Dissertation

submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany

for the degree of
Doctor of Natural Sciences

Put forward by
Henrik von der Emde

born in: Marburg (Lahn), Germany
Date of the oral examination: 26 November 2024

Molecule Localization with STED

Referees: Prof. Dr. Dr. h.c. mult. Stefan W. Hell
Prof. Dr. Joachim P. Spatz

Für meine Familie

Abstract

I present a re-implementation of the super-resolution method MINSTED, which enhances the stimulated emission interaction probability of the STED (stimulated emission depletion) process and therefore reduces the optical power required by approximately one order of magnitude. This results in an enhanced localization performance, achieving uncertainties in single-molecule localization below 1 nm. This molecular-scale localization precision is applied to biological specimens upon the use of the DNA-PAINT labelling scheme. The imaging of the nuclear pore complex in fixed HeLa cells is demonstrated and the colocalization of synaptic vesicle proteins in a multi-colour experiment in fixed rat hippocampal neurons was investigated. The MINSTED localization approach is additionally shown to be used for single molecule tracking at nanometre-millisecond spatio-temporal precision. The biological relevance is highlighted by tracking of the motor protein kinesin-1, clearly resolving its 16 nm steps with a temporal precision of < 2 ms.

Zusammenfassung

Ich präsentiere eine Weiterentwicklung der hochauflösenden Mikroskopiemethode MINSTED, die die Wechselwirkungswahrscheinlichkeit der stimulierten Emission beim STED-Prozess (stimulated emission depletion) erhöht und daher die erforderliche Lichtleistung um etwa eine Größenordnung reduziert. Dies führt zu einer verbesserten Qualität der Lokalisationen, wobei Unsicherheiten bei der Einzelmoleküllokalisierung von weniger als 1 nm erreicht werden. Die molekulare Präzision wird unter der Verwendung der Markierungsmethode DNA-PAINT an biologischen Proben reproduziert. In Lokalisationsmessungen des Kernporenkomplexes in fixierten HeLa-Zellen und der Untersuchung von Kolokalisation synaptischer Vesikelproteine in einem Mehrfarbenexperiment in fixierten Hippocampus-Neuronen der Ratte wird die biologische Anwendbarkeit nachgewiesen. Darüber hinaus wird gezeigt, dass der MINSTED-Lokalisierungsansatz für die Verfolgung einzelner Moleküle mit einer räumlich-zeitlichen Präzision im Nanometer-Millisekunden Bereich verwendet werden kann. Die biologische Relevanz wird durch die Verfolgung des Motorproteins Kinesin-1 verdeutlicht, dessen 16 nm Schritte mit einer zeitlichen Präzision von < 2 ms klar aufgelöst werden.

CONTENTS

1 Preface **1**

2 From Light to Microscopy **5**
2.1 Classical imaging systems . . . 5
2.1.1 Geometric transformations of an optimal lens . . . 5
2.1.2 Fourier optics . . . 7
2.1.3 Spatial frequency loss in optical systems . . . 9
2.1.4 The 4-f system . . . 10
2.1.5 Extensions . . . 12
2.2 Fluorescence . . . 13
2.3 Diffraction limited microscopy . . . 15
2.3.1 Widefield . . . 15
2.3.2 Confocal microscopy . . . 16
2.3.3 Structured illumination microscopy . . . 18
2.4 Sub-diffraction microscopy . . . 19
2.4.1 Imaging methods . . . 20
2.4.2 Localization methods . . . 23
2.5 Summary . . . 28

3 Localizing Molecules with STED **29**
3.1 Introduction . . . 29
3.2 Theory . . . 31
3.2.1 STED . . . 32
3.2.2 MINSTED . . . 34
3.3 Implementation . . . 38
3.3.1 Optics and electronics . . . 38
3.3.2 Measurement control software . . . 43
3.4 Results and discussion . . . 45
3.4.1 Imaging . . . 46
3.4.2 Tracking . . . 57

4 Conclusion **65**
4.1 Imaging . . . 65
4.2 Tracking . . . 66

A Additional calculations **69**
A.1 Cumulative statistics 69
A.2 Cramér-Rao bound 70

B Technical specificities **73**
B.1 Components blue-shifted MINSTED 73
B.2 Re-designed fibre amplifier 76

C MINSTED measurement parameters **77**

D Data analysis **79**
D.1 Filtering 79
D.1.1 MINSTED imaging 79
D.1.2 MINSTED tracking 79
D.2 Cluster analysis 80

Bibliography **81**

Acknowledgement **85**

CHAPTER
ONE

PREFACE

The first microscopes were already built in the 17$^{\text{th}}$ century [23]. Since then, magnification, resolution, aberration correction and field flatness were steadily enhanced. This was achieved by optimizing the manufacturing process of lenses, combining glasses for dispersion-compensation and finding lens arrangements to minimize distortions. In 1873, the physicist Ernst Abbe, who at that time was working on the optimization of objective lenses together with the entrepreneur Carl Zeiß, formulated the optical resolution limit [1]. No matter how perfect the optical arrangement would be, the resolving power of a microscope would not substantially overcome the minimal distance of

$$d = \frac{\lambda}{2n \sin(\alpha)} \ . \tag{1.1}$$

The optical wavelength λ is herein the defining length measure. It is reduced by the term $2n \sin(\alpha)$, where α is the half opening angle of the light cone collected by the objective lens and n the refractive index of the medium connecting the sample and the objective lens. Even if the opening angle is maximised to $\alpha = 90°$, this limit only allows an improvement in resolution by a factor of $2n$ of the wavelength. As the refractive index is a parameter on the order 1, and not highly scalable, the resolution in optical microscopy remained limited to about 200 nm.

One solution to this problem was the use of shorter wavelengths. In this context, the electron microscope (EM) was developed in the 20$^{\text{th}}$ century [29]. Electrons can be scanned and focussed by static magnetic fields and have short effective wavelengths when accelerated to high speeds, close to the speed of light. Today's EMs yield a resolution of less than 0.1 nm [7]. However, EMs operate in vacuum, often at low temperatures. This requires complicated sample preparation and the technique does not allow the observation of living specimens.

Returning to light microscopy, in 1994, Stefan W. Hell introduced the concept of stimulated emission depletion (STED) microscopy [11], which was proposed to overcome the resolution barrier in far-field optical microscopy. This concept tackles the problem at its root: if neighbouring emitters can not be spatially separated in a microscope, it should be avoided to detect them at the same time. This concept was implemented in fluorescence microscopy, by utilizing the naturally occurring ground and excited states of the fluorescence emitters: OFF and ON. The dark ground state is interpreted as an OFF state, whereas the bright fluorescent excited state is considered as an ON state. By exerting

optical control over not only the OFF to ON transition but also the reverse process, it was possible to effectively separate emitter states spatially on sub-diffraction scales. Using an optical beam driving the OFF transition that features an intensity zero, any emitter except for the one laying exactly at the zero will be switched OFF, when assuming infinite interaction probability. Conceptually, detecting only the emitters remaining ON near the zero, enables unlimited resolution in fluorescence microscopy.

In the late 1990s, STED microscopy was practically demonstrated for the first time [16]. Thereafter, in the beginning of the 2000s, the methods PALM (photo-activated localization microscopy) [3] and STORM (stochastic optical reconstruction microscopy) [28] also achieved sub-diffraction resolution by utilizing the ON/OFF state distinction in order to localize isolated emitters in the ON state. In contrast to STED, these methods did not spatially control the ON/OFF transition; instead, sparse stochastic transitions to the ON state were induced in the field of view. The well-separated molecules populating the ON state were then localized by observing their emission on a camera, from which the centre of emission could be inferred.

STED enabled a resolution on the order of 30 nm, somewhat dependent on the specific sample conditions and the OFF switching intensity profile. As the resolution in STED microscopy is scaling with the inverse square-root of the applied illumination power of the OFF-switching light, the resolution enhancement was accompanied by an increased light dose. This was not arbitrarily scalable, as otherwise the sample and the fluorescent molecules would be damaged. This posed a limit to the achievable resolution in STED microscopy under practical conditions. Conversely, in PALM/STORM, the localization of an emitter position from a noisy, diffraction-limited distribution at a limited photon detection budget presented a challenge too. The achieved resolution was scaling with the inverse square-root of the number of detected photons N and was therefore strongly sample-dependent.

In 2017, a way of localizing isolated emitters without purely relying on the sample to dictate the achievable resolution (scaling with $1/\sqrt{N}$) was introduced [2]. MINFLUX (minimal fluorescence photon fluxes) enabled the control over the resolution scaling by localizing the emitter with an illumination pattern close to a zero. One can imagine that by placing an intensity-zero directly onto an emitter, no photon will be emitted. Nevertheless, the information of no photon arriving at the detector will precisely tell us that the emitter was exactly located at the zero position. Thus spatial information is gained even without any photon detection by precisely controlling the position within the illumination pattern at which the emitter is sampled. This concept was shown to localize single emitters to an uncertainty of about 1 nm with few detected photons. This value is on the order of the size of a single fluorescent molecule and thus marks the ultimate precision aim in fluorescence microscopy. MINFLUX was additionally shown to be capable of tracking the motion of an emitter by following it with the scanning illumination beam.

A few years later, localization sampling was also shown to be possible with STED [37]. An easy way of imagining this process is by recording STED scans of isolated emitters. In the image, the emitters will appear at a sub-diffraction size (the pattern created by each emitter in a STED measurement is considered the effective point spread function, E-PSF). Computing the pattern centres in the image results in a resolution scaling with $1/\sqrt{N}$, as in PALM/STORM [26]. But as the initial distribution – the E-PSF – is already

confined by the STED effect, this reduces the localization uncertainty linearly with the STED confinement. However, efficient STED localization was implemented in a different way: in order to reduce emitter bleaching and scanning times, the localization is performed by sampling the E-PSF only in close proximity to the emitter. This avoids exposing the emitter to high intensities of the OFF switching light and maximizes the information gain by sampling in a high-gradient region of the E-PSF. This method is denoted MINSTED.

I will present a novel implementation of MINSTED microscopy, which overcomes stability issues and therefore improves the localization precision. This improvement was realized by spectrally shifting the wavelength of the OFF switching light in order to increase the STED interaction efficiency. Based on these optimizations, MINSTED enabled the localization of fluorescent emitters at sub nanometre precision in artificial samples as well as in biological specimen. Furthermore, the first demonstration of MINSTED's ability to track the movement of dynamic emitters is presented.

Structure of this thesis After introducing the fundamental concepts of imaging systems and giving a brief introduction about fluorescent emitters, the classic microscopy techniques are outlined by calculating the basic metrics of image formation. Thereafter, the theory of the fundamental super-resolution methods will be covered. This leads to the statistical understanding of emission- and illumination-based localization. With these tools at hand, we will describe MINSTED and detail the influence of the spectrally shifted re-implementation. After giving a technical overview about the implementation, some of the achieved results will be presented, including imaging as well as tracking applications of MINSTED.

CHAPTER
TWO

FROM LIGHT TO MICROSCOPY

2.1 Classical imaging systems

The basics of microscopy can be understood, when becoming aware of the way in which linear shift-invariant imaging systems work. For this purpose, let us consider a simple imaging system: an optimal lens.

2.1.1 Geometric transformations of an optimal lens

A lens of focal length f placed at $z = 0$ is transforming an incoming light field from its front focal plane (object) at $z = -f$ to its back focal plane (image) at $z = f$ in a way that

(A) incoming light with a common propagation direction is focussed to a common position,

(B) incoming light, arising from a common position is transformed to a common propagation direction,

(C) light with a propagation direction parallel to the optical axis is focussed to the optical axis, and

(D) incoming light, arising from the optical axis is transformed to a propagation direction parallel to the optical axis.

This behaviour is shown in **Fig. 2.1**. To construct a general case, we assume different refractive indices n_o and n_i before and after the lens, respectively (object and image). We then define four beams, propagating as visualized in **Fig. 2.2**.

(1) A beam, travelling parallel to the optical axis, with a displacement x in the front focal plane. This results in an angle a to the optical axis in the back focal plane.

(2) A beam, that travels parallel to the optical axis, with a displacement χ in the back focal plane. This requires an angle α to the optical axis in the front focal plane.

(3) A beam, that matches the position of (1) and the angle of (2) in the front focal plane and thus the position of (2) and the angle of (1) in the back focal plane.

(4) A beam, travelling on the optical axis.

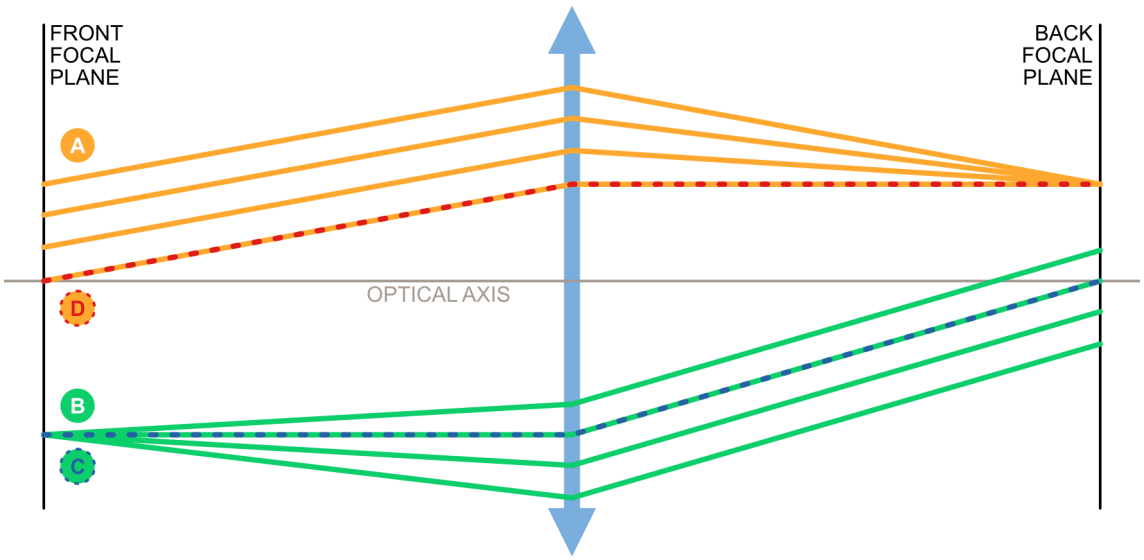

Figure 2.1: Transformation rules of an optimal lens. Four rules apply for the transform of an optimal lens from its front focal plane to its back focal plane: (A) parallel rays are focussed to a common point, (B) rays emerging from a common point are collimated to a parallel ray bundle, (C) rays travelling parallel to the optical axis are focussed to the optical axis and vice versa (D).

For an optimal lens, beams (1) and (4) have to share the same optical path length ($S_1 = S_4$), as they both stem from the same incoming planar wavefront and are transformed to the same outgoing spherical wavefront, which is phase matching in the focus. Same is true for beams (2) and (4). It follows that $S_1 = S_2$. Beams (1) and (3) arise from the same focal point and travel as a tilted wavefront in the back focal plane. For this tilted phase front to come up, S_3 thus has to be longer than S_1 by an amount of $\chi n_i \sin(a)$. Analogous, beams (2) and (3) lead to the condition $S_3 = S_2 - x n_o \sin(\alpha)$. Note that x is a negative value, so that the additional optical path length is positive. As $S_1 = S_2$, we conclude

$$\chi n_i \sin(a) = -x n_o \sin(\alpha) \ . \tag{2.1}$$

Introducing the lateral wave vectors

$$k_x = \frac{2\pi}{\lambda} n_i \sin(a) \quad \kappa_x = \frac{2\pi}{\lambda} n_o \sin(\alpha) \ , \tag{2.2}$$

this relation may also be rephrased as

$$\chi k_x = -x \ \kappa_x \ . \tag{2.3}$$

Rearrangement of eq. (2.1) gives

$$-\frac{x}{n_i \sin(a)} = \frac{\chi}{n_o \sin(\alpha)} = \text{const.} \ , \tag{2.4}$$

where the last statement holds as the left and the right side are fully independent from each other. The constant can be retrieved by considering the paraxial case that has to match the geometric solution (see **Fig. 2.2** inset), we therefore find

$$x = -f \ n_i \sin(a) \quad \chi = f \ n_o \sin(\alpha) \ . \tag{2.5}$$

Replacing the harmonic term by the lateral component of the wave vector results in

$$x(k_x) = -\frac{\lambda f}{2\pi} k_x \quad \chi(\kappa_x) = \frac{\lambda f}{2\pi} \kappa_x \ . \tag{2.6}$$

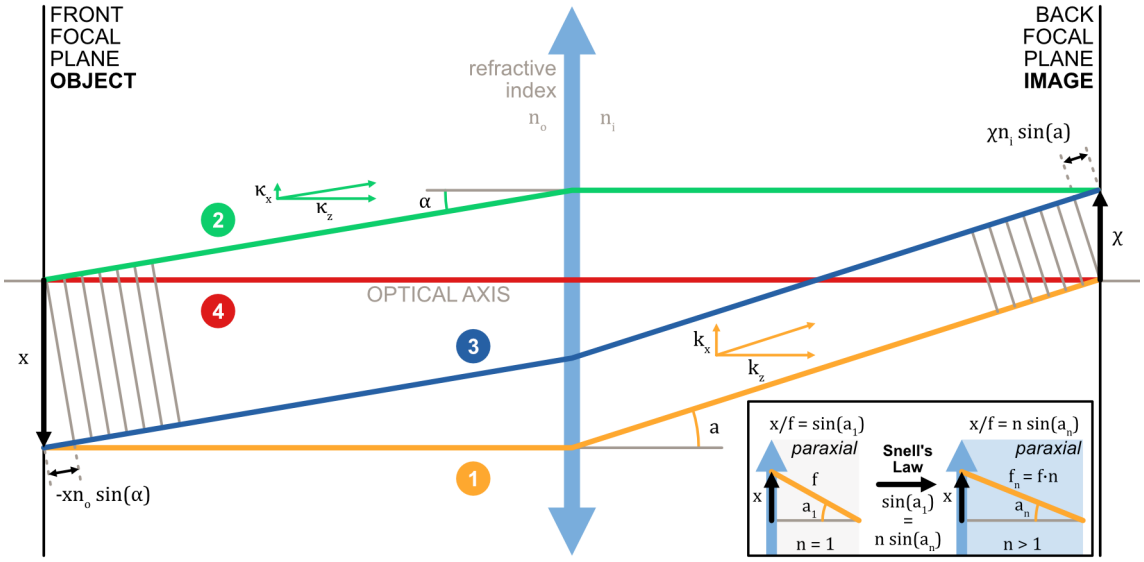

Figure 2.2: Sine-condition of an optimal lens. Four exemplary beams are transformed by an optimal lens from its front focal plane to its back focal plane, following the rules of the optimal lens transform. The arrangement is constructed in a way that beams (1) and (3) are part of the same spherical wave on the left hand side of the lens, which is collimated to a tilted plane wave on the right hand side. Meanwhile, beams (2) and (3) are part of a tilted plane wave on the left hand side, which transforms to a spherical wave on the right hand side. These two findings relate the optical path lengths of (1) and (3), as well as (2) and (3), respectively. The beams (1) and (4) are part of a common non-tilted wavefront, focussed to the optical axis. The beams (2) and (4) are part of the reverse scenario. Thus, the path lengths of (1) and (2) have to match, which leads to the condition $\chi n_i \sin(a) = -x n_o \sin(\alpha)$. The inset shows the geometric relation between an offset x of a beam from the optical axis in the front focal plane and its resulting angle in the back focal plane in the paraxial case, depending on the focal length f and the refractive index n in the back focal plane.

Each position thus transforms to a propagation direction, encoded by the lateral wave vector and vice versa. These quantities are related proportional to each other.

2.1.2 Fourier optics

Having understood these basic principles, we now want to calculate the image light field $E_i(x, y)$ in the back focal plane (at refractive index n_i), resulting from a lens transform of an arbitrary object field $E_o(x, y)$ in the front focal plane (at refractive index n_o). Light from the object point $\mathbf{r} = (r_x, r_y)$ in the front focal plane is creating a plane wave with a wave vector $(\mathbf{k}, k_z) = (k_x, k_y, k_z)$ and a complex amplitude $E_o(\mathbf{r})$ in the image (see **Fig. 2.3**). The component of the image field in the back focal plane at a position $\boldsymbol{\rho}$ produced by this single object point is thus given by[1]

$$E_o(\mathbf{r}) \exp(-i\, \mathbf{k} \cdot \boldsymbol{\rho}) \ . \tag{2.7}$$

[1]The axial component of the plane wave is missing, as the phase of arbitrary tilted plane waves is matching in the focus at $\rho = 0$. This is because all rays arriving in $\rho = 0$ belong to the same converging spherical wave with a common phase relation, resulting from a plane wave propagating parallel to the optical axis in the front focal plane.

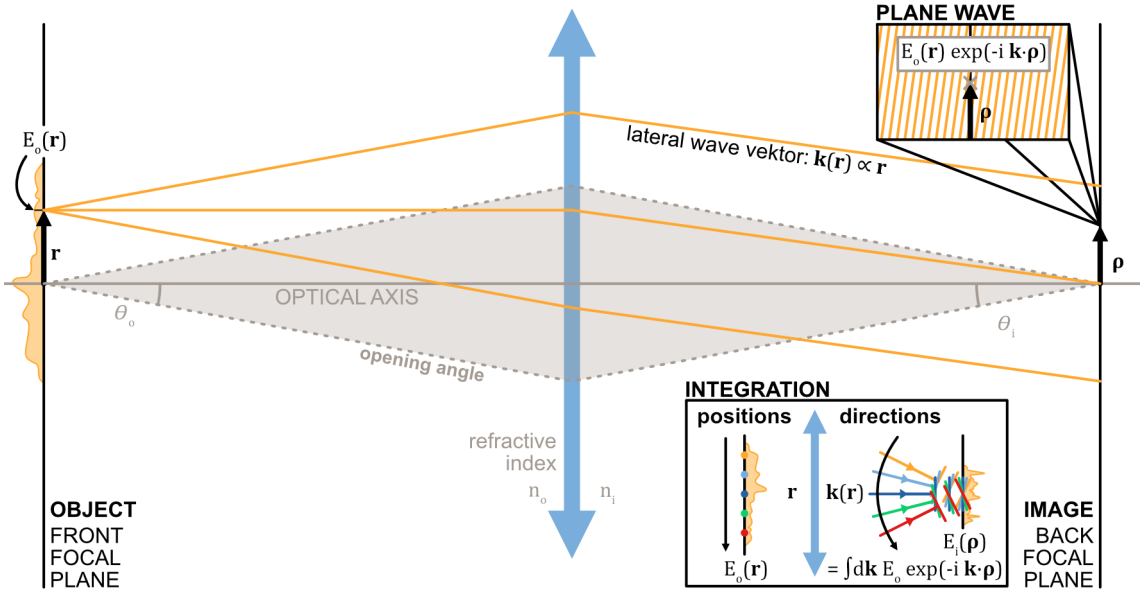

Figure 2.3: Fourier optics. Optical lens transforms an object field $E_o(\mathbf{r})$ to an image field $E_i(\boldsymbol{\rho})$. Each position $\mathbf{r}$ leads to a tilted planar wave of a form as shown in the inset on the top right. Summing over all planar waves emerging from the object plane is constructing the image field. This is sketched in the inset on the bottom right. The lens is thus implementing a Fourier transform. The points in the object plane are converted to spatial frequency components of the image. The opening angle on the image side of the lens is limiting the maximum angular tilts of the constructing planar waves and therefore the spatial details that are contained in the image. Nonetheless, even for a hemispherical aperture, the maximum spatial frequency in the image is limited by the wavelength of the illumination light. Upon transmission in the reverse direction, an equivalent spatial frequency limit applies with respect to the opening angle on the object side.

Integrating over all wave vectors, created by points in the object plane is constructing the full image in the back focal plane

$$\text{Form 1:} \quad E_i(\boldsymbol{\rho}) \propto \int \mathrm{d}\mathbf{k}\; E_o\left(\mathbf{r}(\mathbf{k})\right)\; \exp\left(-i\;\mathbf{k}\cdot\boldsymbol{\rho}\right)\;. \tag{2.8}$$

Additionally, introducing the wave vector $\boldsymbol{\kappa}$, corresponding to the spatial variable $\boldsymbol{\rho}$ and using relation (2.3), we find that

$$\text{Form 2:} \quad E_i\left(\boldsymbol{\rho}(\boldsymbol{\kappa})\right) \propto \int \mathrm{d}\mathbf{r}\; E_o(\mathbf{r})\; \exp\left(i\;\boldsymbol{\kappa}\cdot\mathbf{r}\right)\;. \tag{2.9}$$

We find in both cases E_i to be the two-dimensional (2D) Fourier transform of the object field E_o, while a linear scaling compensates for the wave vector to position coordinate change.

Form 1 of the integral is interpreted in a way that each position in the object creates a tilted plane wave in the image. By adding up all of those waves, the image is constructed by its frequency components and can be evaluated at position $\boldsymbol{\rho}$.

Form 2 of the integral is interpreted in a way that we project the object field onto its spatial frequency components, parametrized by the wave vector $\boldsymbol{\kappa}$. This can be thought as illuminating the object E_o with a tilted plane wave with a lateral wave vector $\boldsymbol{\kappa}$. Decomposing the object into its Fourier components $E_o = \sum_j c_j \exp\left(-i\;\mathbf{s}_j\cdot\mathbf{r}\right)$, for a specific choice of $\boldsymbol{\kappa}$, all components but $\mathbf{s}_j = \boldsymbol{\kappa}$ vanish. This value of the projected frequency

component is then focussed to the position $\boldsymbol{\rho}(\boldsymbol{\kappa})$.

We can thus see a lens as a device to decompose an object into its spatial frequencies or to compose an image from its spatial frequency components, as defined by the object.

Reformulating eqs. (2.8) and (2.9) in terms of the forward and backward Fourier transform

$$\begin{aligned} \mathcal{F}\Big[f(x)\Big](k) &= \int \mathrm{d}x\ f(x)\ \exp\left(-i\ k\cdot x\right) \\ \mathcal{F}^{-1}\Big[f(x)\Big](k) &= \int \mathrm{d}x\ f(x)\ \exp\left(i\ k\cdot x\right)\ . \end{aligned} \tag{2.10}$$

leaves us with

$$\begin{aligned} &\text{Form 1:} \quad E_i(\boldsymbol{\rho}) \propto \mathcal{F}\Big[E_o\big(\mathbf{r}(\mathbf{k})\big)\Big](\boldsymbol{\rho}) \\ &\text{Form 2:} \quad E_i\big(\boldsymbol{\rho}(\boldsymbol{\kappa})\big) \propto \mathcal{F}^{-1}\Big[E_o(\mathbf{r})\Big](\boldsymbol{\kappa})\ . \end{aligned} \tag{2.11}$$

2.1.3 Spatial frequency loss in optical systems

Though, we have to bear in mind that, in contrast to a regular mathematical Fourier transform, our wave vectors $\boldsymbol{\kappa}$ and $\mathbf{k}$ are physical parameters, which are not infinitely scalable. Having a look at eq. (2.11) Form 2, the spatial frequency components of the object field are carried by tilted planar waves with a lateral wave vector $\boldsymbol{\kappa}$, which in addition have to be captured by the lens in order to arrive at the image plane. This sets a limit

$$\kappa_c = \frac{2\pi\ n_o}{\lambda}\ \sin\left(\theta_o\right) \leq \frac{2\pi\ n_o}{\lambda}\ , \tag{2.12}$$

where θ_o defines the maximum angle under which the lens allows light to enter on the object side. Note that the ultimate limit on the right hand side of the equation is assuming a lens with an opening angle spanning a full hemisphere. Also in this case, the spatial frequency is limited by the wavelength of the light itself, which can at maximum create a harmonic oscillation of period λ/n_o.

Same applies to the image plane. As visible from eq. (2.11) Form 1, the image field is constructed by tilted plane waves exiting the lens with a lateral wave vector $\mathbf{k}$. Assuming an opening angle θ_i on this side of the lens, the maximum possible lateral wave vector is given by

$$k_c = \frac{2\pi\ n_i}{\lambda}\ \sin\left(\theta_i\right) \leq \frac{2\pi\ n_i}{\lambda} \tag{2.13}$$

In order to include those limits into the lens transformation integrals, we define an aperture function

$$\mathcal{A}_{s_c}(\mathbf{s}) = \begin{cases} 1, & \sqrt{s_x^2 + s_y^2} \leq s_c \\ 0, & \text{else} \end{cases}\ . \tag{2.14}$$

The limited versions of (2.11) thus read

$$\begin{aligned} &\text{Form 1:} \quad E_i(\boldsymbol{\rho}) \propto \mathcal{A}_{\kappa_c}\big(\boldsymbol{\kappa}(\boldsymbol{\rho})\big)\ \mathcal{F}\Big[E_o\big(\mathbf{r}(\mathbf{k})\big)\ \mathcal{A}_{k_c}(\mathbf{k})\Big](\boldsymbol{\rho}) \\ &\text{Form 2:} \quad E_i\big(\boldsymbol{\rho}(\boldsymbol{\kappa})\big) \propto \mathcal{A}_{\kappa_c}(\boldsymbol{\kappa})\ \mathcal{F}^{-1}\Big[E_o(\mathbf{r})\ \mathcal{A}_{k_c}\big(\mathbf{k}(\mathbf{r})\big)\Big](\boldsymbol{\kappa})\ . \end{aligned} \tag{2.15}$$

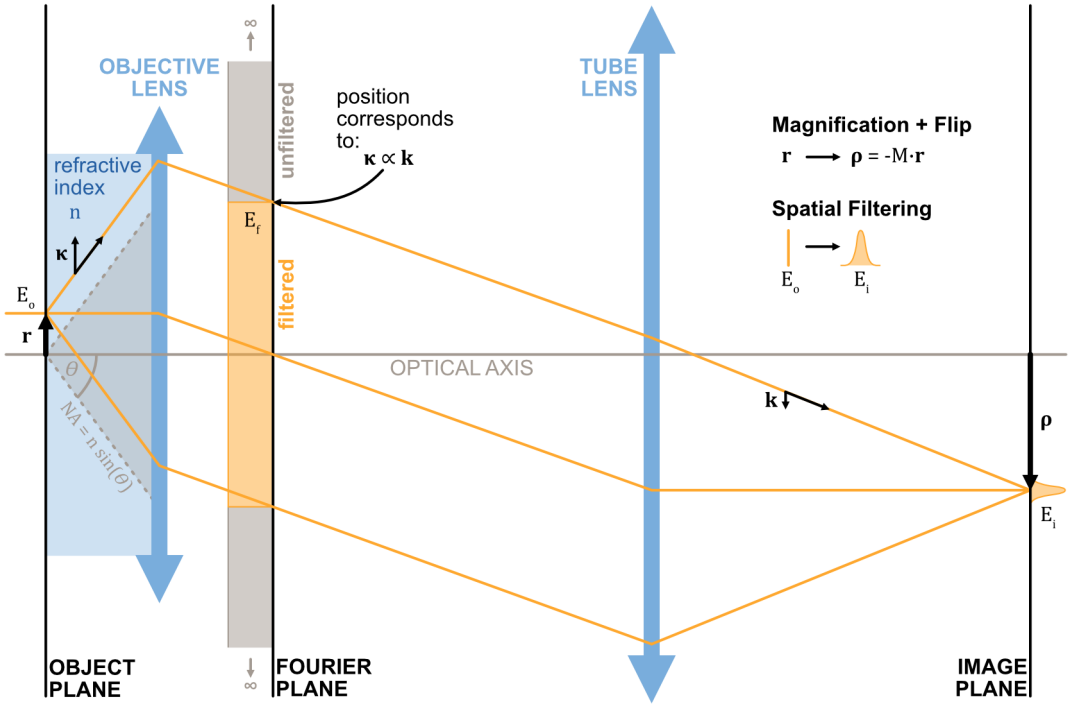

Figure 2.4: The 4-f system. Each point of the object field is transformed to a planar wave in the Fourier plane of the objective lens. The extent of the Fourier field is limited by the opening angle of the objective. The limited support of the Fourier field leads to a spread of each point in the object plane to a broadened light distribution in the magnified and flipped image field. This spread reduces the available resolution in the image field, as this is mixing the information arising from neighbouring points in the object field.

We thus found that spatial frequencies can only be transmitted by lenses up to a limit given by

1. the maximum spatial frequency of the light – the wavelength; and
2. the acceptance angles of the lens.

Yet this limit is not only applying to lenses, but to all cases where light fields are constructed by the overlay of rotated planar waves, which is generally true for all far-field applications.

2.1.4 The 4-f system

A basic microscope can be made of two lenses: an objective lens and a tube lens, which are arranged coaxially such that their facing focal planes fall together as sketched in **Fig. 2.4**. This optical configuration is called a 4-f system or relay system. The three focal planes shall be called the object, Fourier and image plane, and their respective fields E_o, E_f and E_i. The object coordinates are denoted $\mathbf{r}$, while the lateral wave vectors in the object plane, corresponding to positions in the Fourier plane, are denoted $\boldsymbol{\kappa}$. In the image plane, the spatial coordinates shall be denoted $\boldsymbol{\rho}$, while the constructing wave vectors, corresponding to positions in the Fourier plane, are denoted $\mathbf{k}$. Choosing Form 2 for the first transform from object to Fourier plane and Form 1 for the second, from Fourier to image plane, yields

$$\begin{aligned} E_f(\boldsymbol{\kappa}) &\propto \mathcal{A}_{\kappa_c}(\boldsymbol{\kappa})\ \mathcal{F}^{-1}\Big[E_o(\mathbf{r})\ \mathcal{A}_{r_c}(\mathbf{r})\Big](\boldsymbol{\kappa}) \\ E_i(\boldsymbol{\rho}) &\propto \mathcal{A}_{\rho_c}(\boldsymbol{\rho})\ \mathcal{F}\Big[E_f\big(\boldsymbol{\kappa}(\mathbf{k})\big)\ \mathcal{A}_{k_c}(\mathbf{k})\Big](\boldsymbol{\rho})\ . \end{aligned} \tag{2.16}$$

To reduce the number of variables, two of the aperture functions were translated to spatial coordinates. Those limit the field of view (FOV), due to the limited angular transmission upon exiting the objective and entering the tube lens. When assuming an object size that meets this constraint, those aperture functions may be left out of the consideration. Using eq. (2.6), we can relate $\boldsymbol{\kappa}$ and $\mathbf{k}$ as both encode the same position in the Fourier plane. Thus we find

$$f_o\boldsymbol{\kappa} = -f_t\mathbf{k} \quad \Leftrightarrow \quad \boldsymbol{\kappa} = -M\mathbf{k} \,, \tag{2.17}$$

where f_o and f_t are the focal lengths of objective and tube lens, respectively, and $M = f_t/f_o$ is the system's magnification. The objective shall feature a collection angle θ and a refractive index n on the object side. The combined parameter of those two is called the numerical aperture $\mathrm{NA} = n\,\sin(\theta)$. If the numerical aperture of the tube lens on the image side is larger than NA/M, the object side of the objective is the only element that performs relevant spatial frequency filtering, with a maximum lateral wave vector of

$$\kappa_c = \frac{2\pi\ \mathrm{NA}}{\lambda} \,. \tag{2.18}$$

Thus, the image field takes the form

$$E_i(\boldsymbol{\rho}) \propto \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(-M\mathbf{k})\ \mathcal{F}^{-1}\Big[E_o(\mathbf{r})\Big](-M\mathbf{k})\Big](\boldsymbol{\rho}) \tag{2.19}$$

Rescaling the coordinates and exchanging the integration variables, leads to

$$E_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa})\ \mathcal{F}^{-1}\Big[E_o(\mathbf{r})\Big](\boldsymbol{\kappa})\Big](\mathbf{r}) \tag{2.20}$$

Using the convolution theorem, this may be rewritten in the form

$$E_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto \Big(\mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa})\Big] * E_o\Big)(\mathbf{r}) \,. \tag{2.21}$$

The image field is thus a magnified and flipped version of the object field, convolved with the Fourier transform of the frequency filter (here, given by a circular aperture function). Due to this direct relation of object and image, we call the planes connected via a 4-f system to be conjugated. The frequency filter can be thought of as the blurred image of a perfect point emitter – the Green's function g of the optical system. The Fourier transform of the circular aperture function for the 4-f system evaluates to [32]

$$g(\mathbf{r}) \propto \frac{2\ J_1(\kappa_c r)}{\kappa_c r} \,, \tag{2.22}$$

where J_1 is the first order Bessel function and $r = |\mathbf{r}|$. As light fields are not directly observable, but only their intensities, the observable image of a point emitter creates a pattern of the form

$$G(\mathbf{r}) = |g(\mathbf{r})|^2 \,, \tag{2.23}$$

the so-called point spread function (PSF) of the system. This specific PSF is called the Airy disc (see **Fig. 2.5a**). Given a system's PSF, Abbe's resolution criterion assumes that two emitters located at a distance within each others PSFs' full width at half maximum (FWHM) are no longer resolvable. This is found at $\kappa_c d_c = 3.23$, which yields a critical distance d_c of

$$d_c = 0.51\ \frac{\lambda}{\mathrm{NA}} \,. \tag{2.24}$$

This result is considered the classical resolution limit or diffraction limit of imaging systems. Another important quantity of the Airy disc is the diameter of its first circular principal

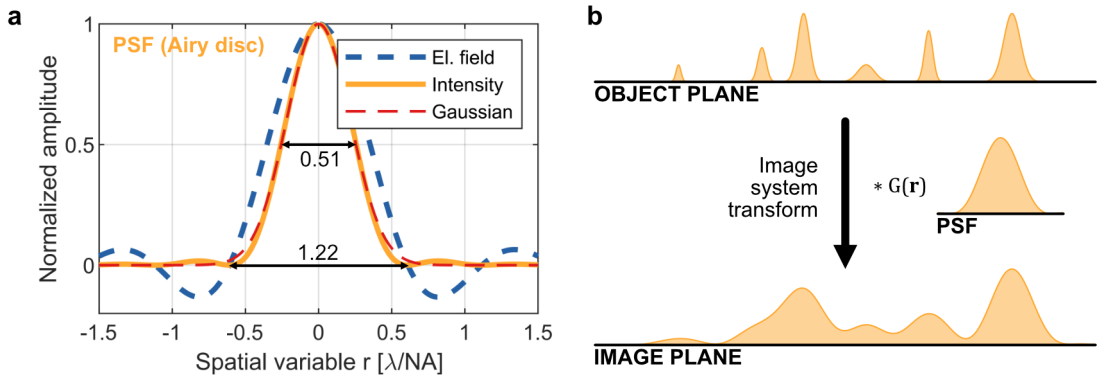

Figure 2.5: PSF of a 4-f system. **a** The normalized Airy disc intensity distribution (orange solid line) and its generating normalized electric field (blue dashed line). The Airy disc is well approximated by a Gaussian (red dashed line) matching its FWHM of 0.51 λ/NA. **b** The image of a linear optical system is retrieved by a convolution of the object field with the PSF.

minimum, which is often used as a measure for the size of the produced light spot. This quantity is called the Airy unit (AU), which amounts to

$$\mathrm{AU} = 1.22\,\frac{\lambda}{\mathrm{NA}}\ . \tag{2.25}$$

The image intensity distribution I_i that is created from a sample containing independent emitters, or which is illuminated with incoherent light, is different to that in the coherent case (eqs. (2.20) and (2.21)). For incoherent imaging, the integration over different sample positions is not allowed before the intensity pattern of each single position was formed in the image. We thus have to interpret the object as consisting of many delta peaks (one at each sample position), that have to be propagated through the system independent from each other. In the incoherent case, the image is thus forming from a convolution of the PSF with the object intensity I_o (see **Fig. 2.5b**)

$$I_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto (G * I_o)\,(\mathbf{r})\ . \tag{2.26}$$

Rephrasing this equation in the form of eq. (2.20) leads to

$$I_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto \mathcal{F}\Big[\tilde{\mathcal{A}}_{\kappa_c}(\boldsymbol{\kappa})\ \mathcal{F}^{-1}\Big[I_o(\mathbf{r})\Big](\boldsymbol{\kappa})\Big](\mathbf{r})\ , \tag{2.27}$$

where the incoherent frequency mask is given by

$$\tilde{\mathcal{A}}_{\kappa_c} = \mathcal{F}^{-1}\left[G\right] = \mathcal{F}^{-1}\left[g^2\right] = \mathcal{F}^{-1}\left[g\right] * \mathcal{F}^{-1}\left[g\right] = \mathcal{A}_{\kappa_c} * \mathcal{A}_{\kappa_c}\ . \tag{2.28}$$

2.1.5 Extensions

In the last sections, we have worked out a rather easy approach to model image generation simply by applying a Fourier transform to the object field. However, this model is missing an essential feature: light does not come as scalar field but as a vectorial quantity $E \to \mathbf{E}$. The light is polarized in the transverse plane with respect to its propagation direction, given by its wave vector. Light arising from the object point $\mathbf{r}$, featuring a vectorial field $\mathbf{E}_o$, is translated by the lens to a plane wave with a wave vector $\mathbf{k}(\mathbf{r})$. The electric field of this planar wave is still oscillating in a direction transverse to its propagation, which

means that for the Fourier integral, the plane wave amplitude has to be rotated by a matrix operation $\hat{R}(\mathbf{r})$ in order for the electric field to lie in the transverse plane of the respective wave vector $\mathbf{k}(\mathbf{r})$. This leads to a reformulated version of eq. (2.8).

$$\mathbf{E}_i(\boldsymbol{\rho}) \propto \int \mathrm{d}\mathbf{k}\ \hat{R}(\mathbf{r}(\mathbf{k})) \cdot \mathbf{E}_o\left(\mathbf{r}(\mathbf{k})\right)\ \exp\left(-i\ \mathbf{k} \cdot \boldsymbol{\rho}\right) \tag{2.29}$$

If additionally, the axial scaling of $\mathbf{E}_i$ is assessed, the further propagation of the plane waves upon variation of the image plane's axial coordinate ρ_z has to be considered.

$$\mathbf{E}_i(\boldsymbol{\rho}) \propto \int \mathrm{d}\mathbf{k}\ \hat{R}(\mathbf{r}(\mathbf{k})) \cdot \mathbf{E}_o\left(\mathbf{r}(\mathbf{k})\right)\ \exp\left(-i\ \mathbf{k} \cdot \boldsymbol{\rho}\right)\ \exp\left(-i\ k_z \rho_z\right) \tag{2.30}$$

This is denoted the Debye-Wolf integral, which provides a good model for calculating the focal field of high-NA objective lenses from the electric field in its pupil. In this formula, the z-component of the wave vector is given by

$$k_z = \sqrt{\frac{2\pi\ n_i}{\lambda} - k_x^2 - k_y^2}\ . \tag{2.31}$$

This extension is increasing the level of complexity and does not strictly help to understand the further considerations in sections 2.3 and 2.4. We will thus stick to the regular 2D Fourier optics model. All of the presented calculations still remain good approximations. However, when it comes to high-NA focal fields, it has to be kept in mind that for a precise model, the vectorial nature of light has to be included.

2.2 Fluorescence

Before moving on to the discussion of the different microscopy methods, the concept of fluorescence, which is a central basis for many microscopy methods, will be introduced. Fluorescence is the spontaneous emission of light by a material after the absorption of an exciting photon. An additional condition for such a process to be considered as fluorescence is that the corresponding excited and ground states are electronic states with an allowed electronic dipole transition [34]. Single molecules, showing such a behaviour are called fluorophores, fluorescent emitters or fluorescent molecules. Throughout this dissertation, I will mostly use the term fluorescent emitter or more general emitter (the last one is not strictly limited to fluorophores). The quantum mechanical details about the states of fluorophores and their transitions are highly complex and will not be discussed as part of this work. However, some basics will be covered in the following.

As mentioned above, all fluorophores possess an electronic ground and excited state. These states are mostly split into various vibrational levels. In the absence of excitation light, the fluorophore is usually in the ground state. Given an ensemble of fluorophores, the vibrational levels are populated depending on their energy with respect to a Boltzmann distribution (exponential) at the given temperature. Exposing the fluorophore to light, it may undergo the transition to the excited state upon absorption of a photon (or multiple for n-photon absorption). The probability of a single-photon absorption is linearly scaling with the photon density (proportional to the intensity) of the light field and the cross-section of the transition, which scales with the overlap of the ground and excited state wave functions [34]. Upon photon absorption, not only the electronic state of the fluorophore

but also its vibrational level may change. Given a certain ground state vibrational level, different excited state vibrational levels are reached depending on the wavelength of the light field. This leads to a wavelength-dependence in the overlap of initial and final state of the transition. Additionally, all possible ground states, given their Boltzmann distributed probability, have to be accounted for in order to find the average transition probability at a given temperature for the wavelength of the illumination field in question.

The spectral excitation probability may be measured by observing the wavelength-dependent absorbance in a fluorophore solution. The absorbance $A = \ln(I_{in}/I_{out})$ is defined as the logarithmic ratio of input and output intensity. Assuming a constant absorption probability in the solution, an exponential decay of the intensity along the trajectory is suspected (Beer-Lambert law) [34]. The exponential features a decay constant that scales with the excitation cross-section. The absorbance A is therefore directly proportional to the cross-section of the excitation process. For most fluorescent molecules, the absorbance spectrum $A(\lambda)$ (usually just called the absorption spectrum) is a well-known distribution, helping to evaluate suitable excitation wavelengths for a given fluorophore (see **Fig. 2.6**).

At photon energies below the electronic gap between ground and excited state, excitation will only take place if the fluorophore was already at a higher vibrational level. The red edge of the absorbance spectrum is therefore usually well approximated by an exponential decay with decreasing photon energy.

After an absorption took place, the fluorophore is vibrationally relaxing towards the lowest vibrational level of the excited state within femto- to picoseconds. At this point, an ensemble of fluorophores in the excited state would again obey Boltzmann distributed vibrational levels. From this, the fluorophore is at some point[2] falling back to any of the vibrational levels of the ground state. Upon transition, the fluorophore is emitting a fluorescence photon with a wavelength corresponding to the respective energy difference of initial and final state. Assuming the vibrational structure of the excited state to exactly replicate that of the ground state, the overlap between two specific vibrational levels upon excitation or emission is equal. This leads to the emission spectrum being a red shifted (Stokes shift) and mirrored replicate of the absorption spectrum. As for most real fluorescent molecules, the vibrational structure is more complex, symmetry is usually broken as visible for the emission spectrum $E(\lambda)$ of the fluorescent molecule Cy3B in **Fig. 2.6**.

The Stokes shift is one of the key properties of fluorescence. As the fluorescence photons have on average a lower energy than those of the excitation light, they can be optically separated from each other. The fluorescence may thus be observed at a high contrast, without being perturbed by the illumination light because it can be isolated by spectral filtering.

While a fluorophore is in the excited state, it can also be forced to undergo a ground state transition via stimulated emission. This process is based on a light field that matches the energy gap of a possible radiative decay of the molecule. This light field stimulates the molecule to undergo the decay while a photon is coherently added to the light field. The cross-section of the stimulated emission process is proportional to the emission spectrum $E(\lambda)$ and a factor of λ^4 as derived in reference [6]. Therefore, not only the excitation may be optically controlled but also the de-excitation.

[2]The average time until the spontaneous fluorescence decay takes place is called the fluorescence lifetime. Lifetimes of commonly used fluorescent molecules are usually on the order of nanoseconds.

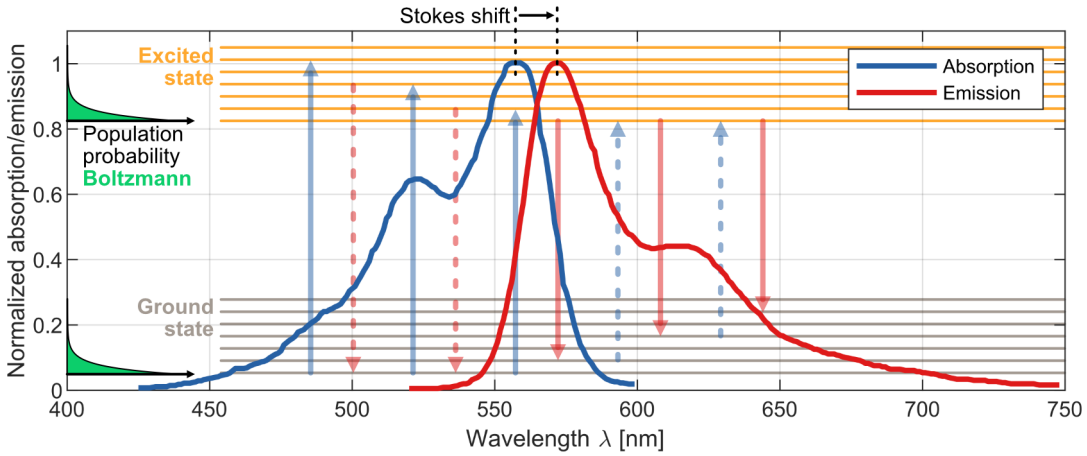

Figure 2.6: Fluorescence spectra and state transitions. The normalized absorption (blue line) and emission (red line) spectrum of the fluorescent emitter Cy3B. A schematic Jablonski diagram of energy states from low energy (bottom) to high energy (top) is underlaid. Arrows schematically indicate possible transitions upon excitation (blue) and emission (red). The exponential population probability of vibrational levels is qualitatively outlined in green.

Besides the ground and the excited state, fluorophores may feature long-lived (lifetime on the order of micro- to millisecond) states with non-radiative decays. Some fluorescent molecules stochastically enter such states, which upon continuous illumination with an excitation field leads to an irregularly interrupted stream of fluorescence photons. This behaviour is called blinking. Some fluorescent molecules also allow an optically triggered transfer to such states.

Apart from electronic states, some fluorescent emitters feature two or more chemical states, such as certain isomers or additional chemical groups that may be split off from the molecular core. Transition between such chemical states (reversible or irreversible) may change the molecule's ability to fluoresce or shift its emission and absorption spectrum. Such state transitions may be spontaneous or optically triggerable and can be used to control the excitability or fluorescence ability at a certain spectral range.

In optical microscopy, fluorescent emitters are used to stain samples in order to get a highly specific contrast, only showing a desired structure of interest. In biological imaging applications, this is a common way of tagging specific proteins in order to extract their position and dynamics from fluorescence microscopy data.

2.3 Diffraction limited microscopy

2.3.1 Widefield

The widefield microscope is the most basic microscopy method. Its fundamental image transform is already explained by the math in section 2.1.4. The sample is illuminated with light of a homogeneous intensity. Each sample point is then 4-f imaged onto a detector plane (or may be observed by eye). When observing independent emitters (fluorescence) or illuminating with incoherent light, neighbouring object positions do not interfere with

each other[3]. Each thus creates its own intensity pattern in the image plane. Hence, the image is formed as a convolution of the object with the intensity PSF (see **Fig. 2.5b**). In the coherent case, the image is formed from the ensemble of emitters. The object is therefore convolved with the Green's function and the absolute square is evaluated at the detector.

In the most straight-forward widefield implementation, a bright field microscope, the sample's absorption is observed on top of the illumination light. Alternatively, several contrast mechanisms may be used to separate the illumination light from that which was affected from its interaction with the sample, for example:

- Phase contrast, in order to observe phase delay created by the sample.
- Dark field in order to observe directional changes induced to the illumination by the sample (scattering).
- Fluorescence, where the emitted red-shifted light may be spectrally separated from the illumination.

2.3.2 Confocal microscopy

In a confocal microscope, an illumination beam is scanned over the sample. The illumination beam itself is limited by diffraction because it is generated in the object plane from a uniform illumination of the Fourier plane (pupil) of the objective. Its electric field in the object plane thus takes the form

$$g_{ill}(\mathbf{r}) = \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa})\Big](\mathbf{r}) = g(r) \ . \tag{2.32}$$

The object field is then given by the multiplication of the illumination field (centered at $\mathbf{r}_{ill}$) and the amplitude-phase transmission function, t_o, specifying the sample

$$E_o(\mathbf{r}) = g_{ill}(\mathbf{r} - \mathbf{r}_{ill})\ t_o(\mathbf{r}) \ . \tag{2.33}$$

This object field may be imaged through the optical system, while we denote the emission Green's function g_{em}. For clarity, we write the convolution with the emission Green's function as an integral and find

$$E_i(\boldsymbol{\rho}, \boldsymbol{\rho}_{ill}) \propto \int \mathrm{d}\boldsymbol{\xi}\ g_{em}(\boldsymbol{\xi})\ g_{ill}(\mathbf{r} - \mathbf{r}_{ill} - \boldsymbol{\xi})\ t_o(\mathbf{r} - \boldsymbol{\xi}) \ , \tag{2.34}$$

where the arguments of the image field refer to the magnified and flipped image coordinates. In confocal microscopy, the signal is imaged onto a pinhole, which is centred on the projected illumination position. All signal transmitting the pinhole is then integrated by a bucket detector and yields the image intensity for the scanning position $\mathbf{r}_{ill}$. The pinhole of radius $M \cdot \mathcal{R}$ may be modelled by an aperture function of the form $\mathcal{A}_{\mathcal{R}}(\mathbf{r} - \mathbf{r}_{ill})$ while the summation of all signal hitting the pinhole is represented by an integration in $\mathbf{r}$

$$E_i(\boldsymbol{\rho}_{ill} = -M\mathbf{r}_{ill}) \propto \int \mathrm{d}\mathbf{r}\ \mathcal{A}_{\mathcal{R}}(\mathbf{r} - \mathbf{r}_{ill}) \int \mathrm{d}\boldsymbol{\xi}\ g_{em}(\boldsymbol{\xi})\ g_{ill}(\mathbf{r} - \mathbf{r}_{ill} - \boldsymbol{\xi})\ t_o(\mathbf{r} - \boldsymbol{\xi}) \ . \tag{2.35}$$

Performing two coordinate transforms

$$\tilde{\boldsymbol{\xi}}(\boldsymbol{\xi}) = \boldsymbol{\xi} - \mathbf{r} + \mathbf{r}_{ill} \quad \mathbf{s}(\mathbf{r}) = \mathbf{r} - \mathbf{r}_{ill} \tag{2.36}$$

[3]Light is never fully incoherent, interference effects at distances below the scale of a single wavelength are thus always present.

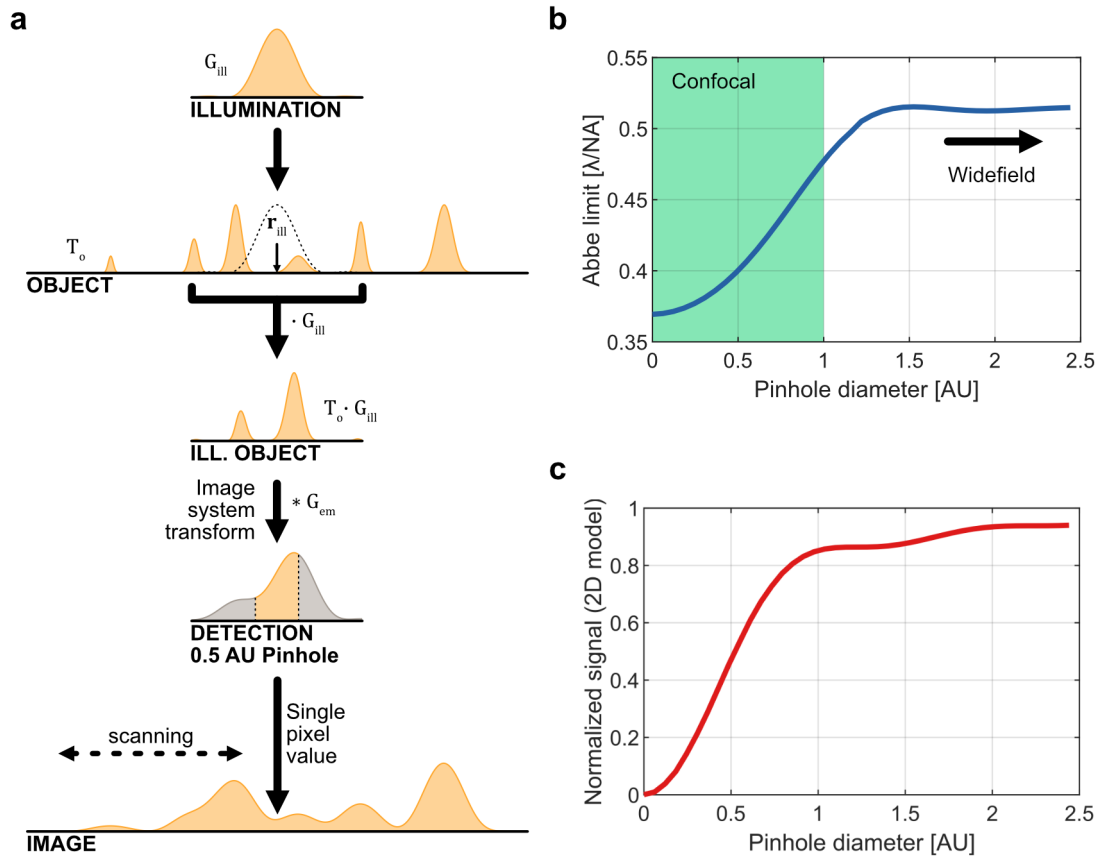

Figure 2.7: Confocal microscopy. **a** Image formation in a confocal microscope. A subset of the object is illuminated at a position $\mathbf{r}_{ill}$. The illuminated is transformed through the imaging system by convolution with the emission PSF. Detecting the central part of the emitted light filters contributions from the periphery of the illumination distribution. Scanning the illumination and the co-aligned filtering pinhole over the sample, is producing the image. **b** Scaling of the Abbe limit with the pinhole diameter. For zero pinhole size, the resolution is theoretically reduced down to 0.37 λ/NA. This goes hand in hand with a reduction of the observed signal as shown in **c**.

leads to

$$E_i(\boldsymbol{\rho}_{ill} = -M\mathbf{r}_{ill}) \propto \int \mathrm{d}\tilde{\boldsymbol{\xi}} \int \mathrm{ds}\ \mathcal{A}_{\mathcal{R}}(\mathbf{s}) g_{em}(\tilde{\boldsymbol{\xi}} + \mathbf{s})\ g_{ill}(-\tilde{\boldsymbol{\xi}})\ t_o(\mathbf{r}_{ill} - \tilde{\boldsymbol{\xi}})\ . \tag{2.37}$$

Using the radial symmetry of our Green's functions, we can rewrite this as

$$E_i(\boldsymbol{\rho}_{ill} = -M\mathbf{r}_{ill}) \propto \Big(\big[(\mathcal{A}_{\mathcal{R}} * g_{em})\ \ g_{ill} \big] * t_o \Big)(\mathbf{r}_{ill})\ . \tag{2.38}$$

The math in the incoherent case yields an analogous result.

$$I_i(\boldsymbol{\rho}_{ill} = -M\mathbf{r}_{ill}) \propto \Big(\big[(\mathcal{A}_{\mathcal{R}} * G_{em})\ \ G_{ill} \big] * T_o \Big)(\mathbf{r}_{ill}) \tag{2.39}$$

The image generation in the confocal case is schematically illustrated in **Fig. 2.7a**. From a mathematical point of view, the confocal microscope looks like a widefield system with a PSF function which is the multiplication of the illumination PSF and the detection PSF. The latter is itself consisting of the convolution of the emission PSF with the pinhole. We find the ultimate resolution limit in confocal microscopy by assuming an infinitely small pinhole, which leads to

$$I_i(\boldsymbol{\rho}_{ill} = -M\mathbf{r}_{ill}) \overset{\mathcal{R}=0}{\propto} \Big((G_{em}\ G_{ill}) * T_o \Big)(\mathbf{r}_{ill}) \approx \left(G_{em}^2 * T_o \right)(\mathbf{r}_{ill})\ . \tag{2.40}$$

Thus, finding the Abbe limit for the squared PSF leads to

$$d_c^{conf} \geq 0.37 \, \frac{\lambda}{\mathrm{NA}} \, , \tag{2.41}$$

which corresponds to an extension of the widefield resolution limit by a factor of $\sqrt{2}$. The scaling of the confocal resolution limit with the pinhole size is presented in **Fig. 2.7b**, while the associated signal loss is shown in **Fig. 2.7c**.

2.3.3 Structured illumination microscopy

In structured illumination microscopy (SIM), the sample is illuminated with a standing wave pattern E_{ill} of lateral frequency $\boldsymbol{\kappa}_{ill}$, satisfying $|\boldsymbol{\kappa}_{ill}| \leq \kappa_c$. Typically, this pattern is introduced under various orientations and phase shifts. Using eq. (2.20), we find

$$\begin{aligned} E_i(\boldsymbol{\rho} = -M\mathbf{r}) &\propto \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa}) \; \mathcal{F}^{-1}\Big[E_{ill}(\mathbf{r}) \; t_o(\mathbf{r})\Big](\boldsymbol{\kappa})\Big](\mathbf{r}) \\ &\propto \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa}) \; \left(\mathcal{F}^{-1}\left[E_{ill}\right] * \mathcal{F}^{-1}\left[t_o\right]\right)(\boldsymbol{\kappa})\Big](\mathbf{r}) \; . \end{aligned} \tag{2.42}$$

Assuming a perfect standing wave illumination, it holds that

$$\mathcal{F}^{-1}\left[E_{ill}(\mathbf{r})\right](\boldsymbol{\kappa}) \propto \delta(\boldsymbol{\kappa} - \boldsymbol{\kappa}_{ill}) \; e^{i\phi/2} + \delta(\boldsymbol{\kappa} + \boldsymbol{\kappa}_{ill}) \; e^{-i\phi/2} \; , \tag{2.43}$$

where δ denotes the Dirac delta distribution and ϕ quantifies the phase shift of the counter-propagating frequency components. Carrying out the convolution with the shifted Dirac peaks, we find

$$E_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa}) \; \Big(e^{i\phi/2}\mathcal{F}^{-1}\left[t_o\right](\boldsymbol{\kappa} - \boldsymbol{\kappa}_{ill}) + e^{-i\phi/2}\mathcal{F}^{-1}\left[t_o\right](\boldsymbol{\kappa} + \boldsymbol{\kappa}_{ill})\Big)\Big](\mathbf{r}) \; . \tag{2.44}$$

Transforming the integration constants, we can easily transfer the wave vector shifts onto the aperture function, which gives us

$$\begin{aligned} E_i(\boldsymbol{\rho} = -M\mathbf{r}) \propto \; & \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa} + \boldsymbol{\kappa}_{ill}) \; \mathcal{F}^{-1}\left[t_o\right](\boldsymbol{\kappa})\Big] \; e^{i\phi/2} \\ & + \mathcal{F}\Big[\mathcal{A}_{\kappa_c}(\boldsymbol{\kappa} - \boldsymbol{\kappa}_{ill}) \; \mathcal{F}^{-1}\left[t_o\right](\boldsymbol{\kappa})\Big] \; e^{-i\phi/2} \; . \end{aligned} \tag{2.45}$$

The first term thus contains spatial frequencies of the object up to $|\boldsymbol{\kappa}_{ill}| + \kappa_c$ (along the direction of the illumination wave vector), while the second term, goes down to $-|\boldsymbol{\kappa}_{ill}| - \kappa_c$. Though, one has to keep in mind that these two fields are overlaid and have to be disentangled in order to obtain the frequency information from the whole range. This may be accomplished numerically by collecting images at different relative phase shifts ϕ.

When using the full spatial frequency range of the illumination, the illumination wave vector may be tuned up to κ_c, which enables (using enough rotation angles and phase shifts) a doubling of the frequency components in each dimension. This leads to a reduction of the widefield resolution limit up to a factor of two

$$d_c^{\mathrm{SIM}} \geq 0.25 \, \frac{\lambda}{\mathrm{NA}} \; . \tag{2.46}$$

When treating the incoherent case, the math is basically analogous, with the exception that the standing wave intensity features a Fourier representation of the form

$$\mathcal{F}^{-1}\left[I_{ill}(\mathbf{r})\right](\boldsymbol{\kappa}) \propto \delta(\boldsymbol{\kappa} - 2\boldsymbol{\kappa}_{ill}) \; e^{i\phi} + \delta(\boldsymbol{\kappa} + 2\boldsymbol{\kappa}_{ill}) \; e^{-i\phi} + 2\delta(\boldsymbol{\kappa}) \; . \tag{2.47}$$

In addition to the negative and positive frequency shifts, the spectrum is overlaid with the non-frequency-shifted version of the widefield image. This makes the disentanglement of the frequency ranges more challenging, but features the same factor of two in resolution enhancement[4].

2.4 Sub-diffraction microscopy

As we have learned in the last section, light can only transport spatial information up to a detail scale of its own wavelength. Using the full frequency support of an illumination field plus the full frequency support of the detection, enables an extension of the classical resolution limit by a factor of two, which is the strongest stretch of the resolution limit that is possible with a far-field system and a linearly responding sample. Everything beyond this limit is achieved by making use of non-linear sample responses. To date, this was extensively shown in fluorescence microscopy.

The concept of n-photon microscopy [18] or the use of special luminescent nano-particles with a strong non-linear emission (of degree n) with respect to the incident intensity [19] extends the diffraction barrier by a factor $\approx \sqrt{n}$. However, multi-photon techniques have to induce high intensities and matching wavelengths to increase the interaction probability. This may induce photo-toxic side-effects in biological specimen. Additionally, multi-photon microscopy is hardly scalable beyond $n \approx 3$. Using exotic nano-particles, with a strong non-linear response to the incident excitation intensity is quite sensitive to the applied optical parameters and not either flexibly scalable to arbitrarily large values of n. Moreover, such particles are on the order of tens of nanometres in size, which renders them unsuitable for a wide range of labelling schemes.

Beyond such approaches, the field of super-resolution with a number of sub-diffraction microscopy methods emerged since the invention of STED in 1994 [11]. Several of those methods will be introduced in the following sections. What all of them have in common is that they principally enable unlimited resolution because they rely on the strongest possible non-linear sample response – the transition between two states, a bright and a dark one, ON and OFF.

The super-resolution methods may be separated into two subclasses: the super-resolution imaging methods, such as

- **STED**: STimulated Emission Depletion microscopy [11]
- **RESOLFT**: REversible Saturable Optical Linear Fluorescence Transitions microscopy [13]
- **GSD**: Ground State Depletion microscopy [10],

and the super-resolution localization methods, which basically divide into the emission localization methods, such as

- **STORM**: STochastic Optical Reconstruction microscopy [28]

[4]Please note, that the addend $2\boldsymbol{\kappa}_{ill}$ in the intensity is not extending the frequency spectrum by a factor of three. The intensity image does in any case feature frequency components of up to $2\kappa_c$, which come about by squaring of the image field, see eq. (2.28).

- **PALM**: Photo-Activated Localization Microscopy [3]
- **PAINT**: Point Accumulation for Imaging in Nanoscale Topography [14],

and the illumination localization methods

- **MINFLUX**: MINimal fluorescence photon FLUXes microscopy [2]
- **MINSTED**: MINimal STimulated Emission Depletion microscopy [37],

while there exists one mixed form of emission and illumination localization, **SIMFLUX** [5], an emission localization method with an additional structured illumination.

2.4.1 Imaging methods

The imaging methods in super-resolution need the ON to OFF transition of their emitters to be optically triggerable. In order to investigate the influence of a light pattern I on an emitter ensemble featuring two states, ON and OFF, we set up a simple model for the emitter state population given by a rate equation of a two-level system[5]

$$\dot{P}_{\mathrm{ON}} = P_{\mathrm{OFF}}\ k_{\mathrm{ON}}(I,\lambda) - P_{\mathrm{ON}}\ k_{\mathrm{OFF}}(I,\lambda)\ . \tag{2.48}$$

The dot represents the first derivative with respect to time t, P_{ON} and P_{OFF} the respective state population probabilities while k_{ON} and k_{OFF} denote the ON and OFF switching rates, respectively, dependent on the light intensity I of wavelength λ. Assuming $P_{\mathrm{ON}} + P_{\mathrm{OFF}} = 1$ and as initial condition $P_{\mathrm{ON}}(t=0,\mathbf{r}) = P_0(\mathbf{r})$, we can solve (2.48) analytically and find

$$P_{\mathrm{ON}} = \frac{k_{\mathrm{ON}}}{k_{\mathrm{ON}}+k_{\mathrm{OFF}}} + \left[P_0 - \frac{k_{\mathrm{ON}}}{k_{\mathrm{ON}}+k_{\mathrm{OFF}}}\right]\exp\left(-(k_{\mathrm{ON}}+k_{\mathrm{OFF}})t\right)\ . \tag{2.49}$$

Assuming purely light-driven transitions, the rates are generally described by (formula abstracted from reference [17])

$$k_i(I,\lambda) = \frac{I\lambda\sigma_i(\lambda)}{hc}\ , \tag{2.50}$$

with the wavelength dependent ON/OFF cross-sections σ_{ON} and σ_{OFF}, the Planck constant h and the speed of light c. Herein, we neglect spontaneous decay from ON to OFF, which presupposes that the lifetime of the ON state is much longer than the integration time describable with the given model. We hereby assume the optical interactions to take place rather quickly, which means that we are describing optically triggered transitions by pulses with durations well below the lifetime of the ON state. Using eq. (2.50), we obtain

$$P_{\mathrm{ON}}(t) = \frac{\sigma_{\mathrm{ON}}}{\sigma_{\mathrm{ON}}+\sigma_{\mathrm{OFF}}} + \left[P_0 - \frac{\sigma_{\mathrm{ON}}}{\sigma_{\mathrm{ON}}+\sigma_{\mathrm{OFF}}}\right]\exp\left(-(k_{\mathrm{ON}}+k_{\mathrm{OFF}})t\right)\ . \tag{2.51}$$

[5]Defining this model only in terms of two states may appear insufficient. Considering for example the ground and excited state of a fluorescent molecule, it is evident that additional vibrational levels are necessary to capture the characteristics of the state dynamics upon light—fluorophore interaction. However, these additional states are short-lived compared to the time scale of interest and may thus be considered unpopulated. They still influence the system by inducing the shape of the absorption and emission spectra. Consequently, upon appropriate choice of the ON and OFF switching rate spectra, the presented calculation effectively contains two arbitrarily complex bands of states (with an infinitely short lifetime), while only considering two populated states explicitly.

For a non-zero illumination, the signal levels off at a constant value given by the first term. We denote this value as the background (steady state)

$$B(\lambda) = \frac{\sigma_{\mathrm{ON}}(\lambda)}{\sigma_{\mathrm{ON}}(\lambda) + \sigma_{\mathrm{OFF}}(\lambda)} \ . \tag{2.52}$$

Additionally, we rewrite

$$k_{\mathrm{ON}} + k_{\mathrm{OFF}} = f(\lambda)\ I(\mathbf{r})\ , \ \text{with} \quad f(\lambda) = \frac{\lambda(\sigma_{\mathrm{ON}}(\lambda) + \sigma_{\mathrm{OFF}}(\lambda))}{hc} \ . \tag{2.53}$$

Rewriting the rates with respect to the space dependent intensity, we find

$$P_{\mathrm{ON}}(\mathbf{r}, t) = B(\lambda) + [P_0(\mathbf{r}) - B(\lambda)] \exp\left(-f(\lambda)\ I(\mathbf{r})\ t\right) \ . \tag{2.54}$$

This solution approximates the ON state population upon illumination with light of arbitrary wavelength λ and intensity distribution $I(\mathbf{r})$. For low light intensities, this function approaches a linear behaviour. In super-resolution imaging, we mostly rely on two optically triggered transitions: the OFF to ON and the ON to OFF transition. Usually, a first beam switches emitters in a desired region ON while a successive second beam triggers the reverse process in order to leave a narrow ON volume close to points of minimal intensity in the OFF switching light. We assume an illumination PSF of the form $\propto G_{ill}(\mathbf{r})$, that triggers the OFF to ON transition from an initial state $P_{\mathrm{ON}} = 0$. In the linear regime (at low intensities), this shall lead to a ON probability of $G_{ill}(\mathbf{r})$, which serves as the initial condition $P_0 = G_{ill}$ for the successive OFF switching process. This OFF switching is the fundamental condition for super-resolution imaging. To this end, an OFF switching illumination pattern of wavelength λ_{OFF} featuring at least one intensity minimum (optimally zero) is used. For our analytical consideration, we now approximate the shape of the intensity, close to such a minimum, by a parabola with an offset ϵ and an intensity scaling factor I_0

$$I(\mathbf{r}) \approx I_0 \left(\epsilon + ar^2\right) \ , \ \text{with} \quad r = |\mathbf{r}| \ . \tag{2.55}$$

This leads to a Gaussian shape of the ON probability after an OFF switching illumination time $t = T$

$$\begin{aligned} P_{\mathrm{ON}}(\mathbf{r}, T) = \ & B(\lambda_{\mathrm{OFF}}) \\ & + [G_{ill}(\mathbf{r}) - B(\lambda_{\mathrm{OFF}})] \exp\left(-f(\lambda_{\mathrm{OFF}}) I_0 T\ \epsilon\right) \exp\left(-f(\lambda_{\mathrm{OFF}})\ I_0 T\ ar^2\right) \ . \end{aligned} \tag{2.56}$$

The two main parameters that we can extract from this analytical calculation, are the FWHM and the visibility of the ON distribution. We define the visibility $\mathcal{V}$ as the ratio of the dynamic range of the distribution (the amplitude of the second term) and the constant background (the first term)

$$\mathcal{V} = \frac{G_{ill}(\mathbf{r} = 0) - B(\lambda_{\mathrm{OFF}})}{B(\lambda_{\mathrm{OFF}})} \exp\left(-f(\lambda_{\mathrm{OFF}})\ I_0 T\ \epsilon\right) \ . \tag{2.57}$$

The first term is limited by the initial maximum ON probability and by the contrast ratio of the ON and OFF switching cross-sections. The wavelength choice may influence this parameter. Additionally, if the minimum position of the OFF switching intensity pattern is featuring a residual intensity of ϵI_0, this leads to an exponential decrease of the visibility

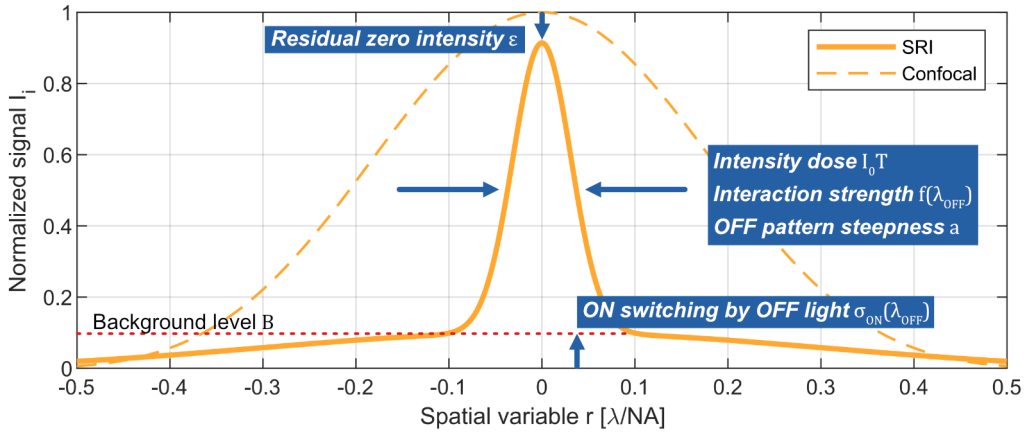

Figure 2.8: E-PSF in super-resolution imaging. From the confocal E-PSF (dashed line), using an OFF transition to a dark state of an emitter, the width of the respective E-PSF (solid line) may be decreased to arbitrarily small values. A narrower E-PSF is achieved by increasing the intensity dose of the OFF-switching light, the interaction strength of the OFF transition or the steepness of the optical pattern that triggers the OFF transition. Meanwhile, non-perfect zero intensity in the OFF-switching optical pattern reduces the signal also. When the OFF light is also partially triggering the reverse (ON) transition, this induces a background B to the E-PSF.

for increasing illumination time T. Meanwhile, the FWHM of the peak (measured from the background offset) is given by

$$d = \frac{d_c^{ill}}{\sqrt{1 + \frac{{d_c^{ill}}^2}{4\ln(2)}\, f(\lambda_{\mathrm{OFF}})\, I_0 T\, a}}\,, \tag{2.58}$$

where d_c^{ill} denotes the FWHM of the initial distribution (which is for this calculation approximated as a Gaussian). The FWHM d is scaling with the inverse square-root of

- the interaction strength $f(\lambda_{\mathrm{OFF}})$,
- the intensity dose $I_0 T$ (proportional to the pulse energy of the OFF switching pulse),
- and the steepness of the minimum a.

It becomes obvious that this FWHM may become arbitrarily small when increasing the quantities under the square-root. While regularly most of the parameters are rigidly defined by the imaging system and the sample, the maximum intensity I_0 (and sometimes also the illumination time T) are freely tunable. This means, that by increasing the light dose onto the system, the peak region (where emitters are allowed to be in the ON state with a high probability), can be decreased to as small values as desired[6].

Thus, by optically driving an ON to OFF transition with an intensity minimum, an arbitrarily small ON volume may be achieved. In order to obtain a good visibility, the ON to OFF transition at λ_{OFF} should additionally be much stronger than the OFF to ON transition. Scanning such an ON distribution across the sample and reading out the emission (which can only arise from emitters in the ON state) is probing the respective volume for the presence of emitters. The total emitted signal arising from such a measurement is thus

[6]Please note that the definition of d is measuring the FWHM of the second term in eq. (2.56), without worrying about the background term.

Table 2.1: Overview of the state and switching modalities of some super-resolution imaging methods, namely STED [11], RESOLFT [13] and GSD [10].

	STED	RESOLFT	GSD
OFF	Fluorescence ground state	Non-fluorescent chemical form	Non-fluorescent long-lived triplet state
ON	Fluorescence excited state	Fluorescence ground state	Fluorescence ground state
OFF→ON	Fluorescence excitation	Optically triggered chemical switch	Spontaneous decay
ON→OFF	Stimulated emission	Optically triggered chemical switch	Optically triggered transition
Read out	Spontaneous emission	Fluorescence excitation	Fluorescence excitation

giving rise to the emitter density within the ON volume. Assuming a pinhole to suppress any off-center signal, we can adapt the math from the confocal system as in eq. (2.39). Therefore, the image that is emerging, when performing a scan and integrating the signal among a pinhole in each scanning position (in object coordinates) is given by

$$I_i(\mathbf{r}) \propto \Big(\big[(\mathcal{A}_\mathcal{R} * G_{em}) \;\; P_{\mathrm{ON}} \big] * T_o \Big)(\mathbf{r}) \;, \tag{2.59}$$

while T_o denotes the object's emitter density, $\mathcal{A}_\mathcal{R}$ defines the pinhole and G_{em} stands for the emission PSF. This shape is shown in **Fig. 2.8**. The distribution P_{ON} takes on the roll of the illumination PSF in the confocal case. We thus call P_{ON} the effective illumination PSF (illumination E-PSF), while the multiplication with the detection PSF is considered as the E-PSF. Thus, an image with a resolution of

$$d_c^{\mathrm{SRI}} = \frac{d_c^{conf}}{\sqrt{1 + \frac{{d_c^{conf}}^2}{4\ln(2)} \; f(\lambda_{\mathrm{OFF}}) \; I_0 T \; a}} \;, \tag{2.60}$$

is forming[7]. Super-resolution imaging (SRI) was implemented using different OFF switching transitions. The theoretical calculations concerning the E-PSF remain the same, but the implementation differs. A few of the implemented SRI methods and their states and switches are presented in **Tab. 2.1**.

2.4.2 Localization methods

The super-resolution localization methods rely on spatially isolated emitters in the ON state. Isolated means that single emitters in the ON state are well separable from each other on a camera or point detector at each time – their distance should thus be much larger than the diffraction limit. Hence, emitters in the vicinity should be switched OFF at that time. Given this condition, the single isolated emitter (or multiple) may be localized by repeatedly measuring its position. Each position estimate $\mathbf{r}_i$ is localizing the emitter to an uncertainty s_i (in each dimension[8]). From those position estimates, we may consider the mean position

$$\mathbf{r} = \frac{1}{N} \sum_{i=0}^{N} \mathbf{r}_i \;, \tag{2.61}$$

[7] SRI is short for super-resolution imaging.

[8] We assume a constant uncertainty s_i in all spatial dimensions.

as the result of the localization procedure. We denote the number of measurements by N. As derived in appendix A.1, the uncertainty of the mean position estimate is retrieved from the sum of squares of the single measurement uncertainties

$$\sigma^2 = \frac{1}{N^2} \sum_i s_i^2 \ . \tag{2.62}$$

For N measurements of the same precision $s_i = s_j = s$, we find an uncertainty

$$\sigma = \frac{s}{\sqrt{N}} \ . \tag{2.63}$$

This relation shows that the precision of a localization is scaling with the inverse square-root of the number of measurements.

Emission localization

The easiest case of a localization is appearing when observing the emission pattern of an isolated emitter. Each photon may be considered as a measurement whose position is randomly drawn from the distribution of the emission. The uncertainty s of each single photon is thus matching the standard deviation of the emission pattern. For this emission localization (EL), we thus find a resolution analogue

$$d_c^{\mathrm{EL}} = \frac{d_c}{\sqrt{N}} \ , \tag{2.64}$$

dependent on the Abbe resolution limit d_c. The methods PALM, STORM and PAINT implement this kind of localization by observing isolated images of fluorescent emitters on a camera. While the localization technique remains unchanged, these methods use three different switching modalities, in order to ensure the sparsity of active fluorescent emitters. In PALM [3], photo-activatable emitters are used. In their original form, they do not show fluorescence emission upon illumination with the excitation light. Yet, those molecules may be optically switched to the active (fluorescent) form, in which they can be localized. The density of active emitters is thus controlled by the applied optical dose of the switching light (mostly blue or ultraviolet). In STORM [28], stochastically blinking molecules are used. Those feature an active and a non-active form, which feature a given stochastic switching dynamics, based on the chemical environment. The density of active emitters is thus controlled by the surrounding chemical conditions. In DNA-PAINT [14], the emitters are attached to short single-stranded DNA strands (so-called imager strands), which are freely diffusing and may stochastically bind to the structure-bound complementary single-stranded DNA docking strands. The binding dynamics may be controlled by the chemical conditions, the imager strand density and the imager-docking strand affinity, based on the implemented nucleotide sequence. The local time-integrated signal from the bound imager strands is herein much higher than that of the fastly diffusing unbound imager strands, which is mimicking a switching behaviour in the camera images. The PAINT concept in general is not limited to DNA as a binding carrier of the emitter; however, DNA-PAINT is the most widely used version of PAINT.

In addition to the pure emission based localization methods, a method called SIMFLUX was introduced [5]. This approach combines emission based localization with a phase modulating structured illumination. Thus, besides the spatial information from the emission, the signal modulation upon shifting the illumination pattern is yielding additional

information. Effectively, this reduces the purely emission localization uncertainty by a factor of two, just like SIM reduces the widefield resolution by a factor of two. We can thus find the resolution analogue for SIMFLUX to match

$$d_c^{\mathrm{SIMFLUX}} \geq \frac{d_c^{\mathrm{SIM}}}{\sqrt{N}} \ . \tag{2.65}$$

Illumination localization

In addition to the emission localization, where the center of an emission profile in the image plane is estimated, we can also localize an emitter by probing its position with an illumination beam. Modelling the detected signal with the well-known illumination profile allows to determine the emitter's position. One could simply perform a confocal scan on an isolated emitter and compute its position from the measured image. This would result in a resolution similar to eq. (2.64), when replacing d_c with d_c^{conf}. However, the main benefit of the illumination localization over the emission type is that the probing position of the illumination pattern with respect to the emitter's position is an adjustable parameter. In the case of emission localization, the full PSF is inevitably sampled, whereas probing with an illumination pattern allows choosing the sampling positions freely. This may not sound advantageous at first, but the key underlying principle is that measurements at different positions of the illumination PSF feature different information contents. Targeting the measurements at high information regions is therefore reducing the measurement uncertainty at equal photon number.

We assume some E-PSF $P(\mathbf{r})$ (as a result from both the illumination and detection modalities), which shall carry the dimension "number of detected photons", where $\mathbf{r}$ is the position of the E-PSF with respect to the emitter. Assuming Poissonian statistics at the detector, the number of detected photons n is following the conditional probability

$$f\left(n \mid \mathbf{r}\right) = \frac{P(\mathbf{r})^n}{n!}\ e^{-P(\mathbf{r})} \ , \tag{2.66}$$

given the E-PSF sampling at position $\mathbf{r}$. From this conditional probability, we may compute the Fisher information, which is quantifying the information content of the random variable n about the parameter $\mathbf{r}$ [15][9]

$$\mathcal{I}(\mathbf{r}) = \int dn \left(\frac{\partial}{\partial \mathbf{r}}\ \ln\big[f\left(n \mid \mathbf{r}\right)\big]\right)^2 f\left(n \mid \mathbf{r}\right) \ . \tag{2.67}$$

Calculating the derivative and re-ordering leads to

$$\mathcal{I}(\mathbf{r}) = \left(\frac{P'}{P}\right)^2 \int dn\ (n-P)^2\ f\left(n \mid \mathbf{r}\right) \ , \tag{2.68}$$

where $P' = \partial P/\partial \mathbf{r}$. The integral term is the variance of the random variable n under the given condition and is thus matching $P(\mathbf{r})$ in the Poissonian case, leaving us with

$$\mathcal{I}(\mathbf{r}) = \frac{(P')^2}{P} \ . \tag{2.69}$$

The information content per photon is retrieved when dividing this result by P.

$$\mathcal{I}_1(\mathbf{r}) = \left(\frac{P'}{P}\right)^2 \tag{2.70}$$

[9]The notation with respect to the vectorial parameter $\mathbf{r}$ shall be interpreted component-wise.

Measurements at positions with low intensity and high gradient are thus most informative. The fact that precision is inversely scaling with the gradient is rather intuitive, as the signal change for a given parameter change is increased. The inverse scaling in P on the other hand, may be interpreted in a way that the noise arising from the Poissonian detection is minimal when the overall signal is low. Signal differences are thus better pronounced over the noise level, when measuring at low overall signal level. From the (single-photon) Fisher information, we can directly relate to the so-called Cramér-Rao bound (derived in appendix A.2), which defines a lower limit to the uncertainty of the positional measurement

$$s(\mathbf{r}) \geq \frac{1}{\sqrt{\mathcal{I}_1}} \ . \tag{2.71}$$

The scaling of the single-photon Cramér-Rao bound for some exemplary E-PSFs is shown in **Fig. 2.9**.

Not only probing a single position but a sampling pattern $\{\mathbf{r}\}$, the single-photon information has to be averaged over the sampling positions

$$\mathcal{I}_1(\{\mathbf{r}\}) = \left\langle \left(\frac{P'}{P}\right)^2 \right\rangle_{\{\mathbf{r}\}} \ . \tag{2.72}$$

Assuming an n_d-dimensional spatially symmetric E-PSF, and a sampling pattern, consisting of positions that are symmetrically arranged around the emitter position at a distance R, leads in each dimension to a single-photon information which reduces by a factor of n_d

$$\mathcal{I}_1(R) = \frac{(P'(R)/P(R))^2}{n_d} \quad \Rightarrow \quad s(R) \geq \sqrt{n_d}\ \left|P(R)/P'(R)\right| \ . \tag{2.73}$$

Thus, the precision of the illumination localization (IL) scheme after N independent measurements (see eq. (2.63)), assuming the above mentioned symmetric arrangement of sampling positions in n_d dimensions at a constant sampling radius R yields

$$\sigma_c^{IL} \geq \frac{|P(R)/P'(R)|}{\sqrt{N/n_d}} \ . \tag{2.74}$$

There are now two ways of minimizing this uncertainty:

1. For a given E-PSF, the sampling close to a zero position is optimal, as it is visible in **Fig. 2.9a**. This is valid as the degree of P is always larger than that of P', meaning that for $R \to 0$, it holds that $\sigma \to 0$. If due to background, no perfect zero exists, an optimal sampling radius may be found further towards higher intensities as shown in **Fig. 2.9b**.

2. When taking into account that the size of the E-PSF may be tuned, as it is done in the super-resolution imaging techniques, P' may be increased via this procedure, as shown in the right side of **Fig. 2.9**.

When a linear response of the emitter to the illumination light is used for probing, the first minimization procedure applies. Hence, P/P' is minimized by sampling the emitter close to a local minimum (zero) of the illumination pattern. As this method thus leads to a minimization of the photon flux from the sample, it is denoted MINFLUX [2]. On the

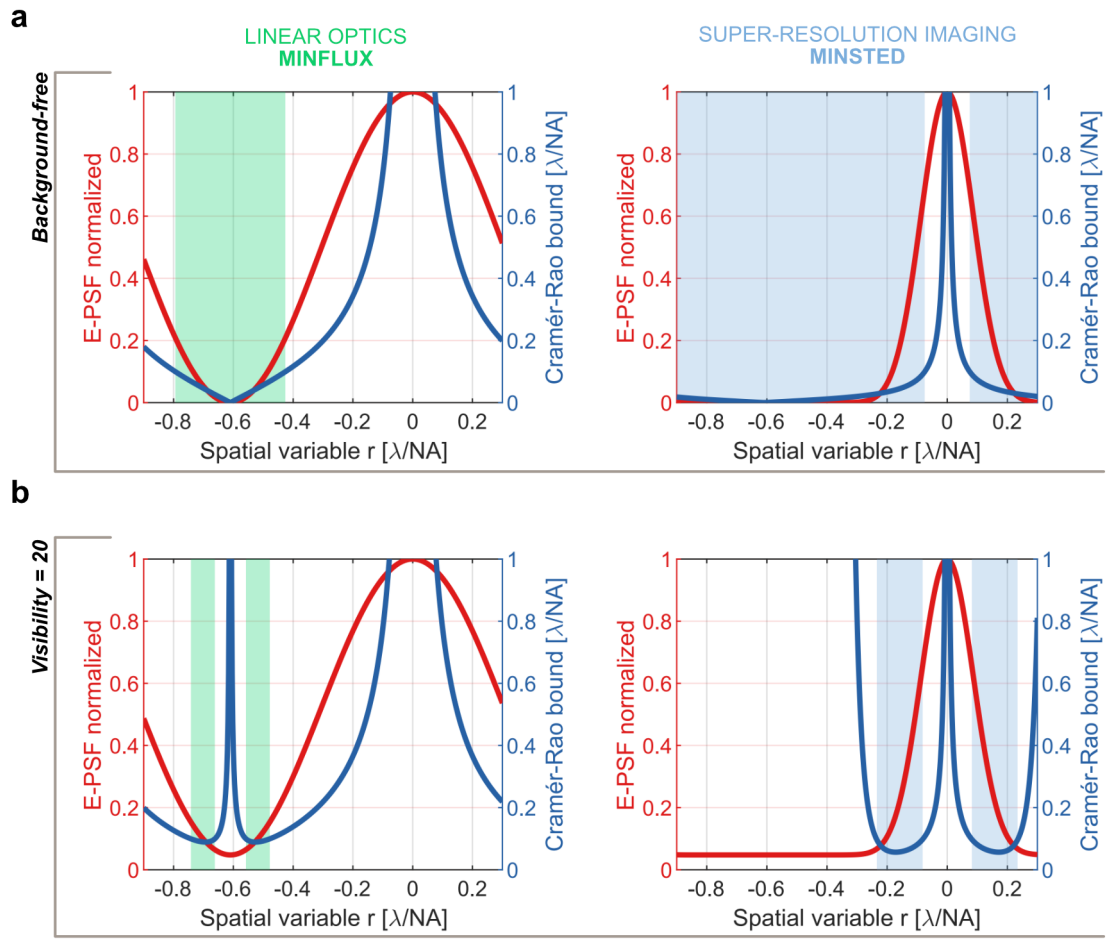

Figure 2.9: Single-photon precision. Photons arising from the interaction of the emitter with different regions of an E-PSF (red lines) carry a variable amount of information. This is related to a corresponding minimal spatial uncertainty, or precision, given by the Cramér-Rao bound (blue lines). **a** In a background-free scenario, optimal precision is achieved at the zero of the E-PSF both for the linear E-PSF (left) as well as the narrowed SRI E-PSF (right). The regions with a Cramér-Rao bound < 0.1 λ/NA are shaded in green and blue, respectively, for the linear and the SRI case. **b** The addition of background leads to the divergence of the single-photon precision at the minima. This shifts the attractive region in the linear case to a narrow band close to the minimum. In the SRI case, the information content is maximized around the steep E-PSF flanks. This figure explains the two tactics in maximizing the single-photon information in the two methods MINFLUX and MINSTED.

other hand, when using STED to shrink the E-PSF to a desired size before optimizing the sampling procedure, we speak of MINSTED [37].

In order to quantify the precision scaling in MINFLUX, we assume a parabolic shape of the minimum, and a constant offset in order to account for possible background contributions

$$P(R) = aR^2 + b \;\Rightarrow\; P'(R) = 2aR \;, \text{ with } \quad b = \frac{aR^2}{\mathrm{SBR}} \;, \tag{2.75}$$

where the signal-to-background ratio (SBR) is defined by the ratio of the signal at the sampling position and the background. This leads to a localization precision of

$$\sigma_c^{\mathrm{MINFLUX}} \geq \sqrt{n_d}\,\frac{R}{2\sqrt{N}}\left(1 + \frac{1}{\mathrm{SBR}}\right) \;. \tag{2.76}$$

It is noteworthy that in contrast to any of the resolution measures from the former sections, the precision in MINFLUX is not depending on any diffraction limited parameter. The scaling is fully controlled by the sampling procedure and thus mainly influenced by the sampling radius R. Reducing the sampling radius is usually limited by the SBR, which is decreasing at reduced signal strength towards the minimum of the E-PSF.

Doing an analogous computation for MINSTED, assuming a Gaussian shaped E-PSF with standard deviation σ_c^{SRI}, we find the localization precision

$$\sigma_c^{\mathrm{MINSTED}} \geq \sqrt{n_d}\, \frac{\left(\sigma_c^{\mathrm{SRI}}\right)^2}{R\sqrt{N}} \left(1 + \frac{1}{\mathrm{SBR}}\right) . \tag{2.77}$$

The precision may thus be enhanced by reducing the E-PSF size or increasing the sampling radius. The latter procedure is again limited by the SBR, which typically decreases for increasing sampling radius R.

As MINFLUX and MINSTED rely on isolated emitters, the same switching methods as mentioned for the emission localization may be implemented here.

2.5 Summary

This chapter told the story of resolution in optical microscopy. Starting with the theoretical background by motivating the theory of Fourier optics, the diffraction limit and its origin were explained. Subsequently, the basics of fluorescence were introduced, before proceeding to an exposition of the theory of image formation in the classic (diffraction limited) microscopy methods. This shows how strongly the resolution limit may be stretched, ending up with the method SIM that gains a factor of two in resolution with respect to the original Abbe limit. From there, the field of super-resolution microscopy is introduced. The precision scaling in super-resolution imaging and in localization microscopy was derived, ending up with the statistical treatment of the illumination localization microscopy methods MINFLUX and MINSTED. The latter two mark a milestone in localization microscopy as both introduce a way of improve localization precision without solely relying on the number of detected fluorescence photons.

This dissertation primarily focuses on MINSTED microscopy. The metrics of MINSTED are mostly analytically assessable and should be readily understandable and motivated given the rigorous theoretic foundation from this chapter.

CHAPTER

THREE

LOCALIZING MOLECULES WITH STED

3.1 Introduction

As outlined in the previous chapter, localization means to find an emitter's position from repeated positional measurements. The underlying spatial sampling can be implemented with a STED E-PSF. This is advantageous, because the localization uncertainty is decreasing for smaller widths of the E-PSF. Hence, in addition to the $1/\sqrt{N}$ scaling of the localization precision with the number of detected photons N (which is a strongly sample-dependent quantity), the E-PSF size serves as a controllable parameter to optimize the precision, without requiring more signal from the sample.

MINSTED is an implementation of a STED localization scheme, proposed and first demonstrated in reference [37]. During the sampling procedure, the STED E-PSF is circled around the current position estimate of the fluorescent emitter. If the circle center position is coinciding with the emitter's position, it is clear that (for a symmetric E-PSF) the probability of detecting a photon is constant along the sampling trajectory. Yet, when the emitter's position is shifted, the probability of detecting a photon is enhanced at trajectory angles in close proximity to the shift direction, while lower on the opposite side. Hence, by recording where on the trajectory a photon was detected, information about the emitter position is obtained. MINSTED uses this information immediately by updating the circle center position with each photon detection towards the direction of the detected photon. This is on average moving the circle center onto the position of the emitter. Thus, each circle center position serves as a position estimate, which is updated with every single photon detection. The elegance in this approach comes with the fact that the localization algorithm is directly implemented into the sampling procedure. Post-processing only needs to calculate the mean position from the recorded center position estimates to obtain the localization result.

The original work [37] was implemented using the standard pulsed STED laser wavelength 775 nm and a 635 nm pulsed excitation. These wavelengths are quite far apart from each other, which implies that the stimulated emission cross-section at 775 nm of most emitters with an efficient absorption at 635 nm has already decayed to a few percent of its maximum value (**Fig. 3.1**). This small cross-section of the stimulated emission process has to be compensated by a higher power of the STED light. The disadvantages of high optical powers are its photo-toxicity to biological specimen, which might restrict the live-cell compatibility of a microscopy method, and the dissipation of heat in the

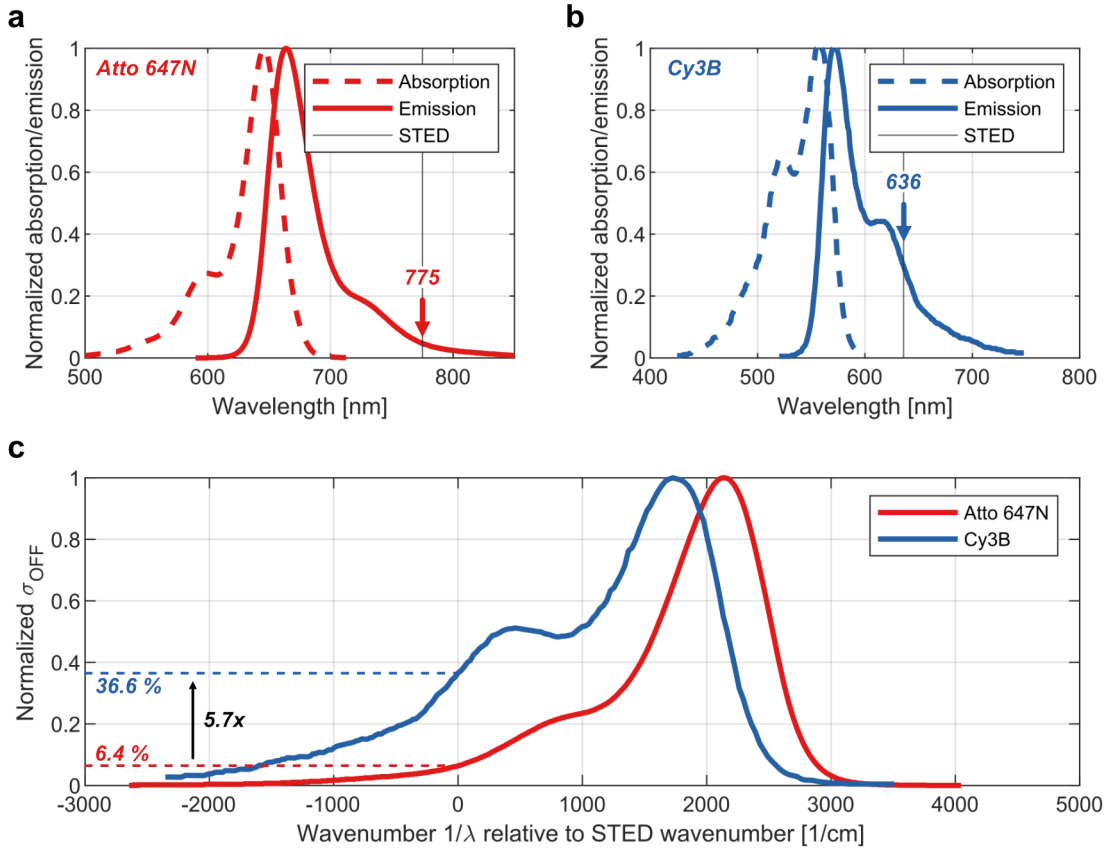

Figure 3.1: Stimulated emission cross-section. The emission (dashed lines) and absorption (solid lines) spectra of the fluorescent emitters Atto 647N (**a**) and Cy3B (**b**) are presented. **c** Normalized stimulated emission cross-section with respect to the wavenumber, calculated as derived in reference [6]. The wavenumbers are shifted to yield zero for the respective STED wavelength 775 nm or 636 nm for Atto 647N (red) or Cy3B (blue), respectively. This leads to a 5.7-fold increase of the normalized stimulated emission cross-section in the blue-shifted case.

sample and the immersion liquid. These drawbacks limit the localization precision of MINSTED: Given a constant number of detected photons and a certain SBR, the way of reducing the localization uncertainty is to further confine the E-PSF, which requires higher STED powers. However, the deposition of heat to the sample and the immersion induces mechanical instability and changes the refractive indices, which worsens the measurement precision. As it turned out in reference [37], the localization precision of MINSTED was therefore limited to about 1 nm in this implementation. A rather straight-forward way to reduce the heat load would have been to reduce the repetition rate of the laser pulses. However, this is directly connected to the available detection rate of the fluorescent emitter and would slow down the measurement by a similar factor as the average power is decreased.

Instead, a blue-shifted STED wavelength can be used to increase the stimulated emission cross-section [35] and therefore allowing for smaller E-PSFs without inducing significant thermal issues. But why is this straight-forward solution not generally applied in STED applications to avoid unnecessary optical load on the sample and the use of costly high power lasers? While blue-shifting the STED wavelength benefits from an increased stimulated emission cross-section (see **Fig. 3.1c**), it also suffers from an increasing excitation cross-section leading to a worsened SBR in the measurement, because the STED beam is no

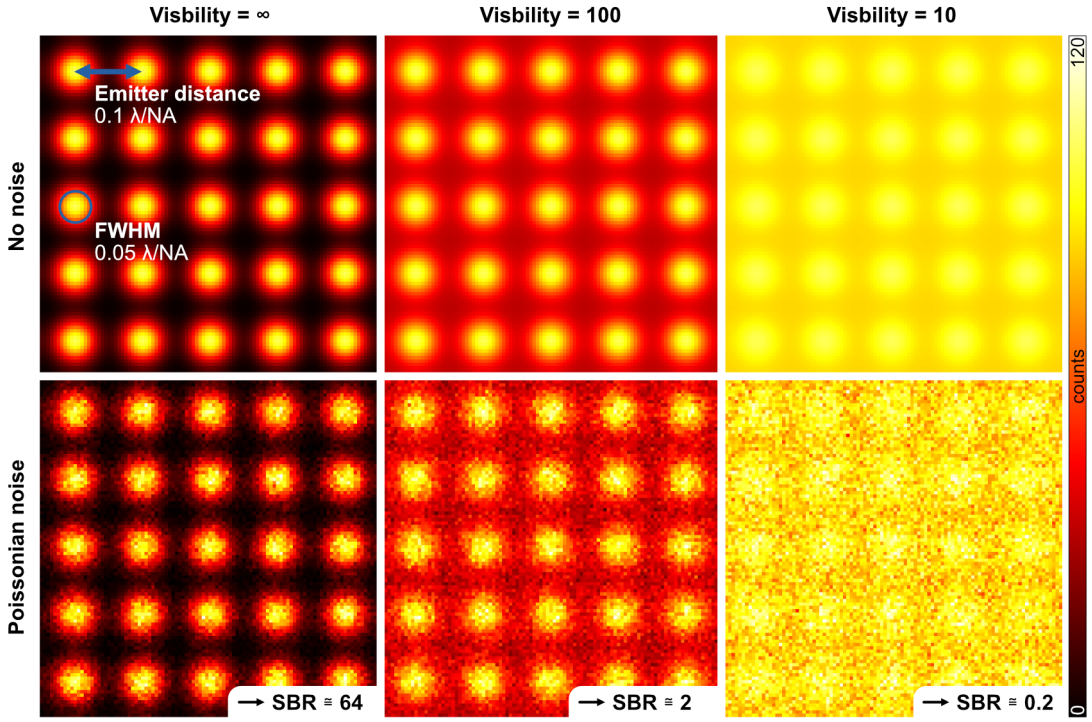

Figure 3.2: SBR in super-resolution imaging. The image of a grid arrangement of point emitters is retrieved by a 2D convolution with the SRI E-PSF. The images are rescaled to each show a maximum number of 100 counts. Random noise is added numerically, assuming a Poissonian distribution. The lattice constant of the object is chosen to 0.1 λ/NA and the FWHM of the E-PSF a factor of two smaller, 0.05 λ/NA. Reducing the visibility of the E-PSF leads to a strong reduction in the resulting SBR. This shows that the E-PSF visibility is a highly sensitive parameter, when imaging dense emitter configurations. This figure was inspired from reference [38].

longer exclusively switching molecules OFF. Residual ON switching creates an E-PSF with a central peak surrounded by a significant background (see **Fig. 2.8**). This has a drastic effect in STED imaging and spectroscopy applications [27, 20], because the highly confined center peak volume of the E-PSF is orders of magnitude smaller than the confocal volume, from which the background is collected. This effect is strongly dependent on the employed FWHM and the emitter density. For example, using a 2D E-PSF, featuring a FWHM of 0.05 λ/NA, and emitters on a square lattice at a distance 0.1 λ/NA, an E-PSF visibility of 100 is already reducing the SBR to 2, as shown in **Fig. 3.2**. This problem is not as severe when localizing isolated emitters as in MINSTED. Here, the background region, no matter how large it may be, only acts on the single fluorescent emitter to be localized. Hence, the SBR (ratio of signalling and background photons) throughout a localization measurement approaches the visibility of the E-PSF. The volumetric extent of the background region does not weigh in if the localization process avoids it. It turned out that by blue-shifting the STED wavelength, the thermal load could be reduced despite the further narrowing of the E-PSF, which allowed MINSTED to reach the sub-nanometre localization precision range [38].

3.2 Theory

Most of the theory relevant for MINSTED is discussed in section 2.4. Nevertheless, for a proper understanding of the blue-shifted MINSTED, important findings will be revisited

and contextualized.

3.2.1 STED

As derived in section 2.4.1, the E-PSF in STED microscopy is found as the multiplication of the probability P_{ON} for an emitter to be in the fluorescent state (the so-called effective illumination PSF) and the probability G_{det} of detecting the emitter's signal (the so-called detection PSF). The effective illumination PSF results from the sequential ON and OFF transitions as defined by the illumination sequence, whereas the detection PSF is due to diffraction and the spatial filtering from the confocal pinhole. We denote the E-PSF as P and find in analogy to eq. (2.59)

$$P(\mathbf{r}) = G_{det}(\mathbf{r})\ P_{\mathrm{ON}}(\mathbf{r}, T)\ . \tag{3.1}$$

For the E-PSF to be directly proportional to P_{ON} at a specific time T, it has to be valid that for times $t \geq T$, the shape of P_{ON} is no longer changing (except for the fluorescence decay) and that the signal arriving before the time T is ignored. Hence, we are talking of a gated detection, which is initiated after all optically triggered state transitions have been completed[1]. The ON probability as given in eq. (2.56) assumes an optical sequence of the following kind:

1. An excitation pulse with a spatial profile $G_{ill}(\mathbf{r})$ that triggers a linear transition to the ON state, meaning that the resulting ON probability is directly proportional to the excitation pattern.

2. A STED pulse of duration T and wavelength λ_{OFF}, inducing both OFF and ON transitions depending on the emitter's cross-sections at the STED wavelength. The shape of the STED pattern close to its central minimum is assumed parabolic with a residual intensity ϵI_0.

This parabolic shape of the STED beam intensity and the ratio of its ON and OFF transition rates results in an ON probability with a central Gaussian peak and a surrounding plateau of constant background B, as given in eq. (2.56). The peak width is controlled by the light dose $I_0 T$ (proportional to the pulse energy) as well as the light—emitter interaction strength $f(\lambda_{\mathrm{OFF}})$ The residual intensity in the minimum is inducing a signal loss in the peak. For convenience, the equation is repeated below.

$$\begin{aligned} P_{\mathrm{ON}}(\mathbf{r}, T) = {} & B(\lambda_{\mathrm{OFF}}) \\ & + [G_{ill}(\mathbf{r}) - B(\lambda_{\mathrm{OFF}})] \exp\left(-f(\lambda_{\mathrm{OFF}}) I_0 T\ \epsilon\right) \exp\left(-f(\lambda_{\mathrm{OFF}})\ I_0 T\ ar^2\right) \end{aligned} \tag{3.2}$$

The detection PSF arises from a convolution of the pinhole transmission function with the emission PSF

$$G_{det}(\mathbf{r}) = \left(\mathcal{A}_{\mathcal{R}} * G_{em}\right)(\mathbf{r})\ . \tag{3.3}$$

The E-PSF's dependence on the STED wavelength is visualized in **Fig. 3.3**. The two main parameters of the E-PSF are its visibility and the FWHM of the E-PSF's central peak. The visibility can be approximated by the ratio of the dynamic range of the ON probability and its background.

[1] If this would not be the case, the E-PSF would have to be written as an integral over altering ON probabilities weighted with the fluorescence decay factor. If e.g. the detection would not be gated, also fluorescence from before the STED illumination would be detected, which is arising from a diffraction-limited ON region. This signal would have to be added to signals arriving at later times, when STED has already reduced the size of the ON region. In the extreme case of continuous wave (cw) STED illumination, the ON region would change permanently – getting sharper for larger times due to the stimulated emission but also less intense due to the ongoing fluorescence decay.

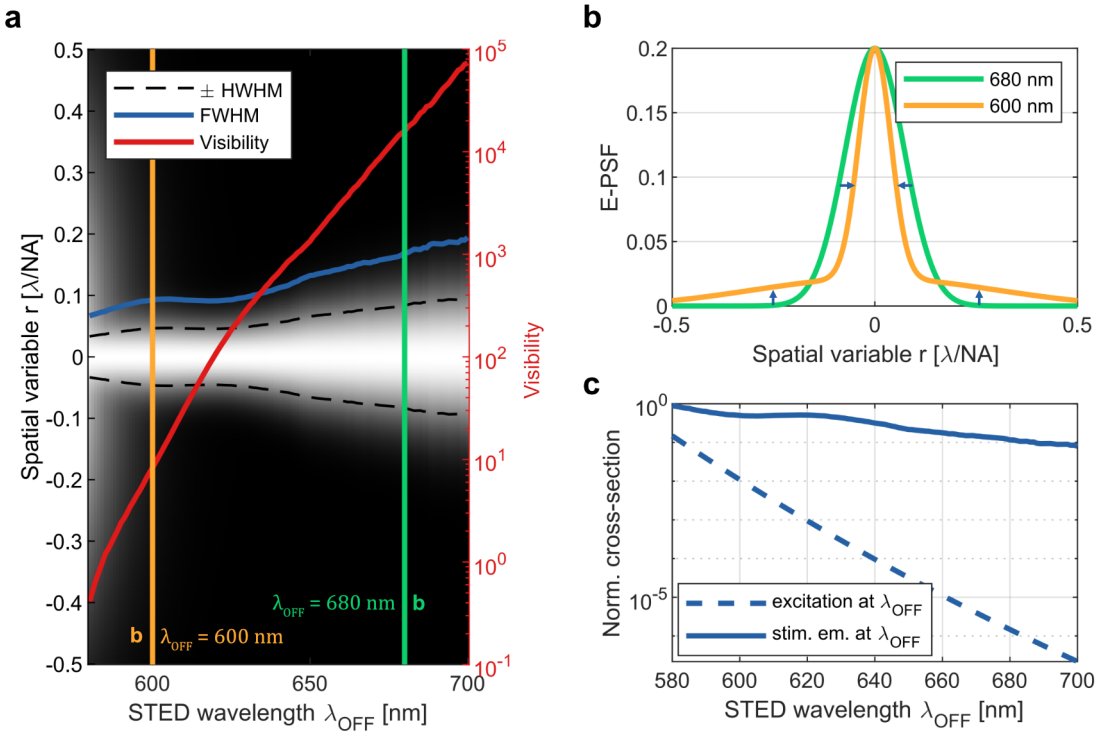

Figure 3.3: STED wavelength influence on the E-PSF. a Analytically estimated E-PSF shape for varying STED wavelength for the fluorescent emitter Cy3B shown as heat map. The E-PSF corresponds to P as defined in eq. (3.1). The ON probability and the detection PSF are calculated as in eqs. (3.2) and (3.3), respectively. The illumination and emission PSFs were both assumed as Airy distributions of wavelength λ with aperture NA. The pinhole diameter was set to 0.5 AU. For the illumination PSF, a maximum excitation probability of 0.2 was assumed, which should still be in the linear regime. For the Gaussian in P_{ON}, $f(\lambda_{\mathrm{OFF}})$ was assumed to be linearly scaling with λ_{OFF} and the sum of the normalized ON and OFF cross-sections (see eq. (2.53)). A reasonable proportionality constant was chosen, which also contained the term $I_o T a$. The residual minimum intensity ϵ was set to zero. The background term B was scaling with the ratio of ON and the sum of ON and OFF cross-sections as in eq. (2.52). The resulting half width at half maximum (HWHM, dashed black line), FWHM (solid blue line) and visibility (solid red line) are plotted on separate axes. Two cuts are displayed at STED wavelengths 600 nm and 680 nm in orange and green, respectively. These refer two exemplary E-PSFs, presented in **b**. **c** Normalized cross-sections for excitation (dashed blue line) and stimulated emission (solid blue line) of Cy3B at wavelength λ_{OFF}. The normalized stimulated emission cross-section is computed from the emission spectrum in analogy to reference [6]. The excitation cross-section is assumed in proportion to the absorption spectrum of Cy3B. On the red edge of the spectrum, the data is continued by modelling an exponential decay (in energy units), which assumes Boltzmann distributed emitters in the vibrational levels of the ground state. In all calculations for this figure, we assume a common maximum cross-section for the ON and the OFF process.

$$\mathcal{V} = \frac{G_{ill}(\mathbf{r}=0) - B(\lambda_{\mathrm{OFF}})}{B(\lambda_{\mathrm{OFF}})} \exp\left(-f(\lambda_{\mathrm{OFF}})\ I_0 T\ \epsilon\right) \tag{3.4}$$

The FWHM of the E-PSF's central peak is

$$d = \frac{d_c^{conf}}{\sqrt{1 + \frac{{d_c^{conf}}^2}{4\ln(2)}\ f(\lambda_{\mathrm{OFF}})\ I_0 T\ a}}\ , \tag{3.5}$$

which we denote d for sake of simplicity. Translating the FWHM to a Gaussian standard

deviation, we find $\sigma_P = d/\sqrt{8\ln(2)}$. The visibility relies on the background level (which is in principle reducible when shifting the STED wavelength further to the red, see **Fig. 3.3a**), on the residual minimum intensity as well as the excitation power, which controls the amplitude of G_{ill}. The E-PSF's FWHM, on the other hand, sharpens for increasing STED intensity as well as an increase in the light—emitter interaction parameter $f(\lambda_{\mathrm{OFF}})$. This usually arises when blue-shifting the STED wavelength (see **Fig. 3.3a**). The choice of a STED wavelength thus balances between a small FWHM at low STED powers and a high visibility.

3.2.2 MINSTED

MINSTED in its 2D implementation is sampling a single fluorescent emitter by encircling it at a sampling radius R with an E-PSF of FWHM d. The circle center after detecting the i^{th} photon shall be denoted $\mathbf{C}_i$. With each photon detection, the center position is shifted by a fraction α of the sampling radius R into the direction, where the E-PSF was located when the photon was registered

$$\mathbf{C}_i = \mathbf{C}_{i-1} + \Delta\mathbf{C}_i \ , \ \text{with} \quad |\Delta\mathbf{C}_i| = \alpha R \ . \tag{3.6}$$

In contrast to earlier assumptions, with this kind of update rule, the resulting position estimates $\mathbf{C}_i$ are no longer independent from each other. The position updates lead to a correlation of consecutive position estimates with a correlation length of about

$$N_c = \frac{\sqrt{2}s}{\alpha R} \ , \tag{3.7}$$

(derivation follows later) with the single-photon spatial uncertainty[2]

$$s = \sqrt{2}\ \frac{\sigma_P^2}{R}\ \left(1 + \frac{1}{\mathrm{SBR}}\right) \ , \tag{3.8}$$

which was abstracted from eq. (2.77). The correlation length is inversely dependent on the stepping fraction – the shorter the steps are chosen, the more correlation is induced to the data. Correlated and thus redundant position estimates suggest an information loss because only N/N_c position estimates may be considered independent. Yet, the correlation also leads to a narrowing of the distribution of position estimates with respect to the single-photon uncertainty. The process of correlation may actually be interpreted as performing a rolling average over N_c consecutive photon detections, leading to a decreased spread of the distribution of center positions. The resulting uncertainty reads

$$s_c = \frac{s}{\sqrt{N_c}} \ , \tag{3.9}$$

while the number of remaining independent bins is given by N/N_c, amounting to a full precision of

$$\begin{aligned} \sigma &= \frac{s_c}{\sqrt{N/N_c + 1}} \\ &= \frac{s}{\sqrt{N + N_c}} \ . \end{aligned} \tag{3.10}$$

The addition in the denominator on the right hand side is assuming that all N center positions are actually distributed with respect to the correlated uncertainty. Starting a

[2]We assume the single-photon spatial uncertainty to match its lower limit, the Cramér-Rao bound.

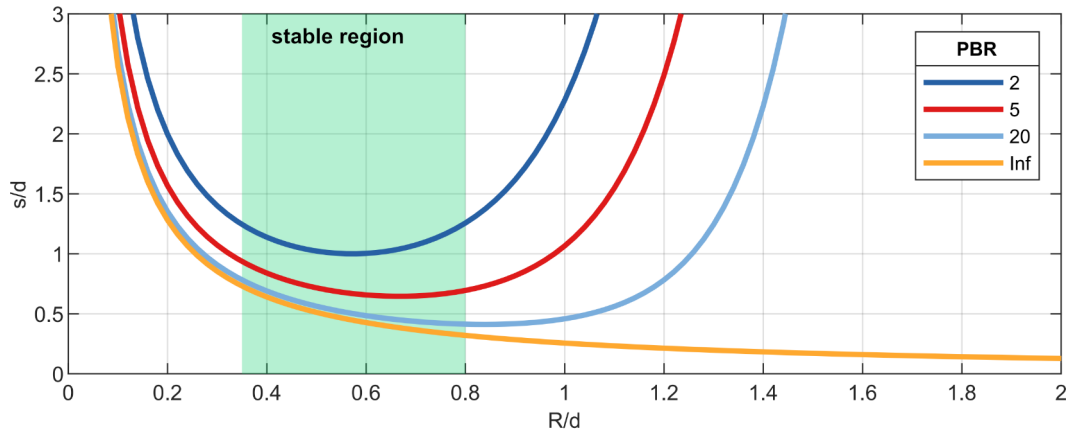

Figure 3.4: Radius influencing the MINSTED single-photon precision. Single-photon precision in MINSTED scaling with the sampling radius for PBR of 2 (blue), 5 (red), 20 (light blue) and infinite PBR (orange). Good precisions for all PBRs are achieved in a region $0.35 < \frac{R}{d} < 0.8$ (green shaded area). This figure was inspired from reference [37].

measurement at $i = 0$, it would take N_c photons until the center positions have converged to this correlated distribution. Yet, discarding the photons prior to convergence, we end up with the precision estimate as stated above. Thus, the overall precision is conserved (except for the additional term in the denominator), which is consistent with the fact that the data is still relying on independent stepping directions.

Looking at eq. (3.8), it seems appealing to choose a large sampling radius R in order for the single-photon uncertainty to decrease. This is not the full truth, as the SBR is no independent variable, but strongly R-dependent, as with increasing radii the signal drops accordingly. Assuming an average sample dependent background B_s and an E-PSF shape as derived previously, we find the radius-dependent SBR (signalling photons divided by background photons) as

$$\mathrm{SBR}(R) \approx \underbrace{\frac{\mathcal{V}}{1 + B_s/B(\lambda_{\mathrm{OFF}})}}_{=\mathrm{PBR}} \exp\left(-0.5\,\frac{R^2}{\sigma_P^2}\right) , \tag{3.11}$$

where we define the first part, the E-PSF's visibility reduced by the sample background, as the PBR (peak-to-background ratio). Given this relation, the scaling of s/d, with respect to R/d is shown in **Fig. 3.4**. Over the whole range of PBR's, choices of $\beta = R/d$ between 0.35 and 0.8 seem reasonable. Since the single-photon precision in this range is rather flat, the exact choice of the radius is not strongly influencing MINSTED's precision.

So far, we have assumed a constant sampling radius and FWHM throughout the localization protocol. This assumes a good initial guess of the emitter position on the order of the size of the radius R, in order to have the emitter sitting inside of the sampling circle. This prior information could e.g. be extracted from a fast STED scan, collecting a few photons from an E-PSF of size d. This is problematic, as most emitter bleaching in STED microscopy is induced by the high intensity regions far away from the zero position. Without prior information, those would also have to be applied to the emitter in order to find it. Instead, MINSTED avoids applying the high intensity STED regions to the emitter by sequentially reducing the FWHM and the sampling radius, starting at a confocal size

$d_0 = d_c^{conf}$ and a respective radius of $R_0 = \beta d_0$. With each detected photon i, the radius and FWHM are reduced to a fraction γ of the former values (usually $\gamma = 0.97$)

$$R_i = \gamma \; R_{i-1} \qquad d_i = \gamma \; d_{i-1} \; . \tag{3.12}$$

Meanwhile, the STED intensity is increased with every photon, according to eq. (3.5), in order to ensure the proper FWHM d_i. The circle center position is updated as above, with respect to the current radius

$$\mathbf{C}_i = \mathbf{C}_{i-1} + \Delta\mathbf{C}_i \; , \text{ with } \quad |\Delta\mathbf{C}_i| = \alpha R_{i-1} \; . \tag{3.13}$$

During this zoom-in process, the single-photon precision is scaling exponentially with the number of photons i

$$s_i \propto \frac{d_c^{conf}\gamma^i}{\beta}\left(1 + \frac{1}{\mathrm{SBR}_i}\right) \; , \tag{3.14}$$

when assuming $\mathrm{SBR}_i = \mathrm{SBR}$. This is not always valid, but for the measurements that were conducted as part of this thesis, the net SBR was increasing for larger STED powers, which is inducing a positive effect on the single-photon precision down-scaling. Thus, by using the information after each photon detection in order to confine the sampling in the next iteration (which again increases the photon information accordingly), an exponentially reducing localization precision may be achieved, which has to be contrasted with the $1/\sqrt{N}$ scaling of the emission localization approaches.

The zoom-in stops as soon as the target parameters R and d are reached. From there on, the localization is continued as described above. The total number of photons observed during both zoom-in and localization phase is denoted M, while we call N the number of photons after the convergence (when the zoom-in stopped). Thus, all of the above mentioned relations, which only deal with the converged part of the trace, are still valid. This kind of zoom-in as well as the subsequent sampling protocol, assure that the emitter is not exposed to high STED intensities, as it is mostly kept at a constant STED intensity region close to the position βd_i of the E-PSF. This "minimization" of STED exposure during the localization process is eponymous for the method MINSTED.

Correlation and dynamics

The number of correlated photons N_c was introduced in eq. (3.7) without any derivation or motivation, so as not to interrupt the line of thought in the above consideration. Next, this gap is filled and the derivation will show a direct connection to the localization of dynamically moving emitters. A similar derivation was already presented in reference [30] and was initially set up by Michael Weber. Let us start, assuming a situation, in which our circle center is displaced by a distance Δ from the actual molecule position in x-direction without loss of generality. We assume the circular MINSTED sampling with radius R, step fraction α and E-PSF standard deviation σ_P, but assume an infinite SBR. The probability of detecting a photon within an angle segment between φ and $\varphi + d\varphi$ is given by

$$p(\varphi)d\varphi = \frac{d\varphi \; \exp\left(-\frac{\left(R\cos(\varphi)+\Delta\right)^2+\left(R\sin(\varphi)\right)^2}{2\sigma_P^2}\right)}{\int_0^{2\pi} d\varphi \; \exp\left(-\frac{\left(R\cos(\varphi)+\Delta\right)^2+\left(R\sin(\varphi)\right)^2}{2\sigma_P^2}\right)} \; , \tag{3.15}$$

which is the value of the normalized E-PSF in the respective direction, when assuming a Gaussian shape of the E-PSF. A photon detection in this direction would lead to an x-displacement of

$$\alpha R \cos(\varphi) \ . \tag{3.16}$$

The mean single-detection step length $\delta(\Delta)$ in x-direction is thus given by integrating over all possible angles

$$\delta(\Delta) = \int_0^{2\pi} d\varphi \ \alpha R \cos(\varphi) \ p(\varphi) \ . \tag{3.17}$$

This step length leads to an average change of the positional offset per time

$$\dot{\Delta} = k\delta(\Delta) \ , \tag{3.18}$$

with k denoting the average photon detection rate. This is easy to visualize: at zero offset, when the sampling circle is located on the molecule, the average x-displacement is zero – all directions are equally probable. The more the sample circle is displaced, the more likely steps in the opposite direction are, leading to an increase in the average single-photon step length in negative x-direction. Thus, solving the dependence of the average single-photon step length δ with respect to the offset Δ allows to extract the time-dependent trajectory $\Delta(t)$. Evaluating the arguments of the Gaussians from eq. (3.15) and cancellation results in the step length taking the form

$$\delta(\Delta) = \alpha R \ \frac{\int d\varphi \cos(\varphi) \exp\left(-\frac{\Delta R}{\sigma_P^2} \cos(\varphi)\right)}{\int d\varphi \exp\left(-\frac{\Delta R}{\sigma_P^2} \cos(\varphi)\right)} \ . \tag{3.19}$$

Those integrals are solvable in terms of the modified Bessel functions of the first kind I_i of orders $i = 0$ and $i = 1$

$$\delta(\Delta) = -\alpha R \ \frac{I_1\left(\frac{\Delta R}{\sigma_P^2}\right)}{I_0\left(\frac{\Delta R}{\sigma_P^2}\right)} \ . \tag{3.20}$$

The quotient of these two indefinite functions is shown in **Fig. 3.5**. In the range of $|\Delta| \lessapprox R$, this is well approximated by the linear relation

$$\delta(\Delta) \approx -\alpha R \ \frac{\Delta R}{2\sigma_P^2} = -\frac{\alpha R}{\sqrt{2}s} \ \Delta \ , \tag{3.21}$$

given the single-photon (background-free) spatial uncertainty s. We thus find the solution to eq. (3.18) as an exponential

$$\Delta(t) = \Delta_0 \exp(-t/\tau) \ , \tag{3.22}$$

with a decay constant

$$\tau = \frac{\sqrt{2}s}{\alpha R k} \ . \tag{3.23}$$

It thus takes the time τ for the offset to be decayed to $1/e$ of its original value.

Let us interpret this result: we imagine to prepare an ensemble of measurements which start at an offset Δ_0 at time $t = 0$. For a fully random process, all subsequent positions would be fully independent of Δ_0, thus after $t = 0$, the average offset of the ensemble would instantly vanish. This is different for the MINSTED sampling. The mean offset and thus

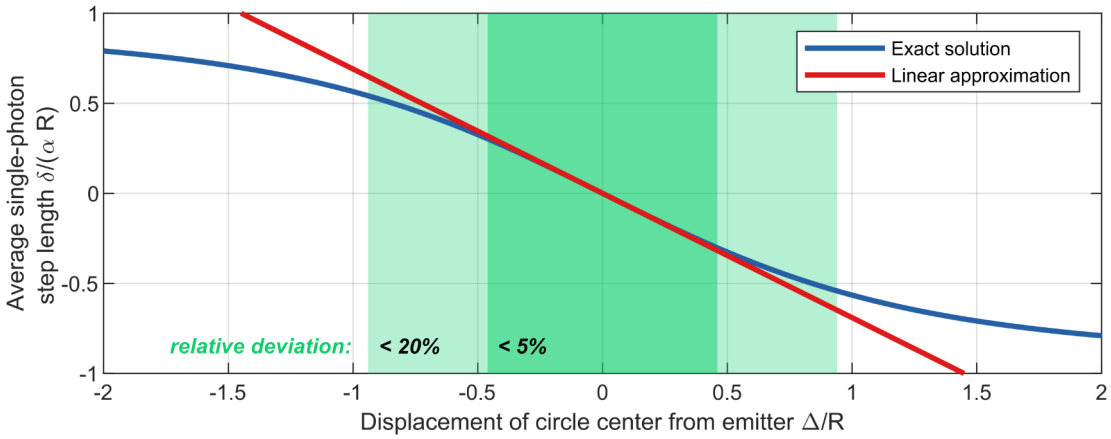

Figure 3.5: Mean dynamics in MINSTED. Average single-photon step length scaling with the displacement of the sampling circle from the emitter. The exact solution (blue line) is following eq. (3.20) while the red line represents the respective first order approximation. The quantity $\delta/(\alpha R)$ may be interpreted as the probability for each photon detection to trigger a motion of the sampling circle into the "right" direction (minimizing the displacement). For zero displacement, there is no preferred direction of the dynamics. As the absolute displacement increases, the average step length in the reverse direction approaches a value of αR. The relative deviation of the linear approximation remains tolerable in a region $|\Delta| \lessapprox R$.

the correlation strength is following the above form $\Delta(t)$ of an exponential decay, reaching Δ/e after a time τ. During this time, a mean number of

$$N_c = k\tau = \frac{\sqrt{2}s}{\alpha R} \tag{3.24}$$

photons is detected. This value is thus quantifying the number of correlated photons in a MINSTED localization measurement.

Reducing a systematic spatial offset between the sampling centre and the emitter is precisely the task when tracking dynamic emitters. We can thus interpret eq. (3.22) as the mean response function of the system to an instantaneous displacement Δ_0. The decay constant τ is therefore quantifying the temporal precision. Yet, the above derived relations are only valid for spatial offsets of about $|\Delta| \lessapprox R$, larger offsets will respond slower.

3.3 Implementation

3.3.1 Optics and electronics

The optical setup of the blue-shifted MINSTED microscope was planned and built by myself. Yet, the optical and electronic design was similar to the prior work from Michael Weber and Marcel Leutenegger [37]. Most of the electronics was in-stock material, some in-house designed custom parts from Udo Gemm, Frank Meyer and Volker Westphal. Those parts could easily be utilized for the application.

The blue-shifted MINSTED [38] as schematically shown in **Fig. 3.6** was implemented as a beam-scanning confocal microscope with two parallel illumination paths:

1. The **localization path** is implementing the scanning via electro optic deflector (EOD) crystals to perform the fast circular sampling motion of the co-aligned excitation and

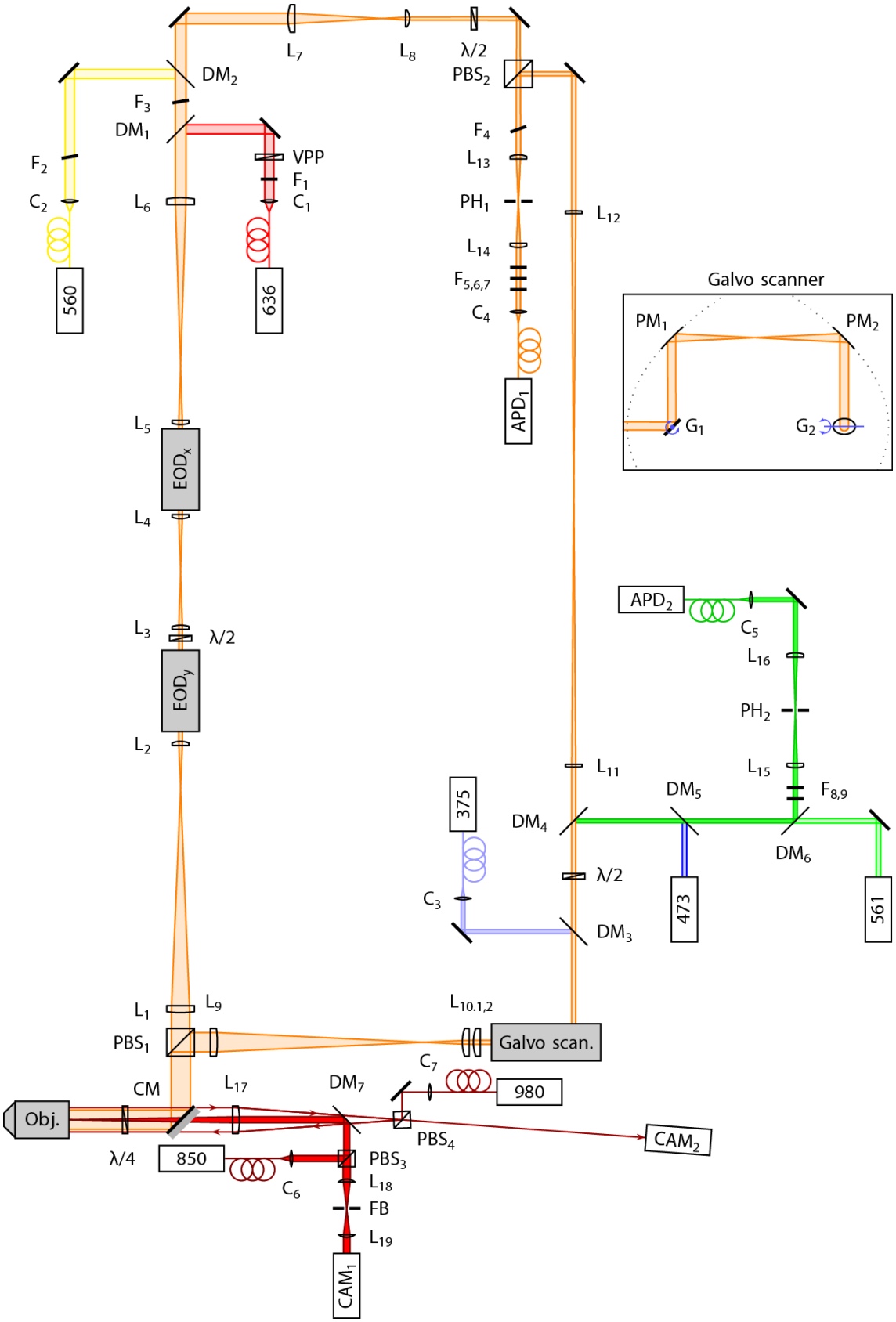

Figure 3.6: Schematic of the blue-shifted MINSTED optical setup. This figure was taken from reference [38].

STED beams. The circling frequency is 125 kHz (full circle within 8 µs). This beam path only supports a scanning range of approx. $2 \times 2\,\mu m^2$ (in sample coordinates), which is defining the field of view (FOV) of the MINSTED localization measurements.

2. The **overview path** is equipped with a galvanometric (galvo) scanner, in order to perform confocal overview scans with a minimum pixel dwell time of 5 µs and a maximum FOV of approx. $80 \times 80\,\mu m^2$. During a localization measurement, the galvo system is not following the quick circling motion of the localization path, but is instead placed at the circle center position, which updates after each photon detection.

Before entering the objective, the two paths are merged via a polarizing beam splitter (PBS). The PBS also splits the fluorescence emission that is captured by the objective and transmits the optical system in the reverse direction. The fluorescence in both paths is isolated via long-pass dichroic mirrors (DMs) and the two fluorescence polarization components are merged again by a second PBS before entering a shared detection unit. To actively stabilize the sample position during a measurement, a near infrared (NIR) illumination and observation system is implemented. In the following paragraphs, the sub-parts of the microscope are described. To ensure proper readability, the equipment and the manufacturers will not be fully specified in the text. Instead, all the components are listed in appendix B.1.

Localization path

The localization path transmits the STED (636 nm) and the excitation beam (560 nm). The excitation enters the system from a single mode polarization maintaining fibre (PM fibre). After collimation and a spectral clean-up, the excitation is coupled to the path via reflection from the long-pass DM_2. The STED light also enters the system via a PM fibre, is collimated and spectrally filtered before transmitting a vortex phase plate (VPP). The VPP features a spatial dependence of its substrate thickness in the lateral plane. Upon transmission, the element is therefore manipulating the phase profile of the illumination light. The pattern of the VPP is shaped like a spiral staircase: Around its center, it features angular regions of an linearly increasing height. After a full revolution, the maximum induced total difference in optical path length is equal to one wavelength (636 nm) – a phase shift of 2π. This design assures that for every lateral position in the light field, there exists a counterpart on the opposite side with a relative phase shift of π. Thus, upon focussing, these opposing beams interfere destructively, creating a zero point in the center of the diffraction pattern of the VPP. Choosing proper polarization, this arrangement produces a donut shaped beam with a perfect central zero in all polarization axes at the focal position of the objective lens. The STED beam is reflected into the main path with the long-pass DM_1. The merged excitation and STED beams are then demagnified by a factor of 3 with a 4-f system consisting of lenses L_6 and L_5. At this position, both beams are s-polarized and enter the x-EOD, which is performing the deflection in x direction. By means of a quadrupole electric field, the EOD creates a linear gradient of the refractive index within its crystalline material along the x-axis. This bends the beam in x-direction, guiding it to a circular trajectory within the crystal. The EOD is placed to match its effective scanning plane[3] with the focal plane of the preceding lens. The next relay (1:1 4-f system of lenses L_4 and L_3) images this effective scanning plane to the effective scanning

[3]The effective scanning plane is an axial position at which the input and output beams (when geometrically elongated) intersect. This position should optimally be independent of the scanning angle.

plane of the y-EOD. Before entering the y-EOD, the polarization is flipped to p using an achromatic (AC) half wave plate. The following relay of lenses L_2 and L_1 expands the beam by a factor of 5, before being merged with the overview path via PBS_1. This arrangement of 4-f systems is imaging the VPP phase pattern to the pupil of the objective.

Overview path

The overview path guides three beams of wavelengths 375 nm, 473 nm and 561 nm, which are combined using DM_3 and DM_5. The near ultraviolet (NUV) light (375 nm) is merged with the main path after exiting from a PM fibre and being collimated by C_3. The 473 nm light is not fibre coupled but enters the system straight from the laser output. The 561 nm laser output enters a manual power modulation unit, consisting of a half wave retarder and a Glan Thompson prism, and is thereafter guided to a 1:1 4-f system of two 50 mm lenses with a pinhole in the common focus for a mode clean-up. Besides merging the illumination beams, two detection paths are separated from the main path:

1. Yellow detection light above 561 nm is coupled out by the long-pass DM_4, and
2. green detection light between 473 nm and 561 nm is reflected from the short-pass DM_6.

The combined illumination beams enter the galvo scanner p-polarized. The scanner contains two galvo mirrors (one for the x-axis scanning and one for the y-axis), which are located in conjugated planes, connected by a 4-f system consisting of two off-axis parabolic mirrors. The arrangement is inspired by reference [24] and was chosen to reduce transmission losses (especially of the 375 nm light) when compared to a lens-based scanning system. The following relay of scanning and tube lens projects the image from the y-galvo to the objective pupil and magnifies it by an approximate factor of 5.3. The scanning lens was made of two AC doublets ($L_{10.1}$ and $L_{10.2}$) facing the galvo system with their planar surfaces. This arrangement reduces aberrations at large scanning angles. After collimation by the tube lens L_9, the overview path is merged with the localization path via PBS_1. At the PBS, the illumination from the overview path is s-polarized, which is arising from the geometric transform that the galvo scanner performs.

Light sources

Excitation The excitation light source is a pulsed white light laser with a pulse duration of approx. 200 ps, operated at 10 MHz pulse repetition rate. Using an acousto optic tunable filter (AOTF), a spectral band around 560 nm is picked, which can also be power modulated, before being fibre coupled to the system.

STED The STED light is emitted from an electrically pulsed single-mode laser diode (triggered from the excitation light source) at 636 nm wavelength with a pulse duration of approx. 1.4 ns. The red light is then coupled into a Praseodymium 3+ doped fluoride glass fibre. From the other side, this fibre is pumped with 450 nm cw light from two cross-polarized single-mode laser diodes, using a long-pass DM. The Praseodymium ions have a sharp emission line close to 636 nm wavelength, thus upon transmission through the fibre, stimulated emission into the red laser mode occurs, which amplifies the STED pulses. The red light is thereafter sent to an electro-optic modulator (EOM), to modulate the optical power and is then fibre coupled to the system. The pulse driver for the 636 nm laser diode was modified from a diode driver evaluation board. The circuit design was done by Michael Weber.

Overview The overview lasers, 473 nm and 561 nm, are both single mode cw diode pumped solid state (DPSS) lasers.

NUV The 375 nm is a single mode cw diode laser with an internal power modulation.

NIR The NIR light sources used for the sample stabilization are

1. a pigtailed cw superluminescent light emitting diode (SLED) with a broad spectrum, centered at 850 nm, and

2. a 980 nm emitting cw diode laser, which is additionally power stabilized by a laser power control unit, which measures the power via a photodiode and stabilizes this quantity by coupling to an upstream liquid crystal power modulator via a feedback loop. Thereafter, the light is fibre coupled to the system.

Sample unit

The sample is mounted on a three axis piezoelectric flexure stage, which is controlling the movement on the order of 1 nm positional accuracy. In order to enable for a coarse positioning with a range on the order of 100 µm, this fine stage is stacked onto a two axis piezoelectric stick-slip stage (lateral) and an underlevered single axis spindle stepper motor (axial). The sample is observed with a Leica 100× 1.4 NA oil-immersion objective. Prior to the objective, a visible AC quarter wave plate is circularizing the illumination polarization. This is crucial for the creation of the donut-shaped STED illumination pattern.

Detection unit

The central part of the major detection unit (yellow) is the confocal pinhole PH_1 of 50 µm diameter, which corresponds to 0.5 AU. The pinhole is located in the shared focus of two lenses, L_{13} and L_{14}. After the collimating lens, the detection light is spectrally filtered by $F_{5,6,7}$ and then coupled to a multi-mode optical fibre, connected to an avalanche photodiode (APD). This part of the detection unit is located in the detection box, which is doubly shielded to suppress any room light. Outside of the detection box, a tilted tunable short-pass filter is cutting most of the residual STED light. Prior to this element, the two detection paths (originating from the localization and the overview path) are combined with PBS_2. The overview fluorescence light is imaged to the detection unit by a relay of two 500 mm lenses (L_{11} and L_{12}), while the localization detection path introduces another demagnification by a factor of two by a 4-f system of lenses L_7 and L_8, to approximately match the effective magnification from sample to pinhole of the overview path. The latter two relay systems are not placed in a clean 4-f configuration with respect to the pinhole relay and the scanning systems. This is not a problem because the beams are already de-scanned at this point. For time gating of the detector counts, the path lengths of the localization path and the overview path were equilibrated to a temporal mismatch of about 10 ps. The time gating was performed with in-house developed electronics, based on the design of Udo Gemm and Volker Westphal.

The secondary detection unit (green) is using the spectral filters F_8 and F_9 and focusses the fluorescence light to PH_2 with an effective pinhole size of 0.9 AU and is then also coupled to a multi-mode fibre and sent to the APD.

Drift correction unit

The drift correction unit is divided into the lateral and the axial stabilization system. The lateral stabilization is implemented via a widefield illumination of the 850 nm SLED, which is focussed to the pupil of the objective by the NIR tube lens L_{17} and coupled to the sample unit by transmitting the cold mirror CM. Back scattered light from gold or silver fiducial markers at the sample surface is imaged by the objective and the tube lens and an additional 1:1 relay of lenses L_{18} and L_{19} to the camera CAM_1. In the conjugated pupil, between L_{18} and L_{19}, the back-reflections of the illumination are filtered by a field block FB with a central obstacle. This illumination filtering sets up a dark-field contrast, which enables a clear localization of the fiducials in the camera image. The fiducial position estimates are used to compensate lateral sample movement by a feedback loop coupled to the fine stage. The axial stabilization system is implemented by focussing the 980 nm light off the center of the objective pupil, which leads to a tilted, collimated beam that illuminates the sample. Part of this light is reflected from the coverslip-sample interface and again collected by the objective. For the coverslip perfectly positioned in the focal plane of the objective, the reflex will be focussed to the opposite side of the pupil, in perfect symmetry with the incoming beam. Axially displacing the coverslip surface is then leading to a shift of the reflex focus angle in the pupil, which is encoding the axial sampling position. This angular shift is observed as a positional displacement on CAM_2, which is located behind the tube lens, far away from the conjugated sample plane. Again, keeping the reflex position on the camera constant by coupling to the sample fine stage via a feedback-loop is stabilizing the axial sample position.

This stabilization procedure is relying on the fact, that observed displacements on CAM_1 or CAM_2 actually arise from sample movement and not from variations within the NIR unit. High stability and shielding of the optics must thus be ensured. Same must be true, especially for the localization path. When instabilities occur in here, the probed sample positions are varying, even if the sample is perfectly stabilized. Since we are aiming for localization precisions on the order of 1 nm and below, the active sample stabilization as well as the passive mechanical stability, shielding from airflow and reduction of vibration disturbances form a highly relevant integral part of the MINSTED microscope.

3.3.2 Measurement control software

Most of the measurement control was carried over from the previous project – the 775 nm MINSTED [37]. The majority of software adaptations that had to be done were implemented by Marcel Leutenegger, who also created the original measurement control. The central part of the measurement control is implemented on a field programmable gate array (FPGA). This device can receive and generate analog and digital voltage signals with well-defined short latencies from tens of nanoseconds to microseconds. An FPGA consists of logic units which enable the implementation of calculations based on the input, output and other registers that are then carried out fully electrically, by transmitting the respective signals through the prepared array. Additionally, this device can receive and emit commands from and to the computer. Besides the laser power control, scanning position control and other signal outputs, the FPGA is executing the MINSTED sampling algorithm. Herein, the scanning motion has to be adapted upon localization with each detected photon. This requires a fast live-processing capability of the control system. The FPGA enables a fully hardware implemented real-time processing of the detector counts that directly feeds to the scanning and laser power outputs, without having to send the signals to the computer, performing the analysis there and responding with the necessary

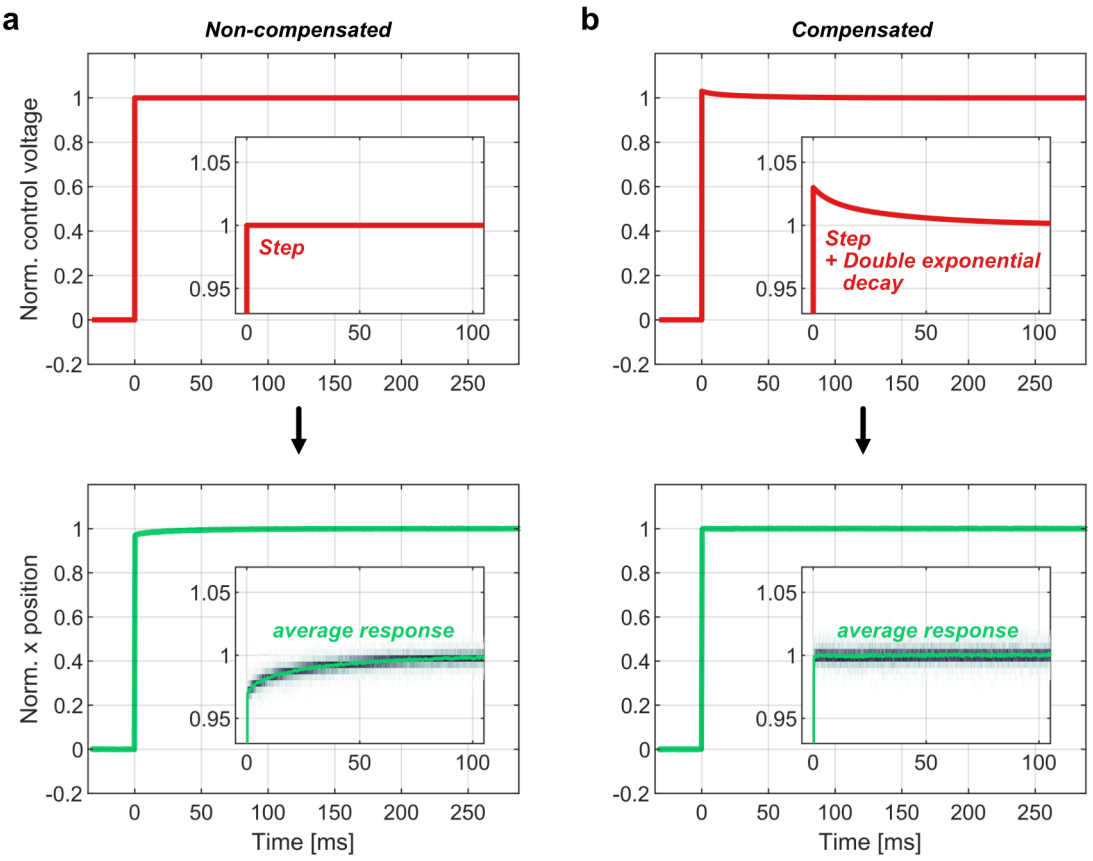

Figure 3.7: Compensation of creeping EOD response. **a** Without compensation, the response to a step-like control voltage change (red) of the x-EOD is an instantaneous deflection to approx. 97 % of the final step length. A creep towards the final position is observed via a position sensitive detector and is well described by a bi-exponential decay. The heat map arises from overlaying multiple measurements of step responses at various step sizes. The average trend is marked in green. **b** Pre-compensating the control voltage by adding a bi-exponential filter to the step is resulting in a clean step in x direction with a residual settling error of less than 1 ‰. This material was adapted from reference [21].

commands.

The control software front-end is implemented in LabVIEW 2017 and MATLAB R2018b. From LabVIEW, the communication with the FPGA is realized. When scanning or localization measurements are initiated, the FPGA takes the full control and executes the measurement independently.

A process that runs in parallel and independently of any measurement routines is the sample stabilization. Herein, the images from the stabilization cameras, which encode both the axial and lateral position of the sample, are analysed and the estimated movements are actively corrected. In the existing implementation, the sample-position-dependent intensity distributions are fitted each with a Gaussian model in order to find their displacement with respect to a reference position. Given the absolute value and the course of the displacement, the sample position is corrected in a proportional–integral–derivative (PID) control loop. The single Gaussian fit per frame for the axial stabilization protocol is performed on the CPU. Meanwhile, the multiple Gaussian models of the selected fiducial images for

the lateral stabilization are executed on the GPU. The fiducial peaks are either selected manually or by a local maxima identification algorithm. A region of interest (ROI) of about 16×16 pixels around each of the fiducials is used for the modelling, leading to an image stack, which is transferred to the GPU, where the fits are performed. Observing the stability of the fiducials over the first 16 frames leads to a trust-score (uncertainty) for each of them. Unstable fiducials are excluded. Thereafter, the weighted mean position of the fiducial images is stabilized by the PID control. The actively stabilized positional uncertainty of all axes was measured to below 1 nm.

From early localization measurements, a creeping of the EOD deflection angle towards its final value after setting the control voltage was observed. Right after applying the voltage, a nearly instantaneous jump of the EOD towards 97 % of the desired step-length was observed (see **Fig. 3.7a**). Thereafter, this initial bias of -3 % reduced to about -1 ‰ during the first 100 ms. The control electronics were ruled out as the cause of the problem. Rather, a material-dependent settling of charge carriers was suspected. A first approach towards this problem was to induce a waiting time of 100 ms in advance to every localization measurement, for the EODs to converge. Yet, as the duration of a localization was itself on the order of 100 ms, this was leading to a significant decrease in measurement speed. A solution to this was found by pre-compensating the positional error by modulating the control voltage opposed the creeping nature of the EODs. This was realized by applying a double exponential filter to the control voltages, which was implemented by Marcel Leutenegger directly on the FPGA. This reduced the error to a to about 1 ‰ of the step length, right after the step, as shown in **Fig. 3.7b**. A detailed description is given in reference [21].

3.4 Results and discussion

The blue-shifted MINSTED microscope was part of five articles, which dealt with different aspects of the method:

1. **Leutenegger, et al. arXiv, 2021** is a technical study on the compensation of the creeping EOD response [21].

2. **Weber, von der Emde, et al. Nature Biotechnology, 2023** presents the blue-shifted MINSTED microscope, quantifies its precision and shows its applicability at one of the standard biological structures for super-resolution microscopy – the nuclear pore complex [38].

3. **Upmanyu, et al. Neuron, 2022** shows the application of blue-shifted MINSTED in neurosciences and presents two-colour MINSTED data, multiplexed using the stochastic switching method DNA-PAINT [33].

4. **Orange kedem, et al. Light: Science & Applications, 2023** presents blue-shifted MINSTED data, which was collected using a new type of VPP to create the donut [25].

5. **Scheiderer, von der Emde, et al. Nature Methods, 2024** proves the capability of MINSTED to be utilized for the tracking of dynamically moving emitters. Its spatio-temporal precision is quantified and the tracking of the motor protein kinesin-1 is presented [30].

Besides the technical publication (1.), which was mentioned above, as part of the methods section, the articles (2.), (3.) and (5.) represent important milestones in the development of MINSTED microscopy, and will be discussed in detail below. All reported measurements, the sample preparation and the data analysis was done by myself, unless stated differently.

3.4.1 Imaging

A smaller E-PSF with less power

MINSTED imaging[4] was first shown in reference [37], using a 775 nm STED wavelength. In the further course of the text, we denote this implementation as 775-MINSTED. In reference [37], the localization of single isolated Atto 647N fluorophores and the caged silicone rhodamine labelled mitochondrial protein Mic60 was presented. The localization precisions were reported to remain above 1 nm, mainly due to STED light induced heating, which caused instabilities during the localization procedure and affected the sample stabilization.

The blue-shifted MINSTED (636-MINSTED) alleviates any heating-induced issues. Comparing the STED E-PSF scaling of the 775-MINSTED, measured on Atto 647N fluorophores with the scaling in 636-MINSTED, using Cy3B emitters, a comparable FWHM was achieved with a factor of ten lower pulse energy (see **Fig. 3.8a**). As the 775-MINSTED applied a pulse repetition rate of 40 MHz, 636-MINSTED at 10 MHz pulse rate required 40 times lower average power for an equivalent E-PSF size. The smallest FWHM $d = 24\,\text{nm}$, which was used in 636-MINSTED imaging measurements, was achieved with an average STED power of 10 mW (**Fig. 3.8b**). These values have to be contrasted with a 775-MINSTED E-PSF of a 60 % larger FWHM $d = 40\,\text{nm}$, for which the 10-fold average power 100 mW was needed.

For these E-PSF measurements, isolated fluorescent Cy3B emitters were immobilized on a coverslip, using the protocol described in reference [38]. First, upon excitation with the cw 561 nm laser, an overview image in a FOV of about $10 \times 10\,\mu\text{m}^2$ was recorded. An automated analysis found the local maxima at which emitters were suspected. The automated measurement procedure then iterated through the identified emitter positions, using the fine three axis piezo stage. At each position, a detail confocal scan with the localization path excitation (STED power $P = 0$) in a FOV of approx. $1 \times 1\,\mu\text{m}^2$ was executed. A Gaussian fit on the resulting image was used for re-centering the FOV for the next step, by offsetting the EODs. After setting the STED power to the desired value, a stack of (around 40) image scans was performed at low pixel dwell time ($\leq 10\,\mu\text{s}$) on a FOV, which was only slightly larger than the expected E-PSF. This choice of a small FOV is inspired from reference [9] in order to keep the STED light exposure to the molecule as low as possible. Summing over the image stack (neglecting images after photo-bleaching of the emitter), re-centering the cumulated image and overlaying the data from many emitters resulted in rather clean E-PSF measurements for different STED powers. Fitting a symmetric 2D Gaussian to each of the E-PSFs led to the FWHMs as shown in **Fig. 3.8a**.

The contrast scaling

In the single emitter measurements, the background remained rather low, resulting in a visibility around > 100 (see **Fig. 3.8b**). The STED-induced background arising from the

[4]Although MINSTED is a localization method, we use the term "imaging" here to distinguish the recording of stationary emitters (imaging) from the recording of moving emitters (tracking).

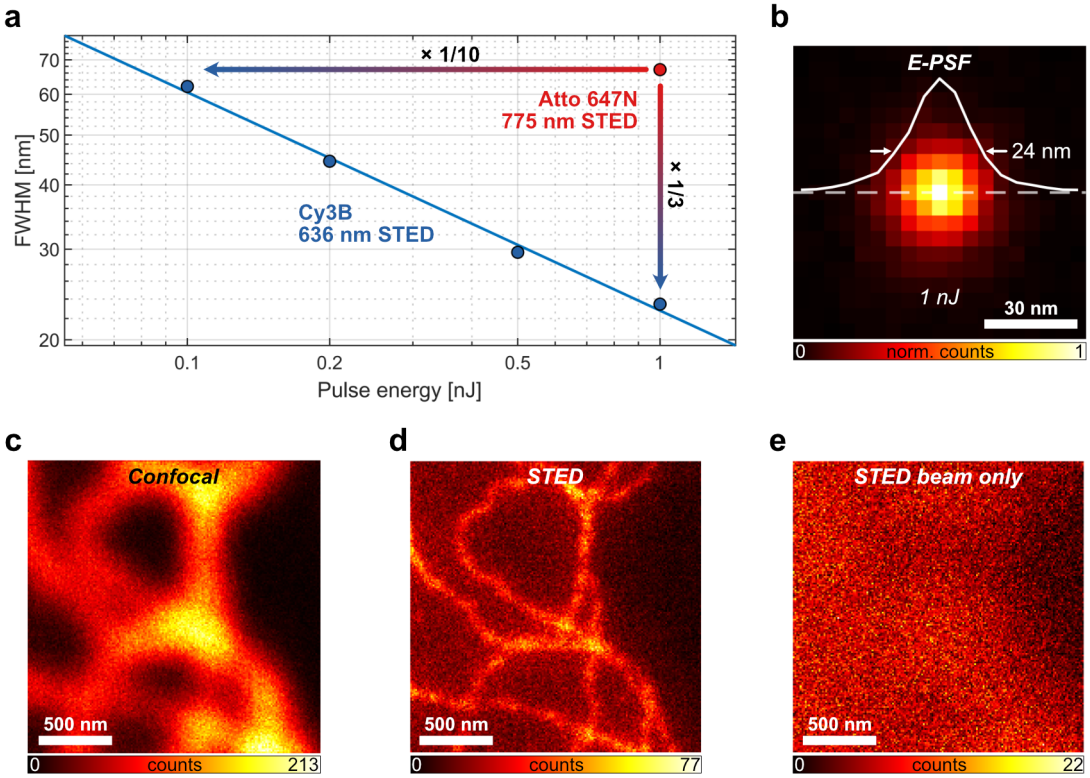

Figure 3.8: E-PSF scaling for a blue-shifted STED wavelength. **a** E-PSF scaling measured from STED imaging of isolated Cy3B emitters, upon pulsed excitation with 560 nm light at approx. 1 µW power, followed by the de-exciting STED pulse at wavelength 636 nm at pulse energies between 0.1 nJ and 1 nJ (blue). For comparison, an E-PSF measurement from the 775-MINSTED, using the emitter Atto 647N in an equivalent measurement, is plotted in red. **b** Exemplary E-PSF from the measurement in **a** with an estimated FWHM of 24 nm at a STED pulse energy of 1 nJ. All relative standard errors of the FWHM remain below 1 % and are thus not further considered. **c-e** Imaging experiment of Cy3B-stained vimentin in U2OS cells. A confocal scan (**c**), a STED scan (**d**) and a scan with only the STED beam active (no excitation, **e**) are shown. A STED pulse energy of 0.5 nJ was applied. This measurement shows a substantial excitation of Cy3B by the STED light, leading to an SBR in **d** of approx. 2.5. This material was adapted from reference [38].

blue-shift was thus far away from disturbing measurements on the single emitter level. Yet, when performing STED imaging, the STED-induced background was clearly visible, as shown in **Fig. 3.8c-e**, where the STED excitation reduced the SBR to about 2.5. The images show the structure of vimentin in U2OS cells with a Cy3B immunolabel staining (performed by Ellen Rothermel and Tanja Koenen). The images were taken as single EOD scans with a STED power of 5 mW, each with and without excitation. As already discussed above, the increased background in the imaging application arises from the fact that for high emitter density, the SBR is scaling not only with the visibility of the E-PSF but also with the ratio of the peak and background volumes. This led to a poor imaging quality for this choice of STED wavelength. It yet remained questionable how the reduced visibility resulting from the blue-shift would affect the MINSTED measurements in practice.

To this end, we decided to use the stochastic repetitive labelling method DNA-PAINT to determine the performance of MINSTED single molecule localization measurements. This allowed us to use the fluorescent emitter Cy3B for all types of labelling. In DNA-PAINT,

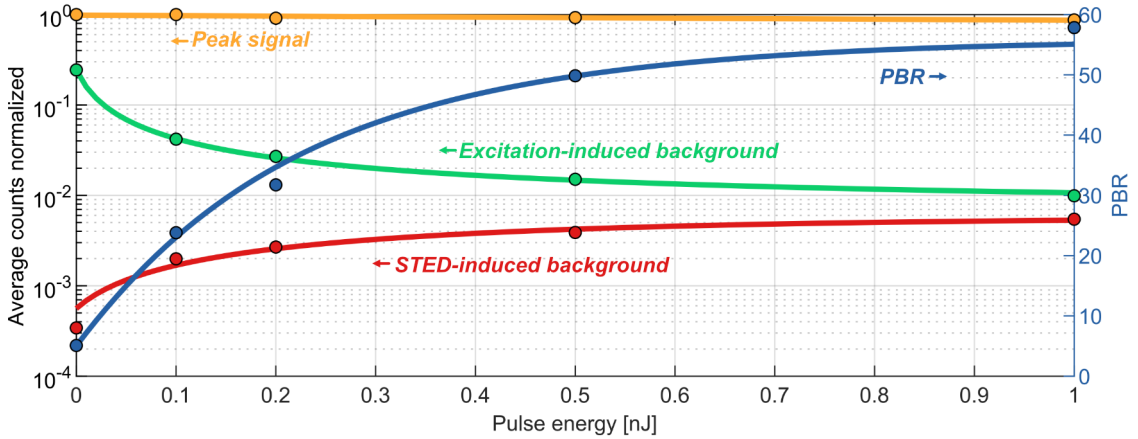

Figure 3.9: PBR scaling. Normalized counts of peak signal (orange), excitation-induced (green) and STED-induced background (red) in a STED scanning measurement of a DNA-PAINT labelled DNA origami sample are shown for varying STED pulse energy. The dots represent actual measurements, while the lines refer to analytical models. The PBR (blue) results from the other three measurements and theoretic models. As this analysis serves as a qualitative measurement, an error analysis was not conducted. This material was adapted from reference [38].

the freely diffusing imager strands are mostly the dominant source of background. For this kind of sample, the STED-induced background is again volume-dependent and not only stemming from the single emitter to be localized. In order to quantify the STED power dependent PBR scaling, an artificial sample, consisting of grid-like DNA origami structures with 3×3 attached docking sites, was prepared (see reference [38]). The origamis were immobilized on the coverslip and mounted in 7.5 nM imager strand solution. Scanning over a $2 \times 2\,\mu m^2$ FOV at different STED powers, with and without excitation, allowed to extract the STED power dependent signal, as well as the excitation and STED-induced background components, which directly enabled an estimation of the STED power dependent PBR. The peak signal was estimated from the scans (with enabled excitation) by detecting local maxima of temporarily bound imager strands and summing over a small surrounding region and normalizing with the 2D integral of a unit-amplitude Gaussian with a FWHM of the expected E-PSF. The sum of STED and excitation-induced background was estimated as the mean signal among the images, recorded with enabled excitation, while the STED-induced background was extracted equivalently from the scans without excitation light. The results are displayed in **Fig. 3.9**. The underlying theoretical model curves arise from the assumptions that:

1. The maximum signal is exponentially decaying with the STED power, due to the residual intensity in the STED donut minimum.

2. The excitation-induced background carries a constant component and one that is connected to the excitation of diffusing emitters. The latter is scaling with the integrated E-PSF signal, which is proportional to d^2 and thus scales with $(1 + kP)^{-1}$, with k being a fit constant.

3. The STED based background is modelled by assuming a constant term plus the solution of the two-level rate equation as given in eq. (2.49), when setting the initial condition to zero. This variable background term is scaling with $(1 - \exp(-kP))$, with a fit constant k. At small powers, this term increases linearly, but saturates to a constant value for large powers.

We find a strong increase of the PBR with increasing STED power. Comparing the confocal and the 10 mW STED case shows an improvement from PBR < 5 to PBR > 50 by more than an order of magnitude. Thus, despite the blue-shift, the STED beam had a net positive effect on the PBR in a DNA-PAINT sample. This illustrates the background suppressing capability of STED, which can be utilized in MINSTED to maintain high localization precision under challenging background conditions or speeding up measurements which rely on stochastic blinking of emitters (for example DNA-PAINT or actual fluorophore blinking) because higher densities of active emitters are tolerable.

Precision scaling in MINSTED

We prepared the same type of sample as above to quantify the precision scaling in MINSTED. We confocally scanned over a $2 \times 2\,\mu m^2$ FOV with a pixel dwell time of 1.6 ms. Whenever, more than 130 counts within four neighbouring pixels were detected, a localization was initiated. We chose a minimum FWHM of $d = 24$ nm and a radius $R = d/2$. The step fraction was chosen to $\alpha = 0.15$ and the downsizing factor of R and d during the zoom-in was set to $\gamma = 0.97$. The localizations were terminated as soon as the number of counts within 20 ms was falling below 16 (details on the applied measurement parameters are given in appendix C). Traces that did not survive the zoom-in phase were interpreted as mistaken localization initiations that do not correspond to an emitter, and were thus discarded prior to the analysis. The remaining localization traces (traces) were filtered with respect to several parameters, in order to remove background localizations and multiple emitter localizations. The filtering procedure is described in appendix D.1. The resulting dataset consisted of 991 localizations in total from approx. 144 binding sites. We computed the precision in two different contexts:

1. The single localization context, where the localization precision is estimated from each single trace, and

2. the binding site context, where the localization precision is estimated from multiple localizations at the same binding site.

The first method seems very convenient, as it provides a precision measure for each single trace. Yet the estimate is bound to the duration of a single localization (on the order of 100 ms). This estimate thus only includes spatial instabilities occurring within this time scale. Any slower processes will not be recognized. On the other hand, the second method includes all types of uncertainties that disturb the measurement:

- MINSTED localization uncertainty (approx. given by single trace localization precision),
- global sample instability (stabilized by NIR path),
- local binding site instability,
- instability of the optical paths and the beam positioning system.

Thus, the second method is a good estimate for the overall measurement precision, including the optical system and sample stability, while the first method estimates mainly the uncertainties arising from the statistics in the localization.

The single trace precision was estimated from the last N photons of the trace (after convergence). We now want to know the mean precision within a subset of length $n < N$

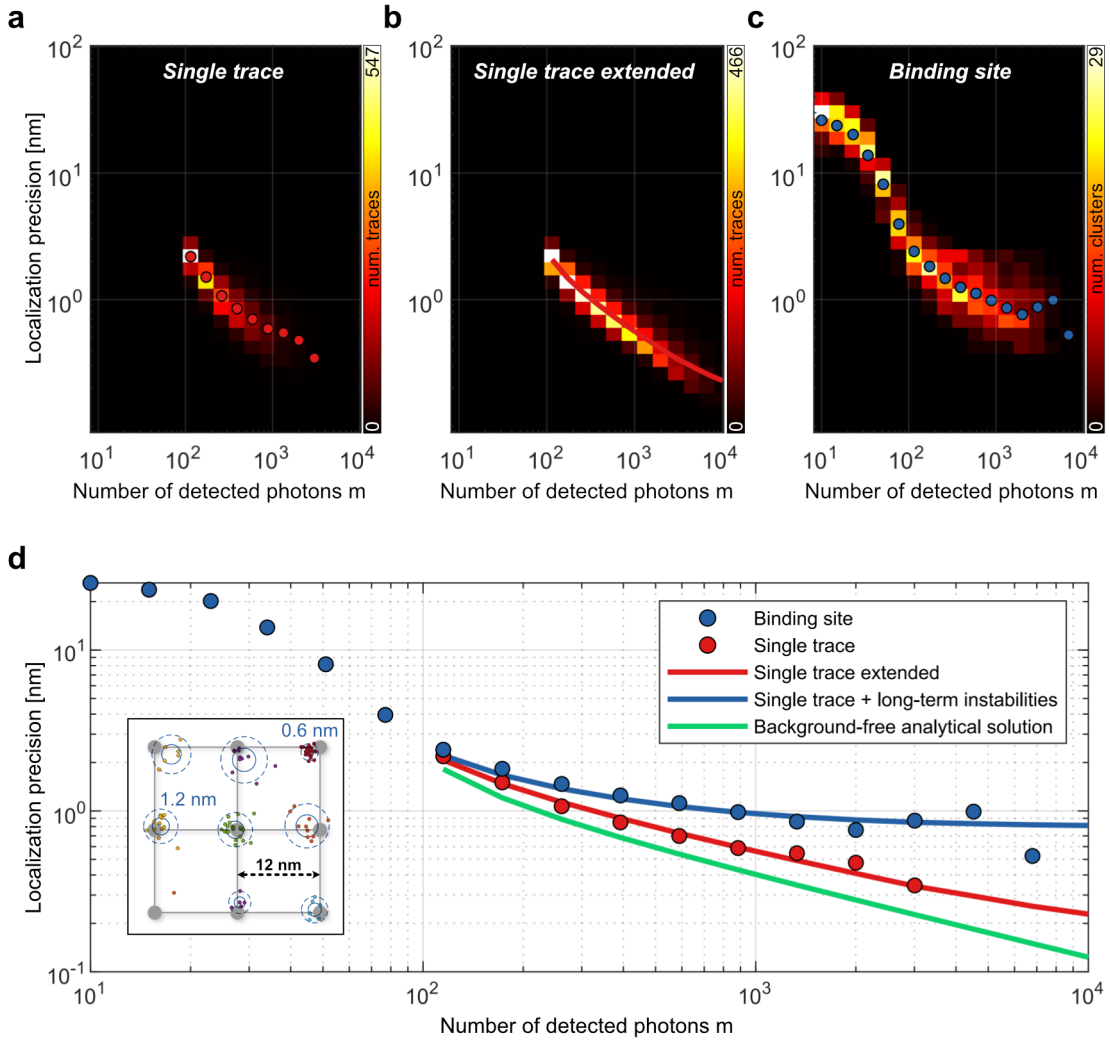

Figure 3.10: MINSTED Precision scaling. **a** A heat map illustrating the single trace localization precision estimate σ_s as a function of the number of detected photons m, containing all filtered traces. The median is indicated by red dots. **b,c** Single trace extended precision estimate σ_s^{ext} (median: red line), and binding site precision estimate σ_c (median: blue dots) are shown in an analogous manner. **d** The medians from **a-c** are again presented here. Additionally, the blue dots are modelled with a function $\sqrt{(\sigma_s^{ext})^2 + (\sigma_{stab})^2}$, which explains the deviation of the single trace estimate from the binding site estimate by an additional uncertainty, $\sigma_{stab} = (0.779 \pm 0.010)\,\mathrm{nm}$, which is interpreted as the overall long-term instability of sample and measurement system (blue line). A background-free analytical solution is presented for comparison (green line). The inset shows the spatial representation of an exemplary DNA origami, 3×3 binding sites on a square lattice with 12 nm next-neighbour distance (expected structure shaded in gray). The clustering of the localizations from separate binding sites is illustrated by differently coloured scatterers. Each cluster's standard deviation (solid lines) and twice its standard deviation (dashed lines) are presented in blue. Two exemplary values for standard deviations are provided in blue font. This material was adapted from reference [38].

of the converged trace. We thus took a moving average of n consecutive positions over the converged trace and computed the standard deviation of the result, which we consider the single trace precision σ_s (see **Fig. 3.10a**). This calculation is valid up to $n \leq N/5$, as for larger n, too few of the averaged blocks are fully independent from each other, which would lead to an underestimation of the precision value. In order to estimate the single trace

precision for even larger values of up to $n = N$, we computed σ_s for multiple increasing values of n and modelled the n-dependent decay of the precision with the expected relation from eq. (3.10). We used a modelling function of the form

$$f(n) = \frac{a}{(b+n)^c} \ , \tag{3.25}$$

with free parameters a, b and c, where we demand $a, b > 0$ and $0 < c < 0.5$. We thus also allow for non-square-root scaling, but restrict this to lower decay powers than 0.5. From the fit model, we could then read out the precision expectation of up to the full converged trace length. We call this the single trace extended estimate σ_s^{ext} (see **Fig. 3.10b**). As the first $M - N$ detections were discarded, claiming a precision σ_s or σ_s^{ext} at photon number n, would embezzle the first detections that are nonetheless needed as prior knowledge, which is gained during the zoom-in. We thus define the photon number $m = n + M - N$, which contains the zoom-in detections (see **Fig. 3.10a,b**).

To estimate the binding site precision, we grouped localizations into clusters of diameter $\lessapprox 10\,\text{nm}$, as further specified in appendix D.2. For the clustering, we used the mean center position of the converged trace as the single trace position estimate. The photon number m dependent precision estimate of the cluster σ_c was then computed as the standard deviation of position estimates over the clustered localizations after m detections. For $m > M - N$, these position estimates are simply given by the mean over the converged trace positions up to the respective number of detections. For $m \leq M - N$, the m^{th} position itself was taken as the position estimate for each trace[5]. We have to bear in mind that the clustered traces do contain different numbers of detections M. Thus, with increasing m, only a subset of the cluster's localizations contains the desired number of detections $m \leq M$. As long as ≥ 5 localizations in a cluster supply the number of m detections, the value σ_c was evaluated (see **Fig. 3.10c**).

The three precision estimates were separately computed in the x- and y-axis before computing their geometric mean to get a combined quantity. The scaling of the three estimates is shown in **Fig. 3.10d**. The single trace precision is decaying to well below 1 nm. The binding site precision on the other hand levels off at about 0.8 nm. By assuming that additional instabilities caused the binding site precision to differ from the single trace precision, we can model this by quadratically adding a stability uncertainty σ_{stab} to the extended single trace precision estimate

$$\sigma_c = \sqrt{(\sigma_s^{ext})^2 + (\sigma_{stab})^2} \ . \tag{3.26}$$

This also relies on the assumption that the instability is independent of the MINSTED uncertainty. Optimizing σ_{stab} to reduce the squared deviation from σ_c, leads to a stability estimate of $\sigma_{stab} = (0.779 \pm 0.010)\,\text{nm}$ (see **Fig. 3.10d**). This value may be interpreted as the average stability of the measurement system and the sample throughout the course of the measurement (about 40 min). Considering that the measurements were conducted at room temperature, this sub-nanometre overall precision is quite remarkable.

The single trace precisions as shown in **Fig. 3.10d** are close to the background-free analytical solution as computed from eq. (3.10). At large photon numbers (> 1000) a

[5] As the uncertainty during the zoom-in is exponentially reduced with each photon, the latest center position is a good position estimate. Including prior center positions with exponentially decaying weights, would only weakly affect this straight-forward estimate.

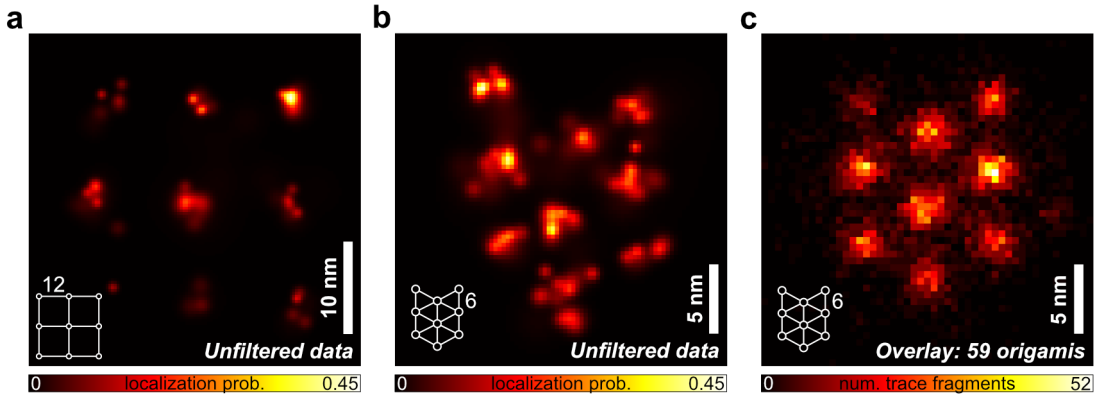

Figure 3.11: MINSTED imaging of DNA origamis. **a** MINSTED localizations rendered as an image of the exemplary 3×3 12 nm origami, already presented as part of **Fig. 3.10d**. **b** MINSTED measurement of 3×3 hexagonal DNA origami structure with 6 nm next-neighbour distance. **a,b** present unfiltered data. In the rendered images, each localization is displayed as a 2D Gaussian distribution with a standard deviation matching the extended precision estimate at the converged photon number. Each Gaussian is normalized so that its integral yields unity. **c** Histogram of overlaid DNA origami localization maps. After filtering as presented in appendix D.1, localizations from 59 origamis were cut in as many fragments of 2000 photons as possible and were translated and rotated for an optimized overlay. This averaging procedure was performed by Jan Keller-Findeisen. The expected structure is well preserved in the measured data. This material was adapted from reference [38].

settling in the data is observed that deviates from the $1/\sqrt{N}$ scaling. This may arise from residual instabilities already during the measurement time of single localization traces. Nevertheless, at 10^4 photons, a median single trace localization precision of < 0.3 nm was estimated. Even if this precision cannot be interpreted as the overall measurement uncertainty, it is a remarkable achievement for the MINSTED localization method itself, being able to reach a precision scale of the size of single atoms. Imagining an experiment under perfect conditions (immobile sample, binding sites and optical paths), the measurement uncertainty would actually match the single trace localization precision. However, under the presented conditions we can assume that our 636-MINSTED instrument (for Cy3B under the employed measurement conditions) yields an overall average measurement precision around 1 nm, which suggests an effective resolution on the single-digit nanometre scale. The latter statement was validated in further imaging experiments.

Imaging applications of blue-shifted MINSTED

DNA origamis As part of the last section, the spatial structure of DNA origamis with a minimum emitter distance of 12 nm were resolved (see **Fig. 3.10d** inset and **Fig. 3.11a**). To further investigate the resolving power of the MINSTED localization microscope, we measured an even smaller DNA origami structure featuring 9 binding sites on a hexagonal lattice with a next-neighbour distance of 6 nm. The sample preparation was equivalent to the one from the previous section, except that we now used an imager strand concentration of 15 nM to speed up the measurement. The applied measurement parameters are detailed in appendix C. An exemplary localization result is shown in **Fig. 3.11b**. In this image, each localization is represented by a Gaussian distribution according to the estimated single trace precision. The amplitude of each Gaussian is chosen in a way that the integrated signal remains constant, irrespective of the precision. Thus, narrow distributions appear

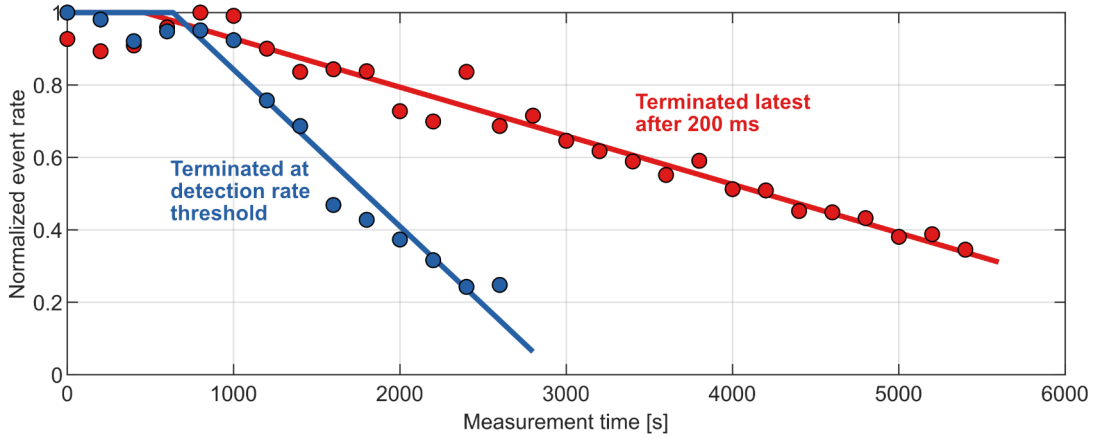

Figure 3.12: Degradation of DNA-PAINT docking sites. The normalized number of localizations per time (event rate) is shown over the course of two MINSTED measurements. For the first dataset, localizations were terminated after the detection rate fell below a set threshold (blue dots). In the second dataset, in addition to the regular termination condition, localizations were aborted after a maximum localization time of 200 ms (red dots). A respective saturated linear decay is presented for both datasets (lines). Assuming the event rate to scale with the number of available binding sites (which should be true at constant imager strand density), this data is indicative of the degradation of DNA-PAINT docking sites over time. As the termination is increasing the average lifetime of docking sites by more than a factor of two, it is suspected that bleaching of fluorescent emitters of bound imager strands might cause damage to the respective DNA-PAINT docking site. Both datasets were acquired with COS-7 cells, labelled with anti-lamin A/C antibodies with attached DNA-PAINT docking strands. Details on the sample preparation are given in reference [38]. The employed measurement parameters are listed in appendix C. This material was adapted from reference [38].

more bright, while broader distributions are rendered less intense. Overlaying the data from 59 origamis recovered the full structure, showing nine clearly resolved distributions in the expected hexagonal arrangement (see **Fig. 3.11c**). The overlay was performed with an algorithm from reference [12] by Jan Keller-Findeisen. These measurements validate the resolving power of the instrument at distances < 6 nm, which was already suspected from the < 1 nm binding site precision from the last paragraph. However, as the resolution is scaling with the inverse square-root of the detected photon number, not every localization is equally precise. As DNA-PAINT enables to measure the same binding site multiple times, the probability to gain high precision information from a substantial number of labelled sites is increased. This enhances the microscope's actual achievable resolving power, resulting in the complete resolution of individual structures as in **Fig. 3.11b** without the need for an ensemble averaging procedure. This has the potential to be a viable approach in biological imaging, as it allows for the observation of molecular spatial relations at each individual occurrence.

As mentioned above, DNA-PAINT principally offers the possibility to collect an arbitrarily high number of binding events from each labelled site. However, we observed a linear decrease in the localization rate throughout the measurement, which we interpreted as resulting from a degradation of docking sites during the measurement time. We suspected that damage to the binding sites was occurring due to an interaction of the docking sites with bleaching-related radicals. By aborting localizations after a maximum duration of 200 ms, bleaching was substantially reduced. From that, we could increase the lifetime of docking sites by more than a factor of two (see **Fig. 3.12**). This early termination of

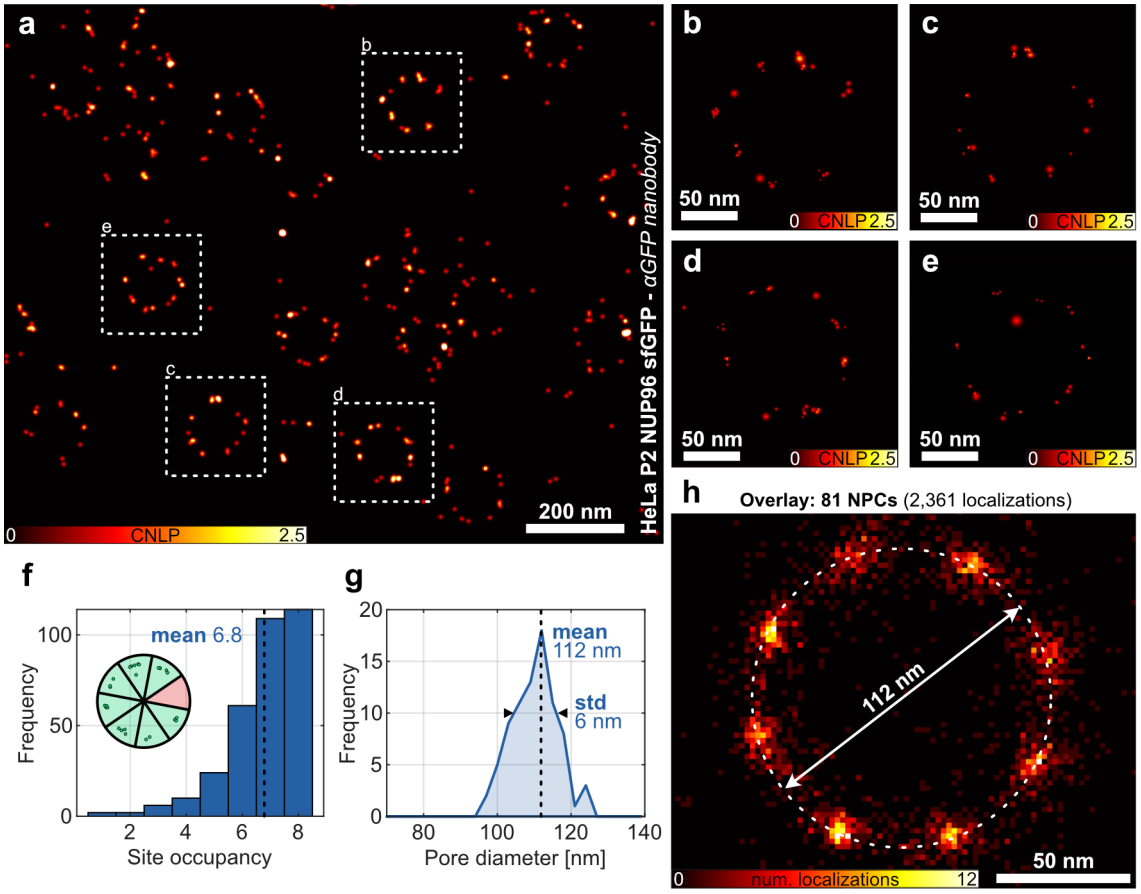

Figure 3.13: MINSTED imaging of the nuclear pore complex. **a** MINSTED localizations of protein arrangement of NUP96, labelled via DNA-PAINT. The image was rendered by illustrating each localization as a 2D Gaussian representing the spatial uncertainty estimate. However, the amplitude of the Gaussians was here set to unity, irrespective of the standard deviation. The pixel values thus represent the cumulative normalized localization probability (CNLP). To improve the visibility of highly precise localizations, the minimum rendered standard deviation was chosen to 3 nm. The image was saturated to a value of 2.5. **b-e** are detail views of **a**, where the localizations were rendered to full precision. **e** From the complete data set, the site occupancy of all identifiable pores (comprising 328 NPCs, representing 78 % of all localizations) was estimated, resulting in an average of 6.8 observed sites per pore. **f** A total of 81 pores were selected based on the criteria of having eight occupied sites and an ellipticity ≤ 1.25. The diameter of these pores is presented in a histogram, with a mean value of 112 nm and a standard deviation of 6 nm. **h** These selected pores were overlaid yielding the presented 2D histogram. The underlying analysis for **f-h** was carried out by Jan Keller-Findeisen. This material was adapted from reference [38].

localizations was applied in all of the following imaging applications of MINSTED.

Nuclear pore complex In order to prove the imaging capability of MINSTED in biological specimens, the nuclear pore protein NUP96, which is part the nuclear pore complex (NPC), was labelled with a DNA-PAINT docking strand in fixed HeLa cells. The sample preparation as well as the generation and characterization of the respective cell line was performed by Philip Gunkel, supervised by Volker C. Cordes. Details about the sample are given in reference [38] while the employed measurement parameters are specified in appendix C. An exemplary rendering of the MINSTED localizations of the labelled NUP96 is shown in **Fig. 3.13a-e**. The average precision is estimated to less than

2 nm. The NPC is known to feature an eight-fold symmetry, which is well preserved in the MINSTED measurements, and even visible in single representations of nuclear pores (see **Fig. 3.13c-e**). Given the full recorded dataset of NPCs from a total area of about 44 µm^2, Jan Keller-Findeisen selected all identifiable pores manually, resulting in 328 NPCs. He analysed those with respect to their site occupancy as described in detail in reference [38], which led to a mean of 6.8 observed sites per selected NPC (see **Fig. 3.13f**). Discarding pores with less than eight occupied sites and an ellipticity of ≥ 1.25 resulted in a set of 81 NPCs. For each of those pores, the diameter was estimated by Jan Keller-Findeisen, which is shown in **Fig. 3.13g**. A mean pore diameter of 112 nm was found. Overlaying the MINSTED localizations from these 81 NPCs was again performed by Jan Keller-Findeisen using the same algorithm as above. The overlay is presented in **Fig. 3.13h** and shows a clean eight-fold symmetric arrangement of the binding sites. However, the single site distributions are significantly larger than the single digit nanometre precision would suggest, which is attributed to the following reasons:

- NUP96 does not show eight single binding sites along the edge of the pore but features four binding sites per edge of the octagonal symmetry. Two of those lie on the top and two at the bottom of the pore (in the axial direction). The laterally projected distance of the single binding sites is approx. 10 nm. This leads, depending on the perspective, to an increased spread, mainly along the angular direction of the pore. This may additionally induce orientation distortions due to the planar projection. These features are not fully represented in the overlay image, as only sparse subsets of the full labelling are observed.

- Despite this higher structural complexity, MINSTED would be capable to distinguish between the four binding sites per edge. Its capability of separating even smaller distances was already shown in the former paragraph. Yet, those binding sites are not directly attached to the protein but to a linker, with a maximum extent of 7 nm (calculated from Volker C. Cordes and Philip Gunkel), that is inducing an offset to the actual protein attachment position. These linkers are usually not oriented in a defined fashion – they may be free to rotate and to expand or compress to a given extent, which obscures the true position of the protein.

- In addition to these effects, the biological structures are probably not completely rigid, but show inhomogeneities in shape, which is also shown by the large standard deviation in the pore diameters of 6 nm (see **Fig. 3.13g**).

The linker flexibility marks a general problem in super-resolution microscopy. Fluorescence is a necessary condition for substantially breaking the diffraction limit, but fluorescence labelling is at the same time a problem, as the label is always displaced from the actual structure of interest. For the above presented data, the instrument yields an average precision of less than 2 nm, even including sample instabilities, while each single binding site is biased from the protein of interest by an unknown extent of up to 7 nm. This leads to an effective resolution of 14 nm or even more, which is conducted by the sample itself. On this scale, the fine resolving power of a method like MINSTED is not reasonably utilised. It could be assumed that due to the free movement of the linker, the position estimates should average out, leading to a blurred yet unbiased estimate of the protein attachment site. However, if and to which extent or on which time scales the linker is free to move is not clear and could in principle vary from binding site to binding site. It thus remains an important aim to develop labelling strategies which facilitate a well-known positional relation between linker and protein attachment site.

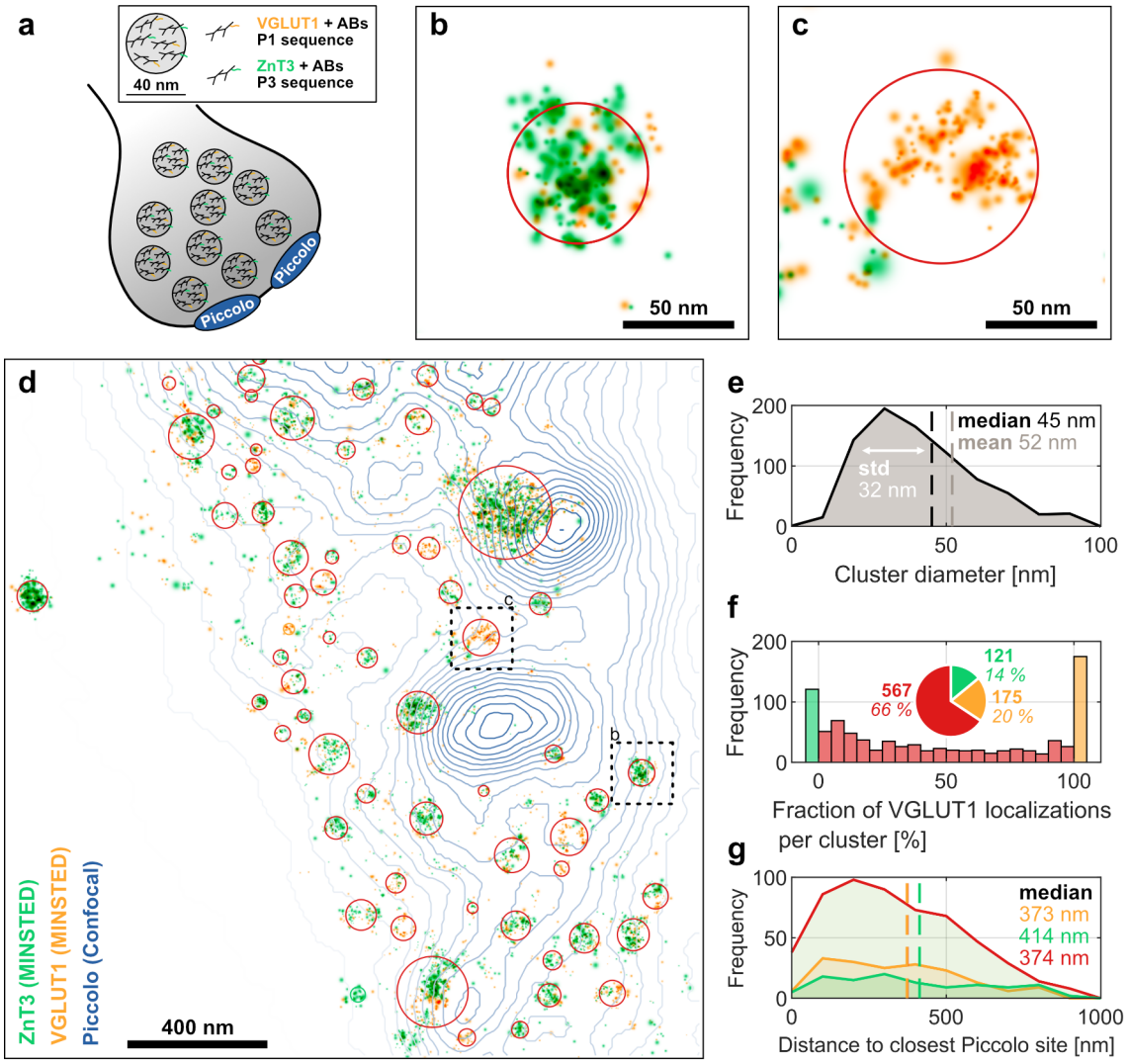

Figure 3.14: Two-color MINSTED imaging of synaptic vesicles. **a** Schematic of the immuno-DNA-PAINT labelled synaptic vesicle proteins VGLUT1 (orange) and ZnT3 (green) as well as the Alexa Fluor 488 immunolabelled protein Piccolo (blue). **d** Two-color MINSTED rendering using a minimum rendered Gaussian standard deviation of 0.5 nm and an amplitude of unity for each displayed localization. Clusters exhibiting solely VGLUT1 localizations are indicated by an orange circle, clusters comprising exclusively ZnT3 localizations are indicated by a green circle (none in the image), and mixed clusters are indicated by a red circle. The confocally imaged Piccolo distribution is underlaid as a contour plot in blue. Two detail views are presented in **b,c**. **e** Cluster diameter histogram for the full dataset. Cluster diameters are approximated as four times the spatial standard deviation of the contained localization centres. **f** Histogram of the color mixture (VGLUT1, ZnT3) among all clusters from the full dataset. The majority of clusters 66 % show colocalization of the two proteins. **g** Distance distribution of the clusters (pure VGLUT1 orange, pure ZnT3 green, mixed red) to their closest Piccolo site. No significant connection with respect to the contained vesicle proteins is found. This material was adapted from reference [33].

Synaptic vesicles An additional biological application of blue-shifted MINSTED was presented as part of the study in reference [33]. For the measurements, we applied the MINSTED localization parameters as specified in appendix C. The sample preparation and all necessary previous steps were performed by Sivakumar Sambandan. For these measurements, two types of proteins that are incorporated in the membrane of synaptic

vesicles – VGLUT1 and ZnT3 – were labelled in fixed cultured rat hippocampal neurons with two different docking strand sequences. In addition, the protein Piccolo was labelled with the fluorescent emitter Alexa Fluor 488. The labeling scheme is illustrated in **Fig. 3.14a**. By exchanging the DNA-PAINT imager solution during the measurement, the two targets could be sequentially observed, which enabled multi-colour MINSTED localization of these two protein distributions. Performing a "colour blind" cluster analysis (details given in reference [33]), the co-localization of these vesicle proteins could be investigated on the single molecule scale. The clustering used a Delaunay tesselation approach for localization-based data similar to the method published in [22], which was independently implemented by Marcel Leutenegger, who therefore contributed to the cluster analysis. An exemplary image is presented in **Fig. 3.14d**, showing the two-colour MINSTED data together with an underlying confocal image of the Alexa Fluor 488 labelled protein Piccolo. The cluster analysis enabled to assign the majority of localizations to a cluster, representing a synaptic vesicle (two examples in **Fig. 3.14b,c**). From the full dataset, a total of 863 clusters was identified. Each cluster's diameter was estimated as four times the geometric mean of the x and y spatial standard deviation of its assigned localization positions. A median cluster diameter of 45 nm with a strong spread of ±32 nm was observed (see **Fig. 3.14e**). Besides the biological inhomogeneity, this spread may also arise from wrong clustering, low labelling efficiency and large label offsets due to linker size, which was already discussed above. Due to the immunolabelling used for this measurement, the linker length was even greater than for the NPC imaging, leading to large possible displacements of the linker from the protein attachment site of more than 10 nm. However, assuming larger distances between the clusters than the linker offset, which is mostly true, the cluster analysis still remains valid and meaningful. We therefore extracted the distribution of the incorporated fraction of the two proteins (VGLUT1 and ZnT3) in the vesicle membranes (see **Fig. 3.14f**). The majority (66 %) of clusters contained both proteins, which suggests a colocalization of VGLUT1 and ZnT3, which was in line with further findings presented in reference [33]. The distance of clusters to their closest Piccolo site turned out to be broadly distributed between zero and 1 µm and did not show any significant dependence on the protein composition of the clusters (see **Fig. 3.14g**). As the labelling efficiency remains an unknown parameter, it might be that the colocalizing fraction of clusters was in fact higher than 66 % and that some of the clusters which were assumed to be pure VGLUT1 or ZnT3 vesicles were actually misclassified. The biological implications are not the primary focus of this study, but this application demonstrates the potential of MINSTED microscopy to provide insights into biological specimens at the molecular level through the labelling of multiple target structures. Nevertheless, labelling-related issues such as unknown labelling efficiency and linker length, significantly reduce the informative value of the data.

3.4.2 Tracking

In addition to the imaging applications, MINSTED was also shown to track the dynamics of emitters in motion [30]. In our implementation of the MINSTED sampling procedure, the system responds to each single photon detection by moving (on average) in the direction of the emitter. Therefore, it inherently follows the position of the emitter, even if it is subject to motion.

For the 636-MINSTED tracking project, the STED light source was updated to a more powerful version (see appendix B.2) of the laser fibre amplifier, which was designed by Michael Weber. The new design supported up to 1 nJ of pulse energy at an increased repetition rate of 40 MHz. The higher pulse rate helped to enhance the temporal precision

by increasing the available fluorescence detection rate from the emitter. The new light source contained an in-house designed pulse driver for the single-mode 636 nm laser diode. For the amplification, the re-design incorporated a customized double-clad fibre with an inner core containing the doped material where the amplification is performed, and an outer core, guiding the pump-light. The larger acceptance region for the pump light enabled the use of multi-mode pump diodes with a significantly larger power, which enhanced the available amplification gain.

Tracking performance measurements

In order to investigate the tracking performance, we used a DNA origami sample with an imager strand concentration of 5 nM, prepared as described in reference [30]. During each localization of a statically bound imager strand, the x-EOD was repeatedly displaced by $\Delta = 16\,\mathrm{nm}$, every $T = 15\,\mathrm{ms}$. This displacement altered the localization process and mimicked an instantaneous step of the emitter to the side. The subsequent convergence of the sampling center position towards the original molecule position reported the instrument's response function – and therefore the spatio-temporal precision in MINSTED tracking. These measurements were assisted by Mira Hesselink. The spatio-temporal precision of the 636-MINSTED was estimated by collecting 4×4 datasets in total, with four different excitation powers and four STED powers (details on measurement parameters are given in appendix C). We performed the following data analysis:

1. **Filtering 1:** All traces with less than $N = 100$ post-convergence photons and those with a converged duration, shorter than 31 ms were discarded.

2. **Segmentation:** We then segmented the single steps with their preceding and following plateaus within a time interval $[-T, T]$ around each step.

3. **Zeroing:** We then set the mean x-position C_x on the zeroing time interval $t \in Z = [-2T/3, 0]$ to zero, ending up with a set of steps $X(t)$, defined on a temporal range $[-T, T]$, jumping from $X = 0$ to $X = \Delta$ at $t = 0$. The choice for the zeroing interval to start at $t = -2T/3$ excluded the decay from the preceding step after $t = -T$.

4. **Filtering 2:** Thereafter, we discarded all steps with a detection rate below 70 % of the median among all localizations as well as steps that featured displacements larger than $\Delta + 2s_c$. The standard deviation of center positions s_c in this filtering step was approximated by the median standard deviation of the y-coordinate C_y among all localizations.

5. **Overlay:** All remaining steps were then mapped to a common temporal grid of times $\tilde{t}$, introducing a regular sampling of $T/5000$. The mapped positions $\tilde{X}(\tilde{t})$ were chosen to match the respective center position X that the system actually took at time $\tilde{t}$, which is the one referring to the most recent photon detection time t. With this, we arrived at a set of comparable steps $\tilde{X}(\tilde{t})$ and calculated the average x-position for all times $\tilde{t}$, which we denote the mean step response.

The corresponding response time τ was estimated by modelling an exponential decay. We then refined the zeroing interval to $Z = [-T + 5\tau, 0]$, which made sure that the mean step response from the former step was decayed to below 1 %, and repeated the analysis. The single-photon positional uncertainty was given by the mean among all steps of the standard

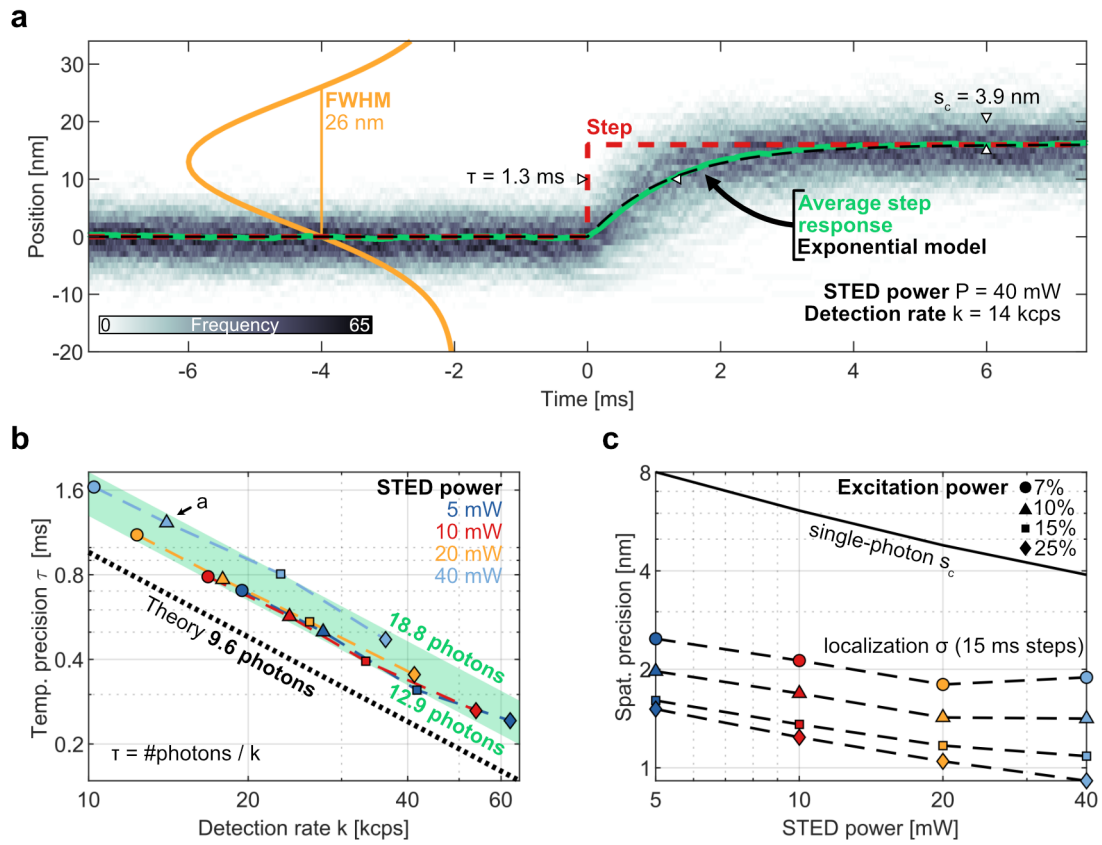

Figure 3.15: Tracking performance of MINSTED. a Exemplary step response measurement in 636-MINSTED at a STED power of 40 mW (at a pulse repetition rate of 40 MHz) and a mean detection rate of 14 kcps (kilo counts per second). The heat map represents a 2D histogram of 389 step responses of the sampling circle center as a function of time with respect to the stepping time at zero. The step (dashed red line), the mean step response (solid green line) and the corresponding exponential model (dashed black line) are overlaid on the data. The modelled response time $\tau = 1.3\,\text{ms}$ and the estimated single-photon spatial uncertainty $s_c = 3.9\,\text{nm}$ are indicated in black, while a Gaussian representing the applied E-PSF shape of FWHM 26 nm (orange) is displayed to clarify the spatial proportions. **b** Temporal precision as estimated from the average step response is plotted as a function of detection rate. The data points correspond to a total of 16 measurements at four different STED powers (5 mW (blue), 10 mW (red), 20 mW (orange), 40 mW (light blue)) and four different excitation powers (7 % (circles), 10 % (triangles), 15 % (squares), 20 % (diamonds)) in percent of approx. 30 µW. The region between minimal (12.9 photons) and maximal photon number (18.8 photons) detected on average during the response time, is shaded in green, while an analytical estimate of the temporal precision (dashed black line) corresponds to a response number of photons of 9.6. **c** Single-photon precision (solid black line) and inter-step localization precision (scatterers) are displayed as a function of STED power. As the scaling of the single-photon precision with the excitation power is negligible, the mean value for each STED power is presented. Relative standard errors of temporal and spatial precisions in **b,c** remain below 1 % and 1 ‰, respectively, and are thus not shown. This material was adapted from reference [30].

deviation of post-step positions

$$s_c = \langle \text{std}(X(t))_{5\tau<t<T} \rangle_{\text{steps}} \,. \tag{3.27}$$

This quantifies the positional noise on the trace. An exemplary step response is shown in **Fig. 3.15a** together with the extracted spatio-temporal uncertainties. The response function is closely following an exponential decay as theoretically predicted.

Analysing all 16 datasets, we find the scaling of the step response time τ (temporal precision) with respect to the detection rate k as shown in **Fig. 3.15b**. The temporal precision is clearly following a k^{-1} behaviour as theoretically expected. The temporal precision estimates are all $< 2\,\mathrm{ms}$, while for the smallest STED power, a temporal precision of $< 300\,\mu\mathrm{s}$ was achieved in the measurement with the largest excitation power. The proportionality factor between the temporal precision and the inverse detection rate is the number of photons which are detected during the time τ

$$\tau = \frac{N_c}{k} \ . \tag{3.28}$$

In the presented measurements, this photon number ranges from $N_c = 12.9$ to 18.8, thus upon detection of less than 20 photons, the system assumes the new position after a 16 nm step. The theoretical prediction of 9.6 is independent of the STED power, as this measure is only related to the fraction of FWHM and radius. For the three smallest STED powers, the measured values for N_c are about 40 % larger than the theoretical prediction, but remain rather stable irrespective of the STED power as theoretically predicted. A strong systematic deviation of N_c is recorded for the largest STED power, which is increased by about a factor of 2 with respect to the analytical solution. This is explainable, as the theoretical assumptions may be considered a valid approximation only in a range $\Delta/R < 1$. This has to be contrasted with a fraction $\Delta/R_{40\,\mathrm{mW}} \approx 1.2$. At the largest STED power, we are thus in a regime, where the process is not analytically explained.

The single-photon spatial uncertainty as shown in **Fig. 3.15c** is decaying with a power law, connected to the STED power, irrespective of the excitation power. Yet, calculating the error of the inter-step average position, we find a detection rate dependence. The localization precision of the average inter-step position, was estimated as the standard deviation over all steps of the average post-step position

$$\sigma = \frac{\mathrm{std}\big(\langle X(t)\rangle_{5\tau<t<T}\big)_{\mathrm{steps}}}{\sqrt{2}} \ . \tag{3.29}$$

The factor $\sqrt{2}$ accounted to the fact that the zeroing shifted the pre-step uncertainty towards the post-step mean. Assuming both uncertainties being equal leads to the given factor. Excluding the factor $\sqrt{2}$ would yield the precision of the mean estimate of the step size. Some of the presented curves show an optimum localization precision not at the maximum STED power but below. This is due to the fact that increasing the STED power is not only reducing d but also k (which is also visible in **Fig. 3.15b**). Thus, the number of position estimates to be averaged upon decreases with increasing STED power. Yet, for 15 ms inter-step durations, a localization precision slightly below 1 nm was achieved.

In tracking, the key to achieve better spatial precision is usually the temporal binning of spatial information, which allows for the smoothing of spatial uncertainties. Thus, the enhancement of spatial precision is typically accompanied with a loss of temporal information. However, upon observation of stepping emitters and when the single-photon spatial precision is sufficient to resolve individual stepping events (as for the presented data), the step detection may be performed right on the raw data at highest temporal precision. Thereafter, the averaging over each of the step plateaus may be performed, in order to yield spatial localization precisions on the scale of below 2 nm at full temporal precision of below 2 ms.

This requires to achieve a spatial precision that can distinguish the defining bits in the emitter's motion. In MINSTED, single-photon spatial precision is controlled via the E-PSF size, the sampling radius and by the number of correlated photons N_c. The latter parameter is actually implementing a kind of smoothing on the spatial data stream, directly during the measurement. The stronger the smoothing (higher N_c), the better spatial precision is achieved while worsening the temporal precision, which scales in proportion to N_c. This behaviour is tuned via the parameter α controlling the step length αR of individual sampling center displacements. The correlation number of photons is inversely scaling with α. Increasing the step length, spatial precision is lost while temporal precision is gained and vice versa. Therefore, the step fraction α is a parameter that is (like the post-processing smoothing length) linked to both spatial and temporal precision and serves to balance the two qualities. A distinctive attribute of MINSTED, however, is that its spatial precision may in principle be tuned independent from its temporal precision and even without compromising the SBR (not so in MINFLUX where reducing the sampling radius is directly worsening the SBR). This is achieved by varying the E-PSF size, which is a free parameter, theoretically not affecting the temporal precision (given constant R/d), which is predominantly controlled by the detection rate k. However, the presented measurements show that detection rate and STED power are not fully independent of each other as already stated above. In general, the detection rate is observed to decrease with higher STED power. The precise mechanism by which this behaviour arises and whether it is solely attributable to the residual intensity within the STED beam minimum remains unclear and should be further investigated to improve the quality of MINSTED tracking.

Nevertheless, given the stepsizes that shall be resolved, the MINSTED E-PSF may be tuned accordingly, to enable the step detection from the raw trace. The presented data for s_c (see **Fig. 3.15c**) suggests a single-photon resolution of minimal distances down to about 8 nm at highest STED power. This potentially makes MINSTED a valuable tool to investigate biological dynamics on the $\gtrapprox 10$ nm scale at millisecond temporal precision. This was validated, as shown in the next section, by tracking of the motor protein kinesin-1.

Tracking of kinesin

Kinesin-1 is a motor protein that transports cargo in eukaryotic cells by performing a stepping motion on microtubule strands that span a filamentous network within the cell body [4, pp. 987]. The directed stepping arises from the energy consumption of the motor by converting adenosine triphosphate (ATP) to adenosine diphosphate (ADP). Each conversion induces a stepping event. Depositing isolated microtubules onto a coverslip surface and having fluorescently labelled kinesin-1 proteins freely diffusing in the surrounding solution, their stepping may be observed in a fluorescent microscope. A respective model system assay was created by Lukas Scheiderer in the context of the work in reference [39], supported by Miroslaw Tarnawski and the Protein Core Facility at the Max Planck Institute for Medical Research in Heidelberg. In this assay, the ATP concentration was a tunable parameter to modulate the stepping rate of the kinesin-1. For the MINSTED measurements, Lukas Scheiderer prepared Cy3B- and Atto 647N-labelled kinesin-1 proteins in solution at a physiologically inspired ATP concentration of 1 mM (details on the sample preparation are given in reference [30]). We performed tracking measurements both on the 775-MINSTED and the 636-MINSTED instruments by scanning a FOV on the order of $2 \times 2\,\mu m^2$ and waiting for kinesin motors to bind to a microtubule, initiating a localization to follow its progression. Michael Weber and Lukas Scheiderer performed the measurements on the 775-MINSTED, while I operated the 636-MINSTED (details in appendix C). Four

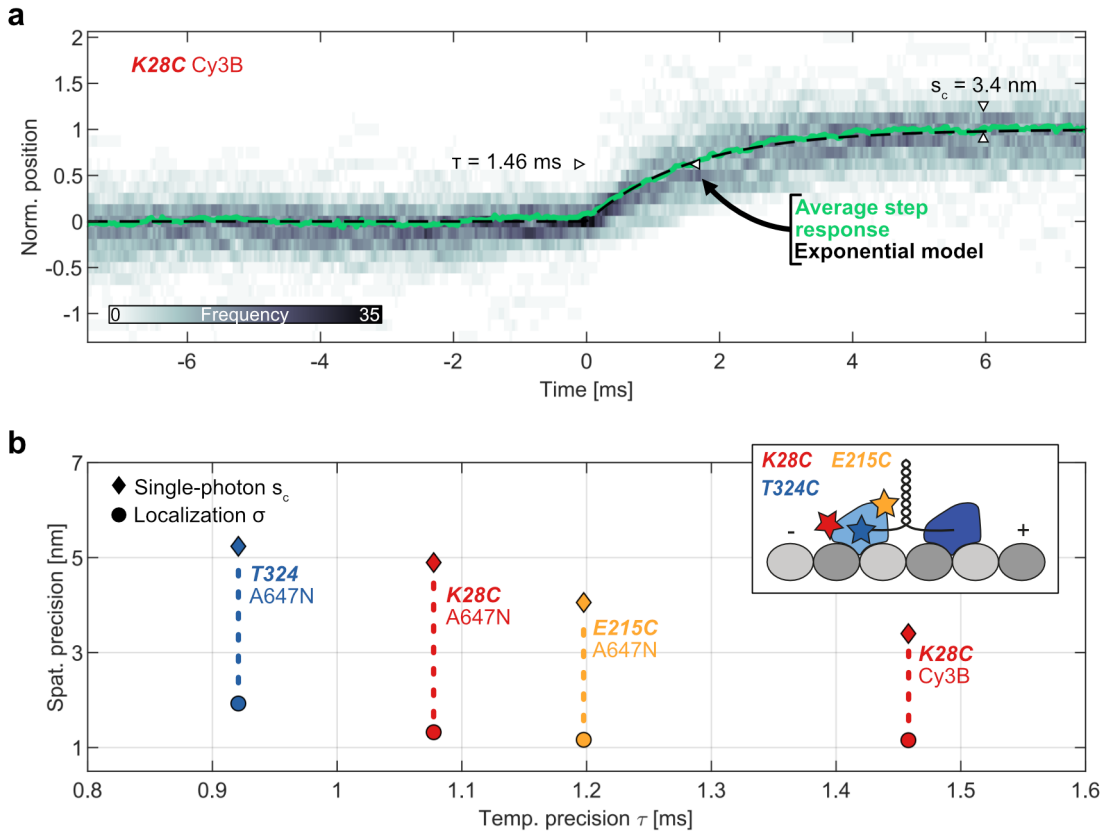

Figure 3.16: Spatio-temporal precision in MINSTED tracking of kinesin-1. **a** Overlay of kinesin steps from the dataset K28C Cy3B, measured with the 636-MINSTED at 40 mW of STED power (at a pulse repetition rate of 40 MHz). The average step response (solid green line) gives rise to an exponential behaviour (dashed black line) with a temporal response time of 1.46 ms and a single-photon spatial uncertainty of 3.4 nm. **b** Spatio-temporal precision of the four measured datasets. These correspond to different binding sites on the kinesin protein, as shown in the inset: K28C (red), T324C (blue) and E215C (orange). The label A647N stands for the fluorescent emitter Atto 647N. The Atto 647N datasets were acquired with the 775-MINSTED. In addition to the single-photon spatial precision s_c (diamonds), the inter-step spatial localization precision σ is displayed. The latter is calculated in analogy to the extended single trace precision estimate (as described in section 3.4.1) from the trace fragments in between each two steps. The temporal and spatial standard errors of the presented data remain below 31 µs and 0.26 nm, respectively, and are thus not shown. This material was adapted from reference [30].

datasets were recorded in total: Three types of kinesin-1 Atto 647N-labelled constructs (where the fluorescent label was sitting on different positions of the protein) were measured with the 775-MINSTED as well as one dataset that was recorded with the 636-MINSTED, using a Cy3B-labelled kinesin-1 construct (see inset in **Fig. 3.16b**).

The recorded localizations were filtered as described in appendix D.1 in order to extract the traces containing actual stepping events. The remaining traces were then analysed as follows:

1. **Coarse rotation:** We rotated the traces to walk in positive x-direction, by fitting a linear model to estimate the microtubular orientation.

2. **Step detection 1:** Using the function findchangepts (MATLAB 2021b), we identified steps both in x- and y-direction (separately). Herein, the penalty parameter

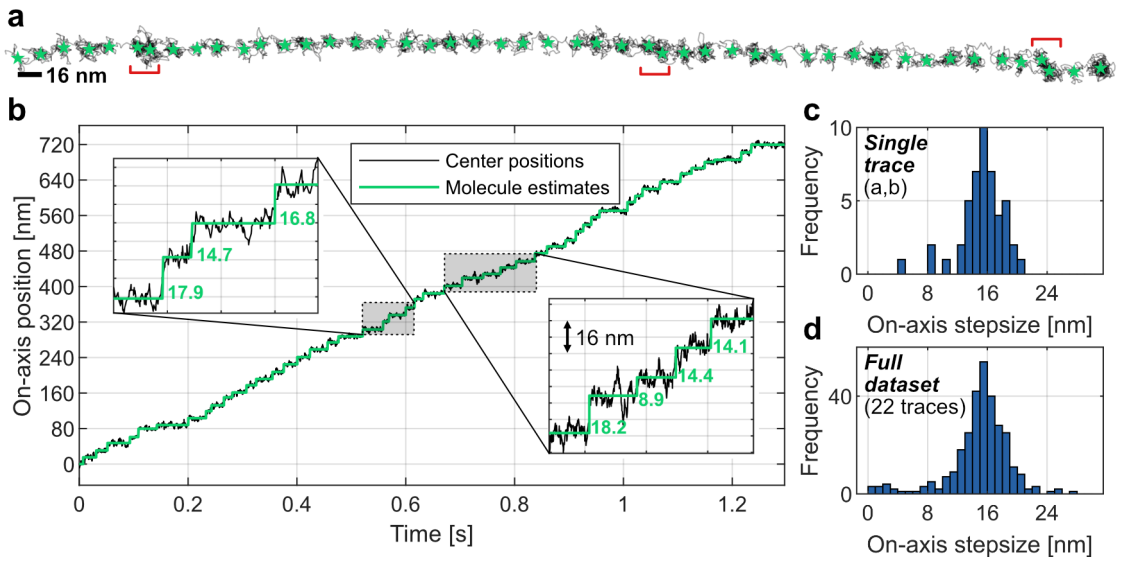

Figure 3.17: Exemplary kinesin-1 trace. **a** 2D representation of the sampling center position following the kinesin stepping motion. This trace is part of the E215C Atto 647N dataset. Average inter-step plateau positions are marked in green. Red square brackets mark short axial steps with a step length around 8 nm combined with off-axial displacements. **b** Axial position as a function of time of the same trace shown in **a**. The average inter-step plateau position is indicated as green line. The insets show zoom-ins of two regions. Stepping distances are indicated in green. **c** Histogram over the stepping distances of the single trace from **a,b** and of the full E215C Atto 647N dataset (**d**). Both histograms show a distribution of step lengths around 16 nm as physiologically expected. In addition, some steps covering half the expected distance, around 8 nm, are observed. However, given the width of the distribution, it is possible that these rare observations may also be attributed to the uncertainty of the step detection. This material was adapted from reference [30].

MinThreshold was chosen to $130 s_c^2$, where s_c was approximated by the median standard deviation in y-direction among all filtered traces. Steps in y-direction which were detected in close proximity (less than 20 data points deviation) to an x-step were discarded. Thereafter the 2D stepping distance was computed and all steps with a distance below 5 nm were discarded.

3. **Fine rotation:** For a better estimate of the microtubule axis, we computed the connecting vectors between the inter-step plateaus. We then computed their angles towards the horizontal and rotated the data in a way for the median angle to vanish. This should prevent off-axial stepping (steps along the y-direction, perpendicular to the microtubule orientation) to disturb the identification of the microtubule axis.

4. **Step detection 2:** The step detection was repeated with the refined rotation.

5. **Spatial precision estimation:** The single-photon spatial precision was estimated for each step plateau as geometrical mean of the standard deviations in both dimensions over the inter-step center positions. On each plateau, the localization precision was estimated by using the same algorithm as for the single-trace extended localization precision estimation in MINSTED imaging (see section 3.4.1).

6. **Temporal precision estimation:** For the temporal precision estimation, we overlaid the steps in a similar procedure as described for the artificial step measurements above. First, the axial position data was mapped onto a regular common temporal grid in a range of ±10 ms around each estimated step time, using time bins of 40 µs.

The data was then normalized by setting the pre- and post-step average to 0 and 1, respectively. Fitting a step model of the form

$$\max\left[1 - \exp\left(-(t-a)/b\right); 0\right] , \tag{3.30}$$

which is modelling a step from 0 to 1 initiated at time $t = a$ and exponentially decaying with a time constant b. According to the fit-parameter a, the data was temporarily shifted in order for the step initiation times to match. This was done for all steps with $b < 5\,\text{ms}$. The rest was considered as failed fits or having arised from fast sequences of multiple steps, closer than the segmentation interval. Overlaying all remaining steps from each dataset led to an average trace that could be modelled with an exponential decay, giving rise to the decay constant τ, as shown in **Fig. 3.16a**.

The mean single-photon spatial precision in the datasets varied between 3 nm and 6 nm, while the mean inter-step localization precision was estimated to $< 2\,\text{nm}$ in all datasets at an average temporal precision of $< 1.5\,\text{ms}$ as shown in **Fig. 3.16b**. An exemplary trace with highlighted step plateaus is shown in **Fig. 3.17a,b**. From both the single trace as well as the full dataset, we find a distribution of step sizes around the expected value of 16 nm (**Fig. 3.17c,d**). Furthermore, due to the single nanometre spatial precision and a temporal blur of sub 2 ms in the raw trace, detailed features of the kinesin's stepping motion were observable. Besides the anticipated 16 nm steps, smaller steps with a length of approximately 8 nm are also observed. In the examplary trace from **Fig. 3.17**, these coincide with off-axial stepping (see red markers in **Fig. 3.17a**). As microtubules are composed of multiple parallel protofilaments along which the kinesin is moving, these off-axis movements may be indicative of the kinesin switching from one protofilament to another. These observations are not statistically sound, yet the highly precise spatio-temporal precision of MINSTED allows the observation of molecular movement for each individual case.

For now, the temporal precision is not sufficient to resolve the trajectory of the foot of kinesin during the step. In the presented measurements, the step is a quasi instantaneous event. The limitations to this end arise from the fluorescent emitters. In principle, it would have been possible to increase the excitation power in the kinesin measurements, in order to boost the fluorescence detection rate. However, this would have resulted in more bleaching, leading to shorter traces, a reduction of observed steps and less statistics. In addition to that, single fluorescent molecules are generally limited in their fluorescent rate. Taking it to the extreme, one or two orders of magnitude could be gained in fluorescence detection rate. Assuming 10 to 20 photon detections within the instrument's response time, the ultimate limit of temporal precision of fluorophore-based tracking would amount to around 10 µs. However, achieving such high detection rates would require cw excitation and different detectors, which is not compatible with MINSTED.

Nevertheless, MINSTED is capable of precise tracking at the nanometre-millisecond level while exhibiting robust resilience against background disturbances. This comes from the inherent nature of the background-supressing STED beam. Especially its high SBR renders MINSTED tracking a high potential method for molecular biology observations in challenging environments, for example at high background in living cells.

CHAPTER
FOUR

CONCLUSION

4.1 Imaging

This dissertation presented a way of alleviating the stability issues in MINSTED microscopy, which enabled to achieved an overall measurement precision of about 1 nm in fluorescence localization microscopy. On the scale of a single localization, MINSTED was even shown to localize to an uncertainty of < 0.3 nm for 10^4 detected photons. This precision close to the atomic scale renders the potential measurement uncertainty that could be achieved under perfect conditions (stable sample and optics) in MINSTED microscopy. Overall spatial instabilities from the sample as well as the optical path were estimated to approx. 0.8 nm, which confirms the performance of the NIR active stabilization system as well as the passive stability of the localization unit. This was estimated from repeated measurements of the same binding sites, utilizing the DNA-PAINT labelling scheme. Despite the blue-shifted STED wavelength and the additional background from diffusing DNA-PAINT imager labels, the MINSTED SBR remained high, the background therefore did not significantly deteriorate the measurements. The SBR even increased for higher STED powers, which highlights a key strength of the MINSTED localization scheme. The nanometre localization precision permitted the resolution of distances of 6 nm and even below, besides the full recovery of a 3×3 binding site origami from 58 single measurements. These results were transferred to fixed biological cells, where the nuclear pore complex was imaged. The overlay of segmented single pores showed the expected eightfold-symmetric arrangement. Nevertheless, due to incomplete labelling and the linker flexibility, the overlay did not fully capture the expected structural arrangement. In fixed rat hippocampal neurons, the ability of MINSTED to perform multi-color imaging via DNA-PAINT multiplexing was shown. All imaging applications featured a mean localization precision below 2 nm and may thus be considered to yield information on the molecular scale.

Label offset We noticed that the fluorescent labelling procedures induces substantial offsets between the protein attachment site and the position of the label. Flexibility and rotational degrees of freedom of the linker enlarge the distribution of possible emitter positions to a region of about 10 nm around the attachment site, even for the smallest possible linkers, such as nano-bodies or snap- and halo-tag. In light of these considerations, a measurement precision at the single nanometre level may not be reasonable for these types of labelling in biological applications. In any case, this calls for novel labelling strategies towards shorter linkers with predictable positioning in order to clarify the spatial relation

between protein and emitter. For example, attaching three labels to one protein each – in a well-known positional and directional arrangement – would enable the resolution of each protein's orientation and position in space. However, this is not feasible yet given the limitations of current labelling efficiencies. Moreover, such dense and complicated labelling strategies further separate light microscopy from its major advantage to be a quick and flexible measurement tool for a large range of samples. If this is not provided by an optical microscope, the question arises why not to use EM. There are two reasons for that: the live-cell compatibility of optical microscopes (which I will discuss in the following paragraph) and the specificity of fluorescent labelling – not seeing everything, as in EM, may also be a big advantage. The second point is of significant importance. Both EM and optical microscopy have distinct strengths, and it is unlikely that optical microscopy will surpass EM in its ability to provide a vast amount of contextual information about a sample – information that cannot be colour coded in a labelling strategy. It is therefore my view that optical high-resolution imaging for fixed cells should concentrate on developing correlative approaches in conjunction with EM in order to benefit from both the contextual information obtained by EM and the clarity afforded by fluorescence imaging. At present, optical imaging has reached a point where correlative imaging can operate at a precision scale that enables the observation of biology on the molecular level.

Live-cell compatibility One of the key advantages of light microscopy is its ability to perform live-cell or even living tissue imaging, which is not achievable with other microscopy techniques. And if one would imagine an optimal microscope, it would exactly enable this: the observation of dynamics on all relevant length and time scales. Most of the novel high-resolution microscopy techniques have alienated from live-cell applicability – being theoretically live-cell compatible but practically not suitable. In the case of MINSTED, photo-toxicity is one potential issue in live-cell imaging (besides the use of DNA-PAINT in the presented applications). Blue-shifting the STED wavelength enabled a drastic reduction of the STED power by more than an order of magnitude at equal or even better resolution. Trading off some of the performance by allowing the localization uncertainty to increase three- to four-fold, one could push MINSTED to the sub-milliwatt STED power regime. Nevertheless, both MINSTED and MINFLUX imaging remain slow because emitters are imaged sequentially, one at a time. This presents a challenge as the living cell dynamics would make it hard to determine actual structural relationships of localizations while the structure is evolving. Speeding up would be possible via parallelization of MINSTED or MINFLUX. This has yet to be demonstrated and presents a number of significant challenges. In beam scanning illumination localization, stochastic switching (ON/OFF) schemes may be employed. This breaks down when parallelizing via a fixed optical pattern. This necessitates the utilization of optically controlled OFF-switching, which limits the number of available high-performing labels, or requires the implementation of an optical configuration that facilitates the adaptable and autonomous relocation of individual zero positions.

4.2 Tracking

Besides the imaging applications, we have shown the ability of MINSTED to track moving emitters on the nanometre and millisecond spatio-temporal precision scale. A comprehensive measurement of the MINSTED's response to instantaneous steps clarified the characteristics of the instrument's spatio-temporal filtering, which adds to the actual motion of the emitter. This characterization was performed for varying parameters of the

STED and excitation power and was mainly in line with the theoretical expectations. The stepping of the motor protein kinesin-1 was observed at a mean spatio-temporal precision of less than 6 nm per 1.5 ms, while the average inter-step positions were localized to a mean positional uncertainty of below 2 nm.

The ability of observing biologically relevant motion at such low uncertainties opens up many possibilities for applications in live sciences. It has to be noted that the temporal precision in tracking of single fluorescent emitters is, however, generally limited to about 10 µs when considering possible detection rates that are achievable with high-performance fluorophores (to date). This precision still remains above the timescale of single conformational changes of proteins under physiological conditions. This implies that the precise procedure of rapid molecular biology mechanisms may from the current perspective not be unravelled through fluorescence microscopy. Scattering may support much higher information rates but comes with the problem of large scattering particles and the lack of contrast.

Nevertheless, many biological processes occurring above this critical temporal scale may be tackled by fluorescence microscopy and could benefit from the presented spatio-temporal precision. As MINSTED additionally features a high SBR, even tracking under challenging conditions, for example at high background in living cells should be manageable. In this context, MINSTED is demonstrably superior to any wide-field tracking techniques and to MINFLUX also. TIRF illumination can drastically enhance the SBR in camera-based emission tracking, but severely limits to asses 3D dynamics. The latter is of great importance when it comes to biologically relevant processes inside cells, as the third dimension is crucial to unravel the real spatial relationships. The important further development in MINSTED tracking is thus, firstly, the realization of localization in all three dimensions, and secondly, its extension to at least one additional colour for simultaneous tracking. High-speed information on the molecular interaction of multiple biological compounds in a living specimen could facilitate the understanding of biological processes. Fortunately, MINSTED could tolerate rather close packing of the simultaneously labelled sites (even without switchable emitters) because fluorescence from neighbouring sites further away than the E-PSFs extent would be strongly depleted. In conclusion, an improved MINSTED microscope is expected to offer significant potential for the investigation of dynamics in molecular biology in living specimens.

APPENDIX

A

ADDITIONAL CALCULATIONS

A.1 Cumulative statistics

The following derivation is inspired from a lecture in statistical physics by Peter Sollich, that I attended in 2019/2020 [31].

We assume a random variable x that is distributed with respect to the probability density $P(x)$ with $\int dx\ P(x) = 1$. We define a function $\chi(\theta)$ as the expectation value of $\exp(i\theta x)$ with respect to P

$$\chi(\theta) = \langle \exp(i\theta x) \rangle = \int dx\ P(x)\ \exp(i\theta x)\ , \tag{A.1}$$

which is basically the Fourier transform of P. We call χ the characteristic function of P. By Taylor expansion of the exponential, we find the characteristic function to be a sum over the moments[1]

$$\chi(\theta) = \sum_n \frac{(i\theta)^n}{n!}\ \langle x^n \rangle\ . \tag{A.2}$$

A simple calculation shows that the k^{th} derivative of χ, evaluated at $\theta = 0$ is proportional to the k^{th} moment

$$\frac{1}{i^k}\ \frac{\partial^k \chi}{\partial \theta^k}\bigg|_{\theta=0} = \langle x^k \rangle\ . \tag{A.3}$$

We define a second function as

$$K(\theta) = \ln(\chi(\theta)) \tag{A.4}$$

and also write it in the form of a Taylor expansion with unknown pre-factors c_n

$$K(\theta) = \sum_n \frac{(i\theta)^n}{n!}\ c_n. \tag{A.5}$$

These factors are again found from the derivatives of K

$$c_k = \frac{1}{i^k}\ \frac{\partial^k K}{\partial \theta^k}\bigg|_{\theta=0}\ . \tag{A.6}$$

[1] The n^{th} moment of a distribution is given by the expectation value of the n^{th} power of the distributed variable $\langle x^n \rangle$.

Calculating the first factor, we find

$$c_1 = \frac{1}{i}\frac{\partial}{\partial\theta}\big[\ln(\chi)\big]\Big|_{\theta=0} = \frac{1}{i}\,\frac{\chi'}{\chi}\Big|_{\theta=0} = \frac{\langle x^1\rangle}{\langle x^0\rangle} = \langle x^1\rangle\ , \tag{A.7}$$

where the 0^{th} moment being unity is following from the normalization of P. The second factor reads

$$c_2 = \frac{1}{i^2}\frac{\partial}{\partial\theta}\left[\frac{\chi'}{\chi}\right]\Big|_{\theta=0} = -\left[-\frac{\chi'^2}{\chi} + \frac{\chi''}{\chi}\right]\Big|_{\theta=0} = \langle x^2\rangle - \langle x\rangle^2\ . \tag{A.8}$$

Thus, while c_1 is equal to the mean, c_2 yields the variance – the two first derivatives of K (and actually also the higher ones) thus yield the respective cumulants of P. We thus call K the cumulant generating function.

With this in mind, we now assume N independent random variables $x_1, ..., x_N$ sampling the distributions $P_1, ..., P_N$. We now want to find the statistics of the combined random variable

$$z = \frac{\sum_n x_n}{N}\ , \tag{A.9}$$

expressing the mean over the x_n. For that, we evaluate the cumulant generating function of z

$$K_z(\theta) = \ln\left(\left\langle \exp\left(i\frac{\theta}{N}\sum_n x_n\right)\right\rangle\right) = \ln\left(\int dz\ P_z(z)\ \exp\left(i\frac{\theta}{N}\sum_n x_n\right)\right). \tag{A.10}$$

For independent measurements, we can write P_z as the product of the P_n and find

$$K_z(\theta) = \ln\left(\prod_n \underbrace{\int dx_n\ P_n(x_n)\ \exp(i\theta x_n/N)}_{=\chi_n(\theta/N)}\right)\ , \tag{A.11}$$

where the integral is equal to the characteristic function of the n^{th} random variable. Pulling the product out of the logarithm, we obtain

$$K_z(\theta) = \sum_n K_n(\theta/N)\ , \tag{A.12}$$

and directly evaluate the first two cumulants of z to

$$\begin{aligned}\langle z\rangle &= \frac{1}{i}\frac{\partial K_z}{\partial\theta}\Big|_{\theta=0} = \sum_n \frac{1}{i}\frac{\partial K_n(\theta/N)}{\partial\theta}\Big|_{\theta=0} = \frac{1}{N}\sum_n\langle x_n\rangle \\ \mathrm{var}(z) &= \frac{1}{i^2}\frac{\partial^2 K_z}{\partial\theta^2}\Big|_{\theta=0} = \sum_n \frac{1}{i^2}\frac{\partial^2 K_n(\theta/N)}{\partial\theta^2}\Big|_{\theta=0} = \frac{1}{N^2}\sum_n \mathrm{var}(x_n)\ .\end{aligned} \tag{A.13}$$

A.2 Cramér-Rao bound

The following derivation is inspired from a lecture in theoretical biophysics by Jörg Enderlein, that I attended in 2019/2020 [8]. It was additionally part of my Master's thesis [36].

We assume the outcome of an experiment $\mathcal{E}$, which is conducted on a system $\mathcal{S}$. From the experiment, we infer the value $\lambda(\mathcal{E})$, which corresponds to a system variable λ^*. The conditional probability to observe an experiment $\mathcal{E}$, given the system $\mathcal{S}$, shall be denoted $P(\mathcal{E}|\mathcal{S})$. For an unbiased estimate of the systems variable, we assume that when averaging over all possible realizations of the experiment, the deviation between $\lambda(\mathcal{E})$ and λ^* should vanish

$$\langle[\lambda(\mathcal{E}) - \lambda^*]\rangle = \int d\mathcal{E}\ [\lambda(\mathcal{E}) - \lambda^*]\ P(\mathcal{E}|\mathcal{S}) = 0\ . \tag{A.14}$$

Taking the derivative with respect to λ^* gives

$$\int d\mathcal{E}\left([\lambda(\mathcal{E}) - \lambda^*]\ \frac{\partial P(\mathcal{E}|\mathcal{S})}{\partial \lambda^*} - P(\mathcal{E}|\mathcal{S})\right) = 0\ . \tag{A.15}$$

As we demand a normalized distribution P, the second term gives unity. Meanwhile, we can rewrite the derivative and find

$$\int d\mathcal{E}\ [\lambda(\mathcal{E}) - \lambda^*]\ P(\mathcal{E}|\mathcal{S})\frac{\partial\ \ln\left(P(\mathcal{E}|\mathcal{S})\right)}{\partial \lambda^*} = 1\ . \tag{A.16}$$

By using the Cauchy-Schwarz inequality for infinitely dimensional vector spaces, we can split the integral in two terms and find

$$\left\langle[\lambda(\mathcal{E}) - \lambda^*]^2\right\rangle \underbrace{\left\langle\left[\frac{\partial\ \ln\left(P(\mathcal{E}|\mathcal{S})\right)}{\partial \lambda^*}\right]^2\right\rangle}_{=\mathcal{I}(\lambda^*)} \geq 1\ . \tag{A.17}$$

The second term equals the definition of the Fisher information $\mathcal{I}(\lambda^*)$. For the measurement uncertainty of $\lambda(\mathcal{E})$, we thus find

$$\sigma(\lambda^*) = \sqrt{\langle[\lambda(\mathcal{E}) - \lambda^*]^2\rangle} \geq \frac{1}{\sqrt{\mathcal{I}(\lambda^*)}}\ , \tag{A.18}$$

which is called the Cramér-Rao bound. Thus, from the experiment $\mathcal{E}$, the inverse square-root of the contained Fisher information about the unknown variable λ defines a lower bound to the attainable uncertainty.

If our measurement is constructed of independent sub-measurements $\mathcal{E} = \{\mathcal{E}_1, ..., \mathcal{E}_N\}$, we may write

$$P(\mathcal{E}|\mathcal{S}) = \prod_n^N P(\mathcal{E}_n|\mathcal{S})\ . \tag{A.19}$$

Inserting this to the definition of the Fisher information, the product is turning to a sum in the logarithm, which may be excluded from the integration

$$\mathcal{I}(\lambda^*) = \left\langle\left[\frac{\partial\ \ln\left(P(\mathcal{E}|\mathcal{S})\right)}{\partial \lambda^*}\right]^2\right\rangle = \sum_n \left\langle\left[\frac{\partial\ \ln\left(P(\mathcal{E}_n|\mathcal{S})\right)}{\partial \lambda^*}\right]^2\right\rangle = \sum_n \mathcal{I}_n(\lambda^*)\ . \tag{A.20}$$

The total information that is contained in independent sub-measurements is thus given by summing over the Fisher information of each sub-measurement. This validates the normalization in eq. (2.70) to obtain the single-photon Fisher information.

APPENDIX

B

TECHNICAL SPECIFICITIES

B.1 Components blue-shifted MINSTED

This component list is an extended version of the one from reference [38].

Light sources

375nm Diode laser iPulse-375-S-3V1 (18 mW cw), Toptica, Gräfelfing, Germany.

473nm Diode laser MBL-III-473-50 (50 mW cw), Changchun New Industries Optoelectronics, Changchun, China.

560nm Pulsed fiber laser SuperK Extreme EXU-6 PP (600mW visible, 80 MHz, pulse picker), NKT Photonics, Köln, Germany. Modulation via AOTF (AA Opto Electronics, Orsay, France).

561nm Diode laser 85-YCA-010 ($<$ 10 mW cw), Melles Griot, Rochester, NY, USA.

636nm Diode laser HL63391DG (633-643nm, 200 mW), Ushio Opto Semiconductors, Tokyo, Japan; amplified in Praesodymium-doped fiber 201218/TB305 (Le Verre Fluoré, Bruz, France). Pumped with 450 nm cw single mode laser diode PLPT450B, Osram, Berlin, Germany. Pulse driver: modified EPC9144, Efficient Power Conversion, El Segundo, CA, USA. Modulation via: LM0202 electro optic modulator, Linos Photonics, Göttingen, Germany; driven with a high-voltage amplifier PZD700A-1-H-SHV-CE ($\pm$700 V, 200 mA, 150 kHz), Trek, Lockport, NY, USA.

850nm Super-luminescent LED module EBD273106-13 ($<$ 5mW cw), Exalos, Schlieren, Switzerland.

980nm Diode laser LuxX-980-150 ($<$ 150 mW cw), Omicron-Laserage Laserprodukte, Rodgau-Dudenhofen, Germany. Power stabilization via NIR laser power controller, Brockton Electro-Optics Corp, Bridgewater, MA, USA.

Filters

F_1 VersaChrome edge tuneable long-pass filter TLP01-628, Semrock, Rochester, NY, USA.

F_2 Excitation filter ZET561/10x, Chroma Technology, Bellows Falls, VT, USA.

$F_{3,7}$ BrightLine single-band band-pass filter FF01-582/75, Semrock, Rochester, NY, USA.

F_4 VersaChrome edge tuneable short-pass filter TSP01-628, Semrock, Rochester, NY, USA.

F_5 Razor-edge ultrasteep long-pass filter LP02-561RE, Semrock, Rochester, NY, USA.

F_6 Razor-edge ultrasteep short-pass filter SP01-633RU, Semrock, Rochester, NY, USA.

F_7 BrightLine multiphoton short-pass filter FF01-680, Semrock, Rochester, NY, USA.

F_8 EdgeBasic long-pass filter BLP01-488R, Semrock, Rochester, NY, USA.

F_9 Emission band-pass filter HQ525/50, Chroma Technology, Bellows Falls, VT, USA.

Collimators

C_1 Fibre collimator 15 mm, 60FC-A15-4-01, Schäfter+Kirchhoff, Hamburg, Germany.

C_2 Fibre collimator 18 mm, 60FC-A18-0-01, Schäfter+Kirchhoff, Hamburg, Germany.

C_3 Achromatic doublet 15 mm, 84-328, Edmund Optics, Mainz, Germany.

C_4 Fibre collimator 7.5 mm, 60FC-A7.5-0-01, Schäfter+Kirchhoff, Hamburg, Germany.

C_5 Fibre collimator 4.5 mm, 60FC-A4.5-0-01, Schäfter+Kirchhoff, Hamburg, Germany.

C_6 Achromatic doublet 20 mm, NIR, AC080-020-B, Thorlabs, Newton, NJ, USA.

C_7 Achromatic doublet 10 mm, 322206000, Linos Photonics, Göttingen, Germany.

Lenses

$L_{1,9}$ Achromatic doublets 400 mm, 322275322, Linos Photonics, Göttingen, Germany.

L_2 Achromatic doublet 80 mm, 322323000, Edmund Optics, Mainz, Germany.

$L_{3,5,15,16}$ Achromatic doublets 100 mm, 322345000, Linos Photonics, Göttingen, Germany.

L_4 Achromatic doublet 100 mm, 49-333, Edmund Optics, Mainz, Germany.

L_6 Achromatic doublet 300 mm, 322273322, Linos Photonics, Göttingen, Germany.

L_7 Achromatic doublet 150 mm, 322331322, Linos Photonics, Göttingen, Germany.

L_8 Achromatic doublet 50 mm, 322339000, Linos Photonics, Göttingen, Germany.

$L_{10.1,10.2}$ Achromatic doublets 150 mm, AC254-150-B, Thorlabs, Newton, NJ, USA.

$L_{11,12}$ Achromatic doublets 500 mm, 322329000, Linos Photonics, Göttingen, Germany.

$L_{13,14}$ Achromatic doublets 100 mm, 322236000, Linos Photonics, Göttingen, Germany.

L_{17} Achromatic doublet 300 mm, AC254-300-B, Thorlabs, Newton, NJ, USA.

$L_{18,19}$ Achromatic doublets 50 mm, 322265525, Linos Photonics, Göttingen, Germany.

Mirrors

CM Cold mirror M254C45, Thorlabs, Newton, NJ, USA.

DM_1 Dichroic mirror Z670SPRDC, Chroma Technology, Bellows Falls, VT, USA.

$DM_{2,4}$ Dichroic mirror FF560-fDi02-t3, Semrock, Rochester, NY, USA.

DM_3 Dichroic mirror H405LPXR, Chroma Technology, Bellows Falls, VT, USA.

DM_5 Dichroic mirror 495DCXR, Chroma Technology, Bellows Falls, VT, USA.

DM_6 Dichroic mirror 555DCXRU, Chroma Technology, Bellows Falls, VT, USA.

$PM_{1,2}$ Off-axis parabolic mirror MPD129-P01, Thorlabs, Newton, NJ, USA.

Rest MaxMirror ultra-broadband mirror, Semrock, Rochester, NY, USA.

Polarization optics

$PBS_{1,2}$ Polarizing beam splitter cube 49-002, Edmund Optics, Mainz, Germany.

PBS_3 Polarizing beam splitter cube PBS202, Thorlabs, Newton, NJ, USA.

PBS_4 Polarizing beam splitter cube PTW 2.20, Bernhard Halle, Berlin, Germany.

$\lambda/4$ Achromatic quarter-wave retarder plate RAC 3.4.10, Bernhard Halle, Berlin, Germany.

$\lambda/2$ Achromatic half-wave retarder plates RAC 3.2.10, Bernhard Halle, Berlin, Germany.

Special optics

FB Field block R1DF100, Thorlabs, Newton, NJ, USA.

PH_1 Pinhole diameter 50 µm, Linos Photonics, Göttingen, Germany.

PH_1 Pinhole diameter 100 µm, Linos Photonics, Göttingen, Germany.

VPP Vortex phase plate V-660-20-1, Vortex Photonics, München, Germany.

Additional material

$APD_{1,2}$ Single-photon counting module SPCM-NIR-3X-W4-FC, Excelitas, Wiesbaden, Germany.

CAM_1 USB3 sCMOS camera pco.panda 4.2 (2048×2048 pixels $6.5 \times 6.5\,\mu m^2$, 16 bits, 100 fps), PCO, Kelheim, Germany.

CAM_2 USB3 CMOS camera xiQ MQ013RG-ON (1280×1024 pixels $4.8 \times 4.8\,\mu m^2$, 8/10 bits, 210 fps), Ximea, Münster, Germany.

$EOD_{x,y}$ Electro-optic deflectors 311A (AD*P, free aperture 2 mm, 200 mm long, $7\,\mu rad\,V^{-1}$, 180 pF, ±500 V), Conoptics, Danbury, CT, USA; high-voltage drivers PZD700A-1-H-SHV-CE (±700 V, 200 mA, 150 kHz), Trek, Lockport, NY, USA.

FPGA PCIe board 7852R with drivers and software LabVIEW 2017, National Instruments, Austin, TX, USA.

$G_{1,2}$ Galvo-mirrors 6 mm × 10 mm, galvo 6215H, servo drivers 671, Cambridge Technology, Bedford, MA, USA.

Objective HCPL APO CS2 100x, 1.4 NA oil-immersion, Leica Mikrosysteme, Wetzlar, Germany.

Stages Three-axis piezo stage P-562.3CD with controller E-727.3CD on two-axis piezo stage M-686.D64 with controller C-867.260 and single-axis piezo stage M-230.25 with controller C-863.11, Physik Instrumente, Karlsruhe, Germany.

B.2 Re-designed fibre amplifier

The re-designed fibre amplifier was seeded with the same 636 nm single mode laser diode as before (HL63391DG, Ushio Opto Semiconductors, Tokyo, Japan). The < 1.5 ns optical pulsing was achieved with an in-house developed electronic pulse driver designed by Michael Weber and Frank Meyer. The amplification fibre (TB376, Le Verre Fluoré, Bruz, France) featured a single-mode 5 µm diameter Praseodymium 3+ doped core, guiding the seed light, and a multi-mode D-shaped cladding (22.5 µm × 25 µm, guiding the pump light. We used two polarization-combined multi-mode beams from 2.2 W laser diodes (PLPT5 450KA, Osram, Berlin, Germany) to pump the Praseodymium 3+ doped core. The seed and pump beams were coupled from opposite sides and the amplified pulses were split from the pump light path, using a short-pass DM (DMSP505T, Thorlabs, Newton, NJ, USA). To avoid damage at the seed diode, a Faraday isolator blocked the residual pump light that did not get absorbed in the fibre. All laser diodes were cooled via thermoelectric cooling and their beam shapes symmetrized with an anamorphic prism pairs. The modulation of the amplified 636 nm pulses was implemented as before by an electro optic modulator (LM0202, Linos Photonics, Göttingen, Germany) that was driven with a high-voltage amplifier (PZD700A-1-H-SHV-CE, Trek, Lockport, NY, USA).

APPENDIX

C

MINSTED MEASUREMENT PARAMETERS

In advance to each MINSTED localization, an active and stably bound fluorophore was searched in a confocal scan over a FOV of about $2 \times 2\,\mu m^2$ with a pixel dwell time of T_i. As soon as the counts of four neighbouring pixels exceeded a value N_i, a localization was initiated. In all presented localizations, a step fraction of $\alpha = 0.15$ and a zoom-in downsizing factor of $\gamma = 0.97$ were chosen. The zoom-in stopped when the final STED power P (at repetition rate r), FWHM d and sampling radius R were reached. Localizations were terminated when less than 16 detections were registered during a termination time window of duration T_t. For the presented datasets, the parameters were set to the values in **Tab. C.1**.

Table C.1: Overview of the measurement parameters, applied in the presented MINSTED localization measurements.

Dataset	P [mW]	r [MHz]	d [nm]	R [nm]	T_i [µs]	N_i	T_t [ms]
Imaging							
3x3 12 nm grid	10	10	24	12	1600	100-140	20
3x3 6 nm grid	10	10	24	15	1600	100-140	8
Lamin	5	10	36	18	1600	100-140	30
Nuclear pores	5	10	30	15	1600	80-140	6
Synaptic vesicles	5	10	30	15	1600	100-140	6
Tracking 636							
DNA origami	5	40	54	27	1600	40-80	8
	10	40	41	21	1600	40-80	8
	20	40	33	17	1600	40-80	8
	40	40	26	13	1600	40-80	8
K28C-Cy3B	40	40	20	10	60	4	8
	40	40	24	12	60	4	8
Tracking 775							
E215C-Atto 647N	100	40	36	18	60	5	16
T324C-Atto 647N	100	40	36	18	60	5	16
K28C-Atto 647N	100	40	36	18	60	5	30
	50	40	50	25	60	5	30

APPENDIX
D

DATA ANALYSIS

D.1 Filtering

D.1.1 MINSTED imaging

MINSTED imaging traces were filtered with respect to three parameters to meet the following conditions:

- at least N_{min} post-convergence photons,
- a standard deviation of center positions in both axes smaller than $s_{c,max}$,
- and a mean detection rate between k_{min} and k_{max}.

This filtering was performed after discarding all traces that did not reach the minimum FWHM. For the MINSTED data presented in this dissertation the filtering parameters in **Tab. D.1** were applied.

Table D.1: Overview of the filtering parameters, applied in MINSTED imaging.

Dataset	N_{min}	$s_{c,max}$ [nm]	k_{min} [kcps]	k_{max} [kcps]
3x3 12 nm grid	100	4.2	-	-
3x3 6 nm grid	-	-	-	-
3x3 6 nm grid overlay	100	5.5	-	16.0
Nuclear pores	100	6.5	5.0	16.5
Synaptic vesicles	100	10.0	4.0	22.0

D.1.2 MINSTED tracking

MINSTED tracking traces of kinesin stepping were filtered with respect to five parameters to meet the following conditions:

- a duration of at least T_{min},
- a covered distance of at least D_{min},
- an off-axis standard deviation between $s_{c,min}$ and $s_{c,max}$,
- a minimum ratio of spatial standard deviations in on- and off-axis direction r_{min},

- and a brightness fluctuation of at most σ^k_{max}.

This filtering was performed after discarding all traces that did not reach the minimum FWHM. The brightness fluctuation is defined as the standard deviation of the smoothed detection rate over the time of the converged localization. The momentary detection rate was estimated by applying a moving mean of length 20 to the time differences between consecutively detected photon detection times

$$k(t) = \frac{1}{\text{movmean}_{20}(\Delta t(t))} \ . \tag{D.1}$$

Thereafter, $k(t)$ was further smoothed with a kernel size 50 before computing its standard deviation, which was compared to σ^k_{max} for filtering. For the kinesin tracking MINSTED datasets, presented above, the filtering parameters in **Tab. D.2** were applied.

Table D.2: Overview of the filtering parameters, applied in kinesin MINSTED tracking.

Dataset	T_{min}	D_{min}	$\{s_{c,min/max}\}$	r_{min}	σ^k_{max}
E215C-Atto 647N	50 ms	64 nm	$\{2, 10\}$ nm	4	2.0 kcps
T324C-Atto 647N	50 ms	64 nm	$\{2, 10\}$ nm	4	2.0 kcps
K28C-Atto 647N	50 ms	64 nm	$\{2, 10\}$ nm	4	2.0 kcps
K28C-Cy3B	50 ms	64 nm	$\{2, 10\}$ nm	4	1.8 kcps

D.2 Cluster analysis

For this clustering algorithm, the localizations (mean of center positions after convergence) are mapped to a pixelated image (2D histogram) and are afterwards smoothed with a 2D Gaussian with a standard deviation $D_{max}/2$. D_{max} is a free parameter, controlling the size of the clusters. For the clustering of single binding sites on the 3x3 12 nm origami, it was chosen to 5 nm. In this smoothed image, we detect local maxima using the function imregionalmax (MATLAB R2021b) – those are regarded as potential cluster centres. Potential cluster centres with an amplitude (smoothed image) $\leq 1\,‰$ with respect to the global maximum are discarded. We then iterate through the remaining ones: If the number of localizations in a region $< D_{max}$ around the cluster center is larger than N_{min} (the minimum number of localizations, here $N_{min} = 5$), those localizations are considered a cluster. If at least one of these localizations is already part of a cluster, the latter is appended. This should lead to clusters of an approximate diameter $2D_{max}$, which are clearly separated and contain a minimum number of N_{min} localizations. Clusters which are not well separated, will be merged and those, carrying too few localizations, will not be considered at all.

BIBLIOGRAPHY

[1] E. Abbe. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468, December 1873.

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

[3] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science*, 313(5793):1642–1645, September 2006. Publisher: American Association for the Advancement of Science.

[4] Alberts Bruce, Heald Rebecca, Johnson Alexander, Morgan David, Raff Martin, Roberts Keith, and Walter Peter. *Molecular Biology of the Cell: Seventh International Student Edition with Registration Card*. W.W. Norton & Company, June 2022. Google-Books-ID: ISdiEAAAQBAJ.

[5] Jelmer Cnossen, Taylor Hinsdale, Rasmus Ø Thorsen, Marijn Siemons, Florian Schueder, Ralf Jungmann, Carlas S. Smith, Bernd Rieger, and Sjoerd Stallinga. Localization microscopy at doubled precision with patterned illumination. *Nature Methods*, 17(1):59–63, January 2020. Publisher: Nature Publishing Group.

[6] A. V. Deshpande, A. Beidoun, A. Penzkofer, and G. Wagenblast. Absorption and emission spectroscopic investigation of cyanovinyldiethylaniline dye vapors. *Chemical Physics*, 142(1):123–131, March 1990.

[7] R. F. Egerton. An Introduction to Microscopy. In R.F. Egerton, editor, *Physical Principles of Electron Microscopy: An Introduction to TEM, SEM, and AEM*, pages 1–26. Springer International Publishing, Cham, 2016.

[8] Jörg Enderlein. Theoretical Biophysics, Lecture Series, 2019.

[9] Fabian Göttfert, Tino Pleiner, Jörn Heine, Volker Westphal, Dirk Görlich, Steffen J. Sahl, and Stefan W. Hell. Strong signal increase in STED fluorescence microscopy by imaging regions of subdiffraction extent. *Proceedings of the National Academy of Sciences*, 114(9):2125–2130, February 2017. Publisher: Proceedings of the National Academy of Sciences.

[10] S. W. Hell and M. Kroug. Ground-state-depletion fluorscence microscopy: A concept for breaking the diffraction resolution limit. *Applied Physics B*, 60(5):495–497, May 1995.

[11] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, June 1994. Publisher: Optica Publishing Group.

[12] Hamidreza Heydarian, Florian Schueder, Maximilian T. Strauss, Ben van Werkhoven, Mohamadreza Fazel, Keith A. Lidke, Ralf Jungmann, Sjoerd Stallinga, and Bernd Rieger. Template-free 2D particle fusion in localization microscopy. *Nature Methods*, 15(10):781–784, October 2018. Publisher: Nature Publishing Group.

[13] Michael Hofmann, Christian Eggeling, Stefan Jakobs, and Stefan W. Hell. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proceedings of the National Academy of Sciences*, 102(49):17565–17569, December 2005. Publisher: Proceedings of the National Academy of Sciences.

[14] Ralf Jungmann, Christian Steinhauer, Max Scheible, Anton Kuzyk, Philip Tinnefeld, and Friedrich C. Simmel. Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami. *Nano Letters*, 10(11):4756–4761, November 2010. Publisher: American Chemical Society.

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

[16] Thomas A. Klar and Stefan W. Hell. Subdiffraction resolution in far-field fluorescence microscopy. *Optics Letters*, 24(14):954–956, July 1999. Publisher: Optica Publishing Group.

[17] Wulfhard Lange. *Einführung in die Laserphysik*. Wissenschaftliche Buchgesellschaft, 1983.

[18] Adam M. Larson. Multiphoton microscopy. *Nature Photonics*, 5(1):1–1, January 2011. Publisher: Nature Publishing Group.

[19] Changhwan Lee, Emma Z. Xu, Yawei Liu, Ayelet Teitelboim, Kaiyuan Yao, Angel Fernandez-Bravo, Agata M. Kotulska, Sang Hwan Nam, Yung Doug Suh, Artur Bednarkiewicz, Bruce E. Cohen, Emory M. Chan, and P. James Schuck. Giant nonlinear optical responses from photon-avalanching nanoparticles. *Nature*, 589(7841):230–235, January 2021. Publisher: Nature Publishing Group.

[20] Marcel Leutenegger, Christian Eggeling, and Stefan W. Hell. Analytical description of STED microscopy performance. *Optics Express*, 18(25):26417–26429, December 2010. Publisher: Optica Publishing Group.

[21] Marcel Leutenegger, Michael Weber, Henrik von der Emde, and Stefan W. Hell. Compensating the delayed electro-optical deflector response, December 2021. arXiv:2112.01621 [physics].

[22] Florian Levet, Eric Hosy, Adel Kechkar, Corey Butler, Anne Beghin, Daniel Choquet, and Jean-Baptiste Sibarita. SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nature Methods*, 12(11):1065–1071, November 2015. Publisher: Nature Publishing Group.

[23] Douglas B. Murphy and Michael W. Davidson. *Fundamentals of Light Microscopy and Electronic Imaging*. John Wiley & Sons, August 2012. Google-Books-ID: f8qhtqRjxQMC.

[24] Adrian Negrean and Huibert D. Mansvelder. Optimal lens design and use in laser-scanning microscopy. *Biomedical Optics Express*, 5(5):1588–1609, May 2014. Publisher: Optica Publishing Group.

[25] Reut Orange kedem, Nadav Opatovski, Dafei Xiao, Boris Ferdman, Onit Alalouf, Sushanta Kumar Pal, Ziyun Wang, Henrik von der Emde, Michael Weber, Steffen J. Sahl, Aleks Ponjavic, Ady Arie, Stefan W. Hell, and Yoav Shechtman. Near index matching enables solid diffractive optical element fabrication via additive manufacturing. *Light: Science & Applications*, 12(1):222, September 2023. Publisher: Nature Publishing Group.

[26] Sujitha Puthukodan, Eljesa Murtezi, Jaroslaw Jacak, and Thomas A. Klar. Localization STED (LocSTED) microscopy with 15 nm resolution. *Nanophotonics*, 9(4):783–792, April 2020. Publisher: De Gruyter.

[27] Christian Ringemann, Ben Harke, Claas von Middendorff, Rebecca Medda, Alf Honigmann, Richard Wagner, Marcel Leutenegger, Andreas Schönle, Stefan W. Hell, and Christian Eggeling. Exploring single-molecule dynamics with fluorescence nanoscopy. *New Journal of Physics*, 11(10):103054, October 2009.

[28] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, October 2006. Publisher: Nature Publishing Group.

[29] R. Rüdenberg. Elektronenmikroskop. *Naturwissenschaften*, 20(28):522–522, July 1932.

[30] Lukas Scheiderer, Henrik von der Emde, Mira Hesselink, Michael Weber, and Stefan W. Hell. MINSTED tracking of single biomolecules. *Nature Methods*, 21(4):569–573, April 2024. Publisher: Nature Publishing Group.

[31] Peter Sollich. Statistical Physics, Lecture Series, 2019.

[32] Wolfgang Stößel. *Fourieroptik: Eine Einführung*. Springer-Verlag, October 2013. Google-Books-ID: BTLTBgAAQBAJ.

[33] Neha Upmanyu, Jialin Jin, Henrik von der Emde, Marcelo Ganzella, Leon Bösche, Viveka Nand Malviya, Evi Zhuleku, Antonio Zaccaria Politi, Momchil Ninov, Ivan Silbern, Marcel Leutenegger, Henning Urlaub, Dietmar Riedel, Julia Preobraschenski, Ira Milosevic, Stefan W. Hell, Reinhard Jahn, and Sivakumar Sambandan. Colocalization of different neurotransmitter transporters on synaptic vesicles is sparse except for VGLUT1 and ZnT3. *Neuron*, 110(9):1483–1497.e7, May 2022. Publisher: Elsevier.

[34] Bernard Valeur and Mário Nuno Berberan-Santos. *Molecular Fluorescence: Principles and Applications*. John Wiley & Sons, March 2013. Google-Books-ID: TcNu1BYLUyEC.

[35] Giuseppe Vicidomini, Gael Moneron, Christian Eggeling, Eva Rittweger, and Stefan W. Hell. STED with wavelengths closer to the emission maximum. *Optics Express*, 20(5):5225–5236, February 2012. Publisher: Optica Publishing Group.

[36] Henrik von der Emde. Molecule Localization with STED. Master's thesis, Georg-August-Universität Göttingen, Göttingen, January 2021.

[37] Michael Weber, Marcel Leutenegger, Stefan Stoldt, Stefan Jakobs, Tiberiu S. Mihaila, Alexey N. Butkevich, and Stefan W. Hell. MINSTED fluorescence localization and nanoscopy. *Nature Photonics*, 15(5):361–366, May 2021. Publisher: Nature Publishing Group.

[38] Michael Weber, Henrik von der Emde, Marcel Leutenegger, Philip Gunkel, Sivakumar Sambandan, Taukeer A. Khan, Jan Keller-Findeisen, Volker C. Cordes, and Stefan W. Hell. MINSTED nanoscopy enters the Ångström localization range. *Nature Biotechnology*, 41(4):569–576, April 2023. Publisher: Nature Publishing Group.

[39] Jan O. Wirth, Lukas Scheiderer, Tobias Engelhardt, Johann Engelhardt, Jessica Matthias, and Stefan W. Hell. MINFLUX dissects the unimpeded walking of kinesin-1. *Science*, 379(6636):1004–1010, March 2023. Publisher: American Association for the Advancement of Science.

ACKNOWLEDGEMENT

At first, I would like to thank Stefan W. Hell for giving me the opportunity to perform this project, for his trust and the freedom that he gave me to develop my skill and my own ideas.

I additionally thank Ellen Rothermel and Tanja Koenen, who were preparing countless samples for me, without whom a lot of the presented measurements would have not been possible. Moreover thanks to Christian Klaba and Mario Lengauer and the rest of the team from the mechanics workshop, who helped me with mechanical construction and manufactured many custom parts. Thank you, Rainer Pick for designing the parabolic mirror galvanometric scanner unit as well as Frank Meyer for helping with electronic issues.

I acknowledge my collaboration partners Sivakumar Sambandan and Lukas Scheiderer for their patience, biological interpretations and advice regarding data analysis, which led to some beautiful and hopefully meaningful results.

My special thanks goes to Marcel and Michael, who were my main supervisors from day one and introduced me to the field of microscopy. Without you two, I would have not gained this range of skills and knowledge that I could develop during the MINSTED project and further on. You are still my most important advisers, when it comes to tricky questions, whether they are about experimental matters or related to theoretical optics. Thanks for sharing your knowledge and your rich wealth of experience with me.

I acknowledge my colleague Mira, whom I have worked with for the last 1.5 years and who will carry on a lot of things that I was not able to finish. We have had a funny, often very silly time together and I hope I was not the worst teacher. You will bring the pieces together. I honestly wish you all the best!

Thank you Thomas, for being my consultant and for your patience, trust and reflection in countless conversations, about science and a lot more. I'm really sure, that we won't lose sight of each other. Thanks for being a good friend.

As my main pastime activity in the last three years was CLEAR GROUNDS, I want to thank Hendrik, Janik, Nils and Pia for being really good friends, talking so much trash, not taking me too serious and being an insanely nice rock band.

Thank you Bekki, for being my companion, my best friend, for being the one I can talk to about everything. Thank you for not getting tired, listening to my worries, my problems and basically all of my thoughts. I love you.

I deeply acknowledge my family who set the basis for the path that I'm allowed to walk on. Thank you for always encouraging me in my interests and in my curiosity and for supporting me on my way. I know that you would always help me without any compromise. Perhaps I don't show it often enough, but I'm so grateful for that.