DISSERTATION | DOCTORAL THESIS

Titel | Title

Interferometric Beam Shaping for Three Dimensional MINFLUX Localisation Microscopy

verfasst von | submitted by

Dipl.-Ing. Maximilian Korbinian Geismann BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 794 685 490
Dissertationsgebiet lt. Studienblatt | Degree programme as it appears on the student record sheet:	Molekulare Biologie
Betreut von | Supervisor:	Dr. Ing. Francisco Balzarotti

Acknowledgements

I would like to thank Francisco Balzarotti for his supervision and guidance.

I also thank my former and current lab mates: Sebastian Isbaner, who contributed to building the custom MINFLUX system in the early stages, Alba Gomez-Segalas, who was heavily involved in the building process in the later stages and who measured and analysed the dual-emitter tracking data, Alessandro Passera and Anna Deputatova, who designed and folded the DNA-origamis, Mehrta Shirzadian, who provided support for the CRB calculations and Jakub Dokulil, who re-designed the z-stabilisation module. Also, I thank Eva Wiedemann and Karina Arellano Ayala for their support in preparing samples.

I thank my thesis advisory committee, Gerhard Schütz, Ilaria Testa and Thomas Juffmann, as well as the VBC training team, Eva Schmid and Chiara Ceriotti, for their advice and oversight. Furthermore, I thank all the people from the facilities on campus that provided excellent support: Bio-Optics, the Mechanical Engineering Center, Plant Sciences and Facility Management.

A special thanks goes to Marco Incarbone, Mattia Dona and Andreas Hagmüller, as well as my peers from the PhD selection for their non-scientific support.

Most importantly, I wholeheartedly thank my wife for her perpetual support and understanding; despite all the hardships we had to go through.

Contents

Abstract

English Version

Accurately manipulating the shape and position of light beams is essential to a wide variety of applications ranging from optical communications, optical tweezers and microscopes. Especially in the life sciences, fluorescence microscopy is an invaluable tool for its ability to precisely label biological samples –ranging from single proteins to whole living animals– with fluorophores. However, studying dense samples with spatial features below the diffraction limit of light of approximately 200 to 300 nanometres requires super-resolution techniques.

Prominently, a recent method termed MINFLUX demonstrated how the concept of single molecule localisation can be significantly boosted in precision utilising shaped beams: the molecule location is estimated through photon counts collected during an iterative sequence of excitation beams, that usually feature an intensity 'zero', and their positions. Remarkably, MINFLUX allows isotropic single-digit nanometre precision in all three dimensions (3D) which, however, requires axially deflecting beams. This poses a technological challenge within a typical sub-millisecond MINFLUX cycle.

In this work, I study a beam shaping approach based on fast dynamic polarisation interferometry that allows manipulation of beam shapes and their intensity zeros, notably the axial deflection of a bottle beam zero and how it can be used for enabling 3D MINFLUX localisation. Furthermore, I present a versatile realisation and characterisation of such a interferometric beam shaper working on four parallel optical paths, which forms the excitation module of a fully fledged 3D MINFLUX microscope. The setup's capabilities are demonstrated through 3D MINFLUX imaging of DNA origami structures and nuclear pore complexes using novel MINFLUX schemes that mix different beam shapes within one cycle for improved localisation precision as well as 3D MINFLUX tracking of two emitters by rapidly interleaving excitation patterns of two wavelengths.

German Version

Die genaue Beeinflussung der Form und Position von Lichtstrahlen ist für eine Vielzahl von Anwendungen, die von optischer Kommunikation über optische Fallen und Mikroskopen reicht, wichtig. Besonders in den Biowissenschaften ist die Fluoreszenzmikroskopie ein unschätzbares Werkzeug, da es die präzise Markierung biologischer Proben, seien es einzelne Proteine oder gar ganze lebende Tiere, mit Fluorophoren ermöglicht. Die Untersuchung dichter Proben mit räumlichen Strukturen unterhalb der Beugungsgrenze des Lichts von etwa 200 bis 300 Nanometern erfordert jedoch Superauflösungstechniken. Eine kürzlich vorgestellte Methode mit der Bezeichnung "MINFLUX" hat gezeigt, wie das Konzept der Einzelmolekül-Lokalisierung mit Hilfe von geformten Strahlen erheblich verbessert werden kann: die Position des Moleküls wird anhand der Anzahl der Photonen effizient geschätzt, die während einer iterativen Abfolge von Anregungsstrahlen, die in der Regel ein "Nullstelle" im Intensitätsprofil aufweisen, detektiert werden. Bemerkenswert ist, dass MINFLUX eine isotrope Präzision im einstelligen Nanometerbereich in allen drei Dimensionen (3D) erreichen kann, was jedoch axiale Ablenkung der Strahlen erfordert. Dies stellt eine technologische Herausforderung für die Realisierung eines typischen MINFLUX-Zyklus im Sub-Millisekundenbereich dar. In dieser Arbeit untersuche ich einen Ansatz zur Strahlformung, der auf schneller dynamischer Polarisationsinterferometrie basiert. Sie ermöglicht die Manipulation von Strahlformen und deren Intensitätsnullstellen, insbesondere die axiale Ablenkung der Nullstelle eines Flaschenstrahls ("bottle beam") die auch für die 3D MINFLUX Lokalisierung verwendet werden kann. Darüber hinaus präsentiere ich eine flexible Realisierung und Charakterisierung eines solchen inteferometrischen Strahlformers, der die Strahlen vier paralleler optischer Pfade formen kann und das Anregungsmodul eines vollwertigen 3D-MINFLUX Mikroskops bildet. Die Fähigkeiten des Aufbaus werden anhand der 3D-MINFLUX-Abbildung von DNA-Origami-Strukturen und Zellkernporenkomplexen unter Verwendung neuer MINFLUX-Schemata, die verschiedene Strahlformen innerhalb eines Zyklus mischen, um die Lokalisierungspräzision zu erhöhen, sowie die gleichzeitige 3D-MINFLUX Verfolgung zweier Fluorophore durch schnelles Schalten zwischen Anregungsmustern zweier Wellenlängen demonstriert.

1. Introduction

This thesis describes a technological development in the field of light microscopy: a technique to manipulate the shape of (tightly) focused laser beams based on the interference of light –a very fundamental phenomenon in physics. This approach was implemented as part of a real microscope and a variety of beam shape manipulations can be achieved this way. One of those is of particular interest for so-called 'MINFLUX' microscopes as it offers a compelling alternative to current technologies that quickly, i.e., within tens of microseconds, move focused beams "up and down" over a range of a couple of hundred nanometres.

MINFLUX is a rather novel tool in the microscopy world. As elaborated in Section 1.2, MINFLUX generates more informative data when compared to other methods by employing shaped beams of light in specific patterns. Ultimately, MINFLUX allows to localise single fluorescently emitting molecules with nanometre precision within milli- or submilliseconds. The presented technique enables the pattern formation required for acquiring three-dimensional MINFLUX data (Section 2). It is studied thoroughly and implemented as part of a 3D MINFLUX microscope (Section 3) which was used to acquire 3D images of small synthetic as well as cellular structures and 3D trajectories of molecules moving rapidly within a 20 nm region (Section 4).

1.1. Microscopy

Microscopy stems from the Greek words μικρος, 'small', and σκοπειν, 'to observe'; it is the study of things that are smaller than what can be seen with the naked eye. The first microscopes consisted of a single lens[1] and allowed pioneering microscopists such as Antoni Leeuwenhoek and Robert Hooke to make a plethora of discoveries in the 17th century including 'cells' that make up all of known life[1, 2]. The working principle was the following: the studied specimen or sample was put on the 'stage' where it was illuminated by the surrounding light or by a designated light source and the reflected light was gathered by the lens to form a magnified virtual image that would be observed by eye.

Nowadays, light microscopes consist of many more parts and capture images digitally on cameras or other light sensitive detectors and are displayed to the microscopist on computer screens. Furthermore, today's life science researchers make heavy use of fluorescence microscopy[3] where biological structures are studied by tagging them with molecules that upon excitation through (normally laser) light, will themselves emit light through electron transitions; they 'fluoresce'[4, 5]. The emitted light has a shifted 'wavelength' or colour, such that it can be collected separately from its excitation light. Hence, only specific sites that are 'labelled' or 'stained' with fluorescent molecules –also called 'fluorophores'– will be visible through the microscope. Biologists and biochemists have established an astonishing repertoire of methods to label sites of interest within biological specimen. These range from immuno-labelling with antibodies –typically grown in live animals like rabbits– that bind to antigens of interest[6, 7, 8], to genetically encoding fluorophores[9] like the green fluorescent protein[10]: any genetically modifiable organism[2] can be reprogrammed to produce the fluorescent protein attached to structures of interest itself. These methods are applicable to chemically 'fixed', i.e., dead but structurally preserved, as well as living cells. Ultimately, all these traits render fluorescence microscopy an invaluable tool

[1] An optical lens is (usually) a piece of glass that is formed in a specific way to bundle light.
[2] After the advent of CRISPR gene editing[11], barely any limits remain.

for life science and we will assume this modality for the remainder of the work, unless noted otherwise.

1.1.1. The Diffraction Limit

The optical microscope's ability to resolve tiny structures is limited due to diffraction: the bending of light around obstacles that block its propagation. Optical elements, like lenses, have finite size or 'aperture' and thus naturally "block" light that falls beyond their edges. Abbe's diffraction limit[12, 13] states that two points emitting light of some wavelength λ, imaged by a microscope with a numerical aperture NA –a characteristic property usually determined by the objective lens[3]– can be distinguished if they are a minimum distance of

$$d_{\min} = \frac{\lambda}{2\mathrm{NA}} \tag{1.1}$$

apart[4]. For visible light with $\lambda = 400\,\mathrm{nm}$ to $700\,\mathrm{nm}$ and realisable numerical apertures of NA = 1.4 to 1.6 one typically expects the limit to be around 200 nm. For completeness' sake, the other important statement of resolution theory, Rayleigh's criterion[14], shall be mentioned. It demands that two emitters need to be separated by a minimum angle, $\theta_{\min}$, in order to be resolved[5]:

$$\theta_{\min} = 1.22\frac{\lambda}{D}\,, \tag{1.2}$$

where the angle is taken with respect to the centre of a circular aperture of diameter D. Multiplying Equation (1.2) by the focal length of the microscope's objective and reformulating the above equation yields a minimum distance of $d_{\min} = 0.61\frac{\lambda}{\mathrm{NA}}$[6], which is slightly larger than Abbe's limit given in Equation (1.1).

1.1.2. Super-Resolution Microscopy

Going beyond the diffraction limit has been the subject of intense efforts. The invention of the confocal microscope[15, 16] could be considered a first step towards that goal opening the field of 'super-resolution' imaging by slightly improving the lateral resolution over conventional microscopes[17]. It works by scanning a focused excitation beam across the sample and recording the resulting fluorescence intensity for each location on a photosensitive device through a small aperture: the 'pinhole'. The "grey scale"[7] image of the fluorescently labelled structures is then built point by point. Collecting the signal through the pinhole not only improves the lateral (xy-plane) resolution but also allows improved 3D imaging through 'optical sectioning': by blocking out-of-focus light[18], the signal originates from a laterally as well as axially (z-axis) confined region; the 'detection volume'. The 3D image is then generated by stacking 2D image slices –each acquired at a different depth by, e.g., moving the stage up and down– on top of each other. It is noted that the detection volume (just like the focused excitation beam) is always

[3] The objective lens is the main lens of a microscope strongly focusing the excitation light into the sample but also collecting fluorescence that ultimately forms the image.

[4] Abbe derives this result by contemplating about grating structures as samples and how their diffraction orders are captured by the objective.

[5] Rayleigh considers the diffraction patterns generated by the light of a point source propagating through a circular aperture as a function of the angle θ; they are first order Bessel functions of the first kind. The patterns are resolved when the maximum of one lies on top of the first minimum of the other.

[6] $D = 2\mathrm{NA}f$.

[7] Usually, the image is coloured according to the employed fluorophore's emission wavelength. Multicolour imaging is possible but requires multiple fluorophore species to be present in the sample. They must differ in their spectral properties regarding excitation, emission or both.

enlarged along the axial dimension when compared to its lateral extent leading to a worse axial resolution.

The resolution of light microscopes was improved even further –down to roughly 20 nm– by the three classic super-resolution methods: STimulated Emission Depletion (STED)[19, 20], Stochastic Optical Reconstruction Microscopy (STORM)[21] and Photo-Activated Localisation Microscopy (PALM)[22].

The first method, STED, is based on the confocal microscope. In addition to the excitation beam, however, there is an additional infrared 'depletion' beam which is manipulated such that, when focused through the objective, is formed like a doughnut: at the centre the intensity of light is ideally zero, while around it there is high intensity in a ring shaped region. Excited fluorophores exposed to the depletion beam will transition back to their unexcited state without emitting light[8]. The excitation beam is aligned such that its maximum of intensity sits right in the centre of the depletion beam. Consequently, the region from which fluorescence can originate, given by the spatial extent of the excitation beam, is reduced by the depletion beam. The minimally resolvable distance decreases[23]:

$$d_{\text{STED}} \propto \frac{\lambda}{2\text{NA}\sqrt{1 + I_{\max}/I_S}}\,, \tag{1.3}$$

where $I_{\max}$ is the maximum intensity of the employed depletion beam and I_S is the saturation intensity at which the depletion probability is $1/e$. In principle, the resolution can be improved arbitrarily by increasing $I_{\max}/I_S$, however, practical limitations apply: I_S is given by the fluorophore and an increase in $I_{\max}$ is limited by noise, an imperfect doughnut shape and 'photo-bleaching' due to the high intensity, i.e., the fluorophore irrecoverably loses its ability to emit light.

The other two techniques, STORM and PALM follow a different approach termed 'single molecule localisation'. Instead of a point-detector, these methods use cameras on which the light of an emitting fluorophore will appear as a blob. The shape of the blob is called the Point Spread Function (PSF): the single molecule acts as a point source but its image on the detector becomes a blob as the emitted light spreads out due to diffraction experienced while propagating to the detector. While the diffraction limit prevents to distinguish close-by emitters, the position of a single emitting fluorophore can be determined with great precision by fitting the single blob to a model function (typically a Gaussian). The precision of the fitted location is limited by shot noise which is a consequence of the Poissonian emission process: light is emitted in tiny packets or quanta called 'photons'. The amount of photons emitted by a fluorophore exposed to a certain excitation intensity, I_{exc}, is a random number N that follows the Poissonian distribution. The probability for N emitted photons is:

$$P(N; I_{\text{exc}}) = \frac{(\alpha I_{\text{exc}})^N e^{-\alpha I_{\text{exc}}}}{N!}\,, \tag{1.4}$$

where α is a scaling factor, determined by the physical properties of the fluorophore, that connects the intensity to the average count of emitted photons[9]. It follows that the localisation precision, measured by the standard deviation, scales as[24, 25]

$$\sigma \propto \frac{\lambda}{\sqrt{N}}\,. \tag{1.5}$$

Thus, the emitter can be localised with arbitrary precision even though the square-root makes it an inefficient process: doubling the precision requires four times the photons ($\sigma/2 \propto 1/(2\sqrt{N}) =$

[8] They are being "depleted" of their excitation.
[9] The mean of the distribution is given by αI_{exc}.

Cumulative Probability	Diameter (σ)		
	1D	2D	3D
0.50	1.35	2.36	3.08
0.90	3.29	4.29	5.00
0.95	3.92	4.90	5.60
0.99	5.16	6.07	6.74

Table 1.1.: Required length/area/volume in terms of diameter in units of standard deviations (σ) to reach a certain cumulative probability and hence fully enclose a certain percentage of localisations in the case of 1D, 2D and 3D Gaussians.

$1/\sqrt{4N}$). Most importantly though, the fluorophore has a limited photon budget. It can only emit a limited number of photons before it bleaches.

Nonetheless, the achievable precision can be used to distinguish two emitters at a distance below the diffraction limit by localising them sequentially. The localisation precision is, however, not synonymous with resolution! Assuming a Gaussian distribution of localisations around the true emitter position it might take a region with a diameter of several standard deviations to capture enough localisations to clearly distinguish close-by emitters; especially when localising in 3D, see Table 1.1. The fundamental challenge of single molecule localisation is to have only one fluorophore in a (diffraction limited) neighbourhood emit at any given time. This is where the two methods slightly differ: STORM exposes the sample to a chemical environment[10] in which all fluorophores stochastically[11] transition between emitting or on- and non-emitting or off-states[12]. Alternatively, transiently binding fluorescently labelled probes that freely diffuse in the sample can be used[27, 28]. These approaches are termed (DNA-)PAINT and their strength is the infinite pool of fluorophores; if a one gets bleached it will eventually be replaced by another due to transient binding, allowing to reach very high photon numbers and resolution[29]. In both cases, the transitions are tuned such that there are no close-by molecules in the on-state simultaneously (with high probability). PALM, on the other hand, employs photo-activatable proteins as fluorophores: the switch to the on-state is induced by activating them first with Ultra-Violet (UV) light. This is again a random process such that not all fluorophores will turn on immediately. The UV exposure is tuned such that there are no close-by fluorophores in the on-state and the activation step is repeated after enough fluorophores have bleached (due to excitation) such that the next round of fluorophores can be turned on.

In both methods, the camera records a movie with an ever changing but sparse distribution of diffraction limited blobs. Each blob's occurence, spread over several frames of the movie, is used to localise its corresponding fluorophore. Eventually, most fluorophores will have been localised at least once and the resulting image is formed out of dots at the localisation coordinates that are optionally blurred to give the illusion of a continuous image. The observed region can reach tens of microns[13] and is called the Field of View (FOV). Therefore, STORM and PALM are also termed as Wide-Field (WF) techniques as opposed to "single point" systems like the confocal and the STED microscope.

Many other microscopy techniques that break the diffraction limit exist: Structured illumination microscopy uses a sequence of patterned excitation light to improve the classic resolution by a factor of two[30]. Super-resolution optical fluctuation imaging calculates super-resolved images from higher-order cumulants (which are similar to correlation functions) obtained from

[10] This environment typically prevents live cell imaging that is otherwise possible with STED and PALM[26].
[11] Here: synonymous with 'randomly'.
[12] This is also called 'blinking' since the transition from the on-state to the off-state is initially different from bleaching. The molecule typically blinks several times before it bleaches eventually.
[13] One micron equals one micrometre.

tiny fluctuations in the fluorophores' emission[31]. Expansion microscopy[32] is actually not a microscopy technique: the sample is embedded in a gel that captures the structures via a net of polymers. The gel is then isotropically[14] expanded; the structures are pulled apart and thus enlarged by an expansion factor of typically two to ten times. Then, any conventional microscopy technique can be used for imaging. The resolution, however, is enhanced by the expansion factor. Rotating coherent scattering microscopy[33, 34] requires no fluorescence, a so-called 'label-free' technique: it uses speckle patterns[15] that originate from multiply scattered laser light transmitted through the sample. Illuminating the sample from all angles and summing the according speckle patterns yields super-resolved images. Interferometric scattering microscopy[35] is another label-free technique that is predominantly used for sensing single particles and tracking them with high speed. It works by interfering a tiny fraction of the illuminating light that has interacted with the particle[16] with the illumination itself. Finally, it is noted that microscopes do not need to operate with light at all: techniques like electron microscopy[36], where electron beams with tiny wavelengths instead of light are utilised, or atomic force microscopy[37], where a tiny cantilever "mechanically" probes the atomic force field, have long been adapted and used by life science researchers to reveal structures as well as dynamics of bio-molecules and bio-molecular complexes with resolutions down to the Ångström level.

1.2. Minflux

While the classic super-resolution methods have not lost their relevance, a recent single molecule localisation method called 'MINFLUX' was able to open a new chapter in the field of optical super-resolution microscopy[38]. It demonstrated unprecedented spatio-temporal localisation precision for imaging and tracking and many subsequent works have presented impactful advancements and applications ever since [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53] It also inspired a series of related works that can be seen as the combination of MINFLUX and structured illumination yielding a two-fold localisation precision improvement [54, 55, 56, 57, 58].

MINFLUX is a ratiometric localisation strategy for single emitters. Instead of using a camera and centroid fitting as other single molecule localisation techniques, it employs a confocal single point detector to acquire a set of photon counts that are generated by a series of excitations with shaped beams. The location estimation from these photon counts works with remarkable efficiency when compared to camera-based localisation as hinted by the two interpretations of the MINFLUX acronym[38]: 'minimal photon flux' or 'maximally informative luminescence excitation'. Every photon collected through MINFLUX carries more information induced by the shaped excitation beams. Hence, MINFLUX is able to achieve superior resolution in imaging[38, 40], high tracking speeds[17][38, 39] or the use of less excitation power to prevent photo-toxicity and bleaching[18]. Furthermore, MINFLUX' precision is not proportional to the wavelength opposing previously discussed scaling laws (Equations (1.1), (1.3) and (1.5)) and it allows isotropic precision in all three dimensions, hence, also the optical axis, a feature that other microscopy techniques do not present.

The basic concept is similar to triangulation. Consider the following 1D toy example[38]: the goal is to determine the position of a point, x_m, within an interval $[-L/2, L/2]$ of length L, but we are only provided with a multiple of the squared distances between the point and the two edges of the interval $f_0 = \alpha|x_m + L/2|^2$ and $f_1 = \alpha|x_m - L/2|^2$. The solution is straightforward;

[14] 'Isotropic' means 'the same in all directions'.

[15] Speckles are the grainy intensity distribution that can be observed for any coherent light source due to interference.

[16] The interaction is based on the small difference of the refractive index between the particle and its surrounding.

[17] Fluorophores have a given emission rate. By requiring less photons to achieve sufficient localisation precision, the wait time between localisations decreases.

[18] The lower the excitation power the lower the bleaching rate, i.e., the molecule stays on for longer.

if either one of f_0 or f_1 is zero, the point must lie on the corresponding edge, if $f_0 = f_1$, the point lies in the centre of the interval. Any other location can be calculated by forming the ratio f_0/f_1, such that the unknown factor α cancels out, and solving for x_m within the interval:

$$\begin{aligned} \frac{f_0}{f_1} =& \frac{|x_m + L/2|^2}{|x_m - L/2|^2} \\ \Rightarrow x_m =& \frac{L}{2} \frac{(1 \pm \sqrt{f_0/f_1})^2}{(1 - \sqrt{f_0/f_1})(1 + \sqrt{f_0/f_1})} \\ \overset{|x_m| \leq L/2}{\Rightarrow} x_m =& \frac{L}{2} \frac{(1 - \sqrt{f_0/f_1})}{(1 + \sqrt{f_0/f_1})} \end{aligned} \tag{1.6}$$

MINFLUX realises this localisation procedure as follows: a fluorophore is sequentially exposed to two excitation beams with parabolic intensity profiles $I(x) = I_0 x^2$ centred at $\pm L/2$. These beams are usually doughnut shaped as they are well approximated by a parabolic intensity profile close to their centres. The two corresponding photon counts $n_0 = \alpha I(x_m + L/2)$ and $n_1 = \alpha I(x_m - L/2)$ are proportional to the squared distance between itself and the two minima of intensity. Hence, the molecule would be perfectly localised by Equation (1.6) (for $f_0 = n_0$ and $f_1 = n_1$). However, the photon counts are only exact on average due to the Poissonian emission process (Equation (1.4)) which turns the localisation into an estimation problem.

Solving an estimation problem amounts to finding or estimating unknown parameters $\boldsymbol{\theta}$ of a known probability distribution given observations or samples $\boldsymbol{X}$ drawn from said distribution. In our case, the random variables are the photon counts that follow a Poissonian distribution. The parameter to be estimated is the molecule position $\theta = x_m$ that influences the mean of the Poissonians $\lambda_i(x_m)$.

There are many ways to estimate but most commonly one opts for the Maximum Likelihood Estimator (MLE): it chooses the parameter that maximises the likelihood, i.e., the probability, of the observations stemming from the distribution with the estimated parameter:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\theta} \mathcal{L}(\boldsymbol{\theta}|\boldsymbol{X}) \,.$$

It turns out that Equation (1.6) is already in the from of the MLE[59]

$$\hat{x}_m = \frac{L}{2}(1 - \sqrt{n_0/n_1})/(1 + \sqrt{n_0/n_1}) \,.$$

However, instead of the photon counts, one usually defines the so-called 'p-functions' as the random variables:

$$p_i(x_m) = \frac{\lambda_i(x_m)}{\sum_j \lambda_j(x_m)} = \frac{I_i(x_m)}{\sum_j I_j(x_m)} \tag{1.7}$$

The probability to acquire a total of N photons is given by a multinomial distribution:

$$\begin{aligned} P(n_0, n_1, \ldots n_{K-1}|N) =& \frac{\prod_{i=0}^{K-1} P(n_i)}{P(N)} = \\ =& \frac{N!}{n_0! n_1! \ldots n_{K-1}!} \prod_{i=0}^{K-1} p_i^{n_i} \,, \end{aligned}$$

where the p-functions act as the success probabilities.

The MLE for each p-function is

$$\hat{p}_i = \frac{n_i}{\sum_i n_i} = \frac{n_i}{N} \,.$$

Hence, with $\frac{\hat{p}_0}{\hat{p}_1} = \frac{n_0}{n_1}$ and due to functional invariance of the MLE, Equation (1.6) is also the MLE of the molecule position for two exposures with quadratic intensity profiles as a function of the p-functions p_i.

The restriction to a given total photon count that brings along the p-functions parametrisation becomes crucial for calculating the so-called Cramér-Rao Bound (CRB)[60, 61][19] which determines the smallest possible standard deviation achievable by any unbiased estimator: when estimating, for example, the same location of a molecule several times from different realisations of the random variables, the estimator will yield different locations x_m. The CRB determines how large the spread of locations will be, if we had the ideal unbiased estimator. An estimator is unbiased if its expected value (over many estimations) is the correct parameter (here: location). This property is important as, for instance, a constant estimator that always yields zero, $\hat{x}_m^{(\text{constant})} = 0$, would have no spread or infinite precision but would be biased, i.e., not accurate, and useless. However, it might be possible to find biased estimators performing better than the CRB that exhibit negligible bias with respect to their improved precision. The MLE is 'consistent' and 'efficient' meaning it has no bias and approaches the CRB for a sufficiently large number of observations, i.e., photon counts.

The CRB for a vector-valued parameter $\boldsymbol{\theta}$ is given through the covariance matrix Σ_{CRB}. For example, for a 3D molecule-location $\boldsymbol{\theta} = \boldsymbol{r}_m = (x_m, y_m, z_m)^\top$ and

$$\Sigma_{\text{CRB}}(\boldsymbol{r}) = \begin{pmatrix} \sigma_x^2(\boldsymbol{r}) & \sigma_{xy}(\boldsymbol{r}) & \sigma_{xz}(\boldsymbol{r}) \\ \sigma_{xy}(\boldsymbol{r}) & \sigma_y^2(\boldsymbol{r}) & \sigma_{yz}(\boldsymbol{r}) \\ \sigma_{xz}(\boldsymbol{r}) & \sigma_{yz}(\boldsymbol{r}) & \sigma_z^2(\boldsymbol{r}) \end{pmatrix} .$$

The CRB is calculated via the Fisher information $\mathcal{F}_{\boldsymbol{\theta}}$:

$$\Sigma_{\text{CRB}}(\boldsymbol{\theta}) = \mathcal{F}_{\boldsymbol{\theta}}^{-1} ,$$

where

$$(\mathcal{F}_{\boldsymbol{\theta}})_{ij} = \mathbb{E}_{\boldsymbol{\theta}} \left[\left(\frac{\partial}{\partial \theta_i} \ln(\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{X})) \right) \cdot \left(\frac{\partial}{\partial \theta_j} \ln(\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{X})) \right) \right] . \tag{1.8}$$

In our case, the parameters θ_i are the p-functions, where, however, the dimensionality is reduced by one since for K p-functions there are only $K-1$ independent ones because

$$\hat{p}_j = \frac{N - \sum_{i \neq j} n_i}{N} = 1 - \sum_{i \neq j} \hat{p}_i \, . \tag{1.9}$$

Hence, the CRB will be given by a $(K-1) \times (K-1)$ matrix providing lower bounds on the p-function estimates' precision. We are, however, interested in the best possible precision of the localisation in real space. To obtain the corresponding CRB, the Fisher information matrix can be reparametrised through a transformation with the Jacobian matrix, J, as

$$\mathcal{F}_{\boldsymbol{r}_m} = J^\top \mathcal{F}_{\boldsymbol{p}} J \, , \tag{1.10}$$

where

$$J_{ij} = \frac{\partial p_i}{\partial r_j}$$

is a $(K-1) \times 3$ matrix and $\boldsymbol{p}$ is the vector of $K-1$ p-functions (where one is expressed as given by Equation (1.9)).

[19] Sometimes more descriptively referred to as the 'Cramér-Rao lower bound'.

For the 1D case with two exposures with quadratic intensity profiles the CRB is given by

$$\sigma_x(x) = \frac{L}{4\sqrt{N}} \left(1 + \left(\frac{x}{L/2}\right)^2\right) . \tag{1.11}$$

When comparing this to the camera-based localisation precision, Equation (1.5), one thing stands out: while the $1/\sqrt{N}$ scaling is also present here, there is no wavelength dependence. The geometric quantity L, the distance between the two minima of the beams, appears in its place. In contrast to the wavelength, which is normally dictated by the fluorophore through its excitation spectrum, L can be chosen almost arbitrarily[20]. For a molecule at the origin of the interval the CRB, $\sigma_x(0) = L/(4\sqrt{N})$, reaches its minimum and the precision improves linearly when shrinking L. In general, one observes the $\sigma \propto L/\sqrt{N}$ behaviour at the origin also for other MINFLUX schemes that use more exposures to localise in higher dimensions (see the supplementary information of [38] for more examples). Furthermore, there do exist beam shapes that also exhibit a quadratic intensity profile along the optical axis; Equation (1.11) also applies in this case when exchanging x for z. Consequently, MINFLUX is able to localise isotropically in all 3D dimensions in contrast to camera-based localisation techniques[62].

Using the smallest possible L is desirable as the precision increases. Apart from approaches that use pixelated detectors[44], conventional MINFLUX is only able to localise within its excitation pattern (the "interval") with the extent given by L[21]. Therefore, using a small L might not be feasible as the location of the to-be-localised molecule might not be known well enough to position the excitation pattern. This motivates the concept of 'iterative MINFLUX' where the pattern starts with a large L[22] and is successively re-centred and shrunk. The first few iterations typically collect far fewer photons than the last iteration which is the one that produces the localisation featuring the smallest L and thus the highest precision.

In real measurements, the collected photon counts include background photons which are all photons that do not originate from the fluorophore that is being localised. Possible sources are dark counts from the detectors, auto-fluorescence of the sample or the objective's immersion oil, other freely diffusion fluorophores passing by etc. The theoretical treatment becomes more involved. Under the simplifying assumption that the background is the same for all exposures of a pattern and that it is proportional to the employed power of the excitation, the p-functions used for the CRB calculation become

$$p_i(\boldsymbol{r}_m) = \frac{\mathrm{SBR}(\boldsymbol{r}_m)}{1 + \mathrm{SBR}(\boldsymbol{r}_m)} p_i^{(0)}(\boldsymbol{r}_m) + \frac{1}{K} \frac{1}{1 + \mathrm{SBR}(\boldsymbol{r}_m)} , \tag{1.12}$$

where $p_i^{(0)}$ is the p-function as defined in Equation (1.7) and SBR is the signal-to-background ratio. To give an example: in the case of 3000 collected photons where 2000 come from the fluorophore and 1000 from background, SBR = 2. The SBR might change from fluorophore to fluorophore. Therefore, the MLE should regard the SBR as unknown and estimate it as well. To calculate the precision in this case, the Fisher information matrix and hence the CRB matrix are extended by one more dimension. A low SBR (known or unknown) complicates the estimation problem and thus the CRB will be generally higher. However, when shrinking L the acquired signal will diminish as the molecule is exposed to smaller intensities[23] while the background stays

[20] It must not be larger than the range for which the quadratic approximation of the intensity of the beam is valid.

[21] There could be different values for L along different dimensions making the "inside" of an excitation pattern, in general, an ellipsoid. Usually, though, spheres are used.

[22] Usually, the first iteration consists of a pattern of Gaussian beams as they can cover a larger area for the first localisation. MINFLUX with Gaussians (that do not feature a parabolic intensity minimum) is less precise, however. Therefore, the following iterations use doughnut shaped beams.

[23] Assuming a pattern with parabolic beams.

the same (it only depends on the overall power of the excitation). In practice, it is thus necessary to find the right trade-off between SBR and L, see Section 4.1. Still, under high SBR conditions, it has been shown[38] that MINFLUX requires $20\times$ less photons than camera-based techniques to reach a given precision or, alternatively, improves the localisation precision five-fold for the same amount of photons used.

Lastly, some more comments on the general MINFLUX operation principles: for imaging, MINFLUX can be used in the STORM or PALM modalities with the major difference being that the detection only covers a confocal volume. The activating UV light is focused and centred with the detection volume and it is assumed that activated fluorophores appear within a sphere of about 400 nm diameter. In order, to generate images over larger FOVs, the confocal detection volume and the excitation beam needs to move (similar to a confocal microscope) in steps with enough dwell time to allow all or most of the molecules at that point to be activated and localised one-by-one. For tracking molecule trajectories, the fluorophores stay on continuously until they bleach. Tracking in small areas is usually achieved with a static, i.e., non-iterative MINFLUX scheme. However, when following a molecule over larger areas the detection volume is required to move along. Typical emission rates (they depend on the excitation power and the type of fluorophore) are in the range of 50 kHz to 300 kHz. While the MLE is not suited for localising with counts of less than a 100 photons, other estimators exist that can handle these small counts[38]. Therefore, the time to run through the full sequence of exposures of a pattern can be as short as 50 µs to 200 µs and still be feasible which results in a tracking speed of 20 kHz to 5 kHz. Pattern times of 100 µs to 200 µs are typical even in imaging applications; the photon counts of several rounds of exposures are added up or 'binned' (usually several thousand) such that the MLE can be employed. The same procedure can also be used in tracking to improve the localisation precision by sacrificing temporal resolution.

2. Polarisation Interferometry for Beam Shape Manipulation

This section discusses how focused light beams can be dynamically manipulated through the interference of different beam shapes. The change in shape of beams that feature a 'zero', which is a region with no light intensity, can move this zero. Several cases are studied. The interferometric axial movement of the 'tophat' beam's zero, however, is of particular interest because it offers a novel way to realise 3D MINFLUX. Finally, it presents an implementation that is based on a polarisation interferometer.

2.1. Electromagnetic Waves, Polarisation and Interference

2.1.1. Electromagnetic Waves

Maxwell's equations fundamentally describe electromagnetism. In isotropic linear media without any free charges or current, Maxwell's equations can be reduced to the wave equation of the electric field[63]

$$\nabla^2 \boldsymbol{E}(\boldsymbol{x},t) - \frac{\varepsilon}{c^2}\frac{\partial \boldsymbol{E}(\boldsymbol{x},t)}{\partial t^2} = \boldsymbol{0}\,, \tag{2.1}$$

where c is the speed of light in vacuum, ε is the dielectric permittivity, $\boldsymbol{x} = \begin{pmatrix} x & y & z \end{pmatrix}^\top$ is the spatial coordinate, t is the time and $\boldsymbol{E}$ is the electric field. The equation is linear and is readily solved in terms of complex[1] Fourier components in space and time by

$$\boldsymbol{E}(\boldsymbol{x},t) = \boldsymbol{E}(\boldsymbol{k},\omega)\exp(-i(\boldsymbol{k}\cdot\boldsymbol{x} - \omega t))\,, \tag{2.2}$$

with ω being the temporal frequency, $\boldsymbol{k}$ the wave-vector whose magnitude, $|\boldsymbol{k}| = k$, is the spatial frequency along its direction and

$$\boldsymbol{E}(\boldsymbol{k},\omega) = \begin{pmatrix} E_x(\boldsymbol{k},\omega) \\ E_y(\boldsymbol{k},\omega) \\ E_z(\boldsymbol{k},\omega) \end{pmatrix}$$

is the complex amplitude of the electric field. Solutions of the form found in Eq. (2.2) are called plane waves as they have flat phase fronts orthogonal to their propagation direction (given by $\boldsymbol{k}$). Furthermore, $k = \frac{n\tau}{\lambda} = \frac{n\omega}{c}$[2], where $n = \sqrt{\varepsilon}$ is the refractive index and λ is the wavelength. Plane waves have no transversal electric field component due to Gauss's law: take, for example, a plane wave propagating along z, i.e., $k_x = k_y = 0$ and $k_z > 0$. By substituting the solution Eq. (2.2) into to Gauss's law $\nabla \cdot \boldsymbol{E} = 0$, it follows that $E_z = 0$. The convention is to assume that light and light beams mainly propagate along the z axis which is then referred to as the 'optical axis'.

[1] To arrive at physically meaningful, i.e., real quantities, the convention dictates to take the real part of the solution.

[2] $\tau := 2\pi$, see Appendix A.

Another important solution under the paraxial approximation is the Gaussian beam with the complex electric field amplitude along some lateral direction

$$E(\rho, z) = E_0 \frac{w_0}{w(z)} \exp\left(-\frac{\rho^2}{w(z)^2}\right) \exp\left(-ikz - ik\frac{\rho^2}{2R(z)} + i\phi_G(z)\right) , \tag{2.3}$$

where ρ is the radial distance, $E_0 = E(0,0)$ is the base electric field strength, $w(z)$ is the waist width with $w_0 = w(0)$, $R(z)$ is the radius of curvature and $\phi_G(z)$ is the Gouy phase. The lateral electric field strength follows a Gaussian envelope, hence the name. The lateral intensity is also Gaussian but with a different width: $\exp\left(-\frac{\mathbf{2}\rho^2}{w(z)^2}\right)$. Conventionally, the beam size is given by twice its $1/e^2$ intensity width, i.e., $2w_0$. At its waist, $z = 0$, $R \to \infty$ and hence the Gaussian beam behaves similar to a plane wave with a flat phase front (but Gaussian amplitude profile). This plane wave behaviour is usually assumed to hold over the Rayleigh range $z_R = w_0^2 k/2$. For a typical wavelength of 640 nm the Rayleigh range of a beam of the smallest $1/e^2$ diameter used in the system presented in this work of 1 mm is 1.23 m. Image relays or telescopes can usually be found within this range such that beam is thought to be plane-wave like or 'collimated' throughout the system excluding the sections in between the lenses of the telescopes where it is focused.

2.1.2. Polarisation

The polarisation describes the orientation of the electric field of transverse light waves. Plane waves and Gaussian beams fall under this category. As there are no transversal components, their fields can be described as a two dimensional vector, implicitly dropping the trivial component. For a given coordinate system that spans a plane orthogonal to $\boldsymbol{k}$, one speaks of 'vertical' or 's-polarised' and 'horizontal' or 'p-polarised'[3] states. In the context of optical systems, 'horizontal' normally means 'parallel to the optical table' on which the system is built.

The Jones formalism describes the polarisation state as a complex unit vector[63]:

$$\boldsymbol{J} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) e^{i\Delta\varphi} \end{pmatrix} , \tag{2.4}$$

where the angle θ describes the orientation of the main axis within or around which the electric field will oscillate and $\Delta\varphi$ is the relative phase between the two components of the electric field. A plane wave, Eq. (2.2), propagating along z would thus be rewritten as

$$\boldsymbol{E}(\boldsymbol{x}, t) = E_0 \boldsymbol{J}(\theta, \Delta\varphi) \exp(-i(kz - \omega t)) ,$$

where E_0 is the electric field strength. I extend the notion of the Jones vector to non-unit vectors including the electric field strength for the remainder of this work. Common polarisation states are listed and visualised in Table 2.1.

Optical elements that manipulate the polarisation state are represented by matrices in the Jones formalism. If light described by a Jones vector $\boldsymbol{J}^{\text{in}}$ interacts with an optical element described by the Jones matrix $\mathcal{M}$ (usually by propagating through it), the outgoing Jones vector is given by the vector-matrix multiplication (from the right):

$$\boldsymbol{J}^{\text{out}} = \mathcal{M}\boldsymbol{J}^{\text{in}}$$

The compound Jones matrix of multiple elements is given by the chained matrix multiplication in reverse order of traversal (the first element is on the right): $\mathcal{M}_n \dots \mathcal{M}_2 \mathcal{M}_1$. The most common

[3] The 's' and 'p' stem from the German 'senkrecht' and 'parallel' meaning 'orthogonal' and 'parallel' to the plane in which light in an ideal optical system (or part of it) propagates.

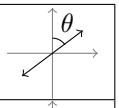
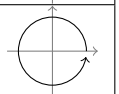
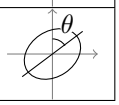
Linear	$\begin{pmatrix}\sin(\theta)\\ \cos(\theta)\end{pmatrix}$	
Circular	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}$	
Elliptical	$\begin{pmatrix}\sin(\theta)\\ \cos(\theta)e^{i\Delta\varphi}\end{pmatrix}$	

Table 2.1.: Polarisation states.

Rotation matrix	$R(\theta) = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$
Half-wave plate	$\text{HWP} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$
Quarter-wave plate	$\text{QWP} = \begin{pmatrix}1 & 0\\ 0 & -i\end{pmatrix}$
Horizontal linear polariser	$\text{HP} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$
Vertical linear polariser	$\text{VP} = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$
Electro-optic modulator	$\text{EOM}(\alpha) = \begin{pmatrix}1 & 0\\ 0 & e^{-i\alpha}\end{pmatrix}$

Table 2.2.: Common Jones matrices.

and important elements are 'retarders' or 'wave plates'. The name stems from the fact that they retard part of the light by imposing a phase usually measured in waves, i.e., rational fractions of the wavelength. The Quarter-Wave Plate (QWP) and the Half-Wave Plate (HWP) induce a phase of $\frac{1}{4}\tau\hat{=}90°$ and $\frac{1}{2}\tau\hat{=}180°$, respectively, along their slow axis. Usually their slow axis is rotated at 45° with respect to a linear input polarisation: the QWP turns the linear input to a circular output while the HWP flips the linear input to the orthogonal linear polarisation state. To capture the rotation a Jones matrix $\mathcal{M}$ is sandwiched between two rotation matrices: $R(\theta)\mathcal{M}R(-\theta)$, where $R(\theta)$ is the rotation matrix of angle θ, see Table 2.2. Another important element is the electro-optic modulator. It is a variable retarder and can act as a HWP, a QWP or any wave plate inbetween; see Section 2.5.1.

Another important tool for working with polarisation is the Poincaré sphere. It translates the parameters of the general Jones vector given in Eq. (2.4) into coordinates on a unit sphere, see Figure 2.1. The Poincaré sphere is equivalent to the Bloch sphere when thinking of the s- and p-polarisation as up- and down-spin, with the exception that the Poincaré sphere is rotated by 90° [4].

2.1.3. Interference

In physics, interference is a general phenomenon that occurs when two or more waves are superimposed coherently. In such a superposition the waves' amplitudes are added with respect to their phase which allows for phenomena like destructive interference, where the amplitudes of opposite phase cancel each other to yield no amplitude at all, or constructive interference, where

[4] The north pole of the Bloch sphere is on the equator of the Poincaré sphere.

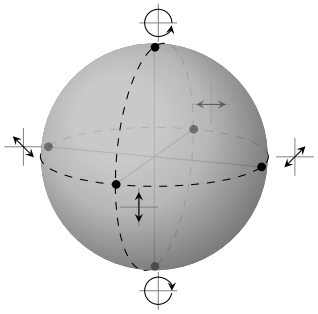


Figure 2.1.: Poincaré sphere.

the resulting amplitude is greater than each of the original amplitudes. In general, a superposition of waves can showcase patterns that drastically differ from its constituents' patterns.

For electromagnetic waves it is the electric field that is added. However, what we can directly observe, e.g., on a camera, is the intensity $I = |\boldsymbol{E}|^2$; a scalar quantity. The simplest case is the superposition of two plane waves oscillating along a common axis but with potentially distinct phases φ_1 and φ_2, which, in terms of intensity, yields:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\varphi_2 - \varphi_1)\,. \tag{2.5}$$

The term $2\sqrt{I_1 I_2}\cos(\varphi_2 - \varphi_1)$ is responsible for the interferometric effects. For example, for $I_1 = I_2$ and a phase difference $(\varphi_2 - \varphi_1) = \frac{1}{2}\tau$ we observe total destructive interference: $I = 0$. In an incoherent superposition the term is not present: the phase relation of the fields is chaotic and any interferometric effects average out even over a short time frame.

2.2. Focusing Beams of Light

In a regular confocal microscope, a collimated beam impinges on the objective lens and is strongly focused into a tiny spot of a shape that can be approximated laterally –in the xy-plane– and axially –along the z-axis– by a Gaussian intensity profile around the focal plane. A general phenomenon is that the profile is elongated axially along the propagation direction. Hence, while the confocal microscope is able to image in 3D by optical sectioning, its axial resolution will always be worse due to the larger excitation profile along z. The ideal beam at the input of the objective lens producing the tightest focus is a plane wave that fills the back aperture of the objective with homogeneous intensity and has a flat phase front. In reality, the flat wave front is approximated by a Gaussian beam that overfills the back aperture of the objective. The bigger the beam width, the better the approximation but the more power is lost due to clipping. The overfilling is quantified by the 'fill factor' $f_0 = w_0/r_{\mathrm{obj}}$, see [64] (Eq. (3.55)), where w_0 is the beam waist radius and r_{obj} is the objective's back aperture radius. Customarily, a fill factor $f_0 = 2$ is used[65] (page 210), which means 30 % of the power is lost.

However, if we alter the incoming phase front spatially, other beam shapes arise. These shapes are also referred to as the 'PSF' of the excitation beam[5]. One example, prominently used in 2D STED and MINFLUX, is the vortex or doughnut shaped beam generated by a spiral or vortex phase. The resulting intensity distribution can be calculated with numerical methods[66] and resembles an axially elongated doughnut. The "hole" is a line coincident with the optical axis along which there is zero intensity; for short: a zero line.

[5] In constrast to the 'emission PSF', here, the point source is the excitation laser and the spreading occurs throughout the system until it is focused into the sample.

Another example, is the tophat, 3D doughnut or bottle beam, prominently used in 3D STED and MINFLUX that is generated by a radial step function with a phase jump of $-\frac{1}{2}\tau$ at some optimal radius of approximately $r_{\mathrm{obj}}/\sqrt{2}$[67, 68]. The focal intensity exhibits a doughnut-like shape laterally with two intense Gaussian-like lobes above and below. At the centre there is a zero point of intensity.

In addition to the one- and zero-dimensional zeros of the vortex and the tophat there exists the 'half-moon' or 'bi-lobed' beam that has a two-dimensional zero; a zero plane. The corresponding phase mask is a 1D step function with a $\frac{1}{2}\tau$ phase jump.

Lastly, it must be mentioned that focused beams in samples with an aqueous environment experience a scaling of their axial size and focus depth due to the refractive index mismatch of the glass coverslip ($n_{\mathrm{glass}} = 1.52$) and water ($n_{\mathrm{water}} = 1.33$)[69]. Here, this amounts to a factor of 0.7[70, 40]. Note that the z-coordinate of all simulations (except the ones found in Appendix H) presented in this work were not scaled by 0.7. The z-coordinates of measurement results (Section 4), however, were scaled by 0.7.

2.3. Interferometric Displacements of Zero Points, Lines and Planes

STED and MINFLUX use beam shapes with zeros of different kinds. Both techniques move these zeros. In STED, the position of the zero determines what subvolume of the sample is currently imaged. The image is built pixel-by-pixel by scanning the zero across the region of interest. MINFLUX on the other hand, localises single emitters by sequentially moving the zero to a predefined set of locations. As MINFLUX can achieve high localisation precision with very few photons, this 'multiplexing' of zero locations can be very fast, with a complete sequence of four to seven positions lasting down to about 100 µs. As discussed in Section 3.3.5, achieving these deflection speeds along the optical axis is technologically challenging, if not impossible, limiting 3D MINFLUX implementations. However, the interferometric manipulation of zeros as described here offers an alternative that puts axial deflection on par with current lateral deflection technologies in terms of speed.

The basic concept is the following: consider a 1D electric field profile with an amplitude given by a linear function $E_{x,1}(x) = xE_1$. The corresponding intensity is a parabola, $I_1(x) = E_1^2 x^2$, with a single zero at the location of the zero crossing $x = 0$. Now, a second electric field with constant amplitude $E_{x,2}(x) = E_2$ is superimposed coherently and in phase such that the fields interfere as $E_x(x) = E_{x,1} + E_{x,2} = xE_1 + E_2 = E_1(x + x_0)$, with $x_0 = E_2/E_1$, which is again a linear function, however, the zero crossing is shifted "to the left". The shift is proportional to the relative amplitude E_2/E_1 of both fields and the new zero location is where the two fields cancel each other; they interfere destructively. The corresponding zero in intensity moves analogously. Hence, by increasing or decreasing the (relative) amplitude we can interferometrically deflect the zero from its original position $x = 0$ to any position on its left $x < 0$. To achieve the same in the opposite direction, $x > 0$, we need to flip the relative phase[6] by $\frac{1}{2}\tau$.

The electric field of a tophat beam, E_{TH}, close to the zero follows a linear amplitude along z and is linearly polarised as well. Similarly, the electric field of Gaussian, $\boldsymbol{E}_{\mathrm{G}}$, in a first order approximation around its maximum, acts as a constant. Therefore, if we can interfere a tophat and a Gaussian beam in the focal plane, control their relative intensities and are able to flip their phase, we have the basis for an implementation that deflects a tophat zero along the optical axis as shown in Figure 2.2. This feat is necessary for 3D MINFLUX and, until this work, has only been achieved by completely defocusing the tophat beam up and down.

[6] Phase is always relative. In the case of two fields it therefore does not matter if the phase is adjusted for only one or both of the fields.

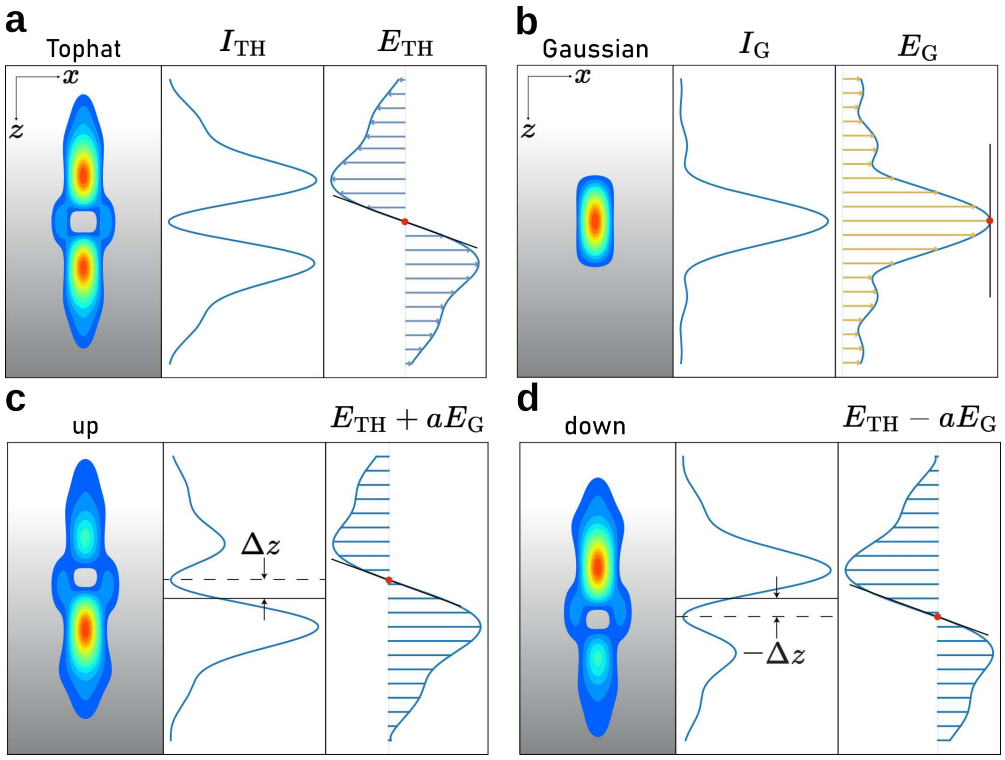


Figure 2.2.: Overview of the axial deflection concept. All subpanels show the xz intensity distribution, the intensity profile at $x = y = 0$ and the electric field profile (ignoring the phase). **a** Shows the tophat beam with a linear zero crossing in the electric field marked by the red dot. The black line indicates the linearity around this point. **b** Depicts the Gaussian. The maximum is marked with a red dot and its derivative is indicated by the black line. **c** and **d** show the deformed tophat with an up- and down-deflected zero. The deflection distance with respect to the undeflected zero (solid line) is marked by the dashed line.

Adding a linear to a constant electric field deflects the zero perfectly. However, the tophat as well as the Gaussian exhibit these properties only locally around the focus where the tophat's zero crossing and the Gaussians maximum are located. Away from focus, it is not only the amplitude but also the phase that differs from the ideal: The so-called 'Gouy phase', $\phi_G^{[m]}$ of Hermite-Gauss or Laguerre-Gauss (LG) beams is given by[71]

$$\begin{aligned}\phi_G^{[m]} &= m\phi_G \\ \phi_G &= \arctan(z/z_R)\,,\end{aligned}$$

where m is the generalised mode order index and z_R is the Rayleigh length. The tophat beam can be thought of as the strongly focused version of an LG beam, $\mathrm{LG}_{p=1}^{l=0}$[72] for which $m = 2p + |l| + 1$[73]. Thus, when going from $z = -\infty$ to $z = \infty$, the Gouy phase is expected to lead to an overall phase shift of $\Delta\phi_G^{[3]} = \frac{3}{2}\tau$ and $\Delta\phi_G^{[0]} = \frac{1}{2}\tau$ for the tophat and the Gaussian, respectively. Note, that the tophat phase exhibits a phase jump of $\frac{1}{2}\tau$ at $z = 0$ due to the zero crossing of the electric field. The remaining mismatch of $\frac{1}{2}\tau$, however, inevitably causes a mismatch of phases. The Gouy phases of the tightly focused simulated Gaussian and tophat are compared in Figure 2.3: the first two rows show the real E_x component for three time points of both beams along the optical axis on the left, and $\arg E_x - k \cdot z$ (assuming $k_z \approx k$) compared to the ideal Gouy phases on the right. The last row showcases the normalised electric field and the Gouy phases of the two beams on the left and right, respectively. The Gaussian is adjusted (see caption) to highlight how the oscillations and phases "drift apart" along z.

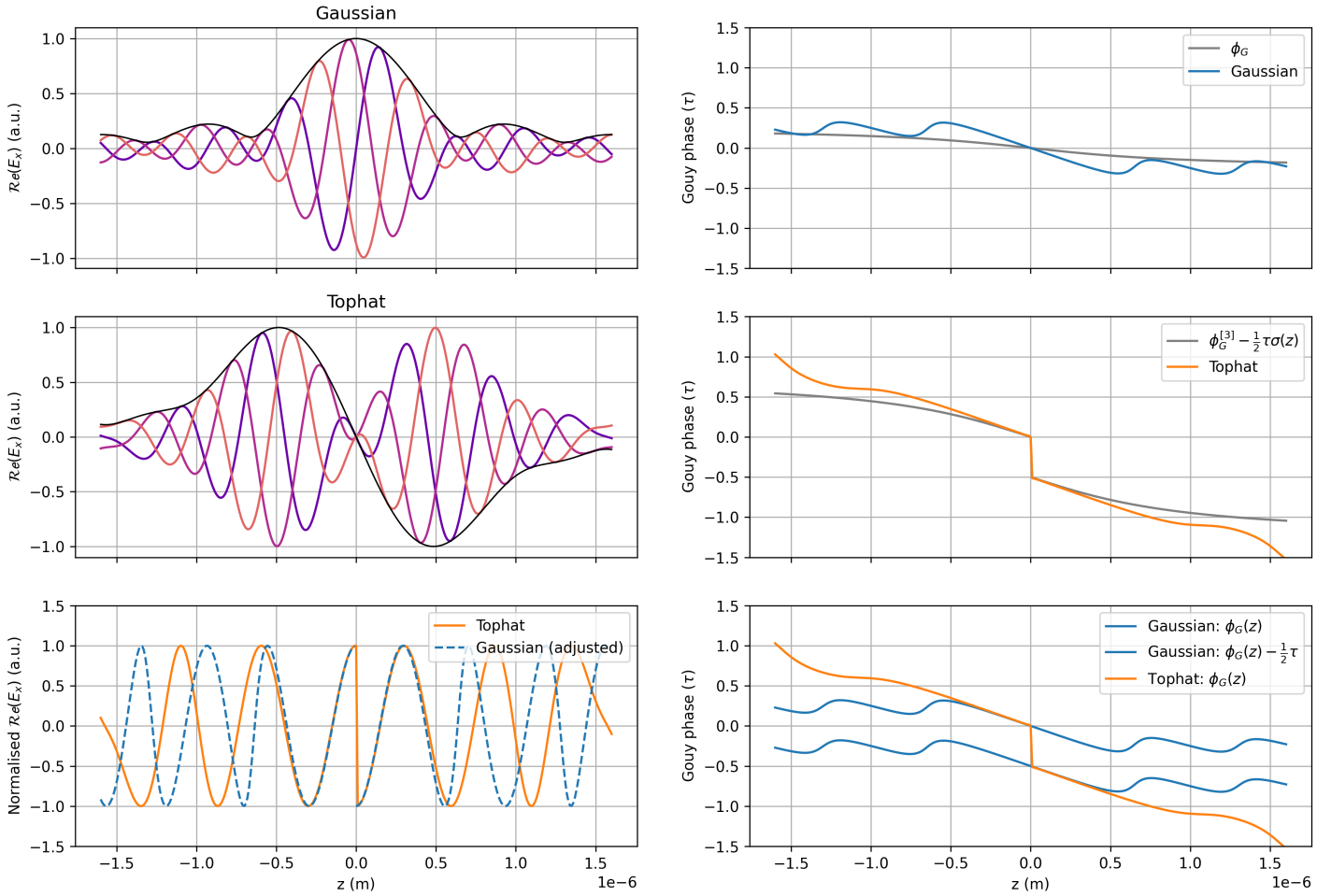


Figure 2.3.: Electric fields and Gouy phase of the Gaussian and the tophat beam (first two rows). The Rayleigh length, z_R, for plotting the ideal Gouy phase was calculated as $z_R = kw_0^2/2$ with $w_0^2 = \frac{2P}{\frac{1}{2}\tau I_{\max}}$ where P is the input power of the beam and $I_{\max}$ is the maximum intensity. The last row shows the direct comparison of the Gaussian and the tophat. The electric field of the Gaussian was shifted in phase by $\frac{1}{4}\tau$ and flipped for $z > 0$ to match that of the tophat.

The consequence is a non-ideal zero deflection, where the intensity zero becomes an intensity minimum due to imperfect destructive interference. Furthermore, interference takes place across the whole (focal) volume and not just around the zero. Thus, the overall beam shape changes. In case of the tophat, the lateral side-lobes as well as the axial side-lobe, towards which the zero moves, decrease in intensity when interfered with a Gaussian. The opposite axial lobe becomes more intense, see Figure 2.2. Thus, while we are primarily concerned with the interferometric manipulation of zeros, this concept can also be thought of as implementing a dynamic beam shaper.

Figure 2.4 shows simulations of six beam combinations in detail. Each row examines another beam combination. The first two columns show the PSFs of the interfering beams, indicate the required polarisation at back focal plane of the objective (Circular Polarisation (CP) and Linear Polarisation (LP)) and their corresponding phase mask. The first three rows show xy planes while the last three show xz planes. Although the single PSFs are shown with equal brightness, the power of the second beam is much lower to produce the result in the third column. The dashed inset indicates the zoom-in region shown in the fourth to sixth columns where, additionally, little arrows indicate the orientation of the electric field projected into the viewing plane. The last two columns show intensities and electric fields for five different relative amplitudes.

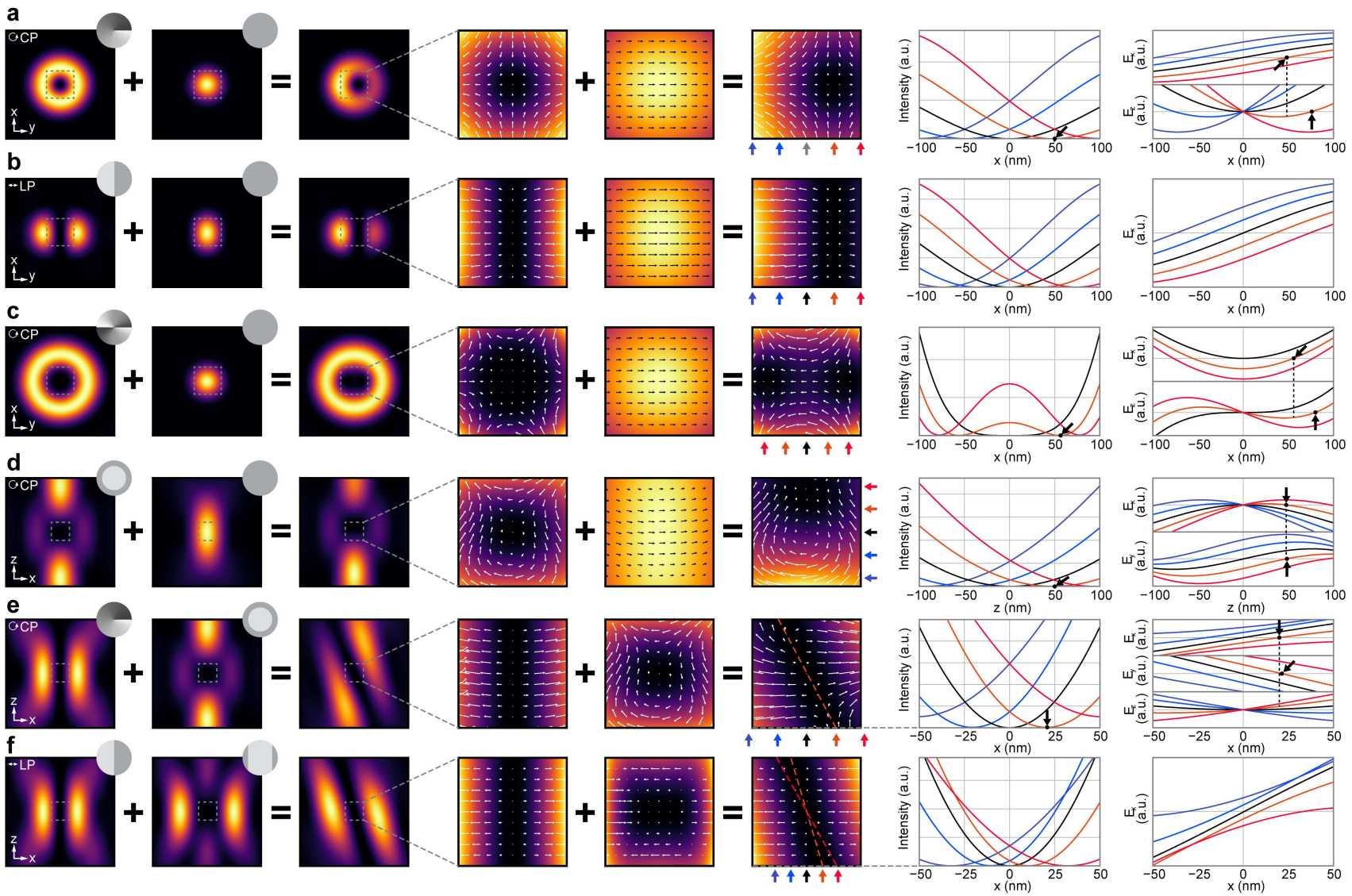


Figure 2.4.: Simulations of six interferometric beam combinations manipulating intensity zeros. The first three columns show the intensity over a **a-c** 800 nm × 800 nm xy-plane and a **d-e** 1200 nm × 1200 nm xz-plane. The dashed squares indicate the zoom-in shown the three middle columns and comprise an area of 200 nm × 200 nm.

In row **a** we see the combination of a vortex and a Gaussian. The result is a deformed vortex with a shifted zero. Similar to the tophat plus Gaussian case, the vortex exhibits a linear field locally around its zero and the locally constant Gaussian shifts the zero crossing. The zero degradation is caused by the mismatch of the positions of the zero crossing of the electric field components in xy and in z, as indicated by the arrows and the dashed line in the last column. The relative amplitude decides the deflection radius. The relative phase changes the angle or direction along which the deflection takes place. Hence, the zero can be deflected freely in 2D if one can control the relative phase in addition to the relative amplitude.

Row **b** shows the 2D deflection of the half-moon. Again a Gaussian plays the role of introducing a locally constant electric field around a zero crossing. It is analogous to the vortex plus Gaussian case, however, the required polarisation is linear. The linear polarisation has to match the orientation of the half-moon's phase mask. Otherwise, there would be no true zero. Note, that the Gaussian shown in the second column is therefore elongated in the direction of the polarisation (see also [64], Figure 3.9).

Row **c** shows the combination of a 2τ-vortex with a Gaussian. The zero splits into two zeros deflected in opposing directions. However, when flipping the phase, e.g., when scanning the relative amplitude from minus to plus the trajectories of the zeros approach the centre, make a 90° turn and move outwards again; the trace of their trajectories is a cross. Again, the angle of deflection is determined by the relative phase. In general, a vortex of higher order $n\tau$ will produce n zeros that move radially along evenly spaced angles.

Row **d** shows the tophat plus Gaussian that was already discussed in detail. The last row highlights the fact that the electric field components in x and in y have the zero crossing at the same location along z. The zero degradation occurs because of the mismatch in Gouy phases and is therefore caused by a residual E_z component at the location of the deflected zero.

Row **e** demonstrates that the combination of a vortex with a tophat can tilt the vortex' zero line. The relative amplitude determines the tip angle (to the optical axis) while the relative phase determines the azimuthal orientation. Viewing the intensity distribution in the tilt plane can be deceiving as one might think the overall 3D shape is that of a tilted doughnut (elongated along z). Figure B.1 shows snapshots of the full rotation from several perspectives revealing that the shape changes in a more complex manner.

Finally, row **f** shows how the zero plane of a half-moon is tilted by interfering it with the 1D analogue of the tophat, which I call the tophat slice: taking a 1D profile –a "slice"– of the tophat phase mask results in a rectangular function with a step height of $\frac{1}{2}\tau$. Extruding this profile orthogonally to fill the phase pupil yields the phase mask necessary to produce the beam shape that tilts the half-moon zero plane[7].

2.4. Interferometric Zero Displacements for (3D) MINFLUX

In the last section, several beam shapes that feature intensity zeros and how they can be interferometrically manipulated were studied. Saliently, the zero point of a tophat beam can be shifted axially as is required by, e.g., 3D MINFLUX. However, the zero quality degrades the farther it is deflected. The question arises if this degradation hinders the applicability of interferometrically deflected zeros in (3D) MINFLUX.

The loss in localisation precision due to zero degradation can be characterised by the CRB that can be calculated numerically for a given MINFLUX scheme[8], number of collected photons

[7] Analogously, the half-moon could be called the 'vortex slice': taking a 1D profile of the vortex phase mask results in a step function with a step height of $\frac{1}{2}\tau$ independent on the angle along which the profile is taken. Extruding this profile orthogonally to fill the phase pupil yields the half-moon phase mask.

[8] Consisting of the geometry of the excitation pattern and the employed beam shape and its weight (proportional to the relative exposure power and time) for each coordinate.

and SBR. Let us consider the original 3D MINFLUX scheme that employs seven tophats: four in the xy-plane evenly spaced on a circle of diameter L_{xy} and three along z with positions $z_i \in \{-L_z/2,\ 0,\ +L_z/2\}$. We choose $L_{xy} = L_z = 100\,\text{nm}$ and SBR = 5; typical values in experiments. Lastly, we set the amount of collected photons to one. It is the canonical choice as the resulting CRB simply needs to be divided by $\sqrt{N}$ to arrive at the CRB for any other photon count N[9].

As conventional techniques move the focus to position the tophats axially, there is no zero degradation. Naturally, this case will act as the reference. In contrast, the zero quality depends on the axial position for the interferometric approach. While the central tophat with its zero at $z = 0$ will showcase no zero degradation, zeros above and below at $z = \pm L_z/2$ do. Furthermore, in the context of iterative MINFLUX, the final localisation takes place after 'converging' onto the emitter position. Hence, the final iteration, that produces the actual localisation, might start with an axial offset, denoted z_b, with respect to the focal plane. In this case the three tophats exhibit different zero degradations and slightly different shapes. Consequently, the CRB not only depends on the emitter position with respect to the centre of the excitation pattern but also on the offset z_b.

Figure 2.5 shows the CRB of the localisation precision in z versus the distance of the molecule to the centre of the excitation pattern $z_m - z_b$ (the x and y coordinates are set to zero). The reference, labelled 'ideal', is shown in grey. The coloured curves show the CRB_z for different values of z_b.

The interferometric approach performs worse the higher the offset. At $z_b = 200\,\text{nm}$, which corresponds to 140 nm in an aqueous sample, the penalty can already be as much as 50 % if the molecule is off centre. However, typical MINFLUX scenarios iteratively converge on the emitter beforehand. The final iteration is often centred within $\pm 10\,\text{nm}$ of emitter position. Hence, for $z_b = 200\,\text{nm}$, a loss between 20 % to 30 % in localisation performance per photon can be expected. Note that the coloured curves are asymmetric. This is a consequence of the asymmetry of the axial lobes of the tophat when interferometrically shifting its zero. The CRB curves corresponding to values $z_b < 0$ are simply the mirrored versions of the $z_b > 0$ curves and are thus not shown.

While the axial deflection has certainly the most attractive prospect, the other zero manipulations presented in the last section might also be of interest. Instead of using conventional lateral deflectors, like Electro-Optic Deflectors (EOD) one could employ a vortex and a Gaussian to achieve lateral zero movement in the microsecond regime. However, it would require an additional EOM that controls the relative phase between the two beams in order to achieve full 2D deflection capabilities. While the cost might be less than that of two EODs, disadvantages remain: the interferometric deflection range is limited to around 400 nm compared to EODs that in our system reach a range of 800 nm only limited by the maximum voltage output of the driving amplifiers, and the zero degrades the farther the beam is deflected. Also, without EODs one loses a valuable option for measuring intensity distributions of the excitation beam. Similarly, there is the option to laterally shift the zero plane of a half-moon beam as discussed in 3.3.7: the half-moon shape is expected to perform better for MINFLUX localisations in 2D. However, the orthogonal half-moon beams require orthogonal linear polarisation states to form a true zero. This requires again an additional EOM. In contrast to the previous case, its purpose is to flip the linear polarisation state right before the objective. Finally, tilting a zero line or plane might be used to generate 3D information as it introduces an axial gradient that was not there before.

[9] An example: if the single photon CRB is 50 nm, the equivlant CRB for $N = 100$ is $\frac{50}{\sqrt{100}}\text{nm} = 5\text{nm}$.

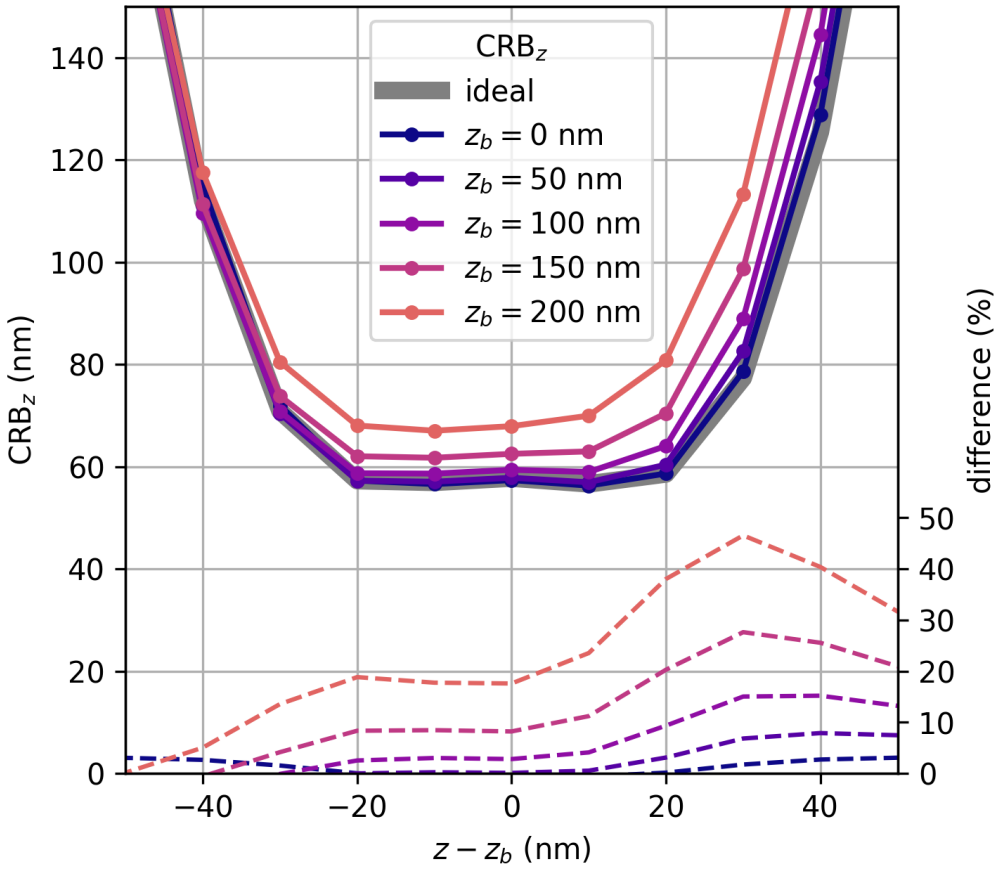


Figure 2.5.: Comparison of CRB_z for SBR = 5, $L = 100\,\mathrm{nm}$ and a single collected photon, $N = 1$. The xy-tophats have double the weight for improved localisation isotropy. The x-axis values indicates the distance between emitter and the origin of the excitation pattern in a range of $\pm 50\,\mathrm{nm}$. The CRB_z of the ideally defocused tophat beams is shown in grey. For the interferometrically deformed tophats several offsets z_b, with respect to the focal plane, are shown (coloured curves). The dashed curves show relative difference to the ideal case on the right y-axis.

2.5. A Fast and Versatile Polarisation Interferometer

The last section discussed the applicability of the interferometric zero deflection to 3D MINFLUX despite the inevitable zero degradation. This section describes how the concept can be realised.

We base our implementation on a dynamic polarisation interferometer that offers many advantages for implementing our strategy:

1. It naturally controls the relative amplitude of the two superimposed beams in a way that yields a deflection that has a close to linear dependence on the control voltage.
2. The required phase flip for deflecting in the opposite direction is automatically taken care of.
3. It operates in a common path geometry; no active stabilisation of interferometer arms is necessary.
4. The driving element, an Electro-Optic Modulator (EOM), allows deflection speeds down to 5 µs and less.
5. The interference occurs at the output of the interferometer far away from the sample and results in a single, linearly polarised beam making further handling through the rest of the system straightforward.

The three main components are

1. A polarisation rotator.
2. A Spatial Light Modulator (SLM) in a double pass configuration.
3. A polarisation interferometer interfering s- and p-polarisation states.

In short, the rotator controls the relative intensity/amplitude of the s- and p-polarisation state at its output. The SLM double pass imprints the phase to generate a tophat on one polarisation state and a Gaussian on the other. Finally, the polarisation interferometer projects the sum of the two patterned polarisation states into a new s-state at its output. This beam, when focused through an objective, results in the described beam shape (with a deflected zero) since focusing is a linear operation[10].

2.5.1. Polarisation Rotator

The polarisation rotator, shown in Figure 2.7 **a**, employs an Electro-Optic Modulator (EOM). These devices consist of birefringent crystals that change their refractive index along one of the crystal's axes –the 'extraordinary' or 'slow'[11] axis– linearly with the electric field strength; the so-called 'Pockels' or 'linear electro-optic effect'. In practice, the crystals sit between two parallel plates with electrical contacts. Thus, the field strength inside the crystals is proportional to an applied voltage. EOMs can be utilised in many ways, for example, as phase, amplitude or polarisation modulators. For our purpose, the extraordinary axis is oriented at $\frac{1}{8}\tau \widehat{=} 45°$ such that the Jones matrix is given by:

$$\mathrm{EOM}(V) = R(45°) \begin{pmatrix} 1 & 0 \\ 0 & e^{i2\alpha(V)} \end{pmatrix} R(-45°) \tag{2.6}$$

$$\alpha(V) = \frac{\tau}{2} \frac{V - V_0}{V_{\lambda/2}}\,, \tag{2.7}$$

[10] It does not matter if the superposition happens in sample or in 'phase' space.

[11] 'Slow' refers to the fact that a medium with an increased refractive index causes the light to propagate more slowly through it.

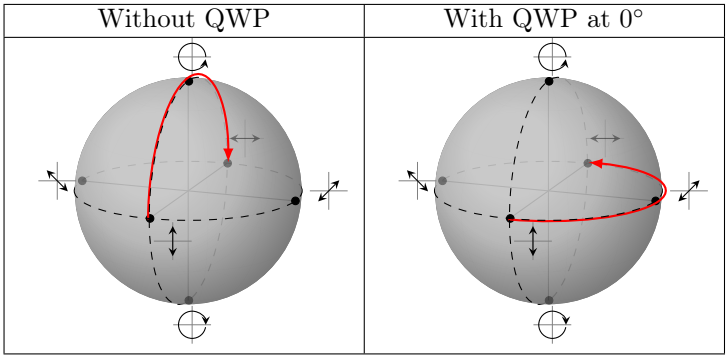


Figure 2.6.: Two options for polarisation "rotations": Left, going through the circular polarisation at the pole. Right, rotating while keeping the relative phase at zero keeps the polarisation linear only changing the orientation angle.

where $R(\theta)$ is the rotation matrix at angle θ, V_0 is the offset voltage required to make the refractive index of the ordinary and the extraordinary axis equal[12], and $V_{\lambda/2}$ is the voltage at which an incoming linear polarisation state is flipped to its orthogonal linear polarisation state after propagating through the EOM. Note, that these parameters are wavelength-dependent and the explicit notation was dropped here.

For an incoming vertically polarised beam, one gets

$$\mathrm{EOM}(V)\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix} i\,\sin\left(\frac{\alpha(V)}{2}\right)\\ \cos\left(\frac{\alpha(V)}{2}\right)\end{pmatrix} \tag{2.8}$$

after some algebraic manipulations including the multiplication by a common phase. The result shows that while the relative amplitude is continuously modulated when scanning the voltage (and hence α), the relative phase of the outgoing s- and p-states, $\Delta\varphi_{\mathrm{sp}}$, is constant for any α: $e^{\Delta\varphi_{\mathrm{sp}}=\frac{1}{4}\tau} = i$. The corresponding trajectory on the Poincaré sphere goes from the initial linear vertical state to circular when $\alpha = \frac{1}{4}\tau$ and then to the linear horizontal state for $\alpha = \frac{1}{2}\tau$, see Figure 2.6.

To get a linearly polarised output for any α, the relative phase of $\frac{1}{4}\tau$ has to be adjusted. This can be done by adding a quarter-wave plate with its fast or slow axis at 0° to the vertical input state. The result is a linearly polarised output beam where the polarisation axis rotates with $\alpha(V)/2$. While we initially chose this configuration because it might be conceptually easier to think in terms of a rotating linear polarisation, the QWP is not necessary. Even without the QWP the polarisation state is rotated with the rotating axis being the major axis of the intermediate elliptical states, see Figure 2.6. The extra –but constant– phase can be easily compensated by the spatial light modulator. Not using the QWP is advantageous since it avoids an extra optical element with all its imperfections.

No matter the configuration, the polarisation's rotation angle determines the relative amplitude that we require for the interference. Furthermore, the relative phase flip necessary for, e.g., the axial deflection of a tophat zero, occurs automatically. Starting from the vertical state ($\alpha = 0$) in Equation (2.8) and applying a positive voltage $> V_0$ makes both entries of the vector positive, applying a negative voltage $< V_0$ flips the sign of the $\sin(\cdot)$ while keeping the $\cos(\cdot)$ positive. Thus, the phase relative flipped by $\frac{1}{2}\tau$.

[12]The offset voltage varies from device to device. In our case the V_0 was in the range of 20 V to 90 V for $V_{\lambda/2} \approx 220$ V.

2.5.2. SLM Double Pass

For generating the two beam shapes that will interfere, we employ a phase-only liquid crystal spatial light modulator, simply referred to as the 'SLM'. It is similar to a liquid crystal display consisting of pixels that can be individually addressed. However, instead of a back-light, one illuminates the front with a light beam and each pixel induces an individual phase. All pixels together hence imprint a 'phase mask', $\varphi(x, y)$, onto the spatial profile of the incoming beam. While transmissive devices exist, we employ reflective SLMs. An essential property is that the SLM only affects one of the polarisation states along a predefined axis; in our case it is the vertical axis. Furthermore, it is common practice to add a blazed phase grating on top of the desired phase mask such that the reflected beam will diffract into two orders: the zeroth order, i.e., the direct reflection off the SLM surface that will be discarded, and the first diffraction order that carries the patterned beam. With an amplitude corresponding to a full wavelength of the incoming light, most of the intensity will be send to the first diffraction order at an angle that depends on the pixel size and the period of the grating[13]. The grating spatially separates patterned from residual unpatterned light that would otherwise pollute the generated beam shape[74] and filters out crosstalk between the two SLM passes as will be discussed in Section 2.5.4.

Instead of using two SLMs for generating the two interfering beams, the SLM is divided into subregions and the input beam passes the SLM twice, where each pass is on a different SLM region. This approach is feasible as current SLMs offer plenty of pixels arranged over a large enough rectangular area[14]. The setup's geometry is well known and has been used in many previous works, for instance, in 3D STED microscopes[75, 76] where a vortex and tophat are generated by the two passes and superimposed incoherently to form a 3D depletion volume in sample space. Our implementation, however, requires the coherent superposition of the two passes.

The double pass and its working principle is visualised In Figure 2.7 **b**. For pedagogic reasons, the depiction is "unfolded" and does not correspond to the actual geometry (see Section 3). The s-state of the input beam, E_1, is patterned on the first pass and is diffracted under some small angle. The p-state, E_2, remains unaffected as the SLM only patterns the vertical polarisation and continues as a direct reflection. Then, both orders pass through a quarter-wave plate and a lens twice (being reflected by a mirror in between). This flips the polarisation states of the two orders, such that the patterned s-state in order one becomes p-polarised. Vice versa, the unpatterned zeroth order flips from the p- to the s-polarisation. Hence, the second pass –which again affects only the vertical polarisation state– patterns the new s-state in the zeroth order. This way each pass patterns exactly one of the two polarisation states of the input beam. The lens is placed such that it forms an image or 4f system from one pass to the other. This ensures that all the diffraction orders have equal optical path length –effectively following a 'common path'– and overlaps the planes of the SLM regions or passes virtually on top of each other. Again, a blazed grating is used on the second pass such that the zeroth order of the first pass is collinear with the direct reflection of the first pass' first order. In the end, the output of interest is a single beam where the s- and p-polarisation states are individually patterned. Still, there is the residual light of the zeroth orders, however, it is blocked by an iris within a telescope situated after the polarisation interferometer. It is called the 'order selector' which resembles a spatial filter.

[13] When the amplitude of the blazed grating for a given wavelength is exactly one full wave or τ, the unwrapped grating is a linear phase along one spatial dimension. In terms of optical path length, it is equivalent to a mirror at an angle given by the steepness; some regions of the incoming beam's spatial profiles are reflected "earlier" and hence accumulate less phase through propagation than beam regions that are reflected "later".

[14] Our SLM has 1920 by 1152 pixels with an area of 17.7 mm by 10.6 mm.

2.5.3. Polarisation Interferometer

Finally, the two polarisations of the double pass output beam are interfered. As depicted in Figure 2.7 **c**, the polarisation interferometer consists of a half-wave plate at 22.5° before a linear polariser: a Polarising Beam Splitter (PBS) cube that is rotated such that the vertical polarisation is transmitted[15] is employed.

In general, a half-wave plate rotates linear polarisation states by twice its angle measured between the vertical polarisation and the fast axis:

$$\mathrm{HWP}(\theta) = R(\theta) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R(-\theta) = R(2\theta)\,.$$

As any polarisation state can be expanded in terms of the basis of the two linear and orthogonal s- and p-states, the rotation applies to both components. As the HWP is at 22.5° the rotation of the components is 45°, the rotated states project equally onto the (vertical) polarisation axis of the PBS. Therefore, the new vertical output state contains the superposition of the two patterned states interfering their phase profiles generated by the phase masks on the SLM. Note, that the two states are also projected onto the horizontal state of the PBS (although with flipped relative phase) and hence the output beam carries only half the power of the input.

The HWP angle θ does not need to be at 22.5°. Other angles, however, will tip the balance towards one or the other polarisation state in the superposition. This can be beneficial for certain applications. For instance, one could have one of the passes more prevalent in order to reduce the 50 % power loss of the associated beam shape if the interferometric manipulation is still feasible with high power loss for the other pass. For the axial deflection of the tophat zero we stuck to an equal mixture as the calibration of voltage to deflection is almost linear in this case as seen in Figure 2.8.

2.5.4. Extended Jones Formalism

In this section, the entire double pass polarisation interferometer is modelled with the Jones formalism to analyse effects due to imperfections of the optical elements. As the double pass operates on two diffraction orders, the formalism is trivially extended to track the two orders and their crosstalk.

Consider a perfect vertically polarised input beam: it is patterned and diffracted on the first pass and then flipped to the horizontal state, leaving it unaffected by the second pass. The "interference" at the PBS consists only of one component (the other is zero); we get the pure beam shape corresponding to the first pass' phase mask as it would be used, for example, in the first Minflux iteration that gathers axial information which uses an undeflected tophat (assuming it has no initial offset in z).

When the QWP (or any of the other optical components) does not act ideally, however, the polarisation will not be completely flipped from one pass to the other. Consequently, even an input with perfect vertical polarisation will be patterned partially on the second pass potentially polluting the pure beam shape with another. This is referred to as 'crosstalk'. As it turns out, though, the diffraction gratings play a crucial role to prevent crosstalk.

Only the first order of each pass is allowed to propagate through the rest of the system to the objective. However, any crosstalk generated by, e.g., an imperfect QWP is already diffracted once from the first pass. It is diffracted a second time on the second pass into a "second order" which is blocked by the order selector. Any such crosstalk is therefore conveniently filtered out. Of course it desirable to reduce the crosstalk to a minimum as it directly translates into a power loss.

[15] The polarisation at the output of the transmitted path is cleaner compared to the reflected output.

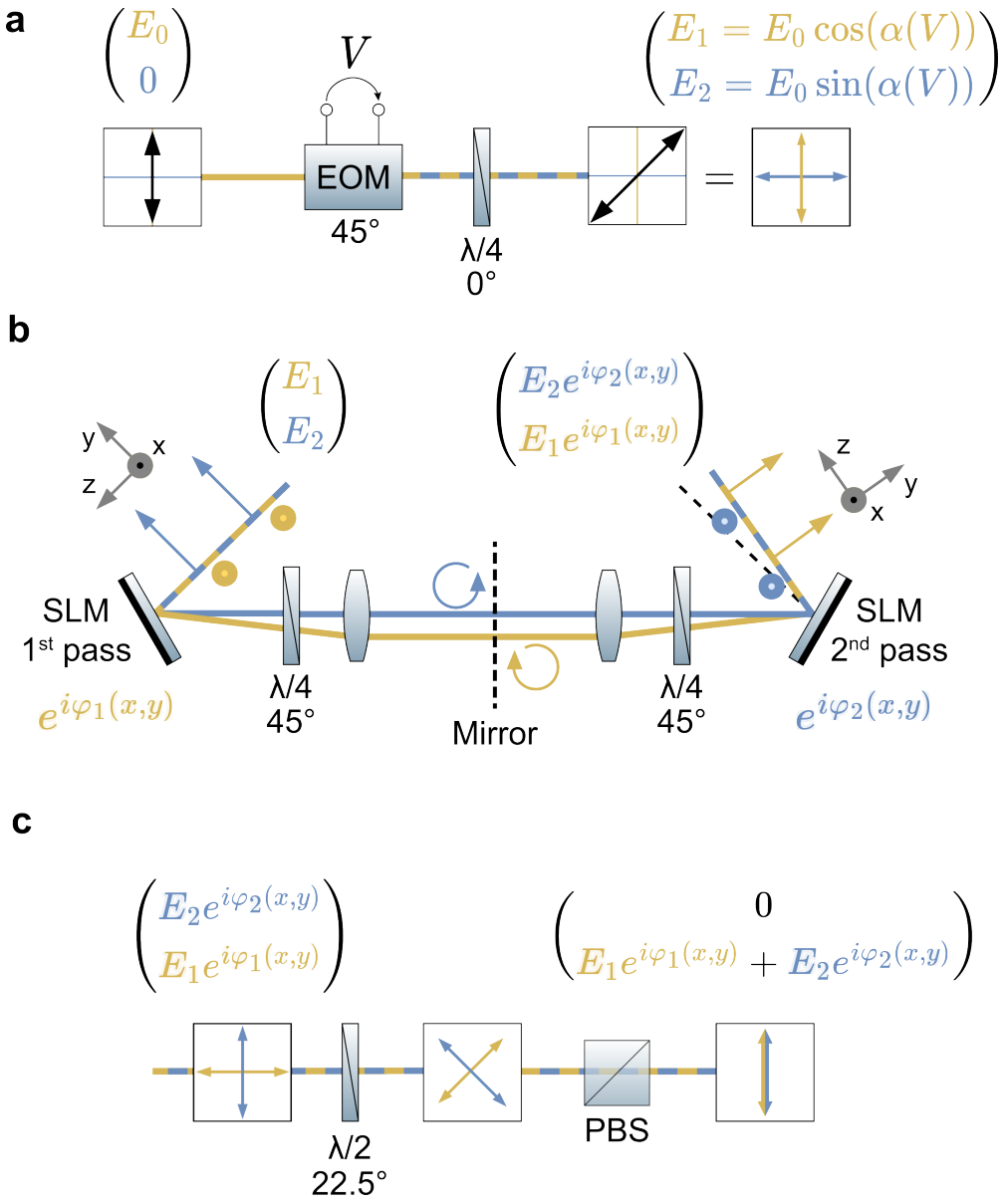


Figure 2.7.: Schematic diagrams of the dynamic beam shaper implementation: **a** polarisation rotator, **b** SLM double pass and **c** polarisation interferometer.

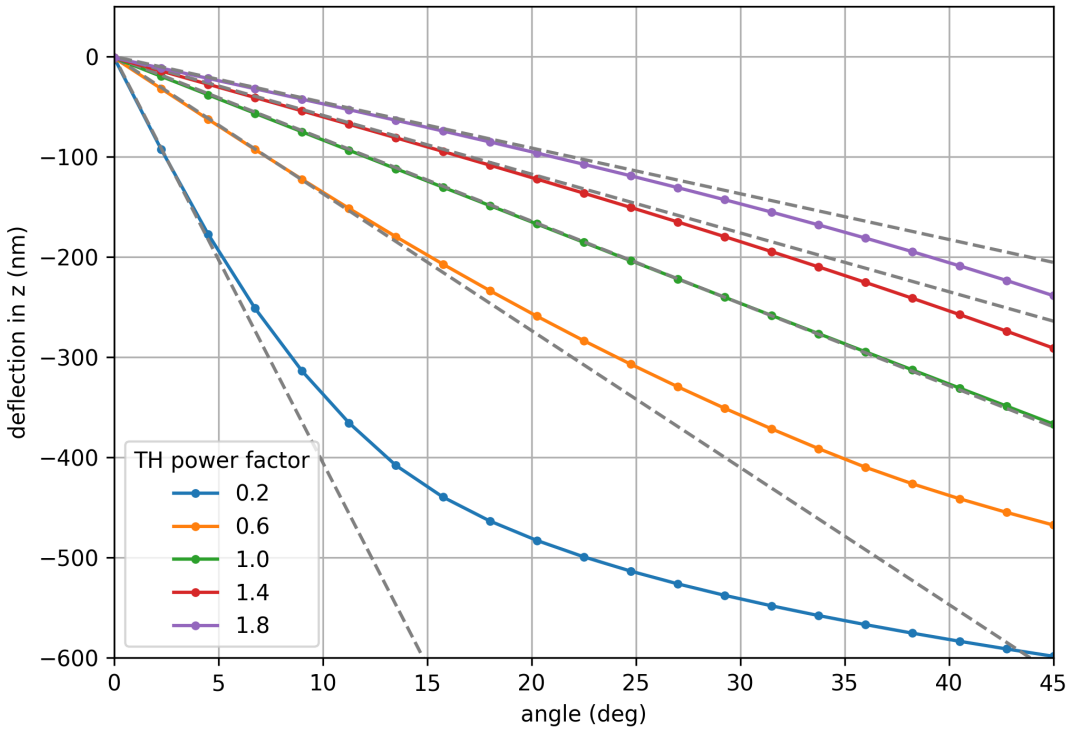


Figure 2.8.: Deflected distance versus polarisation angle (coloured curves). The grey dashed lines visualise the steepness at the origin of their corresponding deflection curves highlighting the their degree of linearity.

To cast these observations into formulae, the double pass is modelled with a chain of Jones matrices such that the output Jones vector is given by:

$$\boldsymbol{J}^{\text{out}} = \text{PBS}\,\text{HWP}\,\text{SLM}_2\,\overline{\text{QWP}}(-\theta_{\text{QWP}}, \Delta\psi)\,\overline{\text{L}}\,\overline{\text{M}}\,\overline{\text{L}}\,\overline{\text{QWP}}(\theta_{\text{QWP}}, \Delta\psi)\,\text{SLM}_1\,\boldsymbol{J}^{\text{in}}\,. \tag{2.9}$$

The following table describes all terms:

$\boldsymbol{J}^{\text{in}} = \begin{pmatrix} \sin(\alpha) E_2 \\ \cos(\alpha) E_1 e^{i\Delta\varphi_{\text{sp}}} \end{pmatrix}$	Input Jones vector.
E_1	Input electric field amplitude, s-polarisation.
E_2	Input electric field amplitude, p-polarisation.
α	Rotation angle of linear input polarisation ($\alpha = 0°$ is vertical).
$\Delta\varphi_{\text{sp}}$	Relative phase between s- and p-polarisation.
$\text{SLM}_1 = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \\ 0 & 0 \\ 0 & (1-\varepsilon) e^{i\varphi_1(x,y)} \end{pmatrix}$	Extended Jones matrix for the first SLM pass.
ε	Fraction of vertically polarised light that is not diffracted by the SLM. In the following we assume $\varepsilon = 0$.
$\text{SLM}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & e^{i\varphi_2(x,y)} & 0 & 0 \end{pmatrix}$	Extended Jones matrix for the second SLM pass.
$\varphi_1(x, y)$	Phase mask of first pass.
$\varphi_2(x, y)$	Phase mask of second pass.
$\text{QWP}(\theta_{\text{QWP}}, \Delta\psi) = R(\theta_{\text{QWP}}) \begin{pmatrix} 1 & 0 \\ 0 & -ie^{i\Delta\psi} \end{pmatrix} R(-\theta_{\text{QWP}})$	Jones matrix of an imperfect quarter-wave plate.
θ_{QWP}	Rotation angle of the quarter-wave plate.
$\Delta\psi_{\text{QWP}}$	Deviation from the ideal retardation of the quarter-wave plate.
$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$	Rotation matrix.
$L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Jones matrix of the lens. In our model we assume it has no effect.
$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	Jones matrix for the metallic mirror. In our model we neglect the angle of incidence.
$\text{HWP}(\theta_{\text{HWP}}) = R(\theta_{\text{HWP}}) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R(-\theta_{\text{HWP}})$	Jones matrix for an ideal half-wave plate.
$\text{PBS} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	Jones matrix for a perfect polariser.
$\overline{X} = \mathbb{1} \otimes X = \begin{pmatrix} X & \mathbb{0} \\ \mathbb{0} & X \end{pmatrix}$	The $\overline{\cdot}$ extends a Jones matrix via the Kronecker product $\otimes$ as described below.
$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	2×2 unit matrix.
$\mathbb{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	2×2 zero matrix.

Table 2.3.: List of symbols used in the extended Jones formalism.

A second Jones vector is introduced for keeping track of the polarisation of both diffraction orders generated by grating on the first SLM pass, see Figure 2.9. The 4×1 extended Jones vector is the concatenation of both Jones vectors

$$\begin{pmatrix} \boldsymbol{J}_0 \\ \boldsymbol{J}_1 \end{pmatrix} = \mathrm{SLM}_1 \boldsymbol{J}^{\mathrm{in}} ,$$

with the rectangular 4×2 matrix SLM_1. Both orders propagate independently until the second SLM pass. Therefore, the Jones matrices of the lens, the QWP and the mirror are extended to block matrices with copies of themselves on the diagonal via the Kronecker product with the unit matrix. On the second pass each order diffracts again into two new orders for a total of four beams. Two of them, however, are collinear which leaves three beams such that SLM_2 is a 6×4 matrix. Nevertheless, the order selector can only be passed by the order that was diffracted exactly once. Hence, one can drop the superfluous orders by reducing SLM_2 since all orders propagate independently after the second pass:

$$\begin{pmatrix} \cancel{\boldsymbol{J}_0^{\mathrm{out}}} \\ \boldsymbol{J}_1^{\mathrm{out}} \\ \cancel{\boldsymbol{J}_2^{\mathrm{out}}} \end{pmatrix} = \mathrm{SLM}_2 \begin{pmatrix} \boldsymbol{J}_0' \\ \boldsymbol{J}_1' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \varepsilon & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & (1-\varepsilon)e^{i\varphi_2(x,y)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{J}_0' \\ \boldsymbol{J}_1' \end{pmatrix} .$$

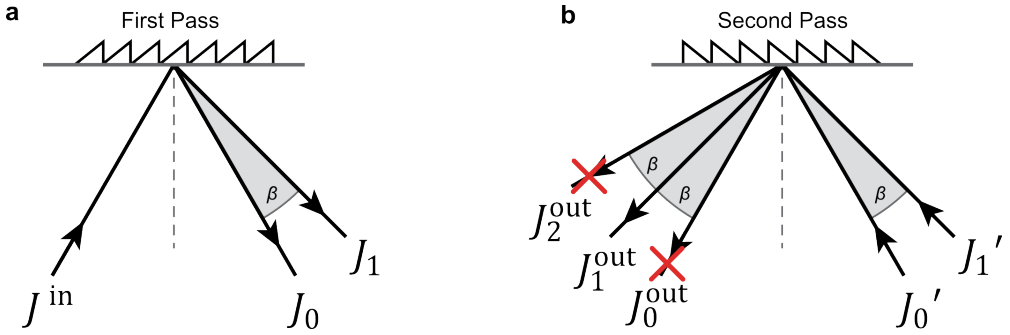


Figure 2.9.: **a** The first pass produces two diffraction orders at an angle β. **b** At the second pass the two orders of the first pass are diffracted into four orders. However, two of them are collinear. The remaining two orders are subsequently blocked by the order selector.

Equation (2.9) is evaluated with focus on the effects of an imperfect QWP through non-ideal rotation and/or retardation. Therefore, one sets $\varepsilon = 0$ which is also justified under the assumption that the undiffracted vertically polarised light of the first pass is negligible since only its crosstalk, i.e., the part that stays vertical after passing the QWP twice (which is also assumed to be small), affects the output. For the s-polarisation after the PBS, $J_1^{\mathrm{out},(1)}$, one arrives at:

$$\begin{aligned} J_1^{\mathrm{out},(1)} &= (\sin(2\theta_{\mathrm{QWP}})\cos(\Delta\psi_{\mathrm{QWP}})) \times \\ &\times \left(E_1 e^{i\varphi_1(x,y)} e^{i\Delta\varphi_{\mathrm{sp}}} \cos(\alpha)\cos(2\theta_{\mathrm{HWP}}) - E_2 e^{i\varphi_2(x,y)} \sin(\alpha)\sin(2\theta_{\mathrm{HWP}}) \right) . \end{aligned} \tag{2.10}$$

The common factor $\sin(2\theta_{\mathrm{QWP}})\cos(\Delta\psi_{\mathrm{QWP}})$ shows that a deviation from the ideal QWP rotation angle, $\theta = \frac{1}{8}\tau$, or a deviation of the ideal QWP retardation, $\Delta\psi_{\mathrm{QWP}} \neq 0$, only leads to an

overall power loss. The second term, which contains the interfering fields, remains unaffected. Only the relative phase difference $\Delta\varphi_{\mathrm{sp}}$ of the input beam appears which is readily compensated by $\varphi_1(x,y) = \tilde{\varphi}_1(x,y) - \Delta\varphi_{\mathrm{sp}}$. This phase will naturally occur as every mirror in the path between the polarisation rotator and the double pass induces adds to it. Furthermore, when using the polarisation rotator without the QWP, $\Delta\varphi_{\mathrm{sp}}$ will have an offset of $\frac{1}{4}\tau$. Any other constant relative phase that is accumulated during the propagation through the double pass can be compensated likewise.

Under ideal conditions and for a specific set of phase masks, Eq. (2.10) has an elegant interpretation. For $E_0 = E_1 = E_2$, $\Delta\varphi_{\mathrm{sp}} = 0$, $\theta_{\mathrm{QWP}} = 45°$, and $\theta_{\mathrm{HWP}} = 22.5°$:

$$J_1^{\mathrm{out},(1)} \hat{=} E_{\mathrm{out}}^{(1)}(x,y) = \frac{E_0}{\sqrt{2}}\left(e^{i\varphi_1(x,y)}\cos(\alpha) + e^{i\varphi_2(x,y)}\sin(\alpha)\right). \tag{2.11}$$

Moreover, for any two phase masks φ_1 and φ_2 of the form

$$\varphi_1(x,y) = \begin{cases} \frac{1}{2}\tau & \text{if } (x,y) \in \mathcal{D}_1 \\ 0 & \text{if } (x,y) \in \mathcal{D}_2 \end{cases} \tag{2.12}$$

$$\varphi_2(x,y) = \frac{1}{4}\tau\,, \tag{2.13}$$

where $\mathcal{D}_1$ and $\mathcal{D}_2$ are two non-overlapping domains[16] within the phase mask, Eq. (2.11) becomes

$$E_{\mathrm{out}}^{(1)}(x,y) = \frac{E_0}{\sqrt{2}}\left(\mp\cos(\alpha) + i\sin(\alpha)\right).$$

Rewriting the above as

$$E_{\mathrm{out}}^{(1)}(x,y) = \frac{E_0}{\sqrt{2}}\begin{cases} -e^{-i\alpha} & \text{if } (x,y) \in \mathcal{D}_1 \\ e^{+i\alpha} & \text{if } (x,y) \in \mathcal{D}_2 \end{cases}, \tag{2.14}$$

shows that when the two interfering beams have phase profiles that differ by $\pm\frac{1}{4}\tau$ for any given point (x,y) of the lateral profile, the resulting beam's phase profile is described by two opposing phasors that, depending on whether they are in $\mathcal{D}_1$ or $\mathcal{D}_2$, rotate clock- or counter-clockwise with α. As $|\mp e^{\mp i\alpha}| = 1$ the amplitude stays constant for any α. The phase landscape –modulo the common phase– of the outgoing beam, $\arg E_{\mathrm{out}}^{(1)}(x,y)$, consists of steps at the boundary (or boundaries) between the two domains. The step height is determined by the angle between these two phasors $\Delta\varphi_{\mathcal{D}}(\alpha) = \frac{1}{2}\tau - 2\alpha$ as depicted in Figure 2.10 **a**.

The phase masks required for generating the tophat and the Gaussian are compatible with Equations (2.12) and (2.13) where

$$\mathcal{D}_1 = \{(x,y)|\sqrt{x^2+y^2} \le r_{\mathrm{tophat}}\}$$
$$\mathcal{D}_2 = \{(x,y)|\sqrt{x^2+y^2} > r_{\mathrm{tophat}}\}\,.$$

The outgoing phase profile is shown in Figure 2.10 **b**. It is a radial step function. When α is increased (blue) the phasors approach each other and the step shrinks and vice versa. A beam with a tophat phase, where the phase jump is not exactly $\frac{1}{2}\tau$ thus corresponds to a deformed tophat with an axially shifted zero.

It is the phase space analogue of our previous notion of interfering the locally constant maximum of a Gaussian with the locally linear zero crossing of the tophat in sample space. However, the phase space interpretation makes the phase relations explicit through Equations (2.12) and

[16] Each domain does not have to be connected.

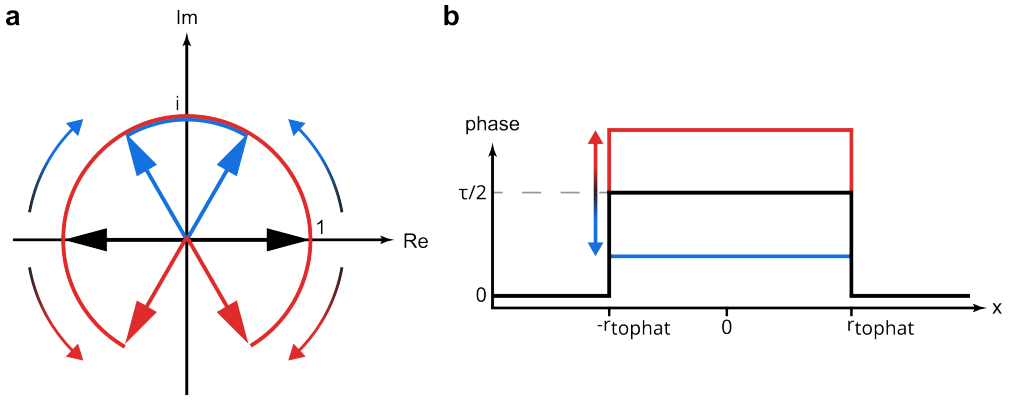


Figure 2.10.: **a** Phasors of the two domains $\mathcal{D}_1$ and $\mathcal{D}_2$. Their relative phase is continuously modulated by the polarisation rotation angle. **b** Tophat phase mask profile. The height of the step is determined by the relative phase between the phasors shown in **a**.

(2.13). Take for example the interference of a Gaussian and a tophat but $\varphi_2 = 0$ instead of $\varphi_2 = \frac{1}{4}\tau$. Choosing α to yield an equal mixture of tophat and Gaussian, results in an annular intensity profile because the $\frac{1}{2}\tau$-shifted inner region of the tophat interferes destructively. The corresponding beam shape is Bessel-like[17]; a shape far from a deformed tophat.

2.6. Alternative Implementations

The order of the polarisation rotator and the double pass can be flipped. This might be beneficial when the implementation is scaled to several optical paths as presented in Section 3. In contrast to the previously described implementation, the SLM double pass input consists of a constant and equal mixture of s- and p-polarisation with arbitrary relative phase $\Delta\varphi_{\mathrm{sp}}$ (see previous section), e.g., a linear polarisation at 45°. Hence, both passes are always "active". The output polarisation is at 45° as well. Then, with the polarisation rotator placed after the double pass, one can control the relative amplitude of the interference by rotating the states into or out of the vertical axis which the PBS transmits: no rotation for an equal mixture, a ±45° rotation to obtain either one of the passes or anything in-between. The HWP in front of the PBS is not necessary in this setup. Furthermore, this variant might be able to compensate an imperfect input that is not an equal mixture of s- and p-states by attenuating the diffraction efficiency via the height of the diffraction grating, see Section 3.4.

Another more simple and cost-efficient implementation might be realisable with birefringent phase masks out of glass which are commercially available. These masks work similar to the SLM. Every "pixel" induces a predefined phase along a polarisation axis which, in contrast to the SLM, can be oriented freely during manufacturing. Here, however, all of them are required to be parallel. To realise the tophat plus Gaussian interference, the mask would induce some baseline phase, φ_0 for the unaffected polarisation due to propagation through the glass and $\varphi_0 + \frac{\tau}{4}$ for the inner part and $\varphi_0 + \frac{\tau}{4} + \frac{\tau}{2}$ for the outer part of the tophat mask[18]. The two polarisations of the incoming beam are independently patterned. As such, the mask would replace the entire double pass and can be combined the polarisation rotator before or after. However, it is feasible

[17] The ideal Bessel beam is generated by an infinitely thin and infinitely intense annulus.
[18] These relations are equivalent to the ones in (??) modulo τ. They are provided in this way as the glass cannot induce negative phase.

to rotate the mask to 45°; something that would not be possible with our SLMs. Hence, the input is not required to be diagonally polarised, if the polarisation rotator is placed after the mask. While this implementation does not employ an SLM for beam shaping, it might be required nonetheless for aberration correction. Furthermore, the glass is manufactured for a fixed wavelength and cannot be adjusted or calibrated like an SLM. Lastly, the same argument against using a single SLM pass with the zeroth order can be made here: any residual light that should have been patterned, but was not, will pollute the resulting beam shape.

3. Interferometric 3D Multi-Channel MINFLUX

We implemented a custom 3D MINFLUX system from scratch. The design follows that of an earlier 3D MINFLUX system [40]. However, instead of an electro-optic/varifocal lens, we employed the beam shaping technology described so far to enable 3D operation. Furthermore, we scaled the excitation module to have four parallel optical paths or channels; two with 640 nm and two with 560 nm wavelength. This highly flexible arrangement is utilised for new MINFLUX schemes that mix beam shapes within one excitation cycle to improve the localisation precision and the demonstration of dual emitter tracking in 3D by rapidly switching between wavelengths (see Section 4).

3.1. Optical Setup

The full system is found in Figure 3.1. Each element is labelled with a code that will be mentioned when appropriate throughout the rest of this section.

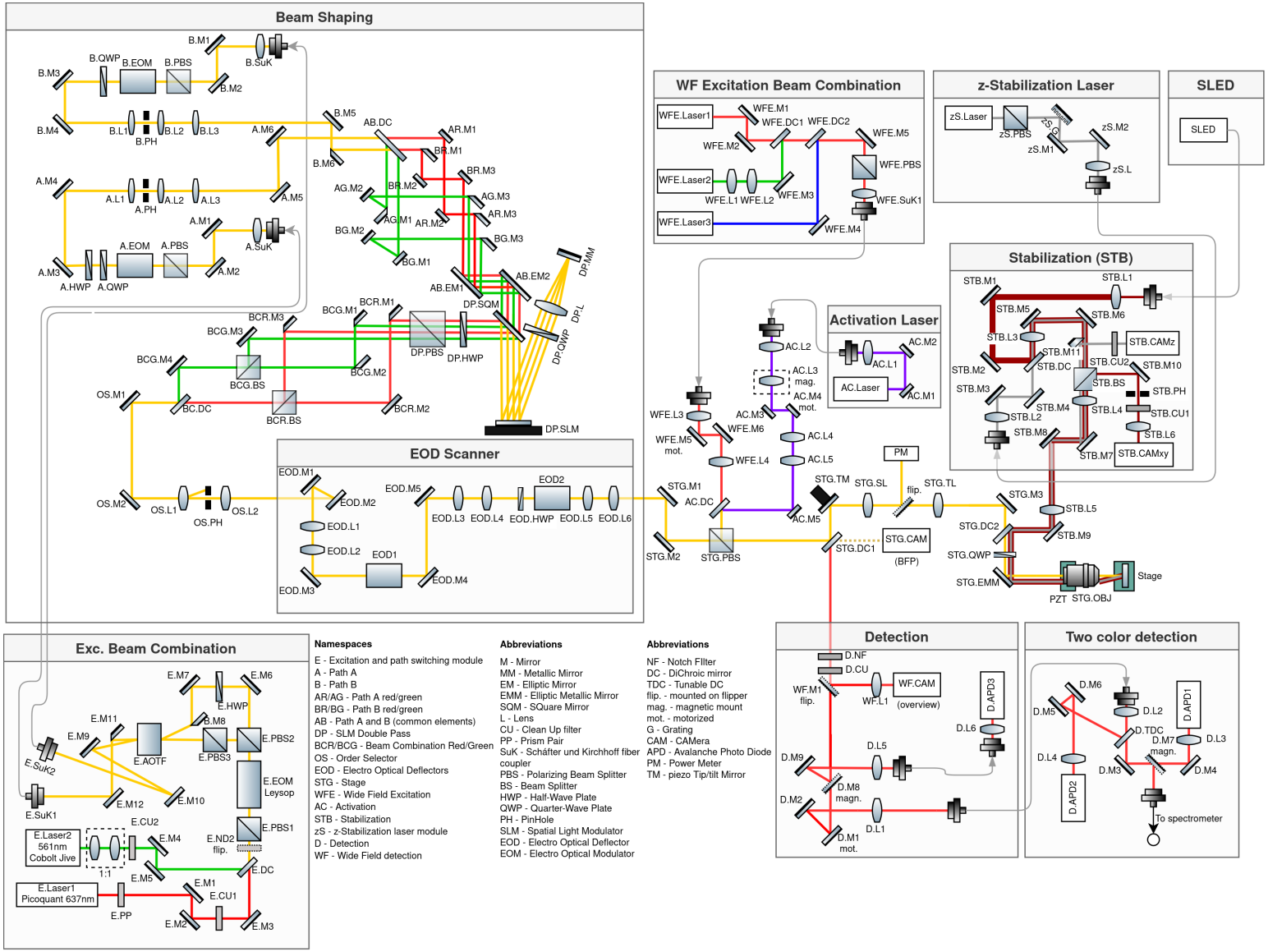


Figure 3.1.: Complete optics diagram of the custom MINFLUX system.

3.1.1. Excitation Source and Path Switch

Two main excitation wavelengths 640 nm 560 nm are employed. While the 560 nm laser (E.Laser2) is Continuous-Wave (CW) only, the 640 nm laser (E.Laser1) also offers pulsed operation for acquiring fluorophore lifetimes. Both laser beams are cleaned spatially, spectrally and in terms of their polarisation and are then combined via a DiChroic mirror (DC) into a common path. Next, a path switcher consisting of an EOM (E.EOM) and a PBS (E.PBS2) allows to rapidly switch between the reflected or transmitted path of the PBS by flipping the incoming vertical polarisation. The polarisation of the reflected output is cleaned by another PBS (E.PBS3). Both paths go through an Acousto-Optic Tunable Filter (AOTF) (E.AOTF) that is capable of modulating the power through diffracting part of the input power into a first order while keeping the path collinear for both wavelengths. Because the AOTF works with vertical polarisation only, a HWP (E.HWP) flips the polarisation of the horizontally polarised path. The AOTF is necessary as changing the laser power directly affects its lasing wavelength and is hence not an option for our interferometric scenario because it would require on-the-fly recalibration of the deflection voltages and the SLM. In addition to being the main power modulator, the AOTF also improves the extinction ratio of the path switch which is essential to suppress unwanted interferometric effects. More details are found in Section 3.5.

3.1.2. Beam Shaping Module

After the AOTF, the two paths (each carrying two wavelengths) are coupled into polarisation maintaining single mode fibres. The fibres provide flexibility and a clean spatial profile at the output. Then, two polarisation rotators (A/B.EOM), each with a QWP, are used. Again, conceptually the QWP is not necessary. Nonetheless, the QWP was adjusted to provide a clean polarisation at the input of the SLM double pass (although the exact setting was a trade-off between the two wavelengths). As seen in the full diagram, Figure 3.1, this adjustment can be necessary as there are many optical elements before the polarisation rotator's output reaches the double pass. In one of the paths an extra HWP (A.HWP) was added and adjusted as otherwise the polarisation input was not satisfyingly clean.

A remark: I propose to swap the order of polarisation rotator and the SLM double pass (as discussed in Section 2.6) for future systems of this kind. Only one rotator would be required to deal with all paths after they are combined.

A zoom or three-lens telescope with spatial filter follows the polarisation rotator. In theory, the three-lens telescope allows to adjust the beam size while keeping the beam collimated at the output by only moving the second lens, see Appendix C. In practice, however, the second and the third lens had to be moved. The beam size was measured on a camera and set to result in a fill factor of 1.75. The collimation was measured through the Zernike defocus term acquired through a Shack-Hartmann wavefront sensor (WFS20-K1/M, *Thorlabs Inc., Newton, NJ, USA*). The spatial filter with a 50 µm pinhole (A/B.PH) cleans the spatial profile the propagation through the EOM. A single DC (AB.DC) separates the two wavelengths for each path. A series of mirrors arranges the four optical paths in parallel with a spacing of 4.4 mm. To get the beams this close to each other, D-shaped mirrors were used. To save on space and reduce diffraction due to a long optical path, fixed mounts were placed on top of small linear translation stages (DT12/M, *Thorlabs Inc., Newton, NJ, USA*). As it turns out, the vertical reflection angle of these mounts vary enough to force height changes of the beam in the preceding part. This has strong effects on the polarisation state of beams with different wavelengths. Hence, I propose to glue the D-shaped mirrors into a pitch-adjustable flexure mount in the future (e.g., PFM1/M, *Thorlabs Inc., Newton, NJ, USA*) to keep the propagation of the paths as horizontal as possible.

The four paths impinge on the SLM double pass that is arranged as described in Section 2.5. In order to scale the setup to four paths, the SLM (DP.SLM) is divided into four pairs of two

passes, i.e., a total of eight regions. The first pass is on the upper row of the SLM. This row is above the optical axis of the imaging system between passes. Hence, the beams reflected off the first pass arrive on the lower row for the second pass. Similarly, the paths are mirrored laterally such that, e.g., the top left (A1) region corresponds to the lower right region (A2), see Figure 3.2. Each region can display arbitrary phase masks and aberration corrections (see Section 4.1) independently. They are furthermore calibrated for their respective wavelengths. The usual arrangement is shown in Figure 3.2: the red and green Gaussians, used for the initial iterations and for deflecting the tophat, are located in the first row marked with A1 and B1. The other two regions of the first row are inactive; only the second pass is used for these paths. The second row displays phase masks for the vortecis (D2, C2) and the tophats (B2, A2). Circular pupils are employed for each region that are relayed and magnified through the rest of the system to match the objective's back aperture. The inside of each pupil is filled with a grating that diffracts an incoming beam horizontally. To avoid contamination by stray light from outside the pupil, vertically diffracting gratings are utilised, see Figure 3.3.

All four paths are picked off by a two-inch square mirror (DP.SQM) glued onto a compact kinematic mount (KMSS/M, *Thorlabs Inc, NJ, Newton, USA*) after the second pass. The mirror is positioned such that the incoming beams can pass above the mirror's upper edge. A 3D CAD model is depicted in Figure 3.4 where arrows indicate the propagation direction for one of the red paths.

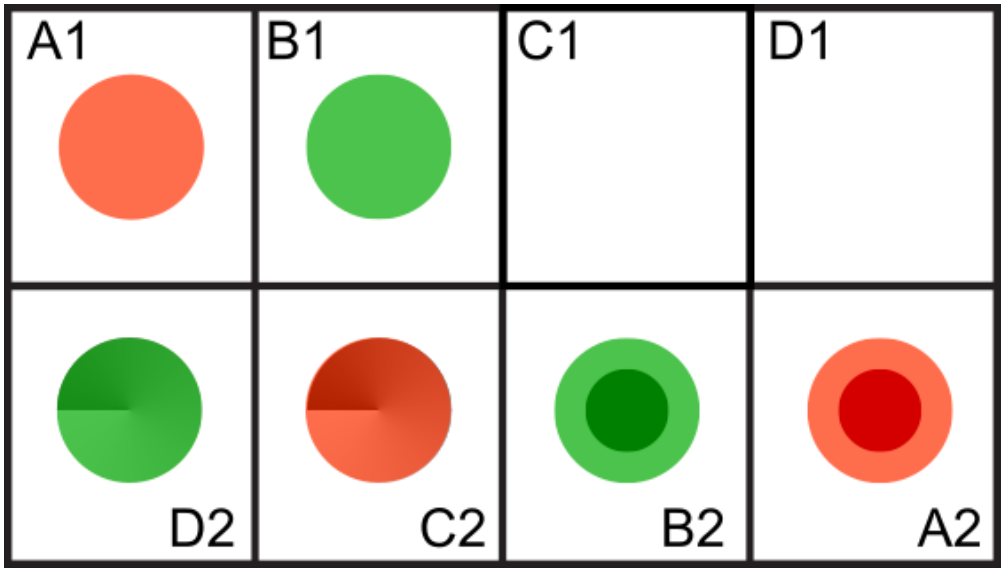


Figure 3.2.: The eight regions of the SLM. A1 and B1 generate Gaussians (flat phase mask). They are paired with A2 and B2, respectively, that use the tophat phase mask. C1 and D1 are typically not in use in the system as it does not require a second set of beams that can be interferometrically shaped (e.g., for lateral deflection). D2 and C2 display a vortex phase mask. Each region is calibrated to its wavelength and displays an aberration compensating phase mask on top (not shown here).

After the mirror, a HWP (DP.HWP) at 22.5° and a PBS (DP.PBS) interfere the two polarisation states of each beam. The PBS is oriented such that the transmitted output is vertical. In order to combine all paths into a single path a series of half-inch mirrors, two regular 50:50 Beam Splitters (BS) (BCR.BS, BCG.BS) and a single DC (BC.DC) are used. As all beams have the same polarisation, the use of BSs –instead of PBSs– is unavoidable for combining paths of

Figure 3.3.: Grating layout of an SLM region: The pupil in the centre has a grating diffracting an incoming beam horizontally. The light outside of the circular pupil is diffracted upwards by a vertical grating.

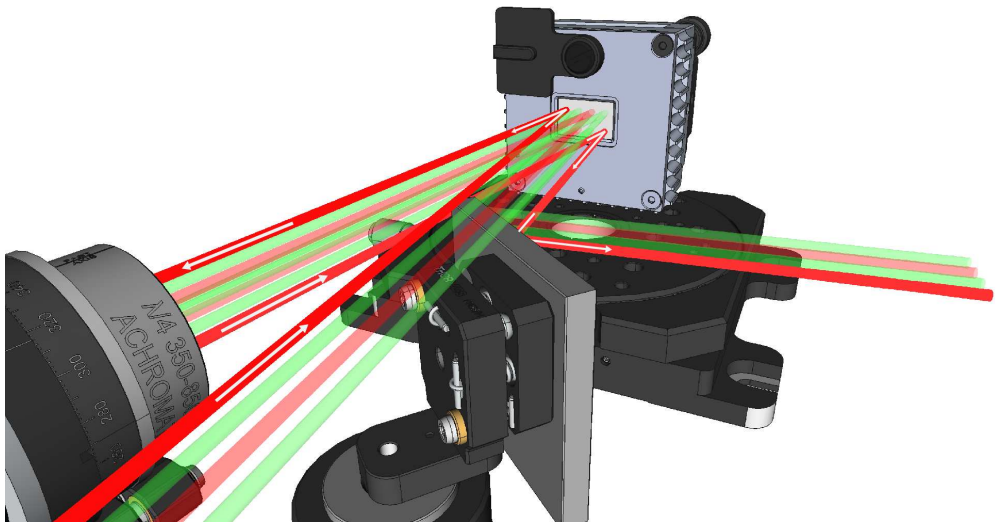


Figure 3.4.: CAD model of the double pass setup. The arrows indicate the propagation directions for one of the red paths.

the same wavelength. This means that half of the power is lost (leading to an overall loss of 75 % together with the interference PBS). However, having two paths of the same wavelength available allows advanced MINFLUX schemes that mix beams of different shapes for improved precision, see Section 3.3.7.

The zeroth order of the SLM has to be blocked and the image of the SLM surface has to be relayed through the rest of system into the back focal plane of the objective. Therefore, the beam combination is followed by a telescope with an adjustable iris (OS.PS) positioned at its focus on an XY-mount: as the diffraction orders have different angles, they will propagate in parallel with a lateral offset after the telescope's first lens. The iris is adjusted such that the first order carrying the interfered beams can pass while the zeroth order and the light originating from the vertical grating outside the pupil is blocked.

3.1.3. Image Relay and Lateral Scanning

The propagation after the double pass through the beam combination part has a path length of close to 900 mm. A regular 4f, relaying the SLM surface to the EOD, system would thus require an inconveniently long focal length of 900 mm for the first lens. However, a conventional telescope also relays images for a continuum of other pairs of points in a 'non-4f' configuration. one can calculate these points using ray optics, see Appendix D. Thus, a lens (OS.L1) with a focal length of 400 mm for the first lens of the telescope right after the double pass was used. To achieve a magnification of $M = 1/2$, the second lens (OS.L2) has a focal length of 200 mm: making the beam smaller is required to pass through the 2 mm aperture of the EODs without clipping the SLM pupil.

In order to deflect the excitation beam laterally in the microsecond regime, we employ EODs. These devices require that the image of the SLM surface is approximately located at half the crystal's length. Because of the non-4f system, another telescope with $M = 1$ (EOD.L1/2) was necessary to pick up the image since it was too close to the second lens (OS.L2).

The SLM surface is again relayed to the second EOD. From there it is further relayed and magnified to the piezo-electric Tip/tilt Mirror (TM) (STG.TM). This device has a larger scan range but operates in the millisecond regime. It is used to jump between regions of interest or do a 'mosaic scan' to generate a large image stitching together multiple detection volumes. Furthermore, it is employed during tracking measurements to centre the MINFLUX pattern (including the detection volume) on an emitter. Finally, the SLM surface is relayed to the back focal plane of the objective with a magnification chosen to match the SLM pupil with the back aperture of the objective.

3.1.4. Widefield Illumination and Activation

Three lasers with wavelengths of 560 nm, 640 nm and 488 nm (WFE.Laser1/2/3) are combined through DCs (WFE.DC1/2) and coupled into a single mode polarisation maintaining fibre. At the output, the 405 nm path is added via a DC (AC.DC) and the beam is focused onto the TM from where it is relayed to the back focal plane of the objective. The combination with the main excitation path is achieved via a PBS (STG.PBS).

For MINFLUX measurements using the STORM or PALM modalities, a point like activation filling the detection volume is required. Hence, a 405 nm laser is coupled into a single mode polarisation maintaining fibre. The output is collimated and relayed by a 4f system where a motorised mirror (AC.M4) is positioned one focal length away from the the first lens (AC.L4): it enables convenient lateral re-alignment of the focused activation beam in the sample plane[1].

[1] As the collimated beam on the mirror is relayed to the back focal plane of the objective, a change in angle translates to lateral movement in the sample.

An optional lens (AC.L3) on a magnetic mount focusing on the motorised mirror can be placed to switch WF activation.

3.1.5. Stage and Stabilisation

After the TM the last 4f system follows. Even though it consists of regular achromatic two-inch lenses, they are referred to as the Scan Lens (SL) and the Tube Lens (TL). A flippable mirror is placed between SL and TL such that the main path is redirected and focused on a power meter. An old metallic mirror was chosen such that the measured power matches the power after the objective within 5 %.

The last QWP (STG.QWP) making the light circularly polarised was placed before the DC coupling in the infrared lasers used for stabilisation and the objective. All measurements presented in this work were done with the last QWP optimised for the 640 nm [2]. However, the QWP, which sits on a magnetic mount, can be exchanged for another optimised for 560 nm. The last mirror (STG.EMM) is metallic as it affects the polarisation state less for different wavelengths. It points upwards towards the objective. The objective is placed on a coarse z-stage that is controlled manually to find the focus roughly. The stage is a stack of a coarse xy-stage and a fine xyz piezo stage with sub-nanometre precision.

The active stabilisation system consists of a dark-field microscope imaging fiducial gold particles and a focus-lock working with the reflection of the glass-water interface of coverslip and sample medium. For xy, the dark-field image of scattered infrared light (840 nm) from a superluminescent light emitting diode (SLED) is acquired on a camera (STB.CAMxy) and continuously monitored: all of the imaged gold particles' locations are fit via a Gaussian and fed into a feedback loop controlling the stage. The stage compensates drift such that the averaged movement of the fiducials is zero. For z, the oblique reflection of the coverslip-water interface is imaged onto a camera (STB.CAMz). Any movement in z causes a lateral shift on the camera. Again a Gaussian fit quantifies the drift and uses the stage to compensate. The z-stabilisation laser (zS.Laser) experienced mode hopping which caused the monitored Gaussian to move and hence induce jumps of the fine stage in z. Therefore, the laser was stabilised via an external cavity made from a diffraction grating (zS.G) in the Littrow configuration: a narrow spectral width is fed back into the laser improving the stability.

3.1.6. Detection

The fluorescence is descanned with the TM and transmitted through the main DC. A quad-notch (D.NF) and infrared filter (D.CU) block residual excitation and infrared light, respectively. A mirror on a motorised flipper (WF.M1) can redirect the light to a camera (WF.CAM) used for widefield detection. Before a MINFLUX measurement, sites of interest in the sample are chosen on the camera. During the measurement, the mirror is flipped down such that the fluorescence is directed towards one of the confocal detectors and the TM moves the confocal detection volume and the excitation to these sites.

The main confocal detector consists of a 150 mm lens focusing into an anti-reflection coated multimode fibre with 5 µm core diameter acting as the confocal pinhole, a collimating lens on the other side of the fibre, a tunable DC (D.TDC) that separates the fluorescent signal into two spectral bands and, finally, two APDs (D.APD1, D.APD2) –one per spectral band– with a lens focusing on the sensor. A motorised mirror (D.M1) before the lens focusing on the fibre is used to correct the alignment such that the detection volume is centred on the focused excitation in the sample. The tunable DC is installed on a manual rotation mount (RP01/M, *Thorlabs Inc, NJ, Newton, USA*) together with a mirror in a configuration that allows to rotate the contraption

[2] Optimisation through rotation but also tilt.

without changing the alignment of the reflected light, i.e., the incidence of the outgoing beam[3]. This way, the cut-off wavelength can be easily set while measuring the spectrum[4] without having to change the alignment for the APDs afterwards.

The detection volume of the main confocal detector has a lateral extension of about 600 nm FWHM. While this volume allows for good detection efficiency over a typical MINFLUX localisation volume of around 400 nm it is not ideal for PSF measurements. When acquiring a PSF by scanning the TM, the detection volume and the excitation light is scanned laterally in unison around a fluorescent bead, fiducial or single molecule. The result is the convolution of the detection volume with the excitation PSF. Side lobes outside the detection volume are hence suppressed which is especially noticeable for the tophat PSF along z where the axial side lobes extend over several microns. Therefore, an alternative detection path with a 3 µm lateral FWHM confocal detector was installed. It is activated by inserting a mirror (D.M8) on a magnetic mount into the main detection path where a 30 mm lens focuses onto multimode fibre. A single lens at the output focuses the light directly onto an APD (D.APD3).

3.2. Interferometer Tuning

I established a procedure to optimise the interferometer which is necessary to arrive at a good zero degradation curve. The idea is to tune Zernike terms on the SLM such that destructive interference across the whole beam-profile is achieved. The read-out is the power of the interfered beam that, ideally, goes to zero.

To that end, the angle/voltage of the polarisation rotator, $\alpha(V)$, is set such that the power of both passes are equal, i.e., $\cos(\alpha)\cos(2\theta_{\mathrm{HWP}}) = \sin(\alpha)\sin(2\theta_{\mathrm{HWP}})$ in Equation (2.10). At this stage, the corresponding SLM pupils only display the aberration correction but no tophat phase mask. Next, the relative phase between the passes needs to be compensated; it corresponds to the unknown $\Delta\varphi_{\mathrm{sp}}$ term in Equation (2.10). The 'piston' Zernike term on the first pass (φ_1) is adjusted to match $\Delta\varphi_{\mathrm{sp}}$. For perfectly overlapping passes and homogeneous intensity profiles the interference would yield zero. Realistically, the two passes are never collinear, however, residual misalignments can be decently compensated by adjusting the tip, tilt and defocus terms on the SLM. All SLM adjustments are iteratively improved while measuring the power of the outgoing beam on the power meter.

The procedure was automated and run regularly. The automation finds the right polarisation rotator voltage by scanning the voltage while only one pass is active, yielding two power curves (one for each pass). Their crossing is calculated through linear fits and corresponds to the desired voltage. Then, while both passes are active, the procedure scans the piston term of the first pass coarsely over a whole period to find the initial value. Next, it iteratively scans the defocus, the tip/tilt and piston (finely) terms. Each time, a power curve is measured and fitted to a parabola to find its minimum. The location of the minimum serves as the initial point for the next iteration that scans the values around that point with a reduced range. The sequence is repeated three times. Less naive approaches using a gradient descent algorithm and including additional Zernike terms might perform better.

The destructive interference is quantified by the ratio of the power in the second pass and the remaining power measured when both passes are active; typical values were around 8 % for the 640 nm laser and around 2 % for the 560 nm laser. Afterwards, a piston of $\frac{1}{4}\tau$ is added to the second pass which is required to yield the correct phase relation between the Gaussian and the tophat (see Section 2.5.4). Then, the tophat phase mask is applied on the second pass, the polarisation rotator voltage is reset and, as a control, z-profiles (as seen in Figures 3.10 and 3.11) are acquired to see if the zero degradation stays flat.

[3] The light transmitted through the tunable DC is assumed not to move.
[4] For this purpose, a clip-on book light with LEDs was used as the light source.

Further improvements could be achieved by compensating the spatially inhomogeneous relative phase between the two passes, i.e., the relative phase per pixel. Some attempts were made in this direction: while both passes carried the same power, the relative phase (piston term on second pass) was scanned in four steps and the resulting intensity modulation was measured on the back focal plane camera (STG.CAM). Equivalent to the pupil segmentation described in Section 4.1, the phase could be extracted for each camera pixel. The resulting phase landscape was smoothed and mapped onto the SLM pupil by utilising the theoretical magnification from the SLM to the camera, the camera pixel size and, finally, manual offsetting to minimise the residual power. While the destructive interference could be improved the deflected profiles did not change too much; the approach was not further pursued. However, with a proper and automated mapping[5] it could be a worthwhile addition to the setup routine.

3.3. Characterisation

This section demonstrates the beam shaping capabilities and characterises the speed as well as the zero degradation through PSF measurements. Furthermore, it presents a CRB calculation for a novel 3D MINFLUX scheme that improves the CRB over the conventional 3D MINFLUX scheme. It utilises laterally deflected vortex beams for localisation in xy and axially deflected tophats (or tophat zeros) for localisation in z.

3.3.1. PSF Measurements via Back-Scattering on Gold Particles

For the PSF gallery (Figures 3.5 and 3.6) and the speed characterisation (Fig. 3.7), back-scattered excitation light of gold nanoparticles with 40 nm diameter was used. In contrast to fluorescent beads that can saturate and bleach, the obtainable signal is unlimited because the signal photons come from the excitation laser itself as they are reflected or back-scattered by the gold particles. The sample preparation is described in Appendix G. A PSF measurement on gold particles differs from a PSF measurement on fluorescent beads in two points. First, one has to remove the excitation filter in the detection path and practice caution to not burn the APDs accidentally. Second, the optical adhesive used to embed the gold particles should match the refractive index of the glass coverslip. Otherwise, there are strong artefacts in the measured PSF stemming from direct reflections of the excitation laser on the glass-water interface. This, however, means that the ability to use the stabilisation system for axial locking is lost since the infrared laser is also not reflected anymore. Locking in xy is still possible though.

3.3.2. PSF Measurements via Single Molecules

Gold particles and Fluorescent beads employed in this work had a diameter of 20 nm to 100 nm. Measured PSFs are thus the convolution of the excitation beam's intensity distribution with the shape of the particle or bead. Especially in the case of gold particles, the shape can introduce significant artefacts to the measured PSF. Also, it is not possible to probe the zero of a beam as the spatial extension of the particle or bead will integrate signal from the surrounding intensity. Thus, for measurements focused on the zero[6], single molecule samples were employed: a single ATTO647N molecule was statically bound to a DNA origami. For most measurements bleaching was not a hindrance; the molecule stayed on for several minutes. Measurements like the deflected z-profiles in Figures 3.10 and 3.11 as well as bleaching experiments to determine the

[5] For example, as an initial step, one could display a grid on the SLM pupil and fit it on the camera to determine the centre coordinate as well as the magnification.

[6] The single molecule samples can also be used for "regular" PSF measurements capturing the whole shape of the beam.

zero quality were acquired through single molecule samples. Furthermore, they were employed for overlapping the the zeros of the four paths through fine adjustments of the tip, tilt and defocus terms on the SLM.

3.3.3. Beam Shaping

The beam shaping capabilities of the system are demonstrated for several beam shapes including most of the cases shown in Figure 2.4. Figures 3.5 and 3.6 show measurements of the excitation PSF acquired through back-scattering on gold nanoparticles. They were acquired by scanning the excitation beam with the EODs over a range of 800 nm by 800 nm in xy and 800 nm by 3000 nm in xz, where the scan in z was achieved via the stage. The voltage of the polarisation rotator was linearly scanned in equidistant steps to acquire a stack of 33 frames, each showing the interference for a different relative amplitude of the two interfering beams. The acquisition of all frames took about 3.5 min. For the half-moon measurement shown in Figure 3.5 **e** the last QWP (sitting on a magnetic mount) was removed to obtain linear polarisation at the input of the objective (otherwise, there is no true zero).

The xy-PSFs was acquired while the system was locked in xy. However, for the xz-PSFs this was not possible: the xy locations of the fiducials that are locked depend on their position in z. As the stage is moved to acquire the PSF along the optical axis, using the xy lock during such an acquisition would result in slanted PSFs. To prevent drift in xy, the xy-lock was set before the acquisition and periodically activated for a short time at a given z position during the scan.

Panels **a-d** show the interference of a $n\tau$-vortex or charge-n Doughnut with a Gaussian with $n = [1, 2, 3, 4]$, respectively. the central zero splits into n zeros that move outward. The higher the charge the more sensitive the PSF is to remaining aberrations that our approach cannot compensate. For $n = 4$, panel **d**, the acquired PSF –especially around the zeros– degrades significantly. Panels **e** and **f** show the lateral deflection of the half-moon zero plane. **e** shows the half-moon with the appropriate linear polarisation state. **f**, on the other hand, shows the same for circular polarisation that produces a bad zero, even though it is hard to see here.

3.3.4. Speed

Again, back-scattering on gold particles was used for measuring the transition of the deflected tophat zero through a series of PSFs captured at different time points during the transition. While the high signal of the scattering was useful to measure good looking PSFs as seen in the previous Section 3.3.3, it was essential for this measurement.

The large 3 µm detector was used (see Section 3.1.6). As the transition happens in the microsecond regime, each pixel of the xz-PSF was sampled during the transition by integrating the signal over a time window of 200 ns [7] at time points 400 ns apart. The laser power was set to yield a count rate of approximately 8 MHz, which is quite high compared to typical MINFLUX count rates of 50 kHz to 300 kHz but far away from the saturation point of 37 MHz specified for the APD.

The measurement deviated from regular PSF measurements in that it was using two consecutive tophat exposures: one at the starting point of the deflection and the second at the deflected target location. The APD counts of both exposures were recorded. All counts of the first exposure were collected while the second exposure was gated only using counts occurring during a 200 ns time window. This time window was situated within the time frame during which the system waits for the axial deflection to be completed and the laser is off. Here, however, the laser is kept on and the time window is shifted in steps of 400 ns after each completed xz-scan to probe the deflection state at different time points. Ultimately, two xz-PSFs of the tophat were

[7] This is the minimum window length that the FPGA allows because it operates on a $1/200\,\text{ns} = 500\,\text{MHz}$ clock.

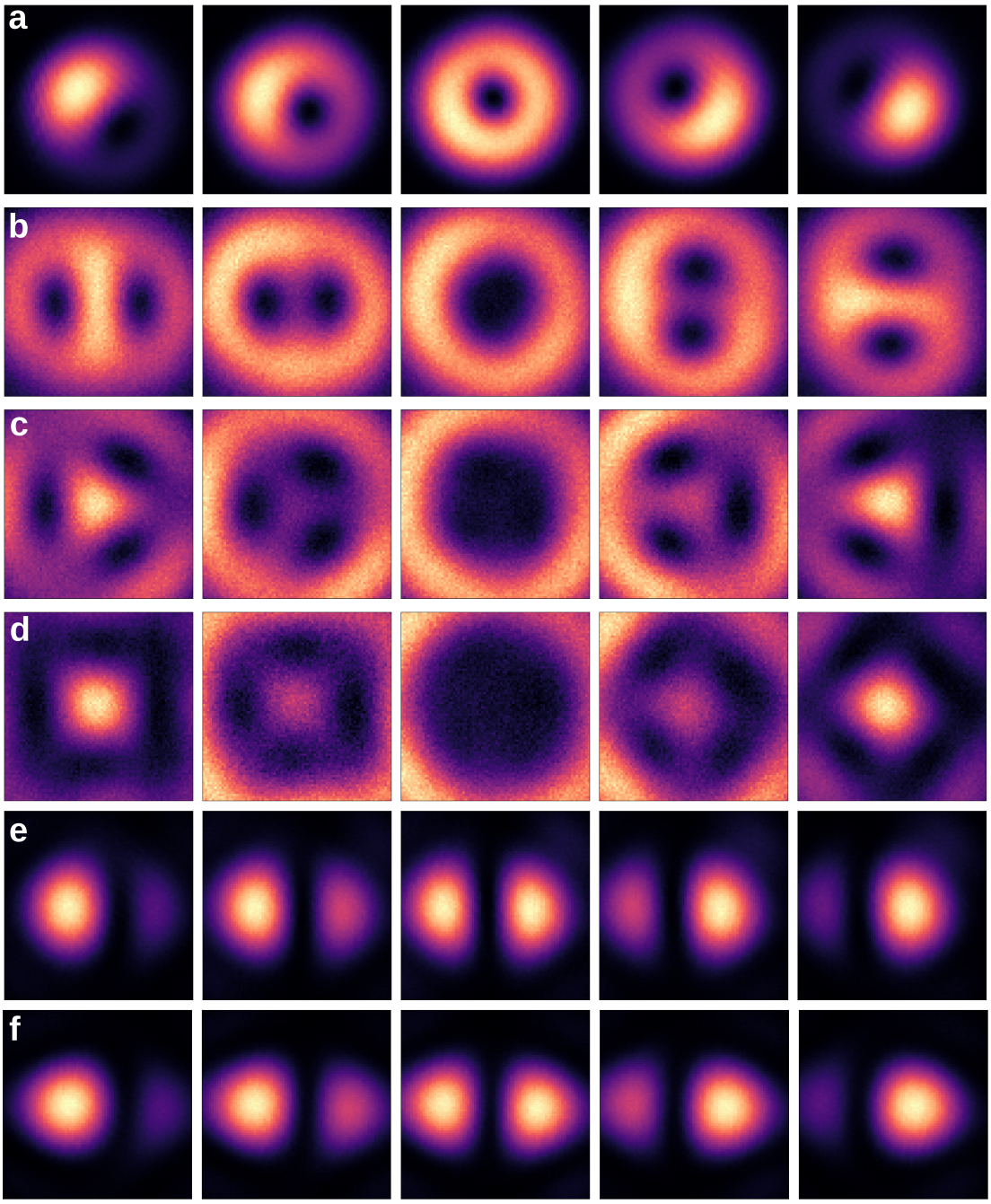


Figure 3.5.: Measured interferometrically shaped PSFs in xy (800 nm×800 nm) acquired through backscattering of 40 nm gold particles. **a** Vortex plus Gaussian; the relative phase was set via the SLM such that the deflection occurs diagonally. **b** 2τ-Vortex plus Gaussian. **c** 3τ-Vortex plus Gaussian. **d** 4τ-Vortex plus Gaussian. **e** Half-moon plus Gaussian, linearly polarised. **f** Half-moon plus Gaussian, circularly polarised.

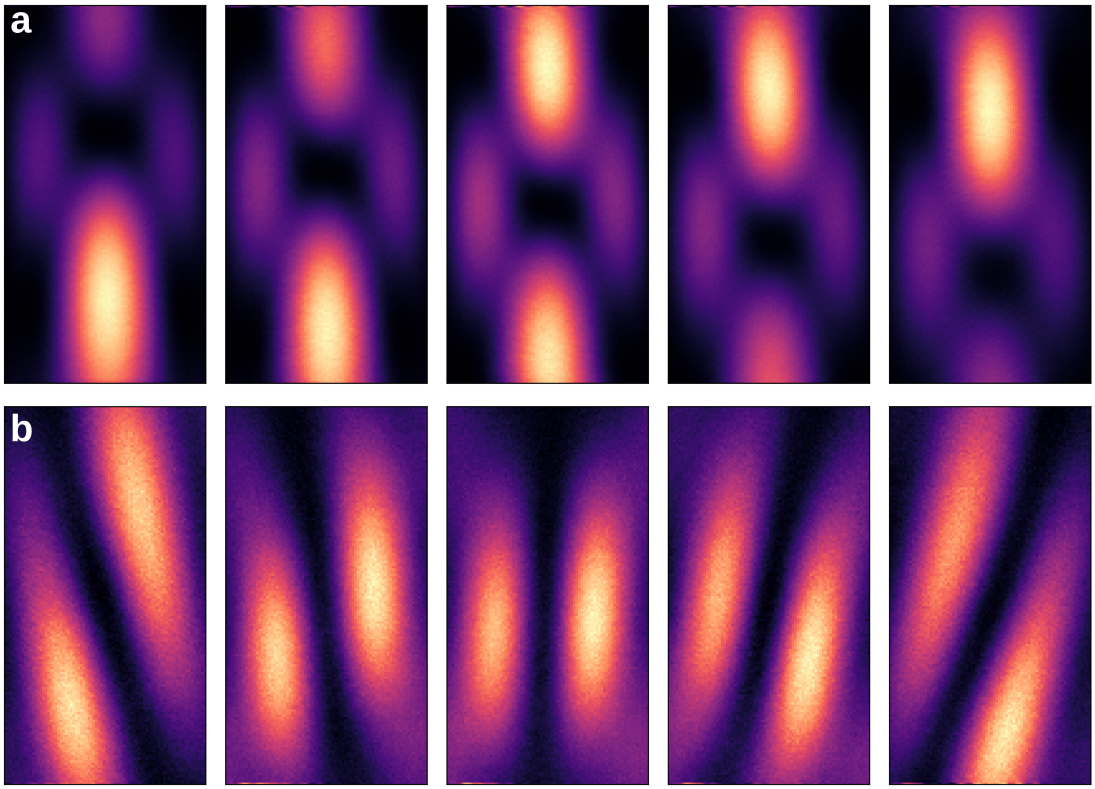


Figure 3.6.: Measured interferometrically shaped PSFs in xz (800 nm × 3000 nm) acquired through backscattering of 40 nm gold particles. **a** Tophat plus Gaussian. **b** Vortex plus tophat.

recorded for several frames: a static PSF with high signal and a dynamic PSF with low signal due to gating.

In order to track how the tophat zero moves along z, the dynamic PSF's zero location was retrieved by fitting a 1D parabola along z around its minimum. The results are shown in Figure 3.7. The deflection was set to $L_z = 271\,\text{nm}$ corresponding to 190 nm in an aqueous sample. The measurement was repeated three times and the resulting deflection trajectories and their average are shown in panel **a**. Some of the dynamic xz-PSFs for indicated time points are shown below in panel **b**. As the measurement was done on gold particles in a refractive-index matched medium and in xz, there was no possibility to lock the sample. Hence, to compensate for potential drift in z, the z location of the zero of the static PSF was fitted directly measuring the drift. This drift trajectory was subtracted from the trajectory acquired for the dynamic PSF. One sees that the transition is completed within 5 µs which is chosen to be the wait time between two MINFLUX exposures in the system. Nevertheless, there is a delay of about 1.4 µs before the transition starts which could be reduced or eliminated by sending the signal initiating the deflection earlier.

Similarly, Figure 3.8 shows axial trajectories for different deflection distances. All trajectories reach their target in the same time, independent of the deflected distance. The curves resemble the charging curve of a capacitance

$$U_C(t > 0) = U_s \left(1 - e^{-\frac{t}{\tau_{\text{RC}}}}\right), \tag{3.1}$$

with the capacitor voltage U_C, the source voltage U_s and the time constant τ_{RC}. Dividing Equation (3.1) by U_s gives the percentual charging curve that is independent of U_s. This hints at that the main limitation of the speed is not the amplifier but the capacitance.

The driving amplifier provides up to 100 mA output current which limits the steepness of the voltage change at the EOM when modelled as a capacitor ($I_C = C\dot{U}_C$). Hence, larger deflections, i.e., higher voltage jumps, might saturate the current during the transition which slows it down. However, this is not the case as all transitions exhibit the same duration. Therefore, in order to further increase the speed one should use shorter cables to connect the amplifier and the EOM[8].

3.3.5. Duty Cycle

The interferometric axial deflection is as fast as lateral deflection through EODs employed in this system, which also takes 5 µs. Other technologies for 3D MINFLUX cannot compete in terms of speed: A varifocal lens can refocus the excitation beam in the nominal MINFLUX range of 400 nm in about 50 µs[40] and a high-speed deformable mirror as used in [43] is specified to have a mechanical response of less than 40 µs [77]. These slow transition times can have a significant impact on the duty cycle of a 3D MINFLUX cycle; particularly in high speed tracking applications.

The duty cycle for is given by

$$D = \frac{\sum_i t_{\text{on},i}}{\sum_i t_{\text{on},i} + t_{\text{off},i}}$$

where $t_{\text{off},i}$ is the wait time before the i-th exposure, meaning that the laser is turned off, and $t_{\text{on},i}$ is the exposure time, i.e., the time during which the laser is on.

To compare the duty cycle of a conventional 3D MINFLUX microscope with a slower axial deflection and the presented implementation, assume the following scenario: the conventional system has a 30 µs wait time before each tophat exposure along z; a favourable assumption

[8] In the presented setup the total capacitance is roughly 232 pF: 82 pF as specified for the EOM and approximately 150 pF from the coaxial cables of 1.5 m length (100 pF per metre).

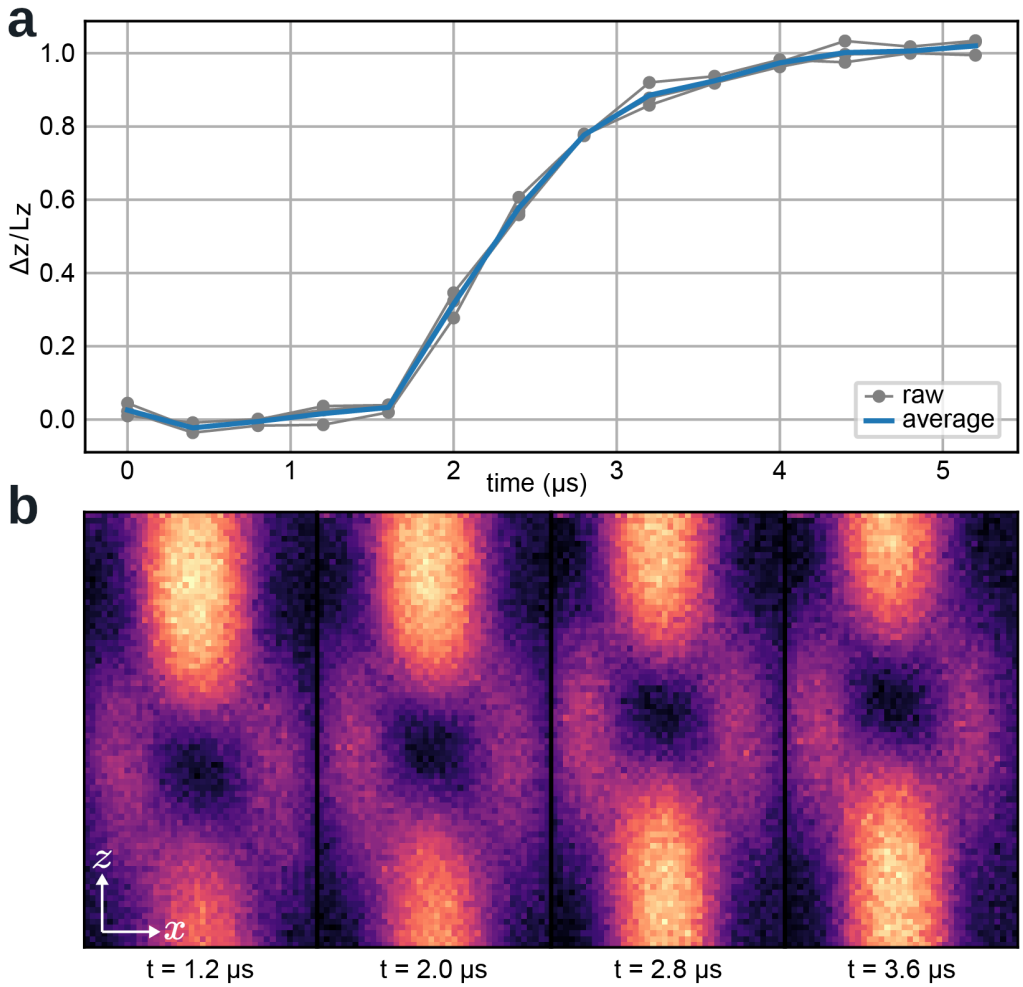


Figure 3.7.: Transient of the interferometrically deflected tophat zero. The commanded deflection is a step of $L_z = 271\,\mathrm{nm}$. **a** raw (grey) and average (blue) transient curves tracking the zero position obtained from fitting PSF measurements (examples are shown in bottom panel). The y-axis is normalised such that $y = 0 \widehat{=} -135.5\,\mathrm{nm}$ and $y = 1 \widehat{=} 135.5\,\mathrm{nm}$. **b** xz-PSFs of the moving tophat for four time points of the transition.

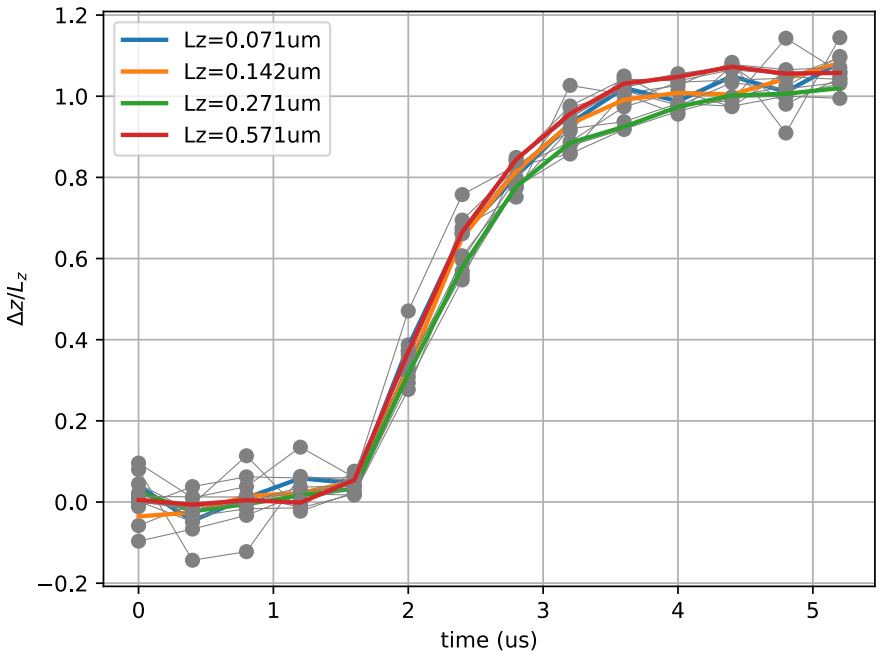


Figure 3.8.: Comparison of deflection transients of the tophat zero for a range of L_z. The y-axis is again normalised to the deflection range revealing that the transient time is independent of L_z. This points to the capacitance of the electrical connection as the speed-limiting factor. The red curve is clearly settling above $y = 1$, i.e., deflecting more than $L_z = 571\,\text{nm}$ ($= 400\,\text{nm}$ in water), revealing that the calibration was not ideal during this measurement.

imagining it used a deformable mirror. The wait time between xy exposures is 5 µs. The interferometric system, on the other hand, has an actual wait of 5 µs no matter the exposure. Both systems use a mixed beam scheme, as presented in Section 3.3.7: it employs three tophats along the optical axis and four vortecis around the centre in xy. The vortex exhibits higher curvature in xy than the tophat along z. Hence, in order to achieve isotropic localisation precision the tophat exposure weights are doubled with respect to the vortex exposures. Increasing the weight can either be done through a prolonged exposure time or an increased laser power (or a mixture of both). For time-critical applications, however, adjusting the power (if possible) is the only reasonable option.

Figure 3.9 shows the results of the comparison. The upper plot of panel **a** shows how the duty cycle changes as a function of the MINFLUX cycle frequency. The blue curve corresponds to the interferometric system while the orange curve is that of a system with a 30 µs wait time for every axial deflection. The on-times were chosen to be equal for each exposure and such that the total length of the two MINFLUX cycles are the same. At typical 6.477 kHz corresponding to a cycle length of 154.4 µs, the interferometric system exhibits a duty cycle of 76.2 % which drops to 27.6 % for the conventional system. The lower plot shows the power multiplier necessary to keep the emitted photon count (for a given fluorophore) constant: as the time windows during which the laser is on are shorter, the higher power compensates by increasing the emission rate. The duty cycle for the conventional system drops to zero and the power multiplier diverges when the total cycle length is equal to the sum of its off-times which is 111.8 µs. At the corresponding frequency of 8.945 kHz, our system still achieves a duty cycle of 67.1 % with a power multiplier of 1.57. Figure 3.9 **b** shows the timing diagram of the power modulation of the exposures for the two cases: the left column corresponds to $f = 6.477\,\text{kHz}$, the right column corresponds to $f = 8.945\,\text{kHz}$. Notice the doubled weights for the tophat exposures (exposure 1, 6 and 7) and the infinite power multiplier compensating the infinitely short on-time in the upper right panel of **b**.

Consequently, the interferometric system can unlock 3D tracking speeds beyond what is possible with slow axial deflectors like varifocal lenses and deformable mirrors. But even before that point the duty cycle is important: the amount of collected photons per cycle will decrease diminishing the localisation precision, if the power is not increased. Exciting the fluorophores with more power can compensate the reduced duty cycle to some extent but the bleaching probability increases, leading to shorter traces, and –in extreme cases– the fluorophore might saturate.

3.3.6. Zero Quality

Figure 3.10 juxtaposes simulated (solid) and measured z-profiles (dashed) of a deflected tophat zero. The upper panel shows the overview including the maximum of the axial lobes of the tophat. The lower panel shows a zoom-in of the grey shaded area of the top panel highlighting the local shape of the deflected zeros.

The measured PSFs were acquired on a single fluorescent molecule to avoid shape artefacts of spatially extended beads or gold particles. First, the tophat zero was centred on the molecule and the xyz position locked in. Then, the stabilisation system was unlocked to perform a 2D scan with the stage and voltage of the polarisation rotator. This was repeated three times for averaging out noise. The resulting profiles show the intensity of the tophat along the optical axis for several deflection offsets, i.e., polarisation rotator voltages. The simulated profiles were scaled in intensity to match the maximum of the measured profiles. The simulation showcases a perfect zero for the undeflected tophat at $z = 0$ while all measured profiles are offset by background and the imperfect zero of the tophat. Still, one can see that there is no significant degradation over a deflection range of ±200 nm. For a ±300 nm deflection the zero clearly degraded similar to the simulation where, however, the degradation is very slight. Furthermore, all profiles are parabolic close to their minimum.

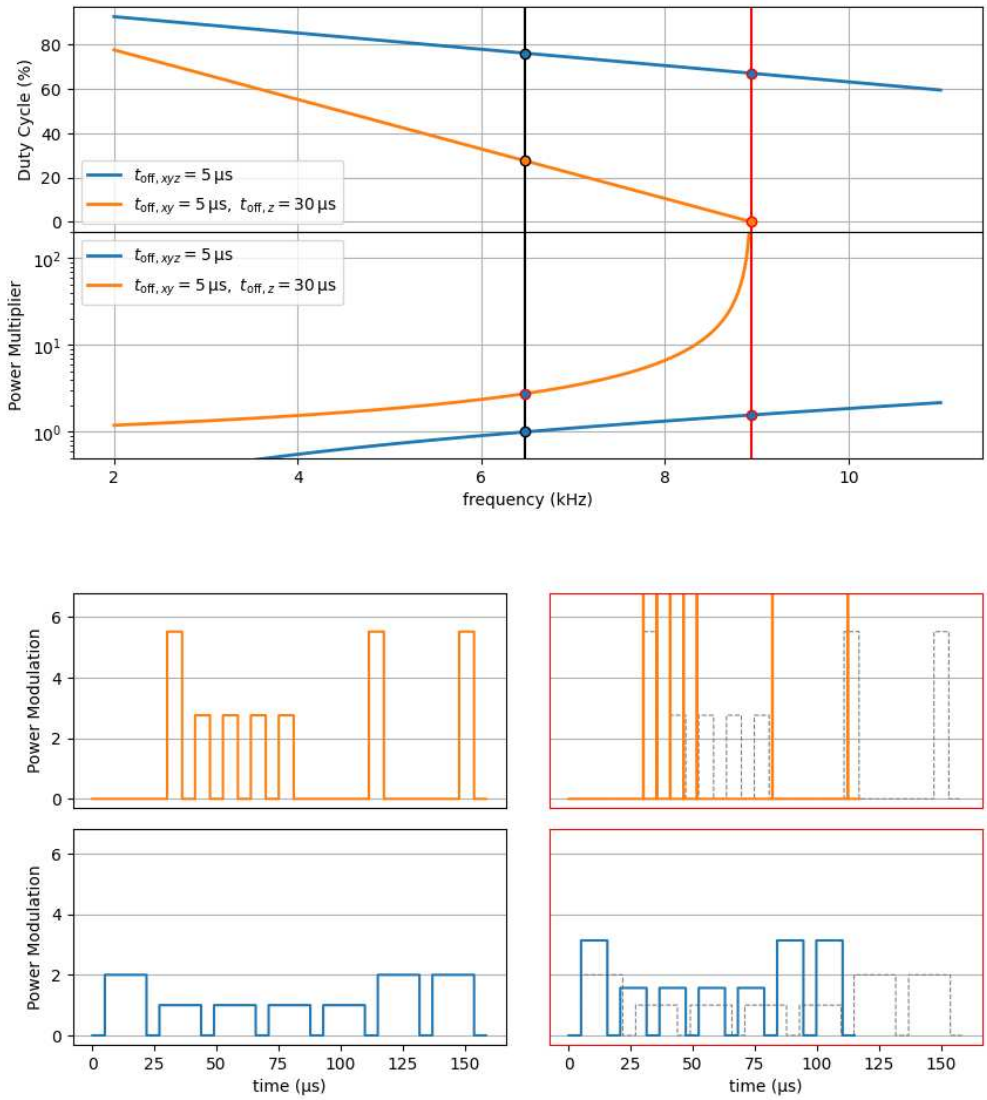


Figure 3.9.: **a** Duty cycle and power multiplier as a function of the MINFLUX cycle length expressed in frequency f (how many cycles/localisations per second). The blue dots with black edges mark the typical 3D cycle of the presented system with a cycle length of 154.4 µs. **b** Timing diagram of the power modulation for the two cases marked in **a**: left column $f = 6.477\,\text{kHz}$, right column $f = 8.945\,\text{kHz}$.

Figure 3.11 is equivalent to Figure 3.10 with the important distinction that the excitation laser was switched to pulsed mode. The same observations regarding the profiles can be made, however, the zero degradation is more pronounced. Most likely due to the fact that the reduced coherence length of the pulses affects the interference negatively. Nonetheless, it demonstrates that our interferometric deflector is capable of handling pulsed excitation sources enabling 3D MINFLUX with lifetime information.

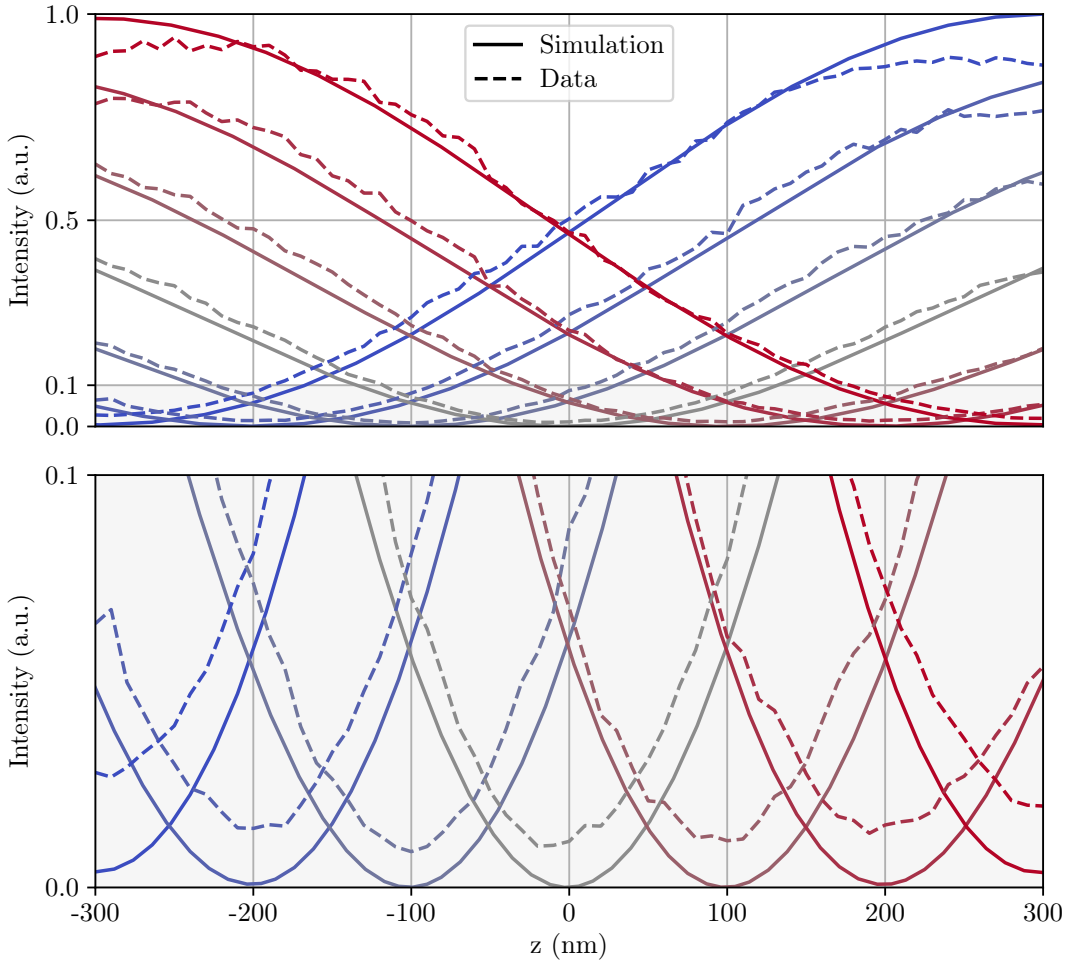


Figure 3.10.: Intensity profiles along z for several deflected tophat zeros. Simulations (solid) and averaged measurements (dashed), acquired on a single molecule with 640 nm excitation laser, are shown. The intensity was normalised to its maximum. The simulation has no background added which, however, is present in the measurement. The lower panel is a zoom-in of the grey shaded area in the upper panel. The deflection minima are approximately on the same level over a range ±200 nm.

3.3.7. Mixed Beam Schemes for 3D Minflux

So far, 3D MINFLUX systems solely relied on the tophat beam shape to produce information in all three dimensions[40, 43], or employ other modalities like graphene energy transfer where the

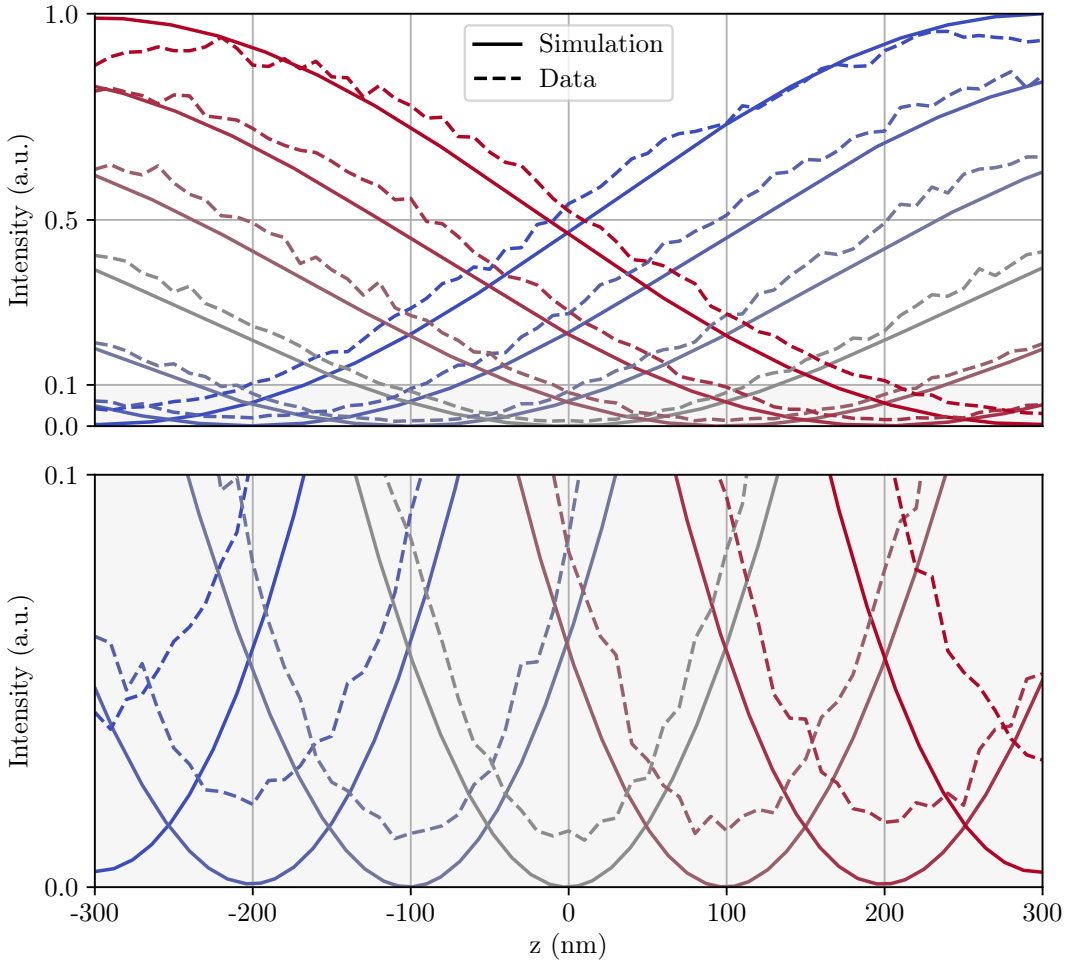


Figure 3.11.: Same as Figure 3.10 except that the excitation laser was used in pulsed mode demonstrating that the interferometric deflection is compatible. Still, the zero degradation is worse, e.g, for $z = -200\,\mathrm{nm}$, than its CW counterpart; potentially due to reduced coherence.

lifetime of the emitter depends on its axial distance to a bottom graphene layer[78].

However, the tophat is not the ideal choice for producing lateral MINFLUX localisations: compared to the vortex, the curvature across xy is lower and the intense axial lobes will produce more background. These axial lobes, however, make the tophat a good choice for localisations in z. Irrespective of the dimension, the tophat is more sensitive to aberrations that degrade its zero quality[79, 80] than the vortex[81, 82].

3D STED systems have long since employed an incoherent mixture of a vortex and a tophat for depletion to facilitate good confinement in all three dimensions[83]; also using the SLM double pass geometry discussed in Section 2.5[75, 76]. The same principle applies to 3D MINFLUX but due to its sequential nature it is not the incoherent mixture but a mixed sequence of vortex and tophat excitation beams.

The mixed beam MINFLUX cycle employed mostly throughout this work positions four vortecis laterally and three tophats axially and is denoted as 'VTH'. It provides better localisation precision as the conventional tophat-only scheme, denoted '7TH', as Figure 3.12 demonstrates: it shows CRB calculations comparing the two schemes. The x-axis specifies the radial distance of the emitter from the pattern origin. The y-axis corresponds to CRB values: the maximum and the minimum of the eigenvalues of the 3×3 (xyz) CRB matrix are shown. Instead of lines the plot shows bands, i.e., for a given radial distance we observe a range of possible values. The CRB eigenvalues depend on all spherical coordinates, not only the radial distance. Therefore, the band encompasses all values occurring on a spherical surface with a radius equal to the given radial distance (between emitter and pattern origin). The yellow and green bands correspond to the 7TH and VTH pattern, respectively, for a $L = 50\,\text{nm}$, SBR $= 5$ and a single collected photon. As can be seen, the VTH scheme outperforms the 7TH scheme. Especially when the emitter is (close to) centred on the pattern, the worst (maximal) CRB eigenvalues of the VTH lie below the best (minimal) CRB eigenvalues of the 7TH.

Figure 3.13 visualises different MINFLUX excitation schemes that the presented system enables. For a first rough localisation in 2D, a seven Gaussian (blue dot) excitation pattern, '7G', shown in **a**, is employed. Compared to a vortex only pattern which works for an L_{xy} of up to 200 nm to 250 nm, it is able to localise over a wider region. This allows to converge on emitters that are initially close to the edge of the MINFLUX detection volume. **b** showcases the conventional 3D MINFLUX pattern consisting of seven tophats (red, orange, yellow). Because of the interferometric deflection, the shapes change along z as indicated in the corresponding plots below. **c** shows the mixed beam pattern, VTH, where vortecis (purple) are used in the xy plane instead of tophats and, optionally, adds a vortex exposure at the origin; this variant is called 'VTH+'. The system is capable of switching between the two shapes through the EOM based path switch within 5 µs, i.e., the typical wait time between exposures. Hence, there is full flexibility during a MINFLUX cycle to change shapes arbitrarily for each exposure. Together with the polarisation rotator, one can address two pairs of beam shapes, so up to four in total, and –optionally– their interferometric intermediates for each of the two wavelengths. **d** pictures the dual colour excitation scheme used to track two emitters in 3D as presented in Section 4: the system switches between a mixed beam excitation pattern in the green and then in the red wavelength, exciting and localising a green and a red dye in rapid succession –"quasi-simultaneously".

The presented implementation is further extensible. For example, it could be scaled to have fewer or more than four paths and assign distinct wavelengths to each path. Of course, for every duplicate wavelength the beam combination induces a hefty power loss as the use of BSs is inevitable. Furthermore, just like the VTH and VTH+ schemes presented here decouple z and xy parts of the information, it is also possible to realise mixed beam schemes that further decouple x and y by employing half-moon beams that feature x- and y-oriented zero planes. This scheme is expected to improve the lateral localisation precision even further –similar to[48]. However, the use of half-moon beams requires to switch the polarisation right before the objective between linear vertical and linear horizontal for the two half-moons, and circular for the tophat. An

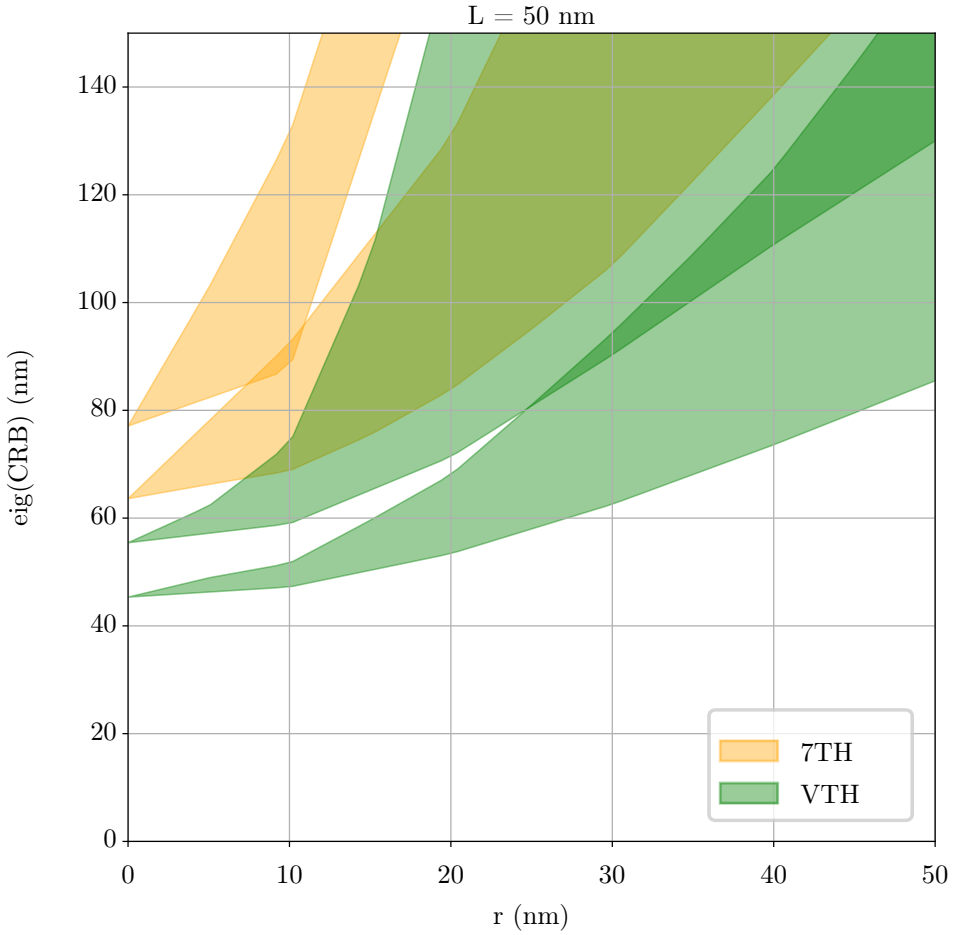


Figure 3.12.: CRB eigenvalues vs radial distance between emitter and the centre of the excitation pattern. The two bands are the maximum and the minimum eigenvalues of the 3×3 CRB (or covariance) matrix. The mixed beam scheme with four lateral vortecis and three axial tophats achieves lower (=better) values.

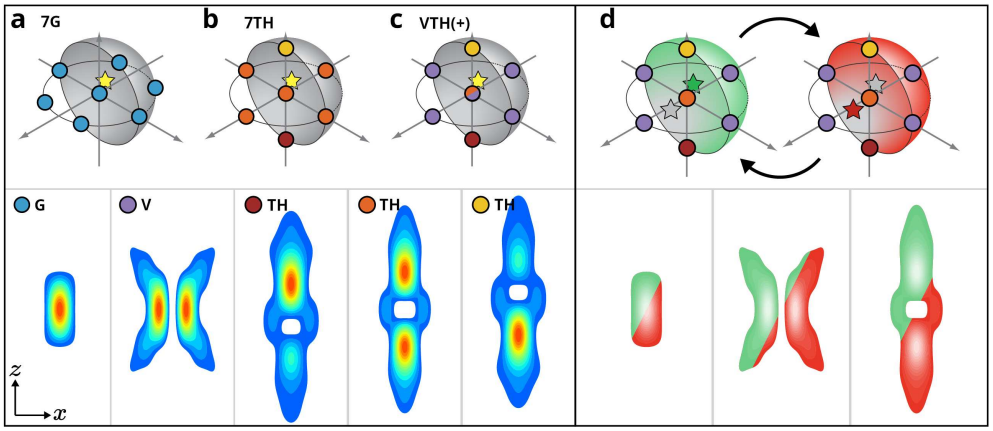


Figure 3.13.: Different excitation schemes enabled by the presented system. **a** 7G: seven Gaussian for a first rough 2D localisation over a larger area than a vortex scheme would allow. **b** 7TH: conventional 3D MINFLUX scheme consisting of only tophats. **c** VTH(+): mixed beam scheme for improved precision; vortecis and tophats are employed for the lateral and axial localisation, respectively. Optionally, a central vortex can be added (VTH+). **d** Rapidly switching between two excitation patterns with distinct wavelengths allows dual emitter tracking.

EOM replacing the last QWP could achieve this switching. However, we refrained from such an implementation to avoid extra complexity.

3.4. Polarisation and Coherence Considerations

3.4.1. Polarisation Effects

Every mirror in the system has an effect on the polarisation state. So far it was assumed, that only the relative phase between the s- and p-states, $\Delta\varphi_{\mathrm{sp}}$ of the reflected beam would be affected which can be easily compensated by the SLM as discussed in Section 2.5.4. In reality, however, the reflectivity is different for s- and p-states[84, 85] for angles of incidence other than 0°. Therefore, a beam with a mixed polarisation state will not only acquire some $\Delta\varphi_{\mathrm{sp}}$ but also a rotation[9] of its polarisation axis (the major axis of the ellipse describing the state). The consequence is that there might be no state of the polarisation rotator that translates into a clean input to the SLM double pass which is necessary to exclusively address one of the two passes. Another difficulty are slight height changes of the optical path: when two mirrors reflect beam from one plane parallel to the optics table to a higher parallel plane, the eigenstates of the mirrors are not the same anymore; the s- and p-states get scrambled such that a pure s-state at the input will be a mixtures of s- and p-states at the output[86]. The achromatic wave-plates A.QWP, A.HWP and B.HWP were not only adjusted by rotation but also by tip and tilt probably compensating these effects to some extent. However, it is impossible to perfectly compensate these polarisation changes for both colours this way because

[9]Rotations exclude "flips" by 90°. Increasing $\Delta\varphi_{\mathrm{sp}}$ from 0 to $\frac{1}{2}\tau$ of an originally linear polarisation at 45° without changing the relative amplitudes of the s- and p-states will do the following: first, the polarisation state becomes more and more elliptical with its polarisation axis 45° until it is circular; there is no clearly defined polarisation axis any more. After passing through the circular state, it will transition back from elliptical to linear, however, with its polarisation axis flipped to −45°.

the reflection coefficients are wavelength dependent; especially in the case of dielectric mirrors that exhibit a strong dependence[87].

An imperfect input state might be compensated when opting for the alternative implementation discussed in 2.6: if the polarisation rotator is situated after the double pass, the input beam is required to be circular or linear at 45°. A relative phase is easily compensated via the SLM as discussed. However, in this scenario the SLM grating height –that determines the diffraction efficiency into the first order– could be used to compensate the relative amplitude of the incoming s- and p-states. If this approach is feasible, it would also alleviate the wavelength dependence of the polarisation state mixing since every path can be adjusted individually. Nevertheless, the optical elements combining the four paths before the polarisation rotator will have the same opportunity to affect the polarisation state badly.

Besides the polarisation state mixing, the accumulation of relative phase might become problematic if it causes a delay that exceeds coherence length of the light source degrading the interference quality, see Section 3.4.2. In theory, it is possible to rectify this delay by flipping the polarisation states with a HWP at 45° positioned half-way such that the phase accumulated in the first half is compensated by the phase accumulated in the second half. In the presented system, this HWP could be placed, for example, after B.M5 and B.M6 approximately balancing the four mirrors in the first and five mirrors in the second half. This approach, however, neglects the influence of other elements, like AB.DC. Furthermore, the induced delays by the mirrors in units of waves are of order one which should not be a problem for the coherence length of the employed lasers as shown in the next section.

3.4.2. Coherence

The coherence length of a laser plays an important role when realising interferometric systems. When interfering a beam with itself, e.g., in a Michelson interferometer, and adding a time delay τ (resulting in a relative phase) between the two interferometer arms, the resulting intensity will go through maxima (constructive interference), $I_{\max}$, and minima (destructive interference), $I_{\min}$. The quality of the interference is measured by the visibility, $\gamma(\tau)$, defined by[88]

$$\gamma(\tau) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}},$$

where $I_{\max}$ and $I_{\min}$ are the extrema at delays closest to the argument τ. The coherence time, τ_{coh}, is the time for which $\gamma(\tau_{\text{coh}}) = 1/e$. The coherence length for a Lorentzian optical spectrum is given by[89, 88]:

$$l_{\text{coh}} = c\tau_{\text{coh}} = \frac{c}{\frac{1}{2}\tau\Delta\nu} \approx \frac{\lambda_0^2}{\frac{1}{2}\tau\Delta\lambda_0},$$

where c is the speed of light in vacuum, $\Delta\nu$ is the linewidth, i.e., the FWHM of the optical bandwidth, λ_0 is the peak wavelength and $\Delta\lambda_0$ is the spectral width in terms of wavelength.

The initial main red excitation laser was an Omicron QuixX® 642-140 (*Omicron, Rodgau-Dudenhofen, Germany*) with a measured $\Delta\lambda_0 = 1.5\,\text{nm}$ resulting in a coherence length of about 87 µm or 136 waves. While the axial deflection could be achieved, the zero degradation was quite pronounced. The laser was eventually exchanged for an old iBEAM-SMART-640 Ver.: G1 (Nov./2015) (*TOPTICA Photonics AG, Munich, Germany*) and all results in this work have been produced with this as the main red excitation laser. According to the manufacturer, this model[10] features has bandwidth of 0.11 nm resulting in coherence length of 1.19 mm or 1860 waves; more than 10× compared to the Omicron laser. The green excitation laser, a Cobolt

[10]Newer models of the same series unfortunately use different diodes: $\Delta\lambda_0 = 0.3\,\text{nm}$ to $0.5\,\text{nm}$ giving $l_{\text{coh}} = 0.26\,\text{mm}$ to $0.43\,\text{mm}$.

Jive™ 100-561 (*Cobolt AB, Solna, Sweden*), features a spectral width of $\Delta\nu < 500\,\mathrm{kHz}$ and hence a coherence length of 191 m.

3.5. Parasitic Mach-Zehnder Interferometer

The polarisation Extinction Ratio (ER) of the path switch is important. If it is insufficient the inactive path will carry some residual power. Since the laser source is the same for both paths, they are coherent with respect to each other. The result is a 'parasitic Mach-Zehnder Interferometer' (pMZI) that is formed between the path switch at the first PBS and the BS of the beam combination. The two arms encompass the whole path length in between; including the fibres and the entire beam shaping module.

Imagine the VTH mixed beam scheme which set up such that the second pass of 'path A' is used to generate a vortex beam, while 'path B' displays the phase masks for a Gaussian and a tophat on the first and second pass, respectively. Path A is active but path B still carries some power due to an insufficient ER. The residual power will correspond to a Gaussian or tophat shape or any mixture inbetween depending on what the voltage of path B's polarisation rotator is. In both cases, this shape will interfere at the beam combination BS (BCR/G.BS)9 with the active path A. Consulting Figure 2.4 reveals that this will either cause path A's vortex' zero to be shifted laterally (interference with Gaussian) or tilted axially (interference with tophat). The residual power, P_r, might be small but the underlying electric field strength scales adversely with the square root. Consider the case of a 0.1 % residual power:

$$\begin{aligned}\frac{P_r}{P_0} &= \frac{\int |E_r|^2(\boldsymbol{r})\mathrm{dA}}{\int |E_0|^2(\boldsymbol{r})\mathrm{dA}} = 0.001\\ \frac{\int |E_r|^2(\boldsymbol{r})\mathrm{dA}}{\int |E_0|^2(\boldsymbol{r})\mathrm{dA}} &= \frac{|E_r|^2}{|E_0|^2}\frac{\int g(\boldsymbol{r})\mathrm{dA}}{\int g(\boldsymbol{r})\mathrm{dA}}\\ \Rightarrow \frac{|E_r|}{|E_0|} &= \sqrt{0.001} = 0.032\,,\end{aligned}$$

where it is assumed that both beams only differ in their phase profile. The residual field strength which is the significant quantity in the interference is 3.2 %. A static shift or tilt of the zero could be tolerated as the pattern origin is subject to frequent calibrations. However, the pMZI arms are not stable with respect to their relative phase: the polarisation maintaining fibres induce a temporal dependence on the phase resulting in the rotation of the vortex minimum around its actual origin at a radius determined by the residual power. In the case of a residual tophat, the zero tilt axis rotates making the radius of the rotation dependent on z. The relative phase and as such the rotation angle changes randomly; the motion goes back and forth and completing a full circle is accomplished within minutes. As a typical MINFLUX cycle takes less than a millisecond, all the vortex exposures during a cycle are affected by the same shift/tilt zero as if the origin of the pattern moved. MINFLUX localises with respect to where this origin is located and hence the localisations of a fixed emitter will slowly draw out a circle (instead of a blob).

Figure 3.14 shows the rotation radius for the vortex plus a residual Gaussian as a function of η (Left) and the vortex plus a residual tophat as a function of z and η (Right) calculated from PSF simulations. For $\eta = 1 \times 10^{-5}$ the rotation radius is already 50 nm (a diameter of 100 nm). The residual power that was measured before improving the path switching (see Section 3.5.1) would typically result in $\eta = 6 \times 10^{-4}$ and rotation diameter of about 5 nm (vortex plus residual Gaussian). Thus, the simulation does not match the observations by several orders of magnitude. It could be that only part of the residual light is coherent thus reducing the deflection effect (the incoherent part would worsen the zero quality).

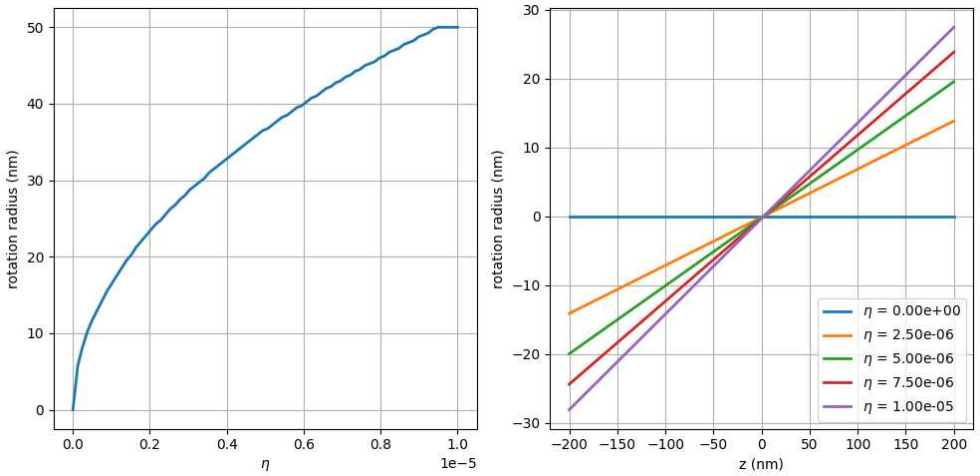


Figure 3.14.: Simulated rotation radii for different residual power factors η. Left: vortex plus a residual Gaussian. Right: vortex plus a residual tophat; the rotation radius depends linearly on z due to the induced tilt of the zero.

The pMZI went unnoticed for quite some time. The initial iteration of the system had a single polarisation maintaining fibre connecting the laser source module to the rest of the system with the path switch coming after the fibre. In that setup, the pMZI arms did not include any fibre. Therefore, only a static/shift was affecting the vortex and localising a static emitter over long periods of time would still yield a single, but slightly offset, blob.

Only after a redesign that moved the path switch onto the laser source module introducing two fibres (one per path), the dynamic behaviour appeared and was noticed. While not ideal, the effects on measurements can be small due to the slow dynamics. A typical MINFLUX iteration extracting 10,000 photons in STORM mode takes 200 ms on average; depending on the processing these photons are split into multiple localisations with fewer photons. Hence, the repeated localisation of single molecule until it bleaches is not affected. Localisations of distinct molecules close to each other that occur within a short time frame should not see their relative distances distorted by the rotating zero. Similarly, a typical single molecule tracking measurement lasts less than a minute.

To minimise the effects we employed a temporary solution to partially mitigate the problem by setting up the system in such a way that instead of a residual Gaussian, the vortex was interfering with a residual tophat meaning that it was slightly tilted along a slowly rotating axis. The zero location at the focal plane is, however, static no matter the tilt. By carefully focusing or refocusing during a measurement the rotation diameter could be reduced to a minimum. The data presented in Section 4, except the averaged transitions of the fast axial oscillator, was acquired under this condition.

3.5.1. AOTF Assisted Path Switching

The main path switch consists of an EOM and two PBSs. The EOM is used to flip the vertical input polarisation between a vertical or horizontal output polarisation. The first PBS either transmits or reflects the beam. The second PBS, cleaned up the polarisation in the reflected path: the employed PBSs are specified to have an ER of $< 1:1000$ for the transmitted light but only around $1:100$[90] for the reflected light. To counteract this contamination, the second PBS transmitting the reflected output of the first PBS was added.

The ER of this configuration turns out to be insufficient. Attempts to improve the ER by, e.g., aligning precisely through the centre of the PBS did not help. Furthermore, imperfections of the EOM output state and small misalignments of the polarisation axes with respect to the PBS will always cause some of the light to be present in the inactive path. The naive solution is to add two power modulators –one per path– such as AOTFs or acousto-optic modulators. The drawback is the increased cost and added complexity of additional devices. We could, however, devise a way to improve the ER with a single AOTF by re-arranging the path switching module.

An AOTF usually operates on a single incident path carrying multiple wavelengths. Each wavelength is assigned a channel which drives the crystal inside the AOTF with a frequency and amplitude that only allow a narrow spectral band, within which the corresponding wavelength falls, to be diffracted by the acoustic waves inside the crystal. The output is the first diffraction order that exits the device at an angle (the zeroth order is blocked). The salient feature of the AOTF is that the first order is collinear for all wavelengths and that it can modulate its power within microseconds.

In the improved scheme, the AOTF follows the path switching and both paths go through the AOTF: one at normal incidence, the other at an angle. The angled incidence shifts the required frequency making it possible to independently address[11] and diffract only one of the two paths despite them having the same wavelength. Together with the path switch the ER was high enough that the power of the switched off path was below the noise level of the power meter (PM).

[11] This requires an input channel per path and not just per wavelength. Here: four channels.

4. 3D MINFLUX Imaging and Tracking

In order to demonstrate that the interferometric deflection technology is practical for 3D MINFLUX we did proof-of-concept measurements: this section showcases 3D imaging and tracking. For imaging, flat DNA-origami and nuclear pore complexes in fixed cells are presented. Furthermore, 3D tracking of dynamic DNA-origamis is shown, including the first demonstration of dual emitter tracking in 3D. The data processing follows previous works[40, 38] but the software was re-implemented in Python from scratch.

4.1. System Setup

The weekly maintenance routine is as follows: slight alignment adjustments before the SLM double pass are only necessary if the Gaussian intensity maximum is not well centred on the SLM pupil (this can be checked by observing the back focal plane on STG.CAM). The alignment after the double pass makes use of phase masks with centred horizontal and vertical $\frac{1}{2}\tau$ phase jumps imprinting a zero-intensity cross pattern onto the outgoing beam. Custom made magnetic plates with centred pinholes of 1 mm and 1.5 mm diameter are used to probe and adjust the alignment through the telescopes that are all mounted on cage systems: the plates snap on to custom magnetic cage plates centring the pinhole with the cage system. Then, the centre of the cross pattern of the beam must propagate through the pinhole; again STG.CAM is used for this purpose except for the alignment through STG.SL and STG.TL (where it is done by eye). To overlap the four paths, a second camera is positioned in one of the focal planes of the telescopes after the double pass. For that, a BS, mounted on a construction that clips into the top two rods of a cage system diverts half the light to be focused on the extra camera and transmits the other half which is then monitored on STG.CAM. Having two images in conjugate planes allows to match the offset and angle of the four paths and thus making them collinear. The position of the excitation beam is marked on the WF camera (WF.CAM) by moving a fluorescent bead with stage through its focus (Gaussian shape) and maximising the signal intensity. The detection and activation (if necessary) volumes are overlapped with the excitation via a 2D scan of the TM of a bead with the WF or activation laser on and matching their PSF's position via the motorised mirrors (D.M1 and AC.M4) with the excitation beam centre. For axial overlap, AC.L2 is moved for the activation or the SLM defocus terms are adjusted for matching the excitation with the detection volume. To compensate for aberrations, a pupil segmentation process is employed[91]: it measures the phase difference of a point within the SLM pupil with respect to its centre by only displaying two gratings in small circular regions at the centre and at one of the said points. The resulting PSF is a sinusoidal interference pattern captured via a 2D bead scan. The relative phase between the gratings is changed in four steps and the resulting offsets in the sinusoidal pattern reveal the relative phase. This procedure is repeated for sufficiently many points across the SLM pupil and a smooth phase landscape is interpolated from the acquired phases. Then, the interferometer was tuned as described in Section 3.2.

In order to further tune the vortex and the tophat beams and to choose the L parameter, convergence[1] is tested on a single molecule sample (Section G.4). For this, the VTH+ scheme is used (irrespective of which scheme employed during measurement): it allows to extract information

[1] Convergence of the final excitation pattern onto the emitter positions after going through several iterations.

about the central p-functions for xy, $p_0^{(xy)}$, and z, $p_0^{(z)}$, separately by using the corresponding ratios of photon counts of the five vortecis and the three tophats, respectively. The index 0 is reserved for p-functions and photon counts related to a central exposure, so $\hat{p}_0 = n_0 / \sum_j n_j$, where n_0 is the photon count generated by the central vortex ($\hat{p}_0^{(xy)}$) or central tophat ($\hat{p}_0^{(z)}$) exposure. The central p-functions can yield valuable information: in the ideal case with a perfectly centred[2] excitation pattern, they become zero. However, due to background and imperfect intensity zeros they will exhibit non-zero values in practice making these quantities tangible. Ultimately, the system's performance can be tuned by minimising $p_0^{(xy)}$ and $p_0^{(z)}$. The two p-functions can be monitored live while localising a static molecule repeatedly giving immediate feedback on parameter changes.

It can be shown that the zero quality of, e.g., a vortex or tophat beam, is absorbed by the effective SBR notated as SBR*:

$$
\begin{aligned}
p_i =& \frac{\lambda_i}{\sum_j \lambda_j} = \\
=& \frac{\lambda^B + \lambda^Z + \lambda_i^S}{\sum_j \lambda^B + \lambda^Z + \lambda_j^S} = \\
=& \frac{\lambda^B + \lambda^Z + \lambda_i^S}{K\lambda^B + K\lambda^Z + \sum_j \lambda_j^S} = \\
\overset{1/\sum_j \lambda_j^S}{=}& \frac{\frac{1}{K}\left(\overbrace{\frac{K\lambda^B}{\sum_j \lambda_j^S}}^{:=1/\mathrm{SBR}} + \overbrace{\frac{K\lambda^Z}{\sum_j \lambda_j^S}}^{:=1/\mathrm{SZR}}\right) + \overbrace{\frac{\lambda_i^S}{\sum_j \lambda_j^S}}^{:=p_i^{(0)}}}{\frac{K\lambda^B + K\lambda^Z}{\sum_j \lambda_j^S} + 1} = \\
=& \frac{\frac{1}{K}\overbrace{(1/\mathrm{SBR} + 1/\mathrm{SZR})}^{=:1/\mathrm{SBR}^*} + p_i^{(0)}}{1/\mathrm{SBR} + \mathrm{ZSR} + 1} = \\
=& \frac{p_i^{(0)}\mathrm{SBR}^*}{1 + \mathrm{SBR}^*} + \frac{1}{K}\frac{1}{1 + \mathrm{SBR}^*},
\end{aligned}
$$

where λ^S, λ^B and λ^Z are the mean photon counts coming from the signal, the background and an imperfect zero quality, respectively, and K is number of exposures. The resulting equation has the same form as Equation (1.12)[3] with SBR being replaced by

$$
\begin{aligned}
\mathrm{SBR}^* =& (1/\mathrm{SBR} + 1/\mathrm{SZR})^{-1} = \\
=& \frac{\mathrm{SBRSZR}}{\mathrm{SBR} + \mathrm{SZR}} = \\
\overset{1/SZR}{=}& \frac{\mathrm{SBR}}{1 + \underbrace{\mathrm{SBR}/\mathrm{SZR}}_{:=ZBR}} = \\
\mathrm{SBR}^* =& \frac{\mathrm{SBR}}{1 + \mathrm{ZBR}}\,. \qquad (4.1)
\end{aligned}
$$

[2] Centred on the emitter.

[3] The derivation is identical when $\lambda^Z = 0$.

Hence, assuming a well centred excitation pattern,

$$p_0 \approx \frac{1}{K}\frac{1}{1+\mathrm{SBR}^*}\,. \tag{4.2}$$

Consequently, with Equation (4.1), the value of p_0 decreases (=improves) for a lower ZBR and therefore it can be utilised to improve the zero quality. While the system localises a static molecule repeatedly, $p_0^{(xy)}$ and $p_0^{(z)}$ are monitored live that are estimated from the photon counts corresponding to the five ($K = 5$) or three ($K = 3$) vortex or tophat exposures of the VTH+ scheme, respectively. For the tophat, p_0^z is minimised by adjusting the inner diameter of the $\frac{1}{2}\tau$ phase step[4]. For the vortex, it is p_0^{xy} which is minimised by introducing a slight astigmatism term (amplitude and angle). Afterwards, the ZBR can be measured by acquiring a series of 2D (vortex) or 3D (tophat) PSF scans in a small (100 nm side length) region around the zero on a single molecule until it is bleached. The frames with the unbleached molecule can then be fitted to quadratic function to estimate the zero level which is at $m_1 = Z + B$. The frames after bleaching can be used to determine the background level[5] $m_2 = B$. The ZBR is then $\mathrm{ZBR} = \frac{m_1 - m_2}{m_2}$. Typical values were between 1.5 to 2.5.

As discussed in Section 1.2, the SBR will decrease for smaller L. However, the CRB can still improve even for the limit $L \to 0$, see Appendix F. L was decreased to aim for the somewhat arbitrary threshold of $\mathrm{SBR}^*(L) \approx 1$. Then, according to Equation (4.2), the target values are $p_0^{(z)} \approx 0.167$ and $p_0^{(xy)} \approx 0.100$. As the p-functions were monitored in single molecule samples, we aimed for slightly better (=lower) values to have an extra margin for "real" samples where the background is expected to be higher (e.g., fixed cells). Typical values were $L_{xy} = 50\,\mathrm{nm}$ and $L_z = 75\,\mathrm{nm}$ hinting at a worse SBR for the tophat. Note, that for very small values, e.g., $L < 50\,\mathrm{nm}$, additional iterations with intermediate Ls might be necessary to guarantee proper centring.

4.2. 3D Imaging

For a first 3D imaging demonstration we designed a flat DNA origami arranging twelve Alexa Fluor 647 dyes in a regular 3×4 grid with 20 nm spacing. Figure 4.1 **a** shows an acquired image of the structure in 2D histograms: each localisations was estimated from 2000 photons, filtered and clustered, see Appendix I. Even though two sites are missing[6], the grid is clearly resolved. While the origami is designed to be flat, the 3D imaging reveals its actual structure as can be seen by the slight bending across the xz plane shown in **b**. The upper and lower panels of **c** quantify the localisation precisions of each site along the central row with coordinate in x' (as indicated in **a**) and z, respectively. Gaussian fits are shown as black envelopes of the orange localisation histogram. Their Full Width at Half Maximum (FWHM) is indicated by the annotated inwards pointing arrows. They lie in the range of 2.52 nm to 2.94 nm in x' and 2.90 nm to 4.33 nm in z which translates to a standard deviation of $\sigma_{x'} = 1.07\,\mathrm{nm}$ to 1.25 nm and $\sigma_z = 1.23\,\mathrm{nm}$ to 1.84 nm. The discrepancy between localisation precision in xy and z can be explained by the employed excitation pattern: the L_{xy} and L_z parameters were independently adjusted as described in Section 4.1 to yield $\mathrm{SBR}_{xy} \approx \mathrm{SBR}_z \approx 1$. The procedure yielded $L_{xy} = 50\,\mathrm{nm}$ and $L_z = 75\,\mathrm{nm}$, theoretically resulting in a 50 % worse precision in z. The red curves in the panels below show the Gaussian distributions of the uncertainty of the mean which

[4] It is the parameter with the strongest influence on the tophat's zero quality.

[5] The beam power in the frames that are used to determine the zero and the background level must be the same. Sometimes it useful to increase the power to induce bleaching faster and one should be careful to put it back to its original value after bleaching, or, if the molecule stayed on for a couple of frames despite the higher power, leave the power as it is and select the frames accordingly.

[6] A common problem due to labelling efficiency.

scales inversely proportional with the square root of the number of localisations. Again, the FWHM is annotated and lies below a single nanometre demonstrating that each site is localised within the Ångström regime. Indeed, the designed spacing of 20 nm is well reproduced in the data as shown by the annotated arrows at the top of **c**.

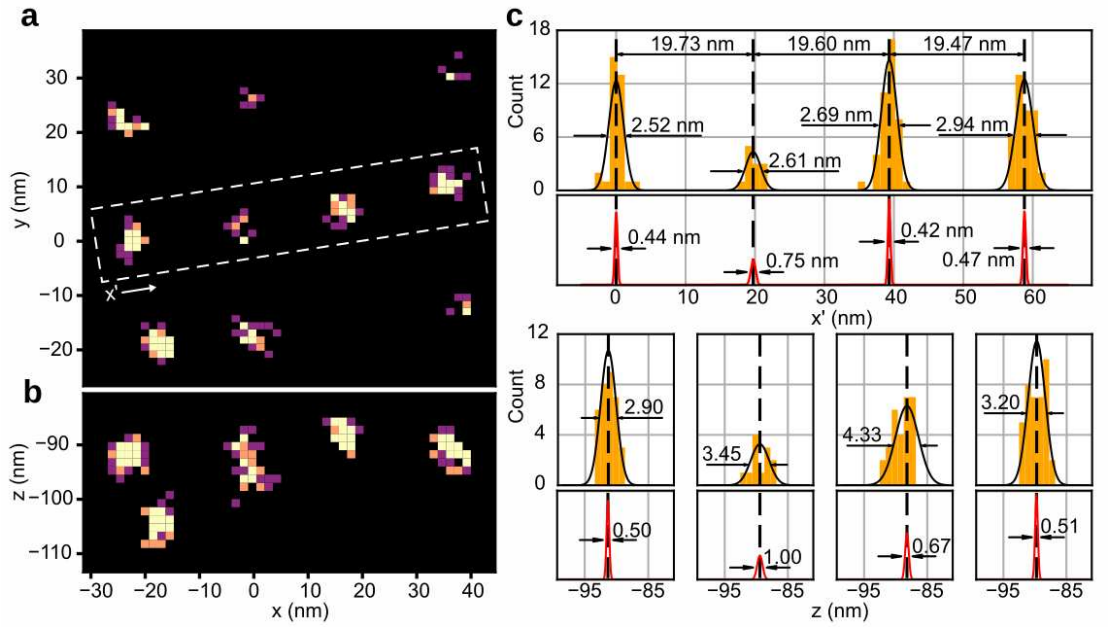


Figure 4.1.: 3D imaging of DNA origami arranging 12 Alexa Fluor 647 dyes in a 3×4 grid. Orange histograms visualise the spread of localisations. The annotations of inwards pointing arrows report the FWHM of the Gaussian fits (in nm). The red Gaussians represent the uncertainty of the mean.

To prove the technique in a more biologically relevant context we imaged SNAP-labelled Nup96 in fixed U2OS cells. A large Field Of View (FOV) of several microns is shown in Figure 4.2 as well as two close up of Pores from a different measurement. The localisation precision of the large FOV is quantified in Figure 4.3: after filtering all localisations are clustered and all clusters' standard deviation in x, y and z are plotted in the histograms.

4.3. 3D Tracking

4.3.1. Axial Oscillator

To test the 3D tracking capabilities of our system, we devised a dynamic DNA-origami in shape of a 50 nm tall vertical pillar with a single ATTO 647N dye molecule attached to a flexible arm, see Figure 4.4 **a**. The arm is anchored on the side of the pillar and oscillates between two catching sites 14 nm apart; one above and one below. 16 biotin-modified DNA staple strands ensure it stands upright on a streptavidin-functionalised surface. In **b**, one sees the histogram of all localisations calculated from all photons within consecutive 10 ms time windows –the so-called 'binning'– resulting in roughly 1500 photons per localisation. Panel **c** shows time traces with 1 ms temporal resolution with precisions of $\sigma_x = 4.9$ nm, $\sigma_y = 4.7$ nm and $\sigma_z = 4.1$ nm. To compute these standard deviations, the traces were low pass filtered and classified through k-means clustering; the resulting distribution and Gaussian fits are shown in purple on the right.

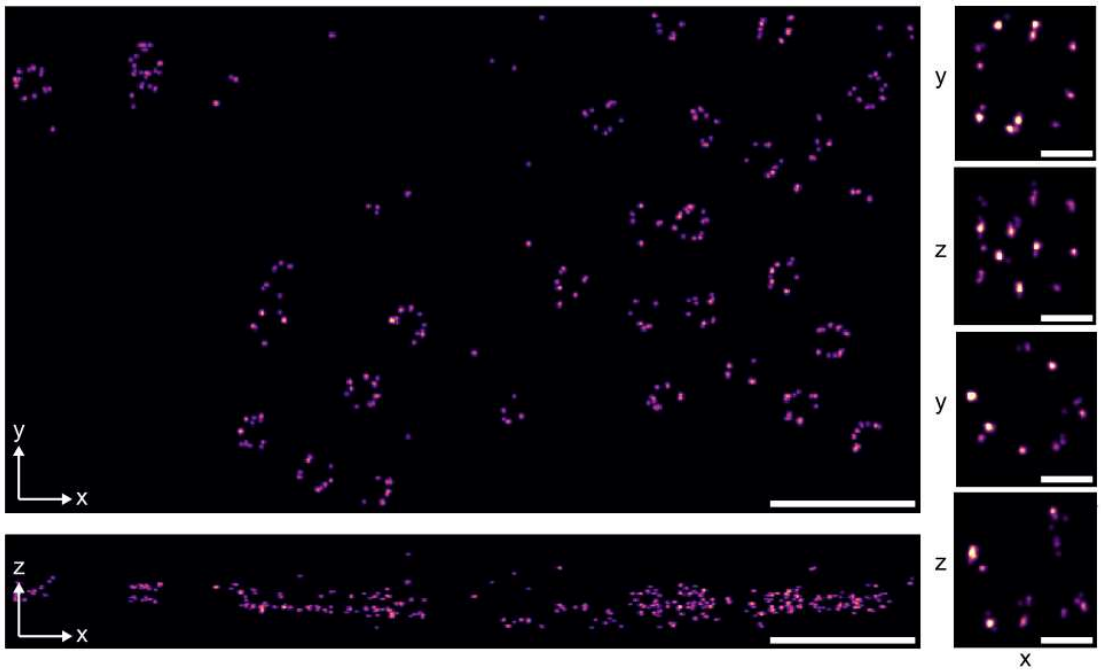


Figure 4.2.: 3D imaging of SNAP-tagged Nup96 labelled with Alexa Fluor 647 in fixed U2OS cells. Scale bars: 500 nm for large FOV and 50 nm for single pores. 3D imaging of SNAP-tagged Nup96 labelled with Alexa Fluor 647 in fixed U2OS cells. Scale bars: 500 nm for large FOV and 50 nm for single pores.

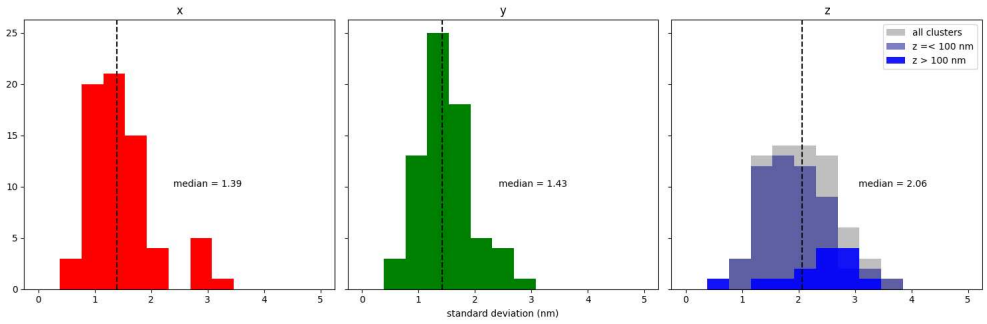


Figure 4.3.: Histograms of the standard deviations for x, y and z across clusters that feature more than five localisations.

The measured axial distance is (14.9 ± 1.7) nm after grouping the clustered traces into groups of 20 localisations and calculating their distance.

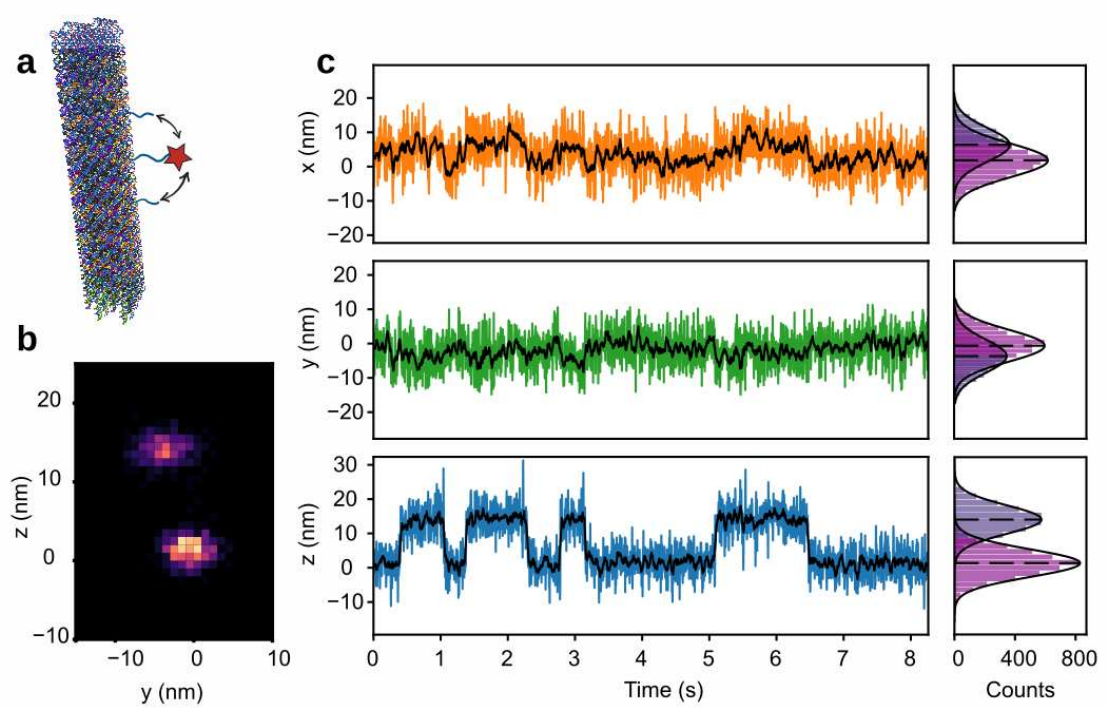


Figure 4.4.: 3D tracking of an ATTO647N dye attached to a flexible arm on the side of a vertical DNA origami pillar. **a** 3D rendering of the structure. **b** Image of the two catching sites in xz (10 ms temporal resolution). **c** Localisation traces at 1 ms temporal resolution. The purple distributions show Gaussian fits of the two states.

4.3.2. Fast Axial Oscillator and Transition Averaging

To exhaust the speed offered by our interferometric axial deflection strategy, a modified version of the axial oscillator oriagmi was tracked in z only: one base pair of the catching sites was removed such that the binding time was reduced and hence transition rate increased. Therefore, many more transition traces could be extracted, aligned and averaged to see if the dye molecule can be caught moving between sites. The system was operated at maximum power to receive the highest possible photon count rate. The last iteration consisted of two axially offset tophat exposures, $L_z = 100$ nm, allowing a cycle time of 42 µs or a localisation frequency of 23.81 kHz. One can argue about the applicability of such a scenario but combined with the preceding centring it delivers information in xy and might be interesting, e.g., for tracking cargo entering or exiting membranes. The trace considered was given by $\Delta\hat{p}(t) = \hat{p_1}(t) - \hat{p_0}(t) = (n_1(t) - n_0(t))/(n_0(t) + n_1(t))$ and was classified to be either one of two states (up and down) for each time point. The classification was done using the python package `pomegranate v0.14.8`: a hidden Markov model was fitted with the `baum-welch` algorithm and the classification into the two states was done through the `predict()` method using the `map` algorithm. The resulting classification introduces artefacts as it fits signatures in the noise that make a transition more apparent but are not real. The result is a slight dip downwards before the transition and an overshoot right after. Figures H.1, H.2 and H.3 compare different algorithms and their effects

on the averaging of simulated data. The resulting traces were filtered to contain at least five points before and after the transition. All transitions were aligned in time (horizontally) such that the last points before the each transition overlap. Furthermore, the averaged level before the transition of each trace was taken as a vertical offset to deal with residual drift in z. To go from $\Delta\hat{p}$ to z, the analytic MLE was employed, $\hat{z}$, which is obtained by solving

$$\frac{\partial \log(\mathcal{L})}{\partial z} = \frac{n_0}{p_0}\frac{\partial p_0}{\partial z} + \frac{n_1}{p_1}\frac{\partial p_1}{\partial z} \stackrel{!}{=} 0\,,$$

for z, where the p_i are given by Equation (1.12). The solution is

$$\hat{z} = \frac{L}{2}\frac{1+\mathrm{SBR}}{\mathrm{SBR}}\frac{\Delta p}{1+\sqrt{1-\left(\frac{1+\mathrm{SBR}}{\mathrm{SBR}}\Delta p\right)^2}}\,.$$

The estimator requires a known SBR. With only two exposures and hence one independent p-function (see Section 1.2) there is not enough information to estimate the SBR as well. However, the SBR was extracted roughly by recording the bleaching step and measuring the ZBR of approximately two.

Figure 4.5 shows the result in the upper plot with the raw traces in grey and the averaged trace in red. The x-axis marks the time before and after the transition. The band indicates the $\pm 3\sigma$ uncertainty of the mean calculated by dividing the standard deviation of the samples (for any given time point) by the square-root of the number of averaged traces. The number of the averaged traces is shown in the lower plot: it has a maximum around $t = 0$ since all traces contain at least five points before and after the transition but traces with long "tails" are increasingly less common. Besides the remaining dip and overshoot that are clearly artefacts, there is no visible transient. It could be the that the classification somehow hides the transient but more likely the time resolution of 42 µs is not enough catch the molecule while it is unanchored. Nonetheless, this sets a threshold to the expectable speed of this transition.

4.3.3. Dual Emitter Tracking

To make full use of the four channel system we designed and measured a dual emitter origami for 3D tracking shown in Figure 4.6 **a**: similarly to the axial oscillator, a Janelia Fluor 549 and a ATTO 647N dye molecule are attached on arms oscillating between two catching sites each. We quasi-simultaneously tracked both molecules in 3D by switching between a green and a red mixed beam excitation pattern every 255.2 µs; utilising six of the eight regions of our SLM (Figure 3.2). The binning was adjusted to yield about 1200 photons per localisation resulting in temporal resolutions of 20 ms and 25 ms for the ATTO 647N and the Janelia Fluor 549 molecules, respectively[7]. Spatial resolutions, as shown in Figure 4.6 **b**, are $\sigma_x = 2.2\,\mathrm{nm}$, $\sigma_y 2.3\,\mathrm{nm}$ and $\sigma_x = 1.8\,\mathrm{nm}$ for ATTO 647N, and $\sigma_x = 1.8\,\mathrm{nm}$, $\sigma_y = 1.7\,\mathrm{nm}$ and $\sigma_y = 1.6\,\mathrm{nm}$ for Janelia Fluor 549. The localisation traces are shown in **c**. One can clearly distinguish the two states for each trace. While this is the first demonstration of dual emitter tracking in 3D[8], the data has some shortcomings: from the xy view it appears that two binding sites overlap and that the two binding site pairs are in two different z-planes which does not conform to the design. As seen in Figure 4.1, DNA origamis are not always flat and exhibit a 3D structure which could partly play a role here. However, other origamis exhibits similar discrepancies and a more probable cause are imprecise calibrations[9] and/or an imperfect beam overlap of the red and green excitation patterns.

[7] The difference stems from the different brightnesses of the molecules.
[8] Dual emitter tracking in 2D has been achieved based on lifetime information recently [50].
[9] The deflection through the TM varies with the wavelength. This difference needs to be compensated by the EODs.

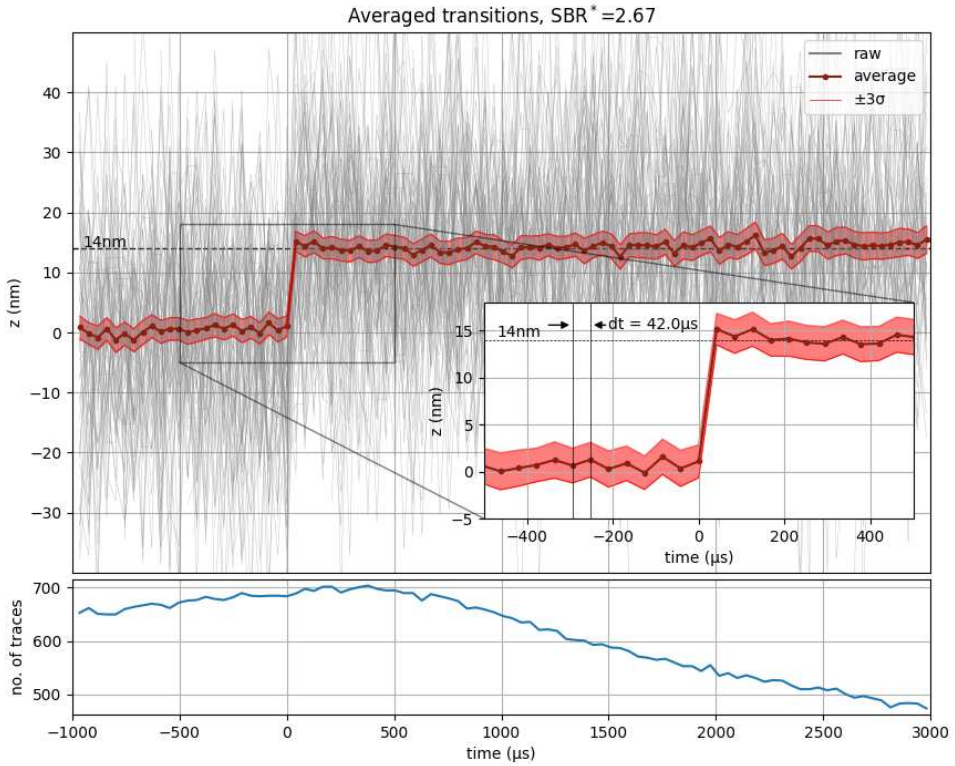


Figure 4.5.: Averaged and aligned transitions of the fast axial oscillator. The upper panel shows every tenth raw trace in gray and the average of all traces in red. The band indicates the $\pm 3\sigma$ uncertainty. The averaged transition jumps from one point to the next to the upper state; there is no transient visible hinting at a transition speed significantly below 42 µs. The lower panel shows the number of averaged traces for every time point.

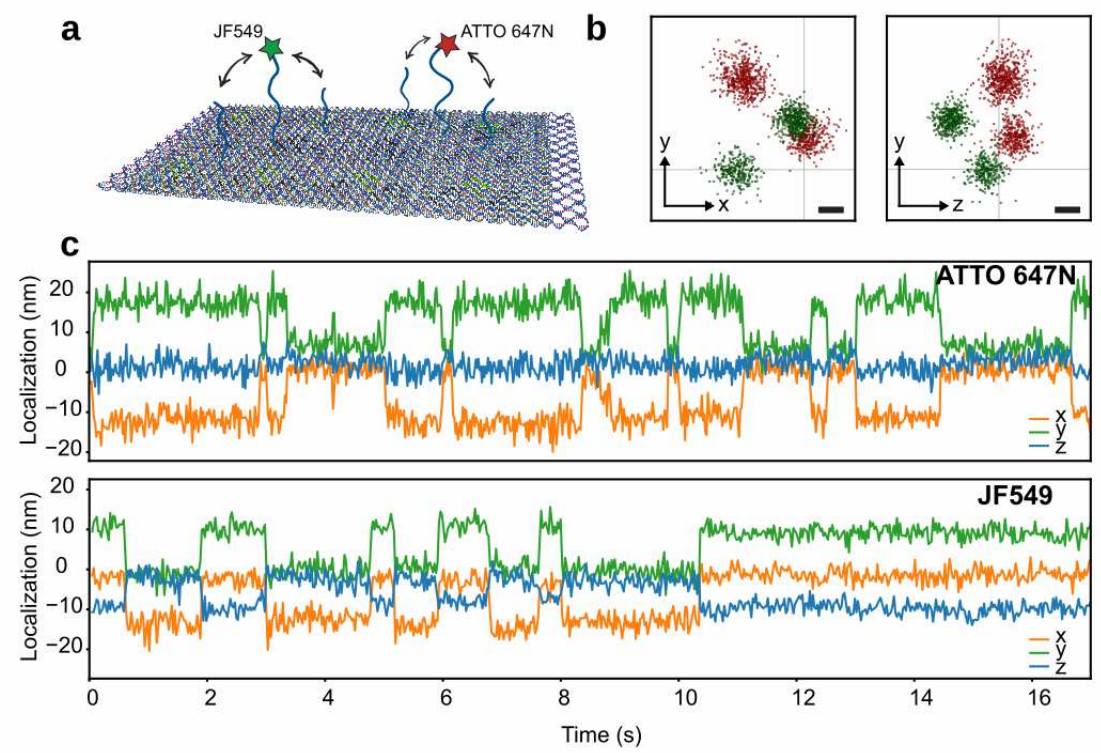


Figure 4.6.: First demonstration of 3D MINFLUX dual emitter tracking. **a** 3D render of the dynamic structure: a Janelia Fluor 549 and a ATTO 647N dye molecule are conjugated to single stranded DNA arms that can oscillate between two catching sites. **b** Images of the four sites in xy and in yz. **c** Localisation traces for both molecules at 20 ms and 25 ms temporal resolution. Scale bar: 5 nm.

5. Conclusion

I presented an interferometric approach to manipulate beam shapes that allows fast axial scanning of the intensity zero of a tophat beam enabling 3D MINFLUX operation. I implemented and demonstrated its capabilities in the context of 3D MINFLUX imaging and tracking experiments on a custom MINFLUX system that was built from scratch.

The interferometric approach relies on the superposition of electric fields of two different beam shapes. Here, the focus was on beams featuring intensity zeros and how these zeros can be interferometrically manipulated because of their importance in MINFLUX. Especially for 3D MINFLUX, it allows an alternative way for axial scanning that, so far, could only be realised with varifocal lenses or deformable mirrors; complex devices that are inferior in terms of deflection speed.

The case of the axially deflected tophat zero can intuitively be explained by a first order approximation of the electric fields around focus: the axial zero crossing of the tophat is approximated by a linear function of z while the coinciding maximum of the Gaussian is approximately constant. A linear plus a constant function is a linear function with a shifted zero crossing. Thus, the intensity zero (located at the zero crossing) of the tophat will move axially when interfered with a Gaussian depending on their relative amplitude. Several other beam combinations were studied that resulted in deflected and tilted zero lines and planes, offering a range of possibilities. In this work, however, only the axial deflection was employed for an actual application; 3D MINFLUX.

The interferometric strategy has limitations. Instead of defocusing the entire tophat beam, it continuously deforms the beam shape to move the zero, limiting the deflection range to within the extent of the tophat's point spread function. Furthermore, the deflected zero loses in quality, i.e, the theoretical intensity null becomes an intensity minimum. The intensity of this minimum increases with the deflection distance due to a mismatch of Gouy phases of the Gaussian and the tophat. This causes a loss of localisation precision: For example, the localisation in z of a molecule at the origin of the excitation pattern, which is 200 nm (140 nm in aqueous samples) above or below the focal plane, exhibits a Cramér-Rao bound that is increased by 20 % when compared to the same scenario with defocused tophats.

The implementation of the beam shaper is based on polarisation interferometry and has three building blocks: a polarisation rotator, consisting of an electro-optic modulator and an optional quarter-wave plate, that controls the relative amplitude between the s- and the p-state, a spatial light modulator (SLM) double pass configuration that patterns each polarisation independently, and, finally, a polarisation interferometer, consisting of a half-wave plate and a polarising beam splitter, which mixes the patterned s- and p-states and projects their superposition into a new s-polarised state. This arrangement has several advantages: it is a common path interferometer and thus requires no active stabilisation of interferometer arms. The driving element is an electro-optic modulator which operates as fast as the connected electronics allow. The interference does not occur in the sample and results in a single vertically polarised beam that is convenient to handle throughout the rest of the system. Furthermore, it is scaleable: the presented MINFLUX system features four optical paths; two carrying red and the other two carrying green light. This configuration provides four beam shape combinations for interferometric shaping and can thus address a total of eight distinct beam shapes. Other arrangements with fewer or more paths and wavelengths are conceivable.

Alternative implementations are possible as well. For instance, birefringent phase masks

made of glass are commercially available and could achieve the same effect as the SLM at lower cost. Contrary to the SLM, it cannot be calibrated for multiple wavelengths and incorporate aberration corrections. Furthermore, the use of diffraction gratings displayed on the SLM makes the SLM-based setup robust against imperfections of the wave plates and relative phases as is shown via a Jones model: unpatterned light, that would pollute the interference, and crosstalk between polarisation states ends up in the zeroth or second diffraction order, which are blocked.

The interferometric beam shaper was implemented and an automated setup procedure for tuning the interference was devised. Its ability to realise the zero manipulations studied in simulations was confirmed through point-spread function measurements of the excitation beam via backscattering on gold particles. The axial deflection of the tophat zero was probed by acquiring intensity profiles along z for several deflection offsets on a single molecule: the profiles stay parabolic around their minima across the measured range of 600 nm. While the zero degradation is clearly visible, it showcases a flat profile over ± 200 nm. Similar profiles with slightly worse zero degradation were acquired for a pulsed excitation beam demonstrating the feasibility of the technique for future applications involving lifetime information. Furthermore, the deflection transient was probed, revealing a < 5 µs transition time.

A peculiarity of the initial implementation was uncovered: a parasitic Mach-Zehnder interferometer. Its two arms stretch from the polarising beam splitter used for path switching to the regular beam splitter(s) that combine the two paths of the same wavelength. Even a small residual power of 0.06 % on the path that is supposed to be turned off will have strong interferometric effects on the beam shape of the active path and might effect the zero quality. This becomes even more problematic as soon as the relative phase between the arms is not stable resulting in intensity zeros that are not only deflected by a couple of nms but also rotate on the order of minutes. It is hence essential when scaling the beam shaper to several optical paths to ensure perfect path switching. Here, an acousto-optic tunable filter acting on both paths that are transmitted at different angles, was employed to completely turn the inactive path off.

3D Minflux imaging and tracking measurements were performed to prove the interferometric beam shaper's feasibility in real applications. The achieved localisation precisions match Minflux' state of the art while the 3D tracking speeds are unmatched by current Minflux systems. Furthermore, dual emitter tracking in 3D was demonstrated for the first time.

For 3D imaging, DNA-origamis positioning fluorophores in a regular 3×4 grid with 20 nm spacings were measured. The sites were clearly resolved with a localisation precision of < 1.3 nm in xy and < 1.9 nm in z. The average position of all available localisations belonging to a site resulted in uncertainties of 0.42 nm to 0.75 nm in xy and 0.50 nm to 1.00 nm in z. Furthermore, NUP96-SNAP labelled with Alexa Fluor 647 in fixed U-2OS cells were imaged: over a large field-of-view the achieved median localisation precision was ≤ 1.43 nm in xy and 2.03 nm in z.

For 3D tracking, dynamic DNA-origamis were employed. The first sample was a vertical pillar with a single ATTO 647N fluorophore attached to a DNA strand that could oscillate axially between two binding sites. The 3D trajectory for a temporal resolution of 1 ms –translating to approximately 1500 photons per localisation– featured precisions of 4.9 nm, 4.7 nm and 4.1 nm in x, y and z, respectively. To demonstrate the speed of the axial deflection, a slightly modified origami that oscillates faster was employed to acquire z-only traces with many transitions at an acquisition rate of 23.81 kHz. The traces were classified, aligned and averaged; the resulting trace shows a clear step between neighbouring time points providing a pessimistic upper bound for the transition time of 42 µs. Finally, two fluorophores, one ATTO 647N and one JF549, that have individual emission and excitation wavelengths were tracked quasi-simultaneously by switching between two excitation patterns with distinct wavelengths every 255.2 µs: the two molecules were attached to two oscillating arms of a planar DNA-origami featuring four catching sites –two per molecule.

The interferometric beam shaper might find applications outside the realm of Minflux. For instance, in [92] a contrast improving Bessel droplet beam is employed for high-throughput neu-

ronal imaging. The Bessel droplet is generated through an SLM with a phase mask of two annuli and features repeating maxima and minima –the droplets– along z when focused. To generate a uniform excitation profile, the SLM is periodically updated to change the relative phase between the two annuli by $\frac{1}{2}\tau$ which swaps the positions of maxima and minima. This function could be realised through beam shaper interfering two copies of the two annuli phase mask. The polarisation rotator would continuously shift the relative phase and hence the droplets (Figure B.2) –potentially in the nanosecond regime when using resonant electro-optic modulators– resulting in an even more homogeneous profile. The same concept applies to the optical trap presented in [93], that uses Bessel droplets to axially confine multiple particles. Similarly, a photophoretic trap that is capable of loading and unloading particles by opening and closing the tophat beam might be realisable[94]. Moreover, the axial deflection of the tophat zero is ideally suited for enhancing STED imaging of small volumes[95]. In the same vein, a novel technique called 'MINSTED'[96, 97, 98] that combines Gaussian-beam MINFLUX with a STED depletion beam, could be extended to the third dimension.

Nevertheless, the interferometric shaper is projected to be especially valuable for future MINFLUX systems. As shown in this work, it delivers state-of-the-art 3D performance in terms of localisation precision while its speed is unrivalled. Its versatility, as demonstrated by the four channel implementation, opens new avenues for advanced MINFLUX schemes like mixed beam shape and multi-wavelength excitation patterns. Besides the axial deflection, other interferometric manipulations could be of interest: the lateral deflection of a vortex or half-moon zero could replace electro-optic deflectors in scenarios where they are not feasible. The tilting of the vortex zero line or the half-moon zero plane has the potential to generate axial information without the tophat. All these properties makes the interferometric beam shaper a compelling building block that could facilitate broader adoption of 3D MINFLUX and its capabilities for studying nanometric biological structures and fast dynamic processes.

Bibliography

[1] A. van Leeuwenhoek, *Arcana Naturae*. Royal Society, 1695. https://echo.mpiwg-berlin.mpg.de/ECHOdocuView?url=/permanent/library/1S6ZUCDM/pageimg&start=11&pn=13&mode=imagepath.

[2] R. Hooke, *Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses, with observations and inquiries thereupon*. Royal Society, 1665. https://www.gutenberg.org/ebooks/15491.

[3] R. Yuste, "Fluorescence Microscopy Today," *Nature Methods* **2** no. 12, (Dec., 2005) 902–904.

[4] J. F. W. Herschel, "IV. 'Aμόρφωτa, No. I.— on a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless," *Philosophical Transactions of the Royal Society of London* **135** (Jan., 1997) 143–145.

[5] M. Renz, "Fluorescence Microscopy—A Historical and Technical Perspective," *Cytometry Part A* **83** no. 9, (2013) 767–779.

[6] A. H. Coons, H. J. Creech, R. N. Jones, and E. Berliner, "The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody1," *The Journal of Immunology* **45** no. 3, (Nov., 1942) 159–170.

[7] A. H. Coons and M. H. Kaplan, "Localization of Antigen in Tissue Cells," *The Journal of Experimental Medicine* **91** no. 1, (Jan., 1950) 1–13.

[8] E. H. Beutner, "Immunofluorescent Staining: The Fluorescent Antibody Method," *Bacteriological Reviews* **25** no. 1, (Mar., 1961) 49–76.

[9] K. Thorn, "Genetically Encoded Fluorescent Tags," *Molecular Biology of the Cell* **28** no. 7, (Apr., 2017) 848–857.

[10] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, "Green Fluorescent Protein as a Marker for Gene Expression," *Science* **263** no. 5148, (Feb., 1994) 802–805.

[11] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, "A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity," *Science* **337** no. 6096, (Aug., 2012) 816–821.

[12] E. Abbe, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv für Mikroskopische Anatomie* **9** no. 1, (Dec., 1873) 413–468.

[13] E. Abbe, "The Relation of Aperture and Power in the Microscope (Continued).*," *Journal of the Royal Microscopical Society* **2** no. 4, (1882) 460–473.

[14] L. Rayleigh, "On the Theory of Optical Images, with Special Reference to the Microscope," *Journal of the Royal Microscopical Society* **23** no. 4, (1903) 447–473.

[15] M. Minsky, "Microscopy apparatus US patent 3013467," *USP Office, Ed. US* **658** (1961) .

[16] C. Sheppard and A. Choudhury, "Image Formation in the Scanning Microscope," *Optica Acta: International Journal of Optics* **24** no. 10, (Oct., 1977) 1051–1073.

[17] C. R. Sheppard, "Super-resolution in confocal imaging," *Optik (Stuttgart)* **80** no. 2, (1988) 53–54.

[18] D. K. Hamilton, T. Wilson, and C. J. R. Sheppard, "Experimental Observations of the Depth-Discrimination Properties of Scanning Microscopes," *Optics Letters* **6** no. 12, (Dec., 1981) 625–626.

[19] S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Optics Letters* **19** no. 11, (Jun, 1994) 780. http://dx.doi.org/10.1364/ol.19.000780.

[20] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, and S. W. Hell, "Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission," *Proceedings of the National Academy of Sciences* **97** no. 15, (Jul, 2000) 8206–8210. http://dx.doi.org/10.1073/pnas.97.15.8206.

[21] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods* **3** no. 10, (Aug, 2006) 793–796. http://dx.doi.org/10.1038/nmeth929.

[22] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science* **313** no. 5793, (Sep, 2006) 1642–1645. http://dx.doi.org/10.1126/science.1127344.

[23] V. Westphal and S. W. Hell, "Nanoscale Resolution in the Focal Plane of an Optical Microscope," *Physical Review Letters* **94** no. 14, (Apr., 2005) 143903.

[24] N. Bobroff, "Position Measurement with a Resolution and Noise-limited Instrument," *Review of Scientific Instruments* **57** no. 6, (June, 1986) 1152–1157.

[25] R. E. Thompson, D. R. Larson, and W. W. Webb, "Precise Nanometer Localization Analysis for Individual Fluorescent Probes," *Biophysical Journal* **82** no. 5, (May, 2002) 2775–2783.

[26] R. Henriques, C. Griffiths, E. Hesper Rego, and M. M. Mhlanga, "PALM and STORM: Unlocking Live-Cell Super-Resolution," *Biopolymers* **95** no. 5, (2011) 322–331.

[27] A. Sharonov and R. M. Hochstrasser, "Wide-Field Subdiffraction Imaging by Accumulated Binding of Diffusing Probes," *Proceedings of the National Academy of Sciences* **103** no. 50, (Dec., 2006) 18911–18916.

[28] J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, "Super-Resolution Microscopy with DNA-PAINT," *Nature Protocols* **12** no. 6, (June, 2017) 1198–1228.

[29] S. C. M. Reinhardt, L. A. Masullo, I. Baudrexel, P. R. Steen, R. Kowalewski, A. S. Eklund, S. Strauss, E. M. Unterauer, T. Schlichthaerle, M. T. Strauss, C. Klein, and R. Jungmann, "Ångström-resolution fluorescence microscopy," *Nature* **617** no. 7962, (May, 2023) 711–716. https://www.nature.com/articles/s41586-023-05925-9. Number: 7962 Publisher: Nature Publishing Group.

[30] M. G. L. Gustafsson, "Surpassing the Lateral Resolution Limit by a Factor of Two Using Structured Illumination Microscopy," *Journal of Microscopy* **198** no. 2, (2000) 82–87.

[31] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, Background-Free, 3d Super-Resolution Optical Fluctuation Imaging (SOFI)," *Proceedings of the National Academy of Sciences* **106** no. 52, (2009) 22287–22292. http://dx.doi.org/10.1073/pnas.0907866106.

[32] F. Chen, P. W. Tillberg, and E. S. Boyden, "Expansion Microscopy," *Science* **347** no. 6221, (Jan., 2015) 543–548.

[33] P. von Olshausen and A. Rohrbach, "Coherent Total Internal Reflection Dark-Field Microscopy: Label-Free Imaging beyond the Diffraction Limit," *Optics Letters* **38** no. 20, (Oct., 2013) 4066–4069.

[34] F. Jünger, P. v Olshausen, and A. Rohrbach, "Fast, Label-Free Super-Resolution Live-Cell Imaging Using Rotating Coherent Scattering (ROCS) Microscopy," *Scientific Reports* **6** no. 1, (July, 2016) 30393.

[35] J. Ortega-Arroyo and P. Kukura, "Interferometric Scattering Microscopy (iSCAT): New Frontiers in Ultrafast and Ultrasensitive Optical Microscopy," *Physical Chemistry Chemical Physics* **14** no. 45, (Oct., 2012) 15625–15636.

[36] H. R. Saibil, "Cryo-EM in Molecular and Cellular Biology," *Molecular Cell* **82** no. 2, (Jan., 2022) 274–284.

[37] D. P. Allison, N. P. Mortensen, C. J. Sullivan, and M. J. Doktycz, "Atomic Force Microscopy of Biological Samples," *WIREs Nanomedicine and Nanobiotechnology* **2** no. 6, (2010) 618–634.

[38] F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynnå, V. Westphal, F. D. Stefani, J. Elf, and S. W. Hell, "Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes," *Science* **355** no. 6325, (Feb., 2017) 606–612. https://science.sciencemag.org/content/355/6325/606. Publisher: American Association for the Advancement of Science Section: Research Article.

[39] Y. Eilers, H. Ta, K. C. Gwosch, F. Balzarotti, and S. W. Hell, "MINFLUX Monitors Rapid Molecular Jumps with Superior Spatiotemporal Resolution," *Proceedings of the National Academy of Sciences* **115** no. 24, (June, 2018) 6117–6122.

[40] K. C. Gwosch, J. K. Pape, F. Balzarotti, P. Hoess, J. Ellenberg, J. Ries, and S. W. Hell, "MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells," *Nature Methods* **17** no. 2, (Jan, 2020) 217–224. http://dx.doi.org/10.1038/s41592-019-0688-0.

[41] J. K. Pape, T. Stephan, F. Balzarotti, R. Büchner, F. Lange, D. Riedel, S. Jakobs, and S. W. Hell, "Multicolor 3D MINFLUX Nanoscopy of Mitochondrial MICOS Proteins," *Proceedings of the National Academy of Sciences* **117** no. 34, (Aug., 2020) 20607–20614.

[42] L. A. Masullo, F. Steiner, J. Zähringer, L. F. Lopez, J. Bohlen, L. Richter, F. Cole, P. Tinnefeld, and F. D. Stefani, "Pulsed Interleaved MINFLUX," *Nano Letters* **21** no. 1, (Dec, 2020) 840–846. http://dx.doi.org/10.1021/acs.nanolett.0c04600.

[43] R. Schmidt, T. Weihs, C. A. Wurm, I. Jansen, J. Rehman, S. J. Sahl, and S. W. Hell, "Minflux Nanometer-Scale 3d Imaging and Microsecond-Range Tracking on a Common Fluorescence Microscope," *Nature Communications* **12** no. 1, (2021) 1478. https://doi.org/10.1038/s41467-021-21652-z.

[44] E. Slenders and G. Vicidomini, "ISM-FLUX: MINFLUX with an Array Detector," *Physical Review Research* **5** no. 2, (Apr., 2023) 023033.

[45] L. M. Ostersehlt, D. C. Jans, A. Wittek, J. Keller-Findeisen, K. Inamdar, S. J. Sahl, S. W. Hell, and S. Jakobs, "DNA-PAINT MINFLUX Nanoscopy," *Nature Methods* **19** no. 9, (Sept., 2022) 1072–1075.

[46] J. Zähringer, F. Cole, J. Bohlen, F. Steiner, I. Kamińska, and P. Tinnefeld, "Combining pMINFLUX, Graphene Energy Transfer and DNA-PAINT for Nanometer Precise 3D Super-Resolution Microscopy," *Light: Science & Applications* **12** no. 1, (Mar., 2023) 70.

[47] T. Deguchi, M. K. Iwanski, E.-M. Schentarra, C. Heidebrecht, L. Schmidt, J. Heck, T. Weihs, S. Schnorrenberg, P. Hoess, S. Liu, V. Chevyreva, K.-M. Noh, L. C. Kapitein, and J. Ries, "Direct Observation of Motor Protein Stepping in Living Cells Using MINFLUX," *Science* **379** no. 6636, (Mar., 2023) 1010–1015.

[48] J. O. Wirth, L. Scheiderer, T. Engelhardt, J. Engelhardt, J. Matthias, and S. W. Hell, "MINFLUX Dissects the Unimpeded Walking of Kinesin-1," *Science* (Mar., 2023) .

[49] F. Semmer, M.-C. Emperauger, C. Lopez, C. Bogicevic, F. Treussart, K. Perronet, and F. Marquier, "3D Real-Time Two-Photon Microscopy Device for Single-Particle Holographic Tracking (3D-Red Shot)," *ACS Photonics* **10** no. 9, (Sept., 2023) 3426–3434.

[50] F. Cole, J. Zähringer, J. Bohlen, T. Schröder, F. Steiner, M. Pfeiffer, P. Schüler, F. D. Stefani, and P. Tinnefeld, "Super-Resolved FRET and Co-Tracking in pMINFLUX," *Nature Photonics* **18** no. 5, (May, 2024) 478–484.

[51] J. D. Rickert, M. O. Held, J. Engelhardt, and S. W. Hell, "4Pi MINFLUX Arrangement Maximizes Spatio-Temporal Localization Precision of Fluorescence Emitter," *Proceedings of the National Academy of Sciences* **121** no. 11, (Mar., 2024) e2318870121.

[52] L. Scheiderer, J. O. Wirth, M. Tarnawski, and S. W. Hell, "Dual-Color MINFLUX: Kinesin-1 Takes Chassé-Inchworm Steps," bioRxiv, Mar., 2024.

[53] M. K. Geismann, A. Gomez-Segalas, A. Passera, M. Shirzadian, and F. Balzarotti, "A Fast Interferometric Beam Shaper for Multi-Emitter 3D MINFLUX," bioRxiv, Feb., 2024.

[54] L. Reymond, J. Ziegler, C. Knapp, F.-C. Wang, T. Huser, V. Ruprecht, and S. Wieser, "Simple: Structured Illumination Based Point Localization Estimator With Enhanced Precision," *Optics Express* **27** no. 17, (Aug, 2019) 24578. https://doi.org/10.1364/oe.27.024578.

[55] P. Jouchet, C. Cabriel, N. Bourg, M. Bardou, C. Poüs, E. Fort, and S. Lévêque-Fort, "Nanometric axial localization of single fluorescent molecules with modulated excitation,". http://dx.doi.org/10.1101/865865.

[56] L. Gu, Y. Li, S. Zhang, Y. Xue, W. Li, D. Li, T. Xu, and W. Ji, "Molecular resolution imaging by repetitive optical selective exposure," *Nature Methods* **16** no. 11, (Sep, 2019) 1114–1118. http://dx.doi.org/10.1038/s41592-019-0544-2.

[57] J. Cnossen, T. Hinsdale, R. Thorsen, M. Siemons, F. Schueder, R. Jungmann, C. S. Smith, B. Rieger, and S. Stallinga, "Localization microscopy at doubled precision with patterned illumination," *Nature Methods* **17** no. 1, (Dec, 2019) 59–63. http://dx.doi.org/10.1038/s41592-019-0657-7.

[58] L. Gu, Y. Li, S. Zhang, M. Zhou, Y. Xue, W. Li, T. Xu, and W. Ji, "Molecular-Scale Axial Localization By Repetitive Optical Selective Exposure," *Nature Methods* **18** no. 4, (2021) 369–373. https://doi.org/10.1038/s41592-021-01099-2.

[59] K. C. Gwosch, *Nanometer resolution imaging with minimal fluorescence photon fluxes*. PhD thesis, Ruperto-Carola University of Heidelberg, June, 2017. https://katalog.ub.uni-heidelberg.de/cgi-bin/titel.cgi?katkey=68190521.

[60] Harald Cramér, *Mathematical Methods of Statistics (PMS-9), Volume 9*. No. 9 in Princeton Mathematical Series. Princeton University Press, Princeton, 1946.

[61] C. R. Rao, "Information and the Accuracy Attainable in the Estimation of Statistical Parameters," in *Breakthroughs in Statistics: Foundations and Basic Theory*, S. Kotz and N. L. Johnson, eds., pp. 235–247. Springer, New York, NY, 1992.

[62] L. von Diezmann, Y. Shechtman, and W. E. Moerner, "Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking," *Chemical Reviews* **117** no. 11, (June, 2017) 7244–7275.

[63] G. A. Reider, "Electrodynamic Theory of Light," in *Photonics: An Introduction*, G. A. Reider, ed., pp. 1–37. Springer International Publishing, Cham, 2016.

[64] L. Novotny and B. Hecht, "Principles of Nano-Optics,". http://dx.doi.org/10.1017/cbo9780511813535.

[65] Pawley, "Handbook Of Biological Confocal Microscopy,". http://dx.doi.org/10.1007/978-0-387-45524-2.

[66] M. Leutenegger, R. Rao, R. A. Leitgeb, and T. Lasser, "Fast focus field calculations," *Opt. Express* **14** no. 23, (Nov, 2006) 11277–11291. https://opg.optica.org/oe/abstract.cfm?URI=oe-14-23-11277.

[67] J. Arlt and M. J. Padgett, "Generation of a Beam with a Dark Focus Surrounded by Regions of Higher Intensity: The Optical Bottle Beam," *Optics Letters* **25** no. 4, (Feb., 2000) 191–193.

[68] M. Belsley, J. Soares-de-Oliveira, and A. J. Pereira, "Bridging the Resolution-Sectioning Gap in Stimulated Emission Depletion Microscopy: Theory and Experiment," bioRxiv, Jan., 2024.

[69] S. Hell, G. Reiner, C. Cremer, and E. H. K. STELzer, "Aberrations in Confocal Fluorescence Microscopy Induced by Mismatches in Refractive Index," *Journal of Microscopy* **169** no. 3, (1993) 391–405.

[70] J. Pape, *Multicolor 3D MINFLUX Nanoscopy for Biological Imaging*. Doctoral thesis, Georg-August-Universität Göttingen, Nov., 2020.

[71] S. Feng and H. G. Winful, "Physical Origin of the Gouy Phase Shift," *Optics Letters* **26** no. 8, (Apr., 2001) 485.

[72] N. Matsumoto, T. Ando, T. Inoue, Y. Ohtake, N. Fukuchi, and T. Hara, "Generation of High-Quality Higher-Order Laguerre-Gaussian Beams Using Liquid-Crystal-on-Silicon Spatial Light Modulators," *JOSA A* **25** no. 7, (July, 2008) 1642–1651.

[73] X. Gu, M. Krenn, M. Erhard, and A. Zeilinger, "Gouy Phase Radial Mode Sorter for Light: Concepts and Experiments," *Physical Review Letters* **120** no. 10, (Mar., 2018) 103601.

[74] C. Maurer, A. Jesacher, S. Bernet, and M. Ritsch-Marte, "What spatial light modulators can do for optical microscopy," *Laser & Photonics Reviews* **5** no. 1, (Dec, 2010) 81–101. `http://dx.doi.org/10.1002/lpor.200900047`.

[75] M. O. Lenz, H. G. Sinclair, A. Savell, J. H. Clegg, A. C. N. Brown, D. M. Davis, C. Dunsby, M. A. A. Neil, and P. M. W. French, "3-D Stimulated Emission Depletion Microscopy with Programmable Aberration Correction," *Journal of Biophotonics* **7** no. 1-2, (2014) 29–36.

[76] S. Tu, X. Liu, D. Yuan, W. Tao, Y. Han, Y. Shi, Y. Li, C. Kuang, X. Liu, Y. Yao, Y. Xu, and X. Hao, "Accurate Background Reduction in Adaptive Optical Three-Dimensional Stimulated Emission Depletion Nanoscopy by Dynamic Phase Switching," *ACS Photonics* **9** no. 12, (Dec., 2022) 3863–3868.

[77] "Deformable Mirrors for Wavefront Control," Boston Micro Machines, 2024. `https://bostonmicromachines.com/products/deformable-mirrors/standard-deformable-mirrors/`.

[78] J. Zähringer, F. Cole, J. Bohlen, F. Steiner, I. Kamińska, and P. Tinnefeld, "Combining pMINFLUX, Graphene Energy Transfer and DNA-PAINT for Nanometer Precise 3D Super-Resolution Microscopy," *Light: Science & Applications* **12** no. 1, (Mar., 2023) 70.

[79] S. Deng, L. Liu, Y. Cheng, R. Li, and Z. Xu, "Investigation of the Influence of the Aberration Induced by a Plane Interface on STED Microscopy," *Optics Express* **17** no. 3, (Feb., 2009) 1714–1725.

[80] T. J. Gould, D. Burke, J. Bewersdorf, and M. J. Booth, "Adaptive Optics Enables 3D STED Microscopy in Aberrating Specimens," *Optics Express* **20** no. 19, (Sept., 2012) 20998–21009.

[81] P. Török and P. R. T. Munro, "The Use of Gauss-Laguerre Vector Beams in STED Microscopy," *Optics Express* **12** no. 15, (July, 2004) 3605–3617.

[82] S. Deng, L. Liu, Y. Cheng, R. Li, and Z. Xu, "Effects of Primary Aberrations on the Fluorescence Depletion Patterns of STED Microscopy," *Optics Express* **18** no. 2, (Jan., 2010) 1657–1666.

[83] B. Harke, C. K. Ullal, J. Keller, and S. W. Hell, "Three-Dimensional Nanoscopy of Colloidal Crystals," *Nano Letters* **8** no. 5, (May, 2008) 1309–1313.

[84] G. A. Reider, "Wave Propagation in Matter," in *Photonics: An Introduction*, G. A. Reider, ed., pp. 39–100. Springer International Publishing, Cham, 2016.

[85] K. Crabtree, "Polarization Critical Optical Systems: Important Effects and Design Techniques," 2007. `https://wp.optics.arizona.edu/optomech/wp-content/uploads/sites/53/2016/10/Tutorial_Paper-Crabtree.pdf`.

[86] G. Anzolin, A. Gardelein, M. Jofre, G. Molina-Terriza, and M. W. Mitchell, "Polarization Change Induced by a Galvanometric Optical Scanner," *JOSA A* **27** no. 9, (Sept., 2010) 1946–1952.

[87] A. V. Tikhonravov, P. W. Baumeister, and K. V. Popov, "Phase Properties of Multilayers," *Applied Optics* **36** no. 19, (July, 1997) 4382–4392.

[88] G. A. Reider, "Optical Interference," in *Photonics: An Introduction*, G. A. Reider, ed., pp. 157–196. Springer International Publishing, Cham, 2016.

[89] B. Saleh, M. Teich, and R. E. Slusher, "Fundamentals of Photonics," *Physics Today* **45** no. 11, (Nov, 1992) 87–88. `http://dx.doi.org/10.1063/1.2809878`.

[90] "Broadband Polarizing Beamsplitter Cubes." Thorlabs. `https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=739`.

[91] P. Gao, B. Prunsche, L. Zhou, K. Nienhaus, and G. U. Nienhaus, "Background suppression in fluorescence nanoscopy with stimulated emission double depletion," *Nature Photonics* **11** no. 3, (Jan, 2017) 163–169. `http://dx.doi.org/10.1038/nphoton.2016.279`.

[92] W. Chen, X. Ge, Q. Zhang, R. G. Natan, J. L. Fan, M. Scanziani, and N. Ji, "High-Throughput Volumetric Mapping of Synaptic Transmission," *Nature Methods* (June, 2024) 1–8.

[93] X. Song, J. Wei, A. Teng, and A. Zhao, "Controllable Axial Optical Chain Beams Using a Holographic Method," *Optics Express* **29** no. 11, (May, 2021) 17304–17315.

[94] A. Turpin, V. Shvedov, C. Hnatovsky, Y. V. Loiko, J. Mompart, and W. Krolikowski, "Optical Vault: A Reconfigurable Bottle Beam Based on Conical Refraction of Light," *Optics Express* **21** no. 22, (Nov., 2013) 26335–26340.

[95] A. Barbotin, I. Urbančič, S. Galiani, C. Eggeling, M. Booth, and E. Sezgin, "z-STED Imaging and Spectroscopy to Investigate Nanoscale Membrane Structure and Dynamics," *Biophysical Journal* **118** no. 10, (May, 2020) 2448–2457. `https://www.sciencedirect.com/science/article/pii/S000634952030309X`.

[96] M. Weber, M. Leutenegger, S. Stoldt, S. Jakobs, T. S. Mihaila, A. N. Butkevich, and S. W. Hell, "MINSTED Fluorescence Localization and Nanoscopy," *Nature Photonics* **15** no. 5, (May, 2021) 361–366.

[97] M. Weber, H. von der Emde, M. Leutenegger, P. Gunkel, S. Sambandan, T. A. Khan, J. Keller-Findeisen, V. C. Cordes, and S. W. Hell, "MINSTED Nanoscopy Enters the Ångström Localization Range," *Nature Biotechnology* **41** no. 4, (Apr., 2023) 569–576.

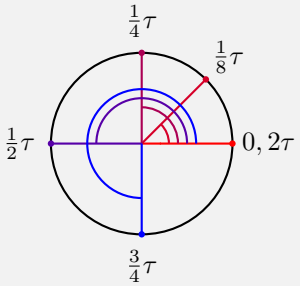
[98] L. Scheiderer, H. von der Emde, M. Hesselink, M. Weber, and S. W. Hell, "MINSTED Tracking of Single Biomolecules," *Nature Methods* **21** no. 4, (Apr., 2024) 569–573.

[99] M.-S. Yeh, S.-G. Shiue, and M.-H. Lu, "First-Order Analysis of a Three-Lens Afocal Zoom System," *Optical Engineering* **36** no. 4, (Apr., 1997) 1249–1258.

A. Use of τ

This document uses τ instead of π to denote a quantity of radians.

$$\tau := 2\pi$$



B. Beam Shape Transitions

B.1. Vortex Tilt Axis Rotation

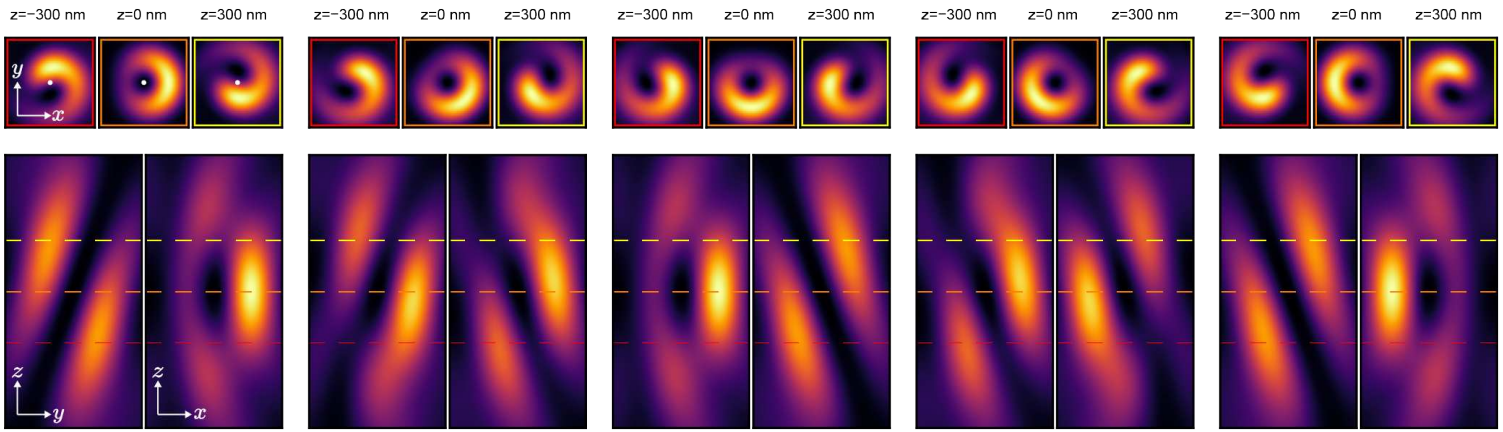


Figure B.1.: Five different relative phases for the vortex plus tophat combination: the tilt axis is rotated by 180°. The top row shows $800 \times 800\,\mathrm{nm}$ xy-planes for three values of z; red: $-300\,\mathrm{nm}$, orange: $0\,\mathrm{nm}$, $300\,\mathrm{nm}$. The white dots mark the origin to make the orientation of tilt axis, oriented along y, more visible. The intensity around the tilt axis shows an azimuthal dependence with a maximum that winds around the optical axis. The bottom row shows xz- and yz-planes, each $800 \times 1600\,\mathrm{nm}$.

B.2. Bessel Droplet Axial Scan

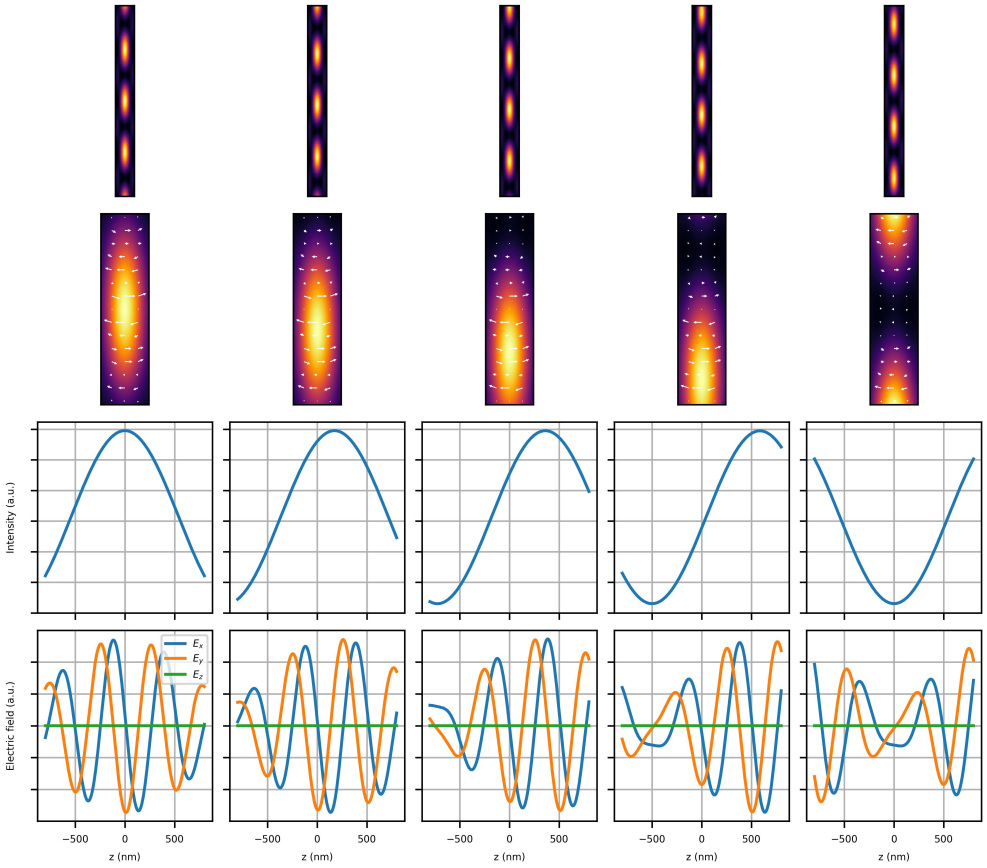


Figure B.2.: Five different relative amplitudes for the interference of two Bessel droplet beams where one Bessel droplet phase mask's inner annulus is shifted by $\frac{1}{2}\tau$. The radii of the annuli were at 36 % and 71 % –with 2 % thickness each– of the SLM pupil. Top row: xz 0.8 µm × 8 µm. Bottom row: xz 0.4 µm × 1.6 µm. axis. The bottom row shows xz- and yz-planes, each 800 × 1600 nm.

C. Three-Lens Telescope

The motivation behind a three-lens telescope is to adjust the magnification without changing the collimation state; something that is impossible with a two-lens system. The spacings between lenses –as shown in Figure C.1– in a three lens afocal zoom-system[1] is given by[99]:

$$\begin{aligned} d_1 =& f_1 + f_2 + \frac{f_1 f_2 M}{f_3} \\ d_2 =& f_2 + f_3 + \frac{f_2 f_3}{f_1 M}\,, \end{aligned}$$

where d_1/d_2 are the distances between the first/second and the second/third lens, f_i is the focal length of the i-th lens and M is magnification of the system which we define to be the multiplicative inverse of the definition used in [99], $M = \frac{h_3}{h_1}$.

The position of the first lens is fixed at the origin. The positions of the second and third lenses are given by d_1 and $d_{12} = d_1 + d_2$, respectively; writing them as a function of M gives:

$$\begin{aligned} d_1(M) =& f_1 + f_2 + \frac{f_1 f_2 M}{f_3} \\ d_{12}(M) =& f_1 + 2f_2 + f_3 + \frac{f_1 f_2 M}{f_3} + \frac{f_2 f_3}{f_1 M}\,. \end{aligned}$$

$d_1(M)$ is a linear function of M: the position of the second lens needs to be adjusted proportional to the desired change in magnification. However, Setting the derivative, d'_{12}, of the position of the third lens with respect to M, to zero, yields:

$$\begin{aligned} & d'_{12} \stackrel{!}{=} 0 \\ \Rightarrow & \frac{f_1 f_2}{f_3} - \frac{f_2 f_3}{f_1 M^2} = 0 \\ \Rightarrow & M = \frac{f_3}{f_1}\,. \end{aligned}$$

Consequently, when aiming for a base magnification M_0, f_1 and f_3 should be chosen such that: $M_0 = \frac{f_3}{f_1}$. Then, in a first order approximation, M can be adjusted by only moving the second lens while the position of the third lens stays constant. In practice, however, both lenses' positions were moved to keep the beam well collimated.

The total length of the three-lens telescope for the above case is:

$$d_1(M_0) + d_2(M_0) = f_1 + \mathbf{4} f_2 + f_3\,.$$

Therefore, the three-lens telescope is always longer than its two-lens counterpart unless one uses concave lenses with negative focal lengths. As our system works with multiple wavelengths, all lenses are required to be achromatic. However, achromatic concave lenses were not easily available. Furthermore, for compactness' sake, f_2 should be kept small –30 mm in our case.

[1] An afocal zoom system maps a collimated input beam to a collimated output beam.

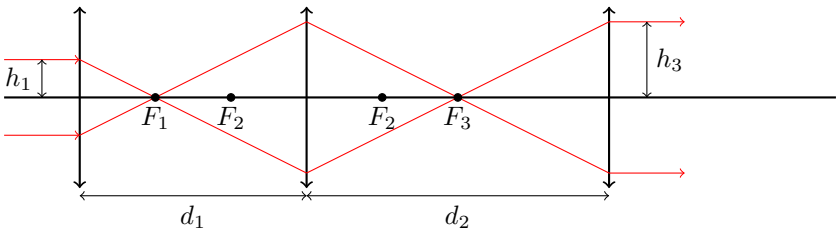


Figure C.1.: Three-lens telescope. A parallel input ray at height h_1 becomes a parallel output ray at height h_3; the magnification is $M = \frac{h_3}{h_1}$. The distances between the lenses are d_1 and d_2. The focal planes of the lenses are indicated by F_i. Here: $f_1 = f_2$, $f_3 = 2f_1$ and $M = 2$.

D. Non-4f Image Relay

In ray optics, a telescope consisting of two thin lenses with focal lengths f_1 and f_2 that are $f_1 + f_2$ apart is modelled by the matrix

$$T(f_1, M) = \begin{pmatrix} -M & (M+1)\ f \\ 0 & -\frac{1}{M} \end{pmatrix}$$

where the parametrisation with f_1 –the focal length of the first lens– is chosen and $M = \frac{f_2}{f_1}$, the magnification.

To calculate where an image is formed after the telescope, we take the ray vector starting from a point on the optical axis ($y = 0$) at an arbitrary angle ϑ. We let it propagate through free space by some distance z, then the telescope, and then some other distance z'.

$$\mathrm{FS}(z')T(f_1, M)\mathrm{FS}(z) \begin{pmatrix} y = 0 \\ \vartheta \end{pmatrix} = \begin{pmatrix} y' = & (M+1)\ f\,\vartheta - \frac{z'\,\vartheta}{M} - M\,z\,\vartheta \\ \vartheta' = & -\frac{\vartheta}{M} \end{pmatrix}, \tag{D.1}$$

where $\mathrm{FS}(z)$ is the free space propagation matrix for a distance z:

$$\mathrm{FS}(z) = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix}.$$

The image of the initial point is formed where its ray hits the optical axis for any ϑ, i.e., all rays converge into a single point. This amounts to solving $y' = 0$ for z' from Eq. (D.1). The result is:

$$z' = M(f_1 - M\Delta z),$$

with $z = f_1 + \Delta z$. See Figure D.1.

The case $z = f_1$ constitutes the regular 4f system such that the image is formed at the distance $z' = f_2$ behind the second lens of the telescope. If the image to be relayed is farther away, i.e., $\Delta z > 0$, we can increase Δz up to f_1/M before the image is positioned on the plane of the second lens ($z' = 0$) which we choose to be the limiting case. In theory it should be possible to pick up the image even it if lies before the second lens. For practical purposes we rather have the image 'in free space' some distance from the second lens (as opposed to inside or before the lens).

This means, that for a telescope with $M = 1$ we can move the original image plane a distance of up to $z = f_1 + \frac{f_1}{1} = 2f_1$ away from the first lens of the telescope. For $M = 1/2$ it would be a distance of $z = f_1 + \frac{f_1}{1/2} = 3f_1$. To relay an image at $z = 900\,\mathrm{mm}$ with an $M = 1/2$ telescope, one could choose $f_1 = 400\,\mathrm{mm}$ (and $f_2 = 200\,\mathrm{mm}$), meaning $\Delta z = 900\,\mathrm{mm} - 400\,\mathrm{mm} = 500\,\mathrm{mm}$ to achieve a maximum range of 1200 mm and form the image at $z' = 75\,\mathrm{mm}$ behind the second lens. This makes the optical setup more wieldable compared to a 4f system as the telescope length and the distance between the relayed images shrinks from 1350 mm to 600 mm and 2700 mm to 1575 mm, respectively.

The drawback, however, is that the image moves closer to the second lens. In the above example, the 75 mm might be too close to allow further manipulations with other optical elements (like electro-optic deflectors) before the next image relay. Hence, one might need to install an extra telescope, .e.g, in a regular 4f configuration that only serves the purpose of picking up the image close to the lens and allows enough space around the image plane at its output. Still, the added flexibility can pay off especially when space on the optical table is scarce.

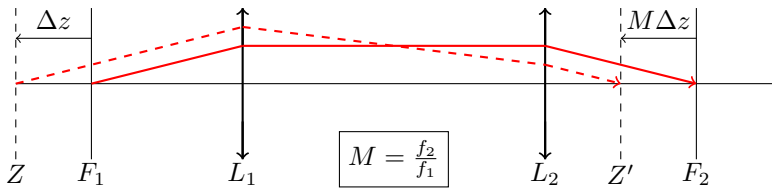


Figure D.1.: Ray optics diagram of a telescope. A 4f system in the conventional configuration relays an image at plane F_1 at a distance equal to the focal length of the first lens f_1 to the plane F_2 at a distance f_2 from the second lens. The image of a plane Z at $f_1 + \Delta z$ is located at Z', $M\Delta z$ closer the second lens (if $\Delta z > 0$).

E. Optical Setup

E.1. List of Components

Lasers	
E.Laser1	LDH-IB-640-B (PicoQuant, Berlin, Germany)
E.Laser2	Cobolt Jive™ 100-561 (Cobolt AB, Solna, Sweden)
WFE.Laser1	Omicron QuixX® 642-140 (Omicron, Rodgau-Dudenhofen, Germany)
WFE.Laser2	Cobolt Jive™ 100-561 (Cobolt AB, Solna, Sweden)
WFE.Laser3	iBEAM-SMART-488-S-HP (Toptica, Munich, Germany)
AC.Laser	LBX-405 (Oxxius, Lannion, France)
zS.Laser	MDL-XS-940 (CNI, Changchun, China)
SLED	EXS210153-01 (Exalos AG, Schlieren, Switzerland)

Beam Modulation	
E.AOTF	AOTFnC 400.650-CPCh-TN (AA Opto Electronic, Orsay, France)
E.EOM	EM200K (Leysop Ltd., Essex, England) + A-304 high-voltage amplifier (A.A.Lab Systems Ltd., Ramat Gan, Israel)
A.EOM, B.EOM	LM 0202 ADP (Excelitas Technologies, Waltham, MA, USA) + A-304 high-voltage amplifier (A.A.Lab Systems Ltd., Ramat Gan, Israel)
DP.SLM	1920 x 1152 Spatial Light Modulator (Meadowlark Optics, Inc., Frederick, CO, USA)

Scanning	
EOD	2 mm aperture Electro-Optical Deflector (Leysop Ltd., Essex, England) + WMA-300 high-voltage amplifier + WMA-IB-HS phase inverter (Falco Systems BV, Katwijk aan Zee, The Netherlands)
STG.TM	PSH 10/2 + NV 200/D controller (piezosystem jena GmbH, Jena, Germany)
Coarse Stage Z	L-505.013212F (Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany)
Stage	(Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany)
Linear actuator	M-229.25S
Piezo stage	P-545.3D8S + E-727 controller

Polarisation and Beam Transport	
BS	50:50 beamsplitter (Thorlabs Inc., Newton, NJ, USA)
PBS	polarising beamsplitter (Thorlabs Inc., Newton, NJ, USA)
HWP	achromatic half-wave plate (either Thorlabs Inc., Newton, NJ, USA or Edmund Optics Inc., Barrington, NJ, USA)
QWP	achromatic quarter-wave plate (either Thorlabs Inc., Newton, NJ, USA or Edmund Optics Inc., Barrington, NJ, USA)
Fibers	polarisation maintaining single mode fiber PM-S405-XP and PM780-HP, and multimode fiber M50L022S-A (Thorlabs Inc., Newton, NJ, USA)
SuK	fiber collimator 60FC-* (Schäfter+Kirchhoff, Hamburg, Germany)

Lenses and Mirrors (Thorlabs Inc., Newton, NJ, USA if not specified otherwise)	
STG.OBJ	Olympus UPLXAPO 60X (Evident Europe GmbH, Hamburg, Germany)
L	achromatic lens with VIS or NIR AR coating
AB.EM	1-inch elliptical dielectric mirror
DP.MM	1-inch metallic mirror
DP.SQM	2-inch dielectric square mirror
STG.EMM	1-inch elliptical metallic mirror
mot.	mirror on N-480 mount (Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany)
E.PP	prism-pair PS875-A

Dichroic Mirrors and Filters	
E.DC	Di03-R561 (Semrock Inc., Rochester, NY, USA)
E.CU1	clean-up filter LD01-640/8 (Semrock Inc., Rochester, NY, USA)
E.CU2	clean-up BrightLine 561/14 (Semrock Inc., Rochester, NY, USA)
E.ND	NE20A (Thorlabs Inc., Newton, NJ, USA)
AB.DC, BC.DC	Di03-R561 (Semrock Inc., Rochester, NY, USA)
WFE.DC1	ZT561rdc (Chroma Technology Corp., Bellows Falls, VT, USA)
WFE.DC2	ZT488rdc (Chroma Technology Corp., Bellows Falls, VT, USA)
STB.CU1	FBH850-40 (Thorlabs Inc., Newton, NJ, USA)
STB.CU2	950-20 bandpass filter (Edmund Optics Inc., Barrington, NJ, USA)
STB.DC	ZT860lpxr (Chroma Technology Corp., Bellows Falls, VT, USA)
AC.DC	ZT442rdc (Chroma Technology Corp., Bellows Falls, VT, USA)
STG.DC1	Di03-R405/488/561/635 (Semrock Inc., Rochester, NY, USA)
D.NF	quad-notch NF03-405/488/561/635E-25 (Semrock Inc., Rochester, NY, USA)
D.CU	FESH0800 (Thorlabs Inc., Newton, NJ, USA)
STG.DC2	T800DCSPXR (Chroma Technology Corp., Bellows Falls, VT, USA)

Detectors	
D.APD	SPCM-AQRH-43-BR1 (Excelitas Technologies, Waltham, MA, USA)
WFD.CAM	DMK 33UX249 (The Imaging Source Europe GmbH, Bremen, Germany)
STB.CAM	DMK 33UX174 (The Imaging Source Europe GmbH, Bremen, Germany)

Control Hardware and Computers(not shown in the diagram)	
PC	2 personal computers running Windows 10 (Microsoft Corp., Redmond, WA, USA) and LabVIEW 2020 (National Instruments, Austin, TX, USA)
DAQ	NI PCIe-6363 (National Instruments, Austin, TX, USA) + USB-3114 (Measurement Computing Corporation, Norton, MA, USA)
FPGA	NI USB-7856R (National Instruments, Austin, TX, USA)

Lens	Focal length (mm)
E.SuK1, E.SuK2	6.2
A.SuK, B.SuK	7.5
A.L1, B.L1	50
A.L2, B.L2	30
A.L3, B.L3	80
DP.L	200
OS.L1	400
OS.L2	200
EOD.L1	150
EOD.L2	150
EOD.L3	200
EOD.L4	200
EOD.L5	125
EOD.L6	400
WFE.SuK1	6.2
WFE.L1	15
WFE.L2	25
WFE.L3	100
WFE.L4	200
AC.L1	4
AC.L2	30
AC.L3	80
AC.L4	200

Lens	Focal length (mm)
STG.SL	180
STG.TL	300
zS.L	6.2
STB.L1	40
STB.L2	6.2
STB.L3	150
STB.L4	200
STB.L5	300
STB.L6	80
WFD.L1	100
D.L1	150
D.L2	30
D.L3	19
D.L4	30
D.L5	19

Table E.1.: Lens and their focal lengths

Retarder	Slow Axis Rotation
E.EOM	45°
A.EOM	45°
B.EOM	45°
A.HWP	$\sim 0°$
A.QWP	0°
B.QWP	0°
DP.QWP	45°
DP.HWP	22.5°
EOD.HWP	45°
STG.QWP	45°

Table E.2.: Rotation angles of retarders.

F. Tuning L

The Fisher information matrix, Equation (1.8), for known SBR and a d-dimensional molecule location, $\boldsymbol{r}$ (transform via Equation (1.10)), is explicitly given by[70]:

$$\mathcal{F}_{\boldsymbol{r}} = \frac{1}{N}\sum_{i=0}^{K-1}\begin{pmatrix} \left(\frac{\partial p_i}{\partial r_0}\right)^2 & \cdots & \frac{\partial p_i}{\partial r_0}\frac{\partial p_i}{\partial r_{d-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial p_i}{\partial r_{d-1}}\frac{\partial p_i}{\partial r_0} & \cdots & \left(\frac{\partial p_i}{\partial r_{d-1}}\right)^2 \end{pmatrix}.$$

Defining a reference MINFLUX scenario with K exposures, a given L_{ref} and $\mathrm{SBR}^*_{\mathrm{ref}}(\boldsymbol{r}=\boldsymbol{0}) = \mathrm{SBR}^*_{0,\mathrm{ref}}$, we can express $\mathrm{SBR}^*_0(L)$ for the same scenario with differing L as

$$\begin{aligned}\mathrm{SBR}^*_0(L) =& \mathrm{SBR}^*_{0,\mathrm{ref}}\frac{\sum_i I_i(0)}{\sum_i I_{i,\mathrm{ref}}(0)} = \\ =& \frac{L^2\mathrm{SBR}^*_0}{L_0^2}\,,\end{aligned}$$

where I_i and $I_{i,\mathrm{ref}}$ are the parabolic intensities for each exposure. The SBR for an arbitrary molecule position, $\boldsymbol{r}$ is

$$\mathrm{SBR}^*(L) = \mathrm{SBR}^*_0(L)\frac{\sum_i I_i(\boldsymbol{r})}{\sum_i I_{i,\mathrm{ref}}(\boldsymbol{r})}\,.$$

After substituting $\mathrm{SBR}^*(L)$ for SBR in the p-functions (Equation (1.12)) and assuming a total photon count that contains a single collected signal photon, $N = 1 + 1/\mathrm{SBR}$, one can evaluate the CRB covariance matrix at the origin $\Sigma_{\mathrm{CRB}}(\boldsymbol{0}) = \mathcal{F}^{-1}_{\boldsymbol{r}(\boldsymbol{0})}$. Then, taking limit $L \to 0$ for the diagonal elements of the square root of the covariance matrix

$$\sigma_{\mathrm{CRB},0} = \lim_{L\to 0}\sqrt{\Sigma_{\mathrm{CRB},ii}}\,,$$

we arrive at the lowest CRB possible, $\sigma_{\mathrm{CRB},0}$, for a given SBR that is derived from a reference scenario (MINFLUX scheme with L_{ref}) for which we assume its according SBR, $\mathrm{SBR}^*_{\mathrm{ref}}$, is known (or can be measured). The ratio $\sigma_{\mathrm{CRB}}(L)/\sigma_{\mathrm{CRB},0}$ yields an expression that does not depend on $\mathrm{SBR}^*_{\mathrm{ref}}$:

$$\frac{\sigma_{\mathrm{CRB}}}{\sigma_{\mathrm{CRB},0}}(L) = \sqrt{\frac{K}{2^d}\mathrm{SBR}^*(L) + 1}\,,$$

with K the number of exposures and $d \in \{1, 2\}$ the dimension of the localisation e.g., $d = 2$ for localising in xy. For 3D localisation, we assume that the axial is decoupled from the lateral localisation such that $d = 1$ for z.

When adjusting L to give $\mathrm{SBR}^* = 1$, $\frac{\sigma_{\mathrm{CRB}}}{\sigma_{\mathrm{CRB},0}}$ evaluates to 1.58 for a 1D localisation ($d = 1$ and $K = 3$) and 1.5 for a 2D localisation ($d = 2$, $K = 5$) meaning the achievable precision could still be improved by $\sim 50\,\%$ to $60\,\%$. However, it is not clear how the MLE might perform in situations with $\mathrm{SBR}^* < 1$; the $\mathrm{SBR}^* = 1$ threshold hence remains arbitrary until further studies.

G. Sample Preparation

G.1. Coverslips and Flow Chambers

All glass coverslips were cleaned by sonicating them in a water-bath with a 2 % solution of Hellmanex III for 10 minutes, repeated three times. Between rounds the coverslips were rinsed with MilliQ water. Before storage, they were dried with compressed nitrogen.

All samples, except the cell samples, used flow chambers consisting of a coverslip on top of two parallel strips of double sided scotch tape applied along the long edges of a microscope slide.

G.2. Fluorescent Beads

The fluorescent bead samples were prepared on a microscope slide with a flow chamber via the following steps:

Flow in	20 µL	gold particles 90 nm in 1x PBS	1:5
Wait	5 minutes		
Wash	20 µL	1x PBS	
Flow in	20 µL	Tetraspeck in 1x PBS	1:250
Wait	10 minutes		
Wash	20 µL	1x PBS	
Flow in	20 µL	1x PBS	
Seal coverslip		Epoxy glue	

G.3. Gold Particles

Microscope slides with flow chambers as described in the last section were used. The gold nanoparticles with 40 nm, instead of 90 nm, were flown in. Next, NOA 74 optical adhesive (*Norland Products Inc., Jamesburg, New Jersey, USA*) was flown in and cured with a handheld UV lamp for 15 to 20 minutes.

G.4. Single Fluorescent Molecule DNA-Origamis

The DNA-origami samples were prepared as described in [38]. Again, the sample was prepared on a microscope slide with a flow chamber:

Flow in	20 µL	gold particles 90 nm in 1x PBS	1:5
Wait	5 minutes		
Wash	20 µL	buffer A	
Flow in	20 µL	0.5 mg neutravidin in buffer A	
Wait	5 minutes		
Wash	20 µL	buffer A	
Wash	20 µL	buffer B	
Flow in	20 µL	DNA-origami in buffer B	30 pM
Wait	8 minutes		
Wash	20 µL	buffer B	
Flow in	20 µL	ROXS (tracking) or blinking buffer (imaging)	
Seal coverslip		Twinsil	

Buffer A consists of 10 mM Tris-HCl, pH 7.5, 100 mM NaCl and 0.05 % Tween 20. Buffer B consists of 5 mM Tris-HCl pH 8, 1 mM EDTA, 12.5 mM MgCl2 and 0.05 % Tween. The ROXS buffer consists of 1X PCA (1 mM), 1X PCD (10 nM), 1X MV (1 mM), and 1X Trolox (1 mM) in buffer B. The blinking buffer consists of 1X GLOX (0.4 mg ml-1 glucose oxidase, 64 µg ml-1 catalase), and 1-10 mM MEA in buffer B + 10 % w/v glucose.

H. Transition Averaging: Two State Classification

The figures in this section show the same transition averaging presented in Section 4.3.2, however, the raw traces are simulated: photon counts were drawn from two Poissonian distributions according to two parabolic intensities for an $L_z = 100\,\mathrm{nm}$. The curvature was chosen such that the average photon counts more or less matched the experimental photon counts ($n_1 \approx n_0 \approx 5$, $N \approx 10$). The molecule position was a trace jumping between two positions 14 nm apart with a 2 nm offset simulating imperfect centring of the excitation pattern. The transition probability was chosen to be 5 % per time step. The two figures compare the use of the `viterbi` and `map` algorithms provided by the `predict()` method of the `HiddenMarkovModel` class found in the python package `pomegranate v0.14.8`. The `viterbi` (forward only) algorithm leads to artefacts in the average, see Figure H.1. It detects transitions that shortly before transition go down and shortly after transition overshoot. Intuitively, this makes sense as it renders the transition "more obvious". Unfortunately, this signature can be found in noise and consequently the `viterbi` classification fails to catch the real transition. The `map` (forward-backward) algorithm reduces this artefact significantly but cannot get rid of it entirely, see Figure H.2. Using the ground truth for classification yields the ideal case with no under/overshoots, see Figure H.3.

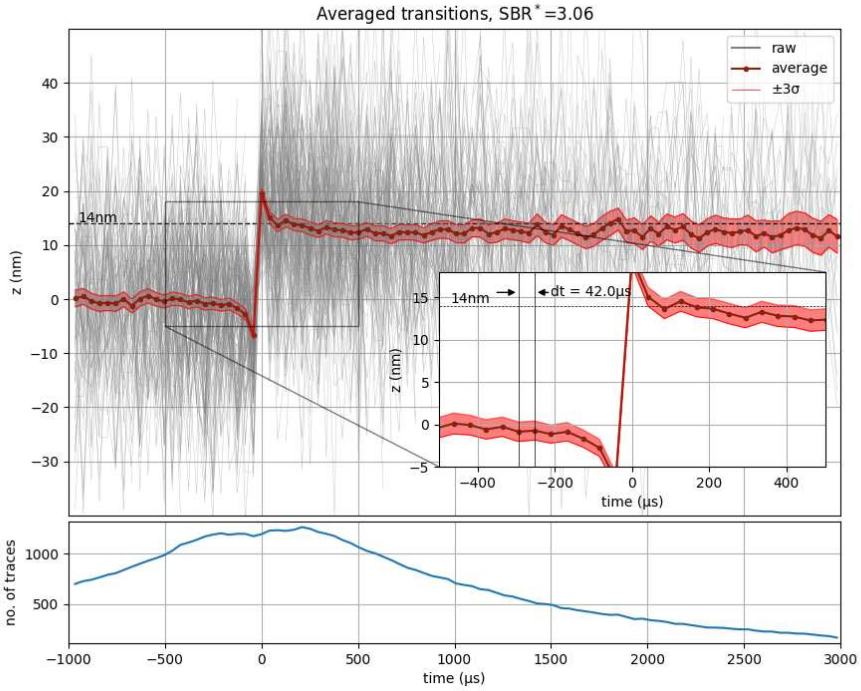


Figure H.1.: Simulated traces classified via the `viterbi` algorithm and the resulting average. One clearly sees an undershoot/overshoot around the transition caused by how the HMM classified a signature in noise; the raw traces in grey align in a particular way which causes the average to exhibit these artefacts.

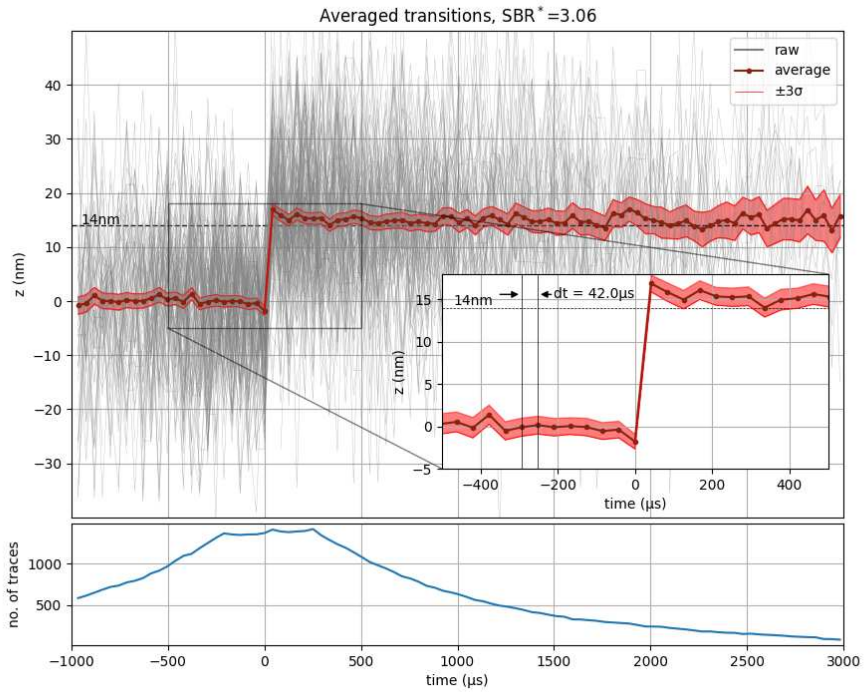


Figure H.2.: Simulated traces classified via the `map` algorithm and the resulting average. Even though the artefacts are severely reduced they are still visible.

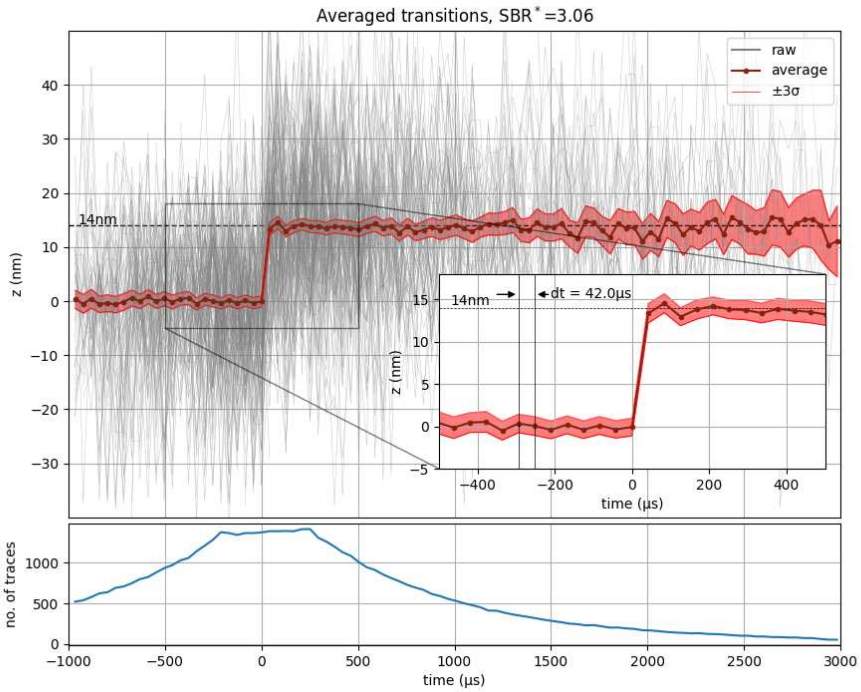


Figure H.3.: Simulated traces classified via the ground truth and the resulting average; the ideal case.

I. Measurement Configuration

I.1. Imaging

Excitation Scheme	L_{xy} (nm)	L_z (nm)	**Photons**	**Live estimator** β_0	**Live estimator** β_1
3x4 DNA origami					
7G	400	-	200	-	-
4G	300	-	200	-	-
5V	150	-	120	0.90	7.18
5V	100	-	280	0.57	10.80
3TH	-	350	280	0.40	2.00
3TH	-	150	280	0.40	1.50
3TH	-	100	320	0.40	1.20
VTH+	50	75	10000	0.80	22.00
Nuclear Pore Complex					
7G	400	-	200	-	-
4G	300	-	200	-	-
5V	150	-	120	0.90	7.18
5V	100	-	280	0.57	10.80
3TH	-	350	400	0.40	2.00
3TH	-	150	400	0.40	1.50
VTH+	75	75	10000	0.80	11.00

Table I.1.: Excitation schemes and live estimator parameters for the 3×4 DNA origami and the Nuclear Pore Complex sample.

Each localisation was computed from 2000 photons. After filtering, see I.3, the localisations were clustered via the DBSCAN algorithm provided by the `scikit-learn.cluster v1.2.0` python package. The arguments for the minimum number of points and the maximum distance (in nm) were

- `minPts = 3`, `eps = 5` for the DNA-origami sample and
- `minPts = 5`, `eps = 6` for the NPC sample.

I.2. Tracking

	Excitation Scheme	L_{xy} (nm)	L_z (nm)	**Termination condition**	**Live estimator** β_0	**Live estimator** β_1
Axial oscillator						
Centring	7G	400	-	t_{max}: 200 ms d: 20 nm	-	-
	5V	150	-	N: 400 t_{max}: 100 ms	0.90	7.18
	3TH	-	350	d: 5 nm	0.40	2.00
	VTH	100	100	-	-	-
Static tracking	3TH	-	350	d: 5 nm	0.40	2.00
	VTH	100	100	-	-	-
Dual-colour oscillator						
Centring (640 nm)	7G	400	-	t_{max}: 200 ms d: 20 nm	-	-
	5V	150	-	N: 400 t_{max}: 100 ms	0.90	7.18
	3TH	-	350	d: 5 nm	0.40	2.00
Static tracking	VTH (640 nm)	100	100	-	-	-
	VTH (560 nm)	100	100	-	-	-

Table I.2.: Excitation schemes, termination conditions, and live estimators for Axial and Dual-colour oscillator tracking.

I.3. Filtering

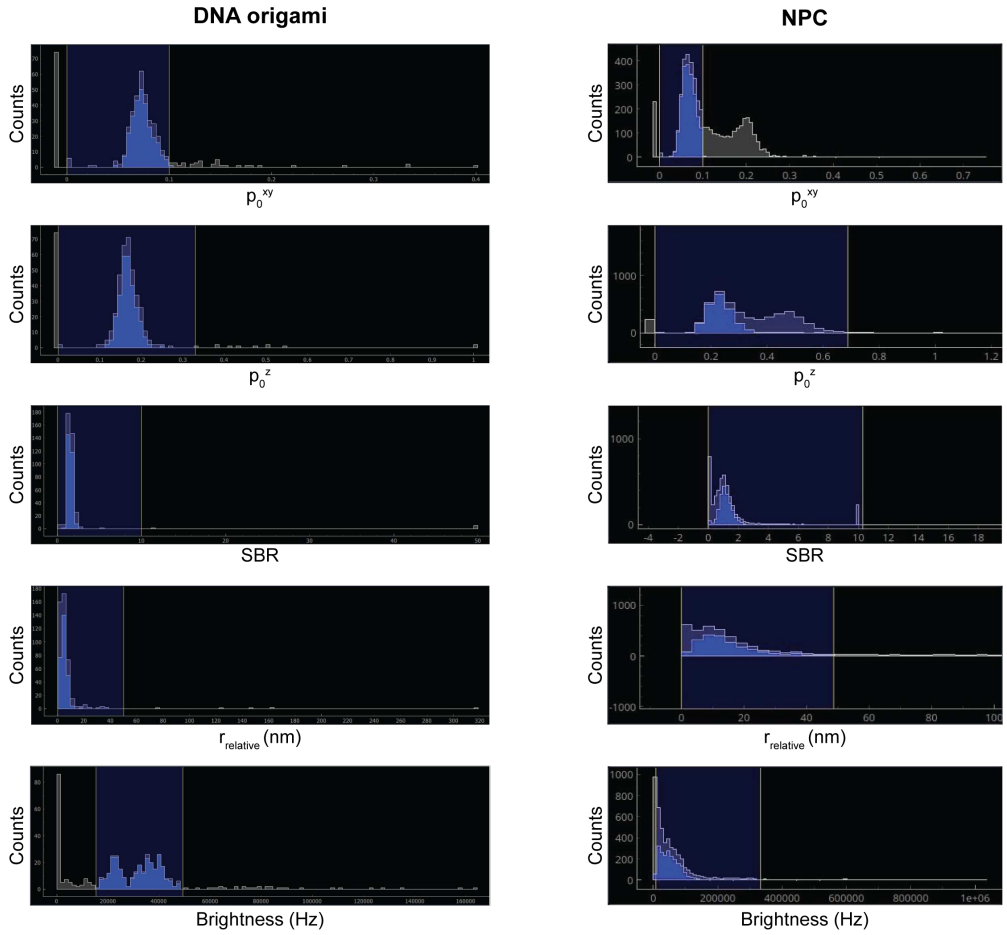


Figure I.1.: Histograms of the filtered parameters for imaging data. The coloured bands indicate the filtering window with min/max values as given in the tables I.3 and I.4.

p_0^{xy}	$[0, 0.10]$
p_0^z	$[0, 0.33]$
SBR	$[0, 10]$
r_{relative}	$[0, 50]$nm
Brightness	$[1.54 \times 10^4, 4.94 \times 10^4]$ counts per second

Table I.3.: Imaging data filtering: DNA origami.

p_0^{xy}	$[0, 0.10]$
p_0^z	$[0, 0.33]$
SBR	$[0, 19.7]$
r_{relative}	$[0, 37.5]$nm
Brightness	$[2 \times 10^3, 1 \times 10^5]$ counts per second

Table I.4.: Imaging data filtering: NPC.