteriorates. In this section, we evaluate the calibration results and localization performance from the proposed method at an imaging depth of 5 μm above the coverslip, using a PSF model calibrated on interpolated calibration data evaluated at $d = 5$ μm. Because the proposed method requires calibration data at four depths around the evaluation depth, we also evaluate the localization performance of calibrating a PSF model at the depth nearest to the evaluation depth. Calibration data were simulated at multiple imaging depths around the evaluation depth of 5 μm. Specifically, calibration bead stacks were generated at depths of 4.2 μm, 4.6 μm, 5.4 μm, and 5.8 μm above the coverslip, following the same procedure as described in Section 2.4. For the nearest-depth calibration, a cspline model was calibrated using the bead data at 4.6 μm, which is the closest available depth to the evaluation depth. For the interpolated calibration, a cspline model was constructed as described in Section 2.5 by interpolating between the calibration data at 4.2 μm, 4.6 μm, 5.4 μm, and 5.8 μm, and evaluating the resulting interpolation at 5 μm (see Figure 5b).

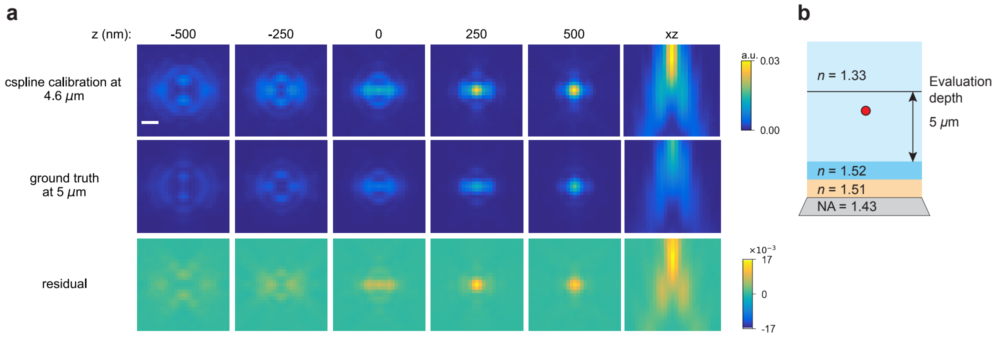


Fig. 4 Cspline calibration at 4.6 μm above the coverslip. (a) PSFs of the cspline model calibrated at the nearest depth to the evaluation depth of $d = 4.6$ μm above the coverslip (top). The ground truth PSFs at the evaluation depth of $d = 5$ μm above the coverslip (middle). The residual between a cspline model calibrated at the nearest depth and data at the evaluation depth is computed as a pixel-wise difference (bottom). Scale bar: 500 nm; a.u. = arbitrary units. (b) The calibration is obtained from a bead (red circle) at 4.6 μm above the coverslip. The evaluation depth is 5 μm above the coverslip.

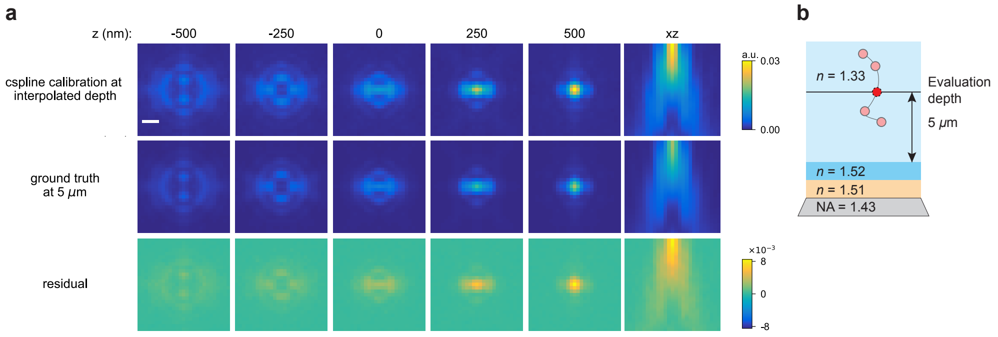


Fig. 5 Cspline calibration at interpolated calibration evaluated at 5μm above the coverslip.(a) PSFs of the cspline model calibrated on interpolated calibration data evaluated at an imaging depth of 5 μm. (top). The ground truth PSFs at the evaluation depth of $d = 5$ μm above the coverslip (middle). The residual between the cspline model calibrated at the interpolated depth and the data at the evaluation depth is computed as a pixel-wise difference (bottom). Scale bar: 500 nm; a.u. = arbitrary units. (b) The calibration is obtained by interpolating between bead data from beads at 4.2 μm, 4.6 μm, 5.4 μm, and 5.8 μm above the coverslip (opaque red beads), and evaluating the interpolation at an imaging depth of $d = 5$ μm above the coverslip (red bead).

First, we evaluate the calibration results for the two proposed methods, as illustrated in Figure 4 and Figure 5. Figure 4a shows the PSFs generated by the cspline model calibrated at 4.6 μm. Compared to the coverslip-calibrated model, the maximum pixel-wise residual decreases from 29×10^{-3} to 17×10^{-3}, indicating an improved PSF representation at this depth. Next, we assess the calibration results of the interpolated model. The corresponding PSFs are shown in Figure 5a. Compared to the cspline models calibrated at the coverslip and at the nearest depth of 4.6 μm, the maximum pixel-wise residual for the interpolated model is 8×10^{-3}, which is the lowest among all evaluated models. In addition, the residual map is spatially more homogeneous compared to the nearest-depth calibration. This suggests that the interpolated calibration provides a more faithful overall representation of the PSF shape at the evaluation depth.

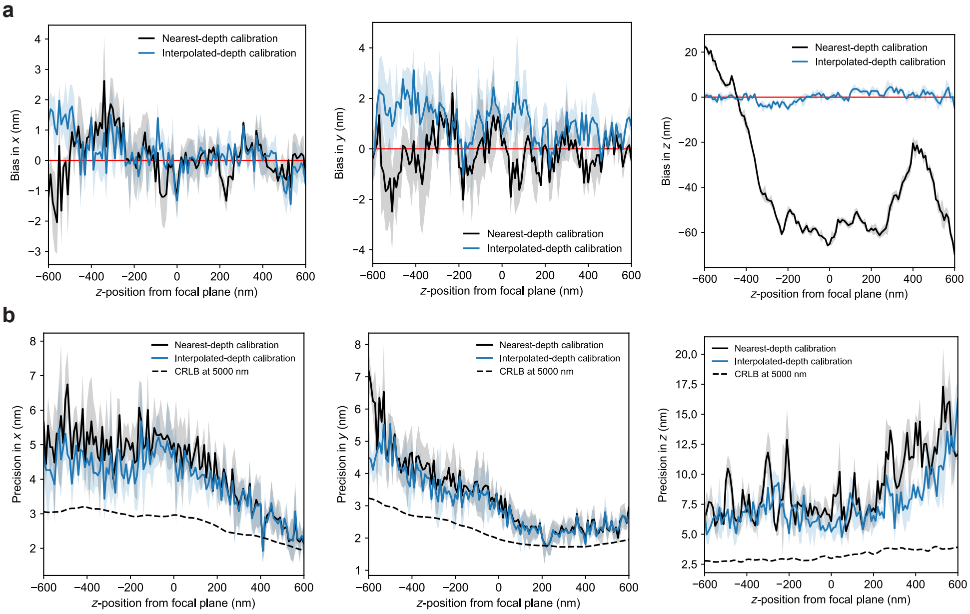


Fig. 6 Localization performance of a cspline model calibrated at the nearest depth of $d = 4.6$ μm above the coverslip and a cspline model constructed from interpolated calibration data evaluated at 5 μm above the coverslip. (a) Localization bias in x, y, and z. Black lines and shaded areas in localization plots represent the mean and standard deviation of the biases. Red lines indicate zero bias. (b) Localization precision in x, y, and z alongside the CRLB. The errors are significantly biased, especially in z. Compared to the localization performance of a coverslip-calibrated model at 5 μm above the coverslip, calibration at the nearest and interpolated depth improves both localization bias and precision.

Next, we evaluate the localization performance of both the cspline model calibrated at 4.6 μm and the interpolated calibration when fitting to data simulated at an evaluation depth of 5 μm. As shown in Figure 6a, when fitting with a cspline model calibrated at the nearest depth, the mean absolute axial bias across the axial range $z \in [-600, 600]$ nm is reduced to 43 nm. This corresponds to a reduction of 86% compared to the mean absolute axial bias of 294 nm obtained with the coverslip-calibrated model. The cspline model calibrated at the interpolated depth reduces the mean absolute bias even more, to a value of 1.8 nm over the same axial range, corresponding to a reduction of 99%. The mean bias across the evaluated z-range is 0.1 nm, while the mean absolute axial bias is 1.8 nm, indicating that the residual axial bias is nearly eliminated.

Figure 6b shows that the lateral localization precision of both models improves compared to the coverslip-calibrated model. The mean precision in x and y decreases to 4.4 nm and 3.3 nm for the cspline model calibrated at 4.6 μm, corresponding to improvements of 58% and 59%, respectively, compared to the coverslip-calibrated model. For the cspline model calibrated at the interpolated depth, the mean localization precision is 4.0 nm in x and 3.1 nm in y. This corresponds to an improvement of 62% and 61% relative to the coverslip-calibrated model. Additionally, for the cspline model calibrated at the interpolated depth, the localization precision of x and y closely follows the CRLB, with differences not exceeding 2.3 nm and 2.2 nm for x and y. For localization precision in z, the difference did not exceed 11.6 nm, and was 3.9 nm on average.

For both models, axial localization precision does not improve compared to the coverslip-calibrated model. The mean axial localization precision is 9.0 nm and 7.5 nm for the cspline model calibrated at the nearest and interpolated depth, respectively, whereas the coverslip-calibrated model yielded an apparent mean axial precision of 2.5 nm. At first glance, the axial localization precision of the coverslip-calibrated model appears to be lower. However, this apparent superior precision of the coverslip-calibrated model is misleading, as it is accompanied by a substantial axial bias. The low precision in this case reflects systematic localization error rather than better localization precision.

Finally, the model quality for both methods was assessed using the χ^2 statistic. The expected value remains at $K = 8.8 \times 10^4$. The mean χ^2 for the cspline model calibrated at the nearest depth was 2.5×10^5 with a standard deviation of 1.2×10^4, corresponding to a difference of 185% to the expected value. This represents an improvement compared to the coverslip-calibrated model, which showed a difference of 222%, indicating improved agreement between the fitted model and the simulated data. For the cspline model calibrated at the interpolated depth, the mean χ^2 value was 2.2×10^5 with a standard deviation of 1.1×10^4, which is 150% higher than the expected value. This is an improvement over both the coverslip-calibrated model and the nearest-depth calibrated model, indicating that the cspline model at the interpolated depth provides the best fit to the simulated data.

In this section, we have shown that the cspline models calibrated at the interpolated depth and at the corresponding nearest depth perform comparably in terms of localization precision. In this specific scenario, however, the interpolated calibration performs substantially better in terms of axial bias, reducing the mean absolute axial bias from 294 nm to 1.8 nm, whereas the nearest-depth model only reduces it to 43 nm. In Section 3.3, we therefore investigate how the spacing between calibration depths affects localization performance, for which conditions nearest-depth calibration may provide comparable performance, and at what interpolation distances the interpolated calibration begins to deteriorate.

3.3 Distance between interpolation depths

In this section, we investigate how the spacing between calibration depths used for interpolation affects localization performance. In contrast to the previous simulation experiments, which were held to demonstrate the effect of depth-dependent model mismatch and the benefit of interpolation at a representative aberration setting, the purpose of this section is to assess the robustness of the proposed interpolation approach across different calibration spacings and under experimentally fitted aberration coefficients. For this reason, the simulations in this section use aberration coefficients based on fitted Zernike aberrations from the experimental bead stack provided by the Ries Group (see Figure E7 for the fitted aberrations). The Zernike coefficients were obtained by fitting a vectorial PSF model to an experimentally measured bead stack using the aberration-fitter implemented in VectorialPSF by van Velde [24]. Calibration stacks are generated at depths $\{d-2\delta,\ d-\delta,\ d+\delta,\ d+2\delta\}$, interpolated and evaluated to obtain a PSF at an intermediate evaluation depth. The interpolation distances were $\delta \in [100, 200, 400, 800, 1000, 1500]$ nm.

In Section 3.2, we showed that interpolation with $\delta = 400$ nm at an imaging depth of $d = 5$ μm above the coverslip already significantly improves localization precision and reduces axial bias compared to the coverslip-calibrated model. In this section, we change the evaluation depth to both 4 μm and 6 μm above the coverslip, and the simulation parameters can be found in Tables E1– E2 in Appendix E. The results at an evaluation depth of 6 μm above the coverslip can be found in Appendix E. Simultaneously, for each interpolation distance δ, we also evaluate the localization performance of a cspline model calibrated at the nearest depth to the evaluation depth. For example, for an evaluation depth of $d = 4000$ nm and $\delta = 100$ nm, interpolation is performed between stacks at depths $d_i \in \{3800, 3900, 4100, 4200\}$ nm.

Figure 7 summarizes the mean localization precision and mean absolute bias across the evaluated axial range $z \in [-600, 600]$ nm for the interpolated model and the two nearest-depth calibrations, as a function of the interpolation distance δ, at an evaluation depth of 4 μm. The corresponding full localization curves as a function of z are shown in Figures E4–E6 in Appendix E. These results show that the localization precision remains largely unaffected by the interpolation distance δ. Compared to the corresponding nearest-depth calibrations, the mean precision across the evaluated z-range remains within approximately 1 nm in x and y, and within 3–4 nm in z, for all models, with overlapping standard deviations.

In contrast, the mean absolute axial bias is consistently and substantially reduced when using the interpolated-depth calibration. For interpolation distances $\delta \geq 200$ nm, the interpolated model consistently outperforms both nearest-depth calibrations, typically by a factor of two or more. At $\delta = 800$ nm, the mean absolute axial bias is reduced from approximately 60 nm for nearest-depth calibration to 7.1 nm for interpolated calibration. A similar trend is observed for the lateral bias in y for small to moderate interpolation distances. For $\delta \leq 600$ nm, the interpolated model yields substantially lower mean absolute bias compared to both nearest-depth calibrations. For example, at $\delta = 400$ nm, the mean absolute bias in y is reduced from approximately 13 nm (nearest below) and 12 nm (nearest above) to 4.4 nm for the interpolated model. However, for larger interpolation distances ($\delta \geq 800$ nm), the lateral bias

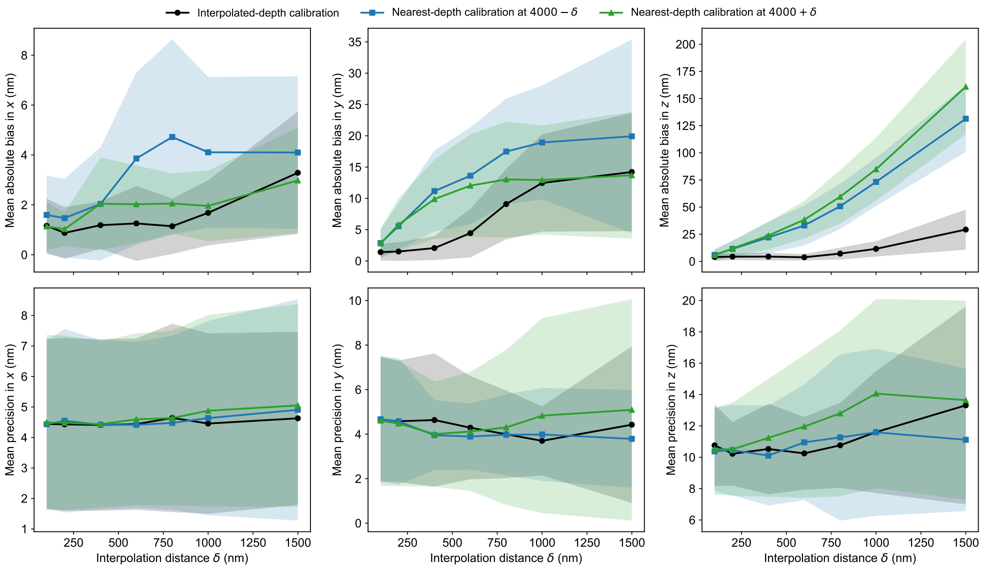


Fig. 7 Mean localization precision and absolute bias in x, y, and z as a function of the calibration depth offset relative to the evaluation depth of 4 μm. Three models are plotted alongside each other, the cspline model calibrated at the interpolated depth with interpolation distance δ, and the cspline models calibrated at the nearest depths at 4 μm $\pm\delta$ nm. Lines denote the mean, and shaded areas denote the standard deviation of the bias and precision across the evaluated z-range.

increases and becomes comparable to or larger than that of the nearest-depth calibrations. For smaller interpolation distances ($\delta \approx 100$ nm), the axial bias of interpolated and nearest-depth calibration is comparable, with nearest-depth calibration occasionally yielding slightly lower bias. This trend persists across both evaluated depths of 4 and 6 μm above the coverslip, where for an imaging depth of 6 μm, the full localization curves are shown in Figures E8–E10, and the summary plot in Figure E11 in Appendix E.

These results indicate that nearest-depth calibration and interpolated-depth calibration perform similarly in localization precision, with no clear precision advantage for either approach. The main difference between the two methods is therefore the axial bias. Interpolation provides lower axial bias than nearest-depth calibration once the nearest calibration stack is more than about 200 nm away from the evaluation depth. For smaller offsets, around $\delta = 100$ nm, nearest-depth calibration and interpolation yield comparable results, and nearest-depth calibration can occasionally give lower axial bias. For $\delta \geq 200$ nm, however, calibration at the interpolated depth consistently outperforms nearest-depth calibration in axial bias, typically by a factor of two or more.

Lastly, interpolation exhibited low axial bias over a broad range of interpolation distances. Across the two simulated evaluation depths, the mean absolute axial bias remained below approximately 5 nm for $\delta \leq 600$ nm, and remained on the order of

4–10 nm up to approximately $\delta = 800$ nm. For larger interpolation distances, the axial bias increased gradually, with clearer degradation observed at $\delta = 1000$ nm and substantial degradation at $\delta = 1500$ nm (21.8 nm at 4 μm and 29.3 nm at 6 μm). We therefore conclude that interpolation is robust up to approximately 800 nm in the simulated setting, while $\delta \leq 600$ nm provides a conservative range for consistently minimal bias.

4 Experimental Results

In this section, we validate the proposed depth-dependent calibration method on experimentally acquired data provided by the Ries Group (European Molecular Biology Laboratory, EMBL). After performing the data preprocessing and depth assignment steps described in Sections 2.6.1–2.6.2, bead stacks at various assigned depths were obtained (see Table H3 in Appendix H).

Figure 8 shows representative calibration and localization results for beads acquired at depths $d \in \{1305, 2667, 3313, 4448, 4816\}$ nm. To evaluate the proposed method, the bead stack at an assigned depth of 3313 nm was excluded from the interpolation procedure and used as the evaluation dataset. We compare the following calibration strategies: calibration at the coverslip, calibration at the nearest available depth (2667 nm), and calibration using interpolated calibration data evaluated at 3313 nm.

In Figure 8a, the PSFs of the different cspline models are shown at several axial positions. The corresponding localization performance is presented in Figure 8b. Consistent with the simulation results, the coverslip-calibrated model exhibits a substantial axial bias, with a mean absolute bias of approximately 76.4 nm.

Both the cspline models calibrated at the nearest and interpolated depths reduce the mean absolute axial bias. The mean absolute axial bias is reduced to approximately 15.4 nm and 24.7 nm for the interpolated and nearest-depth calibrations, respectively, corresponding to reductions of approximately 80% and 68% relative to the coverslip-calibrated model.

The localization precision remains largely unchanged across calibration strategies (coverslip, nearest, and interpolated). Consistent with the simulation results, the coverslip-calibrated model appears to yield a lower axial localization precision, however, this is accompanied by a substantial bias and therefore reflects systematic localization errors rather than improved localization precision.

To assess the performance across different experimental conditions, multiple evaluation cases were considered, corresponding to different target depths and interpolated calibration stacks. An overview of all evaluated cases, including the evaluation depth, interpolated calibrations, and corresponding nearest-depth distance δ, is provided in Table H3 in Appendix H. In the majority of the evaluated cases, interpolation yields lower axial bias than nearest-depth calibration. This is consistent with the trends observed in simulation. However, in a small number of cases, nearest-depth calibration slightly outperformed interpolation. This deviation from the simulation results is expected, as the experimental setting introduces additional factors that were not present in the simulation, including non-uniform spacing of calibration depths and variations in the interpolation distance δ. As a result, the experimental cases do not isolate the effect of δ alone.

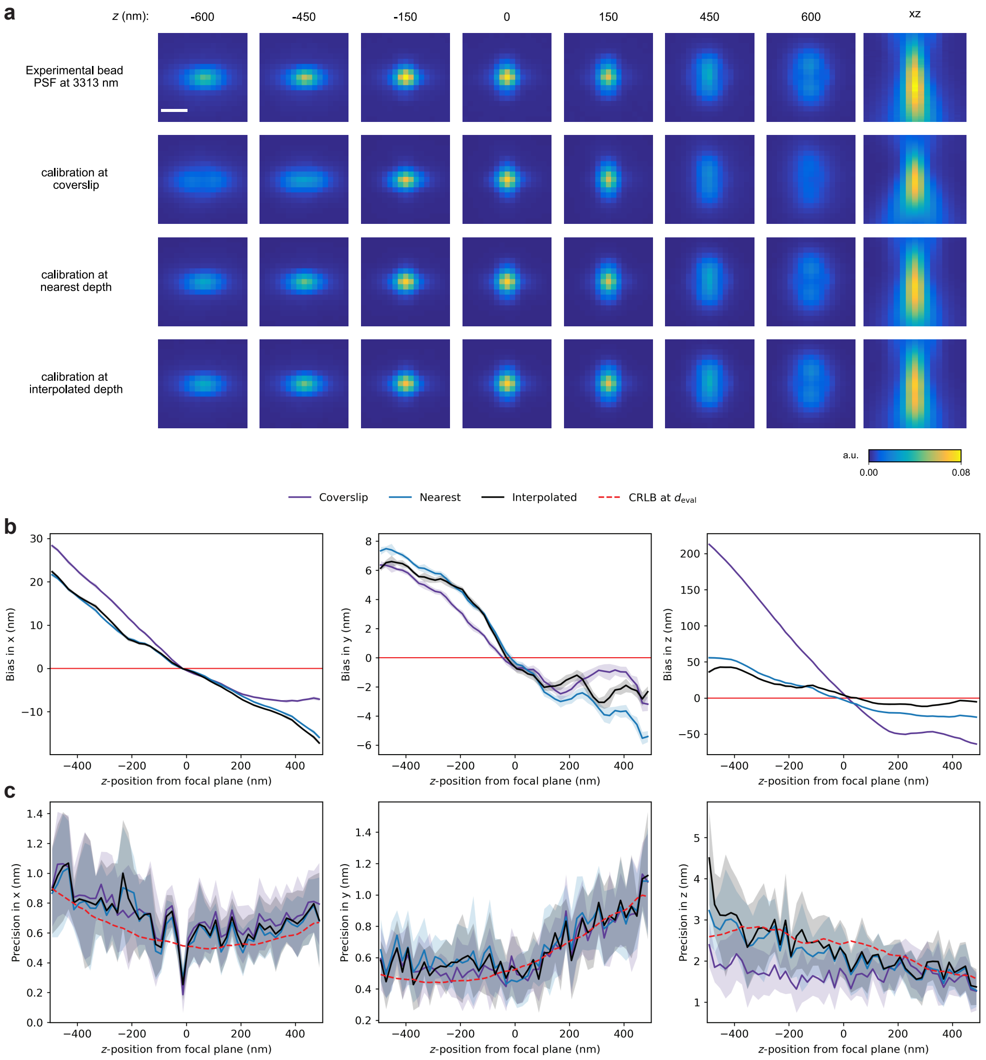


Fig. 8 (a) PSF of the Experimental bead data at an assigned depth of 3928 nm and PSFs of a cspline model calibrated at the coverslip, nearest depth of 3794 nm, and interpolated depth. (b) Localization bias in x, y, and z of a cspline calibrated at the coverslip, nearest, and interpolated depth, when fitting to experimental bead data. Lines and shaded areas in localization plots represent the mean and standard deviation of the biases of each model. Red lines indicate zero bias. The x, y and z biases versus across z are defined as in Equations D28-D29 in Appendix D. (c) Localization precision in x, y, and z alongside the CRLB of each at the evaluation depth. Scalebar: 500 nm; a.u. = arbitrary units.

A possible explanation for this behavior is that, although Catmull–Rom splines support non-uniformly spaced control points, their performance depends on the distribution of the available calibration data. When calibration depths are asymmetrically distributed or contain large gaps, the interpolation may be influenced by calibration stacks that are relatively far from the evaluation depth. In such cases, a single nearby calibration stack can provide a better local approximation of the PSF than interpolation. Because both the interpolation distance δ and the spacing between calibration depths vary across the evaluated cases, the individual contributions of these factors cannot be disentangled. These results therefore suggest that interpolation performance depends on both the distance to the nearest calibration depth and the distribution of the available calibration depths.

For example, in the case with $d_{\text{eval}} = 3817$ nm and $\delta = 136$ nm (using $d_i \in \{3376, 3465, 3953, 5080\}$ nm), nearest-depth calibration yielded a mean absolute axial bias of 16.0 nm, whereas interpolation resulted in 28.0 nm. This agrees with the simulated results, where calibration stacks within about 100 nm of the evaluation depth can already provide performance comparable to or better than interpolation. In contrast, for the case with $d_{\text{eval}} = 2915$ nm and $\delta = 540$ nm (using $d_i \in \{1970, 2375, 3465, 3817\}$ nm), nearest-depth and interpolated calibration gave 6.3 nm and 9.6 nm, respectively, showing that nearest-depth can also outperform interpolated calibration at larger offsets under certain distribution of calibration depths. The remaining exception at $d_{\text{eval}} = 2883$ nm and $\delta = 584$ nm (using $d_i \in \{1381, 2225, 3467, 4025\}$ nm), where nearest-depth calibration slightly outperformed interpolation (13.0 nm to 15.2 nm), is again likely explained by the strongly asymmetric depth distribution of the available calibration stacks.

These cases highlight that interpolation performance is influenced by both the distance to the nearest calibration stack and the remaining distribution of the interpolated calibration stacks. In particular, non-uniform spacing between calibration depths can reduce interpolation accuracy when large gaps exist between neighboring control points, as the interpolation becomes less local and incorporates information from more distant depths. For example, in the first case, the spacing between $d_3 = 3953$ nm and $d_4 = 5080$ nm is substantially larger than between $d_1 = 3376$ nm and $d_4 = 3465$ nm. This uneven spacing causes the interpolation to include the influence of distant calibration depths.

Overall, the experimental results confirm the trends observed in the simulation: interpolation substantially reduces axial bias while maintaining consistent localization precision. While nearest-depth calibration may occasionally yield comparable or slightly better results in specific configurations, interpolation provides a robust and generally reliable calibration strategy across varying depths. The remaining localization results for all evaluated cases are provided in Figures F12-F13 in Appendix F.

5 Conclusion

In this work, we proposed a depth-dependent calibration approach that enables accurate PSF modeling at arbitrary imaging depths by interpolating between calibration stacks acquired at discrete depths. This approach addresses depth-induced model mismatch and improves localization performance at depths where a conventional coverslip calibration fails.

Using simulated data, we showed that a conventional cspline model calibrated at the coverslip fails to accurately represent the PSF at larger imaging depths, resulting in a mean absolute axial bias of 294 nm at 5 μm. In contrast, the proposed interpolated calibration approach significantly improved localization performance, reducing the axial bias by 99% to 1.8 nm and improving lateral precision by over 60%.

We further investigated how the distance between the evaluation depth and the nearest available calibration stack affects localization performance. The results showed that interpolation provides stable localization precision across a wide range of distances, while substantially reducing axial bias as this distance increases.

The method was also validated on experimental data, where interpolation reduced axial bias by up to 80% compared to coverslip calibration while maintaining consistent localization precision. In most cases, interpolation also outperformed nearest-depth calibration. Deviations from this trend were observed in scenarios with both non-uniform and asymmetric distributions of the calibration stacks, where nearest-depth calibration occasionally yielded slightly lower bias. These findings indicate that interpolation performance depends not only on the distance to the nearest calibration depth, but also on the distribution of the remaining interpolated calibration stacks.

These results provide a practical guideline for calibration in 3D SMLM with PSF engineering. When a calibration stack is available within approximately 100–200 nm of the evaluation depth, nearest-depth calibration is sufficient. For larger offsets, interpolation should be preferred. For optimal performance, calibration stacks should be acquired densely in depth and distributed as uniformly as possible around the imaging region of interest. When such data are available, interpolation enables accurate and robust localization across extended imaging depths.

Future work could focus on improving the robustness of the calibration pipeline by automatically selecting well-distributed calibration subsets from densely sampled bead stacks based on the evaluation depth. In addition, the Catmull–Rom formulation provides C^1-continuity, which may limit smoothness compared to higher-order spline methods. Future work could therefore investigate higher-order interpolation schemes, such as quintic Catmull–Rom splines, to achieve C^2-continuity. Lastly, validation on full SMLM datasets would provide insight into the performance of the method under realistic experimental conditions.

References

[1] Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F.: Imaging intracellular fluorescent proteins at nanometer resolution. Science **313**, 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**, 793–795 (2006) https://doi.org/10.1038/nmeth929

[3] Hess, S.T., Girirajan, T.P.K., Mason, M.D.: Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophysical journal **91**, 4258–72 (2006) https://doi.org/10.1529/biophysj.106.091116

[4] Huang, B., Wang, W., Bates, M., Zhuang, X.: Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. Science **319**, 810–813 (2008) https://doi.org/10.1126/science.1153529

[5] Pavani, S.R.P., Thompson, M.A., Biteen, J.S., Lord, S.J., Liu, N., Twieg, R.J., Piestun, R., Moerner, W.E.: Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. Proceedings of the National Academy of Sciences **106**, 2995–2999 (2009) https://doi.org/10.1073/pnas.0900245106

[6] Shechtman, Y., Sahl, S.J., Backer, A.S., Moerner, W.E.: Optimal point spread function design for 3d imaging. Physical Review Letters **113**, 133902 (2014) https://doi.org/10.1103/PhysRevLett.113.133902

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

[8] Booth, M.J., Neil, M.A.A., Wilson, T.: Aberration correction for confocal imaging in refractive-index- mismatched media. Journal of Microscopy **192**, 90–98 (1998) https://doi.org/10.1111/j.1365-2818.1998.99999.x

[9] Petrov, P.N., Moerner, W.E.: Addressing systematic errors in axial distance measurements in single-emitter localization microscopy. Optics Express **28**, 18616 (2020) https://doi.org/10.1364/OE.391496

[10] Burke, D., Patton, B., Huang, F., Bewersdorf, J., Booth, M.J.: Adaptive optics correction of specimen-induced aberrations in single-molecule switching microscopy. Optica **2**, 177 (2015) https://doi.org/10.1364/OPTICA.2.000177

[11] Mlodzianoski, M.J., Cheng-Hathaway, P.J., Bemiller, S.M., McCray, T.J., Liu, S., Miller, D.A., Lamb, B.T., Landreth, G.E., Huang, F.: Active psf shaping and

adaptive optics enable volumetric localization microscopy through brain sections. Nature Methods **15**, 583–586 (2018) https://doi.org/10.1038/s41592-018-0053-8

[12] Siemons, M., Hulleman, C.N., Thorsen, R., Smith, C.S., Stallinga, S.: High precision wavefront control in point spread function engineering for single emitter localization. Optics Express **26**, 8397 (2018) https://doi.org/10.1364/oe.26.008397

[13] Liu, S., Kromann, E.B., Krueger, W.D., Bewersdorf, J., Lidke, K.A.: Three dimensional single molecule localization using a phase retrieved pupil function. Optics Express **21**, 29462 (2013) https://doi.org/10.1364/oe.21.029462

[14] HANSER, B.M., GUSTAFSSON, M.G.L., AGARD, D.A., SEDAT, J.W.: Phase-retrieved pupil functions in wide-field fluorescence microscopy. Journal of Microscopy **216**, 32–48 (2004) https://doi.org/10.1111/j.0022-2720.2004.01393.x

[15] Xu, F., Ma, D., MacPherson, K.P., Liu, S., Bu, Y., Wang, Y., Tang, Y., Bi, C., Kwok, T., Chubykin, A.A., Yin, P., Calve, S., Landreth, G.E., Huang, F.: Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval. Nature Methods **17**, 531–540 (2020) https://doi.org/10.1038/s41592-020-0816-x

[16] Li, Y., Wu, Y.-L., Hoess, P., Mund, M., Ries, J.: Depth-dependent psf calibration and aberration correction for 3d single-molecule localization. Biomedical Optics Express **10**, 2708 (2019) https://doi.org/10.1364/boe.10.002708

[17] Stallinga, S., Rieger, B.: Accuracy of the gaussian point spread function model in 2d localization microscopy. Optics Express **18**, 24461 (2010) https://doi.org/10.1364/OE.18.024461

[18] Smith, C., Huisman, M., Siemons, M., Grünwald, D., Stallinga, S.: Simultaneous measurement of emission color and 3d position of single molecules. Optics Express **24**, 4996 (2016) https://doi.org/10.1364/oe.24.004996

[19] Siemons, M.E., Kapitein, L.C., Stallinga, S.: Axial accuracy in localization microscopy with 3d point spread function engineering. Optics Express **30**, 28290 (2022) https://doi.org/10.1364/oe.461750

[20] Velde, P., Rieger, B., Hindsdale, T., Cnossen, J., Fan, D., Hung, S.-T., Grunwald, D., Smith, C.: Zimflux: Single molecule localization microscopy with patterned illumination in 3d. Optics Express **31**, 42701 (2023) https://doi.org/10.1364/oe.505958

[21] Wilson, T.: Confocal microscopy. In: Yacobi, B.G., Holt, D.B., Kazmerski, L.L. (eds.) Microanalysis of Solids, pp. 219–232. Springer, Boston, MA (1994). https://doi.org/10.1007/978-1-4899-1492-7_7

[22] Török, P., Higdon, P.D., Wilson, T.: Theory for confocal and conventional microscopes imaging small dielectric scatterers. Journal of Modern Optics **45**, 1681–1698 (1998) https://doi.org/10.1080/09500349808230662

[23] Stallinga, S.: Effect of rotational diffusion in an orientational potential well on the point spread function of electric dipole emitters. Journal of the Optical Society of America A **32**, 213 (2015) https://doi.org/10.1364/josaa.32.000213

[24] Velde, P.: VectorialPSF: Simulation of vectorial point spread functions. GitHub. Accessed: 2023-11-30 (2023)

[25] Li, Y., Mund, M., Hoess, P., Deschamps, J., Matti, U., Nijmeijer, B., Sabinina, V.J., Ellenberg, J., Schoen, I., Ries, J.: Real-time 3d single-molecule localization using experimental point spread functions. Nature Methods **15**, 367–369 (2018) https://doi.org/10.1038/nmeth.4661

[26] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C (2nd Ed.): the Art of Scientific Computing. Cambridge University Press, USA (1992)

[27] Hohlbein, J.: Single-molecule localization microscopy as an emerging tool to probe multiscale food structures. Food Structure **30**, 100236 (2021) https://doi.org/10.1016/j.foostr.2021.100236

[28] Babcock, H.P., Zhuang, X.: Analyzing single molecule localization microscopy data using cubic splines. Scientific Reports **7** (2017) https://doi.org/10.1038/s41598-017-00622-w

[29] Ries, J.: Smap: a modular super-resolution microscopy analysis platform for smlm data. Nature Methods **17**, 870–872 (2020) https://doi.org/10.1038/s41592-020-0938-1

[30] Liu, S., Chen, J., Hellgoth, J., Müller, L.R., Ferdman, B., Karras, C., Xiao, D., Lidke, K.A., Heintzmann, R., Shechtman, Y., Li, Y., Ries, J.: Universal inverse modeling of point spread functions for smlm localization and microscope characterization. Nature Methods **21**, 1082–1093 (2024) https://doi.org/10.1038/s41592-024-02282-x

[31] Huang, F., Schwartz, S.L., Byars, J.M., Lidke, K.A.: Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. Biomedical Optics Express **2**, 1377 (2011) https://doi.org/10.1364/BOE.2.001377

[32] Kay, S.M.: Fundamentals of Statistical Signal Processing, Volume II: Detection Theory. Fundamentals of Statistical Signal Processing, vol. 2. Prentice Hall, Upper Saddle River, NJ (1998)

[33] Barry, P.J., Goldman, R.N.: A recursive evaluation algorithm for a class of

catmull-rom splines. ACM SIGGRAPH Computer Graphics **22**, 199–204 (1988) https://doi.org/10.1145/378456.378511

[34] IEEE Standard for Floating-Point Arithmetic. IEEE

[35] Boor, C.: A Practical Guide to Splines. Applied Mathematical Sciences, vol. 27. Springer, New York, NY (1978)

[36] Piegl, L., Tiller, W.: The NURBS Book, 2nd edn. Monographs in Visual Communication. Springer, Berlin, Heidelberg (1997). https://doi.org/10.1007/978-3-642-59223-2

[37] Farin, G.: Curves and Surfaces for Computer-Aided Geometric Design: A Practical Guide, 5th edn. Academic Press, San Diego, CA (2002)

[38] Bos, A.: Parameter Estimation for Scientists and Engineers. Wiley, Hoboken, NJ (2007). https://doi.org/10.1002/9780470173862

[39] Abraham, A.V., Ram, S., Chao, J., Ward, E.S., Ober, R.J.: Quantitative study of single molecule location estimation techniques. Optics Express **17**, 23352 (2009) https://doi.org/10.1364/OE.17.023352

[40] Mortensen, K.I., Churchman, L.S., Spudich, J.A., Flyvbjerg, H.: Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nature Methods **7**, 377–381 (2010) https://doi.org/10.1038/nmeth.1447

[41] Small, A., Stahlheber, S.: Fluorophore localization algorithms for super-resolution microscopy. Nature Publishing Group (2014). https://doi.org/10.1038/nmeth.2844

Appendix A Spline interpolation

A.1 Uniform Catmull–Rom splines

A 1D Catmull-Rom spline $f_i(x)$ is defined on the interval $[t_i, t_{i+1}]$ by interpolation between p_i at t_i and p_{i+1} at t_{i+1}, with prescribed tangents p'_i and p'_{i+1} at those points. A Catmull–Rom spline defines the tangents using points p_{i-1} and p_{i+2} consecutive to points p_i and p_{i+1} respectively. The tangent at p_i is determined by the vector $p_{i+1} - p_{i-1}$, and the tangent at p_{i+1} from $p_{i+2} - p_i$.

For each segment, the spline is constrained to interpolate the endpoints p_i and p_{i+1}, while the endpoint tangents are chosen from neighboring control points:

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix}, \tag{A1}$$

for $i = 1, 2, \ldots, N$, with N being the number of pixels in the image along one dimension. The values $f_i(0)$ and $f_i(1)$ correspond to the start and end of the spline segment, respectively. For a cubic Catmull-Rom spline, each segment is represented by a cubic polynomial:

$$f_i(r) = a_i r^3 + b_i r^2 + c_i r + d_i, \tag{A2}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, and $t_i \leq x < t_{i+1}$. Substituting this polynomial expression into the constraints from Equation A1 yields the following fully determined system of equations to obtain 4 coefficients:

$$\begin{bmatrix} f_i(0) \\ f'_i(0) \\ f_i(1) \\ f'_i(1) \end{bmatrix} = \begin{bmatrix} d_i \\ c_i \\ a_i + b_i + c_i + d_i \\ 3a_i + 2b_i + c_i \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}}_{\mathbf{C}} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix}. \tag{A3}$$

To obtain the coefficients a_i, b_i, c_i and d_i, we solve the system using matrix inversion:

$$\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \mathbf{C}^{-1} \begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix} = \underbrace{\begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -0.5 & 0 & 0.5 \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{A4}$$

where we expressed the tangents through the quadruplet of the points p_{i-1}, p_i, p_{i+1}, p_{i+2} as follows:

$$\begin{bmatrix} p_i \\ \frac{p_{i+1}-p_{i-1}}{2} \\ p_{i+1} \\ \frac{p_{i+2}-p_i}{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -0.5 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}. \tag{A5}$$

Finally, substituting these coefficients back into Equation A2 gives the matrix form of a cubic Catmull–Rom spline segment:

$$f_i(r) = \begin{bmatrix} r^3 & r^2 & r & 1 \end{bmatrix} \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{A6}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, with $t_i \leq x < t_{i+1}$. The spline segment defined on this interval interpolates the points p_i and p_{i+1}, such that

$$f(t_i) = p_i, \quad f(t_{i+1}) = p_{i+1}.$$

A.2 Non-uniform Catmull–Rom splines

The standard formulation of Catmull–Rom splines introduced in the previous section implicitly assumes uniform spacing of the parameter values t_i. In the context of this thesis, calibration PSF stacks are not guaranteed to be obtained at uniformly spaced imaging depths. Therefore, we introduce the formulation proposed by Barry and Goldman [33], which allows the construction of Catmull–Rom splines for non-uniformly spaced parameter values.

Barry and Goldman derive the spline formulation using Neville's algorithm for Lagrange interpolation. In this formulation, the tangent at point p_i is expressed as a weighted combination of the slopes of the two neighboring intervals. Specifically, the tangent is given by:

$$\begin{aligned} p_i' &= \frac{t_{i+1}-t_i}{t_{i+1}-t_{i-1}}\left(\frac{p_i-p_{i-1}}{t_i-t_{i-1}}\right) + \frac{t_i-t_{i-1}}{t_{i+1}-t_{i-1}}\left(\frac{p_{i+1}-p_i}{t_{i+1}-t_i}\right) \\ p_{i+1}' &= \frac{t_{i+2}-t_{i+1}}{t_{i+2}-t_i}\left(\frac{p_{i+1}-p_i}{t_{i+1}-t_i}\right) + \frac{t_{i+1}-t_i}{t_{i+2}-t_i}\left(\frac{p_{i+2}-p_{i+1}}{t_{i+2}-t_{i+1}}\right) \end{aligned} \tag{A7}$$

.

Next, we define the following notations to shorten the expression in Equation A7:

$$\begin{cases} \Delta_i = t_{i+1} - t_i \\ \Delta_{i-1} = t_i - t_{i-1} \\ \Delta_{i+1} = t_{i+2} - t_{i+1} \\ \Delta_{i,i-1} = \Delta_i + \Delta_{i-1} = t_{i+1} - t_{i-1} \\ \Delta_{i,i+1} = \Delta_i + \Delta_{i+1} = t_{i+2} - t_i \end{cases} \tag{A8}$$

After substituting the notations in Equation A8 into the tangent definition in Equation A7 and after some rewriting we obtain the following expressions for the tangents:

$$\begin{aligned} p_i' &= \frac{\Delta_{i-1}}{\Delta_{i,i-1}\Delta_i}\, p_{i+1} + \left(\frac{\Delta_i}{\Delta_{i-1}} - \frac{\Delta_{i-1}}{\Delta_i}\right)\frac{1}{\Delta_{i,i-1}}\, p_i - \frac{\Delta_i}{\Delta_{i,i-1}\Delta_{i-1}}\, p_{i-1} \\ p_{i+1}' &= \frac{\Delta_i}{\Delta_{i,i+1}\Delta_{i+1}}\, p_{i+2} + \left(\frac{\Delta_{i+1}}{\Delta_i} - \frac{\Delta_i}{\Delta_{i+1}}\right)\frac{1}{\Delta_{i,i+1}}\, p_{i+1} - \frac{\Delta_{i+1}}{\Delta_{i,i+1}\Delta_i}\, p_i. \end{aligned} \tag{A9}$$

The system of equations to obtain the spline coefficients, expressed in a vector containing the quadruplets of the points $p_{i-1}, p_i, p_{i+1}, p_{i+2}$ is as follows:

$$\begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \mathbf{C}^{-1} \begin{bmatrix} p_i \\ p_i' \\ p_{i+1} \\ p_{i+1}' \end{bmatrix} = \begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \beta_1 & \beta_2 & \beta_3 \end{bmatrix} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{A10}$$

where α_i and β_i for $i \in 1, 2, 3$ are defined as follows

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\Delta_{i-1}}{\Delta_i} \\ \frac{\Delta_i}{\Delta_{i-1}} \end{bmatrix} \frac{1}{\Delta_{i,i-1}}, \tag{A11}$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\Delta_{i-1}}{\Delta_i} \\ \frac{\Delta_i}{\Delta_{i+1}} \end{bmatrix} \frac{1}{\Delta_{i,i+1}}. \tag{A12}$$

Lastly, substituting these coefficients into Equation A2 gives the matrix form of a cubic non-uniform Catmull–Rom spline:

$$f_i(r) = \begin{bmatrix} r^3 & r^2 & r & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 2 & 1 & -2 & 1 \\ -3 & -2 & 3 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ \alpha_1 & \alpha_2 & \alpha_3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \beta_1 & \beta_2 & \beta_3 \end{bmatrix}}_{\mathbf{D}'} \begin{bmatrix} p_{i-1} \\ p_i \\ p_{i+1} \\ p_{i+2} \end{bmatrix}, \tag{A13}$$

where $r = (x - t_i)/(t_{i+1} - t_i)$ normalizes the interval $[t_i, t_{i+1}]$ to $[0, 1]$, with $t_i \leq x < t_{i+1}$. The spline segment defined on this interval interpolates the points p_i and p_{i+1}, such that

$$f(t_i) = p_i, \quad f(t_{i+1}) = p_{i+1}.$$

In this formulation, the arbitrary parameter spacing is explicitly taken into account when determining the interpolation tangents, while the underlying Catmull–Rom spline is preserved. The matrix formulation of the non-uniform Catmull–Rom spline in Equation A13 can be written to the following compact expression:

$$f_i(r) = \sum_{m=0}^{3} \sum_{n=0}^{3} r^{3-m} A_{m,n} q_n, \tag{A14}$$

where $A = \mathbf{HD'}$, and $q = \begin{bmatrix} p_{i-1} & p_i & p_{i+1} & p_{i+2} \end{bmatrix}^T$.

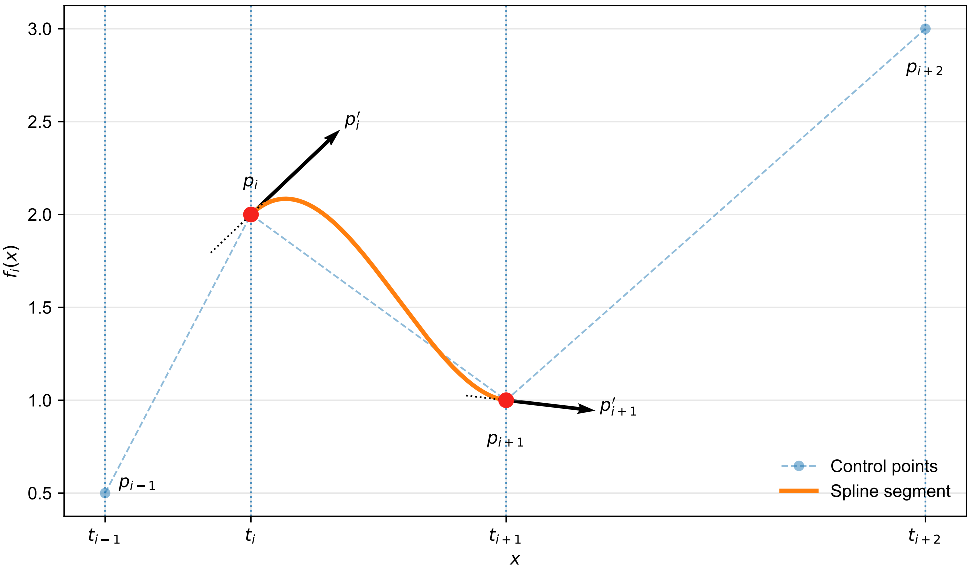


Fig. A1 Example of a 1D non-uniform Catmull–Rom spline segment defined by the four points $p_{i-1}, p_i, p_{i+1}, p_{i+2}$ with values 0.5, 2, 1, and 3, respectively. The corresponding parameter values are non-uniformly spaced at $t_{i-1} = 0.0$, $t_i = 0.8$, $t_{i+1} = 2.2$, and $t_{i+2} = 4.5$. The spline segment interpolates the points p_i and p_{i+1}, and the tangent vectors p_i' and p_{i+1}' are determined according to the non-uniform Catmull–Rom formulation.

A.3 Catmull–Rom interpolation of PSF stacks along depth

To interpolate between three-dimensional calibration stacks acquired at non-uniformly spaced imaging depths, Catmull–Rom interpolation is applied along the depth dimension. Each voxel value is treated as a function of the depth coordinate d, and interpolated using the non-uniform Catmull–Rom formulation from Equation A13.

Let $W_{i,j,k,l}$ denote the sampled PSF value at voxel (i, j, k) and calibration depth d_l. For each fixed voxel location (i, j, k), interpolation along the depth dimension is given by:

$$f_{i,j,k}(d) = \sum_{p=0}^{3} \sum_{p'=0}^{3} \left(\frac{d - d_l}{\Delta d} \right)^{3-p} A_{p,p'}\, W_{i,j,k,l+p'-1}, \tag{A15}$$

where $\Delta d = d_{l+1} - d_l$ with $d \in [d_l, d_{l+1}]$, and $A = \mathbf{HD'}$ is the Catmull–Rom coefficient matrix from Equation A13. This interpolation is performed independently for each voxel.

Here, (i, j, k) denote the voxel indices in the lateral and axial directions, and l indexes the calibration depths. The PSF is sampled on a discrete voxel grid in (x, y, z), and interpolation is therefore performed only along the depth dimension for each fixed voxel.

The resulting formulation is local in the depth dimension. To interpolate PSF stacks between two control depths d_l and d_{l+1}, only four neighboring calibration stacks, namely those at d_{l-1}, d_l, d_{l+1}, and d_{l+2}, are required. For example, if each calibration stack is represented by a cropped ROI of size $25 \times 25 \times 150$, then interpolation within a single depth interval requires access to

$$4 \cdot 25 \cdot 25 \cdot 150 = 3.75 \times 10^5$$

sampled values, corresponding to approximately 1.5 MB in memory in single-precision floating-point format (4 bytes per value) [34].

In contrast, a full four-dimensional cubic spline representation of the PSF over (x, y, z, d) requires storing $4^4 = 256$ coefficients per 4D voxel, resulting in

$$25 \cdot 25 \cdot 150 \cdot 4 \cdot 256 = 9.6 \times 10^7$$

coefficients, or approximately 384 MB in memory.

The exponential scaling of the number of coefficients with the spline dimensionality for cubic splines can be alleviated when using multivariate B-splines. B-splines represent the function as a linear combination of shifted basis functions, where each basis function is associated with a coefficient [35, 36]. In this case, the number of coefficients scales on the order of the number of sampled voxels.

However, in contrast to Catmull–Rom splines, B-splines do not directly interpolate the sampled data and typically require a fitting procedure to determine the coefficients [37]. Furthermore, while B-splines can be defined on non-uniform grids through the use of non-uniform knot vectors [36], their construction and evaluation become more involved in this setting. This makes their application less straightforward for depth-dependent PSF calibration compared to the Catmull–Rom approach.

Appendix B Cramér-Rao lower bound

For an unbiased estimator and assuming the likelihood of the data satisfies the regularity conditions, the CRLB provides a theoretical lower bound on the variance of the estimator [32, 38].

The CRLB is related to a mathematical measure called the Fisher information, which quantifies the sensitivity of an observable quantity (our noisy SMLM image) to changes in its underlying parameters (e.g., emitter position). The CRLB provides a lower bound on the variance for a scalar parameter θ and is computed as follows [32]:

$$\mathrm{CRLB}(\theta) = \frac{1}{-\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\ln p(\vec{c};\theta)\right]} \tag{B16}$$

where $\ln p(\vec{c};\theta)$ denotes the log-probability density function of observing the data $\vec{c}$, given the parameter θ. If $\hat{\theta}$ is an unbiased estimate of θ, the lower bound of the variance cannot exceed the CRLB:

$$\mathrm{var}(\hat{\theta}) \geq \mathrm{CRLB}(\theta). \tag{B17}$$

In the case a vector of parameters $\vec{\theta} \in \mathbb{R}^p$ is estimated, the lower bound on the variance of each parameter θ_i is given by the corresponding diagonal element of the inverse of the FIM, $I^{-1}(\vec{\theta})$:

$$\mathrm{Var}(\hat{\theta}_i) \geq [I^{-1}(\vec{\theta})]_{ii}. \tag{B18}$$

where the FIM is a $p \times p$ matrix defined as:

$$\left[I(\vec{\theta})\right]_{ij} = -\mathbb{E}\left[\frac{\partial^2 \ln p(\vec{c};\vec{\theta})}{\partial\theta_i\partial\theta_j}\right] \tag{B19}$$

for $i, j = 1, 2, \ldots, p$. In the scalar case (B16), $p = 1$ and therefore $I(\theta) = -\mathbb{E}\left[\frac{\partial^2 \ln p(\vec{c};\theta)}{\partial\theta^2}\right]$ is a 1×1 FIM making it a scalar.

Combining the log of the expression for the likelihood in Equation 7 and the expression for the FIM in (B19), we obtain the following expression for the FIM applied to SMLM:

$$\left[I(\vec{\theta})\right]_{ij} = -\mathbb{E}\left[\sum_k (c_k - \mu_k)^2 \frac{1}{\mu_k^2}\frac{\partial\mu_k}{\partial\theta_i}\frac{\partial\mu_k}{\partial\theta_j}\right]. \tag{B20}$$

In Subsection 2.1 we defined the image formation model assuming a Poisson process of for the arrival of photons at pixels, where μ_k is the expected photon count at the k^{th} pixel and c_k is the observed photon count. For a Poisson distribution, the variance is equal to the mean, such that $\mathbb{E}\left[\sum_k (c_k - \mu_k)^2\right]$ simply becomes μ_k. Substituting this into (B20) leads to the following expression for the FIM:

$$\left[I(\vec{\theta})\right]_{ij} = \sum_k \frac{1}{\mu_k}\frac{\partial\mu_k}{\partial\theta_i}\frac{\partial\mu_k}{\partial\theta_j}. \tag{B21}$$

In Appendix C, we provide the analytical expressions for the partial derivatives of μ_k w.r.t each emitter parameter. Given an image formation model, signal-level, background, and noise models, the CRLB provides the theoretical lower bound on the achievable localization precision. In practice, several works have shown that well-designed localization algorithms based on the MLE approach this bound under ideal conditions [7, 39, 40]. However, this optimal precision can only be achieved when the assumed image formation model accurately describes the measured data. Model mismatches, resulting from depth-dependent aberrations caused by refractive-index mismatches, may therefore lead to localization precision that deviates from the CRLB [41] as the regularity conditions imposed on the likelihood between the model and data are violated.

Appendix C Partial derivatives of μ_k

In Subsection 3, the CRLB is calculated from the image formation model described in Equation 1. The PSF model in the image formation model is given by the cubic spline model described in Equation 12. Substituting the PSF model into expression for the image formation model yields:

$$\mu_k = \theta_N \sum_{m=0}^{3} \sum_{n=0}^{3} \sum_{o=0}^{3} A_{i,j,k,m,n,o} \left(\frac{(x-\theta_x)-t_i}{\Delta t} \right)^m \left(\frac{(y-\theta_y)-t_j}{\Delta t} \right)^n \left(\frac{\theta_z - u_k}{\Delta u} \right)^o + \theta_b. \tag{C22}$$

The partial derivatives of μ_k with respect to each emitter parameter in $\vec{\theta} = \begin{pmatrix} \theta_x & \theta_y & \theta_z & \theta_N & \theta_b \end{pmatrix}$ are as follows [28]:

Partial derivative w.r.t θ_x

$$\frac{\partial \mu_k}{\partial \theta_x} = -\theta_N \sum_{m=0}^{2} \sum_{n=0}^{3} \sum_{o=0}^{3} \frac{(m+1)}{\Delta t} A_{i,j,k,m,n,o} \left(\frac{(x-\theta_x)-t_i}{\Delta t} \right)^m \left(\frac{(y-\theta_y)-t_j}{\Delta t} \right)^n \left(\frac{\theta_z - u_k}{\Delta u} \right)^o \tag{C23}$$

Partial derivative w.r.t θ_y

$$\frac{\partial \mu_k}{\partial \theta_y} = -\theta_N \sum_{m=0}^{3} \sum_{n=0}^{2} \sum_{o=0}^{3} \frac{(n+1)}{\Delta t} A_{i,j,k,m,n,o} \left(\frac{(x-\theta_x)-t_i}{\Delta t} \right)^m \left(\frac{(y-\theta_y)-t_j}{\Delta t} \right)^n \left(\frac{\theta_z - u_k}{\Delta u} \right)^o \tag{C24}$$

Partial derivative w.r.t θ_z

$$\frac{\partial \mu_k}{\partial \theta_z} = \theta_N \sum_{m=0}^{3} \sum_{n=0}^{3} \sum_{o=0}^{2} \frac{(o+1)}{\Delta u} A_{i,j,k,m,n,o} \left(\frac{(x-\theta_x)-t_i}{\Delta t} \right)^m \left(\frac{(y-\theta_y)-t_j}{\Delta t} \right)^n \left(\frac{\theta_z - u_k}{\Delta u} \right)^o \tag{C25}$$

Partial derivative w.r.t θ_N

$$\frac{\partial \mu_k}{\partial \theta_N} = \sum_{m=0}^{3} \sum_{n=0}^{3} \sum_{o=0}^{3} A_{i,j,k,m,n,o} \left(\frac{(x-\theta_x)-t_i}{\Delta t} \right)^m \left(\frac{(y-\theta_y)-t_j}{\Delta t} \right)^n \left(\frac{\theta_z - u_k}{\Delta u} \right)^o \tag{C26}$$

Partial derivative w.r.t θ_b

$$\frac{\partial \mu_k}{\partial \theta_b} = 1. \tag{C27}$$

These partial derivatives represent the terms in the expression for the FIM in Equation B21. The diagonal elements of the inverse of the FIM yields the Cramér-Rao lower bound to the localization precision for each emitter parameter.

Appendix D Evaluation of localization on experimental bead data

In this section, we describe how the localization errors of each cspline model in Section 4 were computed for the experimental bead data. Since the true lateral bead position is not known exactly in the experimental dataset, the lateral and axial errors were evaluated differently.

For the axial direction, the localization error was computed as the difference between the fitted axial position and the known axial stage position of each slice in the bead stack:

$$e_z(z) = \hat{z}(z) - z_{\mathrm{GT}}(z), \tag{D28}$$

where $\hat{z}(z)$ denotes the fitted axial position and $z_{\mathrm{GT}}(z)$ denotes the ground-truth axial position, taken from the stage position of the corresponding slice. Here, z indexes the axial slices in the bead stack.

For the lateral directions, the absolute ground-truth position are not available. Therefore, following Liu *et al.* [30], the fitted lateral positions were centered by subtracting their median value over the full axial stack. The lateral errors were defined as

$$e_x(z) = \hat{x}(z) - \operatorname*{median}_z \big(\hat{x}(z)\big), \qquad e_y(z) = \hat{y}(z) - \operatorname*{median}_z \big(\hat{y}(z)\big). \tag{D29}$$

Here, $\hat{x}(z)$ and $\hat{y}(z)$ are the fitted lateral positions obtained by applying MLE to each image slice in the axial bead stack.

Appendix E Supplementary Figures

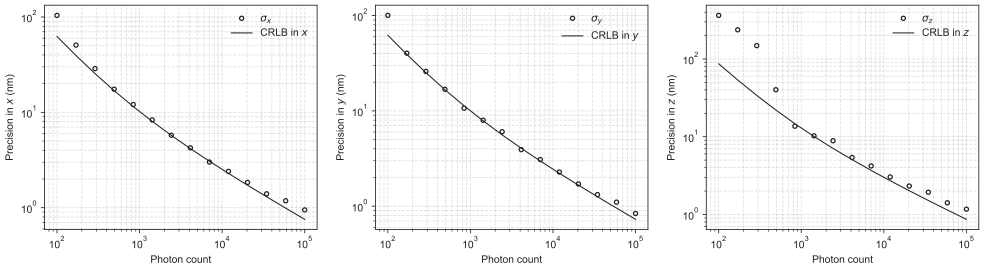


Fig. E2 Localization precision and Cramér-Rao lower bound (CRLB) in fitted x, y and z position for a cspline model calibrated at the coverslip as a function of increasing photon count. The fitted images were simulated with increasing photon counts and a fixed background photon count of 5 photons per pixel. The magnitude of astigmatic aberration was $Z_2^2 = 80$ (rms amplitude). The precision is evaluated across 100 realizations for its fitted results for $z = 0$. These plots show that the cspline model reaches the theoretical limit on the uncertainty for a photon count above ˜4×10^2 in estimating the x-position. When fitting the y-position, the theoretical limit for the uncertainty is reached at a photon count above ˜4×10^2 and for the z-position this is approximately for a photon count of 1×10^3 onwards.

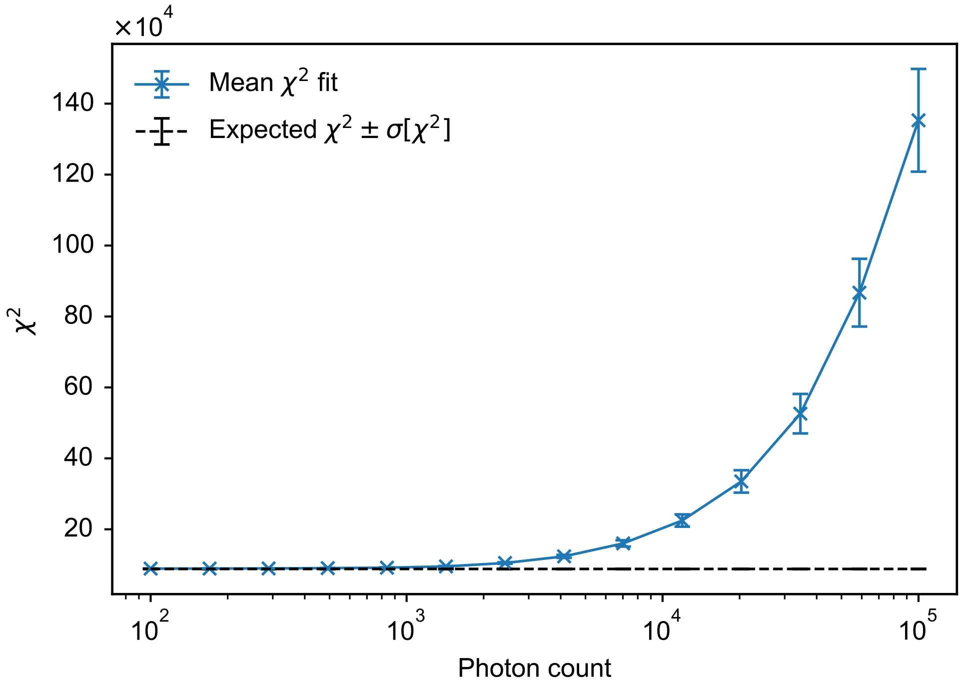


Fig. E3 The χ^2 values of the fit for a cspline model calibrated at the coverslip to simulated images with increasing photon count at the coverslip. 100 realizations of 27×27 pixels single-molecule images were simulated across an axial range of $z \in [-600, 600]$ nm with a step size of 10 nm. The expected χ^2 is equal to the number of statistical independent measures $K = 27 \times 27 \times 121 = 8.8 \times 10^4$. For a photon count $> 10^3$, the mean χ^2 values start to deviate from the expected χ^2 value.

E.1 Interpolation and nearest depth evaluation at 4 μm

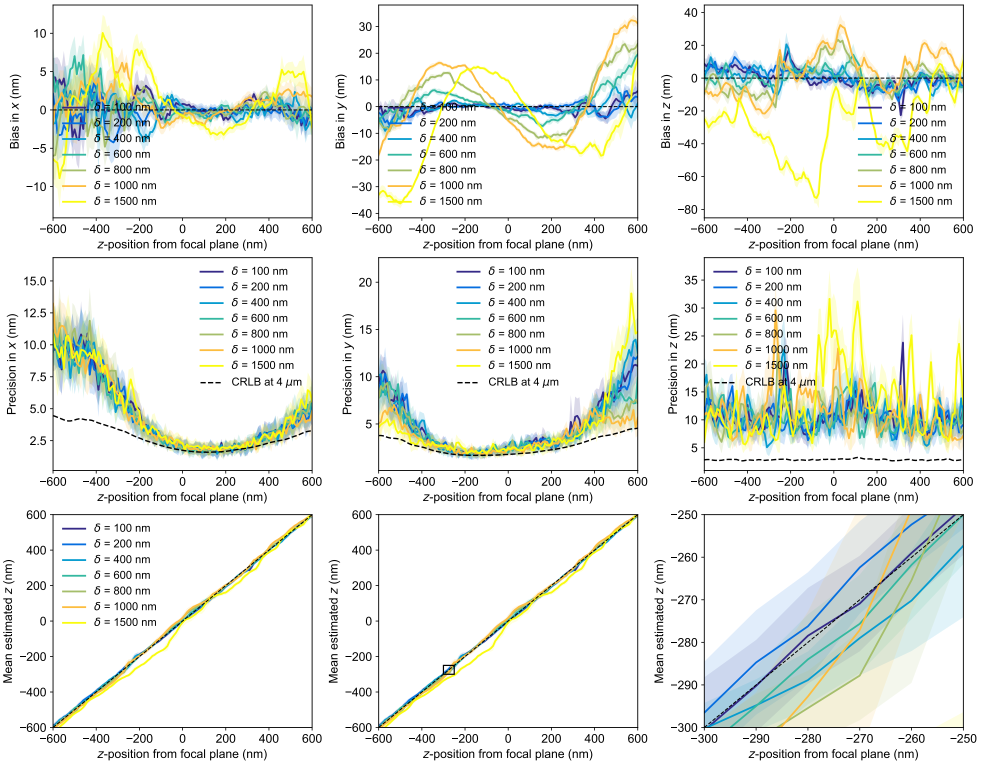


Fig. E4 Effect of interpolation distance δ on localization performance at an evaluation depth of 4 μm (evaluated range $z \in [-600, 600]$ nm). The chosen values are $\delta \in [100, 200, 400, 600, 800, 1000, 1500]$. (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 4 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

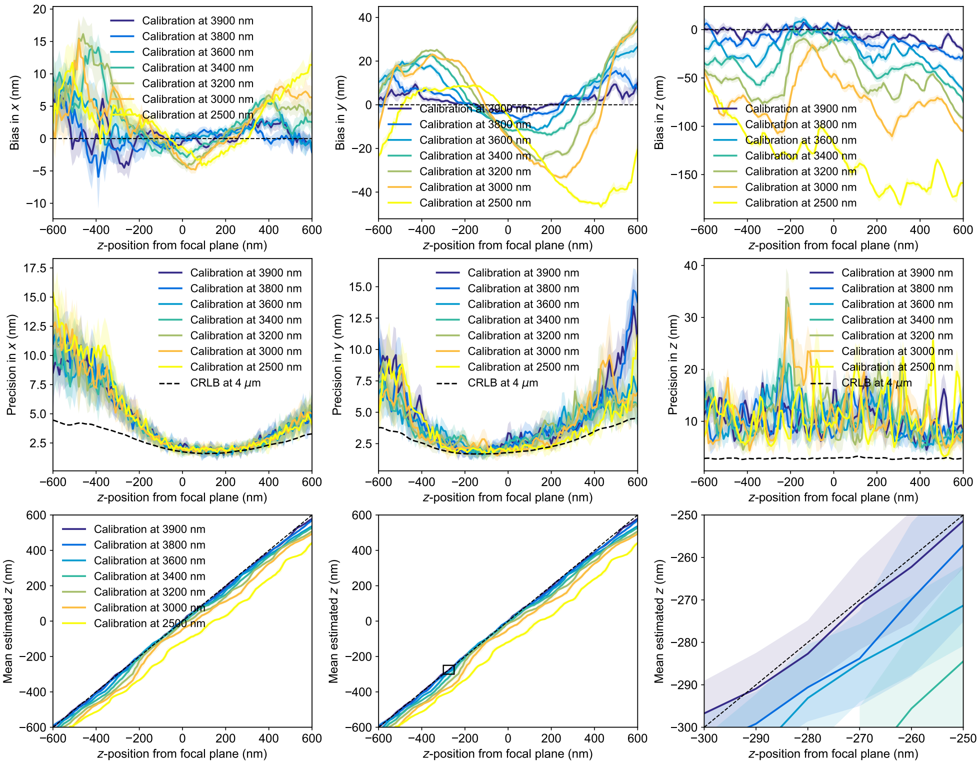


Fig. E5 Localization performance at an evaluation depth of 4 μm for cspline models calibrated at increasing nearest depths from the evaluation depth $d - \delta$ (evaluated range $z \in [-600, 600]$ nm). (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases. (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 4 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

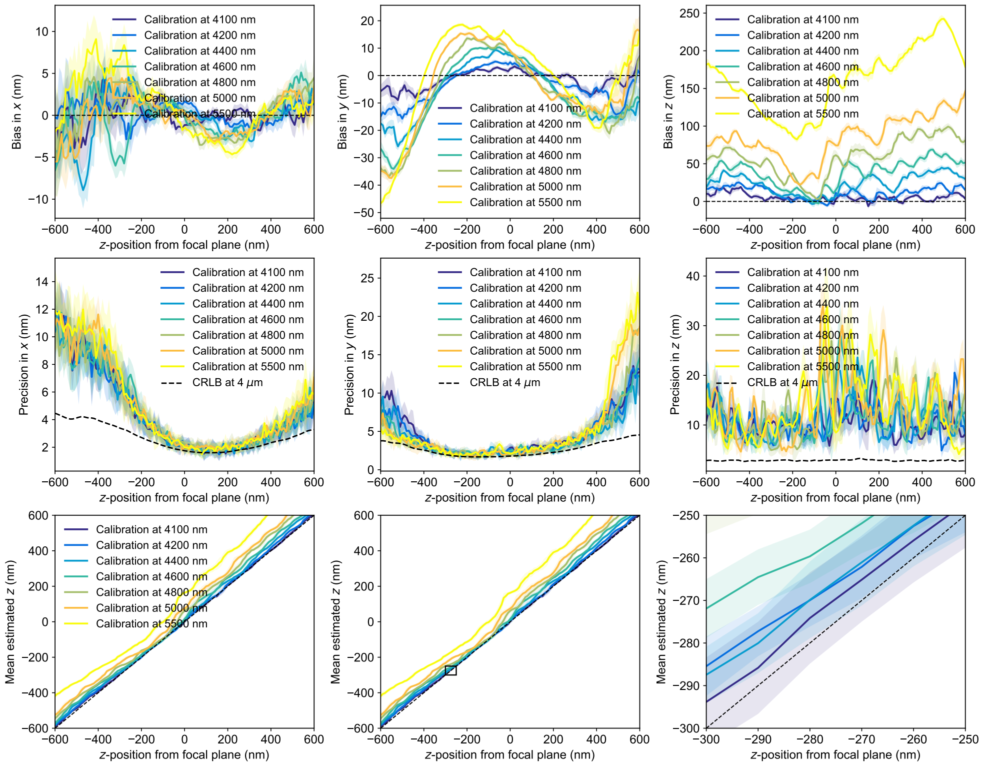


Fig. E6 Localization performance at an evaluation depth of 4 μm for cspline models calibrated at increasing nearest depths from the evaluation depth $d + \delta$ (evaluated range $z \in [-600, 600]$ nm). (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 4 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

Table E1 Optical and imaging parameters used in the simulated experiment at an imaging depth of 4 μm above the coverslip. The parameters are adapted from Li *et al.* [16].

Parameter	Value
Wavelength λ	680 nm
Numerical aperture (NA)	1.43
Refractive index medium n_{med}	1.33
Refractive index coverslip n_{cov}	1.52
Refractive index immersion oil n_{imm}	1.51
Pixel size	100 nm
ROI size	27 × 27 pixels
Emitter photon count	20000 photons
Background count	5 photons/pixel
Focal plane depth in water (FWD)	120000 nm
Pupil sampling (N_{pupil})	52
Imaging depth	4000 nm
Evaluation range in z	$[-600, 600]$ nm

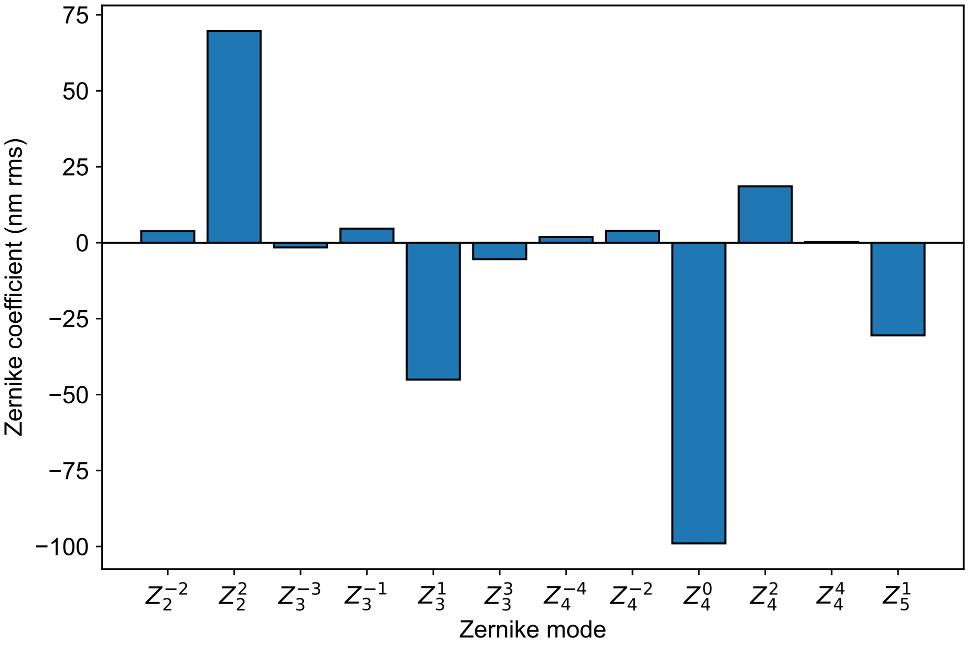


Fig. E7 Zernike aberration coefficients used in the simulated experiment. The magnitudes and modes were chosen based on the fitted Zernike aberrations of a bead from an experimental dataset acquired by the Ries Group at the European Molecular Biology Laboratory (EMBL) also used by Li *et al.* [16]. The coefficients are expressed in nanometers of wavefront optical path difference and, due to the normalization applied in the VectorialPSF implementation by van Velde [24], correspond to rms-normalized Zernike coefficients.

E.2 Interpolation and nearest depth evaluation at 6 μm

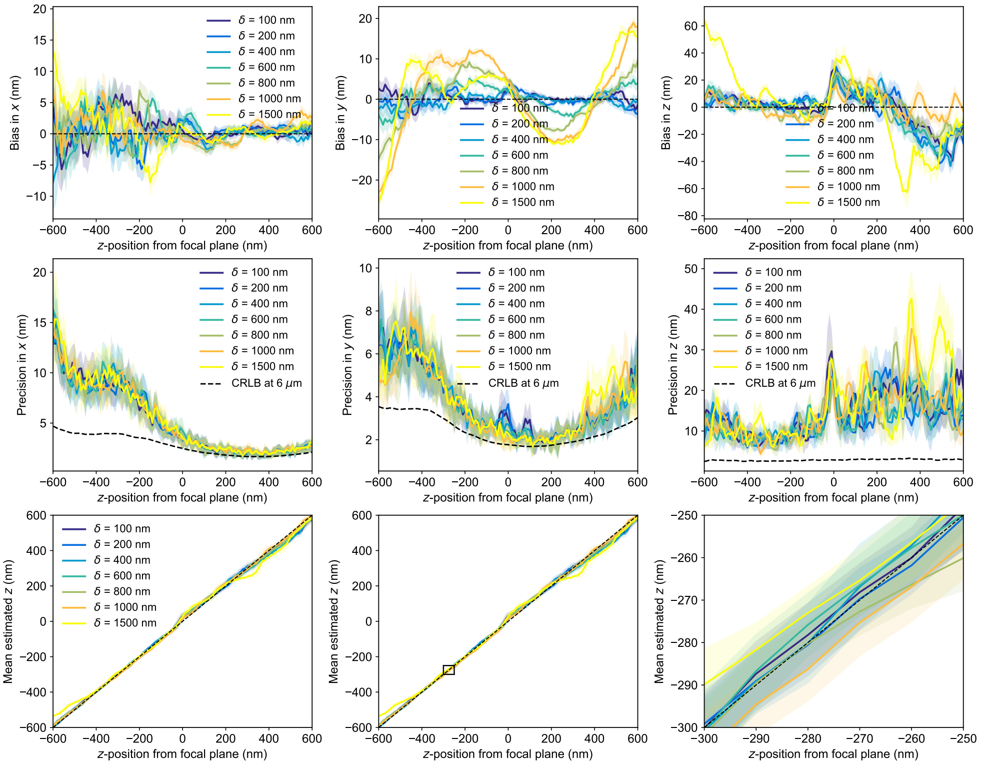


Fig. E8 Effect of interpolation distance δ on localization performance at an evaluation depth of 6 μm (evaluated range $z \in [-600, 600]$ nm). The chosen values are $\delta \in [100, 200, 400, 600, 800, 1000, 1500]$. (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 6 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

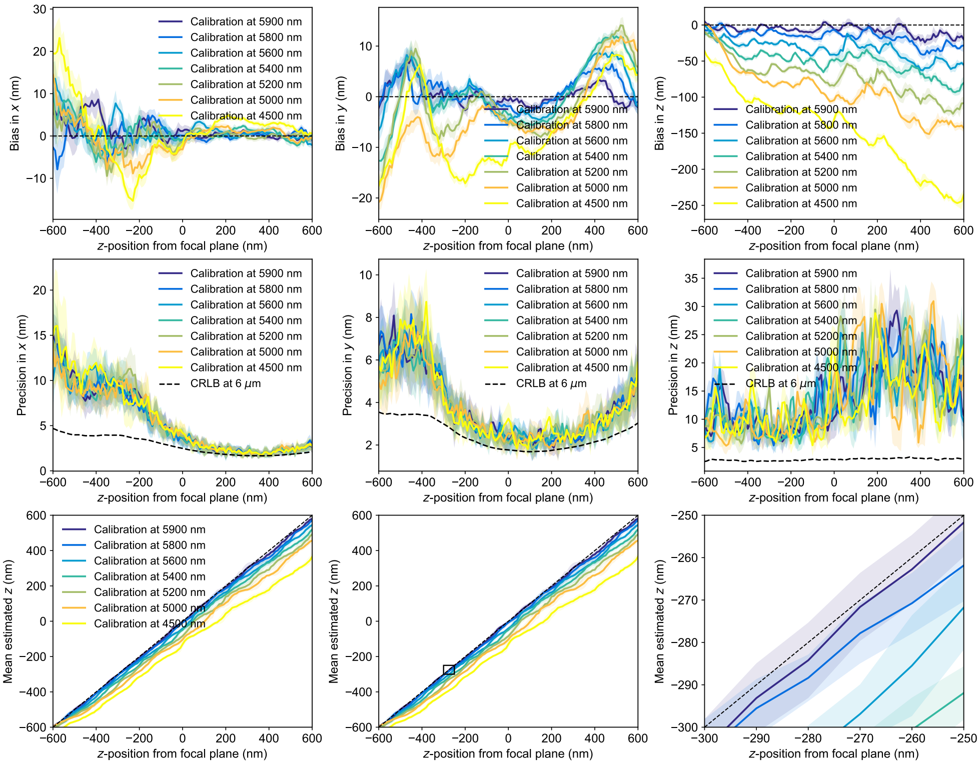


Fig. E9 Localization performance at an evaluation depth of 6 μm for cspline models calibrated at increasing nearest depths from the evaluation depth $d - \delta$ (evaluated range $z \in [-600, 600]$ nm). (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 6 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

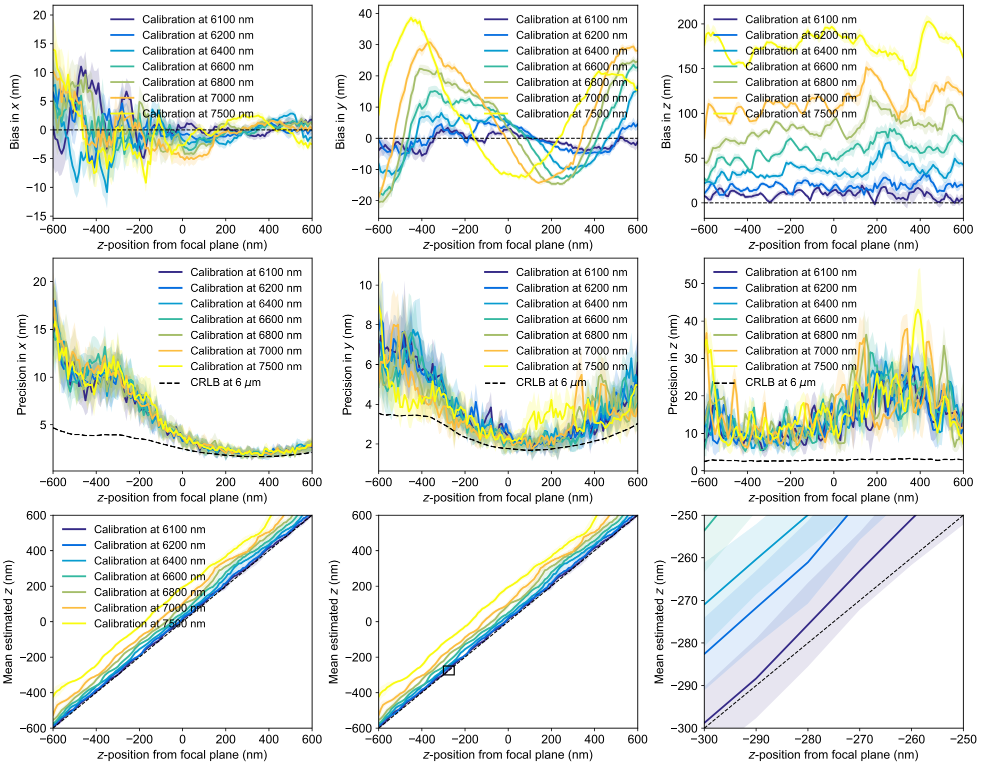


Fig. E10 Localization performance at an evaluation depth of 6 μm for cspline models calibrated at increasing nearest depths from the evaluation depth $d + \delta$ (evaluated range $z \in [-600, 600]$ nm). (Top) Localization bias in x, y, and z for different δ. Shaded areas in the plot indicate the standard deviation of the calculated z biases (Middle) Localization precision in x, y, and z. The dotted black curve represents the CRLB obtained from a cspline model calibrated directly at 6 μm. (Bottom) Mean estimated axial position as a function of the z-position from the focal plane (with zoomed view of the black box). Shaded areas indicate the standard deviation of the calculated mean estimated z-positions.

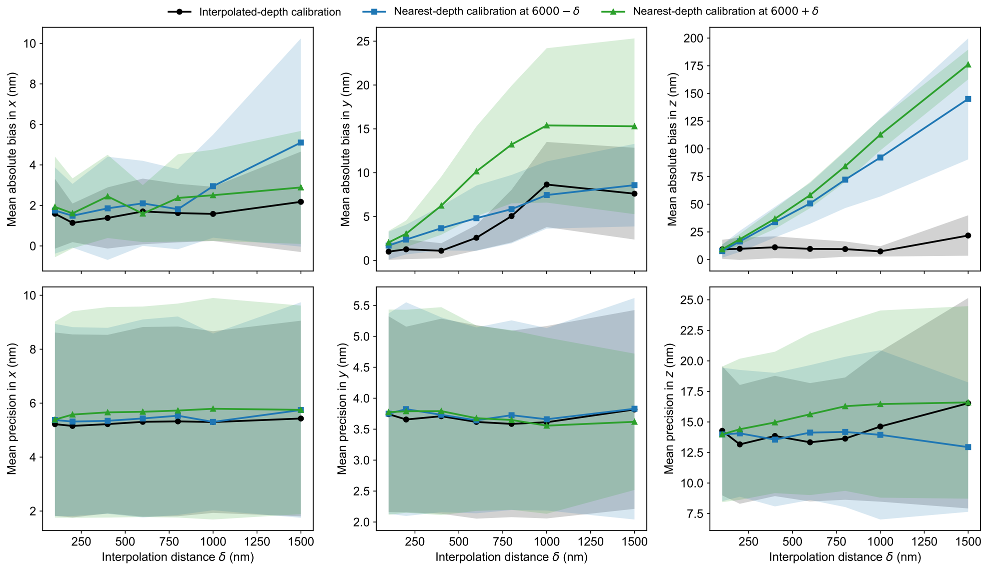


Fig. E11 Mean localization precision and absolute bias in x, y, and z as a function of the calibration depth offset relative to the evaluation depth of 6 μm. Three models are plotted alongside each other, the cspline model calibrated at the interpolated depth with interpolation distance δ, and the cspline models calibrated at the nearest depths at 6 μm $\pm\delta$. Lines denote the mean and shaded areas denote the standard deviation of the bias and precision across the evaluated z-range. When the interpolation distance δ and subsequently the distance for the nearest-depth calibration to the evaluation depth, increases localization precision remains largely unaffected. However, for $\delta \geq 200$, the mean absolute axial bias of the interpolated model is superior to the nearest-depth models by at least 2 fold. Similar to imaging at 4 μm above the coverslip, now at $\delta \geq 1000$ the interpolated approach starts to gradually degrade.

Table E2 Optical and imaging parameters used in the simulated experiment at an imaging depgh of 6 μm.

Parameter	Value
Wavelength λ	680 nm
Numerical aperture (NA)	1.43
Refractive index medium n_{med}	1.33
Refractive index coverslip n_{cov}	1.52
Refractive index immersion oil n_{imm}	1.51
Pixel size	100 nm
ROI size	27 × 27 pixels
Emitter photon count	20000 photons
Background count	5 photons/pixel
Focal plane depth in water (FWD)	120000 nm
Pupil sampling (N_{pupil})	52
Imaging depth	6000 nm
Evaluation range in z	$[-600, 600]$ nm

Appendix F Localization results of the evaluated experimentally acquired bead data

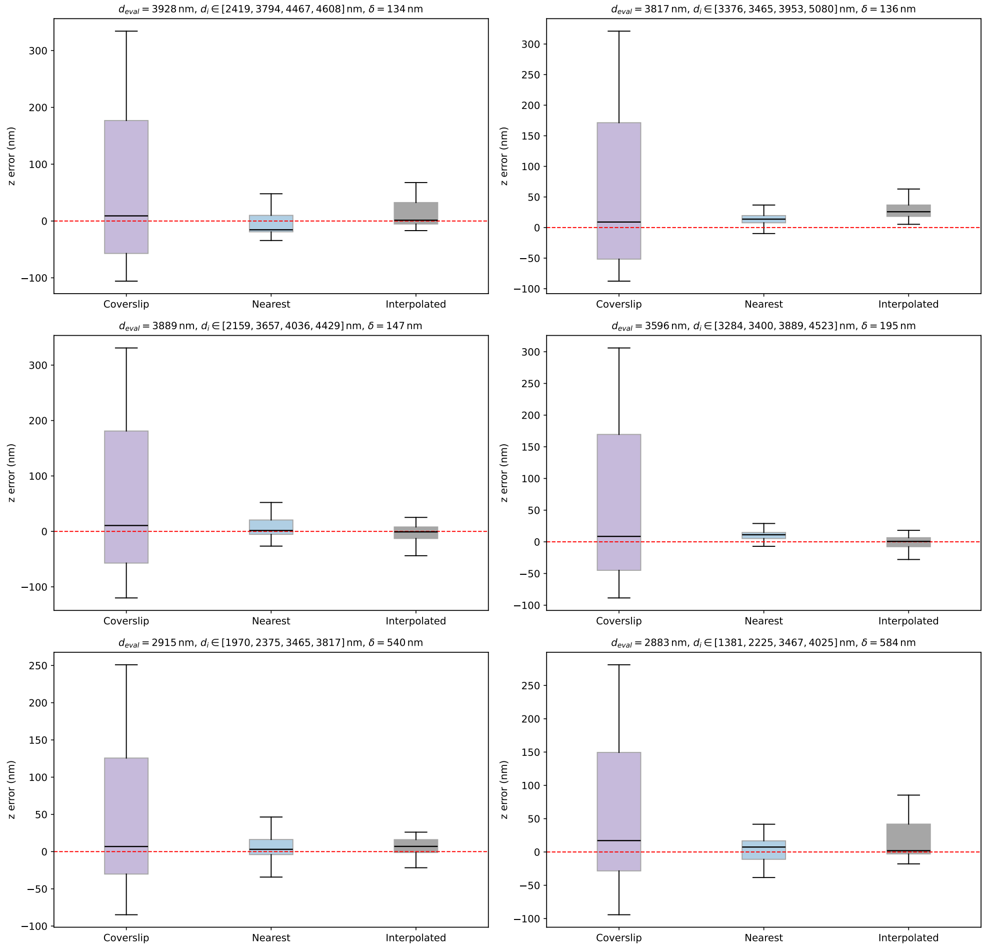


Fig. F12 Boxplots of the axial localization error $z_{\text{fit}} - z_{\text{true}}$ for the first subset of experimentally evaluated bead cases. For each case, the distributions are shown for the coverslip, nearest-depth, and interpolated calibrations. Each boxplot is constructed from the axial localization errors over all fitted axial frames and repeated realizations for each case. The central line indicates the median, the box spans the interquartile range (IQR), and the whiskers indicate the range of the non-outlier data, with outliers omitted. The red dashed horizontal line marks zero error.

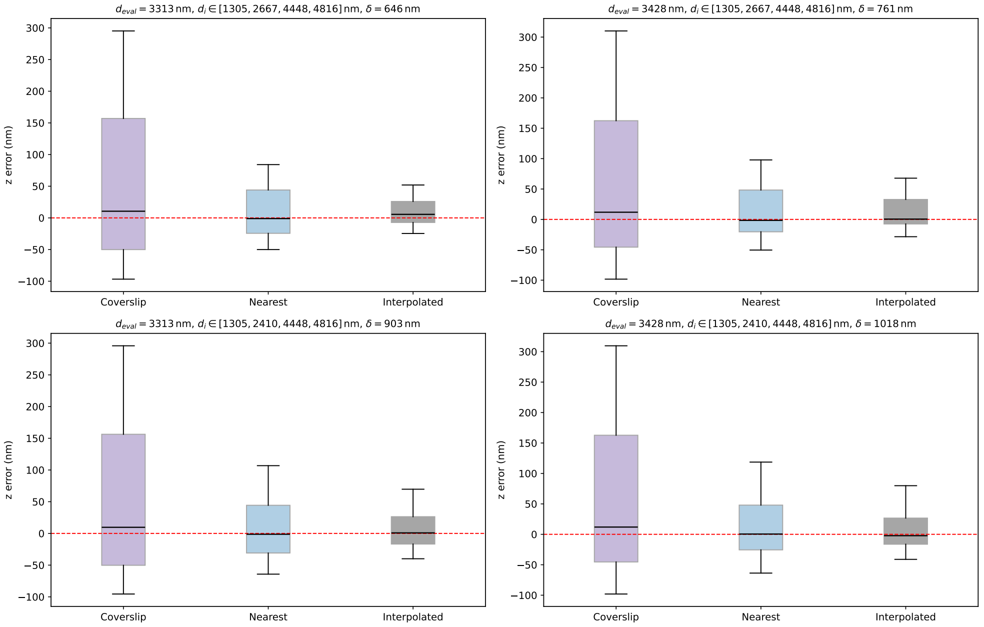


Fig. F13 Boxplots of the axial localization error $z_{\mathrm{fit}} - z_{\mathrm{true}}$ for the remaining experimentally evaluated bead cases. For each case, the distributions are shown for the coverslip, nearest-depth, and interpolated calibrations. Each boxplot is constructed from the pooled axial localization errors over all axial fitted frames and repeated realizations for each case. The central line indicates the median, the box spans the interquartile range (IQR), and the whiskers indicate the range of the non-outlier data, with outliers omitted. The red dashed horizontal line marks zero error.

Appendix G Peak detection algorithm

To detect valid calibration beads in the experimental dataset, we used the local-maximum detection procedure described by Huang *et al.* [31]. Detection was performed on a two-dimensional maximum-intensity projection (MIP) of each experimental bead stack along the axial (z) dimension, after camera offset correction. Since the camera conversion factor from ADU to photons was not available in the metadata, all detection steps were performed on offset-corrected ADU values.

Let I denote the MIP image. First, a bandpass-like filtered image A_1 was computed as the difference between two uniformly smoothed versions of I:

$$A_1 = \text{uniform}\left(I,\, 2\sigma_{\text{PSF}} + 1\right) - \text{uniform}\left(I,\, 2(2\sigma_{\text{PSF}} + 1)\right), \tag{G30}$$

where $\text{uniform}(I, q)$ denotes a uniform filter with a square kernel of size $q \times q$, replacing each pixel value by the average intensity within the corresponding neighbourhood. The filter sizes were chosen based on an estimated lateral PSF width of $\sigma_{\text{PSF}} = 1.5$ pixels.

Next, a second filtering step was applied by computing a maximum filter image A_2 from A_1 as follows:

$$A_2 = \text{maxfilter}(A_1,\, q)\,, \tag{G31}$$

where $\text{maxfilter}(I, q)$ assigns to each pixel the maximum value within a $q \times q$ neighbourhood. The kernel size was chosen as

$$q = \text{odd}(5\sigma_{\text{PSF}})\,, \tag{G32}$$

where $\text{odd}(\cdot)$ denotes rounding to the nearest odd integer, ensuring that the filtering kernel is symmetric and centered around each pixel. For $\sigma_{\text{PSF}} = 1.5$ pixels, this becomes $q = 9$.

A binary local-maximum mask A_3 was then obtained by comparing A_1 and A_2 pixelwise in the following manner:

$$A_3(x, y) = \begin{cases} 1, & A_1(x, y) = A_2(x, y), \\ 0, & \text{otherwise.} \end{cases} \tag{G33}$$

This A_3 step identifies pixels that are local maxima in the filtered image A_1. Candidate bead positions were selected from this mask by additionally requiring:

$$A_1(x, y) > T, \tag{G34}$$

where the detection threshold was chosen adaptively as:

$$T = \mu_{A_1} + 2.5\,\sigma_{A_1}, \tag{G35}$$

with μ_{A_1} and σ_{A_1} denoting the mean and standard deviation of A_1, respectively. This step removes low-intensity local maxima, retaining only peaks that are sufficiently bright to correspond to candidate beads. After this detection step, the candidate peaks were further filtered. Peaks were discarded if a square ROI of size 27 × 27 pixels centered at the peak would intersect the image boundary. Second, peaks were discarded if their center lay within 27 pixels of an already accepted peak, in order to retain only well-isolated beads for calibration.

Appendix H Processed beads

Table H3 Overview of the evaluated experimental cases, including the target evaluation depth, the interpolation depths used for reconstruction, and the corresponding nearest-depth distance δ.

Bead	d_{eval} (nm)	Interpolated depths d_i (nm)	δ (nm)
1	2883	1381, 2225, 3467, 4025	584
2	2915	1970, 2375, 3465, 3817	540
3	3313	1305, 2667, 4448, 4816	646
4	3313	1305, 2410, 4448, 4816	903
5	3428	1305, 2667, 4448, 4816	761
6	3428	1305, 2410, 4448, 4816	1018
7	3596	3284, 3400, 3889, 4523	195
8	3817	3376, 3465, 3953, 5080	136
9	3889	2159, 3657, 4036, 4429	147
10	3928	2419, 3794, 4467, 4608	134

Appendix I Depth assignment

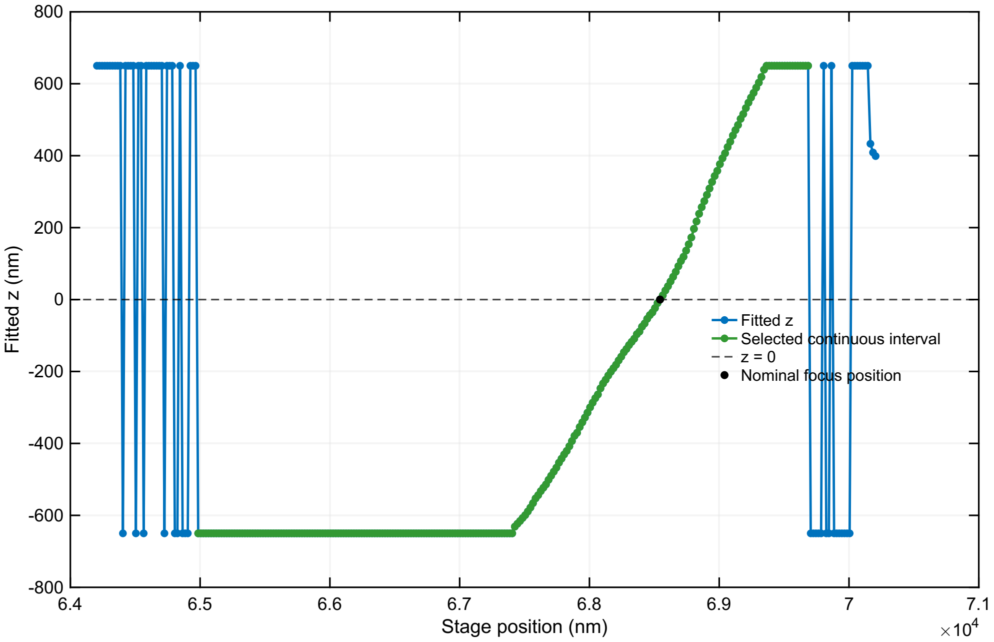


Fig. I14 Depth assignment procedure for an experimental bead. The fitted axial coordinate z_{fit} is plotted as a function of stage position. The bead stack was acquired at axial stage positions ranging approximately from $z \in [64, 70]$ μm. The cspline model used to fit this stack was provided by the experimental dataset. The axial range spans 6 μm, and the calibration spans a range of $z \in [-650, 640]$ nm. As such, this model is mismatched to most of the data it is fitting to. Consequently, mutiple sign changes are present, and large jumps in the fitted z-values are present. This is due to the fitting procedure being a nonlinear optimization scheme with multiple local minima per axial position. Therefore, the longest continuous interval is detected first, before estimating the nominal focal position by interpolation around the zero crossing. For this simulated bead, $z_{\text{fit}} = -9.68$ nm at a stage position of 68525 nm, whereas $z_{\text{fit}} = 0.67$ nm at 68545 nm. Linear interpolation between these two values yields a nominal focal position of $z_0 = 68543$ nm for this bead.

Conclusion

In this chapter, the thesis is summarized and recommendations for future work are provided. The goal of this thesis was to investigate whether interpolation through calibration stacks can yield an accurate calibration for a cspline model at arbitrary imaging depths.

5.1 Summary

In three-dimensional single-molecule localization microscopy (SMLM), the position of a single-molecule emitter relies on fitting a model of the Point Spread Function (PSF) to the single-molecule images. When modeling the PSF with cubic splines, conventional approaches calibrate the model using bead data acquired at the coverslip and assume that this PSF representation remains valid at larger imaging depths. However, refractive index mismatches among the immersion, coverslip, and sample-mounting medium introduce depth-dependent spherical aberrations, causing the PSF shape to vary with imaging depth. As a result, performing localization with a coverslip-calibrated model at greater depths leads to model mismatch, compromising the achieved localization bias and precision.

In this thesis, we investigated whether interpolation between calibration stacks acquired at different depths can be used to obtain accurate cspline models at arbitrary depths. We propose a depth-dependent calibration approach based on Catmull–Rom spline interpolation. By interpolating between calibration stacks and evaluating the interpolation at a desired depth, an interpolated calibration is obtained, which can subsequently be used to calibrate a cspline model. This method yields an accurate PSF model at any depth, between those present in the calibration.

Our simulated results show that a coverslip-calibrated cspline model fails to accurately represent the PSF at large imaging depths. At an imaging depth of 5 μm, this resulted in a mean absolute axial bias of 294 nm. In contrast, the proposed interpolation-based calibration substantially improved localization performance, reducing the axial bias by 99% to 1.8 nm and improving lateral localization precision by more than 60%. Furthermore, the achieved precision closely followed the CRLB, with differences of no more than 2.3 nm, 2.2 nm, and 11.6 nm in x, y, and z, respectively.

We further investigated the effect of the distance between the evaluation depth and the nearest available calibration stack. The results showed that interpolation provides stable localization precision across a wide range of distances, while still significantly reducing axial bias as this distance increases. When a calibration stack is available within approximately 100–200 nm of the evaluation depth, nearest-depth calibration provides comparable or better performance. For larger offsets, interpolation consistently yields lower axial bias.

The proposed method was also validated on experimentally acquired bead data. Interpolation reduced axial bias by up to 80% compared to coverslip calibration while maintaining consistent localization precision. In most evaluated cases, interpolation also outperformed nearest-depth calibration. However, in a small number of cases, nearest-depth calibration yielded slightly lower bias. These deviations from the simulated results were associated with non-uniform and asymmetric distributions of the available calibration depths, indicating that interpolation performance depends not only on the distance to nearby calibration data, but also on the overall distribution of calibration stacks.

These results provide a practical guideline for calibration in 3D SMLM. When a calibration stack is available within approximately 100–200 nm of the evaluation depth, nearest-depth calibration is sufficient. For larger offsets, interpolation should be preferred. For optimal performance, calibration stacks should be acquired densely in depth and distributed as uniformly as possible around the imaging region of interest. Under these conditions, interpolation enables accurate and robust localization across extended imaging depths without requiring additional hardware modifications.

5.2 Limitations and recommendations for future work

The proposed method performs interpolation directly on experimentally acquired calibration stacks. As a result, the interpolated calibration includes not only the PSF shape, but also background fluorescence signal and Poisson photon shot noise. In this work, these effects were minimized by using high signal-to-noise ratio (SNR) calibration data, with photon counts on the order of 10^4 in simulations and 10^5 in the experimental data, such that the PSF structure dominates the observed signal. However, future work could explore the alternative approach by interpolating processed PSF representations rather than raw calibration stacks. By developing a preprocessing pipeline that includes background subtraction, intensity normalization, and smoothening similar to that of Li *et al.* [41], the microscope's PSF can be isolated for interpolation. This may improve the robustness of the method, particularly in the case for lower SNR calibration data.

Furthermore, the current Catmull–Rom formulation provides C^1-continuity, which may limit smoothness compared to higher-order spline formulations. Future work could aim to incorporate C^2-continuity by employing alternative interpolation schemes such as the quintic Catmull–Rom spline [39].

Lastly, while the method was validated on simulated data and experimentally acquired bead stacks, future work could focus on validating the method on SMLM data.

These limitations and recommendations reflect the trade-offs between practical applicability and model accuracy in the proposed approach. While the method provides a hardware-free strategy to address depth-dependent PSF mismatch in 3D single-molecule localization microscopy, several opportunities for improvement remain. Future research should focus on developing interpolation schemes with higher-order continuity, reducing the influence of background and noise through preprocessing of calibration data, and designing an automated calibration pipeline that extracts well-distributed calibration subsets from a deep and densely distributed bead stack. Such developments would further improve the applicability of our proposed depth-dependent calibration method.

Bibliography

[1] E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9:413–468, 12 1873.

[2] Anish V. Abraham, Sripad Ram, Jerry Chao, E. S. Ward, and Raimund J. Ober. Quantitative study of single molecule location estimation techniques. *Optics Express*, 17:23352, 12 2009.

[3] David A. Agard, Yasushi Hiraoka, Peter Shaw, and John W. Sedat. Chapter 13 fluorescence microscopy in three dimensions. In D. Lansing Taylor and Yu-Li Wang, editors, *Fluorescence Microscopy of Living Cells in Culture Part B. Quantitative Fluorescence Microscopy—Imaging and Spectroscopy*, volume 30 of *Methods in Cell Biology*, pages 353–377. Academic Press, 1989.

[4] Hazen P. Babcock and Xiaowei Zhuang. Analyzing single molecule localization microscopy data using cubic splines. *Scientific Reports*, 7, 12 2017.

[5] David Baddeley, Mark B. Cannell, and Christian Soeller. Three-dimensional sub-100 nm super-resolution imaging of biological samples using a phase ramp in the objective pupil. *Nano Research*, 4:589–598, 6 2011.

[6] 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*, 313:1642–1645, 9 2006.

[7] Julie S Biteen, Michael A Thompson, Nicole K Tselentis, Grant R Bowman, Lucy Shapiro, and W E Moerner. Super-resolution imaging in live caulobacter crescentus cells using photoswitchable eyfp. *Nature Methods*, 5:947–949, 11 2008.

[8] Charles Bond, Adriana N. Santiago-Ruiz, Qing Tang, and Melike Lakadamyali. Technological advances in super-resolution microscopy to study cellular processes. *Molecular Cell*, 82(2):315–332, 2022. Focus on technology.

[9] M. J. Booth, M. A.A. Neil, and T. Wilson. Aberration correction for confocal imaging in refractive-index- mismatched media. *Journal of Microscopy*, 192:90–98, 1998.

[10] Martin Booth, Débora Andrade, Daniel Burke, Brian Patton, and Mantas Zurauskas. Aberrations and adaptive optics in super-resolution microscopy. *Microscopy*, 64:251–261, 8 2015.

[11] Benjamin P. Bratton, Joshua W. Shaevitz, and Thomas Abraham. Simple experimental methods for determining the apparent focal shift in a microscope system. *PLoS ONE*, 10, 8 2015.

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

[13] John J Cargille. Immersion oil and the microscope. Technical report.

[14] Ching-Ya Cheng and Chia-Lung Hsieh. Background estimation and correction for high-precision localization microscopy. *ACS Photonics*, 4:1730–1739, 7 2017.

[15] Anindita Dasgupta, Joran Deschamps, Ulf Matti, Uwe Hübner, Jan Becker, Sebastian Strauss, Ralf Jungmann, Rainer Heintzmann, and Jonas Ries. Direct supercritical angle localization microscopy for nanometer 3d superresolution. *Nature Communications*, 12:1180, 2 2021.

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

[17] Erin E. Diel, Jeff W. Lichtman, and Douglas S. Richardson. Tutorial: avoiding and correcting sample-induced spherical aberration artifacts in 3d fluorescence microscopy. *Nature Protocols*, 15:2773–2784, 9 2020.

[18] Elisa Dultz, Matthias Wojtynek, Ohad Medalia, and Evgeny Onischenko. The nuclear pore complex: Birth, life, and death of a cellular behemoth. *Cells*, 11:1456, 4 2022.

[19] Johann Engelhardt, Jan Keller, Patrick Hoyer, Matthias Reuss, Thorsten Staudt, and Stefan W. Hell. Molecular orientation affects localization accuracy in superresolution far-field fluorescence microscopy. *Nano Letters*, 11(1):209–213, 2011. PMID: 21133355.

[20] Gerald Farin. *Curves and Surfaces for Computer-Aided Geometric Design.* Morgan Kaufmann, 2002.

[21] Robert S. Fischer, Yicong Wu, Pakorn Kanchanawong, Hari Shroff, and Clare M. Waterman. Microscopy in 3d: a biologist's toolbox. *Trends in Cell Biology*, 21(12):682–691, 2011.

[22] Sarah Frisken Gibson and Frederick Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *Journal of the Optical Society of America A*, 8:1601, 10 1991.

[23] Joseph W. Goodman. *Introduction to Fourier Optics.* McGraw-Hill, 1996, 2nd edition, 2005.

[24] B. M. HANSER, M. G. L. GUSTAFSSON, D. A. AGARD, and J. W. SEDAT. Phase-retrieved pupil functions in wide-field fluorescence microscopy. *Journal of Microscopy*, 216:32–48, 10 2004.

[25] Eugene. Hecht. *Optics*. Pearson, 2017.

[26] Mike Heilemann, Sebastian vandeLinde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie International Edition*, 47:6172–6176, 8 2008.

[27] Richard G. Held, Jiahao Liang, and Axel T. Brunger. Nanoscale architecture of synaptic vesicles and scaffolding complexes revealed by cryo-electron tomography. *Proceedings of the National Academy of Sciences*, 121, 7 2024.

[28] Samuel T Hess, Thanu P K Girirajan, and Michael D Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91:4258–72, 12 2006.

[29] Laurent Holtzer, Tobias Meckel, and Thomas Schmidt. Nanometric three-dimensional tracking of individual quantum dots in cells. *Applied Physics Letters*, 90, 1 2007.

[30] Bo Huang, Sara A. Jones, Boerries Brandenburg, and Xiaowei Zhuang. Whole-cell 3d storm reveals interactions between cellular structures with nanometer-scale resolution. *Nature Methods*, 5:1047–1052, 2008.

[31] Bo Huang, Wenqin Wang, Mark Bates, and Xiaowei Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319:810–813, 2 2008.

[32] Fang Huang, Tobias M P Hartwich, Felix E Rivera-Molina, Yu Lin, Whitney C Duim, Jane J Long, Pradeep D Uchil, Jordan R Myers, Michelle A Baird, Walther Mothes, Michael W Davidson, Derek Toomre, and Joerg Bewersdorf. Video-rate nanoscopy using scmos camera–specific single-molecule localization algorithms. *Nature Methods*, 10:653–658, 7 2013.

[33] John F. Hughes, Andries van Dam, Morgan McGuire, David F. Sklar, James D. Foley, Steven K. Feiner, and Kurt Akeley. *Chapters 22 & 23: Splines and Subdivision Curves/Surfaces*, pages 595–614. Addison-Wesley, 3rd edition, 2014.

[34] Hellen C. Ishikawa-Ankerhold, Richard Ankerhold, and Gregor P.C. Drummen. Advanced fluorescence microscopy techniques-frap, flip, flap, fret and flim, 4 2012.

[35] S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection theory*. Fundamentals of Statistical Signal Processing. Prentice-Hall PTR, 1998.

[36] Ismail M. Khater, Ivan Robert Nabi, and Ghassan Hamarneh. A Review of Super-Resolution Single-Molecule Localization Microscopy Cluster Analysis and Quantification Methods. *Patterns*, 1(3), June 2020. Publisher: Elsevier.

[37] Ted A. Laurence and Brett A. Chromy. Efficient maximum likelihood estimator fitting of histograms, 5 2010.

[38] Mickaël Lelek, Melina T. Gyparaki, Gerti Beliu, Florian Schueder, Juliette Griffié, Suliana Manley, Ralf Jungmann, Markus Sauer, Melike Lakadamyali, and Christophe Zimmer. Single-molecule localization microscopy, 12 2021.

[39] Juncheng Li and Chengzhi Liu. A quintic polynomial spline with local shape parameters unifying approximation and interpolation. *Mathematics and Computers in Simulation*, 241:582–590, 2026.

[40] Yiming Li, Markus Mund, Philipp Hoess, Joran Deschamps, Ulf Matti, Bianca Nijmeijer, Vilma Jimenez Sabinina, Jan Ellenberg, Ingmar Schoen, and Jonas Ries. Real-time 3d single-molecule localization using experimental point spread functions. *Nature Methods*, 15:367–369, 4 2018.

[41] Yiming Li, Yu-Le Wu, Philipp Hoess, Markus Mund, and Jonas Ries. Depth-dependent psf calibration and aberration correction for 3d single-molecule localization. *Biomedical Optics Express*, 10:2708, 6 2019.

[42] Jeff W. Lichtman and José Angel Conchello. Fluorescence microscopy, 12 2005.

[43] Yu Lin, Jane J. Long, Fang Huang, Whitney C. Duim, Stefanie Kirschbaum, Yongdeng Zhang, Lena K. Schroeder, Aleksander A. Rebane, Mary Grace M. Velasco, Alejandro Virrueta, Daniel W. Moonan, Junyi Jiao, Sandy Y. Hernandez, Yongli Zhang, and Joerg Bewersdorf. Quantifying and optimizing single-molecule switching nanoscopy at high speeds. *PLoS ONE*, 10, 5 2015.

[44] Sheng Liu, Emil B. Kromann, Wesley D. Krueger, Joerg Bewersdorf, and Keith A. Lidke. Three dimensional single molecule localization using a phase retrieved pupil function. *Optics Express*, 21:29462, 12 2013.

[45] R. Loudon. *The Quantum Theory of Light*. OUP Oxford, 2000.

[46] Michael J. Mlodzianoski, Paul J. Cheng-Hathaway, Shane M. Bemiller, Tyler J. McCray, Sheng Liu, David A. Miller, Bruce T. Lamb, Gary E. Landreth, and Fang Huang. Active psf shaping and adaptive optics enable volumetric localization microscopy through brain sections. *Nature Methods*, 15:583–586, 8 2018.

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

[48] Robert P J Nieuwenhuizen, Keith A Lidke, Mark Bates, Daniela Leyton Puig, David Grünwald, Sjoerd Stallinga, and Bernd Rieger. Measuring image resolution in optical nanoscopy. *Nature Methods*, 10:557–562, 6 2013.

[49] Goro Nishimura and Mamoru Tamura. Artefacts in the analysis of temporal response functions measured by photon counting. *Physics in Medicine and Biology*, 50:1327–1342, 3 2005.

[50] Robert J. Noll. Zernike polynomials and atmospheric turbulence*. *Journal of the Optical Society of America*, 66:207, 3 1976.

[51] Petar N. Petrov and W. E. Moerner. Addressing systematic errors in axial distance measurements in single-emitter localization microscopy. *Optics Express*, 28:18616, 6 2020.

[52] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C (2nd ed.): the art of scientific computing.* Cambridge University Press, USA, 1992.

[53] Richard F. Riesenfeld. The development of b-splines for cad. *IEEE Computer Graphics and Applications*, 42:90–100, 3 2022.

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

[55] P. Sarder and A. Nehorai. Deconvolution methods for 3-d fluorescence microscopy images. *IEEE Signal Processing Magazine*, 23(3):32–45, 2006.

[56] José Sasián. *Introduction to Aberrations in Optical Imaging Systems.* Cambridge University Press, 12 2012.

[57] Lothar Schermelleh, Alexia Ferrand, Thomas Huser, Christian Eggeling, Markus Sauer, Oliver Biehlmaier, and Gregor P. C. Drummen. Super-resolution microscopy demystified. *Nature Cell Biology*, 21:72–84, 1 2019.

[58] Joerg Schnitzbauer, Maximilian T Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. Super-resolution microscopy with dna-paint. *Nature Protocols*, 12:1198–1228, 6 2017.

[59] Alexey Sharonov and Robin M. Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences*, 103:18911–18916, 12 2006.

[60] Peter Shirley and Steve Marschner. *Chapter 15: Curves.* CRC Press, Taylor & Francis Group, 4th edition, 2016.

[61] M. Siemons, C. N. Hulleman, R. Ø. Thorsen, C. S. Smith, and S. Stallinga. High precision wavefront control in point spread function engineering for single emitter localization. *Optics Express*, 26:8397, 4 2018.

[62] Alex Small and Shane Stahlheber. Fluorophore localization algorithms for super-resolution microscopy, 2014.

[63] Carlas Smith. *Optimum Single Molecule Localization Microscopy From Single Molecule Detections to Nanoscopic Observations.* PhD thesis, Delft University of Technology, 2016.

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

[65] Sjoerd Stallinga. Effect of rotational diffusion in an orientational potential well on the point spread function of electric dipole emitters. *Journal of the Optical Society of America A*, 32:213, 2 2015.

[66] Sjoerd Stallinga and Bernd Rieger. Accuracy of the gaussian point spread function model in 2d localization microscopy. *Optics Express*, 18:24461, 11 2010.

[67] George Gabriel Stokes. On the change of refrangibility of light. *Abstracts of the Papers Communicated to the Royal Society of London*, 6, 12 1854.

[68] Kurt Thorn. A quick guide to light microscopy in cell biology. *Molecular Biology of the Cell*, 27:219–222, 1 2016.

[69] P. Török, P. D. Higdon, and T. Wilson. Theory for confocal and conventional microscopes imaging small dielectric scatterers. *Journal of Modern Optics*, 45:1681–1698, 8 1998.

[70] Ton van den Boom and Bart De Schutter. Optimization for systems and control, September 2022. Lecture notes for course SC42056.

[71] Adriaan van den Bos. *Parameter Estimation for Scientists and Engineers.* Wiley, 6 2007.

[72] Pieter van Velde, Bernd Rieger, Taylor Hindsdale, Jelmer Cnossen, Daniel Fan, Shih-Te Hung, David Grunwald, and Carlas Smith. Zimflux: Single molecule localization microscopy with patterned illumination in 3d. *Optics Express*, 31:42701, 12 2023.

[73] Michel Verhaegen, Paulo Pozzi, Oleg Soloviev, Gleb Vdovin, and Dean Wilding. Basic optical concepts. In Oleg Soloviev, editor, *Control for High Resolution Imaging*, chapter 2, pages 19–40. Delft University of Technology, Delft Center for Systems and Control, Delft, The Netherlands, 2024. Lecture notes for the course SC42030.

[74] C. C. De Visser. *Global Nonlinear Model Identification with Multivariate Splines: Application to Aerodynamic Model Identification of the Cessna Citation II.* PhD thesis, Delft University of Technology, Delft, Netherlands, 2011. OCLC: 840446321.

[75] Lexy von Diezmann, Yoav Shechtman, and W. E. Moerner. Three-dimensional localization of single molecules for super-resolution imaging and single-particle tracking, 6 2017.

[76] T. Wilson. *Confocal Microscopy*, pages 219–232. Springer US, 1994.

[77] Yicong Wu, Ryan Christensen, Daniel Colón-Ramos, and Hari Shroff. Advanced optical imaging techniques for neurodevelopment. *Current Opinion in Neurobiology*, 23(6):1090–1097, 2013. Development of neurons and glia.

[78] J. C. Wyant and K. Creath. Basic wavefront aberration theory for optical metrology. In R. R. Shannon and J. C. Wyant, editors, *Applied Optics and Optical Engineering*, volume 11, pages 2–53. Academic Press, New York, 1992.

[79] Fan Xu, Donghan Ma, Kathryn P. MacPherson, Sheng Liu, Ye Bu, Yu Wang, Yu Tang, Cheng Bi, Tim Kwok, Alexander A. Chubykin, Peng Yin, Sarah Calve, Gary E. Landreth, and Fang Huang. Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval. *Nature Methods*, 17:531–540, 5 2020.

[80] Sofia V. Zaichick, Kevin P. Bohannon, and Gregory A. Smith. Alphaherpesviruses and the cytoskeleton in neuronal infections. *Viruses*, 3:941–981, 6 2011.

[81] Ping Zhu, Jun Liu, Julian Bess, Elena Chertova, Jeffrey D. Lifson, Henry Grisé, Gilad A. Ofek, Kenneth A. Taylor, and Kenneth H. Roux. Distribution and three-dimensional structure of aids virus envelope spikes. *Nature*, 441:847–852, 6 2006.

Glossary

List of Acronyms

EM-CCD	electron-multiplying charge-coupled device
sCOMS	scientific complementary metal-oxide semiconductor
LM	Levenberg-Marquardt
CRLB	Cramèr-Rao lower bound
PSF	Point Spread Function
OTF	optical transfer function
AO	Adaptive Optics
MLE	maximum likelihood estimation
ROI	region of interest
cspline	cubic spline
NA	numerical aperture
SMLM	single-molecule localization microscopy
PALM	photoactivated localization microscopy
STORM	stochastic optical reconstruction microscopy
dSTORM	direct stochastic optical reconstruction microscopy
PAINT	point accumulation in nanoscale topography
FIM	Fischer information matrix
SNR	signal-to-noise ratio