STED microscopy with two-photon photoactivation in the visible range & Application of generative adversarial networks for image reconstruction in STED microscopy

Dissertation

for the award of the degree
"Doctor rerum naturalium" (Dr. rer. nat.)
of the Georg-August-Universität Göttingen

within the doctoral program Physics
of the Georg-August University School of Science (GAUSS)

Submitted by
Jan-Erik Bredfeldt
from Weingarten, Germany

Göttingen, 2022

Thesis Committee

1. **Prof. Dr. Dr. h. c. mult. Stefan W. Hell**, Department of NanoBiophotonics, Max Planck Institute for Multidisciplinary Sciences, Göttingen
2. **Prof. Dr. Tim Salditt**, Research Group Structure of biomolecular assemblies and x-ray physics, Institute for X-ray Physics, Georg-August University, Göttingen
3. **Prof. Dr. Claus Ropers**, Research Group Nano-Optics and Ultrafast Dynamics, IV. Physical Institute, Georg-August University, Göttingen

Members of the Examination Board

1. First Reviewer: **Prof. Dr. Dr. h. c. mult. Stefan W. Hell**, Department of NanoBiophotonics, Max Planck Institute for Multidisciplinary Sciences, Göttingen
2. Second Reviewer: **Prof. Dr. Tim Salditt**, Research Group Structure of biomolecular assemblies and x-ray physics, Institute for X-ray Physics, Georg-August University, Göttingen
3. Further members of the Examination Board
    - **apl. Prof. Dr. Alexander Egner**, Department of Optical Nanoscopy, Institut für Nanophotonik Göttingen e.V., Göttingen
    - **Prof. Dr. Claus Ropers**, Research Group Nano-Optics and Ultrafast Dynamics, IV. Physical Institute, Georg-August University, Göttingen
    - **Prof. Dr. Stefan Jakobs**, Research Group Mitochondrial Structure and Dynamics, Max Planck Institute for Multidisciplinary Sciences, Göttingen
    - **Prof. Dr. Silvio Rizzoli**, Department of Neuro- and Sensory Physiology, University Medical Center, Göttingen

Date of the oral examination: 05.12.2022

This dissertation is dedicated to my family and Julia.
For their endless support and encouragement.

Abstract

The resolution limit of classical optical far-field microscopes was broken with the invention of super-resolution microscopy. Today, a variety of these techniques achieving different degrees of resolution exist. Image formation with most of these methods can be easily affected by either photobleaching of fluorescent dyes due to the use of high-intensity lasers or out-of-focus fluorescence signal from dyes in different imaging planes. The resulting loss in contrast reduces the possibility to distinguish individual structures or to localize individual fluorophores, ultimately limiting the gain in resolution. This is especially the case when imaging axially extended, densely labeled biological samples, where the out-of-focus signal can be quite substantial.

Two-photon excitation as well as two-photon photoactivation (2PA) are promising strategies to mitigate this problem. The sharply confined two-photon active volume facilitates imaging with significantly reduced background signal. The use of STED compatible photoactivatable dyes enables the application of leading-edge imaging modes including the nanometer-resolution MINSTED nanoscopy. The goal of this thesis is to present efficient 2PA in the visible spectrum with multiple photoactivatable dyes. The broad application of this technique, together with its benefits compared to regular one-photon activation, is demonstrated for different microscopy techniques.

On the other hand, in cases without 2PA, a deteriorated signal-to-background ratio in measurements that are affected by a loss in contrast can be potentially recovered through post-processing. The second part of this thesis investigates how the image processing can be optimized to make informed decisions in cases where the underlying structure is well known from previous experiments. Neural networks, which are trained with simulated data of microtubules, are used to recover the information that is present in any acquired low signal-to-background image. The optimal training parameters are determined and the application on experimental data is presented, outperforming classical algorithms.

Contents

List of Abbreviations

CNN	convolutional neural network
EM	electron microscopy
EPSF	effective point spread function
FWHM	full width at half maximum
GAN	generative adversarial network
GDD	group delay dispersion
GVD	group velocity dispersion
mSSIM	mean structural similarity index measure
NA	numerical aperture
NIR	near infrared
PALM	photoactivated localization microscopy
PSF	point spread function
ReLU	rectified linear unit
SBR	signal-to-background ratio
SiR	silicon rhodamine
SMLM	single molecule localization microscopy
STED	stimulated emission depletion
STORM	stochastic optical reconstruction microscopy

2PA two-photon activation

UV ultraviolet

VIS visible

WLC worm like chain

1. Introduction

The first description of the cell [1] and some larger bacteria was achieved in the 1600s by Anton von Leeuwenhoek through biological imaging with the use of an optical microscope. His simple microscope contained just a single lens which was placed between the sample and the eye. Even today, 360 years after this discovery, optical imaging still plays a decisive role in uncovering biological processes through the visualization of the underlying structure. Over time microscopes have improved significantly and have developed from the use of single lenses to complex optical systems. However, the imaging of individual sub-cellular structures closer than $\sim 200\,\mathrm{nm}$ remained impossible due to a fundamental limit postulated by Ernst Abbe in 1873 [2]. This limit is based on the refraction of light and limits the maximum resolution of visible far field microscopes with a wavelength of λ to a lateral size $d_{xy_{\mathrm{min}}}$

$$d_{xy_{\mathrm{min}}} \approx \frac{\lambda}{2\mathrm{NA}} = \frac{\lambda}{2n\sin\Theta} \tag{1.1}$$

With NA being the numerical aperture of the objective lens, n the refractive index, and Θ the half opening angle of the objective lens. The resolution $d_{xy_{\mathrm{min}}}$ is here defined as the full-width at half-maximum (FWHM) of the point spread function (PSF) of the microscope.

Electron microscopy (EM) can deliver resolution down to the atomic scale with the use of electrons, which possess a wavelength which is orders of magnitude lower compared to visible light. Due to limitations such as imaging in vacuum, *in vivo* imaging of biological samples is not possible with EM. Therefore, fluorescence microscopy is employed, where the biomolecules of interest are tagged with markers that emit light. This allows for a specific labeling of different proteins in a single cell with much simpler sample preparation compared to EM where the specific staining of proteins is not possible. With fluorescence microscopy samples can be

observed in living cells, allowing for the imaging of cellular dynamics.
In fluorescence microscopy the physical limit described above, set by the wave-nature of light, was broken with the invention of stimulated emission depletion (STED) in 1994 [3] and other super-resolution techniques like stochastic optical reconstruction microscopy (STORM) [4] and photoactivated localization microscopy (PALM) in 2006 [5]. Fundamentally, all three methods rely on distinguishing individual fluorophores based on their electronic states. In these methods emitters are typically switched between their bright ON-state and a dark OFF-state, where no fluorescence is emitted. The individual implementation differs between these techniques but can be summarized in two groups, the coordinate-targeted techniques such as STED and coordinate-stochastic techniques such as STORM and PALM.
In coordinate-targeted methods the location of the fluorophore is determined by the position of the fluorescence suppressing STED beam (see section 1.1), which is scanned over the sample. In single molecule localization microscopy (SMLM) methods on the other hand, the spatial separation is achieved by a separation in time with a stochastic switching of single fluorophores. In the following, these single fluorophores are localized individually with sub-diffraction limited localization precision.

1.1 STED microscopy

In STED microscopy the fluorophores are transferred to a high vibrational excited state $S_{1,\mathrm{vib}}$ through the illumination with a diffraction-limited Gaussian-shaped excitation beam (see blue line in Fig. 1.1 D). From there molecules transition quickly to a lower excited state through internal relaxation. The off-switching of the fluorophore is achieved through stimulated emission by a second beam illuminating the sample together with the excitation beam. Instead of releasing a photon by spontaneous emission (fluorescence) the STED beam forces the molecule to a higher level ground state by stimulated emission of a photon with the same wavelength. To achieve good spectral filtering of the fluorescence and stimulated emission and to avoid STED induced re-excitation, the STED beam is shifted to the red tail of the emission spectrum. Therefore, the energy gap between the relaxed ground state and the excited state is bigger than the STED photon energy. However, re-excitation is non-zero since the transition from the first excited state to a higher

excited state could still be induced by the STED light. From these higher excited states the molecule can enter a triplet state through inter-system crossing. The relaxation from this state is non-fluorescent and of significantly longer life-time than the simple excited state.
Photobleaching of the dye is irreversible and can occur from any excited and triplet state. This ultimately leads to a decrease in image quality through a loss of observable contrast. The specific pathway is often very complex [6], but triplet states are assumed to play a major role [7, 8]. The excitation of the molecule to higher electronic levels through the STED beam must therefore be kept at a minimum.

The focused STED beam for two-dimensional imaging is usually patterned through the use of a vortex-phase plate. In the focus this forms a doughnut shape with a minimal intensity in the center. When overlapped with the Gaussian-shaped excitation beam, only fluorophores in the center of the STED beam stay in their ON-state (see Fig. 1.1 A, B). Moving outward, the number of molecules in their ON-state is decreasing until saturation is reached where all fluorophores are turned off. The size of the effective PSF (EPSF) can be reduced by the manipulation of this saturation point. This is achieved through an increase in the STED laser intensity and the effective resolution can be described with a modified version of the resolution formula by Abbe

$$d_{xy} \approx \frac{\lambda}{2\mathrm{NA}\sqrt{1 + I/I_S}} \tag{1.2}$$

with I being the applied STED intensity and I_S the STED intensity where half of the molecules are brought back to the ground state within the excited state lifetime.
For the resolution improvement in axial direction, an annular phase plate is introduced in the beam path, generating a so-called bottle beam in the focus (see Fig. 1.1 C). The axial resolution can be calculated similarly to Eq. 1.2 with

$$d_z \approx \frac{\lambda}{n - \sqrt{n^2 - n^2 \sin^2(\Theta)} \cdot \sqrt{1 + I/I_S}} \tag{1.3}$$

Both beams for 2D- and axial-STED can be overlapped incoherently to achieve resolution enhancement in both directions (see Fig. 1.1 C).

It has been shown that the addition of time-gating the detected fluorescence signal

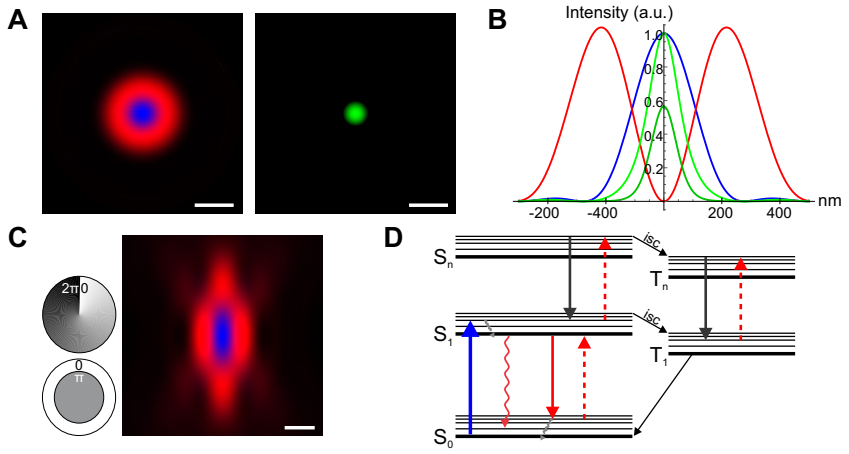

Figure 1.1: STED principle of resolution improvement. A Gaussian excitation beam (blue) overlapped with the doughnut-shaped STED beam (red). The resulting effective STED PSF is shown in green on the right, scale bars 400 nm. **B** Cross-section through the PSF of STED, Gaussian excitation and the resulting EPSF with the normalized intensity on the y-axis. Colors are the same as in A. The gating of the fluorescent signal can improve the resolution but decreases the obtained signal, shown in dark green. **C** Axial xz-view of the resulting profile of the 3D-STED doughnut. This intensity profile in the focus of the beam is achieved with the generation of a 2D-doughnut through a vortex phase plate, shown on the left top, and a bottle-beam through an annular phase plate, left bottom. Scale bar 400 nm. **D** Jablonski diagram of the energy levels of the fluorescent molecule. The excitation beam is shown in blue, the STED beam as dashed-red lines. The fluorescence photon by a red-squiggly line, the stimulated emission by a red solid line. Vibrational relaxation is shown in dark grey, non fluorescent conversions with a black arrow. S_0 is the ground state, S_1 the excited state and S_n the higher excited states. T_1 and T_n are triplet states reached by inter-system crossing (isc).

to time-windows later in the fluorescence lifetime can further improve the resolution (g-STED [9, 10]). Early on, directly after the illumination with STED, not all molecules in the vicinity of the doughnut minimum are yet switched off immediately and some are still able to release photons before an effective stimulated emission. These photons would lead to a reduction of the achieved resolution and are therefore discarded in the application of g-STED, where only photons carrying the maximum resolution information are considered. This effect is also visualized in Fig. 1.1 B where the effective g-STED PSF is shown in dark green. In comparison to regular STED, the resolution is increased exhibiting smaller side lobes and a lower base. However, the overall collected signal is decreasing with an increase in the gating time.

Theoretically, STED nanoscopy can deliver single nanometer resolution by simply increasing the applied STED intensity (see Eq. 1.2). In practice however, with a non-perfect doughnut this is hard to achieve. Any residual STED intensity in the doughnut minimum is resulting in a decreased detected fluorescence signal. This can be quite substantial for very high resolution measurements since only very little STED power is needed for the initial suppression of the majority of the fluorescence signal. Additionally, scanning over the sample with the high intensities in the doughnut crest is resulting in photobleaching of not yet imaged fluorophores. Due to these reasons, it is in practice impossible to achieve single nanometer resolution in a classical scanning STED microscope.

1.2 MINFLUX and MINSTED microscopy

Recent advances in super-resolution microscopy make true molecular resolution not only theoretically but also technically achievable. The new concepts of MINFLUX [11, 12] and MINSTED [13] combine the ideas of the coordinate-targeted and coordinate-stochastic approaches. Both methods, MINFLUX and MINSTED, are based on single molecule localization. The sparsity of the probed structure is ensured by stochastic on-switching of single photoactivatable dyes that are present in the sample.

In MINFLUX the localization is achieved with a doughnut-patterned excitation beam which is scanned through several positions around the estimated molecular coordinate. The fluorescence photon counts at each MINFLUX iteration position are used to localize the molecule and to update the scanning pattern accordingly. The initial guess is based on a scan with a diffraction limited Gaussian excitation beam. The position and the shape of the excitation beam profile, which are recorded at the beginning of the measurement and are well-known values, encode the information about the individual fluorophore's position.
In the theoretical limit of a background free measurement, the most photon efficient localization is achieved when the doughnut minimum is positioned right on top of the fluorophore, resulting ideally in a localization information without emitted photons. Because of this, MINFLUX is typically implemented with an iterative scheme where the area that is covered with the excitation beam scanning

pattern is reduced progressively according to an increased knowledge about the position of the imaged fluorophore. Based on the previous iteration the center of the new scanning pattern is positioned on top of the estimated location of the molecule. This effectively increases the overall photon-efficiency by scanning more tightly around the target fluorophore.

MINSTED follows a similar concept where instead of a single doughnut-shaped excitation beam the conventional pairing of a Gaussian excitation beam and STED doughnut are used for the localization of the fluorophore. After successful activation of a single molecule, its rough position is estimated with an excitation-only overview scan. The precise location of the emitter is then determined by encircling it with the STED EPSF. The radius of the scan pattern is progressively reduced to circle in on the molecule of interest. At the same time the FWHM of the effective STED PSF is decreased until it reaches a minimum d_{min}. This allows nanometer scale localization precision with only ~ 1000 detected photons even under relatively large background conditions [13].

1.3 Contrast loss in super-resolution microscopy

A microscope is generally tasked with three objectives: providing magnification, resolution and contrast. Good magnification and high resolution are meaningless in a microscope without enough contrast to distinguish objects of interest. The well resolved features blend into the background and are therefore invisible.

A loss in contrast in super-resolution fluorescence microscopy might be due to background signal originating from different parts of the imaged sample. In SMLM methods like MINFLUX a high background signal significantly reduces the chances for good fluorophore localizations due to a potentially reduced switching event detection and reduced localization precision. Similarly, STED in thick, densely labeled biological samples suffers from out-of-focus signal which reduces the overall image quality.
Another source of reduced contrast is a higher amount of photobleaching, induced for example by the high power STED beam. During extended imaging sessions, e.g. movie acquisition or 3D imaging, photobleaching can often become a signifi-

cant drawback, especially when a high resolution is required.

This work discusses two different approaches to resolve the issue of contrast-loss in super-resolution microscopy. The first part discusses the use case of two-photon activation in the visible regime for its application with different (super-resolution) microscopy techniques. The second part investigates to which extend post-processing of very low contrast STED images can be used to reconstruct the underlying structure. The gained knowledge about these approaches would allow to make informed decisions about the imaging parameters that should be used for a given imaging task.

2. Two-photon activation

Two-photon microscopy has become a very popular tool for the imaging of thick biological samples due to its optical sectioning and therefore enhanced field of depth [14–16]. For example, it is widely used in neurosciences due to the enhanced penetration depth of the used IR wavelengths and the reduced fluorescence background created through scattered photons [14, 15]. The main focus of these applications lies on two-photon excitation of already fluorescent dyes.
Two-photon photoactivation of photoactivatable fluorescent dyes has, similar to two-photon excitation, a potential to deliver a better signal-to-background contrast compared to regular imaging due to the restriction of the activated two-photon volume, see Fig. 2.1 D. This limits the fluorescence readout from active molecules and the subsequent photobleaching to only a small region, making two-photon activation (2PA) a great tool for extended three dimensional structures.

2.1 Photoactivatable fluorescent dyes

Photoactivatable fluorophores offer the ability to further control the emitted fluorescence spatially as well as temporally. This additional control enables imaging and spectroscopic measurements that were otherwise inaccessible. They are used for diffusion measurements and measurements investigating cellular dynamics [17–19], photolysis experiments [20] and for super-resolution microscopy [5, 11, 13, 21].

Photoactivatable fluorescent dyes are transferred from an inactive non-fluorescent state to an active fluorescent form after being illuminated with light at an activation wavelength. The obtained active version of the fluorophore is subsequently illuminated at the exciting wavelength which delivers the desired fluorescent signal (see Fig. 2.1 A).

In general the process of fluorescence suppression in the non-fluorescent state can be achieved either by reducing the absorption at the excitation wavelength in the ground state S_0 or by increasing the non-radiative decay from the first excited state S_1 to the ground state S_0 [22].
Similarly, the method of activation can differ between different photoactivatable dyes. Typically the activation light induces a cleavage of one or more covalent bonds, which is followed by a rearrangement of the distribution of the remaining electrons in the molecule. This process is irreversible since it permanently separates the dye from its caging group(s).
The actual method of fluorescence suppression as well as the activation mechanism for each dye can be studied in the corresponding literature [13, 21, 23].

2.2 Two-photon microscopy

It has been shown early-on by Maria-Göppert Mayer [24] that an excitation, induced by the absorption of a single photon, can also be achieved by the simultaneous absorption of two photons of lesser energy. In the basic form two photons with the same energy interact with a molecule; in turn producing the same effect as a single photon with double the energy and half the wavelength. Similar to two-photon excitation, 2PA is possible where two photons interact with a photoactivatable fluorophore to transfer it into its active state (see Fig. 2.1 A). For this to happen the photons have to interact almost simultaneously with the molecule on time-scales of $\sim 5 \times 10^{-15}$ s or less [25]. This makes the probability for a 2PA event extremely small and is the reason why the experimental application of the theoretical proposal had to wait decades to become practical.

By focusing the laser beam with a high NA objective, the photon density and therefore the chance for successful 2PA can be increased significantly. In the resulting focal spot the photon flux is much higher than far away from the focal spot. The probability for two-photon absorption is proportional to the square of the laser intensity and therefore limits the two-photon active region to a small area in the center of the PSF, where the intensity is high enough to cross the threshold. This is the basis of the intrinsic three-dimensionality of this non-linear two-photon process (see Fig. 2.1 B right).
In comparison, in the regular one-photon case the density of activated/excited

molecules is similarly the highest in the focal point. But when moving axially away from the focal point the number of activated/excited molecules is not decreasing drastically but is more and more spread out laterally. Therefore, the sum of the total lateral signal detectable after 1PA or similarly the signal which is generated during one-photon excitation is constant over the entire area in the axial dimension. Therefore, there is no optical sectioning after the application of 1PA (see Fig. 2.1 B left).
This intensity squared behavior of 2PA is present throughout the whole sample and illumination beam path. Therefore the FWHM of the effective two-photon excitation PSF, for a beam with the wavelength λ in a medium with the refractive index n focused with high NA, can be calculated using the work of [26, 27]

$$\begin{aligned} \mathrm{FWHM}_{xy} &= \sqrt{2\ln(2)} \cdot \frac{0.33\lambda}{\mathrm{NA}^{0.91}} \\ \mathrm{FWHM}_{z} &= \sqrt{2\ln(2)} \cdot 0.53\lambda \left(\frac{1}{n - \sqrt{n^2 - \mathrm{NA}^2}} \right) \end{aligned} \tag{2.1}$$

This holds true in the absence of two-photon excitation saturation. When reaching excitation saturation in the center of the PSF, further increases in the photon flux will only result in an addition of excited molecules at the outer region of the PSF. This leads to an increase of the two photon active volume and a flattened effective two-photon excitation PSF [28]. This is also a concern in 2PA where a saturation of the activated molecules in the center in order to achieve a high fluorescence signal, can lead to an increase in the apparent two-photon activation volume and a reduction of the optical sectioning capability.
Therefore, there exists an optimal activation intensity which delivers a high amount of fluorescent molecules while preserving most of the spatial resolution gain.

Two ways to elevate the chances of two photons interacting with the molecule simultaneously are to focus the laser beam tighter or to increase the incident laser power. However, instead of simply increasing the power of a continuous wave (cw) laser, the photon density is temporarily increased with the application of pulsed lasers. Ultra-short ($\sim$ fs) pulsed laser increase the photon concentration and deliver a high number of photons over a very short period of time to the sample. While having high peak power, this process keeps the required average laser power relatively low. The expected number of photons that are absorbed by the molecule depend on the peak power P_{peak}. The peak power relates to the

average power P_{avg} through $P_{\text{peak}} = g_p \frac{P_{\text{avg}}}{f\tau}$ with the repetition rate f, the pulse length τ and a pulse shape dependent parameter g_p (0.66 for a Gaussian-shaped pulse). This results in a number of photons n_p that are absorbed by a molecule with a two photon absorption cross section of σ_2 [29]

$$n_p \approx \frac{P_{\text{avg}}^2 \sigma_2}{f^2 \tau} \left(\frac{\text{NA}^2}{2\hbar c \lambda} \right)^2 \tag{2.2}$$

From this follows a direct influence of the pulse length on the activation probability.

Due to the broader spectral width of ultra-short laser pulses, compared to longer pulses, the chromatic dispersion in the material that is being passed can cause frequency dependent phase changes in the illumination wavefront. This can therefore lead to a change in the length of the laser pulse, typically to a broadening, which reduces the chances for a successful two-photon event. Meanwhile, the spectral shape of the pulse remains the same.

After passing the media, the electric field of the laser pulse can be described with $E(t,z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(\omega) exp(i(\omega t - \phi(\omega, z)))d\omega$. The added phase delay originating from the medium $\phi(\omega, z)$ is dependent on the frequency. This added frequency term can be described with a Taylor expansion around the carrier frequency of the pulse $\phi(\omega) \approx \phi_0 + (\omega - \omega_0)(\frac{d\phi(\omega)}{d\omega})_{\omega_0} + \frac{1}{2!}(\omega - \omega_0)^2(\frac{d^2\phi(\omega)}{d\omega^2})_{\omega_0} + \ldots$. The first term describes an added constant phase, whereas the second term describes the group delay adding an overall time delay to the pulse. The second order term is denoted as the group delay dispersion (GDD). Different spectral components of the pulse are delayed differently which change the pulse temporally. The group velocity dispersion (GVD) is closely related to the GDD through $\phi(\omega) = k(\omega)L$ ($\text{GVD} = \frac{d^2k}{d\omega^2}$), with L being the length of the material that is being passed. After Fourier transform the new pulse duration τ can be obtained from the original pulse length τ_0 with

$$\tau(L) = \tau_0 \sqrt{1 + \left(\frac{4\ln(2) d^2\phi(\omega)/d\omega^2}{\tau_0^2} \right)^2} = \tau_0 \sqrt{1 + \left(\frac{4\ln(2)\text{GVD} \cdot L}{\tau_0^2} \right)^2} \tag{2.3}$$

The total GDD which is introduced by the material is based on its chromatic dispersion. The wavelength dependent refractive index can be calculated with a modified Sellmeier formula. The GVD at the wavelength of interest is then given by the second derivative of the index of refraction with respect to the wavelength

GVD $= \frac{\lambda^3}{2\pi c^2}\frac{d^2n}{d\lambda^2}$ [30].
Typical optical materials such as glass introduce positive dispersion. As a result, long wavelength components of the pulse are traveling faster than the short wavelength components, leading to a positive chirp and a pulse broadening.
This dispersive pulse broadening can be compensated by additionally introducing anomalous dispersion with $\frac{d^2n}{d\lambda^2} < 0$. Prisms as well as gratings introduce angular dispersion which produces a negative GDD [31]. Typically, a four prism pulse compressor is constructed. Angular dispersion introduced by the first prism is corrected by the second one and the second set of prisms is required to eliminate spatial dispersions that were introduced. Simpler configurations are possible where the four prism setup is folded, with the easiest utilizing one [32] or two prisms [33] where the beam is passing the prisms multiple times. Prisms are aligned such as the incoming beam which is incident near the Brewster's angle and minimum deviation angle. The GDD introduced by the prism sequence is based on the wavelength dependent path length $P(\lambda)$ with GDD $\propto \frac{d^2P(\lambda)}{d\lambda^2}$ [34]. It can be shown that the total wavelength dependent optical path length introduced results to $P(\lambda) = 2l\cos\beta(\lambda)$ [33, 35] where l denotes the distance between the prism apices. The angle $\beta(\lambda)$ is the angle between the ray of wavelength λ traveling through both prisms and the line connecting their apices. According to [34, 36, 37] the total GDD introduced can be shown to result to

$$\text{GDD} \approx \frac{\lambda^3}{2\pi c^2}\left(-4L_{\text{sep}}\left(\frac{dn}{d\lambda}\right)^2 + L_{\text{prism}}\frac{d^2n}{d\lambda^2}\right) \tag{2.4}$$

The first term is always negative and is a result of the angular dispersion. The second term is positive and accounts for the positive dispersion introduced by passing through a length of L_{prism} through the prisms. The total GDD can easily be varied by the separation of the two prisms and their placement within the beam path. The use of high dispersive glass can help to reduce the separation needed between the prism pairs to compress the pulses adequately.
The minimal achievable pulse length is limited by its time-bandwidth product, the product of the FWHM of its pulse length and the spectral width FWHM. Very short pulses posses a larger spectral width and vice versa. The product for bandwidth-limited Gaussian pulses is ≈ 0.44 [38] giving a lower bound to the pulse length for a given spectral width.

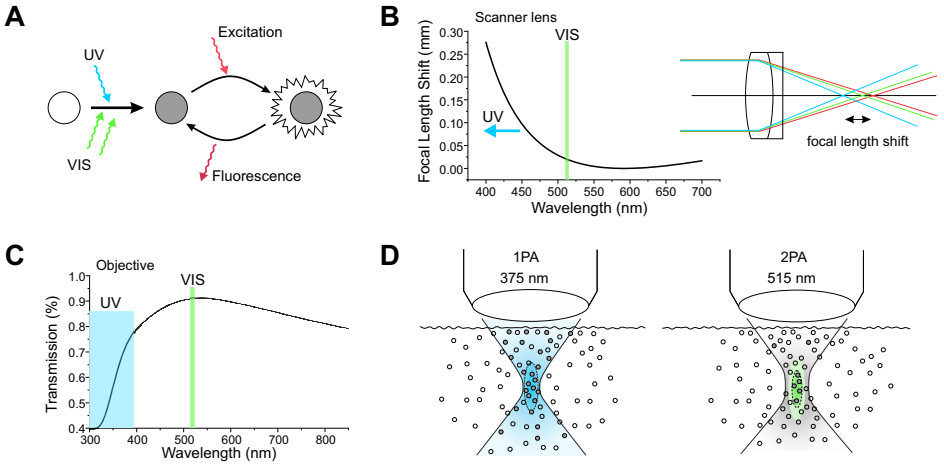

Figure 2.1: Two-photon activation principle. **A** A photoactivatable fluorophore can be activated either using one photon UV (< 400 nm) or two photons in the visible (VIS, this case 515 nm). The activated fluorophores (dark grey) can in turn be excited with another wavelength to emit fluorescence with a third, slightly red shifted, wavelength. **B** Chromatic aberrations from the scanning lens used in this work as focal length shift in mm as the function of the wavelength in nm. The aberrations are increasing drastically when moving the wavelength to UV. The focal length shift for 515 nm is indicated in green. A cartoon of the chromatic focal shift is depicted on the right. **C** Transmission of laser light in percent for different wavelengths in case of the objective lens that was used in this work. Transmission of light below 400 nm is much lower than at 515 nm indicated in green. **D** The activation of fluorophores in a sample with UV leads to continuous activation along the beam path with the highest density in the focal plane (left, FWHM of 375 nm indicated with black dashed outline). The 2PA in the visible has a very restricted activation volume in the focal region (right). The FWHM of the effective 2PA volume is drawn in bright green with a dashed outline, the FWHM of the 515 nm beam is shown in light green.

2.3 Scope of this chapter

STED microscopy of thick biological samples is challenging because of the increased out-of-focus fluorescence in densely labeled structures. This leads to the loss of contrast, described in section 1.3, reducing the overall image quality. Similarly, enhanced photobleaching of fluorescent dyes during 3D STED is adding to that contrast loss. Together, these factors have remained to be the major drawbacks of high-resolution STED imaging.

Two-photon excitation [39–41] is commonly used in confocal microscopy to reduce

background contributions to the signal. Two-photon excitation can also be implemented together with STED [42].
However, similar issues as described above (out-of-focus signal, reduced imaging depth) also affect novel nanometer-scale SMLM techniques such as MINFLUX and MINSTED. These methods make use of photoactivatable fluorescent dyes, which need to be switched to their excitable state before imaging. Consequently, the parasitic emission of activated dyes outside of the focal volume in regular one-photon activation is reducing the number of high-precision localization events.
Similar to two-photon excitation the 2PA can restrict the activation to a small volume in the focal point of the objective. Contributions to the background from activated fluorophores outside of the focal plane are therefore limited in this approach. Additionally, since not yet activated molecules are not getting bleached during the exposure with high intensity excitation and STED laser light, total fluorescence outside of the activation volume is preserved. A successful application of 2PA is therefore not only beneficial for STED, but can also be used in SMLM such as MINFLUX and MINSTED.

Typically, the activation wavelength used for conversion of photoactivatable fluorophores is in the range of UV ($< 400\,\mathrm{nm}$). However, the transmission of light through glass is limited in this wavelength range, leading to substantial losses at every optical element. An example is shown in Fig. 2.1 C. Only 51 % of 375 nm light can pass through the objective lens that was used in this work, compared to a transmission of 85 % for a wavelength in the visible range at 515 nm. Furthermore, common lenses exhibit severe chromatic aberrations in the UV wavelength region, complicating the practical application of this laser light in a microscope. Fig. 2.1 B shows the chromatic focal shift of the used achromatic scanning lens as an example. The deviation of the focal length is changing significantly for wavelengths below 400 nm.
Therefore in addition to the restricted activation volume, a successful application of 2PA at longer wavelengths than UV, through a multiphoton process, could alleviate the aforementioned problems.

In previous work, the application of 2PA has been largely limited to the uncaging of chemicals [43, 44] and optogenetics [45, 46]. Only little work focuses on the photoactivation of dyes or their photo-conversion.
Previous work discussing the 2PA of coumarin fluorophores [47] or of photoacti-

vatable GFP [48] as well as the conversion of fluorescent cyanine-based dyes [49] limits itself to the classical two-photon wavelength regime above 700 nm in the near infrared (NIR). These long wavelengths show enhanced penetration of biological specimens due to lower absorption and reduced scattering [50]. A downside of using these long wavelengths is the relatively large effective two-photon PSF. Additionally, some photoactivatable dyes, especially those designed for STED at 775 nm, are specifically engineered to exhibit a very low two-photon cross section for their unactivated compounds at NIR wavelengths. This is done to minimize unfavorable photoactivation induced by the high power STED beam while imaging [21, 51]. Because of these reasons, a shorter wavelength outside of the NIR range has to be found for 2PA.

Recently, 2PA has been reported at shorter wavelengths for confocal microscopy at 560 nm [52]. Unfortunately, most photoactivatable far red dyes used for super-resolution microscopy such as STED and MINSTED show already a non-negligible absorption at this wavelength. This is potentially resulting in significant bleaching due to the high laser power used for 2PA ($1\,\mathrm{MW/cm^2}$).
To overcome this problem and to use 2PA together with super-resolution microscopy the activation wavelength needed to be shifted. For this reason different photoactivatable dyes were tested in this thesis with a new shorter 2PA wavelength, still in the visible regime, at 515 nm (see Fig. 2.1). Since successful two-photon absorption events are heavily reliant on many parameters as described in section 2.2, optimal activation parameters were determined for all dyes. Different activation schemes were implemented depending on the individual fluorophore's strengths (e.g. activation time or high stability).
Finally, the application of 2PA with all its benefits at this wavelength was demonstrated as a platform for a broad range of microscopy modalities from confocal over STED down to single nanometer-scale SMLM such as MINSTED.

2.4 Methods

2.4.1 STED-Microscope

The STED measurements in this thesis were performed on a custom-built confocal scanning STED microscope (see Fig. 2.2).
The microscope setup features a pulsed STED laser (OneFive Katana, NKT Photonics, Birkerød, Denmark) operating at a wavelength of 775 nm with a pulse duration of 640 ps and repetition rate of 40 MHz. The STED beam is expanded and subsequently split into two different paths with a polarizing beamsplitter cube (PBS). In one path, the formation of the 2D doughnut is achieved by a vortex phase plate (VPP-1a, RPC Photonics, W. Henrietta, NY, USA); a custom-made annular phase plate is placed in the second path for the formation of the 3D doughnut. A second polarizing beamsplitter is recombining both beams incoherently. The ratio of the intensities between both paths can be controlled together with a half-wave-plate ($\lambda/2$) placed before the first PBS.
The microscope is equipped with one LDH-D-C-640 pulsed diode laser (PicoQuant, Berlin, Germany) as an excitation light source at 640 nm which is triggered by the STED laser.

For the imaging with photoactivatable compounds, the microscope has two activation laser wavelengths. For the conventional 1PA, a continuous wave (cw) laser operating at 375 nm (Lambda Beam, rgb lasersystems, Kelheim, Germany) is used where the output power can be modified electronically. For 2PA, the microscope is equipped with a femtosecond pulsed laser with the central wavelength of 515 nm and an output pulse duration of 179 fs operating at a rate of 19.23 MHz (Halite, Fluence technology, Warsaw, Poland). For fast on-off switching of the laser and power control, an AOTF (AOTF-A3-400.700-Pv-T5w4-TN, AA Sa, Orsay, France) was introduced. In order to compensate for the pulse broadening of the 2PA laser, by the optics of the microscope, a compressor system based on two prisms (SF11, Thorlabs) is implemented. The laser beam is passing the two prisms and gets split spectrally during the first pass. The beam components are reflected back at a different height to reverse the spectral splitting. The resulting beam is co-aligned with the input beam and can be picked with a D-shaped mirror.
Based on the GDD introduced by the individual optical elements in the path of the 515 nm activation laser beam, a total GDD of the microscope was estimated, see

Table 2.1. Material and thickness of the individual Thorlabs elements (Thorlabs Inc., Newton, NJ, USA) is well documented on their respective website. The GDD of the remaining elements is based on published data or estimated to a reasonable extent. The total GDD estimate of the microscope results to $\sim$16 628 fs^2. The initial distance between the two prisms was calculated based on this number to be around 40 cm. Optimization of the prism distance was done with measurements of the activation rate of the dyes, see subsection 2.4.2.
For measurements investigating the effect of different pulse lengths on the activation efficiency, optical fibers (P1-405BPM, Thorlabs) with different lengths, 10 cm and 100 cm, were placed after the prism pre-compressor. The corresponding pulse broadening due to the silica core of the fiber is calculated based on Eq. 2.3, with a GVD of 68.706 fs^2/mm at 515 nm [53].

Table 2.1: GVD of elements in the activation beam pathway. The material of the components is listed, if known, and the GVD in fs^2/mm at a wavelength of 515 nm corresponding to it. Together with the thickness, the resulting GDD in fs^2 is calculated.

Element	Material	GVD in fs^2/mm	l in mm	GDD in fs^2
AC254-030	N-SF5/N-SF6	270.22/432.568	6	2108.364
LA1509	NBK7	84.347	3.6	303.65
AC100	NBK7	84.347		$\sim$200
AC245-200	N-K5/N-SF57	95.27/492.72	5/3	1954.51
3xTRH254-040	N-F2/N-K5	213.84/95.27	5/17	2688.79
Tube lens				$\sim$200
Pol-Plate	Quartz/MF2	78.545/24.792 [54]	1.57	123.79
Objective				$\sim$3876 [55]

All beams were scanned over the sample via two galvanometer mirrors (Cambridge Technology, United Kingdom) in a 4f-relay arrangement with $f = 40$ mm triplet lenses (TRH254-040, Thorlabs) and passed a quarter-wave-plate ($\lambda/4$) before the objective to convert linear to circular polarized light. For the imaging of cells, neurons and hepatocytes, a Leica 100x/1.4NA (HCX PL APO, Leica Microsystems GmbH, Wetzlar, Germany) oil objective lens was used. An Olympus 60x/1.35NA silicone oil objective was used for the imaging of tissue samples. The lateral positioning of the sample could be adjusted with a manual xy-stage, whereas focusing was achieved by moving the objective with a z-piezo (PI-725-CDD, Physik Instrumente, Karlsruhe, Germany).

The detection consists of a PMT (H10723-XX, Hamamatsu Photonics, Hamamatsu City, Japan) for the characterization of the PSF. For this, the reflected signal from the coverslide is deflected by a removable pellicle (BP145B1, ThorLabs) and focused on the PMT. An APD (SPCM-AQR-13, Excelitas Technologies, Waltham, MA, USA) is used for the confocal detection of the fluorescence signal. Two fluorescent filters, a 775 nm notch filter (Holographic NotchPlus Filter 775/10°, Kaiser Optical, USA) and a 670/40 beamsplitter (AHF Analysetechnik GmbH, Tübingen, Germany) are placed in front of the APD.
At a later stage, an active axial drift control was implemented. A 980 nm laser (Pegasus diode laser, Pegasus Lasersysteme, Wallenhorst, Germany) was coupled to the system in total internal reflection (TIRF) mode. The spot of the laser which was reflected at the boundary of the coverglass was imaged with the help of a CCD (DMK23UP1300, The Imaging Source, Bremen, Germany). The center of intensity of the reflection on the camera was calculated with LabVIEW and held at a constant position by adjusting the positioning of the objective lens. The correction signal was calculated with a proportional–integral–derivative (PID) controller in software.

All hardware components were controlled with a custom software written in LabVIEW 2019 (National Instruments, Austin, TX, USA). A field-programmable gate array (FPGA, PCIe-7852, National Instruments) is programmed to generate the scanning pattern and to output all electronic signals at the appropriate time. This assures high timing accuracy, independent of the operating system. The FPGA served furthermore as a photon-detection counter and as a generator of timing signals (pixel clock, line clock, etc.) for external appliances.
A host program is used to update the FPGA with new parameters and to read out the number of collected photons per pixel. For the gating of the detected signal, the setup includes a SPC-150 time-correlated single photon-counting module (Becker& Hickl, Berlin, Germany), which is synchronized to the system via the generated clock signals. The gating of the signal can be controlled with the host program.

2.4.2 Prism distance

Based on the estimation of the total GDD without pre-compensation of the pulses, the pulse duration would result to a minimum of 314 fs after passing the dispersive

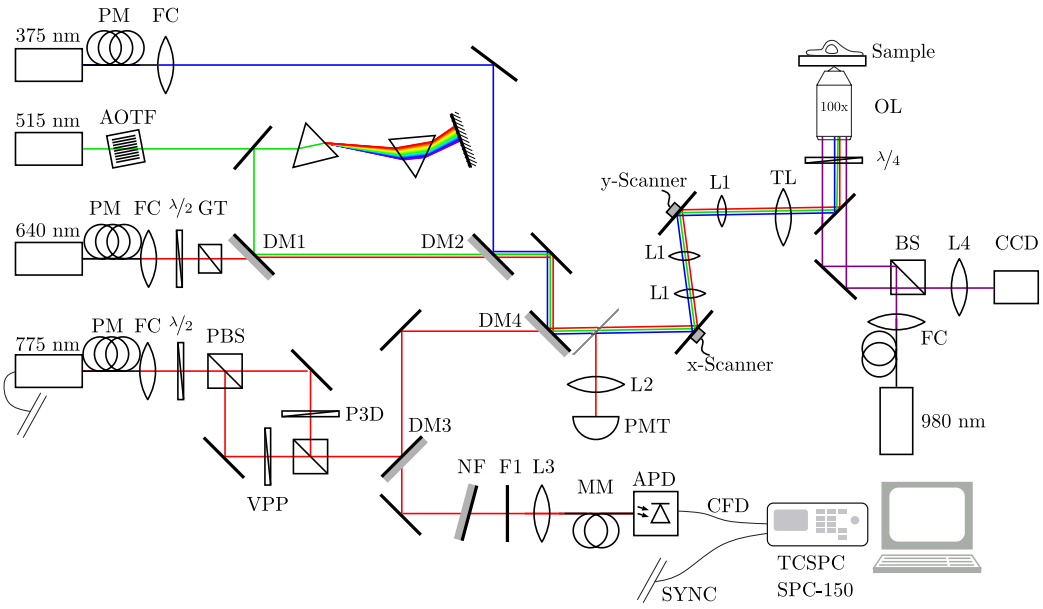

Figure 2.2: STED Microscope with one- and two-photon activation. The microscope is equipped with four laser sources for imaging and activation. The 375 nm laser beam is cleaned through a polarization maintaining single mode fiber (PM) and coupled into the system via a dichroic mirror (415RDC, DM2). The 515 nm beam is passing an AOTF for power control and subsequently a two prism compressor. The excitation laser at 640 nm is cleaned through a PM fiber and passes a half-waveplate and a glan-thompson prism (GT) for manual power control. The green beam is coupled to the red excitation beam with the use of a dichroic mirror (532RDC, DM1). The 775 nm STED laser is cleaned and passes a half-waveplate and a polarizing beam splitter (PBS) to control the laser intensity in the two arms. One arm is forming the 2D doughnut with the help of a vortex phase-plate (VPP) whereas the other arm is forming the 3D doughnut with the help of a custom-made 0-π phase plate. Both beams are combined and coupled into the microscope with the dichroic mirrors 750SPRDC (DM3) and 635RDC (DM4). All beams are scanned via two galvanometric mirrors arranged in a 4f-system with $f = 40$ mm triplet lenses (L1). FC: fiber collimator, PBS: polarizing beam splitter, TL: tube lens, NF: notch filter, F1: 670/40 filter, MM: multimode fiber, APD: avalanche photodiode, PMT: photo-multiplier tube, OL: objective lens, DM: dichroic mirror, CFD: constant fraction descriminator, TCSPC: time correlated single photon counting.

material. The reduction of the pulse length leads to a reduction of the average laser power needed for successful activation of the compounds. The positioning of the two prisms is optimized by observing the average laser power needed for successful activation.

The initial distance between both prisms was estimated to $\approx 40\,\mathrm{cm}$ using Eq. 2.4. The displacement is varied and the activation rate measurements described in subsection 2.4.3 is performed on fresh samples. The final distance between the prisms of 43.5 cm corresponds to the minimum of the average laser power needed for successful activation.

2.4.3 Activation rates

For determination of the activation- and bleaching-rates a special scanning sequence was implemented. The experimental sequence is depicted in Fig. 2.3 A. Initially, a reference confocal image is collected to measure the fluorescence signal present in the sample and to get a baseline signal of already thermally activated fluorophores. This is achieved by scanning the sample with the excitation beam (640 nm, $\sim 100\,\mathrm{ps}$, P = 8 mW, f = 40 MHz, dwell = 30 µs, "Frame 0"). Then, the compound is activated by scanning the same field of view (FOV) with activation light (515 nm, $\sim 200\,\mathrm{fs}$, P = 0 mW − 20 mW, f = 15 MHz, dwell = 100 µs, "Activation"), during which the detector is switched off. In the following scan, fluorescence images are detected by scanning the same FOV with the excitation beam (parameters as previously, "Frame 1"). The sequence of activation and subsequent read out is repeated 38 times for each activation power. For each power, a fresh area of the sample is used with a different confluency and therefore total fluorescence signal. To compare the results, the average fluorescence signal is calculated per pixel for pixels with signal above the noise level (excluding empty regions).
The registered dependency can be modeled by a 3-state system (Fig. 2.3 C), corresponding to the fluorophore in its unactivated N_u, fluorescent N_f and bleached N_b state respectively.

$$N_u \xrightarrow{k_{\mathrm{act}}} N_f \xrightarrow{k_{\mathrm{bl}}} N_b$$

$$\dot{N}_u(t) = -k_{act} \cdot N_u(t) \tag{2.5}$$

$$\dot{N}_f(t) = k_{\mathrm{act}} \cdot N_u(t) - k_{\mathrm{bl}} \cdot N_f(t) \tag{2.6}$$

$$\dot{N}_{\mathrm{bl}}(t) = k_{\mathrm{bl}} \cdot N_f(t) \tag{2.7}$$

The transition of the unactivated to the activated state can be described with Eq. 2.5 with k_{act} being the activation rate. The corresponding transition of the fluorescent state can be described with Eq. 2.6 with k_{bl} as the photobleaching rate of N_f. Therefore, the transition to the bleached state is described with Eq. 2.7.

We also assume that the detected signal is proportional to the amount of activated fluorophores and that the initial fluorescence before activation is minimal compared to the fluorescence detected after activation. The system of differential equations therefore is easily solvable and the obtained signal scales with

$$N_f(t) = \frac{k_{\text{act}}}{k_{\text{act}} - k_{\text{bl}}}(\exp(-k_{\text{bl}} \cdot t) - \exp(-k_{\text{act}} \cdot t)) \tag{2.8}$$

This function can be fitted to the experimental data to extract the parameters k_{act} and k_{bl}.

2.4.4 Bleaching measurements

For bleaching and confocal PALM measurements (see subsection 2.4.8) the microscope setup described in subsection 2.4.1 is equipped with an additional high-power cw laser diode at 636 nm (USHIO OTPO Semiconductors, INC. Type HL63391DG) where the laser output power can be regulated electronically.
Measurements to estimate the amount of photobleaching are performed in a similar way to the measurement to determine the activation rates described in subsection 2.4.3. After every activation frame the activated molecules are excited and the fluorescence is detected. The majority of fluorophores are then bleached using the laser diode with a varying laser power from 2 mW to 5 mW for a dwell time of 125 µs. The remaining fluorescence after bleaching is measured to get an estimate of the remaining active molecules.
Measurements for each activation power are performed on at least three different areas with the laser beam configurations as described in subsection 2.4.3 with a FOV of $(10\,\mu\text{m})^2$. The samples are flat homogeneously stained samples (either collagen samples or LaminB stained in U2OS cells).
To get an estimate of the total fluorescence achieved at a certain activation power level, the average fluorescence per pixel is calculated to average small differences in dye concentration over the sample. A background level is estimated from the measurements after bleaching, which is then subtracted from the remaining fluorescence measurements. The total number of activatable dyes is calculated by the integral of detected fluorescence after background correction.

2.4.5 STED image acquisition

Regular 2D STED images are acquired after a whole frame scan with either the 1PA or 2PA light source. They are taken in series and can feature different scanning parameters. Typically, pixel sizes of 16 nm to 20 nm with pixel dwell times of 100 µs to 130 µs are employed for STED imaging. Samples stained with ONB-2SiR and HCage 620 where activated using 515 nm with a power of 3 mW to 4 mW and a dwell time of 100 µs for pixels with a size of 150 nm. The 2PA laser power for images with pPA-SiR was set to 1 mW to 2 mW.
Confocal images were acquired before the STED images to avoid bleaching effects. This is the case for samples with ONB-2SiR or HCage 620. Confocal images for pPA-SiR are taken after the STED imaging due to the pixel-activation mode necessary for this dye, which is described in subsection 2.4.6.

2.4.6 Pixel-/line-activation

For protected STED [56] or pixel-to-pixel activation measurements all beams are overlapped on a gold bead sample before every measurement to ensure the best result.
The scan logic is adjusted such that at the beginning of a pixel measurement the 515 nm light can be switched on immediately for a variable duration. Because of the delay of the electric signal as well as the delayed response of the AOTF, the signal for the activation window needs to be adjusted carefully by introduction of a negative time offset measured with an oscilloscope. A user-selectable time-gap is bridging the time until the subsequent imaging of the pixel either in confocal or STED mode. The total dwell time of the imaging period can be adjusted accordingly. The APD is gated and is only registering photon counts during the selected detection window.
This is repeated for every pixel until the image is completed. A small schematic of the process is depicted in Fig. 2.7 B. Usual imaging parameters are: 0.6 mW of 515 nm, pixel size of 24 nm with 10 µs activation time. 775 nm STED with 90 mW for an imaging dwell time of 100 µs to 130 µs with a gap time in the low two digit µs regime.

For line activation measurements an entire line is scanned with activation light. Multiple line scans with excitation and STED light as well a detection can be

performed in the following. The amount of lines imaged after each activated line can be set by the user. This procedure is repeated until the entire frame is scanned.

2.4.7 3D measurements

In order to measure the bleaching induced by the 3D doughnut the following measurement scheme was applied. In a sample (U2OS immunolabeled for microtubules with HCage 620) three different layers with equal spacing in z are of interest. Two of the three layers (middle and lower) are activated at the beginning of the measurement either with 1PA or 2PA. The spacing of the three layers at 600 nm is chosen such that it is bigger than the axial size of the two photon volume but small enough to fall into the high power lobes of the 3D doughnut applied. After the initial activation, the central layer is imaged with excitation light together with a 3D-doughnut with 200 mW. After the STED imaging scan the remaining fluorescence signal is detected with only the excitation laser in all three different layers. The last two steps are repeated 7 times until the fluorescence signal is reduced significantly. Finally, the last of the three layers (top) which wasn't activated initially is activated using either 1PA or 2PA and the fluorescence signal is detected. The fluoresce intensity of each layer is normalized to the initial max value at the start of the bleaching series for comparison.

For 3D STED images of LaminB in neural stem cell colonies, different 2D images were acquired with a 3D STED doughnut with a total power of 140 mW where 40 mW are in the 2D-doughnut (2D FWHM $\sim$50 nm). The different layers are separated by a distance of 200 nm. Each layer was activated individually with the corresponding activation light before imaging. After complete acquisition of the volumetric stack, the images are further processed using Fiji [57] to reconstruct a three dimensional view and maximum signal projections along the optical axis. The average fluorescence per pixel above noise level was calculated for every individual 2D slice.

2.4.8 Confocal PALM-like imaging

For single molecule localization microscopy (SMLM) measurements small regions with a FOV of $(10\,\mu m)^2$ are scanned with a pixel size of 78 nm. This ensures at least three pixels per detection FWHM in accordance with the Nyquist criterion.

To avoid broken or incomplete PSFs, multiple fast imaging scans of the same area (typically 5 to 8 frames with a pixel dwell time of 10 µs) are performed following the activation scan of one entire frame. They are subsequently combined before they are used for post-processing. All fluorophores that remain active after the read-out are bleached with the 636 nm laser diode before the next activation step. The activation power is chosen for each dye such that only a small subset of molecules are active at the same time within the acquisition region.

The stack of binned images is further analyzed with a multi-emitter 2D fitting algorithm DAOSTORM [58] (maximum number of iterations: 30) written in Python. The super-resolved image is reconstructed by imposing Gaussians with a variable σ corresponding to the localization precision at the center point of the fit. For comparison confocal images are generated by summation of the counts over all detected frames.
Fourier ring correlation (FRC) maps area calculated with the Nanoj-SQUIRREL plugin for Fiji [59].

2.4.9 MINFLUX, MINSTED and confocal laser scanning microscopy

MINFLUX measurements were performed on a yet unpublished setup which is based on the original MINFLUX publication [12]. For 2PA activation the 515 nm laser from subsection 2.4.1 was placed in the corresponding setup. MINFLUX images were acquired with an iterative MINFLUX scheme. A Gaussian and doughnut beam at 647 nm with 80 µW and 350 µW respectively were employed for the localization with 5 steps and an scan diameter L of 100 nm in the last iteration. The position of the active fluorophore was estimated and the scan centered on it before moving to the next iteration. The last iteration step was concluded after 10000 recorded photons at most. A total FOV of (1.5 µm to 6 µm)^2 was scanned with a photon countrate detection threshold of 5 kHz to 10 kHz. Typically, the activation laser power for 2PA was set to 30 µW to 60 µW and 10 nW for 1PA at 405 nm.
Localizations were rendered after post-processing of the localization traces with tools from [60].

MINSTED images were acquired on a previously published setup [13] with the addition of the 2PA laser. Images were acquired with a final STED FWHM of

50 nm in a FOV of $(1.2\,\mu m)^2$. The scanning period was set to 60 µs. The photon threshold was set to 8 for the start of a localization sequence and was terminated if less than 16 counts were detected in 30 ms period.
For visualization the localizations are rendered with their corresponding localization precision as the σ of a Gaussian profile. Only localization events with photon counts above 200 and localizations where the standard deviation of their center positions after reaching the maximal STED power were below 20 nm were chosen as valid localizations (see [13]).
Confocal overview scans of the MINSTED area were acquired with the EODs in a FOV of $(1.3\,\mu m)^2$ with a pixel size of 13 nm and a dwell time of 20 µs. Bigger overview images of whole cells were taken with the galvo-scanners with a varying FOV of $(60\,\mu m$ to $80\,\mu m)^2$ with a pixel size of 60 nm to 80 nm and a dwell time of 20 µs.

For large confocal overview scans of brain tissue sections, a TCS SP8 confocal laser scanning microscope (Leica Microsystems, Wetzlar, Germany) was used. Cross sectional confocal images of tissue sections were acquired using a previously reported STED system [61] with increased axial scanning range.

2.4.10 Gold and fluorescent microspheres

For daily calibration of the excitation, activation and STED beam, gold bead samples were prepared with 50 nm colloidal gold (Aldrich, St. Louis, Missouri, United States). For precise control of the spatial position of the STED beam in relation to the excitation as well as the accurate timing of the STED beam a reference sample with fluorescent microspheres (FluoSpheres®, 0.02 µm, dark red fluorescent; Thermo Fischer Scientific, Waltham, MA, USA) was prepared.
Coverslips were cleaned with MiliQ water and acetone. Afterwards, they were coated with Poly-L-Lysine (100 µl, 0.01 % Poly-L-Lysine solution, Sigma-Aldrich, St. Louis, MO, USA) for 5 min. The gold bead/fluorescent microsphere solution was sonicated for 15 min. Coverslips were incubated with 100 µl of the corresponding solution. After an incubation time of 5 min the coverslip was washed with Mili-Q water and mounted on the coverslide with either PBS or immersion oil. The sample was sealed using nail polish.

2.4.11 Animals

Described animal procedures were conducted in accordance with Directive 2010/63/EU of the European Parliament and the Council on the protection of animals used for scientific purposes as well as with the German Animal Protection Law (Tierschutzgesetz der Bundesrepublik Deutschland, TierSchG).
All mice were housed with a 12 hours light/dark circadian cycle and ad libitum access to food and water.

2.4.12 Cells

U2OS cells (ECACC, cat. 92022711) were cultured in McCoy's 5A (modified) medium (Thermo Fisher, cat. 16600082) supplemented with 10 % FBS (Bio&Sell, cat. S0615), 1 % Sodium Pyruvate (Sigma, cat. S8636) and 1 % Penicillin-Streptomycin (Sigma, cat. P0781) in a humidified 5 % CO_2 incubator at 37 °C. Cells were seeded on coverslips 24 h before fixation.

The protocol of hepatocyte isolation and culture was modified from [62]. Briefly, an adult female mice of CD1 background were sacrificed and immediately perfused via the hepatic portal vein with EGTA solution (pH 7.4) followed by collagenase solution. The perfused liver was transferred into a pre-warmed collagenase solution, mechanically dissociated and subsequently filtered through a 100 µm cell strainer. Cells were plated on coverslips pre-coated with collagen type I (Corning, cat. 354236) and cultivated in William's Medium E (Gibco, cat. A12176-01) supplemented with 4 % Cell Maintenance Cocktail-B (Gibco, cat. A13448) and 10 µM Dexamethasone (Gibco, cat. A13449). Cultures were kept in a humidified 10 % CO_2 incubator at 37 °C.

Neural stem/progenitor cells were cultured as previously described in [63]. In brief, cortexes were dissected from C57BL/6N mouse embryos at embryonic day E12.5, digested by 0.25 % trypsin (Gibco, cat. 15090046) for 20 min at 37 °C and subsequently mechanically dissociated with a fire-polished, culture-medium-coated Pasteur pipette. Cells were plated on coverslips pre-coated with 0.2 % gelatin in PBS and cultivated in KnockOut DMEM/F-12 medium (Gibco, cat. 12660012) supplemented with 2 % StemPro Neural Supplement (Gibco, cat. 10508-01), 1x GlutaMax (Gibco, cat. 35050061), 2 µg/µl Recombinant Human Basic Fibroblast

Growth Factor (Gibco, cat. PHG0024), 2 µg/µl Recombinant Human Epidermal Growth Factor (Gibco, cat. PHG0314) and 0.5x Antibiotic-Antimycotic (Gibco, cat. 15240062) in a humidified 10 % CO_2 incubator at 37 °C.

The protocol for hippocampal neurons isolation has been described previously in [64]. Briefly, hippocampi were isolated from C57BL/6N mice of mixed gender at postnatal day P0-P1, digested with 0.25 % trypsin (Gibco, cat. 15090046) for 25 min at 37 °C and subsequently manually dissociated. Cells were plated on coverslips pre-coated with poly-L-ornithine (100 µg/ml; Sigma-Aldrich, cat. P3655) and laminin (1 µg/ml; BD Bioscience, cat. 354232) and cultured in Neurobasal medium (Gibco, cat. 21103049) supplemented with 2 % B27 serum-free supplement (Gibco, cat. 17504044), 1x GlutaMax (Gibco, cat. 35050061) and 0.5 % Penicillin-Streptomycin (Gibco, cat. 15140122). Cells were kept in a humidified 5 % CO_2 incubator at 37 °C.

For imaging nuclear pore complexes we used U2OS cells which express the nuclear pore protein NUP96 endogenously tagged with HaloTag (U-2OS-CRISPR-NUP96-Halo clone #252 (330448, CLS GmbH) [65]). They were cultured on coverslips in McCoy's medium (Thermo Fisher Scientific, 16600082) which was supplemented with 10 % [v/v] fetal bovine serum (Bio&SELL, S0615), 1 % [v/v] non-essential amino acid solution (Thermo Fisher Scientific, 11140050) and 1 % [v/v] penicillin-streptomycin (Sigma Aldrich, P0781). The cells were fixed with 2.4 % [w/v] paraformaldehyde in PBS for 15 min. The PFA autofluorescence was quenched by incubating the sample in 100 mM NH_4Cl in PBS for 5 min. We permeabilized with 0.5 % [v/v] Triton X-100 in PBS. The sample was washed with PBS between the steps.

2.4.13 Stereotaxic Injections and slice preparation

The protocol of stereotaxic injection has been described previously in [61, 66]. Briefly, an adult mouse of C57BL/6J background was anaesthetised by 1.0 %-2.0 % isoflurane (Forene, Abbvie) in oxygen-enriched air (47.5 % oxygen, 50 % nitrogen, and 2.5 % carbon dioxide; Westfalen AG, Münster, Germany) and fixed into a stereotaxic apparatus (SG-4N, Narishige, Tokyo, Japan). The scalp was incised and a craniotomy was performed on the parietal bone above the visual cortex of the left hemisphere. A prepulled, tapered, borosilicate glass injection capillary

(World Precision Instruments, Sarasota, FL, USA) was filled with a solution of pAAV-hSyn-Lifeact-EYFP virus [67] diluted 1:5 in sterile artificial cerebrospinal fluid (ACSF; NaCl 126 mM, KCl 2.5 mM, $CaCl_2$ 2.5 mM, MgCl 1.3 mM, HEPES 27 mM, glucose 30 mM; pH 7.4). The capillary was subsequently lowered ca. 500 µm into the brain with a vertical angle of 20° and the volume of 250 nl-500 nl of virus-containing solution was injected using a pressure system (30 ms pulses delivered with 20 ψ on manual command, TooheySpritzer, Toohey Company).
The scalp was surgically closed by polyamide surgical suture (6.0/697H, Ethilon) immediately after retraction of the capillary and the animal was allowed to recover. During the surgery perioperative analgesia was achieved by subcutaneous injection of Buprenovet (0.1 mg/kg mouse body weight, Bayer) and incision sites were analgised by subcutaneous injection of 2 % Lidocaine (Xylocaine, AstraZeneca). Through the whole surgery eyes were protected from dehydration by application of ointment (Bepanthen, Bayer) and a custom build heating plate was used to maintain mouse body temperature. Analgesic and anti-inflammatory post-surgical care was achieved by subcutaneous administration of Carprofen (5 mg/kg, Zoetis) every 24 h for a minimal period of 2 days following surgery.

Intracardial perfusion took place 3–4 weeks after stereotaxic viral injection. The mouse was anaesthetised by intraperitoneal injection of an overdose of Ketamine (Medistar)/Xylasine (Dechra) and subsequently transcardially perfused with PBS, followed by 4 % [w/v] PFA in PBS. The brain was dissected and post-fixed by an overnight incubation in 4 % PFA/PBS at 4 °C. Afterwards the fixative was removed, the brain was transferred into PBS and sliced into 60 µm thick consecutive coronal sections using a vibratome (VT1200S, Leica).

2.4.14 Immunohistochemistry

Fixed slices were successively washed 2 times with Tris Buffer (TB), Tris Buffer Saline (TBS) and Tris Buffer Saline containing 0.5 % [v/v] Triton-X 100 (TBST) (all with pH 7.6) for 15 min each at room temperature (RT). The slices were subsequently blocked with 10 % [w/v] normal goat serum and 0.25 % [w/v] bovine serum albumin (BSA) in TBST (blocking solution) for 1.5 h at RT. Sections were incubated for 72 h at 4 °C on a rocking plate with primary antibodies: (I) rabbit anti-GFP (Abcam, cat. 6556) diluted 1:300 and (II) rabbit anti-Tubulin (Abcam, cat. 18251) diluted 1:100 in blocking solution. Afterwards, the slices were rinsed

4 times with TBST for 15 min each at RT and incubated with HCage 620 [21] conjugated goat anti-rabbit secondary antibodies diluted 1:10 in TBST overnight at 4 °C on a rocking plate. Slices were successively washed 2 times with TBST, TBS and TB before being mounted with PBS.

U2OS cells were fixed with ice-cold methanol for 5 min at −20 °C and subsequently blocked with 2 % BSA in PBS (blocking solution) for 10 min at RT. Cells were then incubated for 1 h at RT with primary rabbit anti-Tubulin antibody (Abcam, cat. 18251) diluted 1:100 in blocking solution. Afterwards the coverslips were washed with blocking solution for 10 min at RT and incubated for 1 h at RT with secondary anti-rabbit antibodies conjugated with: (I) ONB-2SIR [51], (II) HCage 620 [21] and (III) pPA-SiR [23] diluted in blocking solution. Cells were then washed 2 times with PBS and subsequently mounted with PBS.

Primary hepatocytes at the 15^{th} day of in vitro (DIV) culture were fixed with ice-cold methanol for 5 min at −20 °C and subsequently blocked with 10 % BSA in PBS at RT. Afterwards, cells were incubated with primary rabbit anti-actin antibodies (Sigma, cat. A2668) diluted 1:100 in 2 % BSA in PBS. Cells were then successively washed with 0.5 % Triton-X in PBS and 2 % BSA in PBS for 5 min each at RT and subsequently incubated with secondary anti-rabbit pPA-SiR conjugated antibodies diluted 1:20 in 2 % [v/v] BSA in PBS for 1 h at RT. Coverslips were washed with PBS 2 times 15 min each at RT and subsequently mounted with PBS.

Neural stem/progenitor colonies at the 5^{th} DIV were fixed with ice-cold methanol for 10 min at −20 °C and subsequently blocked with 2 % BSA in PBS (blocking solution) for 10 min at RT. Cells were then permeabilized with 0.5 % Triton in PBS for 10 min at RT and incubated with primary rabbit anti-LaminB1 (Abcam, cat. 16048) antibody for 1 h at RT. Afterwards, colonies were washed with blocking solution 2 times 10 min each and subsequently incubated for 1 h at RT with secondary anti-rabbit HCage 620 conjugated antibody diluted 1:20 in blocking solution. Coverslips were then washed with PBS 2 times 15 min each at RT and subsequently mounted with PBS.

Primary hippocampal neurons at the 7^{th} DIV were washed with PBS and subsequently fixed with 4 % PFA in PBS for 10 min at RT. Cells were then quenched with 10 mM NH_4Cl in PBS and permeabilized with 0.1 % [v/v] Triton-X in PBS

for 5 min at RT each. Afterwards cells were blocked with 1 % [w/v] BSA in PBS for 30 min at RT and incubated with primary anti-β-Spectrin II mouse antibody diluted 1:200 in PBS for 1 h at RT. Specimens were then washed 2 times with 0.1 % Triton-X in PBS for 5 min at RT each and then incubated with secondary anti-mouse pPA-SiR conjugated antibodies diluted 1:10 in PBS. Coverslips were subsequently washed with PBS 2 times 15 min each at RT and subsequently mounted with PBS.

For MINSTED measurements U2OS cells were washed with PBS and subsequently fixed with 8 % [w/v] PFA in PBS for 5 min at 37 °C. Afterwards cells were permeabilized with 0.5 % [v/v] Triton-X in PBS for 5 min at RT and subsequently blocked with 2 % BSA in PBS (blocking solution) for 10 min at RT. Cells were then incubated with primary rabbit anti-Caveolin-1 (Cell Signaling) antibody diluted 1:200 in blocking solution for 1 h at RT. After that cells were washed with blocking solution and incubated for 1 h at RT with secondary anti-rabbit antibodies conjugated with: (I) ONB-2SiR and (II) HCage 620 diluted in blocking solution. Cells were subsequently washed with the blocking solution and counterstained for 1 h at RT with tertiary anti-goat antibody conjugated with Alexa 647 diluted 1:500 (Invitrogen) in blocking solution. Cells were subsequently washed with PBS.

2.4.15 Sample preparation

For simple confocal or STED measurements, samples were mounted directly on the coverslide with PBS pH 7.4 as a buffer using picodent twinsil® (picodent®Dental-Produktions- und Vertriebs-GmbH, Wipperfürth, Germany) or two component epoxy resin.
For MINSTED measurements, samples were incubated with diluted silica shelled silver nanoplates for 20 min (nanoComposix, SPSH1064-1M; 1 µg/ml in PBS). After washing the sample with PBS pH 7.4 twice, they were mounted in PBS and sealed using nail polish.
For MINFLUX measurements, gold nanorods (A12-25-980-CTAB-DIH-1-25, Nanopartz Inc., Loveland, CO, USA) were diluted with twice the volume PBS and sonicated for 15 min. In the following the sample was incubated for 15 min at room temperature with the nanorod dilution. After washing with PBS for three times, the cell-sample was mounted on a single-well microscope slide which was filled with

a buffer (either PBS pH 7.4 or 20 mM HEPES pH 7, 150 mM NaCl). The sample was then sealed using twinsil.

For experiments investigating the influence of buffer systems on the behavior of the dyes, a standard STORM buffer was prepared [68]. The buffer consists of 0.4 mg/ml glucose oxidase (G2133, Sigma Aldrich, St. Louis, MO, USA), 64 µg/ml catalase (C100-50MG, Sigma Aldrich), 50 mM TRIS/HCl pH 8.0/8.5, 10 mM NaCl, 10 mM to 200 mM cysteamine hydrochloride (MEA) (M6500, Sigma Aldrich) and 10 % [w/v] glucose [69].

2.5 Results

After the initial screening of possible candidates, three different silicon rhodamine (SiR) based dyes remain: ONB-2SiR [51], HCage 620 [21] and pPA-SiR [23] (see Table 2.2). All of them showed a strong increase in fluorescence after illumination with 515 nm laser light. Two of the three dyes, ONB-2SiR and HCage 620, exhibit two caging groups which are cleaved after the typical irradiation with UV light (<400 nm). This leads to a change in conformation and to its excitable state (see Fig. A.1 A). The activation of the third dye, pPA-SiR, is mechanistically different as it doesn't contain caging groups but is activated through light induced protonation (see Fig. A.1 A). Detailed mechanistic descriptions of the activation process can be obtained from the corresponding literature [21, 23, 51].
NHS esters of all three dyes can be coupled to an antibody of choice. Two of the three dyes, HCage 620 and pPA-SiR, also exist as HaloTag or SNAP-Tag conjugates for their use with these self-labeling proteins.
All suitable fluorophores show, in their unmodified state, a similar spectrum with no absorption at 515 nm before and minimal absorption after activation (spectrum of ONB2-SiR in Fig. A.1 B). Therefore, the activation wavelength is located between their absorption maximum for the non-activated compound in the UV range and the excitation maxima of the activated fluorophore above ∼550 nm.
The spectrum of pPA-SiR however seems to be strongly influenced by the applied labeling mechanism. Versions of pPA-SiR with HaloTag and SNAP-tag show a much higher absorption cross-section at 515 nm before and after activation compared to the NHS ester variant. This leads to a significantly higher bleaching during activation, making these variants unusable for 2PA at this wavelength (see Table 2.2).

Other dyes that were tested, like Abberior CAGE 635, show some fluorescence increase after illumination with 515 nm, but the activation is not efficient enough to use this fluorophore further in the 2PA experiments. PA Janelia Fluor 646 [70] shows no increase in fluorescence at any power level of the 2PA laser. On the contrary, PaX560 [71] shows successful activation with the two-photon femtosecond laser at 515 nm. However, illumination of this dye with a regular cw laser at 514 nm (see Table 2.2) also shows a similar increase in detectable fluorescence signal, pointing to a linear 1PA at this wavelength.

Table 2.2: List of dyes tested for activation with 515 nm together with their different labeling conjugates. If activation with the femtosecond laser at 515 nm was successful for the given compound the cell is marked with *yes*. Otherwise, if the activation was unsuccessful the cell is marked with *no*. HaloTag and SNAP-tag conjugates marked with - either don't exist or weren't tested if the activation of the NHS ester variant was unsuccessful.

Fluorophore	NHS ester	HaloTag	SNAP-tag
ONB-2SiR [51]	yes	-	-
HCage 620 [21]	yes	yes	yes
pPA-SiR [23]	yes	no	no
Abberior Cage 635	yes (weak activation)	-	-
PA Janelia 646 [70]	no	-	-
PaX 560 [71]	yes (but not 2PA)	-	-

2.5.1 Activation-rate scaling

The successful first activation of the three individual dyes with the wavelength of 515 nm during the initial screening process indicates the presence of a multi-photon activation process. Next, the order of this multi-photon activation process needed to be determined.
For this reason we imaged microtubules in U2OS cells that were immunostained with one of the selected dyes respectively (see subsection 2.4.14). The measurement procedure described in subsection 2.4.3 was applied and is depicted in Fig. 2.3 A. Typically, the initial signal of the sample before activation was negligible for all the measurements. An example of microtubules stained with ONB2-SiR before activation is shown in Fig. 2.3 A (lower left). The signal before activation is particularly low upon comparison to the signal obtained after activation ($\sim 1\,\%$). An image of the same area is shown in Fig. 2.3 A (lower right) with the same color range and same excitation power as before the activation. The very low count of already activated fluorophores in the beginning of the measurement is in good agreement with the assumptions made for solving the differential equations in subsection 2.4.3.

An example of the experimental data obtained from measurements with ONB-2SiR can be seen in Fig. 2.3 B. The measured average fluorescence signal per pixel above background level is shown here, depicted as filled circles, as a function of the frame number in which it was acquired.
The initial activation scans increase the detectable fluorescence in the sample con-

tinuously as only a fraction of the total molecules present is activated during a single scan with 515 nm. This is especially true for low activation power levels in the range of 1 mW to 5 mW. Upon reaching a maximum, photobleaching of already activated molecules becomes dominant and the signal begins to decline with further exposure to the activation light. For a higher activation laser power >5 mW the maximal fluorescence intensity is already obtained after the first activation scan and additional activation scans result only in a decrease in fluorescence.
Repeated scanning with only the excitation laser, as a control, results in no increase or decrease of the detected fluorescence signal. The detectable activation and bleaching of molecules is therefore only a result of the activation laser.

To extract the activation rate from such a measurement, the expression in Eq. 2.8, based on the three state system shown in Fig. 2.3 C, is fitted (solid and dashed lines in Fig. 2.3 B) to the experimental data. The model fits very well to the obtained data-points, especially at low to moderate activation power levels. For high activation power the fit doesn't converge for the two dyes ONB-2SiR and pPA-SiR as can be seen by the last data-points in Fig. 2.3 D, F. The rates can be gathered directly from the fit and are visualized according to the used activation power in Fig. 2.3 D-F (squares: activation rate, circles: bleaching rate) in a log-log plot. The activation rate scales with the laser power as $k_{\mathrm{act}} \sim P_{\mathrm{act}}^{n}$, where n can be extracted directly from a linear fit.
For ONB-2SiR the activation order results in $n = 1.96 \pm 0.16$. This nearly quadratic scaling of the activation rate based on the applied laser power is consistent with a two-photon process. The bleaching rate is five-folds smaller and shows a fractional power dependency $k_{\mathrm{bl}} \sim P_{\mathrm{act}}^{1.54 \pm 0.12}$, suggesting a mixture of a linear and higher-order processes. Following the same procedure, similar $n > 1$ activation rate dependencies are observed for the two other dyes HCage 620 with $n = 2.95 \pm 0.03$ and pPA-SiR $n = 2.48 \pm 0.13$. A slightly higher ($n > 2$) activation rate scaling for pPA-SiR is still consistent with a two-photon process. HCage 620 however, shows a clear higher-order dependency of the activation rate on the average activation laser power which is indicative for a multi-photon process (higher than 2PA). They also show both a lower power dependency of their corresponding bleaching rates with $n = 1.00 \pm 0.05$ for HCage 620 and a fractional dependency of $n = 1.66 \pm 0.84$ for pPA-SiR.
The average 515 nm laser power needed for successful activation of pPA-SiR is notably lower (<1 mW) than that needed for ONB-2SiR and HCage 620 (>1 mW).

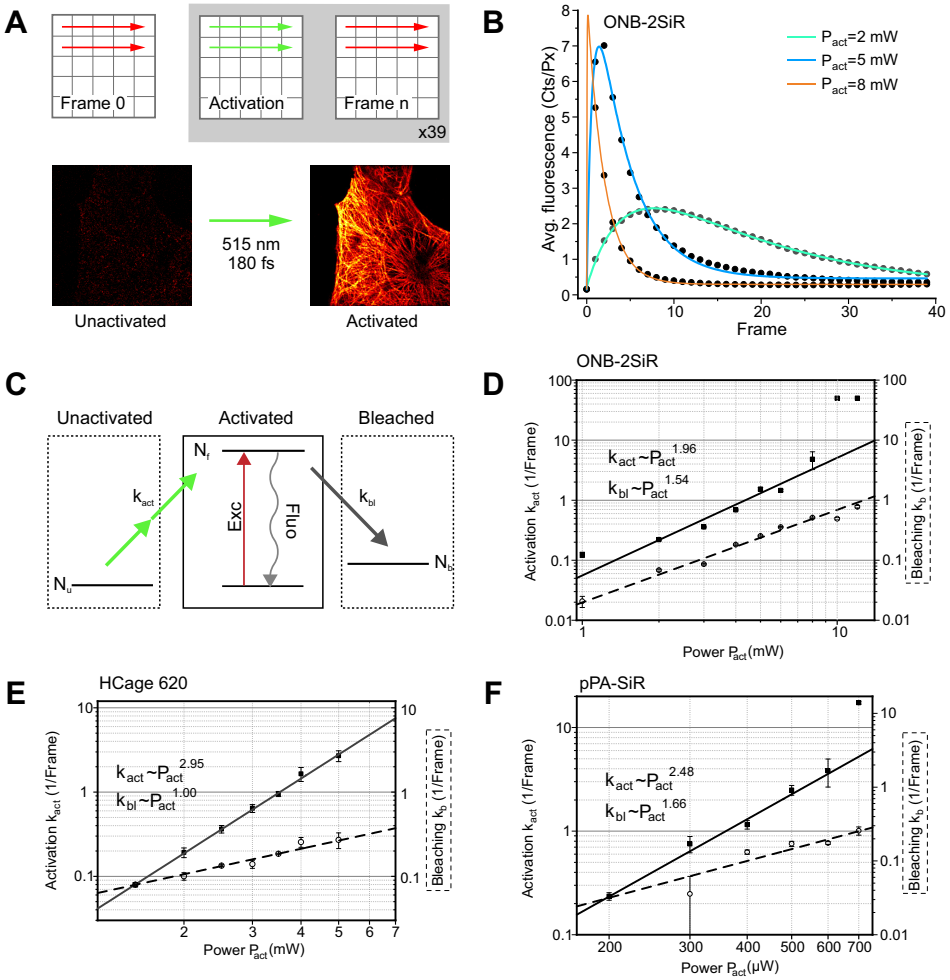

Figure 2.3: Activation rate measurements of the fluorophores reveal the two-photon process. A Activation rate measurements as described in subsection 2.4.3. The first frame is obtained as a baseline reference. The following procedure of successive activation and fluorescence readout is repeated 38 times. An example of a U2OS cell with microtubules immunostained with ONB-2SiR before and after 2PA activation is shown below. **B** Examples of experimental data of activation rate measurements (points) of ONB-2SiR with their fits (lines) for three different activation power levels. **C** Three state system of the fluorophore in its unactivated state N_u, fluorescent state N_f and bleached state N_b with the activation rate k_{act} and bleaching rate k_{bl}. **D-F** Activation (squares) and bleaching (circles) rates extracted from the corresponding fits for the three different dyes ONB-2SiR, HCage 620 and pPA-SiR in a log-log plot as a function of the 515 nm activation laser power. The order of the process is determined from their linear fits (solid line for activation, dotted line for bleaching).

It was noticeable that the behavior of some dyes changed over time, starting some hours after the mounting of a sample. To achieve a consistent fluorescent signal after activation, a significant reduction in the required 2PA power was needed. For time spans longer than 5 h after mounting the sample, the required activation power was reduced significantly (from 4 mW to 2 mW) for the two dyes ONB-2SiR and HCage 620. Changing the imaging buffer or introducing oxygen scavenging systems had no effect. This behavior is possibly due to a partial thermal uncaging where the fluorophore is losing one of the two caging groups without illumination of the activation light. Due to this reason special care was taken to image freshly mounted samples for those two dyes. No change of required activation power over time for the pPA-SiR was detected. Additionally, the activation laser power was randomly selected to avoid any time dependent effects.

With pulse lengths of 200 fs for the 515 nm laser the activation rate of HCage 620 shows a dependency very close to $n = 3$. To check whether this is in fact a three-photon activation process the activation rates are measured with different pulse lengths of the activation laser. For this purpose pulses are stretched to different lengths by introducing different optical fibers (10 cm and 100 cm) into the beam path as described in subsection 2.4.1. The scheme for measuring the activation rates remains unchanged.
The experimental data shows a reduced activation rate of $1/(2.01 \pm 0.28)$ from measurements with no fiber compared to ones with the short fiber of length 10 cm. By replacing the short fiber with the fiber of length 100 cm, the activation rate gets reduced by another factor of 4.33 ± 0.94, see Table 2.3. Similar measurements are performed with ONB-2SiR, the dye which shows the clearest two-photon behavior based on the activation rate measurements. For this fluorophore the data shows a reduced activation rate of 1.79 ± 0.15 from no fiber to the short fiber of length 10 cm and of 3.17 ± 0.29 to the 100 cm fiber. The comparison of the activation rate changes of both dyes shows a similar trend with comparable values for both pulse length increases.
The theoretical pulse stretching can now be calculated with Eq. 2.3 from the estimated total GDD of the microscope and an additional GVD of 68.706 fs^2/mm as a result from the introduced fibers of different length. This results in an increase of the original pulse length of 200 fs by a factor of 1.911 to $\sim$ 382 fs after introduction of the 10 cm fiber in the optical path. The replacement with the longest fiber (100 cm) is increasing the pulse length additionally by another factor of 3.14

Table 2.3: Scaling of the activation rates with the pulse length. Activation rates (in 1/frame) of HCage 620 and ONB-2SiR with fibers of different lengths placed in the beam path of the 515 nm activation laser. Measurements are performed as described before. The ratios in the lower table show the ratio of two n measured with two fibers of different length.

	k_{act}, 0 cm fiber	k_{act}, 10 cm fiber	k_{act}, 100 cm fiber
HCage 620	1.24 ± 0.16	0.62 ± 0.01	0.12 ± 0.01
ONB-2SiR	1.33 ± 0.08	0.72 ± 0.02	0.23 ± 0.02

	0 cm/10 cm	10 cm/100 cm
HCage 620	2.01 ± 0.28	4.33 ± 0.94
ONB-2SiR	1.79 ± 0.15	3.17 ± 0.29

to 1200 fs.

Comparing the results of the theoretical pulse prolongation to the measured changes of the activation rate of both dyes they appear to follow a similar trend. For ONB-2SiR the activation rate scales almost exactly with the inverse of the calculated pulse-length $1/\tau_{theo}$. This is the expected outcome for a two-photon dependent process. The changes of the activation rate are slightly higher for HCage 620. The measured data in Table 2.3 however does not show a dependency of $1/\tau_{theo}^2$, which would be the expected behavior for a three-photon activation process.

As a result we can conclude that the activation of HCage 620 is, despite showing a activation rate dependency of $n = 3$, likely not a three-photon process. Based on the experimental data it is much closer to a 2PA behavior. This leaves the possibility of a mixed process with a higher-order and dominant 2PA or a 2PA combined with an additional step.

2.5.2 Photobleaching estimates

Flat samples with a homogeneous dye distribution are prepared by either embedding the dyes in a collagen matrix or labeling the membrane of the nucleus (see subsection 2.4.14). The samples of the nuclear envelope are chosen because the reported thickness of ~ 30 nm [72] is small enough to eliminate effects of different activation volumes between 2PA and 1PA. These samples are used to estimate the amount of photobleaching of the 2PA on the individual fluorophores. The comparison of the total detectable fluorescence after 2PA (at usual imaging activation power for each dye) to the one after 1PA shows a significantly higher signal in

the 1PA case. 2PA imaging parameters are chosen for the three dyes in a way that delivers the highest fluorescence signal after one frame activation (515 nm power of 4 mW for ONB-2SiR and HCage 620 and 1 mW for pPA-SiR, remaining parameters described in subsection 2.4.5). For plasma membrane samples stained with ONB-2SiR the signal after 1PA is 4.0 ± 1.1 times higher compared to 2PA. Similarly for HCage 620 embedded in collagen samples which are spread flat on a coverglass the signal after 1PA is 4.3 ± 0.3 times higher than after 2PA. Activation measurements on flat cell areas with pPA-SiR show a higher signal after 1PA as well with a ratio of 3.3 ± 1.6 compared to 2PA.
The measurement procedure to asses the activation rates as described earlier, together with its model, can, by design, only include molecules which reside in their fluorescent state during the detection scan. Molecules that are bleached before this scan are never detected and accounted for. Similarly, the amount of active fluorophores that are added after each activation step are observed together with the amount that remains from the previous activation. Both individual quantities are unknown in this measurement. To get a better understanding of the number of newly added fluorescent molecules after every activation step a different measurement procedure described in subsection 2.4.4 is employed. The sample is repeatedly activated, but after every detection of the fluorescence almost all active fluorophores are bleached by a high intensity laser exposure.

Two typical results for lower and higher activation power are shown in Fig. 2.4 where the amount of fluorescence per pixel is plotted against the frame number. Every alternating frame shows the remaining fluorescence after bleaching. This is used in turn for a baseline correction of the remaining data. The overall curve looks similar to the data obtained in subsection 2.5.1, but each point after activation here reflects only the new fluorescence signal which was added. This allows for a better estimation of the expected number of activated fluorophores after a single 2PA activation scan at different power levels. The area under the curve after background subtraction equals to the total number of molecules that could be activated with this light dose. The first measurement point in relation to the total sum gives an estimate of how much of the achievable fluorescence is gained in the first activation for a given imaging configuration.
For HCage 620 a total of (43 ± 6) % is expected to be active after the first frame with an average 515 nm power of 4 mW (other parameters as described in subsection 2.4.4). If we compare this to the measurements on flat samples, the 2PA

resulted in $\sim 23\,\%$ of the total fluorescence that was detectable after 1PA. Based on the estimate of the first frame activation we estimate a total bleaching of roughly $\sim 47\,\%$ for these specific imaging conditions compared to 1PA. At the same imaging configuration the first frame of 2PA for ONB-2SiR results in a subset of $33\,\%$ of active molecules. This results in a 2PA bleaching of $(25 \pm 10)\,\%$ compared to 1PA. For pPA-SiR with an average power of 0.6 mW $(39 \pm 4)\,\%$ of all two-photon activatable fluorophores are in their active state. This results in a bleaching compared to 1PA to $(22 \pm 4)\,\%$.

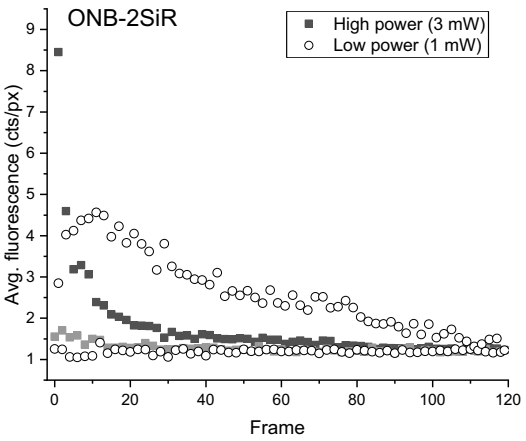

Figure 2.4: 515 nm power dependent photobleaching. Photobleaching measurements for two different 515 nm power levels. Measurements on Lamin-B with ONB-2SiR are shown with the average fluorescence per pixel as a function of the acquisition frame. Low power activation laser measurements shown with circle and higher power with filled squares. The alternating measurements in dark grey show the fluorescence before and in light grey after bleaching of the activated molecules.

The bleaching of molecules is changing with the applied light dose. The higher the used 2PA activation power the higher the bleaching. This is already visible in Fig. 2.4 where the total area under the curve for the low power measurement is bigger than the one for higher average activation power.
To test this, the bleaching measurements were carried out at different activation laser powers as described earlier. The integral of the measured signal is representative of the total fluorophore count achievable for this activation configuration. The total signal is dropping with higher light doses indicating a higher level of bleaching (see Fig. A.4). Measurements were carried out on homogeneous stained flat samples on at least three different areas per activation light dose (examples shown in Fig. 2.4). The addition of an antioxidant (ascorbic acid) had no effect on the bleaching behavior.

Since every point before the bleaching in the plots in Fig. 2.4 is representing new, previously not imaged, fluorophores every new activation delivers additional information about the imaged structure. Therefore it is beneficial for every activation laser power, even high ones, to successively activate multiple frames to combine them. This is however increasing imaging times and might be unfavorable for dynamic structures. If longer acquisition times are feasible it may be better to image multiple frames at low activation light doses, where the overall bleaching is lower, than a single one with high activation laser power.

2.5.3 Activation extent and STED-characterization

The total size of the 2PA volume is expected to be smaller than its diffraction limited spot due to the nonlinear nature of the process. The two-photon spot size can be assessed by imaging the 515 nm PSF which was scattered on gold beads. Due to the intensity squared dependency the effective two-photon activation volume is calculated by dividing the measured FWHM with $\sqrt{2}$. The lateral cross section of the 2PA laser beam PSF is depicted in Fig. 2.5 A with the corresponding FWHM of ~ 210 nm. Therefore the 2PA size is calculated to ~ 148 nm, very close to the theoretical minimum of 145 nm. Similarly the measured axial FWHM is 554 nm depicted in Fig. 2.5 A (right), resulting in a 2PA size of 392 nm which is again close to the theoretical value of 345 nm.
The effective experimental 2PA PSF however can be influenced by saturation effects and can differ from the one measured on gold beads. To get an estimate of the dimensions of the effective PSF a homogeneous sample of diluted dye in a collagen matrix was activated with a big scanning step size. The ramp pattern was confirmed to be a step wise function with measurements on an oscilloscope. The subsequent fluorescence was read out together with 2D STED at a power of 90 mW. The fluorescent spot exhibits a FWHM of (191 ± 4) nm which is bigger than the estimated number before. In the same way the axial size can be measured with a z-doughnut where the EPFS has an axial FWHM of (400 ± 79) nm slightly above the calculated extent.

The experimental linear PSF of the 2PA laser-beam as well as the 3D doughnut shown in Fig. 2.5 A, B are imaged on gold beads.
Both cross-sections of the 515 nm PSF show no significant aberrations and are of diffraction-limited size.

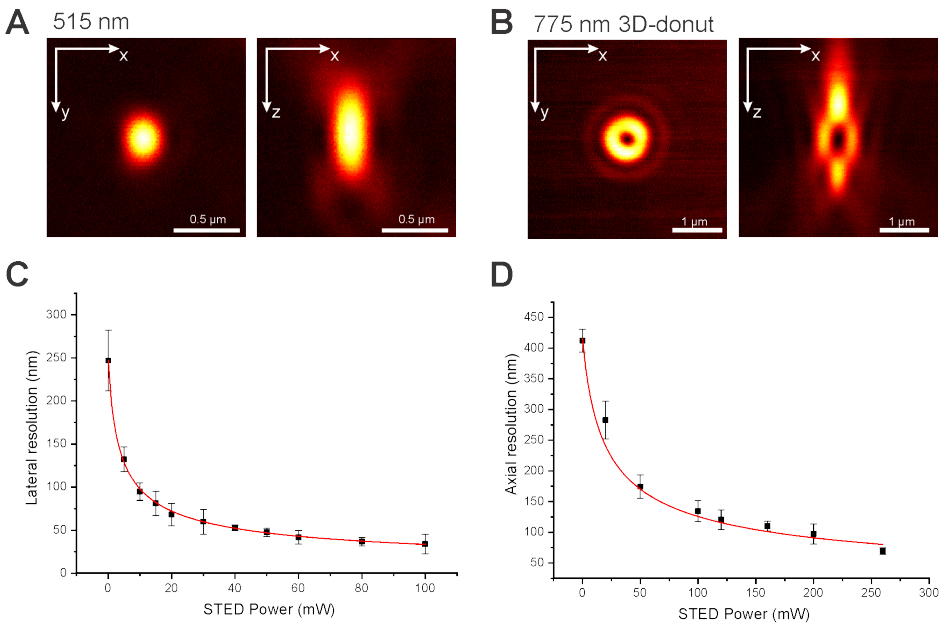

Figure 2.5: Characterization of the STED and 2PA PSF as well as STED resolution scaling. **A** PSF measured on 50 nm gold beads of 2PA laser at 515 nm in xy (left) and cross-section xz (right), scale bars 0.5 µm. **B** PSF of the combined 3D-STED doughnut in lateral xy-scan (left) and cross-sectional scan (right), scale bars 1 µm. **C** Lateral resolution scaling measurements of the 2D-STED measured on Atto647N single molecules as a function of applied average STED laser power (squares). The behavior follows the expected square root law which is shown as a red line [73]. **D** Axial resolution scaling measurements of the z-STED doughnut as a function of applied average STED laser power (squares). Theoretical fit shown in red.

The 3D STED doughnut shows slight astigmatic aberrations. This is visible in the slightly uneven intensity distribution of the 2D doughnut in the xy-scan, in Fig. 2.5 B on the left, and the 3D doughnut displays a slightly higher intensity in the top lobe (Fig. 2.5 B right). However, both small inhomogeneities are irrelevant for the applied use cases. The residual intensity in the center of the STED doughnut is minimal and measured at the same level as the background. The lateral resolution scaling of the STED microscope with applied STED power in the 2D-doughnut, measured on fluorescent microspheres, is depicted in Fig. 2.5 C and shows the typical square root law [73] with achieving a STED FWHM of (48 ± 5) nm at an average STED power of 50 mW. Fig. 2.5 D shows the axial resolution scaling when employing a 3D STED doughnut, reaching a FWHM of (97 ± 16) nm at an average

STED power of 200 mW.

2.5.4 2D STED imaging

To test if reasonable 2D STED super-resolution images could be acquired after 2PA of the dyes, we first imaged U2OS cells immunostained for microtubules.
Good STED images could be obtained by following the staining protocols described in subsection 2.4.14, as can be seen in the example of microtubules stained with ONB2-SiR in Fig. 2.6 A (right side). Enough fluorophores are activated after a single activation scan with a moderate power of 4 mW of the 2PA laser beam to render a continuous, well resolved representation of the sample. The pixel dwell time for activation was set to 100 µs and the activation pixel size to ~ 150 nm.
After acquisition of the 2PA STED image, the FOV is again targeted with the 1PA in order to record a reference image. The STED image obtained after 1PA shows significantly higher background in areas where the imaged cells are thicker in axial direction, mainly around the nucleus Fig. 2.6 A (left side). In direct comparison, after 2PA, the STED image shows better contrast across the entire cell compared to the STED image of the same cell after 1PA. This is also visible when observing the enlarged part of Fig. 2.6 A in Fig. 2.6 B. The intensity profile in Fig. 2.6 D shown for both 1PA STED and 2PA STED along the same line depicted in Fig. 2.6 B shows clear distinctive peaks in the 2PA case, where some of the peaks are masked by higher background level in the 1PA case. This background is furthermore increasing from left to right with the progression to thicker parts of the cell, where it stays constant in the 2PA case. As a result, not all microtubles are clearly visible in the profile line of the 1PA STED image, because they are lost in the higher overall background. Similar results of contrast improvement for the STED acquisition after 2PA compared to 1PA can be observed for the other two dyes HCage 620 and pPA-SiR (see Fig. A.2).
Comparison of the enlarged images in Fig. 2.6 B of the confocal scan after 2PA (top panel) and the two images taken with STED (middle panel 1PA and lower panel 2PA) shows a clear resolution improvement. The size of individual microtubules at a moderate STED power level of 50 mW is almost identical between both cases with an average diameter of (75 ± 13) nm for 2PA STED and (74 ± 19) nm for 1PA STED measured on six different individual filaments of the sample shown in Fig. 2.6 A. An exemplary intensity line profile for both STED cases and confocal is shown in Fig. 2.6 C, the arrows in Fig. 2.6 A indicate the position (additional

line profiles shown in Fig. A.3).

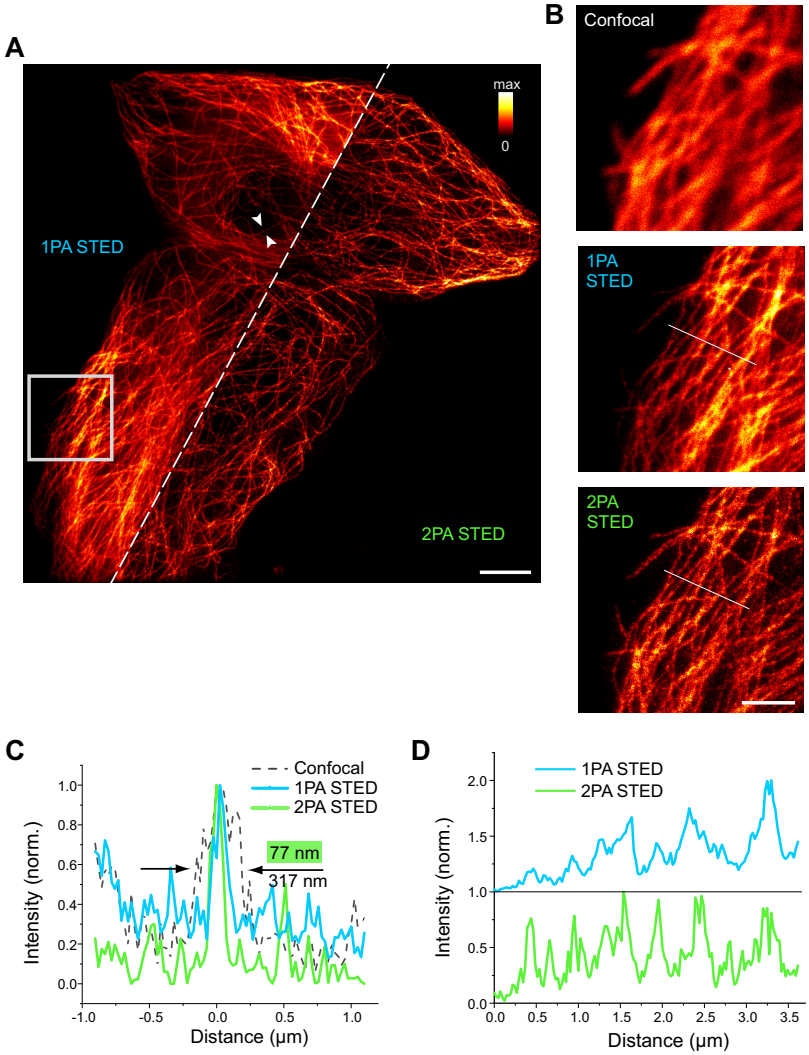

Figure 2.6: STED after 2PA of ONB-2SiR delivers higher signal to background ratio on simple cells. A Example of STED images acquired after 1PA (left) and 2PA (right). U2OS cells with microtubules immunostained with ONB-2SiR, scale bar 5 µm. **B** Enlarged area of selected image region from A (middle). Confocal and STED image after 2PA of the same area are shown for the reference (top and bottom, scale bar 2 µm). **C** Intensity line profile indicated by arrows from both sides from A. Corresponding confocal and 1PA STED profile lines added for reference (images not included) and showing the resolution improvement by STED nanoscopy. Resolution of the acquired image is similar for the 1PA and 2PA case. **D** Intensity profile lines indicated by arrows and line in B showing contrast improvement for 2PA with constant background.

The lateral size of a single two-photon activated spot is smaller than its diffraction limited spot due to the nonlinear nature of the activation process which was measured with STED in subsection 2.5.3. Due to the very targeted nature of the 2PA and its relatively small spot size, scanning configurations reminiscent of the protected STED or MINFIELD approach (see subsection 2.4.6) are possible [56, 74]. By only activating a small region in the sample, the majority of active fluorophores are not exposed to the high intensity STED laser. Scanning around the activated region with the center part of the doughnut reduces the bleaching significantly. The activation and following readout of every pixel individually is tested for all three dyes.
The results of this pixel activation mode for pPA-SiR are shown in Fig. 2.7 A. Different parts of a single U2OS cell immunostained for microtubules were imaged in different activation modes (pixel-activation vs. frame activation). The inset in the top left shows that the pixel activation imaging mode is much more favorable for pPA-SiR compared to the activation of a whole frame before imaging (left). The pixel-activation at equal activation power results in up to three times higher fluorescent intensities in both confocal and STED. The total detection time window of 130 µs as well as the color scale and STED power of 90 mW are the same in both activation cases presented. The parameters for the pixel-activation scan were chosen to be 0.6 mW of 515 nm, a pixel size of 24 nm with 10 µs activation time.

The other two caged dyes (ONB-2SiR and HCage 620) however, show a different behavior with a reduced detected signal when being imaged in pixel-activation mode.
Fig. 2.7 C shows alternating confocal scans of fluorescence detection during pixel-activation scans with (green points) and without (red points) the activation laser (gap time 880 µs, 1.5 mW of 515 nm, a pixel size of 24 nm with 10 µs activation time). A fluorescence detection without the activation laser after a previous pixel-activation scan is equivalent to the frame activation case. The acquisition scans with the 2PA laser light, during the pixel-activation mode, show consistently a reduced detected fluorescence signal compared to the same scan afterwards without the 515 nm light. This reduces the resulting image quality of the pixel-activated image significantly.
The fluorescence signal that can be detected during the pixel-activation scan is highly dependent on the length of the gap time between the activation and imag-

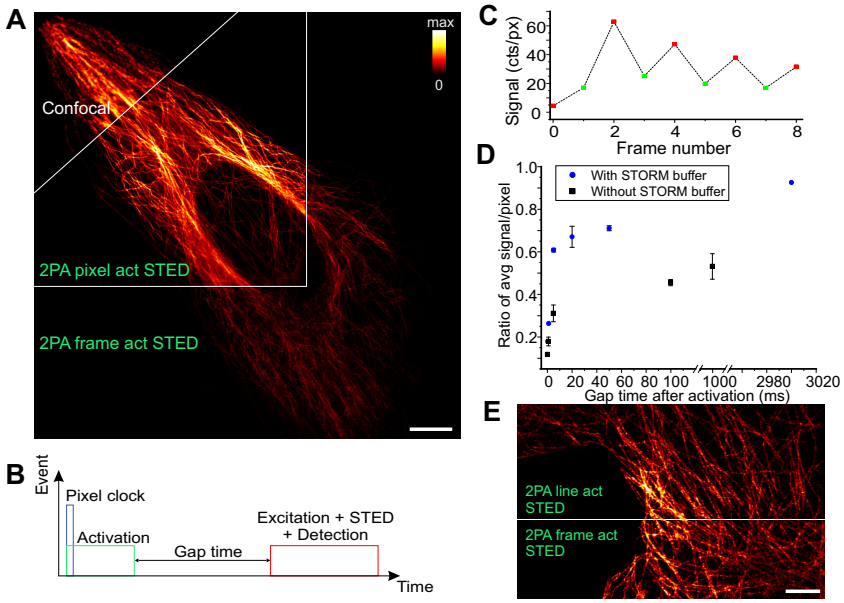

Figure 2.7: Pixel-activation greatly improves fluorescence detected during STED with pPA-SiR. A Confocal image during pixel-activation scan, STED image during pixel-activation and STED image after regular frame-activation. Pixel-activation scans show 3-times signal increase compared to regular frame-activation. U2OS cell with microtubules immunostained with pPA-SiR, scale bar 10 µm. **B** Schematic of the timing of events during one pixel acquisition. **C** Average signal per pixel of alternating pixel-activation scans with 2PA light (green) and with the activation light blocked (red). Data from measurements with ONB-2SiR. **D** Detected average signal per pixel during a pixel-activation scan divided by the signal that could be detected after a scan. Shown as a function of the gap time. **E** Activation of a whole line with following STED imaging shows no increase in detected fluorescence signal. U2OS cell immunostained with ONB-2SiR, scale bar 5 µm.

ing/detection. This is visualized for ONB-2SiR in Fig. 2.7 D where the average signal per pixel is calculated during and after the pixel-activation scan on a similar sample as for pPA-SiR. The ratio of the values of the signal before and after activation is plotted against the length of the gap time in ms. Very long gap times, more than 1 s are needed during the scan to reach 80 % of the fluorescent intensity that is detectable after the pixel-activation scan. This can be reduced somewhat by changing to a STORM buffer but gap times of 50 ms still result in a total acquisition time of ~ 14 h for a $(1024\,\mathrm{px})^2$ STED image in pixel acquisition mode.

Line-activation scanning measurements, as described in subsection 2.4.3, were im-

plemented to find a middle ground between the scanning speed and a potential signal increase due to bleaching protection. However no difference in detectable signal compared to regular frame-activation can be found in this scanning mode with ONB-2SiR or HCage 620. Therefore, no special activation mode is necessary for these two dyes.
This concludes the optimization of activation parameters. The optimal activation strategy for pPA-SiR is pixel-activation with immediate read-out of the fluorescence signal. For ONB-2SiR and HCage 620 the optimum is whole frame activation with following fluorescence read-out.

2.5.5 3D STED imaging

The 2PA volume is not only restricted in size laterally, but also in the axial dimension (see subsection 2.5.3). This is demonstrated by activation measurements using the 515 nm femtosecond laser in $\sim$ 50 µm thick mouse brain slices. To this end, cortical regions of the coronal brain section were immunolabeled for tubulin using HCage 620 (see subsection 2.4.14). Two individual focal planes of the slice were activated with both the 375 nm 1PA laser (power=130 µW, pixel size=150 nm, pixel dwell time=100 µs) and 515 nm 2PA laser (average power=4 mW) right next to each other. One focal plane located at a shallow depth of 6 µm above the coverslide and one deep in a distance of roughly 46 µm from the surface.
The successful activation is visible in the confocal overview scan. Fig. 2.8 A (left) shows a part of the whole coronal brain tissue section with both activated regions highlighted by a white outline. A detailed view onto the highlighted region is shown in Fig. 2.8 B with the 1PA region being on the left and the 2PA on the right. The image was acquired with both regions side by side with the same imaging parameters and is visualized with the same color scale. Even in the enlarged confocal scan in Fig. 2.8 C, at the depth of 6 µm, a pronounced difference in contrast is visible. The 2PA recording shows a better distinction between the structure and background with a much sharper delineation of the activated section.
Cross-sectional scans of this part are shown in Fig. 2.8 D. Two clearly distinct fluorescent layers are visible at the targeted activation depth for the 2PA case (right), illustrating the optical section capabilities of the nonlinear two-photon process. In the 1PA case (left), the fluorescent signal is significantly more spread out in the axial dimension, following the whole path of the beam.
Overall the fluorescence signal is decreasing with imaging depth. The signal is

also minimal in the center of the axial scan of the slice, possibly due to sample preparation related effects.

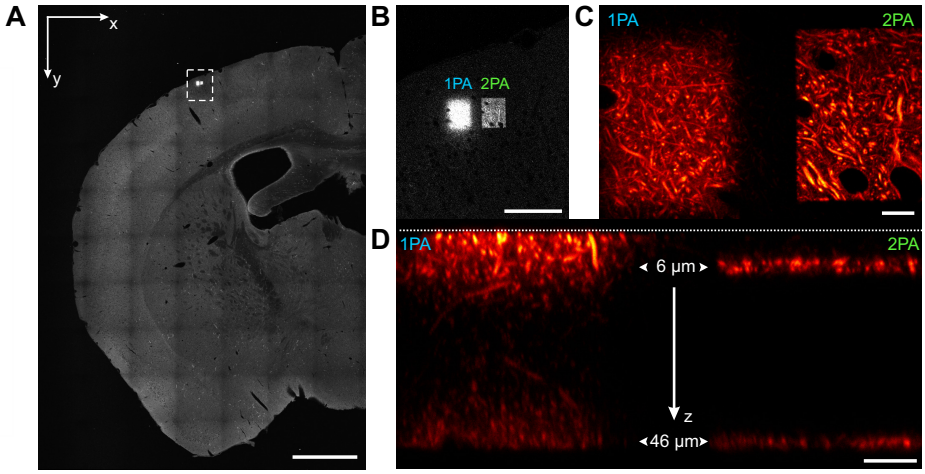

Figure 2.8: 2PA allows improved optical sectioning. **A** (left) Confocal overview scan of a cortical mouse brain section. Tubulin is immunostained with HCage 620. The two activated regions are visible within the white square. The 1PA region is located on the left, the 2PA region on the right, scale bar 1000 µm. **B** Enlarged area of the two activated regions, scale bar 100 µm. **C** Confocal scan shows better contrast in the area after 2PA, scale bar 10 µm. **D** Cross-sectional confocal scans of the region marked in A and shown in B,C. The coverglass position is indicated with a white dotted line. Two focal planes were targeted for activation, as can be seen in the annotations, one 6 µm and another 46 µm deep inside the tissue section, scale bar 10 µm.

The optical sectioning of the 2PA together with activatable dyes can be used, analogue to the two dimensional case, for an axial protected STED-like imaging approach. Unactivated molecules are not affected by the STED beam and are therefore not bleached.
Bleaching measurements of already activated layers with a 3D doughnut are performed as described in subsection 2.4.7. An example of the results for measurements with HCage 620 used to immunostain microtubules in U2OS cells is depicted in Fig. 2.8 A. The average signal per pixel is visualized as a function of the acquired frame number. For the layer that was additionally activated at the start of the measurement a clear drop in detected fluorescence signal is visible (bright green). This layer is continuously bleached by the high power in the 3D STED doughnut lobes. The layer that remained un-activated in the 2PA case shows very low initial

fluorescence signal, which is barely changing over time since almost no molecules reside in the activate state (dark green). After eight frames all STED imaging scans are completed and this previously not activated layer can be activated to show the full fluorescence signal in frame 9. The signal of the similar layer (in another area) for the 1PA case shows a high fluorescence signal already at the start of the measurement (frame 1, blue). This signal also drops with the following successively acquired STED images and cannot be fully recovered after all these bleaching scans.
This underscores that 2PA can be used together with 3D STED to achieve a low-bleaching, pseudo protected STED imaging mode (for the axial dimension).

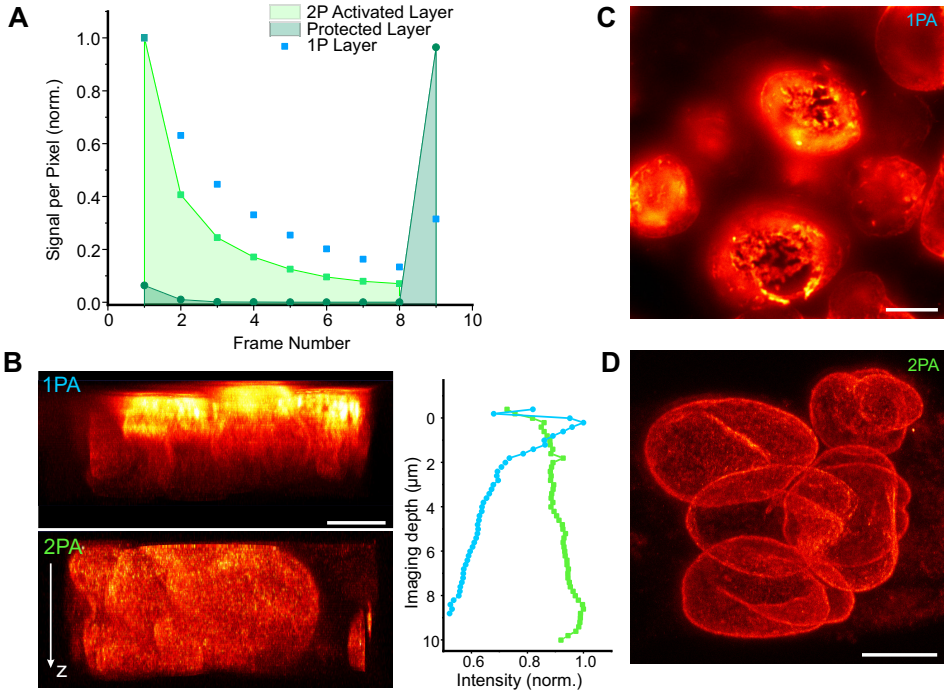

Figure 2.9: 2PA allows for better 3D imaging with preserved axial signal. **A** Results of the bleaching experiment described in subsection 2.4.7. The average fluorescence signal per pixel is plotted as a function of the frame number collected. **B** Reconstructed side view of acquired image stack of LaminB stained with HCage 620 in nuclei of neural stem cell colonies for 2PA and STED (top) and 1PA and STED (bottom), scale bar 5 µm. The graph to the right shows the corresponding normalized intensity as a function of the imaging depth of the measurements on the left. The maximum intensity projections of the image stacks are shown for 1PA in **C** and for 2PA in **D** on the right, scale bars 5 µm.

This advantage of 2PA can be used when imaging extended structures in 3D. To demonstrate this, LaminB of nuclei in neural stem cell colonies are immunostained with HCage 620 and imaged with a 3D doughnut with a total power of 140 mW where 40 mW are in the 2D doughnut (2D FWHM $\sim$ 50 nm, 3D FWHM $\sim$ 134 nm). Individually imaged layers are spaced 200 nm apart and repeatedly activated using either 1PA (160 µW for 10 µs with a pixel size of 20 nm) or 2PA (3 mW for 10 µs with a pixel size of 20 nm).
The cross-sectional reconstruction in Fig. 2.9 B shows a homogeneous distribution of the detected fluorescent signal for the case of 2PA-STED. This is not the case for 1PA-STED where a clear gradient is visible due to successive bleaching of the following layers. The reason for this non-beneficial behavior is that those layers already got activated even though they don't coincide with the focal plane of the 375 nm laser light. The average fluorescence per pixel in each layer from this measurement is depicted in the Graph in Fig. 2.9 B as a function of the imaging depth. The almost flat signal over the imaging depth for the 2PA case illustrates its possibility to preserve fluorescence signal compared to 1PA.
The maximum intensity projection of the imaged stacks allows for a top down view through the nuclei in Fig. 2.9 C, D. The significantly increased background and inhomogeneous distribution of signal in the 1PA-STED case is also visible here. In comparison, the 2PA-STED case allows for a three dimensional like impression through this reconstruction due to its constant signal with low background throughout the imaged slices.

2.5.6 Biological application

The evaluation of the optimal scanning modalities and activation parameters for each dye enables the use of 2PA-STED for biological applications. In the following section, I present some potential applications, which highlight the benefits of combining 2PA together with STED. I will demonstrate that, due to improved optical sectioning, 2PA-STED allows the acquisition of high-contrast data of challenging thick biological samples with reduced photobleaching.

As an example, Fig. 2.10 A, B show 2PA-STED imaging of a thick brain tissue slice. Here, actin in dendrites was stained with HCage 620 (see sec subsection 2.4.14). 2PA-STED allows for a resolution gain sufficient for capturing the individual spine necks and heads. The images depicted here are taken within the tissue slice at

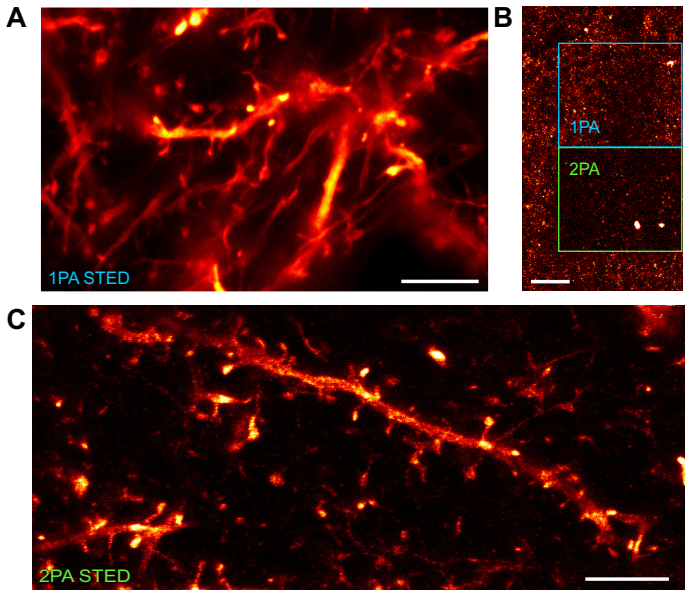

Figure 2.10: 2PA allows acquisition of STED images in brain slices with improved contrast. **A, C** Actin stained in dendrites in tissue samples with HCage 620 as described in subsection 2.4.14. **C** 1PA and STED shows much more background and haze surrounding the imaged structure. **A** 2PA and STED has a better contrast with much lower overall background. Scale bars 5 µm. **B** Unstained mouse brain slices fixed with PFA. The part which was illuminated with 1PA light, marked in blue, shows no change in signal. The part exposed to 2PA laser light, marked in green, shows less fluorescence. Scale bar 10 µm.

a relatively shallow depth to avoid aberration issues. However, even at a depth of 4 µm the reduced and more homogeneous background in the 2PA STED image shown in C is visible compared to the 1PA STED image presented in A. This is due to the restricted activation of the targeted layer and a reduction of auto-fluorescence generated from the PFA fixation of the tissue sample. This PFA induced fluorescence background can be quite effectively bleached by illumination with a normal 2PA laser power of ~ 3 mW. An example is visible in Fig. 2.10 B where an unstained PFA fixed tissue sample was illuminated with 375 nm 1PA light (top) and the 515 nm femtosecond 2PA laser (bottom). The region of the latter has a significantly lower fluorescence after illumination compared to the sample before the exposure (outside of the marked regions), whereas the region of 1PA shows no change in detected signal. Tissue that was not fixed with PFA shows no difference in the fluorescence signal before and after 2PA illumination. This supports the presumption that the change in background is due to the bleaching

of PFA induced signal.

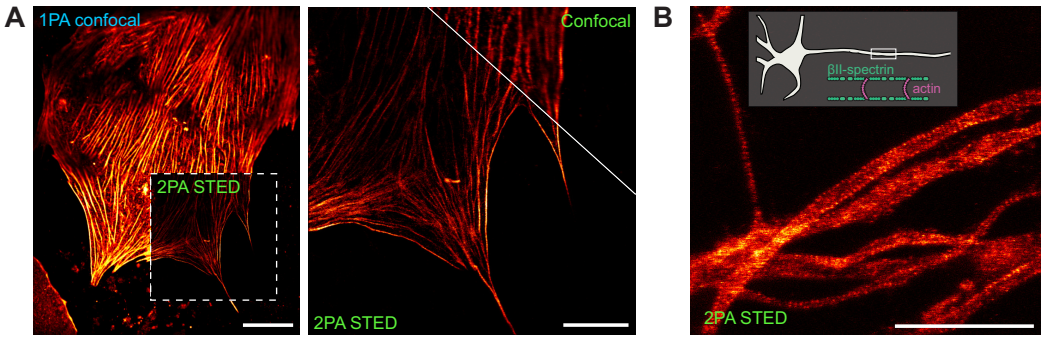

Figure 2.11: 2PA allows acquisition of STED images in challenging thick and flat samples. A Actin in hepatocytes immunostained with pPA-SiR. Overview shows confocal image after 1PA after the imaging of the marked region which is 2PA and STED, scale bar 20 µm. The marked region is shown enlarged on the right with the confocal data, scale bar 10 um. **B** β2-spectrin immunostained with pPA-SiR, scale bar 5 µm. STED images with pPA-SiR were taken in a pixel activation mode. The cartoon in the inset shows the underlying periodic structure.

Another relatively thick sample, which is prone to out-of-focus background during imaging, are hepatocytes which are depicted in Fig. 2.11 A. The hepatocytes were immunostained with pPA-SiR (see subsection 2.4.14) and visualized by 2PA-STED. The enlarged section to the right shows individual actin fibers with a diameter of ~ 90 µm with overall very low background. The insert of the confocal picture demonstrates the resolution improvement by 2PA-STED. The confocal picture shows a non-continuous structure in some areas which is due to the fact that it was acquired after the STED image due to the pixel-to-pixel activation image mode required with pPA-SiR.
Fig. 2.11 B shows STED images of neurons where β2-spectrin was immunostained with pPA-SiR. The STED images of β2-spectrin reveal the underlying periodic axonal cytoskeleton which is shown in the cartoon in the inlet in Fig. 2.11 B. STED recordings show the characteristic striped pattern with an average spacing of (191 ± 16) nm between to consecutive rings of β2-spectrin.

The STED images acquired after 2PA show greatly improved signal-to-background ratio, independent of the imaged structure. The three tested dyes are compatible with STED offering a great versatility when it comes to labeling different biological

structures. Imaging these dyes with 2PA-STED can deliver images with decent quality at the required STED resolution.

2.5.7 Single molecule activation and localization

Activation of the fluorophores in bulk enables confocal and STED microscopy as described in the previous examples. However, 2PA is not restricted to an ensemble activation but could also be extended to the activation of single molecules. This opens the application of 2PA to another set of super resolution microscopy techniques such as SMLM. Especially with the newer SMLM techniques described earlier (see subsection 2.4.9) this could result in nanometer scale resolution, which also benefits from an increased signal-to-background ratio. For successful application of 2PA for SMLM a sparse activation of individual molecules has to be guaranteed.

Preliminary tests for the estimation of the required 2PA laser power were performed on the confocal scanning microscope (see subsection 2.4.1). After their activation the individual molecules were localized (subsection 2.4.8). The resulting number of individually localized dyes is depicted as a function of the applied laser power in Fig. 2.12 for ONB2-SiR (A), HCage 620 (B) and pPA-SiR (C). All these dyes show a similar behavior of a low number of detectable localizations at low light doses. At a power level of $\sim 7\,\mu W$ the number of localizations show a clear increase in the case for ONB-2SiR Fig. 2.12 A. This number is steadily increasing with higher applied laser power until reaching a plateau at $30\,\mu W$. Results from HCage 620 and pPA-SiR show a slightly different behavior, since the number of localizations shows a slow increase until it reaches a similar plateau at a maximum laser power of around $50\,\mu W$.
When the activation beam is blocked while performing the same type of measurement, no increase in the number of localized molecules can be observed. This is falsifying a purely time dependent effect.
To avoid a cumulative effect of the activation a random selection of the power was implemented.

Based on these results, an activation laser power range of $30\,\mu W$ to $50\,\mu W$ was chosen for the following single molecule measurements, depending on the specific dye that was used. A laser power in this range ensures to deliver a steady state of

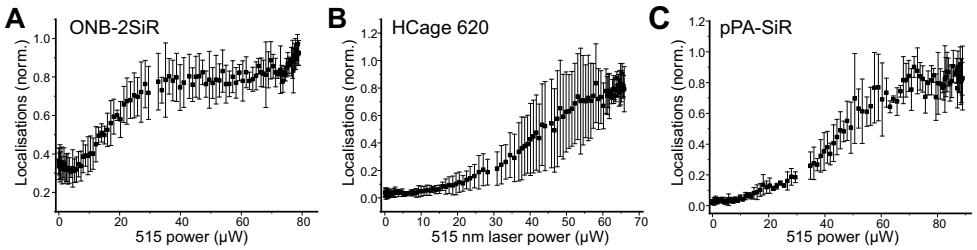

Figure 2.12: Average normalized number of localization. Number of activated ONB-2SiR (**A**), HCage 620 (**B**) and pPA-SiR (**C**) fluorophores as a function of the activation laser power. An increase in localizations for ONB-2SiR is visible starting from ∼7 µW. The number of localized molecules is increasing with higher laser power until it reaches a plateau.

active fluorophores that can be localized.

Successful 2PA of single molecules allows the application of a PALM-like imaging mode, where single activated fluorophores are localized to form a super-resolved structure, as described in subsection 2.4.8. Initial scans of relatively big and extended structures like microtubules immunostained with ONB-2SiR in U2OS cells showed a working implementation of this mode at an activation power level of 40 µW (see Fig. 2.13 A). The individual microtubules can be clearly distinguished from the background in the single molecule localization image and resolve closely spaced microtubules which are indistinguishable in the confocal picture. In line with this, Fourier ring correlation (FRC) calculations indicate a, better than confocal, mean resolution of 117 nm.

Similarly, SMLM images can be acquired of samples that were previously shown in STED in subsection 2.5.6. Fig. 2.13 B shows β2-spectrin in neurons immunostained with pPA-SiR. The super resolved picture shows the characteristic striped pattern, similar to the example shown before for 2PA-STED. The distance of the individual stripe, as measured along the marked white line Fig. 2.13 B, is similar to the distance estimated on STED recordings with a mean peak-to-peak distance of (196 ± 15) nm.

Due to the two-photon nature of the activation, the acquisition is not only restricted to the coverslide. Similar to the STED examples in subsection 2.5.6 images were acquired in mouse brain tissue samples stained with pPA-SiR. Successful

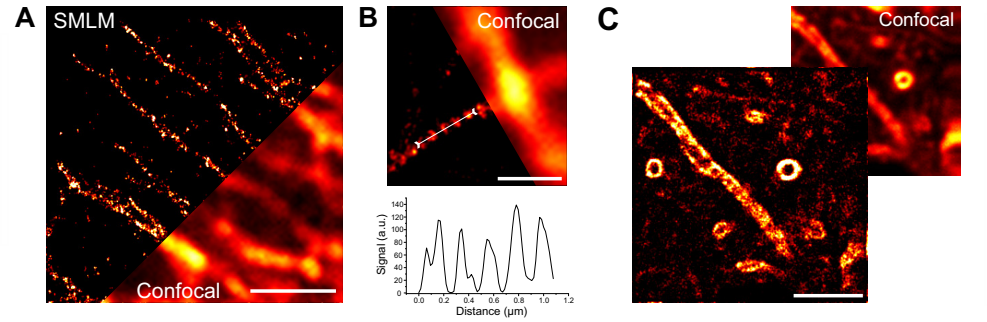

Figure 2.13: PALM-like SMLM with confocal microscope after 2PA. **A** Single molecule activation and localization of ONB-2SiR (top left) versus the confocal image (bottom right). The imaged structure are microtubules in U2OS cells. The SMLM image can clearly resolve multiple strands of microtubules next to each other which blend in the confocal picture. Scale bar 2 µm. **B** β2-spectrin immunostained with pPA-SiR. The SMLM image resolves the periodic structure with a distance of 180 nm to 200 nm, scale bar 1 µm. The intensity distribution along the line profile indicated with the white line in the top image is plotted in the bottom Graph. **C** SMLM of dendrites in mouse brain tissue slices where actin was stained with pPA-SiR. The image was acquired at a depth of ~ 3 µm. Scale bar 3 µm.

activation, detection and localization of individual fluorophores was possible at a depth of around 3 µm. This demonstrates SMLM together with 2PA as a possible technique for tissue imaging at depth in the sample.

To showcase high resolution 2PA SMLM imaging, we stained NUP96, a protein in the nuclear pore complex (NPC) with HCage 620 using the HaloTag. Imaging of this structure however was unsuccessful with very few localization events during the time of the measurement (30 min to 40 min). The main reason for this was, the confocal scanning system was not equipped with an active lateral drift control. Without this stabilization system the imaging is limited in terms of time and FOV (due to an extended acquisition time).

To overcome this issue, the measurements of NPCs were repeated on a MINFLUX setup with an active stabilization system to achieve high-resolution images acquired over an extended time period.
Individual fluorophores could be successfully activated via 2PA (60 µW) and subsequently localized in an iterative MINFLUX scheme (see Fig. 2.14 A, B). The majority of localizations over an extended imaging time were converging success-

fully, as can be seen from the left skewed histogram in Fig. 2.14 D. This histogram shows the difference $r_{\text{est-rel}}$ between the estimated position of the FPGA during the scan and the localization position determined with the maximum likelihood estimator during post-processing. Good localization is also possible due to the low background in the sample, which is visible in the distribution of the ratio of the fluorescence detected at the central MINFLUX position relative to the other three positions p_0, with the majority of localizations being well below 0.2 (Fig. 2.14 E). The experimental localization precision results to $\sigma_x = 2.12\,\text{nm}$ and $\sigma_y = 2.20\,\text{nm}$ with a minimum number of 10000 photons in the last iteration step (Fig. 2.14 C).

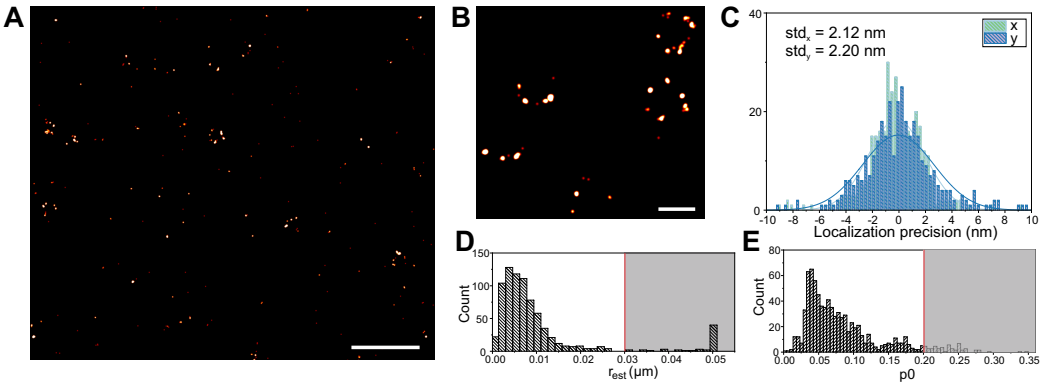

Figure 2.14: MINFLUX images of NUP96 after 2PA. A Overview of a MINFLUX scan with 2PA (60 µW). NUP96 in U2OS-cells are labeled via Halo-tag with HCage 620, scale bar 500 nm. Image contains very low number of nuclear pore complexes with most of them severely underlabeled only showing a maximum of six subunits. **B** Representative enlarged image of a different area showing the individual NUPs, scale bar 100 nm. **C** The experimental localization precision as a histogram for each direction. The standard deviations of the localization precision are calculated for x to 2.12 nm and for y to be 2.20 nm. The data is filtered according to the criteria $p_0 < 0.2$ and $r_{\text{est}} < 0.03\,\mu\text{m}$. The corresponding histograms of a slightly larger FOV are shown in **D, E** with their cutoff as a red line. Localizations to the left of the dividing line, shown in red, are used for the rendering of the image. Gaussians for both images A & B are drawn with a fixed σ of 3 nm.

However, not all nuclear pore complexes in the whole FOV show the expected eight-fold symmetry and appear mostly under-sampled with six or less sub-units visible. As similar results are obtained when using 1PA with 405 nm (10 nW), this observation points to an incomplete labeling of the HaloTag. Therefore, different labeling methods were employed such as SNAP-tag, combined with HCage 620, or labeling

of NUP96-GFP with an anti-GFP-nanobody/antibody with HCage 620. Likewise, we performed immunostaining of nuclear pore sub-complexes with pPA-SiR. All of them resulted in mostly incomplete NUPs in MINFLUX images with both activation wavelengths. Additionally, the variation of the buffer (PBS, HEPES, STORM buffer) or the modification of imaging parameters had no effect on the overall result. Independent of the labeling strategy, successful 2PA and stable localization of molecules could be achieved on top of the U2OS cell's nucleus at an imaging depth of $\sim 3\,\mu m$.

First experiments at shallow depths in mouse brain tissue slices with actin immunostained against GFP in dendrites, showed as well that localization of single activated molecules is possible in challenging samples with MINFLUX using the 2PA approach.

The combination of photoactivatable dyes which are also STED compatible makes this method a good candidate for MINSTED imaging.
For this purpose, Caveolin-1 in simple U2OS cells was immunostained with ONB-2SiR and HCage 620 (see subsection 2.4.14). Imaging and rendering parameters are as described in subsection 2.4.9. As can be seen in Fig. 2.15 C/E the internal structure of the caveolas can be resolved by MINSTED. It is worth noting that the acquired two-dimensional MINSTED image shows a projection of the three-dimensional object which is depicted in the cartoon in Fig. 2.15 D.
The comparison to the confocal overview scan of the caveolas is shown in Fig. 2.15 B. As a note, the confocal image does not show the unresolved fluorophores from the MINSTED measurements, but the corresponding counter-staining of Alexa 647 (see subsection 2.4.14).
The diameters of the clusters of localizations in C are in the range of 50 nm to 70 nm which is the expected size of caveolas, as reported previously [75].

The complete MINSTED image in Fig. 2.15 C consists of a total of 1462 localizations. The histogram of the number of localizations with the corresponding estimated localization precision is shown in Fig. 2.15 F. The histogram shows a clear left skewed distribution, with the majority of localizations at the lower bound and a median localization precision of 2.6 nm. The number of localizations with their corresponding standard deviation at the minimal FWHM of the STED iteration is visualized in a histogram together with the separation line (red) used for filtering

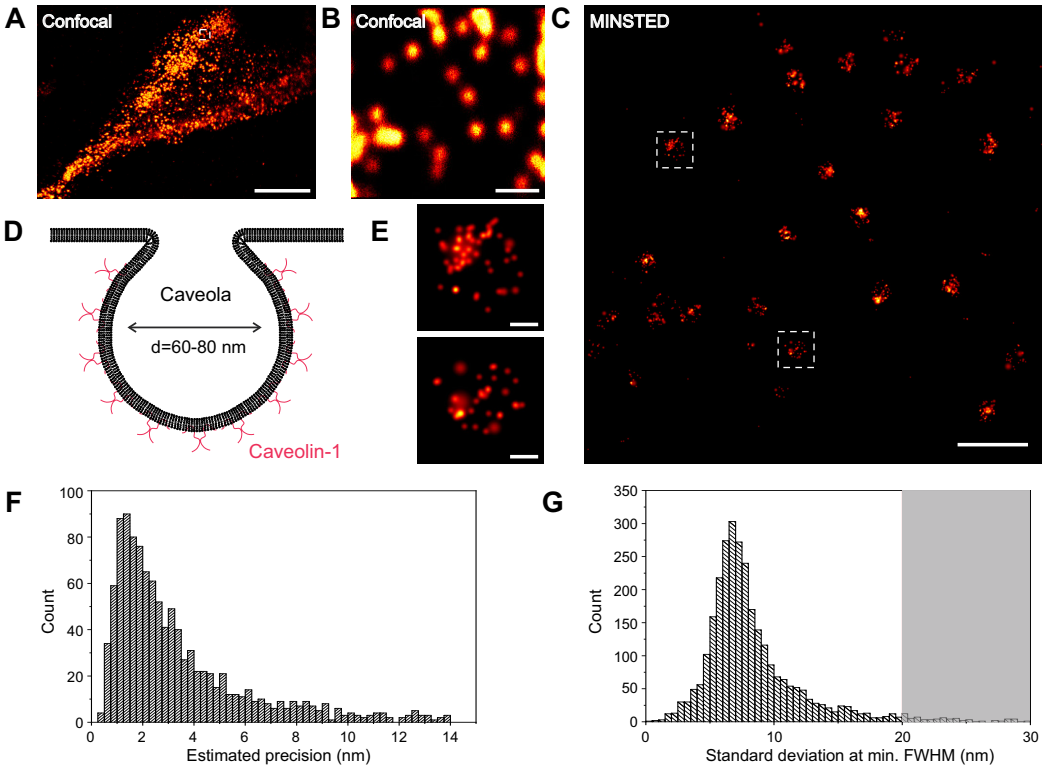

Figure 2.15: MINSTED images of Caveolin-1 after 2PA. A Confocal overview of a U2OS cell where Caveolin-1 was immunostained with ONB-2SiR, scale bar 10 µm. **B** Enlarged confocal region of the cell. Position indicated in A by the white outline. Scale bar 0.3 µm. **C** MINSTED image of the same confocal region shown in B. The MINSTED image resolves the internal structure of the individual caveolas, scale bar 0.2 µm. The scanning FOV of the MINSTED image is slightly smaller than the confocal image in B. This is the reason why caveolas in the region of the edge are not resolved. **D** Cartoon of the caveola with the targeted Caveolin-1 protein shown in red. The average diameter of the caveola is equal to 60 nm to 80 nm with a diameter of the neck of 10 nm to 50 nm [75]. **E** Enlarged areas of the MINSTED image from C. The positions in C are highlighted with white outlines. The upper image corresponds to the upper outline. Scale bar 20 nm. **F** Histogram of the estimated precision of the individual localizations in nm. **G** The standard deviation of the localization measurements when reaching the minimal FWHM of the STED. The events are filtered as described in subsection 2.4.9. The threshold is shown as a red line. All data points right of the line are discarded.

(localizations right of line are filtered out, marked with grey shade) in Fig. 2.15 G. The distribution shows a good convergence of the localizations, with the majority of them within the non-filtered region. The medium amount of detected photons

per ONB2-SiR fluorophore after 2PA is equal to 1556 ± 160. Well above the threshold for filtering. The data is comparable to previously reported data after 1PA [21]. The number of photons per fluorophore after 1PA 1269 ± 200 doesn't deviate from the one measured after 2PA.

Similar measurements are possible with Caveolin-1 immunostained with HCage 620. For HCage 620 however the number of photons is significantly lower with 774 ± 165 detected photons after 2PA and 965 ± 241 photons after 1PA.
MINSTED measurements with 2PA are successful at different imaging depths, delivering super resolved images with good localization precision up to a distance of $\sim 2\,\mu m$ from the coverglass, which is roughly equivalent to the height of a U2OS cell.

2.6 Discussion

This thesis demonstrates that 2PA improves the contrast in confocal and various types of super-resolution microscopy techniques. The main reason for this is that the 2PA restricts the conversion of activatable dyes to a small volume in the focal spot. Thereby 2PA counteracts out-of-focus signal, especially in thick specimens. Three different dyes (ONB2-SiR, HCage 620, pPA-SiR) were selected for this study, all of which are suitable for 640 nm excitation and STED at 775 nm wavelength. The three dyes show two different activation mechanisms and were found to be efficiently activatable at a wavelength of 515 nm with short-pulsed laser light without strong premature bleaching. All of them show a multi-photon activation mechanism. This could be shown by the change of the activation rate which scales with the squared average applied activation laser power. Additionally, a direct inverse change of the activation rate with a variation in pulse length ($\frac{1}{\tau}$) was observed. This is the expected behavior for a two-photon process (see Eq. 2.2), whereas the regular 1PA is unaffected by pulse length changes. Among the three dyes ONB-2SiR shows the clearest two-photon uncaging behavior. HCage 620 exhibiting an activation order higher than two reveals to be either a mixed process with predominantly two-photon behavior or a two-photon activation together with an additional step (similar to the reported 2PA of DsRed2 [76]).
The successful activation of HCage 620 opens up the potential of 2PA at 515 nm for the remaining dyes published in [21], if the absorption of the uncaged dye at this wavelength is small enough. Similarly, color shifted variants of ONB-2SiR with the same caging group have a high chance of being activated by 515 nm light through a two-photon process.

This work shows that bulk activation of all three dyes allows for subsequent STED imaging with overall enhanced contrast. The needed high applied average laser power however is inducing bleaching due to the high peak power of the 515 nm femtosecond laser. Therefore, for single activation measurements an optimal activation configuration was established for each dye.
For a pulse length of $\sim$ 200 fs and a pixel size of 150 nm with a dwell time of 100 µs the optimal average laser power for ONB-2SiR and HCage 620 is equal to 4 mW, which is well in range of previously reported 2PA power levels for different dyes at other wavelengths [49]. This activation configuration delivers the highest amount of activated fluorophores in one scanned frame, resulting in the best achievable

single frame STED images. Since only a subset of all available fluorophores is activated after a single activation scan, a repeated activation and detection scanning, is advisable, if the associated longer acquisition times are manageable.
It was shown that not all dye molecules survive the activation process. The total amount of bleached fluorophores is dependent on the activation laser power and slightly different for each of the three dyes. Bleaching pathways can be very complex. However, a bleaching via the formation of a cation and solvated electron pair, which lead to irreversibly destroying the molecule through collisions with the solvent molecules, can be considered unlikely since the addition of ascorbic acid as an antioxidant had no effect on the amount of detected bleaching [77]. Additionally, pPA-SiR shows a fluorescence quenching effect after activation due to a reaction with the surrounding water molecules resulting in a reduction of detected fluorescence for normal imaging modes (described in [23]). Because of this, a special scanning mode was implemented, where the fluorescence is read out immediately after the activation with 0.6 mW of 515 nm. With this approach, an efficient activation can be achieved, allowing efficient STED nanoscopy at 775 nm for this fluorophore.

An additional indirect proof that the activation with 515 nm is in fact a non-linear process is the resulting small activation volume. The finite activated region is expected for higher order activation mechanisms.
With this precise control over the location of the activation it is now possible to image in a pseudo-protected-STED imaging mode.
A lateral protected STED reminiscent technique was shown to work for pPA-SiR as described above but showed not to be usable for the two caged dyes ONB-2SiR and HCage 620. The detectable fluorescence is increasing over long periods of time (up to s) after an illumination with the 515 nm laser. This might be due to a slower activation process and/or transitions to a long lived dark state induced by the femtosecond laser light.
However, this imaging mode is especially promising for axially extended structures that are imaged with a 3D doughnut, where 2PA delivers a significantly improved signal retention in the third dimension compared to regular 1PA at 375 nm. The confinement of the axial activation volume is additionally decreasing out-of-focus fluorescence which could further degrade image quality. Even on very thin samples with a thickness of $\sim 3\,\mu m$ the application of 2PA together with STED is beneficial, resulting in a much better contrast due to the optical sectioning.

The examples shown in subsection 2.5.5 demonstrate the effectiveness of the activation of a single plane deep in brain tissue slices. This optically-sectioned activation can also deliver much improved STED images in thick samples with overall lower background, as demonstrated before. STED images at a shallow depth are presented as an example in this work. For high-resolution deep-tissue STED imaging, adaptive optics (AO) is needed in order to correct for aberrations introduced by the sample [78–80], as well as the use of a focusing piezo with a bigger range.
It could be argued that two-photon excitation could deliver similar results for bulk STED imaging [42, 81]. However, it has been shown that 2PA together with two-photon excitation has the potential to drastically increase the imaging depth compared to regular excitation and two-photon excitation on its own [82].

The successful application of 2PA in mouse brain tissue samples together with STED, showing a significantly reduced background compared to regular 1PA, makes this activation mechanism a good candidate for MINFLUX or MINSTED image acquisition in comparable challenging samples. Due to its use of an excitation beam only, MINFLUX is susceptible to background and its suppression greatly improves the localizations.
As a ground work, different 2PA single molecule localization techniques (PALM, MINFLUX, MINSTED) near the coverslide were shown to work here. Localizing single molecules with high precision, deep in tissue samples, requires special care and a special microscope system with adaptable optics and a stabilization system supporting the imaging depths. This is open for future work.
Nonetheless, the acquired images show as a proof of concept that the 2PA presented in this work could be applied to any imaging technique, ranging all the way from confocal over STED to the newest nanometer-scale SMLM techniques.

Unfortunately, the selected dyes show some unfavorable characteristics when it comes to single molecule imaging of sparse samples such as nuclear pore complexes. All three dyes seem to exhibit very poor labeling efficiency, independent of the labeling mechanism (antibody, nanobody, HaloTag, SNAP-tag). Furthermore, the fluorophores are not reversibly switchable. This results in a loss of localization targets, if some of the fluorophores are activated either thermally or due to premature exposure to light, e.g. during staining or transportation, since initially fluorescent molecules are bleached at the start of the MINFLUX or MINSTED measurement. In a similar way, activated fluorophores that couldn't be localized

or multiple activated fluorophores too close to localize them individually are lost as a result. These cumulative effects can result in under-sampled images of the underlying structure. However, the results of undersampled structures is not an effect of the usage of 2PA with SMLM but a general issue in SMLM. In fact, the results presented here are comparable to already published data on pPa-SiR [23] (NPCs with STORM) and HCage 620 [21] with classical 1PA in UV, with the exception that 2PA delivers much more precise control over the location of activation and reduced background signal. A potential solution to the problem of sparsity could be an approach which uses a DNA-paint [83] imaging scheme with these dyes.

The experimentally measured size of the two photon PSF at the regular activation power levels used for STED imaging in this work is slightly bigger than the minimal PSF predicted by the theory (see Eq. 2.1). This is somehow expected since the actual dimensions are highly dependent on the activation power which is applied [48]. This can result easily in a saturated effective two-photon PSF [28] with a larger overall volume. Reducing the activation power could therefore result in a minimal activated volume (as discussed in section 2.2), with the downside of not activating enough molecules in the region of interest to acquire a high quality STED image. The balancing of the number of activated molecules with the bleaching of molecules and the effective size of the two-photon PSF can be different for every imaging mode. For the significantly reduced activation power for single molecule imaging, the two-photon PSF can be assumed to be minimal.

The activation at this wavelength resolves a number of issues that existed with the classical UV activation of these dyes [84]. Using visible light at 515 nm results in less chromatic aberrations induced by the optics since it resides in the optimized range of both achromatic and apochromatic lenses. Together with the significantly increased transmission through optical elements, this makes building a microscope with this activation wavelength much easier. The ability to easily characterize the PSF on gold beads simplifies the lateral and axial overlapping of all individual beams in the microscope significantly.
Together, the 515 nm laser provides a very compact solution with an easy to use turn key system which can fit on almost any setup.

On the other hand, the relatively short activation wavelength is not ideal when it comes to deep tissue imaging. The wavelength resides outside of the typical

optical window used for tissue imaging from 650 nm to 1350 nm where the absorption of light is minimal [50]. Furthermore, the scattering of light is increasing significantly at shorter wavelengths, making it less practical for deep imaging in biological structures.
Still, this option remains more practical than the alternative 1PA in the UV region. This regular 1PA wavelength window is, in the case of ONB-2SiR and HCage 620, explicitly shifted to lower wavelengths (below 380 nm) to avoid a two-photon activation by the 775 nm STED beam [21, 51]. For this reason a three-photon activation would be needed for the application of wavelengths in the optical window. However, this would set much more stringent requirements for the dispersion compensation compared to 2PA. Especially third order dispersion compensation introduced by the prism compressor could be ignored in the 2PA case but would be relevant for three photon activation to achieve the required shortest pulses possible [85].

A drawback of the two caged dyes is the overall biocompatibility. Due to the extended caging system, both dyes, ONB-2SiR and HCage 620, are not able to penetrate the cell membrane efficiently, rendering them not live cell compatible. Unfortunately, the remaining dye, pPA-SiR, exhibits a changed spectrum of the unactivated compound with a higher absorption at 515 nm for its live cell compatible Halo- and SNAP-tag conjugates. This is not unexpected as it has been previously reported that these self-labeling tags can have a significant influence on the properties of SiR dyes [86], but here it unfortunately increases the bleaching for a 2PA application. As a result *in vivo* stainings with these dyes are not possible. The potential benefit of 2PA at 515 nm to be less photo-toxic than the regular activation with high power UV photons was not tested due to this reason. This could be confirmed with a new set of biocompatible dyes.

Overall, 2PA at 515 nm exhibits a great potential for microscopy, independent of the method that is used. The ability to control a small activation volume precisely is beneficial for a variety of techniques ranging from confocal to nanometer-scale SMLM. 2PA at 515 nm can replace the classical 1PA below 400 nm for photoactivatable (far-)red dyes, which improves imaging conditions drastically as well as resolving many constraints during the microscope construction. With the combination of AO, 2PA offers the potential of nanometer-scale SMLM acquisition deep in the specimen.

3. Image reconstruction in low signal STED microscopy

3.1 Introduction

Classical image processing concerned itself previously with task specific tailored algorithms. With the advent of machine learning however, deep neural networks have become a popular and powerful tool in data-driven analysis and image processing. These neural networks can be used to find any mapping of an input data distribution to a target distribution, without the need of hard-coding specific mathematical operations. If the input and output are both images, such as in denoising [87] or image reconstruction [88] applications, this task is known as image-to-image translation. Explicit pre-determined image filters are no longer needed as the network is trained to learn the weights itself. Different variations of these neural networks have already found their application in microscopy for signal-to-background ratio (SBR) improvements [87, 89, 90] or deconvolution applications [90].

In its most basic form neural networks are used to find a mapping with the use of multivariate functions from a source to a target distribution. It has been shown by the universal approximation theorem [91] that a neural network with at least three layers can approximate almost any function connecting two datasets if the number of weights is allowed to be arbitrarily large. By presenting data from both distributions to the network, the weight parameters w are learned in the process of training. Typically, the network is trained on a sub-sample of the source data and the performance is then validated on the remaining data from the set. The training data is used to determine the fit of an unknown function between both distributions. The use of activation functions, such as the rectified linear unit (ReLU),

between different layers of the network is introducing important non-linear elements to the overall network design, allowing the network to learn more complex relationships [92].
A calculated loss value is used to determine how well the fit of the network is approximating the true mapping. Popular loss functions are the L1-norm or L2-loss (Euclidian norm, see Eq. 3.2). The aim of the training is finding a mapping between the input and output distributions that minimizes the selected loss function. This is typically achieved through either stochastic gradient descent or with the use of adaptive gradient methods such as Adam [93] or AdamW [94].
Generative adversarial networks are an interesting network architecture for image-to-image translation. They rely on convolutional neural networks (CNNs), especially down/up-sampling CNNs. Both types will be shortly discussed in the following sections.

3.1.1 Convolutional Neural Networks

When processing images in neural networks the naive approach of assigning weights to the relations of individual pixels quickly has its limits. For images described with an $n \times m$ matrix the mapping of all input pixels to all output pixels is increasing the number of weights needed substantially with an increase in image size ($\propto (n \times m)^2$). At the same time this would require a corresponding increase in the available training data to find a good approximation for all unknown weights. Additionally, the increase in compute power and memory needed becomes unfeasible for regular workstations very quickly.

A potential solution to keep the number of weights small was introduced with the application of convolutional neural networks (CNN). The reasoning behind the use of convolution is based on the assumption that pixels that are far apart in an image are less correlated than pixels that are spaced closer together. This is assumed to be true for the entirety of the image.
A sliding window, named convolution kernel, with a finite size of $k \times l$ pixels is moved over the input matrix, learning only a small set $k \times l$ of weights. Only that set of weights is used to determine all entries of the output. It is worth noting that the number of weights here is independent of the input image dimension. Every single output pixel is only influenced by a small, constant fraction of the input image. For higher dimensional data, the kernel used can be extended to spread

over all dimensions or multiple kernels can be used in different directions. The weights of the kernels, also described as filters, are learned from the input data.
Down-sampling of the image can be achieved by implementation of a stride, where the sliding window is not moved from pixel to pixel continuously but with a higher step size, therefore reducing the output dimension. At the same time, the number of output channels (third dimension of the matrix) is usually increased. However, this is done in a way that keeps the total size of the generated matrix smaller than the input matrix. Thereby, the network learns a compact mapping for the given source data.
Similarly up-sampling of the data can be achieved through transposed convolution [95] with an according reduction of the third dimension.
The addition of layers to normalize the data after every convolution step, before the activation function, has been shown to increase the chances of good convergence in the optimization process [96].

3.1.2 Generative Adversarial Network

The invention of generative adversarial networks (GAN) [97] represents a breakthrough for the use of CNNs for image translation tasks [98].
In the supervised image translation task a GAN consists of two different network parts, one encoder-decoder network called the *generator* and a separate encoder network as the *discriminator* (see Fig. 3.1). The generator describes a CNN based network that is used for converting an input image to an output image.

The encoder part of the generator usually consists of convolutional layers with a stride of two and an increase of the number of channels of 2. In this way, the effective overall size of the matrix is halved in every convolution step. The number of encoding down-sampling layers varies in different implementations, where each layer is usually followed by a batch normalization and activation function. The abstract representation of the source data is extracted and fed through the lowest dimensional layer, the bottleneck, from the encoder to the decoder. In the following the decoder is up-sampling the data, decreasing the number of channels while increasing the output image dimensions.
The U-Net [99] represents a special version of the generator architecture, which uses skip connections, concatenating the output of the encoder to the input of the decoder with the appropriate dimensions (see Fig. 3.2 A). This resolves the issue

of the vanishing gradient problem [100] for deep CNN-networks in a similar way to ResNet [101] and is a way to transfer the high frequency signal in the raw data directly to the output.

The second part of the network, the discriminator, is usually a simple CNN network with a small number of down-sampling convolutional layers. The objective is, similar to a regular classification system, to distinguish between two sets of data. Depending on the specific GAN-model either the "real" input/output-pair (for a conditional GAN, cGAN [102]) that is being used to train the generator network or just the "real" input data (for a simple GAN) is supposed to be classified as true. The generated data by the generator network is trained to be labeled as fake by the discriminator.

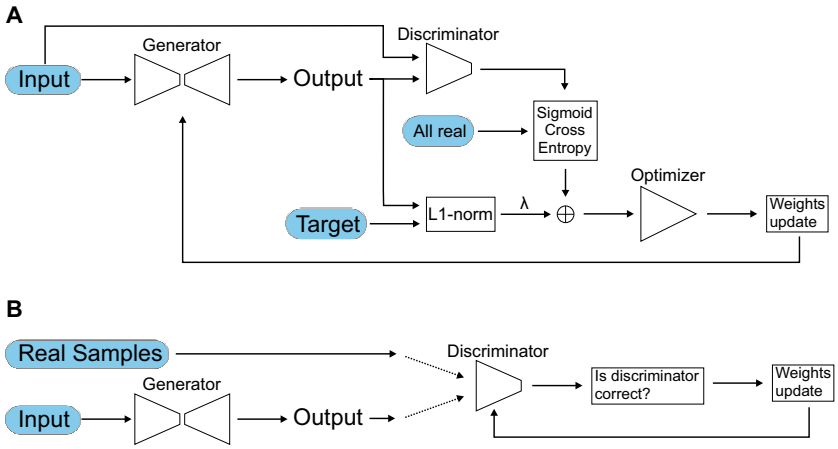

Figure 3.1: Training of a conditional GAN network. **A** The input of random or noisy data is fed into the Generator which is producing a corresponding output. The output is functioning together with the target as the input for the L1-loss function. Additionally the output from the generator is passed together with the original input into the discriminator. The output from the discriminator is passing a sigmoid cross validation with a vector of only ones and the value is added together, after a scaling with λ, with the L1-loss to yield the total loss. The optimizer is using the total loss to generate an update to the internal weights of the generator. **B** The discriminator is fed real samples from the data set and fake samples from the generator. The weights of the discriminator are updated based on the correctness of its classifications.

During training the generator G is tasked to approximate the true data distribution and to trick the discriminator D to falsely classify the generated data as "real",

based on an input image x from the distribution p_t and a random noise vector z from the distribution p_z. On the other hand the discriminator is trained to accurately distinguish between "fake" data $p_g(x)$ originating from the generator and "real" data $p_t(x)$. A scheme of the training of the generator is presented in Fig. 3.1 A. The weights of the generator are updated based on the discriminator output as well as a regularization parameter λ, estimating how closely the predicted output matches the target. Both models are trained alternately where the weights of the discriminator are updated with the generated data from the generators previous step (see Fig. 3.1 B). Both networks are growing stronger over time while competing with each other. This minimax game can be described by optimizing the following total loss function [103]

$$\begin{aligned} \min_G \max_D L(D,G) &= \mathbb{E}_{x\sim p_t(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \\ &= \mathbb{E}_{x\sim p_t(x)}[\log D(x)] + \mathbb{E}_{x\sim p_g(x)}[\log(1-D(x)] \end{aligned} \tag{3.1}$$

The loss function relates to the Jensen-Shannon (JS) divergence $D_{JS}(p\|q)$ with the loss function being $L(G,D^*) = 2D_{JS}(p_t\|p_g) - 2\log 2$ [103]. With an optimal discriminator D^* the loss function describes the similarity between the two distributions of real data and generated data. Since the JS divergence is bound by $[0,1]$ an optimum of the training can be achieved when $L(G^*,D^*) = -2\log 2$.
Due to the competing nature of the alternating training process the training of GANs can become unstable very easily where the individual loss functions start to oscillate. Different methods to stabilize and improve GAN training can be obtained from literature [103, 104].
The issue of potential mode collapse is where the generator is not learning anything, but continuously generates the same image for different input data to fool the discriminator. Mode collapse can be circumvented by adding an additional parameter to the loss function described in Eq. 3.1 which describes some distance of the generated to the target data. Typically the addition of L1-norm in the loss function of the generator with some weight λ is sufficient (see Fig. 3.1) [98].

Since the first publication on GANs many new GAN architectures with slightly different objective functions are reported for special use cases or for better training performances. An overview can be obtained from the literature [105].

3.1.3 Scope of this chapter

Weak contrast can be a significant drawback of super-resolution microscopy, especially in thick specimens. However, in many cases the underlying biological structures which are observed, are already known. This is for example the case in time series where the change in the structure (the movement) is of primary interest or in studies where merely the presence of the target is checked. In these cases the existing prior knowledge about the structure can be used to make informed decisions during the post-processing of the experimentally obtained data. Exploiting prior knowledge constrains the object space and can significantly reduce the uncertainty of estimated parameters. This is of special interest when the acquired images exhibit poor signal levels with potentially high background (low SBR), for example due to high STED intensities, fast imaging acquisition or simply due to a dim fluorophore/fluorescence labeling.
Classical image reconstruction tasks with the help of feature detection algorithms are mostly dependent on tuning a variety of parameters to achieve good results. These can differ between different images and are usually fixed on a single structure. Additionally, wavelet based filtering methods for ridge detection tasks can become quite resource intensive with long compute times for each image.

Neural networks have become a popular tool in many different fields [106, 107] including microscopy [87–90, 108]. They offer higher flexibility with parameter free application.
Therefore, in this chapter the application of neural networks is discussed as a tool for image reconstruction in STED microscopy. A neural network is trained to uncover the underlying, previously known, biological structure from simulated and experimentally acquired STED images. The limit of this approach is tested with a variation of the STED imaging parameters. Optimal imaging conditions to extract the most information from the data are determined.

3.2 Methods

3.2.1 Neural Network

The neural network that was employed was an adaptation of the conditional GAN named *pix2pix* [98] that is widely used for image translation tasks and showed the best initial results.

Different architectures of the GAN network can be used. In the most basic pix2pix network the generator consists of the classical U-net structure with eight two dimensional convolutional downsampling layers with a kernel size of 4 and stride of 2. The number of filters used in each step is increased accordingly by a factor of 2. A two-dimensional batch normalization layer follows each convolutional layer which is then passed through a leaky rectified linear unit (leaky ReLU) activation function with a negative slope of -0.2. This describes the architecture of the encoder. The decoder passes the data after the last convolutional layer through a series of eight, two dimensional transposed convolutional layer for upconversion purposes. A normalization layer (usually BatchNorm) and a ReLU activation function follow after each transposed convolution. The last layer is an exception which only features a tanh activation function without any normalization. The final output is the generated picture with the same dimensions as the input image. The filter-set before each transposed convolution step is concatenated with the filters originating from the corresponding down-convolution step. The upconversion is reducing the filter number by a factor of 2. A visual representation of the basic network architecture can be found in Fig. 3.2 A.

To resolve limitations with GPU memory, the image can also be cropped to a certain patch size before entering the encoder. Multiple of these patches can be stacked to form a batch. They are used together in one training iteration.

The individual convolutional blocks can be exchanged with ResNet blocks to form a classical ResNet architecture with variable depth [109].

Additionally, a three dimensional convolution network was built for image stacks. The previously described architecture was implemented with three dimensional convolutional and three dimensional transposed convolution layers. For similar reasons a long short term memory (LSTM) convolutional network was constructed by replacing the individual convolutional layer with a variant of ConvLSTM blocks [110].

The discriminator (see Fig. 3.2 B) is unchanged to the PatchGAN discriminator

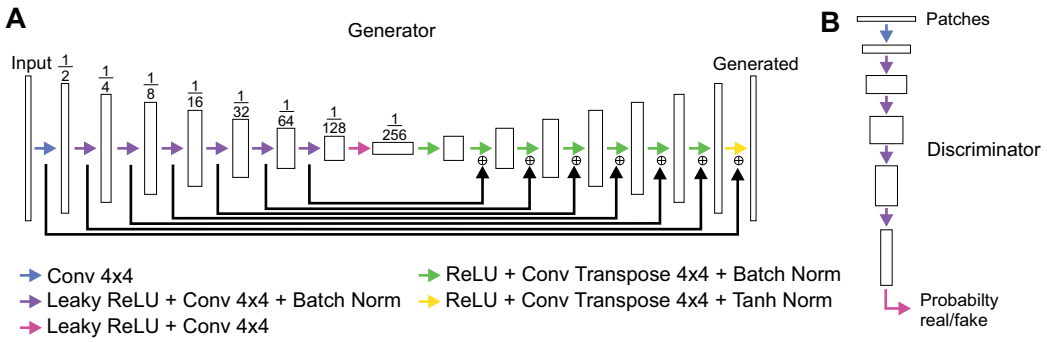

Figure 3.2: Architecture of the applied neural network. A Classic architecture of the U-net pix2pix network with purely convolutional blocks [98]. The image passes several convolutional layers and encodes the information in the filter set. After passing the lowest dimension the encoded data is fed into the decoder with convolutional transpose layers. The filter-set of the corresponding step in the encoder is concatenated to the output before passing the next layer. The decoder outputs a generated picture which is in turn fed into the discriminator **B**. Small patches of the generated image are passed through several convolutional layers before making the prediction if the passed image is real or artificially generated. **A, B** Individual convolutional blocks can be exchanged with three dimensional convolutional, LSTMConv or ResNet layers.

from the original publication [98] with a total of five layers.

Proposed images from the generator are passed to the discriminator and additionally compared to the original data during training. The comparison delivers one loss term (either L1 or L2-loss, see Eq. 3.2) which can be scaled freely with a parameter λ before the addition to the loss of fooling the discriminator (GAN-loss). AdamW was chosen as the optimizer with an optimizer momentum $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Detailed options for the optimizer can be obtained from the ConvNeXt framework [111].

For training the image batch size was varied from 1 to 10 with a random cropping of the images to sizes of ($256\,\mathrm{px}^2$ to $512\,\mathrm{px}^2$). The training was performed for a total of 400 epochs with half of them at a stable learning rate of 0.0002 and a linear decay of the learning rate for the second half.

Before the images were fed into the neural network they were augmented with the use of the Python package *Augmentor*. The dataset was overall increased five-fold with random rotations and random flipping enabled.

Normalization of the data could be achieved by either stretching the contrast, rescaling the intensities to the values between the 2 and 99.9 percentile, or by

changing the contrast locally in parts of the images.
Different metrics were calculated for validation. Typical metrics include mean structural similarity (mSSIM) index [112], which was calculated using the skimage library for Python, least absolute deviations (L1, or norm) and Least Square Errors (L2, Euclidian norm). L1 and L2 are just calculated as the sum of differences or differences squared from the individual pixels in the predicted image $\hat{y}_i$ to the target picture pixel y_i weighted by the overall number of pixels N (see Eq. 3.2).

$$\begin{aligned} \text{mSSIM} &= \frac{1}{M}\sum_{i}^{M}\text{SSIM}(x_i, y_i) = \frac{1}{M}\sum_{i}^{M}\frac{(2\mu_{x_i}\mu_{y_i}+c_1)(2\sigma_{x_i y_i}+c_2)}{(\mu_{x_i}^2+\mu_{y_i}^2+c_1)(\sigma_{x_i}^2+\sigma_{y_i}^2+c_2)} \\ L_1 &= \frac{1}{N}\sum_{i}^{N}|y_i-\hat{y}_i| \\ L_2 &= \frac{1}{N}\sum_{i}^{N}(y_i-\hat{y}_i)^2 \end{aligned} \tag{3.2}$$

N is the total number of pixel in the image. For the calculation of mSSIM the image was divided into M small sections where the SSIM is calculated with the mean value of all pixels μ the standard deviation σ and the covariance σ_{xy}. The constants c_1 and c_2 are dependent on the dynamic range of the input images. Typically they are defined as $c_1 = (0.01 \cdot L)^2$ and $c_1 = (0.03 \cdot L)^2$ where L is the dynamic range of the pixels ($L = 255$ for 8-bit and $L = 1$ when normalized) [113]. The typical metric of the normalised root mean square error (NRMSE) is just the square root of the L2-norm. The two-dimensional Wasserstein distance (also known as transportation theory in mathematics) calculates the cost function of transporting the pixels in the predicted image to the location of the pixels in the target image.

3.2.2 Hardware and Software

Training of the network, evaluation and predictions were run on two different systems. One workstation was equipped with 96 GB of RAM, an Intel Core i9-9820X, 3.30 GHz CPU and an NVIDIA GeForce RTX 3080 Ti with 12 GB of memory, the other one with 16 GB of RAM, an Intel Core i7-8650U, 1.90 GHz CPU and an NVIDIA GeForce RTX 3090 with 24 GB of memory.
Training of the network was performed with python 3.9.7 using PyTorch 1.10.0 and CUDA 11.3.

3.2.3 Simulation of microtubules

Microtubules were simulated using a two dimensional worm like chain (wlc) model in Python. The model includes a persistence length l_p and a total length of the microtubule. The persistence length is a measure of the stiffness of the generated chain, the chain is rigid in short term but flexible over long length scales. Starting from a vector in the xy-plane the following segment equals the previous segment except for a rotation by a random angle Θ. This angle is drawn from a probability distribution described with

$$p(\Theta) = \sqrt{\frac{l_p}{2\pi}} \exp\left(\frac{-l_p \Theta^2}{2}\right) \tag{3.3}$$

The sum of all vectors is resulting in the final chain.
One chain with length l_{tot} is generated for an image with size of n^2 pixel. The chain is folded onto the the image using periodic boundary conditions, generating multiple different strands of microtubules from one wlc chain. The density of the individual lines can therefore be controlled with the length of the generated wlc chain.
The single pixel wide structures are saved as the ground-truth data for the machine learning model and were increased in diameter for the generation of artificial microscopy pictures. The microtubules were convolved with a 2D Gaussian peak with FWHM = 4 px corresponding to the size of an effective STED-PSF. This results in a representation of a perfect STED image without any background or noise. To simulate different imaging intensities, photons were randomly drawn from a Poisson distribution where the expected number of events $\lambda(x, y)$ on the structure is modified by a variable sampling rate.
Additionally, an imaging background was added by drawing again photons from another Poisson distribution with a position independent mean λ_{BG}. STED induced background close to the structure was accomplished by first a convolution with another slightly larger PSF and then a random drawing of photons.
This way the signal from the microtubules and the background can be adjusted independently to for example reach a target SNR and SBR.

3.2.4 Ridge detection

Classic analysis of the data was achieved by ridge detection using the concept of α-molecules based on the work from [114]. As a generator function the mexican hat wavelet, the second derivative of a negative Gaussian, was chosen. The resulting system of α-molecules is then translated, scaled and rotated. The maximum feature width was set to 7, the minimum to 2. Additional parameters resulted to: scalesPerOctave = 2, number of orientations = 32, $j^0 = 1$, $\beta = 20$. Only ridges with positive contrast were considered as valid.
The thinned skeleton of the detected ridges was used for comparison.

3.2.5 Experimental data acquisition

Microtubules in living human fibroblast cells were stained with CP610-CTX at different concentrations.
Images were acquired with the setup described in subsection 2.4.1. The pixel dwell time was varied from 1 µs to 3 µs. Fast scans were performed in a bidirectional scanning mode where the laser beam was scanning over the sample in both directions. Detection of fluorescence was possible without scanner flyback. The timing adjustments necessary due to the response time of the galvo scanner system were performed on fluorescent bead samples. The timing was set such that the resulting image showed no distortions.

3.3 Results

The neural network that was chosen in this work is an adaptation of the pix2pix network which is fully described in subsection 3.2.1. This network architecture has been used as a basis in similar use cases in previous work [87–90] and delivered the most promising results during the initial trials.

The specific architecture design was modified to improve the training process. Different convolutional blocks were used (U-net vs. ResNet), different loss functions (L1, L2, etc.) and their corresponding weights showed to have an influence on the stability of the training. Classical hyperparameters that were changed include the batch size, patch size, learning rate and epoch count.
To optimize the network, a validation metric needs to be determined, which guides the selection of optimal parameters. Typical metrics for comparing two images, such as L1, NRMSE, Wasserstein and mSSIM are described in subsection 3.2.1. When comparing the different metrics on a single validation set, they all show a similar behavior with a fast improvement of the parameter and a subsequent flattening of the learning curve after roughly 100 training epochs. To select the metric of choice for further network optimization, a subset of five images was randomly selected at five different training epochs and judged by three humans to select the best reconstructed image. The human selection was equivalent with choosing the best images based on the mSSIM score. This score (as defined in Eq. 3.2) was therefore used further on for upcoming validation calculations.
The training of the network was performed on 500 unique images. These images were augmented as described in subsection 3.2.1. Calculation of the validation metrics was also performed on a set of 500 images.

Due to the fact that the underlying biological structure served as the target of the image reconstruction the validation metrics could not be calculated on experimentally obtained data. This could have been only possible if the target of the reconstruction was a STED image itself. Therefore, the image data of microtubules was simulated as described in subsection 3.2.3 for the following training optimization steps.

In the beginning the influence of different pre-processing steps was investigated on the outcome and stability of the neural network training. As demonstrated in

Fig. 3.3 A, stretching the contrast of the input data results in an unstable training which is fluctuating heavily. The overall trained final model is worse compared to the unmodified data-set. Similarly, adaptive equalization, where the histogram of the image is stretched in small patches of the image, leads to a worse training result with a significant drop in the last few epochs. Normalizing the data as well as averaging the data with an effective mean filter has little effect on the training and is not resulting in an increased performance. The difference in patch size, edge length of 512 px vs. 256 px, seems to have no influence on the training either. Therefore, for further trainings no pre-processing of the data was applied and the patch size was kept at $(256\,\mathrm{px})^2$.

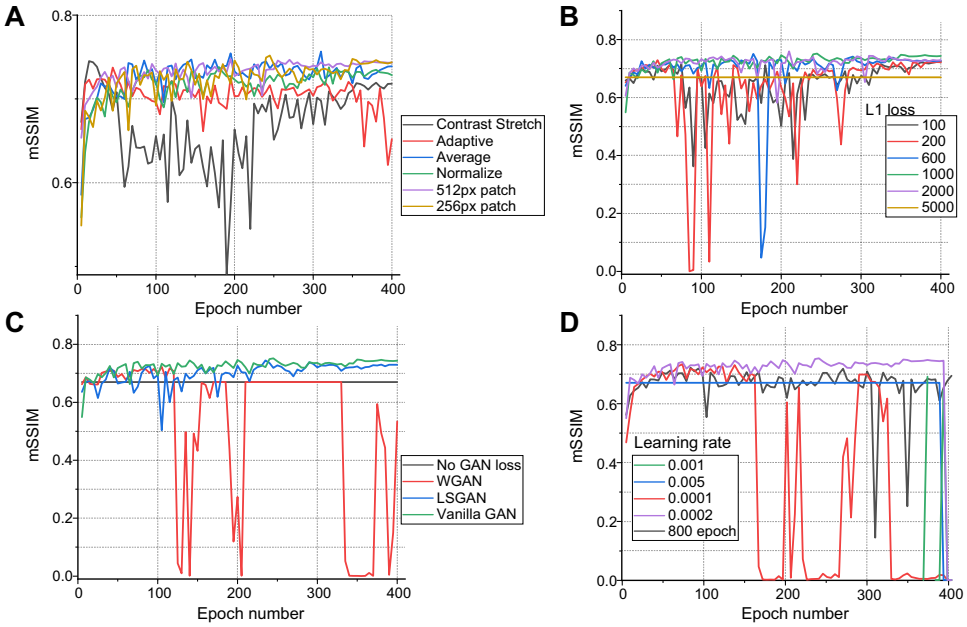

Figure 3.3: Effect of different parameters on the neural network training. **A** Effect of pre-processing of the data on the training performance. The mSSIM is shown as a function of the training epoch. Different pre-processing methods don't seem to show a significant increase in the training performance. **B** The training efficiency is tested with different L1-loss multipliers. An increase of the loss multiplier above 600 results in a stabilization of the training, as long as the value is kept below 5000. **C** A change in the GAN-loss away from the vanilla GAN is not resulting in a better training performance. **D** Change of the learning rate shows highest evaluation metric for a learning rate of 0.0002.

Next the weight of the L1-loss within the overall loss function of the GAN was in-

vestigated and optimal parameters determined (see Fig. 3.3 B). Network training with L1 multipliers of 100 (black), 200 (red) and 600 (blue) show a lot instabilities in early epochs (< 300). Results with these multipliers can often show some degradation of the trained structure, especially on challenging datasets. Usually this manifests itself as hallucinated periodic structures, starting at the edges of the images, creeping into the center with progressing epoch count. Loss multipliers of 1000 (green) and 2000 (purple) work to stabilize the training and to eliminate the described unwanted effects by focusing more on the penalization of the deviation of the proposed structure from the target. A higher loss multiplier of 5000 (brown) leads to no training and a constant loss which is not increasing. For further trainings the L1 loss multiplier of $\lambda = 1000$ is chosen.
The constant evaluation metric of a loss multiplier of 5000 is due to the fact that the model is predicting all images with identical gray pixel values. This delivers higher evaluation metrics than false structure predictions, but this metric is stable and not increasing since the model is trapped in a local minimum which requires higher losses for its escape.
Similarly the GAN loss function can be exchanged. Without any GAN loss the training is ineffective, which can be seen in Fig. 3.3 C in black. Changing to Wasserstein GAN (WGAN) makes the training very erratic and reduces the performance substantially. Least squares GAN (LSGAN) results in a stable but slightly worse training performance than the so called vanilla GAN, which is the loss function of the original GAN publication. Therefore the vanilla GAN is kept for subsequent network trainings.

The optimal learning rate has to be evaluated, together with the best epoch length. With a low learning rate the model needs longer to converge, whereas a learning rate that is too high might fail to find the optimum. Different starting learning rates were tested. A learning rate of 0.005 is too high and leads to no learning at all as can be seen in Fig. 3.3 C in blue. A learning rate of 0.0001 leads to some initial success, but then collapses and results in an unfit model. A learning rate of 0.001 is quite stable but performs worse than a model with a learning rate of 0.0002. Even for a higher epoch count up to 800 no real improvement is detectable after $\sim$300 epochs.
Therefore, a training with an initial learning rate of 0.0002 and an epoch length of 400 was chosen. This epoch length seemed to be a good number for reliably achieving a training that already plateaued.

It is worth to note that the change of the batch size had no effect on the training outcome and could therefore be optimized for the individual GPU to maximize training speed.

Real world STED images rarely exhibit the same SBR over different samples, due to different labeling densities or related factors. Similarly, the density of the imaged biological structures are highly variable even within a single sample. Especially when changing the EPFS by increasing the STED beam intensity, the SBR as well as the density of the structure get reduced. The previous optimization steps presented so far were performed on a fixed SBR and a fixed density.
In order to estimate the effect of these two additional variables on the training and subsequent reconstruction performance, the SBR of the training and validation set was changed from a minimum SBR of 1.5 to a maximum of 30. As can be seen in Fig. 3.4 A, the validation metric of mSSIM (color-scale) is increasing for all trainings (x-axis) with a higher validation set SBR (y-axis). The training of a model with input data exhibiting a lower SBR results also in a better performance for predicting the structure on low SBR validation data. An overall increase in performance for lower SBR validation datasets is visible when reducing the training SBR until reaching an SBR of 4. For lower SBR values during the training the validation metric gets slightly worse. Interestingly, for a training SBR of 3 the structure prediction seems to perform worse on low SBR images compared to higher SBR value of 4 training data as well as lower SBR = 2 input data.
Based on these results a training SBR of 4 seems optimal. However, in cases where the expected data for structure prediction has an SBR lower than 3, a model that was trained with an SBR of 2 might result in slightly better performances.

Similarly, the density of the structure occupying the imaged segment can be changed. This is achieved by varying the length of the generated object chain described in subsection 3.2.3. The density for a given STED FWHM can then be described by $\rho = l \cdot \mathrm{FWHM} A^{-1}$ with l being the length of the chain and A the total size of the image ($A = (512\,\mathrm{px})^2$ and FWHM = 4 for simulated dataset). The structure prediction performance of different densities is depicted in Fig. 3.4 B. The SBR that was used for the generation of this dataset was fixed to 20 as it high enough to single out the density only influence. The metrics for successful prediction of the underlying structure gets worse for higher density validation datasets

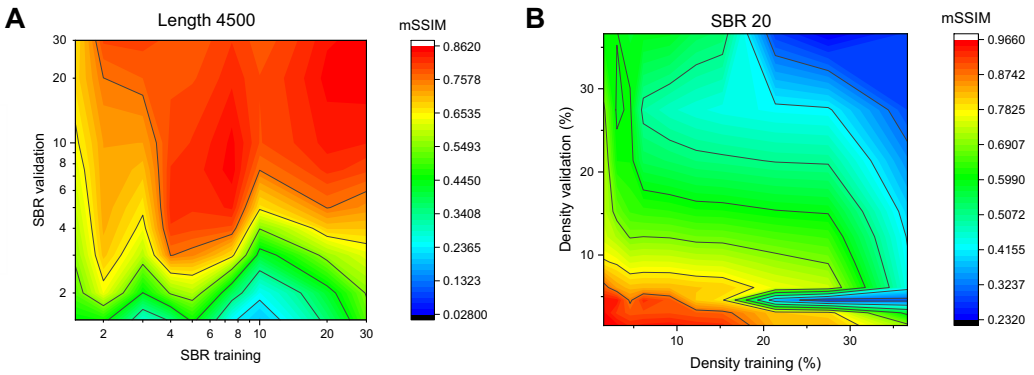

Figure 3.4: Effect of SBR and density of the data on training and prediction. A Effect of the SBR of the input data on the reconstruction performance on validation datasets with different SBR. The x-axis represents the SBR of the training data, the y-axis shows the SBR of the validation data. **B** Effect of the simulated chain length of the input data on the reconstruction performance on validation datasets with different chain lengths. The variation of the worm-like chain length corresponds to a variation in the density of the underlying structure. The x-axis represents the occupation density of the chain in relation to the whole image in percent of the training data, the y-axis shows the density of the validation data. Contour lines correspond to the labeled values at the color scale bar.

for all trainings. Similarly, the metrics drop for all validation datasets for trainings with an increase in the length of the chain and corresponding density of the structure. More dense structures are more difficult to reconstruct. It is interesting however, that a training on less dense samples is resulting in a better performance on more dense samples than models that were trained on equally high dense samples. This suggests, that during training of very dense samples the network is not able to learn the important features of the underlying structure sufficiently. It seems therefore that there is no benefit in training on dense samples when lower density trainings can deliver a better final model.
As a peculiarity, structure prediction on validation data with $\sim 4.5\,\%$ occupancy performs much worse than densities close-by, for models trained on high densities.

An example of the input data and the corresponding predicted structure is shown in Fig. 3.5. The neural network is trained using data with an SBR of 2 and a density of $\sim 6.9\,\%$. The predicted structure in Fig. 3.5 B is closely matching the actual simulated structure in Fig. 3.5 C that is used to generate the input data in Fig. 3.5 A. In cases where the structure is not resolvable with the used FWHM of

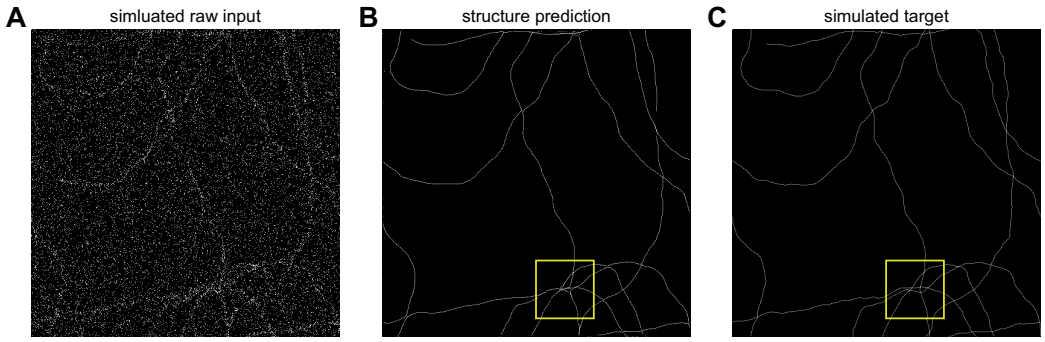

Figure 3.5: Example of structure prediction on simulated noisy data. A Simulated noisy data of microtubules with an SBR of 2 and a density of $\sim 6.9\,\%$. **B** The predicted underlying structure corresponding to the sample presented in A. The Neural Network for the prediction was trained on input data with an SBR of 2. **C** The simulated target data of the structure of the noisy data shown in A. A region where two microtubules run very close is marked in yellow.

the microscope the prediction fails, because the information about the accurate location of the microtubule can not be obtained from the data. This is visible in the lower half where two microtubules run very close (closer than FWHM) side-by-side before splitting again (marked in yellow in Fig. 3.5). Similar image reconstruction performance is achieved on data which exhibits an SBR of 2 with high signal and high background (see Fig. B.1).

Previous variations of the input training data, in terms of SBR and density, show to have an effect on the performance of the final model. Before, we tested only what influence the variation of one of these parameter has on the resulting trained model. In practice however, applying STED power to increase the resolution in the measurement decreases the fluorescence signal. Therefore, the SBR gets reduced due to the smaller EPSF and additional bleaching of the fluorophores, as described earlier. The cumulative influence of the variation of the SBR and density on the training and validation was evaluated. The training is performed on 30 different datasets spanning an SBR from 2 to 30 and sampling the density of the total space covered with microtubules from $1.5\,\%$ to $30.5\,\%$. Each model that was trained with one of these datasets is evaluated on all of them.
The results show, that the overall best performing model was trained on an SBR of 2 and an occupation density of $1.5\,\%$. The best performance is defined as the highest average evaluation score for all predictions on all of the different SBR and

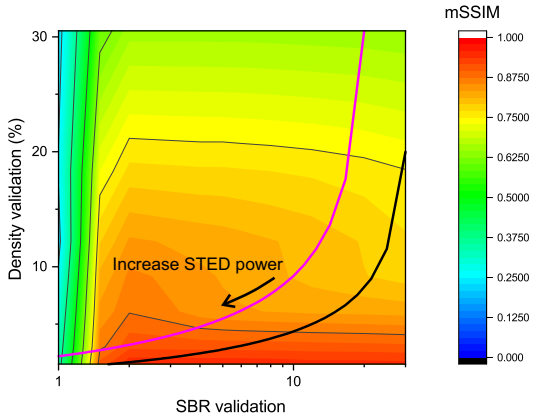

Figure 3.6: Evaluation metrics on different validation data. The neural network was trained with an SBR of 2 and a 1.5 % occupation density of the structure. The evaluation was performed on different validation datasets with different SBRs and densities. The mSSIM is plotted for all validations. The bold black and purple line indicate the trajectory of varying the applied STED laser power on two samples with a different starting occupation and fluorescent labeling density. Contour lines correspond to the labeled values at the color scale bar.

density combinations. The evaluation metrics map, spanning a more densely sampled SBR from 1 to 30, is shown in Fig. 3.6. The model provides almost uniform high evaluation metrics for low densities up to 20 % for all SBR values above 2. The training starts to break down for higher occupation densities. Unsurprisingly, the model, even though trained on low SBR data, is able to predict structures for high SBR almost equally well as models trained with high SBR data. For predictions on low SBR data the models trained with a higher SBR start to fail and produce no or faulty data (see evaluation maps in Fig. B.2).
For input data approaching an SBR of 1 the predictions start to fail. However, predictions from the network for low density data with an SBR of 1.5 are still not completely faulty. The model predicts the majority of the structure with some gaps. The visual inspection of the resulting images shows that the network is marking very few pixels falsely as structure where no structure exists in the target data. The degradation of the evaluation metric seems to be mostly driven by the under assignment of pixels to the structure that is present in the underlying distribution.

The trend of decreasing accuracy for predictions on higher density validation data is visible for all different training variations. The models deliver successively worse

results with a higher density of the structure of interest either in the training or validation data. The evaluation maps for all trained models are shown in Fig. B.2.

By varying the STED power used for imaging both the SBR as well as the occupation density are both decreasing. An increase in STED power is increasing the resolution by a decrease of the EPSF according to the square root law [73], effectively reducing the density of the underlying target structure. At the same time an increase in STED power is resulting in a decrease of the detectable fluorescent signal due to the shrinking of the EPSF and an increased bleaching, resulting from the high intensities in the doughnut crest. However, in this work the bleaching is not considered as it is dependent on many additional factors, such as pixel dwell time, scanning pattern, the specific dye used etc. [74, 115, 116].
Samples can vary in their initial SBR based on the initial fluorescence and in their starting density corresponding to the structure of interest. Two trajectories of a variation of the STED power are indicated in Fig. 3.6 by the bold black and purple line. The density scales according to the square root law $d(P) = d_0 * \frac{1}{1+\sqrt{P_{\mathrm{STED}}}}$, similarly the signal scales with $s(P) = s_0 * \frac{1}{1+P_{\mathrm{STED}}}$.
Walking along this line on the two dimensional plane spanning different SBRs and densities of the structure, an optimum of the used STED power and required resolution in terms of structure prediction performance can be found. Judging by the strong degradation of performance for SBR < 1, the highest STED power is the one that lowers the SBR to 2.

An example of the line through the metrics plane (pink line in Fig. 3.6) is shown in Fig. 3.7. An increase in applied STED power is resulting in an improved resolution, as shown in red. The increase in resolution is first leading to an increase of the evaluation metrics until a point is reached where the SBR is dropping too low and the performance of the model is rapidly decreasing. This point is the optimal STED power that can be used with this sample. The individual STED power that is needed can vary between different samples, depending on the starting fluorescent signal and density of the structure in confocal, as described above.

After training and optimizing the model, it was tested on experimental acquired data. For this purpose microtubules were stained in human fibroblasts with Abberior LIVE 610 and imaged with short dwell times of $\sim 3\,\mu\mathrm{s}$ and a STED power of 80 mW. The gated STED images show only very few counts per frame as can be

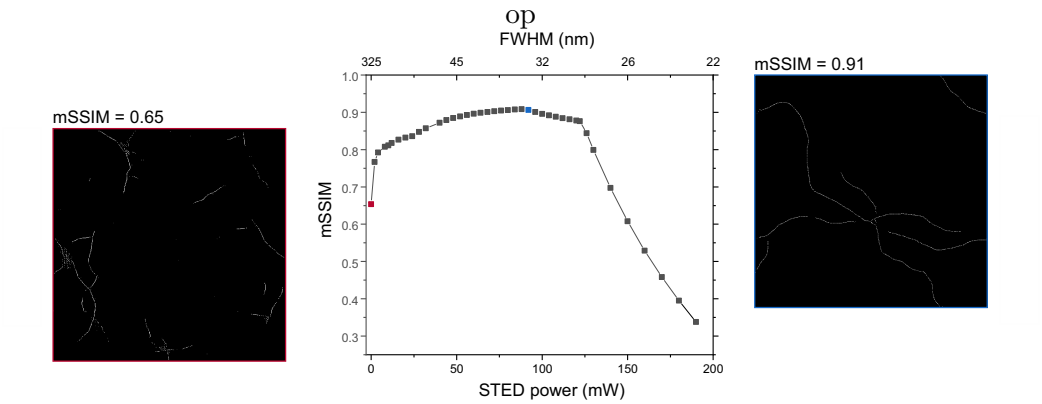

Figure 3.7: Evaluation metrics as a function of the applied STED power. The power is given in mW for an example case with a starting SBR of 20 in the confocal mode. The increase in resolution is shown as the second x-axis on the top. Two prediction examples are shown on the left and right corresponding to two different mSSIM evaluation metrics. The corresponding data-points on the graph are marked in red and blue.

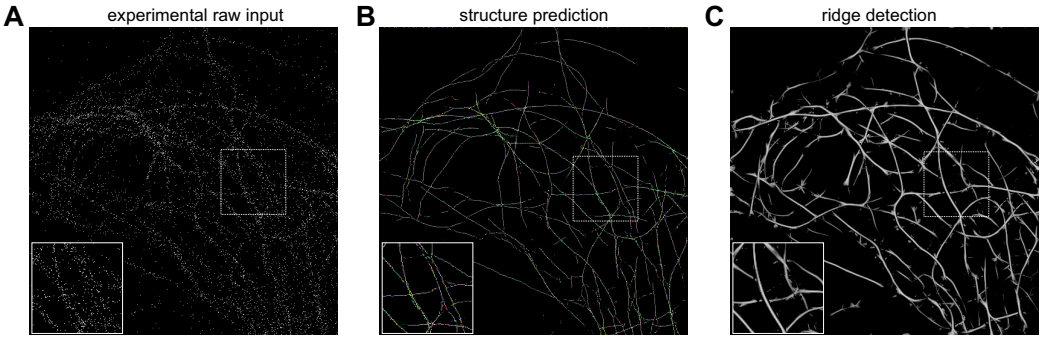

Figure 3.8: Structure prediction on experimental data of microtubules. Microtubules are stained in human fibroblasts with Abberior LIVE 610. Images were acquired with a dwell time of 3 µs and a STED power of 80 mW. **A** Shows one frame of the experimentally acquired STED images. **B** Shows the predicted structure from the neural network. **C** As a comparison an analysis with a special wavelet filter for ridge detection. The location of the enlarged sections shown in the bottom left corner is marked with a dotted white outline.

seen in Fig. 3.8 A. The structure prediction of the neural network model is shown in Fig. 3.8 B. The estimated position of the microtubules seems to be a very reasonable estimate and very few false assignments are made. Since these cells were imaged live there is no ground truth reference.

A classical ridge detection algorithm is used as a comparison (see Fig. 3.8 C). This

wavelet based algorithm is similarly able to find most of the microtubules but performs worse at the boundaries of the localizations. Especially in dense areas of the sample the neural network prediction outperforms the classical approach, being able to find more of the individual microtubule strands.

3.4 Discussion

The goal of this chapter was to determine if previous knowledge about the underlying biological structure in an acquired low contrast STED image can help in image post-processing.
The results show that a reconstruction of the structure is successful using neural networks, in this case an adaptation of the pix2pix architecture. The parameters for training such a CNN based GAN can be optimized to achieve a stable learning with high image translation accuracy. In order to achieve the most accurate prediction for all imaging conditions, it is best to use input training data with an SBR of 2 and a low occupation density of 1.5 %.
The fact that training on a low SBR delivers the best overall performance is not surprising, since a successfully trained model at this SBR is expected to filter the comparatively large noise contribution more effectively than models trained on high SBR data. The low density training data might be helping with a better abstraction of the underlying structure and therefore delivering the best performance for all SBR values. It also seems that once the detection of the structure is successfully trained on low SBR data, the SBR of the data that is used for analysis has little effect. The model does not seem to perform much better on high SBR data. The number of trainable weights and kernels is apparently high enough to achieve a good denoising for the structure reconstruction task. As a result, a higher signal over background noise would have no additional benefit in real experiments.

This thesis demonstrates that the model can be applied to STED imaging. The optimal STED power for a given sample is the power which decreases the SBR below a value of 2. The model shows a good performance at much lower SBR values than previously reported [89]. An increase in resolution is being compensated by a reduction in prediction performance for values below an SBR of 2. The increase in the applied imaging STED power until this point is always beneficial on dense structures, since it reduces the total occupied space by the structure of interest, increasing the model's performance. On structures that are very sparse from the start the benefit is reduced. The increase in STED power is going in tandem with an increase in the information present in the imaged sample, as long as the SBR is high enough.
This optimum in STED imaging power has to be evaluated on each sample individually, since it is highly dependent on the initial fluorescence present in the sample

and the needed resolution improvement to distinguish the objects of interest.

The application of the neural network presented here can also be used to predict the structure in experimentally acquired images. The network is able to identify the structure based on very few counts from a STED image, showing overall very few false predictions, outperforming a wavelet based ridge detection in dense areas. Additionally, the neural network approach is much faster after the initial training is done. Every prediction is done in a few seconds, where the ridge detection based on a wavelet filtering is a multistep algorithm requiring much more time and compute power for every single image.
Classic algorithms, like the ridge detection used here, usually need a careful selection of different parameters for achieving the best result. In contrast, once the neural network model is trained, there are no additional parameters that need to be adjusted for its application. This makes the neural network approach much easier for an untrained end user and large datasets.

For a better applicability, the simulations of microtubules in this work can be further improved. So far, after the convolution of the structure with the STED FWHM, a sample of photons is drawn from a Poisson distribution to simulate different signal levels. More realistically, a random sample of fluorophores attached to the structure could be drawn first to simulate different in-homogeneous labeling densities present in real experimental data. Additionally, simulating a more inhomogeneous background can make the data more realistic. Those adjustments might further improve the application to real acquired data.

After the neural network training was optimized for one structure at different SBRs and densities, the prediction can also be extended to other structures in a possible next step. Currently, the network will predict line segments wherever it can and is not able to generalize to other structures yet. The training of networks with a mixture of structures can easily lead to a confusion of the prediction results. Therefore, the optimal training strategy has to be evaluated. Simple encoding of the target structure in an additional one hot vector, like in segmentation tasks for pix2pix, represents the easiest but the least flexible solution [117, 118].

The network is already at this moment able to predict the structure better than conventional approaches, as shown with the comparison to the ridge detection for

microtubules. But with the ability to identify additional different shapes, a neural network would be able to outperform any classical approach due to its flexibility alone.

Similarly, continuing the previous work to include movies is a possible next step. In a fast acquired movie the structure itself is not changing drastically from one frame to the next. As a result, lower SBR images are enough to predict the movement of the structure since only the difference to the previous frame has to be determined. Additionally, some higher SBR images can be obtained as intermediates as a result of a variable scanning speed. They can serve as anchors to help to improve the predictions to correct slight misjudgments of the network made in-between them (similar to [119, 120]). The optimal STED power in these scenarios has to be determined as a function of the scanning speed and bleaching behavior.
Different networks are optimized for these purposes and already used for analyzing movies. They include feature warping of detected objects to quickly predict the movement of objects [121] or optical flow [122]. Similar principles can be applied for movie structure predictions.

On a similar note, new neural network architectures are published almost on a monthly basis. New architectures utilizing vision transformers [123] or slight modifications to the classical CNN [111] provide better performance on reference datasets such as ImageNet. If this improvement translates to better predictions for input data, as presented in this work needs to be confirmed separately. Faster and more accurate reconstructions can help to image increasingly difficult samples.

4. Conclusion and outlook

This thesis demonstrates two different solutions to deal with loss of contrast in super-resolution microscopy.
The fist part, described in chapter 2, presents the application of 2PA with the subsequent use of different super-resolution microscopy techniques. The contrast loss due to bleaching of a high intensity laser beam during image acquisition can be mitigated by using 2PA. The 2PA with (far-)red dyes is presented in the visible regime at a wavelength of 515 nm. As discussed in the chapter, the method is superior to classical UV 1PA mainly due to its restriction in activation volume. The method can be applied in a variety of biological samples but shows the biggest contrast enhancement in comparison to 1PA in thick, axially extended specimens. If the application of 2PA is not viable, the chapter 3 discusses the use of a cGAN for low signal STED images. The cGAN network is shown to be able to reconstruct the acquired structure, even on low SBR STED image data. The model outperforms classical image reconstruction algorithms and is more universal in its application. A recovery of the structure that is present in low SBR data can help in movie acquisitions or other extended imaging sessions.
The two methods presented in this thesis approach the loss in contrast differently. 2PA tries to retain the signal during the acquisition process, whereas the application of the cGAN network solves the contrast loss during image post-processing. Particularly, since the utilization of the cGAN aided image reconstruction can be done retroactively, it can additionally be applied on previously acquired datasets.

The findings described in this thesis show some means to counteract signal degradation in microscopy. Better dyes and more sophisticated scanning methods to boost the fluorescence signal might help to alleviate some of the problems in the future. However, the presented methods remain to be of interest. Moreover, they resolve a variety of other currently existing drawbacks of classical techniques. 2PA

at 515 nm shows better transmission and less chromatic aberrations compared to UV photoactivation, whereas the application of a neural network reduces the number of user-adjustable free parameters in image post-processing significantly.

Bibliography

[1] Robert Hooke. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses: With observations and inquiries thereupon.* 6th ed. London: Printed by J. Martyn and J. Allestry, 1665.

[2] E. Abbe. "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung". In: *Archiv für Mikroskopische Anatomie* 9.1 (1873), pp. 413–468.

[3] Stefan W. Hell and Jan Wichmann. "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy". In: *Opt. Lett.* 19.11 (June 1994), pp. 780–782.

[4] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)". In: *Nature Methods* 3.10 (2006), pp. 793–796.

[5] Eric Betzig et al. "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313.5793 (2006), pp. 1642–1645.

[6] C. Eggeling et al. "Photobleaching of Fluorescent Dyes under Contitions Used for Single-Molecule Detection: Evidence of Two-Step Photolysis." eng. In: *Analytical chemistry* 70 (13 July 1998), pp. 2651–2659.

[7] Gerald Donnert, Christian Eggeling, and Stefan W. Hell. "Triplet-relaxation microscopy with bunched pulsed excitation". In: *Photochem. Photobiol. Sci.* 8 (4 2009), pp. 481–485.

[8] Gerald Donnert, Christian Eggeling, and Stefan W. Hell. "Major signal increase in fluorescence microscopy through dark-state relaxation." eng. In: *Nature methods* 4 (1 Jan. 2007), pp. 81–86.

[9] Giuseppe Vicidomini et al. "Sharper low-power STED nanoscopy by time gating". In: *Nature Methods* 8.7 (2011), pp. 571–573.

[10] Jeffrey R. Moffitt, Christian Osseforth, and Jens Michaelis. “Time-gating improves the spatial resolution of STED microscopy”. In: *Opt. Express* 19.5 (Feb. 2011), pp. 4242–4254.

[11] Yvan Eilers et al. “MINFLUX monitors rapid molecular jumps with superior spatiotemporal resolution”. In: *Proceedings of the National Academy of Sciences* 115.24 (2018), pp. 6117–6122.

[12] Francisco Balzarotti et al. “Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes”. In: *Science* 355.6325 (2017), pp. 606–612.

[13] Michael Weber et al. “MINSTED fluorescence localization and nanoscopy”. In: *Nature Photonics* 15.5 (Oct. 2021), pp. 361–366.

[14] Fritjof Helmchen and Winfried Denk. “Deep tissue two-photon microscopy”. In: *Nature Methods* 2.12 (2005), pp. 932–940.

[15] Michael Rubart. “Two-Photon Microscopy of Cells and Tissue”. In: *Circulation Research* 95.12 (Dec. 2004), pp. 1154–1166.

[16] Karel Svoboda and Ryohei Yasuda. “Principles of Two-Photon Excitation Microscopy and Its Applications to Neuroscience”. In: *Neuron* 50.6 (2006), pp. 823–839.

[17] Willem F. Veldhuyzen et al. “A Light-Activated Probe of Intracellular Protein Kinase Activity”. In: *J. Am. Chem. Soc.* 125.44 (Nov. 2003), pp. 13358–13359.

[18] Jennifer Lippincott-Schwartz, Nihal Altan-Bonnet, and George H. Patterson. “Photobleaching and photoactivation: following protein dynamics in living cells.” In: *Nature cell biology* (2003), S7–14.

[19] Graham C. R. Ellis-Davies. “Caged compounds: photorelease technology for control of cellular chemistry and physiology”. In: *Nature Methods* 4.8 (2007), pp. 619–628.

[20] Gerard P. Ahern, Shyue-Fang Hsu, and Meyer B. Jackson. “Direct actions of nitric oxide on rat neurohypophysial K+ channels”. In: *The Journal of Physiology* 520.1 (1999), pp. 165–176.

[21] Alexey N. Butkevich et al. “Photoactivatable Fluorescent Dyes with Hydrophilic Caging Groups and Their Use in Multicolor Nanoscopy”. In: *Journal of the American Chemical Society* 143.44 (2021), pp. 18388–182393.

[22] H. Pal, M. Sliwa, and Françisco M. Raymo. “Photoactivatable Fluorophores”. In: *ISRN Physical Chemistry* 2012 (2012), p. 619251.

[23] Michelle S. Frei et al. “Photoactivation of silicon rhodamines via a light-induced protonation”. In: *Nature Communications* 10.1 (2019), p. 4580.

[24] Maria Göppert-Mayer. “Über Elementarakte mit zwei Quantensprüngen”. In: *Annalen der Physik* 401.3 (1931), pp. 273–294.

[25] Robert R. Birge and Brian M. Pierce. “Semiclassical time-dependent theory of two-photon spectroscopy. The effect of dephasing in the virtual level on the two-photon excitation spectrum of isotachysterol”. In: *International Journal of Quantum Chemistry* 29.4 (1986), pp. 639–656.

[26] B. Richards, E. Wolf, and Dennis Gabor. “Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 253.1274 (1959), pp. 358–379.

[27] C. J. R. Sheppard and H. J. Matthews. “Imaging in high-aperture optical systems”. In: *J. Opt. Soc. Am. A* 4.8 (Aug. 1987), pp. 1354–1360.

[28] Gianguido C. Cianci, Jianrong Wu, and Keith M. Berland. “Saturation modified point spread functions in two-photon microscopy”. In: *Microscopy Research and Technique* 64.2 (2004), pp. 135–141.

[29] Winifried Denk, James H. Strickler, and Watt W. Webb. “Two-Photon Laser Scanning Fluorescence Microscopy”. In: *Science* 248.4951 (1990), pp. 73–76.

[30] Pavel Loiko et al. “Sellmeier equations, group velocity dispersion, and thermo-optic dispersion formulas for $CaLnAlO_4$ (Ln = Y, Gd) laser host crystals”. In: *Opt. Lett.* 42.12 (June 2017), pp. 2275–2278.

[31] O. E. Martinez, J. P. Gordon, and R. L. Fork. “Negative group-velocity dispersion using refraction”. In: *J. Opt. Soc. Am. A* 1.10 (Oct. 1984), pp. 1003–1006.

[32] Selcuk Akturk et al. “Extremely simple single-prism ultrashort-pulse compressor”. In: *Opt. Express* 14.21 (Oct. 2006), pp. 10101–10108.

[33] R. L. Fork, O. E. Martinez, and J. P. Gordon. “Negative dispersion using pairs of prisms”. In: *Opt. Lett.* 9.5 (May 1984), pp. 150–152.

[34] J. P. Gordon and R. L. Fork. “Optical resonator with negative dispersion”. In: *Opt. Lett.* 9.5 (May 1984), pp. 153–155.

[35] Hongying Liu et al. “Dispersion compensation based on prism compressor”. In: *Optical Manipulation Conference*. Ed. by Takashige Omatsu. Vol. 10252. International Society for Optics and Photonics. SPIE, 2017, pp. 123–126.

[36] F. J. Duarte. “Generalized multiple-prism dispersion theory for pulse compression in ultrafast dye lasers”. In: *Optical and Quantum Electronics* 19.4 (1987), pp. 223–229.

[37] Ruobing Zhang et al. “Analytical expressions of group-delay dispersion and cubic phase for four-prism sequence used at other than Brewster’s angle”. In: *Optics & Laser Technology* 31.5 (1999), pp. 373–379.

[38] T. L. Koch and J. E. Bowers. “Nature of wavelength chirping in directly modulated semiconductor lasers”. English. In: *Electronics Letters* 20 (25 Dec. 1984), pp. 1038–1040.

[39] Richard K. P. Benninger and David W. Piston. “Two-photon excitation microscopy for the study of living cells and tissues.” eng. In: *Current protocols in cell biology* 59.1 (June 2013), pp. 4.11.1–4.11.24.

[40] M. Sonnleitner, G.J. Schütz, and Th. Schmidt. “Imaging individual molecules by two-photon excitation”. In: *Chemical Physics Letters* 300.1 (1999), pp. 221–226.

[41] Alberto Diaspro and Mauro Robello. “Two-photon excitation of fluorescence for three-dimensional optical imaging of biological structures”. In: *Journal of Photochemistry and Photobiology B: Biology* 55.1 (2000), pp. 1–8.

[42] Gael Moneron and Stefan W. Hell. “Two-photon excitation STED microscopy”. In: *Opt. Express* 17.17 (Aug. 2009), pp. 14567–14573.

[43] Srinivas Kantevari et al. “Two-color, two-photon uncaging of glutamate and GABA”. In: *Nature Methods* 7.2 (2010), pp. 123–125.

[44] Graham C. R. Ellis-Davies. “Two-Photon Uncaging of Glutamate”. In: *Frontiers in Synaptic Neuroscience* 10 (2019).

[45] Dan Oron et al. “Two-photon optogenetics”. In: *Optogenetics: Tools for Controlling and Monitoring Neuronal Activity*. Ed. by Thomas Knöpfel and Edward S. Boyden. Vol. 196. Progress in Brain Research. Elsevier, 2012. Chap. 7, pp. 119–143.

[46] Adam M. Packer et al. “Two-photon optogenetics of dendritic spines and neural circuits”. In: *Nature Methods* 9.12 (2012), pp. 1202–1205.

[47] YuRui Zhao et al. “New Caged Coumarin Fluorophores with Extraordinary Uncaging Cross Sections Suitable for Biological Imaging Applications”. In: *J. Am. Chem. Soc.* 126.14 (Apr. 2004), pp. 4653–4663.

[48] Marc Schneider et al. “Two-photon activation and excitation properties of PA-GFP in the 720-920-nm region.” eng. In: *Biophysical journal* 89 (2 Aug. 2005), pp. 1346–52.

[49] Sheldon J. J. Kwok et al. “Two-photon excited photoconversion of cyanine-based dyes”. In: *Scientific Reports* 6.1 (2016), p. 23866.

[50] Andrew M. Smith, Michael C. Mancini, and Shuming Nie. “Second window for in vivo imaging”. In: *Nature Nanotechnology* 4.11 (2009), pp. 710–711.

[51] Michael Weber et al. “Photoactivatable Fluorophore for Stimulated Emission Depletion (STED) Microscopy and Bioconjugation Technique for Hydrophobic Labels”. In: *Chemistry - A European Journal* 27.1 (2020), pp. 451–458.

[52] Toshiki Kubo et al. “Visible-Wavelength Multiphoton Activation Confocal Microscopy”. In: *ACS Photonics* 8.9 (Sept. 2021), pp. 2666–2673.

[53] I. H. Malitson. “Interspecimen Comparison of the Refractive Index of Fused Silica”. In: *J. Opt. Soc. Am.* 55.10 (Oct. 1965), pp. 1205–1209.

[54] Marilyn J. Dodge. “Refractive properties of magnesium fluoride”. In: *Appl. Opt.* 23.12 (June 1984), pp. 1980–1985.

[55] Jeffrey B. Guild, Chris Xu, and Watt W. Webb. “Measurement of group delay dispersion of high numerical aperture objective lenses using two-photon excited fluorescence”. In: *Appl. Opt.* 36.1 (Jan. 1997), pp. 397–401.

[56] Johann G. Danzl et al. “Coordinate-targeted fluorescence nanoscopy with multiple off states”. In: *Nature Photonics* 10.2 (2016), pp. 122–128.

[57] Johannes Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. In: *Nature Methods* 9.7 (2012), pp. 676–682.

[58] Hazen Babcock, Yaron M. Sigal, and Xiaowei Zhuang. “A high-density 3D localization algorithm for stochastic optical reconstruction microscopy”. In: *Optical nanoscopy* 1.6 (2012).

[59] Siân Culley et al. “Quantitative mapping and minimization of super-resolution optical imaging artifacts”. In: *Nature Methods* 15.4 (2018), pp. 263–266.

[60] Klaus C. Gwosch et al. “MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells”. In: *Nature Methods* 17.2 (2020), pp. 217–224.

[61] Jennifer-Magdalena Masch et al. “Robust nanoscopy of a synaptic protein in living mice by organic-fluorophore labeling”. In: *Proceedings of the National Academy of Sciences* 115.34 (2018), E8047–E8056.

[62] Penny C. Molyneux et al. “A Mouse Primary Hepatocyte Culture Model for Studies of Circadian Oscillation.” eng. In: *Current protocols in mouse biology* 5 (4 Dec. 2015), pp. 311–329.

[63] Tran Cong Tuoc and Anastassia Stoykova. “Er81is a downstream target of Pax6 in cortical progenitors”. In: *BMC Developmental Biology* 8.1 (2008), p. 23.

[64] Elisa D’Este et al. “STED nanoscopy reveals the ubiquity of subcortical cytoskeleton periodicity in living neurons.” eng. In: *Cell reports* 10 (8 Mar. 2015), pp. 1246–51.

[65] Jervis Vermal Thevathasan et al. “Nuclear pores as versatile reference standards for quantitative superresolution microscopy”. In: *Nature Methods* 16.10 (2019), pp. 1045–1053.

[66] Julien Guy et al. “Intracortical Network Effects Preserve Thalamocortical Input Efficacy in a Cortex Without Layers.” eng. In: *Cerebral cortex (New York, N.Y. : 1991)* 27 (10 Oct. 2017), pp. 4851–4866.

[67] Katrin I. Willig et al. “Nanoscopy of filamentous actin in cortical dendrites of a living mouse.” eng. In: *Biophysical journal* 106 (1 Jan. 2014), pp. L01–3.

[68] Nicolas Olivier et al. “Resolution Doubling in 3D-STORM Imaging through Improved Buffers”. In: *PLOS ONE* 8.7 (July 2013), pp. 1–9.

[69] Sebastian van de Linde et al. “Direct stochastic optical reconstruction microscopy with standard fluorescent probes”. In: *Nature Protocols* 6.7 (2011), pp. 991–1009.

[70] Jonathan B. Grimm et al. “Bright photoactivatable fluorophores for single-molecule imaging.” eng. In: *Nature methods* 13 (12 Dec. 2016), pp. 985–988.

[71] Richard Lincoln et al. “A general design of caging-group free photoactivatable fluorophores for live-cell nanoscopy”. In: *bioRxiv* (2021).

[72] Anna M. Chizhik et al. “Three-Dimensional Reconstruction of Nuclear Envelope Architecture Using Dual-Color Metal-Induced Energy Transfer Imaging”. In: *ACS Nano* 11.12 (Dec. 2017), pp. 11839–11846.

[73] Benjamin Harke et al. “Resolution scaling in STED microscopy”. In: *Opt. Express* 16.6 (Mar. 2008), pp. 4154–4162.

[74] Fabian Göttfert et al. “Strong signal increase in STED fluorescence microscopy by imaging regions of subdiffraction extent”. In: *Proceedings of the National Academy of Sciences* 114.9 (2017), pp. 2125–2130.

[75] Zhenjia Wang et al. “Size and dynamics of caveolae studied using nanoparticles in living endothelial cells”. eng. In: *ACS nano* 3.19919048 (Dec. 2009), pp. 4110–4116.

[76] Mikhail Drobizhev et al. “Multiphoton Photochemistry of Red Fluorescent Proteins in Solution and Live Cells”. In: *J. Phys. Chem. B* 118.31 (Aug. 2014), pp. 9167–9179.

[77] P. S. Dittrich and P. Schwille. “Photobleaching and stabilization of. fluorophores used for single-molecule analysis. with one- and two-photon excitation”. In: *Applied Physics B* 73.8 (2001), pp. 829–837.

[78] Travis J. Gould et al. “Adaptive optics enables 3D STED microscopy in aberrating specimens”. In: *Opt. Express* 20.19 (Sept. 2012), pp. 20998–21009.

[79] Brian R. Patton et al. “Three-dimensional STED microscopy of aberrating tissue using dual adaptive optics”. In: *Opt. Express* 24.8 (Apr. 2016), pp. 8862–8876.

[80] Stéphane Bancelin et al. “Aberration correction in stimulated emission depletion microscopy to increase imaging depth in living brain tissue.” eng. In: *Neurophotonics* 8 (3 July 2021), p. 035001.

[81] Kevin T. Takasaki, Jun B. Ding, and Bernardo L. Sabatini. “Live-Cell Superresolution Imaging by Pulsed STED Two-Photon Excitation Microscopy”. In: *Biophysical Journal* 104.4 (2013), pp. 770–777.

[82] Zhixing Chen et al. “Extending the fundamental imaging-depth limit of multi-photon microscopy by imaging with photo-activatable fluorophores”. In: *Opt. Express* 20.17 (Aug. 2012), pp. 18525–18536.

[83] Joerg Schnitzbauer et al. “Super-resolution microscopy with DNA-PAINT”. In: *Nature Protocols* 12.6 (2017), pp. 1198–1228.

[84] Winfried Denk, David W. Piston, and Watt W. Webb. “Two-Photon Molecular Excitation in Laser-Scanning Microscopy”. In: *Handbook of Biological Confocal Microscopy*. Ed. by James B. Pawley. Boston, MA: Springer US, 1995, pp. 445–458.

[85] Deano M. Farinella et al. “Improving laser standards for three-photon microscopy”. In: *Neurophotonics* 8.1 (2021), p. 015009.

[86] Roman S. Erdmann et al. “Labeling Strategies Matter for Super-Resolution Microscopy: A Comparison between HaloTags and SNAP-tags.” eng. In: *Cell chemical biology* 26 (4 Apr. 2019), pp. 584–592.

[87] Romain F. Laine, Guillaume Jacquemet, and Alexander Krull. “Imaging in focus: An introduction to denoising bioimages in the era of deep learning”. In: *The International Journal of Biochemistry & Cell Biology* 140 (2021), p. 106077.

[88] Bowen Yao et al. “Image reconstruction with a deep convolutional neural network in high-density super-resolution microscopy”. In: *Opt. Express* 28.10 (May 2020), pp. 15432–15446.

[89] Mengzhou Li et al. “Deep adversarial network for super stimulated emission depletion imaging”. In: *Journal of Nanophotonics* 14.1 (2020), pp. 1–17.

[90] Jiji Chen et al. “Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes”. In: *Nature Methods* 18.6 (2021), pp. 678–687.

[91] Franco Scarselli and Ah Chung Tsoi. “Universal Approximation Using Feedforward Neural Networks: A Survey of Some Existing Methods, and Some New Results”. In: *Neural Networks* 11.1 (1998), pp. 15–37.

[92] Bekir Karlik and A. Vehbi Olgac. “Performance analysis of various activation functions in generalized MLP architectures of neural networks”. In: *International Journal of Artificial Intelligence and Expert Systems* 1.4 (2011), pp. 111–122.

[93] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization.* 2014.

[94] Ilya Loshchilov and Frank Hutter. *Decoupled Weight Decay Regularization.* 2017.

[95] Vincent Dumoulin and Francesco Visin. *A guide to convolution arithmetic for deep learning.* 2016.

[96] Sergey Ioffe and Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* 2015.

[97] Ian Goodfellow et al. “Generative Adversarial Networks”. In: *Commun. ACM* 63.11 (Oct. 2020), pp. 139–144.

[98] Phillip Isola et al. “Image-to-Image Translation with Conditional Adversarial Networks”. In: *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on.* 2017.

[99] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.* Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241.

[100] Sepp Hochreiter. “The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions”. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 06.02 (1998), pp. 107–116.

[101] Kaiming He et al. *Deep Residual Learning for Image Recognition.* 2015.

[102] Mehdi Mirza and Simon Osindero. *Conditional Generative Adversarial Nets.* 2014.

[103] Martin Arjovsky and Léon Bottou. *Towards Principled Methods for Training Generative Adversarial Networks.* 2017.

[104] Tim Salimans et al. “Improved Techniques for Training GANs”. In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee et al. Vol. 29. Curran Associates, Inc., 2016.

[105] Jie Gui et al. *A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications*. 2020.

[106] Oludare Isaac Abiodun et al. “State-of-the-art in artificial neural network applications: A survey”. In: *Heliyon* 4.11 (2018), e00938.

[107] A. Shanthini, Gunasekaran Manogaran, and G. Vadivu. “Deep Convolutional Neural Network Applications and Visualization Tools”. In: *Deep Convolutional Neural Network for The Prognosis of Diabetic Retinopathy*. Singapore: Springer Nature Singapore, 2023, pp. 47–56.

[108] Yair Rivenson et al. “Deep learning microscopy”. In: *Optica* 4.11 (Nov. 2017), pp. 1437–1443.

[109] Kaiming He et al. “Deep Residual Learning for Image Recognition”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778.

[110] Xingjian Shi et al. *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting*. 2015.

[111] Zhuang Liu et al. *A ConvNet for the 2020s*. 2022.

[112] Zhou Wang et al. “Image quality assessment: from error visibility to structural similarity.” eng. In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 13 (4 Apr. 2004), pp. 600–12.

[113] Allison H. Baker, Alexander Pinard, and Dorit M. Hammerling. *DSSIM: a structural similarity index for floating-point data*. 2022.

[114] Rafael Reisenhofer and Emily J. King. “Edge, Ridge, and Blob Detection with Symmetric Molecules”. In: *SIAM Journal on Imaging Sciences* 12.4 (2019), pp. 1585–1626.

[115] Joanna Oracz et al. “Photobleaching in STED nanoscopy and its dependence on the photon flux applied for reversible silencing of the fluorophore”. In: *Scientific Reports* 7.1 (2017), p. 11354.

[116] Maria J. Sarmento et al. “Exploiting the tunability of stimulated emission depletion microscopy for super-resolution imaging of nuclear structures”. In: *Nature Communications* 9.1 (2018), p. 3415.

[117] Kedar Potdar, Taher S. Pardawala, and Chinmay D. Pai. “A comparative study of categorical variable encoding techniques for neural network classifiers”. In: *International journal of computer applications* 175.4 (2017), pp. 7–9.

[118] Wei Ouyang et al. “Deep learning massively accelerates super-resolution localization microscopy”. In: *Nature Biotechnology* 36.5 (2018), pp. 460–468.

[119] Haotian Liu et al. *YolactEdge: Real-time Instance Segmentation on the Edge*. 2020.

[120] Long Zhuo et al. *Fast-Vid2Vid: Spatial-Temporal Compression for Video-to-Video Synthesis*. 2022.

[121] Daniel Bolya et al. “YOLACT: Real-Time Instance Segmentation”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2019.

[122] Tak-Wai Hui, Xiaoou Tang, and Chen Change Loy. “LiteFlowNet: A Lightweight Convolutional Neural Network for Optical Flow Estimation”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2018.

[123] Sharif Amit Kamran et al. “VTGAN: Semi-Supervised Retinal Image Synthesis and Disease Prediction Using Vision Transformers”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*. Oct. 2021, pp. 3235–3245.

A. Supplementary figures for two-photon activation

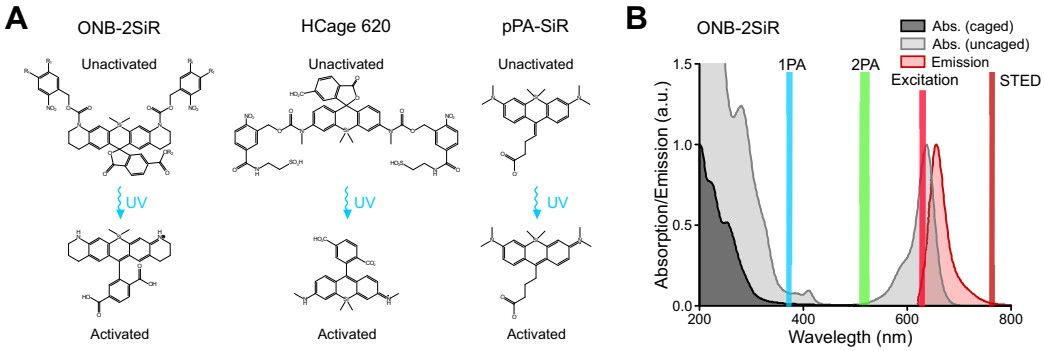

Figure A.1: Structure of the dyes used with a representative spectrum. **A** Structure of the three different dyes, ONB-2SiR, HCage 620 and pPA-SiR, used in this thesis before (up) and after activation (bottom). The classic activation with UV is depicted here. **B** A for the three dyes representative absorption spectrum of ONB-2SiR. Dark grey shows the spectrum of the caged compound. Light grey shows the absorption of the activated fluorophore. The emission spectrum of ONB-2SiR is shown in red. The typical 1PA wavelength, the 2PA wavelength used in this thesis as well as the excitation and STED wavelengths are highlighted with the colored bars.

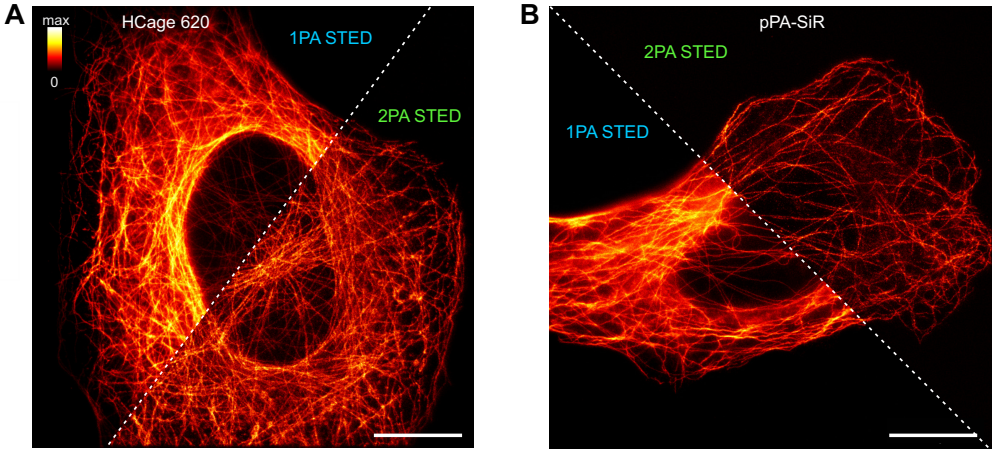

Figure A.2: 2D STED comparison after 1PA and 2PA for HCage 620 and pPA-SiR. STED of U2OS cells immunostained with **A** HCage 620 and **B** pPA-SiR. The STED image after 2PA shows better contrast compared to the STED image acquired after 1PA. Scale bars 10 µm.

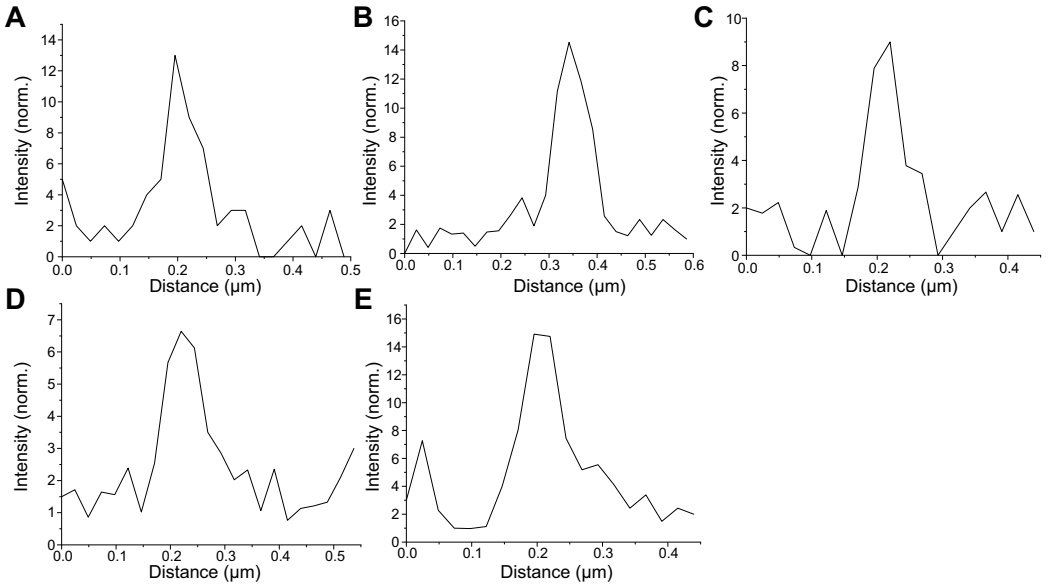

Figure A.3: Line profile of five additional microtubules from Fig. 2.6. Line profiles measured on the sample shown in Fig. 2.6 A at five additional locations to the one presented in Fig. 2.6 C.

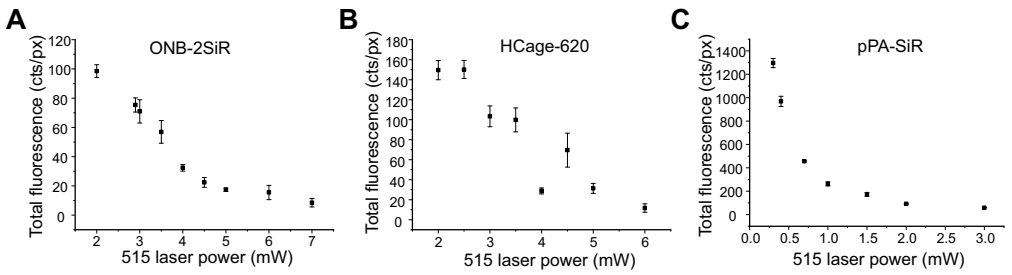

Figure A.4: Total 2PA obtainable fluorescence signal for different average activation laser power. The total integrated fluorescence signal per pixel obtained from the bleaching measurements in subsection 2.4.4 is depicted as a function of the average 2PA laser power that was used. Measurements are shown for all three dyes **A** ONB-2SiR, **B** HCage 620 and **C** pPA-SiR. All show a reduction of the total signal with an increase of the laser power, indicating a higher contribution of bleaching.

B. Supplementary figures for low count STED

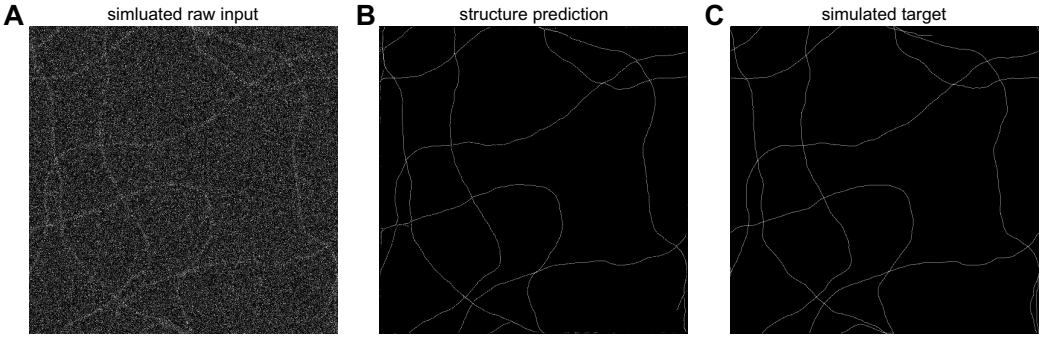

Figure B.1: Example of structure prediction on simulated noisy data. A Simulated noisy data of microtubules with an SBR of 2 and a density of $\sim 6.9\,\%$. **B** The predicted underlying structure corresponding to the sample presented in A. The Neural Network for the prediction was trained on input data with an SBR of 2. **C** The simulated target data of the structure of the noisy data shown in A.

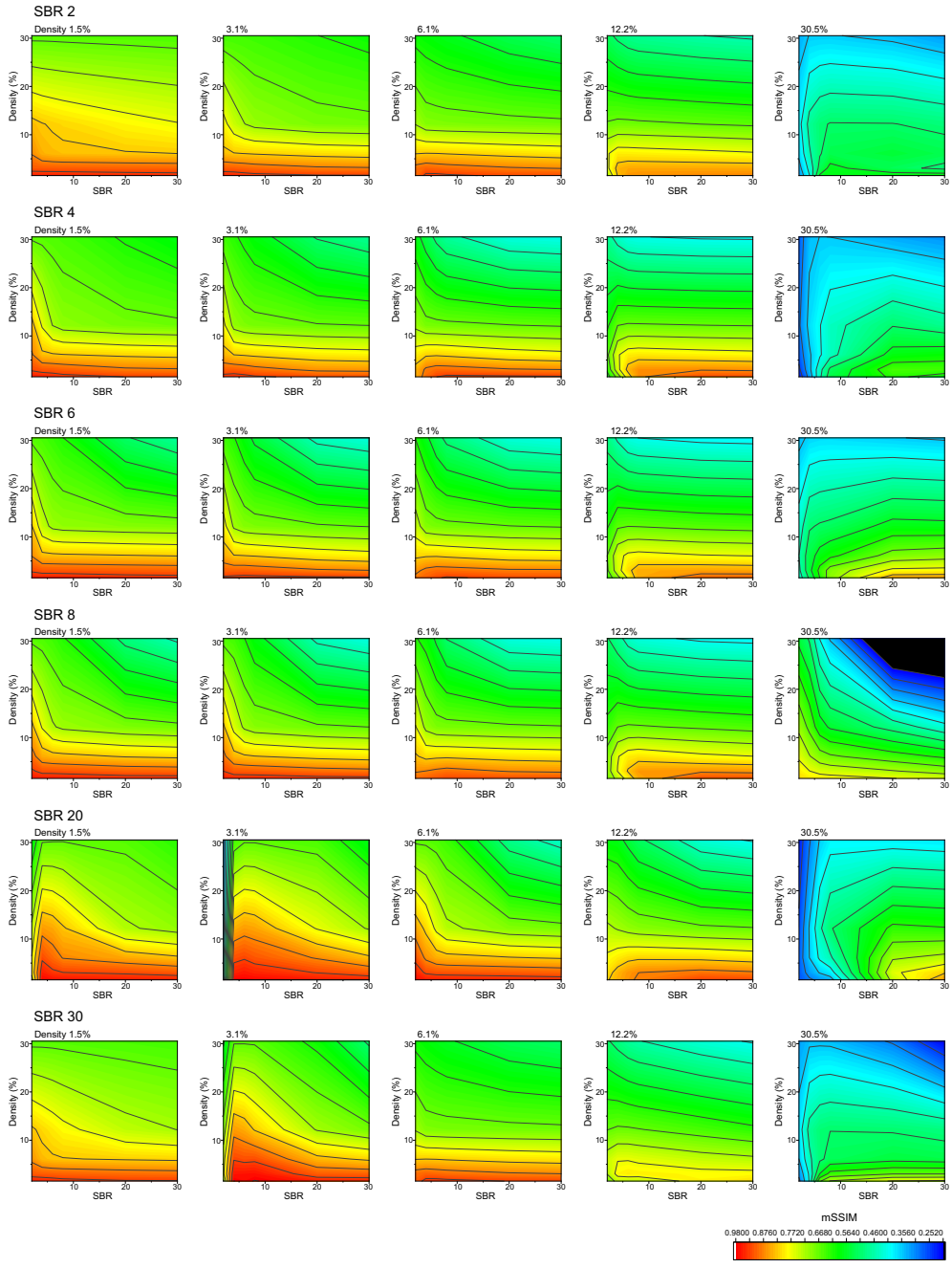

Figure B.2: Evaluation metrics for different SNRs and densities. The training of the neural network was performed with input data varying in SNR and density of the structure of interest. For each combination of parameters the validation is performed on itself and the remaining datasets. The scale bar for the validation metric, the mSSIM, is the same for all graphs. All rows of graphs had the same training SNR as indicated above each row. Above every graph the corresponding training density is shown.

List of Figures

List of Tables

Acknowledgments

I want to express my gratitude to the many people without this dissertation would not have been possible. I would like to thank them for their support provided over these times.

First, I would like to thank Stefan W. Hell for the opportunity to perform this dissertation in his lab and the many fruitful discussions we had over the times. Working in his department has built a very solid foundation not only for my future job but also for my personal life. I am very grateful for all the insights I have gained over the years.

Additionally, I would like to thank Tim Salditt for agreeing to be the second reviewer of my thesis and to be a member of my thesis advisory committee (TAC). Thank you also to the other members of my TAC and examination board, Claus Ropers, Alexander Egner, Stefan Jakobs and Silvio Rizzoli.

A special thank you goes to Joanna Oracz who started the two-photon journey with me and provided a lot of valuable input regarding many aspects inside and outside of work. Without her support and transfer of knowledge, this work would not have been possible.
Similarly, I sincerely thank Kamila Kiszka for handling all animal related tasks and for being very patient while trying many different biological models and exploring many staining techniques. Furthermore, also for being a great office mate with many interesting and enlightening discussions which provided moral support in difficult times.
Thank you as well to Mark Bates for accompanying me to numerous lunches, providing buffer solutions and kickstarting my LabVIEW career.

I am very grateful to Jan Keller-Findeisen for corrections and valuable comments on this thesis. He additionally provided a lot of feedback and guidance for programming-related tasks and the formulation of a research objective for the second part of this thesis.
On a similar note, I would like to thank Marcel Leutenegger for his comments on this thesis. Additionally, he was very instrumental in debugging LabVIEW related problems and in getting the FPGA part of this project running.
Till Stephan gave very valuable feedback for this thesis and I would like to thank him for his time and comments.
I would like to thank Steffen Sahl for his collaboration on finishing and submitting a manuscript to publish the described work about 2PA.

I am very thankful for the continuous support regarding different stainings and biological sample preparations by Ellen Rothermel and Tanja Koenen. The same is true for Marco Roose and Simon Piechota who supplied me with the much needed technical material for this work. I would like to thank the workshop for electronics and precision mechanics, as well as the facility for synthetic chemistry.

Much needed synthesized dye samples were provided in different forms by Alexey Butkevich (HCage 620), Michelle Frei (pPA-SiR) and Taukeer A. Khan & Michael Weber (ONB-2SiR).
I would also thank Thea Moosmayer and Volker Westphal for the opportunity to collaborate with them on the MINFLUX setup and Michael Weber and Marcel Leutenegger for providing access and introduction to the MINSTED setup.